RETRIEVAL OF PATTERN-BASED INFORMATION FROM GIGA-CELLS RASTERS - CONCEPT AND NEW SOFTWARE

Jaroslaw Jasiewicz
Adam Mickiewicz University, Institute of Geoecology and Geoinformation, Dziegielowa 27, 60-680 Poznan, jarekj@amu.edu.pl

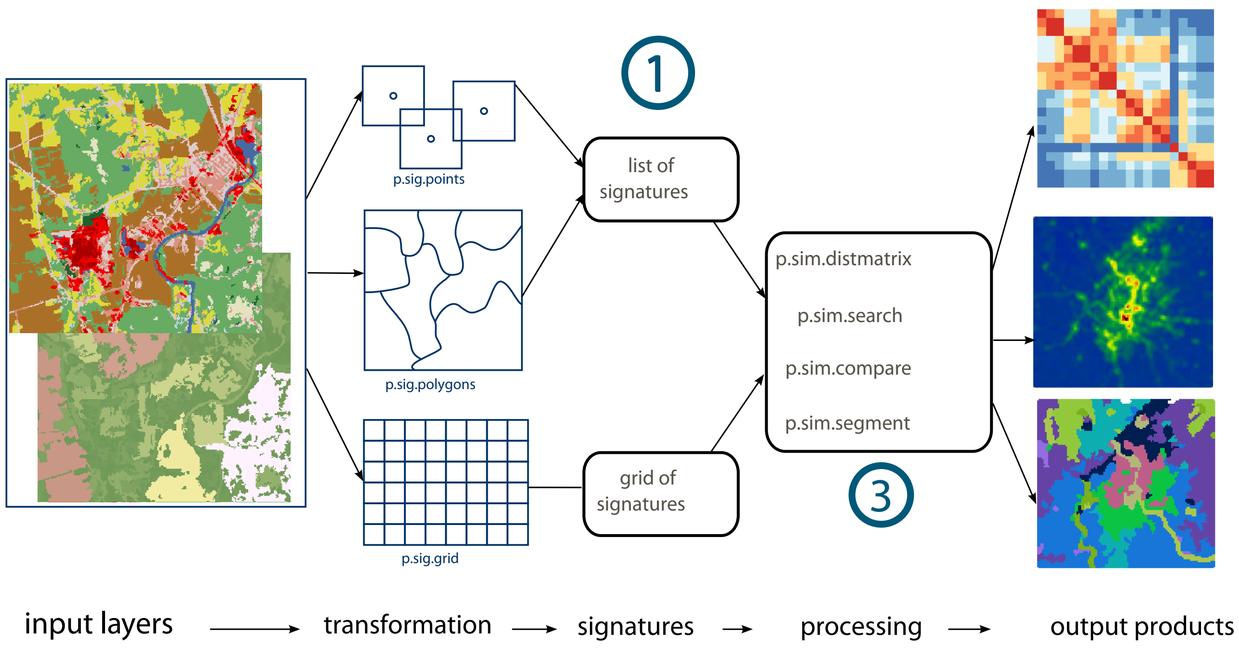
Tomasz F. Stepinski
University of Cincinnati, Space Informatics Lab, Department of Geography, Cincinnati, OH 45221-0131, USA, sepintz@uc.edu

Pawel Netzel
University of Wroclaw, Department of Climatology and Atmosphere Protection, Kosiby 6/8, 51-621 Wroclaw, Poland, pawel.netzel@uni.wroc.pl

INTRODUCTION

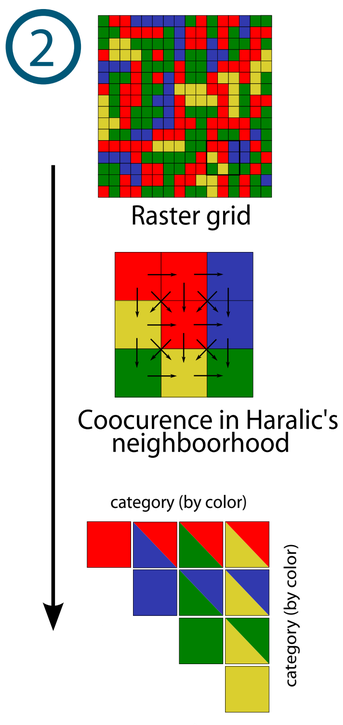
Rapid development of computer technology together with the growing availability of giga-scale data sources brings new possibilities to geo-spatial analysis. Giga-scale datasets are those having size exceeding 10 cells, regardless of their physical scale. They may represent local regions at ultra-high resolution or global mosaics of satellite imagery/DEMs at medium resolution. In the last case the giga-rasters are frequently categorical - products derived from processing of original multispectral image or DEM. They store important information at the level of patterns of the categories. Urban structures, plant habitats, geomorphological surfaces, and landscapes are examples of such patterns.

They have collective functions and meaning and thus contain valuable information that cannot be inferred at the level of cell-based analysis.
Analyzing these patterns in giga-scale datasets is not feasible by means of visual inspection and needs to be done by parsing the data by an ``intelligent" algorithm. In this paper we present the GeoPAT (Geospatial Pattern Analysis Toolbox) - a conceptual framework and software for retrieval of pattern-based information from giga-scale datasets. GeoPAT is applicable to all categorical datasets. Its use is llustrated here using a raster containing categories of landform elements derived from a DEM.

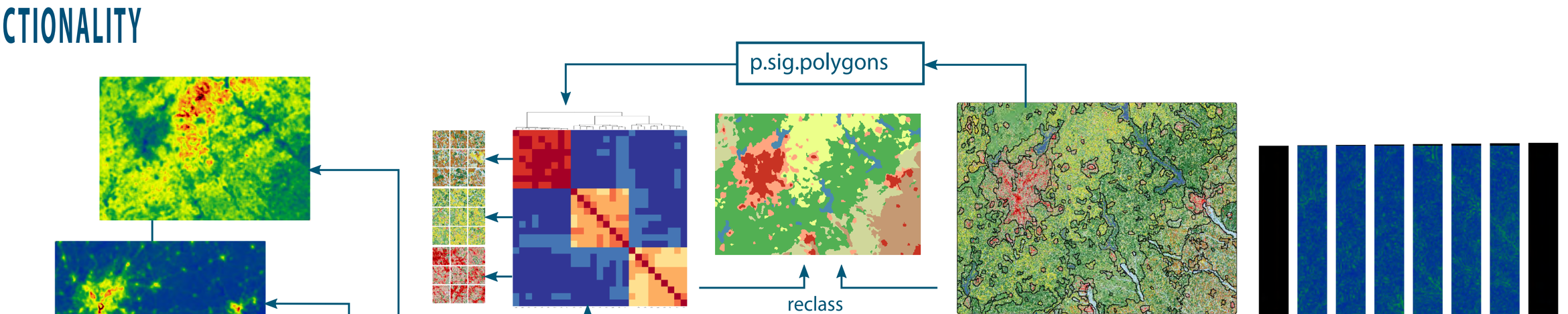


BASIC CONCEPT

A scene contains a local pattern of categories. A scene is small relative to raster but large relative to raster cell. Grid-of-scenes is a lattice of square scenes that covers the entire raster. Scene signature is defined as histogram of pattern features; different patterns may be described by different features. Similarity between patterns is calculated as a similarity between their signatures. Different similarity measures may be selected for signatures built using different features



FUNCTIONALITY



SEARCHING

Searching results in creation of new layer of information, having the same spatial extent and granularity as the grid-of-scenes. Each cell of this layer contains a similarity value between a user-selected query scene and all scenes included in the dataset.

CLUSTERING

Clustering reveals regularity (or lack of it) in a set of the scenes.Clustering tool p.sim.distmatrix calculates a similarity/ distance matrix between all the scenes. Further analysis, such as, various forms of clustering or visualization can be performed using that matrix.

SEGMENTATION

Segmentation delineates regions containing scenes with similar patterns. It can be utilized for finding LULC landscape types from land cover/land use datasets or physiographic provinces from landform datasets. The segments can be subsequently classified.

COMPARISON

Calculates similarity between colocated scenes in two different coregistered grids of scenes to assess change in local patterns. Can also be utilized for assessment of two categorical products created with different parameters.

This work was founded by the Polish National Centre of Science under GRANT DEC-2012/07/B/ST6/01206 and NSF GRANT BCS-1147702