DECIDUOUS-CONIFEROUS TREE CLASSIFICATION USING DIFFERENCE BETWEEN FIRST AND LAST PULSE LASER SIGNATURES

X. Liang*, J. Hyyppä, L. Matikainen

Finnish Geodetic Institute, Department of Remote Sensing and Photogrammetry, Masala, Finland
(Xinlian.Liang, Juha.Hyyppa, Leena.Matikainen)@fgi.fi

Commission III

KEY WORDS: Leaf-off, First-last pulse, Species identification, Airborne laser scanners

ABSTRACT:

In this paper, a deciduous-coniferous tree classification mechanism is proposed, tested and analyzed using solely laser scanner data. The data were acquired under leaf-off conditions by Toposys II system. Under such circumstance, sources of last pulse hits of deciduous and coniferous are different, which allows concise discrimination between these two species. Tree positions were located from first pulse DSM, species were identified by the difference between two pulse data and field measurements were used for validation. The classification results demonstrate that first-last pulse laser data, under leaf-off condition, is ideal for deciduous and coniferous trees classification; and also indicate that the data collected for high accuracy DEM production is also suitable for forest investigation.

1. INTRODUCTION

Airborne laser scanners (ALS), providing small footprint diameters (10 – 30 cm), allow accurate forest information estimation (e.g. Næsset 1997; Magnussen and Boudewyn, 1998). Two main approaches in deriving forest attributes using laser scanner data have been those based on laser canopy height distribution and on individual tree detection. In former approach, percentiles of laser canopy heights distribution are used as predictors to estimate forest characteristics. Næsset (2002), Lim et al. (2003) and Holmgren and Jonsson (2004) have shown that this approach produces highly reliable estimates of stand variables. If the number of laser pulses is increased to several measurements per square meter, individual trees can be recognized (Hyyppä and Inkinen, 1999; Hyyppä et al. 2001, Persson et al., 2002; Popescu et al., 2002; Leckie et al. 2003). From individual tree, height, crown diameter and even species can be derived using laser scanner data. Then, more tree and stand attributes, e.g. timber volume, can be quite reliably estimated using existing forest models based on height, diameter and specie information (Hyyppä and Inkinen, 1999).

Tree species is an essential index in forest studies, inventories, managements and other forest applications. In practice, species classification is performed using range and optical/near-infrared data, together or individually.

In Brandtberg (2002), features describing branch structure, crown shape and color were extracted from high spatial resolution color infrared aerial photographs and then input into a classification system. In Bohlin et al. (2006), spectral values, corresponding to sunlit part of detected crown, were extracted from high spatial resolution color infrared aerial photographs and applied in tree species identification. In the June and October images, 88% and 89% of the detected trees, respectively, could be separated into three classes, pine (Pinus Sylves-tris), spruce (Picea Abies) and deciduous.

The airborne laser scanning data has also been tested for tree species classification. Holmgren and Persson (2004) stated that it is possible to separate pine and spruce using laser scanner data. That approach was tested at individual tree level between Scots pine and Norway spruce. The portion correctly classified trees on all plots was 95%. Moffiet et al. (2005) proposed that the proportion of laser singular returns is an important predictor for the tree species classification. Brandtberg et al. (2003) used laser data under leaf-off conditions for the detection of individual trees. Additionally, classification results of different indices suggest a moderate to high degree of accuracy using single or multiple variables between deciduous trees. Brandtberg (2007) presented a framework to express interactions of the laser beams with individual tree canopy, and proposed species classification strategies for selecting group of laser points, where variables used were quantifications of independent events and statistics/geometric measurements. Overall, 64% classification accuracy is achieved, for three leaf-off deciduous trees, oaks (Quercus spp.), red maple (Acer rubrum) and yellow poplar (Liriodendron tuliperifera).

Persson et al. (2006) identified individual tree species through combining features of high resolution laser data with high resolution multi-spectral images. Classification experiment was conducted in southern Sweden with forest dominated by Norway spruce (Picea abies), Scots pine (Pinus sylvestris), and deciduous trees, mainly birch (Betula spp.). The results implied that by combining structure and spectral features, the classification could be improved (95 % accuracy).

As a summary, laser data has been used for tree species classification successfully, but there are unsolved problems. Firstly, the accurate classification between deciduous and coniferous trees requires the aid of optical or near-infrared data. Secondly, in practical applications, there can only be few training trees for a large area. Therefore, in order to receive

* Corresponding author.

good classification accuracies, more research is needed in deriving good features for tree species classification.

The increasing use of laser scanning for nation-wide elevation data collection also supports use of the data for other purposes. Up to now, airborne laser scanning data can be available for whole countries (Switzerland, Netherland) and for large districts (South Tyrol / Italy, Vorarlberg / Austria, Saxon-Bohemian Switzerland / Germany, Baden-Württemberg / Germany). All this data has been collected with leaf-off conditions in order to get the highest accuracy in DEM. In Finland, the leaf-off data is also the first candidate for the national laser scanning. Forest inventory authorities are, however, planning to have a laser and aerial imaging survey during summer time in order to get more reliable information also for the tree species.

In the present work, we analyzed airborne laser data acquired in a suburban site in 2003 under leaf-off conditions. The objective was to demonstrate that leaf-off laser data is ideal for tree species classification between deciduous and coniferous trees. In previous studies, it has already been demonstrated that deciduous trees can be reconstructed using leaf-off data.

2. MATERIAL AND METHODS

The test site locates in Espoolahti, 15 km west of Helsinki. Toposys II (wavelength of 1.54 µm, maximum scan angle ±7.1°) campaigns were carried out in 14th May 2003. At that time, the leaves were off and in some cases there were small buds depending on the tree species. The flying altitude was 400 m above ground and the beam divergence was 1 mrad giving a footprint of 0.40 m in diameter. In the Toposys II, there are 128 parallel beams (pushbroom type scanner) that are sampled in a fast rate. The point spacing between consecutive beams was 80 cm in the across track direction and between 10-15 cm in the along track direction (depending on the flight speed). Therefore, there was a high autocorrelation between the consecutive hits in the along track direction. We expect that the data corresponded to a nominal pulse density of 4 to 5 pulse per square meter organized in even spacing. Thus, the sampling density of the data was not significantly higher than that used in nation-wide data collection (for example, in Switzerland, the surveying company has collected laser data for their own use with a density of about 4 pulses per square meter).

Reference data included 295 identified trees, which were of direct access and were evenly scattered across the test site. Among them, 176 were coniferous (spruce and pine) and 119 were deciduous trees. Tree species information, coniferous or deciduous, was collected from those trees in summer 2006.

Coordinate transformations, geoid correction, strip adjustment and systematic shift correction were first performed on laser points cloud. Then, last pulse data were classified in TerraScan software (see www.terrasolid.fi) to separate the ground points from others (low and high vegetation). In Terrascan, the ground points were triangulated using TIN densification method developed by Axelsson (2000). The following parameters, cite dependent, were used for the classification: max. building size 100m, terrain angle 75˚, iteration angle 6º, iteration distance 1.2m, and reduce iteration angle when edge length < 5 m. The raster image file corresponding to the Digital Elevation Model (DEM) was created from the classified ground points using the following parameters: lowest hit within 0.5 m grid spacing and gaps filled up to 10 pixels.
The Digital Surface Model (DSM) was calculated for both pulses respectively, with a 0.5 m grid from the highest hits. Gaps were filled up to 3 pixels using interpolation. The final Canopy Height Model (CHM) was then calculated as the difference between the DSM and the DEM.

In the tree top detection, a simplified process to that presented in Hyyppä et al. (2001) was applied. The prefiltering was done with a one step convolution of a 3 x 3 filter

$$\frac{1}{25}\begin{bmatrix} 1 & 3 & 1 \\ 3 & 9 & 3 \\ 1 & 3 & 1 \end{bmatrix} \tag{1}$$

and then the possible tree top position was found by a 5 x 5 maximum filter: if the current pixel was the highest in the 5 by 5 window, it was labelled as possible tree top. Lower local maxima, with height less than 3 m, were neglected from further analyses.

Tree crown radius was estimated using the tree top position and corresponding value of the CHM, which was taken as the tree height. The radius of the tree was assumed, according to Pitkänen et al. (2004), to be

$$R_{Org} = (1.2 + 0.16 \times H) \times 0.5 \tag{2}$$

In our analysis, the radius was reduced by 40%.

$$R_{Est} = R_{Org} \times 0.6 \tag{3}$$

It was assumed that the main difference in first-last pulse signature between coniferous and deciduous trees lies in the crown centre and that smaller radius leads to more reliable estimation for different tree species.

Then, a neighbourhood window $[2R_{Est} + 1, 2R_{Est} + 1]$ was defined as the estimated crown area.

It was expected that, under leaf-off conditions, first pulse signals correspond to reflections from treetops, even with the deciduous trees, as discovered by Brandtberg et al. (2003); and that the source of last pulse hits of deciduous trees is the ground and of coniferous tree, it is the tree top. Based on this assumption, tree species were classified by the absolute height difference between two DSMs, and defined as a function of two thresholds. If the proportion of pixels within the estimated crown area (defined in threshold 1) does not present significant height difference (defined in threshold 2), the tree was classified as coniferous tree. Otherwise, it was identified as deciduous tree.

3. RESULTS

Figure 1 and 2 show points cloud and DSM corresponding to coniferous and deciduous trees respectively. Points are in local coordinates. First and last pulses are marked in green and red, whereas tree top and estimated crown area are marked by cross and circle, respectively.

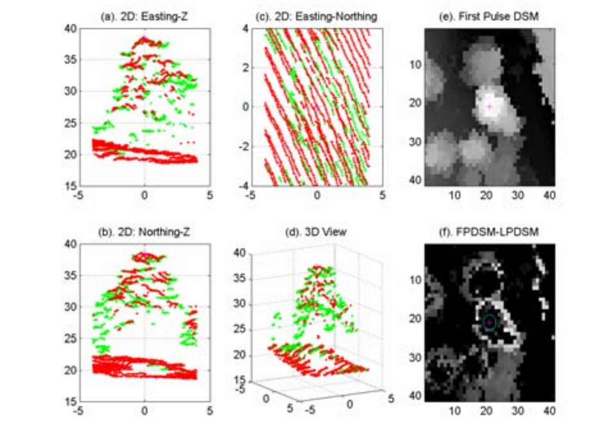

Figure 1. Points cloud and DSM corresponding to a coniferous tree. (a) Points cloud in vertical projection perpendicular to Northing, referred as Easting-Z. (b) Points cloud in vertical projection perpendicular to Easting, referred as Northing-Z. (c) Points cloud in horizontal projection, referred as Easting-Northing. (d) Points cloud in 3D space. (e) DSM based on first pulse laser data. (f) Difference between first and last pulse data.

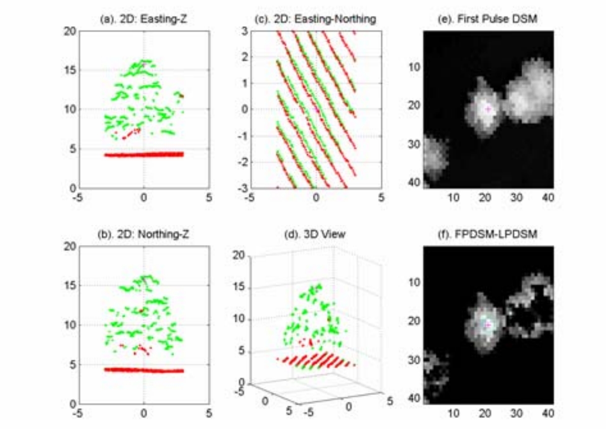

Figure 2. Points cloud and DSM corresponding to a deciduous tree, where a-f corresponds to cases in Fig. 1.

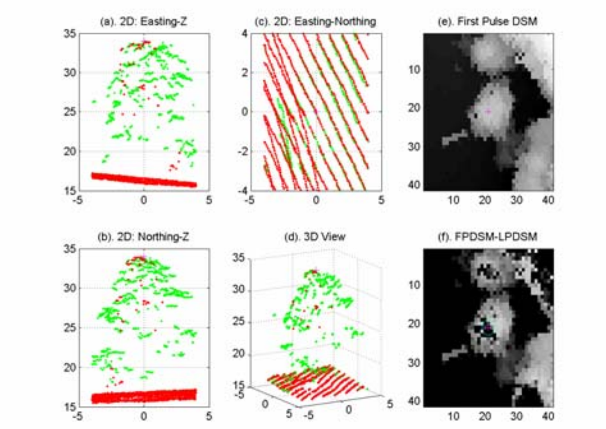

Figure 3. Points cloud corresponding to a deciduous tree, where a-f corresponds to cases in Fig. 1. Obviously, tree crown has been hit by the last pulse.

Figure 3 gives an example of how misclassification could be introduced. In that case, a deciduous tree's crown was hit by both pulses at the centre of the tree, leading to small height difference between first and last pulse data.

The classification results between coniferous and deciduous trees are reported in Table I and Figure.4.

Table I shows the confusion matrix of classification, where Th1 refers to the proportion of pixels within the estimated crown area and Th2 refers to the height difference in meter. The overall accuracy is 89.83%. Figure 4 shows the producer accuracy as a function of parameter Th2, where Th1 equals to 40%.

	Actual Coniferous	Actual Deciduous	Total
Classified as Coniferous	157	11	168
Classified as Deciduous	19	108	127
Total	176	119	295

Table 1. Confusion matrix (Th1 = 40% and Th2 = 0.3)

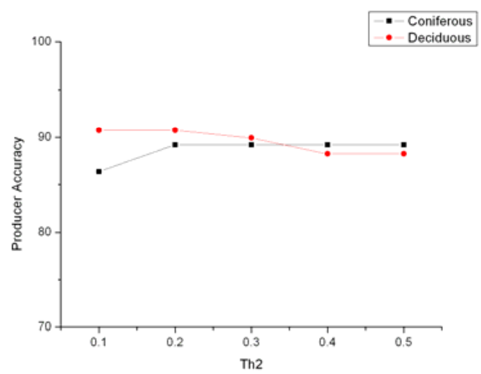

Figure 4. Producer accuracy as a function of threshold 2

4. DISCUSSION

The result indicates that a simple signature, i.e. the range difference between first and last pulse hits under leaf-off conditions, is suitable for deciduous-coniferous tree classification. In order to improve the quality of the classification and to understand where the proposed methodology failed, the misclassified cases were analysed. The reasons for the misclassification could be grouped into categories (1) branch structure, (2) only pulse hits (3) crown shape (4) tree top position, (5) parameter and (6) data processing. Yet, in practice, the misclassification is mostly introduced by several, rather than solely one, factors.

Branch structure: The basic assumption for this study was that, under leaf-off condition, last pulse would penetrate deciduous tree crown and would be reflected from coniferous tree top. However, under certain circumstance, e.g. when coniferous trees are heavily defoliated, this is not always the case. For the deciduous, the denser the crown is, the more last pulse points are reflected from upper branches, which reduces difference between the two pulse modes, and possibly leads to misclassification. On the other hand, small buds also caused that some last pulses were reflected by upper branches.

Number of only pulse hits: In first and last pulse data set, there are some points with same planar coordinates and small, even zero, height differences. We assumed that they corresponded to only pulse hits. In Toposys II, about 1 m vertical difference was needed to separate between first and last pulse mode. Accompanied with incidence angle and crown shape, the only pulse hit may happen to be the ones adopted in DSM generation, thus leading to smaller differences between first and last pulse data. This explains that some deciduous trees were misclassified as coniferous.

Crown shape: particularly, for spruce with cone shaped crown and small open angle, it is possible that both lower and upper crown parts happen fall into the same raster cell, due to relative large cell size. In such case, the height difference, between two pulses, is exaggerated, and then leads to misidentification.

Tree top position: In general, tree top is expected to locate at crown centre and correspond to local height maximum. However, in practice, branches' configuration may be complicated and then makes it hard to define a necessary and sufficient condition of treetops. For instance, tree top may incline to one side and then not be local height maximum; outstretched branches may be higher than real tree top. Therefore, it is possible that some tree tops locate apart from crown centre and some mis-located tree tops exist. In former case, the height difference, around the estimated crown area, may be larger than what is supposed to be for coniferous, due to the edge area. And then the trees would likely be misclassified as deciduous ones. To overcome this problem, it would help to first determine crown area and then find possible, or assumed, tree top position. In the latter case, mis-located trees were neglected from analyses.

Parameter: Figure 4 shows how overall classification accuracy changes according to threshold 2. Clearly, the fluctuation of accuracy is moderate and the classification is not sensitive to parameter TH2. However, selected parameter also plays its own role in the classification. In general, larger Th1 and smaller Th2, which means larger proportion pixels presents smaller height difference between first and last pulse data, lead to higher producer accuracy for the coniferous, smaller producer accuracy for the deciduous, and vice verse. Between the two parameters, the classification is more sensitive to Th1. Considering the overall accuracy, 40% keeps a balance and leads to accuracy around 89% for both species.

Data processing: data processing, which enhances the different reflectance pattern between the deciduous and coniferous, also contributes to identification accuracy amelioration. One example could be the process in last pulse DSM production. When the highest hits are assigned to DSM cells, like in this study, the confusion caused by several last hits reflected by lower coniferous branches does not, generally, introduce large height differences between the DSMs, and the coniferous identification accuracy therefore improves. On the other hand, if the lowest hits were assigned to DSM cells, the confusion caused by a few last hits reflected by upper deciduous branches would be effectively eliminated, which would contribute to high deciduous identification accuracy.

5. CONCLUSIONS

The results in this paper indicate that the difference between first and last pulse is a valuable feature for trees species classification. It reliably (89 % accuracy) gives the difference between coniferous and deciduous trees under leaf-off conditions. In order to conclude the optimal accuracy it could achieve, more experiments based on difference mechanisms are needed.

REFERENCES:

Bohlin, J, Olsson, H., Olofsson, K. and Wallerman, J., 2006. Tree species discrimination by aid of template matching applied to digital air photos. Workshop on 3D Remote Sensing in Forestry, 14th-15th Feb. 2006, Vienna. pp. 199-203.

Brandtberg, T., 2002. Individual tree-based species classification in high spatial resolution aerial images of forests using fuzzy sets. Fuzzy Sets and Systems, 132(3), pp. 371–387.

Brandtberg, T., 2007. Classifying individual tree species under leaf-off and leave-on conditions using airborne lidar. ISPRS Journal of Photogrammetry and Remote Sensing, 61, pp. 325-340.

Brandtberg, T. Warnera, T. A., Landenbergerb, R. E., McGrawb. J. B., 2003. Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density lidar data from the eastern deciduous forest in North America. Remote Sensing of Environment, 85(3), pp. 290-303.

Holmgren, J. and Jonsson, T., 2004. Large scale airborne laser-scanning of forest resources in Sweden. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences. 36(8/W2), pp. 157−160.

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

Hyyppä, J. and Inkinen, M. 1999. Detection and estimating attributes for single trees using laser scanner. The Photogrammetric Journal of Finland, 16(2), pp. 27-42.

Hyyppä, J., Kelle, O., Lehikoinen, M. and Inkinen, M. 2001. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. IEEE Transactions on Geoscience and Remote Sensing, 39, pp. 969-975.

Leckie, D., Gougeon, F., Hill, D., Quinn, R., Armstrong, L., Shreenan, R., 2003. Combined high density LiDAR and multispectral imagery for individual tree crown analysis. Canadian Journal of Remote Sensing, 29(5), pp. 633–649.

Lim, K., Treitz, P., Baldwin, K., Morrison, I., Green, J., 2003. Lidar remote sensing of biophysical properties of tolerant northern hardwood forests. Canadian Journal of Remote Sensing, 29(5), pp. 658–678.

Magnussen, S., and Boudewyn, P. 1998. Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. Canadian Journal of Forest Research, 28(7), pp. 1016-1031.

Moffiet, T., Mengersen, K., Witte, C., King, R., Denham, R., 2005. Airborne laser scanning: exploratory data analysis indicates potential variables for classification of individual trees or forest stands according to species. ISPRS Journal of Photogrammetry and Remote Sensing, 59(5), pp. 289–309.

Næsset, E., 1997. Determination of mean tree height of forest stands using airborne laser scanner data. ISPRS Journal of Photogrammetry and Remote Sensing, 52(2), pp. 49-56.

Næsset, E. 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sensing of Environment, 80(1), pp. 88–99.

Persson, Å., Holmgren, J. and Söderman, U. 2002. Detecting and measuring individual trees using an airborne laser scanner. Photogrammetric Engineering and Remote Sensing, 68(9), pp. 925-932.

Persson, Å., Holmgren, J., Söderman, U., and Olsson, H. 2006. Tree Species Classification of Individual Trees in Sweden by Combining High Resolution Laser Data with High Resolution Near-Infrared Digital Images. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 37(8/W2), pp. 204-206.

Pitkänen, J., Maltamo, M., Hyyppä, J. and Yu, X., 2004. Adaptive methods for individual tree detection on airborne laser based height model. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36 (8/W2), pp. 187-191.

Popescu, S.C., R.H. Wynne and R.E. Nelson, 2002. Estimating plot-level tree heights with lidar: local filtering with a canopy-height based variable window. Computers and Electronics in Agriculture, 37, pp. 71-95.