FUSION OF HYPERSPECTRAL AND LIDAR DATA FOR TREE SPECIES CLASSIFICATION

Tomohiro Matsuki[1], Naoto Yokoya[2] and Akira Iwasaki[2]

[1]Department of Aeronautics and Astronautics, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
matsuki@sal.rcast.u-tokyo.ac.jp
[2]Research Center for Advanced Science and Technology, The University of Tokyo
4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan
{yokoya, aiwasaki}@sal.rcast.u-tokyo.ac.jp

ABSTRACT

Classification of tree species is one of the most important applications in remote sensing. In this study, the authors propose a methodology to classify tree species using hyperspectral and LiDAR data. The method consists of shadow correction, individual tree crown delineation and classification by support vector machine (SVM). Shadows in hyperspectral data are modified by unmixing. Individual tree crown delineation is achieved by a local maxima and region growing method for a LiDAR derived canopy height model (CHM). The input variables of SVM classifiers are principal components of hyperspectral data and the canopy form (height and size). The authors applied this method to the hyperspectral and LiDAR dataset taken over Tama Forest Science Garden in Tokyo and classified the data into 19 classes. As a result, we achived classification accuracy of 68 %, which is 20 % higher than what is obtained by using hyperspectral data only.

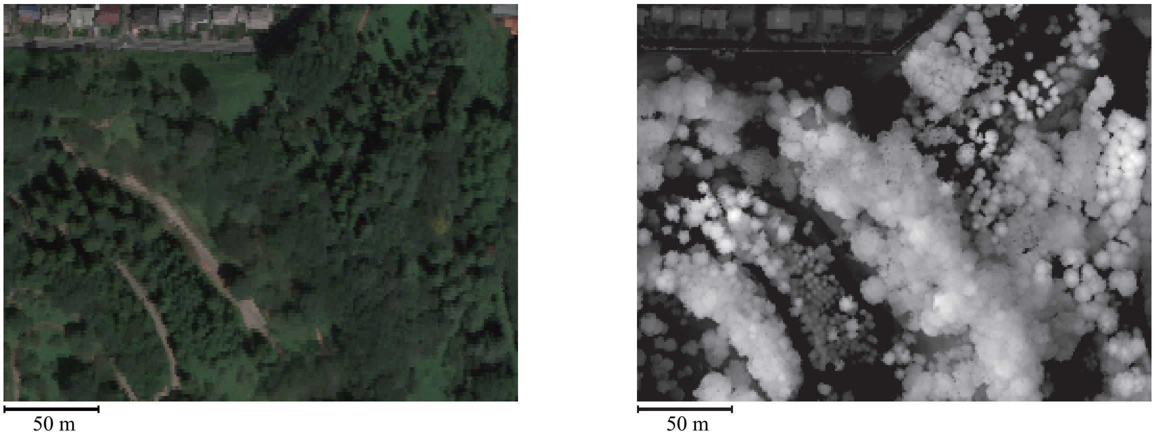
Keywords: classification, forest, data fusion, hyperspectral image processing, LiDAR

1 INTRODUCTION

Tree species classification is an important issue on the study of land cover or forest management. Periodic field surveys by experts, however, cost a large amount of time and human resources. Remote sensing by satellites can observe wide area at one time and provides time series data. For the past several decades, many researchers have performed classification or identification of tree species. Since hyperspectral sensors have hundreds of observation bands and high spectral resolution, we can obtain a continuous spectrum that enables more detailed analysis. In addition to spectral data, light detection and ranging (LiDAR) data is very informative. In recent years, fusion of these two data for classification has been studied (Dalponte 2008). In this study, we propose a methodology to classify tree species using hyperspectral and LiDAR data. We apply the method to forest data and show that it works well.

2 STUDY AREA

The study area is Tama Forest Science Garden in Tokyo in Japan. The hyperspectral and LiDAR data are provided by ERSDAC (Japan Space Systems now). The hyperspectral data are taken by CASI-3 sensor on 10 September, 2009. The observation wavelength range is 400-1050 nm with 72 bands and the spatial resolution is 1 m. The spatial resolution of LiDAR data is 0.5 m. To unify the resolution of the both data, LiDAR data was downsampled. Canopy height model (CHM) is the difference of digital elevation model (DEM) and digital surface model (DSM), and represents the height of trees excluding terrain effect. Figure 1 shows the RGB and CHM images. with the size of 268×207 pixels. Field survey has been conducted in this area and the species and the crown shapes of some trees are known (Odagawa and Okada 2009). There are more than 90 species of



(a) RGB image (b) CHM

Figure 1: Study area

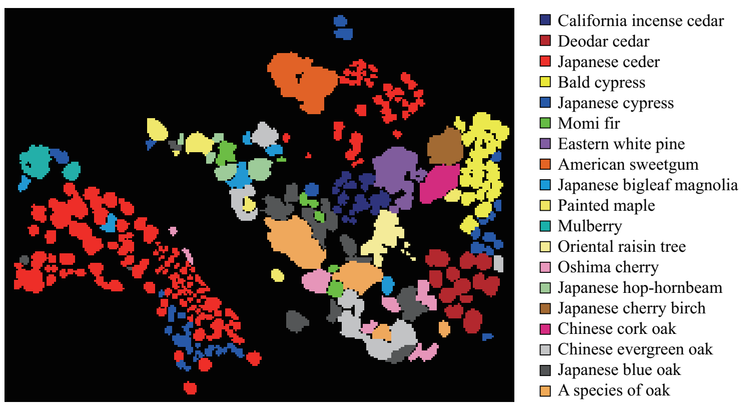


Figure 2: Ground truth

trees in this area. However, many species occupy few pixels. We chose 19 species that have enough ground truth data. Figure 2 shows the distribution map of the ground truth and Table 1 shows the class names and the number of samples.

3 METHODOLOGY

The proposed method consists of three parts: (1) shadow correction, (2) individual tree crown delineation, and (3) classification by support vector machine (SVM). First, we modify shadows in hyperspectral data by unmixing. Next, we delineate individual tree crown in order to use the morphological information of trees. Finally, we classify the data by SVM classifiers. Figure 3 shows the classification flow.

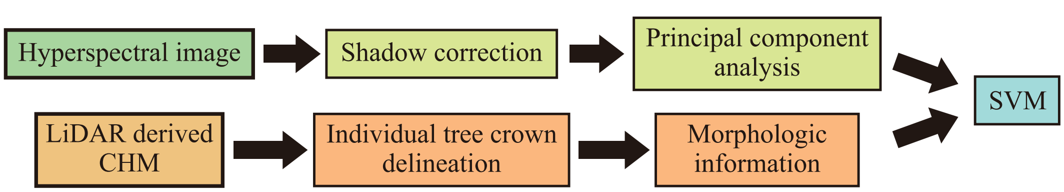


Figure 3: Classification flow

Table 1: Class names and the number of samples

Class	English name	Scientific name	Samples
1	California incense cedar	Calocedrus decurrens	362
2	Deodar cedar	Cedrus deodara	749
3	Japanese ceder	Cryptomeria japonica	3053
4	Bald cypress	Taxodium distichum	875
5	Japanese cypress	Chamaecyparis obtusa	721
6	Momi fir	Abies firma	256
7	Eastern white pine	Pinus strobus	551
8	American sweetgum	Liquidambar styraciflua	723
9	Japanese bigleaf magnolia	Magnolia obovata	308
10	Painted maple	Acer pictum	359
11	Mulberry	Morus bombycis	285
12	Oriental raisin tree	Hovenia dulcis	347
13	Oshima cherry	Cerasus speciosa	478
14	Japanese hop-hornbeam	Ostrya japonica	270
15	Japanese cherry birch	Betula grossa	297
16	Chinese cork oak	Quercus variabilis	297
17	Chinese evergreen oak	Quercus myrsinifolia	876
18	Japanese blue oak	Quercus glauca	991
19	A species of oak	Quercus serrata	959

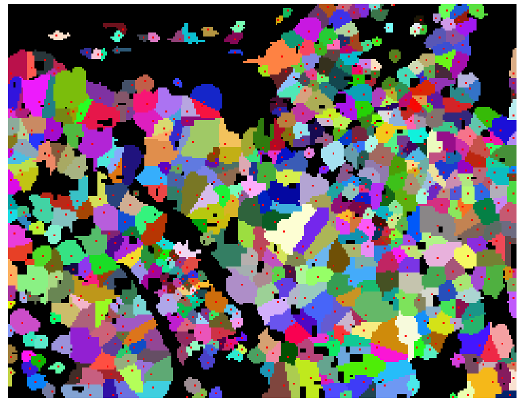
3.1 SHADOW CORRECTION

There are many shadows in forest. They affect classification results. Therefore, shadows in hyperspectral data need to be modified for accurate classification. We used the standard unmixing-based approach for de-shadowing of reflectance data (Boardman 1993). The shadow is defined as a "black" (zero reflectance) endmember and added to endmember spectra detected by the vertex component analysis (VCA). Abundance fractions are estimated using the fully constrained least squares method (FCLS) satisfying sum-to-one and non-negativity constraints. The reflectance spectrum is approximately de-shadowed by dividing by (1−Shadow abundance fraction). Figure 4(a) shows the de-shadowed image of the study area.

3.2 INDIVIDUAL TREE CROWN DELINEATION

Individual tree crown delineation is one of the important techniques to extract information of forests. Many researchers have investigated individual tree crown delineation methods (Erikson and Olofsson 2005). Pollock (1996) developed template matching method and Erikson (2003, 2004) developed region growing method. In this study, we used a simple method based on region growing method for LiDAR derived CHM (Koch et al. 2006). First, to reduce noise, we apply Gaussian smoothing filter to the CHM. The window size is 3×3. Then, we find local maxima of the smoothed image. They include non tree objects. Thus the local maxima whose normalized difference vegetation indices (NDVI) are lower than 0.8 or heights are lower than 3 m are excluded. The rest of them correspond to the tops of trees. Regions corresponding tree crowns grow from the local maxima according to region growing rules. If the neighboring pixels satisfy some conditions, regions grow and the neighboring pixel become the next starting pixel. The conditions are set that the height of neighboring pixel is higher than 1 m and lower than that of the starting pixel so that regions do not expand to the ground. This growing step is repeated until there is no starting pixel. After the growing step, overlapped regions are assigned to the region of the nearest local maximum. Figure 4(b) shows the result of individual tree crown delineation of the forest data. The colors are assigned at random and the red dots represent the tops of trees.

(a) De-shadowed RGB image



(b) Individual tree crown delineation

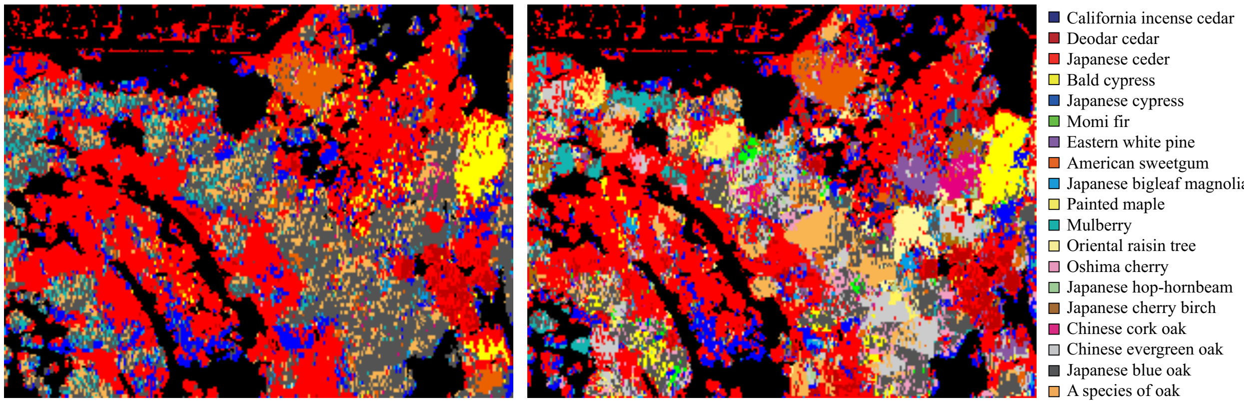
Figure 4: Preprocessed image

3.3 CLASSIFICATION BY SVM

We classify data by kernelized SVM classifiers that are nonlinear classifiers. The input variables of SVM are 15 principal components of the modified hyperspectral data and the morphological data, namely, the canopy heights and sizes. The heights and the sizes are the highest values of CHM and the number of pixels in each crown respectively. If a pixel is not included in any crowns, the height is its own value and the size is 1. The kernel function is radial basis function (RBF) kernel $k(x_1, x_2) = \exp(-\frac{|x_1 - x_2|}{\sigma})$. The bandwidth σ is decided by cross-validation. We coded this method in MATLAB and LIBSVM, which is the library for support vector machine (Chang and Lin 2011). It supports multi-class classification using one-against-one method.

4 EXPERIMENTAL RESULT

We applied our method to the forest data. The training set, whose size is given as a percentage to all ground truth, was selected at random but common to compared methods in one test. Figure 5 shows the classification map when the training set size is 10 %. It is compared with the method using original hyperspectral data only. Black regions, where heights are lower than 1 m and NDVI are lower than 0.7, are thought to be not trees. Our method improves the result of the central area, most of which are misclassified into Japanese blue oak and a species of oak by hyperspectral-only method. Table 2 shows the confusion matrix at the test. Most of the samples are classified into Japanese incense ceder, Japanese blue oak, and a species of oak, and few samples are classified into California incense ceder, Momi fir, Eastern white pine, Japanese hop-hornbeam, and Japanese cherry birch by hyperspectral-only method. In contrast, our method can classify the samples well. This result shows that the morphological information is useful for tree species classification. Nevertheless, species of oak (17, 18, 19) tend to be confound with one other. One reason for the high accuracy of Japanese incense ceder and the low accuracy of California incense ceder, Momi fir, Japanese bigleaf magnolia and Japanese hop-hornbeam is thought to be the difference of the number of samples. Figure 6 shows the average of the accuracy of each class in multiple tests. The accuracy of our method is higher than hyperspectral-only in most of the classes. The overall accuracy is shown in Figure 7, where the training set size were changed from 1 % to 20 %. It is compared with three methods, which are using original hyperspectral data only, using shadow corrected (SC) hyperspectral data only, and using SC hyperspectral and raw CHM data. It shows that the shadow correction increases the accuracy by 4 %, and that the accuracy of our method is 16 % higher than the method using hyperspectral data only. Additionally, our method shows higher performance than the method using hyperspectral and raw CHM data.



(a) Hyperspectral-only method (b) Proposed method

Figure 5: Classification map

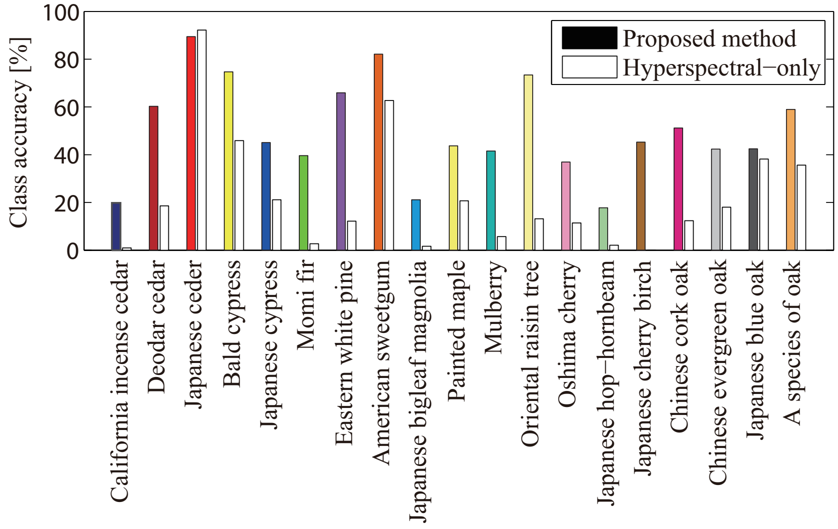


Figure 6: Class accuracy

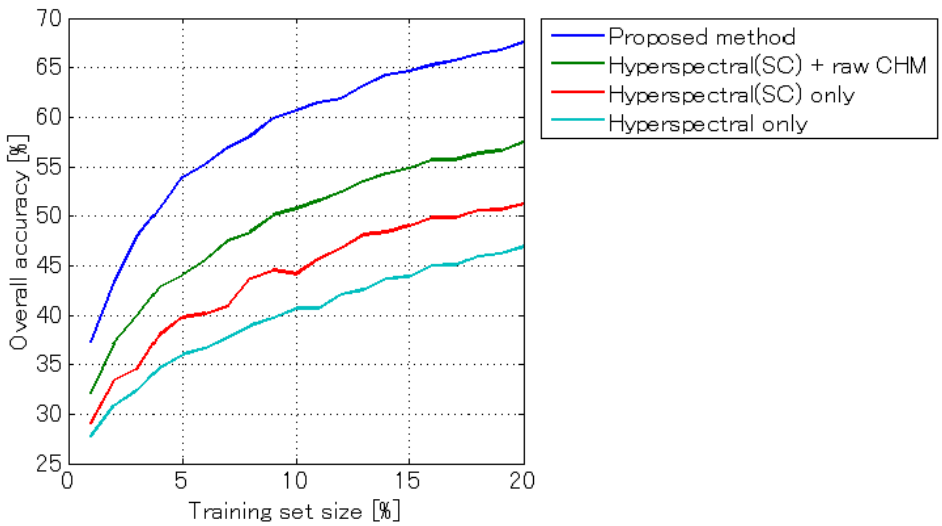


Figure 7: Overall accuracy

Table 2: Confusion matrix

(a) Hyperspectral-only method

Name	Class	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Accuracy
California incense cedar	1	4	4	260	19	30	0	0	22	0	0	0	10	0	0	0	0	6	7	0	1.10
Deodar cedar	2	0	211	507	14	11	0	0	0	0	0	0	0	0	0	0	0	3	3	0	28.17
Japanese ceder	3	3	84	2777	45	40	0	0	15	5	3	0	10	7	0	0	0	18	37	9	90.96
Bald cypress	4	0	12	299	412	37	0	0	10	0	0	0	17	0	0	0	0	36	48	4	47.09
Japanese cypress	5	0	5	415	27	194	0	0	1	0	0	1	13	0	0	0	0	23	19	23	26.91
Momi fir	6	0	0	162	9	9	0	0	8	0	0	0	7	0	0	0	0	17	29	15	0.00
Eastern white pine	7	0	12	358	94	52	0	0	0	1	0	0	16	0	0	0	0	4	13	1	0.00
American sweetgum	8	0	0	96	28	0	0	0	494	0	0	0	1	0	0	0	2	64	32	6	68.33
Japanese bigleaf magnolia	9	0	0	85	4	11	0	0	4	31	4	0	38	2	0	0	0	39	68	22	10.06
Painted maple	10	0	2	43	5	48	0	0	0	10	16	3	16	8	0	0	0	18	166	24	4.46
Mulberry	11	0	4	37	7	4	0	0	2	8	0	16	12	11	0	0	0	20	155	9	5.61
Oriental raisin tree	12	0	2	38	12	19	0	0	0	6	0	0	68	0	0	0	0	42	111	49	19.60
Oshima cherry	13	0	0	93	9	5	0	0	17	16	0	0	48	30	0	0	6	59	117	78	6.28
Japanese hop-hornbeam	14	0	0	26	4	6	0	0	0	1	2	0	3	0	0	0	0	28	148	52	0.00
Japanese cherry birch	15	0	2	67	16	2	0	0	8	5	0	0	31	0	0	0	4	13	103	46	0.00
Chinese cork oak	16	0	0	14	6	0	0	0	16	0	0	0	1	7	0	0	65	31	122	35	21.89
Chinese evergreen oak	17	0	0	89	36	7	0	0	61	6	0	0	32	4	0	0	5	156	331	149	17.81
Japanese blue oak	18	0	0	134	39	9	0	0	67	2	2	1	48	1	0	0	9	154	404	121	40.77
A species of oak	19	0	0	138	26	0	0	0	26	9	0	0	57	8	0	0	5	145	354	191	19.92

(b) Proposed method

Name	Class	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Accuracy
California incense cedar	1	61	8	144	0	45	0	31	23	3	0	0	7	0	0	36	0	0	0	4	16.85
Deodar cedar	2	1	473	221	4	8	0	26	0	1	0	1	0	0	0	1	0	2	3	8	63.15
Japanese ceder	3	13	77	2661	64	88	3	22	10	12	0	6	16	3	1	0	5	8	24	40	87.16
Bald cypress	4	0	14	69	640	59	7	6	2	13	0	0	0	1	9	2	3	25	22	3	73.14
Japanese cypress	5	9	21	177	8	377	5	23	0	0	0	2	10	5	7	34	5	6	21	11	52.29
Momi fir	6	11	2	12	4	11	89	0	2	1	1	2	0	12	0	11	0	15	58	25	34.77
Eastern white pine	7	15	55	76	0	32	0	324	1	5	2	0	7	2	2	5	4	17	3	1	58.80
American sweetgum	8	1	2	23	0	1	0	7	594	1	0	0	0	0	1	1	33	32	18	9	82.16
Japanese bigleaf magnolia	9	10	0	39	8	7	0	12	2	86	3	5	1	0	23	38	17	21	10	26	27.92
Painted maple	10	0	7	7	7	17	4	13	5	5	144	0	0	4	25	9	20	33	36	23	40.11
Mulberry	11	1	0	25	0	4	0	2	0	4	5	159	3	0	4	25	10	10	31	2	55.79
Oriental raisin tree	12	3	1	12	0	1	0	3	0	8	0	0	284	3	0	4	0	1	25	2	81.84
Oshima cherry	13	0	2	62	0	24	0	4	2	0	5	3	13	151	0	0	0	26	156	30	31.59
Japanese hop-hornbeam	14	1	0	8	5	1	2	6	0	12	18	2	5	0	54	19	28	73	31	5	20.00
Japanese cherry birch	15	23	5	9	3	2	0	13	4	11	0	0	9	0	5	152	47	13	1	0	51.18
Chinese cork oak	16	0	0	0	4	0	0	2	11	2	0	0	16	0	7	8	227	20	0	0	76.43
Chinese evergreen oak	17	6	0	44	6	7	0	10	60	13	0	2	1	38	3	26	36	392	156	76	44.75
Japanese blue oak	18	2	0	75	37	31	7	2	36	3	1	15	11	34	3	4	5	108	469	148	47.33
A species of oak	19	4	0	13	0	2	1	0	20	11	2	3	25	41	2	8	6	73	174	574	59.85

5 CONCLUSION

We proposed the novel methodology to classify tree species using hyperspectral and LiDAR data. The method uses spectral information derived from shadow corrected hyperspectral data and morphological information derived from LiDAR data. We applied it to the data taken over Tama Forest Science Garden in Tokyo in Japan. It was showed that shadow correction improved the classification result and our method had higher performance than the methods using hyperspectral data only or using hyperspectral and raw CHM data. Although hyperspectral data have a number of spectral information, it is thought to be difficult to classify tree species in detail using hyperspectral data only. Adding morphological information derived from LiDAR data led to the improvement. Using more suitable individual tree crown delineation method and introduction of parameters that represent more detailed shapes of trees will improve the performance. Additionally, the delineation have not been validated. They are future works.

REFERENCES

[1] Boardman J. W. , "Automating spectral unmixing of aviris data using convex geometry concepts," in *Summaries 4th Annu. JPL Airborne Geoscience Workshop*. JPL Publication 93–26, 1993, vol. 1, pp. 11–14.

[2] Chang C. and Lin C. , "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011, Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm (accessed on 18 August, 2013).

[3] Dalponte M. , Bruzzone L. , and Gianelle D. , "Fusion of hyperspectral and lidar remote sensing data for classification of complex forest areas," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 5, pp. 1416–1427, 2008.

[4] Erikson M. , "Segmentation of individual tree crowns in colour aerial photographs using region growing supported by fuzzy rules," *Canadian Journal of Forest Research*, vol. 33, no. 8, pp. 1557–1563, 2003.

[5] Erikson M. , "Species classification of individually segmented tree crowns in high-resolution aerial images using radiometric and morphologic image measures," *Remote Sensing of Environment*, vol. 91, no. 3—4, pp. 469 – 477, 2004.

[6] Erikson M. and Olofsson K. , "Comparison of three individual tree crown detection methods," *Machine Vision and Applications*, vol. 16, no. 4, pp. 258–265, 2005.

[7] Koch B. , Heyder U. , and Weinacker H. , "Detection of individual tree crowns in airborne lidar data," *Photogrammetric Engineering and Remote Sensing*, vol. 72, no. 4, pp. 357, 2006.

[8] Odagawa S. and Okada K. , "Tree species discrimination using continuum removed airborne hyperspectral data," in *Proceedings of IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing 2009*, Aug. 2009, pp. 1–4.

[9] Pollock R. J. , "The automatic recognition of individual trees in aerial images of forests based on a synthetic tree crown image model," Ph.D. thesis, 1996, AAINN14815.