Preprint
UCRL-JC-150220

Detecting Human Settlements in Satellite Images

S. K. Sengupta, C. Kamath, D. Poland, and J. A. H. Futterman

This article was submitted to :
Optoelectronics at the Lawrence Livermore National Laboratory
Lasers and Applications in Science and Engineering
Photonics West Symposium
SPIE — The International Society for Optical Engineering
San Jose, CA, 25-31 January 2003
SPIE Paper 5001-13

U.S. Department of Energy

Lawrence Livermore National Laboratory

January 15, 2003

Detecting Human Settlements in Satellite Images

S. K. Sengupta, C. Kamath, D. Poland, and J.A.H. Futterman*
Lawrence Livermore National Laboratory

ABSTRACT

The automated production of maps of human settlement from recent satellite images is essential to detailed studies of urbanization, population movement, and the like. Commercial satellite imagery is becoming available with sufficient spectral and spatial resolution to apply computer vision techniques previously considered only for laboratory (high resolution, low noise) images. In this paper we attempt to extract human settlement from IKONOS 4-band and panchromatic images using spectral segmentation together with a form of generalized second-order statistics and detection of edges, corners, and other candidate human-made features in the imagery.

Keywords: land cover, land use, urban boundary, human settlement detection, remotely sensed imagery

1. INTRODUCTION

The automated production of maps of human settlement from recent satellite images can benefit any large-scale application in which the urban land-cover, land-use, or boundary is a consideration. For example, the Landscan[1] database contains estimates of global human population in a grid with cells of 30 arc-seconds (about 1 km) on a side. We seek to create regional or global maps whose resolution captures the boundaries of villages that may be considerably smaller.

Satellite images are now available with sufficient spectral and spatial resolution to make this feasible, in principle. In this paper, we attempt to extract areas containing artificial structures from IKONOS[2] images (4m Ground Sample Distance [GSD], 4-band multi-spectral [MS] and 1m GSD panchromatic [PAN]), with the aim of using these areas as surrogates for regions in which humans may be working or living. In this preliminary work, we will not address demographic issues, but will focus on image understanding.

Following work by Heikkonen and Varfis[3] and Zhang[4] we adopt a multi-stage approach. However, these authors used lower resolution imagery (Landsat TM, ERS-1 SAR, and SPOT Pan), which led them to adopt processing stages and methods specific to their imagery sources. Rather than work with pan-sharpened imagery, we use multi-spectral and panchromatic images of the same scene in different processing stages:

1. We use unsupervised or supervised classification to segment the multi-spectral image into *k* spectral classes.
2. We use second-order statistics derived from the pixel class label co-occurrence matrices to mask the scene into "mixed" and "unmixed" tiles, where the mixed tiles are those likely to contain human settlements.
3. Switching to the panchromatic image, we use the freely distributed SUSAN[5] code to detect edges and corners in the "mixed" tiles.

* Futterman1@llnl.gov, phone 1 925 423-4741, fax 1 925 422-9343, Lawrence Livermore National Laboratory, L-182, PO Box 808, Livermore, CA, USA 94551-0808

4. In the panchromatic image, we again use a tiling scheme to determine whether edges and corners occur with sufficient frequency within a small 20x20 tile and therefore likely to indicate built structure rather than spurious detections.
5. We use morphological techniques to grow regions containing a sufficient density (or high enough frequency) of valid edges and corners.
6. Finally, we find the boundaries of the regions and convert them to a form suitable for importation into a Geographic Information System (GIS).

We discuss these stages in more detail below.

2. MULTI-STAGE IMAGE PROCESSING

We use a novel approach to fuse both first and second order information in the MS imagery [Figure 1], which in the case of IKONOS has four channels (near-IR, red, green, and blue). First, we classify the MS image *pixel*s into *k* (usually five) classes, using the intensities in the MS channels as spectral features. This results in a *labeled* image where the pixel class labels are represented as colors [Figure 2]. For this particular application, the pixel class labels capture all the spectral information we need.

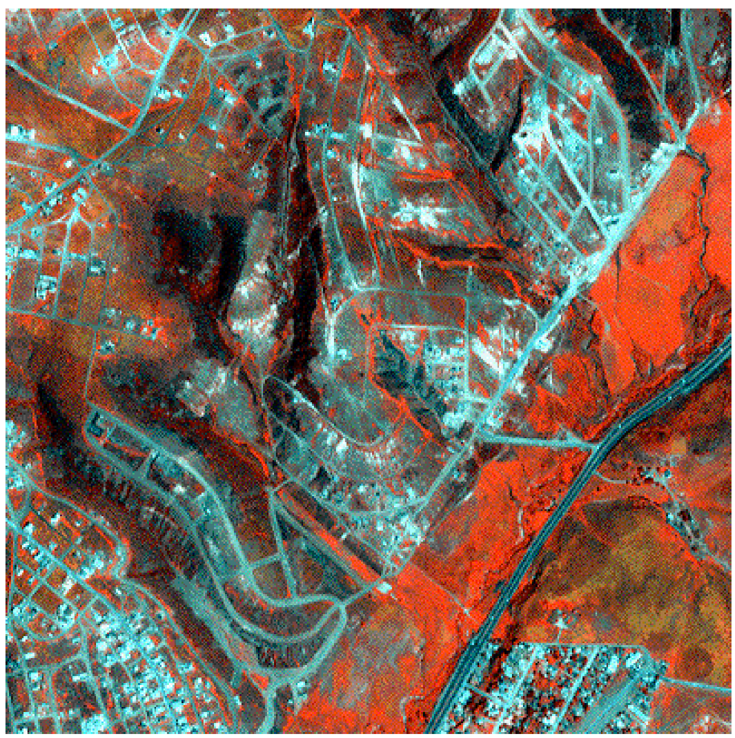

Figure 1. Portion of an IKONOS MS image near the California-Mexico border.
Satellite image by Space Imaging.

Visual inspection of many images suggested to us that, for 4m GSD, built-up areas tended to contain a spatially inhomogeneous mixture of class labels. Therefore, in the second stage, we introduce a kind of cross-spectral second order statistics to capture the spatial variation in spectral information of the neighborhood of each pixel. This is based on the spatial transition rate from one label to another in a single spatial step that occurs in an appropriate sized *tile* (a square sub-image) of pixels. Experience shows that with five classes, a minimum size of 10x10 pixels is necessary for a tile to provide a stable estimate of these rates within the tile. These class label transition rates form the elements of a generalization of the classic Gray-Level Co-occurrence Matrix (GLCM) that we call the *Class-Label Co-occurrence Matrix* (CLCM).

These rates of transition can be *directional* (horizontal, vertical, diagonal, cross-diagonal etc.) and can be made dependent on spatial steps of various sizes. We have considered isotropic and single-step transitions for our problem. The CLCM is of order kxk where k is the number of classes (or labels) defined in the first stage. When the rows are normalized to a sum of one, the CLCM is equivalent to the one-step transition probability matrix.

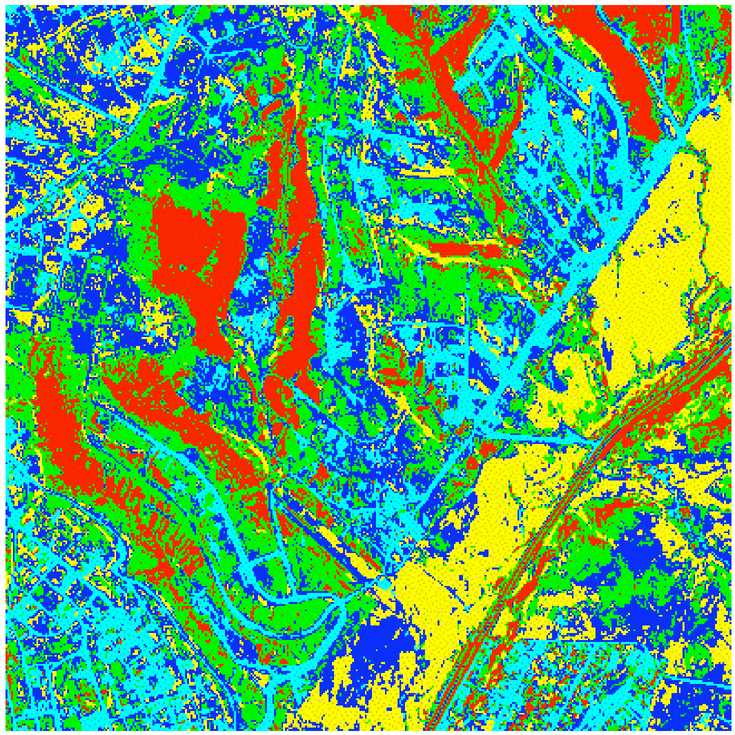

Figure 2: Five-fold unsupervised k-means classification of the image from Figure 1.
Original satellite image by Space Imaging.

We now use second-order statistical quantities derived from the CLCM to classify the tiles into two classes: mixed tiles (class 1 — likely to contain built structures) and relatively homogeneous tiles (class 2 — unlikely to contain built structures). This classification is based on what are essentially textural features derived from the CLCM in the same way that textural features are derived from the GLCM for single-channel images. Features that we found useful in this regard are the angular second moment (also known as energy), entropy and the diagonal entries of the CLCM. The first two are analogous to the definitions provided for the GLCM in the literature.[6,7] For classification purposes, the use of textural features based on the intensity in a *single channel* is abundant in the literature.[8,9,10,11,12]

We then mosaic-ed the tiles from class 1 into a mask (Mask 1), which we then use to demarcate areas of "sufficient" spectral mixture that they are likely to contain built structures [Figure 3]. Obviously, Mask 1 needs refinement before we can consider it to indicate areas containing built structure, as distinct from areas that are spectrally mixed for some other reason. However, it is sufficient to limit the expenditure of downstream computational resources to the areas within the mask. Just as obviously, setting the threshold higher or lower on our mixing criterion will yield a smaller or larger mask.

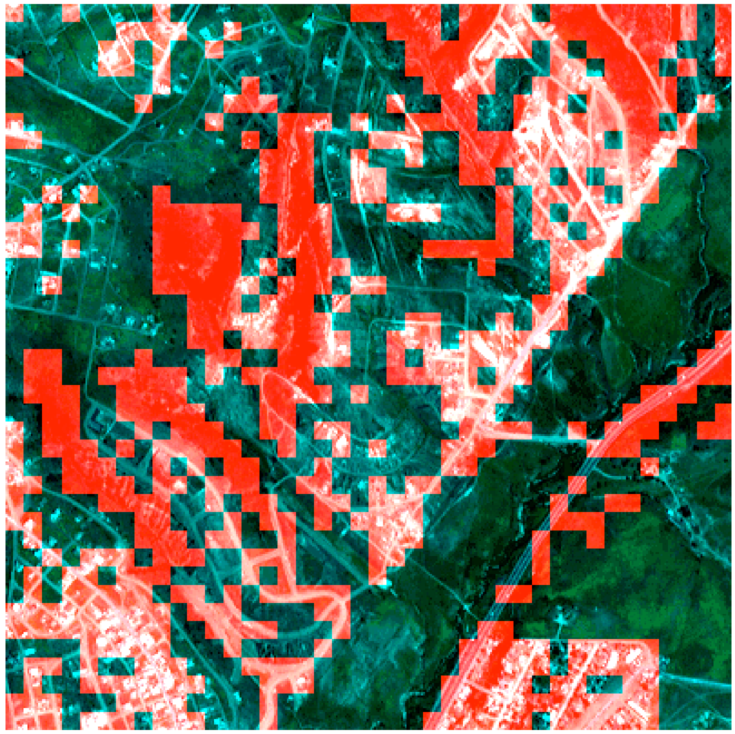

Figure 3. Spectrally mixed areas of Figure1 (Mask 1).
Satellite image by Space Imaging.

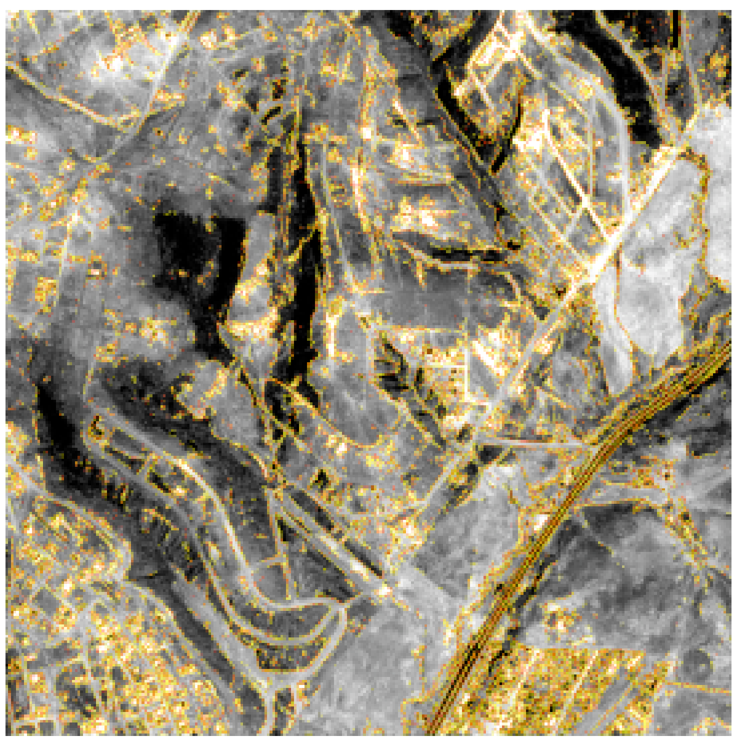

Figure 4. Edges (yellow) and corners (red) detected by SUSAN in same area of PAN image.
Satellite image by Space Imaging.

In the next stages of analysis we refine Mask 1 with panchromatic processing. We do this by transferring Mask 1 (multiplying pixel dimensions by 4 to transform from 4m to 1m GSD) to the IKONOS PAN image, and detecting concentrations of 'corners' and 'edge' pixels [Figure 4]. Specifically, we consider all 20x20 pixel tiles in the panchromatic image, each of which was classified based on the frequencies of the 'corners' and 'edge' pixels as determined by the SUSAN computer program to produce a Mask 2 [Figure 5]. Specifically, we determined a tile should contain several edge pixels and at least two corner pixels to be a candidate for containing built structures. For a 20x20 tile, the respective frequency thresholds were set at 20 for edge pixels and 2 for corner pixels.

SUSAN operates on each pixel in a gray-scale image by calculating brightness similarity values for neighboring pixels, where the neighborhood is defined by a 37 pixel pseudo-disk. Analysis of the centroids and first-order and second-order moments of these values, along with non-maximum suppression and thinning routines, allows for fast and effective smoothing and detection of edges and corners. We found that application of SUSAN smoothing to our images prior to edge and corner detection resulted in a fairly robust algorithm with respect to selection of the brightness threshold. Optimization of the mask size and the spatial threshold for this application, where the corners of interest are predominantly 90°, may improve performance but has not been explored.

We used morphological processing to “clean up” Mask 2 by merging small regions with neighboring large regions, and eliminating small isolated regions. The “cleaned” Mask 2 was then *resized* to match the resolution of the MS image and combined via a logical *AND* operation with Mask 1 to create a final Mask 3, which demarcated regions containing built structures [Figure6]. Finally, we performed a connected component analysis of Mask 3, followed by a morphological gradient operation, in order to find boundaries for export to a Geographic Information System (GIS).

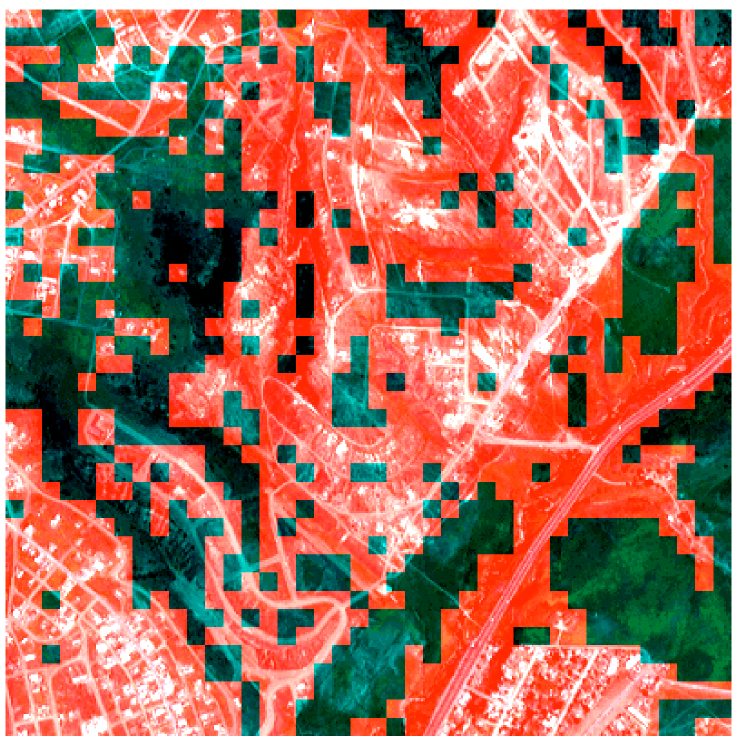

Figure 5. Mask2 of tiles containing edges and corners.
Satellite image by Space Imaging.

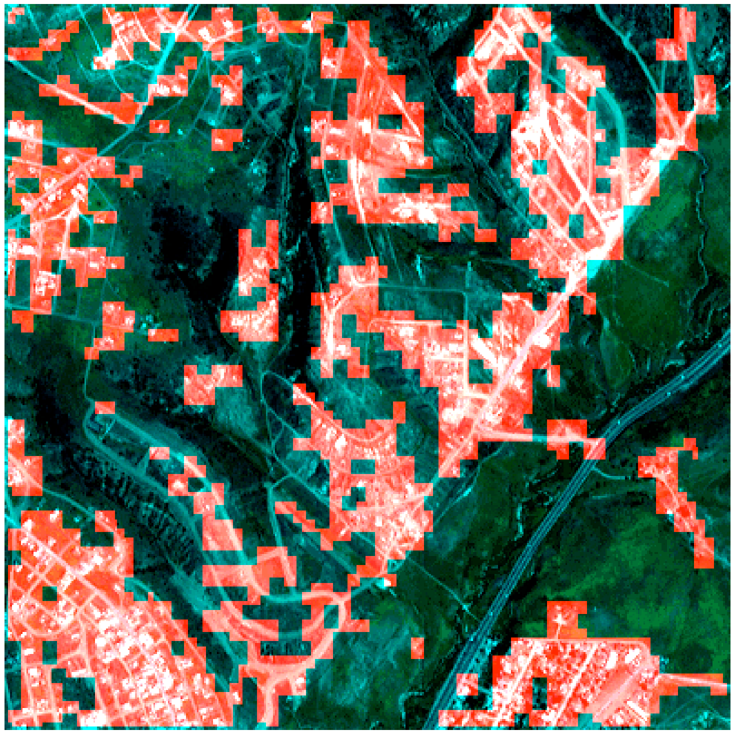

Figure 6. Final mask (Mask 3) containing spectral mixing, edges, and corners (indicating built structures).
Satellite image by Space Imaging.

3. ROBUSTNESS EXPERIMENTS

We also explored the applicability of our techniques to images from different regions of the world. In particular, we are interested in understanding how much of the analysis conducted in one region can be directly applied to another region. For example, can we use the classification of the pixels in the multi-spectral image of one region to build a model that can accurately classify the pixels in another region?

Figure 7: Image A (400x400 px) Nebraska
Satellite image by Space Imaging.

Figure 8: Image B (400 by 400 px) northern Mexico.
Satellite image by Space Imaging.

To investigate this idea, we used two sample images: Image A (Figure 7) is a 400 by 400 pixel image from Nebraska, while Image B (Figure 8) is a 400 by 400 pixel image from northern Mexico. First, we applied unsupervised techniques using the k-means clustering algorithm to the multi-spectral image in Figure 7. A visual inspection indicated that for this image, 6 classes resulted in the best clustering (Figure9). Using representative pixels from these 6 classes, we then created a training set of 4860 pixels, with approximately 800 pixels from each of the 6 classes. We used this training set to build a decision tree model. This model resulted in a ten-fold cross-validation error rate of less than 1%. Next, we used the decision tree to classify all the 160,000 pixels in Image A, resulting in Figure 10. Comparing the original Image A (Figure 7), the result of unsupervised classification (Figure 9), and the result of supervised classification (Figure 10), we observe that for this image, the decision tree model built using a sub-set of the pixels in the image generalizes quite well to the entire image. This indicates that the training set and the decision tree model are an accurate representation of the pixels in Image A.

Next, we applied the decision tree to the pixels in Image B, resulting in the classification shown in Figure 12. Comparing this with the original Image B (Figure 8), and the result of unsupervised classification using 6-class k-means algorithm on Image B (Figure 11), we observe that the decision tree model also works well for pixels in this region. Though the two images A and B are from two different regions of the world, they are similar enough that a model built to classify pixels in the multi-spectral image of one can be used successfully to classify pixels in the multi-spectral image of the other. If the two images were quite different, it is unlikely that such an experiment would have been equally successful. This indicates that it might be possible to build models that would be tuned to different regions of the world. However, depending on the region, it may be necessary to build separate models to account for seasonal variation.

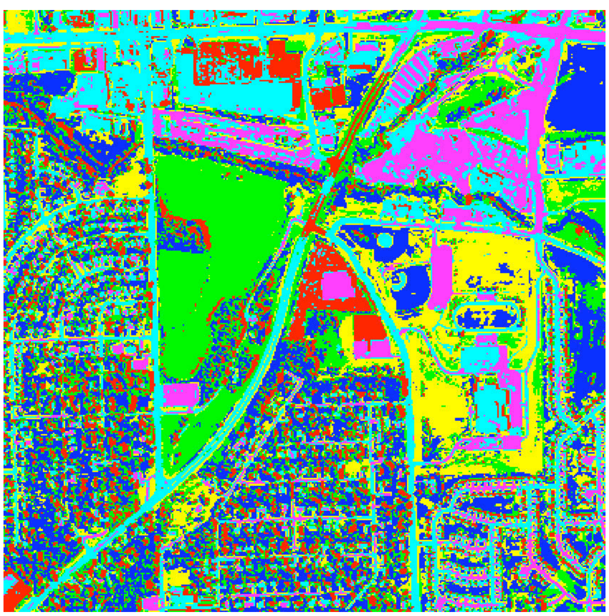

Figure 9: Image A classified using k-means algorithm, with 6 classes.
Original satellite image by Space Imaging.

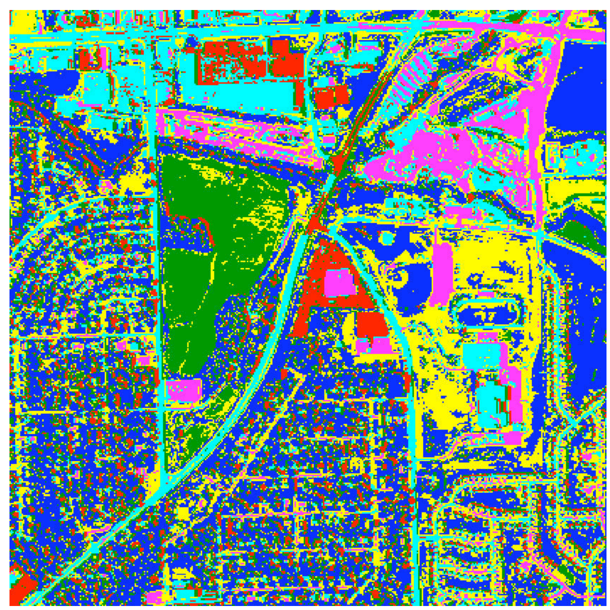

Figure 10: Image A classified using a decision tree algorithm using 4860 pixels from Figure 7
Original satellite image by Space Imaging.

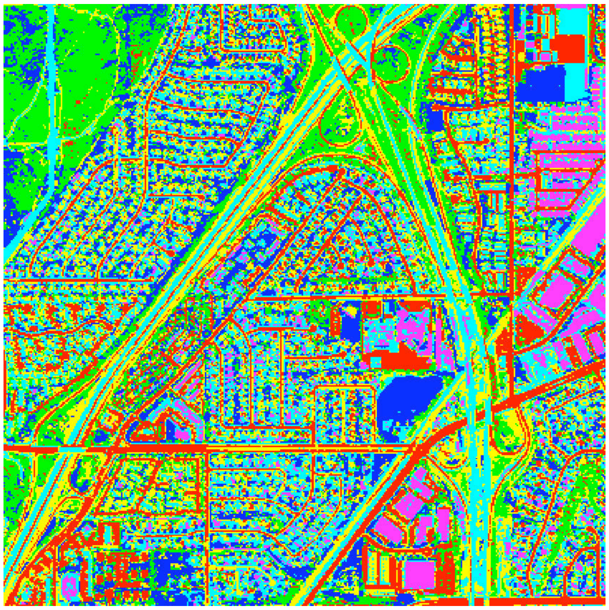

Figure 11: Image B classified using k-means algorithm, with 6 classes.
Original satellite image by Space Imaging.

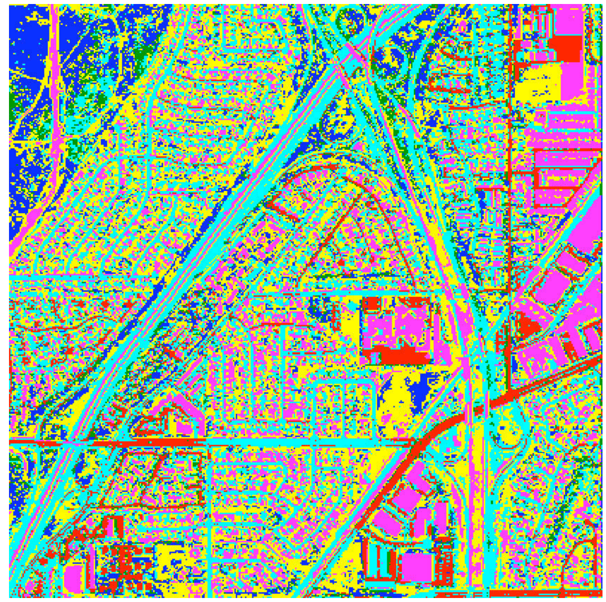

Figure 12: Image B classified by a decision tree algorithm using 4860 pixels from Figure 7.
Original satellite image by Space Imaging.

We explore the robustness of our methodology in further detail in a forthcoming paper.[13]

4. CONCLUSIONS AND FURTHER DIRECTIONS

We have developed a multi-stage method for extracting areas containing artificial structures (buildings, indicating human settlements) from satellite imagery. Because we used only a frequency threshold on the joint occurrence of edges and corners in the textural processing stages, we believe our methodology will prove robust against partial obscuration of buildings. In future work, relative placement of edges, corners, and other features can be used to detect specific types of structures. For this paper we prototyped our algorithms in ENVI/IDL.[14] The supervised classification and robustness work was done using Sapphire,[15] a system developed in part by one of the authors (Kamath). Thus far, we have used spectral mixing and textural (edge, corner) information to mitigate false positives in each type of processing. Clearly, we could detect edges and corners only within the areas of spectral mixing (Mask 1), or detect spectral mixing only in areas containing edges and corners (Mask 2) to reduce the computational burden required to produce the built-up areas (Mask 3). In future work, we hope to do such timing studies, and to explore both supervised learning techniques, especially Decision Tree and Artificial Neural Network classifiers, for various stages of the processing, as well as to examine unsupervised techniques such as isodata[16] and k-means re-clustering.[17,18] We also hope to establish robustness to geo-cultural, seasonal, illumination, and look-angle variations. We anticipate that several parameter sets will be required to produce global settlement maps, with each parameter set optimized for a particular portion of a continent and season.

ACKNOWLEDGEMENT

This work was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48.

REFERENCES

[1] J. E. Dobson, E.A. Bright, P. R. Coleman, R. C. Durfee, and B. A. Worley, "Landscan: A Global Population Database for Estimating Populations at Risk," *Photogrammetric Engineering and Remote Sensing*, **66**, (7), July 2000, pp 849-857.
[2] Space Imaging, Inc. See http://www.spaceimaging.com
[3] J. Heikkonen, and A. Varfis, "Land Cover/Land Use Classification of Urban Areas: A Remote Sensing Approach," *Int'l J. Pattern Recognition and Artificial Intelligence*, **12**, (4), June 1998, pp 475-489.
[4] Y. Zhang, "Optimisation of building detection in satellite images by combining multispectral classification and texture filtering," *ISPRS J. Photogrammetry and Remote Sensing*, **54**, (1), 1999, pp 50-60.
[5] S. M. Smith and J. M. Brady, "SUSAN — A New Approach to Low Level Image Processing," *International Journal of Computer Vision*, **23**, 1997, pp. 45-78.
[6] R. M. Haralick, "Statistical Image Texture Analysis," in *Handbook of Pattern Recognition and Image Analysis*, Eds. T.Y. Young and King-Sun Fu, Academic Press, New York, NY, 1986.
[7] M. D. Levine, *Vision in Man and Machine*, McGraw Hill, Inc., New York, 1985.
[8] P. Miller and S. Astley, "Classification of breast tissue by texture analysis," *Image and Vision Computing*, **10**(5): 277-282,1992.
[9] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray scale and rotation invariant texture classification with Local Binary Patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* **24**, 2002, in press.
[10] P. Rabindra, S. K. Sengupta, and R. M. Welch, "An Interactive Hybrid Expert System for Polar Cloud and Surface Classification," *Environmetrics*, 3, 1992, pp121-147.
[11] R. M. Welch, S. K. Sengupta, and D. W. Chen, "Cloud field Classification Based upon High Spatial Resolution Textural Features Part I: Gray Level Co-Occurrence Matrix Approach," *Journal of Geophysical Research*, **93**, 1988, pp 12663-12681.
[12] See also M. Tuceryan and A. K. Jain, "Basic Methods in Computer Vision: Texture Analysis," and R. Chellappa, *et al*., "Model-Based Texture Segmentation and Classification," in *Handbook of pattern recognition and computer vision (2nd Edition)*, 1999, 1044pp.
[13] C. Kamath, S. K. Sengupta, D. N. Poland and J. A. Futterman, "Use of Machine Vision Techniques to Detect Human Settlements in Satellite Images," *Image Processing: Algorithms and Systems II*, SPIE Proceedings, Vol. **5014**, January 2003.
[14] See http://www.rsinc.com for further details.
[15] See http://www.llnl.gov/CASC/sapphire for details.
[16] G. H. Ball, and D. J. Hall, "A Novel Method of Data Analysis and Pattern Classification, SRI, Menlo Park, CA, 1965.
[17] D. Paglieroni, "A Self-calibrating Multi-band Region Growing Approach to Segmentation of Single and Multi-band Images, *Optical Engineering at the Lawrence Livermore National Laboratory*, SPIE Proceedings, Vol. **5001**, 2003, to be published.
[18] A. Meyer, and D. Paglieroni, "K-means Reclustering: Algorithmic Options with Quantifiable Performance Comparisons," *Optical Engineering at the Lawrence Livermore National Laboratory*, SPIE Proceedings, Vol. **5001**, 2003, to be published.