Large-Scale Geospatial Indexing for Image-Based Retrieval and Analysis

Kenneth W. Tobin[1*], Budhendra L. Bhaduri[2], Eddie A. Bright[2], Anil Cheriyadat[2], Thomas P. Karnowski[1], Paul J. Palathingal[3], Thomas E. Potok[3], Jeffery R. Price[1]

[1]Image Science and Machine Vision Group, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6010
{Tobinkwjr, Karnowskitp, Pricejr} @ornl.gov
http://www.ornl.gov/sci/ismv

[2]Geographic Information Science and Technology Group, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6017
{Bhaduribl, Brightea, Cheriyadatma} @ornl.gov
http://www.ornl.gov/sci/gist/gisthome.html

[3]Applied Software Engineering Research Group, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6085
{Palathingalp, Potokte} @ornl.gov
http://www.epm.ornl.gov/~v8q/

Abstract. We describe a method for indexing and retrieving high-resolution image regions in large geospatial data libraries. An automated feature extraction method is used that generates a unique and specific structural description of each segment of a tessellated input image file. These tessellated regions are then merged into similar groups and indexed to provide flexible and varied retrieval in a query-by-example environment.

1 Introduction

Large geospatial data libraries of remote sensing imagery are being collected today in higher resolution formats both spatially and spectrally and at an unprecedented rate. These libraries are being produced for many applications including hazard monitoring, drought management, commercial land use planning, estuary management, agricultural productivity, forestry, tropical cyclone detection, homeland security, and other intelligence and military applications [1, 2]. While these systems do provide end-users with useful geographic information data products, it is typically required that a user know precise information in a world-oriented dataset regarding a region of study if they are to achieve effective results.

Techniques that facilitate search and retrieval based on image content, for example in a query-by-example environment, can provide an analyst or researcher with a rapid method for searching very large geospatial libraries with minimal query specification.

*Prepared by Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

Content-based image retrieval (CBIR) refers to techniques used to index and retrieve images from databases based on their pictorial content [3, 4]. Pictorial content is typically defined by a set of statistical or semantic features extracted from an image to describe the spectral content, texture, and/or shape of the entire image or of specific image regions. Region-based image retrieval is referred to as RBIR [5].

In a geospatial library environment these searches produce results such as the fraction of queried cover type existing in a defined region, e.g., describing the coverage of city, urban/suburban, or forest content. Many CBIR methods for geospatial data attempt to produce a description of image primitives at the pixel level (e.g., based on local structures, textures, or spectral content) [5, 6]. Yet as the resolution of these data sources increases, the ability to automatically identify cover types by classifying pixels becomes problematic due to the highly-resolved mixture of man-made and natural structures that are present in complex spatial arrangements.

Fig. 1 demonstrates this point through several examples of the high-resolution imagery that will be used throughout this discussion. These image regions represent a wide variety of cover types ranging from mixed deciduous and conifer forest lands to suburban and industrial settings. At these resolutions and with the complex proximities of the various man-made and natural structures, it is difficult to apply pixel classification methods to segment image content.

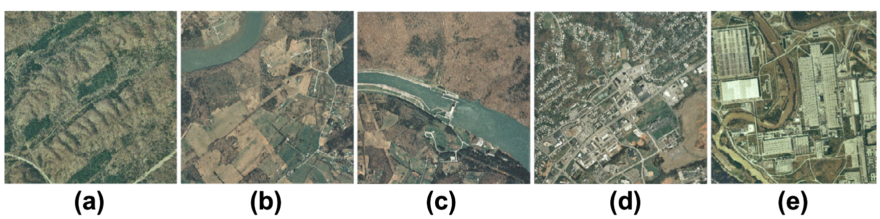

Fig. 1. Examples of a wide variety of spatial data regions that may exist in a large geospatial image database including, (a) forest, (b) agriculture, (c) water structure (locks, dams, etc.), (d) urban/suburban area, and (e) industrial sites. Resolution of these scenes are 0.5m per pixel.

At the Oak Ridge National Laboratory (ORNL) we are developing methods to automatically describe these region types so that a large image library can be efficiently assembled and indexed to perform content-based retrievals that will accommodate searches for specific spatial structure. This system encompasses three main development areas: a software agent architecture to support distributed computing and to gather image content and metadata from the web, a geospatial data modeling component to register the imagery in a consistent world-coordinate system, and a RBIR component to index imagery for search and retrieval. In this paper we will focus primarily on the RBIR aspects of search and retrieval. In Section 2 we give a brief overview of the architecture of the archive generation system that has been developed. In Section 3 we review the critical components of our image region description and indexing approach. Finally, in Section 4 we present and discuss results obtained using the data set represented in Fig. 1, a total indexed land area of approximately 153 km^2 (59 mi^2) at 0.5m per pixel resolution.

2 Overview of Geospatial Library System Architecture

At ORNL we have developed a system and architecture by combining novel approaches from three distinct research areas: software agents, georeferenced data modeling, and content-based image retrieval. The resulting technology represents a comprehensive image data management and analysis system. This system allows us to meet the challenges of organizing and analyzing large volumes of image data, and of automating the image consumption process to populate the database. The overall system approach breaks down into three components: (1) an innovative software-agent-driven process that can autonomously search through distributed image data sources to retrieve new and updated information, (2) a geo-conformance process to model the data for temporal currency and structural consistency to maintain a dynamic data archive, and (3) an image analysis process to describe and index spatial regions representing various natural and man-made cover types.

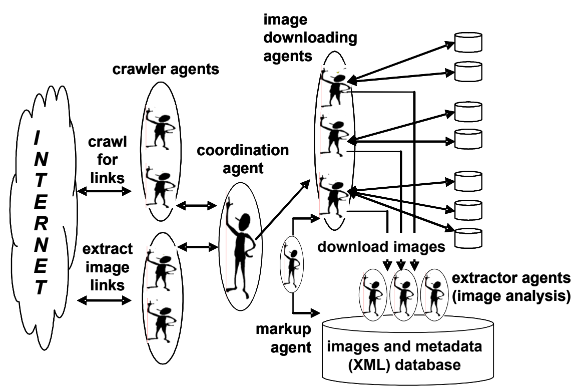

Fig. 2. Schematic representation of the agent architecture.

Fig. 2 represents the agent-based architecture of our design. There are five types of agents that are represented in this system. The Coordination Agent controls the workflow between the different agents. The Crawler Agent performs a depth-first search for image links on potential websites (in our case, only URL's ending with .edu, .gov and .net). The Download Agent downloads images for all the image links generated by the Crawler Agent. The Download Agent coordinates with the image repository to ensure that the image does not already exist in the repository or that the image is newer or has a higher resolution than the existing one.

The fourth type of agent is the Markup Agent. This type of agent creates XML files that have images marked up with their properties and metadata. For each image in the repository, this agent extracts image properties like height, width, bit planes, etc. In addition, this agent extracts geospatial information like the images bounding box coordinates from the accompanying metadata/world file. After collecting this information, it creates an XML file for each image in the image repository using all of the above-deduced properties. The XML files are then stored in a separate XML Repository.

Finally the fifth agent type, Extractor Agents, perform preprocessing of the images. Typically each Extractor Agent runs on a separate processor so that images can be processed in parallel. An image is first segmented into block segments of size 128×128 pixels. Once the image segments are created, a feature vector file describing each segment is created by making use of the image properties in the XML file and the feature extraction methods described below.

To deploy this agent architecture, we used the Oak Ridge Mobile Agent Community (ORMAC) framework [7]. This framework has been under development over the course of several agent-based research projects. ORMAC is a generic agent framework providing transparent agent communication and mobility across any Internet connected host.

3 Image Analysis

Once the imagery has been downloaded by the software agents, our goal is to generate a succinct description of an image-dependent number of contiguous areas. Fig. 3 provides an overview of the process. For this application we are restricting our analysis to a single spatial resolution. In general, for our architecture, multiple resolution data is handled independently and searches can be drilled into by performing a query at one resolution to locate candidate regions, followed by a step-up or step-down in resolution based on longitude and latitude coordinates. Our approach begins with an image tile, for example of the size represented in Fig. 1. These image tiles are 3100x3100 pixels representing a size of 1,750m on a side.

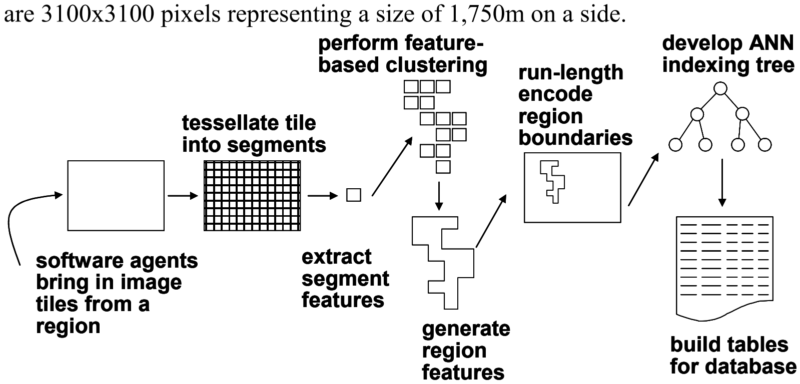

Fig. 3. Process flow shows tessellation of the input tile, feature extraction, segment clustering, indexing, and database building for query-based search and retrieval.

The tiles are tessellated into 128×128 pixel segments corresponding to 64m×64m area. The segment size was determined heuristically by ensuring that various cover structure would be adequately represented in each segment. Fig. 4 shows examples in clockwise order from the upper left of four cover types: agricultural, forested, suburban, and industrial. A number of structure-oriented features are extracted from each segment. These features are reduced using a PCA/LDA method [8] to provide a short-length vector for segment clustering by a region growing procedure to organize similar segments into contiguous groups. Each contiguous group represents a sub-region in the original image tile and a summary feature description is generated for indexing. Also, the region boundary is run length encoded for efficient storage in the database. Finally, an indexing tree is developed using the region features by application of an approximate nearest neighbor (ANN) method as described in Ref. [9]. The indexing tree provides $O[\log_2(n)]$ retrieval efficiency from the database through a query-by-example RBIR.

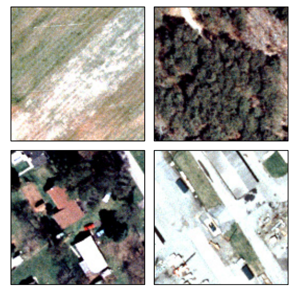

Fig. 4. Example image segments representing four cover types.

3.1 Feature Analysis

For this research, we have experimented with features that measure the texture and structure of the image segments. These features are not intended to be exhaustive in their characterization of these attributes, but rather to demonstrate the feasibility of the approach. The spectral attributes of an image segment are also valuable and have been used by many researchers to classify cover types in geospatial data [5, 6, 10]. But over large geospatial extents, spectral information can unintentionally limit a query-based search to a confined region. This is demonstrated by the Landsat Thematic Mapper (TM) data shown in Fig. 5 (30m per pixel resolution). Although we are not using Landsat TM data for this study, the four regions show agricultural areas over a large geographical distance and include Maine, Virginia, Tennessee, and Florida. Although the same three spectral bands were used to visualize crop regions, the spectral content varies tremendously. To avoid this unintentional bias in our indexing and retrieval process, we have adapted two feature sets that rely primarily on edge information to describe texture and structure.

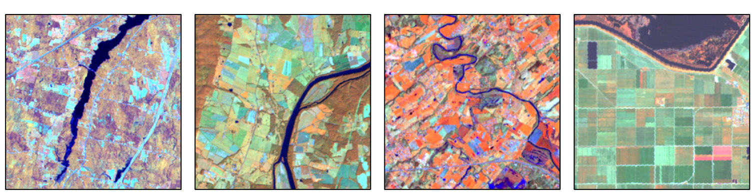

Fig. 5. Examples from the Landsat Thematic Mapper showing variation in spectral response across large geospatial extents. Three of six spectral bands have been selected for display that emphasize variations in crop cover. (UL) Maine. (UR) Virginia. (LL) Tennessee. (LR) Florida.

We characterize image segment texture using local binary patterns (LBP) [11] and local edge patterns (LEP) [12]. In the rotation-invariant LBP texture operator, each

3×3 pixel neighborhood in the intensity image is thresholded by the intensity value of the center pixel. As there are 8 neighboring pixels, each of which can be represented as a 1 or 0 (if above or below the center pixel value, respectively), it is evident that there are 256 (2^8) possible patterns that can result from this thresholding. Since, however, we desire rotational invariance, we note only those patterns that are unique under rotation. For example, the three patterns in Fig. 6 are all (approximately) equivalent under rotation about the center pixel. Applying this equivalence-under-rotation idea, it can be shown that there are only 36 unique patterns. This implies that every pixel in the image can be assigned a number from 1-36 depending upon its LBP. The 36-bin normalized distribution (i.e., histogram) of the LBP values in a given 128×128 image segment hence provides 36 features for that region.

The LEP is computed almost identically to the LBP, except that we examine 3×3 pixel neighborhoods in the image edge map rather than the image intensity values. When considering the edge map, we must also consider the state of the center pixel, which is a 1 if the center pixel is an edge, or a 0 if not. This doubles the number of potential patterns from 36 to 72 so that every pixel in the image can be assigned a number from 1-72 depending upon its LEP. The 72-bin normalized distribution of the LEP values hence provides 72 features.

1	0	0
0		0
1	0	0

0	1	0
1		0
0	0	0

1	0	1
0		0
0	0	0

Fig. 6. Three local binary patterns (LBP) that are equivalent under rotation about the center pixel.

To characterize structure in a segment, we analyze the distribution of edge orientations. The motivation to this approach is that man-made structures generally have regular edge features oriented in only a few directions (usually two for buildings) while natural image regions have randomly oriented edges. Different mixtures of man-made and natural structures will result in a variety of descriptions.

We compute local image orientation at each edge pixel using steerable filters [13]. We then find the 64-bin histogram of the edge orientation over angles from -90 to +90 degrees. The edge orientation distribution for a man-made structure is shown in the top of Fig. 7 (a) and that for the natural image is shown in the bottom of Fig. 7 (a). Note in the top of Fig. 7 (c) that there are two peaks in the edge orientation distribution near -80 degrees and +10 degrees that correspond to the orientations of the building. The distribution for the natural scene in the bottom of Fig. 7 (c) however, is approximately uniform. Since we require that the stored features be invariant to rotations of the source image, we next take the discrete or fast Fourier transform of the 64-point edge orientation histogram and keep the magnitude of the first 32 points as the final features. The magnitude of the DFT or FFT is invariant to circular shifts of the data.

The total number of texture and structure features used at this point is therefore 140 (i.e., 36+72+32). Subsequent to the feature extraction step we apply a PCA and LDA process that results in a reduction from 140 to 8 features per image segment. These features are the basis of geospatial clustering and indexing for search and retrieval.

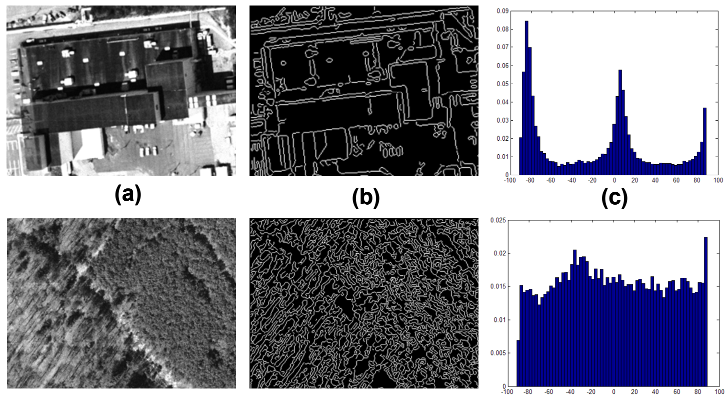

Fig. 7. Image segment in (a), edge map in (b), and edge orientation of (b) in (c). The top row represents a man-made structure and the bottom row a natural scene.

3.2 Geospatial Clustering

Once the features of the image segments have been extracted, it is possible to use this feature vector as an index for retrievals. Since there are generally a large number of contiguous 128×128 segments that define a content-based region (e.g., forested, suburban, etc.), we seek to group neighboring segments with similar features together to form a sub-region within an image tile or tiles. We perform a geospatial clustering procedure using a region growing technique to connect large contiguous and homogeneous segments of similar structure and texture characteristics.

Region growing is initialized by randomly selecting a seed segment at location (x,y), where (x,y) designates a coordinate of the corner or centroid of a segment. A segment with feature vector $\boldsymbol{v}(x,y)$ is merged with a neighboring segment with feature vector $\boldsymbol{v'}$, or with a segment group with mean vector $\langle\boldsymbol{v'}\rangle$ if,

- the coordinate of the neighboring segment or of the closest segment group is an element of the set $\{(x\pm 1, y\pm 1)\}$,
- $|\boldsymbol{v}-\boldsymbol{v'}| < T_1$ or $|\boldsymbol{v}-\langle\boldsymbol{v'}\rangle| < T_1$, where T_1 is a user-specified threshold,
- the resulting variance, σ^2, of the new segment group is less than T_2, where T_2 is a user-specified threshold used to limit the variance.

The merging process is continued until all the segments in the image tile have been tested. Fig. 8 shows typical results of this merging process.

Once the contiguous segment regions have been determined, each segment group (image sub-region) has 16 descriptive features associated with it, i.e., each sub-region is described by vector $\boldsymbol{w} = (\langle f_1\rangle, \langle f_2\rangle, \ldots, \langle f_N\rangle \ldots, \sigma_1^2, \sigma_2^2, \ldots, \sigma_N^2)^t$, for N=8, where $\langle f_n\rangle$ is the average of the n-th feature across the ensemble of segments in that group, and σ_n^2 is the corresponding variance of that feature. It is this sub-region description that is used for indexing and retrieval in the RBIR library.

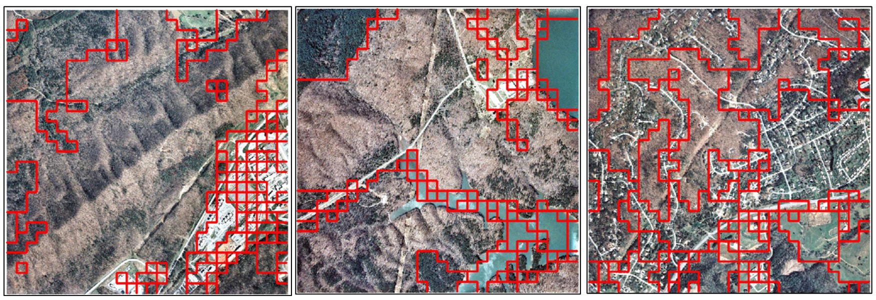

Fig. 8. Region growing results across three tiles from the image library. Each bordered region represents one homogeneous, connected group of segments as determined by their texture and structure features.

4 Results

The results presented here are for a geospatial library composed of 50 image tiles (3500x3500 pixels, 1750m×1750m, 0.5m per pixel) representing approximately 153 km^2 of land in and around the U.S. Department of Energy's Oak Ridge Reservation [14]. For this demonstration, the image regions (i.e., over all tiles) were tessellated into 39,200 segments of size 128×128 pixels. Features were then extracted for each of the segments, which were subsequently clustered as described in Sections 3.1 and 3.2 above. The number of sub-regions developed through geospatial clustering was 4,810, resulting in a reduction of 88% in the number of unique, spatially distinct objects indexed for retrieval.

For demonstration purposes, we have indexed the original 39,200 segments in one descriptive dataset, and the 4,810 sub-regions in another dataset. Fig. 9 shows an example retrieval of several cover types at the image segment level. In the figure, each of the images in the left-hand column represents a query. The remaining five images in each row represent the top five matching results of the query for an industrial complex in (a), suburban area in (b), an agricultural area in (c), and a specific search for striped parking lots and roadways in (d). Note the flexibility of the system to locate imagery of an extreme variety of cover types and detail using only eight features (from the original 140).

Although this type of segment-level query can be useful to an analyst, it can also provide too much redundant information regarding large spatial extents of visually similar imagery. For example, forested or agricultural regions may occupy a large percentage of the data library in a rural region such as this. An ability to collect together these large geospatial regions, or alternatively, to search for small man-made structures of a particular type distributed throughout large contiguous rural areas (i.e., a needle in a haystack) would also be useful. Therefore, in Fig. 10, we show examples of three queries performed on the geospatially clustered regions of the image library. Once again, the query region is represented by the left most column of images. In (a) we see an industrial complex, in (b) a large deciduous forest region, and in (c) an example of large, contiguous agriculture areas.

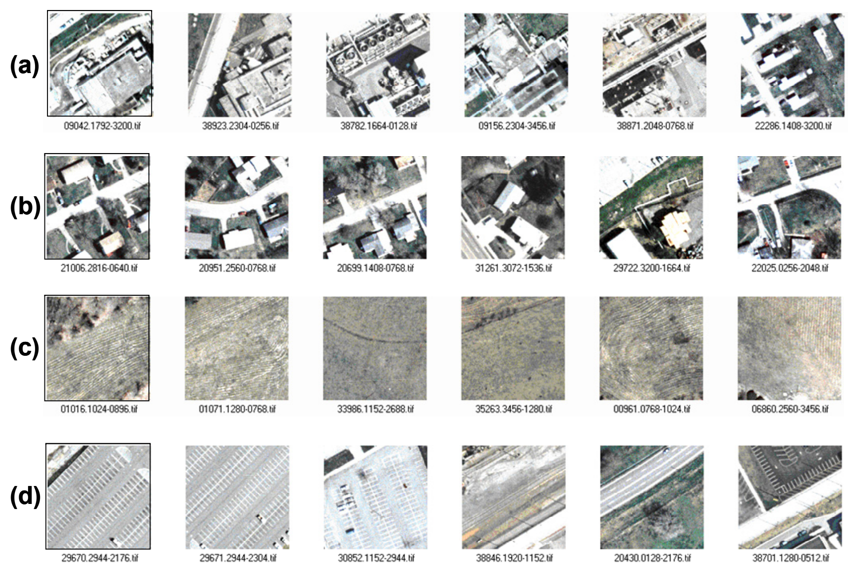

Fig. 9. Examples of the retrieval of image segments. The query images are in the left column. (a) an industrial complex, (b) a suburban setting, (c) agriculture, and (d) striped pavement.

5 Conclusion

In this paper we have presented a novel method for automated feature extraction, spatial clustering, and indexing of a large geospatial image library. Although retrieval experiments were described for a relatively small geospatial data set, the system architecture and processing methodology have been developed to facilitate very large data libraries that can be maintained and updated in a dynamic manner through distributed computing with a software agent architecture. The feature analysis and indexing approach used in this research provides an efficient and flexible method for describing a broad range of cover types while allowing a user to locate very specific structural detail in a query-by-example environment. Future work in this area includes the incorporation of other geographical information metadata into the query process along with the addition of spectrally-based features for augmenting the specificity of local searches within a geospatial region of interest.

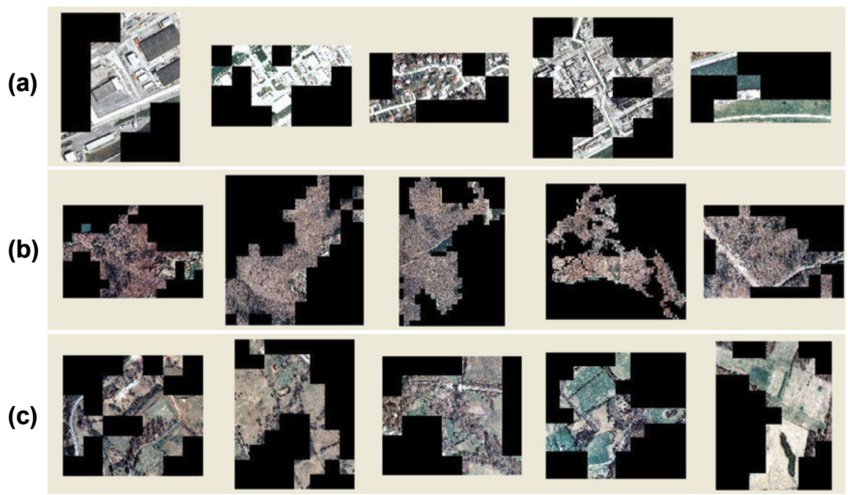

Fig. 10. Examples of the retrieval of geospatial clusters. (a) An industrial complex, (b) a large expanse of deciduous forest, (c) an agricultural region.

References

1. Tapley, B.D., Crawford, M.M., Howard, T., Hutchison, K.D., Smith, S., "A Vision for Creating Advanced Products From EOS Core System Data to Support Geospatial Applications in the State of Texas," International Geoscience and Remote Sensing Symposium, Vol.2, 2001, p. 843-845.

2. You, J., Cheung, K.H., Liu, J., Guo, L., "On Hierarchical Content-Based Image Retrieval by Dynamic Indexing and Guided Search," Proceedings of SPIE, Storage and Retrieval Methods and Applications for Multimedia, Vol. 5307, 2004, p. 559-570.

3. S. Santini, Exploratory Image Databases, Content-based Retrieval, Academic Press, San Fransisco, CA, 2001.

4. E. Vicario, Editor, Image Description and Retrieval, Plenum Press, New York, NY, 1998.

5. Datcu, M., Daschiel, H., Pelizzari, A.,Quartulli, M., Galoppo, A., Colapicchioni, A., Pastori, M., Seidel, K., Marchetti, P.G., D'Elia, S., "Information Mining in Remote Sensing Image archives: System Concepts," IEEE Trans. On Geoscience and Remote Sensing, Vol. 41, No. 12, Dec. 2003, p. 2923-2936.

6. Schroder, M., Rehrauer, H., Seidel, K., Datcu, M., “Interactive Learning and Probabilistic Retrieval in Remote Sensing Image archives,” IEEE Trans. on Geoscience and Remote Sensing, Vol. 28, No. 5, Sept. 2000, p. 2288-2298.

7. Potok, T., Elmore, M., Reed, J. and Sheldon, F.T., "VIPAR: Advanced Information Agents Discovering Knowledge in an Open and Changing Environment," Proc. 7th World Multiconference on Systemics, Cybernetics and Informatics, Orlando FL, pp. 28-33, July 27-30, 2003.

8. Bingham, P.R., Price, J.R., Tobin, K.W., Karnowski, T.P., “Semiconductor Sidewall Shape Estimation,” SPIE Journal of Electronic Imaging, Vol. 13, No. 3, July 2004.

9. Tobin, K. W., Karnowski, T.P., Arrowood, L.F., Ferrell, R.K., Goddard, J.S., Lakhani, F., “Content-based Image Retrieval for Semiconductor Process Characterization”, EURASIP Journal on Applied Signal Processing, Special Issue on Applied Visual Inspection, Vol. 2002, No. 7, 2002.

10. Harvey, N.R., Theiler, J., Brumby, S.P., Perkins, S., Szymanski, J.J., Bloch, J.J., Porter, R.B., Galassi, M., Young, C., “Comparison of GENIE and Conventional Supervised Classifiers for Multispectral Image Feature Extraction,” IEEE Transactions on Geoscience and Remote Sensing, VOl. 40, No. 2, Feb. 2002, p. 393-404.

11. Pietikainen, M., Ojala, T. and Xu, Z. Rotation-invariant texture classification using feature distributions. Pattern Recognition, vol. 33, no.1, pp. 43-52, 2000.

12. Yao, C.-H. and Chen, S.-Y. Retrieval of translated, rotated and scaled color textures. Pattern Recognition, vol. 36, pp. 913-929, no. 4, 2003.

13. Freeman, W. and Adelson, E. The design and use of steerable filters. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 13, no. 9, pp. 891-906, 1991.

14. Tuttle, M., Pace, P., “ORNL Basemapping and Imagery Project: Data Collection, Processing, and Dissemination,” Geographic Information System (GIS) Environmental management conference, Reno, NV, CONF-9603148-1, 1996.