ISPRS Journal of Photogrammetry and Remote Sensing 88 (2014) 119–127

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Automated parameterisation for multi-scale image segmentation on multiple layers

L. Drăguţ [a,*], O. Csillik [a], C. Eisank [b], D. Tiede [b]

[a] Department of Geography, West University of Timişoara, V. Pârvan Blv. 4, 300223 Timişoara, Romania
[b] Interfaculty Department of Geoinformatics – Z_GIS, University of Salzburg, Schillerstraße 30, 5020 Salzburg, Austria

ARTICLE INFO

Article history:
Received 13 April 2013
Received in revised form 24 September 2013
Accepted 24 November 2013

Keywords:
Automation
Imagery
Object
Representation
GEOBIA
MRS

ABSTRACT

We introduce a new automated approach to parameterising multi-scale image segmentation of multiple layers, and we implemented it as a generic tool for the eCognition® software. This approach relies on the potential of the local variance (LV) to detect scale transitions in geospatial data. The tool detects the number of layers added to a project and segments them iteratively with a multiresolution segmentation algorithm in a bottom-up approach, where the scale factor in the segmentation, namely, the scale parameter (SP), increases with a constant increment. The average LV value of the objects in all of the layers is computed and serves as a condition for stopping the iterations: when a scale level records an LV value that is equal to or lower than the previous value, the iteration ends, and the objects segmented in the previous level are retained. Three orders of magnitude of SP lags produce a corresponding number of scale levels. Tests on very high resolution imagery provided satisfactory results for generic applicability. The tool has a significant potential for enabling objectivity and automation of GEOBIA analysis.
© 2013 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS) Published by Elsevier B.V. All rights reserved.

1. Introduction

Geographic object-based image analysis (GEOBIA) has been gaining prominence in the fields of remote sensing and geographic information science (GIScience) over the past decade, especially for the processing of high spatial resolution imagery (Blaschke, 2010). Creating representative image objects with image segmentation algorithms is an important pre-requisite for classification/feature extraction and further integration in geographical information systems (GIS) analysis. Multiresolution Segmentation (MRS) (Baatz and Schäpe, 2000) is probably the most popular algorithm for these purposes. Implemented in the eCognition® software (Trimble Geospatial Imaging), this algorithm quickly became one of the most important segmentation algorithms within the GEOBIA domain. MRS relies on a key control, called the scale parameter (SP), to partition an image into image objects. The SP controls the internal (spectral) heterogeneity of the image objects and is therefore correlated with their average size, i.e., a larger value of the SP allows a higher internal heterogeneity, which increases the number of pixels per image-object (Baatz and Schäpe, 2000; Benz et al., 2004).

Because the average size of image objects critically impacts on the classification accuracy (Gao et al., 2011), the selection of an accurate value of the SP is a crucial decision in segmenting remote sensing imagery (Kim et al., 2011). However, the standard procedure that leads to this decision is a trial-and-error optimisation (e.g. Duro et al., 2012), which is based on a visual assessment of segmentation suitability (Whiteside et al., 2011). While allowing flexibility in incorporating expert knowledge in GEOBIA, this procedure is hardly reproducible and raises important scientific issues with respect to the robustness of the approach (Arvor et al., 2013).

Since the SP is the key control in MRS and heavily impacts on the classification accuracy, making its selection a more objective decision (at least traceable or reproducible) is a hot topic in GEOBIA (Blaschke, 2010). According to Zhang et al. (2008), methods of evaluating the image segmentation quality to identify suitable segmentation parameters can be classified into supervised and unsupervised, aside from the standard visual assessment. Unsupervised methods can lead to the self-tuning of segmentation parameters, which is, thus, automation (Zhang et al., 2008). The concept of local variance (LV) graphs (Woodcock and Strahler, 1987) was introduced to GEOBIA by Kim et al. (2008) to determine the optimal SP for alliance-level forest classification of multispectral IKONOS images. Drăguţ et al. (2010) automated this approach and extended it into multi-scale analysis based on single layers and created a generic tool to detect the scales where patterns occur in the data, which is called the Estimation of Scale Parameters (ESP tool). Espindola et al. (2006) proposed an objective function that obeys the principles of regionalisation, namely, minimising the internal variance while maximising the external difference. Martha et al. (2011) further developed this approach into

* Corresponding author. Tel.: +40 720 163858.
E-mail addresses: lucian.dragut@cbg.uvt.ro (L. Drăguţ), cskovi@yahoo.com (O. Csillik), clemens.eisank@sbg.ac.at (C. Eisank), dirk.tiede@sbg.ac.at (D. Tiede).

0924-2716/$ - see front matter © 2013 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS) Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.isprsjprs.2013.11.018

multi-scale analysis. Johnson and Xie (2011) employed the same heterogeneity measures (weighted variance and Moran's I, respectively) to identify and further refine over- and under-segmented regions within a given scale level.

All of the above-mentioned existing methods require user interpretation, which hinders automation of the segmentation and of the rule-sets in a GEOBIA framework. Udupa et al. (2006) argued that segmentation methods cannot be automatic, which might be true when segmentation does necessarily include object recognition. In GEOBIA, however, segmentation is instead regarded as a pre-processing step (Castilla and Hay, 2008), and its results, namely, image objects, are rarely envisaged as end products. The process of endowing the image objects with meaning is a complex one (Castilla and Hay, 2008) and usually takes place in the classification step. From this perspective, automation of the segmentation process is a necessary step toward the automation of image processing in GEOBIA. While some degree of automation in segmentation has been achieved for specific tasks, for example, the extraction of tree-crown objects (Ardila et al., 2012), generic solutions are rare. Esch et al. (2008) developed a segmentation optimisation procedure that was based on spectral similarity between image objects at two scales in a hierarchy. Although MRS was employed to generate the two scales, segmentations were conducted without optimising the data; thus, the results do not directly depend on the segmentation itself but instead depend on the statistics of the arbitrarily-generated parent/children image objects. Drăguţ and Eisank (2012) proposed a concept for automating the optimisation of the SP, which has been successfully applied for automated object-based classification of topography from SRTM data. However, this approach works on a single layer, which hinders applications on multi-spectral data. In brief, a generic solution to automate the parameterisation in MRS is still missing, which is considered to be a disadvantage of GEOBIA (Whiteside et al., 2011) and a priority for further research (Jakubowski et al., 2013).

Building on the work of Drăguţ et al. (2010) and Drăguţ and Eisank (2012), this study introduced a fully automated methodology for the selection of suitable SPs relative to the patterns encoded within the data. Compared to previous approaches, this work considers multiple layers and implements a three-level hierarchy concept. Woodcock and Harward (1992) showed that a single-scale segmentation is an unrealistically simple scene model. On the one hand, some landscape elements are structured in nested hierarchies, for example, a forest composed of forest stands and individual trees (Woodcock and Harward, 1992). This concept is accommodated in eCognition® by building parent/child relationships when choosing the 'hierarchy' option in segmentation. On the other hand, visible features in a landscape are often multidimensional (e.g., small buildings coexisting with large agricultural fields), and each feature class is best represented at a certain scale (Martha et al., 2011). This issue is technically tractable by combining image objects of different sizes, which are created with the 'non-hierarchy option'. In any scenario, multi-scale segmentation is more suitable than single-scale to model image objects in a scene (Woodcock and Harward, 1992).

We demonstrated the performance of the tool in three test cases, on very high spatial resolution (VHR) multispectral imagery, in different applications scenarios. For these applications scenarios, we used expert delineated polygons and quantitative measures (Clinton et al., 2010) to evaluate the results of the segmentations.

2. Methods

The methodology comprises the computation of LV on multiple layers (Section 2.2), to allow optimal SPs to be selected automatically (Section 2.3). The workflow was implemented using eCognition Network Language (CNL), within the eCognition® 8.7.2 software, as a customised algorithm that is easy and ready to use (Section 2.4). The final outputs of the tool are assessed using quantitative measures (Section 2.5).

2.1. Study areas and data

Various test areas were chosen to assess the behaviour of the tool in diverse situations, ranging from urban settlements to semi-natural landscapes, as described in Table 1. We focused on areas for which we had access to VHR imagery. The first test area (T1) is located in Darfur, Sudan, and covers 2.31 km^2. It represents a semi-arid Sahel landscape that includes wadis, isolated trees and the Zam Zam internally displaced persons camp. Traditional (dark) huts and bright tents are the main dwelling types in the camp. A QuickBird scene, which was acquired on December 20th, 2004, by Digital Globe, Inc., was used for the T1 area. It includes a panchromatic band at 0.6 m spatial resolution with three visible (RGB) bands and one near-infrared (NIR) band at 2.4 m. The image was pan-sharpened to 0.6 m with the Gram-Schmidt spectral sharpening algorithm (Laben and Brower, 2000).

The T2 test area covers 2.25 km^2 in the western part of the city of Salzburg, Austria and includes residential, industrial and agricultural features. The T3 test area represents a semi-natural landscape at the border between Austria and Germany, between the cities of Salzburg and Oberndorf. Extended across 3.05 km^2 and crossed by the river of Salzach, it includes forests, agricultural fields and water bodies. T2 and T3 are covered by WorldView-2 satellite images that were acquired on September 11th, 2010 (T2) and July 9th, 2011 (T3). The original bands were: panchromatic at 0.5 m spatial resolution and multispectral at 2 m spatial resolution, namely, coastal blue, blue, green, yellow, red, red-edge, NIR 1, and NIR 2. The images were pan-sharpened to 0.5 m with the Hyperspherical Colour Sharpening algorithm implemented within ERDAS IMAGINE.

2.2. LV on multiple layers

To take full advantage of the multispectral information, segmentation on multiple layers is desirable. To accomplish this goal, a mean value of LV (*mean*LV) is computed for each image level that was created; the ratio between the sum of the LVs for each layer (LV *1*–LV*n*) and the number of layers (*n*) used in the image segmentation is given in (1):

$$\textit{mean}\ \mathrm{LV} = (\mathrm{LV}_1 + \mathrm{LV}_2 + \cdots + \mathrm{LV}_n)/n \quad (1)$$

The maximum number of layers is not limited because the tool automatically identifies the total number of layers within the scene as well as their names. This computation is implemented through an iterative process, using an index that allows scrolling through all of the layers that are loaded into eCognition®. To derive the mean of LV for the entire scene, each layer is selected; its value of LV is computed, LV(index), and added to the final variable, LV(n), which is divided by the total number of image layers present in the project (Fig. 1). The process is executed as long as the index value is smaller than the number of image layers, as recorded during the iteration. It is important to note that all of the layers included in the project are used to segment the scene into image objects. If a user wants to exclude specific layers/multispectral bands from the analysis, the layers can be loaded after the execution of the tool.

Download English Version:

https://daneshyari.com/en/article/6949731

Download Persian Version:

https://daneshyari.com/article/6949731

Daneshyari.com