Doing Geomorphometry with Pattern Analysis

Tomasz F. Stepinski[1], Jaroslaw Jasiewicz[2,1], Pawel Netzel[3,1], Jacek Niesterowicz[1]
[1]University of Cincinnati, USA; [2]Adam Mickiewicz University, Poland; [3]University of Wroclaw, Poland

Abstract—**We have developed a concept of pattern-based analysis of land-surface where a basic unit of analysis is a pattern of landform elements defined over an arbitrary local region referred to as a scene. The DEM is subdivided into a regular grid of scenes reducing its effective dimension by orders of magnitude. Scenes are described by histograms of local pattern primitive features and similarities between scenes are calculated using histogram distance measures. With scene histograms as cell attribute and histogram distance as a metric the grid-of-scenes can be analyzed much like the ordinary DEM including its segmentation and classification. The result is a framework for efficient and robust automatic classification of topography on continental or global scales using extensive DEM archives. The framework also supports spatial search for similar landscapes. The concept of pattern-based analysis is described and an example pertaining to automatic delineation of physiographic units in Poland is presented. In addition, a GeoWeb application for spatial search of landscapes in Poland is also discussed. To facilitate pattern-based analysis we have developed GeoPAT – a toolbox of GRASS GIS modules intended as a platform for experimentation with the pattern-based analysis of DEMs and other spatial datasets.**

I. Introduction

Geomorphometry is the science of quantitative land-surface analysis. A significant portion of this analysis focuses on surface classification, or more specifically on DEM (digital elevation model) classification as DEM is the most commonly used quantitative representation of the surface. Classification converts DEM into a thematic map of semantically meaningful classes. Landform elements – elementary forms characterized by constant values of morphometric variables – are the most popular target classes of classification [1,2,3]. This stems from a traditional geomorphologic interest in relating land-surface form to physical process. Increased availability of continental and global scales DEMs led to an additional, different rationale for performing DEM classification – an objective algorithmic delineation of different types of topography. This is in-line with interest in other disciplines of geosciences in providing global, objective delineations of geospatial variables such as, for example, land cover or land cover types classes [4] or climate classes [5].

An original approach to algorithmic classification of global topography [6] utilized a cell-based methodology with classification algorithm assigning class labels to individual cells in a DEM. Note that this is fundamentally different from how a human interprets a visualization of a DEM by perceiving the coherence of different landforms on multiple scales simultaneously and assigning a topographic class label to extended tracts (not an individual cell) of the surface on the basis of pattern of different landforms. Therefore cell-based classification algorithms suffer from poor performance especially if applied to high resolution DEMs, where individual cells correspond to small elements of surface and their associated numerical attributes are not sufficient to recognize the topographic class, or, if applied to very large DEMs where the goal of analysis is to retrieve generalized topographic classes (physiographic units).

Object-based classification of topography was developed [7] to alleviate the problems associated with cell-based classification. In the object-based method the DEM is first segmented into “objects” – tracts of surface homogeneous with respect to cell-based morphometric variables – which in turn are classified into topographic classes. Object-oriented algorithms get closer to the way an analyst interpret a DEM but they still suffer from a number of shortcomings First, segmentation itself is a complex and computationally expensive process and there is no single method that performs consistently well (does not under-segment or over-segment portions of a DEM). Second, because objects are, by definition, homogeneous segments of the surface, current object-oriented methods can only classify DEM into very general topographic classes (see [7]) as they are not able to take advantage of the information contained in the pattern of landform elements constituting a landscape.

To get a more robust means of classifying topography from a DEM we have developed a pattern-based method that has proven to be fast and effective on even the largest datasets. In our method a DEM is divided into a regular grid of local blocks of cells (referred to as local scenes) thus converting a large DEM into much smaller grid-of-scenes at very small computational cost. The core ingredients of the method are the mathematical description of a topographic pattern in each scene and a function that calculates a degree of similarity between the patterns. With pattern representation and similarity function defined, the grid-of-scenes can be segmented and classified in a manner similar to an ordinary DEM but at a small fraction of computational cost and at significantly higher degree of information generalization.

In: Geomorphometry for Geosciences, Jasiewicz J., Zwoliński Zb., Mitasova H., Hengl T. (eds), 2015. Adam Mickiewicz University in Poznań
- Institute of Geoecology and Geoinformation, International Society for Geomorphometry, Poznań

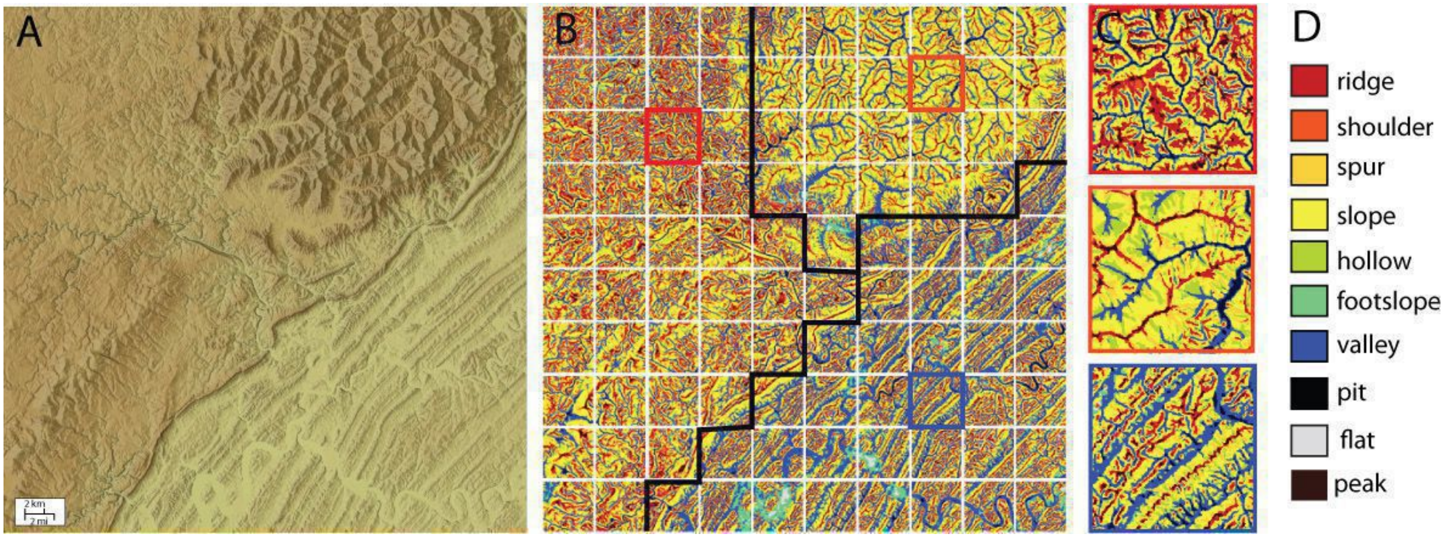

Figure 1. (A) Hillshade rendition of the 2000×2000 cells DEM. (B) Cell-based classification of the DEM into ten landform elements and its sub-division into a grid-of-scenes. (C) Examples of three scenes representing three different physiographic regions present in the region. (D) Landform elements legend.

II. METHODOLOGY

Fig.1 illustrates the concept of pattern-based analysis of DEM. Fig.1A shows a hillshade rendition of a 2000×2000 cells DEM. In the depicted region one can observe at least three distinct physiographic units. The DEM cells are first classified into ten landform classes (Fig.1D) using the geomorphons method [3]. The result of this classification is shown in Fig.1B. The region is then divided into a regular 10×10 grid of hundred scenes (Fig.1B) with each scene containing 4×10^4 cells forming a local, block-bounded pattern of landforms.

To delineate the three physiographic units as seen in the DEM the method segments and/or classifies the coarse grid-of-scenes in a way that is in general analogous to how a cell-based algorithm would perform these tasks on the entire DEM. Significant technical differences in performing these operations on scenes vs. cells stem from differences in mathematical representations of patterns vs. numbers, and from differences in the definitions of a distance between patterns vs. distance between vectors.

Three scenes representative of three different physiographic units are selected from the grid and highlighted by red, orange, and blue frames, respectively. Fig.1C shows close-ups of these scenes showing distinct patterns of landform elements in each scene. Either unsupervised (clustering and/or segmentation) or supervised methods can be used to delineate the three physiographic provinces from the grid-of-scenes. This is schematically shown by black lines on Fig.1B.

To perform pattern-based analysis of DEMs (and other datasets) we have developed the Geospatial Pattern Analysis Toolbox (GeoPAT) - a collection of GRASS GIS modules that integrates the various tools necessary for pattern-based analysis of DEMs including a classification task as described above. GeoPAT integrates into the GIS system procedures for pattern description, pattern similarity, and the search and retrieval of similar patterns. These concepts were originally developed for working with natural images in the context of Content-Based Image Retrieval (CBIR) systems [8] but are now utilized by GeoPAT for the purpose of geospatial analytics. GeoPAT works with DEMs of all sizes but it is designed to be especially effective when applied to giga-cell and larger DEMs. In addition to segmentation and classification GeoPAT supports tasks such as spatial search and scene-by-scene comparison of two grids. GeoPAT is available at http://sil.uc.edu/.

A. Pattern representation

The input to the pattern-based method is not an original DEM but a categorical grid of the same size as the DEM. This grid is a result of classifying cells of the DEM into landform elements. We use the geomorphon method [3] to achieve this pre-processing step but other methods (for example, [6]) can also be used. A concise mathematical description of categorical pattern

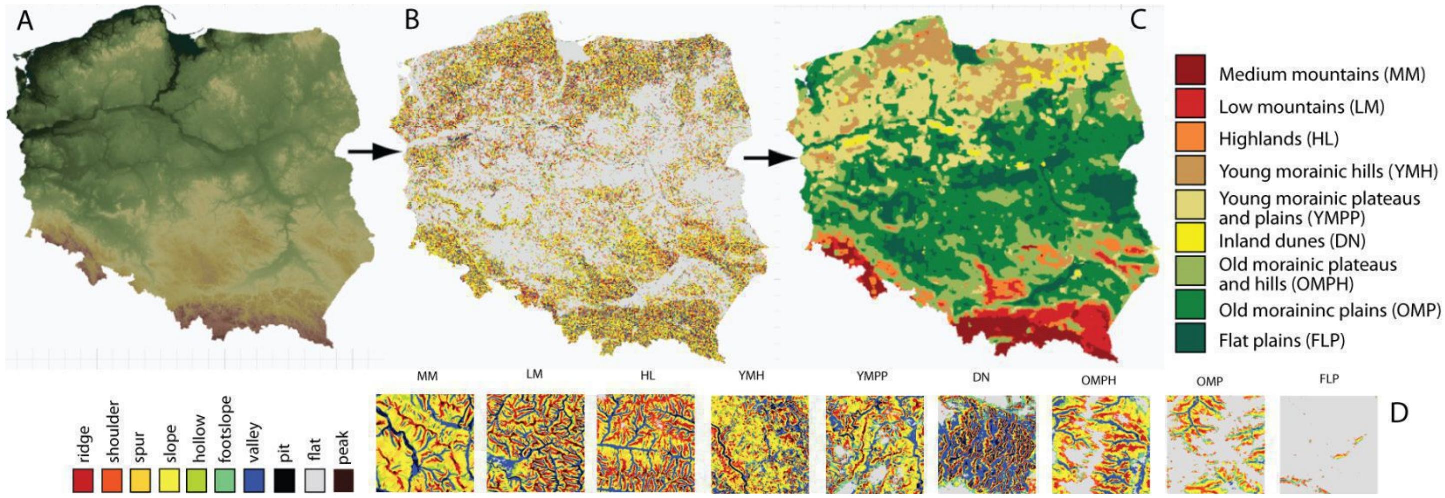

Figure 2. (A) DEM of Poland with 30m resolution. (B) Ten-categories map of landform elements with 30m resolution. (C) Results of pattern-based supervised classification of grid-of-scenes into nine physiographic units. (D) Examples of characteristic patterns of landforms for the nine physiographic units.

of landform elements in a scene is a histogram of pattern "primitive features." Primitive features are simple local elements of a pattern. GeoPAT implements several popular methods of representing pattern by a histogram of primitive features, a co-occurrence method is recommended for working with patterns stemming from topographic data.

The co-occurrence method is a variant of the Gray-Level Co-occurrence Matrix (GLCM) [9] with gray-scale values replaced by landform element classes. Co-occurrence method uses a single primitive feature - a pair of landform elements classes assigned to two neighboring cells. When the DEM is classified by the geomorphons method into 10 landforms elements the co-occurrence histogram has $(10 \times 10)/2 + 5 = 55$ bins. Thus, a pattern in each scene is encapsulated by 55 numbers describing the composition of different landforms and their relative spatial configuration.

B. Pattern similarity/distance

"Distance" between two scenes assesses the degree of dissimilarity between them, it is the opposite of similarity. In our method a distance between scenes is a distance between histograms representing the scenes. When the value of distance is equal to zero identical histograms are indicated, and thus scenes have identical or very similar patterns, whereas large values of the distance indicate very different histograms and scenes having significantly different patterns. Over 40 possible histogram distance measures have been proposed [10]. For topographic data and co-occurrence signature the normalized Wave Hedges distance metric is recommended.

III. Examples

We apply pattern-based methodology to delineate physiographic unit in the country of Poland using supervised approach [11]. The input is the 30m DEM with the size of $21{,}696 \times 24{,}692$ cells (Fig.2A). This DEM is classified using the geomorphons method (Fig.2B). The region is subdivided into grid-of-scenes with 433×493 coarse cells each having size of 50×50 cells (1.5 km scale) and being a center of 15×15 km scene resulting in a significant overlap between scenes.

The territory of Poland exhibits a number of physiographies and we decided, based on the prior knowledge, to map nine selected physiographic units (see legend to Fig.2C). For each of these units a number of representative scenes have been selected as examples; one example for each unit is shown in Fig.2D. The pattern of all scenes was encapsulated using the co-occurrence signature and the Wave-Hedges distance function was used to calculate dissimilarity between scenes. The nearest neighbor supervised classification was used to assign one of nine labels to each cell in the grid-of-scenes. The resultant map shown in Fig.2C is comparable to a manually developed physiographic map of Poland [12].

Another application of pattern-based method is web-based landscape search which enables the discovery of locations having landscapes similar to a specified landscape of interest. Given a query – landscape (topography) of the site of interest – landscape search returns the similarity map which visually shows a degree of similarity to a query at all location throughout the entire study. Similarity map provides much more information than a non-spatial list of top matches to a query. By utilizing spatial organization it simultaneously shows similarity relations between the query and all scenes in the database. Thus, it allows an analyst to concentrate on revealed geomorphic phenomena rather than on similarity between specific scenes. For the Poland data we have implemented the landscape search as TerraEx-PL http://sil.uc.edu/webapps/terraex_pl/ (see Fig.3).

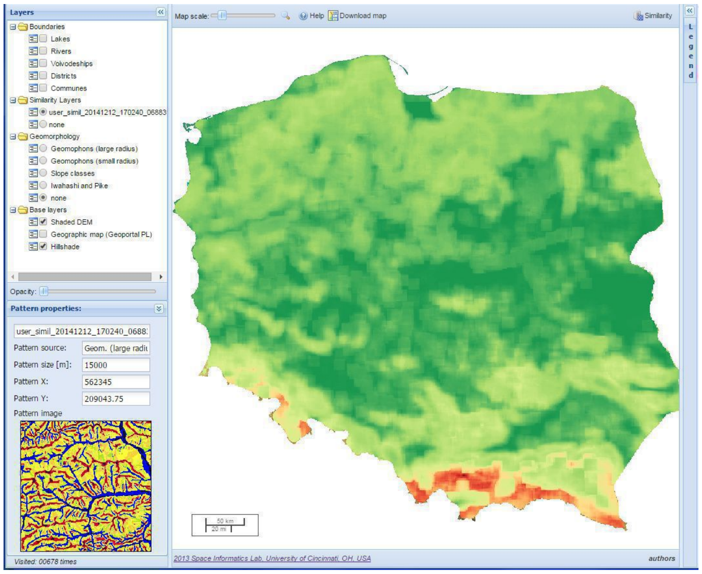

Figure 3. Screenshot of TerraEx-PL GeoWeb application for performing spatial searches for landscapes similar to a user-selected query. The query is shown in the bottom-left corner and the output is the similarity-to-query map with colors red-to-green indicating decreasing similarity of local landscapes to the landscape of the query.

IV. SUMMARY

Our pattern-based approach analyses the land-surface at the high level of generalization with the basic unit of analysis being a local landscape rather than more basic landform element. Such approach works best with very large DEMs and when the object of analysis is exploration or large-scale mapping of topography. The concept of pattern-based analysis of DEMs is new and will require much more work to mature. In particular, the key issues of scene signature and scene distance/similarity needs more study. We have developed the GeoPAT – toolbox of GRASS GIS modules intended as a convenient platform for experimentation with the pattern-based analysis of DEMs and other spatial datasets including datasets having giga-cell and larger sizes. In addition to classification, our pattern-based approach yields itself to spatial search function that can be implemented as web application.

ACKNOWLEDGMENT

This work was supported partially by the National Science Foundation under Grant BCS-1147702, by the National Science Centre (NCN) under Grant DEC-2012/07/B/ST6/01206, and by the University of Cincinnati Space Exploration Institute.

REFERENCES

[1] Wood, J., 1996. “The geomorphological characterisation of digital elevation models”. Ph.D. Thesis, PhD Thesis Department of Geography, University of Lancaster, UK.

[2] Dikau, R.; Brabb, E. & Mark, R. Morphometric landform analysis of New Mexico Zeitschrift für Geomorphologie, 1995, Supplement 101, 109-126

[3] Jasiewicz, J., Stepinski, T.F 2013. “Geomorphons -a pattern recognition approach to classification and mapping of lanforms.” Geomorphology 182, pp. 147-156.

[4] Jin, S., Yang, L., Danielson, P., Homer, C., Fry, J., and Xian, G. 2013. “A comprehensive change detection method for updating the National Land Cover Database to circa 2011”. Remote Sensing of Environment, 132, 159 – 175.

[5] Metzger, Marc J., Robert GH Bunce, Rob HG Jongman, Roger Sayre, Antonio Trabucco, and Robert Zomer. 2013 "A high‐resolution bioclimate map of the world: a unifying framework for global biodiversity research and monitoring." Global Ecology and Biogeography 22, no. 5, 630-638.

[6] Iwahashi, J., Pike, R., 2007. “Automated classification of topography from DEMs by an unsupervised nested-means algorithm and three-part geometric signature.” Geomorphology 86, 409-440.

[7] Dragut, L., Eisank, C., 2012. “Automated object-based classification of topography from SRTM data.”Geomorphology 141-142, 21--23.

[8] Datta, R., Joshi, D., Li, J., Wang, J.Z. 2008. “Image Retrieval: Ideas, Influences, and Trends of the New Age.” ACM Computing Surveys, 40,1—60.

[9] Haralick, Robert M., Karthikeyan Shanmugam, and Its' Hak Dinstein. 1973. "Textural features for image classification." IEEE Transactions on Systems, Man and Cybernetics, 610-621.

[10] Cha, S.-H. 2007. “Comprehensive survey on distance/similarity measures between probability density functions. “International Journal of Mathematical Models and Methods in Applied Sciences, 1(4), 300-307.

[11] Jasiewicz, J., Netzel, P., Stepinski, T.F. 2014. “Landscape similarity, retrieval, and machine mapping of physiographic units”. Geomorphology 221, pp.104-112

[12] Kondracki, J., 2002. ”Geografia regionalna Polski.” 3rd Edition. Wydawnictwo Naukowe PWN, Warszawa.