Classification of Trees and Powerlines from medium resolution Airborne Laserscanner data in Urban Environments

Simon Clode
Intelligent Real-Time Imaging and Sensing Group
The University of Queensland
Brisbane, QLD 4072, Australia
sclode@itee.uq.edu.au

Franz Rottensteiner
School of Surveying and SIS
The University of New South Wales
UNSW SYDNEY NSW 2052, Australia
f.rottensteiner@unsw.edu.au

Abstract

A method for the classification of trees and powerlines in urban areas by using only dual return (first and last pulse) medium resolution Airborne Laserscanner (ALS) data is presented. ALS points with a different first and last pulse return are initially identified and building detection techniques are then used to separate buildings from initial areas of interest. The separation of tree and powerline data is performed by applying a classification method based on the theory of Dempster - Shafer for data fusion. Examples of the classification method are compared against ground truth for a test site in Sydney, Australia.

1. Introduction

1.1. Motivation and Goals

Research on automated object extraction for 3D city models has been fuelled in recent years by the increasing use of geographic information systems (GIS), and the need for data acquisition and update for GIS. The main focus in this context was the detection and reconstruction of buildings [2], [5], [16] and roads [6]. Some existing methods use multiple data sources in order to achieve comprehensive 3D city models. Recently, the use of 3D point clouds generated from airborne laser scanning (ALS) for automatic creation of 3D city models has been gaining importance.

ALS data have several unique properties. Firstly, laser points are not selective and as such do not automatically strike the object required [13]. Secondly, due to the finite spot size of the laser beam, an imperfection pointed out in [12], there might be more than one echo of the laser. Modern ALS systems are capable of collecting both first pulse (FP) and last pulse (LP) data during one flight, and some objects can only be discerned in a FP-LP difference image (Figure 1). Finally, ALS systems deliver the intensity of the returned laser beam, which however is usually undersampled and thus noisy. This is caused by the imbalance of the average footprint size of the laser beam (e.g. 20-30 cm) and the average point distance (e.g. 1 m) [19].

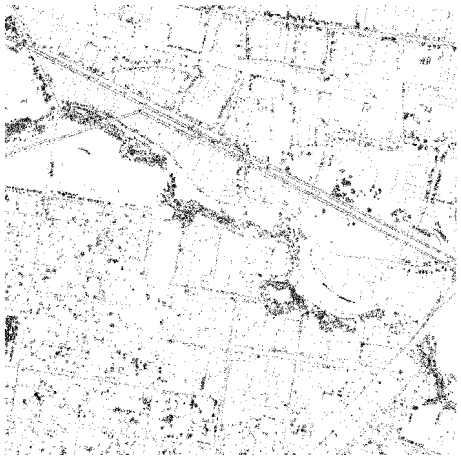


Figure 1. Height differences between pulses

During the early stages of development, ALS was primarily used for topographic mapping of terrain in forested areas in order to generate digital terrain models (DTM's) [1], [2], [10], [16]. As sensor technology has improved, so has the achievable resolution of point clouds from ALS data [9], and methods to extract objects from stand-alone ALS data have emerged. Buildings have been extracted from ALS data using a variety of methods [2], [15], [16]. Roads have been effectively classified using ALS from an urban landscape in [4]. In order to extend the comprehensiveness of 3D city model creation from stand-alone ALS data, other object classes need to be extracted, too. As both trees and powerlines can be easily seen in an unprocessed image of the height differences between the surfaces corresponding to the first and last laser pulses of ALS data (Figure 1), a method to effectively classify these object types was sought. It is the goal of this paper to

- Detect trees and powerlines from stand-alone ALS data in urban areas with resolution of approximately 1 point per square metre
- Improve the ability of creating 3D city models from a single data source, namely ALS data.

This paper presents results of a new method to classify trees and powerlines in an urban area from ALS data. Section 1.2 provides a review of related research. Section 2 describes the conceptual approach, model assumptions and describes our new method for determining areas covered by trees and powerlines. Results from the sample data set are discussed in section 3 whilst conclusions and future work are examined in section 4.

1.2. Related Work

ALS systems have been used in areas covered by trees since their infancy [1]. Initially, DTM's were derived from the ALS data but this soon progressed into canopy height determination which can be used to model canopy volumes and above ground biomasses [11].

In [10], DTMs were created in forested areas with a single last pulse ALS system. The use of ALS in wooded areas was considered very beneficial due to the ability of the laser to penetrate the trees and make contact with the ground. Although it was acknowledged that further filtering and interpolation was required to divide the ALS data into ground and non-ground strikes, it was concluded that the accuracy of the final DTM was comparable to that of DTM's generated in open areas with photogrammetry.

In [5], multispectral imagery and ALS data are combined for the extraction of buildings, trees and grass covered areas. Trees and grass covered areas are easily classified from the multispectral imagery but not easily separated. Similarly, trees and buildings can be separated using the height differences between a digital surface model (DSM) and the DTM. Both data sources are combined in order to identify the three classification types. In [17], classification of land cover into four different classes (building, tree, grass land, and bare soil) is achieved by combining ALS data and multispectral images. The ALS data is initially preprocessed to generate a DTM before building detection is performed by data fusion based on the theory of Dempster - Shafer [8].

The potential of ALS for the detection of individual trees has been explored several times, e.g. [7], [14], [18]. Early experiments were performed in [7] within forests dominated by coniferous trees within boreal forests . The results varied and difficulties were encountered in dense young forest or in groups of deciduous trees. In [14], very high resolution ALS data ($>$10 points/m^2) was used to segment single trees. Local maxima in a DSM are used as seed points in the raw ALS data for the tree identification. The heights derived from the ALS data where consistent with ground truth information but location of the trees did not occur with such consistency. Crown volume and diameter were then calculated. In [18], single trees and their crown parameters are extracted from a DSM from ALS data and optical images. The height difference between first and last pulse data is not used. An evaluation on a per-tree basis using four different data sets achieved completeness between 50% and 96% and correctness between 59% and 86%, depending on the data quality, especially on the state of the tree canopies at the time of flight.

A variety of tools have been used for the classification of ALS data. Basic remote sensing tools were used in [1] to classify both buildings and trees. The authors used a vegetation index called the Normalised Difference (ND), derived from the first and last pulse data DSMs similar to the Normalised Difference Vegetation Index from infrared images. The ND index basically shows if both a FP and LP were recorded in the same pulse and does not distinguish between the type of objects that could cause such a return. As identified in [16] and [3] there are many different object types that could be detected in such a manner, such as powerlines, building edges, and trees.

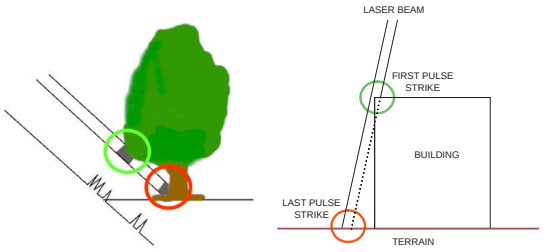
2. The Classification Method

2.1. Conceptual Approach and Method Overview

In our method we assume that buildings have previously been extracted from the ALS data (Figure 4(a)). This is achieved by the method described in [16], evaluating cues such as the relative height of the ALS points above a DTM and the surface roughness of the DSM created from the ALS data. Having done that, we detect trees and powerlines from the ALS data by merely evaluating the height differences between FP and LP data and the ALS intensity values. In Figure 1, all the ALS points in the surveyed region that have registered a different first and last pulse return are displayed. Trees, powerlines and building edges can easily be seen in this unprocessed image. Figure 2 shows how the laser beam interacts with trees and building edges.

Figure 3 gives an overview over the work flow of our method. In our classification model we assume that any difference between FP and LP is caused by either trees, powerlines, or building edges. We first exclude all points on building edges from further processing, making use of the previously detected buildings. After that, we differentiate between powerlines and trees. Trees are characterised by the fact that there will be many points with a large height difference between first and last pulse data in a local neighbourhood. The ALS intensities might be in any range, depending on the tree species and the time of year. Powerlines on the other hand tend to have only few points of a large height difference between first and last pulse data in a local neighbourhood. They also have low intensities of return.

These model assumptions are used to derive several cues for classification which are then combined in a data fusion process based on the theory of Dempster-Shafer [8], [17].



(a) Laser penetrating a tree (b) The edge of a building

Figure 2. Laser Reflections

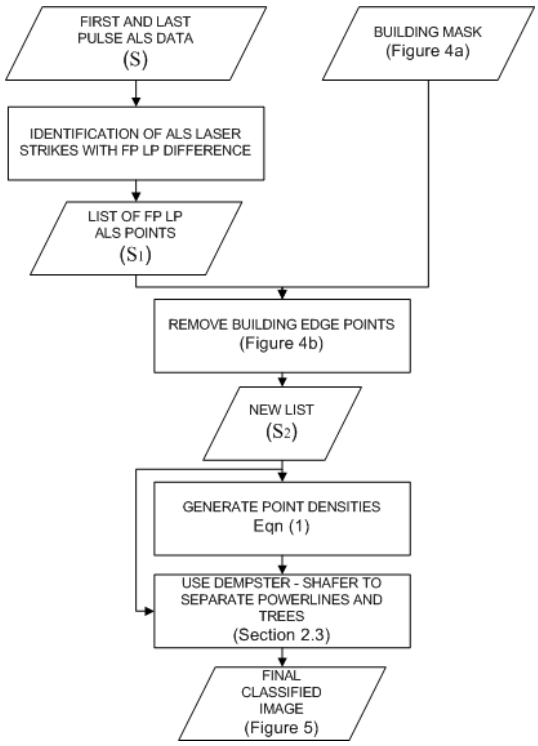


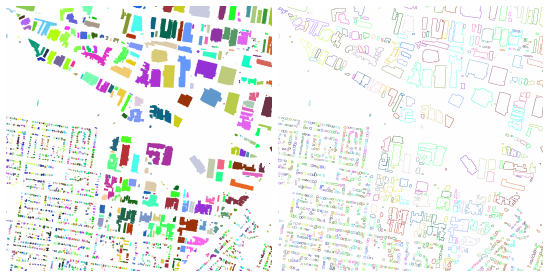
Figure 3. The classification flowchart

In this paper, we will give the results of our method applied to a test data set from Fairfield in Sydney, Australia. The data set was initially collected with an approximate point density of 1 point per 1.3 m^2. Both first and last pulse and the intensity of the reflected laser beam were recorded.

2.2. Classification Cues

For the purpose of this paper, we will describe any ALS data point $\mathbf{p}_k$ as being defined by $\mathbf{p}_k = (lx_k, ly_k, lz_k, li_k, fx_k, fy_k, fz_k, fi_k)$, where the first letter describes the pulse, i.e. $l = last$ and $f = first$, and the second letter describes the pulse 3D coordinate or intensity by either x, y, z or i. Let $\mathbf{S}$ represent the set of all laser points collected, i.e. $\mathbf{S} = \{\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_N\}$, where $\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_N$ are the individual ALS points. A point $\mathbf{p}_k$ is considered to have different first and last pulses where $\Delta H_k > 0$ and $\Delta H_k = fz_k - lz_k$. A set of all points points that have different first and last pulses is described by $\mathbf{S}_1 = \{\mathbf{p}_k \in \mathbf{S} : \Delta H_k > 0\}$.

In a way similar to [20], a band of pixels around each building outline is created. The width of the band is dependant on the original point spacing. As the resolution of the test data (1.3 m) is approximately the same as the pixel size (1 m), a band of width 2 pixels was chosen to form a corridor on either side of the existing building boundary. ALS points which lie inside the initial building outline band and exist in $\mathbf{S}_1$ can be considered as being situated on the building edge and as such need to be removed from the set. We define a new set $\mathbf{S}_2$ where ALS points identified as building edges have been removed from $\mathbf{S}_1$ (Figure 4).



(a) Building Mask (b) Points on building outlines

Figure 4. Removing building edge points

The ALS points contained in $\mathbf{S}_2$ are preprocessed to form inputs into a classification method based on the theory of Dempster-Shafer. Three images are created in this preprocessing step, namely a first pulse laser intensity image I_F, an image ΔH_{FL} containing the height differences between FP and LP, and a local point density image ρ. The pixel values of the images I_F and ΔH_{FL} have to be interpolated from the intensities and height differences of the ALS points. We use an interpolation method based on inverse distance weighting for that purpose. The image ρ describes the ratio of the number of points having different first and last pulse heights (thus, from the data set $\mathbf{S}_2$) to the total number of ALS strikes within a local area. The value of ρ at any position (k) is described by Equation 1, where $||p_k - p_j||_2$ denotes the Euclidean distance between p_k and p_j and d is the radius of the local neighbourhood:

$$\rho_k = \frac{|\{p_j \in S_2 : ||p_k - p_j||_2 < d\}|}{|\{p_j \in S : ||p_k - p_j||_2 < d\}|} \quad (1)$$

2.3. Separating Powerlines and Trees

We start this section with an outline of Dempster-Shafer fusion based on [8]. Consider the classification problem

where the input data are to be classified into n mutually exclusive classes $C_j \in \Theta$. The power set of Θ is denoted by 2^Θ and contains not only the the *original classes* C_j but also all their possible unions (hence called *combined classes*). A probability mass $m(A)$ is assigned to every class $A \in 2^\Theta$ by an "image" (a classification cue) such that $m(\emptyset) = 0$, $0 \leq m(A) \leq 1$, and $\sum_{A \in 2^\Theta} m(A) = 1$, where $\emptyset$ denotes the empty set. Uncertainty in classification from an individual cue can be modelled by assigning a non-zero probability mass to the union of two or more classes C_j. The *support* $Sup(A)$ of a class $A \in 2^\Theta$ is defined as the sum of all masses assigned to A:

$$Sup(A) = \sum_{B \subseteq A} m(B) \tag{2}$$

Dubiety $Dub(A) = Sup(\bar{A})$ is the degree to which the evidence contradicts a proposition, or supports the complementary hypothesis of A: $A \cup \bar{A} = \Theta$. If there are p inputs, probability masses $m_i(B_j)$ have to be defined for each i such that $1 \leq i \leq p$ and $B_j \in 2^\Theta$. The probability masses from several inputs can then be combined to compute a combined probability mass for each class $A \in 2^\Theta$:

$$m(A) = \frac{\sum_{B_1 \cap B_2 \ldots \cap B_p = A} \left[\prod_{1 \leq i \leq p} m_i(B_j)\right]}{1 - \sum_{B_1 \cap B_2 \ldots \cap B_p = \emptyset} \left[\prod_{1 \leq i \leq p} m_i(B_j)\right]} \tag{3}$$

Once the combined probability masses $m(A)$ have been computed, both $Sup(A)$ and $Sup(\bar{A})$ can be calculated. The accepted hypothesis $C_a \in \Theta$ is determined as the class obtaining maximum support.

We apply the Dempster - Shafer theory to the data on a pixel by pixel basis to classify the inputs into one of three classes, Tree (T), Powerline (L) or Other (O). As described in section 2.2, three input cues are used in the classification:

(1) The height differences ΔH_{FL} between FP and LP distinguish powerlines and trees from other objects, without separating these two classes. We thus assign a probability mass $P_{\Delta H} = P_{\Delta H}(\Delta H)$ ascending with ΔH to the combined class $T \cup L$ and $1 - P_{\Delta H}$ to class O.

(2) The density image ρ can be used to separate powerlines from trees because trees cover a larger area and thus there will be more points with FP-LP differences in a local neighbourhood, but only where $\Delta H > 0$. In areas where $\Delta H = 0$, we do not use ρ, which is modelled by assigning a probability mass of 1.0 to Θ. Where $\Delta H > 0$, we assign a constant small probability mass P_O to class O in order to model the fact that not all points with $\rho > 0$ will be points on powerlines or trees. We then assign a probability mass $P_\rho = P_\rho(\rho) - P_O/2$ ascending with ρ to the class T and $1 - P_O/2 - P_\rho$ to class L.

(3) The intensity I_F separates trees from powerlines and other objects, but only in areas where an intensity value has actually been measured by the sensor. Where an intensity value exists, a probability mass $P_I = P_I(I_F)$ ascending with I_F is assigned to class T and $1 - P_I$ to class $L \cup O$. Otherwise, Θ will be assigned a probability mass of 1.0.

The functions for computing the probability masses ($P_{\Delta H}$, P_ρ, P_I) are assumed to be equal to a constant P_1 for input parameters $x < x_1$. For input parameters $x > x_2$, they are assumed to be equal to another constant P_2, with $0 \leqslant P_1 < P_2 \leqslant 1$. Between x_1 and x_2, the probability mass is described by a cubic parabola using $\bar{x} = \frac{x - x_1}{x_2 - x_1}$ and $k \in \{\Delta H, \rho, I\} : P_k(\bar{x}) = P_1 + (P_2 - P_1)(3\bar{x}^2 - 2\bar{x}^3)$. P_1 and P_2 are chosen to be 5% and 95%, respectively, and P_O = 10%. Further, we choose (x1, x2) = (2.5 m, 4.5 m) for $P_{\Delta H}$, (x1, x2) = (0%, 70%) for P_ρ, and (x1, x2) = (0, 7.5) for P_I.

3. Results

The overall classification results from our new method are shown in Figure 5. The areas covered by trees are indicated by the light green pixels and the powerline classification by the black pixels.

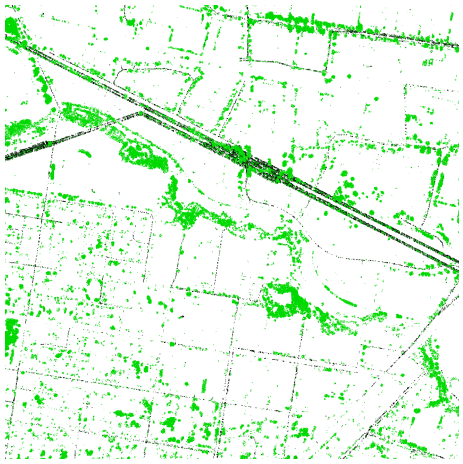


Figure 5. The final classification results

3.1. Accuracy of Tree Detection

In order to evaluate the classification algorithm, ground truth data for areas covered by trees was obtained by manually digitising trees in an orthophoto of the area. The resultant ground truth image is displayed in Figure 6(a). The detected trees are shown in figure 6(b).

In order to assess the quality of the classification, the *completeness* and *correctness* of the results are computed. Completeness is the ratio of the correctly extracted records to the total number of relevant records within the ground truth data, whereas correctness is the ratio of the number of relevant records extracted to the total number of relevant and irrelevant records retrieved:

$$\text{Completeness} = \frac{TP}{TP + FN}$$
$$\text{Correctness} = \frac{TP}{TP + FPO} \quad (4)$$

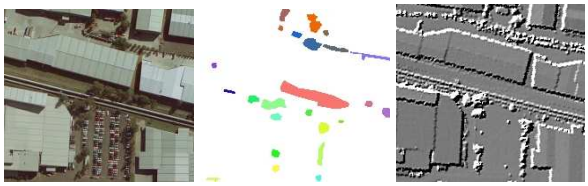
In Equation 4, *TP* denotes the number of True Positives, FN the number of False Negatives (i.e. missing "tree" pixels), and FPO the number of False POsitives (i.e. "tree" pixels not being classified as trees in the reference data). To assist in the analysis of the results, figure 8 shows the spatial distribution of the TP, FN , FPO and True Negatives (TN) pixels in yellow, blue, red and white respectively.

In our test, completeness was determined to be 64% and correctness was 64%. These numbers appear to be too pessimistic. As the aerial image and the ALS data were captured in different epochs, there are many discrepancies between the ground truth and the ALS data. There are two major factors that will effect the quality figures quoted as the data has been collected at different epochs. The first is obviously the time that has elapsed between the collection of the two data sets. Vegetation is a dynamic object class as opposed to buildings and will grow over time. Also, people cut back trees so the opposite effect is also true. The second effect is that seasonal changes can be observed. Trees that loose their leaves in autumn will have a finer canopy during this period as compared to spring. An example of these contradictions between the data sets is shown in Figure 7. Finally, this comparison gives a balance of the area covered by trees that is correctly classified. Errors mostly occur at the tree boundaries. As most trees are relative small objects, these errors at the tree boundaries might contribute up to 20% of the area covered by trees.

(a) Manually digitised trees from the orthophoto. (b) Areas covered by trees as a result of our method.

Figure 6. The results of the tree classification

A visual inspection of Figure 6(b) reveals that there is a misclassification along the thick powerlines running from the West to the East of the image. Another limiting factor to the tree detection will be the limitation of the laser as mentioned in [3]. The failure of the laser to detect a FP if $\Delta H_{FL} < 4.6m$ meant that trees with a height of less than this height would not have been detected but would probably exist in the ground truth data.



(a) Trees exists along the the pipeline in the orthophoto. (b) The manually digitised vegetation ground truth. (c) No trees have been detected in the FP ALS data .

Figure 7. Contradictions between ground truth and ALS data.

Figure 8. The quality summary map

3.2. Accuracy of Powerline Classification

There was no ground truth data available for the powerline classification method and it was considered too difficult to accurately digitise the powerlines in a similar manner to the trees. It was decided that the most effective way to assess the quality of the powerline classification was to compare the classified powerline image visually against the road network as powerlines generally run parallel to roads. Ground truth for the road network was again obtained by manually digitising the orthophoto of the area and can be seen in Figure 9(a). The classification results can be seen in Figure 9(b).

A visual perusal of both images shows that with the exception of the major powerlines that basically run from the west to the east of the image, the overall pattern of the powerline classification matches the road network as expected.

(a) The manually digitised road network of Fairfield. (b) The detected powerlines.

Figure 9. The powerline classification results

4. Conclusions and Future Work

The ability to identify trees in an ALS point cloud by considering the first-pulse/last-pulse differences has been extended by considering the local point density of the occurrence of the these measurements. By using the local point density and intensity of the first pulse return, separation of powerlines and trees has been achieved. A visual check of the results reveals that classification is accomplished with a certain amount of success. The results of tree classification are encouraging, and a formal quantitative analysis was performed but unfortunately the results are not truly representative of the quality of the classification achieved. By ensuring that ground truth data is captured during the same epoch as the ALS data and utilising a newer laser with a smaller "dead spot", the results obtained from the algorithm are expected to provide better quality results. The work presented in this paper is still in progress. Although the results obtained have been very encouraging, future work should be concentrated on quantifying the results against ground truth data that has been captured during the same epoch. This would give a true indication of the actual performance of the classification algorithm. Investigation into methods that will allow vectorisation of the the classified powerlines should also be performed.

Acknowledgements

This research was funded by the ARC Linkage Project LP0230563 and the ARC Discovery Project DP0344678. The Fairfield data set was provided by AAMHatch, Queensland, Australia. (http://www.aamhatch.com.au)

References

[1] H. Arefi, M. Hahn, and J. Lindenberger. LIDAR Data Classification with Remote Sensing Tools. In *ISPRS Joint Workshop on Challenges in Geospatial Analysis, Integration and Visualization II*, pages 131 – 136, Stuttgart, Germany, 2003.

[2] C. Brunn and U. Weidner. Extracting Buildings from Digital Surface Models. In *IAPRS*, volume XXXII / 3-4W2, pages 27 – 34, 1997.

[3] S. Clode, P. Kootsookos, and F. Rottensteiner. Accurate building outlines from als data. In *Proceedings of ASPRS*, Fremantle, Australia, 2004.

[4] S. Clode, P. Kootsookos, and F. Rottensteiner. The Automatic Extraction of Roads from LIDAR Data. In *IAPRSIS*, volume XXXV-B3, pages 231 – 236, 2004.

[5] N. Haala and C. Brenner. Extraction of Buildings and Trees in Urban Environments. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54:130 – 137, 1999.

[6] S. Hinz and A. Baumgartner. Automatic Extraction of Urban Road Networks from Multi-View Aerial Imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58/1-2:83 – 98, 2003.

[7] J. Hyyppa and M. Inkinen. Detecting and Estimating Attributes for Single Trees Using Laser Scanner. *The photogrammetric journal of Finland*, 16:27 – 42, 1999.

[8] L. Klein. *Sensor and Data Fusion, Concepts and Applications*. SPIE Optical Engineering Press, 1999.

[9] K. Kraus. Principles of Airborne Laser Scanning. *Journal of the Swedish Society for Photogrammetry and Remote Sensing*, 1:53 56, 2002.

[10] K. Kraus and N. Pfeifer. Determination Of Terrain Models In Wooded Areas With Airborne Laser Scanner Data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 53:193 – 203, 1998.

[11] K. Lim, P. Treitz, M. Wulder., B. St-Onge, and M. Flood. LiDAR Remote Sensing of Forest Structure. *Progress in Physical Geography*, 27,1:88 106, 2003.

[12] H.-G. Maas and G. Vosselman. Two Algorithms for Extracting Building Models from Raw Laser Altimetry Data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54:153 – 163, 1999.

[13] M. Morgan and A. Habib. Interpolation of LIDAR Data and Automatic Building Extraction. In *ACSM-ASPRS Annual Conference Proceedings*, 2002.

[14] F. Morsdorf, E. Meier, B. Allgower, and D. Nuesch. Clustering in Airborne Laser Scanning Raw Data for Segmentation of Single Trees. In *IAPRSIS*, volume XXXIV - 3/W13, pages 27 – 33, 2003.

[15] F. Rottensteiner and C. Briese. A New Method for Building Extraction in Urban Areas from High-Resolution LIDAR Data. In *IAPSIS*, volume XXXIV / 3A, page 295 301, 2002.

[16] F. Rottensteiner, J. Trinder, S. Clode, and K. Kubik. Building Dectection Using LIDAR data and Multispectral Images. In *Proceedings of DICTA*, pages 673 – 682, Sydney, Australia, 2003.

[17] F. Rottensteiner, J. Trinder, S. Clode, and K. Kubik. Using the DempsterShafer Method for the Fusion of LIDAR Data and Multi-spectral Images for Building Detection. *Information Fusion*, 2004. In print.

[18] B.-M. Straub. A Top-down Operator for the Automatic Extraction of Trees - Concept and Performance Evaluation. In *IAPRSIS*, volume XXXIV-3/W13, pages 34 – 39, 2003.

[19] G. Vosselman. On the Estimation of Planimetric Offsets in Laser Altimetry Data. In *IAPRSIS*, volume XXXIV/3A, pages 375 – 380, 2002.

[20] U. Weidner and W.Förstner. Towards Automatic Building Extraction from High Resolution Digital Elevation Models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 50(4):38 – 49, 1995.