Remote Sensing of Environment 179 (2016) 66–78

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Mapping tree species in tropical seasonal semi-deciduous forests with hyperspectral and multispectral data

Matheus Pinheiro Ferreira [a,*], Maciel Zortea [b], Daniel Capella Zanotta [c], Yosio Edemir Shimabukuro [a], Carlos Roberto de Souza Filho [d]

[a] Remote Sensing Division, National Institute for Space Research, Av. dos Astronautas 1758, 12227-010 São José dos Campos, SP, Brazil
[b] Institute of Informatics, Federal University of Rio Grande do Sul, Av. Bento Gonçalves 9500, 91509-900 Porto Alegre, RS, Brazil
[c] National Institute for Science, Education and Technology, R. Eng. Alfredo Huch 475, 96201-460 Rio Grande, RS, Brazil
[d] Geosciences Institute, University of Campinas, R. João Pandiá Calógeras 51, 13083-870 Campinas, SP, Brazil

ARTICLE INFO

Article history:
Received 9 July 2015
Received in revised form 9 March 2016
Accepted 22 March 2016
Available online 25 March 2016

Keywords:
Brazilian Atlantic Forest
Imaging spectroscopy
WorldView-3
Classification
Individual tree crowns

ABSTRACT

Accurately mapping the spatial distribution of tree species in tropical environments provides valuable insights for ecologists and forest managers. This process may play an important role in reducing fieldwork costs, monitoring changes in canopy biodiversity, and locating parent trees to collect seeds for forest restoration efforts. However, mapping tree species in tropical forests with remote sensing data is a challenge because of high floristic and spectral diversity. In this research, we discriminated and mapped tree species in tropical seasonal semi-deciduous forests (Brazilian Atlantic Forest Biome) by using airborne hyperspectral and simulated multispectral data in the 450 to 2400 nm wavelength range. After quantifying the spectral variability within and among individual tree crowns of eight species, three supervised machine learning classifiers were applied to discriminate the species at the pixel level. Linear Discriminant Analysis outperformed Support Vector Machines with Linear and Radial Basis Function (RBF-SVMs) kernels and Random Forests in almost all the tested cases. An average classification accuracy of 70% was obtained when using the visible/near-infrared (VNIR, 450–919 nm) bands. The inclusion of shortwave infrared bands (SWIR, 1045–2400 nm) increased the accuracy to 84%. Narrow-band vegetation indices (VIs) were also tested and increased the classification accuracy by up to 5% when combined with VNIR features. Furthermore, the spectral bands of the WorldView-3 (WV-3) satellite sensor were simulated for classification purposes. WV-3 VNIR bands provided an accuracy of 57.4%, which increased to 74.8% when using WV-3 SWIR bands. We also tested the production of species maps by using an object-oriented approach that integrated a novel segmentation algorithm that was tailored to delineate tree crowns and label high class membership pixels inside each object. In this scenario, RBF-SVMs produced the best species maps, correctly identifying 84.9% of crowns with hyperspectral data and 78.5% with simulated WV-3 data. The use of a reduced set of hyperspectral bands, which were selected with stepwise regression, did not significantly affect the classification accuracies but allowed us to depict the most important wavelengths to discriminate the species. These wavelengths were located around the green reflectance peak (550 nm), at the red absorption feature (650 nm) and in the SWIR range at 1200, 1700, 2100 and 2300 nm. These encouraging results suggest the feasibility of the proposed approach for mapping pioneering and climax tree species in the Brazilian Atlantic Forest Biome, highlighting its potential use in forest recovery and inventory initiatives.

1. Introduction

Seasonal semi-deciduous forests (SSFs) are tropical forest formations that are subjected to a well-defined dry season, in which a portion of the trees defoliate. In Brazil, this ecosystem originally covered some 500,000 km^2, mainly in the southern and southeastern parts of the country (IBGE, 2004). SSFs are part of the Atlantic Forest Biome (Morellato & Haddad, 2000), which is internationally recognized as a biodiversity hot-spot for conservation priorities (Myers, Mittermeier, Mittermeier, Fonseca, & Kent, 2000). The fragmentation and degradation of the Atlantic Forest and its associated ecosystems have reached alarming levels. For example, the remaining SSF areas in the state of São Paulo in southeastern Brazil are estimated to be only 5.7% of its original extent (SOS/INPE, 2015). This scenario has motivated several ecological restoration initiatives in recent decades (Rodrigues, Lima, Gandolfi, & Nave, 2009). Nevertheless, such efforts face major challenges to become feasible and effective because of their high costs.

* Corresponding author.
E-mail addresses: pferreira.matheus@gmail.com (M.P. Ferreira), mzortea@gmail.com (M. Zortea), daniel.zanotta@riogrande.ifrs.edu.br (D.C. Zanotta), yosio@dsr.inpe.br (Y.E. Shimabukuro), beto@ige.unicamp.br (C.R. de Souza Filho).

http://dx.doi.org/10.1016/j.rse.2016.03.021
0034-4257/

Locating parent trees that provide seeds for seedling production is one of the first steps in forest restoration practices. Information regarding the spatial distribution of these trees may reduce fieldwork costs and improve intraspecific genetic diversity, which is considered a key element to the maintenance and evolution of forest systems (Lesica & Allendorf, 1999).

Currently, hyperspectral remote sensing, particularly in high spatial resolution, has been hailed as a promising technology for mapping tree species at the Individual Tree Crown (ITC) level (e.g. Baldeck et al., 2015, Féret & Asner, 2012, 2013, Naidoo, Cho, Mathieu, & Asner, 2012). Hyperspectral sensors that measure reflected radiation in hundreds of narrow bands can detect subtle variations in the chemical and structural attributes of the forest canopy.

However, discriminating tree species in tropical environments poses a set of technical issues. The high floristic diversity challenges the basic assumption that the spectral variability within species is significantly lower than the spectral variability among species. This assumption was verified at the leaf level (Castro-Esau, Sánchez-Azofeifa, Rivard, Wright, & Quesada, 2006; Féret & Asner, 2011; Ferreira, Grondona, Rolim, & Shimabukuro, 2013) but has been disputed and requires further investigation at the canopy level (Zhang, Rivard, Sanchez-Azofeifa, & Castro-Esau, 2006). Despite this problem, several studies have demonstrated the feasibility of mapping tree species in humid tropical forests (Carlson, Asner, Hughes, Ostertag, & Martin, 2007; Clark, Roberts, & Clark, 2005; Féret & Asner, 2013; Papeş, Tupayachi, Martínez, Peterson, & Powell, 2010; Shafri, Suhaili, & Mansor, 2007), but no previous work has been performed in SSFs. Similarly, recent studies have reported the successful identification of tree species in other ecosystems by using hyperspectral imagery (Alonzo, Bookhagen, & Roberts, 2014; Dalponte, Orka, Gobakken, Gianelle, & Næsset, 2013; Peerbhay, Mutanga, & Ismail, 2013).

In SSFs, the percentage of overstory trees that present total leaf loss varies from 20% to 50% during the dry season (Veloso, Rangel Filho, & Lima, 1991). Trees from the same species can show distinct leaf fall patterns depending on their age and environmental conditions (Morellato, 1991). The effects of this behavior on the spectral variability and, consequently, the classification accuracy are still poorly understood. Moreover, little is known regarding the performance of machine learning algorithms to discriminate tree species in SSFs. Linear Discriminant Analysis (LDA, Duda, Hart, & Stork, 2000) demonstrated good results for tropical and temperate forests (Clark et al., 2005; Dalponte, Bruzzone, Vescovo, & Gianelle, 2009; Féret & Asner, 2013). Non-parametric methods, such as Support Vector Machines (SVMs; Vapnik, 2000) and Random Forest (RF, Breiman, 2001), also proved suitable (Baldeck et al., 2015; Clark & Roberts, 2012; Féret & Asner, 2013; Naidoo et al., 2012).

Image processing techniques that can extract taxonomic information from hyperspectral data are in the early stages of development. Previous investigations have shown that object-oriented approaches improve discrimination accuracies (Clark et al., 2005; Féret & Asner, 2013). However, these approaches require automatic tree crown delineation, which is extremely difficult in tropical forests even when using canopy height models that are derived from Light Detection and Ranging (LiDAR) data (Tochon et al., 2015). Therefore, the production of reliable species maps in high-diversity tropical forests with object-based approaches has not yet been accomplished. The production of species maps is puzzling in those ecosystems, as hundreds of species compose the canopy and a large degree of uncertainty is obtained when classifying all the image pixels. Advances in this research field may provide valuable insights for ecologists and forest managers, contributing to floristic assessments and the monitoring of species of interest.

Hyperspectral sensors operate in different regions of the electromagnetic spectrum. The majority of studies that have been performed in tropical forests used sensors that operated in the visible and near-infrared domain (VNIR, 390–1000 nm) (e.g. Féret & Asner, 2013, Zhang et al., 2006). In this part of the spectrum, the signal-to-noise ratio is higher and atmospheric effects from water vapor absorption are lower compared to the short-wave infrared (SWIR, 1000–2500 nm) region. Even multispectral data that are acquired at the VNIR were demonstrated to be suitable to discriminate tree species. Recently, Cho, Malahlela, and Ramoelo (2015) demonstrated the utility of WorldView-2 imagery for tree species mapping in an African tropical forest. Advances in remote sensing technology enabled the acquisition of SWIR data in spectral bands of relatively narrow widths (<10 nm), which may improve species discrimination by enhancing non-pigment biochemical differences. Non-pigment leaf constituents, particularly water, nitrogen, cellulose and lignin, exhibit specific absorption features in the SWIR because of chemical bonds (Kokaly, Asner, Ollinger, Martin, & Wessman, 2009). Plants of different species may have various concentrations of these constituents, and their reflectance spectra are affected accordingly. SWIR sensing capabilities were also recently incorporated into high-resolution orbital platforms, such as the WorldView-3 satellite, opening new opportunities for vegetation mapping. Thus, evaluating the classification accuracy with VNIR individually and combined with SWIR features is important.

In this research, we performed tree species discrimination and mapping in tropical seasonal semi-deciduous forests by using real airborne hyperspectral and simulated multispectral data in the 450 to 2400 nm wavelength range. Three machine learning methods were tested (LDA, SVMs and RF), as were extensive and selected features from the VNIR range individually and combined with SWIR and narrow-band vegetation indices, to discriminate tree species at the pixel level. The role of SWIR bands in species classification was also assessed by using simulated WorldView-3 data. A novel tree crown segmentation approach that explores the rich spectral information of the hyperspectral data was presented to map tree species. Finally, we investigated the effects of within- and among-species spectral variability on the classification accuracy.

2. Materials

2.1. Study area

The study area is the St. Genebra Forest Reserve, which is located in the municipality of Campinas, São Paulo State, southeastern Brazil (Fig. 1). The reserve comprises 251.8 ha of a well-preserved submontane semi-deciduous forest formation (Oliveira-Filho & Ratter, 1995) that is subjected to a 5-month dry season (May to September) (Leitão Filho, 1982). The local climate is a humid subtropical climate (Cwa) according to the classification system of Köppen, which is characterized by dry winters and rainy summers. Floristic surveys that were performed in the area found >100 tree species within one hectare (Farah et al., 2014; Guaratini, Gomes, Tamashiro, & Rodrigues, 2008). The forest canopy is highly heterogeneous and composed of deciduous and evergreen species (Farah et al., 2014).

2.2. Hyperspectral data

Hyperspectral data were acquired with the ProSpecTIR-VS system (SpecTIR, Inc., USA). This instrument has two sensors that individually cover the visible/near-infrared (VNIR, 450–970 nm) and short-wave infrared (SWIR, 970–2500 nm) wavelength ranges, called AisaEAGLE and AisaHAWK (Spectral Imaging, Inc., Oulu, Finland), respectively. Table 1 summarizes the characteristics of the sensors.

VNIR and SWIR images were acquired synchronously, yielding 357 spectral bands (Table 1). An Inertial Navigation System (INS) provided Global Positioning System (GPS) and line-of-sight information to generate an Internal Geometry Map (IGM) file that contained geocoding information for each raw pixel in the image. The INS was equipped with an Inertial Measurement Unit (IMU) and coupled with a 12-channel GPS that utilized real-time differential corrections to provide a geometric accuracy of less than one meter for each pixel. The IGM files of each