Remote Sensing of Environment 216 (2018) 328–344

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

A fully learnable context-driven object-based model for mapping land cover using multi-view data from unmanned aircraft systems

Tao Liu[a,b,*], Amr Abd-Elrahman[a,b], Alina Zare[c], Bon A. Dewitt[a], Luke Flory[d], Scot E. Smith[a]

[a] University of Florida, School of Forest Resources and Conservation, Gainesville, FL, United States of America
[b] University of Florida, Gulf Coast Research Center, Plant City, FL, United States of America
[c] University of Florida, Department of Electrical and Computer Engineering, Gainesville, FL, United States of America
[d] University of Florida, Agronomy, Gainesville, FL, United States of America

ARTICLE INFO

Keywords:
Multi-view
CRF
DCNN
OBIA
Wetland
Classification
SVM
RF
GMM

ABSTRACT

Context information is rarely used in the object-based landcover classification. Previous models that attempted to utilize this information usually required the user to input empirical values for critical model parameters, leading to less optimal performance. Multi-view image information is useful for improving classification accuracy, but the methods to assimilate multi-view information to make it usable for context driven models have not been explored in the literature. Here we propose a novel method to exploit the multi-view information for generating class membership probability. Moreover, we develop a new conditional random field model to integrate multi-view information and context information to further improve landcover classification accuracy. This model does not require the user to manually input parameters because all parameters in the Conditional Random Field (CRF) model are fully learned from the training dataset using the gradient descent approach. Using multi-view data extracted from small Unmanned Aerial Systems (UASs), we experimented with Gaussian Mixed Model (GMM), Random Forest (RF), Support Vector Machine (SVM) and Deep Convolutional Neural Networks (DCNN) classifiers to test model performance. The results showed that our model improved average overall accuracies from 58.3% to 74.7% for the GMM classifier, 75.8% to 87.3% for the RF classifier, 75.0% to 84.4% for the SVM classifier and 80.3% to 86.3% for the DCNN classifier. Although the degree of improvement may depend on the specific classifier respectively, the proposed model can significantly improve classification accuracy irrespective of classifier type.

1. Introduction

Small Unmanned Aerial Systems (UAS) have rapidly growing roles in precision agriculture and natural resource management (Alsalam et al., 2017; McCabe et al., 2016; Müllerová et al., 2017; Pande-Chhetri et al., 2017), because of several advantages of UAS over other remote sensing platforms. For example, compared to space-borne platforms, UAS can fly at much lower altitudes, and thus are able to generate remote sensing images with sub-decimeter resolution (Rango et al., 2006). This feature is important, because even though civilian remote sensing satellites can collect images with resolution as high as 25 cm (e.g. WorldView-3), this resolution is still insufficient for some natural resource management applications (Lu and He, 2017). Even though piloted aircrafts can collect images with a resolution comparable to UAS images (e.g., 5–6 cm), high cost, operational logistics and pilot safety associated with piloted aircraft missions make UAS adoption for local scale applications desirable (Rango et al., 2006). In addition, flight route and time can be flexibly controlled by the UAS operator, making UAS a preferable remote sensing platform for some natural resource management tasks such as invasive plant species control that require timely and repetitive monitoring of landcovers.

Many natural resource management applications require timely and accurate mapping techniques to monitor landscape scale changes such as non-native plant invasions, insect outbreaks or disease, and to develop and apply management efforts (Mulla, 2013; Zhang et al., 2002). For invasive plant management in particular, accurate maps facilitate monitoring population dynamics as well as precision targeting of infested areas. Thus, UAS technology may fulfill a critical need for invasive plant management but models are needed to accurately determine landcover from UAS-based images.

Object-Based Image Analysis (OBIA) is commonly used for processing the very high-resolution images collected by UASs (Blaschke et al., 2014; Chen et al., 2018; Liu and Abd-Elrahman, 2018a,b; Liu et al., 2018; Pande-Chhetri et al., 2017). Compared to pixel-based

* Corresponding author at: University of Florida, School of Forest Resources and Conservation, Gainesville, FL, United States of America.
E-mail address: taoliu@ufl.edu (T. Liu).

https://doi.org/10.1016/j.rse.2018.06.031
Received 23 January 2018; Received in revised form 19 June 2018; Accepted 24 June 2018
0034-4257/ Published by Elsevier Inc.

approaches, OBIA usually not only creates more visually appealing results in mitigating the salt and pepper effect but also obtains comparable if not higher classification accuracy (Cleve et al., 2008; Fu et al., 2017; Gao et al., 2012; Ma et al., 2017). A common workflow of OBIA for processing UAS images is 1) Generate an orthoimagery from individual UAS-collected images using an off-the-shelf package (e.g., Agisoft, Pix4D or VisualSFM), 2) Segment the orthoimagery into individual objects with appropriately selected parameter values for segmentation (e.g., scale, shape, and compactness for eCognition), 3) Extract features (e.g. mean spectral band values of object pixels) for each object, and 4) Train and apply a classifier (e.g., SVM, random forest) for classification of each object. In this common workflow, all the information used for classification in OBIA comes from orthoimagery alone. One of the primary drawbacks of this approach is that the original information contained in the raw multi-view UAS images is discarded, even though multi-view information has proven useful for land cover classification in several publications (Abuelgasim et al., 1996; Gatebe and King, 2016; Koukal et al., 2014; Su et al., 2007).

Plant functional groups often predictably co-occur (Carranza et al., 2011; Chytrý et al., 2008; Frouz, 1997), providing object context information that may be useful to predict class types. There have been examples of applying context information to improve the classification recently (Albert et al., 2017; Zhao et al., 2017b), but none integrate multi-view data with context-driven models. Markov Random Field (MRF) and Conditional Random Field (CRF) (Sutton and McCallum, 2012) are graphical models that can encode the information to model contextual information and have already been used by few researchers in the remote sensing community (see Table 1). Most previous studies have used pixel-based classification (Kasetkasem et al., 2005; Li et al., 2016; Liu et al., 2008; Zare and Gader, 2009; Zhong et al., 2014). More recently, some of the studies in remote sensing community (Albert et al., 2017; Zare and Gader, 2009; Zhao et al., 2017b) have focused on objects as classification units to utilize the abundant high resolution remote sensing images that have been generated by the technical advancements of remote sensing sensors and platforms.

Even though MRF and CRF have shown potential to improve classification accuracy with varied degree in previous studies, the gains are generally not impressive and have been disproportionate to the greater model sophistication, thus decreasing model utility and accessibility. The limited improvement may be attributed to the fact that previous models usually required the users to empirically determine and input model parameters (see Parameter Determination column in Table 1), which not only adds extra difficulties for using the model, but also prevents the CRF from releasing its full classification improvement power. In addition, various classifiers (e.g., RF, SVM, GMM in Table 1) have been used to generate the unary terms for the MRF or CRF model, but none of them have investigated whether using a different classifier would have an impact on the CRF performance. Furthermore, all the previous studies relied on approximate methods (e.g., iterated conditional modes, alpha-expansion, loopy belief propagation) for model inference to find the collective label configurations of all nodes in a graph (see last column in Table 1), while none of them have tried exact model inference method such as belief propagation to derive object labels via exact marginal distribution.

Given these issues, the research objectives of this study are to:

I). Develop a fully learnable context-driven object-based classification model. This model does not require the user to input any parameter values as all the parameter are automatically learned from the training data. Such automation in parameter estimation can greatly increase our model usability compared to other existing models that usually require the user to empirically determine the model parameters.

II). Develop a method to extract information from multi-view data to be usable in the context driven model. Although the model developed here experiments with multi-view data, it can be generally applied to standard OBIA classification that use the orthoimagery only.

III). Compare context model implementation options, including

- i) Context-driven object-based classification model performance using the DCNN, RF, SVM and GMM classifiers. Such comparison can determine whether consistent improvement can be achieved across the board with different classifiers.
- ii) The belief propagation and the commonly used alpha-expansion model inference methods to investigate whether the type of model inference method affects context-driven model performance for our study site.

2. Study area and materials

2.1. Study area

A study area of 700 m × 500 m was selected because it is large enough to include all the primary land cover types that interest the land managers, while at the same time it is small enough to facilitate quick experiments for this study. The study area is part of a 31,000-acre ranch in Southern Florida, that consists of tropical forage grass pastures, palmetto wet and dry prairies, pine flatwoods and large interconnecting marsh of native grass wetlands. This area also has cabbage palm (*Sabal palmetto*) and live oak (*Quercus virginiana*) hammocks scattered along creeks, gullies, and wetlands.

A portion of the study area was invaded by Cogongrass (*Imperata cylindrica*) (Fig. 1, adapted with permission from Liu and Abd-Elrahman, 2018a,b; Liu et al., 2018), a perennial rhizomatous grass that is highly problematic, because it is not palatable for livestock, decreases native plant biodiversity and wildlife habitat quality, and increases fire hazard (Estrada and Flory, 2015). The U.S Army Corps of Engineers (USACE) is involved in monitoring and treatment of the invasive vegetation in this area. Currently, because there is not a reliable landcover map for this area, to treat the invasive vegetation, the entire area must be assessed for invasive species. To evaluate the treatment effects, considerable effort is required to locate and evaluate the target vegetation. Thus, an accurate landcover map in this area would greatly improve management efficiency.

All other classes, except the shadow class in our study, were assigned according to the standard of vegetation classification for South Florida natural areas (Rutchey et al., 2006). Our objective is to classify the Cogon grass (species level) and five other community-level classes as well as the shadow class as listed in Table 2 (This table was reprinted with permission from Liu and Abd-Elrahman, 2018a,b; Liu et al., 2018).

2.2. UAS image acquisition and preprocessing

The images used in this study were captured by engineers from Surveying and Mapping Branch in USACE-Jacksonville District using the NOVA 2.1 small UAS. A flight mission designed with 83% forward overlap and 50% sidelap was planned and implemented (Table 3, reprinted with permission from Liu and Abd-Elrahman, 2018a,b; Liu et al., 2018). A Canon EOS REBEL SL1 digital camera with a CCD sensor of 3456*5184 pixels was mounted on the UAS to collect images for this study. The images were synchronized with an onboard navigation grade GPS receiver to provide image locations. Five ground control points were established, including four near the four corners and one close to the center of the study area, and used in the photogrammetric solution.

2.3. Orthoimagery creation and segmentation

The UAS images were pre-processed to correct for the change in sun angle during the acquisition period before the orthoimagery was created. Given an original UAS image i with zenith angle θ_i, the original UAS images were corrected as $ImgCorrected_i = ImgOriginal_i \left({}^{\cos(\theta_i)}/_{\cos(75^\circ)} \right)$