ISPRS Journal of Photogrammetry and Remote Sensing 128 (2017) 223–239

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

A Markov random field integrating spectral dissimilarity and class co-occurrence dependency for remote sensing image classification optimization

Leiguang Wang [a,d], Xin Huang [b,*], Chen Zheng [c,d], Yun Zhang [d]

[a] Southwest Forestry University, School of Forestry, China
[b] Wuhan University, School of Remote Sensing and Information Engineering, China
[c] Henan University, School of Mathematics and Statistics, China
[d] University of New Brunswick, Department of Geodesy and Geomatics Engineering, Canada

ARTICLE INFO

Article history:
Received 26 August 2016
Received in revised form 17 March 2017
Accepted 29 March 2017

Keywords:
Remote sensing image classification
Spatial regularization
Markov Random Fields (MRFs)
Spatial energy function
Spectral dissimilarity
Class co-occurrence dependency

ABSTRACT

This paper develops a novel Markov Random Field (MRF) model for edge-preserving spatial regularization of classification maps. MRF methods based on the uniform smoothness lead to oversmoothed solutions. In contrast, MRF methods which take care of local spectral or gradient discontinuities, lead to unexpected object particles around boundaries. To solve these key problems, our developed MRF method first establishes a spatial energy function integrating local spectral dissimilarity to smooth the initial classification map while preserving object boundaries. Second, a new anisotropic spatial energy function integrating the class co-occurrence dependency is constructed to regularize pixels around object boundaries. The effectiveness of the method is tested using a series of remote sensing data sets. The obtained results indicate that the method can avoid oversmoothing and significantly improve the classification accuracy with regards to traditional MRF classification models and some other state-of-the-art methods.

1. Introduction

Remote sensing image classification, which aims to classify a remotely sensed image into a thematic map, is a very active research field. As more and more images with higher spatial resolution became available, advanced classification methods utilize not only the spectral, but also the spatial properties in order to improve classification accuracy. In this context, a large collection of spectral-spatial classification methods (Benediktsson et al., 2005; Blaschke, 2010; Chen et al., 2016; Huang et al., 2014; Li et al., 2013; Zhang and Jia, 2012; Zhang et al., 2006) have been proposed.

Representative spatial features include pixel shape index (Zhang et al., 2006), extended morphological profiles (Benediktsson et al., 2005) and extended morphological attribute profiles (Dalla Mura et al., 2011), among many others (Cheng and Han, 2016; Fauvel et al., 2013). Another group of spatial-spectral methods is known as object-based image analysis (OBIA) (Blaschke, 2010; Walter, 2004; Zheng et al., 2013), which utilizes segments as basic units for extracting features. OBIA can suppress the salt-and pepper noise that is often observed from pixelwise classification results. Besides these hand-crafted spatial features, features learnt automatically from input images, which are known as deep learning-based methods (Chen et al., 2016, 2014; Hinton and Salakhutdinov, 2006), became popular recently. These methods, typically convolutional neural network (Chen et al., 2016), by simulating the processing of the primate visual system through a deep hierarchy, can extract a series of low- and high-level features. Both hand-crafted and learnt features make efforts on integrating spatial information at feature extraction stage or during the classification stage.

Recently, a number of works have developed strategies to integrate the spatial information at the postprocessing stage, such as relearning (Huang et al., 2014), object-based method (Büschenfeld and Ostermann, 2012), filtering based method (Kang et al., 2014), and Markov random fields (MRFs) (Aghighi et al., 2014; Schindler, 2012; Tarabalka et al., 2010). These methods generally rely on the common assumption that neighboring pixels tend to belong to the same class.

Filtering methods impose a kernel on an initial label image using a sliding window, and then assign each center pixel to an

* Corresponding author.
E-mail address: huang_whu@163.com (X. Huang).

http://dx.doi.org/10.1016/j.isprsjprs.2017.03.020

output value obtained according to the existing gray and/or label values in the window. For instance, Kang et al. (2014) proposed to use the guided filter to achieve edge-preserving smoothing of the probabilistic map of an initial label image. Unfortunately, the overall high accuracies are essentially achieved at the cost of over-smoothed objected boundaries. Object-based methods (Büschenfeld and Ostermann, 2012) conduct (weighted) majority voting over each image object to determine the resulting class to be assigned to the object. During the voting process, the initial probabilities of labels and the distance from the current pixel to the corresponding object border are considered. However, the effectiveness of object-based voting is also influenced by the performance of segmentation algorithms. More recently, Huang et al. (2014) have presented a new concept of relearning to smooth the initial classification result. The goal is achieved by iteratively updating the initial result according to the frequency and spatial arrangement of the class labels.

Some MRFs (Aghighi et al., 2014; Li et al., 2012; Moser et al., 2013; Tarabalka et al., 2010) can also be seen as postprocessing methods, as they utilize neighbor label information to produce smoothing effects on initial classification results. Under the MAP-MRF framework, this optimization is formulated as the minimization of the class posterior probability, which is equivalent to minimizing an energy function comprising the feature and label models (Li, 2009). The feature model is related to features used in the classification, and is often initialized by the output of a spectral feature-based pixelwise classification (Schindler, 2012; Tarabalka et al., 2010). Meanwhile, the label model is related to the spatial prior of classes, which is formulated as a MRF (Moser et al., 2013). In contrast, without modeling the feature and label models individually, alternative MRF methods can directly express the class posterior probability as a MRF (Li, 2009; Zhang and Jia, 2012). This group of MRFs is also named as conditional random fields (CRFs).

Therefore, both groups belong to the random field model assumed to exhibit the Markov property, and use the probability function to model the spatial interactions between image sites. Both of them assumes that the class labels and/or feature levels in a neighborhood of the image lattice do not change abruptly. It has been demonstrated that, even with this simple spatial prior, MRF methods can perform quite well in terms of improving classification accuracy (Huang et al., 2014; Schindler, 2012; Zhong and Wang, 2010). However, this generic smoothness prior also lead to oversmoothed solutions when effective, i.e., the class boundaries do not align with real object boundaries (Schindler, 2012; Tarabalka et al., 2010; Zhang and Jia, 2012). The main reason is that the uniform smoothness assumption is often violated at the image boundaries, where abrupt changes of pixel values occur. Therefore, several works have established more complex spatial *a priori* models involving local discontinuities, such as the derivative magnitude (Tarabalka et al., 2010; Yu and Clausi, 2008) or the spectral difference (Moser et al., 2013). These models aim to suppress the smoothness effect when the value of the term becomes larger (often with high probability exactly corresponding to, or near real boundaries). In this way, the models can effectively preserve edges.

Unfortunately, as shown in Figs. 8(a) and 12 (g), these state-of-the-art MRF models involving local discontinuities, still suffer from unexpected and isolated class labels around object boundaries, where a salt-and-pepper noise effect can be appreciated. These pixels located around boundaries have distinct spectral presentations with surrounding pixels. According to the MRF models, with a high possibility, they are labelled as the class with the most similar spectral property, rather than the spatial neighboring classes. Whereas, in order to properly consider spatial dependency among different land classes and obtain better visual inspection, we have strong motivations to divide these pixels with different spectral properties into the surrounding classes. For instance, in Fig. 8, pixels located between different crop types are expected to be recognized as one of the adjacent types, rather than some other types with similar spectral properties but long spatial distance.

From previous literature review, we can find that most of post-processing optimization methods perform quiet well in homogenous regions. A key aspect when utilizing optimizations is to design a proper spatial model, which can deal with the features and labels around object boundaries. An expected spatial regularization method should refine an already classified map, smooth labels in homogenous regions, meanwhile, align the boundaries among different labels with real object boundaries.

In this context, this paper encodes spatial *a priori* assumptions involving both spectral dissimilarity and class co-occurrence dependency into two spatial energy functions, and results in a new MRF method with two-step spatial regularization. The first spatial energy function integrating local spectral dissimilarity is to smooth the initial classification map while preserving object boundaries. The second spatial energy function integrating the class spatial dependency is constructed to further regularize pixels around object boundaries. It is also our main contribution in this research.

The rest of the paper is organized as follows. Section 2 presents background on MRF-based methods intended to achieve spatial smoothness. The proposed MRF method with two-step spatial regularization is presented in Section 3. Experimental results and discussions, analyzing the influence of different spectral dissimilarity metrics and a detailed parameter sensitivity assessment, are presented in Section 4. Comparisons with some other state-of-the-art methods are conducted in Section 5. Conclusions and hints at plausible future research lines are given in Section 6.

2. Background on MRF-based methods to achieve spatial smoothness

2.1. Notations and problem formulation

Let S denote the set of sites over which a remote sensing image $\boldsymbol{I}$ is defined, and let $\Omega = \{1, 2, \ldots, k\}$ denote a set of labels, being k the number of labels. Both the observation random field $\mathbf{Y}$ and the label random field $\mathbf{X}$ are defined on S. The observed image $\mathbf{y} = \{y_i | i \in S\}$ is a realization of the observation random field $\mathbf{Y}$. A label image $\mathbf{x} = \{x_i | i \in S, x_i \in \Omega\}$ is a realization of $\mathbf{X}$, in which each x_i takes a value from Ω denoting the class to which the site i belongs.

The spatial regularization task performed by a pixelwise classification map is formulated as finding an optimal estimation $\widehat{\mathrm{x}}$ that maximizes the posterior $P(\mathbf{X}|\mathrm{y})$ given the observed image y. According to the Bayesian rule and the log-linear property, finding the maximum *a posteriori* (MAP) solution $\widehat{\mathrm{x}}$ of $P(\mathbf{X}|\mathrm{y})$ (also called optimal configuration) is equivalent to minimizing the following two-part energy function (Li, 2009):

$$\widehat{\mathrm{x}} = \arg\max_x(p(\mathbf{y}|\mathbf{x})P(\mathbf{x})) = \operatorname{argmin}_x(E_f + E_l) \tag{1}$$

with

$$E_f = -\log(p(\mathbf{y}|\mathbf{x})), \tag{2}$$

and

$$E_l = -\log(P(\mathbf{x})). \tag{3}$$

With the aforementioned formulation in mind, two issues need to be addressed: (1) how to analytically represent the feature model E_f and the label model E_l, and (2) how to find a solution for the objective function (1).