Remote Sensing of Environment 190 (2017) 137–148

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Region merging using local spectral angle thresholds: A more accurate method for hybrid segmentation of remote sensing images

Jian Yang [a,b,*], Yuhong He [b], John Caspersen [c]

[a] Department of Geography, University of Toronto, 100 St. George Street, Toronto, ON M5S 3G3, Canada
[b] Department of Geography, University of Toronto Mississauga, 3359 Mississauga Rd North, Mississauga, ON L5L 1C6, Canada
[c] Faculty of Forestry, University of Toronto, 33 Willcocks Street, Toronto, ON M5S 3B3, Canada

ARTICLE INFO

Article history:
Received 17 February 2016
Received in revised form 11 November 2016
Accepted 17 December 2016
Available online xxxx

Keywords:
Image segmentation
Region merging
Local SA thresholds
Primitive segments
Multi-band watershed transformation
Geographic object-based image analysis (GEOBIA)

ABSTRACT

Image segmentation is the decisive process in object-based image analysis, but segmenting a landscape scene into meaningful geo-objects remains a challenge. In recent years, there has been growing interest in hybrid methods that combine initial segmentation with subsequent region merging, because they exploit spectral signals across an entire geo-object, including both the boundary signals used to delineate the initial segments, and the interior signals used to merge similar segments. However, existing algorithms commonly use a single, global parameter to control the process of region merging, thereby limiting the goodness-of-fit between segments and geo-objects, since homogeneous and heterogeneous segments are treated equally. To overcome this limitation, we developed a new hybrid segmentation method that employs local spectral angle (SA) thresholds for region merging. We implemented our local SA method in three very different landscapes, then compared our region merging method to the global SA method, as well as the global elevation method used in System for Automated Geoscientific Analyses (SAGA). In all three landscapes, the results revealed that the local SA segmentation provides a better fit to reference polygons than the two global threshold methods, as measured using a modified discrepancy measure for the purpose of geo-object recognition (QR_M). We also found that the local SA method produced segments with a greater variation in size, indicating the method is effective for achieving multi-scale segmentation.

1. Introduction

Geographic object-based image analysis (GEOBIA) is an emerging paradigm in remote sensing and geographic information science (Blaschke et al., 2014) that can improve image classification by making use of geo-object features, spatial concepts, and contextual information (Cleve et al., 2008; Myint et al., 2011; Yan et al., 2006). For example, object-based image classification can be used to reduce spectral variability within geo-objects, and thereby eliminate the salt and pepper noises that confound traditional pixel-based methods (Johansen et al., 2010). Within the GEOBIA framework, image segmentation is considered the decisive procedure for subsequent processing steps, including geo-object recognition, information extraction, and image classification (Liu et al., 2015; Yang et al., 2015b). Segmentation partitions the entire image into a set of spatially contiguous and spectrally homogeneous regions, hereafter referred to as segments.

Segmentation remains one of the most challenging steps in classifying complex images, so a variety of methods have been developed, including edge-based methods and region-based methods. Edge-based methods are designed to detect neighbouring pixels that are spectrally dissimilar, and thus delineate the boundaries between adjacent segments (Canny, 1986; Vincent and Soille, 1991). In contrast, region-based methods are designed to detect neighbouring pixels that are spectrally similar, and hence belong to the same geo-object (Adams and Bischof, 1994; Fukunaga and Hostetler, 1975). Yet, neither method exploits spectral signals across an entire geo-object, including both the boundaries and the region they delineate. Furthermore, both of these methods tend to exhibit segmentation bias – the former towards over-segmentation (Gaetano et al., 2015) and the latter towards under-segmentation (Liu et al., 2015).

A new trend in image segmentation is to employ a split-and-merge strategy, in which the initial segments are first delineated using edge-based methods, then merged using region-based methods (Gaetano et al., 2015; Liu et al., 2015; Wang and Li, 2014; Wuest and Zhang, 2009; Zhang et al., 2014). Such hybrid methods have recently attracted growing interest, not only because they make use of two complementary methods, but also owing to their potential to exploit the spectral signals across an entire geo-object, including both the boundary signals used to delineate the initial segments, and the interior signals used to merge similar segments (Zhang et al., 2014).

* Corresponding author at: Department of Geography, University of Toronto, 100 St. George Street, Toronto, ON M5S 3G3, Canada.
E-mail address: jiangeo.yang@mail.utoronto.ca (J. Yang).

http://dx.doi.org/10.1016/j.rse.2016.12.011
0034-4257/

Many region merging processes employ a single global parameter to control the size and/or number of segments. Two widely-used examples include multi-resolution segmentation in eCognition (Benz et al., 2004) and feature extraction in ENVI (Jin, 2012). Another apt example is the watershed transformation algorithm used to segment height maps in the System for Automated Geoscientific Analyses (SAGA) (Conrad et al., 2015), which requires users to choose an elevation threshold that controls the iterative process of merging segments (basins): if the difference in elevation or height (from seed to saddle) is less than this threshold, then they are merged. While this global threshold gives the user some control over both the degree of bias and the direction of bias (over- versus under-segmentation), it is limited by the fact that the same threshold must be used for each segment, regardless of how homogeneous it is compared to other segments in the image. In order to further improve the quality of segmentation, merging algorithms should allow homogeneous segments to have a higher threshold, since a homogeneous segment is more likely to be part of an adjacent object than a heterogeneous segment, even if the difference (e.g., height or spectral difference) between the two segments and the adjacent object is the same.

Over the past few years, several local threshold methods have been developed for image segmentation (Cánovas-García and Alonso-Sarría, 2015; Chen et al., 2014; Johnson and Xie, 2011). For example, Johnson and Xie (2011) & Chen et al. (2014) made use of local measures to identify segments that were under- and over-segmented at the selected optimal scale parameter, and further refined them by appropriate splitting and merging. This local refinement strategy was effective in terms of improving segmentation quality because it eliminated under- and over-segmentation problems. However, the additional splitting and merging steps pose a challenge when implementing this method in an operational context. Another study reported by Cánovas-García and Alonso-Sarría (2015) divided a large and heterogeneous agricultural area into several plots with different land uses for optimal scale parameter selection. Despite its local parameter optimization relative to the large study area, this is still a global method rather than an intrinsically local method as each plot was treated independently. To the best of our knowledge, no method has been proposed to control the merging process through the utilization of a local threshold.

In this study, we describe a new hybrid segmentation method that combines the advantages of edge- and region-based methods, and employs a local threshold for region merging. In addition to overcoming the limitations of global thresholds, this method employs an intuitive and physically-defined metric (spectral angle: SA) to quantify spectral distance between segments (Yang et al., 2015a; Yang et al., 2014b), unlike other methods that use less intuitive metrics that are not defined in simple physical terms (Benz et al., 2004; Jin, 2012; Zhang et al., 2014). The rest of the paper is organized as follows. Section 2 describes the details of our hybrid segmentation method, especially focusing on the region merging process controlled by local SA thresholds, followed by the description of the study sites and images in the third section. The fourth and fifth sections demonstrate the results and discussion, respectively. The main conclusions are reiterated in the last section.

2. Methods

2.1. Overview

Our hybrid segmentation method can be summarized in general terms as follows. First, initial segments are produced using a multi-band watershed transformation that detects edges based on the spectral distance between neighbouring pixels. Then, a region adjacency graph (RAG) is used to specify which segments are adjacent to one another, and the spectral distance between neighbouring segments is calculated. Finally, neighbouring segments are iteratively merged if their spectral distance is smaller than a local, heterogeneity-dependent threshold. The method is described in more detail the following sections and flowchart (Fig. 1).

2.2. Multi-spectral watershed transformation

Most watershed transformation methods are designed to detect edges using a gradient image derived from a panchromatic image or a single band of a multispectral image (Li et al., 2010; Parvati et al., 2008; Wang et al., 2004). In recent years, however, new methods have been developed to take full advantage of the edge information contained in the spectral contrast of multiple bands. For example, a multispectral scalar gradient can be calculated as the modulus of the vector difference between the multispectral dilation and erosion (Li et al., 2011; Li and Xiao, 2007). For this study, we used a similar, but less computationally-demanding method (Yang et al., 2014a) to calculate the maximum spectral angle (SA) between a pixel and its neighbours:

$$M_{SA}(x,y) = \max_{a,b \in W(x,y)} \theta_{ab} \tag{1}$$

$$\theta_{ab} = \cos^{-1} \frac{\sum_{i=1}^{n} a_i b_i}{\sqrt{\sum_{i=1}^{n} a_i^2 \sum_{i=1}^{n} b_i^2}} \tag{2}$$

where the magnitude of a given pixel, $M_{SA}(x, y)$, is equal to the maximum SA between any pair of two pixels within a moving window W of a fixed size centering on the pixel (x, y), n is the number of spectral bands, and a_i and b_i represent the digital numbers (DNs) of two different pixels in band i, respectively.

The resulting gradient image was then segmented using the watershed algorithm proposed by Vincent and Soille (1991). We did not use a filter to remove noise prior to segmentation, as is often done to avoid over-segmentation (Carleer et al., 2005; Chen et al., 2006; Sun and He, 2008). Our rationale for not doing so was two-fold. First, smoothing the gradient image reduces the spectral contrast that we seek to detect at the edge of segments (Gaetano et al., 2015). Second, our aim was to produce initial segments for subsequent merging, in which case over-segmentation is considered a good starting point.

2.3. Calculating the spectral distance between adjacent segments

We used a RAG to represent the spatial relationships between pairs of segments, with nodes and arcs representing segments and their

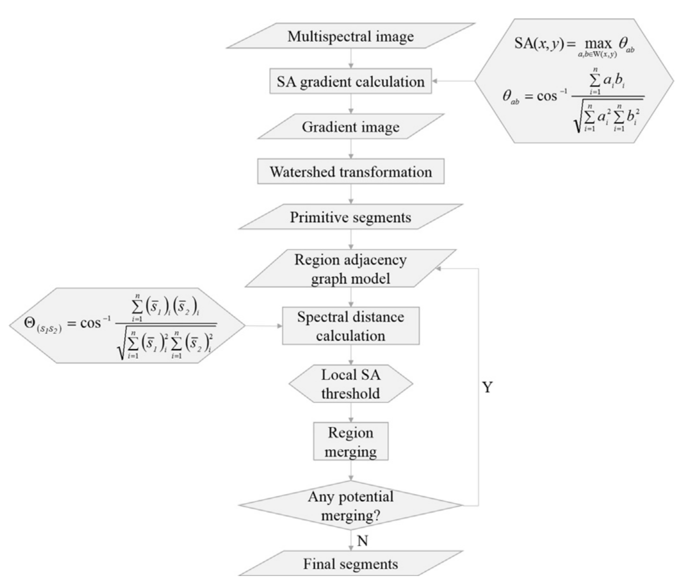

Fig. 1. Hybrid method for multi-scale segmentation using spectral angle thresholds for local region merging.