COMPARING SPECTRAL AND OBJECT BASED APPROACHES FOR CLASSIFICATION AND TRANSPORTATION FEATURE EXTRACTION FROM HIGH RESOLUTION MULTISPECTRAL IMAGERY

Sunil Reddy Repaka, Research Assistant
Dennis D. Truax, Professor
2351 Walker Engineering Building
Department of Civil Engineering
Mississippi State, MS 39762
srr52@msstate.edu, truax@engr.msstate.edu
Eric Kolstad, Research Assistant, Dept. of Computational Engineering
Charles G. O'Hara, Consortium Manager, NCRST-E
Engineering Research Center
2 Research Boulevard
Mississippi State, MS 39762
eric@erc.msstate.edu, cgohara@erc.msstate.edu

ABSTRACT

An increasing need exists to update older transportation infrastructure, land use/land cover, environmental impact assessment and road network layer maps. Planning and development rely on accurate data layers for new construction and changes in existing routes. Recent developments in commercial satellite products have resulted in a broader range of high quality image data, enabling detailed analysis. This information differs in its characteristic features (e.g. spatial resolution and geolocational accuracy) as well as utility for particular tasks.

Transportation features have historically been difficult to accurately identify and structure into coherent networks; prior analyses have demonstrated problems in locating smaller features. Roadways in urban environments are often partly obscured by proximity to land cover or impervious objects. Recent research has focused on object-based methods for classification and different segmentation techniques key to this approach. Software packages such as eCognition have shown encouraging results in assessing spatial and spectral patterns at varied scales in intelligent classification of aerial and satellite imagery.

In this study we examine 2.44m QuickBird and 4m Ikonos multispectral imagery for a 7.5' quad (Dead Tiger Creek) near the Mississippi Gulf Coast. Both spectral and object-based approaches are implemented for pre-classification, after which road features are extracted using various techniques. Results are compared based on a raster completeness model developed. Challenges include intricate networks of smaller roads in residential zones and regions of tall/dense tree cover. Observations for these sites will assist in developing a larger-scale analysis plan for the CSX railroad corridor relocation project.

INTRODUCTION

Over the past decade, methods utilizing remotely sensed data for extraction of transportation features have been considered for application in updating cartographic products and databases. Remotely sensed imagery is a readily accessible global resource that provides wider variety and greater quantity of information relative to traditional mapping data. It has particular utility for transportation applications such as land use and land cover change detection, transportation infrastructure management, environmental impact assessment and creation of high-accuracy base maps for intelligent transportation planning. With recent increases in spatial resolution, however, feature extraction poses more of a challenge. Obtaining maximum benefit from these resources will require research involving not only improved data analysis approaches, but also development of tools for the manipulation of such data. Data extraction and analysis steps are of particular importance.

An overall methodology can be developed in the process of examining change in transportation features. In this paper, we examine imagery from different satellite sources, using four methods of classification to evaluate road features. Analysis, comparison and development of statistics based on the extraction and classification of transportation features were performed with spectral and object-based analyses for an ortho-quad from the

Mississippi coastal corridor area. One 7.5' quad region (Dead Tiger Creek, shown in Figure 1) is the focus of this study. This area was selected based on the regional data available from two commercial high-resolution satellite data vendors: DigitalGlobe (QuickBird imagery) and Ikonos. QuickBird's high spatial resolution provides a means to more effectively identify man-made objects and has an additional advantage in mapping and planning activities when analyzed in conjunction with Ikonos data, a similarly high-resolution source.

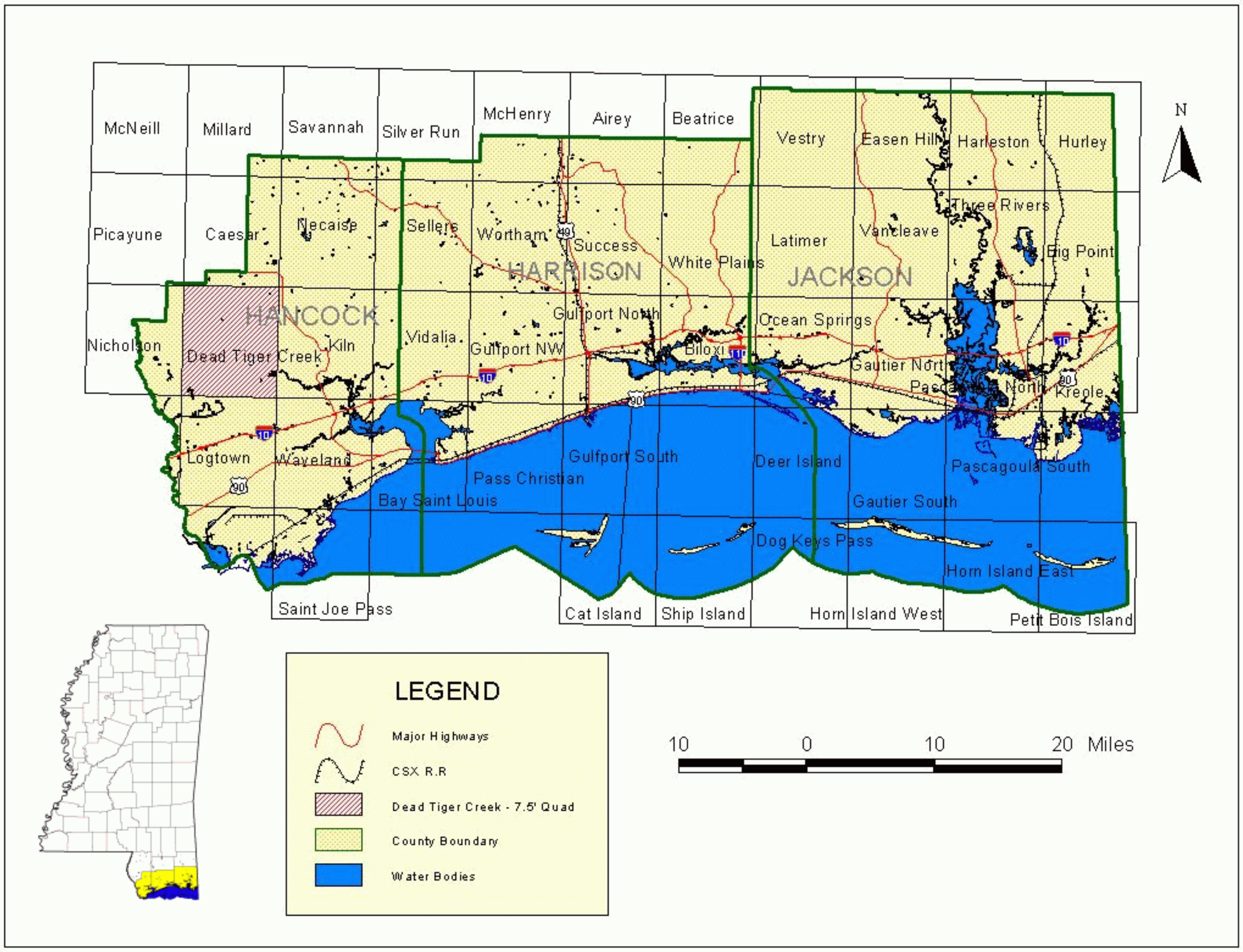

Figure 1: Mississippi Gulf Coast with major transportation features and study area

ANALYSIS BACKGROUND

A preliminary object in examining large volumes of multispectral image data was to reduce it to a manageable number of subsets meaningful to our analysis. These classes may be further categorized into land cover types through *image segmentation*. Varied well-known classification algorithms were considered, several of which were selected for this investigation. An overview of these methods is as follows.

Spectral-Based Approaches

Simple spectral or pixel-based classifiers were first developed in the 1970's for use with multispectral data. A *spectral-based* algorithm utilizes spectral pattern value combinations associated with different feature types, each assigned a unique Digital Number (DN), evaluating spectral reflectance and emittance values present within each pixel to find meaningful patterns. A pixel's class is determined from the image data's overall DN value. Statistics are derived from the spectral characteristics of all pixels in the image. Both *unsupervised* and *supervised* spectral-based approaches are routinely applied to remotely sensed data based on spectral or pixel-based schemes.

Unsupervised classification algorithms are optimal in cases where detailed knowledge such as ground truth data is not readily available for a study region. Based on user-defined parameters, unknown image pixel data is iteratively grouped together in clusters until either some proportion of pixels' class values remain unchanged or a maximum number of iterations has been reached. Classes can be determined by spectral distinctions inherent in the data (Lillesand and Kiefer 2002). Three major unsupervised classification methods generally used are:

- **K-Means** – A number of clusters, *n*, is specified by the analyst. An arbitrary set of *n* cluster centers is generated in the multidimensional measurement space, after which these are iteratively repositioned until optimal spectral separability is achieved based on distance-to-mean.
- **Fuzzy C Means** – A method similar to K-Means which also incorporates fuzzy logic in later processing.
- **ISODATA** – The Iterative Self-Organizing Data Analysis Technique begins with either an existing set of cluster means or an arbitrary one evenly distributed in the data space. Pixels are iteratively classified using a minimum spectral distance formula. At the end of each iteration, current class means are recalculated. A new pixel classification is then performed relative to the new mean locations. This cycle repeats until the number of pixels in each class changes by less than a user-specified convergence limit (e.g. 90% of the pixels remain in the same cluster after two iterations), or a maximum number of iterations is reached. At the end of the process, the user marks which class or classes correspond to a desired feature.

Supervised algorithms rely on user-defined training sets defined by Regions Of Interest (ROIs), image areas whose classifications are assumed to be known. These regions are selected so as to be reasonably representative of only one class. Several significant methods of supervised classification include (Hodgson, 1988):

- **Maximum Likelihood** – This approach assumes that statistics for each class in each band are normally distributed. Probabilities that a given pixel belongs to an arbitrary class (Gaussian value distribution) are computed, and the pixel is assigned to the most likely of these within a probability threshold, if specified.
- **Minimum Distance** – Using the mean vectors of each ROI, Euclidean distance from each unknown pixel to the mean vector for a class is computed. Pixels are associated with the closest ROI class unless the user specifies standard deviation or distance threshold criteria (which may leave certain pixels unclassified).
- **Mahalanobis Distance** – This method is a direction-sensitive distance classifier that retains unique statistics for each class. While similar to the Maximum Likelihood classification, it assumes all class covariances to be equal, making it somewhat faster in execution. Pixels are likewise assigned to the closest ROI class, where a distance threshold may be set to limit classification.
- **Spectral Angle Mapper Classification** – The Spectral Angle Mapper (SAM) method involves a physically based spectral classification that uses an *n*-dimensional angle to match pixels to reference spectra. Spectral similarity between two spectra is determined by calculating the "angle" between spectra distribution's primary axes, treating them as vectors in a space with dimensionality equal to the number of bands.

Object-Based Classification

Object-based classification attempts to describe relationships in terms of several categories of characteristics, each of which may be assigned weighting factors. Object features include properties such as color, texture, shape, area, and scale. Class-related features involve a connection to nearby objects, e.g. super- and sub-objects in a hierarchy. Other relationships include spatial terms such as "nearest neighbor" or statistical similarities. Panchromatic aerial photographs for instance might be classified with additional information concerning texture and relational descriptions. In order to classify objects, the source image must be partitioned in some respect.

Image segmentation seeks to classify objects by partitioning or grouping related regions based on shared attributes (Asano *et al*., 1996). Groups of pixels within each region can be viewed as a statistical sampling of collected multispectral or feature values, or alternately as derived information including as band ratios, textural features, or a region's shape. Derived data can provide further cues for labeling regions, e.g. as ground cover or land use categories. Standard segmentation can be further augmented by *knowledge-based partitioning* and the construction of *sub-objects* for special classification tasks. Knowledge-based partitioning utilizes previously created classifications as additional cues for the merging of objects. Segmented sets of one class can be merged at the same level or grouped beneath a new, higher level. Relationship-based classification is possible as each object is aware of its neighbor, sub- and super-objects. Ensuring a consistent hierarchical structure requires that: 1) object borders of higher levels are inherited by the sub-levels, and 2) the segmentation process is bounded by super-object borders (Willhauck, 2000). As the segmentation process is comparable to the construction of a database with information of each image object, the classification process can be viewed as a form of database query (Tilton, 1997).

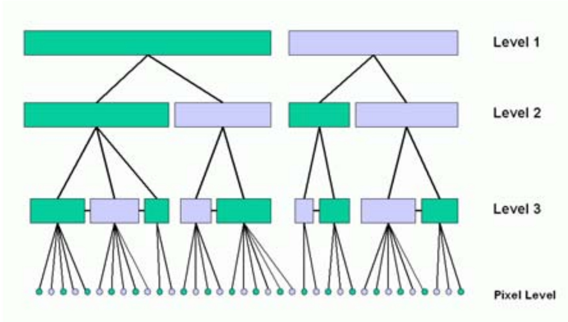

Figure 2. Hierarchical network of image objects (after Willhauck, 2000)

Partitioning is done via image segmentation techniques such as *clustering*, *boundary detection* and *region-growing*. *Fuzzy logic* functions permit complex, adjustable-detail classifications. A continuous range of values [0...1] represents degree of certainty, as opposed to binary (0/1) decision. *Membership functions* or a *nearest neighborhood* classifier are used to translate the range of most feature types into adaptable fuzzy logic expressions such as AND (max), AND (mean), OR, IF and ELSE. All expressions that form part of one class are combined to reach the final result (Mayer *et al*., 1993).

Clustering algorithms such as K-Means and ISODATA make decisions based on individual pixels and/or pixel regions; this is sometimes ineffective because it does not exploit spatial information. Boundary (edge) detection exploits spatial information by locating local edges throughout the image. Simple, noise-free images result in predictable boundaries, although more complex or noisier images can produce missing or extraneous edges and region boundaries that don't necessarily form closed connected curves around the joined regions. Region growing is preferred because it exploits spatial information and guarantees the formation of closed, connected regions of comparable value in factors such as homogeneity or boundary sharpness. However, spectrally similar but spatially disjoint regions do not get associated. Region growing is also computationally intensive, and it's unclear at what point processing should stop. Hybrid region growing and spectral clustering methods such as the *connectivity-preserving relaxation-based segmentation* (aka *active contour*) model have been suggested (Tilton, 1997).

Method selection

Image interpretation and understanding always involve object recognition and scene analysis tasks (Duda and Hart, 1973). Presently, object recognition has been accommodated by both unsupervised and supervised classification techniques (Richards, 1986), with spectral similarity as its methodological core. This spectral-based approach implies user efforts in recognizing image objects (Wilkinson and Burril, 1991) either prior to the algorithm (when selecting training areas in supervised classification) or after the algorithm (for validation of the clustering results), meaning that user reasoning guides the classification process. The challenge is to arrive at the set of spectral data clusters that maximizes the mapping to the conceptual and knowledge user's models. Scene analyses representing structural relations among image objects are mainly affected by user reasoning, though some work has been developed to accommodate it (Wilkinson and Burril, 1991).

In this study, both an unsupervised and supervised clustering algorithm are examined. ERDAS Imagine was chosen to implement the ISODATA algorithm (incorporating a modified K-means statistical clustering process). The maximum likelihood technique was selected for supervised classification tests. As this algorithm assumes a Gaussian distribution of pixel values within each training class, it tends to be somewhat more accurate in regions with high surficial variability. Such regions are anticipated for our study area, which covers a variety of neighboring ground cover types (e.g. roads, forest, grasslands). Image pixels that fall within some standard deviation of the training class mean will thus be assigned to that class. This method also allows for a weighting factor; image pixels are less likely to be classified as ground cover types with lower probabilities of occurrence within the scene.

The object-based classification approach evaluated utilized Definiens' software product, eCognition, which permits object-oriented, multi-scale image analysis. eCognition provides a patented feature called "Multiresolution Segmentation," an extraction technique that segments panchromatic and multispectral images into highly homogeneous image objects at a specified resolution, forming a hierarchical network of these objects. These image objects represent the image information in a format allowing for further classification steps. They can contain attributes including spectral information as well as color, texture, shape and area, along with context and relationship data concerning other object layers. Adjacent regions in an image will in theory be separated as long as they have significant contrast to each other - even in cases where the regions are contain certain textural or noise components (eCognition User Guide).

STUDY PREPARATION

Among varied computational GIS software packages available for data manipulation, analysis and presentation, two leading software packages were selected specifically for their strengths in evaluating the two methodologies outlined above. Utilized in this research were the image-analysis based ERDAS Imagine (version 8.6), and Definiens' object-oriented package eCognition (version 3.0). ERDAS Imagine includes tools for image mosaicking, 3-D visualization, surface interpolation, data modeling, vector and raster support, image interpretation, rectification and spatial analysis. Definiens' eCognition similarly has features which include high-performance image analysis functions, object-based data fusion and adaptive approaches to intelligent discrimination of objects within a scene.

This study focuses its efforts on an area within the Mississippi coastal corridor. This region consists of different land cover type like forests, transition zones, wetlands and marshes, residential area and industrial area. One quad section was selected via query based on local availability of QuickBird and Ikonos imagery. DigitalGlobe's QuickBird satellite provides high (sub-meter) resolution imagery in both panchromatic and multispectral formats, with geolocational accuracy and a relatively large imaging footprint. In this study we have utilized Ikonos 4m multispectral and 1m panchromatic imagery with a CE_{90} (circular error with 90% certainty) measure of 15 meters, and QuickBird 2.44m multispectral and 0.6m panchromatic imagery with a CE_{90} of 23 meters.

To proceed with comparative evaluation, distinct feature elements must first be identified on the land use map. Elements such as roads, buildings, and other land use types are very diverse in space and spectrum. Some areas are particularly vulnerable to damage; frequently these are urban landscapes. Urban landscapes are composed of varied materials (concrete, asphalt, shingles, metal, plastic, glass, water, grass, shrubs, trees, and soil) arranged by humans in complex ways to build houses, transportation systems, utilities, commercial buildings, and recreational landscapes (Swerdlow, 1998). To remotely sense these urban phenomena, it is necessary to appreciate the urban attributes' temporal, spectral and spatial resolution characteristics. A proper understanding of the temporal development cycle, i.e. the changes that occur in a certain period of the urban phenomena, is necessary to prevent embarrassing and costly interpretation mistakes. When extracting urban/suburban information from remotely sensed data, it is more important to have high spatial resolution than a large number of multispectral bands. A minimum spatial resolution of from $\leq$ 0.25 meter to 5 meters is necessary to detect or distinguish between the types of individual buildings (Jensen and Cowen, 1999). This range can be achieved with both QuickBird and Ikonos imagery.

IMAGE PROCESSING

Scenes registered to the same geographic datum and map projection were utilized in research and analysis. The QuickBird image products were received partitioned into smaller pieces due to media storage and analysis limits; DigitalGlobe provides tiles at either 8,000x8,000 or 16,000x16,000 pixels. Larger data files reaching 2 GB or more can prove difficult to import and process with available GIS software. More significant is the fact that study area images were acquired across a broad range of dates. Data availability limited the selection to 02/17/02 and 5/23/02 in the case of Quickbird, and 11/08/00 and 10/05/01 for Ikonos images. Regardless of image source, definite problems arise when attempting to mosaic imagery of varied temporal origin; there is no one "correct" approach. Visual discrepancies make it difficult to produce mosaic with histogram matching alone, or via color balancing with an intensity adjustment algorithm. Varied parameter combinations were tried in an attempt to obtain usable mosaics for classification. Feathering techniques were opted for in order to match edges, forming an overall seamless image.

While panchromatic images provide a usable classification basis within eCognition's object-based approach, examining these via a pixel-based method "is usually not recommendable to apply such enhancement methods (pansharpening), since they falsify the spectral properties of the objects and thus might complicate the subsequent classification" and "the objects' spectral statistics become modified, which could lead to a poorer separability" (Hofmann, 2000). After mosaicking, a considerable change was observed in individual pixels' DN values after color balancing and histogram matching. Figure 3b shows two examples of composited results. Images were therefore classified separately prior to the mosaic process to avoid uncertainties caused by variation in spectral values. Image registration with a ground truth layer was performed after classification; points for co-registration are generated from the original multispectral image and then used for the classified image. This also helps avoid spectral value changes introduced in image resampling, keeping error to a minimal level.

Figure 3a. Ikonos scene mosaic, AOI #1

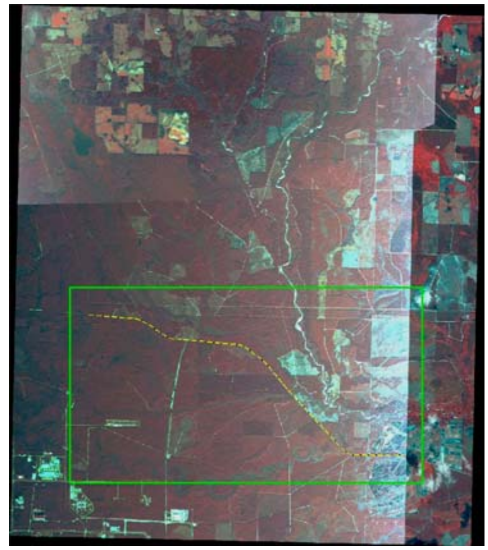

Figure 3b. QuickBird scene mosaic, AOI #1

ERDAS Imagine 8.5 was used to mosaic the 7.5-minute Digital OrthoQuads. A series of steps found to give acceptable results is as follows. Initially, image resampling was performed using Nearest Neighbor interpolation, with default grid sampling density and an RMS tolerance of 0.10 pixels. The appropriate resampling method varies depending on the application. Nearest neighbor interpolation is most effective where original pixel intensity values are required to remain intact, e.g. in assessment of discrete surfaces or vegetative health. Other applications interested more in the visual appearance may use methods such as cubic convolution, which produce smoother images while maintaining edge detail.

A histogram match was done using 1> Automatic color balancing technique with a parabolic [surface specified in "mosaic preferences"], and 2> pixel value with no image matching performed. For overlapping intersection regions, no cutlines were used. Individual images were stitched together using feathering. The mosaic was created either with output defined by an AOI or an exact quad boundary. It is recommended that the complete mosaic be retained, as this provides overlap and onlap information when merging 2 or more quads together. Due to large variations in the Digital Number values of individual image pixels, however, the composited results were not immediately used in classification. Instead, separate scenes were first classified individually, at which point these were mosaicked (in the case of Ikonos imagery) or classified and registered (in case of QuickBird imagery).

CLASSIFICATION AND ANALYSIS

Ground truth data for road networks and GCPs for the study were difficult to obtain at this point of study, due to various project constraints. For this reason Ikonos panchromatic imagery was used as a reference layer. QuickBird image scenes were first classified and then registered, using the model information from registration done to rectify QuickBird multispectral imagery with Ikonos pan imagery. Thereafter the image tiles were mosaicked, and subsetted to form the final classified image of the Dead Tiger Creek quad. In case of Ikonos imagery, the classified image scenes are directly mosaicked and subsetted to produce the study quad image output. The conventional method of accuracy assessment was not utilized here, as it does not properly represent the classification. For this purpose we propose a new model termed "Raster Completeness". In this case, the content of road polygon regions is compared to the total number of pixels marked as roads in the classified image, giving a percentage of raster values which characterize the results. In selecting the different roadways/regions for analysis we considered using road networks with different alignments in terms of direction. This would aid in evaluating directional extraction.

Two classification procedures were examined as part of this investigation. For the spectral-based approach, ERDAS Imagine was used for classification. We initially conducted an unsupervised classification for 100 classes, using the ISODATA statistical clustering procedure. These classes were thereafter combined to get a set of desired class categories. Using knowledge gained from the unsupervised classification, signature sets were generated for different classes and we conduct a supervised maximum likelihood classification. In this supervised classification all classes were used by default. One could however use only specific bands during classification. As such, an attempt was made to use a three-band combination, eliminating the blue band (which generally carries lot of haze and atmospheric dispersion); results were similar to those obtained with a complete set of four bands. When creating more homogenous signatures for supervised classification to increase the accuracy and extent of classification, there were large conflicts where areas were being misclassified. These signatures were overlapping one another and in turn creating conflicts in classification. The effective limit was around 20-30 distinct signatures, depending on the image being examined. After different sets of signatures for the classes are generated, the classes could be labeled and aggregated. The resulting signature sets were tested for reusability with the same type of image to assess whether some proportion of the results would match. The image classification however suffered a major setback when trying to use signatures developed for one scene on the other scenes. Ground cover type and land use classes for classification which gave distinct classes using unsupervised and supervised classification are: Water Bodies, Road Networks / Built-up, Sparsely Vegetated, Forested Upland, Barren Land.

Definiens' eCognition was used to evaluate object-based classification. eCognition has two main types of classifiers, one involving a membership function and the other a nearest neighbor approach. Classification-based segmentation involved three methods: class-based fusion of image objects, border optimization and image object extraction. Segmentation occurs in two stages, initially defining a structure of group sets which is followed by object generation. Efforts focused on ascertaining which function would perform better for our classification and what parameters would work optimally.

Classification decisions group sets of unique objects into classes, which members share a common feature. In standard classification, the objects are single pixels with attributes for Value, Position and Size. The pixels line up in arrays which collectively form an image. A digital image contains only implicit information about the objects in the scene. Based upon object models, it is possible to discern individual entities in a seemingly unstructured collection of pixels. In a per-field analysis or a "pixel-in-polygon" analysis, pixel information is already linked to a spatial database built up in a digitizing session. In the spatial database, besides the explicit information, there is still a huge amount of implicit information available (Sester, 2000).

Traditionally, image analysis takes place in three basic domains of image data, namely image space, spectral space and feature space (Landgrebe, 1999). Object-based analysis uses the 'image object' or 'local pixel group' as a basis. The measured electromagnetic energy per pixel serves as the base for the pixel-oriented multispectral classification. Each pixel is characterized by a special reflectance in the multidimensional feature space (spectral signature). The pixel will be associated with a certain class according to this spectral signature (Leiss, 1998).

Each image scene was segmented using the best suitable criteria developed from a test region subjectively chosen from a common area in Ikonos and QuickBird imagery. A major advantage of this type of segmentation process is the use of high-resolution panchromatic images in combination with relatively lower resolution multispectral imagery. This offers a certain advantage with object-based classification vs. pixel-based methods, which often utilize only multispectral imagery. In segmentation, grayscale images are handled as color images having "lightness" coordinates. Tests are carried out on each pixel, followed with histogram thresholding, and region growing is done in places where one or more matching pixels are detected. Table 3 relates the best results obtained for the varied data sets and methods used to compare object-based feature extraction.

Table 3. eCognition Segmentation criterion for QuickBird and Ikonos

Imagery Source	Segmentation Level in Object Hierarchy	Scale Parameters	Homogeneity criteria		Shape Ratio		Segmentation Mode (Algorithm)
			Color %	Shape %	Compactness	Smoothness	
QuickBird	Level 1	40	80	20	0.7	0.3	normal
Ikonos	Level 1	30	80	20	0.7	0.3	normal

In this step the above segmentation levels were utilized to create five land cover classes: Water Bodies, Roads, Sparsely Vegetated, Forested Upland, and Barren Land. Nearest neighborhood classification was used to create one training set and a membership function was used to create the second set. In the case of membership functions, two classes were defined: Roads and Not Roads. Different criteria based on object features related to layer values such as mean of different layers, brightness, standard deviation, ratios and other statistical measures were used in combination to develop the membership classification.

Figure 4. Dead Tiger Creek: Areas of Interest

Evaluation results

Imagery acquired at different times of the year noticeably influenced the extraction process results. The imagery was inconsistent, as leaf-off conditions generally permit good classification of roads (although having potential to over-classify), whereas leaf-on conditions can produce reasonable classifications but may have tree canopies blocking the aerial view. Tall trees along the road corridors also cover a significant percentage of the roads, causing obstruction where satellite images' off-nadir view angles are the case of non-vertical scene areas.

Remotely sensed data from QuickBird and Ikonos were classified and the accuracy of each of the different techniques tested. Figure 4 (at left) highlights the five study areas (AOIs) selected in the overall scene. Figures 5a-5d show example results obtained via supervised and unsupervised classification (via ERDAS Imagine), membership functions and a Nearest Neighbor approach (using eCognition) for AOI #3. Tables 4-7 on the following page relate summary statistics concerning respective methods' analysis obtained for the various regions.

Figure 5a. Supervised classification, ERDAS Imagine (Ikonos multispectral image source)

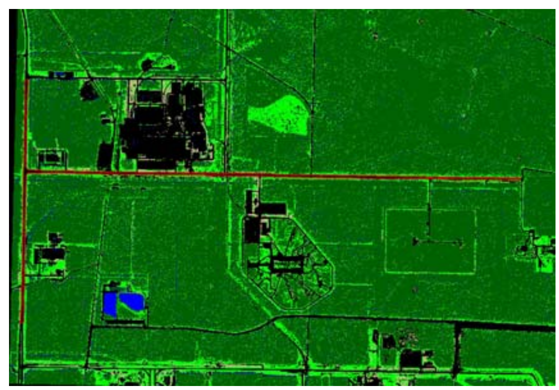

Figure 5b. Unsupervised classification, ERDAS Imagine (Ikonos multispectral image source)

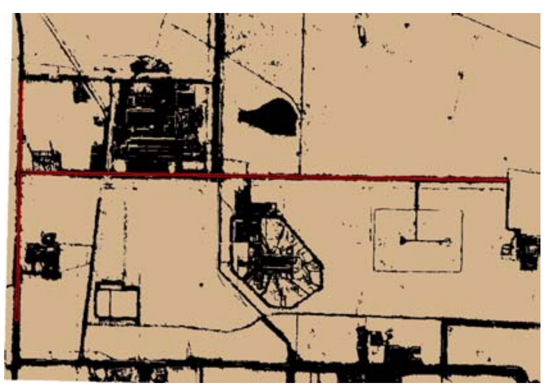

Figure 5c. Membership function, eCognition (QuickBird multispectral & panchromatic sources used)

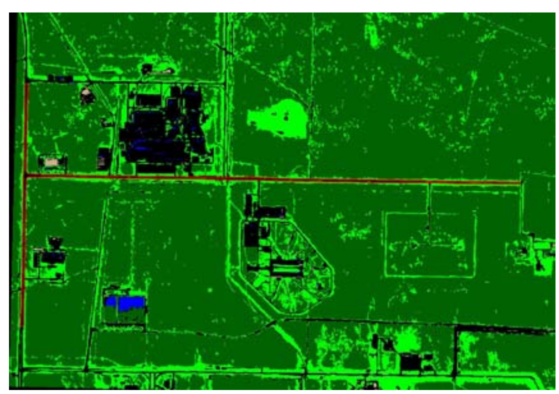

Figure 5d. Nearest Neighbor classification, eCognition (QuickBird multispectral & panchromatic sources used)

Table 4. **Extraction results from Supervised/Unsupervised classification for Ikonos imagery**

Road Poly AOI	Total # of Pixels in AOI	Percentage Classified		Approximate Length (m)	Roadway Orientation
		Supervised	Unsupervised		
1	9035	67.93	72.86	9600	NW/SE
2	3228	70.38	66.88	3600	N-S
3	5958	84.12	74.38	4950	E-W
4	3996	83.03	86.46	2500	NSEW (Intersection)
5	1346	72.36	83.43	2000	NE/SW
Overall classification %		75.18	75.34		

Table 5. **Extraction results from Supervised/Unsupervised classification for QuickBird imagery**

Road Poly AOI	Total # of Pixels in AOI	Percentage Classified		Approximate Length (m)	Roadway Orientation
		Supervised	Unsupervised		
1	12233	63.73	67.39	9600	NW/SE
2	4025	64.07	66.36	3600	N-S
3	9698	70.11	73.12	4950	E-W
4	7075	82.65	79.59	2500	NSEW (Intersection)
5	1769	64.10	67.32	2000	NE/SW
Overall classification %		69.42	71.35		

Table 6. **Extraction results from Membership/Nearest Neighbor (N.N.) classification for Ikonos imagery**

Road Poly AOI	Total # of Pixels in AOI	Percentage Classified		Approximate Length(m)	Roadway Orientation
		Membership	N.N		
1	118040	77.17	72.60	9600	NW/SE
2	42548	47.72	46.63	3600	N-S
3	78786	91.70	89.02	4950	E-W
4	56561	62.33	81.62	2500	NSEW (Intersection)
5	17327	69.07	62.54	2000	NE/SW
Overall classification %		73.70	74.28		

Table 7. **Extraction results from Membership/Nearest Neighbor classification for QuickBird Imagery**

Road Poly AOI	Total # of Pixels in AOI	Percentage Classified		Approximate Length(m)	Roadway Orientation
		Membership	N.N		
1	185224	69.51	62.08	9600	NW/SE
2	58212	62.98	66.25	3600	N-S
3	156275	99.89	79.15	4950	E-W
4	115271	94.44	82.55	2500	NSEW (Intersection)
5	25564	62.06	62.22	2000	NE/SW
Overall classification %		82.55	71.84		

CONCLUSIONS AND FUTURE WORK

In this study we have examined certain spectral and object-based classification methods for transportation feature extraction. Quantitative results from this study give a good overview for understanding the differences in their performances on the same data and also a comparison of the different data types used. We observed considerable variability in the performances of these methods, which use completely different approaches to perform the task of classifying and extracting road features from high-resolution multi-spectral imagery.

The specific Ikonos imagery utilized yielded somewhat better feature extraction results relative to the QuickBird image set, as processed. This is perhaps due to adjustments required for the specific images on-hand when constructing an overall result. Some variation may be caused by inherent differences in the data sources, but as the number of images were limited and acquisition dates varied broadly (including in terms of range for respective sources) it is not possible to make definite statements. Ikonos supervised classification gave results similar to unsupervised classification; results were similar in the case of QuickBird imagery. In this case, we found that use of membership function/classifier in eCognition improved the quality of results when used with the QuickBird imagery. Our Ikonos image set seemed reasonably effective for feature extraction; slightly more varied results and a longer image processing/analysis time were observed for the QuickBird image set. In general, East-West aligned roads tended to produce better extraction results compared with North-South roads. That difference may be a result of relative satellite travel path, position and angle at which images were acquired at a given time.

A secondary objective in this study was to examine possibilities for fine-scale feature extraction and differentiation with respect to road surfaces, e.g. urban streets and highway features. This proved extremely difficult using either a spectral approach or via object-based classification. Although different detection characteristics were observed with the respective methods, in both cases tall tree cover along road corridors often resulted in discontinuous clusters of pixels or segmented object blocks. Relaxing shape or color/texture-based constraints in eCognition allowed better merging to form continuous objects. However, unrelated features such as dry streambeds sometimes ended up aggregated within overall merged regions in our test cases. Increasing the level of detail for segmentation does not in itself assure that smaller features detected will be relevant.

An added form of classification (e.g. spectral-based analysis) can offer means to discriminate which regions are appropriate. Finer-scale object features with some potential to be roads were initially segmented. Image export of the classified region map intersected with spectral-based pixel classification marks otherwise uniform-class object areas with scattered values. Pixels denoting roads for instance now lie within formerly homogenous regions; such regions would stand out when statistically based sub-object (sub-region) classes are created in eCognition.

Unsupervised classification with roughly 20 classes might be compared to the more intense 100-class version to determine appropriate processing formats. Examining regions with larger variability in land cover type would be of interest in building higher-detail classification levels. Differentiation of road features into street-related subclasses would focus on extracted object feature characteristics such as connectivity and length/width ratio.

It is clear that these high-resolution image sources offer an advantage over traditional data sources such as Landsat in performing tasks involving finer-scale feature extraction, which are not feasible with less detailed imagery. These observations are supported by a study conducted on Landsat imagery for the counties in the coastal Mississippi (Repaka, 2004). Although extraction methods performed on different image types yielded slightly different accuracy, there is little variability with respect to their results. Recognition accuracy realized for the two data sets is noteworthy considering the complexity of the task. However, there was considerable variability in eCognition's object classification based on combined choice of segmentation and texture methods. Image segmentation via "split and merge" was the most resource-intensive; the "combined feature set" option appears to offer some advantage in classification, however, and this may be explored in future study. Detailed review of patterns of classification error and cases where mistakes were made should improve overall imaging results.

Initial attempts were made to resample the multispectral images to higher resolution using panchromatic band imagery, so as to evaluate the classification process after pan-sharpening. Prior to pixel-based classification, multispectral imagery might be resampled to a higher resolution (e.g. panchromatic scale) by applying a wavelet-based pan-sharpening process. Efforts were hindered due to large image file sizes, which available software had difficulty in processing. In some cases segmentation problems arose; eCognition can generate a greater number of objects than can be referenced via Windows 32-bit address space. One workaround is to focus on smaller subsets.

In examining varied image sources for potential to detect and delineate features (such as roads/highways and railroads, varied land cover types etc.) it is often useful to examine spatial accuracy. Obtaining a sufficient set of regional Ground Control Points and road centerlines via a Real-Time Kinematic GPS unit would allow geo-referencing (geocoding) of the mosaic to permit detailed comparison of feature-classified results with ground truth information. In a separate study we have conducted an initial Assessment of DigitalGlobe Quickbird Imagery for an area near the Mississippi coast. Systematic errors related to instrument accuracy and differences due to sensor position and/or orientation is something that can be estimated, and potentially eliminated or substantially reduced.

REFERENCES

Asano, T., Chen, D.Z., Katoh, N., and Tokuyama, T. (1996). Polynomial-time solutions to image segmentation. Proceedings of the 7th Annual SIAM-ACM Conference on Discrete Algorithms, January 1996, pp. 104-113.

Duda, R. and Hart, P. (1973). *Pattern Classification and Scene Analysis*. Wiley & Sons, New York.

eCognition Professional 3.0 User Guide. <http://www.definiens-imaging.com/down/index.htm>

Greenwalt, C.R., and Schultz, M.E. (1968). Principles of Error Theory and Cartographic Applications. ACIC Technical Report No. 96, Aeronautical Chart and Information Center, U.S. Air Force, St. Louis, Missouri.

Hodgson, M.E. (1988). Reducing the computational requirements of the minimum-distance classifier. *Remote Sensing of Environment*, vol. 24.

Hofmann, Peter (2000). Detecting buildings and roads from IKONOS data using additional elevation information. Definiens Imaging GmbH, Germany.

Jensen, J.R. and Cowen, D.C. (1999). Remote Sensing of Urban/Suburban Infrastructure and Socio-Economic Attributes. *Photogrammetric Engineering and Remote Sensing*, 65: 611-624.

Landgrebe, D.A. (1999). Some fundamentals and methods for hyperspectral image data analysis. SPIE International Symposium on Biomedical Optics (Photonics West), San Jose, California, January 23-29, 1999. In *Proceedings of the SPIE*, vol. 3603.

Leiss, I. (1998). Landnutzungskartierung mit Hilfe multitemporaler Erdbeobachtungs – Satellitendaten. Dissertation. Geographisches Institut Zürich, Zürich, Switzerland.

Lillesand, T., and Kiefer, R. (2000). *Remote Sensing and Image Interpretation (4th ed.)*, Wiley & Sons, New York.

Mayer, Mechler, Schlindwein, and Wolke (1993). *Fuzzy Logic*. Addison Wesley.

Repaka, Sunil R. (forthcoming, 2004). Comparing Spectral and Object-Based Approaches for Extracting and Classifying Transportation Features Using High Resolution Multi-spectral Satellite Imagery. Unpublished master's thesis, Dept. of Civil Engineering, Mississippi State University, Mississippi.

Sester, M. (2000). Knowledge acquisition for automatic interpretation of spatial data. *International Journal of Geographical Information Science*, vol.14, no. 1, pp. 1-24.

Swerdlow, Joel L. (1998). Making Sense of the Millennium. *National Geographic*, 193 (1), January 1998, pp. 2-11.

Tilton, J. C. (1997). Utilization of New Data Products based on Hierarchical Segmentation of Remotely Sensed Imagery Data in Earth Science Enterprise Applications. National Aeronautics and Space Administration, Goddard Space Flight Center.

Wilkinson, G. and Burrill, A. (1991). Integration of GIS-Derived Spatial Data with Satellite Imagery for Improved Land Cover Mapping using an Expert System. In *Proceedings of EGIS91, European Conference on Geographical Information Systems*, pp. 1241-1249.

Willhauck, G. (2000). Comparison of object oriented classification techniques and standard image analysis for the use of change detection between SPOT multispectral satellite images and aerial photos. *ISPRS*, vol. XXXIII, Amsterdam, The Netherlands.