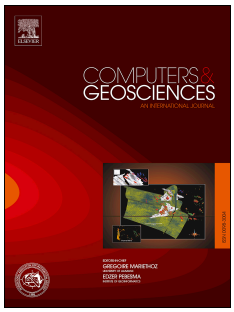
Accepted Manuscript



Multi-scale segmentation algorithm for pattern-based partitioning of large categorical rasters

Jarosław Jasiewicz, Tomasz Stepinski, Jacek Niesterowicz

PII: S0098-3004(17)30520-4

DOI: 10.1016/j.cageo.2018.06.003

Reference: CAGEO 4143

To appear in: *Computers and Geosciences*

Received Date: 7 May 2017

Revised Date: 25 May 2018

Accepted Date: 7 June 2018

Please cite this article as: Jasiewicz, Jarosł., Stepinski, T., Niesterowicz, J., Multi-scale segmentation algorithm for pattern-based partitioning of large categorical rasters, *Computers and Geosciences* (2018), doi: 10.1016/j.cageo.2018.06.003.

Multi-scale segmentation algorithm for pattern-based partitioning of large categorical rasters

Jarosław Jasiewicz[a,b,1], Tomasz Stepinski[b,2,*], Jacek Niesterowicz[b,2]

[a]*Institute of Geoecology and Geoinformation, Adam Mickiewicz University in Poznan, Krygowskiego 10, 60-680 Poznan*
[b]*Space Informatics Lab, Department of Geography,University of Cincinnati, Cincinnati, OH 45221-0131, USA*

Abstract

Analyzing large Earth Observation (EO) data on the broad spatial scales frequently involves regionalization of patterns. To automate this process we present a segmentation algorithm designed specifically to delineate segments containing quasi-stationary patterns. The algorithm is designed to work with patterns of a categorical variable. This makes it possible to analyze very large spatial datasets (for example, a global land cover) in their entirety. An input categorical raster is first tessellated into small square tiles to form a new, coarser, grid of tiles. A mosaic of categories within each tile forms a local pattern, and the segmentation algorithm partitions the grid of tiles while maintaining the cohesion of pattern in each segment. The algorithm is based on the principle of seeded region growing (SRG) but it also includes segment merging and other enhancements to segmentation quality. Our key contribution is an extension of the concept of segmentation to grids in which each cell has a non-negligible size and contains a complex data structure (histograms of pattern features). Specific modification of a standard SRG algorithm include: working in a distance space with complex data objects, introducing six-connected "brick wall" topology of the grid to decrease artifacts associated with tessellation of geographical space, constructing the SRG priority queue of seeds on the basis of local homogeneity of patterns, and using a content-dependent value of segment-growing threshold. The detailed description of the algorithm is given followed by an assessment of its performance on test datasets representing three pertinent themes of land cover, topography, and a high-resolution image. Pattern-based segmentation algorithm will find application in ecology, forestry, geomorphology, land management, and agriculture. The algorithm is implemented as a module of GeoPAT – an already existing, open source toolbox for performing pattern-based analysis of categorical rasters.

Keywords: segmentation, spatial patterns, categorical rasters, land cover, topography

1. Introduction

The goals and the means of analyzing Earth Observation (EO) data depend on the spatial scale of the data. Most analyses are performed on the scale of a single image which, in geographical terms, depicts a city or a non-urban region of an equivalent spatial extent. At such scale, the goal of the analysis is either to identify individual objects (like, for example, buildings) if the resolution of an image is sufficiently high, or to generalize an image into a map showing land use/land cover classes (LULC). The means for achieving these goals are either pixel-based classification (Li et al., 2014) or the Object-Based Image Analysis (OBIA) (Lang, 2008; Blaschke, 2010), which combines pixel-based segmentation with segments classification.

However, neither individual objects nor LULC classes are useful when analyzing EO data on broad spatial scales (province, country, continent, the entire Earth surface). This point is illustrated in Fig.1(A) which shows a fragment of

*Corresponding author
Email addresses: jarekj@amu.edu.pl (Jarosław Jasiewicz), stepintz@ucmail.uc.edu (Tomasz Stepinski), niestejk@mail.uc.edu (Jacek Niesterowicz)

[1]Design of the algorithm, lead programmer, early draft of the manuscript
[2]Team leader, conception, writing and revising the paper
[3]Co-programmer, testing, applications