OBJECT-BASED URBAN ENVIRONMENT MAPPING WITH HIGH SPATIAL RESOLUTION IKONOS IMAGERY

Ruiliang Pu, Assistant Prof.
Department of Geography, University of South Florida
4202 E. Fowler Avenue, NES 107, Tampa, FL 33620 USA
Email: rpu@cas.usf.edu

Shawn Landry, Interim Director
Florida Center for Community Design and Research
4202 E. Fowler Avenue, HMS 301, Tampa, FL 33620 USA
Email: landry@arch.usf.edu

Qian Yu, Assistant Prof.
Department of Geosciences, University of Massachusetts
611 N Pleasant St., Amherst, MA 01003 USA
Email: qyu@geo.umass.edu

ABSTRACT

Advances in remote sensing such as increasing spatial/spectral resolutions have strengthened its ability of urban environmental analysis. Unfortunately, high spatial resolution imagery also increases internal variability in land-cover / use unit, which can cause consequent classification result showing a "salt and pepper" effect. To overcome this problem, region-based classification has been used. In such a classification, image-object (IO) is used rather than pixel as a classification unit. Using IKONOS high spatial resolution imagery, in this study, we propose to test whether the IO technique can significantly improve classification accuracy when applied to urban environmental mapping with high spatial resolution imagery compared to pixel-based method in Tampa Bay, FL, USA. We further evaluate the performance of artificial neural network (ANN) and Maximum Likelihood Classifier (MLC) in urban environmental classification with high resolution data and test the effect of number of extracted IO features on urban classification accuracy. Experimental results indicate that, in this particular study, a statistically significant difference of classification accuracy is proved between using pixel-based and IO-based data; ANN outperforms MLC when both using 9 features pixel-based data; and using more features (30 vs. 9 features) can increase IO classification accuracy, but seems not statistically significant at the 0.9 confidence level at this study.

INTRODUCTION

Timely and accurate information on the status and trends of urban ecosystems and biophysical parameters is critical to developing strategies for sustainable development and to improving urban residential environment and living quality (Yang et al., 2003; Song, 2005). Therefore, developing techniques and enhancing the ability for monitoring and mapping urban land use / land cover (LULC) are important for city modeling and planning. One of the most common applications of remote sensing images is the extraction of LULC information for digital image base maps. Such information is useful to city governments to seek better planning and management approaches to deal with numerous problems associated with increasing urbanization (e.g., LULC change / attribution and storm water planning / mitigation) (Shackelford and Davis, 2003). During the last decade, satellite remote sensing has advanced in increasing spatial resolution (e.g., IKONOS multispectral images at 4-m resolution and panchromatic band at 1-m resolution) and spectral resolution (e.g., Hyperion hyperspectral sensor at 10 nm spectral resolution). High spatial resolution commercial satellite imagery (e.g., IKONOS) has been shown to be a cost-effective alternative to aerial photography for generating digital image base maps (Davis and Wang, 2003).

With the advent of high spatial resolution satellite sensors, new challenges did arise for automatic classification of LULC on intra-urban areas. Traditional pixel-based classifiers are quite limited for classification of images acquired with those sensor systems, which can capture details of very heterogeneous urban scenes with a large internal class variation. The LULC information extracted by the multispectral pixel-based classification proves to have more complexity owing to the internal variation increase in the land cover unit (Carleer and Wolff, 2006; Kux

and Pinho, 2006). With the increase in spatial resolution, LULC classes tend to be represented by spatial units of heterogeneous spectral reflectance characteristics and their statistical separability is reduced when the traditional pixel-based classification approaches are used. Consequently, classification accuracy is reduced and the results usually show a salt-and-paper effect due to individual pixels classified differently from their neighbors. Previous studies have proved that a decrease in LULC classification accuracy is likely to occur as the spatial resolution of the image is improved, when other sensor characteristics are kept unchanged (Townshend and Justice, 1981; Latty et al., 1985, Martin et al., 1988; Gong and Howarth, 1990; Treitz and Howarth, 2000). Classification accuracy is particularly problematic in urban environments that typically consist of mosaics of small features made up of materials with different physical properties (Mathieu et al., 2007). To overcome this problem region-based or image-object (IO)-based classification can be used. Image segmentation, before classification, produces regions that are more homogeneous in themselves than with nearby regions and represent discrete objects or areas in the image (Carleer and Wolff, 2006). Such a region or IO is then used rather than a pixel as a unit for classification.

Reasons that IO-based classification strategy can potentially improve classification accuracy compared to pixel-based classification may include: (1) partitioning an image into IOs is similar to the way humans conceptually organize the landscape to comprehend it (Hay and Castilla, 2006); (2) besides spectral features, IOs also enable the use of texture and contextual (relations with other objects) features and some shape/geometric features (e.g., form, size and geomorphology) (Hay and Castilla, 2006; Yu et al., 2006); and (3) the objects of interest to be extracted from a certain scene can be associated to different abstraction levels (i.e., different scales) and these levels can be represented in an analysis system (Kux and Pinho, 2006). Many previous researchers have demonstrated such advantages of IO-based classification (Ton et al., 1991; Johnsson, 1994; Hill, 1999; Herold et al., 2003; Carleer and Wolff, 2006; Kong et al., 2006; Kux and Pinho, 2006; Marchesi et al., 2006; Yu et al., 2006; Mathieu et al., 2007).

After reviewing the literature it is apparent that more work is needed to evaluate object-based classification approaches with high resolution imagery, especially the efficiency of such approaches on urban environmental land cover classification. Using IKONOS high spatial resolution imagery, we propose to further test the ability of the IO technique to significantly improve classification accuracy compared to pixel-based methods when applied to urban detailed land cover mapping in Tampa Bay, FL, USA. Therefore, the substantial objectives of this analysis consist of (1) testing whether the IO technique can significantly improve classification accuracy when applied to urban environmental classification with high spatial resolution imagery compared to pixel-based method; (2) comparing the performance of ANN and MLC in urban environmental classification with high resolution data; and (3) evaluating the effect of number of extracted IO features on urban classification accuracy. Some limitations of using the object-based classification approach will also be discussed.

STUDY AREA AND DATA SETS

Study Area

The study area is a 100 km^2 area within the City of Tampa. Tampa is the largest city on the west coast of Florida consisting of approximately 285 km^2. During the last three decades, the city had experienced continuous growth in population and expansion in extent. The population is increasing and is currently estimated at approximately 335,000 people (www.tampagov.net accessed on Nov. 26, 2007). The city is located at approximately 28° N and 82° W (Figure 1). Tampa's climate is considered subtropical with an annual average temperature of 22ºC. The city has two distinct seasons related to annual precipitation, wet (June-October) and dry (November-May) that can be punctuated by periods of wet weather during December- February (Campbell and Landry, 1999). Historically, the natural plant communities of the Tampa Bay region included pine flatwoods, cypress domes, hardwood hammocks, high pine forests, freshwater marshes, and mangrove forests. Based on the City of Tampa Urban Ecological Analysis (Andreu et al., 2008), current tree canopy cover is 28.1% with the ten most common tree species accounting for approximately 84% of all stems. Other green vegetation areas are occupied by shrubs, grass/lawns in varying sizes, golf courses, and crops. Man-made materials for buildings and roofs in the city are concrete, metal plate and brick tile, etc. Various impervious road surfaces are covered by asphalt, concrete and rail track.

Data Sets

IKONOS imagery. High resolution IKONOS satellite imagery (GeoEye, Inc., USA) was acquired for the study area on April 6, 2006. Georeferenced 1-m resolution panchromatic (Pan, 526 – 929 nm) and 4-m resolution multispectral (MS) images (four bands, blue (445 – 516 nm), green (506 – 595 nm), red (632 – 698 nm), and NIR

(757 – 853 nm)) were acquired. The IKONOS imagery including Pan and MS images is major data set for this object-based classification analysis.

Digital aerial photographs. A set of true color digital aerial photographs was taken in January, 2006 (SWFWMD, 2006). The aerial photographs included three visible bands (blue, green and red) at 0.3-m spatial resolution. They were used as reference to define training, test and validation areas / samples.

Ground plot measurements. Measurements from approximately 60 ground plots in the study area, including LULC cover type and percentage, plant species, dbh (diameter at breast height), crown width, etc were provided from the project of The City of Tampa Ecological Analysis 2006-2007 (Andreu et al., 2008). Ground plot measurements were used as reference for determining training and test areas.

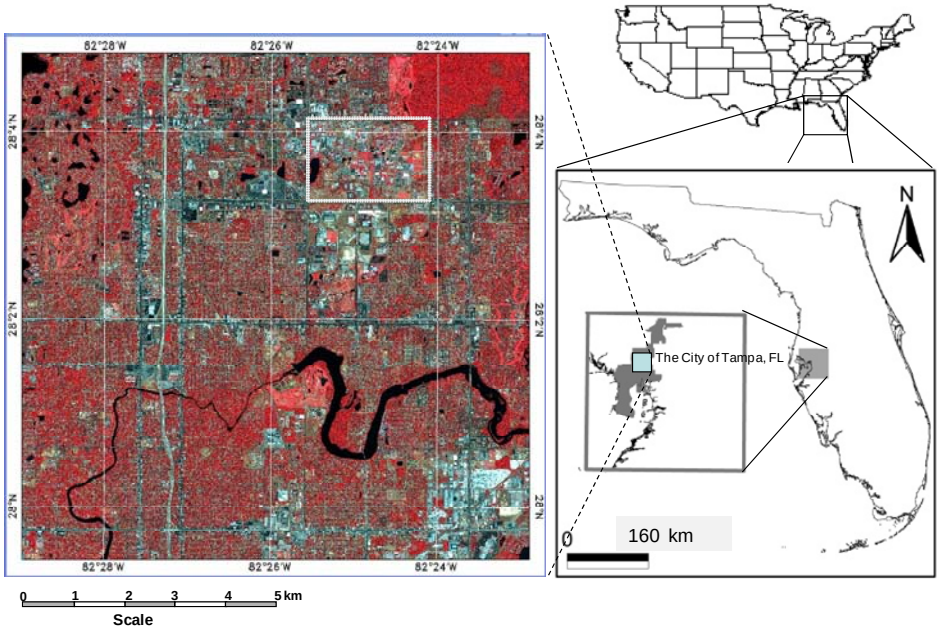

Figure 1. Location map of the study area.

METHODOLOGY

Figure 2 presents a flowchart of this analysis procedure of the urban environmental classification using high resolution IKONOS imagery with both pixel-based and IO-based classification strategies. In this analysis, after IKONOS imagery data were preprocessed, including radiometric correction and calibration and data fusion, nine basic pixel-based image layers were prepared, comprised by 4 pan-sharpening (PS) bands, 3 Hue-Intensity-Saturation (HIS) indices, 1 soil adjusted vegetation index (SAVI) and 1 texture image (created from PS band 4 with co-occurrence and homogeneity parameters from ENVI (ITT, 2006)). The 9 pixel-based image layers were then used for testing pixel-based classification approach and creating image objects (IOs) for testing IO-based urban surface component classification. After IOs were generated from the 9-image layers (themselves form 9 features), 27 more features were extracted from IOs used for object-based classification analysis.

Image Segmentation

The object-based image analysis software used in this research was Definiens eCognition 5.0. eCognition uses a multi-resolution segmentation approach which is a bottom-up region merging technique starting with one-pixel objects. In numerous iterative steps, smaller image objects are merged into larger ones (Baatz et al., 2004). The merging criterion minimizes the average heterogeneity of image objects weighted by their size in pixels (Baatz and Schape, 2000; Benz et al., 2004). Quantitatively, the definition of heterogeneity takes into account both spectral variance and geometry of the objects (Yu et at., 2006). The outcome of the segmentation algorithm is controlled by a scale factor and a heterogeneity criterion. The heterogeneity criterion controls the merging decision process, and is computed using spectral layers (e.g. multispectral images) or non-spectral layers (e.g. thematic data such as elevation) (Mathieu et al., 2007). The heterogeneity criterion includes two mutually exclusive properties: color and

shape. Color refers to the spectral homogeneity whereas shape considers the geometric/geomorphologic characteristics of the objects. Shape is further divided into two equally exclusive properties: smoothness and compactness (Baatz et al., 2004).

The optimum segmentation parameters depend on the scale and nature of the features to be detected. These can be determined using a systematic trial and error approach validated by the visual inspection of the quality of the output image objects, i.e. how well the image objects matched feature boundaries in the image (Mathieu et al., 2007) for a particular application. Once an appropriate scale factor was identified, the color and shape criterion are modified to refine the shape of the image objects. Most previous studies had found that more meaningful objects are extracted with a higher weight for the color criterion (e.g., Laliberte et al., 2004; Mathieu et al., 2007). In this application with input of nine data layers (4 PS bands, 3 HIS indices, 1 SAVI and 1 texture image) with each layer's pixel values rescaled to [0, 10,000] for urban environmental LC mapping, the color criterion was assigned with a weight of 0.7, whereas the shape received the remaining weight of 0.3. Further, the compactness was assigned with a weight of 0.3 and smoothness with remaining weight of 0.7. After visually checking their matching degree of image objects to feature boundaries of LC types in the study area, we employed the image objects (IOs) created with a scale of 70 in following IO-based classification analysis.

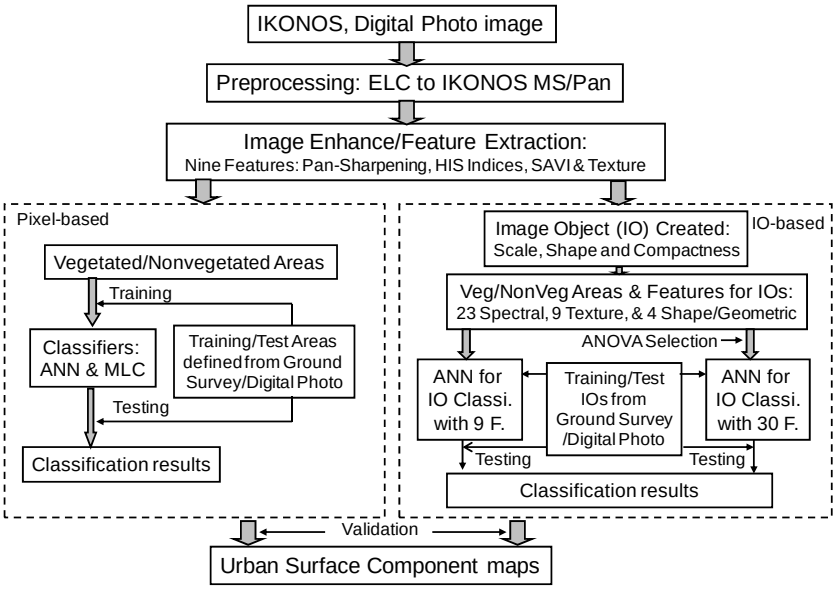

Figure 2. Flowchart of the analysis procedure of the urban land cover mapping, consisting of pixel-based and IO-based classification strategies. In the figure, Veg/ NonVeg areas mean vegetated and non-vegetated areas; Classi. and F. are classification and features, respectively. For all other abbreviations, see their full names in the text.

Feature Extraction and Selection

Besides the nine features used for creating IOs, 27 more feature variables were generated from each IO. A total of 36 features (23 spectral features, 9 texture features and 4 shape/geometric features) were generated for this object-based classification analysis and listed in Table 1. The consideration and determination of these features were selected based on previous studies by Haralick et al., (1973), Carleer and Wolff, (2006), Kong et al., (2006), and Yu et al., (2006), etc.

To reduce redundancy, it is necessary to select a subset of features from the total 36 spectral variables prior to object-based classification of urban enviornmental LC classes. In this analysis, a one-way ANOVA analysis was performed. This was done based on greater feature separability between any two LC classes (paired-class) of the 14 LC classes (at Level III in Table 2) using the SPSS statistical package (www.spss.com 2007). For any paired-class from the 14 classes, all training and test IOs for the paired-class were used to conduct the ANOVA analysis across the 36 feature variables (Table 1). Then based on the degree of feature separability of each feature variable between the paired-class, a statistical frequency was calculated at probability levels $p \leq 0.01$ and $p \leq 0.05$ for each feature

variable. For this analysis, a maximum frequency at either p≤0.01 or p≤0.05 is 91 (because of ${}_{14}C_2 = \frac{14 \times 13}{2} = 91$, i.e., total possible combinations of picking any 2 classes from a total of 14 classes).

Classification Strategies

To improve the urban surface component classification accuracy, a hierarchical classification system (Table 2) was adopted for the study. One advantage of this methodology is that it matches the logical structure of most LULC classification schemes utilized by previous researchers (e.g., Townsend and Walsh, 2001; Pu et al., 2008). The hierarchical classification scheme was constructed in three levels (Table 2). At Level 1, vegetated and non-vegetated are separated using a SAVI threshold of 0.19. SAVI values greater than 0.19 were assigned as vegetation. With this threshold, the two sub-areas could be clearly separated within the study area. The vegetated and non-vegetated areas are further subdivided into five vegetated and four non-vegetated classes at Level II. The five vegetated types include Broad-leaf trees (BT), Needle-leaf trees (NT), Palm trees (PT), Shrub (Sh) and Grass/lawn (GL). The four non-vegetated classes include Building/roof (BR), impervious area (IA), Sand/soil (SS) and Water (Wa). Level II classes' descriptions were presented in Table 2. At Level III, only one vegetated class, Broad-leaf trees (BT), is further subdivided into two classes, High NIR reflectance (BT1) and Low NIR reflectance (BT2). This processing was consideration of significant difference of NIR reflectance between sand live oak and most other BT species due to differences of their biological characteristics (e.g., deciduous vs. even green). For four non-vegetated classes, Building/roof (BR) and Impervious area (IA) are future subdivided into High, Medium and Low albedo (BR1, BR2, and BR3; IA1, IA2, and IA3), respectively. Classification operations were carried out at Level III separately for each Level I area (vegetated/non-vegetated) using pixel-based features or IO-based features with ANN and MLC algorithms (Figure 2). The final classification results at Level II were obtained through merging BT1 and BT2 in BT, BR1 through BR3 into BR, and IA1 through IA3 into IA. Some accuracy indices are calculated at Level II.

Two supervised classification algorithms were employed for the urban environmental LC classification: A non-parametric artificial neural network (ANN) and a parametric Maximum Likelihood Classifier (MLC). In this analysis, a feed-forward ANN algorithm was used for classifying the 14 classes at Level III. The network training mechanism is an error-propagation algorithm (Rumelhart et al. 1986; Pao, 1989). An MLC classifier was also used to classify the 14 classes with input of 9 pixel-based feature variables to compare with the pixel-based classified results by ANN. The ENVI software was used (ITT, 2006) for MLC classification. MLC is a standard classifier that has been extensively used in LULC classification practice (e.g., Gong et al., 1997; Fung et al., 1998; van Aardt & Wynne, 2001).

Assessment and Validation

The training and test samples were determined from pixel-based and object-based image data by referencing 0.3-m resolution digital aerial photographs and available ground plot measurements. The pixel-based training/test samples were relatively easily defined whereas the IO-based training/test samples were not easy and were listed in Table 2, corresponding the 14 classes at Level III. About 2/3 of the samples were used for training and about 1/3 of the samples were used as test samples. This procedure was repeated three times (runs) to obtain three different sets of test samples (but training sets with a part overlaid between any two training sets). Finally, an average accuracy (AA) or overall accuracy (OAA) and Kappa index were calculated from a confusion matrix constructed with the test samples using ANN and MLC at Level II. To validate urban environmental classification results mapped with IKONOS image data, a system sampling approach with a 500-m grid was applied to both 0.3-m resolution aerial photographs and urban LC classification maps created with either pixel-based or IO-based IKONOS imagery. A total of 441 points (cross points of the 500-m grids) each representing about 4 m^2 were visually identified and interpreted from both the digital aerial photographs and resultant urban LC maps. An OAA value and Kappa index can be calculated from the 441 paired-points and used for assessing the accuracy of urban environmental LC classification maps produced using either pixel-based or object-based IKONOS image data with either ANN or MLC algorithm.

RESULTS AND ANALYSIS

ANOVA

After the nine input data layers were input to eCognition with scale = 70 and other parameters, according to the definitions for spectral variables listed in Table 1, a total of 36 feature variables were extracted. A one-way ANOVA

analysis was first performed for all the extracted feature variables from which a subset of feature variables was selected. Figure 3 shows the frequency distribution of ANOVA of all the feature variables separating any paired-class (from left to right for greatest to least frequency). By analyzing the frequency variation across all 36 variables, these variables can be divided into two groups with a frequency threshold of 60 at the significant level of 0.01. Hence, by the figure, a total of 30 feature variables were selected (SAVI through GLCMD from left to right).

Among the 30 feature variables, all 22 spectral features were selected, which imply that those spectral features make a substantial contribution to separating most of the 14 classes. The remaining 8 feature variables consist of 6 textural features and 2 shape/geometric features. From the selected spectral features, it is obvious that the ability to separate any paired-class mainly relies on the variation of pixel itself spectral information extracted from and characterizing IOs.

Table 1. Image-object (IO) features used in this analysis.

Feature name	Description
Band1	Mean of pan-sharpening IKONOS band1(blur), input pixel layer.
Band2	Mean of pan-sharpening IKONOS band2 (green), input pixel layer.
Band3	Mean of pan-sharpening IKONOS band3 (red), input pixel layer.
Band4	Mean of pan-sharpening IKONOS band4 (NIR), input pixel layer.
Hue	Mean of Hue image processed from pan-sharpening IKONOS bands 3,2,1, input layer
Sat	Mean of Saturation image processed from pan-sharpening IKONOS bands 3,2,1, input layer
Val	Mean of Value (Intensity) image processed from pan-sharpening IKONOS bands 3,2,1, input layer
SAVI	Mean of soil adjusted vegetation index: 1.5(band4-band3)/(band4+band3+0.5), input layer
Tex	Mean of texture information of co-occurance homogeneity extracted from band4, input layer
SDB1	Standart deviation of Band1.
SDB2	Standart deviation of Band2.
SDB3	Standart deviation of Band3.
SDB4	Standart deviation of Band4.
SDH	Standart deviation of Hue.
SDS	Standart deviation of Sat.
SDV	Standart deviation of Val.
SDVI	Standart deviation of SAVI.
SDTX	Standart deviation of Tex.
Ratio1	Band1 mean divided by sum of band1 through band4 means.
Ratio2	Band2 mean divided by sum of band1 through band4 means.
Ratio3	Band3 mean divided by sum of band1 through band4 means.
Ratio4	Band4 mean divided by sum of band1 through band4 means.
Bright	Brightness, average of means of bands 1 through 4.
GLCMH	GLCM homogenity from band4, $\sum_{i,j=0}^{N-1} \frac{p_{i,j}}{1+(i-j)^2}$
GLCMCON	GLCM contrast from band4, $\sum_{i,j=0}^{N-1} p_{i,j}(i-j)^2$
GLCMD	GLCM dissimilarity from band4, $\sum_{i,j=0}^{N-1} p_{i,j} \mid i-j \mid$
GLCME	GLCM entropy from band4, $\sum_{i,j=0}^{N-1} p_{i,j}(-\ln p_{i,j})$
GLCMSD	GLCM standard deviation from band4, $\sigma_{i,j}^2 - \sum_{k=0}^{N-1} p_{i,j}(i, j-\mu_{i,j}) \quad where \ \mu_{i,j} = \sum_{k=0}^{N-1} p_{i,j} / N^2$
GLCMCOR	GLCM correlation from band4, $\sum_{i,j=0}^{N-1} p_{i,j}\left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}}\right]$
GLDVA	GLDV angular second moment from band4, $\sum_{k=0}^{N-1} V_k^2$
GLDVE	GLDV entropy from band4, $\sum_{k=0}^{N-1} V_k(-\ln V_k)$
GLDVC	GLDV contrast from band4, $\sum_{k=0}^{N-1} V_k K^2$
Compact	Compactness, the product of the length and the width of the corresponding object and divided by the number of its inner pixels.
CompactP	Compactness, the ratio of the area of a polygon to the area of a circle with the same perimeter.
ShapeI	Shape index, the border length of the IO divided by four times the square root of its area, i.e., smoothness.
NumP	Number of edges,the number of edge that form the polygon.

Note: i is the row number and j is the the column number, $V_{i,j}$ is the value in the cell i,j of the matrix, $p_{i,j}$ its the normalized value in the cell i,j, N is the number of rows or columns.

Pixel-Based Classification

After the 14 regions of interests (ROIs) were delineated from PS bands 4, 3 and 2 with a reference of 0.3-m digital aerial photographs and available ground plot measurements, the pixel-based classification results were produced using MLC and ANN algorithms. As the aforementioned, 6 vegetated classes and 8 non-vegetated classes were first separately classified with the two algorithms. The 14-class classification results were then merged to 9 classes at Level II for the two sets of results, respectively. Based on the high average accuracy (AA) and Kappa value calculated from test samples, a set of ideal structure parameters of ANN for pixel-based classification with nine features were adopted (learning rate (η) = 0.2, momentum coefficient (α) = 0.8 and number of nodes in a hidden layer (h1) = 12 or 10). Their results were presented in Figure 4 (top). After checking their 1:1 zoom-in maps, it is apparent that ANN has a higher ability to deal with the "salt-and-pepper" effect, a common phenomenon caused by using high spatial resolution and pixel-based image data with a traditional classification method MLC.

Table 3 summarizes the pixel-based classification results produced by using the two algorithms with test samples. In the table, by comparing all accuracy indices (AA, OAA, and Kappa) between the ANN and MLC, we can see that all accuracy indices except OAA are higher for ANN than for MLC. So, in general, the performance of ANN is better than that of MLC by comparing AA and Kappa values between the two algorithms.

Table 2. Unban land cover classes, definitions and number of training/test image-object (IO)s used in this analysis.

Level I	Level II			Level III		No of training / test IOs
	Name	Abbreviation	Description	Abbreviation	Description	
Vegetated area	Broad-leaf trees	BT	All broadleaf tree species canopies	BT1 BT2	High NIR refletance Low NIR refletance	171 169
	Needle-leaf trees	NT	All conifer tree species canopies	NT	-	82
	Palm trees	PT	All palm tree species canopies	PT	-	71
	Shrub	Sh	All shrub, bush, including some bush in wetland, pond & lake side.	Sh	-	86
	Glass/lawn	GL	All grassland, golf course and lawns	GL	-	90
Non-vegetated area	Building/roof	BR	All different size building or roofs with different materils coverd	BR1 BR2 BR3	High albedo Medium albedo Low albedo	181 147 143
	Impervious areas	IA	All impervious surface areas, e.g., road, park lots, etc.	IA1 IA2 IA3	High albedo Medium albedo Low albedo	135 148 143
	Sand/soil	SS	All bare sand/soil and/or very dry/dead glasslands.	SS	-	99
	Water	Wa	All different types of water bodies.	Wa	-	73
Total of training IOs						1738

Table 3. Accuracy of urban environmental classification using different classification units (pixel-based and IO based) and different algorithms (MLC and ANN) with 9 or 30 features (bands).

Algorithm	Number of features	Accuracy (%)				Kappa value			
		Pixel-based		IO-based		Pixel-based		IO-based	
		AA	OAA	AA	OAA	Kappa value	Variance	Kappa value	Variance
ANN	9	73.58	73.82	76.69	78.48	0.6956	0.000030	0.7371	0.000454
	30	n/a	n/a	80.51	81.19	n/a	n/a	0.7795	0.000415
MLC	9	68.25	74.01	n/a	n/a	0.6819	0.000002	n/a	n/a

Note: AA = average accuracy, OAA = overall average.

Object-Based Classification

To find the better ANN structure parameters for the IO-based classification with inputs of either 9 or 30 features, various combinations of learning rate (η), momentum coefficient (α) and number of nodes in a hidden layer (h1) were tested using the first training/test data set. For the input of 9 features, the better ANN structure parameters are: η = 0.8 or 0.7, α = 0.2 or 0.1, and h1 = 15 or 12. For the input of 30 features, the better ANN structure parameters are: η = 0.7, α = 0.2 or 0.1, and h1 = 20 or 25. The IO-based classification results with input of either 9 (Figure 4 bottom) or 30 features using ANN algorithm were created. By checking Figure 4 1:1 zoom-in maps, we can see that the classification result created with the IO-based image data better than that with the pixel-based, especially two vegetated classes: Broad-leaf trees (BT) and Needle-leaf trees (NT).

The results shown in Table 3 were calculated by averaging the three sets of results produced from test samples. The results in Tables 3 indicate that the results using more feature variables are better than those using less feature variables.

Validation

An OAA value and Kappa value were calculated from the 441 paired-points and used for validating the accuracy of urban detailed environmental classification maps produced using either pixel-based or object-based IKONOS image data with either ANN or MLC algorithm. Due to a lack of NT (Needle leaf trees) identified at grid points (possibly due to a low frequency of NT in the study area), accuracy AA was not calculated for the validation. The validation results indicate that the IO-based result (OAA = 76.64%, Kappa = 0.70713) is better than that created with pixel-based features by either ANN (72.79%, 0.66865) or MLC (67.80, 0.60254) in terms of both accuracy indices using the 9-feature image data. In comparing the pixel-based classifier with the same number of features, we can find the ANN outperforms MLC from the validation result. When we compare the effectiveness of different numbers of feature variables on the IO-based classification accuracy by ANN, mapping urban environmental LC with more features (30) is better that with less features (9). These validation results are basically consistent with those analyzed with accuracy indices derived from test samples from the previous section.

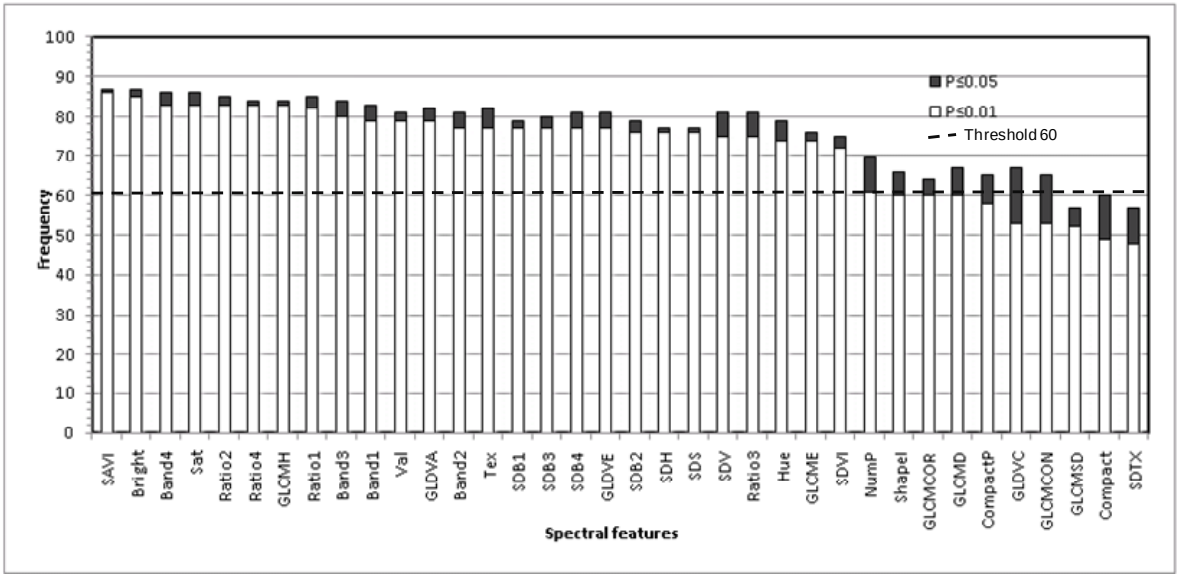

Figure 3. Frequency distribution of ANOVA of all 36 spectral variables for every two classes across the 14 classes. The possible maximum frequency is 91 and the bar represents number of pair-class between which a spectral variable difference is significant at 0.01 (empty bar) and 0.05 (filled bar) probability levels. The 1st 30 spectral variables (SAVI through GLCMD from left to right) were used as input to ANN for urban land cover classification.

Comparison

Based on the accuracy indices derived from test samples (averaged from the three sets of test samples) and validation results derived from the 441 grid points, the comparison analysis was conducted from the following three aspects. Firstly, in considering the two types of classification units (pixel-based and object-based units), we can compare the result produced from test samples of 9-feature pixel-based with that from IO-based image data by ANN algorithm and compared their corresponding validation results derived from the 441 grid points. Table 3 shows that all accuracy indices (AA, OAA, and Kappa) produced with IO-based data were consistently higher than those with pixel-based image data. Secondly, in considering two algorithms' performance with pixel-based image data only, from three accuracy indices in Table 3, ANN mostly outperforms MLC except OAA that was explained by MLC in a little bit of favor of classifying the 8 non-vegetated classes with relatively large test samples. Thirdly, when comparing the effects of different numbers of features on the urban environmental classification, we can see that all accuracy indices, from Table 3 with 30 IO-based features are higher than those with 9 IO-based features. To test whether these differences (between the two classification units, between the two algorithms, and between the two numbers of features) are statistically significant, Table 4 lists Z-statistics calculated with Kappa and corresponding variance derived from test samples. According to the Z-statistics in the table, the difference of results between ANN and MLC created with pixel-based image data is significant at 0.95 confidence level, and the difference of ANN results produced with different classification units (pixel-based and object-based) is significant at 0.90 confidence

level, whereas the difference of results created with different numbers of features (9 vs. 30) is not significant at 0.90 confidence level although the absolute accuracy indices derived from 30-feature are all higher than those from 9-feature. This implies that there still exists a lot redundant information among the 30 feature variables, which does not proportionally improve classification result as the number of feature variables involved increase.

DISCUSSION

In this study, our experimental results demonstrated that the IO-based urban component classification outperformed the pixel-based classification. When we used the same input feature variables (9 features), classifier (ANN) and same training/test data, the improvement of classification result with IO-based classification unit is statistically significant at 0.90 confidence level compared to pixel-based unit (AA increasing 3.1%, OAA increasing 4.7% and Kappa increasing 0.04; from the validation result, OAA increasing 3.8% and kappa increasing 0.04). Unlike pixel-based techniques which only use the layer pixel values, the IO-based techniques can also use shape and contextual information of a scene covering the study area. Therefore, if we consider more features (30 features), including some textural and shape/geometric features unique and only available to IOs, an improvement of classification result can be achieved compared to pixel-based and fewer features as input (AA increasing 6.9%, OAA increasing 7.4% and Kappa increasing 0.06; from the validation result, OAA increasing 10.7% and kappa increasing 0.12). Some features derived from IOs, in this analysis, were efficient to identify grass/lawn from tree canopies and to separate building/roof and impervious surface area. Such a conclusion derived from our experiment, in fact, was not surprising when compared to the few previous studies conducted by other researchers. For example, Shachelford and Davis (2003), Yu et al. (2006), and Guo et al. (2007) used object-based high spatial resolution imagery (airborne or satellite image data) to obtain similar conclusions, including: improving the identification of urban surface components; increasing accuracy of vegetation community classification; and mapping more accurate oak tree mortality. In addition, when some researchers compared IO-based techniques with pixel-based techniques to change detection analysis, including deforestation analysis and other LULC change analysis, the analysis accuracies were improved significantly (Al-Khudhairy et al., 2005; Desclée et al., 2006; Zhou et al., 2008). Thus all those previous studies and our experimental results did demonstrate the efficiency of using object-based classification techniques over that of using pixel-based approaches.

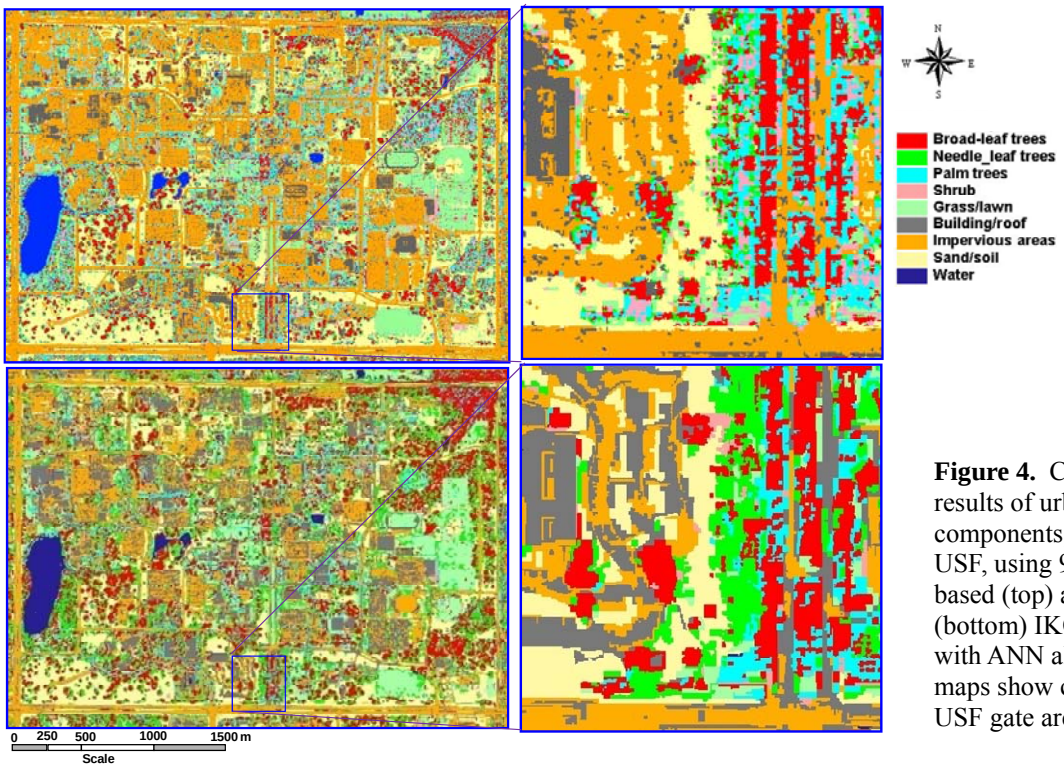

Figure 4. Classification results of urban surface components, showing part of USF, using 9 features of pixel-based (top) and IO-based (bottom) IKONOS imagery with ANN algorithm. Right maps show corresponding USF gate area at 1:1 scale.

We also tested the performance of ANN (nonparametric algorithm) and MLC (parametric algorithm) with the nine pixel-based features for classifying the 9 urban LC classes. The preliminary test result indicates that ANN outperforms MLC and the difference of accuracies between the two algorithms is statistically significant at 0.95 confidence level. In this analysis, the ANN is capable of handling spectral/textural confusions among Sand/soil (SS), Impervious surface area (IA) and Building/roof (BR) and among BT, NT and PT due to its ability to more efficiently deal with some possible nonparametric features than MLC. We know that only four PS band features are parametric and therefore applicable to using MLC, but 3 HIS transfer indices and other SAVI and textural features are probably not full parametric features. Compared to 0.3-m aerial photo interpretation, the mapped results for BR, IA, BT and PT produced with ANN apparently are better than those with MLC (maps not shown in the paper). A second reason to explain ANN generating better result is that it can efficiently make use of subtle spectral differences in the four PS band images due to its multilayer structure. These two reasons may help explain why the ANN algorithm is more capable of mapping the urban environmental LC classes than MLC. In fact, there have been many cases reported that ANN has produced better classification results with remote-sensing data when compared with traditional methods (e.g., linear discriminate analysis and MLC) (e.g., Gong et al., 1997; Erbek et al., 2004).

Using the ANN algorithm, the difference of results created with different numbers of features (9 vs. 30 features) is not significant at 0.90 confidence level although the absolute accuracy indices derived from 30-feature are all higher than those from 9-feature. This may be explained by two reasons. Apparently, firstly, the variances with both 9 and 30 features, relative to those with pixel-based, are large (e.g., 0.000030 and 0.000454 for the 9 features), which tends decreasing the significant level of difference of classification accuracy. Secondly, although we had conducted an ANOVA analysis to select a subset from all candidate feature variables, it is possible that there still exists much redundant information among the 30 feature variables, which leads to not proportionally improving classification result with number of feature variables. This is because the ANOVA statistically tests the ability of individual features to differentiate every two classes rather than measuring the relationship between any two feature variables. In addition, shadow might also influence the selection of features. In this study, four ratio features (i.e., Ratio1 through Ratio4 selected into the subset of 30 features), based on their definition, should weaken the effect of the shadow on selection of features and classification results. However, the effect of the shadow was evident from the mapped results.

Table 4. *Z*-statistic tests calculated from Kappa-variance of classification results of test samples, generated with different classification units (pixel-based and IO-based) and using different algorithms (ANN and MLC) with 9 or 30 features.

Scheme	Pixel-based	IO-based
Z(ANN(30 features vs. 9 features)	n/a	1.4383
Z(ANN(pixel vs. IO each with 9 features))	1.8864*	
Z(ANN vs. MLC each with 9 features)	2.4218**	n/a

Note: *--difference between classification accuracies by two sets of input features is significant at 0.90 confidence level.

**--difference between classification accuracies by two sets of input features is significant at 0.95 confidence level.

$Z = \frac{|k_1 - k_2|}{\sqrt{v_1 + v_2}}$ where, k1 and k2 are kappa vappa values of corresponding input feature 1 and input feature 2, respectively, and v1 and v2 are corresponding variances.

CONCLUSIONS

The experimental results indicate that in classifying urban surface components with the high spatial resolution IKONOS pan-sharpening data, the difference of accuracies produced with pixel-based and IO-based image data is statistically significant. This is because object-based input features eliminate the effect of the "salt-and-pepper" on classification through image segmentation to create IOs, using features extracted from IOs including spectral, textural/contextual and shape/geometric features. In evaluating the performance of the two algorithms, ANN outperformed MLC when using 9 features pixel-based image data, possibly because ANN can handle nonparametric image features such as SAVI, and textural features. And in this particular test, using more features (30 vs. 9 features) could increase IO classification accuracy, but was not statistically significant at 0.9 confidence level. This might be attributed to redundant information existing among the selected features and possibly the impact of the shadow. After this analysis, we can suggest some issues related to image segmentation worthy of greater attention, including: how to select the appropriate criteria to create ideal IOs to achieve accuracy for a particular application;

how to evaluate whether edge and shape of IOs overlap (coincide) the boundaries landscape (LULC type/patch) through justifying scales; and operationally what relationship exists between IOs and ecological basis. These issues should be continuously considered by us in developing object-based techniques with high spatial resolution imagery in the future.

ACKNOWLEDGEMENTS

This work was partially supported by University of South Florida (USF) under the New Researcher Grant (Grant #: 18300). USF graduate student, Mr. John Kunzer is greatly appreciated for his help in the field trip.

REFERENCES

Al-Khudhairy, D.H.A., I. Caravaggi, and S. Giada, 2005. Structural damage assessments from ikonos data using change detection, object-oriented segmentation, and classification techniques, *Photogrammetric Engineering and Remote Sensing*, 71(7), 825-837.

Andreu, M.G., M.H. Friedman, S.M. Landry, & R.J.Northrop, 2008. *City of Tampa Urban Ecological Analysis 2006-2007*, Final Report to the City of Tampa, April 24, 2008. City of Tampa, Florida.

Baatz, M., & A. Schape, 2000. Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation, *Angewandte Geographische Informations- Verarbeitung XII* (J. Strobl, T. Blaschke, and G. Griesebner, editors), Wichmann Verlag, Karlsruhe, pp. 12–23.

Baatz, M., U.Benz, S. Dehghani, M. Heynen, A. Höltje, P. Hofmann, I. Lingenfelder, M. Mimler, M. Sohlbach, M. Weber, & G. Willhauck, 2004. *eCognition Professional User Guide*, Definiens Imaging GmbH: München, Germany.

Benz, U.C., P. Hofmann, G. Willhauck, I. Lingenfelder, M. Heynen, 2004. Multi-resolution, objectoriented fuzzy analysis of remote sensing data for GIS-ready information, *ISPRS Journal of Photogrammetry and Remote Sensing*, 58, 239–258.

Campbell, K., & S. Landry, 1999. *City of Tampa Urban Ecological Analysis*, The Florida Center for Community Design and Research, University of South Florida.

Carleer, A.P., & E. Wolff, 2006. Region-based classification potential for land-cover classification with very high spatial resolution satellite data, In *Proceedings of 1st International Conference on Obgect-based Image Analysis (OBIA 2006),* Salzburg University, Austria, July 4-5, 2006. Vol. XXXVI, ISSN 1682-1777.

Davis, C.H., & X. Wang, 2003. Planimetric accuracy of Ikonos 1m panchromatic orthoimage products and their utility for local government GIS basemap applications, *International Journal of Remote Sensing*, 24(22), 4267–4288.

Desclée, B., P. Bogaert, & P. Defourny, 2006. Forest change detection by statistical object-based method, *Remote Sensing of Environment*, 102, 1-11.

Erbek, F.S., C. ÖZkan, & M. Taberner, 2004. Comparison of maximum likelihood classification method with supervised artificial neural network algorithms for land use activities,. *International Journal of Remote Sensing*, 25(9), 1733-1748.

Fung, T., F.Y. Ma, & W.L. Siu, 1998. Hyperspectral data analysis for subtropical tree species recognition, *Symposium Proceedings, IGARSS'98*, Sensing and Managing the Environment, Vol. 3 (pp. 1298– 1300). New York, NY, USA: IEEE International Geoscience and Remote Sensing, Seattle, WA, USA, July 6– 10, 1998.

Gong, P., & P.J. Howarth, 1990. The use of structural information for improving spatial resolution and classification accuracy land-cover classification accuracies at the rural–urban fringe, *Photogrammetric Engineering and Remote Sensing*, 56, 67–73.

Gong, P., R. Pu, & B. Yu, 1997. Conifer species recognition: An exploratory analysis of in situ hyperspectral data, *Remote Sensing of Environment*, 62, 189-200.

Guo, Q., M. Kelly, P. Gong, & D. Liu, 2007. An object-based classification approach in mapping tree mortality using high spatial resolution imagery, *GIScience & Remote Sensing*, 44(1), 24–47.

Haralick, R. M., K. Shanmugam, & I. Dinstein, 1973. Textural features for image classification, *IEEE Transactions on Geoscience and Remote Sensing*, 3 (6), 610-621.

Hay, G.J., & G. Castilla, 2006. Object-based image analysis: strengths, weaknesses, opportunities and threats (SWOT), In *Proceedings of 1st International Conference on Obgect-based Image Analysis (OBIA 2006),* Salzburg University, Austria, July 4-5, 2006. Vol. XXXVI, ISSN 1682-1777.

Herold, M., X.H. Liu, & K.C. Clarke, 2003. Spatial metrics and image texture for mapping urban land-use, *Photogrammetric Engineering & Remote Sensing*, 69(9), 991–1001.

Hill, R.A., 1999. Image segmentation for humid tropical forest classification in Landsat TM data, *International Journal of Remote Sensing*, 20(5), 1039–1044.

ITT, 2006. ENVI Version 4.3. Boulder, CO, ITT Industries, Inc.

Johnsson, K., 1994. Segment-based land-use classification from SPOT satellite data, *Photogrammetric Engineering & Remote Sensing,* 60(1), 47–53.

Kong, C., X. Kai, & C. Wu, 2006. Classification and extraction of urban land-use information from high-resolution image based on object multi-features, *Journal of China University of Geosciences,* 17(2), 151-157.

Kux, H.J.H. & C.M.D. Pinho, 2006. Object-oriented analysis of high-resolution satellite images for intra-urban land cover classification: Case study in São José Dos Campos, São Paulo State, Brazil, In *Proceedings of 1st International Conference on Obgect-based Image Analysis (OBIA 2006),* Salzburg University, Austria, July 4-5, 2006. Vol. XXXVI, ISSN 1682-1777.

Laliberte, A.S., A. Rango, K.M. Havstad, J.F. Paris, R.F. Beck, R. McNeely, A.L. Gonzalez, 2004. Object-oriented image analysis for mapping shrub encroachment from 1937 to 2003 in southern New Mexico, *Remote Sensing of Environment*, 93, 198–210.

Latty, R.S., R. Nelson, B. Markham, D. Williams, D. Toll, & J. Irons, 1985. Performance comparison between information extraction techniques using variable spatial resolution data, *Photogrammetric Engineering and Remote Sensing*, 51, 1159–1170.

Marchesi, A., R. Colombo, & P. Valentini, 2006. Application of high spatial resolution satellite imagery for urban environment mapping, In *Proceedings of 1st International Conference on Obgect-based Image Analysis (OBIA 2006),* Salzburg University, Austria, July 4-5, 2006. Vol. XXXVI, ISSN 1682-1777.

Martin, L.R.G., P.J. Howarth, G. Holder, 1988. Multispectral classification of land use at the rural–urban fringe using SPOT data, *Canadian Journal of Remote Sensing*, 14, 72–79.

Mathieu, R., J. Aryal, & A.K. Chong, 2007. Object-based classification of ikonos imagery for mapping large-scale vegetation communities in urban areas, *Sensors*, 7, 2860-2880.

Pao, Y. (1989). *Adaptive Pattern Recognition and Neural Networks*, New York, Addison and Wesley.

Pu, R., M. Kelly, G.L. Anderson, & P. Gong, 2008. Using CASI hyperspectral imagery to detect mortality and vegetation stress associated with a new hardwood forest disease, *Photogrammetric Engineering and Remote Sensing*, 74(1), 65-75.

Rumelhart, D.E., G.E. Hinton, & R.J. Williams, 1986. Learning internal representations by error propagation, In *Parallel Distributed Processing-Explorations in the Microstructure of Cognition*, Vol. 1 (Cambridge, MA: MIT Press), pp.318-362.

Shackelford, A.K. & C.H. Davis, 2003. A combined fuzzy pixel-based and object-based approach for classification of high-resolution multispectral data over urban areas, *IEEE Transactions on Geoscience and Remote Sensing*, 41(10), 2354–2363.

Song, C., 2005. Spectral mixture analysis for subpixel vegetation fractions in the urban environment: How to incorporate endmember variability?, *Remote Sensing of Environment*, 95, 248-263.

SWFWMD, 2006. 2006 Natural Color Imagery for Hillsborough County, Southwest Florida Water Management District. Brooksville, FL, Accessed January 2, 2007, http://www.swfwmd.state.fl.us/data/gis/doqq_search.htm.

Ton, J.C., J. Sticklen, & A.K. Jain, 1991. Knowledge-based segmentation of landsat images, *IEEE Transactions on Geoscience and Remote Sensing*, 29(2), 222–232.

Townsend, P.A. & S.J. Walsh, 2001. Remote sensing of forested wetlands: Application of multitemporal and multispectral satellite imagery to determine plant community composition and structure in southeastern USA, *Plant Ecology*, 157, 129-149.

Townshend, J., & Justice, J. (1981). Information extraction from remotely sensed data, a user view. *International Journal of Remote Sensing*, 2, 313–329.

Treitz, P. and P.J. Howarth, 2000. High spatial resolution remote sensing data for forest ecosystem classification: An examination of spatial scale, *Remote Sensing of Environment,* 72, 268–289.

van Aardt, J.A.N. and R.H. Wynne, 2001. Spectral separability among six southern tree species, *Photogrammetric Engineering and Remote Sensing*, 67(12), 1367-1375.

Yang, L., G. Xian, J.M. Klaver, B. Deal, 2003. Urban lad-cover change detection through sub-pixel imperviousness mapping using remotely sensed data, *Photogrammetric Engineering & Remote Sensing*, 69, 1003-1010.

Yu, Q., P. Gong, N. Clinton, G. Biging, M. Kelly, & D. Schirokauer, 2006. Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery, *Photogrammetric Engineering and Remote Sensing*, 72(7), 799-811.

Zhou, W., A. Troy, & M. Grove, 2008. Object-based land cover classification and change analysis in the Baltimore metropolitan area using multitemporal high resolution remote sensing data, *Sensors* , 8, 1613-1636.