

# Unsupervised Delineation of Urban Structure Types Using High Resolution RGB Imagery

J. Niesterowicz<sup>1</sup>, T. F. Stepinski<sup>1</sup>, J. Jasiewicz<sup>1,2</sup>

<sup>1</sup>Space Infomatics Lab, University of Cincinnati, 401 Braunstein Hall, Cincinnati, OH 45221-0131, US  
Email: {niestejk; stepintz}@mail.uc.edu

<sup>2</sup>Institute of Geocology and Geoinformation, Adam Mickiewicz University, Dziegiełowa 27, Poznan, Poland  
Email: jarekj@amu.edu.pl

## Abstract

We present a method for delineating Urban Structure Types (USTs) using only high resolution RGB images. As the method is unsupervised, it does not require training; the interpretations of delineated USTs are assigned a posteriori. The method utilizes freely available software and performs delineation in a short time even for very large images. A 1-meter resolution image of the entirety of Los Angeles is delineated as an example. We have found seven distinct USTs which were given interpretations based on examination of their patterns. These interpretations are validated by population statistics. The method aims at broaden the usage of USTs delineations for applications in urban and social studies.

## 1. Introduction

Urban Structure Type (UST) is a distinct spatial pattern of the urban structure at the neighborhood scale, which can be interpreted in terms of the type of activity or of residential pattern. Classification of a city into USTs complements standard land cover/land use classification by working at a scale that is significantly coarser than an individual pixel. Fairly extensive literature exists on how to delineate USTs from remotely sensed data (for example, see Heiden *et al.* 2012), but, because these works focus on supporting effective urban planning, they use multisource data and supervised learning. This means that they are restricted to very few places where this data exists and where the significant cost of supervised analysis is justified by the need. There also exists extensive literature on using single-source data (RGB or multispectral images) but only in the context of separating two specific types of USTs – formal from informal settlements (slums); for example see Graesser *et al.* (2012). Algorithms proposed there are restricted to this single purpose; they also are predominantly based on supervised learning.

In this paper we present an approach to delineation of USTs that uses only RGB images (many of which are freely available online) as input, delineates an exhaustive set of USTs, is based on training-free, data-driven unsupervised principles, and can process very large input data in a reasonable time. In addition, our method relies only on existing public domain software. Our motivation is to make the delineation of USTs more broadly accessible to analysts from different disciplines. The methodology is described and applied to a ~2 billion pixel 1 m-resolution image covering the greater Los Angeles area.

## 2. Methodology

Our method is based on the concept of Complex Object-Based Image Analysis (COBIA) (Vatsavai 2013; Stepinski *et al.* 2015). In COBIA a raster (not necessarily restricted to an image) is divided arbitrarily into a grid of local blocks of cells. We refer to these blocks as motifs –

they encapsulate a local pattern (motif) of raster variable. Motifels are much larger than pixels and they are also much more complex; motifels are used as elementary units for further analysis. COBIA is a well-suited approach to the problem of delineation of USTs from an RGB image. Individual pixels in the RGB image may not have an unambiguous interpretation, but we expect that the composition and arrangement of different colors within a motifel can be associated unambiguously with the specific UST. We utilize the GeoPAT toolbox (Jasiewicz *et al.* 2015) – an open source collection of GRASS-GIS modules for pattern-based geoprocessing – to implement COBIA. GeoPAT works with categorical rasters, thus, as the preprocessing step, we first quantize an RGB image to obtain a raster of class color labels.

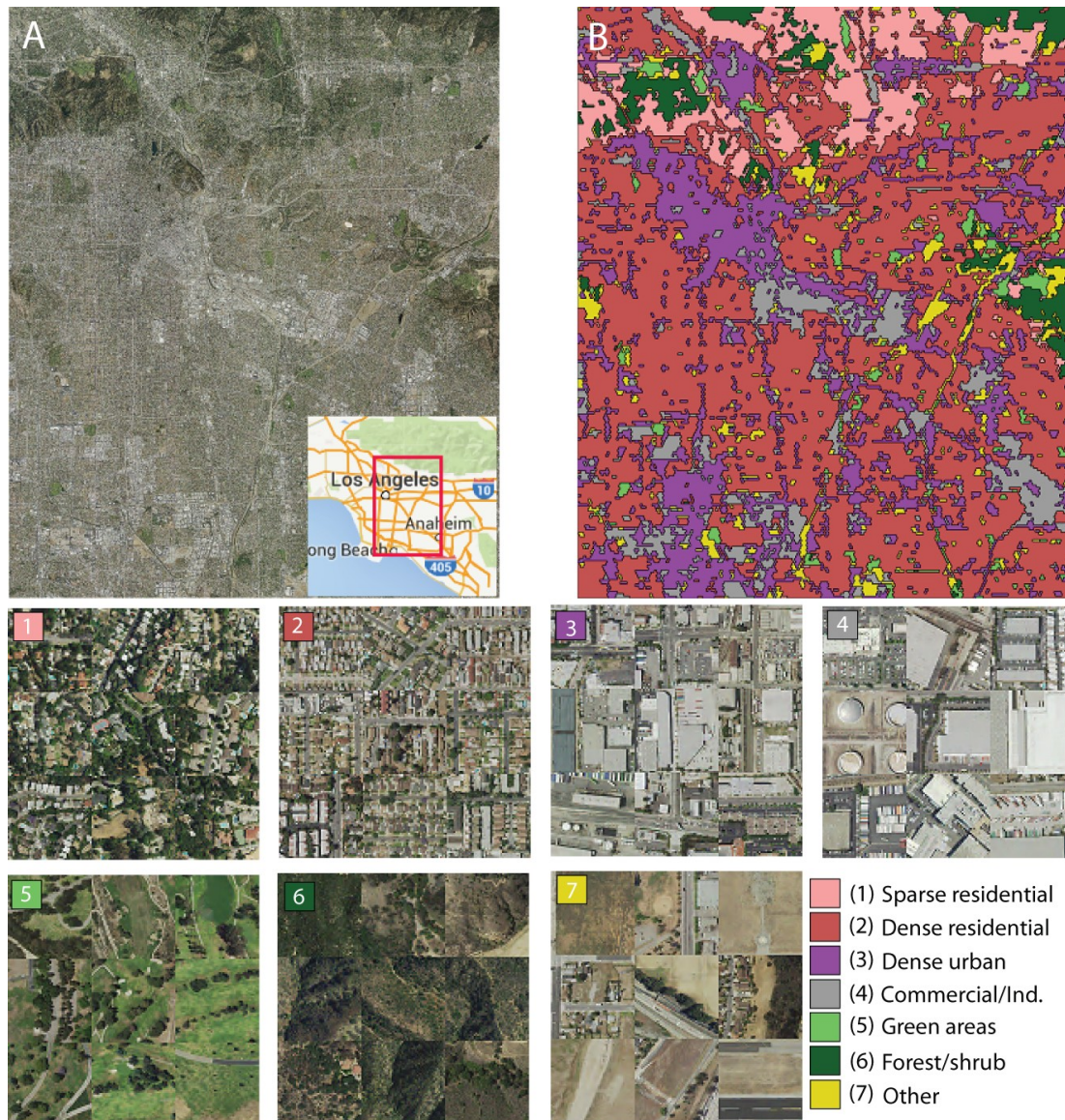
We encapsulate the complexity of the motifel structure by a color co-occurrence histogram and calculate the degree of dissimilarity between two motifels using the Jensen-Shannon Divergence (JSD) (Lin, (1991)) between their corresponding histograms. A grid of motifels having dimensions much smaller than an original image, is first segmented into local areas of homogeneous patterns. Segmentation provides enhanced spatial cohesion to the final UST classification and reduces the dimensionality of the subsequent clustering step. As spatially distinct segments may have very similar patterns, in the final step we cluster the segments into a small number of USTs. Both segmentation and clustering steps utilize JSD as the measure of distance. We use a segmentation algorithm which is custom-designed for grids of motifels. This algorithm is a part of the GeoPAT toolbox and is described in Jasiewicz *et al.* (2016) (see also Jasiewicz *et al.* contribution to this conference); its only parameter is the motifel's size. We use R implementation of the Partitioning Around Medoids (PAM) algorithm to cluster the segments.

### 3. Delineating USTs in the Los Angeles area

To demonstrate our methodology we delineated USTs in the greater Los Angeles area (see inset in panel A of Fig.1) using 1-meter resolution RGB (3-band) aerial imagery freely distributed by the U.S. Geological Survey (<http://viewer.nationalmap.gov/launch>). We downloaded 36 individual images (available in jpeg2000 format) from USGS and mosaicked them into a single image (shown in Fig.1 panel A) having dimensions of 41600×50200 pixels. This image was quantized into 27 color class labels to prepare it for GeoPAT processing. The quantization procedure first converts RGB color into the CIE L\*a\*b\* color space, in which the numerical distance between two colors corresponds to the perceptual distance between them. Then, the CIE L\*a\*b\* color space is divided into cubic blocks of a selected size. Non-empty blocks are the quantized colors; 27 colors were obtained with the size of the block equal to 25 units. Next, we have chosen the size of the motifel to be 200 pixels (meters) which roughly corresponds to the size of a city block. With such a choice the image was transformed into a grid of 208×251 motifels. The segmentation step resulted in 4841 segments which were clustered into 7 USTs. The number of USTs was determined experimentally so the homogeneities of patterns in USTs are maximized while the number of USTs is kept to the minimum.

Because our method is unsupervised, an interpretation of each UST must be made a posteriori. To arrive at an interpretation we constructed a synthetic image for each UST. A synthetic image of an UST consists of 400 motifels selected randomly from its extent and organized into 20×20 array. Smaller versions of synthetic images for all USTs are shown in the last two rows in Fig.1. Based on our examination of these synthetic images we gave the USTs interpretations as listed in the legend in Fig.1. Two of the USTs (labeled as 5 – green areas and 6 – forest/shrub) are interpreted as undeveloped and uninhabited areas. Two others (labeled as 1 – sparse residential and 2 – dense residential) are interpreted as residential areas with detached

housing. Two UST classes (labeled as 3 – dense urban and 4 – commercial/ind) are characterized by their high percentage of impervious surfaces with class 3 appearing to be a mixture of multi-level apartments, shopping centers, and urban infrastructure; class 4 appears to contain purely commercial structures. The seventh UST class (labeled as 7 – others) consists of large construction areas, barren land, and partially, of sparsely populated barren land.



**Figure 1: Unsupervised delineation of USTs in the Los Angeles area. (A) Original 1-meter resolution image (inset shows the location of the image). (B) Map of seven USTs found in this area. The two lowest rows show a series of seven synthetic images each consisting of 9 randomly selected motifs from each UST. The legend to USTs is given in the lower right corner.**

In order to validate our interpretations we used the newly available online resource SocScape (available at [http://sil.uc.edu/webapps/socscape\\_usa/](http://sil.uc.edu/webapps/socscape_usa/)) which provides high resolution (30-meters) gridded population data for the entire U.S. Unlike census data, which gives population counts aggregated to areal units, SocScape data distinguishes between inhabited and uninhabited

areas and can be used to calculate population density and the percentage of inhabited area in each UST class. The results are given in Table 1 and confirm our interpretations.

**Table 1. Population statistics for the UST classes.**

	1	2	3	4	5	6	7
Population density (people per ha)	20.0	50.2	40.3	8.9	2.9	0.6	10.8
% inhabited area	82.4	86.3	55.2	22.3	10.3	14.2	35.4

## 4. Conclusions

We presented a method for fast, unsupervised delineation of USTs from high resolution RGB images. We also demonstrated how this method works using an image of the Los Angeles area as an example. For this large image the processing time on the 8-core I7 computer was ~30 min, of which ~20 min was color quantization preprocessing. The results indicate that the method gives reasonable and valuable results using only an easy-to-obtain RGB image and our software which is in the public domain. The method works even if an image is in a compressed jpeg2000 format. This is in contrast to other works on delineation of USTs which either use multisource, difficult to obtain data or concentrate on the extraction of a single UST (slums), in both cases using proprietary software. We noticed that some USTs are more difficult to distinguish from an RGB image than others when using our method. In particular, dense urban environments mix with some (but not all) infrastructure and light commercial environment (class 3). Increasing the number of clusters does not help to resolve this problem because these environments are indeed characterized by similar patterns in an image. As 1-meter RGB images are available from the USGS for the entire U.S., our method can be used for comparative studies among major U.S. metropolitan areas.

## Acknowledgements

This work was supported by the University of Cincinnati Space Exploration Institute, and by Grant NNX15AJ47G from NASA.

## References

- Graesser, J., Cheriyyadat, A., Vatsavai, R.R., Chandola, V., Long, J., Bright, E., 2012. Image Based Characterization of Formal and Informal Neighborhoods in an Urban Landscape. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(4), pp. 1164–1176
- Heiden, U., Heldens, W., Roessner, S., Segl, K., Esch, T., Mueller, A., 2012. Urban structure type characterization using hyperspectral remote sensing and height information. *Landscape and Urban Planning*. 105, pp. 361–375
- Jasiewicz, J., Netzel, P. & Stepinski, T., 2015. GeoPAT: A toolbox for pattern-based information retrieval from large geospatial databases. *Computers & Geosciences*, 80, pp.62–73.
- Lin, J. 1991, Divergence measures based on the Shannon entropy. *IEEE Transactions on Information Theory*, 37(1):145–151.
- Stepinski, T.F., Niesterowicz, J. & Jasiewicz, J., 2015. Pattern-based Regionalization of Large Geospatial Datasets Using Complex Object-based Image Analysis. *Procedia Computer Science*, 51, pp.2168–2177.
- Vatsavai, R.R., 2013. Object based image classification: state of the art and computational challenges. In *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data*. pp. 73–80.