GeoPAT – Pattern-based GIS Software for understanding content of large Earth science datasets

Tomasz Stepinski, Jaroslaw Jasiewicz, Pawel Netzel, Jacek Niesterowicz

Space Informatics Lab
http://sil.uc.edu
University of Cincinnati

ESTF 2016
Example of a dataset we are interested in

NLCD2011
categorical map
16 land cover categories
104,242 x 161,190 pixels

How to analyze such dataset in its entirely?
What relevant information can we extract from such dataset?
What are other, similar datasets?
What are the applications for such analysis?
Some answers…
Tools relevant to analyzing NLCD in its entirety:

Segmentation into regions of homogeneous land cover patterns

Spatial search

Temporal change

Las Vegas

2001-2006

2001

2006

change
Basic concepts

Grid of motifels

Motifel – elementary unit of analysis - a square block of pixels representing local landscape

Motifel is represented by histogram of features (co-occurrence, decomposition)
Distance between motifels is a distance between their histograms (Jensen-Shannon Divergence)
Difference between our method and object-based image analysis (OBIA)

**object-based image analysis**
- works best for small scenes
- identifies individual objects
- well-developed

**complex object-based image analysis**
- works best for large scenes
- identifies types of neighborhoods
- our project

irregular objects having simple, homogeneous content

regular objects having complex, inhomogeneous content
GeoPAT

Geospatial Pattern Analysis Toolbox (GeoPAT) is a collection of GRASS GIS modules, written in C, for carrying out pattern-based geospatial analysis of large categorical spatial datasets.

We are developing GeoPAT 2, a stand alone version of our software, which does not require GRASS or Linux.

The new version has expanded functionalities, features brand new segmentation module, and will be optimized to work with Xeon Phi co-processor.
GeoPAT segmentation algorithm is based on the principle of seeded region growing (SRG) but introduces a number of novel features in order to make the SRG applicable to the task of pattern-based segmentation of categorical rasters:

1. Motifels instead of pixels (new representation and distance)
2. Brick-wall topology of grid
3. Novel method of constructing a priority queue for potential seeds and adoption of locally determined growth-stopping criteria for regions growing from these seeds
Segmentation 3 of 8

region growing

- priority queue of motifs
  - take top motif
  - found? (N) stop
  - Y create new segment
    - add motif to segment
    - add neighbours to perimeter
    - add to segment the perimeter's motif best fitting the growing criterion
    - found? (Y) stop

main process

- building brick wall grid
- converting grid to histogram data
- calculating local homogeneities
- seeded segment growing
- segments merging
- removal of small segments
- swapping unmatched motifs
- segmented raster

region merging

- grow segments
- build graph on segments
- find most similar pair of segments
- can be merged? (N) change merging conditions
- Y merge
- recalculate affected pairs
- merge
- more than one segment? (Y) N N
- still small segments? (Y) stop
All calculations were performed on computer with Intel 3.4GHz, 4-core processor and 16 GB of memory running the Linux operating system.
Segmentation 6 of 8
Multi-scale segmentation by changing the size of motifel
We calculate regionalization of world climates using a concept of climate as time series (temporal pattern) and utilizing the Dynamic Time Warping (DTW) as similarity function.
GeoPAT is used to implement pattern-based searches as GeoWeb applications:

1. LandEx – search for similar patterns of land cover over the U.S.  
   sil.uc.edu/webapps/landex_usa/
2. ClimateEx – search for similar climates over the world. sil.uc.edu/webapps/climatex/
3. TerraEx – search for similar landforms over the world (beta)  
   sil.uc.edu/webapps/terraex/
Search 2 of 5

Search of similar urban landscapes in Houston, TX

all examples computed online by LandEx (sil.uc.edu)
query – Washington DC

search results for Washington

query – Miami, FL

search results for Miami

all examples computed online by ClimateEx (sil.uc.edu)
Saudi Arabia Empty Quarter

all examples computed online by TerraEx (sil.uc.edu) (beta)
all examples computed online by TerraEx (sil.uc.edu) (beta)
A degree of dissimilarity between the same locations (motifels) at two different times is a measure of change in pattern.
Change in land cover patterns over conterminous U.S. 2001-2006
Climate change 2000 - 2070

all examples computed online by ClimateEx (sil.uc.edu)
Climate change -6000 -2000

all examples computed online by ClimateEx (sil.uc.edu)
Stand-alone version of GeoPAT

Present version of GeoPAT requires GRASS and (in most cases) Linux because GRASS needs to be compiled together with GeoPAT.

GeoPAT 2 is GRASS-free. All computations can be performed in terminal mode and exported to GIS software of choice (ArcGIS, QGIS, etc).

GeoPAT 2 will be provided as source code and as executable for Linux and Windows.

We have already converted several GeoPAT modules GRASS environment to stand-alone and demonstrated to work on Windows.

GeoPAT 2 will be open and free software
Summary

GeoPAT fills a niche devoted to analysis of large rasters (starting with GeoPAT 2 they don’t need to be categorical).

Public data:

1. Land cover (NLCD, GLC30, GlobCover, CORINE, EOSD)
2. Topography (NED, SRTM)
3. Croplands (USDA CropScape)
4. Urban Structure Types (UST) (National Map)

Applications:

1. Ecology (delineation of ecoregions)
2. Forestry (conservation planning and management)
3. Agriculture (identify spatio-temporal patterns of usage of U.S. croplands)
4. Climate (climate change)
Roadmap for the next 12 months

GeoPAT:

1. Finish development and release GeoPAT 2 for GRASS
2. Finish development and release the stand-alone GeoPAT 2
3. Finish development and release GeoPAT 2 optimized for Xeon Phi

Online applications:

1. Finish TerraEx

Testing GeoPAT on public data

1. Calculate and validate regionalization of US into physiographic units
2. Calculate and validate regionalization of US into landscape types
3. Validate utility of GeoPAT for delineation of UST in American cities