DISSERTATION

KLASSIFIZIERUNG UND DETEKTION VON ÄNDERUNGEN UNTER VERWENDUNG VON PUNKTWOLKEN

Ausgeführt zum Zwecke der Erlangung des akademischen Grades eins

Doktors der technischen Wissenschaften (Dr.techn.)

Unter der Leitung von

Univ.-Prof. Dipl.-Ing. Dr.techn. Norbert Pfeifer

E 120.7

Department für Geodäsie und Geoinformation

Forschungsgruppe Photogrammetrie

Eingereicht und der Technischen Universität Wien

Fakultät für Mathematik und Geoinformation

von

Tran Thi Huong Giang

Matrikelnummer: 1328711

Wien, im Mai 2018 .....

DISSERTATION

CLASSIFICATION AND CHANGE DETECTION USING POINT CLOUDS

A thesis submitted in fulfillment of the academic degree of

Doktors der technischen Wissenschaften (Dr.techn.)

under the supervision of

Univ.-Prof. Dipl.-Ing. Dr.techn. Norbert Pfeifer

E 120.7

Department of Geodesy and Geoinformation

Research Group Photogrammetry

Research conducted at TU Wien

Faculty of Mathmetics and Geoinformation

by

Tran Thi Huong Giang

Matriculation number: 1328711

Vienna, May 2018 .....

Supervisor/Reviewer: Prof. Dr. Norbert Pfeifer
Department of Geodesy and Geoinformation
Research Group Photogrammetry
Vienna University of Technology
Gusshausstraße 27-29, 1040, Vienna, Austria.
E-Mail: Norbert.Pfeifer@geo.tuwien.ac.at

……

Reviewer: Prof. Dr. Jantien Stoter
Department of Urbanism
Research Group 3D Geoinformation
Delft University of Technology
Julianalaan 134, Delft 2628BL, the Netherlands.
E-Mail: j.e.stoter@tudelft.nl

……

Reviewer: Prof. Dr. Wolfgang Kainz
Department of Geography and Regional Research
Vienna University
Universitätsstraße 7/5, 1010, Vienna, Austria
E-Mail: wolfgang.kainz@univie.ac.at

……

ERKLÄRUNG ZUR VERFASSUNG DER ARBEIT

AUTHOR'S STATEMENT

Hiermit erkläre ich, dass ich diese Arbeit selbstständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen -, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

I hereby declare that I independently drafted this manuscript, that all sources and references are cited, and that the respective parts of this manuscript – including tables, maps, and figures – which were included from other manuscripts or the internet either semantically or syntactically are made clearly evident in the text and all respective sources are correctly cited.

……..

Tran Thi Huong Giang

ACKNOWLEDGEMENTS

I would like to thank all people who made this dissertation possible. Special thanks go to my best supervisor Norbert Pfeifer for his great supervision during the time I am doing my PhD in TU Wien university, to Johannes Otepka, Gottfried Mandlburger for their great support in technique during the time I do my research, to Camillo Ressl, Markus Hollaus, Milutin Milenkovic, Di Wang for their brilliant co-authorship, to Markus Pöchtrager for his enthusiasm to help me translate my English into German, to Nan Li, Ana Lyra, Ana-Maria Loghin, Livia Piermattei, Sabine Horvath, Stephan Muckenhuber,… for the time we share together, to the entire Photogrammetry Research Group for unique work environment and extraordinary team work, to all the colleagues of the Department of Cartography, Hanoi University of Mining and Geology for support working and researching, and to my family and friends who supported me throughout the years.

I want to thank also all people I haven't mentioned personally for valuable discussions and team work on conferences, workshops, meetings or elsewhere, and all scientists who I haven't met personally but who paved the road for this thesis.

Last but not the least, my sincere thanks to my beloved father Tran Van Son, my beloved mother Nhu Thi Lan Huong, my soul parent in law, Nguyen Van Nghia and my warm-heart mother in law Bui Thi Quy, my dearest husband Nguyen Ba Duy, for their continuing support and patience during this period. Thank you for their encouragement and believing in me, which helped me to pursue my research study towards the end.

KURZFASSUNG

Punktwolken haben sich zu einer mächtigen Datengrundlage für die Erstellung von 3D Modellen entwickelt und finden dadurch Anwendung in verschiedenen Disziplinen. Zum Beispiel in der Verwaltung natürlicher Ressourcen, dem Umweltmanagement sowie im städtebaulichen Kontext. 3D-Punktwolken spielen eine wesentliche Rolle in der Bereitstellung und Aktualisierung von räumlichen Informationen, während herkömmliche Rasterbilder zweidimensionalen Beschränkungen unterliegen. Mittlerweile wurde ein breites Spektrum an Sensoren und Plattformen für die Aufnahme von Punktwolken entwickelt. Das Laserscanning und das Dense-Image-Matching sind die zwei wichtigsten Aufnahmetechniken für die großflächige Erfassung von Punktwolken. Die Anzahl der veröffentlichten wissenschaftlichen Beiträge im Bereich der Klassifizierung von Punktwolken und der multitemporalen Veränderungsdetektion (Change Detection) ist stetig steigend. Ein Großteil der Beiträge zu Klassifizierung und Veränderungsdetektion konzentriert sich jedoch eher auf 2D Rasterdaten, als auf Punktwolken. Des Weiteren steht das Interesse an einer Klassifizierung von Punktwolken aus Image Matching jenem einer Klassifizierung von ALS-Punktwolken noch deutlich nach.

In dieser Dissertation werden Klassifizierung und Veränderungsdetektion mittels rasterbasierter und punktbasierter Ansätze bei unterschiedlichem Detaillierungsgrad untersucht. Diese Arbeit beschäftigt sich des Weiteren mit der Erkennung geeigneter Attribute für die Klassifizierung und Veränderungsdetektion, dem Austausch von Schwellwerten für die Attribute zwischen unterschiedlichen Datensätzen und Testgebieten, sowie der Evaluierung von Machine Learning Ansätzen und deren Nutzen für die Klassifizierung und Veränderungsdetektion. Die Forschungsfragen reichen von der Wahl der Messmethode über die Merkmalsextraktion bis hin zu Prozessierungsmethoden und wurden in vier Forschungsartikel untersucht und bewertet. Die präsentierten Studien wurden, einem peer-review Prozess unterzogen, in facheinschlägigen Journalen und in Konferenzbeiträgen veröffentlicht. Die Beiträge I und II untersuchen die Klassifizierung von (i) Full-Waveform ALS-Punktwolken und (ii) Punktwolken aus Image Matching, unter der Verwendung von einfachen Entscheidungsbäumen und Machine Learning Methoden. Die vorgestellten Ansätze

zeigen großes Potential für die Klassifizierung verschiedener Objekttypen im städtischen Gebiet. Beitrag III untersucht die Veränderung des Baumbestandes unter Verwendung von herkömmlicher rasterbasierter Differenzbildung. Die entwickelte Methode findet neue Merkmale in der LiDAR Punktwolke, welche entscheidend für die Veränderungsdetektion von einzelnen Objekten im Waldgebiet sind. Der abschließende Beitrag IV untersucht die simultane Detektion und Klassifikation von Veränderungen für mehrere Objekttypen im städtischen Gebiet, auf Basis von ALS-Daten.

Die vorgestellten Beiträge zeigen, dass die Punktwolke, egal ob aus Airborne Laserscanning oder Image Matching, eine effektive und zweckdienliche Datenquelle für eine großflächige und präzise Klassifikation und Veränderungsdetektion darstellt.

ABSTRACT

The point cloud is a very powerful source for deriving 3D models which are widely applied in natural resource, environmental management, and urban domain. Point cloud classification and change detection are used in the context of Earth observation to monitor and assess the status and change of the natural and built environment. They have an essential role in providing and updating information in three dimensions compared to the provision of 2D information from traditional raster images. There are a number of sensors and platforms that acquire point clouds at different resolutions and spaces, in those, airborne laser scanning (ALS) and image matching (IM) are two main sources which allow to collect point clouds over large areas. The number of published research articles regarding to point cloud classification and change detection is increasing. Many studies uses ALS data on classification and change detection, but concentration on raster, and fewer publication on point clouds. In addition, image matching point cloud classification draw a less attention so far compared to ALS data.

The objectives of this dissertation are focused on point cloud classification and change detection based on raster-based and point-based approaches to consider advantages they bring in different levels of details and types of datasets. This includes finding effective attributes for classifying and detecting changes, transferring attribute thresholds between different data sets and locations, and evaluating the benefit of machine learning in classification and change detection. The study questions range from measurement technology via feature derivation to processing methods are investigated and evaluated in four research articles. The presented studies and are published in peer-reviewed journals and a conference paper. Article I and II investigate the classification using (i) full-waveform airborne laser scanning, and (ii) an image matching point cloud based on simple decision tree and machine learning method. The presented approaches show high potential for classifying multiple objects over urban areas. Article III investigates the reduction of individual trees in forested area using traditional image differencing method. The presented method finds new features of the LiDAR point cloud, which are useful for detecting single object change in wooded areas. Finally, Article IV investigates the integration of detecting and classifying changes simultaneously for multi-objects change detection in urban area based on airborne laser scanning data.

The presented studies prove, that the point cloud, either acquired by airborne laser scanning or by image matching, is an effective and practicable data source for accurate classification and change detection in large areas.

LIST OF PUBLICATIONS

Parts of the present work have already been published in several peer-reviewed journal papers and at a conference, which include:

I Tran, T.H.G.; Nguyen, B.D.; Milenkovic, M.; Pfeifer, N. Potential of full waveform airborne laser scanning data for urban area classification – Transfer of classification approaches between missions. In Proceeding of The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Volume XL-7/W3, 2015, Berlin, Germany. DOI: 10.5194/isprsarchives-XL-7-W3-1317-2015.

II Tran, T.H.G.; Otepka, J;D.; Wang, D.; Pfeifer, N. Classification of image matching point clouds over urban area. *International Journal of Remote Sensing* **2018**, 39:12, 4145-4169, DOI: 10.1080/01431161.2018.1452069.

III Tran, T.H.G.; Hollaus, M.; Nguyen, B.D.; Pfeifer, N. Assessment of Wooded Area Reduction by Airborne Laser Scanning. *Forests* **2015**, 6(5), 1613-1627. DOI:10.3390/f6051613.

IV Tran, T.H.G.; Ressl, C.; Pfeifer, N. Integrated Change Detection and Classification in Urban Areas Based on Airborne Laser Scanning Point Clouds. *Sensors* **2018**, 18(2), 448. DOI: 10.3390/s18020448.

The contribution of Tran Thi Huong Giang to the papers included in this thesis as follows:

I Developed the method, implemented the classification workflow, performed the analyses and validation, and wrote the major part of the manuscript.

II Generated the image matching point clouds, implemented the workflow, performed the analyses, manual classified the reference data used for validation, and wrote the major part of the manuscript.

III Developed and implemented the change detection method, performed the analysis and validation, and wrote the major part of the manuscript.

IV Developed, implemented the change detection and classification method, performed the analyses and manual classified the reference data, and wrote the major part of the manuscript.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND ACRONYMS

2D	Two-Dimensional
2.5D	Two-and-a-Half-Dimensional
3D	Three-Dimensional
ALS	Airborne Laser Scanning
DEM	Digital Elevation Model
DIM	Dense Image Matching
DSM	Digital Surface Model
DTM	Digital Terrain Model
LiDAR	Light Detection And Ranging

1. INTRODUCTION

Humankind, assisted by technology, has made significant headway in natural resources management, environment protection and other such prerequisites for sustainable development. The advent of remote sensing technology is a vital turning point, enabling easy, rapid and large-scale collection of spatial data in 2D as well as 3D.

To overcome the limitations of 2D maps in simulating the real world with an additional dimension in height, 3D geoinformation plays an essential role in human's society. It contains an immense information potential, which brings in advantages in a variety fields of application in public and private sectors. Altmaier and Kolbe (2003) named the uses of 3D in the areas of architecture, urban and transport planning, surveying and mobile telecommunications. 3D models become increasingly important in the fields of marketing, video games compelling, tourism, navigation, community protection (noise and pollutant dispersion, flood protection, disaster management), real estate management (banks, brokers, assurances), and facility management. Final derivatives from collected 3D data include Digital Elevation Models (DEM), extracted human-made features (buildings, power lines, roads) and natural features (vegetation areas, single trees, vertical tree structures). The third dimension represents one of the opportunities of contemporary cartography, which is widely spread for very different purposes (Popovic et al. 2017). With the fast development of the Internet, lightweight 3D geographic scenes visualization systems are expected to play a more important role in web-based GIS, which is useful for cartographic analysis, destination query, data query and map mark with the system in browsers (Miao et al. 2017). Automatic 3D cartography and modeling have gained immense interest in the scientific community, due to the ever-increasing demand for landscape analysis for different popular applications coupled with the recent advanced in 3D data acquisition technologies (Aijazi et al. 2013).

Due to the extraordinary technology development of sensors, platforms and algorithms for 3D data acquisition and generation, accurate 3D data is more available and accessible. Terrestrial LiDAR and low-cost, compact, lightweight LiDAR system are more affordable. Besides, the recent development of automated image geo-referencing (Lee and Oh 2014, Pierrot Deseilligny and Clery 2011, Tao 2000) and advanced dense image matching (DIM) techniques have dramatically increased the availability of image-based

3D information (Remondino et al. 2014) with upgraded quality. Image data from UAV (unmanned aerial vehicle) and satellite platforms can be efficiently processed automatically, moving from images to point clouds and digital surface models (DSM). LiDAR is an alternative measurement method to automatically produce comparable dense point clouds within a reasonable processing time (Gehrke et al. 2010). Additionally, developed optical satellite sensors enable obtaining large scale (even multi-view) stereo images with sub-meter spatial resolution (such as Worldview, GeoEye images), with short revisit cycles. Nanosatellite systems (Barnhart et al. 2007), which coordinate a series of low-cost optical satellites in orbit, can continuously acquire high-resolution images with a global coverage on a daily basis. These sources make the 3D information more available and thereby have motivated considerable interest in using such 3D data for classification and change detection problems.

A 3D point cloud has clearly more advantages compared to the limitation of 2D (or 2.5D) image-based classification and change detection with respect to perspective distortion, higher spectral variability, and lack of volumetric information (Qin et al. 2016). Thus, 3D point clouds have drawn a lot of attention because of their essential role in providing and updating information of the real world at large-scale. The number of published research articles regarding these topics is increasing (Aijazi et al. 2013, Brodu and Lague 2012, Niemeyer et al. 2014, Qin and Gruen 2014, Yan et al. 2015). A great variety of classification and change detection methods addressing different methodologies and related questions can be found in the literature, which can be categorized into raster-based, point-based, and object-based approaches.

Raster-based and point-based methods have limitation in salt-and-pepper effects compared to object-based classification. However, they are simpler and not affected by over-segmentation or under-segmentation errors (Liu and Xia 2010). Furthermore, there are still gaps to investigate their utilities in classification and change detection (i.e., transferring classification methods between missions, finding new features for change detection, integrate methods for detecting change and classification).

The goal of this thesis is to focus on raster-based and point-based classification as well as change detection of point clouds in different levels of details and point cloud datasets. The raster has advantage in processing time and is the traditional approach. The point cloud has information in height, but requires large storage and longer processing time. The thesis is carried out cumulative, and brings together four published articles regarding

this topic. In chapter 1 an introduction and the statement of the research problem are presented. In Chapter 2, specific objectives of this thesis are presented. Chapter 3 includes a summary of the research articles. Finally, the results are discussed and conclusion are drawn in chapter 4.

1.1. Statement of the research problem

This thesis focuses on using point clouds for classification and change detection. This raises a lot of questions.

- Which advantages and disadvantages do raster-based and point-based approaches bring in classification and change detection?
- What efforts and gains does machine learning bring in classification and change detection?
- How to transfer rules for classification across different modalities (scale, location, point vs raster)?

All the questions are answered by the publication of research articles, and these built the core of this thesis. These include (i) urban area classification based on full-waveform airborne laser scanning data and transfer classification approaches between missions, (ii) classification of image matching point clouds over urban area, (iii) assessment of wooded area reduction by discrete return airborne laser scanning, (iv) integrated change detection and classification in urban area based on ALS data.

An essential contribution of this thesis is to present the potential of point cloud for classification and change detection in different datasets and different dimensions (2D and 3D) and to discuss the benefits and drawbacks of the presented methods.

It shall be noted, that the concentration in this thesis is not in comparing different machine learning approaches or developing them further. Thus, one machine learning method, random forests (Breiman 2001) is used in this thesis.

1.2. Background of the scientific publications

This thesis and the included publications are presented in the context of Earth observation (EO), photogrammetry and monitoring of forestry and urban areas. An important aspect of point cloud classification and change detection is the acquisition and interpretation of laser scanning data, image matching data and the sequential extraction of features from it. These tasks are related to EO, photogrammetry, and machine learning.

EO is the gathering of information about planet Earth's physical, chemical and biological systems via remote sensing technologies supplemented by earth surveying techniques, encompassing the collection, analysis, and presentation of data. Earth observation is used to monitor and assess the status of, and changes in, the natural environment and the built environment (Solimini, 2016).

Photogrammetry allows to reconstruct the position, orientation, shape, and size of objects from pictures, for example, high-resolution images (Kraus et al. 2007). Classes and changed information, extracted from the ALS and image matching point cloud, are of interest for the urban and forestry domain.

Machine learning is a field of computer science that gives computer systems the ability to "learn" with data, without being explicitly programmed (Samuel 1959). Machine learning drew a great attention in classification. The whole process of classification and change detection is vital for the task of storing up to date the information.

2. OBJECTIVE

The objective of this thesis is research on point cloud classification and change detection from raster-based (2D and 2.5D) to point-based (3D) data. The developed methods are applied to different dataset (full-waveform ALS, discrete return ALS, image matching point clouds), and different study areas. The questions range from measurement technology via feature derivation to processing methods.

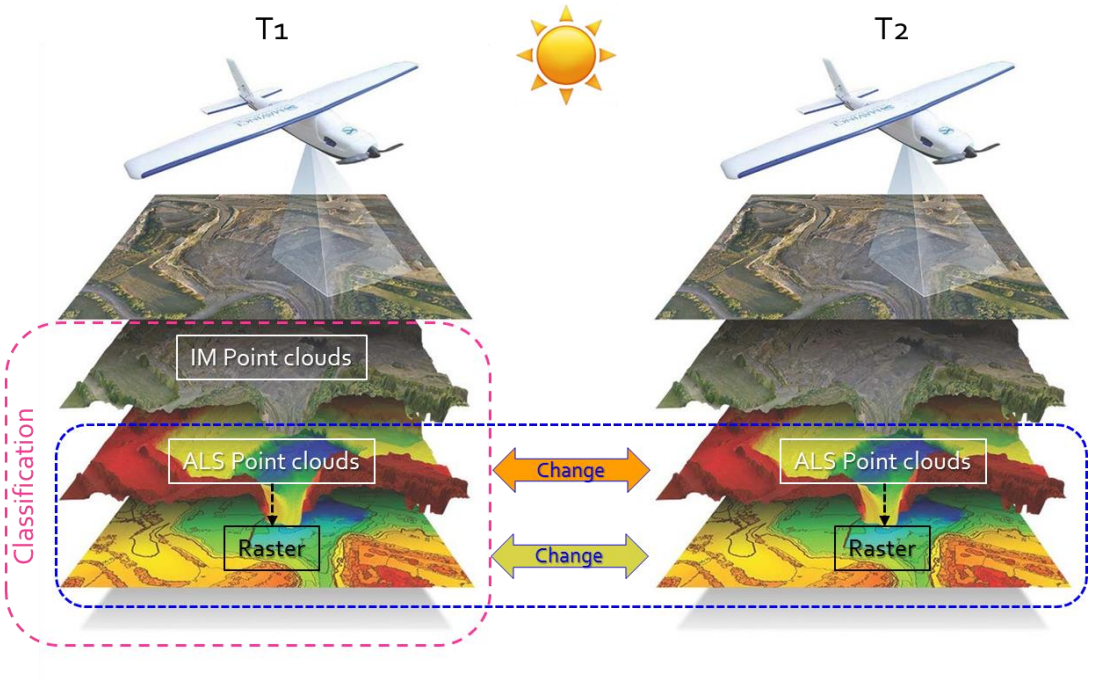

Figure 1. Classification and change detection using different types of point cloud data (adapted from Topcon Aerial Mapping http://www.topconsolutions.com)

The specific objectives, as well as the novelties of the approaches for the Articles I-IV, are:

I. Raster-based classification using full-waveform laser scanning

(a) Investigate the potential of full waveform ALS data in urban classification.

(b) Demonstrate that high classification accuracy can be reached by a simple decision trees.

(c) Study if this classification approach using the selected features and the thresholds can be transferred to another dataset.

(d) Suggest a method to re-compute the echo width threshold for different missions acquired for urban full waveform point clouds.

Novelty of the approach: Development of a decision tree approach for classifying four main classes (building, vegetation, ground, and water) in urban areas based on attributes derived from full-waveform ALS data (i.e., Echo Ratio, Echo width, Sigma0, nDSM and point density value). The classification tree is transferred between different study areas. We suggested a simple normalization method for attributes which have different value ranges in different missions.

II. Point-based classification using image matching point cloud

(a) Generate the colorized 3D point clouds from aerial images by utilizing the dense image matching technique semi-global matching (SGM).

(b) Investigate the potential of 3D image matching point cloud for urban classification.

(c) Develop different classification methods (automatic machine learning classifier and manual decision tree classifier) for different levels of point density and study sites.

(d) Consider the advantage 3D brings over raster approaches in urban scene.

(e) Investigate the maintainable features for different study sites and pyramid levels (i.e., resolution levels).

Novelty of the approach: Focus on studying image matching point cloud classification, and on the classification results by a change in resolution. The advantages of machine learning or an operator-based decision tree approaches have relative to each other over complex urban scenes are also investigated.

III. Single object change detection

(a) Investigate the reduction of single object change from two different ALS datasets by using image differencing.

(b) Find new features of LiDAR point clouds, which are useful for change detection, and show low dependence of the sensor characteristics.

Novelty of the approach: Developing a robust methods in the presence of different ALS mission parameter to detect change of forest area. The lost tree areas are detected based on different factors (DSM, slope adaptive echo ratio (sER), and Sigma0). New factors are investigated, which are significant for change detection in forested area but not influenced

by geo-referencing problem. This allows new applications and simplifies existing applications of discrete return ALS data in forest change detection and management.

IV. Multi-objects change detection and classification

(a) Investigate features useful for point cloud change detection.

(b) Investigate the method to detect the change in multi-objects immediately.

(c) Develop a new approach combining change detection and change classification in one step.

Novelty of the approach: The integration of change detection and classification, which has not been done so far. The method provides a separation whether there is a change or no change at the location of the point, as well as individual class information for each point.

A graphical representation of the thesis, especially on the four articles, is given in Figure 2. The common ground of all four articles are collecting the data in optical domain with the ground sample distance (GSD) range from 6 to 16 cm, and analysis in static state. All article solved the problems of classification or change detection task of the point clouds which are acquired by ALS or image matching. The point clouds are then applied in raster-based and point-based approaches to classify and detect the change of single or multiple objects.

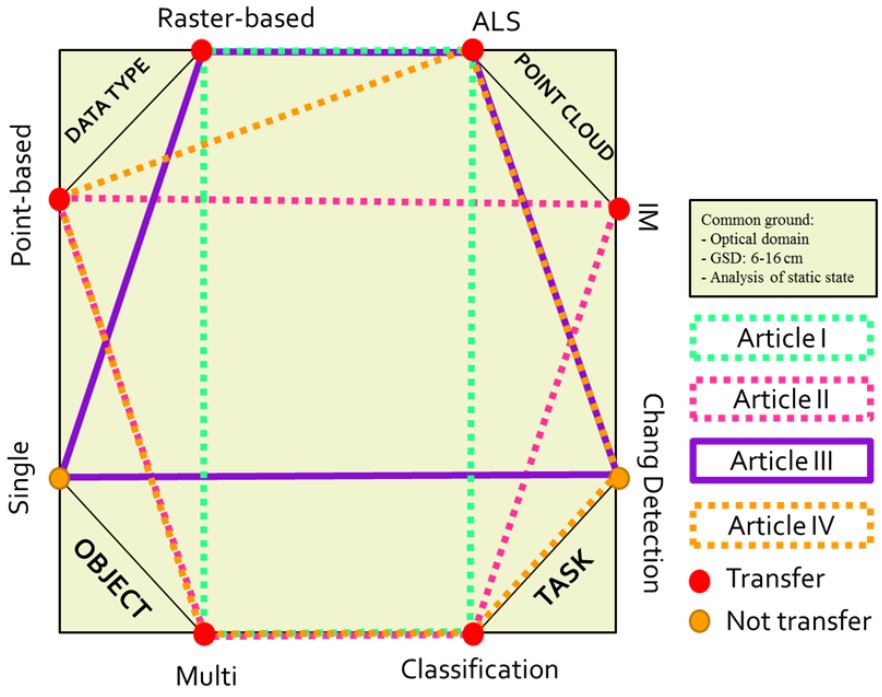

Figure 2. Graphical overview of thesis objectives, and relation to articles.

3. SUMMARY OF PUBLICATIONS

Four publications focus on classification and change detection of point cloud data. Paper I focuses on the abilities of traditional raster-based classification, explores the potential of additional attributes of full-waveform data in classification and a transfer approach between missions. Publication II shows the point-based classification of image matching point clouds using two supervised methods for different locations and scales. Publication III investigates single object change detection using the traditional image differencing method. Publication IV detects the changes of multi-objects directly in the point cloud using machine learning.

3.1. Article I: Potential of Full Waveform Airborne Laser Scanning Data for Urban Area Classification – Transfer of Classification Approaches between Missions

Article I investigates the advantages of full-waveform (FWF) LiDAR data for urban classification and the transfer of the classification rules between different study sites using a traditional rater-based approach. The paper was published in May 2015 in the Proceeding of The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences conference.

FWF data, exploiting signal processing methods, provides additional information next to the point coordinates to overcome many drawbacks of classical multi-echo LiDAR data (discrete return) on reflecting object characteristics. The classification methods applied for FWF data reach from simple decision tree to support vector machines. However, the transferability of classification approaches between different FWF LiDAR datasets has received less attention. To study the transferability of classification approaches between different full waveform LiDAR data sets, this paper quantified the transferability of full-waveform airborne LIDAR-based parameters (e.g., echo width, Signa0, echo ratio, and nDSM) and produced corresponding built up area classification maps.

Two full-waveform airborne LiDAR data were available for the two study sites: Eisenstadt and Vienna. Both datasets were acquired with a Riegl LMS-Q560 sensor, with density and a diameter laser-footprint of 8 points/m^2 and 60 cm (Eisenstadt) and 12 points/m^2 and 30 cm (Vienna), respectively. Both raw full-waveform data sets were

processed in a similar way using the software OPALS and sensor manufacturer software. As input the georeferenced point cloud with features derived from the measurement process were used. Additionally, a number of attributes were computed for each point, using the paradigm of point cloud processing. From these attributes different images are computed ("gridding") at a pixel size of 1m. A terrain model is derived also. Then, a decision tree is applied to classify each pixel into one of the four classes: building, vegetation, ground, and water body. Water is first identified by its low point density. Buildings and trees are distinguished from each other by height (nDSM), surface roughness, and ER. All pixels not classified so far are considered as ground. Image algebra (e.g., morphological operations) is used in between to refine the results. The quality of the results is assessed using the completeness and the correctness measure reaching value more than 93% for building and vegetation.

An initial assumption was that the thresholds for the decision tree derived for one data set can also be used for the other data set. Density is a physical measure (points per square meter) and the overall shot density was similar (8 vs. 10 points per square meter). Height above ground (nDSM) is a measure independent of the measurement device and also independent of the sampling distance. Echo Ratio is by definition a relative measure and should therefore adapt itself to the data distribution. Sigma0 is the local plane fitting accuracy. For data sets of similar measurement accuracy (same sensor model used for both areas) and similar neighborhoods, both number of neighbors and spatial extent, it should deliver comparable values. Echo width obviously depends on the width of the emitted pulse (same sensor model used for both data sets), but may also depend on the footprint diameter (which was different in the two data sets investigated) or other effects. Due to the doubts of echo width transferability, a method to normalize echo width is suggested. Weak, low amplitude echoes typically lead to a poor determination of echo width. Thus only stronger echoes (larger amplitude) are used for deriving the echo width normalization parameters. Normalized echo width from one dataset is applied for classifying the other data set. The Eisenstadt values applied to Vienna led to a result of lower quality, but the Vienna thresholds applied to Eisenstadt did produce qualitatively a very similar classification. The loss in completeness and correctness does not change for the building class and is below 5% for the vegetation class.

The results of the paper demonstrate that high raster-based classification accuracy can be reached with decision trees. The study also demonstrate that this classification approach using the selected features and thresholds can be transferred to another data set. Finally a

method was suggested to re-compute the echo width threshold for different missions acquired by urban full waveform point clouds.

3.2. Article II: Classification of Image Matching Point Clouds over Urban Area

Article II investigates the classification in urban area from image matching point clouds. The improvement of this article is the increase in dimension of point cloud classification from 2D to 3D. It is studied how classification is affected by a change in resolution, if the thresholds of attributes can be transferred between sites and pyramid levels, if the 3D approach allows higher accuracy compare to 2.5 raster-based approach, and which advantages machine learning or operator-based decision tree approaches have relative to each other over complex urban scenes. The paper was published in the International Journal of Remote Sensing.

There are two sources for point clouds over large urban areas: airborne laser scanning and dense matching of aerial images. So far, little research on using the point clouds from high density image matching for classification of urban scenes was conducted. Because of the regular availability of aerial images (e.g. yearly photo flights), classified point clouds derived from image matching would allow frequent updates of derived products. It is, therefore, interesting to study the feasibility of classifying 3D point clouds from dense image matching and assess the accuracy that can be obtained.

The point clouds are generated from high resolution aerial images taken by UltraCam-xp camera. Dense image matching technique semi-global matching is applied for different pyramid levels (0, 1, and 2, corresponding to GSD of 6 cm, 12 cm, and 24 cm). Those data sets are then evaluated regarding classification accuracy, completeness and time consumption. The pyramid level 0 point cloud is used to create a digital terrain model (DTM) used for feature computation at all levels. In order to focus the research, three particular test areas were chosen which reflect the variability of urban environments.

Additional geometric and spectral features of a point and its neighborhoods are computed in the point cloud of each pyramid level. Those features are then input for two supervised classification methods: (1) simple (subjective) decision tree method which uses thresholds determined by a human operator, and (2) machine learning method based on the random forest algorithm, which automatically generates the best threshold to distinguish various urban objects.

For a simple decision tree method, the features are rasterized for empirical threshold selection to construct a rule-based hierarchical decision tree. Firstly, ground points (sealed surface and grassland) are extracted from the point cloud. Then elevated points are separated into lower objects (cars and low trees), medium height objects (small buildings and medium trees) and higher objects (high buildings and high trees). Finally, the four height classes are split into the mentioned sub-classes. Thresholds of the simple decision tree of one region are transferred to the other two regions. Similarities for different pyramid levels are also investigated.

Machine learning classifiers infer the classification rules from the annotated training data with minimal human intervention. For each study site, sampling points for each class are taken manually. The training samples contain "pure" data and required careful selection. To evaluate, if the 3D point-based approach offers a gain in comparison to a raster-based approach, the same machine learning method is applied to the point cloud of level 0 after gridding. Finally, reference classifications were thoroughly done by an operator and compared with classification results on the urban test data scenes.

Both methods obtain the highest classification accuracy in the point cloud of pyramid level 0, which is reduced evenly per level. In the same level, machine learning method always gets higher accuracy due to its higher number of input parameters and more complex decision tree structure: around 87.2% while for the simple manual decision tree it is 84.1%. For higher pyramid levels (GSD 12 cm and 24 cm) the accuracy of the classification drops per level by 4% for machine learning and 7% for the simple decision tree, respectively. Machine learning applied to rasterized data provided slightly lower accuracy than the point-based method for the pyramid level 0, namely by 5%.

3.3. Article III: Assessment of Wooded Area Reduction by Airborne Laser Scanning

This paper investigates the single object change detection using discrete return ALS data. The article is published in Forests.

Forests are an important factor in maintaining the balance in the Earth system. With the advantage of penetrating the canopy through small gaps, Airborne Laser Scanner (ALS) is a potential technique for monitoring vegetation changes compare to optical remote sensing and photographic techniques. Using 3D point clouds from ALS, the change in both coverage and height can be detected. Although ALS data holds a high promise in vegetation change detection, thus far, research using multi-temporal ALS to detect forest change cover has not yet been fully explored. In this research, we investigate the ability

of forest reduction detection from two different ALS datasets by using image differencing, a traditional pixel-based change detection method. Our aim was to find features of LiDAR point clouds, that are, as much, as possible independent of the sensor characteristics.

The study is done for a mountain forest in Vorarlberg, Austria. The total covered area is about 1.5 km^2 of mountainous region, with the elevations ranges from 1225 m above sea level (a.s.l) in the valleys, to a maximum of 1786 m a.s.l. In general, in this area, approx. 24% is covered by forest and the dominating tree species is Norway spruce (Picea abies). The ALS data sets were acquired in 2005 and 2011 using an Airborne Laser Terrain Mapper systems (ALTM 1225) and a Trimble Harrier 56 system, respectively. All ALS data sets were acquired under snow-free and leaf-off conditions and were available as georeferenced 3D-point clouds.

In this publication, we supposed that forest reduction up to individual trees can be observed by three different LiDAR derived models: Digital Surface Model (DSM), Slope-adaptive Echo Ratio (sER), and "Sigma0" (a local roughness measure), as well as their combinations. The DSMs demonstrate the change in height and, thus, indicate that tall objects were removed. sER demonstrates the change in vertical penetrability and indicates that layered objects (e.g., understory and canopy) were removed. Sigma0 demonstrates the change in the vertical dispersion of the points and indicates that objects distributed in height (e.g., trees, bushes) were removed. As the decreases of the elevation of the canopy surface indicate the loss of trees, the DSM_{2005} is subtracted from the DSM_{2011}. The thresholds for all change values (i.e., ΔDSM, ΔsER, and ΔSigma0) are assessed empirically. To acquire the threshold for each combination of two models, ΔsER versus ΔDSM, ΔsER versus ΔSigma0, and ΔDSM versus ΔSigma0, the feature space is used to distinguish changed areas. Finally, six change detection outputs were established: DSM only, sER only, Sigma0 only, DSM combined with sER, sER combined with Sigma0 and DSM, and sER and Sigma0. Based on image interpretation of aerial orthophotos with additional use of 3D point cloud viewing of the raw ALS data, the referent data were derived. The six final change detection results are then compared to the referent based on completeness and correctness.

The combination of sER and DSM achieved the highest correctness and completeness of 92% and 85% respectively. While many studies have, thus far, used DSM and its change, we found that sER is a good single predictor for tree cover change. This study opens up

a new application of discrete return ALS data in forest change detection and, therefore, in forest management.

3.4. Article IV: Integrated Change Detection and Classification in Urban Areas based on Airborne Laser Scanning Point Clouds

Article IV suggests a new approach in change detection (CD) in 3D point clouds. It combines classification and CD in one step using machine learning. The point cloud data of both epochs are merged for computing features. All these features are merged in the points and then training samples are acquired in order to create the model for supervised classification, which is then applied to the whole study area. The paper was published in Sensors.

Many approaches suggested in the literature demonstrate the high potential of LiDAR point clouds for change detection. Most studies apply two steps: first, detect the change and then, second, classify the change, or alternatively first, classify objects for both periods and then, second, detect changes between the classifications. Both approaches, consequently, will be influenced by sequence error, i.e., the accuracy of classified changes depends on the change detection method and the classification method. Furthermore, most of those studies focus only on one class (typically either building or trees). However, a change typically does not happen in a single class only, but leads to changes in multiple classes. We are therefore suggesting to investigate the possibilities of performing change detection and classification of all the main classes (building, tree, and ground) simultaneously in one step.

First, outliers are removed from the data. Second, the data of both epochs are merged to compute features of four types: features describing the point distribution, a feature related to height above the terrain, features specific for the multi-target capability of ALS, and features combining both epochs to identify the change. Training data are taken manually, and machine learning is applied to compute a model for the classification. Finally, based on the additional attributes of each point, change types are computed. Each point cloud is classified and investigated for change by an individual machine learning step. The Leopoldstadt District, in Vienna, Austria, is taken as the study area. The experimental region, covering an area of about 3 km × 1.5 km, is generally flat. It contains complex objects, containing a number of old-fashioned and modern high-rise buildings, a suburban area with mainly single-dwellings, an open-wide area (including a stadium), water, overpasses, an amusement park, and a variety of other objects. Two sets of LiDAR data

are available, which were obtained in 2007 (from 07.12.2006 to 05.02.2007) and 2015 (09.11.2015 to 24.11.2015). These data have average point densities of 12 points/m^2 measured with a Riegl LMS_Q560 and 16 points/m^2 measured with a Riegl LMS_Q680i, respectively.

The highlight of our change detection method is the combination of the steps of change detection and change classification in one single step based on the "stability" value combined with other attributes in order to classify all objects into different classes, comprising: unchanged points (ground, building, and tree), and changed points (new building, lost buildings, new tree, lost tree, ground changed in height, ground changed into other objects). The final results are then evaluated using a point-based accuracy evaluation by comparing the automatic change detection classification results with the manual change detection classification. The overall accuracy achieved is 91% and 92% in the 2007 data set and the 2015 dataset, respectively. The traditional two steps change detection is also applied for the dataset. All the point features are rasterized. DSM difference of 1 m is chosen to mask the changed and unchanged regions. Training samples are also rasterized and used to create models for the next step classification. Unchanged and changed objects are classified based on the created models. The finals results gain 78% overall accuracy for both datasets (i.e., 2007 and 2015).

4. DISCUSSION AND CONCLUSION

4.1. Which advantages and disadvantages do raster-based and point-based approaches bring in classification and change detection?

Four publications exploited the aspects related to raster-based and point-based approaches in classification and change detection. Paper I and III focus on raster, which is easier to process in a large area and more (GIS-) tools are available. Here, points were only used to derive features which were stored in raster files. The classification method (simple decision tree) and traditional change detection method (image differencing) were applied effectively based on ALS features. For classification, the same simple decision tree classifier with empirical thresholds was applied for two different regions and achieved a high classification quality (>93%). For change detection, the combination of different models (i.e. DSM (digital surface model) and ER (Echo Ratio)) increased the accuracy of detecting changed in forest area compared to single detectors. However, the final classification and change detection results of the raster-based approach lost the 3D content of the point cloud. Moreover, the results also depend on interpolation and the applied morphological algorithms.

Paper II and IV used point-based approach to overcome the aforementioned limitations of raster-based approaches. The comparison of raster-based and point-based method was also performed in these two articles. In both cases the accuracy using point based classification was higher than with raster data (Table 1).

	II **Classification** **Overall accuracy**	**IV** **Change detection** **Overall accuracy**
Point-based	87%	91%
Raster-based	82%	78%

Table 1. The comparison of classification and change detection results between point-based and raster-based data.

Using the results shown in Table 1, it is concluded that point-based approaches lead to higher quality in both, classification and change detection. Additionally, the following

advantages were identified using the point cloud. Firstly, the point-based approach preserves the full 3D content of the point cloud, in contrast to raster approaches, e.g. vegetation above other objects such as roofs, ground, and water (Höfle et al. 2009, Kraus and Pfeifer 1998), bridges above ground (Sithole and Vosselman 2006), or vertical walls and parcellation borders (Filin et al. 2009). Secondly, the point cloud is the original data. Thus, if we work on the point cloud, there are no interpolation required. Thirdly, the final attributes of classification and change detection results are stored in points, and it is easy to convert those attributes to other data types such as raster, voxel, mesh, or TIN … However, point-based approach required a longer processing time, more storage and therefore a stronger computer configuration than the raster-based approach.

4.2. What efforts and gains does machine learning bring in classification and change detection?

Machine learning based on the Random Forest algorithm is well suited for point cloud classification and change detection because it is efficient and requires a moderate amount of training data. Machine learning is applied in paper II and IV for the purpose of classification and change detection.

In Article II, machine learning is used to classify urban objects, represented in an image matching point cloud. Here, machine learning is compared to the simple decision tree method. The results showed that machine learning brings more advantages compared to a human-based operator approach, being less subjective and having a higher degree of automation. In the same resolution level, the machine learning methods always gain higher accuracy than the simple decision tree. Furthermore, due to the higher number of input parameters and more complex decision tree structures, machine learning was able to detect more classes than the simple decision tree such as water. However, it requires selecting samples for each region (and pyramid level) individually, which takes more time than designing a manual decision tree (approximately twice as much time). The accuracy assessment results proved the ability of machine learning for accurate image matching point cloud classification over urban areas. This aims at providing information of urban classification annually.

In Article IV, the application of machine learning is extended in combining change detection and classification. With the feature set that was required for the categorization of the classes and the correlation between epochs, machine learning based on the samples, created the models which were applied for classifying changed and unchanged classes.

This simplified the two-step process of detecting and classifying changed in previous studies. With the flexibility in feature selection, machine learning allows doing change detection and classification for different data depending on the features given to identify the classes. Numerous features can be used for a machine learning classifier. Depending on the classification tasks, the number of features can be spread out or decreased. Because of using machine learning, taking samples is a required step (Table 2). The results of classification strongly depend on the training samples. Especially for a complex urban area, it is required to consider various types of objects. Thus, to select the samples of each class required careful observation and selection. For each type of point cloud (i.e. ALS or image matching), an operator needs to understand the point distribution of different data to select the samples correctly, this required a first step of understanding data and point distribution of each object type. Thus, in order to obtain a balanced training data set, that equally considers the variety within the class, a trial-and-error process of learning the model from sample data, applying it, evaluating the results (e.g. visually) and refining the sample data is needed. Table 2 shows the time spent for sampling of classification base on image matching point cloud, and change detection based on ALS point cloud. The area of change detection is larger ten times than the area for the classification task, and the number of sample points is six time larger. Because of big study area in change detection, the study area was subset into smaller parts for easier viewing and selecting samples. For that reason, the required time (second/1000 points) is longer. Furthermore, the number of classes in change detection task is also more than in the classification task.

The final results of change detection and classification reach an overall accuracy of 91%. This proved that the integration of change detection and classification is a potential approach for updating information and management in urban regions.

Task	**Area**	**Number of points**	**Time required**	**Seconds/ m^2**	**Seconds/ 1000 points**
Classification	315 m × 140 m	835720	3h	0.2	13
Change detection	3 km ×1.5 km	5251883	25h	0.02	17

Table 2. Time required for training data used for random forest machine learning in different tasks.

The time spent for manual selecting threshold by a human saved half of the spent time compared to the time spent for taking sample in machine learning. However, the overall accuracy of machine learning method always results in a higher accuracy than the simple

decision tree method at the same pyramid level. At lowest pyramid level, (GSD of 6 cm) its accuracy is approximately 87.2% while for the simple manual decision tree, the accuracy is 84.1%. For higher pyramid levels (GSDs of 12 cm and 24 cm), the classification accuracy drops per level by 4% for machine learning and 7% for the simple decision tree.

4.3. How to transfer rules for classification across different modalities (scale, location, point vs raster)?

The studies presented in article I and II investigate the ability of transferring classification rules in raster-based and point-based data. Both studies used a simple decision tree for this investigation. The rules and approaches are transferred between scale, between locations (i.e. space), and between point and raster domain.

Related to transfer among different locations, two simple decision trees are created using empirical threshold selection in the two Articles I and II for the study of transferring thresholds between different missions (Full-waveform ALS data) or within the same mission (image matching data). The result of Article I showed that the method which was used for one dataset could be transferred to another dataset in a different missions. For the two different study areas (Vienna and Eisenstadt) the thresholds applied for Echo Ratio, Sigma0, and nDSM were maintained between regions. Echo width values were shown to depend on the flight mission and processing parameters and were normalized to make it transferable between missions. The tree structure for the classifier is simple and easy to understand. From this study, it can be seen that the simple decision tree raster-based approach can be transferred between different missions and different areas. However, there are many factors that need to be considered, for example seasons (leaf-on and leaf-off seasons), geographical conditions, corresponding objects, sensor acquisition, etc., which influence the threshold selection. This has implication for future research with respect to finding new features which are independent of the sensor characteristics and investigate improved models for normalizing echo width and generally feature values.

In Article II, a simple decision is applied for point-based classification among three different test areas. Different sites required individual thresholds. Using only three attributes, namely, $Z_{\text{Normalised}}$ (height above terrain), GR (Green Ratio) and $O_{\mathbf{T}}$ (Omnivariance), a simple decision tree classification could be performed for two sites (Sites 2 and 3). Site 1 needed an additional attribute, namely, $P_{\mathbf{T}}$ (Planarity), to achieve

satisfactory results. The thresholds for grassland, sealed surface, low tree, and car could be maintained for 3 sites. The thresholds of O_T needed to be varied from site to site because of the differences in tree arrangements. The threshold for distinguishing between trees and buildings depends on the heights of the trees and buildings in the region.

Regarding threshold transfer among different scales and resolutions, Article II investigated a simple decision tree on three pyramid levels (0, 1, and 2). The same thresholds as were used in pyramid level 0 are considered and evaluated in the higher pyramid levels. The threshold of pyramid level 0 can be applied to pyramid level 1 but not pyramid level 2. In pyramid level 2, only the high building threshold delivers satisfactory results; other thresholds needed to be changed. $Z_{Normalised}$ thresholds did not need to be changed, which suggests that this feature depends more strongly on the object type than on the parameters of image acquisition or image matching.

Regarding the transfer between point and raster, the same machine learning method using the same reference data was applied for the image matching point cloud in the highest resolution (pyramid level 0). The main advantage of the 2.5D raster approach compared to point cloud classification lies in its simpler processing because its neighbourhoods are implicitly given and the processing time is shorter because of the faster neighbourhood search. However, the raster is less sharp compared with the point cloud (Figure 3), and the accuracy of raster classification is slightly less than that of the point cloud (approximately 5%).

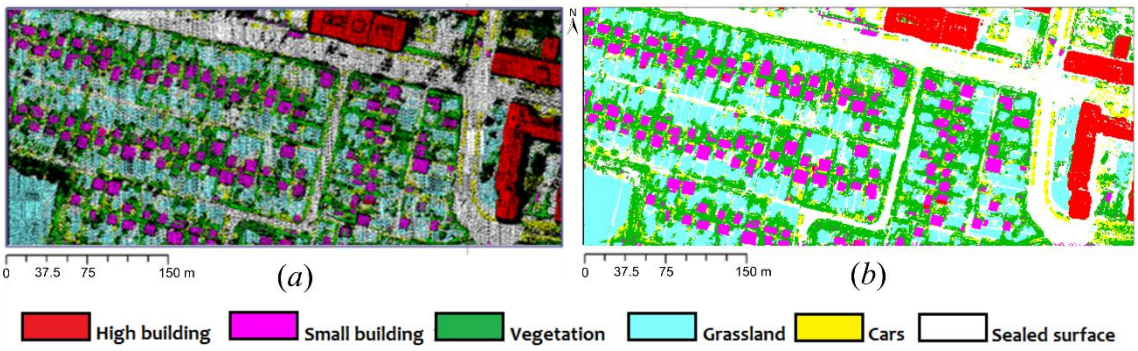

Figure 3. Machine learning classification results: (a) Point-based; (b) Raster-based

4.4. General conclusions

The 3D point cloud, especially derived from ALS and image matching, opens up new research avenues in classification and change detection. These results provide useful 3D information to management and plans for sustainable development. This thesis synthesizes methods (i.e., raster-based and point-based) for classifying and detecting

changes of point clouds from various types of data sources, especially: ALS (discrete-return and full wave-form) and image matching. All three types of data sources provide a 3D point cloud over large areas and are beneficial to study urban and forested regions.

Raster-based and point-based approaches exhibit different advantages and disadvantages in classification and change detection. While the raster-based approach is the traditional method with quick processing and more available software, point-based approach persevere the 3D content of the point cloud at the final results. At the same dataset, the point-based approach gains a higher accuracy compared to the raster-based approach in both classification (approximately 5%) and change detection (approximately 12%).

Machine learning based on Random Forest algorithm is an efficient method with respect to classification of multiple objects in urban areas. In evaluation of the final classification results in different resolutions of point cloud data, machine learning always achieved a higher accuracy than the simple decision tree because of the automatic threshold selection and abundant predefined attributes. Separated training samples are required for different data sets. The accuracy of the samples play an essential role for the final outcomes. The combination of change detection and classification in one step using machine learning is a big step of this study and needs to be investigated more in future to gain a higher accuracy and quicker change detection and classification results.

Some features of full-waveform ALS point cloud can be transfer between regions. The features which are independent of the flight mission parameter, such as Echo Ratio, Sigma0, and nDSM, could be maintained between regions. Other parameters, which depend on the flight parameter, could be normalized (in one example) to transfer the value between missions. Also in the experiment with the image matching point cloud, thresholds in the simple decision tree could be maintained, especially the threshold of height above the terrain ($Z_{\text{Normalised}}$), which is the point cloud correspondance to the nDSM of the raster-world. The GR (Green Ratio, difference of spectral values divided by their sum), which is as the ER (Echo Ratio, number of points in a 3D neighborhood divided by the number of 2D neighbors) a feature that has a normalization built in. The found results are conducted in Austria, where typical European terrain with modern and contemporary architectures in the city and a high share of coniferous forest can be found. The results, therefore, can be applied to regions with the same condition in European countries. For different types of urban (i.e., urban with dense and small houses in Asian countries) and forest (e.g., mangrove forest) the transfer between rules has not been explored, and this

can be investigated in further studies. Concerning machine learning, only random forests were studied, and other classification methods may lead to different accuracy. However, it is expected, that the relative performance of different approaches is not affected strongly by the choice of the learning method.

In this Dissertation, point cloud classification and change detection are investigated. The point cloud has proven as a valuable data source for a three-dimensional map and model of the real world. With the scenario of new powerful sensors and advanced platforms, it offers numerous point cloud sources with additional attributes which can be used for classification and change detection in the future. These meet the demands for the task of storing and updating the information in the context of Earth observation and Photogrammetry. The products of this thesis will be a beneficial for a further application of point clouds in management, planning, visualization and analysis in 3D web browsers.

BIBLIOGRAPHY

Aijazi, A.K., Checchin, P. and Trassoudaine, L., 2013. "Automatic removal of imperfections and change detection for accurate 3D urban cartography by classification and incremental updating." *Remote Sensing* 5(8), 3701-3728.

Altmaier, A. and Kolbe, T.H., 2003. "Applications and solutions for interoperable 3D geo-visualization.", in Fritsch, D (ed.) *Proceedings of the Photogrammetric Week 2003 in Stuttgart*, Wichmann Verla.

Barnhart, D.J., Vladimirova, T. and Sweeting, M.N., 2007. "Very-Small-Satellite Design for Distributed Space Missions." *Journal of Spacecraft and Rockets* 44(6), 1294-1306.

Breiman, L., 2001. "Random forests." *Machine Learning* 45(1), 5-32.

Brodu, N. and Lague, D., 2012. "3D terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology." *ISPRS Journal of Photogrammetry and Remote Sensing* 68(1), 121-134.

Filin, S., Borka, A. and Doytsher, Y., 2009. "Generation of parcelation proposals aided by lidar derived spatial cues." *Computers Environment and Urban Systems* 33(4), 278-284.

Gehrke, S., Morin, K.W. and Downey, M., 2010. "Semi-global matching: an alternative to LiDAR for DSM generation?" In: Proceedings of the "2010 Canadian Geomatics Conference and Symposium of Commission I (vol. 2)". ISPR, Canada, pp. 6, Calgary, AB.

Höfle, B., Mücke, W., Dutter, M., Rutzinger, M. and Dorninger, P., 2009 "Detection of building regions using airborne LiDAR - A new combination of raster and point cloud based GIS methods" In A. Car, G. Griesebner, & J. Strobl (eds). Geospatial crossroads @ GI_Forum '09: proceedings of the geoinformativs forum Salzburg, Geoinformatics on stage, July 7-10, 2009, pp. 66-75.

Kraus, K., Harley, I. and Kyle, S., 2007. "Photogrammetry: Geometry from Images and Laser Scans." Walter de Gruyter. ISBN: 3110892871, 9783110892871.

Kraus, K. and Pfeifer, N., 1998. "Determination of terrain models in wooded areas with airborne laser scanner data." *ISPRS Journal of Photogrammetry & Remote Sensing* 53, 193-203.

Lee, C.N. and Oh, J.H., 2014. "LiDAR chip for automated geo-referencing of high-resolution satellite imagery." *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography* 32(4), 319-326.

Liu, D. and Xia, F., 2010. "Assessing object-based classification: Advantages and limitations." *Remote Sensing Letters* 1(4), 187-194.

Miao, R., Song, J. and Zhu, Y., 2017. "3D geographic scenes visualization based on WebGL." In proceeding of Agro-Geoinformatics, 2017 6th International Conference, 7-10 Aug. 2017, Fairfax, VA, USA. DOI: 10.1109/Agro-Geoinformatics.2017.8046999.

Niemeyer, J., Rottensteiner, F. and Soergel, U., 2014. "Contextual classification of lidar data and building object detection in urban areas." *ISPRS Journal of Photogrammetry and Remote Sensing* 87, 152-165.

Pierrot Deseilligny, M. and Clery, I., 2011. "Apero, an open source bundle adjusment software for automatic calibration and orientation of set of images." In proceeding of ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XXXVIII-5/W16, 2011, pp.269-276.

Popovic, D., Govedarica, M., Jovanovic, D., Radulovic, A. and Simeunovic, V., 2017. "3D Visualization of Urban Area Using Lidar Technology and CityGML." In proceeding of IOP Conference Series: Earth and Environmental Science, vol 95. doi :10.1088/1755-1315/95/4/042006.

Qin, R. and Gruen, A., 2014. "3D change detection at street level using mobile laser scanning point clouds and terrestrial images." *ISPRS Journal of Photogrammetry and Remote Sensing* 90, 23-35.

Qin, R., Tian, J. and Reinartz, P., 2016. "3D change detection – Approaches and applications." *ISPRS Journal of Photogrammetry and Remote Sensing* 122, 41-56.

Remondino, F., Spera, M.G., Nocerino, E., Menna, F. and Nex, F., 2014. "State of the art in high density image matching." *The Photogrammetric Record* 29(146), 144-166.

Samuel, A.L., 1959, "Some Studies in Machine Learning Using the Game of Checkers." *IBM Journal of Research and Development* 3(3), 210-229.

Sithole, G. and Vosselman, G., 2006 "Bridge detection in airborne laser scanner data." *ISPRS Journal of Photogrammetry and Remote Sensing* 61(1), 33-46.

Solimini D., 2016. "Understanding Earth observation." Springer International Publishing. DOI:10.1007/978-3-319-25633-7.

Tao, C.V., 2000. "Semi-automated object measurement using multiple-image matching from mobile mapping image sequences." *Photogrammetric Engineering and Remote Sensing* 66(12), 1477-1485.

Yan, W.Y., Shaker, A. and El-Ashmawy, N., 2015 "Urban land cover classification using airborne LiDAR data: A review." *Remote Sensing of Environment* 158, 295-310.

SCIENTIFIC ARTICLES

ARTICLE I

POTENTIAL OF FULL WAVEFORM AIRBORNE LASER SCANNING DATA FOR URBAN AREA CLASSIFICATION - TRANSFER OF CLASSIFICATION APPROACHES BETWEEN MISSIONS

G. Tran [a,b,*], D. Nguyen [a,c], M. Milenkovic [a], N. Pfeifer [a]

[a] Department of Geodesy and Geoinformation, Vienna University of Technology, Austria
[b] Department of Cartography, Hanoi University of Mining and Geology, Vietnam – tranthihuonggiang@humg.edu.vn
[c] Department of Photogrametry and Remote Sensing, Hanoi University of Mining and Geology, Vietnam

KEY WORDS: LiDAR, Full-waveform, Urban classification, geophysical features, transfer

ABSTRACT:

Full-waveform (FWF) LiDAR (Light Detection and Ranging) systems have their advantage in recording the entire backscattered signal of each emitted laser pulse compared to conventional airborne discrete-return laser scanner systems. The FWF systems can provide point clouds which contain extra attributes like amplitude and echo width, etc. In this study, a FWF data collected in 2010 for Eisenstadt, a city in the eastern part of Austria was used to classify four main classes: buildings, trees, waterbody and ground by employing a decision tree. Point density, echo ratio, echo width, normalised digital surface model and point cloud roughness are the main inputs for classification. The accuracy of the final results, correctness and completeness measures, were assessed by comparison of the classified output to a knowledge-based labelling of the points. Completeness and correctness between 90% and 97% was reached, depending on the class. While such results and methods were presented before, we are investigating additionally the transferability of the classification method (features, thresholds …) to another urban FWF lidar point cloud. Our conclusions are that from the features used, only echo width requires new thresholds. A data-driven adaptation of thresholds is suggested.

1. INTRODUCTION

Airborne LiDAR 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). Generally, two different generations of receiver units exist: discrete echo recording systems, which are able to record multiple echoes on-line and typically sort up to four echoes per laser shot (Lemmens, 2009) and full-waveform (FWF) recording systems capturing the entire time-dependent variation of the received signal power with a defined sampling interval such as 1ns (1 nanosecond) (Mallet and Bretar, 2009; Wagner et al., 2006). With signal processing methods, FWF data provide additional information which offers the opportunity to overcome many drawbacks of classical multi-echo LiDAR data on reflecting characteristics of the objects, which are relevant in urban classification.

Airborne LiDAR data have been used in various applications in urban environments, particularly aiming at mapping and modelling the city landscape in 3D with its artificial land cover types such as buildings, power lines, bridges, roads. Moreover, as urban environments are active regions with respect to alteration in land cover, urban classification plays an important role in update changed information (Matikainen et al., 2010). If FWF data is available, amplitude, echo width, and the integral of the received signal are additional information. Furthermore, a higher number of detected echoes has been reported for FWF data in comparison to discrete return point clouds. These additional attributes were successfully used in classification (Alexander et al., 2010). The classification methods applied reach from simple decision trees to support vector machines (SVM). (Ducic et al., 2006) applied a decision tree based on amplitude, pulse width, and the number of pulses attributes of full-waveform data in order to distinguish the vegetation points and non-vegetation points. (Rutzinger et al., 2008) used a decision tree based on the homogeneity of echo width to classify points from full-waveform ALS data to detect tall vegetation - trees and shrubs. (Mallet et al., 2008) used SVM to classify four main classes in urban area (e.g. buildings, vegetation, artificial ground, and natural ground). In these studies the parameters of the classification (threshold values, etc.) are set by expert knowledge or learned from training data. Thus, these values are optimal for the investigated data set.

The transferability of classification approaches between different full waveform LiDAR data sets has received less attention so far (Lin, 2015). The aim of this paper is therefore to:

- demonstrate that high classification accuracy can be reached with decision trees, and to
- study, if this classification approach using the selected features and the thresholds can be transferred to another data set, and finally to
- suggest a method to re-compute the echo width threshold for different missions acquiring urban full waveform point clouds.

In this study, the following attributes are used.

- echo width: full waveform attribute, describing variation of the target along the ranging direction,
- Sigma0: local smoothness,

* Corresponding author.

- echo ratio: a measure of surface penetration, and
- nDSM: normalized digital surface model, height above ground.

Four main classes are derived for the built up areas of Eisenstadt and Vienna: buildings, vegetation, water body, and ground. They are classified based on decision tree method using OPALS (Pfeifer et al., 2014).

To quantify the transferability of parameter between the different regions/data sets, the parameters are applied for the Eisenstadt set and then applied to the Vienna set. This indicates which parameter is stable for various study areas and which need conversion.

2. STUDY AREA AND DATA USED

2.1 Study area

Eisenstadt is a town in the south eastern part of Austria. It is characterized by buildings of medium size. The centre of Eisenstadt was selected for the analysis located on lat. N 47°50'51", long. E 16°31'5".

Vienna is the capital of Austria and characterized by old large buildings in the centre, but also open park areas and trees along a boulevard. The center of study area located on lat. N 48°12'26", long. E 16°21'52".

2.2 Data

The full-waveform airborne LiDAR data were available for the two mentioned cities. Eisenstadt area was scanned with a Riegl LMS-Q560 sensor in April 2010. The resulting point density was approximately 8 points/m^2 in the non-overlapping areas, while the laser-beam footprint was not larger than 60 cm in diameter. The Vienna city-center area was scanned with the same model of the scanner, in January 2007. The resulting point density was 12 points/m^2 in the non-overlapping area, and the laser-footprint was not larger than 30 cm. The investigated area covers 2.5 km^2 for Eisenstadt and 1.4 km^2 for Vienna.

Both raw full-waveform data sets were processed in the same way using the software OPALS and sensor manufacturer software. First, Gaussian decomposition (Wagner et al., 2006) was applied to extract geometrical (range) and full-waveform (amplitude and echo with) attributes per echo. No additional information on how echo width was specified (FWHM, std.dev.) was available for this research. Then, considering additionally the trajectory information (GPS and INS information), direct georeferencing was performed for each strip. The output of this procedure was strip-wise georeferenced point clouds, stored in the OPALS datamanager (ODM) format and projected in ETRS89/UTM zone 33N. Each ODM file includes point attributes: X-, Y-and Z-coordinate, Echo Number, Number of Echoes, Amplitude, Echo Width, and strip identifier as the primarily acquired ("measured") attributes of each echo. The ODM does allow storage of freely defined attributes at each point and provides spatial access, e.g. used in neighborhood queries for computing additional point attributes (see below).

Additionally to the LiDAR data, RGB Orthophotos - projected in the same coordinate system - were used for visually interpretation.

3. METHODOLOGY

First, a number of attributes is computed for each point, using the paradigm of point cloud processing (Otepka et al., 2013). From these attributes different images are computed ("gridding") at a pixel size of 1m. A terrain model is derived also. Then, a decision tree is applied to classify each pixel into one of the four classes: building, vegetation, ground, and water body. Image algebra (e.g., morphological operations) is used in between to refine the results. The quality of the results is assessed using the completeness and the correctness measure.

Mallet et al. (2008) showed that for urban area classification from Lidar data a combination of attributes should be used to obtain classification results of high quality. In their analysis of feature (attribute) importance, it was demonstrated that attributes considering the local dispersion of the point cloud, attributes describing geometric properties, and the echo width of FWF Lidar should be used together. This was used in the selection of attributes for the present study.

3.1 DTM creation

The Digital Terrain Model (DTM) give important geometric information about objects in urban area, e.g. object heights, and thus, they were directly derived from the LiDAR data. To calculate the DTM, first the LiDAR ground points were selected by applying the robust filtering algorithm (Kraus and Pfeifer, 1997; Pfeifer and Mandlburger, 2008) implemented in the software SCOP++. Then, the DTM was interpolated from the selected ground points using the moving plane interpolation implemented in OPALS.

3.2 Attributes for the classification

Prior to attribute computation in each point, the LiDAR point clouds are checked in order to remove erroneous points which influenced to the accuracy of further processing steps. The relative height of each point above the DTM, nH = z (point)-z(DTM), was computed. All points with nH below -1m and above > 40m are removed. For the Vienna data set the highest buildings are approx. 100m, but also no erroneously high points were found in the data. Thus only the lower threshold was applied for Vienna.

The value nH defines the attribute nDSM, i.e. normalized surface model (object height). The nDSM represents, as written above, the height of points above the terrain. In the classification it is used to distinguish all the point above the terrain such as buildings and vegetation from the ground points.

To distinguish buildings and vegetation points the Echo Ratio (Höfle et al., 2009; Rutzinger et al., 2008) is used. The echo ratio (ER) is a measure for local transparency and roughness and is calculated in the 3D point cloud. The ER is derived for each laser point and is defined as follows:

$$Echo\ Ratio\ ER[\%] = n_{3D} / n_{2D} * 100.0 \quad (1)$$

n_{3D} = Number of points within distance measured in 3D (sphere).

n_{2D} = Number of points within distance measured in 2D (unbounded vertical cylinder).

In building and ground, the ER value reach a high number (approximately 100%), but for vegetation and permeable object

ER < 100%. ER is created by using *OpalsEchoRatio* module, with search radius is 1m, slope-adaptive mode. For the further analyses the slope-adaptive ER is aggregated in 1m cells using the mean value within each cell.

The attribute Sigma0 is the plane fitting accuracy (std.dev. of residuals) for the orthogonal regression plane in the 3D neighborhood (ten nearest neighbors) of each point. It is measured in meter. Not only the roofs, but also the points on a vertical wall are in flat neighborhoods. Echo Ratio and Sigma0 both represent the dispersion measures. Concerning their value they are inverse to each other (vegetation: low ER, high Sigma0). What is more, Sigma0 is only considering a spherical neighborhood and looks for smooth surfaces, which may also be oriented vertically. The ER, on the other hand, considers (approximately) the measurement direction of the laser rays (vertical cylinder). Those two attributes play an importance role in discriminate trees and buildings. Using *OpalsGrid* module with moving least square interpolation the $Sigma_0$ image with the grid size of 1m was created.

The Echo Width (EW) represents the range distribution of all individual scatterers contributing to one echo. The width information of the echo pulse provides information on the surface roughness, the slope of the target (especially for large footprints), or the depth of a volumetric target. Therefore, the echo width is narrow in open terrain areas and increases for echoes backscattered from rough surfaces (e.g. canopy, bushes, and grasses). Terrain points are typically characterized by small echo width and off-terrain points by higher ones. The echo width also increases with increasing width of the emitted pulse. It is measured in nano seconds. OpalsCell module is used to create the EW image with the final gird size of 1m.

The local density of echoes can be used for detecting water surfaces. As demonstrated by (Vetter et al., 2009) water areas typically feature areas void of detected echoes or very sparse returns. It is measured in points per square meter. Density was also computed for 1m cells.

The attributes used for classification are thus: nDSM, Echo Ratio, Sigma0, Echo Width, and Density.

3.3 Object classification

First each pixel is classified using the decision tree shown in Fig. 1 including the threshold values. After the first 2 classes, water and building (candidates) are extracted, mathematical morphology is applied to refine the building results. The pixels not classified are then tested for fulfilling the vegetation criteria. If they are not in vegetation, they are considered to be ground.

Water is first identified, based on the low point density. As mentioned above, water has very low backscatter, and often no detected echo.

Building objects are distinguished from other objects by height (above 3m) and surface roughness. ER is used to distinguish buildings from tree objects. However, with various shapes of building roof and some buildings being covered by high trees, only ER is not sufficient and would include vegetation in the building class. Thus, EW is used to detect only hard surfaces. Buildings are contiguous objects and have typically a minimum size. This is considered by analysing all the pixels classified as buildings so far with mathematical morphology. A closing operation is applied first to fill up all small holes inside the buildings, and then opening is performed to remove few pixel detections ("noise") from the building set. This also makes the outlines of buildings smoother.

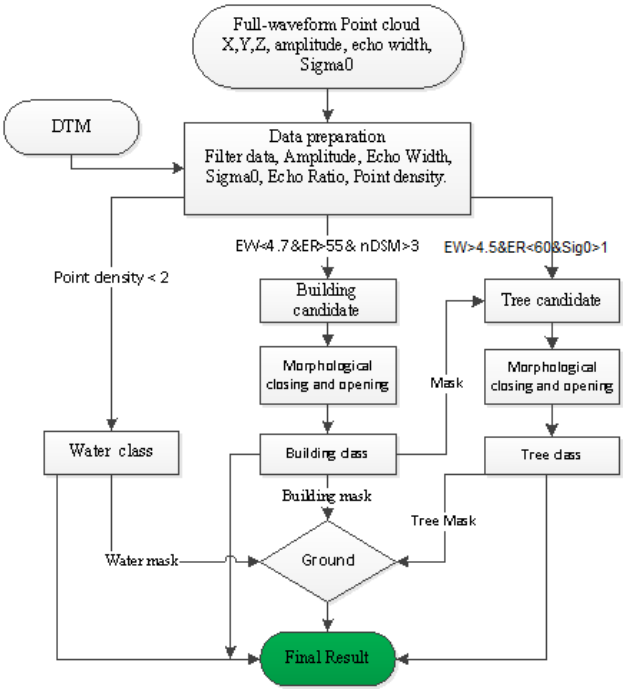

Figure 1. Decision tree for the classification.

ER, Sigma0 and EW are then used to classify trees. The building mask is applied to classify only pixels not classified before. Also this result is refined with image morphological operations. Finally, all pixels not classified so far are considered ground.

3.4 Echo Width normalisation

An initial assumption was that the thresholds for the decision tree derived for one data set can also be used for the other data set. The rational was that:

- Density is a physical measure (points per square meter) and the overall shot density was similar (8 vs. 10 points per square meter).
- Height above ground (nDSM) is a measure independent of the measurement device and also independent of the sampling distance.
- Echo Ratio is by definition a relative measure and should therefore adapt itself to the data distribution.
- Sigma0 is the local plane fitting accuracy. For data sets of similar measurement accuracy (same sensor model used for both areas) and similar neighbourhoods, both number of neighbours and spatial extent, it should deliver comparable values.
- Echo width obviously depends on the width of the emitted pulse (same sensor model used for both data sets), but may also depend on the footprint diameter (which was different in the two data sets investigated) or other effects.

Due to the doubts of echo width transferability, a method to normalize echo width is suggested. Weak, low amplitude echoes typically lead to a poor determination of echo width. Thus only stronger echoes (larger amplitude) are used for deriving the echo width normalization parameters.

Assuming that each data set contains some bright, flat surfaces (orthogonal to the incident Lidar signal), a minimum echo width, EW_{min}, was chosen based on single echoes (i.e. extended targets) of high amplitude and narrow width. A maximum echo width, EW_{max}, was chosen based on the assumption that in each data set tree crowns can be found. Those cause large echo width. Thus, strong, first-of-many echoes with a large width were chosen for a maximum echo width. One way to find specific values of EW_{min} and EW_{max} is to use quantiles of the distribution of echo width and amplitude. Using quantiles is suggested because of their robust stochastic properties.

The normalized value of EW for the two datasets can then be computed using:

$$NorEW = \frac{EW - EW_{min}}{EW_{max} - EW_{min}} \quad (2)$$

It is noted that this can lead to negative normalized EW, which may be left as they are or set to zero. Also values larger than 1 can appear, e.g. for very wide echoes not considered in the normalization due to low amplitude.

A different method to normalize EW value is proposed by (Lin, 2015) which used concept of Fuzzy Small membership.

4. RESULT AND DISCUSSION

4.1 Classification results

The thresholds for the classification were set manually, based on exploratory analysis of the data sets and on expectation of the objects. This was done for both data sets independently.

The main properties of ER, EW, Sigma0, nDSM, and Density values for both Eisenstadt and Vienna are summed up in Table 1. From that properties and combining with empirical selection, the threshold for each parameter was set in the Table 2.

	ER [%]	EW [ns]	Sigma0 [m]	nDSM [m]	Density [pt/m^2]
Eisenstadt	4.2-100	0-29	0-19	-1.52 – 39	0-72.4
Vienna	2.5-100	.,003-66	0-3474	-1 – 884	0-63

Table 1. The range of ER, EW, Sigma0, nDSM, Density for Eisenstadt and Vienna.

	Building			Tree			Water
	ER	EW	nDSM	ER	EW	Sig.0	Density
Eisenstadt	>55	<4.7	>3	<60	>4.5	>1	<2
Vienna	>55	<9.8	>2	<60	>9.6	>1	<2

Table 2. The threshold values using for decision tree classification of buildings, trees and water body region for Eisenstadt and Vienna.

The results were evaluated quantitatively and qualitatively. Based on the point density characteristic of water region it produces a good result. All the water bodies in the interested area are classified. However, some small parts of the study area where the laser signal could not reach the ground because of occlusion by high buildings, are misclassified. This could possibly be improved with the overlap of another strip.

While buildings in general can be classified well, very complex roof shapes and walls cause difficulties. It was observed that selecting threshold conservatively the shape of the building is maintained, while its size is reduced slightly.

The tree class includes high trees but also lower vegetation (bushes, etc.), also at heights below 3m. Especially for the latter category EW proofed helpful in distinguishing between vegetation and building edges and also in identifying single trees. For very tall trees, Sigma0 and ER allow reliable detection.

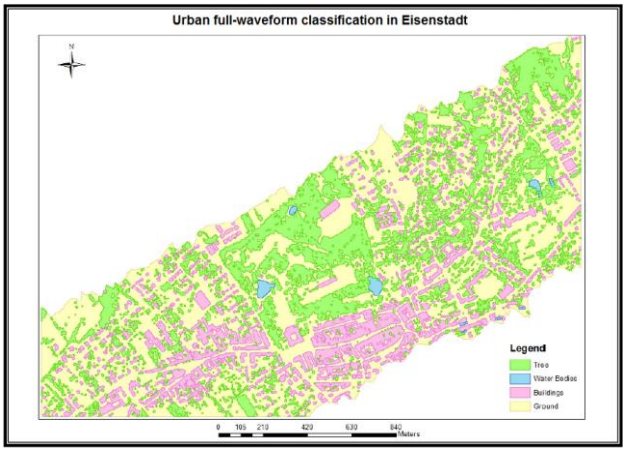

Figure 2. Urban full-waveform classification in Eisenstadt.

Figure 3. Urban full-waveform classification in Vienna.

Ground includes all objects such as: roads, grass land, car park, fields… A further split into artificial and natural ground was explored but finally not performed. Both Sigma0 and Amplitude were considered candidates for this separation. Natural ground tends to have higher Amplitude than artificial ground. However, while valid locally, no global thresholds could be found in the data sets studied.

The final classification results were then assessed based on Correctness and Completeness (Heipke et al., 1997). Some buildings, trees and water bodies are digitized manually as reference data. Comparing the results of the automated extraction to reference data, an entity classified as an object that also corresponds to an object in the reference is classified as a True Positive (TP). A False Negative (FN) is an entity corresponding to an object in the reference that is classified as background, and a False Positive (FP) is an entity classified as an object that does not correspond to an object in the reference. A True Negative (TN) is an entity belonging to the background both in the classification and in the reference data.

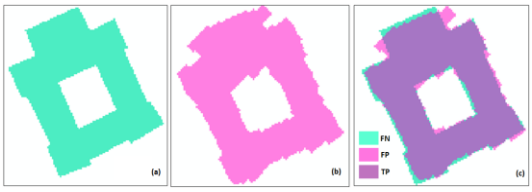

Figure 4. (a) Ground truth data; (b) classified result; (c) accuracy assessment

The Completeness and Correction for building, tree, and water class are given in Table 3. It is also illustrated for one building in Figure 4. The two main classes of building and tree feature values above 93%.

	Comp	Corr	Quality
Building	97.3%	96.0%	93.7%
Tree	97.8%	93.9%	92.0%
Water	89.0%	90.7%	81.6%

Table 3. Accuracy assessments of Building, Tree and Water classes in Eisenstadt region.

4.2 Echo width normalisation

After estimate the threshold values, a comparison of the used thresholds for both regions is carried out to find which parameters keep stable through different dataset and which required to be normalised. As can be seen in the table 2, the threshold of ER, nDSM, Sigma0 and Density can be applied for both Eisenstadt and Vienna. In other words, those values can be transferable between different regions. However, the EW threshold is notably different. Thus, the normalization suggested in Sec. 3.4 was applied to evaluate its usability.

The Figure 5 and Figure 8 show the distribution of EW for the two regions. The ranges of EW are unexpected wide, from 4.003ns to 66.877ns for Vienna, and 0 to 29.000ns for Eisenstadt, given the emitted pulse width of approx. 4ns. However, more than 96% of EW values fall in a more narrow range, from approx. 7ns to 18ns for Vienna, and from approx. 3ns to 10ns for Eisenstadt. This demonstrates the reason for normalization.

As suggested in Sec. 3.4, the minimum EW, EW_{min}, value is the 5% quantile of single, strong echoes. Strong echoes are those that have amplitude more than 1% of the highest amplitude found in the data set. Thus, only 5% of all "strong" echoes have a shorter EW than this EW_{min}. (See Figures 6 and 9). The maximum EW is chosen as the 99% quantile of EW from the strongest third of the first-of-many echoes with the highest amplitude (Figures 7 and 10). The thresholds for normalizing EW for Eisenstadt and Vienna are summed up in Table 4:

	EW			
	Min	Max	Mean	Std
Eisenstadt	4.4000	9.1000	4.632	0.295
Vienna	8.9480	16.7190	9.538	0.404

Table 4. The EW thresholds for Eisenstadt and Vienna

Applying the normalization, the threshold for normalized EW in Eisenstadt and Vienna are presented in Table 5 While the normalization brings those values closer together (buildings have normalized EW below 6% and 11% respectively, and trees have normalized EW above 3% and 9% respectively), they are not as close together as for the other thresholds (Table 2).

	Building	Tree
	EW	EW
Eisenstadt	<0.064	>0.027
Vienna	<0.106	>0.085

Table 5. Normalized EW thresholds for Eisenstadt and Vienna.

	Comp	Corr	Quality
Building	97.3%	96.9%	94.4%
Tree	93.5%	93.3%	87.6%

Table 6. Accuracy assessments of Building, Tree classes in Eisenstadt region after Echo Width normalisation.

Another way of evaluating the normalization is to apply the thresholds on normalized echo width from one dataset for classifying the other data set. The Eisenstadt values applied to Vienna led to a result of lower quality, but the Vienna thresholds applied to Eisenstadt did produce qualitatively a very similar classification. The completeness and correctness measures are shown in Table 6 . The loss in completeness and correctness does not change for the building class and is below 5% for the tree class.

5. CONCLUSION

This study used full-waveform LiDAR data to classify urban areas, i.e. Eisenstadt and Vienna. Four classes were built: Water bodies, Buildings, Trees and Ground. The computations were executed in OPALS, ArcGIS and FugroViewer. Overall, a high accuracy (>93%) could be achieved.

Full-waveform LiDAR with its additional attributes is an advanced data to classify urban area. The echo width proofed valuable in classifying vegetation and buildings reliably. The other attributes used were Echo Ratio, Sigma0, nDSM, and Density.

Applying the classification thresholds, i.e. those with normalized Echo Width, derived for the Vienna dataset to the Eisenstadt dataset demonstrated that thresholds are, indeed, transferable between missions, resulting in a minor loss of accuracy (5%) in comparison to the classification tailored for the Eisenstadt mission.

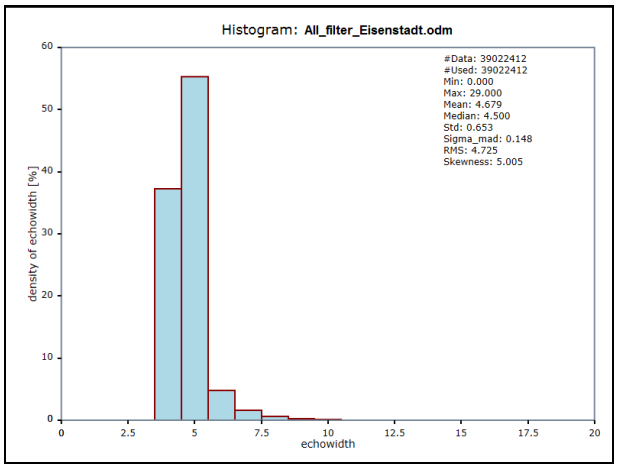

Figure 5. Histogram of echo width value in Eisenstadt region

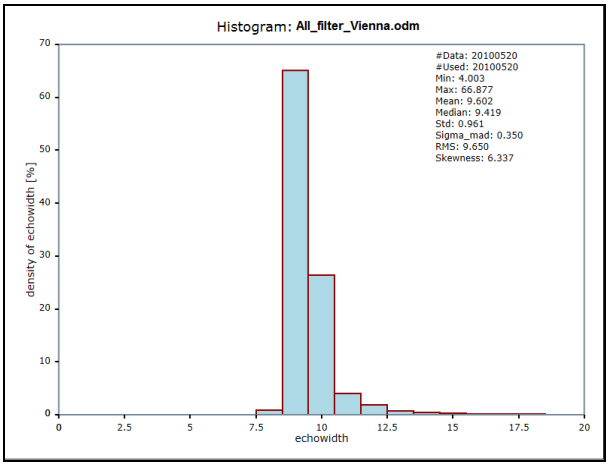

Figure 8. Histogram of echo width value in Vienna region

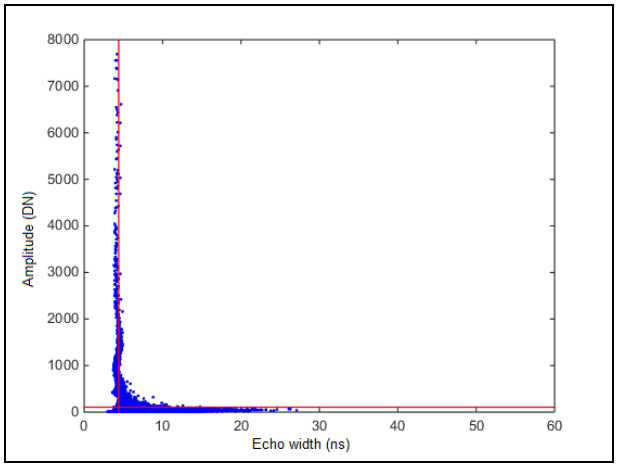

Figure 6. Eisenstadt dataset: Scatterplot of Echo Width versus Amplitude for single echoes. Echoes with more than 1% of the highest Amplitude are placed above the red horizontal line. The vertical line shows the minimum echo width EW_{min} which is 5% quantile of the single echoes above the Amplitude threshold.

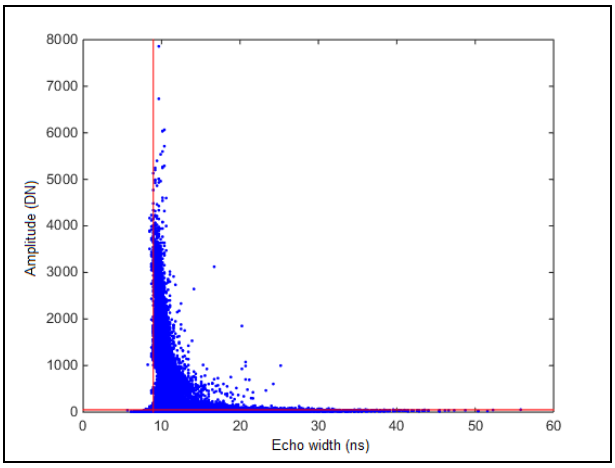

Figure 9. Vienna dataset: Scatterplot of Echo Width versus Amplitude for single echoes. Echoes with more than 1% of the highest Amplitude are placed above the red horizontal line. The vertical line shows the minimum echo width EW_{min} which is 5% quantile of the single echoes above the Amplitude threshold.

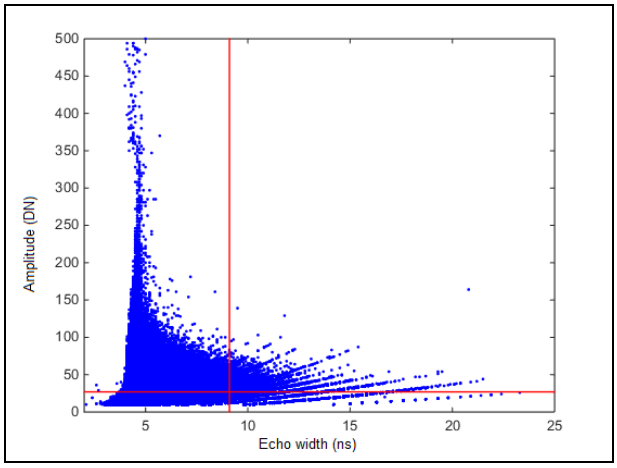

Figure 7. Eisenstadt dataset, Scatterplot of Echo Width versus Amplitude for the first-of-many echoes. The red horizontal line separates the weak (66.6%) from the strong echoes ("highest third"). The vertical line shows the maximum echo width EW_{max} which is the EW at the 99% quantile of the strong echoes.

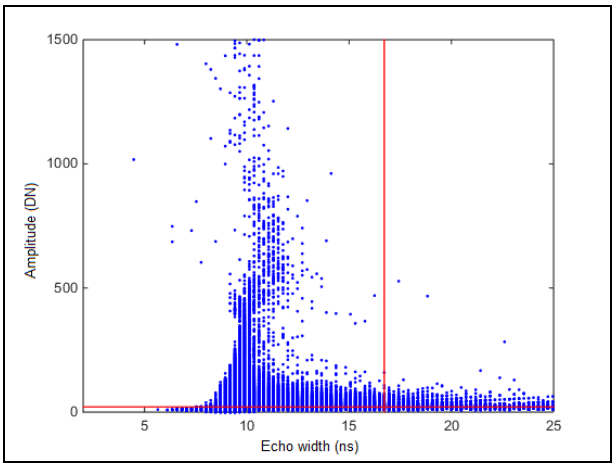

Figure 10. Vienna dataset: Scatterplot of Echo Width versus Amplitude for the first-of-many echoes. The red horizontal line separates the weak (66.6%) from the strong echoes ("highest third"). The vertical line shows the maximum echo width EW_{max} which is the EW at the 99% quantile of the strong echoes.

For the two different study areas the threshold applied for Echo Ratio, Sigma0, and nDSM were the same. Echo Width was shown to depend on the flight mission parameters. The cause was not studied, but the footprint size may have influence. It is noted that using the differential cross section, instead of echo width would not necessarily change this. The differential cross section is obtained by deconvolving (Jutzi and Stilla, 2006; Roncat et al., 2011) the received signal with the emitted pulse shape (more precisely the system waveform).

A simple model for normalizing echo width was suggested. Improvements of this model, e.g., choice of minimum and maximum echo width for normalization, could be investigated, e.g. histogram matching.

ACKNOWLEDGEMENTS

This work is partly founded from the European Community's Seventh Framework Programme (FP7/2007–2013) under grant agreement No. 606971, the Advanced_SAR project.

The authors would also like to thank VIED (Vietnam International Education Development) and OeAD (Österreichischer Austauschdienst) scholarships for their support in studying period.

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, 423-432.

Ducic, V., Hollaus, M., Ullrich, A., Wagner, W., Melzer, T., 2006. 3D Vegetation mapping and classification using full-waveform laser scanning, in: Forestry., P.o.t.W.o.D.R.S.i. (Ed.), EARSeL/ISPRS, Vienna, Austria, pp. 211–217.

Heipke, C., Mayer, H., Wiedemann, C., 1997. Evaluation of automatic road extraction, IAPRS "3D Recontruction and Modeling of Topographic Objects", Stuttgart, pp. 151-160.

Höfle, B., Mücke, W., Dutter, M., Rutzinger, M., 2009. Detection of building regions using airborne lidar - A new combination of raster and point cloud based GIS methods study area and datasets, International Conference on Applied 395 Geoinformatics. Proceedings of GI-Forum 2009, Salzburg, Austria, pp. 66-75.

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

Kraus, K., Pfeifer, N., 1997. A new method for surface reconstruction from laser scanner data, ISPRS Workshop, Haifa, Israel.

Lemmens, M., 2009. Airborne Lidar Sensors. GIM International 2, 16-19.

Lin, Y.-C., 2015. Normalization of Echo Features Derived from Full-Waveform Airborne Laser Scanning Data. Remote Sensing 7, 2731-2751.

Mallet, C., Bretar, F., 2009. Full-waveform topographic lidar: State-of-the-art. ISPRS Journal of Photogrammetry and Remote Sensing 64, 1-16.

Mallet, C., Soergel, U., Bretar, F., 2008. Analysis of Full-waveform LiDAR data for classification of urban areas, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Beijing, pp. 85-92.

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, 1217-1248.

Otepka, J., Ghuffar, S., Waldhauser, C., Hochreiter, R., Pfeifer, N., 2013. Georeferenced Point Clouds: A Survey of Features and Point Cloud Management. ISPRS International Journal of Geo-Information 2, 1038--1065.

Pfeifer, N., Mandlburger, G., 2008. Filtering and DTM Generation, in: Shan, J., Toth, C. (Eds.), Topographic Laser Ranging and Scanning: Principles and Processing. CRC Press, pp. 307--333.

Pfeifer, N., Mandlburger, G., Otepka, J., Karel, W., 2014. OPALS – A framework for Airborne Laser Scanning data analysis. Computers, Environment and Urban Systems 45, 125-136.

Roncat, A., Bergauer, G., Pfeifer, N., 2011. B-spline deconvolution for differential target cross-section determination in full-waveform laser scanning data. ISPRS Journal of Photogrammetry and Remote Sensing 66, 418-428.

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, 4505-4528.

Vetter, M., Hö fle, B., Rutzinger, M., 2009. Water classification using 3D airborne laser scanning point cloud. Österreichische Zeischrift für Vermessung & Geoinformation 97(2), 227-238.

Vosselman, G., Maas, H.-G., 2010. Airborne and terrestrial laser scanning. Whittles Publishing, Dunbeath.

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, 100-112.

ARTICLE II

International Journal of Remote Sensing

ISSN: 0143-1161 (Print) 1366-5901 (Online) Journal homepage: http://www.tandfonline.com/loi/tres20

Classification of image matching point clouds over an urban area

Thi Huong Giang Tran, Johannes Otepka, Di Wang & Norbert Pfeifer

To cite this article: Thi Huong Giang Tran, Johannes Otepka, Di Wang & Norbert Pfeifer (2018) Classification of image matching point clouds over an urban area, International Journal of Remote Sensing, 39:12, 4145-4169, DOI: 10.1080/01431161.2018.1452069

To link to this article: https://doi.org/10.1080/01431161.2018.1452069

Published online: 23 Mar 2018.

INTERNATIONAL JOURNAL OF REMOTE SENSING, 2018
VOL. 39, NO. 12, 4145–4169
https://doi.org/10.1080/01431161.2018.1452069

Taylor & Francis
Taylor & Francis Group

Classification of image matching point clouds over an urban area

Thi Huong Giang Tran[a,b], Johannes Otepka[a], Di Wang[a] and Norbert Pfeifer[a]

[a]Department of Geodesy and Geo information, TU Wien, Vienna, Austria; [b]Department of Cartography, Hanoi University of Mining and Geology, Hanoi, Vietnam

ABSTRACT

Airborne laser scanning (ALS) and image matching are the two main techniques for generating point clouds for large areas. While the classification of ALS point clouds has been well investigated, there are few studies that are related to image matching point clouds. In this study, point clouds of multiple resolutions from high-resolution aerial images (ground sampling distance, GSD, of 6 cm) over the city of Vienna were generated and investigated with respect to point density and processing time. Three different study sites with various urban structures are selected from a bigger dataset and classified based on two different approaches: machine learning and a traditional operator-based decision tree. Classification accuracy was evaluated and compared with confusion matrices. In general, the machine learning method results in a higher overall accuracy compared to the simple decision tree method, with accuracies of 87% and 84%, respectively, at the highest resolution. At lower-resolution levels (GSDs of 12 cm and 24 cm), the overall accuracy of machine learning drops by 4% and that of the simple decision tree by 7% for each level. Classifying rasterized data instead of the original point cloud resulted in an accuracy drop of 5%. Thus, using machine learning on point clouds at the highest available resolution is suggested for classification of urban areas.

ARTICLE HISTORY
Received 19 May 2017
Accepted 26 February 2018

1. Introduction

Distribution of land cover has an intense effect on the global climate and environment (Pielke 2005). Different scales, from global to regional to local land cover mapping, play an important role in providing effective monitoring of the changing environment (Mahmood et al. 2010; Yan, Shaker, and El-Ashmawy 2015). Satellite remote sensing has been utilized as a functional tool in mapping the Earth's relief and land cover over large areas. Improving the spatial resolution of mapping products, which are required for diverse large-scale urban tasks, e.g. traffic management (Youn et al. 2008), is one of the current challenges in remote sensing investigations (Momm, Easson, and Kuszmaul 2009; Myint and Lam 2005; Yan, Shaker, and El-Ashmawy 2015). Thanks to advances in technology, point clouds became a standard product (Lemmens 2014), and their 3D

CONTACT Thi Huong Giang Tran tranthihuonggiang@humg.edu.vn Department of Geodesy and Geoinformation, Vienna University of Technology, Gusshausstrasse 27-29, 1040, Vienna, Austria

content makes them interesting for studying complex scenes such as urban areas. Point cloud classification is a significant step in information extraction (Vosselman 2013).

'Point cloud' is a term that describes a set of unorganized points, which are often georeferenced in an Earth-fixed coordinate system (Otepka et al. 2013). In addition to three coordinates, (*x, y, z*), each point may have additional attributes (e.g. normal vector components and 'colour' in different spectral bands). Thus, point clouds can be used not only to visualize a scene, e.g. Nebiker, Bleisch, and Christen (2010), but also to infer quantitative information (Otepka et al. 2013). Point cloud data became important for urban studies such as land cover classification (Antonarakis, Richards, and Brasington 2008; Yan, Shaker, and El-Ashmawy 2015), digital terrain modelling, building modelling (Dorninger and Pfeifer 2008; Haala and Kada 2010; Rottensteiner et al. 2007), canopy detection and vegetation analysis (Antonarakis, Richards, and Brasington 2008; Rutzinger et al. 2008; Van Leeuwen, Coops, and Wulder 2010; Wang, Weinacker, and Koch 2008), building damage and change detection (Kang and Lu 2011; Pang et al. 2014), and in the broader applications of land cover change detection or monitoring, urban structure type mapping (Heiden et al. 2012), urban planning, and disaster management (Biasion et al., 2005). A special advantage of classified point clouds is that they allow the relevant points to be chosen for different modelling tasks. One example is the digital surface model (DSM) for visibility analysis, in which points on wires or mobile objects (cars, etc.) are disregarded. This method is applied, e.g. by the city of Vienna.

There are two sources for point clouds over large urban areas: airborne laser scanning (ALS) and dense matching of aerial images. Wang et al. (2009) expressed that LiDAR (Light Detection and Ranging) technology is 'rapidly replacing the photogrammetric approach'. However, in practical experiments, Veneziano, Hallmark, and Souleyrette (2002) and Choma, Ratcliff, and Frisina (2005) were less impressed by LiDAR and perceived a continuing role for stereo photogrammetry. With the advantages of high point density, 'rich' semantic information (Leberl et al. 2010) and direct observation of linear features (Rau, Jhan, and Hsu 2015), in addition to developments within computer vision, especially dense multi-image matching, point clouds from photogrammetric products are increasingly utilized in science and 'reality of human life' (Maltezos and Ioannidis 2015). The comparative evaluation of point clouds from image matching and ALS reveals subtle differences and indicates a higher reliability of ALS points clouds (Ressl et al. 2016). Little research on using point clouds from high-density image matching for classification of urban scenes has been conducted. Because of the regular availability of aerial images (e.g. yearly photo flights), classified point clouds that are derived from image matching would allow frequent updates of derived products. Therefore, it is interesting to study the feasibility of classifying 3D point clouds from dense image matching and assess the accuracy that can be obtained.

Our aim is to study the classification of point clouds from high-density aerial image matching. Often, image data are acquired at high resolution, e.g. a 10-cm ground sampling distance (GSD), and used for interactive interpretation and visualization purposes. However, classification may not be required at this high level of detail. Thus, it will be investigated whether a lower-resolution point cloud is also appropriate for classification. This would lead to a reduction in storage and processing time, especially if the classification accuracy were independent of the resolution that was used for the

generation of the 3D point cloud. The investigated raw point clouds contain information on coordinates and colour (RGB).

For urban scenes, a special advantage of point clouds is that the 3D content is fully preserved, in contrast to raster approaches (2.5D), e.g. vegetation above other objects such as roofs, ground, and water (Höfle et al. 2009; Kraus and Pfeifer 1998), bridges above ground (Sithole and Vosselman 2006), or vertical walls and parcellation borders (Filin, Borka, and Doytsher 2009). The question arises as to whether 3D point cloud classification has an advantage over a 2.5D (raster) approach with respect to accuracy or if the benefit of the 3D content is counterweighted by a loss of classification accuracy.

Additionally, decision trees are one of the most popular learning methods and are commonly used for data exploration (Rokach and Maimon 2008). In this study, we use a random forest classifier and a traditional operator-based decision tree for complex urban areas to determine which advantages they bring. The former sets thresholds automatically but requires training samples to be acquired at each site. The latter is subjective, as thresholds are set experimentally by the operator. These decision trees are typically simple and, therefore, more transparent. Thus, it is expected that they allow easier transfer of the classification model (i.e. a decision tree) from one site to another and from one resolution level to another. As there was uncertainty as to whether the manual setting of thresholds would lead to results of poor quality, an automatic determination of thresholds for a decision tree was also investigated.

In summary, the main contribution of this article lies in providing information on the classification accuracy that is obtained from image matching point clouds, how it is affected by a change in resolution, whether the 3D approach allows higher accuracy, and which advantages machine learning and operator-based decision tree approaches have relative to each other on complex urban scenes.

2. Related works

ALS can produce highly accurate, reliable, and dense 3D point clouds of ground and objects. It is particularly beneficial for vegetation areas where the laser pulse can partially pass through gaps in the foliage, touches the ground and reflects topographic information. However, using only ALS data for urban detection is challenging for extracting more than four urban object classes according to Rau, Jhan, and Hsu (2015). Lafarge and Mallet (2012) proposed a robust city model reconstruction approach with a 2.5D hybrid representation that includes only buildings, trees, and ground. Mallet et al. (2011) used a support vector machine (SVM) to classify the urban area into three classes: building, vegetation, and ground. Shapovalov, Velizhev, and Barinova (2010) applied non-associative Markov networks to classify five classes. Xiong et al. (2011) used the same aerial dataset from Shapovalov et al. to distinguish five urban objects (ground, buildings, tree, low vegetation, and car), thereby showing how point-based and region-based classifications of LiDAR data can interact in a pairwise conditional random field (CRF). Niemeyer, Rottensteiner, and Soergel (2013) proposed a supervised classification method that incorporated a random forest classifier into a CRF framework to detect asphalt, natural ground, vegetation, buildings, fences, and cars. Najafi et al. (2014) presented a non-associative higher-order graphical model for classifying five objects from ground, building, vehicle, bushes/low vegetation and trees/high vegetation from

the GML-PVC (http://graphics.cs.msu.ru/en/node/922) dataset. Niemeyer et al. (2015) presented a contextual classification methodology that was based on two-stage CRF for classifying data from the Vaihingen LiDAR dataset from the ISPRS Benchmark on Urban Classification and 3D Reconstruction into eight classes (Rottensteiner et al. 2012). They achieved an overall accuracy of 80.5%.

Image matching point cloud classification approaches for urban areas are presented in the following. Liao and Huang (2012) collected 92 images in the visible band, and near-infrared images with a Canon EOS 5D digital camera with an NIR filter. The two types of images were uploaded to the Photosynth website to produce point clouds. The resulting point clouds were combined in one coordinate system to compute Normalized Difference Vegetation Index (NDVI) values. The data were classified using height and colour information by thresholding into several categories: buildings, cement plane 1, cement plane 2, trees and grasses. The results showed 36% commission and omission error. That paper indicated the advantage of using a multispectral point cloud in classification over a LiDAR point cloud when using NDVI to distinguish grassland and cement planes. However, the result depended on the quality of the registration between the pair of point clouds. Gerke and Xiao (2013) demonstrated and compared two methods for 3D scene classification, which were derived from image matching for two different study sites. The point clouds were converted into the voxel representation and segmented according to features such as colour, texture and straight lines. A supervised method, which was based on the 'Random trees' machine learning technique, and an unsupervised method, which applied Markov random field with graph cuts for energy minimization, were used to classify five objects: façade, roof, rubble, sealed ground, and trees. Overall accuracies of 73.1% and 58.9%, respectively, in the Haiti area and 84.7% and 78.3% for the Enschede area were achieved. Debella-Gilo et al. (2013) found object-based image analysis (OBIA) to be a satisfactory approach for land cover mapping in the Nord-Trøndelag county of Norway using 3D point cloud data that were obtained from stereo aerial image matching and spectral data. The Ecognition software was used to segment homogeneous objects in the scenes. Then, four land types (forest, biotic open area, abiotic open area, and mire) were classified based on defined rules. The overall accuracy was approximately 80% and the kappa coefficient (κ) was approximately 0.65. Omidalizarandi and Saadatseresht (2013) performed their segmentation experiment with point clouds of three different levels of density and accuracy using a surface-growing-based approach. They verified that updating plane parameters and robust least-squares plane fitting improves the results of building extraction, especially in case of low-accuracy point clouds. Maltezos and Ioannidis (2015) proposed a method for detecting building points automatically by using two different data sets: LiDAR point clouds and image matching point clouds, which were derived using semi-global matching (SGM) (Hirschmüller 2008). The vegetation was removed by utilizing normal and roughness values, while NDVI was used to extract vegetation in the Colour Infrared (CIR) point clouds. Then, the bare earth was extracted based on morphological operations. The final building point clouds were used to generate 3D city models and perform building change detection. They concluded that point clouds that are derived from high-resolution CIR digital aerial imagery have great potential for classification tasks. Modiri, Masumi, and Eftekhari (2015) also utilized a region-growing technique to classify buildings and vegetation from two images that were acquired by the UltraCam-X camera

based on colour and a vegetation index. Hron and Halounova (2015) demonstrated the proficiency of digital aerial images for updating buildings based on the analysis of coloured point clouds that were created by an automatic image matching technique from regular acquisitions. Time series of point clouds were compared using a differential digital surface model to detect recently developed (133/139) and demolished (68/70) buildings in the test areas. Rau, Jhan, and Hsu (2015) proposed a rule-based hierarchical classification scheme method for automatically classifying a 3D point cloud from oblique aerial imagery (OAI) into various urban objects such as roof, façade, road, tree and grass in two study sites. The point cloud was generated through multi-view stereo dense matching and OBIA was applied to extract five classes. All the feature layers (OAI, object height image (OHI), gradient image, and edge image) were integrated into Ecognition for OBIA classification. They achieved an accuracy between 82.5% and 91.8% for their classification results.

In summary, in much of the investigated classification research, point clouds from matching of oblique images are rasterized or voxelized. Since any kind of rasterization, even building voxels, reduces 3D content, our proposed approach uses the original point cloud for the classification task. Although oblique images are beneficial for urban areas (especially if façade information is required), nadir images, with their different projection properties, are still an important data source in practical applications. Furthermore, little research has been done on accuracy assessments of point cloud classification over urban areas from nadir images, which will be covered in this work.

3. Materials and methods

3.1. *Data*

The entire flight consists of 266 high-resolution aerial images that were taken by the UltraCam-Xp camera in March 2011. Table 1 shows further details of the flight.

3.2. *Study area*

Vienna is located in north-eastern Austria, at the easternmost extension of the Alps, with an elevation from 151 to 542 m. To focus the research, three test areas were chosen, as shown by the yellow rectangles in Figure 1. The study sites were chosen to reflect the variability of urban environments.

Site 1 has a total area of 9.2 ha (440 m × 230 m) and a centre located at 48° 12′43.4″ N; 16°23′03.7″ E and is characterized as model city with new high buildings

Table 1. Data information.

Flight date	March 2011
Camera	UltraCam-Xp
Image format (vertical × horizontal)	11,310 × 17,310 pixel
Pixel size	6 µm in image = 6 cm on Object
Focal length	100.5 mm
Flight height	1035 m
Forward overlap/Side lap	80%/80%
Number of strips	7
Images per strip	38

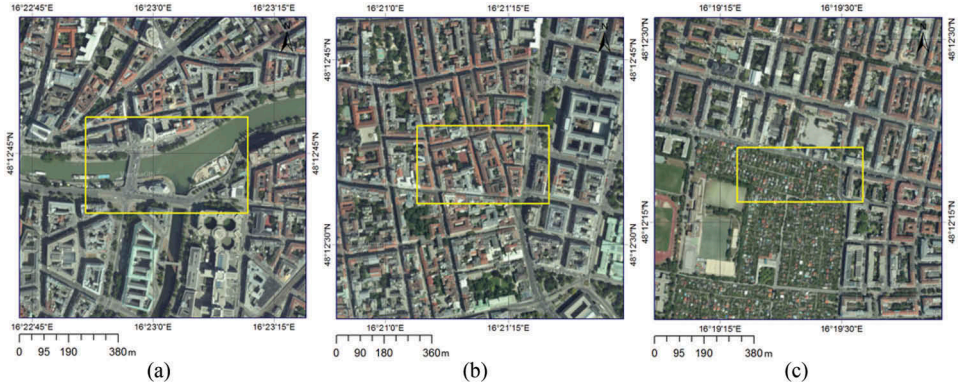

Figure 1. Three study sites in Vienna: (a) site 1; (b) site 2; (c) site 3.

(maximum 82 m), large wide water (Danube channel), a bridge, an avenue, and high trees (about 23 m) along the roads. Site 2 has a total area of 5.6 ha (297 m × 190 m), centre located at 48°12′36.1″ N; 16°21′14.2″ E and is characterized as an old city centre with old apartment buildings (complex roof structures), narrow roads and an avenue with cars and little vegetation cover. Site 3 has a total area of 4.4 ha (315 m × 140 m), centre location 48°12′18.7″ N; 16°19′25.9″ E and is characterized as a suburban area with mainly single-family dwellings and some high buildings, and a wide avenue with vehicles. This area contains various vegetation types, including high trees, hedges, isolated medium-high trees, small trees, and grassland. This is the most complex site in the study, with various kinds of urban objects.

3.3. *Outline of the method*

In this study, colourized 3D point clouds are generated from aerial images by utilizing the dense image matching technique of SGM from the software SURE (Photogrammetric Surface Reconstruction from Imagery) (Rothermel et al. 2012). The point clouds are generated for different pyramid levels (0, 1, and 2, which correspond to GSDs of 6 cm, 12 cm, and 24 cm). Those data sets are then evaluated in terms of classification accuracy, completeness and time consumption. The pyramid level 0 point cloud is used to create a digital terrain model (DTM), which is used for feature computation at all levels. Moreover, additional geometric and spectral features of a point and its neighbourhood are computed in the point cloud of each pyramid level. Those point features include, e.g. the normal vector, roughness, echo ratio (ER), Green ratio (GR) and Green-Red Vegetation Index (GRVI) (Motohka et al. 2010). Six classes are distinguished: grassland, vegetation, high buildings, small buildings, paved surface and other (including cars).

Two distinct decision tree methods are applied for the three study sites at each pyramid level: (1) a simple (subjective) decision tree method, which uses thresholds that are determined by a human operator, and (2) a machine learning method that is based on the random forest algorithm (Breiman 2001), which automatically generates the best threshold for distinguishing various urban objects. Considering all study sites, pyramid

levels and classification methods, the number of classes varies for the individual classifications. For example, water only occurs in study site 1. Moreover, for higher-resolution data sets, certain sub-classes (e.g. roofs and walls of buildings) were introduced to cover the variation of the point cloud features. For the evaluation, however, a common class set (see the end of this section) was defined.

Thresholds of the simple decision tree of one region are transferred to the other two regions to determine which thresholds should be maintained and which need to be adjusted. Similarities among different pyramid levels are also investigated.

To determine whether the 3D point cloud approach offers a performance gain in comparison to a 2.5D approach, the same machine learning method is applied to the point cloud of level 0 after gridding.

Finally, reference classifications were carefully assigned by an operator and compared with the classification results on the urban test data scenes. The classes for the evaluation area are the typical land cover classes: grassland, vegetation (bush and tree), high buildings, small buildings, sealed surface, and other. The applied processes are performed using OPALS (Orientation and Processing of Airborne Laser Scanning data) (Pfeifer et al. 2014) and ArcGIS.

3.4. *Point cloud and DTM generation*

Image matching is always performed on the entire image blocks for different pyramid levels. Additionally, the number of images in which a point needs to be visible is set. The point cloud with the highest resolution was chosen to create a DTM. To remove the off-ground points, the Lasground tool (http://www.cs.unc.edu/~isenburg/lastools/) is used to remove all buildings, trees, cars, and other off-ground points. After this processing, almost all off-ground points are removed. However, misclassified points remain and are removed carefully through a manual process using the CloudCompare software (http://www.cloudcompare.org). Then, the final point clouds are imported into OPALS software to create a DTM with a grid size of 10 cm.

3.5. *Features*

The point clouds from matching are represented by their coordinates (X, Y, Z) and 'reflectance' values (ρ_{Red}, ρ_{Green}, ρ_{Blue}) obtained from the measurements of the camera in the corresponding spectral bands. As no radiometric calibration is applied, this refers to the radiance measured in the camera and stored as digital number (Honkavaara et al. 2009). The estimation of local planes on a point basis is useful for different tasks (e.g. shaded relief) and surface normals are important geometric properties of a surface. Here, the local tangent plane is estimated by computing the best-fitting plane for the ten nearest points. Its normal vector (X_{Normal}, Y_{Normal}, Z_{Normal}) and the standard deviation (σ_{Normal}) of the fit are used as additional descriptions of the points. The distribution of points in the neighbourhood, which contains more helpful information, is derived from the structure tensor $\mathbf{T}$ (Gressin, Mallet, and David 2012). Linearity ($L_{\mathbf{T}}$), planarity ($P_{\mathbf{T}}$), and omnivariance ($O_{\mathbf{T}}$) are three features that are obtained from $\mathbf{T}$. The $L_{\mathbf{T}}$ feature is used to characterize 3D line objects such as power lines or similar structures. $P_{\mathbf{T}}$ is a feature that describes the smoothness of the surface and is related to roughness measures. $O_{\mathbf{T}}$

describes volumetric point distributions as they occur for trees. These features are computed by using the three eigenvalues $\lambda_1 > \lambda_2 > \lambda_3 > 0$ of the matrix **T** (Equations (1)-(3)).

$$L_{\mathbf{T}} = 1 - \frac{\lambda_2}{\lambda_1} \tag{1}$$

$$P_{\mathbf{T}} = \frac{\lambda_2 - \lambda_3}{\lambda_1} \tag{2}$$

$$O_{\mathbf{T}} = \sqrt[3]{\lambda_1 \lambda_2 \lambda_3} \tag{3}$$

Different neighbourhood definitions are used for the attribute computation of the ER, Z_{Rank}, Z_{Range}, and $Z_{\text{Normalised}}$, which can be derived to provide more information of the points. The ER is a measure that describes the vertical point distribution and, thus, the penetrability of the surface (Höfle, Hollaus, and Hagenauer 2012; Rutzinger et al. 2008). Z_{Range} represents the maximum height difference between the points in the neighbourhood, while Z_{Rank} is the rank of the point that corresponds to the height of the point in the neighbourhood. $Z_{\text{Normalised}}$ is the height of the point that is above the DTM ($Z_{\text{Normalised}} = Z - Z_{\text{DTM}}$).

In addition to RGB colour, spectral features such as GR and GRVI (Rau, Jhan, and Hsu 2015) are used:

'GR: considering that the green band value is generally higher than the red and blue bands in the vegetation area, GR is used to detect trees and grass as well as asphalt or cement road.' (Equation (4)) (Rau, Jhan, and Hsu 2015).

$$\text{GR} = \rho_{\text{Green}} / (\rho_{\text{Red}} + \rho_{\text{Green}} + \rho_{\text{Blue}}) \tag{4}$$

'GRVI: similar to NDVI, which utilizes near-infrared (IR) and red spectral bands, the green band is used to replace the near-IR. The GRVI value ranges from −1 to 1 wherein the greenish object GRVI values are larger than 0 and the reddish objects have a GRVI value less than 0.' (Equation (5)) (Rau, Jhan, and Hsu 2015).

$$\text{GRVI} = (\rho_{\text{Green}} - \rho_{\text{Red}}) / (\rho_{\text{Green}} + \rho_{\text{Red}}) \tag{5}$$

Thus, in addition to coordinates, the points contain the following additional parameters: normal vector (X_{Normal}, Y_{Normal}, Z_{Normal}), roughness (σ_{Normal}), ER, Z_{Range}, Z_{Rank}, $Z_{\text{Normalised}}$, $L_{\mathbf{T}}$, $P_{\mathbf{T}}$, $O_{\mathbf{T}}$, ρ_{Red}, ρ_{Green}, ρ_{Blue}, GR, and GRVI.

3.6. *Machine learning classification by random forests*

Machine learning classifiers infer classification rules from annotated training data with minimal human intervention. The implementation of the machine learning algorithm that is used in this study is described by Waldhauser et al. (2014). A small training set is used for building a model of classification rules, which can then by applied to classify the larger, unseen data automatically. The acquired sample must contain the specular characteristic of each class.

Point-based classification: The training samples of the three different study sites at different pyramid levels are carefully chosen. Seven main groups are considered:

grassland, trees (high tree, medium tree, and low tree), high buildings (solid high building roof, walls, roof represented in low quality, and structure on the roof), small buildings (dwellings), sealed surface (road surface and bare ground), water (water surface and structure on the water surface) and other (cars and traffic installations). Due to detail differences in the pyramid levels, more detailed sample classes can be obtained at pyramid level 0.

Raster-based classification: The point clouds are rasterized, including the reference class, where available. The point that is closest to the cell centre is selected to represent the data at the raster cell. The same machine learning classification method is applied to the rasterized data and the original point cloud. The process for running the classification is based on the point cloud.

All processing steps, from selecting training samples, attribute creation, classification model building and applying the classification model, are performed by OPALS and DTMaster (http://www.trimble.com/imaging/inpho.aspx) software.

3.7. *Simple decision tree classification*

Unlike machine learning, which chooses thresholds and rules automatically based on the sample data, a simple decision tree is constructed by an operator, who finds the thresholds of the classification rules by visually analysing the point features within the point cloud. From a practical point of view, it is beneficial to use rasterized images of point features since they show reduced noise and corresponding values can be selected easily in any image viewing software. The features are aggregated in a grid size of 10 cm for the point cloud in pyramid level 0 and a grid size of 1 m in pyramid level 2 (for pyramid level 1, no aggregation was performed). These images were compared visually and those feature images that allowed (visually) the best separation of classes were selected: $Z_{\text{Normalised}}$, GR, $O_{\mathbf{T}}$, and $P_{\mathbf{T}}$. $Z_{\text{Normalised}}$ is useful for distinguishing objects based on their height, such as on-ground vs. off-ground objects, and classifying buildings and trees into categories of low, medium and high, including buildings in the lower category, vegetation, and cars. GR is useful for distinguishing greenish objects from other objects of the same height, such as grassland vs. sealed surface, cars, and low trees, and small buildings vs. medium trees. $O_{\mathbf{T}}$ distinguishes volumetric objects from other objects of the same height, such as high trees vs. buildings. $P_{\mathbf{T}}$ distinguishes planar objects and volumetric objects, such as small buildings vs. medium trees.

After defining the thresholds, a rule-based hierarchical decision tree is constructed. First, ground points (sealed surface and grassland) are extracted from the point cloud. Then, elevated points are separated into lower objects (cars and low trees), medium-height objects (small buildings and medium trees) and higher objects (high buildings and high trees). Finally, the four height classes are split into the defined sub-classes. In site 1, the discrimination of water points was attempted. Within the point clouds, water points and road points exhibit similar planar distributions. Additionally, based on the colour, some points of shaded roads are classified as water.

An automatic decision tree is constructed using the standard classification and regression tree (CART) technique (Breiman et al. 1984). Given that the operator-based simple decision tree can be considered an 'over-pruned' tree with limited nodes, the

depth of the automatic decision tree is controlled to generate a tree with similar structure to the simple tree. Consequently, the corresponding thresholds can be identified and compared to those that are selected by the operator.

3.8. *Reference data and accuracy assessment*

The reference data for the accuracy assessment are classified manually by a human using the DTMaster software and an external high-resolution data source, i.e. the orthophoto. It is necessary to label each point through a thorough visual analysis. Because this is very time-consuming, only two small parts (of areas 1680 m^2 and 1120 m^2) that contain abundant objects in study Site 3 are chosen. These small parts comprised natural objects and artificial objects, which are combined into six classes: ground (grassland area), road (including road and cement surfaces), tree (including small, medium, and high trees), high building, small building, and other (here, cars). The selected area was not used as training data for the classification procedures and is not significantly influenced by shadows, which allows for an objective assessment. Three different pyramid levels of point clouds are manually processed and the time that is required for manual classification is approximately 150 hours with 555,435, 167,258, and 43,344 points in levels 0, 1, and 2, respectively.

The reference that is used for accessing the 2.5D raster classification is also rasterized from the point cloud reference.

A confusion matrix was built from the point cloud files, which contain all classification values (machine learning, simple decision tree, and reference), for reporting the overall accuracy, producer and user accuracies, and κ.

4. Result and discussion

4.1. *Image matching results*

Three different pyramid levels are produced from different image resolutions with SGM (Hirschmüller 2008). Level 0 is the default full resolution, level 1 has a quarter of the pixels, and level 2 a sixteenth. For the study of the computation time, an additional level, namely, level 3, is generated. The pyramid level influences the processing time, point density (Table 2), quality and completeness (Table 3) of the resulting point clouds. An important image matching parameter describes the number of images in which a point

Table 2. Point densities in different pyramid levels. Density is given in [1/m^2] and the column $P(i)/P(i+1)$ refers to the change in density.

	Site 1		Site 2		Site 3	
	Point count/Point density	$P(i)/P(i+1)$	Point count/Point density	$P(i)/P(i+1)$	Point count/Point density	$P(i)/P(i+1)$
Pyramid level 0 – fold 3	12,686,882/125.0	3.55	9,743,175/172.0	3.46	9,816,210/222.6	3.63
Pyramid level 1 – fold 3	3,568,895/35.3	3.68	2,816,211/50.0	3.78	2,702,115/61.3	3.86
Pyramid level 2 – fold 3	969,114/9.6		745,059/13.2		700,253/15.9	

Table 3. Total areas of incompleteness for different pyramid levels and folds.

Pyramid level	Total area of incompleteness (m^2)	Percentage (%)
-pyr 0 -fold 4	2545	5.8
-pyr 0 -fold 3	1094	2.5
-pyr 1 -fold 3	908	2.1
-pyr 2 -fold 3	823	1.9

must be visible to be accepted. In SURE, this is implemented as the number of stereo pairs in which a point must appear and is called the 'fold'. The fold parameter balances precision against completeness.

For the full resolution with a minimum of 3 stereo models, the computation took approximately fourteen days for processing with the Windows Server 2008 R2 Operating system, Opteron 2 × 2.6 GHz – Quad core, and 32G of RAM (Random Access Memory). In contrast, for 1/64 of the number of pixels (pyramid level 3; only computed for processing time analysis), it took only 4 h to run the whole area. Between adjacent pyramid levels, processing time grows quadratically, or at least stronger than linearly, in the number of pixels or object points (Figure 2). The main advantage of the lower pyramid levels lies in the higher point density (Figure 3), with factors that are between 3 and 4 between adjacent levels (Table 2).

Lower pyramid levels, on the other hand, lead to higher incompleteness of the output data. Table 3 lists the losses of information in the total cover area at Site 3 in different levels, which are obtained by counting the total number of no-data pixels in a cell of size 0.5 m. According to Table 3, fold 3 achieves higher completeness than fold 4 in the same

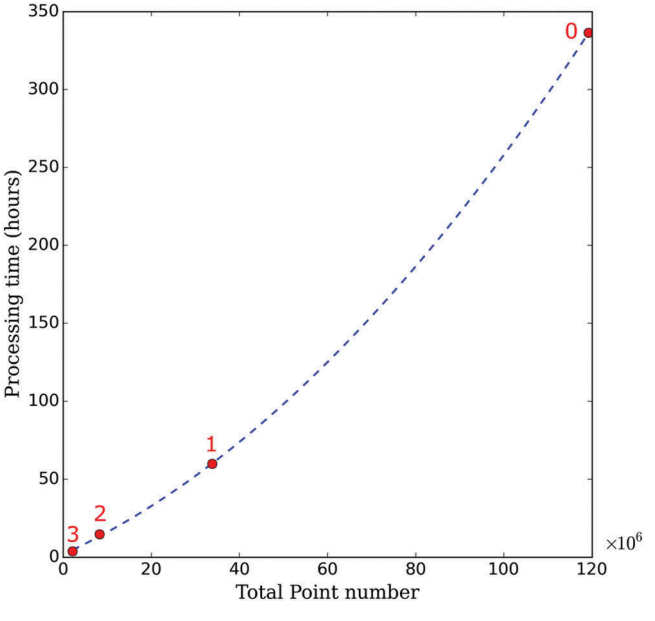

Figure 2. Relationship between processing time and the total number of points in different pyramid levels.

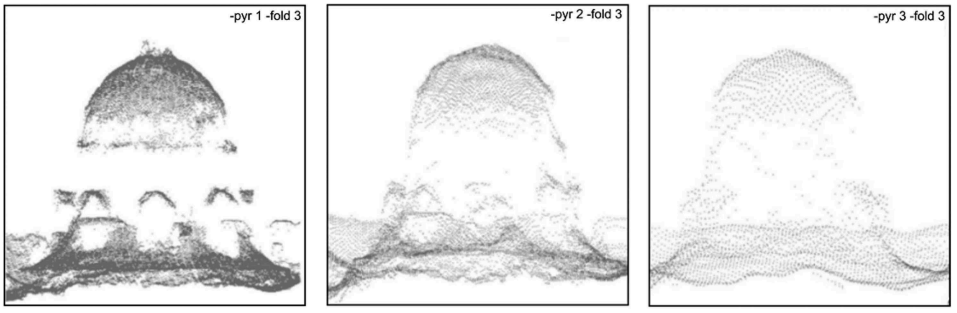

Figure 3. Point densities of different pyramid levels: (a) -pyr 1 – fold 3, (b) -pyr 2 – fold 3, (c) -pyr 3 – fold 3.

pyramid level. The ratio of matched points per pixel is, on average, 0.4 points/ground pixel in Site 1, 0.5 points/ground pixel in Site 2, and 0.7 points/ground pixel in Site 3.

Due to image matching ambiguity, point clouds can show large variation and even no results, especially in shadowy areas and homogeneous areas (e.g. road surfaces and water regions).

4.2. *Machine learning classification using random forests*

First, the results of point-based machine learning classification are described. In the generation of the training data, our aim was to sample the 6 target classes by sub-classes that can be easily separated in the data. The number of (sub-)classes decreases with the reduction in point density and precision (Table 4). The objects with planar surfaces, such as building roofs, grassland, and sealed surface, can be easy distinguished from vegetation objects in pyramid level 0. However, it becomes more difficult at higher pyramid levels. At that state, roofs and sealed surfaces are not planar any more. They become undulated (Figure 4), which leads to difficulties in separating small buildings

Table 4. Groups and classes of samples in pyramid levels 0, 1, and 2.

Group	Pyramid level 0	Pyramid level 1	Pyramid level 2
Sealed surface	Cement surface Roads	Cement surface Roads	Cement surface Roads
Grassland	Grassland	Grassland	Grassland
Vegetation	High trees	High trees	High trees
	Medium trees	Medium trees	Medium trees
	Low trees		
Small building	Dwellings	Dwellings	Dwellings
High building	Solid high building roof	Solid high building roof	Solid high building roof Walls
	Walls	Walls	
	Roof, represented in low quality	Roof, represented in low quality	
	Structure on the roof		
Water	Water surface	Water surface	Water surface
	Structure on the water	Structure on the water	
Other	Cars	Cars	
	Traffic installations		

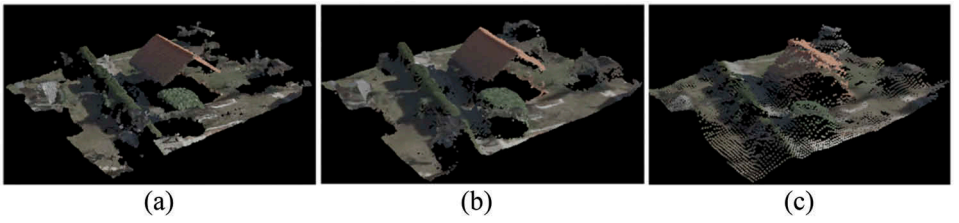

Figure 4. Point distributions in different pyramid levels: (a) -pyr 0, (b) -pyr 1, (c) -pyr 2.

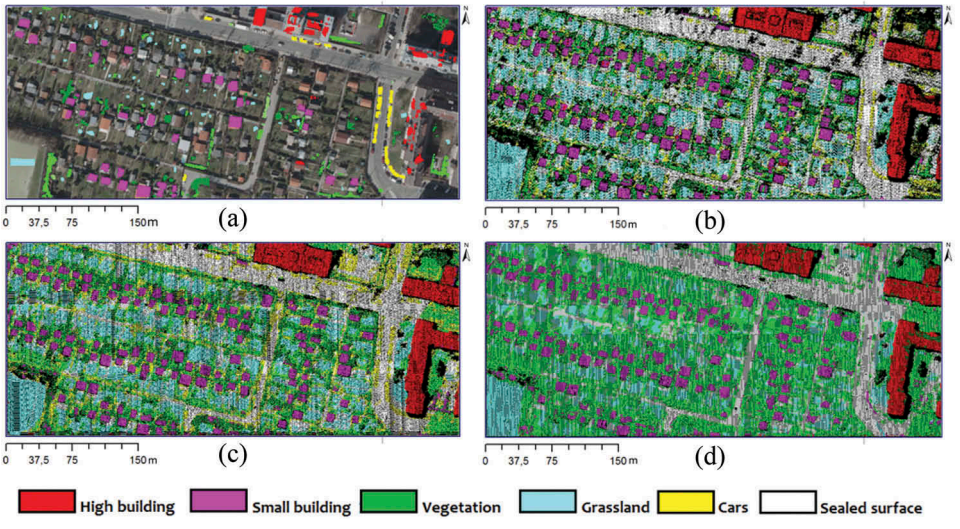

Figure 5. Machine learning classification results for different pyramid levels in site 3: (a) sampling data for level 0 overlaid on the ortho-image, (b) classification result in pyramid level 0; (c) classification result in pyramid level 1, and (d) classification result in pyramid level 2.

from medium vegetation and road from cars. This makes it challenging to manually select training data for use in learning the decision trees.

Figure 5 shows the classification results in pyramid level 0, pyramid level 1 and pyramid level 2 of the machine learning classifier. The training data for pyramid level 0 are shown in Figure 5(a). The number of sample points for the different classes are as follows: grassland – 69,107, trees – 185,060, high buildings – 100,679, small buildings – 204,592, sealed surface – 231,984, and cars – 44,298. The total time that is used to acquire samples is 3 hours for this site. Half of the training data are used for learning and half are used for evaluation of the learned model (but not for the overall evaluation). Visually, the final results are satisfactory in terms of distinguishing the major objects. The classification results strongly depend on the sample point data, not only the number of points but also the distribution of samples. The training samples must cover the full attribute variations within the data set to achieve satisfactory classification results. According to experience, this can be achieved by evenly distributing the training samples across the whole area such that each group fully covers the attribute variations. For example, grassland in this season does not have the same colour for the whole

region, so instead of sampling grassland points in the stadium area only, clusters of grassland in different courtyards need to be sampled. To increase the accuracy of the classification results in higher pyramid levels, the reference objects need to be redoubled. For instance, in pyramid level 0 of Site 3, the small building sample consists of 11 objects. Therefore, in pyramid level 1, the number of small building sample objects should be 22 to ensure appropriate classification results. As shown in Figure 5(b–d), the number of classes that can be distinguished decreases as the pyramid level increases and misclassification increases.

With each study site, sampling points for each class are acquired manually. The training samples contain 'pure' data and require careful selection. The total time that was spent taking samples was 6 hours for the three sites. Figure 7 (left column) displays the classification results of all three sites at pyramid level 0. Site 1 covers high modern buildings with planar roof structure, a wide avenue, and a river. Visually, grassland, vegetation, high buildings, small buildings, water points, and sealed surface are detected proficiently. However, shaded sealed surface points are misclassified into grassland because of their dark colour. Site 2 represents an older part in Vienna, which is dominated by high buildings and narrow roads in between. Building roofs have varying structures and colours, and the scene is strongly influenced by shadows. Thus, in some areas, point clouds show high noise or are missing. This leads to roof points that are distributed similarly to tree points. In addition, trees are affected by shade, which is why their colour is varying. Eventually, this leads to building roofs being misclassified as trees. The wide road, which is homogeneously illuminated, is well classified, whereas narrow roads are wrongly detected because of shadows from high buildings. Site 3 is the most vegetated area and there are many objects compared with the two other sites. Objects are classified effectively. However, the 'other' class is mixed with the low tree class in shaded areas. Grassland is mixed into the sealed surface class in some areas that are covered by dried ground.

As mentioned in the introduction, one aim of this investigation is to determine whether point-based methods outperform raster-based methods that use machine learning classification. The main advantage of the 2.5D raster approach compared to point cloud classification lies in its simpler processing because its neighbourhoods are implicitly given and the processing time is shorter because of the faster neighbourhood search. However, the raster is less sharp compared with the point cloud (Figure 6) because of pixels where no point or multiple points are given. However, to reduce the influence of feature

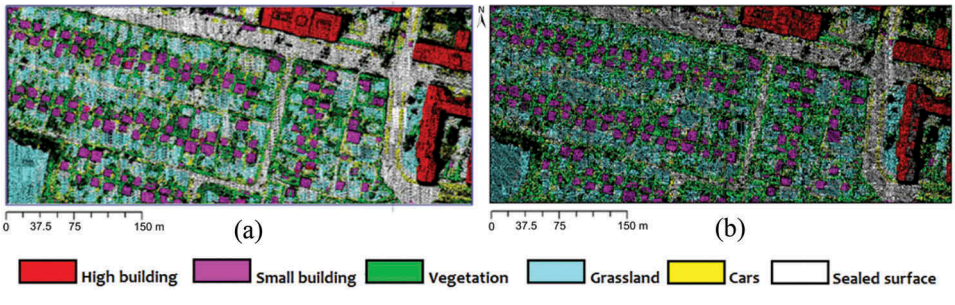

Figure 6. Machine learning classification results: (a) point-based; (b) raster-based.

computation on the comparison, the point cloud features were also gridded. The final result for Site 3 shows that the accuracy of raster classification (Table 12) is slightly lower than that of the point cloud (approximately 5%). Additionally, the 3D content is inevitably lost when applying a raster approach in comparison to the point cloud classification.

4.3. *Simple decision tree classification*

Using only three attributes, namely, $Z_{\text{Normalised}}$, GR and $O_{\mathbf{T}}$, simple decision tree classification could be performed for two sites (Sites 2 and 3). Site 1 needed an additional attribute, namely, $P_{\mathbf{T}}$, to achieve satisfactory results. For this site, point clouds can be classified into eight classes: grassland, sealed surface, low vegetation, medium vegetation, high vegetation, small buildings (dwellings), high buildings and cars. $Z_{\text{Normalised}}$ is the main threshold for distinguishing various objects. GR is used to detect vegetation objects. However, with high trees, GR is not effective because of seasonal and shading effects. Thus, a threshold on the $O_{\mathbf{T}}$ attribute is applied to separate high buildings and high trees.

Different sites required individual thresholds (Table 5, Table 6, and Table 7). However, the thresholds for grassland, sealed surface, low tree, and car could be maintained. The thresholds of $O_{\mathbf{T}}$ are varied from site to site because of the differences in tree arrangements.

The threshold for distinguishing between trees and buildings depends on the heights of the trees and buildings in the region. The final classification results for the simple decision tree method of the three different study sites are shown in Figure 7 (right column). The time that is spent empirically selecting thresholds and designing the decision tree is 3 hours. The same thresholds as were used in pyramid level 0 are considered and evaluated in the higher pyramid levels. Figure 8 shows the results when the same threshold as that used for pyramid level 0 is applied to pyramid levels 1 and 2. The threshold of pyramid level 0 can be applied to pyramid level 1 but not pyramid level 2. In pyramid level 2, only the

Table 5. Thresholds of simple decision tree classification in pyramid level 0 of site 1.

Site 1	$Z_{\text{Normalised}}$ (m)	GR	O_{T}	P_{T}
Grassland	≤0.2	>0.3393	–	–
Sealed surface	≤0.2	≤0.3393	–	–
Low trees	<1.8	>0.3393	–	–
Cars	<1.8	≤0.3393	–	–
Small buildings	>1.8 and ≤ 5.5	–	–	>0.6
Medium trees	>1.8 and ≤ 5.5	–	–	≤0.6
High trees	>5.5 and <23	–	>0.03	–
High buildings	>23	–	<0.03	–

Table 6. Thresholds of simple decision tree classification in pyramid level 0 of site 2.

Site 2	$Z_{\text{Normalised}}$ (m)	GR	O_{T}
Grassland	≤0.2	>0.3393	–
Sealed surface	≤0.2	≤0.3393	–
Low trees	<1.5	>0.3393	–
Cars	<1.5	≤0.3393	–
Small buildings	>1.5 and ≤ 5.5	≤0.348	–
Medium trees	>1.5 and ≤ 5.5	>0.348	–
High trees	>5.5 and <14.5	–	>1
High buildings	>17	–	<1

Table 7. Thresholds of simple decision tree classification in pyramid level 0 of site 3.

Site 3	$Z_{Normalised}$ (m)	GR	O_T
Grassland	≤0.2	>0.3393	–
Sealed surface	≤0.2	≤0.3393	–
Low trees	<1.8	>0.3393	–
Cars	<1.8	≤0.3393	–
Small buildings	>1.8 and ≤ 5.5	≤0.348	–
Medium trees	>1.8 and ≤ 5.5	>0.348	–
High trees	>5.5 and <14.5	–	>0.17
High buildings	>14.5	–	<0.17

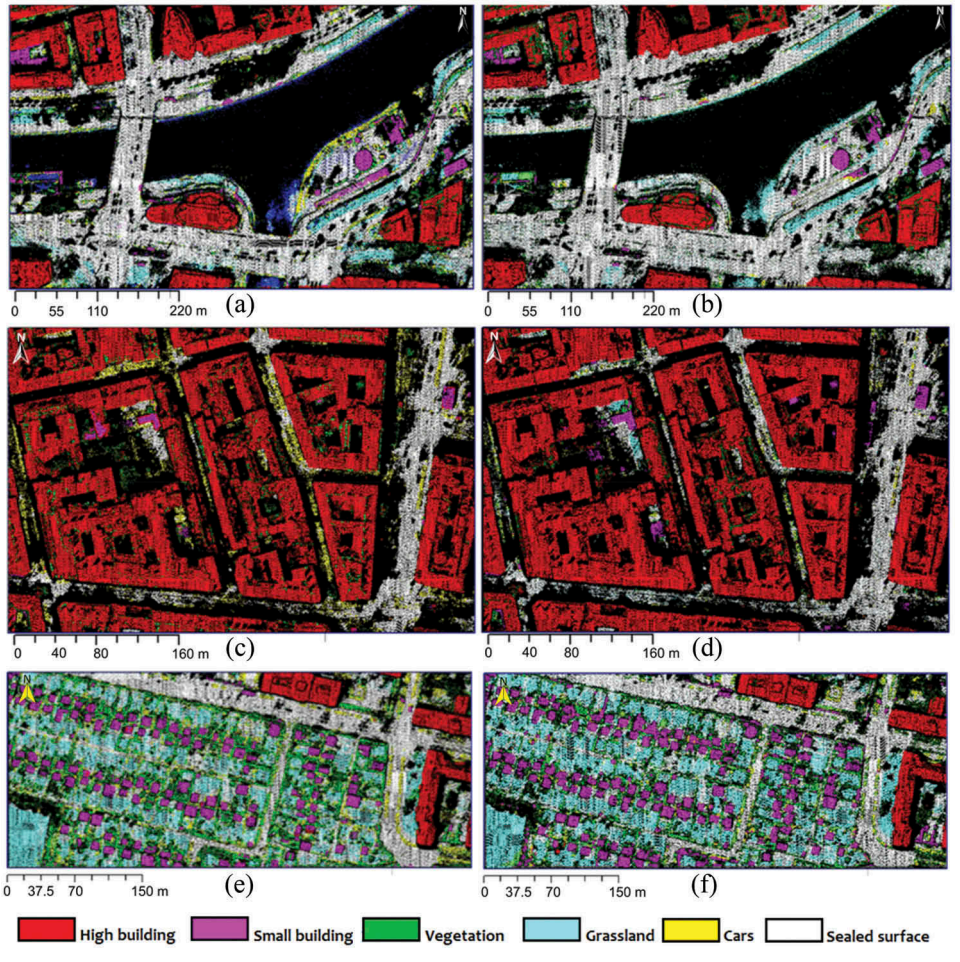

Figure 7. Machine learning (random forest) in (a), (c), (e) and simple decision tree (manual thresholds) for three study sites in (b), (d), (f) for the three study sites.

high building threshold delivers satisfactory results; other thresholds needed to be changed. Hence, an additional step of empirically selecting thresholds is required. After changing thresholds (Table 8), the result improves (Figure 9). Notably, $Z_{Normalised}$ thresholds did not need to be changed, which suggests that this feature depends more strongly on the object type than on the parameters of image acquisition or image matching.

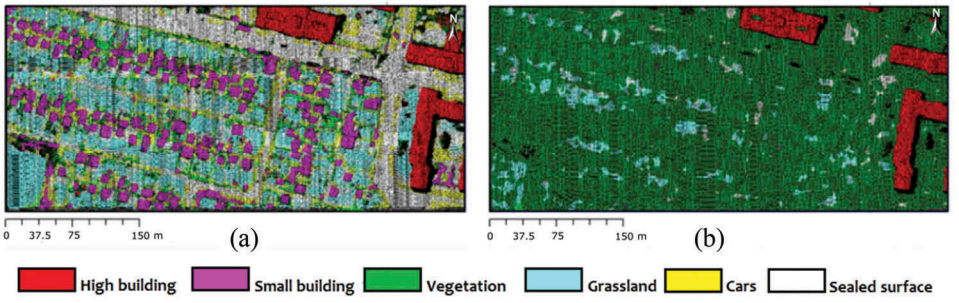

Figure 8. Results of the simple decision tree of level 0 for pyramid levels 1 (a) and 2 (b) of site 3.

Table 8. New threshold for pyramid level 2 of site 3.

Site 3	$Z_{Normalised}$ (m)	GR	O_T
Grassland	≤0.2	>0.341	–
Sealed ground	≤0.2	≤0.341	–
Low trees	<1.8	>0.334	–
Cars	<1.8	≤0.334	–
Small buildings	>1.8 and ≤ 5.5	≤0.36	–
Medium trees	>1.8 and ≤ 5.5	>0.36	–
High trees	>5.5 and <14.5	–	>0.5
High buildings	>14.5	–	<0.5

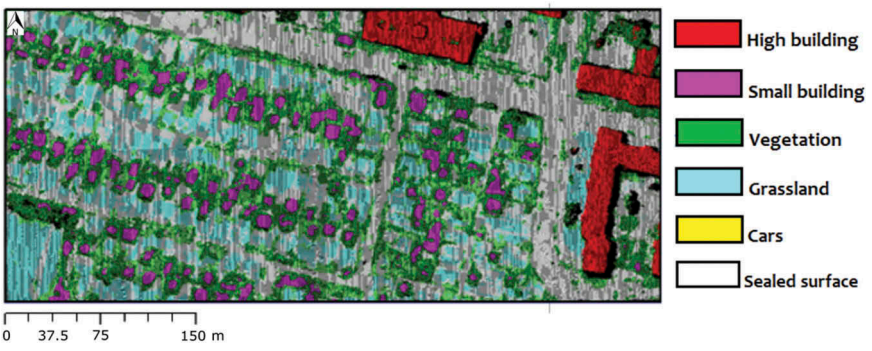

Figure 9. New thresholds applied to pyramid level 2 in site 3.

The automatic method for setting thresholds for a decision tree using the same three features yielded very similar results to the manual setting of threshold values (Figure 7, right column). Thus, the results are not shown separately. The analysis of the confusion matrix indicates that small buildings are classified with higher producer and user accuracies using the automatically selected thresholds.

4.4. *Accuracy assessment*

To compare the classification results between the two methods, namely, machine learning and simple decision tree, the classes are combined into the same six main classes: grassland (Grass), sealed surface (S-surface), vegetation (Veg), high building (H-build), small building (S-build) and other. Site 3 was chosen because its object variety is

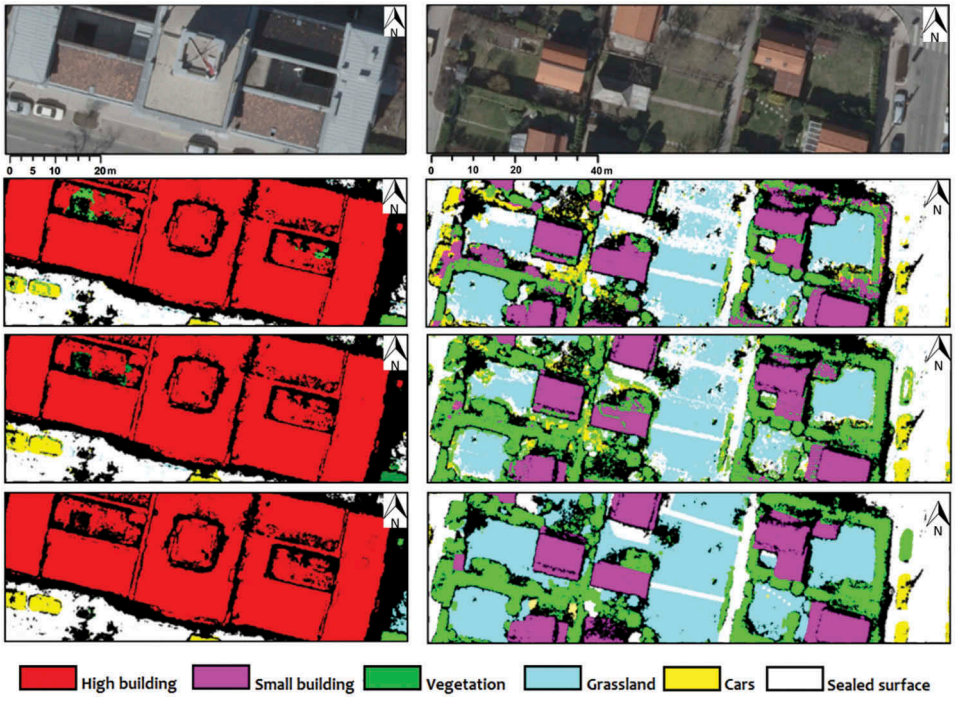

Figure 10. Two different accuracy evaluation areas in site 3: left column shows an area with a high building and street, right column shows small buildings with gardens. From top to bottom: orthophoto, classification result of the simple manual decision tree, classification result of machine learning by random forest, and reference classification.

the highest, compared to the other sites. Figure 10 displays the two reference areas in Site 3.

The overall accuracies of machine learning by random forest and simple manual decision tree for different pyramid levels are summarized in Table 9. In general, machine learning provides better classification results compared with the simple decision tree in the same pyramid level. The reason is that in machine learning, the model (i.e. the set of decision trees) is more complex than a simple decision tree and a thorough examination of the diverse point features is performed, while the operator can only consider a limited set of features when designing a decision tree. The automatic selection of threshold values for a decision tree, which considers four features, leads to slight improvement in performance relative to manual selection. The overall accuracy increases from 84.1% to 84.5% with a κ of 0.80.

Table 9. Overall accuracies of machine learning by random forest and simple manual decision tree for different pyramid levels.

	Machine learning		Simple decision tree	
	Overall accuracy (%)	κ	Overall Accuracy (%)	κ
P0f4	87.2	0.84	84.1	0.80
P1f3	83.2	0.79	77.9	0.73
P2f3	79.2	0.74	68.6	0.61

Table 10. Confusion matrix of pyramid level 0 of simple decision tree (SDT) versus reference classification.

		Reference classification							
	Class	Grass	Veg	H- build	S- build	S-surface	Other	Total	Producer (%)
SDT	Grass	84,152	235	0	0	2349	129	86,865	96.9
	Veg	5306	51,30	1448	5352	4249	4468	72,124	71.1
	H-build	6	148	179,865	0	0	0	180,019	99.9
	S-build	1	16,186	0	50,741	0	13	66,941	75.8
	S-surface	37,150	1369	0	4	96,926	588	136,037	71.3
	Other	1460	7424	0	118	263	4184		31.1
	Total	128,075	76,663	181,313	56,215	103,787	9382	13,449	
	User (%)	65.7	66.9	99.2	90.3	93.4	44.6	555,435	
	Overall accuracy (%)			84.1		κ		0.80	

Table 11. Confusion matrix of pyramid level 0 of machine learning (random forest, ML) versus reference classification.

		Reference classification							
	Class	Grass	Veg	H-build	S-build	S- surface	Other	Total	Producer (%)
ML	Grass	82,973	249	0	1	3433	14	86,670	95.7
	Veg	14,162	69,584	602	6502	4910	1068	96,828	71.9
	H-build	6	146	180,711	0	0	0	180,863	99.9
	S-build	0	2169	0	49,629	0	38		95.7
	S-surface	25,556	1086	0	3	93,293	158	52,136	77.7
	Other	3670	5288	0	79	2013	8092	120,096	42.3
	Total	126,367	78,522	181,313	56,214	103,649	9370	19,142	
	User (%)	65.7	88.6	99.7	88.3	90.0	86.4	555,435	
	Overall accuracy (%)			87.2		κ		0.84	

Both methods perform well in terms of high building (over 99% accuracy) and sealed surface (over 90% accuracy) distinction. Machine learning by random forest is distinctively better for 'other' objects, with an accuracy of 86.4%, which is twice the user accuracy of the simple decision tree (44.6%) (Table 10 and Table 11). It also performs better in separating small buildings from vegetation. The raster-based method, in comparison to the point-based method (Table 12 and Table 11), performs less well in terms of the quality of the discrimination between grass and sealed surface and in discriminating between vegetation and small buildings.

Table 12. Confusion matrix of pyramid level 0 of the Raster-machine learning (random forest, Ras-ML) approach versus reference classification.

		Reference classification							
	Class	Grass	Veg	H-build	S-build	S- surface	Other	Total	Producer (%)
Ras	Grass	61,266	15,635	6	1	44,708	3629	125,245	49.0
-ML	Veg	98	70,356	20	586	642	5342		91.3
	H-build	1	95	181,217	0	0	0	77,026	99.9
	S-build	63	12,970	1479	41,472	0	70	181,313	74.0
	S-surface	1852	3130	0	0	95,010	2238	56,054	93.0
	Other	76	1987	0	1	999	9964	102,230	76.5
	Total	63,356	104,173	182,722	42,060	141,341	21,242	13,027	
	User (%)	96.7	67.4	99.2	98.6	67.2	46.9	554,895	
	Overall accuracy			82.3%		κ		0.78	

Because of the limitations of the image matching method, the point cloud is positioned on top of the canopy of vegetation, so it is difficult to discriminate tree objects because their volumetric structure is not captured. The greenness attribute plays an important role in classifying trees. However, it remains problematic in differentiating artificial objects such as a green car, a green roof, and water. For that reason, machine learning based on point-neighbourhood geometric features such as the normal vector improves the quality of the classification results, as shown in Figure 7, where the classification results of both methods for three sites are compared. Machine learning can recognize water points, which is a hard task, using a simple (manual) decision tree.

Shade is a big problem for photogrammetric imagery. It reduces the point precision and colour information due to the limited textural information. Therefore, the accuracy of the classification results decreases. For instance, shady roads can be misclassified as grassland. For some high building roofs, because of shade, occlusion, diversity of roof structure, and lack of texture, the matching results are deteriorated. This can cause an undulating height variation of planar roof surfaces, which can be misclassified into the high tree class. Seasonal variations can also result in lower classification quality: The images were taken in March (Table 1), when the foliage was not fully developed. Therefore, not all trees were completely green. This caused additional difficulties in distinguishing them from other objects.

Both supervised methods provide high accuracies for urban classification tasks. The highest accuracy and details are obtained for the full-resolution pyramid level 0, compared to other pyramid levels. However, the time that is required for processing is higher by approximately a factor of 4 from one level to the next. This may lead to an arduous task if processing large areas. For such situations, a smaller number of classes and lower pyramid levels are recommended for quicker processing. Both methods perform well in building and sealed surface detection. Thus, they are suitable for urban areas, where buildings and roads occupy a high ratio of the total land cover.

The main advantages of machine learning are the automatic selection of thresholds and its lower subjectivity than the simple manual decision tree threshold selection. However, it requires selecting samples for each region (and pyramid level) individually, which takes more time than designing a manual decision tree (approximately twice as much time). To evaluate our empirical threshold approach, we also used the samples that were used for machine learning, to automatically learn the thresholds for a simple decision tree. The final result of this automatic decision tree yielded an accuracy of 84.5%, which is an improvement of 0.4% compared to our simple manual decision tree result.

Gerke and Xiao (2013) and Rau, Jhan, and Hsu (2015) proposed two methods for image matching point cloud classification in urban areas. Both studies used oblique airborne imagery, which has outstanding properties regarding vertical structures such as building façades and trees. Therefore, their classified urban objects are different from those presented here. They divide buildings into roof and façade classes, whereas we separate buildings into high buildings and small buildings. The classes that they have in common are trees, grass, and road (sealed surface). Moreover, cars were classified in our approach but not in theirs. With respect to the method, both approaches used object-based classification on rasters, while our method is a point-based classification, which is carried out in full 3D. Our overall accuracy of approximately 85% is equivalent to theirs.

5. Conclusions

In this paper, we have demonstrated two supervised classification methods using point clouds that were created from dense image matching technology. The GSD of the investigated images was 6 cm. The point clouds, which include colour information, are extracted by the SGM method in the software SURE. Quality, completeness, processing time, and number of classes are investigated for point clouds of different levels of detail. For the full-resolution pyramid level, the highest quality and precision were achieved, but with lower performance in terms of completeness (area coverage) and processing time. In the three pyramid levels, the point density changes by a factor of 3 to 4 per level for all three study sites. The computation time for dense matching also depends strongly on the pyramid level. Our experiments, in which SURE is used, indicate that the computation time grows more strongly than linearly with the number of pixels between pyramid levels.

Additional features are calculated based on geometric and spectral radiometric properties, such as normal vector (X_{Normal}, Y_{Normal}, Z_{Normal}), roughness (σ_{Normal}), ER, Z_{Range}, Z_{Rank}, $Z_{\text{Normalised}}$, $L_{\mathbf{T}}$, $P_{\mathbf{T}}$, $O_{\mathbf{T}}$, ρ_{Red}, ρ_{Green}, ρ_{Blue}, GR, and GRVI. Those features are used as input for two supervised classification methods. One is machine learning based on a random forest algorithm and the other is a rule-based hierarchical scheme that uses simple decision trees for which thresholds are determined empirically. At pyramid level 0, the most details are available, and the features are the most expressive. Thus, it obtained the highest number of discernible classes. As a result, the lower the level of detail in the point cloud, the smaller the number of classes. Both methods obtain the highest classification accuracy in the point cloud of pyramid level 0, which is reduced evenly per level. In the same level, the machine learning method always achieves higher accuracy due to its higher number of input parameters and more complex decision tree structure; its accuracy is approximately 87.2%, while for the simple manual decision tree, the accuracy is 84.1%. For higher pyramid levels (GSDs of 12 cm and 24 cm), the classification accuracy decreases per level by 4% for machine learning and 7% for the simple decision tree. The application of machine learning to rasterized data provided slightly lower accuracy (by 5%) than the point-based method for pyramid level 0. However, because of the rasterization of the 3D content, objects that are on top of each other are lost.

Point cloud classification is performed in different research regions and levels of detail to evaluate the performances of point-based classification methods and the capability of transferring thresholds between regions and pyramid levels. In machine learning classification, for each study region and pyramid level, it is necessary to select new training points. For the simple decision tree and for different regions, some thresholds can be preserved, but others need to be adjusted to achieve an acceptable classification result. The investigated thresholds at levels 0 and 1 (6 cm and 12 cm GSD) could be maintained, but for level 2 (24 cm GSD), a readjustment was required. $Z_{\text{Normalised}}$ appears to be strongly related to the object and, therefore, mission-independent, whereas features that describe the distribution of points ($O_{\mathbf{T}}$) appear to be more dependent on the parameters of data acquisition or matching.

In this study, point clouds with different levels of detail were created by dense image matching and classified proficiently. Therefore, classified point clouds can be applied in many aspects of urban planning and management.

Acknowledgments

We would like to thank to Institute for Photogrammetry, Stuttgart University, for supplying us with the free license for the SURE software, to generate point cloud data from high-resolution aerial images. The authors would also like to thank VIED (Vietnam International Education Development) and OeAD (Österreichischer Austauschdienst) for support during the study period. The images were provided by the 'Stadtvermessung Wien', the Magistrate of the City of Vienna.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Antonarakis, A. S., K. S. Richards, and J. Brasington. 2008. "Object-Based Land Cover Classification Using Airborne LiDAR." *Remote Sensing of Environment* 112 (6): 2988–2998. doi:10.1016/j.rse.2008.02.004.

Biasion, A., L. Bornaz, and F. Rinaudo. 2005. "Laser Scanning Applications on Disaster Management." In *Geo-Information for Disaster Management*, edited by P. van Oosterom, S. Zlatanova, and E. M. Fendel. Berlin, Heidelberg: Springer.

Breiman, L. 2001. "Random Forests." *Machine Learning* 45: 5–32. doi:10.1023/A:1010933404324.

Breiman, L., F. Jerome, J. S. Charles, and A. O. Richard. 1984. *Classification and Regression Trees*. Boca Raton, FL: CRC press.

Choma, A., C. Ratcliff, and R. Frisina, 2005. "Evaluation of Remote Sensing Technologies for High-Resolution Terrain Mapping." In *SSC 2005 Spatial Intelligence*, edited by I.A.P. Melbourne: National biennial Conference of the Spatial Sciences Institute.

Debella-Gilo, M., K. Bjørkelo, J. Breidenbach, and J. Rahlf. 2013. "Object-Based Analysis of Aerial Photogrammetric Point Cloud and Spectral Data for Land Cover Mapping." *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-1/W1: 63–67. doi:10.5194/isprsarchives-XL-1-W1-63-2013.

Dorninger, P., and N. Pfeifer. 2008. "A Comprehensive Automated 3D Approach for Building Extraction, Reconstruction, and Regularization from Airborne Laser Scanning Point Clouds." *Sensors* 8: 7323–7343. doi:10.3390/s8117323.

Filin, S., A. Borka, and Y. Doytsher. 2009. "Generation of Parcelation Proposals Aided by Lidar Derived Spatial Cues." *Computers, Environment and Urban Systems* 33 (4): 278–284. doi:10.1016/j.compenvurbsys.2009.01.010.

Gerke, M., and J. Xiao. 2013. "Supervised and Unsupervised MRF Based 3D Scene Classification in Multiple View Airborne Oblique Images." *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* II-3/W3: 25–30. doi:10.5194/isprsannals-II-3-W3-25-2013.

Gressin, A., C. Mallet, and N. David. 2012. "Improving 3D LiDAR Point Cloud Registration Using Optimal Neighborhood Knowledge." *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 1 (3): 111–116. doi:10.5194/isprsannals-I-3-111-2012.

Haala, N., and M. Kada. 2010. "An Update on Automatic 3D Building Reconstruction." *ISPRS Journal of Photogrammetry and Remote Sensing* 65: 570–580. doi:10.1016/j.isprsjprs.2010.09.006.

Heiden, U., W. Heldens, S. Roessner, K. Segl, T. Esch, and A. Mueller. 2012. "Urban Structure Type Characterization Using Hyperspectral Remote Sensing and Height Information." *Landscape and Urban Planning* 105 (4): 361–375. doi:10.1016/j.landurbplan.2012.01.001.

Hirschmüller, H. 2008. "Stereo Processing by Semiglobal Matching and Mutual Information." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30 (2): 328–341. doi:10.1109/TPAMI.2007.1166.

Höfle, B., M. Hollaus, and J. Hagenauer. 2012. "Urban Vegetation Detection Using Radiometrically Calibrated Small-Footprint Full-Waveform Airborne LiDAR Data." *ISPRS Journal of Photogrammetry and Remote Sensing* 67: 134–147. doi:10.1016/j.isprsjprs.2011.12.003.

Höfle, B., W. Mücke, M. Dutter, M. Rutzinger, and P. Dorninger. (2009). "Detection of Building Regions Using Airborne LiDAR - A New Combination of Raster and Point Cloud Based GIS Methods." In *Geospatial Crossroads @ GI_Forum '09: Proceedings of the Geoinformatics Forum Salzburg: Geoinformatics on Stage*, edited by A. Car, G. Griesebner, and J. Strobl, 66–75. July 7–10, 2009. Heidelberg, Germany: Wichmann Verlag.

Honkavaara, E., R. Arbiol, L. Markelin, L. Martinez, M. Cramer, S. Bovet, L. Chandelier, et al. 2009. "Digital Airborne Photogrammetry—A New Tool for Quantitative Remote Sensing?—A State-Of-the-Art Review on Radiometric Aspects of Digital Photogrammetric Images." *Remote Sensing* 1: 577–605. doi:10.3390/rs1030577.

Hron, V., and L. Halounova. 2015. "Use of Aerial Images for Regular Updates of Buildings in the Fundamental Base of Geographic Data of the Czech Republic." *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-3/W2: 73–79. doi:10.5194/isprsarchives-XL-3-W2-73-2015.

Kang, Z., and Z. Lu, 2011. "The Change Detection of Building Models Using Epochs of Terrestrial Point Clouds." *Proceedings of 2011 IEEE International Workshop on Multi-Platform/Multi-Sensor Remote Sensing and Mapping (M2RSM)*, 1–6, Xiamen, China.

Kraus, K., and N. Pfeifer. 1998. "Determination of Terrain Models in Wooded Areas with Airborne Laser Scanner Data." *ISPRS Journal of Photogrammetry & Remote Sensing* 53: 193–203. doi:10.1016/S0924-2716(98)00009-4.

Lafarge, F., and C. Mallet. 2012. "Creating Large-Scale City Models from 3d-Point Clouds: A Robust Approach with Hybrid Representation." *International Journal of Computer Vision* 99: 69–85. doi:10.1007/s11263-012-0517-8.

Leberl, F., A. Irschara, T. Pock, P. Meixner, M. Gruber, S. Scholz, and A. Wiechert. 2010. "Point Clouds: Lidar versus 3D Vision." *Photogrammetric Engineering & Remote Sensing* 76 (10): 1123–1134. doi:10.14358/PERS.76.10.1123.

Lemmens, M. 2014. "Point Cloud (1) - the Functionalities of Processing Software." *GIM International* 28 (6): 16–21.

Liao, C. T., and H. H. Huang. 2012. "Classification by Using Multispectral Point Cloud Data." *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XXXIX-B3: 137–141. doi:10.5194/isprsarchives-XXXIX-B3-137-2012.

Mahmood, R., A. I. Quintanar, G. Conner, R. Leeper, S. Dobler, R. A. Pielke, A. Beltran-Przekurat, et al. 2010. "Impacts of Land Use/Land Cover Change on Climate and Future Research Priorities." *Bulletin of the American Meteorological Society* 91: 37–46. doi:10.1175/2009BAMS2769.1.

Mallet, C., F. Bretar, M. Roux, U. Soergel, and C. Heipke. 2011. "Relevance Assessment of Full-Waveform Lidar Data for Urban Area Classification." *ISPRS Journal of Photogrammetry and Remote Sensing* 66 (6): S71–S84. doi:10.1016/j.isprsjprs.2011.09.008.

Maltezos, E., and C. Ioannidis. 2015. "Automatic Detection of Building Points from Lidar and Dense Image Matching Point Clouds." *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* II-3/W5: 33–40. doi:10.5194/isprsannals-II-3-W5-33-2015.

Modiri, M., M. Masumi, and A. Eftekhari. 2015. "Automatic Classification of Point Clouds Extracted from Ultracam Stereo Images." *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-1-W5: 493–496. doi:10.5194/isprsarchives-XL-1-W5-493-2015.

Momm, H. G., G. Easson, and J. Kuszmaul. 2009. "Evaluation of the Use of Spectral and Textural Information by an Evolutionary Algorithm for Multi-Spectral Imagery Classification." *Computers, Environment and Urban Systems* 33: 463–471. doi:10.1016/j.compenvurbsys.2009.07.007.

Motohka, T., K. N. Nasahara, H. Oguma, and S. Tsuchida. 2010. "Applicability of Green-Red Vegetation Index for Remote Sensing of Vegetation Phenology." *Remote Sensing* 2 (10): 2369–2387. doi:10.3390/rs2102369.

Myint, S. W., and N. Lam. 2005. "A Study of Lacunarity-Based Texture Analysis Approaches to Improve Urban Image Classification." *Computers, Environment and Urban Systems* 29: 501–523. doi:10.1016/j.compenvurbsys.2005.01.007.

Najafi, M., S. Taghavi Namin, M. Salzmann, and L. Petersson, 2014. "Non-Associative Higher-Order Markov Networks for Point Cloud Classification." In Proceeding of Computer Vision – ECCV 2014: 13th European Conference, Zurich, Switzerland, 500–515, September 6–12, 2014, Part V. Cham: Springer International Publishing.

Nebiker, S., S. Bleisch, and M. Christen. 2010. "Rich Point Clouds in Virtual Globes – A New Paradigm in City Modeling?" *Computers, Environment and Urban Systems* 34: 508–517. doi:10.1016/j.compenvurbsys.2010.05.002.

Niemeyer, J., F. Rottensteiner, and U. Soergel, 2013. "Classification of Urban LiDAR Data Using Conditional Random Field and Random Forests." *Proceedings of the JURSE. IEEE, São Paulo*. Brazil, 139–142.

Niemeyer, J., F. Rottensteiner, U. Soergel, and C. Heipke. 2015. "Contextual Classification of Point Clouds Using a Two-Stage Crf." *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-3/W2: 141–148. doi:10.5194/isprsarchives-XL-3-W2-141-2015.

Omidalizarandi, M., and M. Saadatseresht. 2013. "Segmentation and Classification of Point Clouds from Dense Aerial Image Matching." *The International Journal of Multimedia & Its Applications* 5: 33–51. doi:10.5121/ijma.2013.5403.

Otepka, J., S. Ghuffar, C. Waldhauser, R. Hochreiter, and N. Pfeifer. 2013. "Georeferenced Point Clouds: A Survey of Features and Point Cloud Management." *ISPRS International Journal of Geo-Information* 2: 1038–1065. doi:10.3390/ijgi2041038.

Pang, S., X. Hu, Z. Wang, and Y. Lu. 2014. "Object-Based Analysis of Airborne LiDAR Data for Building Change Detection." *Remote Sensing* 6: 10733–10749. doi:10.3390/rs61110733.

Pfeifer, N., G. Mandlburger, J. Otepka, and W. Karel. 2014. "OPALS – A Framework for Airborne Laser Scanning Data Analysis." *Computers, Environment and Urban Systems* 45: 125–136. doi:10.1016/j.compenvurbsys.2013.11.002.

Pielke, R. A. 2005. "Land Use and Climate Change." *Science* 310: 1625–1626. doi:10.1126/science.1120529.

Rau, J.-Y., J.-P. Jhan, and Y.-C. Hsu. 2015. "Analysis of Oblique Aerial Images for Land Cover and Point Cloud Classification in an Urban Environment." *IEEE Transactions on Geoscience and Remote Sensing* 53: 1304–1319. doi:10.1109/TGRS.2014.2337658.

Ressl, C., H. Brockmann, G. Mandlburger, and N. Pfeifer. 2016. "Dense Image Matching Vs. Airborne Laser Scannin - Comparison of Two Methods for Deriving Terrain Models." *Photogrammetrie - Fernerkundung - Geoinformation* 2016: 57–73. doi:10.1127/pfg/2016/0288.

Rokach, L., and O. Maimon. 2008. *Data Mining with Decision Trees: Theroy and Applications*. Singapore: World Scientific Publishing Co.

Rothermel, M., K. Wenzel, D. Fritsch, and N. Haala. 2012. ""SURE Photogrammetric Surface Reconstruction from Imagery." In *Proceedings LC3D Workshop*, Berlin.

Rottensteiner, F., G. Sohn, J. Jung, M. Gerke, C. Baillard, S. Benitez, and U. Breitkopf, 2012. "The ISPRS Benchmark on Urban Object Classification and 3D Building Reconstruction." *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Melbourne, Australia. doi:10.5194/isprsannals-I-3-293-2012

Rottensteiner, F., J. Trinder, S. Clode, and K. Kubik. 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: 135–149. doi:10.1016/j.isprsjprs.2007.03.001.

Rutzinger, M., B. Höfle, M. Hollaus, and N. Pfeifer. 2008. "Object-Based Point Cloud Analysis of Full-Waveform Airborne Laser Scanning Data for Urban Vegetation Classification." *Sensors* 8: 4505–4528. doi:10.3390/s8084505.

Shapovalov, R., A. Velizhev, and O. Barinova, 2010. "Non-Associative Markov Networks for 3D Point Cloud Classification." In *Proceeding of IAPRS, Saint-Mande*, 103–108, France. http://www.isprs.org/proceedings/XXXVIII/4-W15/

Sithole, G., and G. Vosselman. 2006. "Bridge Detection in Airborne Laser Scanner Data." *ISPRS Journal of Photogrammetry and Remote Sensing* 61 (1): 33–46. doi:10.1016/j.isprsjprs.2006.07.004.

Van Leeuwen, M., N. C. Coops, and M. A. Wulder. 2010. "Canopy Surface Reconstruction from a LiDAR Point Cloud Using Hough Transform." *Remote Sensing Letters* 1: 125–132. doi:10.1080/01431161003649339.

Veneziano, D., S. Hallmark, and R. Souleyrette. 2002. *Comparison of Lidar and Conventional Mapping Methods for Highway Corridor Studies*. Santa Barbara, CA: Iowa State University, Center for Transportation Research and Education.

Vosselman, G. 2013. "Point Cloud Segmentation for Urban Scene Classification." *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-7/W2: 257–262. doi:10.5194/isprsarchives-XL-7-W2-257-2013.

Waldhauser, C., R. Hochreiter, J. Otepka, N. Pfeifer, S. Ghuffar, K. Korzeniowska, and G. Wagner, 2014. "Automated Classification of Airborne Laser Scanning Point Clouds." In Solving Computationally Expensive Engineering Problems. Springer International Publishing, 97: 269–292. doi:10.1007/978-3-319-08958-0_12

Wang, T., Q. Abdullah, D. Chavez, R. Kandukuri, N. Csanyi-May, and D. Simerlink, 2009. "The Quality of One-Foot Contour Modeling Using Lidar Data versus Photogrammetrically-Derived Solutions." In *proceedings of International Lidar Mapping Forum. Unpaginated CD-ROM*, New Orleans, Louisiana, January 9.

Wang, Y., H. Weinacker, and B. Koch. 2008. "A Lidar Point Cloud Based Procedure for Vertical Canopy Structure Analysis and 3D Single Tree Modelling in Forest." *Sensors* 8: 3938–3951. doi:10.3390/s8063938.

Xiong, X., D. Munoz, J. A. Bagnell, and M. Hebert, 2011. "3-D Scene Analysis via Sequenced Predictions over Points and Regions." In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA11)*, Shanghai, China, 2609–2616.

Yan, W. Y., A. Shaker, and N. El-Ashmawy. 2015. "Urban Land Cover Classification Using Airborne LiDAR Data: A Review." *Remote Sensing of Environment* 158: 295–310. doi:10.1016/j.rse.2014.11.001.

Youn, J., J. S. Bethel, E. M. Mikhail, and C. Lee. 2008. "Extracting Urban Road Networks from High-Resolution True Orthoimage and Lidar." *Photogrammetric Engineering and Remote Sensing* 74: 227–237. doi:10.14358/PERS.74.2.227.

ARTICLE III

Forests **2015**, *6*, 1613-1627; doi:10.3390/f6051613

OPEN ACCESS

forests

ISSN 1999-4907

www.mdpi.com/journal/forests

Article

Assessment of Wooded Area Reduction by Airborne Laser Scanning

Thi Huong Giang Tran [1,2,*], Markus Hollaus [1], Ba Duy Nguyen [1,2] and Norbert Pfeifer [1,*]

[1] Department of Geodesy and Geoinformation, Vienna University of Technology, Gußhausstraße 27-29, A-1040 Vienna, Austria;
E-Mails: Markus.Hollaus@geo.tuwien.ac.at (M.H.); nguyenbaduy@humg.edu.vn (B.D.N.)

[2] Faculty of Surveying and Mapping, Hanoi University of Mining and Geology, Hanoi 10000, Vietnam

* Authors to whom correspondence should be addressed;
E-Mails: tranthihuonggiang@humg.edu.vn (T.H.G.T.); Norbert.Pfeifer@geo.tuwien.ac.at (N.P.);
Tel./Fax: +43-1-58801-12219.

Academic Editors: Juha Hyyppä and Eric J. Jokela

Received: 20 January 2015 / Accepted: 27 April 2015 / Published: 7 May 2015

Abstract: Airborne Laser Scanning (ALS) data hold a great deal of promise in monitoring the reduction of single trees and forests with high accuracy. In the literature, the canopy height model (CHM) is the main input used frequently for forest change detection. ALS also has the key capability of delivering 3D point clouds, not only from the top canopy surface, but also from the entire canopy profile and also from the terrain. We investigated the use of two additional parameters, which exploit these capabilities for assessing the reduction of wooded area: Slope-adapted echo ratio (sER) and $Sigma_0$. In this study, two ALS point cloud data sets (2005 and 2011) were used to calculate Digital Surface Model (DSM), sER, and $Sigma_0$ in 1.5 km^2 forest area in Vorarlberg, Austria. Image differencing was applied to indicate the change in the three difference models individually and in their combinations. Decision trees were used to classify the area of removed trees with the minimum mapping unit of 13 m^2. The final results were evaluated by a knowledge-based manual digitization using completeness and correctness measures. The best result is achieved using the combination of sER and DSM, namely a correctness of 92% and a completeness of 85%.

Keywords: ALS; LiDAR; forest; change detection; single tree harvesting

1. Introduction

Forests are an important factor in maintaining the balance in the Earth system. However, the ecological processes are often affected by human activity [1]. In order to control the change of forests under the impacts of deforestation, wind throw, and diseases, it is required for forest managers to apply techniques supporting monitoring and updating forest information regularly. *In situ* forest inventory and remote sensing technologies are in use for detecting and monitoring these changes. Remote sensing, as one of these techniques, has proven its ability in change detection, automatically, efficiently, and consistently, especially for large areas, this also requiring less manual labor.

Optical remote sensing is a good choice to detect changes of forests as demonstrated in different studies [2–5]. However, optical acquisition techniques are limited by clouds if they are below the platform. Photograhic techniques additionally provide little information in cast shadow, as well as topographic shadow areas, providing little or no texture and, thus, lead to lower accuracy there. Furthermore, the varying shadow conditions for different acquisition times limit an automatic large area derivation of forest changes.

The image matching technique shows its potential use in forestry [6,7]. Photogrammetric imagery can also be exploited for gaining 3D point clouds [8]. Nevertheless, the limitation of this technique is that information is restricted to the point cloud of the upper canopy and, thus, does not provide ground height [9]. The image-based point cloud quality depends on factors like ground sampling distance, radiometric image resolution, stereo-parameters, viewing geometry, sun-angle, and amount of shadows [9].

With the advantage of penetrating the canopy through small gaps, Airborne Laser Scanner (ALS) is a potential technique for monitoring vegetation changes. Being an active technique, ALS emits its own energy for sensing and is consequently not affected by the ambient illumination (cast shadows, shadows of high clouds). Using 3D point clouds from ALS, the change in both coverage and height can be detected [10,11]. Moreover, not only dense forests, but even single trees can be detected from dense point clouds [12], which can be used to estimate forest biomass [13,14] and generate 3D tree models [15,16].

Although ALS data holds a high promise in vegetation change detection, thus far, research using multi-temporal ALS to detect forest change cover has not yet been fully explored. Yu *et al.* [17] detected harvested trees using two small footprint, high sampling density ALS acquisitions based on image differencing. Three-dimensional canopy height models (CHM) were calculated for both data sets using raster-based algorithms. The major change of CHM at the same pixel was acquired by a threshold value. They reported that 61 out of 83 field-checked harvested trees were detected automatically. The undetected trees were mainly smaller trees. St-Onge *et al.* [18] also used the threshold of CHM difference of two medium density LiDAR data acquired in 1998 and 2003 to identify new canopy gaps and assess height growth. With the same data Vepakomma *et al*. [19] expanded their study in accessing the feasibility of small footprint LiDAR to map the canopy gap expansions and canopy gap

closures for the conservation zone of Quebec. Vastaranta *et al*. [20] developed a ΔCHM method for canopy change detection of snow-damaged trees by applying bi-temporal LiDAR data for the period 2006–2010. Næsset and Gobakken [21] estimated a boreal forest growth over two years by using canopy metrics, *i.e.*, measures of spatial distribution of the acquired point cloud. Nyström *et al.* [22] employed histogram matching to calibrate the metrics in order to reduce the difference between two ALS datasets and produce change imagery. They controlled the changes by partial and complete tree removal in selected plots. Hollaus *et al.* [23] assessed the changes in growing stock of 160 km^2 mountain forest by two ALS datasets. The model was established with 184 FI (forest inventory) plots.

The above-mentioned studies used the differences found in (two) datasets over a forest area. Alternatively, processes in the forest can be detected using only one dataset by the traces they leave in the site. For example, Mücke *et al.* [24] used full-waveform ALS data, obtained under leaf-off conditions to detect fallen trees. Here, an echo width model was derived based on the point cloud and normalized echo heights in order to delineate downed stems. Lindberg *et al.* [25] and Nyström *et al.* [26] contributed two different methods highlighting the potential of high density ALS data to detect wind-thrown trees under forest canopy. Lindberg *et al*. [25] used a line template matching method applied directly to the ALS point cloud (69 points/m^2), while Nyström *et al*. [26] used the difference between two elevation models created from the same high density ALS data to detect wind-thrown trees.

In this research, we investigate the ability of forest reduction detection from two different ALS datasets by using image differencing [27]. In comparison to Nyström, who applied histogram matching to account for different sensor characteristics, our aim is to find features of LiDAR point clouds, that are, as much, as possible independent of the sensor characteristics. Unlike other studies mentioned above, our hypothesis is that forest reduction up to individual trees can be observed by the three LiDAR derived models: Digital Surface Model (DSM), Slope-adaptive Echo Ratio (sER), and "$Sigma_0$" (a local roughness measure), as well as their combinations. Our primary interest is, thus, not to demonstrate that ALS can detect forest changes as this was done before [17,22]. Rather, we are interested in finding robust methods in the presence of different ALS mission parameters. The study is done for a mountain forest in Vorarlberg, Austria. Two epochs were acquired with six years difference between the data acquisitions. The DSMs demonstrate the change in height and, thus, indicate that tall objects were removed. sER demonstrates the change in vertical penetrability and indicates that layered objects (e.g., understory and canopy) were removed. $Sigma_0$ demonstrates the change in the vertical dispersion of the points and indicates that objects distributed in height (e.g., trees, bushes) were removed. Using these variables, in different combination, the forest reduction is derived. The results are accessed with their accuracy by using the completeness and correctness measure. All the processes are supported by OPALS [28] and ArcGIS software.

2. Material and Method

2.1. Study Area

Study area is located in the south of Vorarlberg province, Austria, centered on lat. N 47°04′12″, long. E 9°49′12″ (Figure 1). The total covered area is about 1.5 km^2 of mountainous region, with the

elevations ranges from 1225 m above sea level (a.s.l) in the valleys, to a maximum of 1786 m a.s.l. In general, in this area, approx. 24% is covered by forest and the dominating tree species is Norway spruce (*Picea abies*). The forests in the study area are managed by the Stand Montafon Forstfonds (http://stand-montafon.at/forst), which operates the local forest inventory. Based on this inventory the tree heights vary between approximately 6 m to 42 m, with the mean height of 27.5 m and the standard deviation is 6.8 m [29]. Due to the topography the majority of the forests in the study area have a protection function against natural hazards, *i.e.*, snow avalanches. Therefore, exploitations of single trees or group of trees are foreseen in the forest management plan, meaning clear cuttings of larger areas are not allowed.

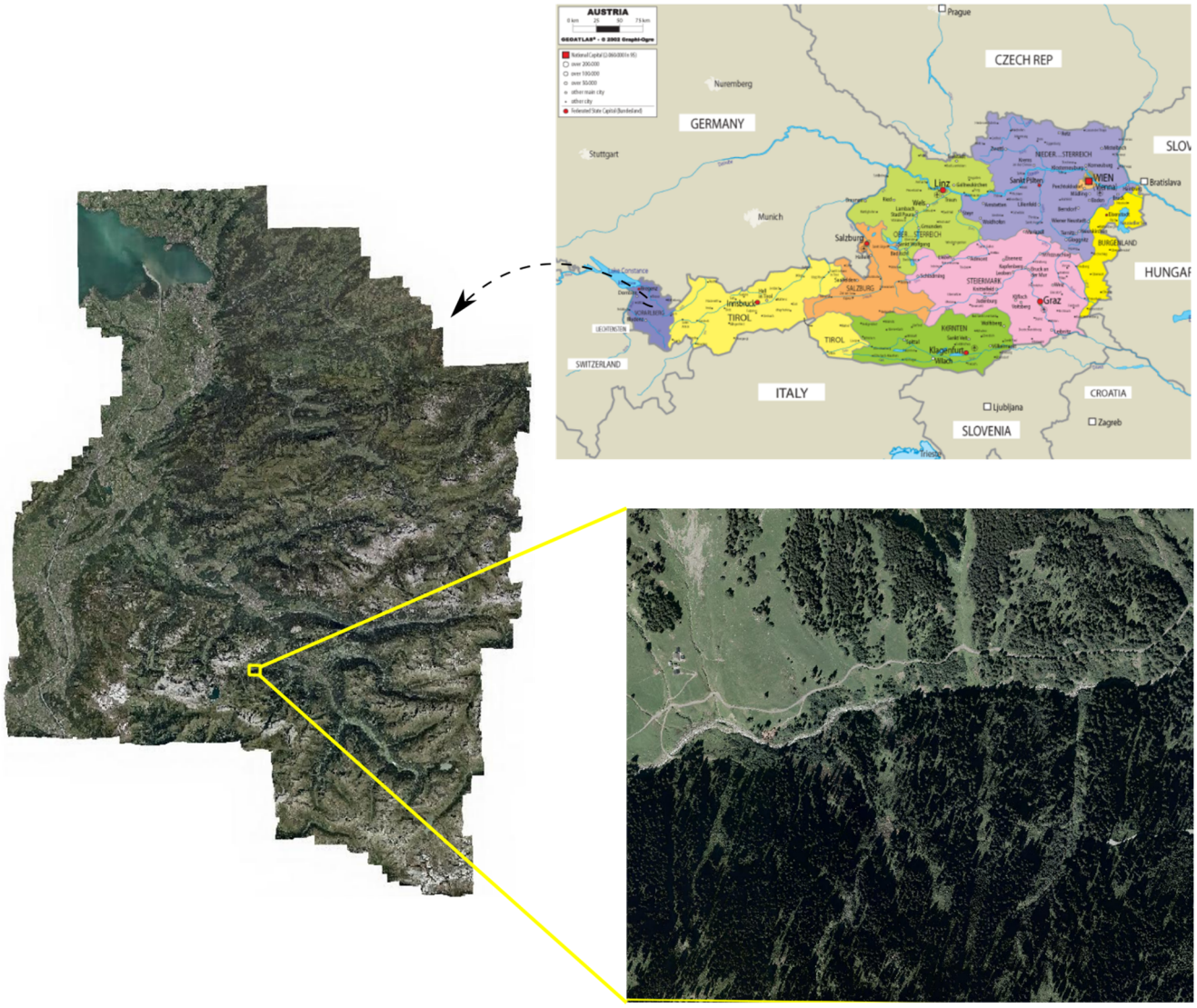

Figure 1. Study area (Orthophoto: office of survey and geoinformation from Vorarlberg, Austria, Political Map of Austria: GEOATLAS.com.).

2.2. Data

The two ALS data sets were provided by the local office of survey and geo-information from the federal state of Vorarlberg, and are subsets of the official federal state-wide ALS data acquisition

campaigns. The ALS data sets were acquired in 2005 and 2011 using an Airborne Laser Terrain Mapper systems (ALTM 1225) and a Trimble Harrier 56 system, respectively. The study area is covered by two flight strips for the first ALS acquisition and three flight strips for the second acquisition. All ALS data sets were acquired under snow-free and leaf-off conditions and were available as georeferenced 3D-point clouds. In Table 1, the relevant sensor characteristics of used ALS systems are summarized.

Table 1. Summary of sensor characteristics of the applied ALS systems.

Sensor Characteristics	Sensors	
	Acquisition Year 2005	**Acquisition Year 2011**
	Optech ALTM 1225	Trimble Harrier 56
Beam divergence	0.3 mrad	<0.5 mrad
Max scan angle (from nadir)	20°	30°
Wavelength	1064 nm	1064 nm
Pulse repetition frequency	<25 kHz	160 kHz
Sensor type	Discrete	Full waveform
Average point density	6 echos/m^2	24 echos/m^2

2.3. Reference Data

The reference data was derived, based on image interpretation of aerial orthophotos [30] with additional use of 3D point cloud viewing of the raw ALS data. Orthophotos with a spatial resolution of 0.12 m (2012) and 0.5 m (2005) are used to overview the forest cover status. Using additionally various visualizations of the ALS data, such as 3D point cloud visualization in FugroViewer software, nDSM values of the same pixels in two epochs, each single 1 m^2 pixel was evaluated and digitized. Due to border effects, a tree crown can be influenced in the DSMs by a slightly different shape. Therefore, small changes of very few m^2 were not considered as relevant and the therefore, the minimum mapping area was set to 13 m^2. This process took roughly 100 h. For the 223 digitized polygons, the minimum polygon area is 13 m^2 and the maximum area is 2351 m^2, the average size of a harvested polygon is 152 m^2 and the standard deviation is 289 m^2. The mean height of harvested polygons is 30.8 m and the standard deviation is 8.2 m. The final digitized output is converted into a binary format with a raster size of 1 × 1 m^2 that is used for accuracy assessment.

2.4. Method

In this study 3D ALS point clouds are used as the basis for deriving the following parameters, which are used for delineating harvested trees:

- DSM,
- sER, and
- $Sigma_0$.

For forest change detection the DSMs are an important input because they describe the height changes of the top most canopy surface. This means a decrease of the DSM indicates the loss of trees. To derive the DSMs from the two ALS data sets, the land cover dependent method described, in

Hollaus *et al*. [31], is applied. This method uses the strengths of different algorithms for generating the final DSM by using surface roughness information to combine two DSMs, which are calculated based on the highest echo within a raster cell, and on moving least squares (mls) interpolation with a plane as functional model (*i.e*., a tilted regression plane fitted through the k-nearest neighbors). For smooth areas (e.g., roof planes, streets, short grassland), noise reduction by moving least squares interpolation is exploited, whereas for rough surfaces (e.g., canopy surface, building edge) the highest point within a raster cell is used ($DSM(X, Y)_{max}$). The input for moving least squares interpolation is a subset of the point cloud (highest points within 0.5 m rater cells), which ensures that the interpolated surface goes through the tree tops. The derived final DSMs have a spatial resolution of 1 m.

More formally, the DSM calculation runs in the following way (Here we use the trinary operator c ? r_1: r_2. Its value depends on the condition c. If c is TRUE, the result is r_1, otherwise the result is r_2.).

$$z[DSM\ (X_i, Y_i)]^{2005} = z[\sigma_z(X_i, Y_i)]^{2005} < 0.5 \text{ or not } z[DSM\ (X_i, Y_i)_{max}]^{2005} \text{ ? } z[DSM\ (X_i, Y_i)_{mls}]^{2005} : z[DSM\ (X_i, Y_i)_{max}]^{2005}$$

$$z[DSM\ (X_j, Y_j)]^{2011} = z[\sigma_z(X_j, Y_j)]^{2011} < 0.5 \text{ or not } z[DSM\ (X_j, Y_j)_{max}]^{2011} \text{ ? } z[DSM\ (X_j, Y_j)_{mls}]^{2011} : z[DSM\ (X_j, Y_j)_{max}]^{2011}$$

$$\Delta DSM\ (X, Y) = DSM^{2011}\ (X, Y) - DSM^{2005}(X, Y) \quad (1)$$

The sER is a measure that describes the vertical point distribution and thus the penetrability of the surface [32,33]. The echo ratio (not slope-adaptive) is defined as the ratio between the number of neighboring echoes in a fixed search distance, measured in 3D (a sphere, n_{3D}, see appendix for exact definition), and all echoes located within the same search distance in 2D (a vertical cylinder, n_{2D}). To guarantee a correct ER on steep slopes, the search radius of the sphere (r_{2D}) has to be extended considering the slope ($r_{3D} = r_{2D}/\cos(\alpha)$) (*i.e*., dividing the initial 3D search distance by the cosine of the slope). Thus, the derived ER is the slope-adaptive echo ratio, sER.

$$sER = \frac{n_{3D}(r/\cos\alpha)}{n_{2D}(r)} \times 100 \quad (2)$$

The sER is computed for each echo in the first processing step and shows for continuous and impenetrable surface (*i.e*., ground and roof surfaces) values of 100% and for tree canopy points of lower value. For further analyses, the sER is aggregated in 1 m cells using the max value within each cell. The modules opalsEchoRatio and opalsCell were used for this computation.

$$\Delta sER\ (X, Y) = sER^{2011}(X, Y) - sER^{2005}(X, Y) \quad (3)$$

For the computation of $Sigma_0$, all echoes are used with a neighborhood size of the ten nearest neighbors. It is derived during the interpolation of a height model using the moving least squares approach. The standard deviation of the residuals in this interpolation is determined at each grid post. This provides a grid congruent with the interpolated heights. In each grid post, this value of $Sigma_0$ indicates how well all of the original points fit to the least squared plane. The grid width was 1 m.

$$\Delta Sigma_0\ (X, Y) = Sigma_0^{2011}(X, Y) - Sigma_0^{2005}(X, Y) \quad (4)$$

As the decreases of the elevation of the canopy surface indicate the loss of trees, the DSM^{2005} is subtracted from the DSM^{2011}. In addition to the elevation changes, the changes in the vertical echo distribution and penetrability described with the sER model and the changes of the surface roughness,

represented by the $Sigma_0$ model, also indicate areas with lost (e.g., harvested) trees. Thus, difference models of the sER and the $Sigma_0$ models are also calculated. To use this information, it is assumed that the sER of point clouds have a lower value for trees than for open ground (e.g., removed trees) and for $Sigma_0$ it receives a higher value.

The thresholds for all change values (*i.e.*, DSM, sER and $Sigma_0$) are assessed empirically. No primary maps of vegetated area were derived for the epochs. In other words, no maps of DSM, sER, and $Sigma_0$ were studied. Instead, we are searching directly for thresholds on the observed differences, thus, in ΔDSM, ΔsER, and $\Delta Sigma_0$.

All three data sets (ΔDSM, ΔsER, $\Delta Sigma_0$) indicate lost tree positions. Additionally, combinations of those input variables, for improving the accuracy of the final result, were investigated. Each pair and the triple of variables are used with newly determined empirical thresholds. The change results can be expected to show small errors, localized in single pixels or very small groups of pixels (*i.e.*, along the border of trees or forests). This is caused by the accuracy of the acquired data (*i.e.*, point density, georeferencing), as well as by the interpolation. The results are, therefore, converted into binary format for applying methods of binary mathematical morphology. Closing and then opening morphology with a circular kernel shape with a diameter of 1 and 2 pixels respectively are applied to all output raster datasets to reduce noise and to smooth object outlines. Finally, six change detection outputs were established: DSM only, sER only, $Sigma_0$ only, DSM combined with sER, sER combined with $Sigma_0$ and DSM, and sER and $Sigma_0$. Early in the analysis it became obvious, that the combination of DSM and $Sigma_0$ offers no increase of the achievable accuracy than the other pairs, thus, it was omitted.

Completeness (Comp) and correctness (Corr) [34] are used for the accuracy assessment of the final results.

$$\mathrm{Comp} = \frac{\|\mathrm{TP}\|}{\|\mathrm{TP}\| + \|\mathrm{FN}\|} \tag{5}$$

$$\mathrm{Corr} = \frac{\|\mathrm{TP}\|}{\|\mathrm{TP}\| + \|\mathrm{FP}\|} \tag{6}$$

The forest reduction area in the reference and change detection results are compared, where a true positive (TP) indicate the change in both datasets, false negative (FN) is labeled in the reference data but has no correspondence in the change detection results, and false positive (FP) is labeled in the change detection results and has no corresponding in the reference data.

3. Results and Discussion

The main properties of the primary models (DSM, sER and $Sigma_0$) are summarized in Table 2. Based on the empirical analyses, the thresholds for each of the different image (ΔDSM, ΔsER and $\Delta Sigma_0$ variables were found and summarized in Table 3. Figure 2 shows the harvested tree detection results of the three variables DSM, sER, and $Sigma_0$ independently, as well as the manual digitized reference map. As can be seen, the downed tree area was detected, more or less, correctly. However, the results are affected by salt-and-pepper type of noise. After applying image morphology operations, the $Sigma_0$ final results are still strongly affected by this and give the worst accuracy (Table 3) compared with the other two variables (Figure 3a–c).

Table 2. Minimum, maximum, mean, standard deviation of DSM, sER, and $Sigma_0$ in the different epochs.

	Min	Max	Mean	Std.dev
DSM^{2005}(m)	1226.2	1817.1	1516.4	124.0
DSM^{2011}(m)	1226.0	2022.1 [1]	1517.1	124.0
sER^{2005}(%)	1.9	100.0	69.9	29.3
sER^{2011}(%)	2.3	100.0	65.5	29.9
$Sigma_0^{2005}$(m)	0.0	28.0	2.9	4.2
$Sigma_0^{2011}$(m)	0.0	220.3 [2]	2.7	4.2

[1] This value is affected by gross errors. Excluding them leads to a max DSM of 1817.9; [2] Excluding gross error the max $Sigma_0^{2011}$ is 27.2.

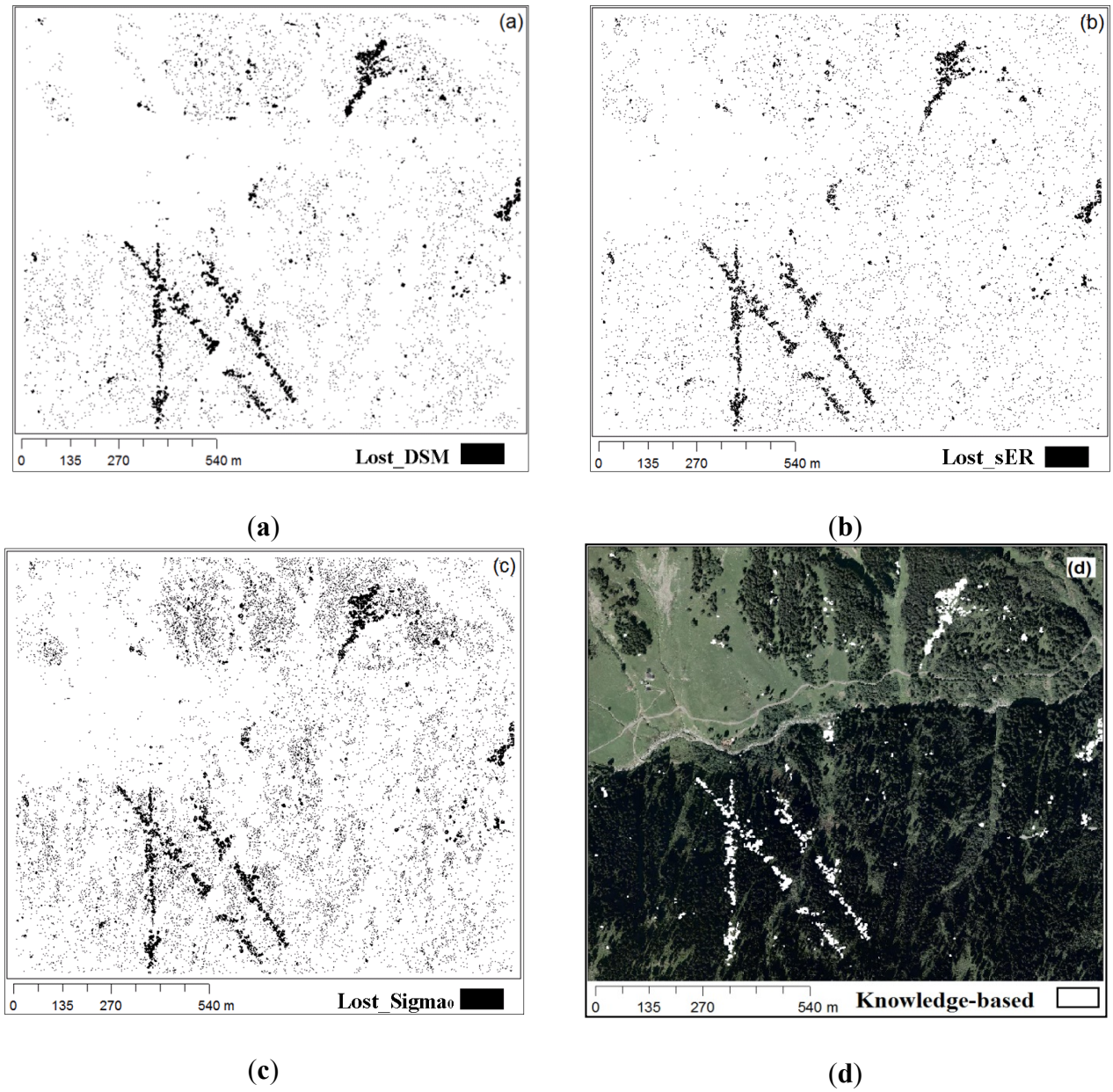

Figure 2. Forest reduction based on the selected thresholds for the (**a**) DSM; (**b**) sER; (**c**) $Sigma_0$; and (**d**) Knowledge-based digitization.

Table 3. Threshold values and accuracy measures.

Change threshold	DSM (m)	sER (%)	$Sigma_0$ (m)	Corr (%)	Comp (%)
ΔDSM	<−7.0	----	----	84.6	90.9
ΔsER	----	>30	----	87.5	87.1
Δ$Sigma_0$	----	----	<−7.0	38.6	56.8
ΔDSM and sER	<−2.0	>27	----	91.9	85.1
ΔsER and $Sigma_0$	----	>27	<−2.0	90.9	80.8
ΔDSM and sER and $Sigma_0$	<−2.0	>25	<−1.0	92.8	82.4
ΔDSM and sER and $Sigma_0$ [1]	<−7.0	>30	<−7.0	96.4	38.2

[1] The result of this combination is not shown in the Figure 3.

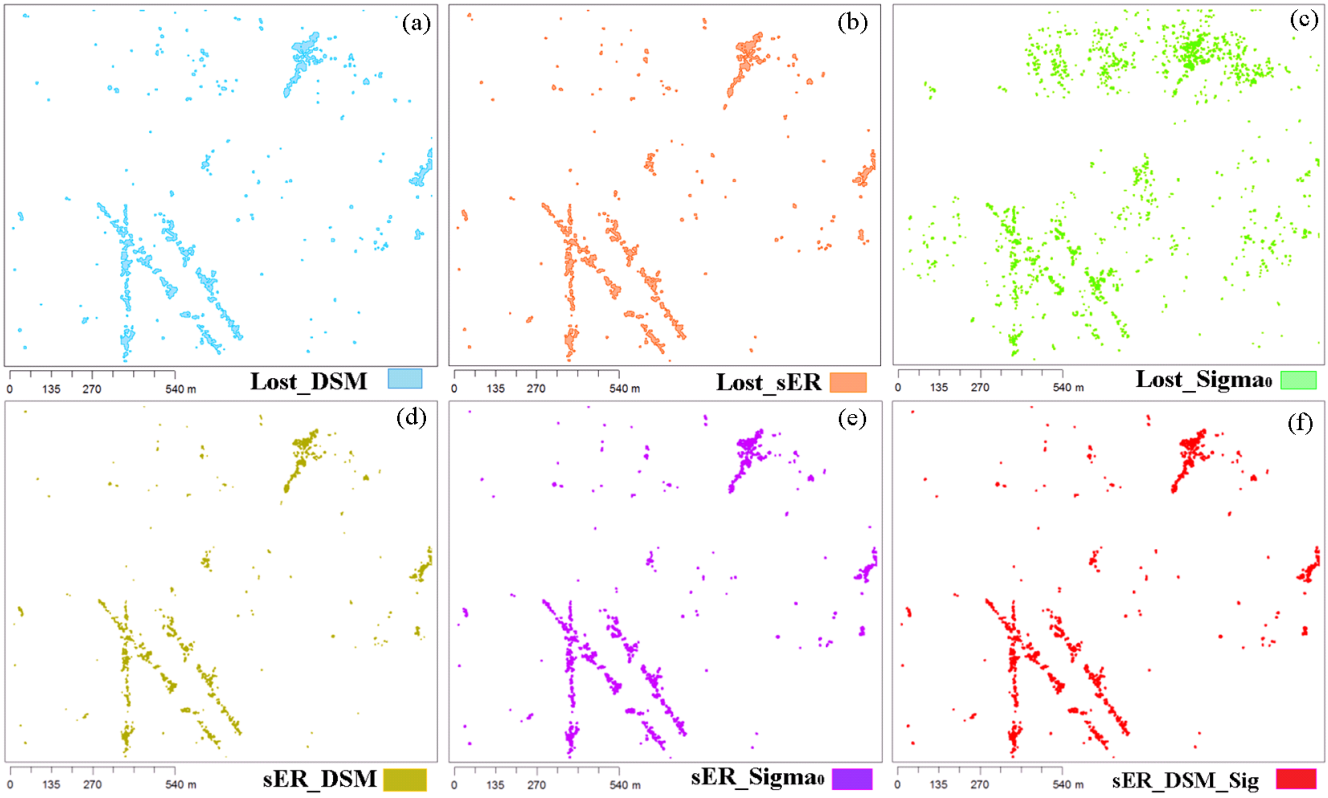

Figure 3. Final forest reduction after morphological operation (**a**) DSM; (**b**) Echo ratio; (**c**) $Sigma_0$; (**d**) Echo ratio and DSM; (**e**) Echo ratio and $Sigma_0$; (**f**) Echo ratio and DSM and $Sigma_0$.

The limitations of DSM and $Sigma_0$ compared to sER are to detect the reduction of low trees. Using a threshold, which is too low in height change, the lost tree cover is easy to be mixed with the unchanged forest cover. Additionally, $Sigma_0$ may depend much more on flying parameters, such as the flying height, which influences the point density and consequently the $Sigma_0$ values. Regarding to sER, because sER is larger influenced by its neighborhood so it has an increased value for the lost trees in a larger area. This leads to a reduction in the capability of detecting lost trees in dense canopy regions.

In order to overcome the limitation of each variable, all variables are incorporated in combinations in order to achieve improved results. To acquire the threshold of each pair combination between ΔsER *versus* ΔDSM, ΔsER *versus* Δ$Sigma_0$, and ΔDSM *versus* Δ$Sigma_0$, a feature space (Figure 4) is used to distinguish changed areas. The area of change is in either case in the upper left part of the feature

space. It deviates from the distribution of unchanged areas, which is centered on (0, 0). As stated above, the pair ΔDSM *versus* ΔSigma$_0$ does not show the discrimination in the feature space compared to other pairs, thus it is not used to detect changes. The median value of ΔDSM, ΔsER and ΔSigma$_0$ of the reference data also calculated and plotted into the feature space for delineating the threshold. From the feature space, the threshold for the ΔDSM and ΔsER combination (ΔDSM < −2 AND ΔsER > 27) and the ΔsER and ΔSigma$_0$ combination (ΔsER > 27 AND ΔSigma$_0$ < −2) are found, as a result, the accuracy of change detection is improved (Table 3, Figure 3). The former provides a correctness of 91.9% and completeness of 85.1% higher than the latter with a correctness of 90.9% and a Completeness of 80.8%. The combination of all three variables awards the highest correctness (92.8%) and a lower completeness (82.4%). As could be expected, using the original thresholds of the single variable classifications shows the highest correctness (96%) at the cost of a lower completeness (only 38.2%). The accuracy assessment of seven final change detection results is shown in Table 3.

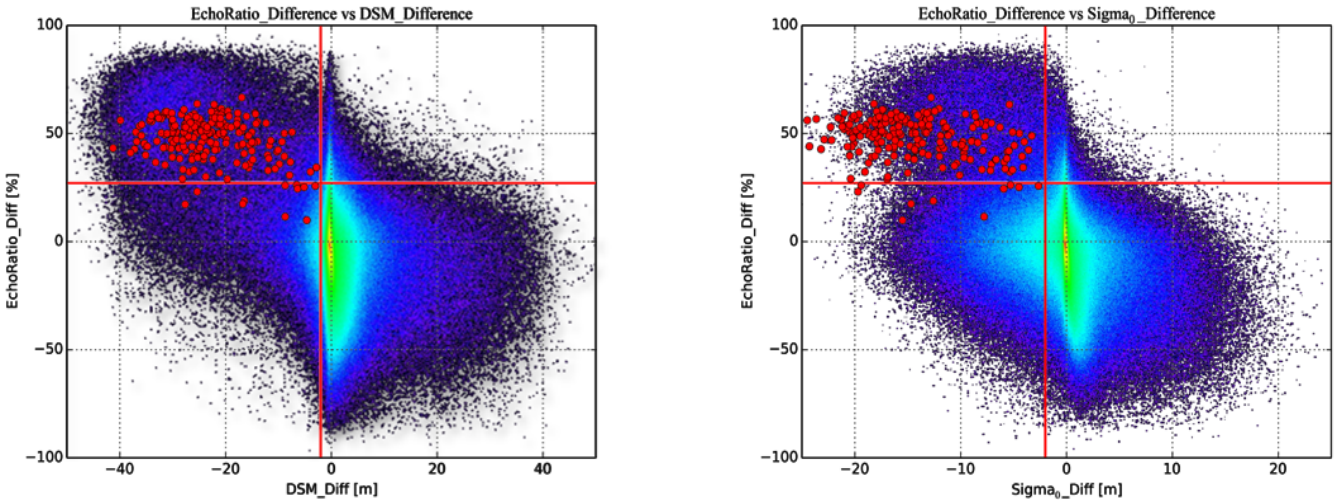

Figure 4. Feature space of (**a**) ΔsER *versus* ΔSigma$_0$; (**b**) ΔsER *versus* ΔDSM. The red lines indicate the threshold values for change detection. Brighter colors denote a higher density of points (from blue, to green, to red). Red dots indicated the median value of ΔDSM, ΔsER, and ΔSigma$_0$ of the reference polygons.

Our final result achieved a accuracy (with a minimum mapping unit of 13 m^2) compared to the research of Yu *et al*. [17], who detected 61 out of 83 harvested trees with an accuracy of 73.5%. St-Onge and Vepakomma [18] used canopy height differences to identify new gaps (especially fallen trees) and the minimum area criterion is 5 m^2. Their data has a density of 3 shots/m^2 in each epoch. Producer and user accuracies are similar to ours, although a bit higher (95%–98%). However, they apparently confirmed the existence of gaps in the reference (optical images) and the LiDAR result, but not the exact spatial location. Small edge localization errors and gap size differences between our reference and LiDAR result add to lower producer and user accuracy in our case. Additionally, the results of Nyström *et al.* [22] can be compared to our results. Their overall accuracy in detection is 88%, thus, somewhat lower than our result. However, in the forest-tundra ecotone, in which their study is set, the geometric signal of changes is lower than in ours in the case of taller trees.

Analyzing the $Sigma_0$ values of the different missions indicated that they depend more strongly on the parameters of the data acquisition. Two point clouds of a single tree are displayed and visualized in 3D, and it is realized that the point distributions and point densities of the same tree are different. This will influence the value of the $Sigma_0$ results. Therefore, in this study, surface roughness ($Sigma_0$) alone is not a reasonable measure to detect forest change.

In Figure 4, there are some sER median values for reference polygons located under the selected threshold. It is explained that, in some dense tree positions, the sER values of downed trees are influenced by points of the other surrounding trees.

In the case of deciduous forests with a dense canopy surface and fully developed foliage, there is no penetration into the canopy and to the ground. It leads to high sER and small $Sigma_0$ values for trees. Harvested trees can be better detected with ΔDSM in this case. On the other hand, errors in vertical geo-referencing directly influence the DSM and, therefore, ΔDSM, but it does not affect sER and $Sigma_0$. Therefore, the combination of sER and DSM will provide the highest quality in detecting the reduction of wooded area.

Scan frequency, flying altitude, scan angle, acquisition time (*i.e.*, leaf-on, leaf-off), and applied methods for preprocessing have an influence on data quality [35,36]. Based on the applied method for DSM calculation, the influences of these properties are minimized [31].

For a detailed assessment of forest biomass changes it is important to differentiate between forest growth and exploitation. For the quantification of the exploitation, detailed information about reduced (*i.e.*, harvested trees) forest area is required. The assessment of forest growth is based on changes in the DSM, which requires robust methods to derive DSM from the ALS point clouds that are, as much as possible, independent from sensor characteristics and data acquisition settings [31]. To minimize influences originating from georeferencing issues, advanced georeferencing, including a strip adjustment, is normally required [23]. For the assessment of the biomass change, regression models can be applied to the individual data sets. For the calibration of these regression models [37], reference data, e.g., forest inventory data, are required. Finally, the biomass change can be derived from the differences between the individual biomass maps.

4. Conclusions

Using the image differencing method, a traditional pixel-based change detection method was applied to detect reduction of forest area. In this study, we used the three variables DSM, slope adaptive echo ratio (sER), and $Sigma_0$, derived from two different ALS data sets, to detect downed trees in a forest. While many studies have, thus far, used DSM and its change, we found that sER is a good single predictor for tree cover change. sER is a local measurement, which means that global height differences, e.g., ALS block geo-referencing problems, do not influence the assessment. It is also noted that the threshold value for sER (Table 2) did not change as much as for DSM and $Sigma_0$ in the combined classifications.

The incorporations of two or more variables always improved the quality of detection results, only the combination of DSM and $Sigma_0$ does not provide an improvement. This study opens up a new application of discrete return ALS data in forest change detection and, therefore, in forest management.

We conclude that the best results were achieved using sER change with additional consideration of DSM change, namely a correctness of 92% and a completeness of 85%.

Acknowledgments

This work is partly founded from the European Community's Seventh Framework Programme (FP7/2007–2013) under grant agreement No. 606971, the Advanced_SAR project.

Author Contributions

Giang Tran processed the data, performed the experiment, and drafted the article. Duy Nguyen analyzed the data. Markus Hollaus reviewed and edited the manuscript. Norbert Pfeifer supervised the research, reviewed and edited the manuscript.

Appendix

For each point of the point set *P*, we computed the number of neighbors $n_{3D,i}$ in a 3D spherical neighborhood and the number of neighbors $n_{2D,i}$ in a cylindrical neighborhood.

$$p_i \in P, \quad i = 1, \dots, n, \quad p_i \in R^3, \quad p_i = (x_i, y_i, z_i)$$

$$n_{3D,i} = \left|\{q_j \in P | \left\|p_i - q_j\right\| \leq r\}\right|,$$

$$n_{2D,i} = \left|\left\{q_j \in P | \left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2 \leq r^2\right\}\right|, j = 1, \dots, n$$

$$n_{3D,i}(r), n_{2D,i}(r)$$

$$sER = \frac{n_{3D}(r/\cos\alpha)}{n_{2D}(r)} \times 100$$

Conflicts of Interest

The authors declare no conflict of interest.

References

1. FAO. *State of the World's Forests*; Food Agriculture Organization of the Uniterd States: Rome, Italy, 2012.
2. Broich, M.; Hansen, M.; Stolle, F.; Potapov, P.; Margono, B.A.; Adusei, B. Remotely sensed forest cover loss shows high spatial and temporal variation across sumatera and kalimantan, indonesia 2000–2008. *Environ. Res. Lett.* **2011**, *6*, 014010, doi:10.1088/1748-9326/6/1/014010.
3. Donoghue, D.N.M.; Watt, P.J.; Cox, N.J.; Dunford, R.W.; Wilson, J.; Stables, S.; Smith, S. An evaluation of the use of satellite data for monitoring early development of young sitka spruce plantation forest growth. *Forestry* **2004**, *77*, 383–396.
4. Huang, X.; Friedl, M.A. Distance metric-based forest cover change detection using modis time series. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *29*, 78–92.

5. Potapov, P.; Hansen, M.C.; Stehman, S.V.; Loveland, T.R.; Pittman, K. Combining modis and landsat imagery to estimate and map boreal forest cover loss. *Remote Sens. Environ.* **2008**, *112*, 3708–3719.
6. Baltsavias, E.; Gruen, A.; Eisenbeiss, H.; Zhang, L.; Waser, L.T. High-quality image matching and automated generation of 3d tree models. *Int. J. Remote Sens.* **2008**, *29*, 1243–1259.
7. Bohlin, J.; Wallerman, J.; Fransson, J.E.S. Forest variable estimation using photogrammetric matching of digital aerial images in combination with a high-resolution dem. *Scand. J. For. Res.* **2012**, *27*, 692–699.
8. Otepka, J.; Ghuffar, S.; Waldhauser, C.; Hochreiter, R.; Pfeifer, N. Georeferenced point clouds: A survey of features and point cloud management. *ISPRS Int. J. Geo-Inf.* **2013**, *2*, 1038–1065.
9. White, J.; Wulder, M.; Vastaranta, M.; Coops, N.; Pitt, D.; Woods, M. The utility of image-based point clouds for forest inventory: A comparison with airborne laser scanning. *Forests* **2013**, *4*, 518–536.
10. Houldcroft, C.J.; Campbell, C.L.; Davenport, I.J.; Gurney, R.J.; Holden, N. Measurement of canopy geometry characteristics using lidar laser altimetry. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 2270–2282.
11. Yu, X.; Hyyppä, J.; Kukko, A.; Maltamo, M.; Kaartinen, H. Change detection techniques for canopy height growth measurements using airborne laser scanner data. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 1339–1348.
12. Lindberg, E.; Eysn, L.; Hollaus, M.; Holmgren, J.; Pfeifer, N. Delineation of tree crowns and tree species classi fi cation from full-waveform airborne laser scanning data using 3-d ellipsoidal clustering. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 3174–3181.
13. Jochem, A.; Hollaus, M.; Rutzinger, M.; Hofle, B. Estimation of aboveground biomass in alpine forests: A semi-empirical approach considering canopy transparency derived from airborne lidar data. *Sensors* **2011**, *11*, 278–295.
14. Kim, S.-R.; Kwak, D.-A.; Lee, W.-K.; Son, Y.; Bae, S.-W.; Kim, C.; Yoo, S. Estimation of carbon storage based on individual tree detection in pinus densiflora stands using a fusion of aerial photography and lidar data. *Sci. China Life Sci.* **2010**, *53*, 885–897.
15. Vosselman, G. 3D recontruction of roads and trees for city modelling. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2003**, *34*, 231–236.
16. Rutzinger, M.; Pratihast, A.K.; Oude Elberink, S.J.; Vosselman, G. Tree modelling from mobile laser scanning data-sets. *Photogramm. Rec.* **2011**, *26*, 361–372.
17. Yu, X.; Hyyppä, J.; Kaartinen, H.; Maltamo, M. Automatic detection of harvested trees and determination of forest growth using airborne laser scanning. *Remote Sens. Environ.* **2004**, *90*, 451–462.
18. St-onge, B.; Vepakomma, U. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. In *Assessing Forest Gap Dynamics and Growth Using Multi-Temporal Laser-Scanner Data*; ISPRS Archives: Freiburg, Germany, 2004; pp. 173–178.
19. Vepakomma, U.; St-Onge, B.; Kneeshaw, D. Spatially explicit characterization of boreal forest gap dynamics using multi-temporal lidar data. *Remote Sens. Environ.* **2008**, *112*, 2326–2340.

20. Vastaranta, M.; Korpela, I.; Uotila, A.; Hovi, A.; Holopainen, M. Mapping of snow-damaged trees based on bitemporal airborne lidar data. *Eur. J. For. Res.* **2012**, *131*, 1217–1228.
21. Nasset, E.; Gobakken, T. Estimating forest growth using canopy metrics derived from airborne laser scanner data. *Remote Sens. Environ.* **2005**, *96*, 453–465.
22. Nyström, M.; Holmgren, J.; Olsson, H. Change detection of mountain birch using multi-temporal als point clouds. *Remote Sens. Lett.* **2013**, *4*, 190–199.
23. Hollaus, M.; Eysn, L.; Karel, W.; Pfeifer, N. Growing stock change estimation using airborne laser scanning data. In Proceedings of the Talk: 13th International Conference on LiDAR Applications for Assessing Forest Ecosystems (SilviLaser 2013), Beijing, China, 9–11 October 2013.
24. Mücke, W.; Hollaus, M.; Pfeifer, N.; Schroiff, A.; Deák, B. Comparison of discrete and full-waveform als for dead wood detection. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2013**, *II-5/W2*, 199–204.
25. Lindberg, E.; Hollaus, M.; Mücke, W.; Fransson, J.E.S.; Pfeifer, N. Detection of lying tree stems from airborne laser scanning data using a line template matching algorithm. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2013**, *II-5/W2*, 169–174.
26. Nyström, M.; Holmgren, J.; Fransson, J.E.S.; Olsson, H. Detection of windthrown trees using airborne laser scanning. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *30*, 21–29.
27. Hussain, M.; Chen, D.; Cheng, A.; Wei, H.; Stanley, D. Change detection from remotely sensed images: From pixel-based to object-based approaches. *ISPRS J. Photogramm. Remote Sens.* **2013**, *80*, 91–106.
28. Pfeifer, N.; Mandlburger, G.; Otepka, J.; Karel, W. Opals—A framework for airborne laser scanning data analysis. *Comput. Environ. Urban Syst.* **2014**, *45*, 125–136.
29. Hollaus, M.; Wagner, W.; Eberhöfer, C.; Karel, W. Accuracy of large-scale canopy heights derived from lidar data under operational constraints in a complex alpine environment. *ISPRS J. Photogramm. Remote Sens.* **2006**, *60*, 323–338.
30. Lillesand, T.M.; Kiefer, R.W.; Chipman, J.W. *Remote Sensing and Image Interpretation*; John Wiley & Sons: Hoboken, NJ, USA, 2008.
31. Hollaus, M.; Mandlburger, G.; Pfeifer, N. Land Cover Dependent Derivation of Digital Surface Models from Airborne Laser Scanning Data; IAPRS: Saint-Mandé, France, 2010; pp. 221–226.
32. Höfle, B.; Mücke, W.; Dutter, M.; Rutzinger, M. International Conference on Applied Geoinformatics. In Detection of Building Regions Using Airborne Lidar—A New Combination of Raster and Point Cloud Based Gis Methods Study Area and Datasets; GI-Forum: Washington, DC, USA, 2009; pp. 66–75.
33. Rutzinger, M.; Höfle, B.; Hollaus, M.; Pfeifer, N. Object-based point cloud analysis of full-waveform airborne laser scanning data for urban vegetation classification. *Sensors* **2008**, *8*, 4505–4528.
34. Heipke, C.; Mayer, H.; Wiedemann, C. 3D Recontruction and Modeling of Topographic Objects. In *Evaluation of Automatic Road Extraction*; IAPRS: Stuttgart, Germany, 1997; pp. 151–160.
35. Hodgson, M.E.; Jensen, J.R.; Schmidt, L.; Schill, S.; Davis, B. An evaluation of lidar- and ifsar-derived elevation models in leaf-on conditions with usgs level 1 and level 2 dems. *Remote Sens. Environ.* **2003**, *84*, 295–308.

36. Friedrich, A. Airborne laser scanning—Present status and future expectations. *ISPRS J. Photogramm. Remote Sens.* **1999**, *54*, 64–67.
37. Hollaus, M.; Wagner, W.; Schadauer, K.; Maier, B.; Gabler, K. Growing stock estimation for alpine forests in austria: A robust lidar-based approach. *Can. J. For. Res.* **2009**, *39*, 1387–1400.

ARTICLE IV

sensors

MDPI

Article

Integrated Change Detection and Classification in Urban Areas Based on Airborne Laser Scanning Point Clouds

Thi Huong Giang Tran [1,2,*], Camillo Ressl [1] and Norbert Pfeifer [1]

[1] Department of Geodesy and Geoinformation, Technische Universität Wien, Gußhausstraße 27-29, 1040 Vienna, Austria; camillo.ressl@geo.tuwien.ac.at (C.R.); norbert.pfeifer@geo.tuwien.ac.at (N.P.)

[2] Faculty of Geomatics and Land Administration, Hanoi University of Mining and Geology, Hanoi 10000, Vietnam

* Correspondence: tranthihuonggiang@humg.edu.vn; Tel.: +43-680-225-8679

Received: 19 December 2017; Accepted: 31 January 2018; Published: 3 February 2018

Abstract: This paper suggests a new approach for change detection (CD) in 3D point clouds. It combines classification and CD in one step using machine learning. The point cloud data of both epochs are merged for computing features of four types: features describing the point distribution, a feature relating to relative terrain elevation, features specific for the multi-target capability of laser scanning, and features combining the point clouds of both epochs to identify the change. All these features are merged in the points and then training samples are acquired to create the model for supervised classification, which is then applied to the whole study area. The final results reach an overall accuracy of over 90% for both epochs of eight classes: lost tree, new tree, lost building, new building, changed ground, unchanged building, unchanged tree, and unchanged ground.

Keywords: LiDAR; change classification; machine learning

1. Introduction

Change detection plays an important role in keeping topographic databases up-to-date, in monitoring, and in planning [1]. One major data source for documenting landscape change are 2D satellite images, especially in large-scale problems as urbanization, forest monitoring, earthquake hazard and risk assessment, etc. [2–7]. For these tasks, many studies used low-to-medium resolution images [8–12], although high resolution images were also employed for change detection at a higher level [13–15]. However, high resolution 2D-based change detection has several limitations such as higher spectral variability [6], perspective distortion [16,17], and lack of volumetric information [18,19]. These limitations complicate 2D-based change detection. With three dimension (3D) geometric information, 3D change detection is not influenced by illumination, perspective distortion and illumination variations as 2D change detection [20]. The third dimension as a supplementary data source (height, full 3D information, or depth) and the achievable outcome (height differences, volumetric change) expand the scope of CD applications [20] in 3D city model updating [21,22], 3D structure and construction monitoring [23,24], 3D object tracking [25,26], tree growth monitoring and biomass estimation [27,28], and landslide surveillance [29].

An established source of detailed and accurate 3D information is airborne LiDAR (light detection and ranging), which provides a point cloud, and is applied in various fields [30–33]. Therefore, airborne LiDAR is creating new possibilities for 3D change detection, especially in urban areas where complex 3D situations prevail [34].

Many approaches suggested in the literature demonstrate the high potential of LiDAR point clouds for change detection (see Section 2). Most studies apply two steps: first, detect the change;

and, second, classify the change; alternatively, first, classify objects for both periods; and, second, detect changes between the classifications. Both approaches, consequently, will be influenced by sequence error, i.e., the accuracy of classified changes depends on the change detection method and the classification method. Furthermore, most of those studies focus only on one class (typically either building or trees). However, a change typically does not happen in a single class only, but leads to changes in multiple classes. We are therefore suggesting to investigate the possibilities of performing change detection and classification of all the main classes (building, tree, and ground) simultaneously in one step.

In this paper, we suggest a new approach in change detection (CD). It combines classification and CD in one step. Additionally, it builds on the point cloud, which is a common data source for high resolution geoinformation from laser scanning and image matching alike. It exploits the power of machine learning [35]. Given two raw point clouds of different epochs, sampled training data is required only once. The method provides a separation whether there is a change or no change at the location of the point as well as individual class information for each point. The method is presented for one site, and its properties are discussed.

2. Related Work

We suggest to classify change detection approaches using ALS data in urban areas into two methods: "post-classification" CD and "pre-classification" CD. In the first category, post-classification, the urban objects are first classified into specific classes and then changes are detected in the classifications. In the second category, pre-classification, the differences between two datasets are detected first and then the change types are classified later. An overview of published approaches is shown in Table 1. The description will only focus on those articles, in which a specifically new aspect was added to the overall approach.

Table 1. List of proposed change detection approaches.

Authors	Year	Data Used			CD Approach	CD Classes
		ALS	Image	Maps		
Matikainen et al. [36]	2004	X	X	X	Post-classification	Building
Matikainen et al. [37]	2010	X	X	X	Post-classification	Building
Stal et al. [38]	2013	X	X		Post-classification	Building
Malpica et al. [39]	2013	X	X		Post-classification	Building
Matikainen et al. [40]	2016	X	X	X	Post-classification	Building
Matikainen et al. [41]	2017	X	X	X	Post-classification	Building, roads
Vosselman et al. [42]	2004	X		X	Post-classification	Building
Tang et al. [43]	2015	X		X	Post-classification	Building
Awrangjeb et al. [44]	2015	X		X	Post-classification	Building
Choi et al. [45]	2009	X			Post-classification	Ground, vegetation, building
Xu et al. [46]	2015b	X			Post-classification	Building
Teo et al. [47]	2013	X			Post-classification/DSM-based	Building
Murakami et al. [48]	1999	X			Pre-classification/DSM-based	Building
Pang et al. [49]	2014	X			Pre-classification/DSM-based	Building
Vu et al. [50]	2004	X			Pre-classification/DSM-based	Building
Zhang et al. [51]	2014	X			Pre-classification	Ground
Xu et al. [34,46]	2015a	X			Pre-classification	Building, tree

In post-classification CD, ALS data can either be combined with other datasets from a different epoch, or two different epochs of airborne LiDAR data are used. The former is a common method, investigated in many studies, incorporating the advantages of the LiDAR height data with either images [36–38] or existing maps for updating information [42–44]. Malpica et al. [39] proposed an approach that employed ALS data and satellite imagery for updating buildings of a geospatial vector database. LiDAR data were used to derive the height above the terrain, which was associated with spectral information and became the input for a support vector machine (SVM) classification. This method proved useful for tall buildings, yet small houses and low buildings surrounded by trees were not well-detected. Teo and Shih [47] suggested a CD method, in which the change in building types were obtained by handling multi-temporal interpolated LiDAR data. Recently,

Matikanen et al. [40,41] demonstrated the potential of multispectral airborne laser in automated classification and change detection. Land cover classification was derived from multispectral ALS data using a random forest classifier. Afterwards, building changes were detected by combination of the land cover classification results with a digital surface model (DSM) and building vectors from a previous epoch. Road changes were detected by comparing road classes from the classification results with the road centerline vectors. The approaches mentioned above enable detecting changes in 2.5D (DSMs) or only in 2D (Maps), both of which may cause loss of information under trees. In contrast, two ALS data epochs facilitate overcoming this issue. Choi et al. [45] based change detection on a DSM subtraction between two ALS epochs to detect change areas. The points within the detected areas were then organized into surface patches, which were subsequently classified into ground, vegetation, and building. The type of the change was determined based on the classes and properties of each patch. Xu et al. [46] detected the changes in buildings from commercial (acquired in the years 2008 and 2010), and residential area (2010 and 2012) by two epoch ALS data. Their "scene classification" used a rule-based classifier combined with the point-to-plane distance between two epochs to distinguish "changed", "unchanged", and "unknown". Then, changed points were re-classified in a second step into different classes (dormers, roofs, constructions on top roof, cars, and undefined objects) with an accuracy in the range of 80% to 90%. They showed that the quality of the classification results will influence the quality of the change detection.

In the literature, 3D change detection using two ALS epochs is more often investigated in the pre-classification scenario. First change is detected and then it is classified. DSM-based methods were employed in most studies. Murakami et al. [48] operated two multi-temporal ALS data to identify changes in buildings by subtracting the DSMs. Likewise, Vu et al. [50] demonstrated an automatic change detection method to detect damaged buildings after an earthquake in Japan. Pang et al. [49] proposed an object-based analysis method to automatically detect building changes by multi-temporal point cloud data in an 8.5 km^2 area. Going beyond DSM-based methods, Zhang and Glennie [51] presented a weighted anisotropic iterative closest point (ICP) algorithm, which determined 3D displacements between two point clouds by iteratively minimizing the sum of the squares of the distances. Xu et al. [34] proposed a three-step point-based method for identifying building and tree changes from two LiDAR datasets. The point cloud data were first registered using the ICP algorithm and filtered to extract non-ground points. Then, the non-ground points were stored and indexed in an octree. At last, the buildings and trees that changed were detected by comparing the two LiDAR point clouds and applying the AutoClust algorithm [52].

The aforementioned pre-classification CD studies [34,47–51] have illustrated the possibility of automatic change detection, which can achieve over 80% of accuracy in urban area. Most methods, however, depend on the DSM quality [38] and are concentrated on building changes. Xu et al. [34] has overcome the limitations of previous studies by proposing a method which does not require a DSM and expands the change types to tree cover change in urban area. However, their method has limitations in detecting of the natural growth of trees, which was classified into newly planted trees. Besides, ground points needed to be filtered out in their study. Of course, also ground can change through time and relevant change information should be supplied. In addition to change types, all the methods have the same process: firstly, to separate the "changed" and "unchanged" points, and, afterwards, classify the change types based on the "changed" detection.

Machine learning can be performed supervised, using training data, or unsupervised, with the aim of clustering points with similar features. As the relevant classes in the urban scene are known, we use supervised learning. If only two classes shall be distinguished, Support Vector Machines [53] could be well used. For point cloud classification this is described by [54–57]. For multiple class problems, Random Forests was suggested [35]. They are efficient and require a moderate amount of training data. Its application to point cloud classification is described, e.g., by [58,59]. Conditional Random Fields (CRF) [60] allow adding context to classification, i.e., considering the relation between neighboring points, and especially were shown to improve the results for classes with fewer points [61,62].

Convolutional neural networks (CNN) were also described for point clouds [63], but they require an extensive set of training data, which is not available in our case. Dictionary learning methods require less training samples but need a long time for processing [64–66]. Thus, the method of random forests for point cloud classification was selected.

The new automatic change detection method we suggest targets changes in and between the classes buildings, trees, and ground. A "Changed" and "Unchanged" separation does not need to be performed. Instead, all change types and unchanged types are detected simultaneously based on machine-learning [67].

3. Methodology

The proposed method is shown in Figure 1. First, outliers are removed from the data (see Section 3.1). Second, the data of both epochs are merged to compute features of four types: features describing the point distribution, a feature related to height above the terrain, features specific for the multi-target capability of ALS, and features combining both epochs to identify the change (Section 3.2). Training data (Section 3.3) are taken manually, and machine learning (Section 3.4) is applied to compute a model for the classification. Finally, based on the additional attributes of each point, change types are computed (see Figure 1). As shown in Figure 1, each point cloud is classified and investigated for change by an individual machine learning step. All processing is performed in OPALS [68] supported by DTMaster [69] and FugroViewer [70].

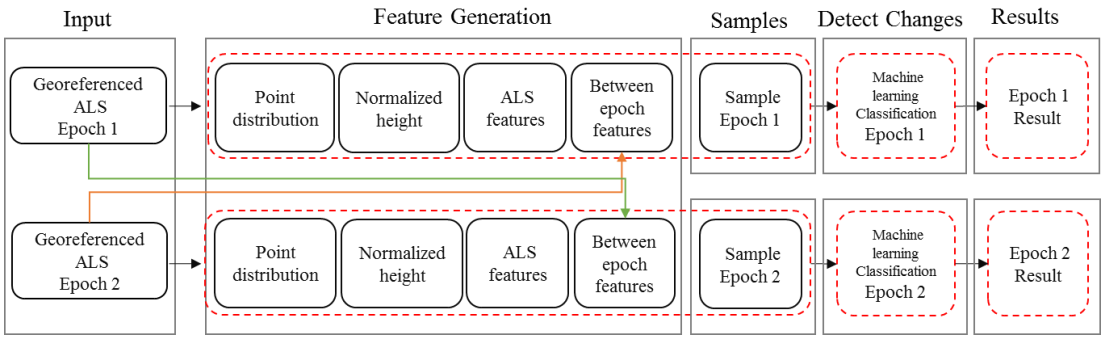

Figure 1. Flowchart of the proposed method.

3.1. Outlier Detection

We assume that the original data are georeferenced already in the same coordinate system and projection. Outlier points, which would cause unpredictable errors in the results, need to be eliminated. Statistics (min, max, mean, and standard deviation) and robust estimators (median and sigMad) of z value of the points are computed to set thresholds for outlier detection. Additionally, isolated points are detected and removed. An isolated point is defined as having no neighbors within a certain search radius.

3.2. Features

The first set of features describes the point distribution [71]. These features are required for the separability of the classes. Estimation of local planes on a point basis is useful for different tasks (e.g., shaded relief) and surface normals are important geometric properties of a surface. Here, the local tangent plane is estimated by computing the best fitting plane for the ten nearest points. Its normal vector (termed NormalX, NormalY, and NormalZ in the following) and the standard deviation (std.dev.) of the fit are used as additional descriptions of the points (termed NormalSigma). The distribution of the points in the neighborhood, which contain more helpful information, are derived from the structure tensor T [72]. Linearity, planarity, and omnivariance are three features obtained from T. The linearity feature (L_T) is used to characterize 3D line objects such as power lines

or similar structures. Planarity (P_T) is a feature which describes the smoothness of the surface and is related to roughness measures. Omnivariance (O_T) describes volumetric point distributions as they occur for trees. These features are computed using three eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$ of the matrix T (Equations (1)–(3)).

$$L_T = 1 - \frac{\lambda_2}{\lambda_1} \tag{1}$$

$$P_T = \frac{\lambda_2 - \lambda_3}{\lambda_1} \tag{2}$$

$$O_T = \sqrt[3]{\lambda_1 \lambda_2 \lambda_3} \tag{3}$$

Different neighborhood definitions are used for the attribute computation of the features EchoRatio, ZRank, and ZRange, which can be derived to provide more information of the points. The EchoRatio is a measure that describes the vertical point distribution and thus the penetrability of the surface [31,73,74]. ZRange represents the maximum height difference between the points in the neighborhood, while ZRank is the rank of the point corresponding to its height in the neighborhood. Thus, the full list of features of the first group is: NormalX, NormalY, NormalZ, NormalSigma, L_T, P_T, O_T, EchoRatio, ZRank, and ZRange.

Secondly, the normalized height is considered as a feature. Mallet et al. [54] have shown that classification of urban areas improves if this feature is considered. However, as we are primarily interested in change detection, the quality of the terrain model is expected to have a lower impact, and thus a simple method [75] is deemed sufficient to compute the DTM, if it is not already available. We use a hierarchic block minimum method (with two iterations). In the first iteration, all the last echo points are selected first. From these points, a raster model is derived by using the "5th lowest" height points in each cell of size 10 m. The height difference of a point and this raster model (nH = z (point) − z (raster)) is then calculated for each point and all the points in a threshold range above or below the cell elevation are filtered out. For the remaining points the same process (raster creation and nH computation) is repeated using smaller cell size and a smaller nH threshold range in order to obtain the final points for DTM interpolation.

The third set of ALS features exploits the multi-target capability of pulsed LiDAR systems, which can return multiple echoes per emitted laser shot. These echoes are measured directly and the point clouds from ALS not only contain the coordinates (x, y, z) but also further echo info: echo number within the shot, and number of echoes of the shot. Both values are used as features of the point.

Finally, the fourth set of features are features between epochs. They are computed for each point by considering the distribution of the neighboring points in the other epoch. In Figure 2, the points of the epoch 1 point cloud E1 are investigated for change relative to the point cloud E2. In each point of E1, we search in 3D to find the number n_{3D} of neighboring points of E2 within a sphere of radius R. If this number is zero, there is most likely a change at that point. This is just enough for detecting changes at building and isolated trees, but not for a dense tree area or trees close to buildings. For example, the right red tree in Figure 2 appears in epoch 1 but not in epoch 2. Most of the points in the tree are detected as changed. Nevertheless, this lost tree is close to another unchanged tree, so in the same search radius, some of the lost tree points are still unchanged points because they can find the nearest neighbor of E2 in the unchanged tree. This will be reduced if we consider also a 2D search around the same point to find the number n_{2D} within a cylinder of radius R. Finally, the ratio of these point numbers in percent is called "stability" (Equation (4)). This is identical to EchoRatio, with the difference that the point of evaluation if from a different point set than the points counted in the 2D and 3D neighborhood.

$$Stability = \frac{n_{3D} \times 100}{n_{2D}} \tag{4}$$

where n_{3D} is the number of points found in a fixed search distance (e.g., 1 m) measured in 3D (i.e., search sphere); and n_{2D} is number of points found in the same distance measured in 2D (i.e., vertical search cylinder with infinite height).

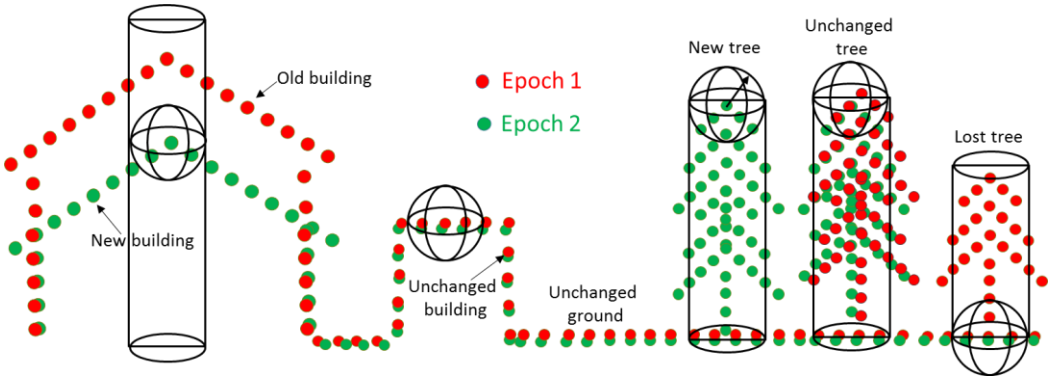

Figure 2. Stability of changed and unchanged objects.

3.3. Training Sample

The training samples provide information to the learning system. The supervised learning algorithm analyzes the training data and produces an inferred function, which can be used for mapping the remaining data. The training sample quality directly influences the classification results, as label noise, i.e., a wrong class decision on a training point, influences the learned function for the classification. All the changed types are taken thoroughly by manual selection. In this study, the change samples follow the rules shown in Table 2. It is not necessary to foresee changes in all classes and in this experiment the class water was not investigated for change.

Table 2. Rules of taking sample for machine learning classification.

Change Objects	Change Types	Description
Buildings	Unchanged high-building	The same high-building is in both epochs
	Unchanged low-building	The same low-building is in both epochs
	New high-building	New building with height >15 m
	Lost high-building	Lost building with height >15 m
	New low-building	New building with height ≤15 m
	Lost low-building	Lost building with height ≤15 m
	New walls	Walls in new building
	Lost walls	Walls in lost building
	Unchanged walls	Walls in unchanged building
Trees	New tree	New planted tree
	Lost tree	Cut tree
	Unchanged trees	The same tree in both periods
Ground	Unchanged ground	The same ground or absolute height differences ≤0.5 m
	Change in height	Ground has absolute height differences >0.5 m
	New ground	Buildings changed to grounds
	Lost ground	Ground changed to buildings
Water	Water	Water points

3.4. Change Types Classification

As a state-of-the-art machine learning algorithm, random forest [35] is used to classify the urban area because of its advantages. It does not overfit, runs fast and efficiently for a large dataset such as LiDAR [58] and it requires a moderate amount of training data. This method is useful for automatic classification of urban objects. All the sample points contain the four sets of features (mentioned in Section 3.2). Random forest selects randomly features for subsets of the sample points to train several

decision trees. Each randomly created decision tree is used to predict the class of a new (unseen) point based on its features and stores this outcome. The highest voted predicted outcome is considered the final classification for each point. The final classification model is then applied to the rest of the point cloud to generate the final change detection classification results.

4. Study Site

The Leopoldstadt District, in Vienna, Austria, is taken as the study area. The experimental region (Figure 3), covering an area of about 3 km × 1.5 km, is generally flat. It contains complex objects, containing several old-fashioned and modern high-rise buildings, a suburban area with mainly single-dwellings, an open-wide area (including a stadium), water, overpasses, an amusement park, and a variety of other objects. Since 2005, this area has been one of the most dynamic areas with respect to changes in land use in Vienna. Old buildings have been rebuilt into new buildings or open ground, new buildings are constructed from bare ground and cut trees, new trees are planted suitable for the landscape, and a new road and a stadium construction was built. All these led to changes in buildings, vegetation, and ground in this area.

Two sets of LiDAR data are available, which were obtained in 2007 (from 7 December 2006 to 5 February 2007) and 2015 (9–24 November 2015). These data have average point densities of 12 points/m^2 measured with a Riegl LMS_Q560 (Riegl, Horn, Austria) and 16 points/m^2 measured with a Riegl LMS_Q680i, respectively. As the datasets are already registered well, no extra steps for registration were required. Ortho-photos from the time of flight were not available, and thus Google earth images of the respective years were used to support interpretation of the main objects. This was necessary for taking training samples for machine learning, as well as a manual classification of the point cloud for the accuracy assessment process at the end.

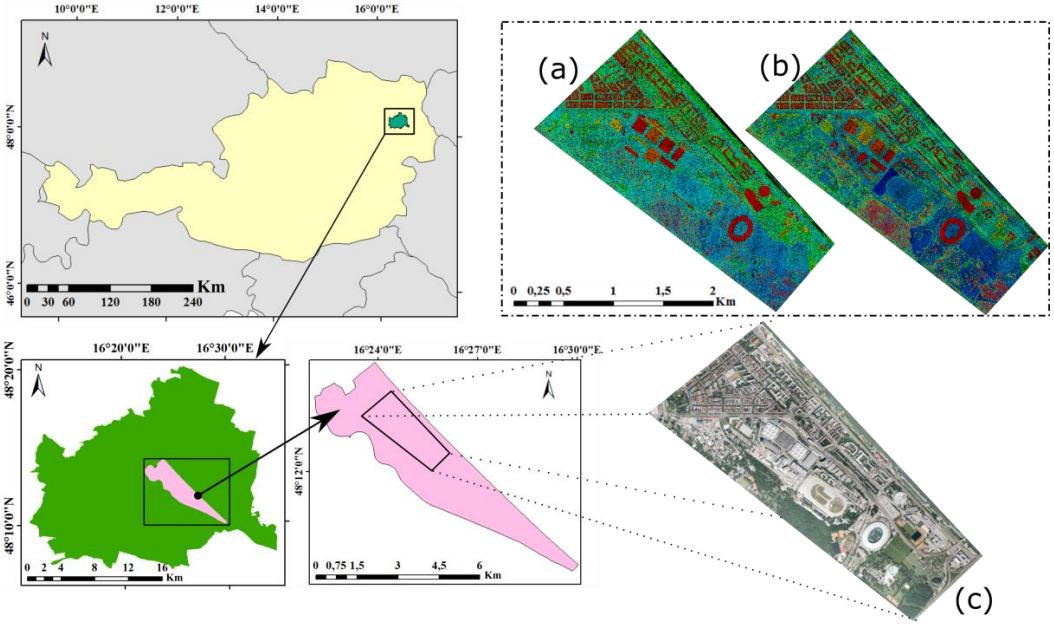

Figure 3. Experiment area: (**a**) LiDAR data for 2007; (**b**) LiDAR data for 2015; and (**c**) Google Earth image of experiment area.

5. Results and Discussion

The highlight of our change detection method is the combination of the steps of change detection and change classification in one single step based on the stability value combined with the other attributes in order to classify all objects into different classes, comprising: unchanged points (ground, building, and tree), and changed points (new building, lost buildings, new tree, lost tree, ground

changed in height, and ground changed into other objects). The final results are then evaluated using a point-based accuracy evaluation by comparing the automatic change detection classification results with the manual change detection classification.

5.1. Stability Feature

Stability (Equation (4)) is a feature which is used to detect change points in this paper. A good estimate for the selection of a suitable search radius is the double of the average point spacing found in the study area. This guarantees a representative number of neighbors, while avoiding too large neighborhoods, which would cause expanded transition zones at the border of two objects with different surface structure [73]. A search radius of 1.0 m is chosen in this paper. In flat open terrain this will result in around 38 neighboring points for 2007 and around 50 points for 2015. If no points of E2 are found by 3D search, the value of stability is 0%. That point is then a changed point. In the case of unchanged points, buildings and ground have low transparency, the number of 3D and 2D neighbors of E2 should be approximately the same, so resulting in a high stability (100%). In contrast, vegetation is transparent to LiDAR shots (to some degree) and thus the laser point density on and within vegetation depends on the density of the branches, the twigs and the leaves. The latter even depends on the vegetation type and the time of year. Consequently, one has to expect lower stability values at vegetation objects.

Figure 4 presents a rasterized image of the stability value for each of both datasets. From these images, it can be seen that the changed and unchanged regions are detected. Figure 4a,b shows the stability value ranges from 0 to 100% for the whole study area in epoch 2007 and 2015, respectively. To be perfectly clear in detail, a small specific area is zoomed in and indicated in height value (Figure 4c,d). Figure 4e,f indicates the stability value of this area. Changed buildings and grounds obtain a very low value (approx. 0%).

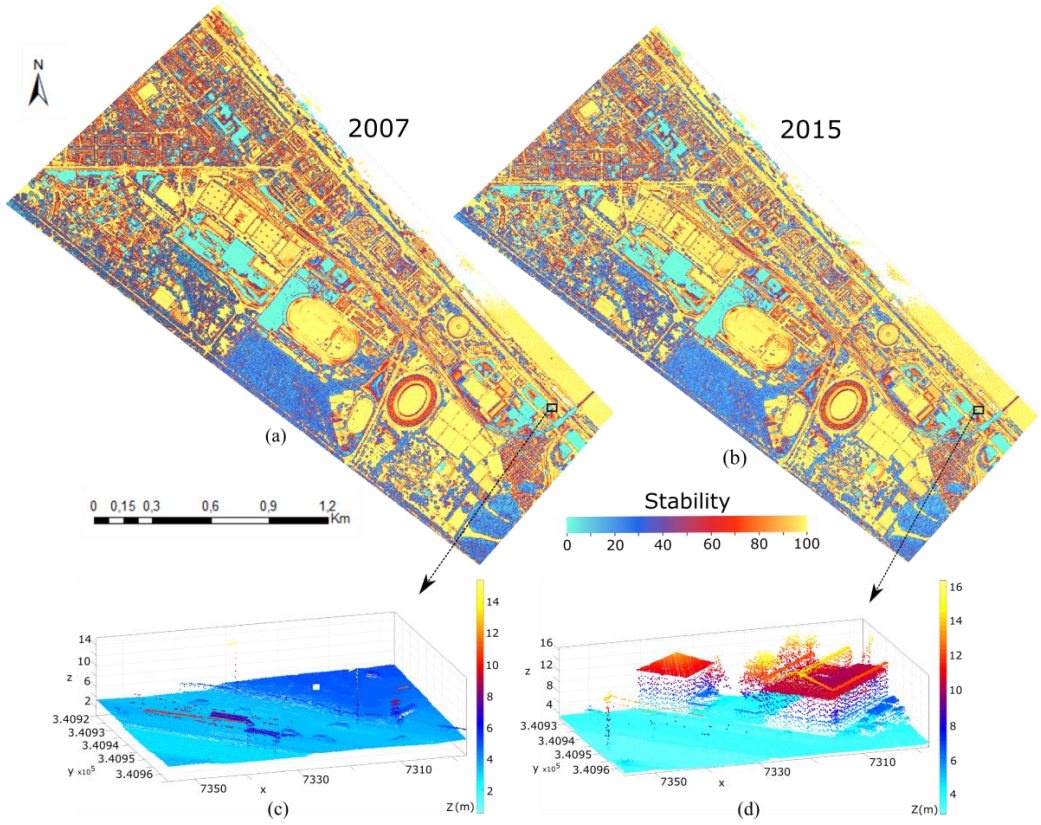

Figure 4. *Cont.*

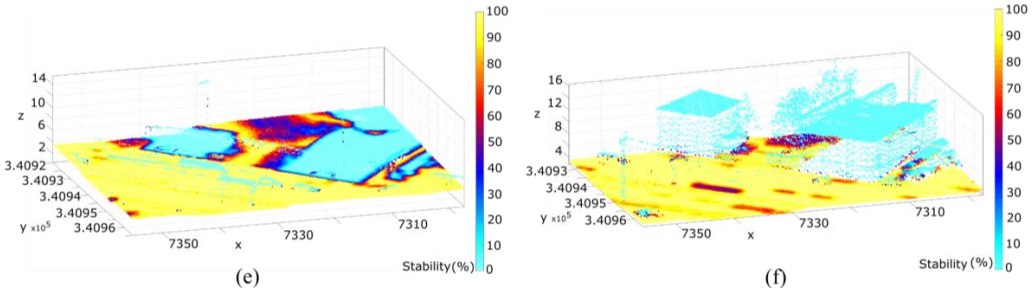

Figure 4. Stability feature for change detection: (**a**) stability shown for 2007 as color value, reaching from 0 (cyan) to 100% (yellow); (**b**) stability shown for 2015; (**c**) point cloud of 2007 shown as height value in an specific area; (**d**) point cloud of 2015 shown as height value in an specific area; (**e**) stability value range in specific area of 2007; and (**f**) stability value range in specific area of 2015.

5.2. Sampling Training Data and Classification

Because of the large area, the sample regions were taken from six small regions, where the changes took place actively. The samples of the unchanged objects (i.e., unchanged grounds, unchanged buildings, and unchanged trees) and grounds changed in height were taken simultaneously for both datasets. "Lost tree" and "Lost buildings" samples are only taken in 2007, whereas "New trees" and "New buildings" are only taken in 2015. The training samples were taken in DTMaster software. Table 3 sums up the number of sample points in each class. Seventy percent of the training data is used for learning, whereas the remaining 30% is used for evaluating of the learned random forest classification model (but not for the overall evaluation—different reference data are used for that; see Figure 5). Figure 5 displays the sample distribution in both datasets over the whole experiment area. The data for overall evaluation are overlaid with a yellow cross (see Section 5.4).

Table 3. Sample points of different change types in 2007 and 2015 datasets.

Change Types	Sample Points 2007	Sample Points 2015
Unchanged grounds	698,323	639,465
Unchanged low buildings	181,022	169,015
Unchanged high buildings	443,891	463,812
Unchanged walls	44,504	43,796
Lost walls	9341	-
New walls	-	62,795
New high building	-	479,565
Lost high building	65,653	-
New low building	-	53,219
Lost low building	189,327	-
Lost tree	193,035	-
New tree	-	138,402
Unchanged trees	184,781	515,326
Ground change in height	113,662	85,766
New ground	-	51,919
Lost ground	373,161	-
Water	2400	40,703

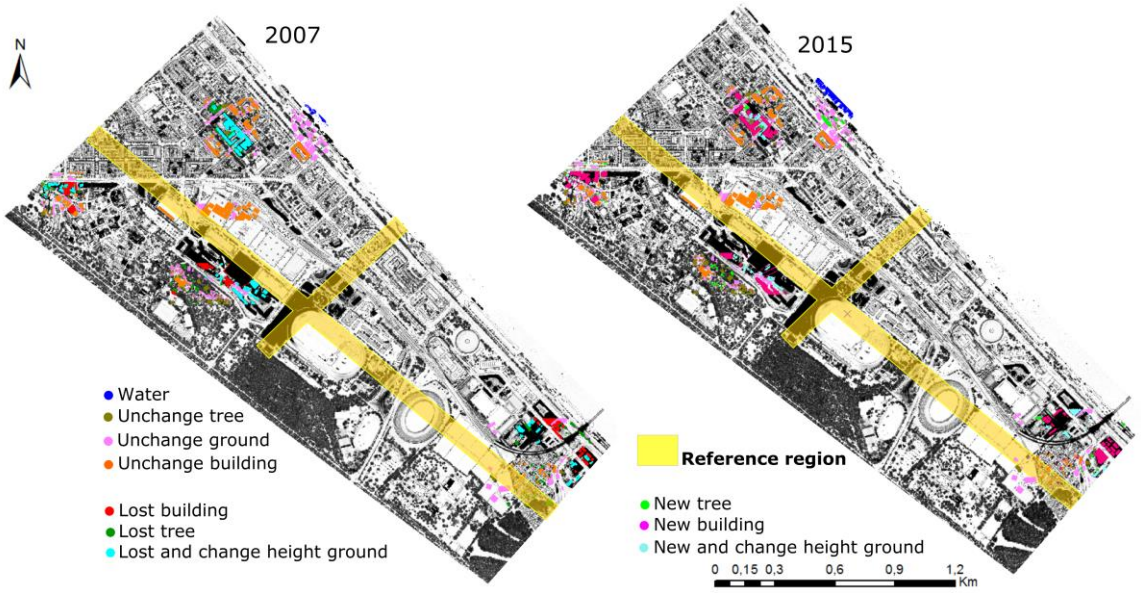

Figure 5. Sample points distribution in both datasets. The area that was used as reference for the overall evaluation, i.e., the accuracy assessment of the classification results, is shown in yellow.

After taking the sample points, they are used for training and creating the classification model for each dataset (one model for 2007 dataset, and one model for 2015 dataset). The models were then applied to the whole area to obtain the final change detection classification results in both datasets (Figure 6). The total number of processed points are 97,809,515 and 117,734,603 points in 2007 and 2015 datasets, respectively. The time for creating the models from the samples and applying the models to the total points in two datasets took 1:41 h for 2007 and 2:49 h for 2015 on a single PC with a Windows 7 Enterprise system (AMD FX ™-6100 Six-Core Processor, 16G RAM) (Singer Computer, Vienna, Austria).

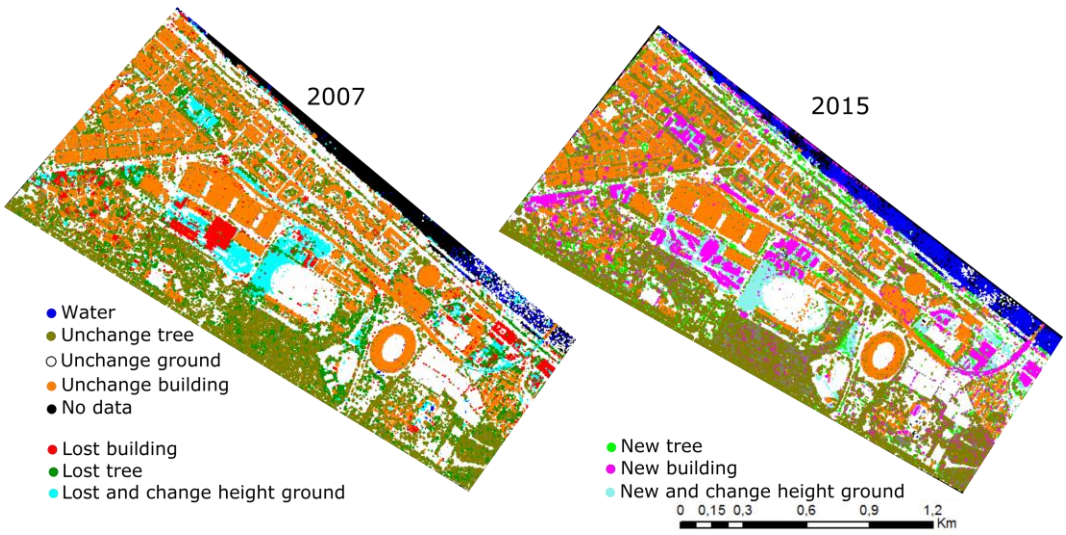

Figure 6. Change detection classification results of 2007 and 2015 datasets.

As can be seen in Figure 6, the results of the proposed method are satisfactory, thus indicating that the method is effective for the complex urban area. All the unchanged and changed objects were detected simultaneously. A visual comparison of both results in Figure 6 shows that the changes in 2007 and 2015 correspond nicely to one another; i.e., where there is a change in 2007 (with respect

to 2015) change in 2015 (with respect to 2007) also appears. The same holds true for the unchanged objects. Figure 7 shows in detail the change type classification results. The first and the second column show the objects in the data 2007 and 2015. Points are colored by elevation blue to red. The third column shows the change detection and classification results of the change types.

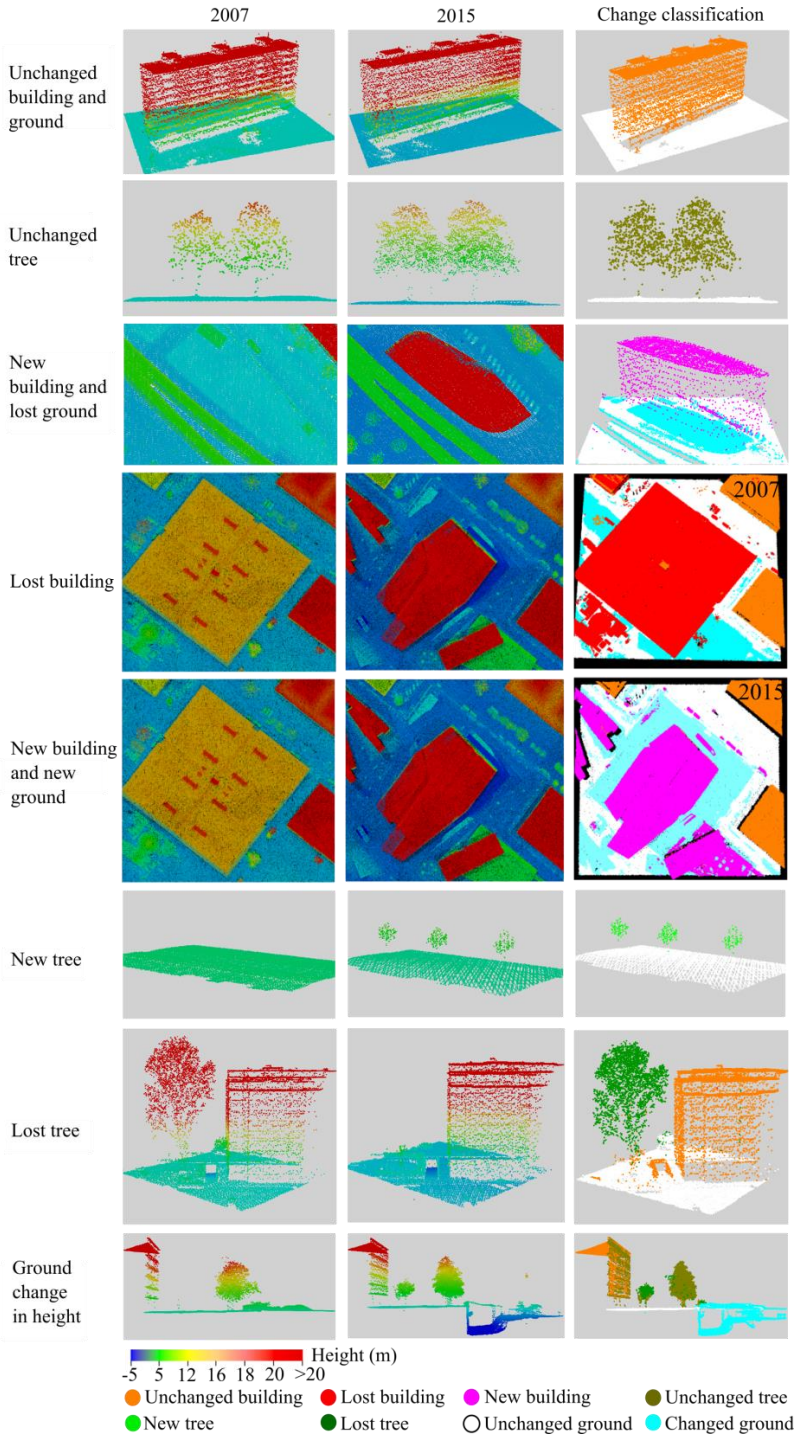

Figure 7. Example of change types of classification results. The first two columns show point clouds from 2007 and 2015, respectively, with color indicating height (legend below images). In the right column both point clouds, 2007 and 2015, are shown with their class label, with exception of the building in row 4 and 5, where the point clouds are shown separately for each year.

5.3. Impact of Using the Raw Point Cloud

The data used are several years apart. Such a time difference is suitable for new constructions. Vegetation objects may change considerably in the long period because of their natural growth. Additionally, each dataset itself contained certain problems, apart from outliers removed beforehand. Because of the duration of the measurement campaign, in the active change areas also changes within the overlap of 2007 LiDAR strips were found. It contained different objects (e.g., difference in ground) at the same position. Figure 8 shows a ground height difference of 4.7 m at the same location. This violates the underlying assumption of a stable object within one epoch and leads to a wrong classification in the ground of the 2007 dataset. In the 2015 dataset, because of a building wall material acting as a mirror, trees and grounds are wrongly recorded inside the buildings (Figure 9). Those wrong measurements could also not be discovered as noise in the outlier removal step. These problems were identified when collecting the reference data by the manual operator. Although all wrong points are removed as high point in the accuracy assessment step (see below), they have an impact on the final results because they influence the real points in the step of calculating attributes which are used for the classification.

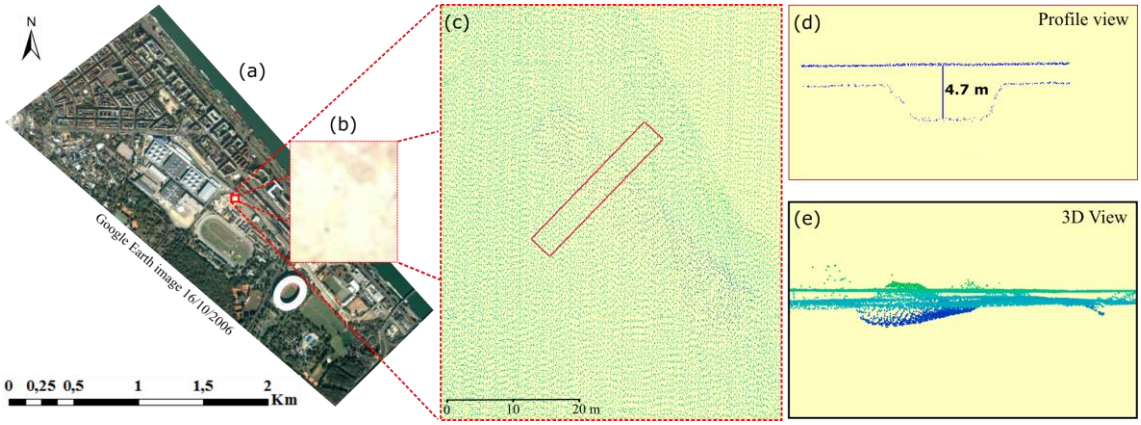

Figure 8. Open ground in the point cloud of 2007. Two contradicting layers of ground are shown due to construction activity within the duration of the 2007 campaign. (**a**) The Google earth image shows the location of the construction area; (**b**) orthophoto showing the selected area; (**c**) ground view of the point cloud indicates the position of the profile shown in sub-figure (**d**); and (**e**) a 3D view of the multilayer point cloud with the difference in height reaching 4.7 m.

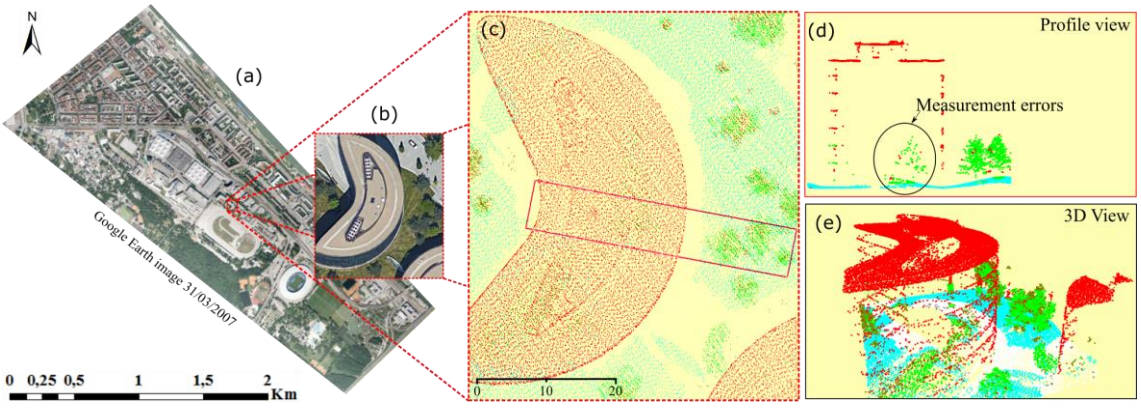

Figure 9. Erroneously measured points inside a mirroring building with a glass façade in the 2015 dataset. (**a**) The Google earth image located the position of the building; (**b**) orthophoto showing the selected area; (**c**) the ground view of the point cloud; (**d**) the profile view displays the erroneously measured point inside the building; and (**e**) 3D view of the point cloud.

5.4. Accuracy Evaluation

To evaluate the effectiveness of the proposed methods for detecting changes in urban areas, a comparative analysis between change detection results and the reference data was conducted. Because no photos were acquired which were collected simultaneously with the ALS data, the reference data were obtained by visual interpretation in 3D and manual classification. Reference point collection was conducted based on the rules:

- Unchanged buildings: The same geometric building in two epochs or buildings which have changes in roof but lower than 1 m (i.e., paying tribute to the chosen search radius).
- Lost buildings: Buildings are available in the older data but not in the later data.
- New buildings: Buildings are available in the later data but not in the older data.
- Unchanged ground: The height of the ground did not change more than 0.5 m.
- Changed ground: The ground has changed in height, changed to other types of land use (i.e., new buildings), or new ground.
- Unchanged trees: Trees at the same position.
- Lost trees: Trees that were cut.
- New trees: Newly planted trees.
- High points: Cars, fences (wooden, concrete, metal, and small bushes), wires, ships on the water, etc.

This manual classification approach is a tough task and time-consuming. However, this approach is more advantageous than using orthophotos in the case of comparing the change in height of ground, which is difficult when using 2D orthophotos. The selected region for doing manual classification, which is shown in Figure 5, was cut out from the datasets. The criteria to choose this area were: (1) select an area where all types of changes occur; (2) avoid the training samples as much as possible to ensure the objectivity of the accuracy assessment; and (3) investigate the entire study region, also for objectivity. The total reference area is about 33.7 ha out of the whole area 376.7 ha (approximately 9%). The time spent for manual classification of this area was about 150 working hours.

To focus on the above-mentioned changed objects only, the "high points" are manually classified but not investigated for change. They contain primarily objects, which are dynamic within a day, and objects for which the sampling is rather poor (thin objects). Those high points also removed from the accuracy evaluation. In addition, the class water is not represented in the confusion matrix. High points and water together add up to 3% of the ground truth data. The evaluated points are grouped into classes according to object change: unchanged ground (UG), changed ground (CG), unchanged building (UB), lost building (LB), new building (NB), unchanged tree (UT), lost tree (LT), and new tree (NT). The confusion matrix documenting the accuracy are shown in Table 4. Both datasets achieved a high overall accuracy of about 91% and 92% in 2007 and 2015, respectively.

From Table 4 it can be seen, that five out of six classes show over 80% correctness in the 2007 dataset. Only the class UT reached 70.7% of correctness because of misclassification as unchanged building (1.1%) and lost tree (0.5%).

There are some specific problems, most relating to special objects in the dataset. Unchanged building points are misclassified as unchanged tree in the case of complex building roofs, especially at the edge of the buildings, and the stadium frame dome where the distribution of points is the same as the tree point distribution.

In the confusion matrix of the 2015 dataset (Table 4), the most striking class is NT (new trees), for which correctness and completeness reach only 58% and 56.5%, respectively. Here, about 39% (1.1/2.8) of the points that in reality are NT were wrongly classified as UT (unchanged trees). The reason for this low completeness can be explained by two reasons. Firstly, some old trees were cut and at the very same position new trees were planted (see Figure 10). Consequently, in the change detection result, these tree points are classified as unchanged trees. Some new small trees grow near big cut trees (Figure 10) and are also mis-detected.

Table 4. Confusion matrix of the classification result for the 2007 and 2015 datasets. The rows correspond to the reference classification, the columns to the automatic classification. EOO, Error of Omission; Comp, Completeness; EOC, Error of Commission; Corr, Correctness; UG, unchanged ground; CG, changed ground; UB, unchanged building; LB, lost building; NB, new building; UT, unchanged tree; LT, lost tree; NT, new tree.

2007	UG	CG	UB	LB	UT	LT	Ref Sum	EOO	Comp
Ref_UG	53.8	1.8	0.1	0	0.1	0	55.8	3.6	96.4
Ref_CG	3.6	10.1	0	0.1	0	0	13.8	26.7	73.3
Ref_UB	0.1	0	16.7	0.4	1.1	0.1	18.2	8.7	91.3
Ref_LB	0	0	0.2	2.9	0	0.1	3.1	8.6	91.4
Ref_UT	0	0	0.4	0	4.1	0.4	4.9	16.1	83.9
Ref_LT	0	0	0.1	0.1	0.5	3.4	4.1	18.2	81.8
Sum	57.5	12	17.4	3.5	5.8	3.9	100	0	100
EOC	6.4	15.6	4.2	17.6	29.3	13.2	0	0	100
Corr	93.6	84.4	95.8	82.4	70.7	86.8	100	100	0
Overall Accuracy: 90.93									
Total number of points: 8,542,450									
2015	**UG**	**CG**	**UB**	**NB**	**UT**	**NT**	**Ref_sum**	**EOO**	**Comp**
Ref_UG	48.3	0.5	0.1	0	0	0	48.9	1.3	98.7
Ref_CG	0.9	10	0.1	0.1	0	0	11.0	9.1	90.9
Ref_UB	0	0	16.5	0.2	0.9	0	17.7	6.9	93.1
Ref_NB	0	0.2	0.1	4.6	0.1	0	5.0	8.4	91.6
Ref_UT	0	0	0.3	0.2	11.1	1.1	12.8	12.9	87.1
Ref_NT	0	0	0.1	0.1	1.1	1.6	2.8	43.5	56.5
Sum	49.2	10.8	17.1	5.3	13.2	2.7	98.2	0	100
EOC	1.8	6.9	3.6	12.4	15.9	42	0	0	100
Corr	98.2	93.1	96.4	87.6	84.1	58.0	100	100	0
Overall Accuracy: 92.05									
Total number of points: 8,636,900									

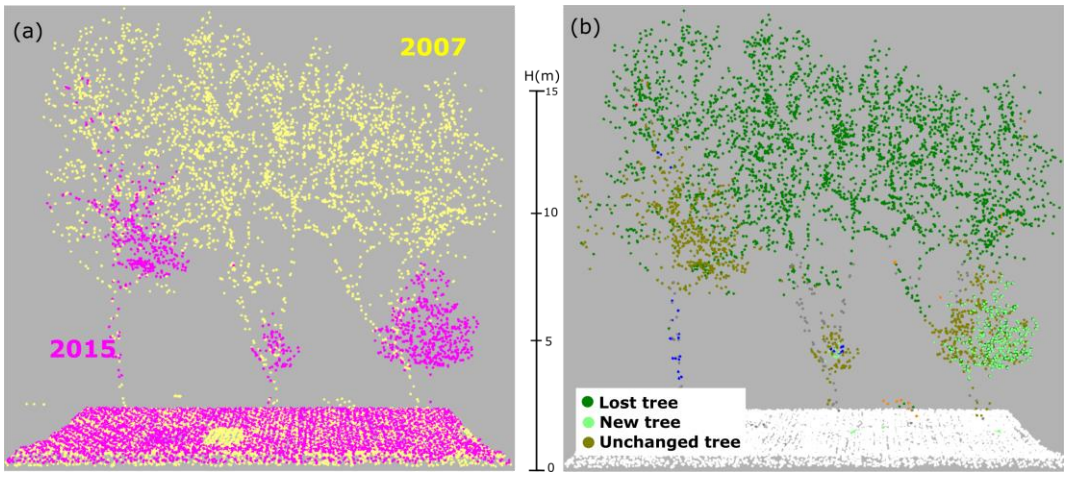

Figure 10. New trees planted at the same location as old lost trees: (**a**) point clouds of 2007 and 2015; and (**b**) change detection and classification results.

Because of roughly eight years apart, trees grow up significantly for the small trees and the grown trees have changed their shape (e.g., branch cut and new branch). Consequently, those growing points are classified into new trees, but in reality, they are parts of the same tree. This leads to a low completeness in new tree points of the 2015 ALS data (Figure 11).

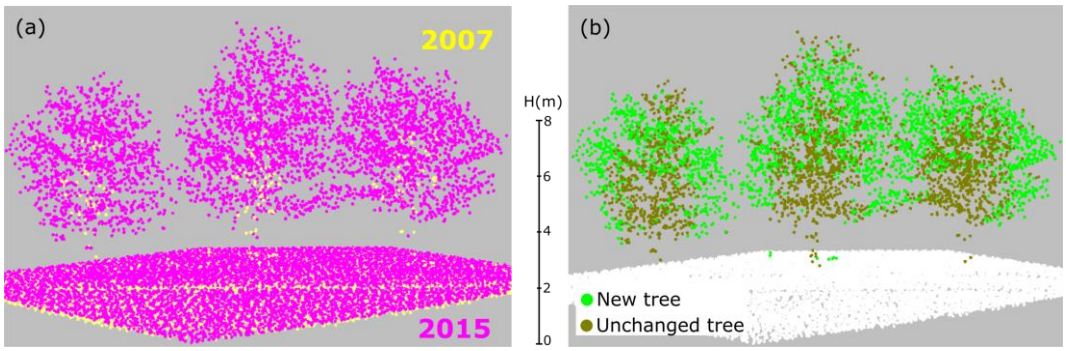

Figure 11. New planted trees at the same location with lost tree: (**a**) point cloud in 2007 and 2015; and (**b**) change detection and classification results).

A visual analysis of the entire change detection result suggests that the following classes can be determined with high reliability: unchanged ground, changed ground, new building, lost building, unchanged building, lost tree. However, this analysis also revealed that the change detection for growing trees constitutes a big challenge, as some unchanged tree points were classified as new buildings. This originates in a very dense canopy cover during the 2015 data acquisition resulting in a planar distribution of points and therefore features which are more similar to building points. This can be seen in the forested areas of 2015 on the southwestern border of the dataset (see Figure 6). By selecting points of the respective classes, we estimated that about 1.5% of unchanged tree points are wrongly classified as new buildings.

5.5. Discussion

Thus far, most studies focused on change detection of buildings only and they achieved very high accuracy. A very recent research which is closest to our study is the one by Xu et al. [34]. Their overall accuracy for change detection of buildings and trees reached 94.8% and 83.8%. However, their method to access the accuracy is different from ours. Their reference data was obtained by visual interpretation of the aerial images and airborne LiDAR data from two dates counting changed objects (not points). Then, their experimental results and the reference data were put together. The correctness and completeness of the buildings are determined visually based on the count and the area respectively. Our method does not only evaluate more objects compared to their method, but also our comparison is based on the classification of corresponding points, not on object count and area. Thus, the accuracy values cannot be compared directly.

The classification results in Figure 6 and the evaluation outcomes in Table 4 demonstrate the good performance of the proposed method in urban areas. This method exploits the ability of extending machine learning classification to perform classification and change detection in one step. In addition to combining change detection steps, the proposed method also is flexible in feature selection as well as in the data used. Firstly, in feature selection, for 3D point clouds numerous features can be used for a machine learning classifier. Weinmann et al. [71] mentioned numerous 2D and 3D features. In our study, we just used some of these features. However, depending on the classification task, the selection of features may be extended or compacted (e.g., using color in point clouds from image matching). In addition, the change detection feature used in this study is "Stability" to detect the changes between two epochs. However, other features, such as difference in roughness value of the local point to the nearest point of the other epoch, surface distance between one point in one epoch to the surface of the nearest points in the other epoch (compare tangent plane ICP correspondence [76]) etc., can be used as alternative sources of change information. To investigate this, new features were investigated: (1) distance to nearest point in the other epoch, (2) difference in roughness of current point and the nearest point in the other epoch. With these features new models were learned and the classification performed for the entire dataset. Using only the distance feature, the overall quality

91% decreases slightly to 89%, using only the difference in roughness it drops to 73%. Using all three features as markers of change, the overall accuracy increases in the order of 0.1% to 91% and 92% for 2007 and 2015, respectively. Secondly, the proposed method has potential for combining different data sources for change detection. With the flexibility in feature selection, our method allows doing change detection and classification for different data depending on the features given to identify the classes. Image matching point clouds recently became one of the important sources used in urban classification, also exploiting the provided color information. This data can be applied in further studies for change detection in urban area where the changes in buildings, trees, and grounds occur frequently.

Although the proposed method obtained a satisfying change detection classification result in one step compared to other studies, there remain some limitations. Firstly, the results of classification strongly depend on the training samples. Especially for a complex urban area, it is required to consider various types of objects. Thus, to select the samples of each class required careful observation and selection. Secondly, in the case of changes where old and new points are too close to each other, the method did not work well. For example, cut trees and a new building are shown in Figure 12. Post classification methods (e.g., label smoothing [77]) may support improvement of the results. Thirdly, as mentioned above, growing tree points are mis-detected as new trees. It is difficult to separate this class (growing tree points) from the new tree class. A solution may require object detection, i.e., single tree detection in this case. Parameters of individual trees could then be compared.

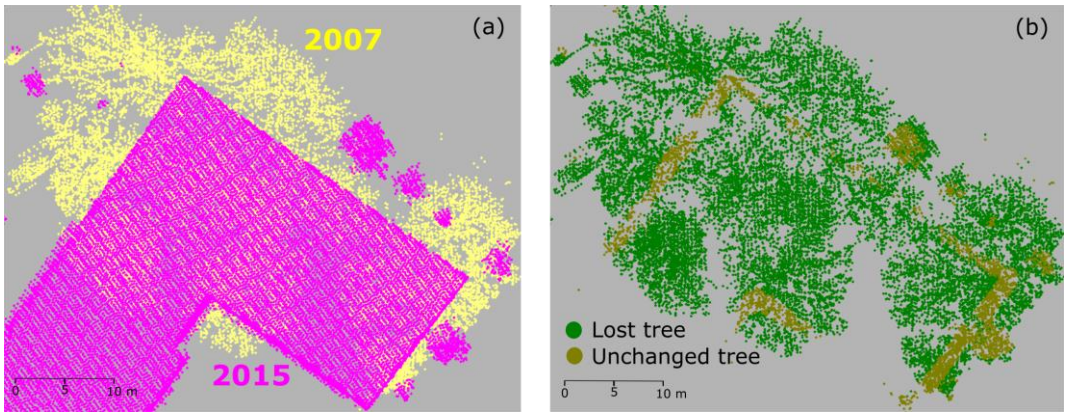

Figure 12. Example of misclassification in the case of object with adjacent old and new data: (**a**) data in 2007 and 2015; and (**b**) misclassification results in 2007. In the intersection of the objects, points are misclassified as unchanged trees.

Finally, we compared our method to a traditional two-step classification approach, i.e., detect the change and then classify the changes. Two DSM models of 2007 and 2015 are derived with a cell size of 1 m. The DSM difference of 1 m is chosen to separate changed and persistent objects. The first three sets of features (i.e., point distribution, height above the terrain, ALS features) are rasterized with the same cell size of 1 m. Those rasterized images are masked into a changed and an unchanged region based on the DSM masks. The training samples are rasterized and divided into changed and unchanged samples for each epoch 2007 and 2015. Based on those samples, the classification is performed. Finally, combining the masks and the classification result, the change detection classification is performed for 2007 and for 2015. This traditional raster-based approach is easy to process and less time is required for processing compared to our point-based method. However, the final results depend on the first change detection step. DSM-based change detection is useful for buildings, but not for trees. Tree growth can be higher than 3 m, given the eight years apart. Therefore, if the DSM difference is 1 m, unchanged trees are classified into new trees. Increasing the DSM difference, the change in ground and small buildings are lost. For this reason, the overall accuracy of this method is only 78% for both the 2007

and the 2015 datasets. Furthermore, the raster-based method does not feature the 3D content anymore, which is available in the point-based method.

6. Conclusions

This paper has presented a fusion of automatic classification and change detection using a supervised point-based machine learning approach to infer changes in the objects building and tree, as well as changes of the ground. The main contribution of the proposed method can be summarized as: (1) the proposed method establishes a complete set of processes to detect and classify changes in buildings, trees and ground; and (2) not only are changes detected, but they are also simultaneously classified, which had not been done before, especially for the major classes ground, building, and tree in one step. The combination of the "Stability" feature with other attributes plays an important role for the automatic change detection and classification of different types of urban objects. The overall accuracy of the final classification of each change type of the 2007 dataset and 2015 dataset reached 90.93% and 92.04%, respectively. Therefore, the proposed method can be used as an alternative method for detecting changes in urban areas in high resolution point clouds.

Acknowledgments: The authors would also like to thank VIED (Vietnam International Education Development) and OeAD (Österreichischer Austauschdienst) support during the studying period. The data were provided by the 'Stadtvermessung Wien', the Magistrate of the City of Vienna.

Author Contributions: Thi Huong Giang Tran processed the data, performed the experiment, and drafted the article. Camillo Ressl prepared the data and accessed the accuracy of the final results. Norbert Pfeifer supervised the research, reviewed and edited the manuscript. All authors contributed to the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hebel, M.; Arens, M.; Stilla, U. Change detection in urban areas by object-based analysis and on-the-fly comparison of multi-view ALS data. *ISPRS. J. Photogramm.* **2013**, *86*, 52–64. [CrossRef]
2. Güler, M.; Yomralioğu, T.; Reis, S. Using landsat data to determine land use/land cover changes in Samsun, Turkey. *Environ. Monit. Assess.* **2007**, *127*, 155–167. [CrossRef] [PubMed]
3. Menaka, E.; Kumar, S.S.; Bharathi, M. Change detection in deforestation using high resolution satellite image with Haar wavelet transforms. In Proceedings of the 2013 International Conference on Green High Performance Computing, ICGHPC 2013, Nagercoil, India, 14–15 March 2013.
4. Brunner, D.; Bruzzone, L.; Lemoine, G. Change detection for earthquake damage assessment in built-up areas using very high resolution optical and SAR imagery. In Proceedings of the 2010 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Honolulu, HI, USA, 25–30 July 2010; pp. 3210–3213.
5. Anniballe, R.; Chini, M.; Pierdicca, N.; Bignami, C.; Stramondo, S.; Noto, F.; Scalia, T.; Martinelli, A.; Mannella, A. Detecting earthquake damage in urban area: Application to COSMO-SkyMed imagery of L'Aquila earthquake. In Proceedings of the SPIE 9642, SAR Image Analysis, Modeling, and Techniques XV, Toulouse, France, 23 October 2015; 96420C.
6. Lu, D.; Mausel, P.; Brondízio, E.; Moran, E. Change detection techniques. *Int. J. Remote Sens.* **2004**, *25*, 2365–2401. [CrossRef]
7. Muchoney, D.M.; Haack, B.N. Change detection for monitoring forest defoliation. *Photogramm. Eng. Remote Sens.* **1994**, *60*, 1243–1251.
8. Ying, Q.; Hansen, M.C.; Potapov, P.V.; Tyukavina, A.; Wang, L.; Stehman, S.V.; Moore, R.; Hancher, M. Global bare ground gain from 2000 to 2012 using Landsat imagery. *Remote Sens. Environ.* **2017**, *194*, 161–176. [CrossRef]
9. Li, G.; Zhang, F.; Jing, Y.; Liu, Y.; Sun, G. Response of evapotranspiration to changes in land use and land cover and climate in China during 2001–2013. *Sci. Total Environ.* **2017**, *596–597*, 256–265. [CrossRef] [PubMed]
10. Lunetta, R.S.; Knight, J.F.; Ediriwickrema, J.; Lyon, J.G.; Worthy, L.D. Land-cover change detection using multi-temporal MODIS NDVI data. *Remote Sens. Environ.* **2006**, *105*, 142–154. [CrossRef]

11. Meyfroidt, P.; Lambin, E.F. Forest transition in Vietnam and its environmental impacts. *Glob. Chang. Biol.* **2008**, *14*, 1319–1336. [CrossRef]
12. Sesnie, S.E.; Gessler, P.E.; Finegan, B.; Thessler, S. Integrating Landsat TM and SRTM-DEM derived variables with decision trees for habitat classification and change detection in complex neotropical environments. *Remote Sens. Environ.* **2008**, *112*, 2145–2159. [CrossRef]
13. Zhao, L.; Yang, J.; Li, P.; Zhang, L.; Shi, L.; Lang, F. Damage assessment in urban areas using post-earthquake airborne PolSAR imagery. *Int. J. Remote Sens.* **2013**, *34*, 8952–8966. [CrossRef]
14. Guida, R.; Iodice, A.; Riccio, D. Monitoring of collapsed built-up areas with high resolution SAR images. In Proceedings of the 2010 IEEE International Geoscience and Remote Sensing Symposium, Honolulu, HI, USA, 25–30 July 2010; pp. 2422–2425.
15. Vakalopoulou, M.; Karatzalos, K.; Komodakis, N.; Paragios, N. Simultaneous registration and change detection in multitemporal, very high resolution remote sensing data. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Boston, MA, USA, 7–12 June 2015; pp. 61–69.
16. Chen, G.; Zhao, K.; Powers, R. Assessment of the image misregistration effects on object-based change detection. *ISPRS J. Photogramm.* **2014**, *87*, 19–27. [CrossRef]
17. Qin, R.; Gong, J.; Li, H.; Huang, X. A coarse elevation map-based registration method for super-resolution of three-line scanner images. *Photogramm. Eng. Remote Sens.* **2013**, *79*, 717–730. [CrossRef]
18. Waser, L.T.; Baltsavias, E.; Ecker, K.; Eisenbeiss, H.; Feldmeyer-Christe, E.; Ginzler, C.; Küchler, M.; Zhang, L. Assessing changes of forest area and shrub encroachment in a mire ecosystem using digital surface models and CIR aerial images. *Remote Sens. Environ.* **2008**, *112*, 1956–1968. [CrossRef]
19. Tian, J.; Nielsen, A.A.; Reinartz, P. Improving change detection in forest areas based on stereo panchromatic imagery using kernel MNF. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 7130–7139. [CrossRef]
20. Qin, R.; Tian, J.; Reinartz, P. 3D change detection—Approaches and applications. *Photogramm. Eng. Remote Sens.* **2016**, *122*, 41–56. [CrossRef]
21. Biljecki, F.; Ledoux, H.; Stoter, J. Generating 3D city models without elevation data. *Comput. Environ. Urban Syst.* **2017**, *64*, 1–18. [CrossRef]
22. Steinhage, V.; Behley, J.; Meisel, S.; Cremers, A.B. Automated updating and maintnance of 3D city models. In Proceedings of the Core Spatial Databased—Updating, Maintenance and Services ISPRS Archive, Haifa, Israel, 15–17 March 2010; pp. 1–6.
23. Golparvar-Fard, M.; Peña-Mora, F.; Savarese, S. Automated progress monitoring using unordered daily construction photographs and IFC-based building information models. *J. Comput. Civ. Eng.* **2015**, *29*. [CrossRef]
24. Golparvar-Fard, M.; Peña-Mora, F.; Savarese, S. Monitoring changes of 3D building elements from unordered photo collections. In Proceedings of the 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops), Barcelona, Spain, 6–13 November 2011; pp. 249–256.
25. Matsumoto, J.; Nishimaru, H.; Ono, T.; Nishijo, H. 3D-video-based computerized behavioral analysis for in vivo neuropharmacology and neurophysiology in rodents. In *Vivo Neuropharmacology and Neurophysiology*; Humana Press: New York, NY, USA, 2017; pp. 89–105, ISBN 978-1-4939-6490-1.
26. Matsumoto, J.; Urakawa, S.; Takamura, Y.; Malcher-Lopes, R.; Hori, E.; Tomaz, C.; Ono, T.; Nishijo, H. A 3D-video-based computerized analysis of social and sexual interactions in rats. *PLoS ONE* **2013**, *8*, e78460. [CrossRef] [PubMed]
27. Meyer, V.; Saatchi, S.S.; Chave, J.; Dalling, J.W.; Bohlman, S.; Fricker, G.A.; Robinson, C.; Neumann, M.; Hubbell, S. Detecting tropical forest biomass dynamics from repeated airborne LiDAR measurements. *Biogeosciences* **2013**, *10*, 5421–5438. [CrossRef]
28. Wu, G.; Jiang, X.; Xie, H.; Park, D.-H. Experimental study on tree growth in XLPE using 3D PD patterns. In Proceedings of the 6th International Conference on Properties and Applications of Dielectric Materials (Cat. No.00CH36347), Xi'an, China, 21–26 June 2000; pp. 558–561.
29. Jaboyedoff, M.; Oppikofer, T.; Abellán, A.; Derron, M.H.; Loye, A.; Metzger, R.; Pedrazzini, A. Use of LiDAR in landslide investigations: A review. *Nat. Hazards* **2012**, *61*, 5–28. [CrossRef]
30. Tran, T.H.G.; Hollaus, M.; Nguyen, B.D.; Pfeifer, N. Assessment of wooded area reduction by airborne laser scanning. *Forests* **2015**, *6*, 1613–1627. [CrossRef]

31. Rutzinger, M.; Höfle, B.; Hollaus, M.; Pfeifer, N. Object-based point cloud analysis of full-waveform airborne laser scanning data for urban vegetation classification. *Sensors* **2008**, *8*, 4505–4528. [CrossRef] [PubMed]
32. Egberth, M.; Nyberg, G.; Næsset, E.; Gobakken, T.; Mauya, E.; Malimbwi, R.; Katani, J.; Chamuya, N.; Bulenga, G.; Olsson, H. Combining airborne laser scanning and Landsat data for statistical modeling of soil carbon and tree biomass in Tanzanian Miombo woodlands. *Carbon Balance Manag.* **2017**, *12*. [CrossRef] [PubMed]
33. Gruno, A.; Liibusk, A.; Ellmann, A.; Oja, T.; Vain, A.; Jürgenson, H. Determining sea surface heights using small footprint airborne laser scanning. In Proceedings of the SPIE 8888, Remote Sensing of the Ocean, Sea Ice, Coastal Waters, and Large Water Regions 2013, Dresden, Germany, 16 October 2013. [CrossRef]
34. Xu, H.; Cheng, L.; Li, M.; Chen, Y.; Zhong, L. Using Octrees to Detect Changes to Buildings and Trees in the Urban Environment from Airborne LiDAR Data. *Remote Sens.* **2015**, *7*, 9682–9704. [CrossRef]
35. Breiman, L. *Machine Learning*; Kluwer Academic: Dordrecht, The Netherlands, 2001; Volume 45, pp. 5–32. ISSN 0885-6125.
36. Matikainen, L.; Hyyppä, J.; Kaartinen, H. Automatic detection of changes from laser scanner and aerial image data for updating building maps. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2004**, *35*, 434–439.
37. Matikainen, L.; Hyyppä, J.; Ahokas, E.; Markelin, L.; Kaartinen, H. Automatic detection of buildings and changes in buildings for updating of maps. *Remote Sens.* **2010**, *2*, 1217–1248. [CrossRef]
38. Stal, C.; Tack, F.; De Maeyer, P.; De Wulf, A.; Goossens, R. Airborne photogrammetry and LiDAR for DSM extraction and 3D change detection over an urban area—A comparative study. *Int. J. Remote Sens.* **2013**, *34*, 1087–1110. [CrossRef]
39. Malpica, J.A.; Alonso, M.C.; Papí, F.; Arozarena, A.; Martínez De Agirre, A. Change detection of buildings from satellite imagery and LiDAR data. *Int. J. Remote Sens.* **2013**, *34*, 1652–1675. [CrossRef]
40. Matikainen, L.; Hyyppä, J.; Litkey, P. Multispectral Airborne Laser Scanning for Automated Map Updating. In Proceedings of the International Archieves of the Photogrammetry, Remote Sensing and Spatial Information Science, XLI-B3, Prague, Czech Republic, 12–19 July 2016; pp. 323–330.
41. Matikainen, L.; Karila, K.; Hyyppä, J.; Litkey, P.; Puttonen, E.; Ahokas, E. Object-based analysis of multispectral airborne laser scanner data for land cover classification and map updating. *ISPRS J. Photogramm. Remote Sens.* **2017**, *128*, 298–313. [CrossRef]
42. Vosselman, G.; Gorte, B.G.H.; Sithole, G. Change detection for updating medium scale maps using laser altimetry. In Proceedings of the International Archives of the Photogrammtry, Remote Sensing and Spatial Information Sciences, Istanbul, Turkey, 12–23 July 2004; Volume 34, pp. 207–212.
43. Tang, F.; Xiang, Z.; Teng, D.; Hu, B.; Bai, Y. A multilevel change detection method for buildings using laser scanning data and GIS data. In Proceedings of the 2015 IEEE International Conference on Digital Signal Processing (DSP), Singapore, 21–24 July 2015; pp. 1011–1015.
44. Awrangjeb, M. Effective generation and update of a building map database through automatic building change detection from LiDAR point cloud data. *Remote Sens.* **2015**, *7*, 14119–14150. [CrossRef]
45. Choi, K.; Lee, I.; Kim, S. A feature based approach to automatic change detection from LiDAR data in urban areas. In Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Paris, France, 1–3 September 2009; Volume 38.
46. Xu, S.; Vosselman, G.; Oude Elberink, S. Detection and Classification of Changes in Buildings from Airborne Laser Scanning Data. *Remote Sens.* **2015**, *7*, 17051–17076. [CrossRef]
47. Teo, T.A.; Shih, T.Y. LiDAR-based change detection and change-type determination in urban areas. *Int. J. Remote Sens.* **2013**, *34*, 968–981. [CrossRef]
48. Murakami, H.; Nakagawa, K.; Hasegawa, H.; Shibata, T.; Iwanami, E. Change detection of buildings using an airborne laser scanner. *ISPRS J. Photogramm. Remote Sens.* **1999**, *54*, 148–152. [CrossRef]
49. Pang, S.; Hu, X.; Wang, Z.; Lu, Y. Object-Based Analysis of Airborne LiDAR Data for Building Change Detection. *Remote Sens.* **2014**, *6*, 10733–10749. [CrossRef]
50. Vu, T.T.; Matsuoka, M.; Yamazaki, F. LiDAR-based change detection of buildings in dense urban areas. In Proceedings of the 2004 IGARSS'04Geoscience and Remote Sensing Symposium, Anchorage, AK, USA, 20–24 September 2004; pp. 3413–3416.
51. Zhang, X.; Glennie, C. Change detection from differential airborne LiDAR using a weighted anisotropic iterative closest point algorithm. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 3338–3346. [CrossRef]

52. Estivill-Castro, V.; Lee, I. Argument free clustering for large spatial point-data sets via boundary extraction from Delaunay Diagram. *Comput. Environ. Urban Syst.* **2002**, *26*, 315–334. [CrossRef]
53. Cortes, C.; Vapnik, V. Support-Vector Networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
54. Mallet, C.; Bretar, F.; Roux, M.; Soergel, U.; Heipke, C. Relevance assessment of full-waveform LiDAR data for urban area classification. *ISPRS J. Photogramm.* **2011**, *66* (Suppl. 6), S71–S84. [CrossRef]
55. Shirowzhan, S.; Trinder, J. Building Classification from LiDAR Data for Spatio-temporal Assessment of 3D Urban Developments. *Procedia Eng.* **2017**, *180*, 1453–1461. [CrossRef]
56. Hernández, J.; Marcotegui, B. Point cloud segmentation towards urban ground modeling. In Proceedings of the 2009 Joint Urban Remote Sensing Event, Shanghai, China, 20–22 May 2009.
57. Lehtomäki, M.; Jaakkola, A.; Hyyppä, J.; Lampinen, J.; Kaartinen, H.; Kukko, A.; Puttonen, E.; Hyyppä, H. Object Classification and Recognition From Mobile Laser Scanning Point Clouds in a Road Environment. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 1226–1239. [CrossRef]
58. Guo, L.; Chehata, N.; Mallet, C.; Boukir, S. Relevance of airborne LiDAR and multispectral image data for urban scene classification using Random Forests. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 56–66. [CrossRef]
59. Ni, H.; Lin, X.; Zhang, J. Classification of ALS point cloud with improved point cloud segmentation and random forests. *Remote Sens.* **2017**, *9*, 288. [CrossRef]
60. Sutton, C.; McCallum, A.; Rohanimanesh, K. Dynamic conditional random fields: Factorized probabilistic models for labeling and segmenting sequence data. *J. Mach. Learn. Res.* **2007**, *8*, 693–723.
61. Niemeyer, J.; Rottensteiner, F.; Soergel, U. Contextual classification of LiDAR data and building object detection in urban areas. *ISPRS J. Photogramm. Remote Sens.* **2014**, *87*, 152–165. [CrossRef]
62. Weinmann, M.; Jutzi, B.; Mallet, C. Sematic 3D scene interpretation: A framework combining optimal neighborhood size selection with relevant features. In Proceeding of the Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Zurich, Switzerland, 5–7 September 2014; Volume 2, pp. 181–188.
63. Socher, R.; Huval, B.; Bhat, B.; Manning, C.D.; Ng, A.Y. Convolutional-recursive deep learning for 3D object classification. *Adv. Neural Inf. Process. Syst.* **2012**, 656–664.
64. Li, N.; Pfeifer, N.; Liu, C. Tensor-based sparse representation classification for Urban Airborne LiDAR points. *Remote Sens.* **2017**, *9*, 1216. [CrossRef]
65. Zhang, Z.; Zhang, L.; Tong, X.; Guo, B.; Zhang, L.; Xing, X. Discriminative-Dictionary-Learning-Based Multilevel Point-Cluster Features for ALS Point-Cloud Classification. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 7309–7322. [CrossRef]
66. Zhang, Z.; Zhang, L.; Tan, Y.; Zhang, L.; Liu, F.; Zhong, R. Joint Discriminative Dictionary and Classifier Learning for ALS Point Cloud Classification. *IEEE Trans. Geosci. Remote Sens.* **2017**, *56*, 524–538. [CrossRef]
67. Gu, Y.; Wang, Q.; Jia, X.; Benediktsson, J.A. A Novel MKL Model of Integrating LiDAR Data and MSI for Urban Area Classification. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 5312–5326.
68. Pfeifer, N.; Mandlburger, G.; Otepka, J.; Karel, W. OPALS—A framework for Airborne Laser Scanning data analysis. *Comput. Environ. Urban Syst.* **2014**, *45*, 125–136. [CrossRef]
69. Inpho. Available online: https://geospatial.trimble.com/products-and-solutions/inpho (accessed on 30 January 2018).
70. FugroViewer. Available online: https://www.fugro.com/about-fugro/our-expertise/technology/fugroviewer (accessed on 18 December 2017).
71. Weinmann, M.; Jutzi, B.; Mallet, C. Feature relevance assessment for the semantic interpretation of 3D point cloud data. In Proceedings of the ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Antalya, Turkey, 11–13 November 2013; Volume II-5/W2, pp. 313–318.
72. Gressin, A.; Mallet, C.; David, N. Improving 3D LiDAR point cloud registration using optimal neighborhood knowledge. In Proceedings of the ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Melbourne, Austria, 25 August–1 September 2012; Volume I-3, pp. 111–116.
73. Höfle, B.; Mücke, W.; Dutter, M.; Rutzinger, M. Detection of building regions using airborne LiDAR—A new combination of raster and point cloud based GIS methods Study area and datasets. In *Geospatial Crossroads @ GI_Forum '09, Proceedings of the Geoinformatics Forum Salzburg, Geoinformatics on Stage, Heidelberg, Germany, 7–10 July 2009*; Car, A., Griesebner, G., Strobl, J., Eds.; Wichmann Verlag: Heidelberg, Germany, 2009; pp. 66–75.

74. Höfle, B.; Hollaus, M.; Hagenauer, J. Urban vegetation detection using radiometrically calibrated small-footprint full-waveform airborne LiDAR data. *ISPRS J. Photogramm. Remote Sens.* **2012**, *67*, 134–147. [CrossRef]
75. Pfeifer, N.; Mandlburger, G. LiDAR Data Filtering and DTM Generation. 2008. Available online: ftp://ftp.ecn.purdue.edu/jshan/Zproject/proofs/11/51423_C011_correctedProof.pdf (accessed on 29 January 2018).
76. Glira, P.; Pfeifer, N.; Briese, C.; Ressl, C. A Correspondence Framework for ALS Strip Adjustments based on Variants of the ICP Algorithm. *Photogramm. Fernerkund. Geoinf.* **2015**, *4*, 0275–0289. [CrossRef]
77. Landrieu, L.; Raguet, H.; Vallet, B.; Mallet, C.; Weinmann, M. A structured regularization framework for spatially smoothing semantic labelings of 3D point clouds. *ISPRS J. Photogramm. Remote Sens.* **2017**, *132* (Suppl. C), 102–118. [CrossRef]