Feature Extraction from High Resolution Satellite Imagery as an Input to the Development and Rapid Update of a METRANS Geographic Information System (GIS)

Final Report

METRANS Project AR 05-02

June 2011

Christopher Lee, Ph.D.

Principal Investigator

Department of Geography, College of Liberal Arts

California State University Long Beach

Long Beach, CA 90840

METRANS Transportation Center

USC | CSULB

Disclaimer

Disclosure

This project was funded in entirety under this contract to California Department of Transportation.

Acknowledgements

Funding for this project was made possible through the support of METRANS. Graduate student researcher Gregory Ziolkowski was instrumental in the completion of this project. The following collaborators at the Port of Los Angeles collaborated on study area selection and expressed interest in this study: Rich Riingen, David Malin, Stanley Yon and Daryl Raasch.

Abstract

This report describes an accuracy assessment of extracted features derived from three subsets of Quickbird pan-sharpened high resolution satellite image for the area of the Port of Los Angeles, CA. Visual Learning System's Feature Analyst and Definiens eCognition Developer software platforms were used for image classification and results were compared. A methodology was developed to determine the accuracy of the feature extraction algorithms using both software packages in the ESRI ArcView GIS interface, where the results of automatic and manual image classification methods were compared. The findings indicate that Definiens Ecognition Developer was overall more successful at identifying ground features in the areas where research was conducted.

Table of Contents

Table of Contents--continued

List of Figures

List of Figures--continued

List of Tables

I. Introduction

The purpose of this project was to obtain and process high-resolution satellite imagery for three test sites in the Los Angeles Port to demonstrate image classification and feature extraction techniques for development and rapid updating of a Geographic Information Systems (GIS). The data acquired, processed and presented provides a foundation that will directly contribute to an understanding of goods movement and trade transportation problems confronted by metropolitan areas.

The Los Angeles and Long Beach Port complex is the largest in the U.S. and the fifth largest in the world. Together their facilities handle approximately one third of all U.S. container traffic with 60.5% of that being imports and 39.7% exports (SCAG, 2005). Because of the two ports' huge economic impact on the United States economy, stakeholders have a special interest in making sure that geographic data about the port is as up to date and as accurate as possible in order to be able to respond quickly and efficiently to any emergency.

II. Background

The cities of Los Angeles and Long Beach as well as the Ports of Los Angeles and Long Beach each have their own Geographic Information Systems (GIS). Each has the need to update various spatial data layers in an increasingly timely fashion. This requires data and software driven techniques that are tested, validated and delivered to users in easily adoptable formats.

Additionally, experience shows that the Port stakeholders are sometimes reluctant to share GIS data with each other, as well as with outside interests, including researchers working on Port issues. In order for METRANS researchers to study spatial and temporal aspects of Port operations they need access to a GIS that will allow them to input and analyze project specific data in the context of the spatial framework of the Port infrastructure.

Creation of a GIS can be accomplished by a variety of methods ranging from manual digitizing of existing maps and air photos to extraction of feature information from data sources such as, for example and demonstrated in this paper high resolution satellite data. The complexity of land cover at the Ports coupled with the need for periodic and rapid updates point to a need for obtaining a means of rapidly creating land cover classification that takes advantage of automated feature extraction from imagery for change detection, while at the same time minimizing costs and maximizing accuracy.

Over the past five years new spatial data products such as high resolution optical imagery from orbital sensors and digital elevation data from airborne LiDAR have become readily available from commercial vendors. At the same time commercial vendors have developed new image processing and GIS software platforms that allow rapid and automated feature extraction and their subsequent integration into a GIS. The combination of detailed spatial data sets and software designed to extract data and process it into information offers exciting possibilities for the creation and rapid updating of information bases through GIS.

III. Project Goals

The main goal of this project is to demonstrate how data obtained through remote sensors can be processed, analyzed, and presented for the purposes of landscape analysis and extraction of intelligent knowledge through the combination of combination of automated computerized techniques. Automated software algorithms make it possible for accessing intelligent analytical extraction techniques from data obtained from complex remote sensing devices, including satellite gathered imagery and Lidar. These data can then be best integrated using GIS.

IV. Study Area

The Port Complex is a large geographical area consisting of a multitude of various landscape types. Remote sensing techniques have been shown to be successful in identifying various types of land cover, but for the purposes of this study three smaller subset areas within the complex were chosen for more detailed investigation. They are named as follows: Pier 400, Area 2, and Urban Interface. These areas of interest (AOI's) are a good representational sample of the Ports land cover at large. As such, Pier 400 is an AOI consisting of mainly Port related features, such as containers and is mostly homogenious. Area 2 is a mixed type of terrain which includes some of the features also found in the Pier 400 AOI, but intermixed with buildings, roads, vegetation, and open land. The Urban Interface AOI is an urban type of landscape, immediately adjacent to the Port and consisting mainly of residential structures, roads, and vegetation. The map below shows the extent of these three areas.

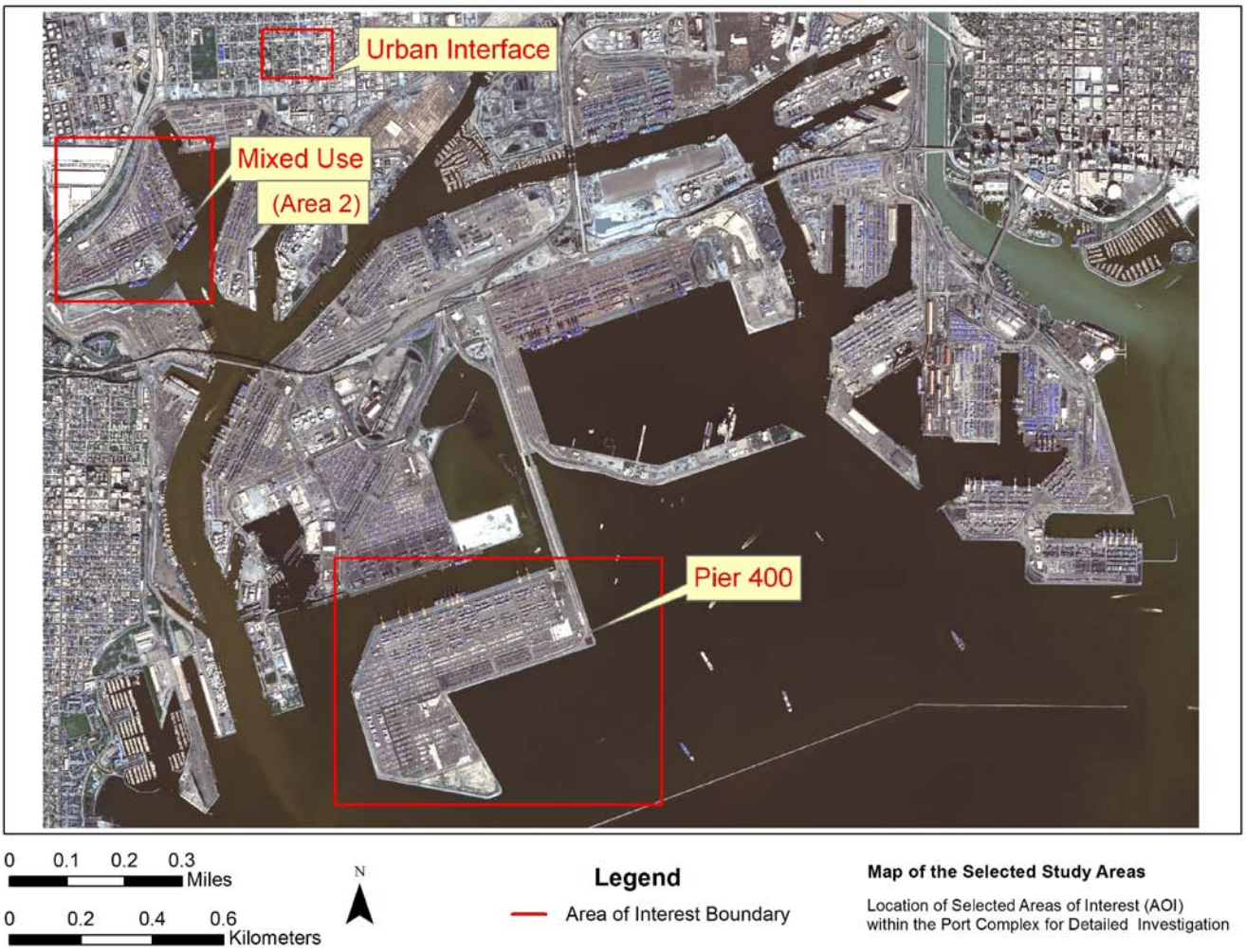


Figure 1: Map Showing the Three Areas of Interest (AOI's) Selected Within the Port Complex for Detailed Study in this Project. They are: Area 2 AOI, and Urban Interface AOI, Pier 400 AOI

V. Pan-sharpening Techniques

a. Pan-Sharpening

Satellite acquired remote sensing imagery is an increasingly important resource for military, commercial, and scientific applications. Current satellites are equipped with several charged coupled devices capable of capturing images at multiple resolutions and in different spectral wavelengths. High resolution images include panchromatic (Pan), where wavelengths from blue through near-infrared (IR) are captured as a single channel of data resulting in a detail rich grayscale image, in the case of Digital Globe QuickBird with spatial resolution of 0.6m. Also available are lower spatial resolutionmultispectral images that contain four channels of data representing red, green, blue (RGB), and IR at a spatial resolution of 2.4m. Pan-sharpening involves using one of many techniques to fuse the lower resolution multispectral imagery with the higher resolution panchromatic image to create a pan-sharpened multi-spectral image. Most of the analysis in this project utilized pan-sharpened Quickbird satellite imagery with a spatial resolution of 0.6 meters. Because pan-sharpening techniques were a basis for resolving detail in the satellite imagery used in this study, it is appropriate to discuss them, and errors that are involved in creating them, in more detail.

b. Pan-Sharpened Images

A pan sharpened image represents a sensor fusion between the multispectral and panchromatic images. It provides an end product in which high spectral resolution and high spatial resolution images are fused together in order to preserve the spectral information from multispectral data.

c. Sources of Error in the Pan-Sharpening Process

Within the interpolative pan-sharpening algorithm, error is often introduced into the new image as a result of the pan-sharpening process. These errors can include aliasing along and near edge boundaries, ringing artifacts near circular or oval objects, and the introduction of other artificial artifacts which can be caused by the bleeding or blurring of the multispectral information into the fused data.

d. *Pan-Sharpening Algorithms*

<u>Intensity-Hue-Saturation Transformation</u>

This is one of the most popular methodologies used to fuse data of different resolutions. According to this system, I (intensity), H (Hue), and S (Saturation) are the three light attributes which define human perception in relation to colors. In this study we used Smith's triangular method of IHS transformation for the purpose of extraction of data from images.

Principal Component Analysis Method

The PCA is useful in image compression, image enhancement, dimensionality reduction, and image fusion. In this method, the first principal component is replaced by the panchromatic image. The inverse PCA transform is computed to go back to the image domain. The PCA sharpening is sensitive to the area to be sharpened. The variance of the pixel values and the correlation among the various bands differ based on the land cover type. Since the PCA involves the computation of covariance matrices, the performance can vary with images having different correlations between the multispectral bands.

Brovey Transform Method

Ratio images are very useful in change detection. The Brovey transform method, named after its author, uses band ratios to sharpen the multispectral (MS) image. In this method, the MS image is normalized and each band of the fused MS image is obtained by multiplying the normalized MS bands with the panchromatic image. The Brovey transform provides excellent contrast in the image domain but affects the spectral characteristics a great deal. The Brovey sharpened image is not suitable for pixel-based classification as the pixel values are changed drastically.

e. Pan-Sharpening Accuracy Assessment

Image quality assessment is paramount in remote sensing applications, where the characterization of spectral and spatial variations within an image are needed. The primary use of quality metrics is to quantitatively measure the quality of an image that correlates with the true/perceived quality. Any good quality metric should be consistent, accurate, and repeatable in predicting the quality of any given image.
In a 2004 study, Vijayaraj performed a study using statistical changes in classification, and feature level changes to assess the quality of pan-sharpened images using a series of different pan-sharpening algorithms. The statistical measures used to assess the images were spectral mean square error, root mean square error, correlation coeffiecients, and histogram based metrics.

f. Error Metrics

Mean Square Error

"The mean square error (MSE) refers to the average of the sum of the squares of the error between two images. The MSE is defined as follows:

Equation 1: *MSE*

$$\sigma^2_{ms} = E[|u(m,n) - v(m,n)|^2],$$

(Jain 1998)

where u (m, n) and v (m, n) are two images of size m, x, n. The average least square error metric, which computed as shown in equation 2, is used as an approximation to the MSE.

Equation 2: *Average Least Square Error*

$$\sigma^2{}_{ls} = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}|\,u(m,n)-v(m,n)\,|^2 \ .$$

(Wald 1997)

The MSE quantifies the amount of difference in the energy of the signals. The mean square error metric has its limitation when used as a global measure of the image quality. However when used as a local measure, it is much more effective in predicting the image quality accurately. The root mean square error (RMSE) is the square root of the MSE. It quantifies the average amount of distortion in each pixel of the image. Both the MSE and RMSE give an account of the spectral fidelity of the image" (Vijayaraj 2004).

Correlation Coefficient

"The closeness between two images can be quantified in terms of the correlation function. The correlation coefficient can range from –1 to 1. A correlation coefficient value of 1 indicates that the two images are highly correlated, i.e., very similar to one another. A correlation coefficient of –1 indicates that the two images are exactly opposite to each other (Eskicioglu 1995). The correlation coefficient is computed from:

Equation 3: *Correlation Coefficient*

$$Corr(A/B) = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(A_{i,j}-\overline{A})(B_{i,j}-\overline{B})}{\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}(A_{i,j}-\overline{A})^2 \quad \sum_{i=1}^{M}\sum_{j=1}^{N}(B_{i,j}-\overline{B})^2}} \ .$$

(Vijayaraj 2004).

In the above formula, A and B are the two images between which the correlation coefficients was computed. The correlation between each band of the multispectral image before and after sharpening was also computed. The best spectral information is available in the multispectral image and hence the pan-sharpened image bands should have close correlation to the multispectral image band, because the spectral quality of the sharpened image is of satisfactory quality only if the correlation values are within a short range to each other. Another set of correlation coefficients was computed between each band of the multispectral image and the panchromatic image. Since the panchromatic image has better spatial resolution, the correlation between the sharpened image bands and the pan-image is expected to increase in comparison to the original multispectral image. An increase in the correlation indicates an increase in the spatial information compared to the multispectral image" (Vijayaraj 2004).

Relative Shift in Means

The mean value of pixels in a band is the central value of the distribution of pixels in that band. The relative shift in the mean value quantifies the changes in the histogram of the image due to processing (Parchardis 2000). The relative shift in the mean computation is defined as:

Equation 4: *Relative Shift in Means*

$$RM = \frac{Outputmean - OriginalMean}{Originalmean}\%.$$

(Eskicioglu 1995)

Change in Standard Deviation

"The standard deviation gives information about the spread of the histogram. The change in the standard deviation of the distribution is considered in addition to the shift in the mean. A combination of these two metrics quantifies changes in the shape of the histogram of each band. The histogram is spread over a large range of pixel DN values if the standard deviation is high. The relative shift in the mean and standard deviation help to visualize the change in the gray level distribution of the image bands" (Vijayaraj 2004).

Entropy and Increase in Information

"Entropy is defined as the amount of information contained in a signal. The entropy of the image can be evaluated as:

Equation 5: *Entropy*

$$H = -\sum_{i=1}^{d} p(d_i)\log_2(p(d_i)),$$

(Parchardis 2000)

where d is the number of gray levels possible and $p(d_i)$ is the probability of occurrence of a particular gray level. d_i is the probability of occurrence of a particular grey level.aP The increase in information metric is the difference in the entropy of each band of the original multispectral and the corresponding band in the sharpened image. The sharpened image should have more information compared to the original multispectral." (Vijayaraj 2004).

g. Pan-Results for Comparisons of Quickbird Imagery

The QuickBird multispectral and panchromatic data are at their original resolution of under three meters and one meter respectively. Following application of the sharpening technique, a new image was created from the original multispectral image.
The error metrics were computed between the sharpened image bands and the resampled multispectral bands. The MSE and RMSE values are shown in Table 1. DN pixel error compared to the resampled multispectral image.

Table 1. MSE and RMSE for QuickBird data: full resolution application

Method	**Band 1**		**Band 2**		**Band 3**	
	MSE	**RMSE**	**MSE**	**RMSE**	**MSE**	**RMSE**
IHS	50992.13	225.81	50926.57	225.67	48960.46	221.27
Brovey	44143.35	210.10	94426.64	307.28	47106.04	217.03
PCA	5528.85	74.35	16225.31	127.37	11584.29	107.63
Wavelet	606.36	24.62	606.37	24.62	603.38	24.56

Correlation between the different spectral bands and the spectral bands and panchromatic image were computed. The values are given in Tables 2 and 3. The wavelet-sharpened bands have spectral correlation similar to the multispectral bands indicating preservation of spectral information. The correlation values computed indicate that the PCA and Brovey sharpened bands have a good spatial information compared to the IHS and wavelet-sharpened bands.

Table 2 Correlation between spectral bands for QuickBird data: full resolution application

Method	**Band 1&2**	**Band 1&3**	**Band 2&3**
Original MS	0.9857	0.9523	0.9815
IHS	0.9693	0.9481	0.9627
Brovey	0.9251	0.8241	0.9668
PCA	0.9761	0.9210	0.9692
Wavelet	0.9799	0.9472	0.9808

Table 3 Correlation between spectral bands and pan for QuickBird data: full resolution application

Method	**Band1&Pan**	**Band2&Pan**	**Band3&Pan**
Original MS	0.7262	0.7451	0.7269
IHS	0.7325	0.7435	0.7099
Brovey	0.8481	0.8682	0.8220
PCA	0.8430	0.8591	0.8362
Wavelet	0.8030	0.8376	0.8171

The relative shift in the mean and variation in standard deviation in the histogram of the sharpened bands were also computed. The means of the Brovey and IHS sharpened bands have shifted by at least 60%. The spread of the histogram is also reduced for both sharpened images. The computed histogram metrics are given in Tables 4 and 5.

Table 4 Relative shift in the mean of spectral bands for Quick Bird data: full resolution application

Method	**Band 1 (%)**	**Band 2 (%)**	**Band 3 (%)**
IHS	75.72	82.97	69.87
Brovey	71.22	67.62	64.44
PCA	10.07	11.90	15.78
Wavelet	0.06	0.05	0.07

Table 5 Standard deviation of spectra bands for QuickBird data: full resolution application

Method	**Band 1**	**Band 2**	**Band 3**
Original MS	108.87	183.68	157.35
IHS	46.28	46.86	46.36
Brovey	41.94	50.63	31.68
PCA	84.27	142.14	121.66
Wavelet	107.60	181.01	154.74

The entropy of each band of the multispectral and sharpened image was also computed. The computed values are shown in Table 6.39. The values indicate a decrease in entropy for the Brovey and IHS sharpened images. The increase in information is more for the wavelet-sharpened bands compared to the PCA technique.

Table 6 Entropy and Increase in information of spectral bands for QuickBird data

Method	**Band 1**	**Increase**	**Band 2**	**Increase**	**Band 3**	**Increase**
Original MS	7.5506	-----	8.5567	-----	8.3619	-----
IHS	7.0746	-0.4761	7.1138	-1.4429	7.1231	-1.2388
Brovey	6.8340	-0.7166	7.1397	-1.4170	6.4840	-1.8779
PCA	7.6740	0.1234	8,5466	-0.0101	8.3826	0.0207
Wavelet	7.9214	0.3708	8.7054	0.1487	8.5575	0.1956

h. Discussion of Results for Pan-sharpening of Quickbird Imagery

Pan-sharpening algorithms used in this study were tested using three different sets of satellite data viz., SPOT, IKONOS, and QuickBird satellite imagery. The multispectral and panchromatic images had different spatial resolutions and spectral ranges. The sharpening algorithms were applied to the various images tested, both at low resolution and at full resolution. The wavelet based sharpening consistently produced sharpening results with better spectral quality compared to other techniques. The spatial quality of the wavelet-sharpened image varied between the data sets. In the wavelet-based method, the spatial detail information is derived from the Panchromatic image and added to the spectral information. For other techniques, such as IHS and PCA, the spectral information was derived form the multispectral image and added to the Panchromatic image. The spatial performance of the wavelet sharpening is heavily dependent on a number of factors, such as, for example, high co-registration accuracy between the multispectral and panchromatic image in the order less than half a pixel, or the type of the re-sampling technique used in the production of the multispectral and panchromatic image.

In this study, quantitative image quality metrics for data extracted from Pan-sharpened images are presented. The MSE and RMSE, correlation coefficients between the spectral bands, correlation coefficients between the spectral bands and the panchromatic image, as well as the relative shift in the mean and variations in standard deviation in the histogram of the image bands are presented. Also shown are the correlation between NDVI values for the multispectral and sharpened images and the resulting increase in information for the sharpened bands (See Tables above). Changes in classification are the quantitative metrics resulting from this improvement in image quality (as will be presented in the next secion) and the improved image of the 0.6 m resolution will be used in a later part of this study for the purposes of data extraction through remote sensing processing software (Feature Analyst and eCognition).

VI. Feature Extraction with High Resolution Satellite Data

This study applies specific software systems (Feature Analyst and eCognition Developer) for object classification using high resolution remotely sensed imagery to evaluate the efficacy of these automated systems in extracting Port-related features. There are presently three approaches used in the identification and classification of targets of interest in remotely sensed imagery: manual identification, task specific automated approaches, and intelligent software systems.

Manual Identification involves the use of a trained image analysts who systematically identifies features of interest using image analysis tools. The drawbacks of this method are that: (1) it requires a skilled human laborert; (2) it is time consuming; and (3) can be expensive because of the time involved. Task specific automated approaches are an attempt to automate the feature extraction process using computer programs, but they too are not without fault. These programs typically involve long development cycles, are slow, and contain complex algorithms that are useful in identifying only a single feature class. Finally there are intelligent software systems which offer a compromise between between automated and manual approaches by streamlining the extraction process through the use of machine learning algorithms. ArcGIS Feature Analyst Extension and eCognition Developer are examples of such software. Here algorithms can be tuned "on-the-fly" through the use of user-friendly "smart" interface, thus automating the process of inputting intelligent variables. These programs present the user with a simplified data analysis extraction process and allow for tasks to be performed by users not trained in complex remote sensing techniques.

a. Object Based Imagery Analysis

Remote sensing has made significant strides in the past several decades and a variety of sensors now provide medium and high resolution data virtually on demand. Many of the existing applications still in use today rely on image processing techniques developed in the seventies: the classification of single pixels in a multi-dimensional feature space with little or no emphasis on the spatial context of these pixels in relationship to each other and the larger geography. Alternatives to pixel based classification are now being developed and refined which take into account shape, form, texture, as well as spectral information (Hay et al., 2006). This technique is called Object-Based Image Analysis (OBIA).

b. Feature Analyst

Feature Analyst is a commercial feature extraction software system that uses inductive learning algorithms and techniques to model the feature recognition process, rather than explicitly writing a specialized code. The user provides a set of sample features within the image whereupon the system then automatically develops a model that correlates known data (such as spatial and spectral signatures) with targeted outputs (i.e. the features or objects of interest). The learning model then automatically classifies and extracts the remaining objects/targets. This approach leverages the natural ability of humans to recognize objects in complex imagery.

Key components of the Feature Analyst system for object recognition and feature extraction from imagery include the following:

- A simple user Interface which masks the complexity of the assisted feature extraction approaches
- State-of-the-art learning algorithms for object recognition and feature extraction
- Post-processing tools for editing and generalizing features
- Assisted feature extraction modeling tools for capturing workflows and automating feature collection tasks

c. Feature Analyst Workflow

The feature analyst workflow includes the following steps:

1. The user digitizes several examples of the feature they are trying to extract

2. The user sets the feature type from the user interface, which automatically sets all of the learning parameters behind the scenes

3. The user extracts features using the one-button approach

4. The user examines the results and, if necessary, provides positive and negative examples to remove clutter using hierarchical learning, whereupon the results can be aggregated.

5. The user then smoothes the aggregation into a final product (Blundell et al., 2006).

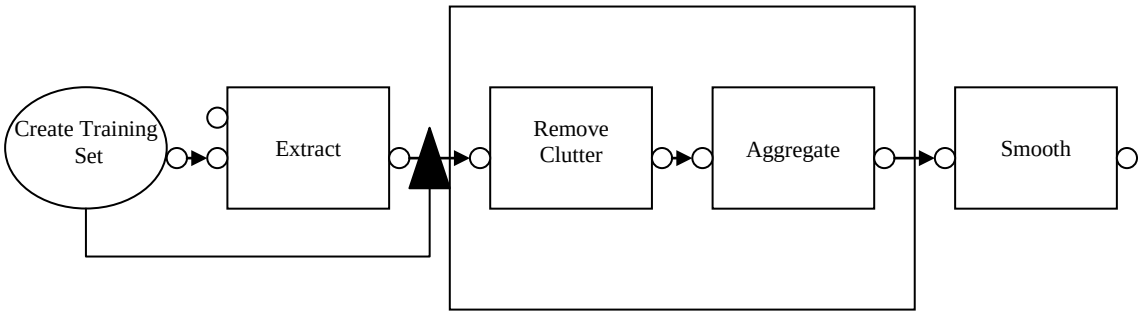


Figure 2: Feature Analyst Software Workflow Chart

<u>Creating a Training Set</u>

In the following paragraphs, the “Learner” being referred to is an intelligent software analytical tool which can be programmed by the user. It’s important to provide the Learner with the best possible examples of the feature you are trying to extract. If the target feature is spread throughout the image, the analyst should make sure that the training polygons are also spread throughout the image. If the target feature runs in all directions in the image, the training polygons should reflect that as well. The final training set should capture spectral differences in the target features based on their spatiality. The training set always dictates the classification outcome, so the analyst should take the time to investigate, analyze, and apply samples based on intelligent

analysis. In any atomized extraction the final results can only be as good as the training set they are built on. The assumption is that the Learner will be provided with quality samples of target features in order to obtain the best possible results. In other words, the quality of results is dependent on the quality of the training set. The following techniques should be applied while collecting training dataset:

- Create representative shapes throughout the image to give the Learner many and varied examples to learn from.

- Zoom in on the image to help better visualize the features being extracted

- Include examples of how feature orientation varies throughout the image.

- Pay close attention to the types of information being included in the training set, for examples including or excluding shadow areas.

Bad training examples consist of the following can be poorly drawn examples or too few examples.

Input Bands

Input bands provide the spectral data needed in automated feature extraction. A panchromatic or black and white image contains one band of spectral data. Multi-spectral or color images consist of a minimum of three bands of spectral data, which comprise the basis of the visible color spectrum. These bands provide reflectance data to the Learner. Feature Analyst automatically detects the number of bands in the image and displays them in the Selected Bands list box on the Input Bands tab on the Set Up Learning dialog box.

Input Representation

One of the most significant differences between the Feature Analyst and other image classification software products is its ability to use of the spatial context of the pixels. The Learner takes into consideration not only the target pixel, but also possesses the capability to analyze the surrounding information as well. Often it is necessary to look at what lies around a feature in order to distinguish it from the background as either the target or clutter. That is how the Feature Analyst tells the difference between rivers and lakes, roads and parking lots, and other features which have similar spectral signatures. By adjusting the spatial context, the software used can significantly affect the resultant feature classification output. *Input Representation* interface determines how each pixel is looked at in relation to its neighbors. There are eight pre-set input patterns, plus two user-defined patterns. Each of the pre-set patterns includes a range of pattern widths. The red outlined pixel in the center represents the “decision” pixel. Feature Analyst looks at the center pixel in relationship to the highlighted surrounding pixels and then decides whether or not the center pixel is a member of the feature class.

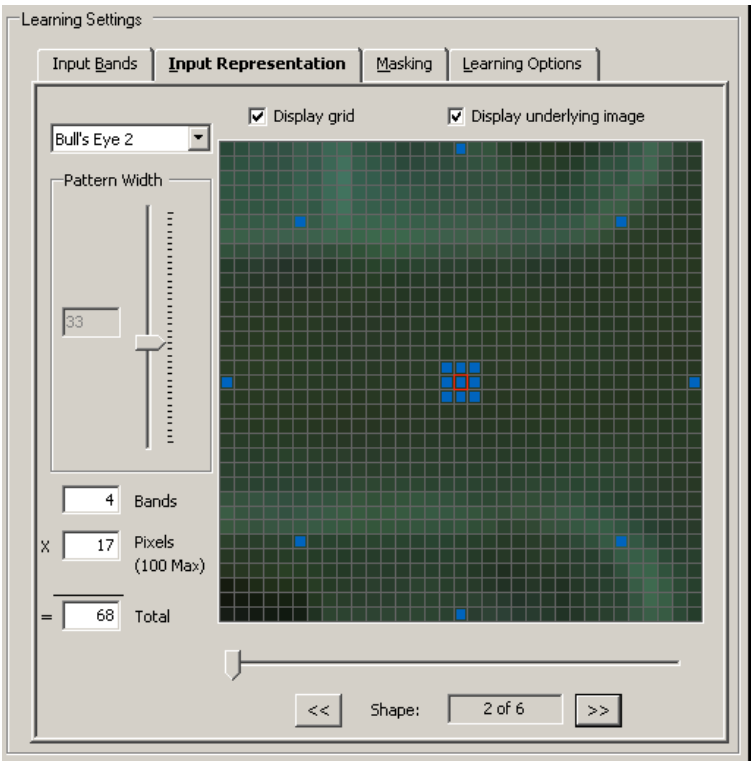
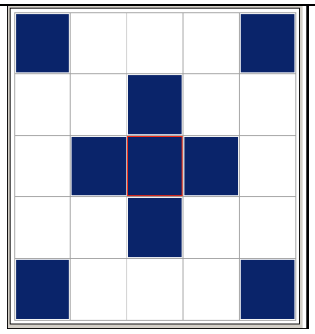
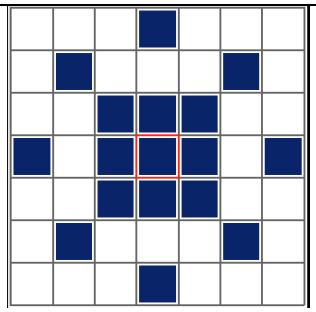
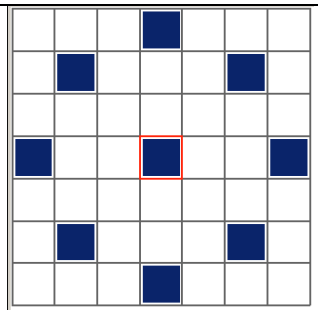
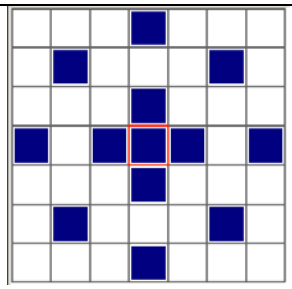
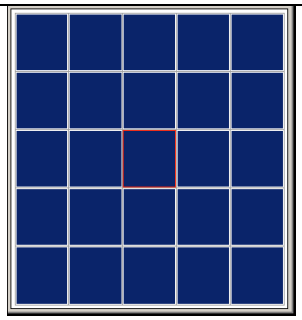
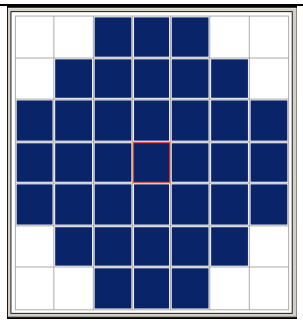
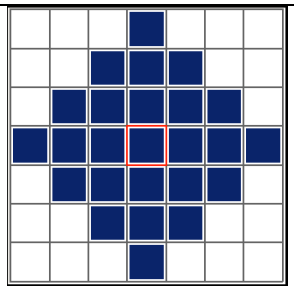


Figure 3: Feature Analyst *Input Representation* Screen Example

Bull's Eye 1 **• Use for narrow linear features, such as: Side walks, Streams, Small man-made features** **Choose which pattern and width depending on the actual feature and on the image resolution.**	
Bull's Eye 2 **• Use for wide linear features,** **such as: Freeways, Rivers, Parking lots**	

Bull's Eye 3 **• Use for: Trees**	
Bull's Eye 4	
Square	
Circle	
Manhattan **• Use for: Lakes 3x3, Land Cover 3x3, Vegetation 3x3, Buildings**	

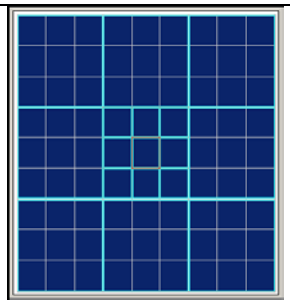
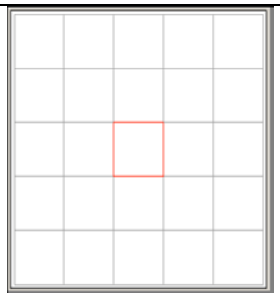
Pre-defined Foveal: **You can create a user-defined Foveal by simply changing the cells highlighted in the pattern.**	
User-defined	

Figure 4: Examples of the Pixel Input Interface Types in the *Input Representation*

Masking

Masking allows the user to include or exclude certain parts of an image during the extraction pass. Masking can speed up the process by limiting how much of the image the Learner has to look at while finding target features. There are three masking options. The default option is No Masking and the software will classify the entire image. Masking by pixel value allows the user to exclude only black or only white pixels. Mask by Layer option allows the user to include or exclude features from other feature classes.

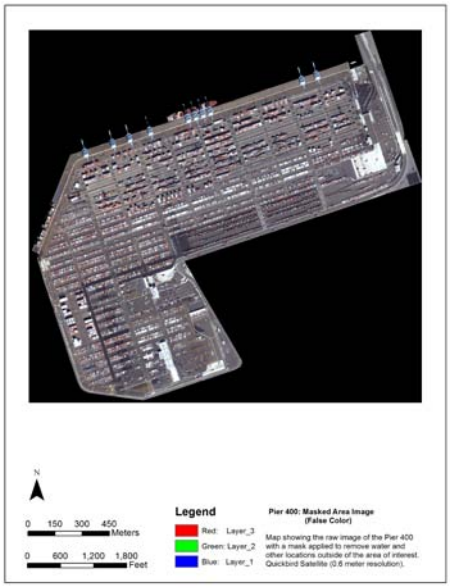


Figure 5: Example of an Application of a Mask to Exclude Areas Outside of AOI

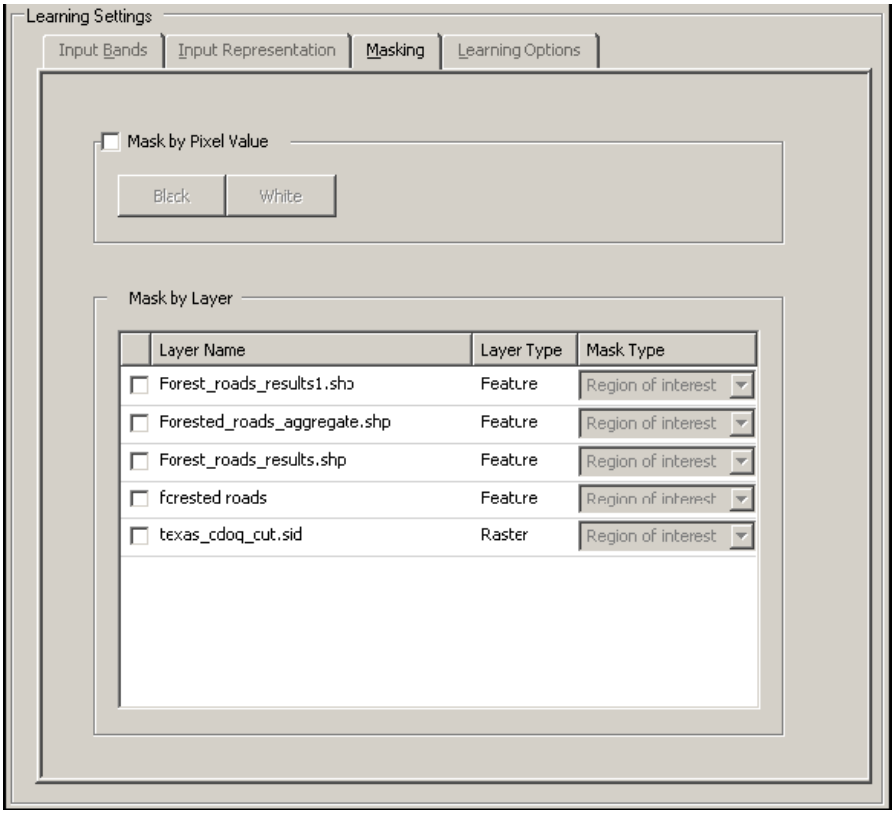


Figure 6: The Learner Settings Screen for *Masking Algorithm*

Learning Options

The Learning Options tab allows the user to select one three learning algorithms. This tab also allows to take into consideration object orientation and fill options.The first Learning Algorithm, Approach 1, is for general purposes. The second, Approach 2, is good for removing clutter. Approach 3 is good for quick extractions. It is recommended that the user applies Approach 1 for majority of extractions. The Learning Options tab also provides two output options: Vector and Raster. The default will match the input layer format type. The Learning Options tab includes five additional learning options:

• Aggregate Areas: informs the system to look for polygons with continuous pixel clusters with no fewer pixels then the number specified by the user. The system eliminates anything smaller than the prescribed number of pixels or it fills “islands” that are equal to or smaller than the aggregate size.

• Smooth Polygons: allows to automatically run a Douglas-Peucker smooth algorithm on the resulting polygons. A threshold value must be specified in the Threshold field.

• Find Rotated Instances of Features: informs Feature Analyst to look for target features running in all directions within the image. Find Rotated Instances option should be used for most extractions. An exception is when the target features have an associated shadow. The results for this type of extraction are often better if the feature is turned off.

• The Wall-to-Wall Classification: becomes available when working on multi-class extractions. This informs Feature Analyst that every pixel in the image will fall into one of the specified classes. When this option is turned off, the system may return unclassified pixels.

• Fill Background regions with shapes: allows the user to return a wall-to-wall classification by creating a class called “background”. This allows you to get a multi-class result without creating a multi-class extraction problem.

Once sample polygons have been created and parameters set, the user is ready to run the extraction of features process and let Feature Analyst find the target features. With the parameters set, the user should One Button Learning on the Feature Analyst toolbar. The system asks to name the feature class that will be created from the feature extraction pass and the results will appear in the Table of Contents.

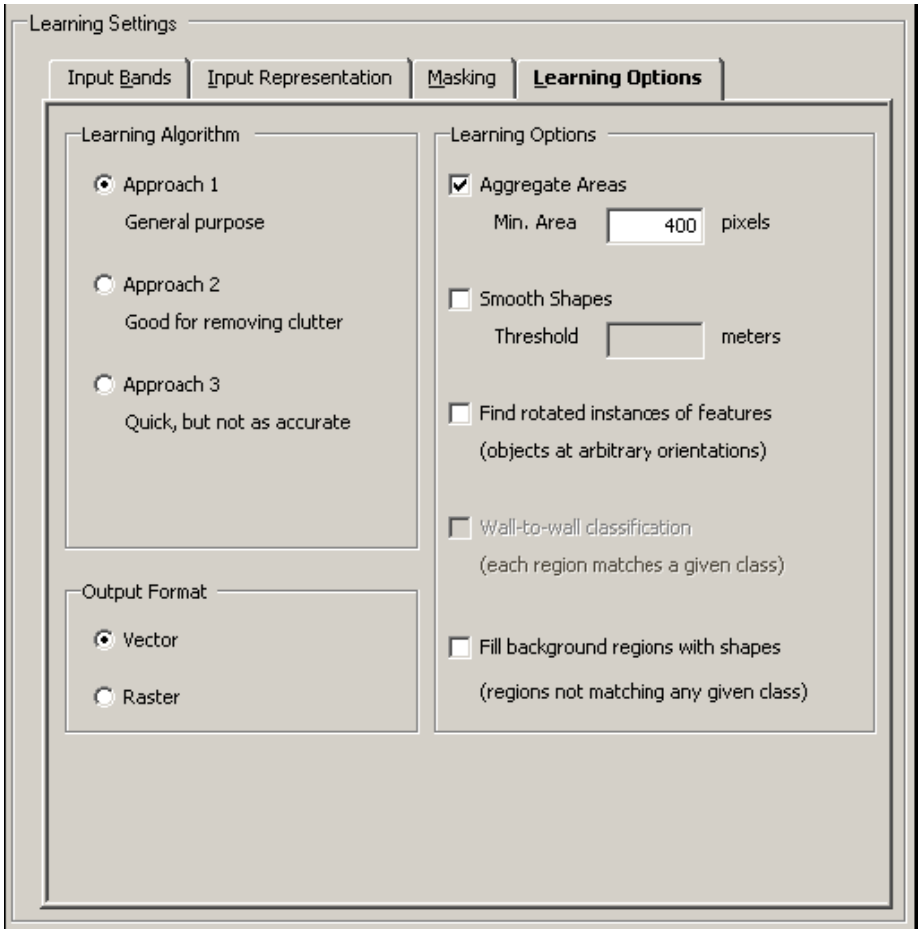


Figure 7: Example of the *Process Extraction Execution* Screen

Removing Clutter

Clutter Removal allows for quick way to eliminate polygons returned in the result layers which were misidentified. The clutter removal tools allow you to identify the correct and incorrect polygons. When identifying correct and incorrect objects, the user must first find the best examples of each for the Learner to learn from. Based on a sampling of correct and incorrect polygons, the Feature Analyst will locate the

remaining correct and incorrect polygons. Feature Analyst also allows for identification of portions of polygons as correct or incorrect.

Icon	Option	Function
	**Select Correct Features**	**The Select Correct Features tool allows you to select entire polygons as correct examples of the target feature. Once you have created your clutter removal layer, select the Select Correct Features tool. Click on a sampling of correct polygons in the feature class. Remember to select the objects that best represent your target features.**
	**Select Incorrect Features**	**The Select Incorrect Features tool allows you to select entire polygons as incorrect examples of the target feature. Once you have created your clutter removal layer, select the Select Incorrect Features tool. Click on a sampling of incorrect polygons.**
	**Digitize Correct Features**	**The Digitize Correct Features tool allows you to select portions of polygons as correct. For example, if you are looking for buildings and the initial results returned the parking lots adjacent to the buildings as correct, you can draw a box around the building, but exclude the parking lot.**
	**Digitize Incorrect Features**	**The Digitize Incorrect Features tool allows you to select portions of polygons as incorrect. For example, if you are looking for buildings and the initial results returned the parking lots adjacent to the buildings as correct, you can draw a box around the parking lot, but exclude the building.**

Figure 8: Example of the Interface for *Removing Clutter*

Remove Clutter by Shape

The second means of clutter removal is Remove Clutter by Shape. This feature allows the user to select from a list of common attributes and use only those features to separate correct and incorrect polygons the results in the feature class. One or more of these options can be used together to get the best possible results.

Option	Function
Compactness	**Use to remove clutter when one category of shapes** **has a significant number of holes and spreads irregularly over a large area, while the other is densely packed and regular. This option is particularly useful when one category of shapes has a profile that is close to a circle and the other spreads out and contains holes, for example, dark colored grasses versus sage brush clumps.**
Second Order of Moments	**This option uses the statistical properties of a shape to describe it. It measures the orientation, elongation, and size of a shape relative to a given axis. Use in scenarios where you are trying to divide a polygon extraction into two classes: large and small.**
Invariants	**Use when shapes of one category share a similar template, but differ in rotation and size. As this option is unaffected by both rotation and scaling, you could potentially use it with a feature dataset that includes both urban and suburban area buildings. In such a dataset, buildings normally would have various sizes and rotations but would have a similar template.**
Number of Vertices	**Use when the correct shapes have a different number of vertices compared to the incorrect shapes. For example, when you are trying to extract only the buildings from surrounding vegetation, which is normally characterized by large shapes and an irregular number of vertices.**
Number of Holes	**Use when one category of shapes has a different number of holes compared to the other. Useful when you are trying to extract soil from forest cover where soil is normally returned by a solid shape and forest cover has a lot of holes.**

Perimeter	**Use when the correct shapes have different perimeters compared to those shapes that have been marked as incorrect. This descriptor is useful in differentiating irregular shapes from regular shapes, for example, vegetation polygons versus buildings.**
Area	**Use when two categories of shapes have different areas. This descriptor can be used to filter out either small or large shapes.**
Bounding Rectangle	**Use when the shapes are irregular and a bounding rectangle best differentiates one set of shapes from the other. This option is similar to the Area option. The system derives the metric by describing a bounding rectangle around the shape.**

Figure 9: Example of the Interface for *Removing Clutter by Shape*

Aggregate Features

Aggregate Features option allows the user to fill-in holes in polygons or to remove polygons that are smaller than a desired size. During the initial extraction pass, Feature Analyst aggregates features based on an area specified in the Learning Options tab Minimum Aggregate Area field. Once the results are returned, the user can aggregate them or separate them for further improvement of results.

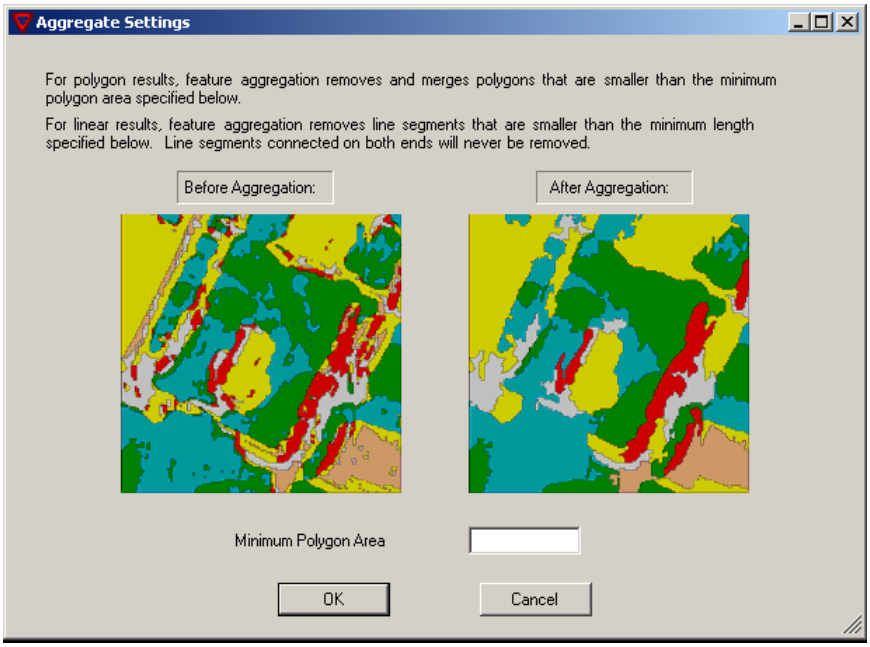


Figure 10: Example of the *Aggregate Feature* Interface

Smooth Features

The Smooth Features feature the user to clean up the edges of the polygons. This tool is especially useful when extracting buildings and the resulting features are wanted to closely resemble the actual features.

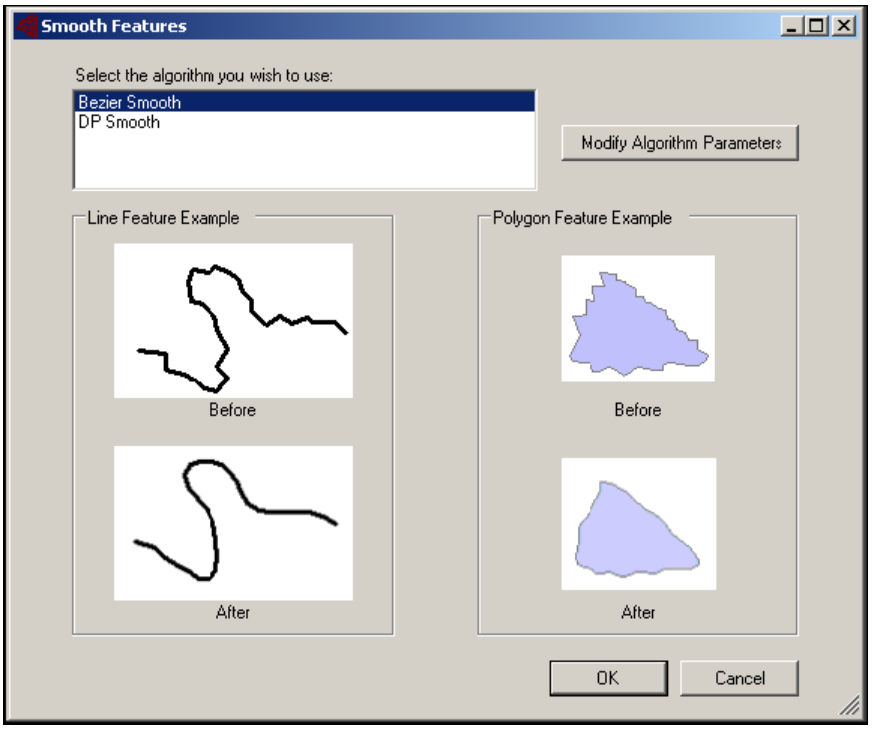


Figure 11: Example of the Interface for *Smoothing Features*

d. Feature Analyst Processing of Quickbird Imagery

The data used in this project is a series of 4 orthorectified images. This terrain corrected product, which was radiometrically calibrated and corrected for sensor and platform-induced distortions and set to a cartographic projection. This product is GIS-ready and can be used as an image base map for a wide variety of applications where a high degree of absolute accuracy is required. The orthorectified Imagery used is an area-based product, meaning that the product is defined by the client's area of interest without reference to scenes. The Quickbird imagery was collected in 5 different spectral bands and 11-bit dynamic radiometric resolution, and was provided to the analyst as a 16 bit product with specifications as listed below:

Product Option	Band	Spectral Range
Panchromatic	Black and White	450 - 900 nm
Multispectral	Blue	450 - 520 nm
	Green	520 - 600 nm
	Red	630 - 690 nm
	Near-Infrared	760 - 900 nm

Figure 12: Spectral Information for Quickbird Satellite Image for All Bands

The 2.44 meter imagery was then pan-sharpened (resampled) to 60 cm using a proprietary MTF Kernel which uses an 8 by 8 pixel window to determine the value of the destination pixel. This resampling method was optimized for the actual MTF response of the QuickBird sensor and produced sharp edge detail.

Imagery Preparation

The four images that made up the study area were combined in ERDAS Imagine using the mosaic function in order to create a single seamless image. Subsets of the study areas corresponding to the areas of interest were then extracted from the single image.

Feature Extraction

This project involved researching feature extraction techniques from high-resolution satellite imagery in a GIS environment. Feature extraction (e.g. building footprints, bridges, and roads) from high-resolution satellite imagery is an ever-growing technology that is crucial in the rapid updating of a GIS for rapid damage assessment of affected areas in case of emergencies, change detection, and efficient land use management. Two approaches exist for the extraction of features, manual and automated. Manual feature extraction simply refers to heads-up digitizing. It is more accurate but also requires a trained analyst, and can be time consuming and costly. On the other hand, automated feature extraction requires specialized software with algorithms designed to perform such tasks. Although specialized software can be faster at extracting features, the accuracy of their results is not necessarily better than that of manually digitized features. Another drawback to the automated methods is that with VLS's Feature Analyst, they come in a "black box" environment and there is no way an analyst can know what type of algorithm the software uses to extract features, nor does the software produce a statistical output that can be evaluated. Several requests to VLS were made on our part to obtain details of the algorithm the software uses but all attempts were declined. Companies such as VLS do not like sharing such information because of competition.

Methodology

This section of the paper focuses on feature extraction from satellite imagery using Visual Learning Systems' (VLS) Feature Analyst software and describes the methodology used as well as results obtained. Multiple extraction passes using different input settings were run on each of the study areas. Numerous learning options and input representations were applied to the study areas and the parameters that provided the highest percentage of correctly identified features for each of the features were selected. The resulting shapefiles were then processed to remove clutter by selecting both correct and incorrect results.These results were in turn aggregated to remove any remaining holes and inconsistencies within the resulting shapefiles.

e. Results of Feature Analyst Extractions

The following figures present the outcome of the classification for the three areas of interest within the Port. These results will be discussed in the final section when compared with those obtained from eCognition Developer software.

Figure 13: Map Demonstrating Results of Automated Building Extraction vs. Hand Digitized Buildings in Area 2 section of the Port

Figure 14: Map Showing Results of Automated Building Extraction of Pervious Surfaces vs. Hand Digitized Pervious Surfaces in Area 2 section of the Port

Figure 15: Map Showing Results of Automated Extraction of Tanks vs. Hand Digitized Tanks in Area 2 section of the Port

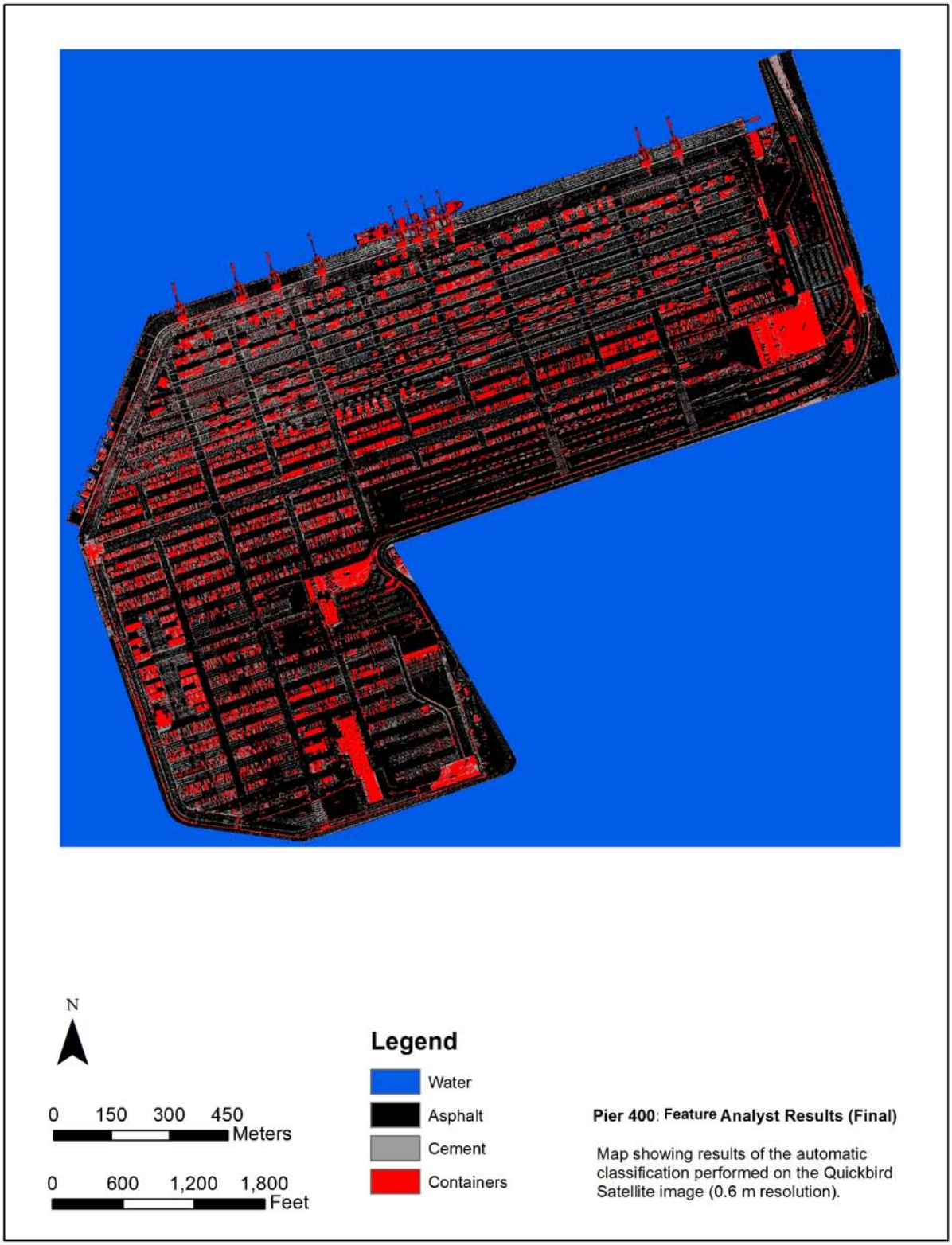


Figure 16: Map Showing Results of Automated Extraction of Features in the Pier 400 Area using Feature Analyst

Figure 17: Map Showing Results of Automated Extraction of Features in the Urban Interface AOI using Feature Analyst

f. Definiens Professional eCognition

Definiens Professional is off-the shelf classification software that focuses on object-based rather than per-pixel analysis. As humans we see an image as an assemblage of 2 dimensional objects rather than an array of pixels. It is our knowledge of the arrangement of these objects within an image that allows us to interpret them in a three dimensional context, to recognize obects, and to recognize spatial relations between them. The concept behind eCognition is to create these image objects and classify them in a learning approach that exploits meaningful spatial information as well as spectral information contained within the image.

g. eCognition Workflow

1. Load data and create project eCognition's object based approach allows incorporation of continuous and thematic data so existing thematic maps may be combined directly in the classification with remotely sensed images. Thematic data can be loaded together with image data as image layers or separately as thematic layers.

When used as thematic layers, raster thematic files can be loaded in a large range of formats.

2. There are four main segmentation algorithms that can be applied to the data i) Multiresolution segmentation, ii) Spectral segmentation, iii) Chessboard segmentation, iv) Quadtree segmentation. For this project only Multiresolution Segmentation was used. The basic work flow in eCognition is simple. Data are loaded into a project, homogeneous image objects are created through segmentation (and are automatically arranged in a hierarchy when more than one object level is created); objects are classified rather than individual pixels. The multiresolution segmentation process starts with single pixels, which are compared to their neighbors using a plane 4 pixel neighborhood function. Using spectral, spatial scale criteria entered by the user, homogeneous objects are grown from the single pixels using a heuristic function. Object size depends upon a scale factor but it is not an absolute determinant of object size.

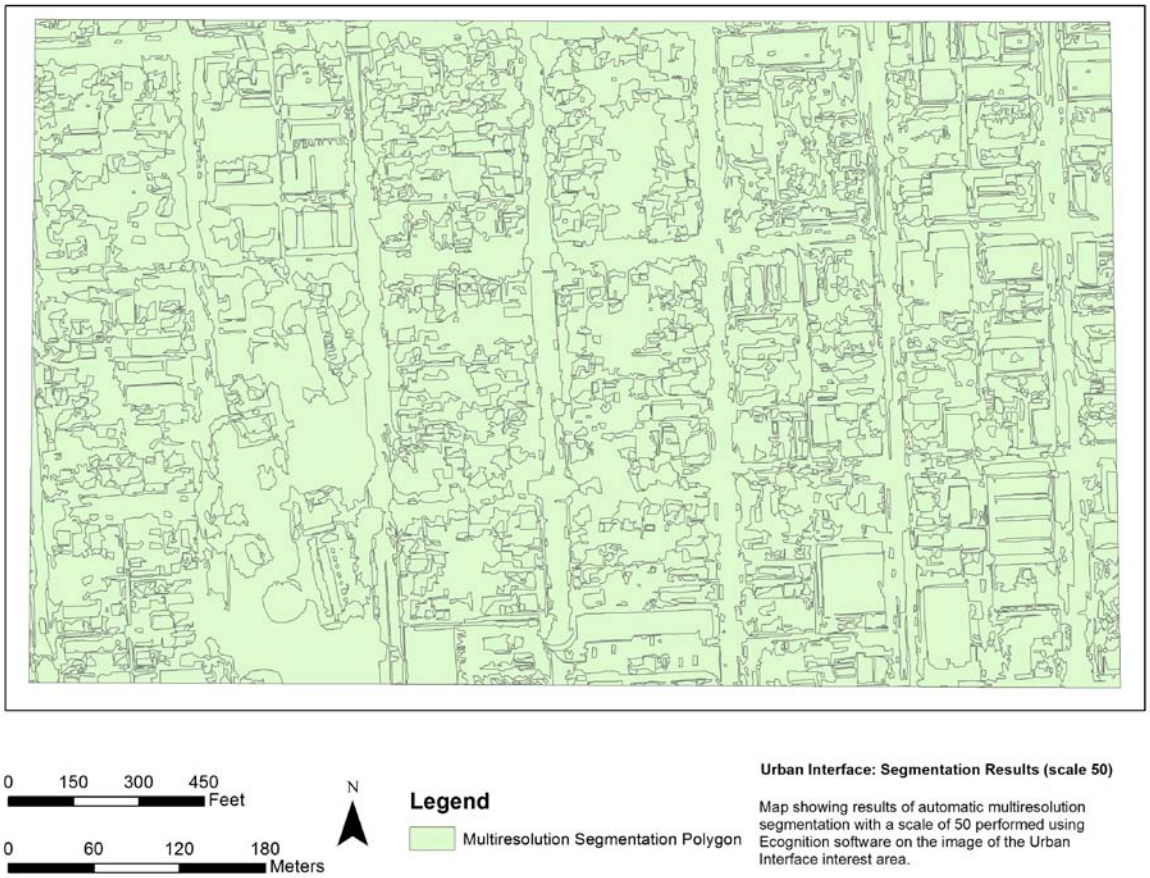


Figure 18: Map Showing Results of Multiresolution Segmentation Results Prior to Classification Attempts in the Urban Interface AOI using eCognition Software

3. The same scale factor can produce different results when used on images with different spatial resolution or when combined with different homogeneity criteria. Object shape depends upon the user-defined homogeneity criteria: a color criterion (hcolor) and a shape criterion (hshape). If used alone, the color criterion results in very fractal image objects so it is used in conjunction with the shape criterion. The shape criterion for spatial homogeneity is defined by the departure of an object from a compact shape

(compactness) and the ratio of the actual border length of the object to the shortest possible border length of the object (smoothness).

4. Many different levels of objects above and beneath existing levels can be created, based upon different scale parameters. Objects on a new level use existing objects on the level immediately below (subobjects) as building blocks and the borders of level-above objects (superobjects) as hard boundaries. Thus smaller subobjects fit exactly inside superobjects – they cannot overlap superobject borders. While subobjects can be the same size as superobjects, they cannot be bigger than superobjects. The following figure shows a typical object hierarchy. The largest objects are contained in the uppermost level – level 3. These were created when a scale parameter of 50 was used. Smaller scale parameters produced smaller objects on the two lower levels, with the smallest objects resident on the lowest level – level 1.

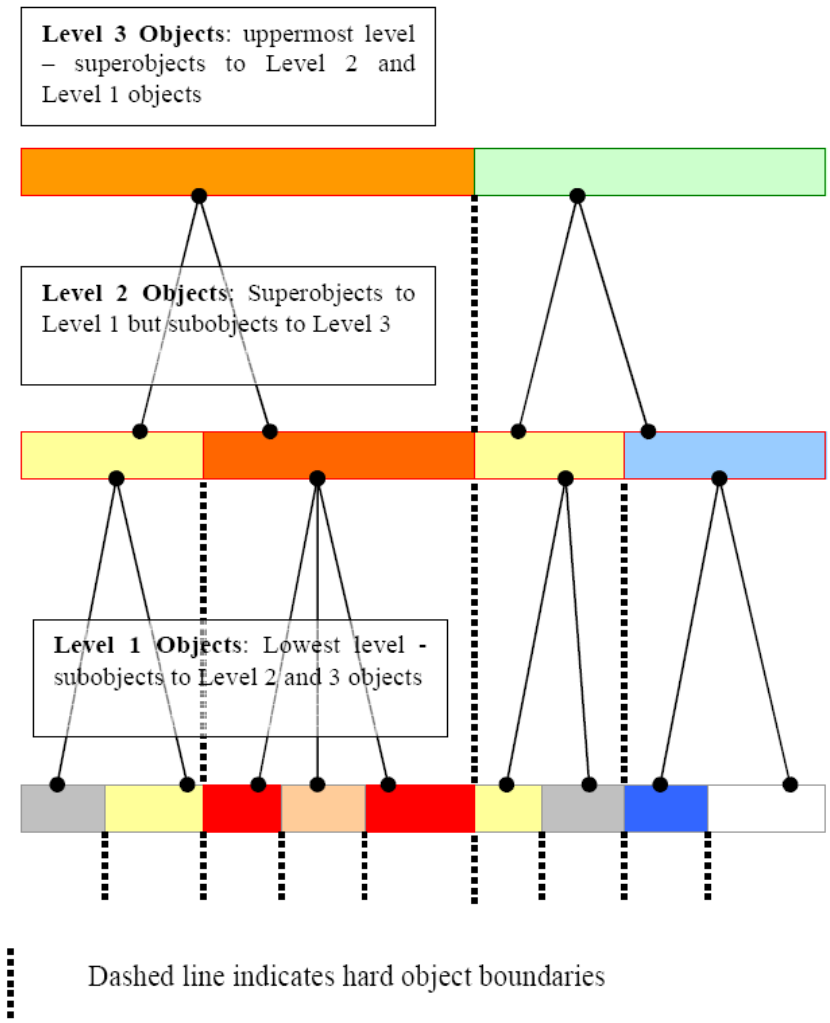


Figure 19: Logic of eCognition Software Object Hierarchy Structure

5. Image Objects have spectral, shape, and hierarchical characteristics – in eCognition, these are called features. There are numerous varieties of features available for use; however, a detailed explanation of these is beyond the scope of this document. Please refer to the eCognition 5 user guide pp. 76 – 81 and the Definiens Professional 5 manual (pp. 33-105) for more information. There are several tools within eCognition that

allow the user to explore features' characteristics and these aid the user in deciding and setting on classification criteria algorithms. They are the Feature Viewer, Sample Editor (with sample selection and sample navigator), Layer histograms, Feature Space Optimizer (FSO) and 2 D Feature Space plot.

6. Classification in eCognition is based on fuzzy logic applied either through membership functions or through the nearest neighbor (NN) classifier. Since the NN is quite straightforward, it was the primary means of classification for this project. Since the user decides the rules of classification, the membership functions allows for a high level of control over feature thresholds for class membership. Definiens recommends membership functions when only a few features are needed to separate a class within a fine spatial resolution image and offers a broad array of functions. It is often the case that objects can be extracted with one or two membership functions alone, as will be demonstrated below. Membership functions are also very useful for exploring the complex interactions of object hierarchy and class membership since they permit the use of class related features. E.g., user can create a classification on one level and refer to it in a classification on a different level. Furthermore, if the intent is to use decision a tree software in the classification process-work (e.g., CART, RandomForest, C5, etc.) then the rules suggested by the decision tree analysis can be easily applied in eCognition as well. Drawbacks include the not infrequent situations which require the user to use a bit of guesswork for pinpointing the function threshold values (not a when using CART). It's also not advisable to use membership functions if the user is not very familiar with the area within the image. Nearest neighbor (NN) is eCognition's rapid classifier. It requires class training and works well when supplied with good samples. The software also allows samples (called TTA masks) to be loaded as separate raster, polygon or point files.

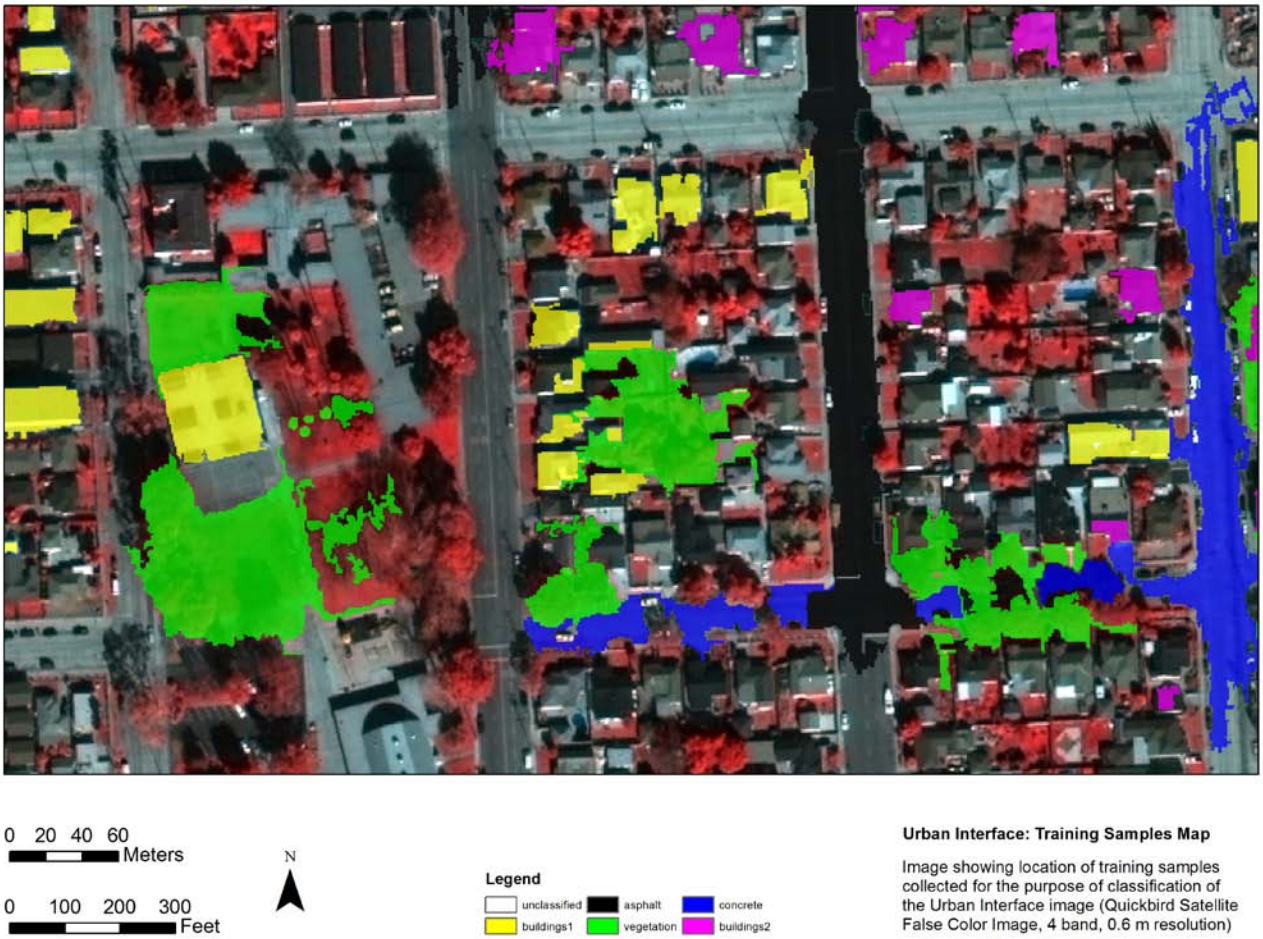


Figure 20: Example of a Training Dataset Collected for the Urban Interface AOI Image Objects Using eCognition Developer Software. Asphalt and Concrete Were Later Merged Into a Single Class Representing Roads. Similarly, Buildings1 and Building2 Classes Were Later Merged Into a Single Class.

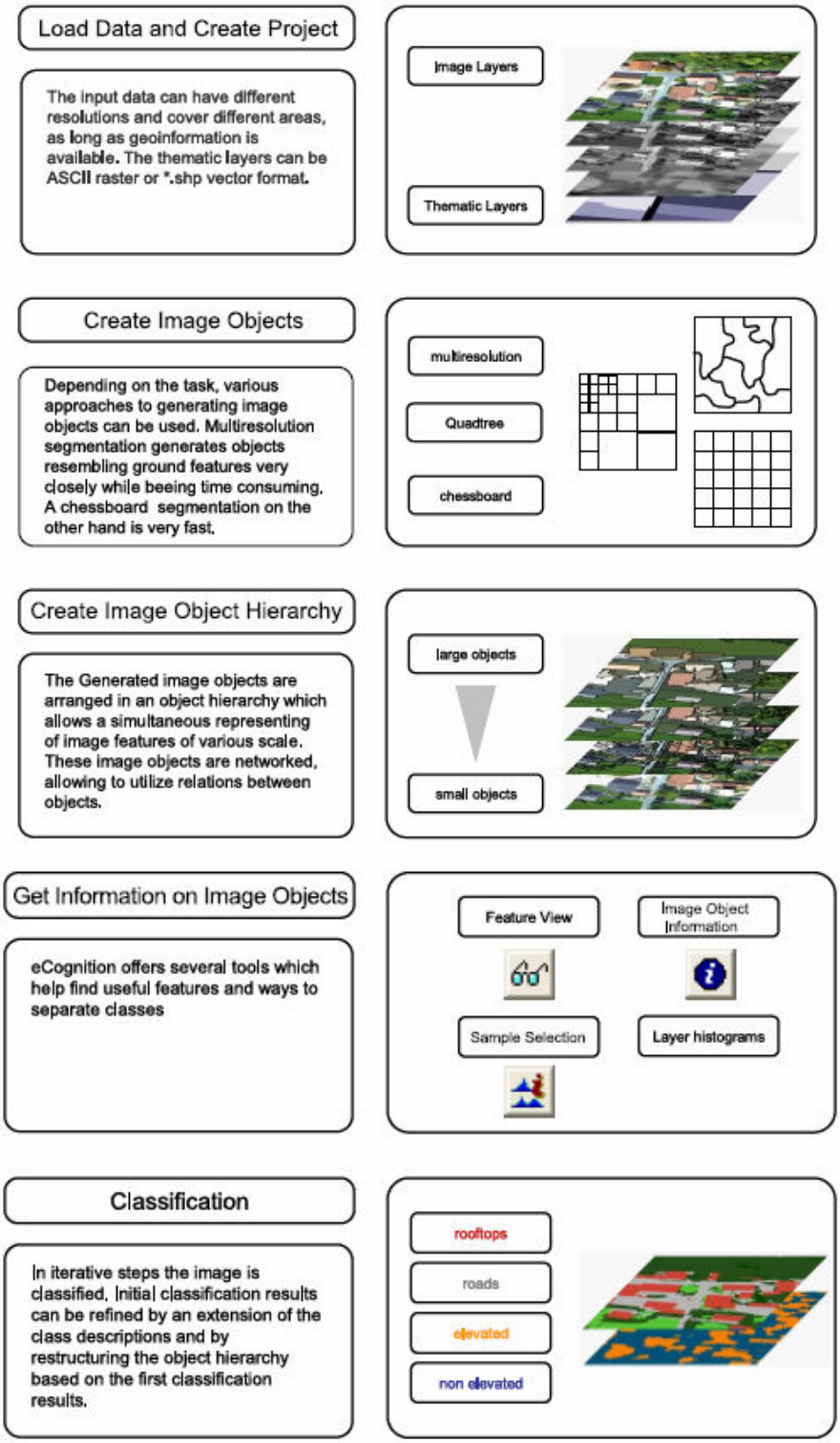


Figure 21: Workflow Chart Demonstrating the Process of Classification using eCognition Developer Software (Definiens, 2006)

h. Results of eCognition Extractions

Methodology

This section describes the steps involved in the process of classifying a single scene (Pier 400 AOI image) for the purposes of demonstrating a typical methodology involved in the image classification process using eCognition Developer software.

A Pan-sharpened 4-band Quickbird image for the AOI was activated in Ecognition Developer 8.0 and was converted to a false color display (Red: Layer1; Green: Layer3; Blue: Layer 2). Linear Equalization of 1.00% was used for the saturation parameter to brighten the image and to increase image contrast. Several nearest neighbor classification algorithms were performed on the image pixels and object areas using various degrees of feature space optimization techniques for both spectral and object oriented algorithm types before a satisfactory result could be achieved.

Originally nine different classes of land cover type were defined by the analyst. They were: asphalt, car, concrete, container, lights, paint, tower, wall, and water. Training samples were collected for each of the nine classes using sub-areas defined based on a multiresolution segmentation algorithm of a scale parameter of 10. Correct classification of containers versus background was the primary goal. Only spectral information of pixels was used for image object feature identification in the first classification attempt. This method produced very poor results and misidentified many features. Defined segmentation areas in the image share similar spectral qualities with other classes and that led to multiple omission and commission errors. For example, many containers appeared white or gray and thus were similar in spectral responses to cement roads, buildings, and asphalt surfaces. The analyst concluded that too many categories and a small size of the training data set in comparison to the overall size of the image area contributed to the poor outcome of the classification.

The image was then converted back to a true color display for the purpose of obtaining a better separation of containers from the background (Red: Layer3; Green: Layer2; Blue: Layer 1). A second attempt at a classification of the Pier 400 area was conducted using only five classes and was based on an image object level created using a spectral difference segmentation algorithm of a scale of 25 (250% fold increase in size of image object area size). Next, a training set data for the five classes (asphalt, cement, container, tower, and water) was collected. In addition to utilizing spectral variations of pixels in the image objects for feature recognition, two geometrically oriented algorithms (shape and size) were introduced into classification. These algorithms specified containers' extent (less than 115 pixels per object) and rectangular fit ratio (less than or equal to 0.65). Again, similarity in spectral signatures of objects resulted in multiple classification errors. The main problem with this result was the fact that many containers were identified as towers, asphalt, or cement.

In the next classification attempt, the tower class was removed and the classification algorithm was executed again, following the collection of a new training dataset, this time only for four unique classes (towers were removed). The number of correctly identified containers increased substantially, but some containers previously identified as towers were this time classified as asphalt. Others, previously identified as asphalt

remained misclassified in the same category. While this was not a classification result which would provide a complete isolation of all containers from the background, it was deemed satisfactory for results expected from a fully automated classification involving only a few steps. The final image of this analysis in presented in Figure 24. It was stipulated that one of the reasons why many containers were omitted and remained classified as background was due to similar spectral responses of the darkest containers, and the mostly automated software algorithm was unable to differentiate between dark containers and background based on the specified criteria. The next step was to farther explore the already introduced manual algorithms.

An attempt was made at a new segmentation technique, based on a larger object segmentation size; the segmentation threshold was increased to a scale of 50. Because most containers resulted in larger polygon objects, the geometrical threshold for container's extent was increased to 150 pixels, but the rectangular fit remained the same (less than or equal to 0.65), based on sample test obtained from subdivided image objects. The same automated classification algorithm was then repeated using a new class training set. It was quickly noticed that at this separation scale many individual containers became merged into single image objects. Also, the separation between different ground features was less successful in identifying similarities and differences between them and thus grouped them incorrectly. It was also noted that many of the buildings remained unclassified, likely because no category was defined for objects of high spectral brightness, such as cement roof-tops. While the larger size of image objects made it easier for the analyst to select a more extensive set of training polygons, the results of this classification were deemed less satisfactory than in the previous attempt. Thus, the previous image was chosen as the best classification result from this trial.

Classification Results

The next three figures present the final classification results for the three areas of interest of this study. These results will be discussed in the next section.

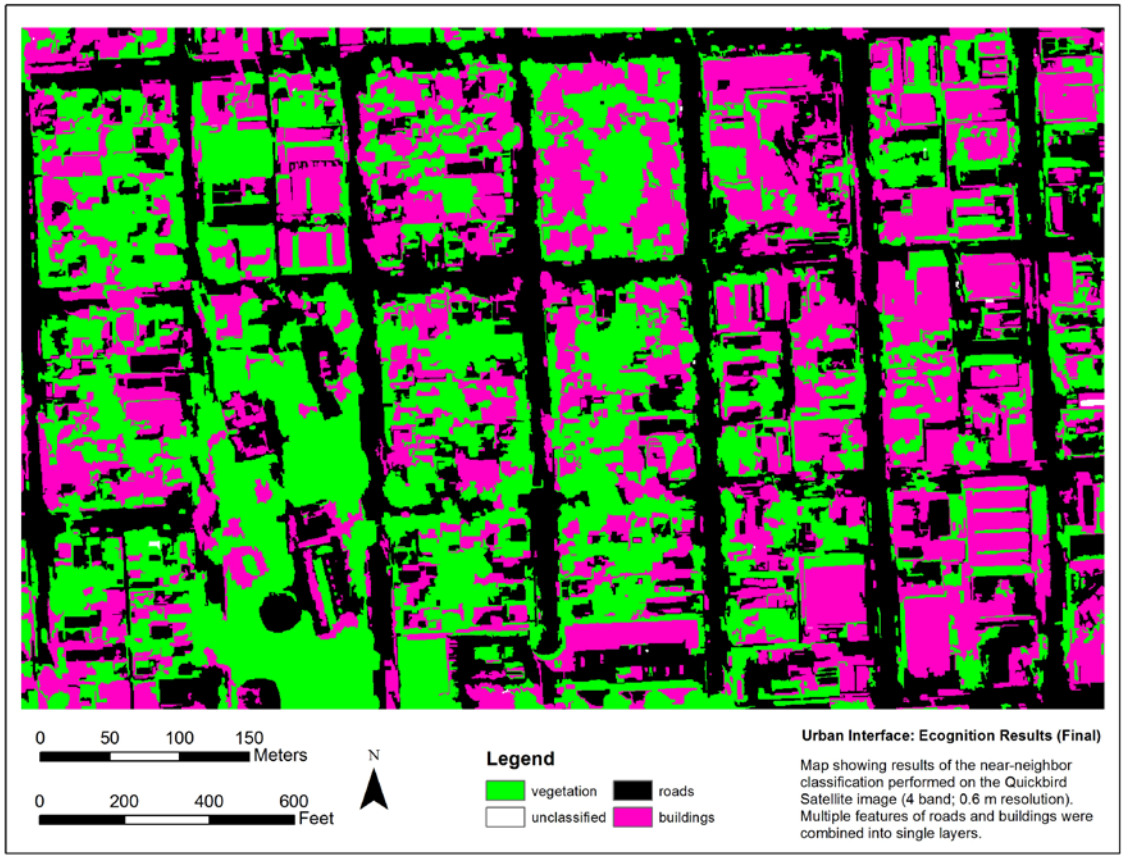


Figure 22: Map Showing Results of Multiresolution Segmentation Results Prior to Classification Attempts in the Urban Interface AOI using eCognition Software

Figure 23: Map Showing Results of Classification for the Area 2 AOI using eCognition Developer Software

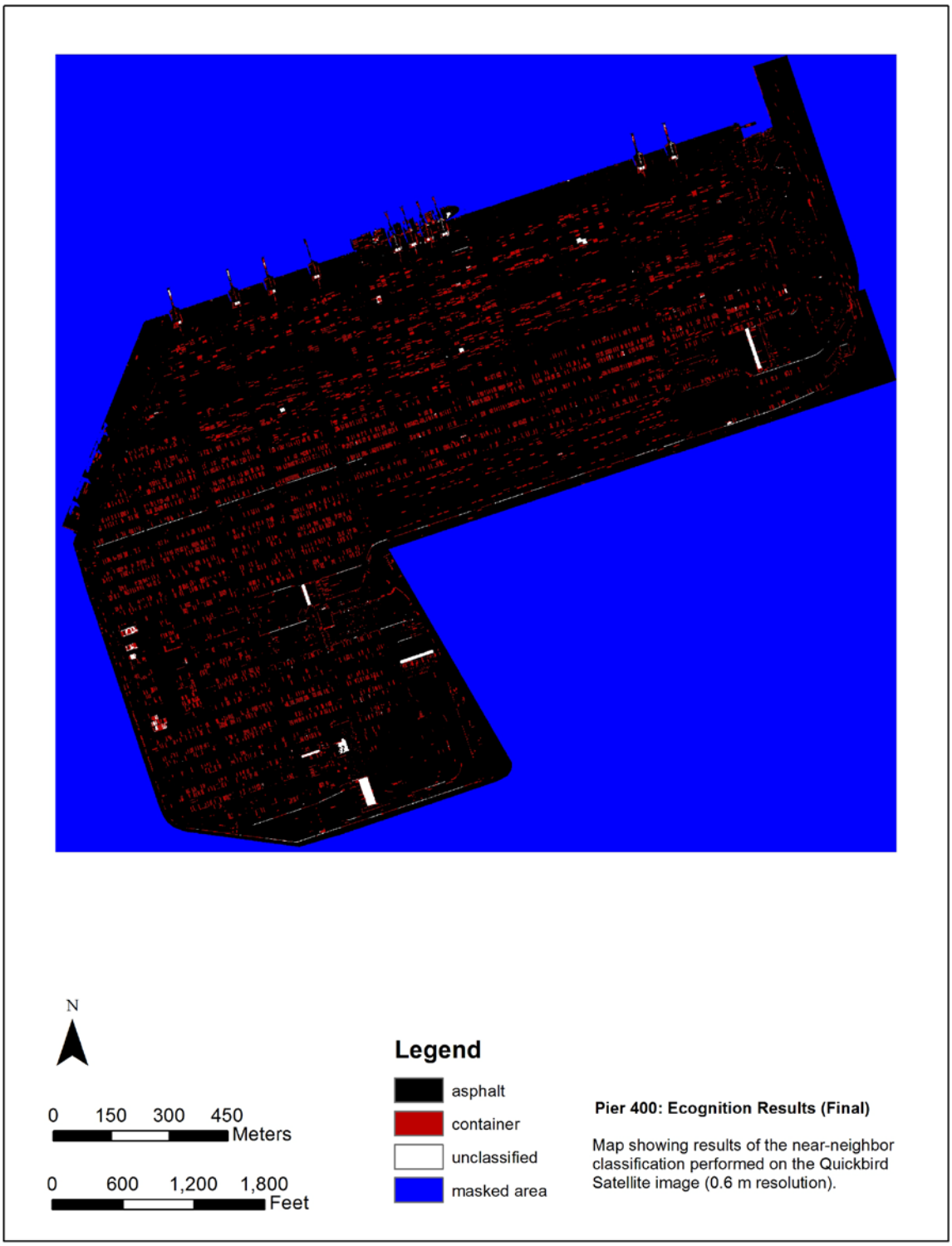
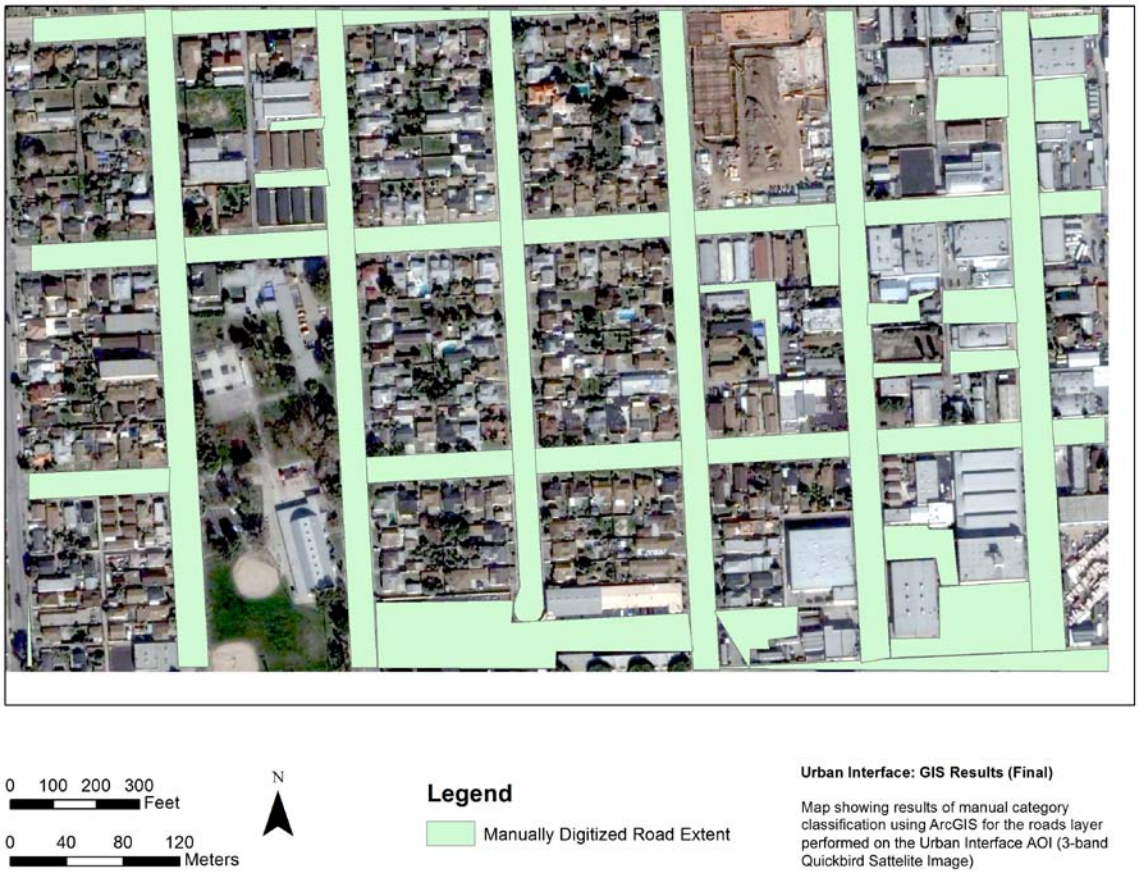


Figure 24: Map Showing Results of Classification for the Pier 400 AOI using eCognition Developer Software

VII. eCogniton vs. Feature Analyst

a. Comparison of Results

Classification results obtained from both the Feature Analyst and eCognition Developer Programs were compared to one another for accuracy assessment. In order to establish the actual extent of the ground features, a manual digitization technique was performed and the true sizes of the areas were calculated using acres (US). These measurements were than compared to the sizes of the feature areas obtained using the automated classification methods for both software platforms, as was discussed previously. The image displayed below portrays the actual extent of roads (asphalt and cement) for the Urban Interface AOI which was manually digitized by the analyst.



.

Figure 25: Map Showing Results of Manual Digitizing of the Roads Extents in the Urban Interface AOI for the Purposes of Later Accuracy Assessment.

The same technique was used for other features in the Urban Interface AOI and Area 2 AOI. No manual digitizing was performed for the Pier 400 AOI due to the complexity of the image objects. The results of the comparison for the feature areas are presented below.

Feature Type	Ecognition Extracted Features	Feature Analyst Extracted Features	Manual Digitizing of Features
	Urban Interface Ecognition	Urban Interface Future Analyst	Urban Interface Manual Digitizing
	Area (Acres)	Area (Acres)	Area (Acres)
Roads	27.8	8.72	26.7
Buildings	27.98	26.32	n/a
	Pier 400 Ecognition	Pier 400 Future Analyst	
	Area (Acres)	Area (Acres)	
Water	589.43	589.44	
Containers	29.81	79.04	
Asphalt	488.77	363.54	
Cement	0	78.87	
Unclassified	2.87	0	
Total	1110.88	1110.89	
	Area2 Ecognition	Area 2 Feature Analyst	Area 2 Manual Digitizing
	Area (Acres)	Area (Acres)	Area (Acres)
Buildings	Buildings 41.2	39.64	44.2
Pervious	Pervious 60.1	16.52	79.57
Tanks	Tanks 11.5	13.86	2.23

Figure 26: Table Listing Area Comparisons for the Three AOI's Chosen Within the Port Complex.

b. Discussion of Results

A comparison of the results above indicates that overall the eCognition software package automated classification was more successful at identifying land cover classes in the study area. In some cases, the Feature Analyst was also very successful; however both programs either overestimated or underestimated actual areas of ground objects in many cases. For the Urban Interface AOI eCognition algorithms estimated both roads and building cover with very good results, while the Feature Analyst was very unsuccessful with identifying road features. This was due mainly to the fact that the algorithm was unable to recognize between two road types: asphalt and concrete. Asphalt and concrete have very different spectral qualities and asphalt, which has high absorbance qualities across the visible

spectrum often appears identical to areas in deep shade. Concrete, on the other hand, can be easily mistaken for building roof tops or pervious areas with no vegetation cover (dirt). The analyst's ability to modify the classification algorithm in eCognition based on objects geometry provided the means to differentiate between these objects with more success. Roads and buildings were also split into separate classes and corresponding training samples obtained for each class, and these classes were later merged together.

The Feature Analyst was also very unsuccessful in separating containers away from the background asphalt in the Pier 400 AOI and largely overestimated the extent of the containers. A large part of this error was due to the software mistakenly identifying containers as cement. eCognition Developer on the other hand slightly underestimated the extent of containers due to its inability to identify the darker containers which were recognized as asphalt. The developer also performed very poorly at identifying concrete, and in fact no objects were classified in this category. For the Area 2 AOI, both eCognition and Feature Analyst performed very well at identifying buildings, but both programs overestimated the extent of tanks. Also, Feature Analyst greatly underestimated pervious ground cover, while eCognition underestimated it by a lesser amount.

VIII. Conclusion

The above results show that both Feature Analyst and eCognition programs can be used for successful object identification for the Port complex, but only through careful analysis. Results should be reviewed and refined until a required degree of accuracy is achieved. Automated classification methods, while they greatly simplify remote sensed image analysis, should always be conducted with caution and final outcome should be inspected by a well- trained analyst and should be compared to the real objects on the ground.

IX. References

Eskicioglu, Eskicioglu, A.M.; Fisher, P.S.; Image quality measures and their performance Communications, IEEE Transactions on, Dec 1995 43 (12); pp 2959 – 2965.

Definiens Imaging, 2006. ECognition Developer User Manual, Munchen, Germany.

Hay et al., 2005 G.J. Hay, G. Castilla, M.A. Wulder and J.R. Ruiz, An automated object-based approach for the multiscale image segmentation of forest scenes, *International Journal of Applied Earth Observation and Geoinformation* **7** (4); pp. 339–359.

Opitz and Blundell, 2008 D. Opitz and S. Blundell, Object recognition and image segmentation: The Feature Analyst approach. In: T. Blaschke, S. Lang and G.J. Hay, Editors, *Object Based Image Analysis*, Springer, Heidelberg, Berlin, New York (2008); pp. 153–168.

Parcharidis, I.; Kazi-Tani, L.M., 2000, Landsat TM and ERS data fusion: a statistical approach evaluation for four different methods Geoscience and Remote Sensing Symposium, 2000. Proceedings. IGARSS 2000. IEEE 2000 International, Honolulu, (5); pp. 2120 - 2122 vol.5.

Yan et al., 2006 G. Yan, J.-F. Mas, B.H.P. Maathuis, Z. Xiangmin and P.M. Van Dijk, Comparison of pixel-based and object-oriented image classification approaches — A case study in a coal fire area, Wuda, Inner Mongolia, China, *International Journal of Remote Sensing* **27** (18) pp. 4039–4055.

Vijayaraj V, C G O'Hara, N H Younan, Quality analysis of pansharpened images, in IEEE International IEEE International IEEE International Geoscience and Remote Sensing Symposium 2004 IGARSS 04 Proceedings 2004 (2004).

Wald, L., 1997. An overview of concepts in fusion of Earth data. In Proceedings of the EARSeL General Assembly,Lyngby, June 1997.

Zhou and Troy, 2008 W. Zhou and A. Troy, An object-oriented approach for analysing and characterizing urban landscape at the parcel level, *International Journal of Remote Sensing* **29** (11) (2008), pp. 3119–3135.