ISPRS Journal of Photogrammetry and Remote Sensing 101 (2015) 145–162

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

A new segmentation method for very high resolution imagery using spectral and morphological information

Jing Liu [a,b], Peijun Li [a,*], Xue Wang [a]

[a] Institute of Remote Sensing and GIS, School of Earth and Space Sciences, Peking University, Beijing 100871, PR China
[b] Department of Natural Resources, Faculty of Geo-Information Science and Earth Observation, University of Twente, 7500 AE Enschede, The Netherlands

ARTICLE INFO

Article history:
Received 16 March 2014
Received in revised form 8 November 2014
Accepted 16 November 2014

Keywords:
Very high resolution image
Image segmentation
Morphological profile
Seeded region growing
Region merging
GEOBIA

ABSTRACT

Image segmentation is a key and prerequisite step for object-based analysis of very high resolution (VHR) imagery. Most existing image segmentation methods use either spectral or spatial information of an image alone. A novel image segmentation method for VHR multispectral images using combined spectral and morphological information is proposed in this paper. The method can be summarized as follows. First, a morphological derivative profile is calculated from an original multispectral image and combined with the spectral bands to quantify spectral-morphological characteristics of a pixel, which are considered as a criterion of homogeneity of neighboring pixels. Image segmentation is then conducted using a seeded region-growing procedure, which is based on the seed points automatically generated from the gradient image and dynamically added and the similarity between a seed pixel and its neighboring pixels in terms of spectral-morphological characteristics. The obtained segmentation result is further refined by a region merging procedure to generate a final segmentation result. The proposed method is evaluated using three VHR images of urban and suburban areas and compared with two existing segmentation methods, in terms of visual inspection, quantitative evaluation and indirect evaluation. Experimental results demonstrate that the joint use of spectral and morphological information outperformed the use of morphological information alone. Furthermore, the proposed image segmentation method performed better than existing methods. The proposed image segmentation method is well applicable to the segmentation of VHR imagery over urban and suburban areas.

1. Introduction

Commercial very high resolution (VHR) images, such as those taken by IKONOS, Quickbird, GeoEye-1 and Worldview-2 satellites, have been increasingly used in diverse applications, in particular in urban applications, such as land cover classification, object recognition, and urban disaster damage assessment (Walker and Briggs, 2007; Durieux et al., 2008; Turker and Sumer, 2008; Li et al., 2011). The increasing availability of VHR data has resulted in the need for more sophisticated image processing and analysis methods. Many studies have showed that, compared with traditional pixel-based methods, object-based methods are more suitable for VHR image processing and analysis (e.g. Baatz and Schäpe, 2000; Laliberte et al., 2004; Benz et al., 2004; Blaschke, 2010). Furthermore, the Geographic Object Based Image Analysis (GEOBIA) framework has been presented and is considered a new paradigm in remote sensing and geographic information science (GIScience) (Hay and Castilla, 2008; Blaschke et al., 2014).

Image segmentation is a prerequisite and decisive step for object-based image analysis (or GEOBIA). The image objects or segments produced by image segmentation form the basis or information carrier for subsequent processing steps (Trias-Sanz et al., 2008), such as object-based image classification. Therefore, the quality of image segmentation directly affects the quality of subsequent object-based image analysis and applications (Neubert and Meinel, 2003; Witharana et al., 2014), and the exploration of new and sophisticated image segmentation methods has been a focus of VHR image processing and analysis (Wang et al., 2010).

Image segmentation involves the partitioning of a given image into a number of homogeneous regions according to a given criterion (Fan et al., 2005). It can also be considered as a pixel labeling process in the sense that all pixels that belong to the same homogeneous region are assigned the same label (Haris et al., 1998). There are primarily four approaches to image segmentation, namely the thresholding technique, edge-based methods,

* Corresponding author.
E-mail address: pjli@pku.edu.cn (P. Li).

http://dx.doi.org/10.1016/j.isprsjprs.2014.11.009
0924-2716/

region-based methods and hybrid methods that combine edge and region criteria (Adams and Bischof, 1994; Shih and Cheng, 2005). Edge-based methods and region-based methods are two commonly used image segmentation strategies for remote sensing (Carleer et al., 2005). Edge-based methods employ the postulate that values of neighboring pixels, such as intensity, color and texture, change rapidly in the boundary between two regions (Adams and Bischof, 1994). However, edge-based methods are sensitive to noise or texture in the images and prone to produce over-segmentation in textured regions, thus are more effective in the detection of homogeneous and contrasting objects (Janssen and Molenaar, 1995). Region-based methods rely on the assumption that neighboring pixels within one region have similar values (e.g., intensity, color and texture). This leads to a class of algorithms known as region growing algorithms. The general procedure of region growing methods is to compare one pixel with its neighbor(s). If a criterion of homogeneity is satisfied, the pixel is considered to belong to the same class as one or more of its neighbors (Adams and Bischof, 1994). Since region-based methods are less sensitive to texture, which is an advantage in the segmentation of VHR images (Pesaresi and Benediktsson, 2001; Carleer et al., 2005; Witharana and Civco, 2014), various region-based methods have been proposed and widely used in analyzing VHR images, e.g., multi-resolution segmentation (MRS) in eCognition (Baatz and Schäpe, 2000; Benz et al., 2004) and SPRING (Câmara et al., 1996).

Seeded region growing (SRG) is one type of region-based methods, proposed by Adams and Bischof (1994). It starts with selected seeds, and grows regions by merging a pixel into its neighboring seed region according to a homogeneity criterion. The SRG method is a robust and rapid segmentation method. It is attractive for image segmentation because it involves high-level knowledge of image components in the seed selection procedure (Adams and Bischof, 1994). However, the SRG method suffers from the problems of automatic seed generation and pixel sorting orders for labeling (Mehnert and Jackway, 1997; Fan et al., 2005). Some improved methods have been proposed (Mehnert and Jackway, 1997; Fan et al., 2005; Shih and Cheng, 2005).

A literature survey showed that most existing region-based segmentation methods for remote sensing images are based on merging neighboring pixels according to user-defined thresholds of their spectral similarity; i.e., spectral information is used alone (Woodcock and Harward, 1992; Lee and Lee, 2010; Bilgin et al., 2011). However, most currently available VHR images have limited spectral resolution. Objects with similar spectral characteristics (e.g. neighboring building and street in urban area) may not be well separated using spectral information alone. Therefore, there has been increasing interest in using spatial information from imagery, such as textural, structural and shape information (Soltanian-Zadeh et al., 2004; Silva and Bigg, 2004), which is abundant in VHR images. However, only few studies have effectively combined spectral and spatial information in region-based image segmentation (Segl et al., 2003; Akcay and Aksoy, 2008), e.g. the MRS method in eCognition (Benz et al., 2004).

In a related study, structural information quantified by a morphological profile was exploited and used in the segmentation of high-resolution images (Pesaresi and Benediktsson, 2001). Application of the method produced promising segmentation results for high-resolution images over urban areas. The morphological profile has also been used to express and extract relevant structures from VHR images ever since (e.g. Jin and Davis, 2005; Chini et al., 2009; Huang and Zhang, 2012). However, there are some problems with the method. First, the method (Pesaresi and Benediktsson, 2001) segments the image according to the scale size at which the derivative morphological profile (DMP) of a pixel reaches its maximum, considering this reflects the overall structural information. However, the assumption that all pixels in a structure (region or segment) have only one significant derivative maximum occurring at the same scale (Pesaresi and Benediktsson, 2001) often does not hold for VHR images (Akcay and Aksoy, 2008). In more complex environments, it is possible that some structures may have more than one significant derivative maximum (Pesaresi and Benediktsson, 2001). Second, the segmentation decision is made without considering spectral information. Two objects with similar morphological features but different spectral features may be segmented into a region (i.e. under-segmentation) using this method. Although some improvements of the method (Pesaresi and Benediktsson, 2001) have been proposed, few studies have effectively combined spectral and morphological information in the segmentation of VHR images (Akcay and Aksoy, 2008). Third, the method was initially proposed for panchromatic images. When used in the segmentation of a multispectral or hyperspectral image, feature extraction methods (e.g. principal component analysis, PCA) are first employed and a few derived features are then used to produce morphological profiles and used in segmentation (Benediktsson et al., 2005; Akcay and Aksoy, 2008). However, in this way, the information of the non-chosen feature images is lost and spectral information may not be fully used.

The main objective of this study is to propose a new segmentation method for VHR multispectral imagery using both spectral and morphological information. The specific objectives are (1) to present an effective method that combines spectral and morphological information representing pixel characteristics and a measure that quantifies the spectral and morphological similarity of neighboring pixels; (2) to propose an SRG method using spectral and morphological similarity and automatically generated seed points; and (3) to propose a region merging method that uses spectral and morphological information.

2. Background

In this study, the morphological profile is adopted to quantitatively express structural information of VHR imagery (Pesaresi and Benediktsson, 2001). Specifically, morphological opening and closing operations with increasing structuring element (SE) sizes were performed on an image to generate morphological profiles for each pixel and to model morphological characteristics of pixel neighborhoods (Serra, 1982; Pesaresi and Benediktsson, 2001). Each opening or closing operation results in a transformation of the original image in which those objects smaller than the SE size are deleted. As a result, opening and closing operations are able to isolate structures that are brighter and darker than their surroundings, respectively. Contrary to opening (respectively, closing), opening by reconstruction (respectively, closing by reconstruction) preserves the shape of the structures that are not removed by erosion (respectively, dilation). In other words, using opening by reconstruction and closing by reconstruction, image structures that cannot contain the SE are removed while other structures remain (Akcay and Aksoy, 2008). A vector containing the pixel values in openings and closings by reconstruction of different scales is called the morphological profile (Bellens et al., 2008). Morphological profiles carry information about the size and shape of objects in the image. The DMP is a vector that contains the differences between subsequent values of opening or closing by reconstruction in the morphological profile. Specifically, the DMP of a pixel x is

$$\mathrm{DMP}(x) = \left\{ \begin{array}{l} \Delta_{\lambda} : \Delta\gamma_{\lambda}, \forall\lambda \in [1, \ldots, n] \\ \Delta_{-\lambda} : \Delta\phi_{\lambda}, \forall\lambda \in [1, \ldots, n] \end{array} \right\}, \tag{1}$$

where $\Delta\gamma_{\lambda}$ denotes the opening derivative of scale λ of pixel x, $\Delta\phi_{\lambda}$ denotes the closing derivative of scale λ of pixel x, and n is the maximum scale of the SE (Pesaresi and Benediktsson, 2001). The SE scale corresponding to the maximum height in the DMP is