# PATTERN-BASED ASSESSMENT OF 2001/2006 LAND COVER CHANGE OVER THE ENTIRE UNITED STATES

*Pawel Netzel**

University of Wroclaw
Dept. of Climatology and Atmospheric Protection
Kosiby 6/8, 51-621, Wroclaw, Poland

*Tomasz F. Stepinski*†

University of Cincinnati
Space Informatics Lab
Cincinnati, OH 45221-0131, USA

***Index Terms***— Land cover change, pattern recognition, NLCD 2001 and 2006

## 1. INTRODUCTION

Remote sensing provides repetitive coverage of the terrestrial landmass with imagery data of relatively high resolution and consistent quality. Availability of such data coupled with advances in image processing algorithms makes possible autodetection of changes in Earth's surface features. Land cover change is the most popular form of surface change analysis. Numerous techniques of land cover change have been developed [1], most of them pertaining to tracking the change at the level of the individual raster cell. Such techniques are useful for assessment of changes over spatially focused study area where cell-to-cell changes provide relevant information, but are less useful for large-scale (continental or global scale) assessment where cell-level details are not pertinent.

We have developed a land cover change method based on comparison of land cover patterns rather then land cover classes. The method is region-based instead of cell-based. Pattern-based analysis addresses the question of structural, and thus semantic similarity between landscapes. Unlike a single cell, a pattern has rich enough content to have functional significance for a system. Thus, pattern-based approach to analyzing land cover dynamics over large spatial extent seems more natural than the cell-based approach. The method relies on multi-date, post-classification comparison. In this paper we compare the National Land Cover Datasets 2001 and 2006 (NLCD 2001 and NLCD 2006). These datasets cover the entire conterminous U.S. with the resolution of 30 m/cell. Each cell is labeled by one of $K = 16$ nominal land cover labels. The result of our assessment is a dissimilarity map that shows a degree of landscape change over the entire U.S. This map is available for browsing at our GeoWeb application DataEye-USA (http://sil.uc.edu/). High values of dissimilarity indicate a bona fide change in land cover class whereas small (but significant) values of dissimilarity indicate some changes in landscape pattern without change to its semantic meaning. Interestingly, extensive areas in the southern U.S. (as well ares in some other regions) exhibit high number of 2001/2006 cell-label changes but no semantic changes in landscape.

## 2. METHODS

For pattern-based approach to the land cover change assessment we utilize local raster tiles. A tile $\mathcal{A}$ is defined as a square-shaped subset of the NLCD having the size $n \times n$ cells; the size of the tile determines the scale over which the degree of change is assessed. In this paper we use tiles with $n = 500$, thus, a land cover pattern and its change are determined on the scale of 15 km. To cover the entire US, 1,684,540 tiles are arranged in a grid with their centers separated by $k = 100$ cells allowing for ample overlapping between neighboring tiles. Change is determined by comparing 2001 vs. 2006 patterns in each local tile. The results of this comparison is stored in the 3 km/cell raster, which, upon visualization, yields the US-wide map of change.

In our approach land cover change is assessed at the level of a tile. A change has occurred, if the spatial pattern of land cover classes within a tile has changed from one motif to another. Note that even extensive cell-to-cell class label changes may not necessarily lead to a change in the pattern motif. The level of change is assessed as a degree of dissimilarity between pattern motifs at two different time steps. Calculating an appropriate dissimilarity value between two different pattern motifs is at the core of our methodology. Our method is based on concepts developed in the context of Content-Based Image Retrieval (CBIR) domain. We have originally introduced [2] this method for query and retrieval of similar land cover patterns in a single-date dataset (NLCD 2006); here it was modified for the purpose of change detection. The method has two components, pattern signature and pattern dissimilarity. Pattern signature is a compact mathematical description of a pattern, and pattern dissimilarity is a function that assigns a numerical value to a pair of patterns on

*The author performed this work while at the University of Cincinnati.
†This work was supported by the National Science Fundation under grant BCS-1147702.


**Fig. 1**. Maps of land cover change. (A) Spatial distribution of density of cells that changed land cover label from 2001 to 2006. Six local regions are highlighted for future discussion. Spatial distribution of Jensen-Shannon divergence: (B) local patterns are characterized by 1D histograms of land cover classes; (C) local patterns are characterized by 2D histograms of land cover classes and clump sizes. Spatial resolution of all maps is 3 km/cell, see main text for details.

the basis of their respective signatures.

For pattern signature $A(\mathcal{A})$ we use a class/clump-size histogram constructed from the cells in the tile. Such signature has advantage of being rotationally and translationally invariant. A simplest signature is a class histogram of the tile. However, such signature accounts only for the overall bulk composition of classes in the tile but not for their spatial arrangement. In order to incorporate some spatial information into our signature we segment the tile into clumps (four-connected region of a single class) using a standard connected components algorithm [3] and calculate a size of each clump in terms of a number of individual cells within it. Clump sizes are numerical data that upon quantization constitute the second component of our class/clump-size histogram. We quantizes clump sizes by assigning them to bins with ranges based on the powers of two (i.e. 1-2, 2-4, 4-8 etc). Each cell inherits its clump-size class from a clump to which it belongs. Because histogram $A(\mathcal{A})$ is normalized to unity it can be thought of as a probability density function (pdf) of a random variable (land cover class, clump-size class).

We use the Jensen-Shannon divergence [4] to calculate dissimilarity between two histograms. We have chosen the Jensen-Shannon divergence because of it robustness and good performance in the side-by-side comparison with other measures (reference). For two histograms $A$ and $B$ the Jensen-Shannon divergence (JSD) measures the deviation between the Shannon entropy of the mixture of the two histograms $(A + B)/2$ and the mean of their individual entropies, and is given by

$$\mathrm{JSD}(A, B) = H\left(\frac{A+B}{2}\right) - \frac{1}{2}\left[H(A) + H(B)\right] \qquad (1)$$

where $H(A)$ indicates a value of the Shannon entropy of the histogram $A$. JSD is always defined, symmetric, bounded by 0 and 1, and equal to 0 only if $A = B$. The value of $H(A)$ reflects a distributional character of histogram $A$, the large value of $H(A)$ indicates $A$ evenly spread between the bins, whereas the small value of $H(A)$ indicates $A$ concentrated in just few bins. JSD measures (in a single number) a difference between distributional characters of $A$ and $B$. Note that if the two tiles, $\mathcal{A}$ and $\mathcal{B}$, have similar histograms, $A(\mathcal{A})$ and $B(\mathcal{B})$, the histogram of their mixture, $(A+B)/2$, is similar to each of the two individual histograms and the value of JSD is small. If the two tiles have dissimilar histograms, the histogram of the mixture is more spread than each of the two original histograms and the value of JSD is large. A maximum difference, JSD=1, is assigned for two histograms where each is having only a singe but different bin (two tiles each having a single but different land cover class).

## 3. RESULTS

Only 1.68% of all NLCD cells are mapped as changed between 2001 and 2006. We first derive a map of density of changed cells. For each tile we calculate a variable $\rho$ = (number of cells that had changed the land cover class label)/Max (number of changed cells), where the Max $\approx$ 200,000 is the maximum over all of the tiles. The variable $\rho$ has the range between 0 and 1. The density map, shown in Fig. 1A, is unlikely to reflect well the true nature of 2001/2006 change in land cover. High values of $\rho$ may or may not indicate a fundamental change in land cover pattern. Fig. 1A suggests that extensive areas in the southern U.S. underwent a change in land cover pattern in just 5 years.

In order to check this suggestion we calculate pattern-based map of change using only class histograms as pattern signature. The results are shown in Fig. 1B. Note that despite having identically looking legends the two maps show different variables, $\rho$ for Fig. 1A and JSD for Figs. 1B and 1C. The map in Fig. 1B indicates that vast majority of tiles across the U.S. experienced very limited change in the bulk composition of land cover classes. Significant change in composition is restricted to isolated regions located mostly in Nevada, Utah, and Colorado. Most strikingly, the area in the southern U.S., that has show enhanced density of changed cells, is characterized by limited change in the bulk class composition. The land cover change in this region appears to have a fundamen-

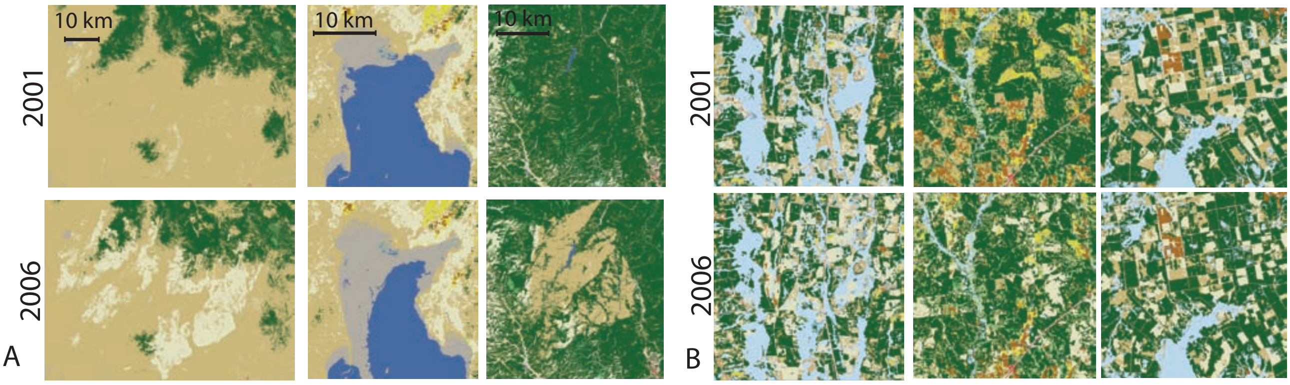


**Fig. 2**. (A) Highlighted regions that underwent change in landscape pattern (from left to right: Nevada, Utah, and Colorado). (B) Highlighted tiles (15 km $\times$ 15 km) characterized by high density of label-changed cells but small change in land cover pattern. (from left to right: Florida, Alabama, North Carolina).

tally different character from that observed in the isolated regions listed above. Rather than an outright change to a different land cover class, the region appears to maintain landscape pattern despite numerous changes to individual cells.

To further analyze changes in landscape patterns we calculate pattern-based map of change using 2D histograms based on the class/clump size signature. Such measure compares motifs of land cover classes including their spatial characteristics and is better suited to reflect a true nature of landscape change than density or bulk composition. The results are shown in Fig. 1C. Interestingly, this map combines the most interesting features of the maps shown in Fig. 1A and Fig. 1B. It shows (in red) an isolated regions with relatively high values of JSD, but it also shows (in light blue) regions with smaller but not negligible values of JSD. Recall that JSD measures divergence of pattern motifs - values of JSD $\geq$ 0.4 indicate significant disparity between two histograms and thus the landscapes, whereas values of JSD $\sim$ 0.1 indicate overall similar landscapes with some changes. Fig. 2A shows three examples of regions which have experiences bona fide change in land cover pattern and are characterized by high values of JSD. Fig. 2B. shows three examples of regions characterized by high density of label-changed cells but small changes in landscape patterns (moderate values of JSD).

## 4. CONCLUSIONS

Presented method is designed for large-scale assessment of land cover change. It is based on comparison of pattern of land cover (that could be succinctly referred to as comparison of landscapes) rather than on cell-by-cell comparison of land cover labels.

Cell-by-cell assessment of land cover change over large-scale regions yields matrix of transformation that shows information on "from class" to "to class" statistics. For the NLCD 2001/2006 datasets the transformation matrix is given in [5]. It provides a useful information but leaves a lot of questions unanswered. For example, the matrix shows that approximately the same number of cells have lost their 2001 "woody wetlands" label as gained it in 2006, resulting in an insignificantly small overall loss of woody wetland designation. However, the matrix does not provide any spatial information. Is the gain and loss of woody wetlands occur in geographically different areas, or more likely, in the same areas as the result of slight re-arrangements of landscape without change to its function and meaning? Similar statistics holds for land cover classes of "herbaceous wetland", "cultivated crops", "barren land", and "open water." Our pattern-based method can answer such questions. A dissimilarity clearly identifies areas where bona fide change in landscape has occurred, and distinguish them from areas where landscape remains unchanged despite large number of label changes at an individual cell level. Of course, the result of our assessment depends on the particular choice of scale (the size of a tile). Choosing a smaller tiles will result in finding more localities characterized by true landscape change, whereas choosing a larger tiles will result in finding less such localities. As the method is explicitly based on comparison of regions, dependence on scale is build-in.

Our comparison of 2001 and 2006 maps of land cover over conterminous U.S. revels (see Fig. 1C) that very few locations underwent significant landscape change on the scale of 15 km. The three of the most prominent such locations are shown in Fig. 2A. It also reveals an existence of many locations (shown in light blue on Fig. 1C) that underwent land cover change on the spatial scale $<<$ 15 km, but maintained the overall character of their landscapes on the scale of 15 km. The three regions with high concentration of such locations are Southern U.S., Pacific Northwest, and Northern Maine.

## 5. REFERENCES


[1] D. Lu, E. Mausel, E. Brondizio, and E. Moran, "Change detection techniques," *Int. J. Remote Sensing*, vol. 25(12), pp. 23652407, 2009.

[2] J. Jasiewicz and T. F. Stepinski, "Example-Based Retrieval of Alike Land-Cover Scenes From NLCD2006 Database," *Geoscience and Remote Sensing Letters*, vol. 10, pp. 155–159, 2013.

[3] A. Rosenfeld and J. L. Pfaltz, "Sequential operations in digital processing," *J. ACM*, vol. 13, pp. 471–494, 1966.

[4] J. Lin, "Divergence measures based on the Shannon entropy," *IEEE Transactions on Information Theory*, vol. 31(1), pp. 145–151, 1991.

[5] J. A. Fry, G. Xian, S. Jin, J.A. Dewitz, C. G. Homer, L. Yang, C. A. Barnes, N. D. Herold, and J. D. Wickham, "Completion of the 2006 National Land Cover Database for the Conterminous United States," *Photogrammetric Engineering & Remote Sensing*, vol. 77(9), pp. 858–864, 2011.