TREE SPECIES IDENTIFICATION USING LIDAR AND OPTICAL IMAGERY

Gints Priedītis, **Ingus Šmits**, **Salvis Daģis**, **Dagnis Dubrovskis**
Latvia University of Agriculture
Gints.Prieditis@llu.lv

Abstract
Tree species identification is important for a variety of natural resource management and monitoring activities especially in forest inventory. The objective of research is to identify tree species using digital aerial photography and LIDAR data in Latvian forest conditions. The study outlines a number of tree species identification posibilities: the ability to identify conifers and deciduous trees; the ability to identify pine and spruce; the ability to identify birch, aspen and black alder. The study site is a forest in the middle part of Latvia at Jelgava district (56º39’ N, 23º47’ E). Aerial photography camera (ADS 40) and laser scanner (ALS 50 II) were used to capture the data. LIDAR resolution is 9 points m^{-2} (500 m altitude). The image data is RGB, NIR and PAN spectrum with 20 cm pixel resolution. During the study a modified region growing algorithm was developed to determine tree canopy and tree species identification using threshold segmentation, Fourier transform, frequency filtering and reverse Fourier transform. Tree species classification of coniferous and deciduous trees is possible in 82% of the cases, the first storey of the trees can be classified correctly in 96% of the cases, but the second storey of the trees only in 49% of the cases. Spruce identification is possible in 81.1% of the cases, for first storey trees in 89.6% of the cases and for the second storey trees in 72.9% of cases. Deciduous tree correct classification is possible in 63% of the cases, birch 75%, black alder 60% and aspen only in 41% of the cases.
Key words: Forest inventory, laser scanning, aerial photography, data fusion.

Introduction

Tree species recognition is economically and socially important task of forest inventory. Traditionally recognition of tree species is conducted by a labor-intensive inventory on the field. However, these methods are costly, time-consuming and not applicable to large areas. Since remotely sensed data became applied in forestry, there have been numerous studies to classify tree species at different forest types (Wen et al., 2008; Korpela et al., 2009; Waser et al., 2010). Many researchers using different data sets have been studying spectral patterns that allow obtaining tree species characterizing data from the high-resolution aerial photography (Andersen et al., 2006; Heinzel et al., 2008). The spectral characteristics of tree species were studied at various scales from a single tree (Heinzel et al., 2008; Kim et al., 2008; Vastaranta et al., 2011) to a stand level (Thomas et al., 2006).

The most commonly used datasets in forest inventory based on remote sensing are RGB (Red, Green, and Blue) and NIR (Near Infrared) spectrum. Different studies have described tree species identification ((Heinzel et al., 2008; Kim et al., 2008) and tree location accurate determination (Hyyppä et al., 2008) using such data structure. Mainly required information is obtained from near-infrared captured image intensity, because it describes tree species specificity the best. Most researchers recognize that the range of intensity values cannot be precisely defined corresponding to individual species (Rossmann et al., 2007; Heinzel et al., 2008).There is a very large number of factors that may cause the value of the infrared spectral differences (Vaughn et al., 2012), for example, the exact time when a picture is taken, the weather conditions at the time the image was captured, the flight direction, exposure settings etc. Despite these problems, various studies have shown that using infrared spectrum it is possible to classify coniferous and deciduous trees successfully (Korpela, 2004; Leppänen et al., 2008).

Finnish researcher I. Korpela points out that, in order to raise tree species identification accuracy high resolution areal images are required (Korpela et al., 2009). The study also shows that in automatic tree species identification main source of error is a low-resolution aerial images (Korpela, 2004; Korpela et al., 2009).

Very promising results are obtained with a multispectral and hyperspectral imaging (Holmgren et al., 2008; Dinuls et al., 2011) where hundreds of spectral bands enabled a finer discrimination of spectral properties and have been applied to identify tree species. Factor analysis methods that reduce the high range of data and process only the part which is characterized by tree species are frequently used (Kim et al., 2008; Waser et al., 2010). The main disadvantages of such equipment use are relatively high cost and the quality of the data acquired. Numerous studies have shown that, despite the large number of spectra, spectral measurements tend to be the same for different species (Kim et al., 2008; Wen et al., 2008; Waser et al., 2010). However, R. Dinuls has conducted research in Latvia forests on species classification capabilities using a multispectral camera, he has been able to classify correctly five economically most important tree species with 97% accuracy (Dinuls et al., 2011). It should be noted that the study was conducted under ideal conditions, but in real practice a number of obstructive factors will influence the classification process and correct classification rate will be lower.

In S. Hatmi study eCognition program was used for automatic tree species identification from multispectral remote sensing data with 2 m spatial resolution. In result the species composition of forest area was determined with 78% accuracy (Hatami, 2012). This paper shows that the tree canopy compartment boundaries and species composition can be determined automatically using specialized software.

Practically for all researchers so far it has been difficult to determine the species in mixed forest stands. Automated identification of species with the individual tree method is still problematic, even in cases with the access to different types of data (Vauhkonen et al., 2008) is available.

Although coniferous and deciduous species can be distinguished using spectral properties, they have their own 3-dimensional crown structures which cannot be detected via passively sensed imagery data (Kim, 2007). Airborne laser scanning (LIght Detection And Ranging, or LIDAR) is one of the active optical remote sensing technologies that can provide highly accurate measurements of both the forest canopy and the ground surface. It provides data that makes it possible to isolate individual trees and using information about 3dimensional crown structures identifies species. However, the use of digital aerial photography and LIDAR data has limitations to distinguish tree species due to the lack of high spectral resolution or large number of spectral bands, or laser scanner parameter settings.

Data processing methods at different conditions work differently, mainly due to forest density, represented tree species and forest diversity, as well as LIDAR and digital aerial cameras technology specifics. Numbers of different methods are used to identify tree species using airborne LIDAR and digital aerial cameras. All of them can be divided into classification, clustering or segmentation approaches which can be subdivided in more details. Frequently used methods are nearest neighbor classification; image segmentation; threshold value segmentation; watersheds segmentation and segmenting of region extensions.

The objective of the research is to identify tree species using digital aerial photography and LIDAR data in Latvian forest conditions. The study outlines a number of tree species recognition variants: the ability to identify conifers and deciduous trees; the ability to identify pine and spruce; the ability to identify birch, aspen and black alder.

Materials and Methods

The study site is a forest in the middle part of Latvia at Jelgava district (56º39' N, 23º47' E). Represented species are Scots pine (*Pinus sylvestris* L.), Norway spruce (*Picea abies* (L.) H.Karst), silver birch (*Betula péndula* Roth), black alder (*Alnus glutinos* L.), and European aspen (*Populus trémula* L.). The area consists of mixed coniferous and deciduous forest with different age, high density, complex structure, various components, composition and soil conditions.

Remote sensing data were obtained using a specialized aircraft Pilatus PC-6, which is equipped with a digital aerial photography camera (ADS 40) and laser scanner (ALS 50 II). A LIDAR digital terrain models (DTM) were estimated from leaf-on data having 9 points m^{-2} at 500 m altitude. Flight was done on 19 May, 2010. The image data is RGB (Red, Green, and Blue), NIR (Near Infrared) and PAN (Panchromatic) spectrum with 20 cm pixel resolution.

Tree species accuracy was tested in 350 sample plots (0.045 ha). Only those trees which were recognized with the remote sensing methods were tested.

Differentially corrected Global Positioning System measurements were used to determine the position of each plot center. The accuracy of the positioning was approximately 1 meter. Tree locations within a plot were measured using a center as the origin and then tree azimuth and distance to the center were determined.

To determine tree canopy, an image segmentation was done. The aim is to find all the pixels that belong to the same tree. The method is based on a single tree segmentation using local maxima as tree tops, and a local minimum of the tree crown boundaries. Before the tree segmentation process, in order to improve the results obtained, it is necessary to perform data filtering, which reduces the shadow effect on the segmentation results. The filter, which was applied in the image preparation phase is modeled using band interaction parameters, the mathematical expression is as follows:

$$b'_1 = \frac{b_1}{\sqrt{b_1{}^2 + b_2{}^2 + b_3{}^2}} \tag{1}$$

where, b1, b2, b3 - a range of image intensity values; b'1 - is the image intensity spectrum of the filtered values.

During the study, a modified region growing algorithm was developed (shown in Figure 1). Algorithm starts with a tree center pixel detection process using a local peak detection method with

Fourier transformation. Fourier transform was performed to each image from the previously prepared data sets. After this process, the texture of image was obtained (noise and disturbance that in further processing may be falsely considered as local maxima were filtered). From each found tree center region extension is performed, which will compete with the neighboring regions, until the end of one of the conditions: no more free pixels; the radius of the region exceeds the maximum; none of the surrounding pixels color intensity values fit in 3 standard deviations area from the segment of the internal pixel color intensity values has been reached.

The result of this process is images with foliar segments (shown in Figure 1). For further use of the data, collected segments were converted to a GIS polygon and put into the database. After transformation data were filtered, allowing to get rid of segments with error.

Within the study species identification possibilities were analyzed using near-infrared and RGB aerial photographies, as well as the LIDAR data. During the

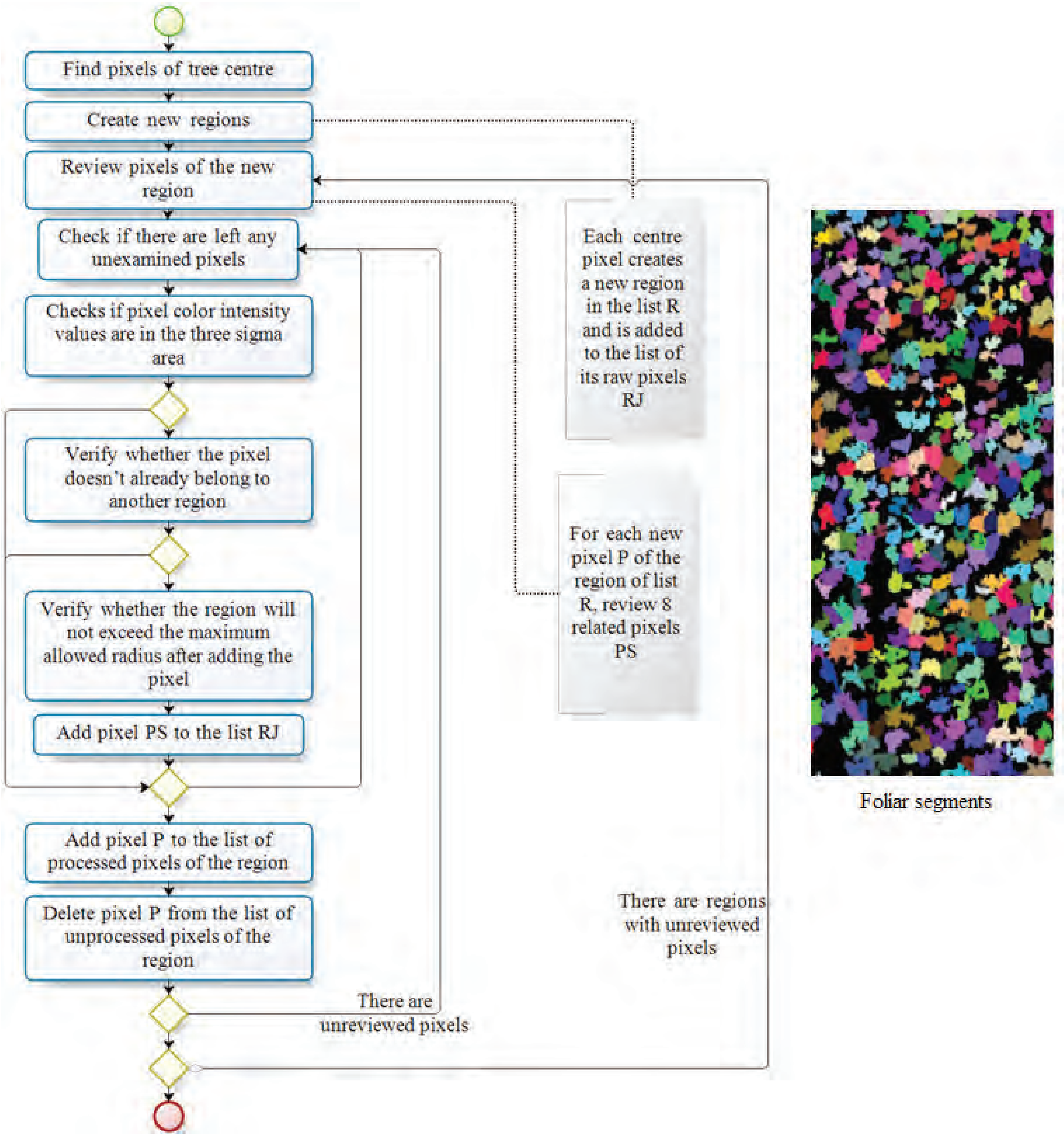

Figure 1. Region growing algorithm to determine tree canopy.

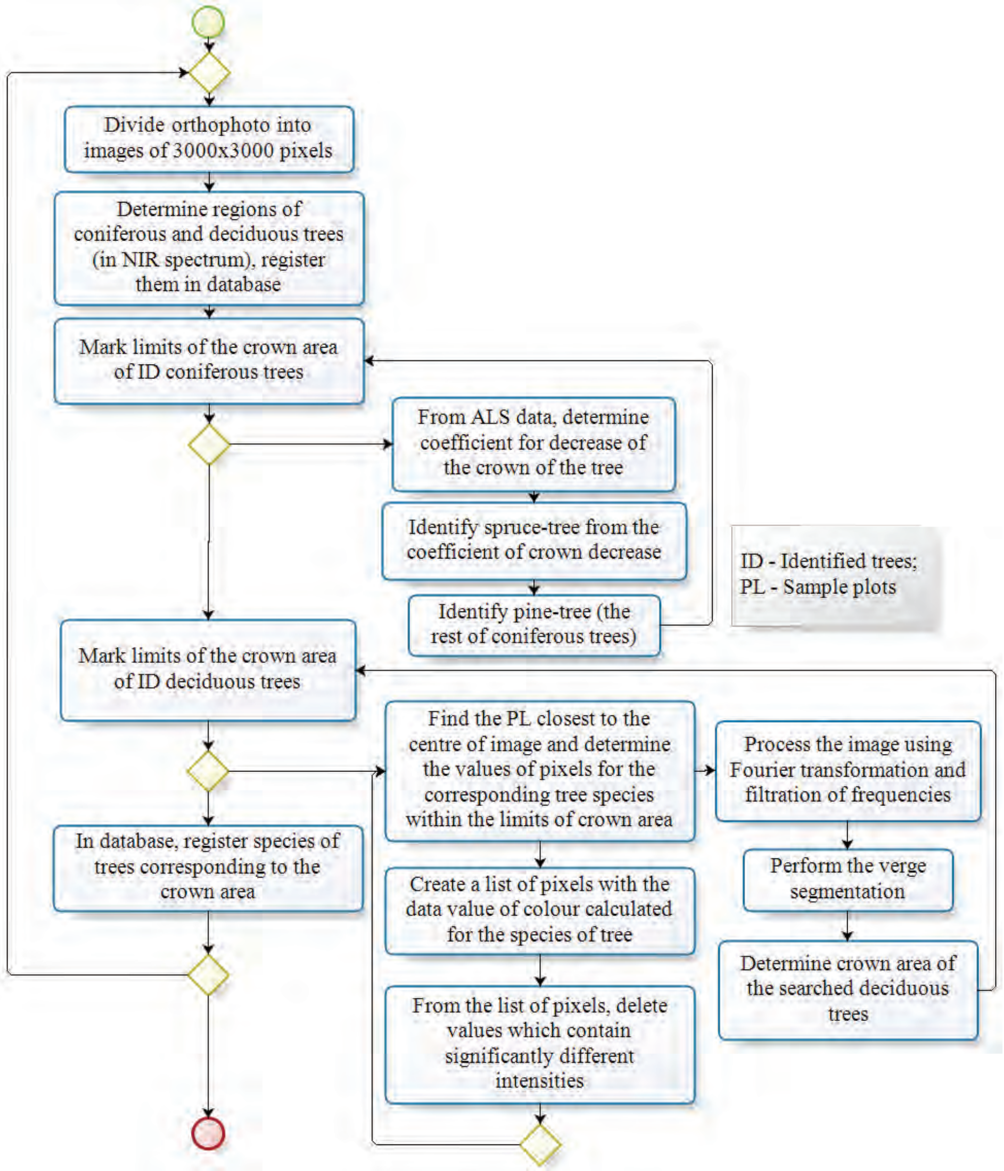

Figure 2. Tree species identification algorithm.

study species identification algorithm was developed (shown in Figure 2.).

Firstly it is necessary to define crown area segments and its boundaries. Then using NIR image coniferous and deciduous regions are defined according to the pixel values in the tree crown segment. Such a distribution of the two groups is simple, because within the image is the observed differences in the intensity of these two groups of species.

To divide coniferous trees in species level LIDAR data were used. All points owned by each tree were found and used to identify the species from tree crown slope factor. Following linear equation was applied to calculate tree crown slope factor:

$$y = mx + b \quad (2)$$

where, y – crown's top slope, degrees; m – slope factor; b – coefficient (in this case, zero).

Tree center point to the other points is considered as the zero point. Slope coefficient describes the height and cross-plane displacement ratio changes, following ratio was used:

$$m = \frac{Y1 - Y2}{X1 - X2} \quad (3)$$

where, Y – plane elevation changes, m; X – plane displacement changes, m.

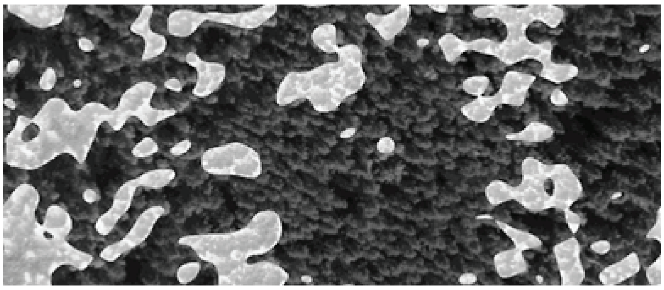

Figure 3. Threshold segmentation resulting image.

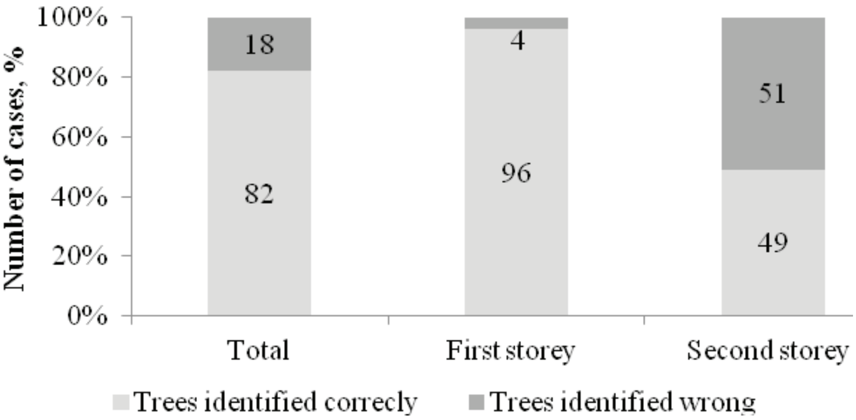

Figure 4. Results of tree classification of coniferous and deciduous trees.

Using a crown slope factor, it is possible to identify spruce, because spruce tops are more pointed than other tree species. This approach has been used for determined tree species spruce, but other conifers were automatically classified as pine.

In species identification within deciduous trees as the output material RGB image was used. The image size is important; too large image can make different color intensity for the same species, thus increasing the possibility of an error. The user is required to mark the sample pixels according to searched species. Output data are taken from the nearest plot against the image center. The resulting sample pixel list is checked and cleared of pixels with very different intensity values, for example, shadow affected and treetop highlighted pixels. Fulfilling all the necessary adjustments the image processing process starts whereby Fourier transform and frequency filtering allows to get rid of the noise. Threshold segmentation is performed to resulting image. The result is a binary (two-color) image (Figure 3.), which represents searched species (segments of the white areas). The threshold value is selected from prepared list of minimum color intensity.

This approach of segmentation allows defining regions with different intensities. In data processing it is necessary to point sample pixels manually according to required species, and this is a big disadvantage for this method.

Results and Discussion

Tree species classification of coniferous and deciduous trees is possible in 82% of the cases, the first storey of the trees can be classified correctly in 96% of the cases, but the second storey of the trees only in 49% of cases. Results of tree classification of coniferous and deciduous trees are shown in Figure 4.

A satisfactory precision level for identification of tree species of coniferous and deciduous trees is 95%, but it can be very difficult to be achieved, especially in mixed stands (Korpela et al., 2009).

The results of the crown slope factor for spruce and pine species level is shown in Figure 5, but crown slope factor for forest storey level (only spruce) in Figure 6. In both figures distribution range of slope factor is shown using amplitude. In this case, attention has been highlighted on the absolute values in order to understand better the properties of spruce crown slope factor. The age class range for pine and spruce is twenty years.

It is possible to recognize tree species using crown shape comparison method from LIDAR data using crown slope factor. Analysis of slope factor values

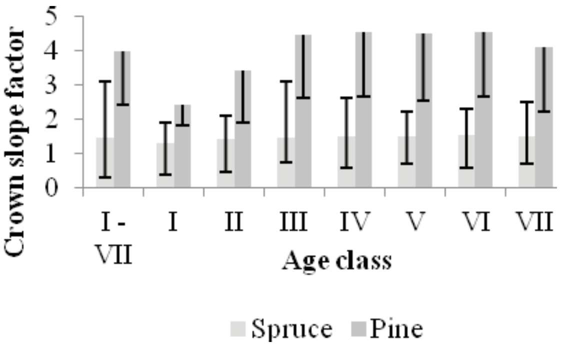

Figure 5. Results of the crown slope factor for spruce and pine species level.

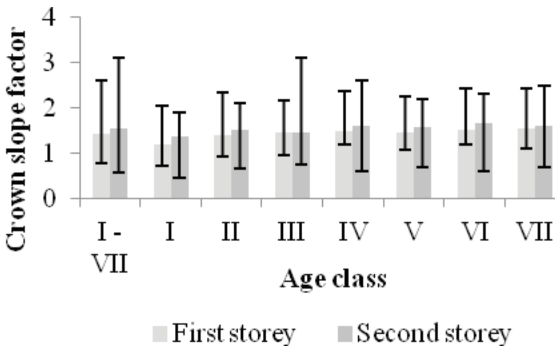

Figure 6. Results of spruce crown slope factor for forest storey level.

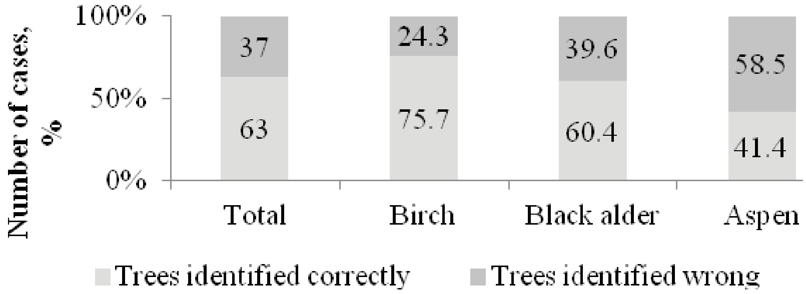

Figure 7. Results of deciduous tree classification.

shows that estimated coefficient at spruce level in 82% is not greater than 1.65; as a result, all trees that are less are considered to be spruce. When discussed young and middle-age stands, the larger range of errors are found. Identified tree species detection results show that method can give correct results in 81.1% of the cases, for the first storey trees in 89.6% of the cases and for the second storey trees in 72.9% of cases. Most trees which are not classified correctly are with small crown dimensions and are located under the dominant tree canopy.

Results of deciduous tree classification are shown in Figure 7. Deciduous tree species classification is possible in 63% of the cases, birch 75%, black alder 60% and aspen only in 41% of cases.

In literature using a similar approach the tree species identification results show variable results. The use of digital aerial photography and LIDAR data has limitations to distinguish tree species due to the lack of high spectral resolution or large number of spectral bands, or laser scanner parameter settings. In many works, the authors mention that the forest type and the dominant species are the main factors that affect tree species identification possibilities (Korpela et al., 2009; Waser et al., 2010). Almost all scientists up until now had problems to identify tree species in a mixed stands using both LIDAR data, as well as digital data of aero photographies. To improve tree species identification level is to perform aero photographing in spring when the forest is less dense and deciduous leafage is with high contrast that can be useful in tree species identification process. Also, either tree crown shape analysis from LIDAR data or multispectral data analysis can be used for better species recognition.

Conclusions

1. Out of all identified trees, classification of coniferous and deciduous tree is possible in 82% of the cases. The first storey trees can be classified correctly in 96% of cases but the second storey trees only in 49% of cases.
2. Deciduous species identification of the available data structure is complicated. Creation of an automated algorithm, during the course of the study that would be able to successfully perform classification of deciduous trees failed, which means that other data collection sources of remote sensing must be searched in order to be able to provide a more diverse information on the research object and to expand possibilities of RGB and CIR data structures.
3. Classification algorithm of deciduous trees showed weak results. Classification of deciduous trees at the level of species works satisfactory with

the dominant trees of the first storey. Results of the classification of species performed during the course of the study shows that 63% of deciduous trees were classified correctly. The species of silver birch was classified correctly in 75%, black alder in 60% but aspen only in 41% of the cases.
4. Other remote sensing data sources need to be used and that could provide more comprehensive information and complete RGB and CIR data structures opportunities.
5. The identification results of identified tree species of the stands for spruce shows that method applied can reach a correct result in 81.1% of the cases, for the first storey trees in 89.6%, but for the second storey trees in 72.9% of the cases.
6. Latvian forest conditions are difficult for single tree remote sensing methods mainly of mixed deciduous and coniferous spaces with high level of the second storey trees in one stand. Mostly trees are close together with high density and homogeneous crown. In most cases, the trees that are identified incorrectly are with small crown dimensions and are growing in the second storey of the stand. One of the error sources are the trees that are not being identified, but with their canopy are creating errors.

References

1. Andersen H.E., McGaughey R.J., Reutebuch S.E., Schreuder G., Agee J. (2006) The Use of Height - Resolution Remotly Sensed Data in Estimating Crown Fire Behavior Variables: Final Report JFSP Project 01-1-4-07. Available at: http://forsys.cfr.washington.edu/JFSP06/FINAL_REPORT_Proj_01_1_4_07.pdf, 16 January 2009.
2. Dinuls R., Lorencs A., Mednieks I. (2011) Performance Comparison of Methods for Tree Species Classification in Multispectral Images. *Electronics and Electrical Engineering*, 111, pp. 119-122.
3. Hatami F. (2012) *Carbon Estimation of Individual Trees Using High Laser Density Data of Airborne LiDAR*: Master's thesis. University of Twente. Enschede (Netherlands), pp. 64.
4. Heinzel J.N., Weinacker H., Koch B. (2008) Full automatic detection of tree species based on delineated single tree crowns - a data fusion approach for airborne laser scanning data and aerial photographs. In: *SilviLaser 2008 conference Proceedings*. Edinburgh (UK), pp. 76-85.
5. Holmgren J., Persson A., Söderman U. (2008) Species identification of individual trees by combining high resolution LiDAR data with multi□spectral images. *International Journal of Remote Sensing*, 29, pp. 1537-1552.
6. Hyyppä J., Hyyppä H., Leckie D., Gougeon F., Yu X., Maltamo M. (2008) Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. *International Journal of Remote Sensing: 3D Remote Sensing in Forestry*, 29, pp. 1339-1366.
7. Kim S., Schreuder G., McGaughey R., Andersen H.E. (2008) Individual tree species identification using LiDAR intensity data. In: *ASPRS 2008 Annual Conference Proceedings*. Oregon: American Society for Photogrammetry and Remote Sensing, pp. 12.
8. Kim S. (2007) *Individual tree species identification using LIDAR- derived crown structures and intensity data*: dissertation. University of Washington, pp. 137.
9. Korpela I., Tokola T., Ørka H.O., Koskinen M. (2009) Small-Footprint Discrete-Return LiDAR in Tree Species Recognition. In: *ISPRS Hannover Workshop 2009 High-Resolution Earth Imaging for Geospatial Information*, pp. 6.
10. Korpela I. (2004) *Individual Tree Measurements by Means of Digital Aerial Photogrammetry*: Academic dissertation. University of Helsinki, pp. 94.
11. Leppänen V.J., Tokola T., Maltamo M., Mehtätalo L., Pusa T., Mustonen J. (2008) Automatic Delineation of Forest Stands from LiDAR Data. In: *GEOBIA 2008 - Pixels, Objects, Intelligence. GEOgraphic Object Based Image Analysis for the 21st Century*. Vol. 38 – Part 4/C1, pp. 6.
12. Rossmann J., Schluse M., Bücken A., Krahwinkler P. (2007) Using Airborne Laser-Scanner-Data in Forestry Management: a Novel Approach to Single Tree Delineation. In: ISPRS Workshop on Laser Scanning 2007 and SilviLaser 2007: conference proceedings. Vol. 36, Part 3/W52, pp. 350-354.
13. Thomas V., Treitz P., McCaughey J.H., Morrison I. (2006) Mapping stand-level forest biophysical variables for a mixedwood boreal forest using lidar: an examination of scanning density. *Canadian Journal of Forest Research*, 36, pp. 34-47.
14. Vastaranta M., Holopainen M., Yu X., Haapanen R., Melkas T., Hyyppä J., Hyyppä H. (2011) Individual Tree Detection and Area – Based Approach in Retrieval of Forest Inventory Characteristics from Low – Pulse Airborne Laser Scanning Data. *The Photogrammetric Journal of Finland*, 22, pp. 1-13.

15. Vaughn N.R., Moskal L.M., Turnblom E.C. (2012) Tree Species Detection Accuracies Using Discrete Point Lidar and Airborne Waveform Lidar. *Remote Sensing*, 4, pp. 377-403.
16. Vauhkonen J., Tokola T., Leppänen V. (2008) Applied 3D texture features in ALS-based tree species segmentation. Available at: http://homepages.ucalgary.ca/~gjhay/geobia/Proceedings/Sessions/Session12/6724_Tokola_Proc.pdf, 16 January 2009.
17. Waser L.T., Baltsavias E., Ginzler C., Küchler M. (2010) Semiautomatic classification of tree species by means of multi-temporal airborne digital sensor data ADS40. In: *ISPRS Technical Commission VII Symposium*. Vol. 38 - Part 7B, pp. 633-638.
18. Wen Z., Baoxin H., Linhai J., Woods M., Courville P. (2008) Automatic Forest Species Classification using Combined LIDAR Data and Optical Imagery. In: *Geoscience and Remote Sensing Symposium*, pp. 134-137.