International Journal of Applied Earth Observation and Geoinformation 59 (2017) 79–91

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

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

Research Paper

Object-based habitat mapping using very high spatial resolution multispectral and hyperspectral imagery with LiDAR data

Alex Okiemute Onojeghuo[a,*], Ajoke Ruth Onojeghuo[a,b]

[a] Department of Surveying and Geoinformatics, Nnamdi Azikiwe University, Awka, Anambra State, Nigeria
[b] Centre for Landscape and Climate Research, University of Leicester, Leicester, LE1 7RH, United Kingdom

ARTICLE INFO

Article history:
Received 15 January 2017
Received in revised form 11 March 2017
Accepted 14 March 2017

Keywords:
QuickBird
Eagle hyperspectral
OBIA
LiDAR
CHM
Intensity
SVM
kNN
Regression forest
Habitat mapping

ABSTRACT

This study investigated the combined use of multispectral/hyperspectral imagery and LiDAR data for habitat mapping across parts of south Cumbria, North West England. The methodology adopted in this study integrated spectral information contained in pansharp QuickBird multispectral/AISA Eagle hyperspectral imagery and LiDAR-derived measures with object-based machine learning classifiers and ensemble analysis techniques. Using the LiDAR point cloud data, elevation models (such as the Digital Surface Model and Digital Terrain Model raster) and intensity features were extracted directly. The LiDAR-derived measures exploited in this study included Canopy Height Model, intensity and topographic information (i.e. mean, maximum and standard deviation). These three LiDAR measures were combined with spectral information contained in the pansharp QuickBird and Eagle MNF transformed imagery for image classification experiments. A fusion of pansharp QuickBird multispectral and Eagle MNF hyperspectral imagery with all LiDAR-derived measures generated the best classification accuracies, 89.8 and 92.6% respectively. These results were generated with the Support Vector Machine and Random Forest machine learning algorithms respectively. The ensemble analysis of all three learning machine classifiers for the pansharp QuickBird and Eagle MNF fused data outputs did not significantly increase the overall classification accuracy. Results of the study demonstrate the potential of combining either very high spatial resolution multispectral or hyperspectral imagery with LiDAR data for habitat mapping.

1. Introduction

Recent advances in Earth Observation (EO) sciences have led to the increased use of multi-sensor derived data for land cover and habitat mapping across diverse landscapes. Hence, the approach of combining spectral information from very high resolution (VHR) multispectral and hyperspectral sensors with Light Detection Radar (LiDAR) information has recently become common practice in land-cover and vegetation mapping (Mücher et al., 2015). The combined use of optical imagery and LiDAR data has allowed for more detailed land-cover and habitat mapping at finer spatial resolutions, thereby meeting specific needs of end-users such as nature conservation managers and other stakeholders. Mücher et al. (2015) noted that the limitation of using only optical imagery was that reliable vegetation height information cannot be retrieved directly. However, the complementary use of LiDAR-derived measures (such as Canopy Height Model (CHM)) to optical EO data allows for the measurement of vegetation structure which in turn increases classification accuracy when used for mapping diverse or single habitat classes (Onojeghuo and Blackburn, 2011; Zhang, 2014).Previous studies have demonstrated the value of combining hyperspectral and multispectral imagery with LiDAR data to improve classification accuracy (Geerling et al., 2007; Hill and Thomson, 2005; Holmgren et al., 2008; Jones et al., 2010; Ke et al., 2010; Mücher et al., 2015; Rapinel et al., 2015; Zhang, 2014, 2015; Zhang and Qiu, 2012; Zhang and Xie, 2014). The benefits of integrating VHR multispectral imagery (such as QuickBird) with LiDAR data for mapping diverse vegetation types like forests (Bork and Su, 2007; Ke et al., 2010) or wetlands (Rapinel et al., 2015) across diverse landscapes have been extensively proven. Ke et al. (2010) evaluated the synergistic use of high spatial resolution QuickBird multispectral imagery (2.4m)and low-posting-density LIDAR data (3 m) for forest species classification using an object-based approach. Zhang (2014) explored a combination of hyperspectral imagery, LiDAR elevation and intensity features for vegetation mapping. The results indicated synergy of hyperspectral imagery with all LiDAR-derived features achieved the best classification results. However, limited studies have explored the integration of both multispectral and

* Corresponding author.
E-mail address: lexisgis@yahoo.com (A.O. Onojeghuo).

http://dx.doi.org/10.1016/j.jag.2017.03.007
0303-2434/

hyperspectral imagery with LiDAR-derived measures using different object-based classifiers and an ensemble of input classifiers.

Previous studies have demonstrated the value of combining very high spatial resolution multispectral (Holmgren et al., 2008; Ke et al., 2010) and hyperspectral imagery (Zhang, 2014) with LiDAR data for habitat mapping. The contribution of combining LiDAR-derived measures such as CHM, intensity and topographic information (DTM mean, maximum and standard deviation) with spectral information has shown to improve classification accuracy (Ke et al., 2010). However, limited studies have compared the performance of contemporary object-based machine learning classifiers and ensemble analysis with VHR multispectral and hyperspectral imagery as performed in this study. The use of LiDAR-derived elevation measures has shown to increase classification accuracy of habitat mapping (Ke et al., 2010; Zhang, 2014). In addition to LiDAR-derived height information, LiDAR intensity data has shown potential for feature identification (Onojeghuo and Blackburn, 2013) particularly whist performing land cover classifications with LiDAR-derived intensity (Song et al., 2002).

LiDAR-derived intensity measures the ratio of reflected light strength to that of emitted light and has shown to be of immense value in distinguishing different materials especially for land cover classification (Song et al., 2002; Zhang, 2014). In addition to LiDAR intensity, topographic statistical features have shown to increase classification accuracy. Given the homogeneous nature of topographic features (such as mean, maximum and standard deviation) within existing objects, issues of within-class variability can be reduced substantially when used as additional layers in the land cover classification. Zhang (2014) noted that topographic information could assist in reducing within-class variability among neighbouring objects resulting from shadows or gaps thus contributing to increased classification accuracy.

Over time, the limitation of sufficient spectral information contained in VHR multispectral imagery has been catered for in the advent of hyperspectral remote sensing. For this study, the AISA Eagle airborne hyperspectral data were used. A key concern in this study is developing a framework that can effectively utilise either multispectral or hyperspectral imagery with LiDAR derived measures using object based classifiers for habitat mapping.

The emergence of hyperspectral remote sensing techniques has shown to be an effective means of vegetation mapping as it provides hundreds of narrow spectral bands throughout the visible and infrared regions of the electromagnetic spectrum (Anderson et al., 2008; Buddenbaum et al., 2005; Christian and Krishnayya, 2009; Dalponte et al., 2012; Jones et al., 2010; Onojeghuo and Blackburn, 2011; Peerbhay et al., 2013; Pu et al., 2008; Vyas et al., 2011; Zhang and Xie, 2014). The application of hyperspectral remote sensing has become an integral approach for wetland (Adam et al., 2010) and forest (Jones et al., 2010)habitat mapping. During the analysis of remote sensing data, each band is considered a feature for class identification purpose. Hence, the salient question is whether all the bands contribute to the discrimination of the classes of interest. Feature reduction or selection of relevant band information is a critical step, particularly when analysing hyperspectral imagery. Since the advent of hyperspectral remote sensing in the mid-1980′s, feature reduction has become a more significant component of image interpretation process. The problems associated with hyperspectral data analysis and classification include high data volume, difficulties in displaying images and reading class statistics, an increase in processing time, biased class statistics estimates for small classes and the Hughes phenomenon (Hughes, 1968). In the classification of hyperspectral imagery, the amount of training data is always limited and the ratio of training pixels to the number of bands is usually small thus resulting in the Hughes phenomenon. Studies have demonstrated the value of dimension reduction techniques used in selecting or compressing hundreds of narrow spectral bands to fewer more relevant bands needed for the classification process (Anderson et al., 2008; Buddenbaum et al., 2005; Christian and Krishnayya, 2009; Jones et al., 2010; Peerbhay et al., 2013; Pu et al., 2008; Vyas et al., 2011). Amongst the dimension reduction techniques reviewed were Stepwise Discriminate Analysis (Vyas et al., 2011), Minimum Noise Fraction (MNF) transformation, Principal Component Analysis (PCA), Spectrally Segmented PCA and Sequential Forward Floating Selection. The MNF transformation, similar to PCA, is based on an orthogonal linear transformation that translates data to a new coordinate system such that the greatest variance is explained by the first set of coordinates, the second dimension variance by the second set of coordinates and so on (Jolliffe, 2002). Basically, the MNF transformation applies two cascaded PCAs, with the first transformation decorrelating and rescaling noise in data and the second transformation creating coherent eigenimages containing useful information and generation of noise-dominated eigenimages. For this study, the MNF transformation technique was used to select relevant spectral information contained in the Eagle hyperspectral imagery for land-cover classification. Studies have highlighted the advantage of the MNF method as it is able to reduce data dimensionality and remove the inherent noise contained in hyperspectral imagery (Zhang, 2014, 2015).

In place of traditional classifiers (such as the Maximum Likelihood and Minimum Distance) commonly used for pixel-based classifications, object-based classifiers were used in this study. The object-based machine learning algorithms included: Random Forest (RF), Support Vector Machine (SVM) and k-Nearest Neighbour (k-NN). These contemporary machine learning algorithms have shown to performed quite well when applied to both VHR multispectral (Dalponte et al., 2012) and hyperspectral (Zhang and Xie, 2014) imagery. In addition to applying the three object-based classifiers, an ensemble of all three classifiers was evaluated in this study.

The overall focus of this study was to investigate the feasibility of utilising spectral information and structural details as contained in spaceborne multispectral and airborne hyperspectral/LiDAR data to develop practical and effective approaches of implementing habitat mapping based on modern object-based classifier algorithms as against traditional pixel-based approach. We focused on developing a framework that combines high spatial resolution multispectral/hyperspectral imagery with LiDAR data for effective habitat vegetation mapping for a selected site in the UK. Past studies have advocated for the use of object based classifiers for habitat mapping, however the viability of adopting an approach that is cost effective in terms of implementation is one this study aimed to address. As earlier noted the, the challenges associated with handling large volume of spectral information contained in hyperspectral imagery is a major concern in land cover mapping related projects.

The use of data fusion and classifier ensembles have become common practice in remote sensing as demonstrated in several review papers (Du et al., 2012; Pohl and Van Genderen, 1998; Solberg, 2006).Hence, this study aimed to investigate the effects of fusing multispectral and hyperspectral datasets with LiDAR-derived measures for habitat mapping using object-based machine learning algorithms and ensemble analysis. The hypothesis of the analysis was that the fusion of either multispectral or hyperspectral data with LiDAR-derived measures using object-based approaches would improve habitat mapping accuracy. Also, an ensemble of all three classifiers used in the study is bound to outperform the performance of individual classifiers as it draws on the strengths of all classifiers used. In order for this to be achieved the key objectives included: