ISPRS Journal of Photogrammetry and Remote Sensing 96 (2014) 1–11

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Accurate mapping of forest types using dense seasonal Landsat time-series

Xiaolin Zhu, Desheng Liu *

Department of Geography, The Ohio State University, Columbus, OH 43210, USA

ARTICLE INFO

Article history:
Received 30 June 2013
Received in revised form 15 April 2014
Accepted 24 June 2014

Keywords:
Forest types
Classification
Landsat
Seasonal time-series
Hierarchical approach
Feature selection

ABSTRACT

An accurate map of forest types is important for proper usage and management of forestry resources. Medium resolution satellite images (e.g., Landsat) have been widely used for forest type mapping because they are able to cover large areas more efficiently than the traditional forest inventory. However, the results of a detailed forest type classification based on these images are still not satisfactory. To improve forest mapping accuracy, this study proposed an operational method to get detailed forest types from dense Landsat time-series incorporating with or without topographic information provided by DEM. This method integrated a feature selection and a training-sample-adding procedure into a hierarchical classification framework. The proposed method has been tested in Vinton County of southeastern Ohio. The detailed forest types include pine forest, oak forest, and mixed-mesophytic forest. The proposed method was trained and validated using ground samples from field plots. The three forest types were classified with an overall accuracy of 90.52% using dense Landsat time-series, while topographic information can only slightly improve the accuracy to 92.63%. Moreover, the comparison between results of using Landsat time-series and a single image reveals that time-series data can largely improve the accuracy of forest type mapping, indicating the importance of phenological information contained in multi-seasonal images for discriminating different forest types. Thanks to zero cost of all input remotely sensed datasets and ease of implementation, this approach has the potential to be applied to map forest types at regional or global scales.

1. Introduction

Forest covers about 40 percent of all land of the Earth's surface and it is very important for the ecosystem and socioeconomic system (Westoby, 1989). For instance, forest land highly impacts carbon dynamics, provides habitats for organisms, conserves soil and water resources, and supports human demand for timber and recreation (Band, 1993; Hummel, 1984). As one of the most important and abundant land cover types, forests have been mapped at regional or global scales in various land cover and land use products (Friedl et al., 2002; Gong et al., 2013; Townshend et al., 2012; Yuan et al., 2005). Those products classified forest as either one land cover class or several broad forest categories using remotely sensed data with medium or coarse resolution. For example, using the global archive of Landsat images, forests have been classified as one land cover class at a resolution of 30 m (Gong et al., 2013; Townshend et al., 2012). Based on the International Geosphere Biosphere Programme (IGBP) scheme, the MODIS land cover type product classified global forest into 5 broad categories including evergreen needleleaf forest, evergreen broadleaf forest, deciduous needleleaf forest, deciduous broadleaf forest, and mixed forest at a resolution of 500 m. Although these products have acceptable accuracy for forest mapping, their very general class scheme, i.e., considering forest as one land cover class or several broad categories, cannot satisfy the needs for understanding forest succession and aiding in the proper management of forest resources, which requires spatially explicit knowledge of more detailed forest types at various scales (Zhang et al., 2009).

Detailed forest types are defined by the dominant overstory tree species, such as Maple-Ash forest, Fir-Beech forest, and Pine forest (Xiao et al., 2002). Mapping detailed forest types using remotely sensed data is still challenging work. Existing studies attempting to classify detailed forest types can be categorized into four groups. First, spaceborne or airborne high-resolution images and aerial photography have been widely used to map forest types (Carleer and Wolff, 2004; Franklin et al., 2000; Johansen and Phinn, 2006;

* Corresponding author. Address: 1036 Derby Hall, 154 North Oval Mall, Columbus, OH 43210-1361, USA. Tel.: +1 614 247 2775; fax: +1 614 292 6213.
E-mail address: liu.738@osu.edu (D. Liu).

http://dx.doi.org/10.1016/j.isprsjprs.2014.06.012
0924-2716/

Katoh, 2004; Key et al., 2001; Kim et al., 2009a; Meyer et al., 1996; Waser et al., 2011; Yu et al., 2006). Due to the high spatial resolution of these images, the textural information is often used as a supplement to the multi-spectral information to indentify forest types at individual tree level by pixel-based or object-based classifiers. Secondly, considering different biochemical properties and leaf morphology among different tree species, airborne and spaceborne hyperspectral optical sensors may capture these subtle differences among tree species, so it is possible to classify tree species with hyperspectral imagery (Asner et al., 2008; Clark et al., 2005; Dalponte et al., 2012; Gong et al., 1997; Martin et al., 1998; Xiao et al., 2004). For example, airborne hyperspectral data with 126 bands between 402.9 and 989.1 nm was used to classify 5 forest types in the southern Alps with 79.3% overall accuracy (Dalponte et al., 2012). Thirdly, Light Detection and Ranging (LiDAR) images have been explored to improve the accuracy of forest mapping from optical images because different tree species may have different vertical structure which can be detected by LiDAR instrument (Dalponte et al., 2012; Heinzel and Koch, 2012; Kim et al., 2009b). For instance, the overall accuracy of forest mapping in the southern Alps can be improved from 79.3% to 89.1% after adding LiDAR images into the classification process (Dalponte et al., 2012). In addition, previous studies suggest that the use of topographic information in combination with the spectral data can also improve the overall classification accuracy of forest types (Dorren et al., 2003; Strahler et al., 1978). For example, an early study (Strahler et al., 1978) shows that accuracies of forest type classification from Landsat imagery can be improved by 27% or more through the incorporation of topographic information. Dorren et al. (2003) also demonstrated that including topography into classification can improve the overall accuracy of forest type mapping from 64% to 73%. Topographic information can be represented by a digital elevation model (DEM) or variables derived from a DEM (Dorren et al., 2003).

Although these explorations using high-resolution data, hyperspectral images, or these sets of data in combination with other ancillary data such as LiDAR have obtained acceptable classification accuracy of forest types, the greatest limitation of these data is their high cost and low availability. It is therefore not possible to use these data to classify forest types at regional or global scales. Alternatively, some studies try to use the temporal information from time-series data to classify forest types because different tree species may show different phenological characteristics (Hill et al., 2010; Kempeneers et al., 2011; Xiao et al., 2002). This idea has been carefully tested in a small area of 157 h using 5 airborne multi-spectral images acquired across the growing season. The results suggest that time-series data is more capable of discriminating between deciduous forest types than a single image in the best season such as colorful autumn (Hill et al., 2010). However, due to technical limitations, remote sensing instruments trade spatial resolution with swath width (Zhu et al., 2010). As a result, in practice, only time-series data with coarse resolution, such as data from the Advanced Very High Resolution Radiometer (AVHRR) and SPOT VEGETATION, has been used to classify forest types (Kempeneers et al., 2011; Xiao et al., 2002). The spatial resolution of these time-series data is often from hundreds of meters to a few kilometers. However, forests are generally managed at a size significantly less than one kilometer (Healey et al., 2008), so satellite data with finer resolution are needed to classify forest types, especially in heterogeneous landscapes. Landsat is the best satellite program that satisfies the spatial resolution requirement for mapping forests at large scale. Landsat has the longest data record (since 1972), a spatial resolution in accordance with the grain of land management (30 m), and is available for free globally (Zhu and Liu, 2013). In fact, a number of earlier studies have used Landsat imagery to map forest types (Dorren et al., 2003; Foody and Hill, 1996; Shao et al., 1996; Wolter et al., 1995; Zheng et al., 1997). However, most of these studies only used a single image in the peak-growing season or very sparse multi-temporal images to classify forest types. Given the availability of dense Landsat time-series, the temporal dimension of Landsat imagery can be better explored to include more phenological information for more accurate classification of forest types (Tottrup, 2004).

The main objective of this study is to explore the use of dense Landsat time-series acquired in different seasons for mapping detailed forest types accurately at large scale. However, using dense time-series data to classify forest types by the commonly used supervised or unsupervised classifiers will face some challenges. The first challenge comes from the high dimension of input features. This can result in the Hughes effect, where classification accuracy decreases as more features given a fixed training set (Pal and Foody, 2010). The common approach used to reduce the Hughes effect is to implement feature reduction techniques to select the most informative features (Bazi and Melgani, 2006; Guyon et al., 2002; Pal and Foody, 2010). The second challenge comes from the limited training samples because reference data of detailed forest types can only be collected from costly fieldwork. However, small training sample size can severely affect the accuracy of classification (Jackson and Landgrebe, 2001). To mitigate the small training problem, several methods have been developed to add unlabeled samples in the classification process through an iterative process (Demir et al., 2011; Jackson and Landgrebe, 2001; Rajan et al., 2008). The last challenge comes from the interference of non-forest land cover types when we classify forest types together with these non-forest land cover types. Classifiers trained for all classes cannot ensure optimal results for forest type classification (Foody et al., 2006). To address this problem, hierarchical classification method divides classification problems into different layers. For each layer, we can train classifiers by different features and samples to increase the discrimination between classes (Lauver and Whistler, 1993; Thompson, 1996).

To overcome above challenges of using dense time-series data to classify detailed forest types, we proposed an operational method to integrate the aforementioned techniques. Specifically, we integrated a feature selection and a training data adding procedure into a hierarchical classification framework. Ground reference collected from the field plots will be used to train the proposed method and validate the results. Considering the data availability of field plots, the performance of the proposed method was tested in a pilot area in southeast Ohio. In addition, considering that topographic information has proved valuable to improve the forest type mapping in previous studies (Dorren et al., 2003; Strahler et al., 1978), we also included topographic information from DEM data into the classification process to explore if it can provide additional information besides Landsat time-series for mapping forest types.

The remainder of this paper is organized as follows: we first provide a detailed description of our proposed method. Then, we test the proposed method in a pilot area and compare its performance with the traditional approach. Finally, we discuss the implications of our approach.

2. Methodology

2.1. Study area and data

In order to test the effectiveness of the proposed method for mapping forest types, a case study has been completed in a pilot area covering Vinton County, Ohio, which is about 1790 square kilometers (Fig. 1). This area is classified as a humid continental climate (United States. Soil Conservation Service, 1985). The original forests (old-growth) in this area were cut or burned one or more