ISPRS Journal of Photogrammetry and Remote Sensing 97 (2014) 219–228

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

A parametric model for classifying land cover and evaluating training data based on multi-temporal remote sensing data

Hunter Glanz [a,*], Luis Carvalho [a], Damien Sulla-Menashe [b], Mark A. Friedl [b]

[a] Department of Mathematics and Statistics, Boston University, Boston, MA 02215, USA
[b] Department of Earth and Environment, Boston University, Boston, MA 02215, USA

ARTICLE INFO

Article history:
Received 14 January 2014
Received in revised form 5 September 2014
Accepted 10 September 2014
Available online 8 October 2014

Keywords:
Maximum likelihood estimation
Image classification

ABSTRACT

Time series of multispectral images are widely used to monitor and map land cover. However, high dimensionality and missing data present significant challenges for classification algorithms that use multi-temporal remotely sensed data. Further, generation and assessment of high quality training data, including detection of outliers and changed pixels in training data, is difficult. In this paper we present a new statistical framework that is based on a parametric model that enables a targeted principal component analysis (PCA) to reduce the dimensionality of multi-temporal remote sensing data. In doing so, the model provides a novel basis for land cover classification and evaluating the nature and quality of training data used for supervised classifications. The methodology we describe uses a Kronecker operator to reduce the spectral dimensionality of multi-temporal images while preserving their temporal structure, thereby providing low-dimensional data that is well-suited for classification and outlier detection problems. As part of our framework, we use an expectation–maximization method to impute missing data, and propose new metrics that characterize the representativeness and pixel-to-pixel homogeneity of training sites used for supervised classification. To evaluate our approach, we use data from NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) and extracted more than 200 training sites where the land cover has been characterized from high spatial resolution imagery. The original input data was composed of 196 features (28 dates $\times$ 7 bands), and the PCA-based approach we describe captured 91% of the variance, in these 7 bands, in 3 components. Results from maximum likelihood classification show that the retained principal components successfully distinguish land cover classes from one another, with classification results that were comparable to supervised machine learning methods applied to the original MODIS data. Analysis of our site composition metrics show that they successfully characterize the homogeneity (or lack thereof) and representativeness of individual pixels and entire sites relative to other training sites in the same class.

1. Introduction

Land cover plays a key role in the Earth's climate and ecological system (Bonan, 2008). Because land surface properties affect biosphere–atmosphere interactions, accurate information related to the properties of global land cover is required to parameterize land surface processes in regional-to-global scale Earth system models (Running and Coughlan, 1988; Bonan et al., 2002; Ek et al., 2003; Sterling and Ducharne, 2008). Land cover also plays a critical role in the global carbon cycle (Pacala et al., 2001; Schimel et al., 2001; Houghton and Goodale, 2003; Hurtt et al., 2006; Ramankutty et al., 2007). The global area of land dominated by humans has expanded rapidly in the last 100 years (Vitousek et al., 1997; Ramankutty and Foley, 1999; Goldewijk, 2001; Sanderson et al., 2002; Ellis and Ramankutty, 2008), and society depends heavily on goods and services provided by terrestrial ecosystems (Foley et al., 2005). Reliable information regarding global land use and land cover is therefore essential.

While regional maps of land cover have been generated from remote sensing for several decades, global land cover mapping from remotely sensed data is relatively new. Prior to the AVHRR era, global land cover data sets were compiled from a mix of sources including physiographic maps and atlases, national mapping programs, and highly generalized biogeographic maps (Matthews, 1983; Olson et al., 1983; Wilson and Henderson-Sellers, 1985). In

* Corresponding author.

http://dx.doi.org/10.1016/j.isprsjprs.2014.09.004

the 1990s, continental and global scale land cover was mapped for the first time from remote sensing using AVHRR data (DeFries and Townshend, 1994; Stone et al., 1994; DeFries et al., 1995; Hansen et al., 2000; Loveland et al., 2000). As newer moderate spatial resolution remote sensing data sources have emerged (e.g., MODIS, SPOT-VEGETATION, MERIS), substantial effort has focused on developing improved methods to exploit these data. The current generation of global land cover products include the GLC2000 product produced from SPOT VEGETATION (Bartholomé and Belward, 2005), the MODIS Collection 5 Vegetation Continuous Fields product (Hansen et al., 2002), the GlobeCover products created using data from MERIS (Arino et al., 2007), and the MODIS Collection 5 Land Cover Type product (Friedl et al., 2010).

In addition to providing high quality multispectral imagery at global scale, moderate spatial resolution sensors such as MODIS provide important information related to multitemporal dynamics within and across land cover types. Indeed, most land cover classification algorithms that use moderate resolution data implicitly or explicitly exploit multitemporal information to help distinguish among land cover types. However, the use of multitemporal data significantly increases the computational and algorithmic complexity of land cover classification methods. Specifically, inclusion of multitemporal features dramatically increases the volume and dimensionality of data to be classified, and also introduces multicollinearity among features. For example, a full year of 8-day surface reflectance data for the MODIS land bands includes 322 features (7 bands, 46 time periods), many of which are highly correlated with respect to both their spectral and temporal properties. This creates significant challenges for both feature selection and classification.

In this paper we present a new methodology that addresses these challenges using a parametric modeling approach that reduces the volume, dimensionality, and multi-collinearity of multitemporal remote sensing data, but which also retains spectral and temporal information required to map land cover. In addition, we demonstrate how the framework we have developed can be used to explore the character and quality of training site data used in supervised classification models. To accomplish this latter goal, we present new statistical metrics that characterize the quality and representativeness of training site data used in supervised classifications.

2. Data and methods

The central challenge that our method addresses is how to best reduce dimensionality in multitemporal data sets without losing information that is useful for land cover classification. We examine this question in the context of parametric classification models, because these models provide a clear statistical basis and justification for class assignments. However, because the number of parameters required for parametric models is at least of quadratic order on the number of features in order to capture variances, parametric models are difficult to use with multispectral multitemporal imagery, which can possess dozens to hundreds of input features. For example, the Collection 5 MODIS Land Cover Type product (MLCT) (Friedl et al., 2010) currently uses 135 features to classify land cover at each 500-m MODIS pixel. To overcome this challenge, the MLCT algorithm uses machine learning methods. This type of approach is computationally efficient, but the algorithms themselves are "black-boxes" and tend to be sensitive to properties of the training data. In this paper we explore a different approach that relies on parametric statistical models. To make this feasible, however, we require a methodology to reduce the dimensionality of multitemporal remote sensing datasets in a manner that allows them to be efficiently and accurately classified using parametric methods.

Principal component analysis (PCA) is probably the most widely used method for reducing dimensionality and multi-collinearity in multispectral (e.g., Crist and Cicone, 1984) or multitemporal imagery (e.g., Eastman and Filk, 1993). Here we also use PCA, but our approach is fundamentally different from previous studies in that while we apply PCA to data that includes *both* multispectral and multitemporal properties, we focus our PCA to target the spectral dimensionality only (i.e., most previous studies have used PCA to reduce dimensionality in data that is either multispectral *or* multitemporal, but not *both*) To do this, we describe a new approach that separates spectral variance from temporal variance in image time series. In doing so, our method provides an effective way to reduce data dimensionality and retain the majority of information contained in the original data without convolving spectral information with temporal information. This approach yields principal components that are similar to the classic 'tasseled cap' components, but also captures multi-temporal and multispectral properties that are specific to individual land cover classes, which are critical for land cover classification.

2.1. Data

To develop and test our methodology we used nadir BRDF-adjusted surface reflectance (NBAR) data from MODIS (Schaaf et al., 2002). Specifically, we utilized NBAR data extracted for land cover training sites created by manual interpretation of high resolution imagery that are used to produce the MODIS Land Cover Type product (Friedl et al., 2010). For the purposes of this analysis, we used a subset of the MODIS Land Cover Type product training site database that includes 204 sites located in the conterminous United States (extending from MODIS tile v04h08 in the northwest to MODIS tile v05h11 in the southeast). NBAR data from these sites encompass 2733 MODIS pixels for MODIS bands 1–7, include all major biomes and land cover types in the lower 48 United States, and consist of time series at each pixel that include 46 8-day periods in 2005 (i.e., 322 total features).

2.2. Parametric model

The data for each pixel consists of a spectral–temporal series with T time points and B spectral bands, where $T = 46$ and $B = 7$. The observations X_v for pixel v are therefore represented by a B-by-T matrix. In our parametric model we characterize each land cover class c by a mean spectral–temporal profile μ_c (also a B-by-T matrix) and a temporal covariance matrix Σ_c of order T. We assume that the spectral covariance matrix Σ_s of order B is common to all land cover classes. Then, given the land cover class $\theta_v = c$ at pixel v, the values of X_v are assumed to follow a matrix normal distribution (Dawid, 1981),

$$\mathrm{vec}(X_v) \mid \theta_v = c \sim N(\mathrm{vec}(\mu_c), \Sigma_s \otimes \Sigma_c), \tag{1}$$

where vec is the operator that concatenates the columns of its matrix argument into a single vector and $\otimes$ is the Kronecker product. This formulation decomposes spectral and temporal variance into orthogonal dimensions, and as a result, the model requires fewer parameters than using a full BT-by-BT covariance. Moreover, as we describe below, separation of spectral and temporal sources of variance allows dimensionality reduction to be focused in the spectral bands (see Section 2.4).

2.3. Model parameter estimation via EM

Because surface reflectance data from MODIS includes missing values, estimation of parameters in (1) is challenging. To overcome this difficulty, we impute missing values using a procedure that is similar to expectation–maximization (EM) (Dempster et al., 1977),