Remote Sensing of Environment 191 (2017) 1–12

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Automatic tree species recognition with quantitative structure models

Markku Åkerblom[a,*], Pasi Raumonen[a], Raisa Mäkipää[b], Mikko Kaasalainen[a]

[a] Department of Mathematics, Tampere University of Technology, P.O. Box 553, 33101Tampere, Finland
[b] Natural Resources Institute Finland (Luke), P.O. Box 18, Vantaa FI-01301, Finland

ARTICLE INFO

Article history:
Received 17 May 2016
Received in revised form 10 November 2016
Accepted 3 December 2016
Available online 19 January 2017

Keywords:
Tree species recognition
Terrestrial laser scanning
Quantitative structure model
Tree reconstruction

ABSTRACT

We present three robust methods to accurately and automatically recognize tree species from terrestrial laser scanner data. The recognition is based on the use of quantitative structure tree models, which are hierarchical geometric primitive models accurately approximating the branching structure, geometry, and volume of the trees. Fifteen robust tree features are presented and tested with all different combinations for tree species classification. The classification methods presented are k-nearest neighbours, multinomial regression, and support vector machine based approaches. Three mainly single-species forest plots of Silver birch, Scots pine and Norway spruce, and two mixed-species forest plots located in Finland and a total number of trees over 1200 were used for demonstration. The results show that by using single-species forest plots for training and testing, it is possible to find a feature combination between 5 and 15 features, that results in an average classification accuracy above 93% for all the methods. For the preliminary mixed-species forest plot testing, accuracy was lower but the classification approach presented potential to generalize to more diverse cases. Moreover, the results show that the post-processing of terrestrial laser scanning data of multi-hectare forest, from tree extraction and modelling to species classification, can be done automatically.

1. Introduction

Large multi-hectare areas of forests with thousands of trees can now be measured quickly with terrestrial laser scanning (TLS) (Calders et al., 2015a). This kind of massive-scale remote sensing of trees requires that most, if not all, post-processing steps are done automatically. In addition to the geometric and volumetric data, an important piece of information that can be determined from the point clouds is the tree species. For example, the species have an effect on the greenhouse gas exchange of a tree (Meier et al., 2016), and measuring the change in biodiversity is related to the number of species and their distribution. Thus, automatic and reliable tree species recognition would be an essential step to make the massive-scale remote sensing from TLS data practical.

There are a number of published studies that use TLS data for tree species recognition: Haala et al. (2004) used the combination of TLS and high-resolution panoramic images to make a comparison of the bark textures of four trees. Their results show that the texture is a candidate for classification as it seems to stay similar in a stem, but differs between stems. However, the approach was not tested on a larger dataset nor was it automatic.

Puttonen et al. (2010) used TLS and hyperspectral data to classify 24 trees of three species with a support vector machine (SVM). The scanning was done indoors, so point cloud segmentation into trees was not required. The classification features included shape parameters computed from the TLS data and averaged reflectance values of the hyperspectral data. With a combination of 2 features from each dataset, the average classification accuracy was over 85% for all species. When using only a pair of TLS data features, the accuracy was over 70% for only 43% of the pairs, but the best classification accuracy was 95.8%.

Puttonen et al. (2011) combined mobile laser scanning (MLS) and hyperspectral data to classify 133 trees of 10 species with SVM. Individual trees were manually isolated from the point cloud. Similarly to Puttonen et al. (2010) , the classification features consisted of MLS-based shape parameters and per channel averaged spectral data. The results showed that MLS features on their own were able to separate coniferous and deciduous trees with 90.5%, and individual species with 65.4% accuracy. For the combination of MLS and spectra the percentages were 95.8% and 83.5%, respectively.

Vauhkonen et al. (2013) tested hyperspectral LiDAR (HSL) in laboratory conditions for classifying 18 spruce and pine trees. The classification accuracies varied between 78% and 89%. Different scans of the same trees were used for training and classification.

---

[*] Corresponding author.

http://dx.doi.org/10.1016/j.rse.2016.12.002
0034-4257/

Othmani et al. (2013) used the 3D texture of the bark of 230 trees of five species. In their approach a 30 cm long patch is manually isolated in a stem and its texture analysed using 2D signal processing techniques and a random forest (RF) classifier. The overall species recognition was 88%.

In a most recent study (Lin and Herold, 2016), 40 trees of 4 species were classified by using SVM and *explicit tree structure* parameters (ETS). In contrast to the shape parameters used by, e.g., Puttonen et al. (2011) , ETS parameters describe the actual shape of the tree stem or crown rather than the distribution of TLS samples. The authors refer to tree isolation details in Holopainen et al. (2013) but fail to state which of the methods was used, and thus the level of automation is unknown. At least the separation of stem and branch points is done interactively. The classification tests were done using the leave-one-out cross-validation (LOOCV) in two different scenarios, maximum and robust, with accuracies 90.0% and 77.5%, respectively. The authors state that the latter scenario is more likely to be suitable for real applications.

The above literature survey shows that the tree species recognition from TLS data has been the topic of only a few studies and in most cases it has been combined with other data sources to achieve sufficient classification accuracies. Furthermore, the sample sizes have been relatively small, and no fully automatic solution has been presented yet.

In this paper, we propose a proof-of-concept for fully automatic species recognition approach from TLS measurements. Rather than using 3D point cloud data directly for classification, trees are first reconstructed as quantitative structure models (QSM) (Calders et al., 2015b; Raumonen et al., 2013). Notice that the QSM reconstruction is done by using only the *xyz*-coordinates of the points and thus no intensity data, spectral information, photographs, or ultra-high resolution scans are required. The classification features are computed from the geometric and topological tree properties stored in the models, which means that we have more than three dimensions to work with. This enables the use of properties that have been hard or impossible to determine directly from the point cloud data. The proposed classification features are listed in Section 2.5.

For the species recognition, we tested three different classification methods with numerous feature sets to show their differences and suitability for the application. Namely, we tested *k*-nearest neighbours, multinomial regression, and support vector machine based approaches. The classification methods are presented in Section 2.4.

It has been shown that QSMs can be automatically computed in massive scale (Raumonen et al., 2015), and when combined with automated feature computations it makes the complete classification procedure fully automatic. To demonstrate the approach, three large, mainly single-species plots from Finland are used. In addition, two mixed-species forest plots, also from Finland, are used to demonstrate preliminary results from more heterogeneous stands. The three species are the most numerous in Finland and represent both deciduous Silver birch (*Betula pendula* Roth) and coniferous Scots pine (*Pinus sylvestris* L.) and Norway spruce (*Picea abies* [L.] Karsten). Forest plot and scanning setup details are presented in Sections 2.1 and 2.2, respectively. We present results in Section 3, and sum up in Section 5.

2. Materials and methods

2.1. The forest plots used for the demonstration

We have three large almost single-species forest plots and two unmanaged mixed-species stands, which have been scanned with TLS. One of the single-species plots is a systemically planted plot with only Silver birch trees and the other two are natural coniferous plots with Norway spruce and Scots pine trees. All three study plots are in Punkaharju, Finland, where annual mean precipitation is 600 mm and effective temperature sum 1300 dd (Merilä et al., 2014).

2.1.1. Silver birch plot

A Silver birch stand used in this study is a field experiment in Punkaharju, Finland (61°48′N, 29°18′E), established in 1999 to study within-stand differences among genotypes (22 genotypes micro-propagated from local trees) (Possen et al., 2014). Trees were planted on agricultural field with a planting distance of 2 × 2 m in 1999. In April 2008, 50% of the trees were harvested. At the time of the scanning in October 2014, the stand density was approximately 1000 trees per ha, height of the trees varied from 18 to 24 m, and diameter at the height of 1.3 m (DBH) was 10–17 cm.

2.1.2. Scots pine plot

The Scots pine dominated study plot in Punkaharju, Finland (61°46′N, 29°20′E) is conventionally managed forest. The latest thinning took place in 1994, thereafter only dead trees are removed. At the time of the TLS in October 2014, the stand age was 95 years, stem number was approximately 500 stems per ha, the DBH was 18–40 cm, and the height of trees 27–32 m. The stand grows on sub-xeric site and the average stem volume growth is 11 $\mathrm{m}^3\ \mathrm{ha}^{-1}\ \mathrm{yr}^{-1}$. Scots pine and Norway spruce dominated study plots belong to the European forest monitoring network established under the UN-ECE ICP programme (Derome et al., 2002; Merilä et al., 2014).

2.1.3. Norway spruce plot

The Norway spruce dominated stand of this study is conventionally managed forest on herb-rich site, where the latest thinning took place in 1994 and since then the site has been a part of forest monitoring programme. The stand is located in Punkaharju, Finland and the density was approximately 400 stems per ha, the DBH was 28–45 cm, and the height of trees 28–33 m. The average stem volume growth is 8.8 $\mathrm{m}^3\ \mathrm{ha}^{-1}\ \mathrm{yr}^{-1}$ (Merilä et al., 2014).

2.1.4. Mixed-species plots

The two mixed-species plots are located in Sipoo (60°28′N, 25°12′E) and Lapinjärvi (60°39′N, 26°7′E) in Southern Finland. These sites are unmanaged Norway spruce dominated forests (>70% of standing volume), where other tree species were also present. At the time of the scanning in 2014, the stand density was approximately 1300 trees per ha in Sipoo and 1000 trees per ha in Lapinjärvi. For further details, see Rajala et al. (2012).

2.2. Terrestrial laser scanning

The scanning of all forest plots was performed with a RIEGL VZ-400 scanner and a 0.04° resolution. The Silver birch (leaf-off) and Scots pine plots were scanned completely on October 21st, 2014. The scanning of the Norway spruce plot was started on the same day and completed on November 26th, 2014. On both days the weather was cloudy with no rain and light wind, and the temperatures were −1°C and +1°C, respectively. The approximate scanning times were 1, 3 and 4 (2+2) h for the Silver birch, Scots pine and Norway spruce plots, respectively. The Sipoo plot was scanned on November 20th, 2014 and Lapinjärvi plot on November 24th, 2014. On both days the temperature was close to 0 ° C. The number of scanning points per forest plot was selected during the measurements based on visibility in order to cover most of the trees in the scans.

Retroreflectors were attached to tree stems to enable co-registration, which was later performed with the RiScan Pro software. The number of points in the scans, initially and after plot restriction and filtering, were the following: 94 and 35 million for the Silver birch plot, 300 and 58 million for the Scots pine plot, and 340 and 116 million for the Norway spruce plot. For Sipoo