Regionalization of the conterminous U.S. into hierarchical landscape pattern types

J. Niesterowicz[1], T. F. Stepinski[1], J. Jasiewicz[2]

[1]Space Informatics Lab, Department of Geography, University of Cincinnati, Cincinnati, OH 45221-0131, USA
Telephone: +1 513 .556.3583
Email: niestjk@mail.uc.edu , stepintz@uc.edu

Adam Mickiewicz University, Dziegielowa 27, 60-680 Poznań
Email: jarekj@amu.edu.pl

Abstract

We present a pattern-based regionalization of the conterminous U.S. – a partition of the country into a number of mutually exclusive and exhaustive regions that maximizes the intra-region stationarity of land cover patterns and inter-region disparity between those patterns. The result is a discretization of land surface into a number of landscape pattern types (LPTs). First, the entire NLCD is partitioned into a grid of 500 × 500 pixels blocks (15 km scale). The NLCD classes of pixels within a block forms a local landscape pattern which is mathematically represented by a histogram of co-occurrence features. Using the Jensen-Shannon divergence as dissimilarity function between the patterns we cluster the local landscape patterns into several LPTs. The broad-extent maps of progressively more generalized LPTs are shown and discussed.

Keywords: land cover patterns, regionalization, segmentation, clustering, NLCD.

1. Introduction

A landscape pattern type (LPT) is defined (Wickham and Norton, 1994) as a spatial unit containing a unique quasi-stationary (small spatial gradient) pattern of land covers classes. Identification and delineation of LPTs in a large area makes possible the creation of broad-extent maps which are of significant interest for conservation, planning, as well as for academic research. Availability of high resolution land cover datasets on continental and even global scales makes possible to approach the problem of regionalization of large areas into LPTs from a computational perspective. In this paper we present a series of regionalizations of the entire conterminous U.S. into progressively more generalized LPTs using recently developed method (Jasiewicz and Stepinski, 2013; Niesterowicz and Stepinski, 2013, Jasiewicz et al., 2015) that subdivides a land cover dataset into an arbitrary grid of local landscape patterns (LLPs) and hierarchically segments that grid into regions consisting of similar LLPs. We consider resultant regions to constitute the sought-after LPTs.

Previous attempts at delineating LPTs algorithmically include the works by Cardille and Lambois (2009) and by Partington and Cardille (2013). The first of these two works has identified 17 different LPTs across the U.S. but did not delineate their spatial extents so no map was produced. The second work was applied to the Canadian province of Quebec and has identified and mapped 5 LPTs. Both methods used a vector of landscape indices to characterize LLPs and the Euclidean distance to cluster LLPs in data space.

The data clusters are taken to be the south-after LPTs. We use a different approach that utilizes histograms of pattern co-occurrence features to represent local landscapes and uses probabilistic dissimilarity measure (Jensen-Shannon divergence) between two histograms to segment and cluster the grid of LLPs into the LPTs.

2. Data and Methods

We use the National Land Cover Dataset (NLCD 2011) as an input. The NLCD has 16 land cover classes and covers the entire conterminous U.S. We divide the NLCD into a regular square grid of 500 × 500 pixels blocks setting the scale of a local landscape at 15 km. A pattern of NLCD classes in each block constitutes a LLP. We calculate co-occurrence pattern features (Chang and Krumm, 1999) from the pixels in the block and construct their histogram. A co-occurrence feature is a pair of classes assigned to two neighboring pixels. Further calculations are performed on a 322 × 209 grid of LLPs with each grid cell characterized by a 136-bins co-occurrence histogram. Dissimilarity between LLPs is calculated using the Jensen-Shannon divergence between the histograms which represent them. The grid is first segmented into 6217 segments using a region growing algorithm. The segmentation step is performed to reduce (by an order of magnitude) the number of objects to be clustered and to assure spatial cohesion of LPTs. In the second step the segments are clustered using a hierarchical clustering algorithm.

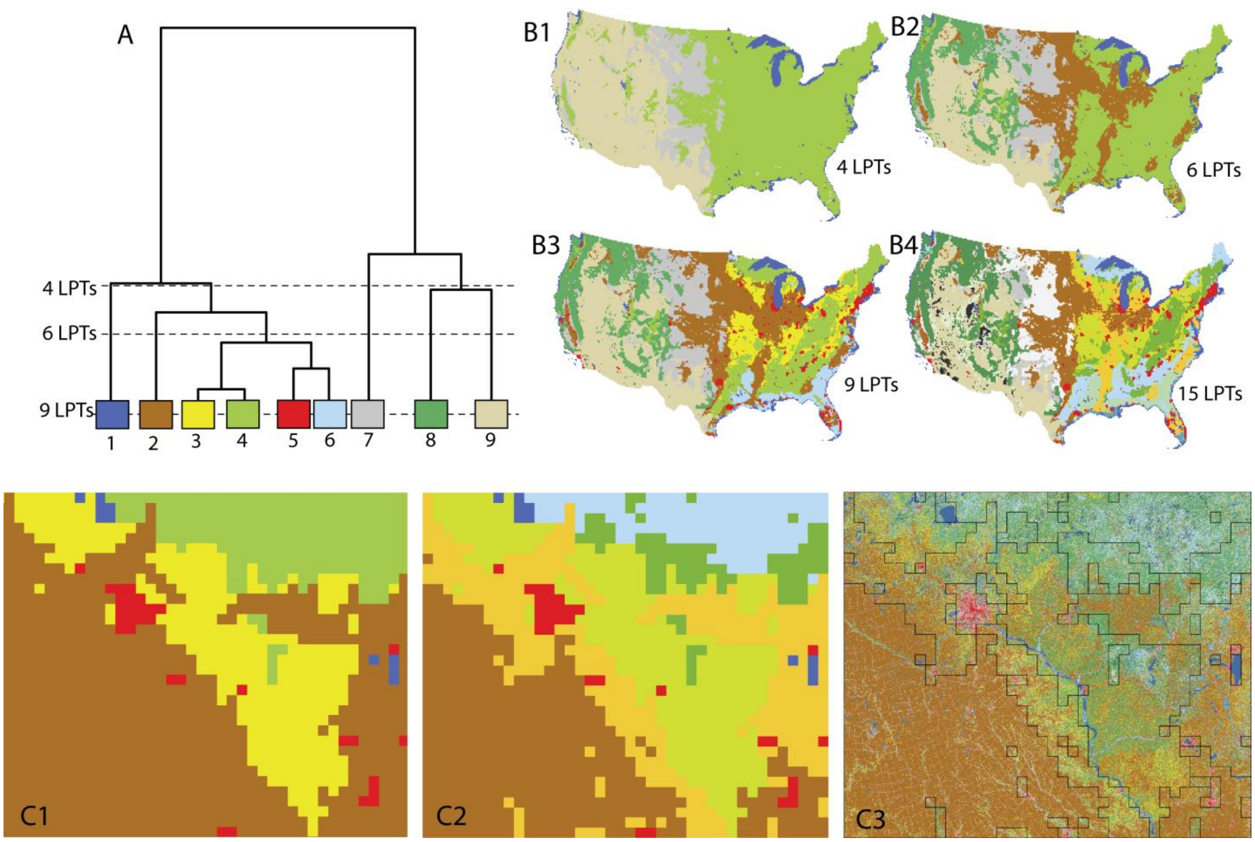

Figure 1. Hierarchical regionalization of LPTs. (A) Dendrogram illustrating merging of LPTs from nine types down to one. (B1 to B4) Regionalization maps for a number of

LPTs as indicated. (C1 to C3) Illustration of hierarchical merging of LPTs. Legends to LPTs and NLCD are given in Fig.2

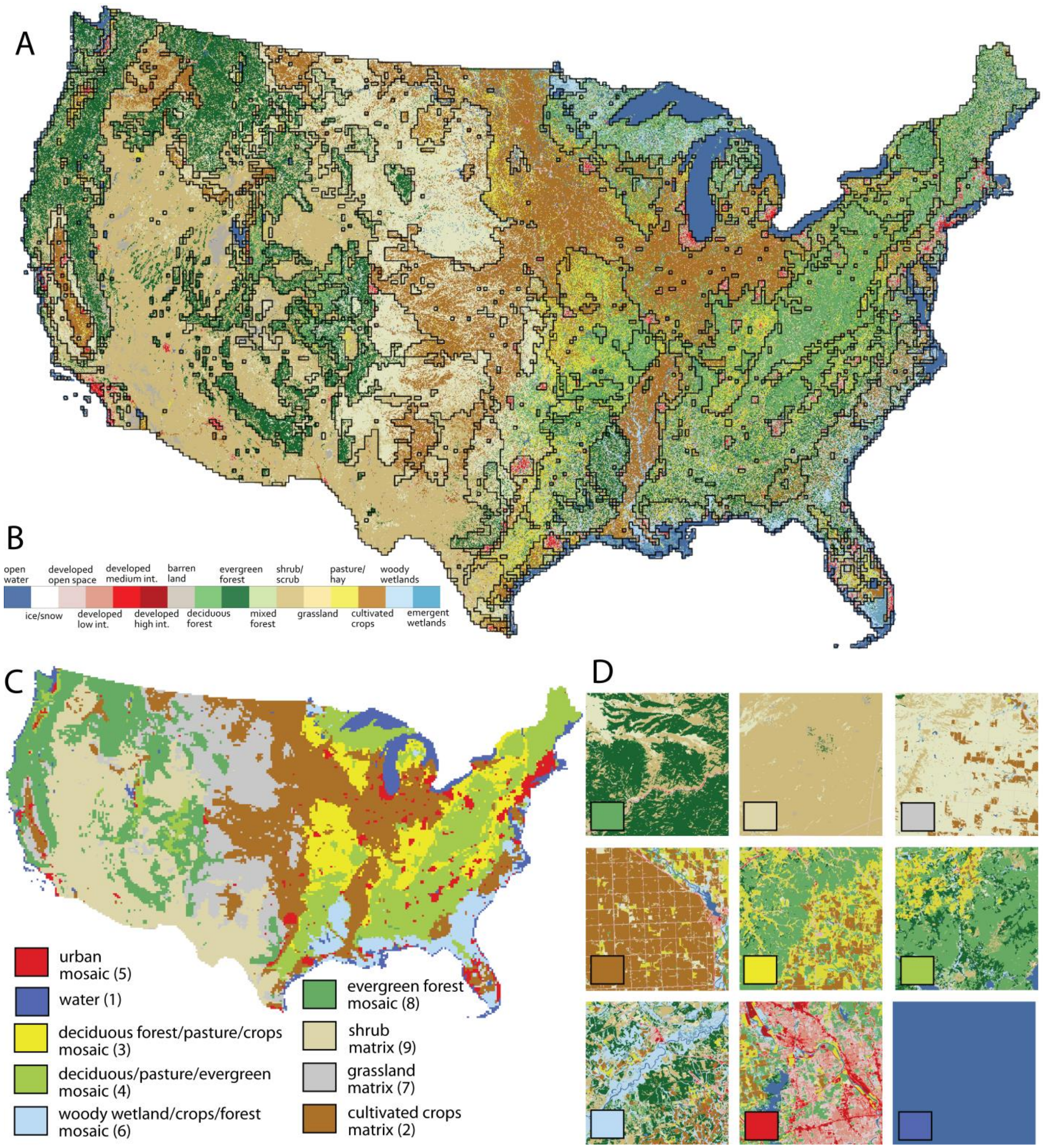

Figure 2. (A) Boundaries of nine LPTs superimposed on NLCD. (B) NLCD legend. (C) Map of nine LPTs with a legend. (D) Characteristic land cover patterns for each LPT.

3. Results

The results of hierarchical clustering are visualized by a dendrogram (Fig. 1A). The dendrogram has been cut at the level of 9 nodes although it extends down to 6217 original segments. Each node represents a LPT. Any number of LPTs, from 1 to 6217 can be selected for mapping, but the smaller the number of LPTs the more generalized patterns they represent. Figs. 1B1 to 1B4 show maps of four, six, nine, and fifteen LPTs, respectively. The legend to nine LPTs is given in Fig. 2B. Each more generalized LPT consists of a number of more specific LPTs. This is illustrated in Figs. 1C1 to 1C2 which show regionalization maps zoomed into a region centered on Minneapolis, MN. A map in Fig.1C1 corresponds to nine LPTs (nationwide) and a map in Fig.1C2 corresponds to delineation of 15 LPTs. Fig.1C3 shows the NLCD with superimposed boundaries of 15-LPTs delineation. More generalized LPTs of cultivated crops matrix (brown) and deciduous forest mosaic (light green) are divided into more specific LPTs.

Fig.2 shows the details of regionalization of NLCD into nine LPTs. The boundaries of nine LPTs are superimposed on the map of NLCD in Fig.2A and the NLCD legend is given in Fig.2B. The map of nine LPTs and their legend is given in Fig.2C. Naming LPTs is difficult because they are complex. We use a term matrix to describe a pattern dominated by a single NLCD class and the term mosaic to describe pattern to which several NLCD classes contribute significantly. Fig.2D shows medoids of all nine LPTs. A medoid is the LLP whose average value of dissimilarity to all other LLPs in the LPT is minimal; it represents a typical landscape in the LPT.

4. Conclusions

We presented an exploratory approach to the analysis of the entire NLCD using an unsupervised machine learning methodology. The novel aspect of our methodology is a conversion of a very large NLCD into a smaller grid of local landscape patterns. This enables an efficient computation of regionalization using a grid of local landscapes instead of raster of land cover classes. The crucial technical know-how is the numerical representation of LLPs and means of measuring dissimilarity between LLPs. Because our method works with categorical data (land cover data) it is computationally efficient and could be applied to large datasets (like the NLCD). A similar methodology has been developed (Vatsavai 2013) to delineate LPTs directly from a multispectral image but at the price of much greater computational cost which limits its applications to relatively small regions.

As the method is exploratory, an analyst may choose the values of free parameters. Here we study landscape patterns on the scale of 15 km. Another choice is a number of LPTs to map. Hierarchical character of our regionalization helps to make this decision, which is subjective and depends on how specific or generalized LPTs are needed to illustrate a given point. An interesting insight is gained by studying a dendrogram (Fig.1A). With only two LPTs selected the U.S. is basically divided into western and eastern parts. With six LPTs selected the western part breaks into evergreen forest and shrub and the eastern part breaks into cultivated crops and deciduous forest. There are also two additional LPTs corresponding to grassland (in the middle of the country) and water. With nine LPTs selected a deciduous forest LPT (located predominantly in the eastern part of the U.S.) breaks into four LPTs one of which includes urban mosaic. We think that nine LPTs reflect well the major patterns of land cover in the U.S. on the scale of 15 km. More specific LPTs can be obtained by analyzing more LPTs. Investigating smaller scales of pattern may reveal additional types pattern types (like small towns) which are just not present at the larger scale.

5. Acknowledgements

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

6. References

Cardille, JA, Lambois, M, 2009, From the redwood forest to the Gulf Stream waters: human signature nearly ubiquitous in representative US landscapes, *Frontiers in Ecology and the Environment*, 8(3): 130-134.

Chang, P, Krumm, J, 1999, Object recognition with color cooccurrence histograms. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society Conference.

Jasiewicz, J, and Stepinski, TF, 2013, Example-based retrieval of alike land-cover scenes from NLCD 2006 database. *IEEE Geosci. Remote Sens. Lett*., 10(1): 155-159.

Jasiewicz, J, Netzel, P., and Stepinski, TF, 2015, GeoPAT: A toolbox for pattern-based information retrieval from large geospatial databases. *Computers & Geosciences*, 80: 62-73.

Niesterowicz, J, Stepinski, TF, 2013, Regionalization of multi-categorical landscapes using machine vision methods. *Appl. Geogr*., 45: 250-258.

Partington, K, Cardille, JA, 2013, Uncovering dominant land-cover patterns of Quebec: Representative landscapes, spatial clusters, and fences. *Land*, 2(4):756-773.

Vatsavai, KK, 2013, Gaussian multiple instance learning approach for mapping the slums of the world using very high resolution imager. Proceedings of the 19th ACM SIGKDD International Conference on Knowledge discovery and data mining: 1419-1426

Wickham, JD, Norton, DJ, 1994, Mapping and analyzing landscape patterns. *Landscape Ecology*, 9(1): 7-23.