2011 Progress Report (UMD Part)

Using Landsat Global Land Survey Data to Measure and Monitor Worldwide Urbanization

Chengquan Huang
Department of Geography
University of Maryland

November, 2011

The "Using Landsat Global Land Survey Data to Measure and Monitor Worldwide Urbanization" project, also called Global Urban Mapping (GUM) project, is a joint effort between NASA/Goddard Space Flight Center (GSFC) and the University of Maryland (UMD). Led by PI Brown de Colstoun at GSFC, this project aims at developing global, 30-m resolution percent impervious data sets for circa 2000 and 2010, and using these data sets to identify urbanization "hot-spots".

The UMD team will focus on the following three aspects of this project: 1) produce surface reflectance products using the GLS 2010 data set and validate those products using MODIS data; 2) assist in deriving reference impervious data sets for training and validation purposes; 3) assist other project teams in modeling and mapping global urban areas. The official start date of the UMD part is 4/15/2011. This report summarizes the progress the UMD team has made since then.

1 Staffing

The UMD part of the GUM project is conducted within the Global Land Cover Facility (GLCF), and will be supported by GLCF members as necessary. In particular, through a MEASURES project titled "Global Forest Cover Change (GFCC)", GLCF has developed the expertise, computer systems, and necessary algorithms for processing global Landsat data sets, which will be highly valuable for achieving many of the goals of the GUM project. In addition, highly automated algorithms will be required for developing large quantities of reference data sets necessary for modeling urban imperviousness at the global scale. We have recruited Guangxiao Zhang, a PhD student from the Computer Science Department to support this project. Ms Zhang will develop and examine new machine learning and computer vision algorithms that can be used to greatly improve the efficiency of retrieving urban features using Landsat and high resolution images (see section 3 for more details).

2 Data and System

GLCF currently has a computer system capable of processing global Landsat data sets, including the Global Land Survey 1975, 1990, 2000, and 2005 data sets. The Landsat surface reflectance data set for 2000, which is required by the GUM project, has already been generated, and has been validated using near simultaneous MODIS observations (Figure 1). With a moderate upgrade to the CPU and storage, this system can be used to process the GLS 2010 data set, which is also required by the GUM project. The GLS 2010 data set is still being produced, although images for certain areas have become available through the GLOVIS website. We have been in close contact with Rachel Headley at USGS and Jeffrey Masek at NASA/GSFC to check on a regularly basis the progress of this data set. Once the production of this data set is complete for one or more continents, we will start to download the images and process them for use in this project.

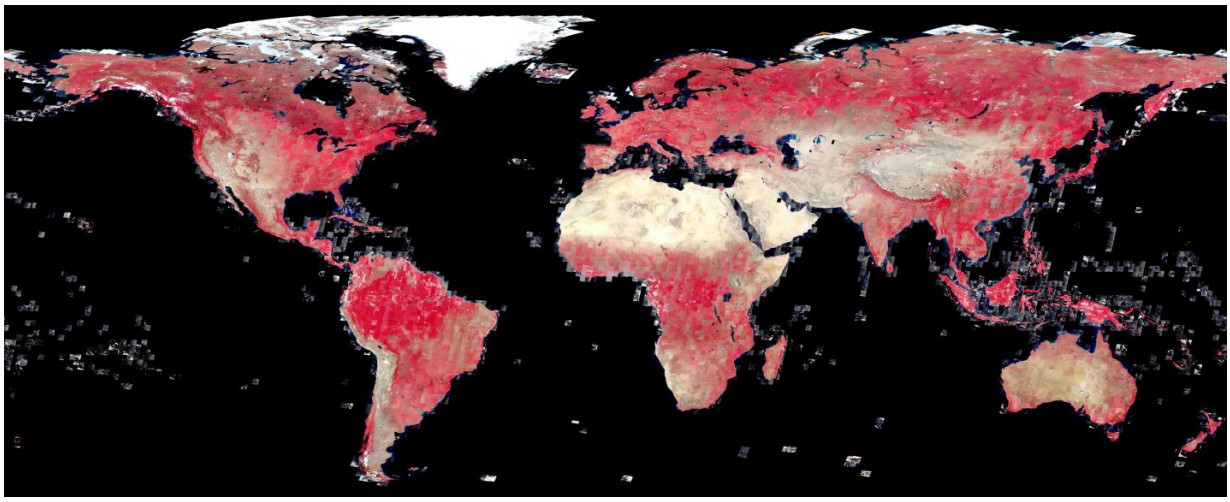

Figure 1. A mosaic of global Landsat surface reflectance images derived using the Global Land Survey 2000 data set.

In addition, GLCF current hosts a large collection of high resolution images (Figure 2), most of which were acquired through NASA's Science Data Purchase program. Many of these images are located within or near urban areas and can be used to derive reference data on percent imperviousness. The data collection will complement a much larger collection of high resolution images available at the NGA, many of which will be used in this project to derive reference data sets on percent imperviousness.

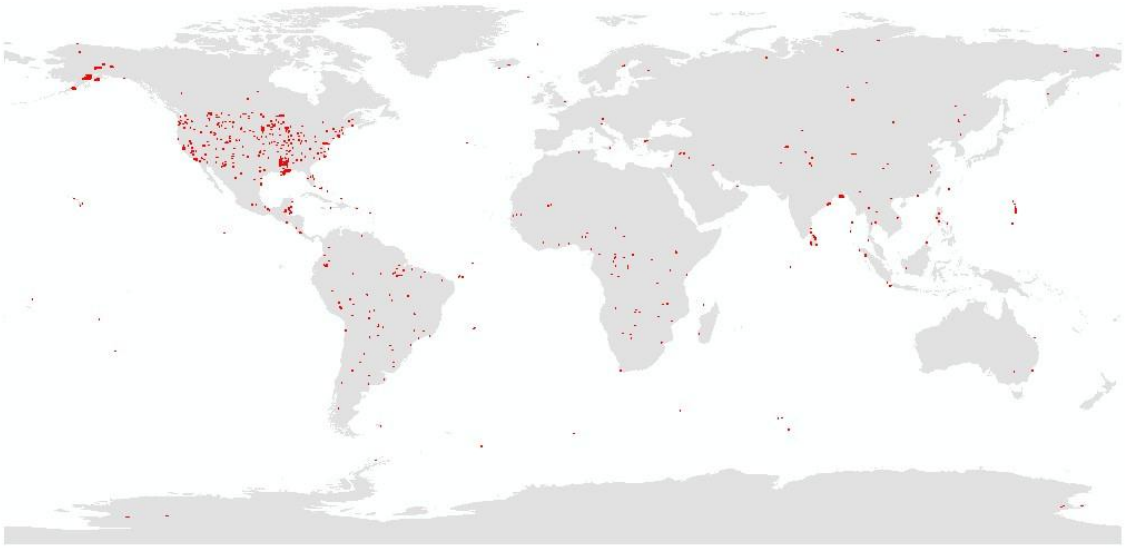

Figure 2. Distribution of over 6000 high resolution satellite images currently available at GLCF, many of which will be used to derive reference data sets on percent imperviousness.

3 Algorithm Development

Among the various land cover classes, urban is unique in that many urban areas consist of a mixture of impervious surfaces, urban tree, grass, and other non-impervious surfaces. As a result, classification of urban is often difficult on a per-pixel basis. This is because

that except for some of the high intensity urban areas, most urban areas are mixtures of impervious and non-impervious surfaces. Therefore, spatial texture and contextual information is critical for urban characterization. We have explored the use of conventional and new texture measures, as well as non-texture measures for identifying urban areas at both the Landsat and high spatial resolutions. The conventional texture measures include gray-level co-occurrence metrics (GLCM). The new textures we are exploring are a family of measures derived based on binary local patterns (BLP), which has been used extensively in computer vision algorithms. Other measures we have examined include color vector and local variance.

Our hypothesis is that relationships between percent imperviousness and spectral data likely will vary substantially from one Landsat image to another because the spectral value of Landsat images are affected by many factors that may be different from one image to another. Therefore, it would be difficult to apply the model derived based on one Landsat image to another. Textures, however, are less sensitive to among-image spectral differences because they are measures of relative differences among pixels within a local neighborhood. Therefore, models derived using texture measures over one image may have a better chance to make good predictions when applied to new images than models derived using spectral values only. We have conducted a series of experiments to test this hypothesis using test data selected from the following 5 sites:

Site 1: p014r032, city: NYC and Philly
Site 2: p015r033, city: DC and Baltimore
Site 3: p041r037 and P042r035, city: LA
Site 4: p024r033, city: St. Louis
Site 5: p037r037, city: Pheonix

In each experiment, samples collected over 4 sites were used to train a classification model, which was then used to classify the samples from the 5^{th} site. Figure 3 shows the accuracies derived using different texture measures and using each of the 5 sites as the test site. It demonstrates that accuracies of over 90% can be achieved when a classification model derived using training data collected over some Landsat images is used to classify an image not used in model development. We will examine more test sites to determine if these results hold true over a broader range of test sites.

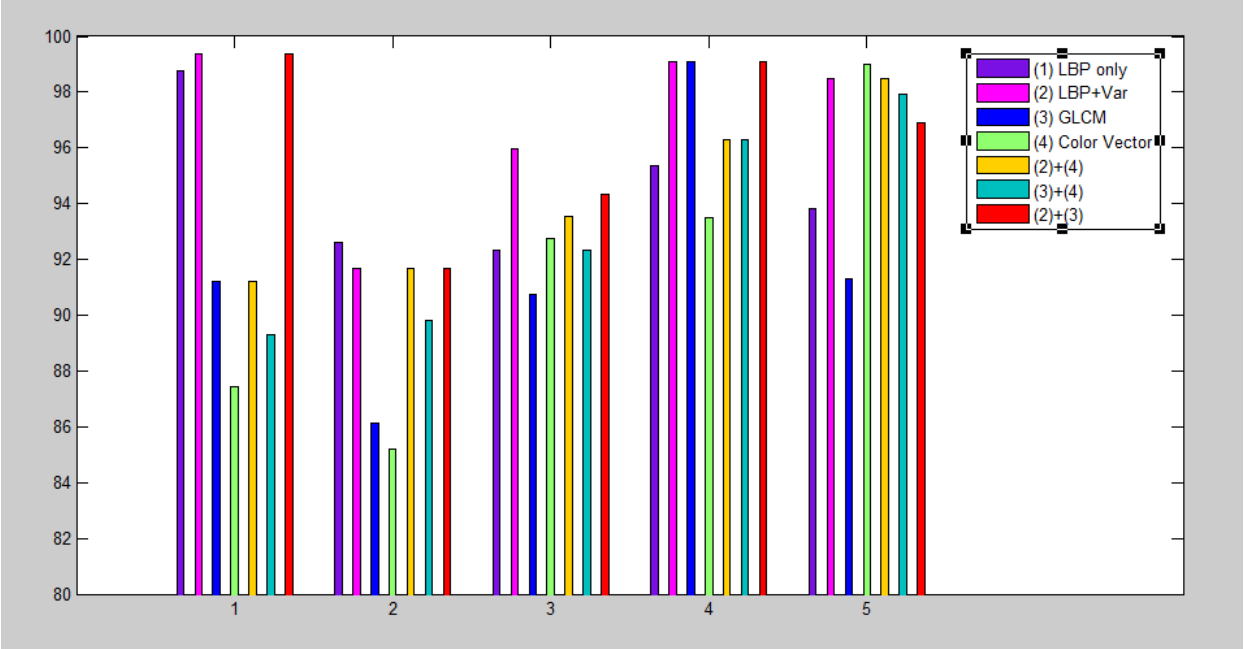

Figure 3. Overall accuracy of urban classifications derived over 5 test sites using texture features. In each experiment, one site is reserved to test the classification model derive using samples collected over the other 4 sites. See text for description of the 5 sites.

We have also started to look into a new segmentation approach called superpixel segmentation to determine whether it can be used to improve urban feature extraction at Landsat and high spatial resolutions.

4 Publications

A poster titled “Fine Scale Assessment of Global Forest Cover Change and Urbanization using Landsat Data”, which provides a high level description of this project, was presented at the 2011 LCLUC workshop held on November 5-9, 2011, in Hanoi, Vietnam.

5 Other Activities

The UMD team participated in regular meetings with the GSFC team in developing project approaches and milestones. We have obtained the segmentation code developed by Co-I Tilton and have started to evaluate its usefulness for deriving impervious surface using Landsat and high resolution images. We have also obtained a copy of the regression tree software called Cubist, which will be used to model percent imperviousness at the Landsat resolution.

6 Plans for the Next Year

The UMD team will continue to work closely with the GSFC team to push the project forward. Major tasks for the UMD team for year 2 include:

- Obtain the GLS 2010 data set, produce surface reflectance and validate the entire data set using near simultaneous MODIS data. We anticipate that production of the GLS 2010 data set will be complete and the entire data set will be made available by mid-2012.
- Continue to develop and test the texture and segmentation algorithms described in section 3 for urban feature extraction.
- Start to derive substantial quantities of reference data using high resolution images and use them to model percent imperviousness over selected regions.