ISPRS Journal of Photogrammetry and Remote Sensing 142 (2018) 162–173

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

A supervised approach for simultaneous segmentation and classification of remote sensing images

Daniel Capella Zanotta[a,*], Maciel Zortea[b], Matheus Pinheiro Ferreira[c]

[a] National Institute for Science, Education and Technology, R. Eng. Alfredo Huch 475, 96201-460 Rio Grande, RS, Brazil
[b] IBM Research, Avenida Pasteur 138–146, 22290-240 Rio de Janeiro, RJ, Brazil
[c] Remote Sensing Division, National Institute for Space Research - INPE, São José dos Campos 12227-010, SP, Brazil

ARTICLE INFO

Keywords:
Object-based image analysis
Segmentation
Supervised classification
Multispectral imaging

ABSTRACT

Object-based image classification is recognized as one of the best strategies to analyze high spatial resolution remote sensing images. This process includes defining scale parameters to form regions sharing similar characteristics such as color, texture, or shape. Traditionally, in an object-based supervised classification setting the image is classified only after the segmentation process is completed. However, when the imaged objects on the ground are heterogeneous and of different sizes, some resulting segments can be appropriate for classification while others are over or under-segmented. This may cause partial failure of the subsequent classification. In this paper, we introduce a simultaneous approach based on the interception of the segmentation stage by provisional classification of under-growing segments. Our proposal is to optimize the classification process by iteratively updating the labels of previously generated regions only if the estimated posterior probabilities of the winning classes in the new segments increase. Experiments with three multispectral datasets acquired by Landsat-5 TM, QuickBird-II, and WorldView-3 in rural and urban areas compare traditional object-based approach based on region growing with the proposed method using well-established classifiers. Our results show that the proposed method becomes much less sensitive to the choice of segmentation parameters and reaches similar, or even better, classification accuracies.

1. Introduction

Image classification is a key technique to derive land use and land cover information from multispectral images collected by remote sensing. Conventional classification approaches usually rely on the spectral information of each single pixel, such as its digital number, to define which of the available thematic classes it belongs to. With the increased availability of high spatial resolution remote sensing imagery collected by airborne and spaceborne sensors, many techniques have been proposed aimed to exploit not only the spectral information but also the texture and morphological features of targets present in the scenes (Neubert and Meinel, 2003; Blaschke et al., 2008; Zehtabian and Ghassemian, 2016; Xiao et al., 2017; Amitrano et al., 2018). These approaches are often referred to as object-based image analysis and usually consist of two separate steps: segmentation and classification.

In the segmentation step, the input image is split into small regions formed by spatially connected pixels sharing similar characteristics (objects), which will later receive a unique label in the classification step (Blaschke et al., 2014; Navulur, 2006; Blaschke, 2010). Object-based classification is highly suited for applications that use medium to high-resolution satellite imagery to map land cover and monitoring land change (Ma et al., 2017; Neubert and Meinel, 2003). This type of image data can resolve fine details of heterogeneous classes present on the ground. However, the spectral variability of the resulting pixels makes it difficult to recognize isolated pixels without considering their neighborhood. The object-based paradigm vanquishes limitations of pixel-based image processing by exploiting information like context, texture, and shape of objects for feature extraction (Navulur, 2006; Kaplan and Avdan, 2017; Khadanga et al., 2016). Thus, much attention has been given to the generation and modeling of objects by image segmentation techniques. Depending on the resolution of the remote sensing image and the application, segmentation techniques can vary from region-based methods (Karoui et al., 2010; Ho and Chen, 2004; Li et al., 2015; An et al., 2011; Zanotta et al., 2015a), watershed algorithms (Cai et al., 2009; Li et al., 2010; Levner and Zhang, 2007), or techniques based on image thresholding (Hu et al., 2016; Xie et al., 2010; Li et al., 2013), among others (Richards, 2013; Ma et al., 2015).

Although many applications have been taking advantage of object-based analysis, image segmentation is not a trivial task. One reason is that the choice of segmentation parameters is usually subjective and

* Corresponding author.
E-mail address: daniel.zanotta@riogrande.ifrs.edu.br (D.C. Zanotta).

https://doi.org/10.1016/j.isprsjprs.2018.05.021
Received 14 January 2018; Received in revised form 29 April 2018; Accepted 31 May 2018

arbitrary, often leading to unsatisfactory results with few image divisions (under-segmentation) or very fragmented images (over-segmentation), with possible negative impacts on the final classification. Many of these undesirable effects are caused by the inability of the segmentation processes to deal with the complexity of the targets present in an image. For instance, in applications such as land cover mapping, the imaged objects can vary from large or small, natural or artificial, homogeneous or heterogeneous, bright or dark, all of them along the same scene. Nixon (2012) argue that any mathematical algorithm for segmentation should be supplemented by heuristics that involve semantic information about the classes under consideration. Such condition often forces analysts to recursively try different segmentation parameters attempting to achieve a suitable outcome for the entire image, a time-consuming and labor-intensive task. Interesting approaches regarding automatic spectral-spatial classification strategies are provided by Ghamisi et al. (2014), Zehtabian and Ghassemian (2017), and Shen et al. (2017). In Zehtabian and Ghassemian (2017), a semi-automatic framework for classification of hyperspectral images in which the main parameters are adaptively tuned is proposed leading to encouraging results.

At the same time, many techniques have been proposed to improve the performance of segmentation algorithms. For example, segmentation evaluation measures are available to drive the search for the most suitable setup in each application (Liu and Xia, 2010; Ming et al., 2009; Peng et al., 2016; Zhang et al., 2015; Löw et al., 2015; Ma et al., 2015). These methodologies are particularly effective in tuning the segmentation parameters. However, many of them rely on information provided by experts, such as human-made reference images, and are designed for specific applications (e.g., high spatial resolution images) (Peng et al., 2016; Zhang et al., 2015; Mafanya et al., 2017).

Adaptive segmentation processes are also suggested to deal with different types of targets in the input image (Navon et al., 2005; Bhanu and Peng, 2000; Judah, 2014; Li and Wan, 2011; Jiang et al., 2010). Most of them were proposed to work on photography and some are specifically designed for remote sensing or aerial images (Baik et al., 2003; Li and Wan, 2011; Judah, 2014). Although apparently robust and efficient, most of these techniques require auxiliary data such as multiresolution images (Judah, 2014), or rely on many step-by-step procedures and specific information related to the application or data type (Li and Wan, 2011; Espindola et al., 2006). Alternative techniques focus on rules that constrain segmentation results to produce objects with a predefined shape (Solberg et al., 2006) or size (Heinzel et al., 2011; Zhong et al., 2014).

The key characteristic shared by all of the above-mentioned methods is that their pattern recognition processing consists of image segmentation followed by classification. Such approaches assume intrinsically that the segmentation phase is able to accurately extract the objects of interest from the background image automatically (Nebti, 2013). However, these assumptions are rarely met in real-world applications, and therefore optimal results are hardly achieved. In another study, the authors found qualitative evidence that good classification maps could be obtained using stepwise segmentation and classification based on the statistical similarity of growing segments (Zanotta et al., 2015b).

In this paper, we introduce a novel concept for supervised object-based classification, proposing the simultaneous application of the segmentation and classification processes. The main advantage of the proposal is that the final classification map becomes much less sensitive to the specific choice of the segmentation parameters, given a certain segmentation method and classifier selected by the user. Our proposal focuses on the classical region growing segmentation algorithm to illustrate the idea (but in principle it could also be extended to other segmentation methods). Differently from the traditional supervised object-based classification approaches, the proposed method aims to recursively classify under-growing regions, with meaningful size, according to predefined classes. By adopting an iterative region-based segmentation strategy, once a region is merged with another, the classification rule is applied to derive a provisional label and a degree of membership (or class membership value) for the corresponding class. These values are derived from the posterior probabilities estimated by the classification process. The provisional label, and its degree of membership, can be updated whenever a new region is formed and an improvement of classification is verified. By using this simultaneous methodology, under-growing regions with high class membership can be early associated to one of the predefined classes, avoiding, for example, further mislabeling of regions during the classification process that is usually applied only after the segmentation phase is finished. Moreover, by considering objects at different scales, their specific characteristics can improve image analysis and increase the likelihood of classification success.

The remainder of the paper is organized into four sections. Section 2 formulates the proposed method. Section 3 demonstrates experimentally how the proposed method performs land cover classification in rural and urban areas using medium and high spatial resolution images. Finally, Section 4 draws the conclusions of the work.

2. Simultaneous segmentation and classification framework

The proposed method aims to improve supervised object-based image classification by iteratively combining the segmentation and classification steps until all pixels in the image are classified. For convenience, we outline the novel approach using the iterative behavior of the well-known region growing segmentation algorithm. The region growing formulation is aimed to produce objects by iteratively merging pixels or small image portions according to their similarities (Navulur, 2006). We start by revising the basic idea behind this classical segmentation method, and then we introduce our interception strategy by supervised classification leading to the proposed joint segmentation-classification approach.

2.1. Region growing formulation principle

Basically, region growing segmentation algorithms start from the pixel level and apply similarity tests based on user-defined input parameters to determine whether two neighboring pixels or regions should be merged or not. The growth of a region should stop when no more neighboring regions satisfy the merging criteria Navulur (2006).

Let $\mathbf{R}$ represent the entire image so that $\mathbf{R} = \{r_i\}_{i=1}^{N}$, where $\mathbf{R}$ is composed by N subregions r_i. Let $\mathbf{V}_i = \{v_i\}_{i=1}^{N}$, where $\mathbf{V}$ is the set of attributes denoting the characteristics of the regions, which contains the attribute(s) of r_i. Then, let T be the threshold applied to each region to define whether or not two adjacent regions must be merged. In classical formulations, the related attribute generally corresponds to differences between digital number (DN) averages within two regions, but it may also contain descriptors of texture, context, and shape of segments. For the sake of simplicity, we assume a similarity test based on the Euclidean distance between the attributes of regions. Thus, two adjacent regions r_A and r_B will be merged if the following condition is satisfied:

$$\|v_A - v_B\| < T \tag{1}$$

This process is repeated until no more regions satisfy the condition in Eq. (1). High values of T force the union of distinct regions, resulting in few regions (under-segmentation). Conversely, low values produce few unions, resulting in many regions (over-segmentation). Other conditional merging factors can be used, such as minimum region size and statistical properties of segments (Zucker, 1976). A review on region growing segmentation algorithm possibilities is out of the scope of this work, the interested readers may refer to Gonzalez and Woods (2009), Revol-Muller et al. (2012), and Mather and Tso (2009) for a comprehensive overview of alternative formulations and detailed rules.