Semantic Information Extraction from Multispectral Geospatial Imagery via a Flexible Framework

Shaun Gleason[1], Regina Ferrell[1], Anil Cheriyadat[1], Raju Vatsavai[1], Soumya De[2]
[1]Oak Ridge National Laboratory
[2]Missouri University of Science and Technology

1. Introduction

The Oak Ridge National Laboratory is developing an algorithmic framework to assist the geospatial image analyst with the tedious process of searching through geospatial image libraries for potential nuclear proliferation activities. Geospatial libraries are continuously collected today in higher spatial and spectral resolution than ever before. The ability to process and comprehend this data is limited by the number of analysts and available software tools. The ultimate goal of this research is to detect a variety of potential nuclear proliferation-related structures and activities. As a first step, development and testing are focused on detecting and identifying nuclear power facilities. First, images are segmented using homogeneity-based techniques including quad-tree image partitioning and region/boundary based techniques. Approximately radiometric-invariant features are extracted and used to quantize the segmented regions, or sub-images. These sub-images are fed into a probabilistic model, which predicts the most likely image "topic," or semantic description of the structures (e.g. nuclear plant, coal plant, industrial site) based on the distribution of identified objects within the image. Initial results on semantic identification of complex structures within imagery show the strong potential of the approach.

2. Data Collection and Storage

Approximately 50 multispectral (panchromatic, RGB, pan-sharpened, and infrared) satellite images have been collected from commercial satellites for U.S. and international nuclear and coal power plant facilities. Pixel data from the images and results from image segmentation and feature extraction methods have been organized and stored in a relational database. In addition to detailed information about the source and location of the images, the database contains information about low level processing results that are available for selective retrieval. This organization of data allows the automated retrieval of images, segmentation results, and feature data that can useful for efficiently generating consistent results across multiple experiments.

3. Method

The method employed within this framework for semantic information extraction consists of three major steps: (1) segmentation, (2) feature extraction, and (3) semantic identification based on a probabilistic model.

3.1. Segmentation

An important task in this research is segmenting the original image into labeled objects, or primitives. Preliminary approaches include a homogeneity-based quad-tree partitioning of the original image into tiles that were then clustered using both supervised and unsupervised techniques, and then spatially grouped to form objects. Although the partitioning scheme works well in terms of identifying the gross structure of image objects, it is limited in terms of its ability to precisely capture the boundary, and therefore, the object shape, so an adaptation of the JSEG segmentation method [1,2] is under development. The JSEG algorithm involves a two-step approach: color quantization and spatial image segmentation. In the first step, the several colors in an image are quantized into several representative classes that are used subsequently to differentiate regions in an image. Each quantized color is then assigned a class label, thus forming the class map of the image. After this, a J value for each pixel location is computed, which provides the measure of the ratio of inter-class to intra-class variability in the class map of the image [3]. One of the limitations of the JSEG approach is its inability to handle spatially varying illumination. In this research, features developed during previous work [4] are being incorporated in the quantization step so as to compute a value (similar to the J metric in JSEG) that will capture local structural information in the image to help obtain a more accurate segmentation result.

3.2. Feature Extraction

Once the image is segmented, either as tiles or irregular regions, the features are extracted to characterize the texture and structural components of the segmented regions. Then, each of the segmented regions is mapped to an associated feature space using a reduced feature vector. The original feature vector is composed of local binary patterns (LBP), local edge patterns (LEP), pixel intensity histograms, and Fourier coefficients of edge orientations (see example in Fig. 1). The LBP and LEP are rotation-invariant local texture operators to characterize pixels based on their local neighborhood intensity and edge associations, respectively. The histograms associated with LBP and LEP values generated from the segmented region are used as features to represent the texture component. Edge-orientations are extracted from the segmented region by filtering the region using steerable filters having different orientations. The resulting edge-orientation histogram is subjected to Fourier analysis to extract the coefficients representing the structure component of the region. Next, the original feature vector is subjected to a dimensionality reduction technique based on principal component analysis and linear discriminant analysis to map the feature vector to a reduced feature space. The reduced feature vectors are vector-quantized to generate “visual words” for the topic modeling algorithm described next.

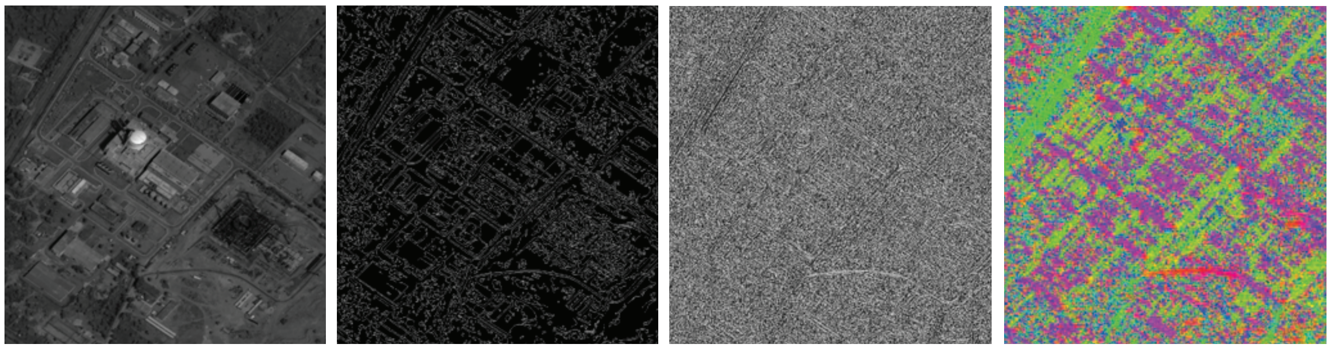

Figure 1. Images left-to-right: original, LEP, LBP, and pseudo-colored edge orientation map (-90 to +90 degrees).

3.3. Semantic Classification

The main objective of semantic classification is to annotate various image patches (regions or objects) into user defined semantic labels. The Latent Dirichlet Allocation (LDA) model [5] is one of the recent approaches used for semantic labeling. Though LDA was first presented as a graphical model for topic discovery in text analytics domain, it was recently adopted for semantic classification of remote sensing images [6]. We have adopted LDA for semantic classification of high-resolution satellite images in the context of nuclear nonproliferation. Figure 2 shows a schematic representation of LDA based classification approach.

Semantic classification process consists of dividing the image into regions using the segmentation strategies presented previously. For each region, the set of features described previously is generated. Each image is then modeled as a pair of elements (**r**, **w**), where the first element $\mathbf{r} = \{r_1, \ldots, r_N\}$ corresponds to a collection of feature vectors associated with image regions and the second element $\mathbf{w} = \{w_1, \ldots, w_M\}$ refers to the collection of 'M' words or labels. These words are then grouped into "visual" words, using vector quantization, such as the K-means algorithm. A generative model, LDA is built over these visual words. In this implementation, it is assumed that the image/label is generated by a finite Gaussian mixture model (GMM). The model fitting consists of two steps: first, the joint probability of image/label is found by marginalizing over a hidden variable z; second, the

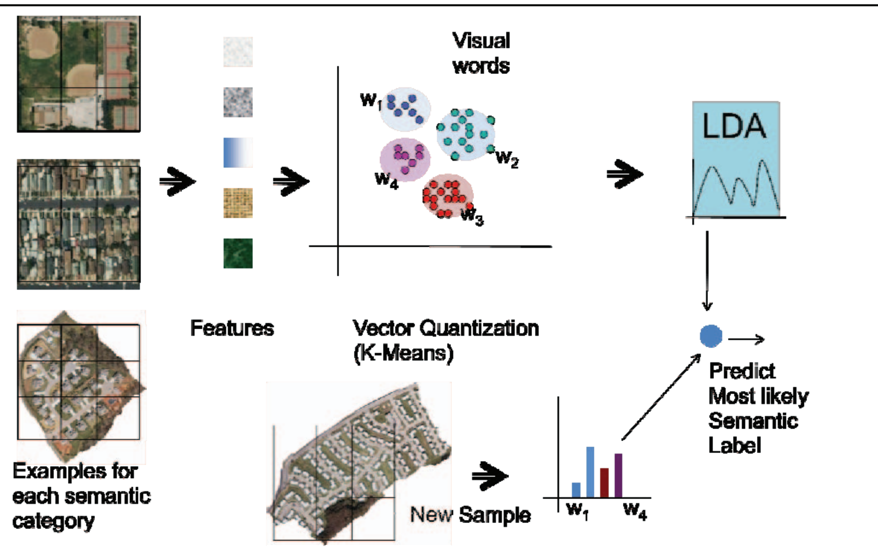

Figure 2. Schematic representation of LDA-based semantic classification of remotely sensed images.

conditional distribution of words given an image is computed using Baye's rule. The first step is performed using the Expectation Maximization (EM) algorithm. Image classification is performed using the perplexity measure.

4. Results and Conclusion

This approach was applied to the collection of high-resolution satellite images described previously, which contained 5 coal plant images and 10 nuclear power plant images. Initial experiments show 83% training accuracy and 66% test accuracy. Though the sample size is small, initial results are promising, and more examples are being collected. In the future, semi-supervised approaches need to be developed as collecting large number of nuclear plant images is very difficult and highly labor intensive. There are near term plans to extend this work by modeling spatial context in the LDA framework.

5. Bibliography

[1] Y. Deng and B. S. Manjunath, "Unsupervised segmentation of color-texture regions in images and video," IEEE Trans. Pattern Analysis and Machine Intelligence 23, pp. 800–810, August 2001.

[2] Celebi ME, Aslandogan YA, Stoecker WV, Iyatomi H, Oka H, Chen X, "Unsupervised border detection in dermoscopy images," Skin Res Technol 2007; 13:1–9.

[3] Duda RO, Hart PE, Stork DG. *Pattern Classification*. Wiley, NY: Wiley Interscience, 2000.

[4] Tobin, K.W., Bhaduri, B.L., Bright, E.A., Cheriyadat, A., Karnowski, T.P., Palathingal, P.J., Potok, T.E., Price, J.R., "Automated Feature Generation in Large-Scale Geospatial Libraries for Content-Based Indexing," Journal of Photogrammetric Engineering and Remote Sensing, 2006. 72 (5).

[5] D.M. Blei, A.Y. Ng, and M.I. Jordan, "Latent Dirichlet Allocation," Journal of Machine Learning Research, vol. 3, 2002, p. 2003.

[6] M. Datcu and D. Cerra, "Semantic Annotation of Satellite Images using Latent Dirichlet Allocation," ESA-EUSC.

Notice: This manuscript has been authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the U. S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. Research sponsored by the NA-22 Office, National Nuclear Safety Administration, U. S. Department of Energy.