International Journal of Applied Earth Observation and Geoinformation 50 (2016) 187–197

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

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

Combining QuickBird, LiDAR, and GIS topography indices to identify a single native tree species in a complex landscape using an object-based classification approach

Lien T.H. Pham[a,b,*], Lars Brabyn[a], Salman Ashraf[c]

[a] The University of Waikato, Department of Geography, Tourism & Environmental Planning, Private Bag 3105, Hamilton 3240, New Zealand
[b] Vietnam National University—Ho Chi Minh City, University of Science, Faculty of Environmental Science, 227 Nguyen Van Cu Str., Dist. 5, HCMC, Viet Nam
[c] GNS Science, P.O. Box 30368, Lower Hutt 5040, New Zealand

ARTICLE INFO

Article history:
Received 9 October 2015
Received in revised form 30 March 2016
Accepted 30 March 2016
Available online 6 April 2016

Keywords:
Object-based classification
Pohutukawa
Random forest
Support vector machine
QuickBird
LiDAR

ABSTRACT

There are now a wide range of techniques that can be combined for image analysis. These include the use of object-based classifications rather than pixel-based classifiers, the use of LiDAR to determine vegetation height and vertical structure, as well terrain variables such as topographic wetness index and slope that can be calculated using GIS. This research investigates the benefits of combining these techniques to identify individual tree species. A QuickBird image and low point density LiDAR data for a coastal region in New Zealand was used to examine the possibility of mapping Pohutukawa trees which are regarded as an iconic tree in New Zealand. The study area included a mix of buildings and vegetation types. After image and LiDAR preparation, single tree objects were identified using a range of techniques including: a threshold of above ground height to eliminate ground based objects; Normalised Difference Vegetation Index and elevation difference between the first and last return of LiDAR data to distinguish vegetation from buildings; geometric information to separate clusters of trees from single trees, and treetop identification and region growing techniques to separate tree clusters into single tree crowns. Important feature variables were identified using Random Forest, and the Support Vector Machine provided the classification. The combined techniques using LiDAR and spectral data produced an overall accuracy of 85.4% (Kappa 80.6%). Classification using just the spectral data produced an overall accuracy of 75.8% (Kappa 67.8%). The research findings demonstrate how the combining of LiDAR and spectral data improves classification for Pohutukawa trees.

1. Introduction

It is often important to map single tree species, such as when a tree has high ecological or cultural significance, and requires intensive management because it is under threat. Pohutukawa (*Metrosideros excelsa* Sol. ex Gaertn) is such a tree in New Zealand because it has been subject to fires and land clearance, and more recently possum browsing (Bylsma et al., 2014). Pohutukawa is a multi-stemmed tree up to 25 m high with large rounded crowns growing in northern coastal regions of New Zealand. Providing accurate information about the distribution of this species is necessary to help managers decide on appropriate conservation strategies.

Remote sensing and image analysis is advancing quickly with the capture of high spatial resolution data, which includes multispectral images as well as LiDAR. There have also been advances in data analysis techniques, including object based image analysis (OBIA), combining GIS terrain analysis, and advanced classifier algorithms. In the past, remote sensing of vegetation has focused on identifying broad vegetation classes, but advances in data and analysis techniques make it possible to identify specific vegetation species. LiDAR data produces accurate information on the vertical vegetation structure, which has been used for tree species classification (Kim et al., 2009; Ørka et al., 2009). LiDAR has also been combined with multispectral information to identify species (Cho et al., 2012; Dalponte et al., 2012).

OBIA has become increasingly popular over the last decade (Blaschke, 2010) because it provides a higher accuracy of classification compared to traditional pixel-based approaches (Ouyang et al., 2011). OBIA integrates spectral properties and spatial and contextual information into the classification process (Blaschke, 2010;

* Corresponding author.
E-mail addresses: lthp1@students.waikato.ac.nz, pthlien@hcmus.edu.vn (L.T.H. Pham), larsb@waikato.ac.nz (L. Brabyn), S.Ashraf@gns.cri.nz (S. Ashraf).

http://dx.doi.org/10.1016/j.jag.2016.03.015
0303-2434/

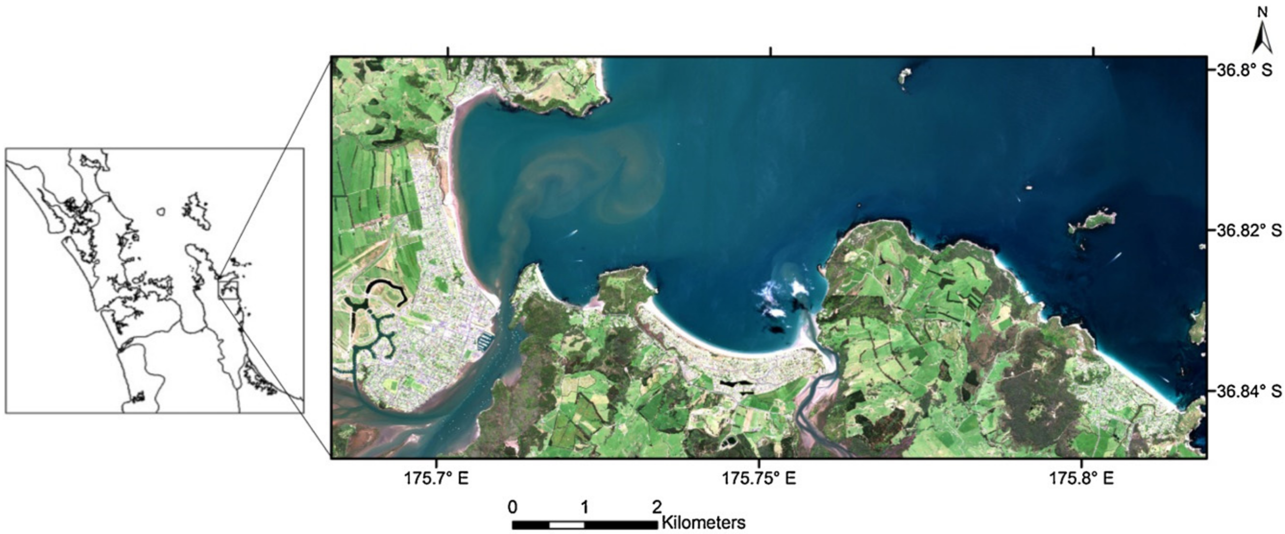

Fig. 1. QuickBird image of the Coromandel study area. The coordinate is in NZTM2000 projection system.

Han et al., 2014), and can be combined with multi-scale analysis to classify at regional and individual tree scales (Blaschke, 2010).

An important first step in OBIA is image segmentation which divides an image into contiguous, separate and homogeneous areas called image objects (Blaschke et al., 2004). For segmenting individual tree crowns, various automated methods have been developed and include: template matching (Korpela et al., 2007; Olofsson et al., 2006); valley following (Leckie et al., 2003); watershed segmentation (Chen et al., 2006); and region growing (Bunting and Lucas, 2006; Zhen et al., 2014). For some crown delineation algorithms the prior detection of treetops is required, which often uses the local maximum filtering technique with fixed or variable window sizes (Chen et al., 2006; Zhen et al., 2014). The local maximum technique is based on the assumption that treetops have the highest reflectance (multispectral images) or the highest elevation value (LiDAR data) within a tree crown. Using variable window sizes to identify treetops provides higher accuracy than a fixed window size (Gebreslasie et al., 2011).

The region growing method for crown delineation has outperformed other methods such as valley-following (Hussin et al., 2014) and template matching (Larsen et al., 2011) in both mixed and dense forests. The region growing method starts with a set of seed pixels (treetops), which are then merged to adjacent pixels that are similar (Ke and Quackenbush, 2011). This process of growing continues until a threshold is reached, and defined by specified homogeneity criteria (Blaschke et al., 2004). In this study, the support vector machine (SVM), a non-parametric classifier, was used because the number of training samples was small. With small training data sets, the SVM is the preferred classifier because it has good generalization ability (Mountrakis et al., 2011). In addition, a non-parametric classifier does not need an assumption of a normal distribution of the dataset; thus it is suitable for the integration of non-spectral data into a classification process (Lu and Weng, 2007). SVM can also produce more accurate classification results than other traditional parametric classifiers in a complex landscape (Dalponte et al., 2009). SVM algorithm finds the best decision boundary that separates the dataset into discrete classes with minimal misclassification (Mountrakis et al., 2011). A SVM can be nonlinear and linear, however, the nonlinear SVM is proving to be more accurate for nonlinear, complex classification problems (Izenman, 2008). An important pre-process for SVM is selecting relevant features, which improves the classification accuracy and computational efficiency (Huang and Wang, 2006).

Although recent studies on tree species mapping have used a combination of multispectral and LiDAR data with OBIA (and produced promising results), the combined technique requires further testing on a range of species, contexts, and input data, including the inclusion of additional GIS generated feature objects, such as the terrain wetness index. It is clear that humans use surrounding context information when manually identifying individual trees from an image, and it is well known that water is a key driver of vegetation distribution. It therefore makes sense that topographical indices well established in the GIS and ecological literature are included. This study therefore combines LiDAR and QuickBird imagery to: (1) develop an OBIA workflow for segmentation and classification of Pohutukawa trees, and (2) identify which object features are important based on classification accuracy. For comparison other broad classes of vegetation are also classified.

2. Materials

2.1. Study area

The research area is the Eastern side of the Coromandel region (see Fig. 1) at between 36°48′30“S to 36°47′30“S latitude, and 175°38′30“E and 175°47′30“E longitude. The total area of the study site is 1277.76 ha. The site is characterized by different land cover types including built-up area, urban parkland/open space, and both coniferous and broadleaf species.

2.2. Field data collection

Details of field data collected are shown in Table 1. The position of 560 trees were randomly selected and the species type recorded. Of these trees, 320 (57%) were used as training data and the remaining were used for accuracy assessment. Tree heights and crown diameters (mean of the N–S and E–W directions) of 90 trees were measured to determine the relationship between these variables, which are used for the treetop algorithm. These crowns were also manually mapped, which was required for assessing the segmentation accuracy.

2.3. Image data

Two main data sets were used–a QuickBird image and a LiDAR point cloud. The QuickBird multispectral image was cap-