USING SPATIAL POINT PATTERN ANALYSIS AS SUPPLEMENT FOR OBJECT-BASED IMAGE CLASSIFICATION OF TREE CLUSTERS

E. L. M. Tañada[1], A. C. Blanco[1,2]

[1]Phil-LiDAR 2, Training Center for Applied Geodesy and Photogrammetry
[2]Department of Geodetic Engineering, College of Engineering
University of the Philippines, Diliman, Quezon City, Philippines – elmtanada@gmail.com, ariel.blanco@coe.upd.edu.ph

KEY WORDS: trees, LiDAR, point pattern analysis, nearest neighbour, object-based image analysis, eCognition

ABSTRACT:

Single tree detection is important for monitoring tree plantations in order to estimate crop yield, which can be used for assessing food and biofuels. In the field of remote sensing, there have been studies regarding single tree extraction using Light Detection and Ranging (LiDAR) point cloud data. Canopy Height Model (CHM) can be simply derived from these LiDAR data where trees can be analysed using different physical characteristics such as tree crown width, and height. However, using these physical characteristics can't be very useful for identifying which trees are within a plantation, a forest, or a single tree. This is the motivation for developing algorithms of automated differentiation of random trees from plantation trees. Object Based Image Analysis (OBIA) was used to acquire the locations of individual trees. Trees were separated from other land cover in the image using a CHM derived from LiDAR point cloud data. Using watershed algorithm in eCognition, trees with good canopy spacing was easily distinguished. A spatial point pattern analysis was done from the extracted locations of trees to differentiate regularly patterned points from randomly distributed points. Nearest neighbour statistics was done to provide measurement of the distribution of the extracted points. The Z-Score parameter from nearest neighbor was then added to the layers in eCognition as supplement to the available height derivatives such as mean, standard deviation, and different textural parameters. Supervised type of classification was used in order to classify the trees on the image. Training samples of each tree from the image were entered in Support Vector Machine (SVM) to classify the different trees. Two sets of validation were done to compare the results of using the object parameters only, and the object parameters and Z-Score. Addition of the Z-Score from the object parameters increased the obtained accuracy on both validation sets. Classification accuracy based on validation set 1 and set 2 increased by about 20% and 6%, respectively. Additional spatial pattern analysis methods and validation can further prove the use of such layers in improving classification of trees.

1. INTRODUCTION

Based on different studies, classification of objects using object-based image analysis is dependent on the characteristics of each object. Mean, standard deviation and other textural parameters such as homogeneity and entropy are some of the information used. The methodology for classifying tree-dominant images used by Rollan, et al. (2015) utilized different LiDAR derivatives such as Normalized Digital Surface Model (NDSM), intensity, number of returns, and other textural data. The study achieved an overall accuracy of 87%. It was also stated that it is important to follow an iterative process in classifying using Support Vector Machine (SVM) while not getting the expected output, and it is important to create sufficient samples for different variations of a specific class.

This study aims to extend the defining characteristic of each tree outside its crown footprint by adding spatial pattern analysis on top of the individual object characteristics in order to classify clusters of trees accurately. Observing the distribution of the trees near the target tree might provide additional information to lessen the confusion on classifying a randomly distributed tree to a spatially patterned tree.

2. RELATED LITERATURE

2.1 Point Pattern Analysis

Point pattern analysis is the evaluation of pattern, or distribution, of a set of points on a surface. The analysis of spatial point patterns came to prominence in geography during the late 1950s and early 1960s (Gatrell, et al., 1995). Spatial point patterns can be characterized as ranging from dispersed to clustered, with random point patterns having elements of both dispersion and clustering (Robeson, et al., 2014). It has been proven that spatial pattern analysis can be used on different problems. Ek (1969) identified three forest patterns, clustered, random and regular, using spatial point pattern analysis. Ek obtained indices of nonrandomness (coefficients of randomization) for the spatial distribution of stems using a method described by Pielou (1959). A study by Li and Zhang (2007) compared different point pattern analysis methods for classifying the spatial distributions of spruce-fir stands. The author presented four methods of spatial point pattern analysis (nearest neigbor, refined nearest neighbour, ripley's K function, and pair correlation function analysis), and performed them on 50 separate plots which were patterned as complete spatial random (CSR), cluster, or regular. Though Li and Zhang successfully compared the four methods, they pointed out that they did not use height and crown measurements of each tree.

2.2 Support Vector Machine

A Support Vector Machine (SVM) is a discriminative classifier formally defined by a separating hyperplane. In other words, given labelled training data (supervised learning), the algorithm outputs an optimal hyperplane which categorizes new examples (Ervin, D., n.d.).

SVM has become widely used in image classification. Also, the performance of SVM is comparable to other supervised classifiers like Multilayer Perceptrons neural networks and k-

nearest neighbour statistical classifier (Roli, et al., n.d.). Mountrakis et al. (2010) stated in his review on SVM in remote sensing that SVMs outperformed even best neural networks when it comes to data with no normal or seen distribution, which is highly possible in remotely sensed images.

3. MATERIALS AND METHODOLOGY

3.1 Study Area

Study area is located in Nasipit, Agusan del Norte, Philippines. Selected area was based primarily on the available data, and the presence of different variations of trees and tree distribution. The study site includes coconut, mango, and banana plantations, and a few other trees (Figure 1).

Figure 1. Orthophotos showing widely spaced coconut trees (upper left), coconut trees (lower left), mango trees (upper right), randomly distributed trees (lower right)

3.2 Datasets

Light Detection and Ranging (LiDAR) point cloud, generated by the DREAM Phil-LiDAR 1 Program, was used for the classification of trees. The point cloud has a density of 2 points per sq. m.. The data was pre-classified into ground and non-ground. Orthophotos, which includes red, green and blue bands, were available but was only used for the selection of training and validation points. This is because the area contains four adjacent orthophotos which are uncalibrated, and using these images will result in confusion in the classification.

3.3 Methodology

3.3.1 LiDAR Derivatives

LiDAR point cloud was converted to a raster image using Lastools. Normalized Digitized Surface Model (NDSM) with a 0.5 meter resolution was derived from the point cloud.

3.3.2 Object Based Image Analysis (OBIA)

NDSM was imported in eCognition and it was used as the main image for OBIA. Roads, water, buildings, and power lines were initially segmented and classified. The following procedure was

then focused on the remaining unclassified objects, ground and trees. Unclassified objects were segmented using multiresolution segmentation with a large scale to extract cluster of trees (Figure 2). The scale parameter used for the segmentation was 100; it was tweaked until the observed objects are homogeneous. Watershed algorithm and local maxima filter in eCognition was used to determine location of trees (Wulder, et al., 2000, Rollan, et al., 2015). The locations of trees produced in eCognition are 1 pixel size each; these points were exported as a point shapefile.

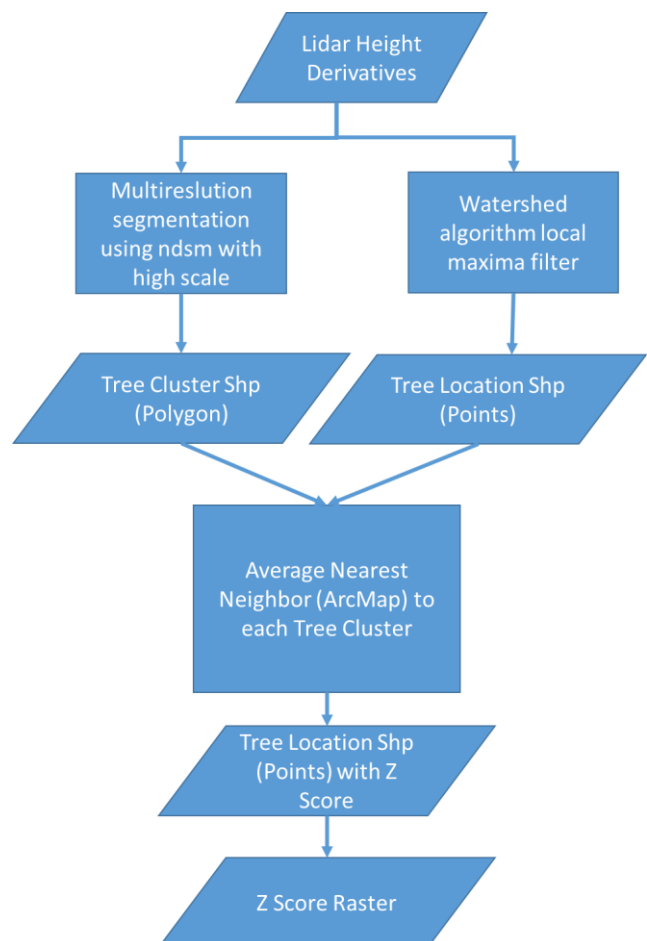Figure 2. Initial segmentation to extract clusters of trees.

Figure 3. Workflow for creating the Z-Score Raster.

3.3.3 Point Pattern Analysis

Tree cluster and tree location shapefiles were then imported in ArcMap. A tool was created using the ArcMap Model Builder to automatically iterate the Average Nearest Neighbor tool. This tool adds attributes in the cluster shapefile which shows the calculated nearest neighbour for each set of tree points inside each cluster. A raster image of Z-Scores of each cluster was created using the calculated attributes in the cluster shapefile (Figure 4), which was then imported in eCognition.
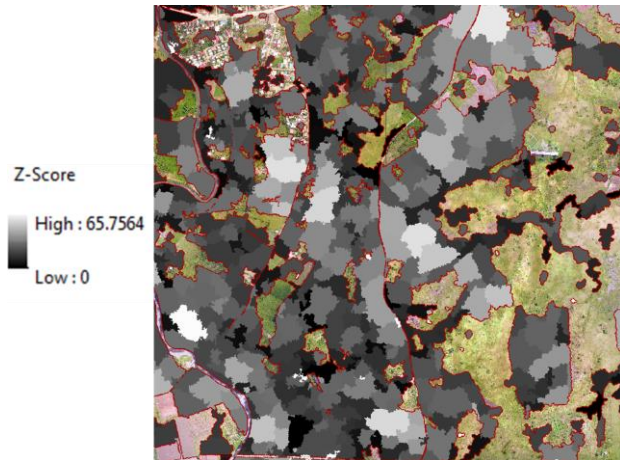

Figure 4. Z-Scores of each tree cluster.

3.3.4 Tree Classification

SVM was used for the classification of trees. The parameters that were used include Radial Basis Function kernel function, 200 for the C parameter, 1/(number of dimensions used) for the Gamma parameter (Carranza, et al., 2014). Training points were selected using the orthophoto. Around 50 samples per tree cluster type (coconut plantation, mango plantation, and randomly distributed) were selected for the classifier. Different combinations between NDSM (mean, standard deviation, homogeneity, and entropy) and Z-Score were tested to see if the addition of Z-Score is beneficial for the classification.

Combination A uses height statistics and height textures. And then, for combination B, Z-Score was added to combination A. For combination C, height texture layers were removed from combination B. Lastly, Z-Score was tested by itself.

Table 1. Combination of layers used in SVM.

Combination of Layers used in SVM	
Combination	Layers
A	NDSM (Mean, Standard Deviation, Homogeneity, Entropy)
B	NDSM (Mean, Standard Deviation, Homogeneity, Entropy), Z-Score
C	NDSM (Mean, Standard Deviation), Z-Score
D	Z-Score

There were two sets of validation samples created. Both sets were done similarly, but made sure the samples were not overlapping each other, and with the training samples. 20 samples for each tree cluster (coconut, mango, and random tree), for a total of 60 samples per set, were randomly selected for validation based on the orthophoto. Difference in accuracy was observed on the different combinations to analyse the addition of Z-Score on the SVM classifier. Also, the results from the two validation sets were compared to each other.

4. RESULTS AND DISCUSSION

Based on initial observation, Z-Score displayed a clear trend. Clusters with low Z-Score are those trees which were randomly distributed, and clusters with high Z-Score are trees arranged in patterns (Figure 5 and 6).
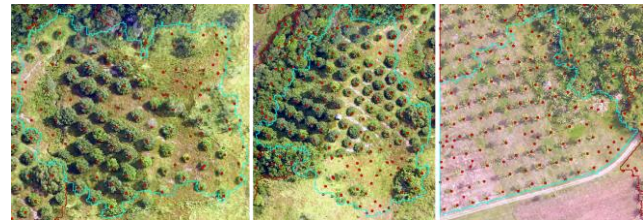

Figure 5. Tree clusters with low Z-Score

Figure 6. Tree clusters with high Z-Score

Table 2 shows the result of the accuracy of the classification using the two validation sets.

For validation set 1, accuracy of combination A reached 73.21%. Accuracy was raised to 93.76% when the Z-Score was added. And then, the accuracy went back down when the height textures was removed from combination B. Lastly, Z-Score alone scored the least accuracy of 74.96%.

Table 2. Overall accuracy of tree cluster classification.

Accuracy Assessment		
Combination	Validation Set	
	1	2
A	73.21%	81.08%
B	93.76%	87.45%
C	77.21%	79.14%
D	74.96%	55.81%

For validation set 2, accuracy of combination A resulted to 81.08%. Adding the Z-Score raised the accuracy to 87.45% which is similar to validation set 1. And then, it went down upon removal of the height textures. Z-Score by itself acquired the lowest accuracy.

Comparing the two validation sets, it was observed that in both of them, combination B reached the highest accuracy. But

using Z-Score alone attained the lowest accuracy. Removing either the height textural layers or the Z-Score from combination B resulted to a decrease in accuracy.

5. CONCLUSION AND RECOMMENDATIONS

The use of information on the distribution of a tree with respect to its neighboring trees, together with the statistical and textural parameters of the object, increased the accuracy of classification using SVM. Aside from the parameters derived from the point cloud data, extending the analysis outside the object can help in identifying and classifying trees by using the Z-Score of the tree clusters. Other point statistics layers may further improve the classification. Additional spatial pattern analysis methods and validation can further prove the use of such layers in improving classification of trees.

REFERENCES

Disaster Risk and Exposure Assessment for Mitigation (DREAM) Program. (2016). *The UP DREAM Program*. Retrieved June 29, 2016, from Disaster Risk and Exposure Assessment for Mitigation (DREAM) Program: https://dream.upd.edu.ph/about/

Ek, A. R., (1969). Stem Map Data for Three Forest Stands in Northern Ontario. Information Report 0-X-113, Canadian Forest Service, Department of Fisheries and Forestry.

Ervin, D. (n.d.). Point Pattern Analysis. Retrieved June 2016, from http://gispopsci.org/point-pattern-analysis/

Gatrell, A. C., Bailey, T. C., Diggle, P. J., Rowlingson, B. S., (1995). Spatial Point Pattern Analysis and its Application in Geographical Epidemiology. Transactions of the Institute of British Geographers, New Series 21 (1): 256-274

Jerez, M. V., Carranza, C. U., Gatdula, N. B., Rollan, T. M., Tañada, E. M., Pagkalinawan, H. M., et al. (2015). Phil-LiDAR 2 Project 1: Agricultural Resources Extraction from LiDAR Surveys (PARMap) Technical Narrative Report. Quezon City, Metro Manila, Philippines.

Li, F., Zhang, L., (2007). Comparison of Point Pattern Analysis Methods for Classifying the Spatial Distributions of Spruce-Fir Stands in the North-East USA. Forestry 80 (3): 337-349

Mountrakis, G., Ogole, C., & Im, J. (2011). Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 247-259.

Pielou, E. C., (1959). The Use of Point-to-Plant Distances in the Study of the Pattern of Plant Populations. The Journal of Ecology 47(3): 607-613

Robeson, S. M., Li, A., & Huang, C. (2014). Point Pattern Analysis on the Sphere. Spatial Statistics 10: 76-86

Roli, F., Fumera, G., (n.d.). Support Vector Machines for Remote-Sensing Image Classification. Retrieved June 2016 from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.11.5830&rep=rep1&type=pdf.

Rollan, T. A., Carranza, C., Jerez M., Blanco A.C., (2015). Combining Watershed Transformation and Local Maxima Approach in Developing a Tree Detection and Counting. Phil-LiDAR 2 Project 1: Agricultural Resources Extraction from LiDAR Surveys (PARMap) Technical Narrative Report. Quezon City, Metro Manila, Philippines.

Wulder, M., Niemann, K. O., Goodenough, D. G., (2000). Local Maximum Filtering for the Extraction of Tree Locations and Based Area from High Spatial Resolution Imagery. Retrieved June 2016 from http://www.sciencedirect.com/science/article/pii/S0034425700001012