THE DEVELOPMENT OF GROUND TRUTH DATA AND ACCURACY ASSESSMENT OF HYPERSPECTRAL IMAGE CLASSIFICATION AND SPECTRAL UNMIXING

By

Carlos Rivera Borrero

A thesis submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

in

ELECTRICAL ENGINEERING

UNIVERSITY OF PUERTO RICO
MAYAGÜEZ CAMPUS

2007

Approved by:

Miguel Vélez-Reyes, Ph.D Member, Graduate Committee	Date
Fernando Gilbes, Ph.D Member, Graduate Committee	Date
Shawn D. Hunt, Ph.D President, Graduate Committee	Date
Roy Armstrong, Ph.D Representative of Graduate Studies	Date
Isidoro Couvertier, Ph.D Chairperson of the Department	Date

THE DEVELOPMENT OF GROUND TRUTH DATA AND ACCURACY ASSESSMENT OF HYPERSPECTRAL IMAGE CLASSIFICATION AND SPECTRAL UNMIXING

By

Carlos Rivera Borrero

2007

Chair: Shawn D. Hunt
Major Department: Electrical and Computer Engineering

This work includes both theoretical and practical aspects of hyperspectral image classification and unmixing assessment. The theoretical aspect includes a study of end-to-end performance of hyperspectral classification and unmixing systems. Specifically, it compares widely used current state-of-the-art algorithms with those developed at the University of Puerto Rico at Mayaguez. These include algorithms for image enhancement, band subset selection, feature extraction, supervised and unsupervised classification, and constrained and unconstrained abundance estimation. The classification algorithms are compared in terms of percentage of correct classification. The unmixing algorithms are compared using a new procedure to evaluate unmixing performance is described in this work and tested using coregistered data acquired by various sensors at different spatial resolutions. Techniques for image complexity analysis currently available for automatic target recognizers are studied and adapted in an attempt to predict the performance of the classifiers for different image classes. Performance is generally specific to the image used. The practical aspect included both acquisition and manual classification of ground truth data. Most important among this work is a 1m ground truth map of the Enrique reef in La Parguera Puerto Rico. In addition, 158 calibrated spectral measurements along with their GPS location at submeter resolution were collected. Finally, a hyperspectral image from Hyperion was coregistered with the 1m ground truth to determine the abundaces for each of the pixels in the low spatial resolution image.

EL DESARROLLO DE DATOS DE VALIDACIÓN Y EVALUACIÓN DE LA EXACTITUD DE CLASSIFICACIÓN DE IMÁGENES HIPERESPECTRALES Y DESMEZCLADO ESPECTRAL.

Por

Carlos Rivera Borrero

2007

Consejero: Shawn D. Hunt
Departamento: Ingeniería Eléctrica y Computadoras

Este trabajo incluye aspectos teóricos y prácticos sobre la evaluación de clasificación y desmezclado de imágenes hiperespectrales. El aspecto teórico incluye un estudio del rendimiento de principio a fin de sistemas de desmezclado y clasificación hiperespectral. Específicamente compara algoritmos de punta ampliamente utilizados con aquellos desarrollados en la Universidad de Puerto Rico en Mayaguez. Estos incluyen algoritmos para realce de imágenes, selección de subconjunto de bandas, extracción de rasgos, clasificación supervisada y no supervisada y estimación de abundancias con restricciones y sin restricciones. Los algoritmos de clasificación son comparados en términos de porcentaje de clasificación correcta. Los algoritmos de desmezclado son comparados utilizando un nuevo procedimiento para evaluar desmezclado descrito en este trabajo y probado utilizando data adquirida por varios sensores a distintas resoluciones espaciales. Técnicas para el análisis de complejidad de imágenes actualmente disponibles para reconocedores automáticos de blancos son estudiadas y adaptadas en un intento de predecir el rendimiento de clasificadores para distintas clases de imágenes. Rendimiento es generalmente específico a la imagen utilizada. El aspecto práctico incluyó adquisición y clasificación manual de la data de validación. Lo más importante de este trabajo es un mapa de clasificación del arrecife Enrique en La Parguera Puerto Rico. En adición, 158 mediciones espectrales calibradas junto a sus localizaciones geográficas con resoluciones menores a un metro fueron recolectadas. Finalmente, una imagen de Hyperion fue coregistrada

junto a la data de validación a un metro, para determinar las abundancias para cada píxel de la imagen hiperespectral de baja resolución espacial.

I dedicate this thesis to my loving God. Thank you for help me in every step of my way.

ACKNOWLEDGMENTS

I thank professor Shawn Hunt for all his counseling for this work and for my graduate studies. Thanks to Dr. Miguel Velez, Dr. Fernando Gilbes, Dr. Roy Armstrong, Dr. James Goodman, Dr. Yahya Masalmah and Mr. Julio Duarte for all their greatly appreciated suggestions. Special thanks to Samuel Rosario, Carmen Zayas, Adrienne Mundorf, Suhaily Cardona, Carolina Gerardino and Marcos the boat man for their great help in the ground truthing of the Enrique Reef in La Parguera Puerto Rico. Thanks to my family, including my girlfriend Cynthia Reyes, for their patience with me during my graduate studies.

Work reported herein was funded primarily by the Center for Subsurface Sensing and Imaging Systems (CenSSIS) sponsored by the Engineering Research Centers Program of the US National Science Foundation under grant EEC-9986821.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AD	Angle Detection
ATR	Automatic Target Recognizer
AVIRIS	Airborne Visible Infrared Imaging Spectrometer
CUI	Correct Unmixing Index
Dirsig	Digital remote sensing image generator
DSRCIC	Different Spatial Resolution Coregistered Image Comparison
ECHO	Enhanced Classification for Homogeneous Objects
ED	Euclidean Distance Classifier
FED	Fuzzy Euclidean Distance
FLD	Fisher's Linear Discriminant
FML	Fuzzy Maximum Likelihood
GIFOV	Ground Instantaneous Field of View
HIAT	Hyperspectral Image Analysis Toolbox
HRI	High Resolution Image
LMM	Linear Mixture Model
LOG	Laplacian of Gaussian
LRI	Low Resolution Image
MD	Mahalanobis Distance
ML	Maximum Likelihood
NN	Neural Networks
PCA	Principal Components Analysis
PCC	Percentage of Correct Classification
PSD	Power Spectral Density
RGB	Red Green Blue
SNR	Signal to Noise Ratio
SVD	Singular Value Decomposition
SVM	Support Vectors Machine
UPRM	University of Puerto Rico at Mayaguez
VTP	Variance Time Plot

LIST OF SYMBOLS

μ	Mean.
λ	Eigenvalue.
σ	Singular value.
σ^2	Variance.
dB	Decibels.
log	Base 10 logarithm.
m	Meters.
nm	Nanometers.
rad/sam	Radians Per Sample.
argmin	Argument of the minimum.
$\| \ \|_F$	Frobenius norm.
$\| \ \|_2$	Two's norm.
$P(i,j)$	= Probability density function for pixel at row i and column j.
$\cdot$	Vector inner product.

CHAPTER 1
INTRODUCTION

The performance of classification algorithms can be expressed in terms of their accuracy, their computational complexity, and their training time, among others. The objective of this work is to test the end-to-end performance of various hyperspectral image algorithms for both classification and unmixing in terms of percentage of correct classification (PCC) and correct unmixing, and using complexity measurements to generalize the results for different types of images. The phrase end-to-end indicates that the images will be processed using a predetermined sequence of algorithms using the Hyperspectral Image Analysis Toolbox (HIAT) for MATLAB as described in the methodology in Chapter 3.

The objective of this thesis work is to compare the classification algorithms developed at the University of Puerto Rico at Mayaguez (UPRM) with the most widely used algorithms in terms of image classification. All the algorithms developed at UPRM have been tested as they were developed, and accuracy improvements of the classifications have been examined for some of them. Still, the improvement in classification for an end-to-end system for many algorithms and combination of algorithms has yet to be done.

The end-to-end processing includes both pre-processing, classification and unmixing algorithms. In all cases, widely used state-of-the-art algorithms are compared to those developed at Mayaguez. For pre-processing, the widely used filter based on principal component analysis (PCA) is used along with a filter based on oversampling theory (OS) developed at UPRM. The state-of-the-art classification algorithms

included in the study are Maximum Likelihood, Euclidean Distance, Fisher's Linear Discriminant, Mahalanobis Distance, Angle Detection, Support Vector Machine, Neural Networks, and K-means clustering. The classification algorithms developed at UPRM are Fuzzy Maximum Likelihood and Fuzzy Euclidean Distance. The new spectral unmixing algorithm Non Negative Sum to One Abundance Estimation and state-of-the-art Non Negative Least Squares Abundance Estimation are also compared. The algorithms developed at UPRM and many of the state-of-the-art algorithms have been implemented in the HIAT toolbox, and will be used for these comparisons.

Obtaining the percentage of correct classification implies that we need to know the true values of the samples acquired, in this case the ground-truth for a remotely sensed image. Obtaining the ground truth for remotely sensed images is difficult because they can cover hundreds of square kilometers, they can contain a variable number of classes and they can be images from hard to reach places. Obtaining the ground truth for a linear spectral unmixing model of an image is even more difficult since the percent coverage of each class in each pixel is required. The method used here for obtaining the ground truth of the unmixing problem is to use a high spatial resolution image coregistered with the low spatial resolution hyperspectral image.

Accuracy testing of hyperspectral image analysis algorithms requires the availability of ground-truth data for verification of the results. This ground-truth data is difficult to obtain for most scenarios. Ground-truth data for spectral unmixing applications is even more difficult to obtain and is required for the accuracy testing.

Previous works [1] have used synthetic data for testing the algorithms. Since the images are generated from known data the correct composition of the pixels are known *a priori*. The problem with this approach resides in the fact that the complexities in classification associated with real and synthetic images are different.

Since most applications use these algorithms to process real hyperspectral images, testing the results with real images is imperative.

A new procedure must be developed to obtain the truth data for spectral unmixing of real images.

The accuracy of the algorithms tested depends on the complexity of the images processed. In order to compare different algorithms, images with the same complexity must be used. Different types of images must also be used to test the accuracy for a wider range of image complexities. An analytical description of image complexity must be used to express the accuracy results for specific image complexities. In addition, image complexity is used to generalize the results for different types of images. The new metrics should correlate with classification and unmixing accuracy.

The importance of this work resides in the fact that it is necessary to have a concrete set of tests that proves the accuracy improvements of the toolbox algorithms as a whole.

1.1 Objectives

- Perform an End to End processing of images using the algorithms included in the Hyperspectral Image Analysis Toolbox of Matlab, using data acquired by various sensors.
- Evaluate the accuracy results for different image complexities.

1.2 Thesis Overview

Chapter two contains a literature review of previous works which aided in the development of the methodologies used in this work.

Chapter three discusses the background theory of hyperspectral image analysis, classification, accuracy testing, image complexity metrics and other concepts used in this work.

Chapter four presents the methodology used for accuracy testing of hyperspectral image analysis algorithms. New methods to analyze spectral unmixing accuracy

are explained in detail and the use of image complexity features to analyze accuracy for distinct types of images is described.

Chapter five shows the experiments and results. The results are tabulated, plotted and explained for better understanding.

Chapter six contains the conclusions, future work and final remarks.

CHAPTER 2
LITERATURE REVIEW

The Hyperspectral Image Analysis Toolbox (HIAT) for Matlab, developed in the University of Puerto Rico at Mayaguez, is a collection of algorithms for the pre-processing and classification of multispectral and hyperspectral images [2]. The toolbox was developed by image processing researchers in the Laboratory for Applied Remote Sensing and Image Processing (LARSIP).

The end-to-end processing of hyperspectral images referred to in the objectives is divided in three steps. The first step is pre-processing. This includes resolution enhancement used to improve the signal to noise ratio of images. This results in an improvement in the accuracy of the classifications [3]. One method is based on oversampling in the spectral dimension, researched at UPRM [4].

Hyperspectral images are composed of hundreds of closely spaced narrow spectral bands which demand high computational effort at processing. Hyperspectral image classification does not typically use all the spectral bands. The next step is either to select a subset from the original bands [5] or to perform some transformation of the bands or feature extraction. Band subset selection is a procedure in which a reduced set of bands of the images is chosen, selecting the least correlated bands. This approach, researched at UPRM, is in contrast to reduction using principal components analysis where the images have to be transformed into a set of uncorrelated bands, and the physical relation to the original spectral information is lost [6–8].

The last step is classification. The HIAT permits two kinds of classification; supervised classification and unsupervised classification. It also has several implementations for abundance estimation in spectral unmixing. The algorithms employed in supervised and unsupervised classification are widely described in the literature.

These algorithms use the Euclidean Distance, Fisher's Linear Discriminant, Mahalanobis Distance, Maximum Likelihood and Angle Detection classifiers.

In supervised and unsupervised classification, it is assumed that a pixel covers a single object that should be assigned to a single class. When processing images with low resolution this assumption may have drawbacks since every pixel can be a mixture of different class spectra. Mixture models for reflectance spectroscopy of chemical mixtures have been studied since 1920 [9]. These mixing models were later adapted to remote sensing applications. Spectral unmixing solves the problem by decomposing a pixel's spectra into the constituent pure features, called endmembers, and their corresponding abundances in the pixel.

It is necessary to repeat the experiments for a diverse set of images. Image complexity measures can be used to ensure that a diverse set of images is available. An unsupervised image complexity metric has been researched since 1990. These image complexity measures have been studied to compare the performance of different automatic target recognizers (ATR). Image features for complexity analysis have been divided in [10] as target dependent and target independent. Complexity features that use complete images for their computation are called global metrics, and those that use segmentations of the images are called regional metrics. Other taxonomies include gray-level, edge or shape dependent features. Image complexity features within these taxonomies are listed in [10].

An approach used since 2000 to study the difficulty of target detection is clutter complexity analysis [11, 12]. For ATR, clutter complexity refers to the similarity of the objects in the background of the scene and the target. A weighted sum of

features for clutter complexity analysis was used in [12] to obtain a single feature that correlates better with ATR performance than each feature alone.

Edge detection was used in [13] to classify images as little complex, more or less complex, and very complex. Edge levels were defined using how much concentration a person must use to detect an edge. Centroids obtained from a fuzzy clustering process were used to establish thresholds for the image complexity classifications.

CHAPTER 3
THEORETICAL BACKGROUND

3.1 Hyperspectral Image Analysis

Hyperspectral data is composed of hundreds of images taken at closely spaced narrow spectral bands. Hyperspectral sensors can have channels that capture data from the visible to the long-wave infrared portions of the electromagnetic spectrum. In remote sensing applications, the energy or *radiance* captured by the hyperspectral sensor at each spectral band corresponds to the same scene, that is, to the same covered area and at the same time.

The contribution of each spectral band can be seen as grayscale images which can be placed together to form a three dimensional structure called a datacube.

Three measures distinguish the quality of a hyperspectral image, these are the spatial resolution, the spectral resolution, and the radiometric resolution. The spatial resolution is the area on the ground that each pixel on the image represents. The spectral resolution is the smallest difference in wavelength between two spectral bands, and the radiometric resolution is the smallest difference in intensity level that can be detected and quantified.

3.1.1 Resolution Enhancement

Like any other measured physical data hyperspectral images contain noise from different sources. The images contain a combination of sensor noise, radiance path noise and quantification noise. Several algorithms have been developed to reduce noise and improve the signal to noise ratio of images. The process of noise reduction is called denoising.

Denoising by Principal Components Analysis

The most widely used denoising method uses principal component analysis (PCA) for image enhancement. PCA use linear transformations to obtain an orthogonal set from the image bands. The hyperspectral image pixels can be modeled as

$$\underline{x} = \underline{s} + \underline{n}, \tag{3.1}$$

where $\underline{x}$ is the pixel vector, $\underline{s}$ is the spectral signature of the pure materials the pixel represents, n is white noise with covariance $\mu^2 I$ and s and n are uncorrelated. The covariance of x is given by

$$\Sigma_x = \Sigma_s + \mu^2 I, \tag{3.2}$$

where Σ_s is the covariance of s with rank $p < n$, where n is the dimension of the covariance matrix of the original signal and p is the dimension of the signal subspace as will be explained in Section 3.1.2.

Let λ_{xi} be the i-th eigenvalue of Σ_x and λ_{si} be the i-th eigenvalue of Σ_s then:

$$\lambda_{xi} = \lambda_{si} + \mu^2, \quad i = 1, 2, \ldots, n. \tag{3.3}$$

The eigenvalues are ordered in descending magnitudes. Since Σ_s is rank p, λ_{si} will be zero for $i > p$. Therefore $\lambda_{xi} = \mu^2$ for $i = p+1, \ldots, n$, that is, eigenvalues greater than p will contain only noise contribution.

A filtered image x_p can be constructed using the linear filter

$$\underline{x}_p = \mathbf{V}_p \mathbf{V}_p^T \underline{x}, \tag{3.4}$$

where $\mathbf{V}_p$ is the matrix of the first p eigenvectors of Σ_x. This increases the signal to noise ratio by discarding the contribution of the n-p lowest principal components which contribute primarily noise. This is called the reduced rank filter (RRF) method [14].

Image Enhancement by Oversampling

A method for resolution enhancement developed at the University of Puerto Rico Mayaguez (UPRM) uses oversampling theory to increase the SNR of images. Oversampling means to sample a signal at a higher rate than the Nyquist frequency, that is, higher than two times the highest frequency component of the signal.

It has been proved in [3] that oversampling filtering improve the classification accuracy .

3.1.2 Dimensionality Reduction

Hyperspectral image data is usually analyzed as vectors in a space with dimensions equal to the number of spectral bands, which in the case of hyperspectral data, the dimension would be in the range of hundreds. The high dimensionality of hyperspectral images demand high storage capacity and high computational efforts from classification algorithms. The number of training samples required for classification is the most important restriction because this amount increases exponentially with the number of bands used. This is called the curse of dimensionality.

The high dimensional data in hyperspectral images usually contain information redundancies due to strong correlation between adjacent spectral bands, and natural spectral dissimilarities of the objects in the scene. Two approaches are commonly used to summarize the information of a hyperspectral data set in order to solve the dimensionality problem. These are band transformation and band selection.

Band Transformation

The most widely used method for dimension reduction using band transformation is principal components analysis (PCA). Let $X = [\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n]$ be a hyperspectral image arranged as a two dimensional matrix , where n is the number of bands and each $\underline{x}_i$ are the bands arranged as column vectors. Using the singular value decomposition (SVD), we can factor X as follows

$$X = U\Sigma V^T, \tag{3.5}$$

where $U = [u_1, u_2, \ldots, u_n]$ and $V = [v_1, v_2, \ldots, v_n]$ are orthonormal matrices containing the left and right singular vectors respectively, $\Sigma = diag\{\sigma_1, \sigma_2, \ldots, \sigma_n\}$, $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$, and the $\sigma_i \geq 0$ are the singular values of X.

The j-th principal component is given by

$$PC_j = \sigma_j \underline{u}_j = X\underline{v}_j, \ j = 1, 2, \ldots, n. \tag{3.6}$$

Due to information redundancy in hyperspectral data it is often the case that for some singular values $\sigma_p >> \sigma_{p+1}$ for $p < n$ which implies that matrix X is near rank deficient. Therefore, most of the information will be summarized in a p-dimensional subspace that in most cases is much less than the original dimension n. Data reduction using PCA is achieved by transforming the image to the PCA space and keeping only the first p principal components which, in many cases, result in a significant reduction of dimensionality.

The percentage of variability explained by each individual principal component is given by

$$\%Var_j = \frac{\sigma_j^2}{\sum_{i=1}^{n} \sigma_i^2}. \tag{3.7}$$

After reducing the dimensionality to p principal components the total percentage of variability will be

$$\%Var = \frac{\sum_{i=1}^{p} \sigma_i^2}{\sum_{i=1}^{n} \sigma_i^2}. \tag{3.8}$$

The PCA method transforms the original hyperspetral image X into a set of orthogonal images and then keeps only those related to the first p principal components. A disadvantage of this method is the loss of physical meaningful spectral data that results from the transformation of the original image to the principal components coordinates [7].

Band Subset Selection

Band subset selection tries to both summarize information for dimension reduction and retain spatial information. Subset selection refers to selecting the most independent columns of a matrix, in this case the two dimensional hyperspectral matrix X.

The subset selection problem can be formulated as retaining the first column vectors from

$$\bar{X} = XP = \left[\bar{X}_1 \ \bar{X}_2\right], \tag{3.9}$$

where P is a permutation matrix which arranges the columns of X such that the first columns $\left(\bar{X}_1\right)$ contain the desired bands to keep. The permutation matrix P is calculated in [6–8] as the pivot matrix that results from a QR factorization of the first p right singular vectors of X. The reduced hyperspectral image is formed by keeping the first columns of $\bar{X}$ as the selected band subset.

3.1.3 Classification

Resolution enhancement and dimensionality reduction algorithms are used in the HIAT as pre-processing for classification applications. Image classification is the process of labeling the pixels of an image as pertaining to one group of pixels of the same type. Those groups are called classes. Image classification algorithms group pixels with similar features into the same class.

Classification algorithms included in the HIAT are divided into supervised classifiers and unsupervised classifiers. Supervised algorithms use training samples to derive classification rules while unsupervised algorithms group the pixels into groups or clusters.

The supervised classification algorithms used in this work are listed in table 3–1.

Table 3–1: Supervised Classifiers

Classifier	Abbreviation
Maximum Likelihood	ML
Fuzzy Maximum Likelihood	FML
Euclidean Distance	ED
Fuzzy Euclidean Distance	FED
Fisher's Linear Discriminant	FLD
Mahalanobis Distance	MD
Angle Detection	AD
Neural Networks	NN
Support Vectors Machine	SVM

The unsupervised classification algorithms included in the HIAT use k-means to obtain the means of the clusters and then use similar methods as those used by the supervised algorithms for the classification. These are listed in table 3–2.

3.2 Lineal Spectral Unmixing

In hard classification, the pixels are assumed to represent the radiance of a single object of a scene. This assumption has problems with low spatial resolution sensors where there can be many objects inside the area of a single pixel. Most of

Table 3–2: Unsupervised Classifiers

Classifier	**Abbreviation**
Maximum Likelihood	UML
Fuzzy Maximum Likelihood	UFML
Euclidean Distance	UED
Fuzzy Euclidean Distance	UFED
Fisher's Linear Discriminant	UFLD
Mahalanobis Distance	UMD
Angle Detection	UAD

the time the pixels really contain a mixture of radiance from different materials, and it is useful to identify these constituent materials and their contributions to the radiance acquired by the sensors.

The simplest model of a mixed spectrum is the linear model. The linear mixing model (LMM) states that a single pixel is constructed from a weighted combination of some pure materials, plus noise [9]. The pure materials are called endmembers. Endmembers represent the spectra acquired from a single material without noise.

Let x_j be the observed spectral signature of pixel j. Let S be the matrix of endmembers corresponding to the different contributing materials where each endmember is organized columnwise. The linear mixing model relates the measured spectra with the endmembers as follows

$$x_j = \sum_{i=0}^{P} S_i a_{ij} + w_j \quad , j = 1, 2, \ldots, N; \tag{3.10}$$

where S_i are the endmembers expressed as column vectors, a_{ij} is the abundance of endmember i for pixel j, w_j is a noise vector, P is the number of endmembers, and N the number of pixels in the image.

Abundance values represents the fraction of a pixel that each endmember covers. They must be non negative and must sum to one. That is

$$a_{ij} \geq 0 \; and \sum_{i=1}^{P} a_{ij} = 1. \tag{3.11}$$

Equation (3.10) can be written in matrix form as

$$X = SA + W, \tag{3.12}$$

where $X = [x_1, \ldots, x_N]$, $A = [a_1, \ldots, a_N]$, and $W = [w_1, \ldots, w_N]$, and N is the number of pixels.

Spectral unmixing is represented in [15] as the following optimization problem

$$\widehat{S}, \widehat{A} = \underset{S_{ij} \geq 0, a_{ij} \geq 0, \mathbf{1}^T A = 1}{\operatorname{argmin}} \|X - SA\|_F^2, \tag{3.13}$$

where $\| \ \|_F$ is the Frobenius norm. This problem is related to the positive matrix factorization problem, for which several different approaches are studied in [15, 16].

When the endmembers of an image are known the unmixing problem reduces to the abundance estimation problem. Algorithms for abundance estimation that consider the constrains in equation (3.11) are included in the HIAT.

3.3 Accuracy Testing

The validating data used for remote sensing images is called the ground truth of a scene. The ground truth is a classmap where each pixel is labeled with the most abundant material present in the area covered by the pixel.

Testing abundance estimation also requires the knowledge of the true abundance values. Ground truth for abundance estimation is difficult or impossible to obtain for some scenarios. Synthetic data has been used previously [1] by constructing abundance matrix A using gradient values according to the constraints in equation (3.11).

The accuracy of classification algorithms depend on the intrinsic complexity of the input images. An accuracy description of these algorithms should be understood for images with different complexities. While synthetic images can be used to test these algorithms including abundance estimation, real images should also be used because synthetic images do not exhibit the complexities associated with real images.

3.4 Image Complexity

Algorithms are typically tested using a variety of images to verify that the algorithm works with a wide class of images. These test images include ground truth so that the performance results can be quantified with numbers such as percentage of correct classification. Performance results are generally specific to the image used, and the percentage of correct classification will vary depending on the image. It would be nice, however to have an idea of the classification or unmixing performance of an algorithm for a specific image, even if ground truth is not available.

An image complexity metric would facilitate the evaluation of different classification algorithms by setting a standard for comparison. Publications on automatic target recognition (ATR) [10, 12] define image complexity as the degree of difficulty to detect and identify a target in a scene. The image complexity metric that this work pursues, however, is based on classification applications. This work defines image complexity as the quantified difficulty to identify the materials that compose an image and their spatial coverage.

Measuring the complexity of an image can be seen as a pattern recognition problem in which the input image is the sample that will be classified as one of several complexity classes. These classes can be discrete or continuous. For example an image can be classified as low complex, mid complex, or high complex. Also the image can be classified as a percentage in which zero will be the occurrence where there is no possibility of error and 100% will indicate that the image only contains white noise.

Several image features have been proposed in [11, 12] to objectively discriminate image complexity. These are summarized in Table 3–3.

Image complexity will be considered in this work from a classification point of view. Images with equal complexity will be those that obtain the same classification accuracy for the same classifier even when the parameters, such as the mean and

Table 3–3: Image Complexity Features

Complexity Feature	**Abbreviation**
Average Histogram Entropy	Ent
Hurst Parameter	H
Edge Detection using:	
Sobel	
Prewitt	
Roberts	
Laplacian of Gaussian	LOG
Zero-Cross	
Canny	

variance of their respective training samples, are different. Images enhanced using pre-processing will be considered as having different complexities than the original images. Variations of image noise were used to obtain images with different complexities. This work studies image noise variance as the cause of image complexity.

Average Histogram Entropy

Average histrogram entropy (Ent) is a texture feature that can be used as a clutter complexity metric and therefore an image complexity metric for classification purposes. Image entropy is defined as

$$Ent = -\sum_{i}\sum_{j} P(i,j)\log P(i,j), \tag{3.14}$$

where i and j are the spatial coordinates of the pixels and $P(i,j)$ is the probability density function of the pixels in one band of the image. $P(i,j)$ is computed using the histogram of the bands digital numbers. The entropy index for the complete hyperspectral image was computed by taking the mean of the entropy indices for each individual band. This is expressed in the following equation

$$Ent = -\frac{1}{B}\sum_{b}\sum_{i}\sum_{j} P(i,j)\log P(i,j), \tag{3.15}$$

where B is the number of bands in the image and $b = 1, \ldots, B$.

Hurst Parameter

The Hurst parameter (H) has been proposed as a feature for image complexity [11, 12]. It is a version of the fractal dimension used in time-series analysis which describes the property of long term correlation or long range dependence. Long range dependence occurs when $\sum_{k=-\infty}^{\infty} \rho(k)$ diverges, where $\rho(k)$ is the autocorrelation coefficient of the data.

The Hurst parameter is currently used for network traffic modeling because it characterizes the self similarity of the information transmitted. It is directly related to the fractal dimension (D) as $D = 2 - H$. Both fractal dimension and Hurst parameter of a surface describes its texture roughness. The Hurst parameter has been defined as follows

$$H = -\log_2 w, \tag{3.16}$$

where w is a parameter obtained from fitting a plot to a Takagi fractal curve, which is a generalized equation for certain curves that observes self-similarity.

There are different methods of calculating the Hurst parameter. The most common methods are: Variance Time Plot (VTP), R/S Analysis, Periodgram, and Whittle's Estimator.

This work uses the VTP method because of its fast computation compared to the other methods. Let $X = (X_t : t = 0, 1, 2, \ldots, N)$ be a covariance-stationary stochastic process with mean $\mu = E(X_t)$ and variance $\sigma^2 = Var(X_t)$ and N be the length of the data vector. Let $m = 1, 2, 3, \ldots, N$ be the size of non-overlapping blocks of data and $X^{(m)} = \left(X_k^{(m)} : k = 1, 2, 3, \ldots\right)$ be the new covariance stationary time series obtained by averaging the original series X non-overlapping blocks of size m. $X_k^{(m)}$ is given by

$$X_k^{(m)} = \frac{1}{m}\left(X_{km-m+1} + \ldots + X_{km}\right), \; k \geq 1, \tag{3.17}$$

where $1 < k < m$. k denotes the elements of the non-overlapping blocks of sizes $m = 1, 2, 3, \ldots, N$.

For each m, $X^{(m)}$ has a triangular form for which $X^{(1)}$ has length equal to N and $X^{(N)}$ has length equal to one. After all $X^{(m)}$ have been calculated the next step is to obtain the slope of the graph

$$f(m) = \left(\log(m), \log\left(Var\left(X^{(m)}\right)\right)\right), \ m = 1, 2, \ldots, N \tag{3.18}$$

plotted in figure 3–1.

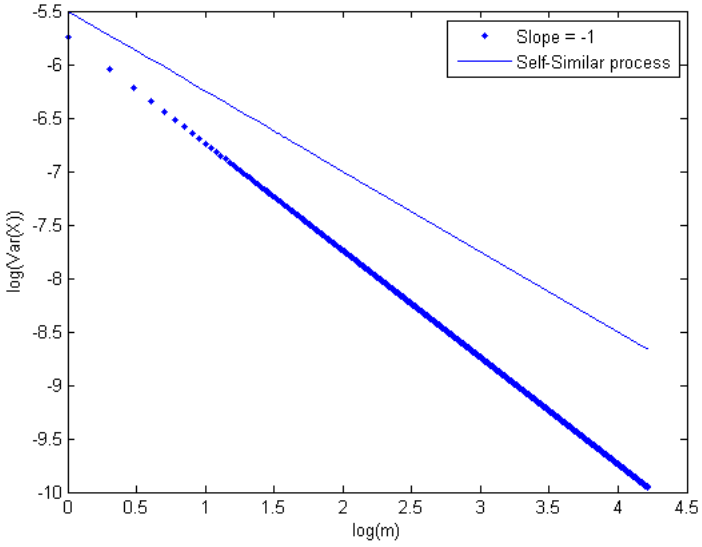

Figure 3–1: log(m) vs log(Var(X))

The slopes of graph 3–1 are used to analyse the self-similarity of a process. For this application the slope is needed to obtain a value β as

$$f'(m) = -\beta. \tag{3.19}$$

The Hurst parameter can be obtained from β as

$$H = 1 - \beta/2. \tag{3.20}$$

Edge detection

Image features that would be expected to correlate with classification accuracy include the number of objects in the scene, the inner variability of the classes, and the general randomness in the image. The number of objects in the scene can be described by the number of pixels considered as edges in the image [10]. The effect of the other two image features mentioned are commonly avoided by typical edge detection applications. Because we are looking for features that correlate with classification errors, this effect is convenient.

Algorithms for edge detection use moving window filters to detect the edges by comparing the result of the filters with some threshold. Some edge detection algorithms use different methods to calculate this threshold automatically. This work compares some of these methods which are listed in Table 3–3, to determine which has a higher correlation with image classification accuracy.

The image complexity feature based on edge detection is defined as follows.

$$Edge\ feature = \frac{1}{NB} \sum_{b} \sum_{i} \sum_{j} I\left(i, j, b\right) \tag{3.21}$$

where i, j are again spatial coordinates, b is the spectral coordinate, B is the number of bands, N is the number of pixels in the image and matrix I is the result of the edge detection algorithms which contain 1's for pixels digital numbers determined as edges and 0's for the others. Matrix I has dimensions equal to the original hyperspectral image.

All these features correlate with the percentage of correct classification. Let A be the matrix formed by the results of the complexity features for each image in the dataset. The features are ordered columnwise and the complexity feature results are organized in the rows of A. Let $\underline{b}$ be a vector containing the PCC values or accuracy

measures for each image in the dataset, using the algorithms under test. Then,

$$A\underline{x} = \underline{b}, \tag{3.22}$$

where $\underline{x}$ contain the weights each feature will receive in order to obtain a combined complexity feature with better correlation with the accuracy measures than each feature alone.

The number of rows in A is equal to the number of images tested in the dataset. The number of columns in A equals the number of complexity features used. The weights x can be obtained using some least squares optimization algorithm.

3.5 Data Description

In this work, multispectral and hyperspectral images from different sensors are used. A description of the IKONOS and Hyperion sensors and characteristics of the images are summarized next.

IKONOS is an earth observation satellite with a panchromatic sensor and a multispectral sensor that consists of four channels [17]. The four multispectral channels correspond to the blue, green, red and near infrared bands of the electromagnetic spectrum, and have four meters of spatial resolution. The panchromatic band has a higher spatial resolution of one meter. This last band can be used to upsample the other four bands to obtain multispectral images completely at one meter of spatial resolution using a method called panchromatic sharpening or pan-sharpening.

Table 3–4 shows a list of IKONOS wavelength ranges and spatial resolution for each spectral band [17].

Table 3–4: IKONOS Band List

Band	**Wavelengths**	**Spatial Resolution**
Blue	445-516 nm	4m
Green	506-595 nm	4m
Red	632-698 nm	4m
NIR	757-853 nm	4m
Panchromatic	450-900 nm	1m

Hyperion is a hyperspectral sensor capable of acquiring images with 220 spectral bands ranging from 400nm to 2500nm and 30 meters spatial resolution [18].

Hyperspectral imaging allows the idea of acquiring information from the spectrum of a scene, sometimes at the cost of sacrificing spatial resolution. Classmaps of higher spatial resolution images can be obtained, and can be used to validate the results of abundance estimation of hyperspectral images. The steps required to test the accuracy of these algorithms for abundance estimation of high spectral/low spatial resolution images is described in detail in the following chapter.

CHAPTER 4
METHODOLOGY

This chapter explains the process of obtaining accuracy measures for different algorithms for hyperspectral image analysis, and how to analyze the results for different image complexities.

4.1 Ground Truth

First, a ground truth has to be obtained to analyze the accuracy of classification and spectral unmixing algorithms of remote sensed images as explained in the following section. Ground truth data is composed of labels of the materials contained in the pixels of remotely sensed images. Their validating data is required to analyze the accuracy of the algorithms under test. Obtaining the ground truth for remotely sensed images is difficult because they can cover hundreds of square kilometers, they can contain a variable number of classes and they can be images from hard to reach places. Obtaining the ground truth for a linear spectral unmixing model of an image is even harder since the percent coverage of each class in each pixel is required.

The method used here for obtaining the ground truth of the unmixing problem is to use a high spatial resolution image coregistered with the low resolution hyperspectral image. This method is referred to as the distinct spatial resolution coregistered images comparison (DSRCIC) method.

4.1.1 High resolution image classification map

A ground truth or classmap of a high spatial resolution image is needed to estimate the abundances of lower resolution image endmembers. The classes of

this classmap must be the same as the classes and endmembers considered by the classification and unmixing algorithms respectively.

4.1.2 Low resolution image up-sampling

A low spatial resolution image will be used as input to the classification and spectral unmixing algorithms. This image is used as base image for the coregistration step. So as not to loose information, the lower resolution image is up-sampled before the coregistration takes place.

The up-sampling method used in this work is to replicate each pixel spectra. Since the original low resolution image is the one to be used as input to the classification algorithms, the up-sampling method used is trivial.

4.1.3 Coregistration

The usefulness of the ground truth image depends greatly on the results of the coregistration step. If the borders of a homogeneous region result in low unmixing accuracy values it could be because of an error in the coregistration. The correct unmixing measure described in this thesis can be used to correct the coregistration.

A polynomial algorithm is preferred for the registration since a rotation, scale, and translation algorithm can not handle the mismatches due to the acquisition angle errors, the sensor type differences, spatial shifts of using a whiskbroom platform, and other phenomena that occur in the acquisition of the images.

The values of the high resolution image classmap contain label numbers that are usually integers or natural numbers. The label numbers should not be interpolated in the coregistration step. Nearest neighbors is the resampling method of preference for this application since the values of the class map registered are not interpolated.

4.1.4 Abundance Estimates

The spatial coverage whose spectra formed a pixel of the low resolution image is also covered by a group of pixels of the high resolution image. A correspondence between the pixels of these two images has to be established. The matching between

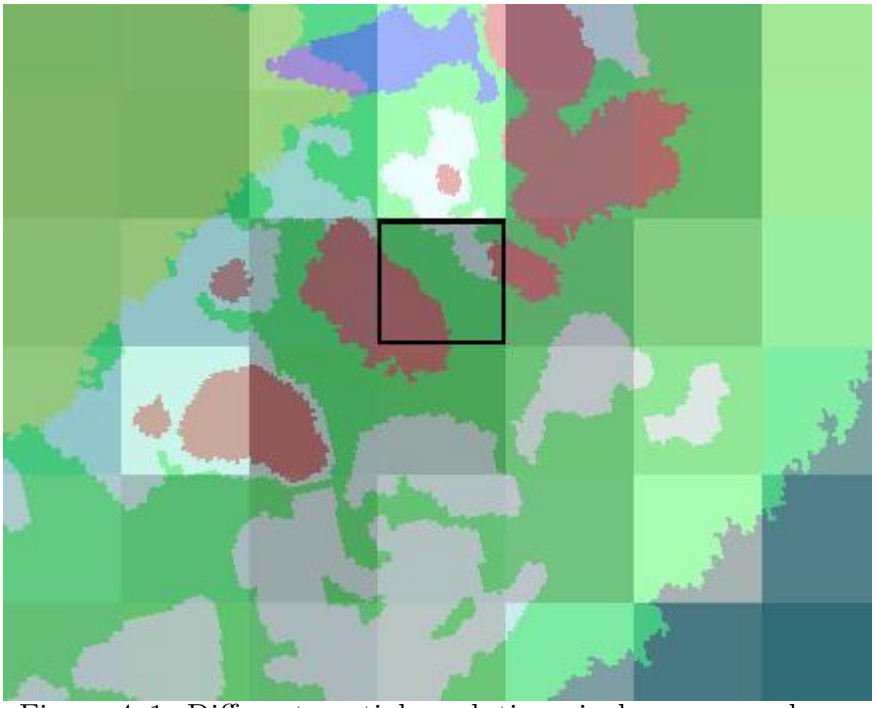

Figure 4–1: Different spatial resolution pixels correspondence

the low and high resolution pixels can be observed in figure 4–1. In the figure is easy to notice the low resolution pixels as big squares containing a high number of high resolution pixels with a more detailed spatial representation of the ground than the low resolution pixels.

It is now straightforward to determine the percentage of each class in each pixel of the lower resolution image from the classified pixels in the high resolution image. The percentage of high resolution pixels of one class inside one of the low resolution pixels forms the estimated abundance for that class. These estimated abundances for each pixel can be joined to compose the ground truth for the low resolution image.

A three dimensional matrix can be created with two dimensions corresponding to the spatial dimensions of the lower resolution image, and a third dimension for the class abundances. This can be viewed as a multispectral image, where each band contains the abundance map for one endmember. Once the ground truth

matrices have been obtained, they can be used to test the accuracy of an end-to-end processing of hyperspectral image algorithms.

4.2 End-to-End Processing

After obtaining the ground truth, the low spatial resolution images are used as input for the hyperspectral image algorithms. These algorithms include image enhancement, band subset selection, supervised and unsupervised classification and spectral unmixing.

4.2.1 Preprocessing

The image enhancement methods include denoising by Principal Components Analysis and by oversampling filtering. A dataset of images is obtained using principal components analysis for different percentage of variability and another dataset is obtained using oversampling for different cutoff frequencies. The accuracy of the classification of these datasets is compared to see which are the best cutoff frequencies and percentage of variability.

4.2.2 Dimension Reduction

The HIAT toolbox contain algorithms for dimension reduction [2] for both band transformation and band selection as explained in section 3.1.2. The band selection algorithms are

1. Information Divergence
2. Singular Value Decomposition Band Selection

The feature extraction methods are

1. Principal Components Analysis
2. Discriminant Analysis
3. Information Divergence Projection Pursuit
4. Optimized Information Divergence Projection Pursuit

This work uses band subset selection by SVD method to obtain reduced versions of the images to form a new dataset. They are treated as different images to the

original dataset because the complexity of the images is changed by the resolution enhancement and dimension reduction steps.

4.2.3 Spectral Unmixing

The endmembers used for the spectral unmixing must be for the same material types as the classes considered when classifying the higher resolution image. The results can then be directly compared to the multidimensional classification maps developed in Section 4.1.4.

The DSRCIC method only works for spectral unmixing algorithms that use the linear mixing model and the abundance non-negative and sum-to-one constraints as in equation (3.11).

4.2.4 Abundance Map Ground Truth

The next step is to quantify the results of the unmixing algorithm. This step is not trivial, as every pixel has an abundance estimate for each class. This means that each pixel will have N numbers associated with it, where N is the number of classes. When quantifying the overall unmixing performance, the abundance for each class at each pixel must be considered. This is done using a vector distance measure. This distance measure is referred to as the Correct Unmixing Index (CUI) and is defined as follows. Let

$$\max \|\underline{a} - \underline{b}\|_2 \leq K \tag{4.1}$$

where $\underline{a}$ and $\underline{b}$ are the vectors corresponding to the same pixel for the ground truth and abundance estimation result respectively. The length of the vectors is equal to the number of classes used. The value of the maximum error is used to normalize the Euclidean distances of the vectors corresponding to all the pixels of the image. That is, for every pixel, the

$$Normalized\ error\ per\ pixel = \frac{\|\underline{a} - \underline{b}\|_2}{K} \tag{4.2}$$

The normalized error per pixel is now in the range from zero to one. The CUI is then calculated as

$$CUI = 1 - Normalized\ error\ per\ pixel \tag{4.3}$$

Notice that if the distance between vectors is small, the CUI is close to one. The maximum error is the maximum Euclidean distance possible in the vector space. Since the values of the matrices are in the range of zero to one, the highest error for each variable of the vectors is one. Since the abundances for one pixel should sum to one it is only possible for two of the elements of the vectors to have distance equal to one. In other words, the maximum error possible occurs when trying to compare the abundances for two different endmembers. That is

$$\left\| \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right\|_2 = \sqrt{2} \tag{4.4}$$

The maximum distance of the two farthest vectors possible will be the square root of two. This value can now be used as the normalization factor K.

The CUI provides an error measure for the abundances of each pixel of the images. An error measure for each pixel is beneficial for spatial applications. This justifies the need for the CUI measure instead of the root mean square used in [1] which provides an error measure for the image as a whole. Error measures for the complete images are still achievable using the CUI, for example by calculating their mean. Other statistics can be derived from the CUI also.

4.2.5 Percentage of Correct Classification

A procedure similar to the DSRCIC method can be used to obtain a simulated truth data for supervised and unsupervised classifications. First, the higher resolution image must be classified and coregistered with the lower resolution image. Then the matrices with the percentage of lower resolution image pixels per class are created that enclose the pixels from the lower resolution image as before.

Now the low resolution image must be classified using the supervised or unsupervised algorithms to be tested. The class types for the training samples of the supervised classification must be the same as the ones used to classify the higher spatial resolution image. The number of clusters for the unsupervised algorithm can be equal or higher than the number of classes in the high spatial resolution image. If the number of clusters is larger than the number of classes, some can be joined with other clusters after the classification to obtain the same number of classes as in the high spatial resolution image.

To obtain a ground truth from the high resolution image, a low resolution classification map can be produced where each pixel in this low resolution map is assigned to the class having the most pixels from the high resolution image.

The ratio of all the matching pixel classes to the amount of pixels in the image gives the percentage of correct classification for the low resolution image.

4.3 Image Complexity Measures

The percentage of correct classification and the correct unmixing index are accuracy metrics that will change according to the complexity of the images processed. An approach to obtain a set of images that can be organized in increasing order of complexity is adding noise to synthetic images.

Synthetic Data

DIRSIG is a computer program developed at Rochester Institute of Technology for the generation of synthetic remote sensing hyperspectral images [19]. This

software provides for sensor, platform and atmosphere modeling. The spectral resolution can be specified and the spatial resolution is calculated based on the altitude of the platform and the acquisition angle.

Previous works using synthetic abundances with gradient values [1] specifies the noise to be added as proportional to the variance of the signal values for each band. In this work noise is added proportional to the standard deviation of the signal images because the scale of the pixel digital numbers is in the range of 10^{-4} and the variances would give values in the scale of 10^{-8} which are almost undetectable. Also, noise proportional to the standard deviation can be linearly related to the resulting image complexity.

DIRSIG also provides the ground truth for the images it generates. After the dataset of synthetic images with noise has been created, and the ground truth for those images is obtained, they can be classified using known testing samples from the ground truth.

The classified images can be evaluated using the ground truth image to obtain the percentage of correct classification for all the images in the dataset. PCC values for each image are obtained by counting the pixels classified correctly and dividing by the number of pixels in the image.

The complexity features mentioned in section 3.4 can be calculated for each image of the datasets. The complexity features can be combined as in equation (3.22) to obtain a single feature with better correlation than each feature alone. Finally, the PCC values can be graphed against the combined feature values to observe the accuracy performance of the algorithms.

4.4 Methodology Summary

The procedure to obtain the ground truth for the spectral unmixing of an image and its use to obtain accuracy measures is summarized in Figure 4–2.

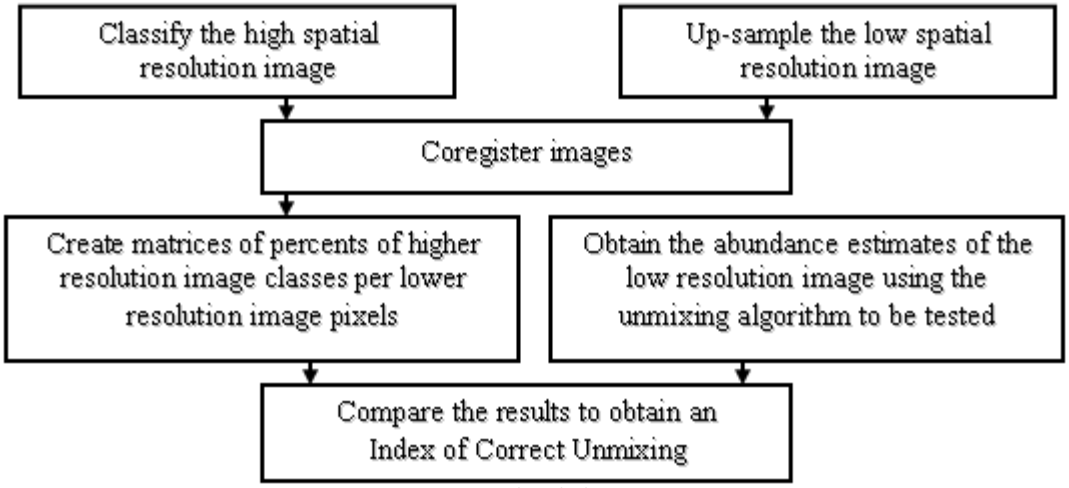

Figure 4–2: Methodology summary

The first step is to obtain the ground truth for the images to be used. In the case of lineal spectral unmixing the steps are described in Section 4.1. The steps are

1. Obtaining a high resolution image classification map.
2. Up-sampling of the low spatial resolution image.
3. Coregistration of these images.
4. Obtaining abundance estimates for the low resolution pixels from the high resolution pixels.
5. Obtaining the ground truth from the abundance maps as explained in section 4.1.4.

The second step is the end-to-end processing of the low resolution image without upsampling. The original images for which the ground truth is available are preprocessed to obtain a new set of images to be further processed and analyzed.

The ground truth is used to verify the accuracy of the classification and unmixing algorithms. The results are expressed in terms of PCC and CUI respectively.

Complexity features for both synthetic and real images are extracted and combined as in equation (3.22). These are related to the accuracy measures PCC or CUI according to the processing involved.

The last step is to analyze the accuracy of the algorithms for increasing image complexities and draw conclusions from the results.

CHAPTER 5
EXPERIMENTS AND RESULTS

This work is divided into three parts. These are the ground truth acquisition using the new method explained in Section 4.1, the end-to-end processing and the image complexity analysis explained in Sections 4.2 and 4.3 respectively. The ground truth acquired in the first experiment is used to obtain the accuracy resulting from the end-to-end processing, and the resulting accuracy to be compared with image complexity measurements.

5.1 Ground Truth Experiment

The data used for this experiment were satellite images of Enrique Reef in Southwestern Puerto Rico. The high spatial resolution image used comes from the IKONOS multispectral sensor and is shown in Figure 5–1.

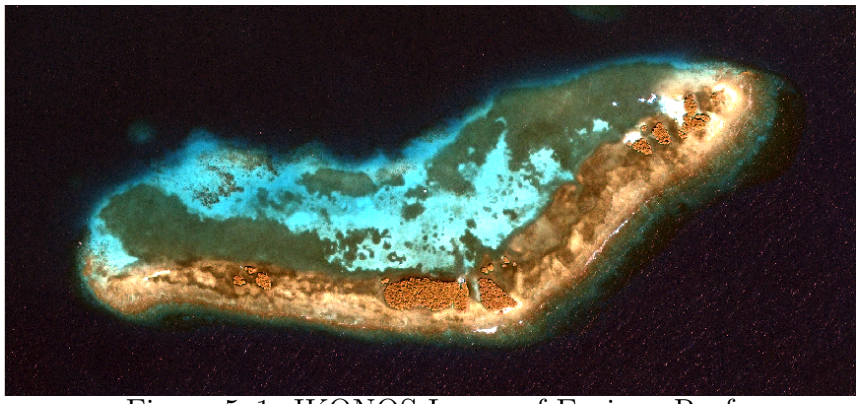

Figure 5–1: IKONOS Image of Enrique Reef

The IKONOS image was acquired in 2006 and contains four layers consisting of red, near infrared, green and blue bands. These bands were originally captured at 4m spatial resolution, but panchromatic sharpening [20] was used to obtain a multispectral image at one meter spatial resolution, as mentioned in Section 3.5.

The low spatial resolution image of Enrique Reef comes from the Hyperion sensor. This image was used as input to the classification and abundance estimation algorithms. This image has 242 bands at 30 meters of spatial resolution. Many of the spectral bands are from the absorption bands and contribute little or no information about the scene. A RGB composite of this image is shown in Figure 5–2.

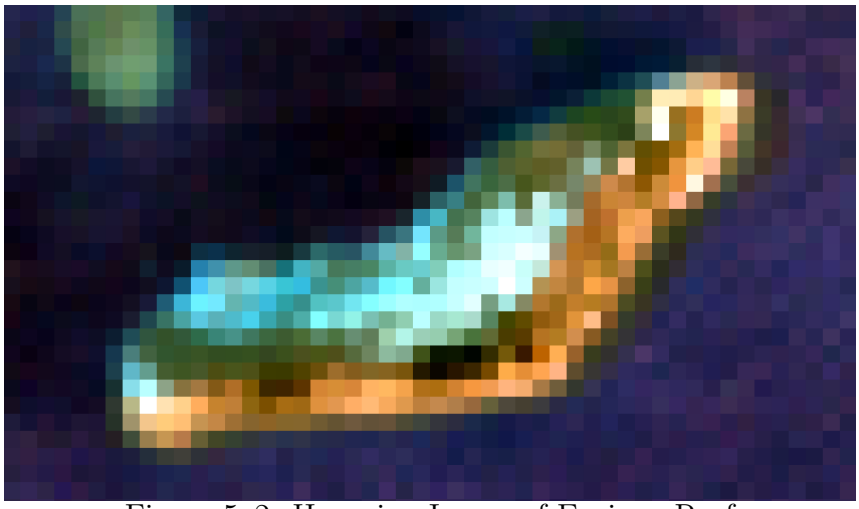

Figure 5–2: Hyperion Image of Enrique Reef

The one meter ground truth map was based on an IKONOS image and is shown in Figure 5–3. The map was used for both classification and unmixing of the hyperspectral images of the region. The classes are red for mangrove, blue for sand, green for sea grass, white for deep water and the coral color is for the reef flat zone. The ground truth map was developed as follows. The regions of interest (ROI) tool within ENVI was used to enclose the pixels that were visually similar within one polygon or ROI. A subjective approach was used to choose the ROI boundaries between one class and the others, using the four bands of the IKONOS image. During this research work, three versions of the ground truth were generated. Most of the experiments in this work used the second version of the ground truth because the third version was not available yet. The third and final version was verified using a GPS receiver by placing the IKONOS image into the GPS handheld and performing visual verification in the field.

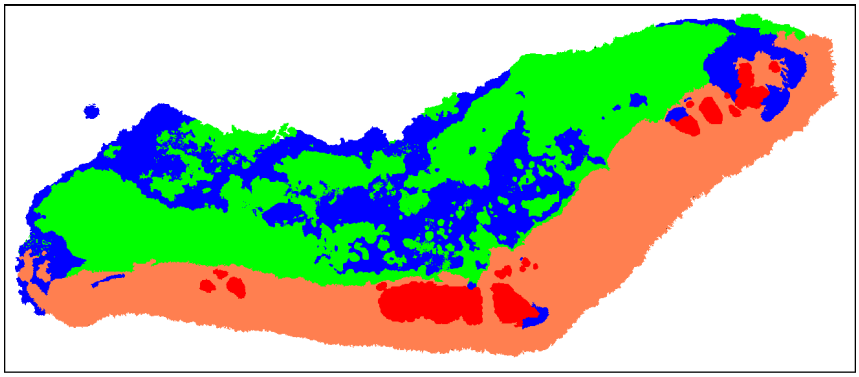

Figure 5–3: Enrique reef classification map

Coregistration

A classification map of the IKONOS classified image is used as the ground truth. The ground truth and the Hyperion images must be coregistered. The Hyperion image was up-sampled first as explained in Section 4.1.2. A rescale factor of 30 by 30 was used to obtain pixels the same size as the ground truth.

The Hyperion image and ground truth were coregistered using a polynomial method within ENVI. The IKONOS image was warped using the Hyperion image as base for the coregistration. The resampling method used is nearest neighbors because it does not interpolate the classmap labels.

The polynomial method used for the coregistration was of degree two. The degree of the polynomial method was low because of the number of ground control points available. Figure 5–4 shows the low resolution image overlaid on the high resolution image after the coregistration.

Abundance Maps

The next step is to form the abundance maps from the classmap pixels enclosed in the Hyperion pixels. The classes and endmembers used for these images are reef flat, sea grass, sand, mangrove and deep water.

Reef flat is the technical name for one of the zones into which a coral reef is divided. It is a flat area behind the fore reef and reef crest. These other two areas protect the reef flat by absorbing the impact of the waves. The reef flat of Enrique

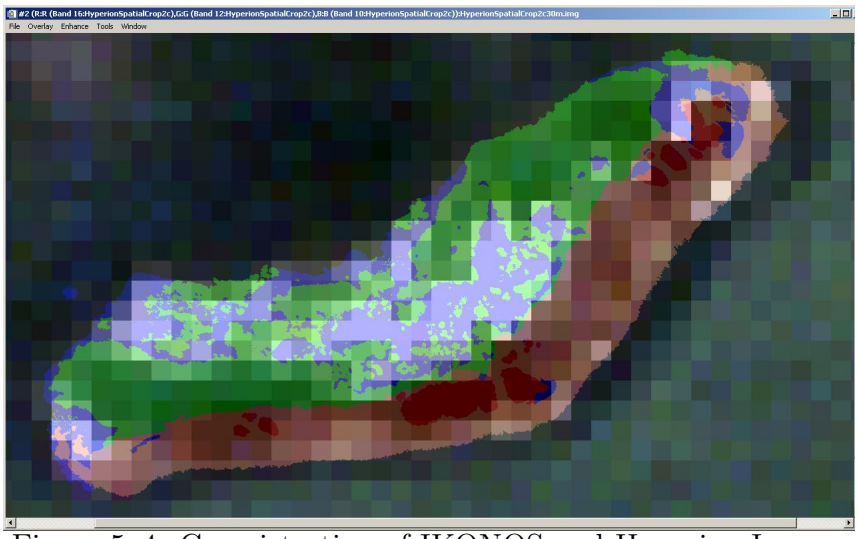

Figure 5–4: Coregistration of IKONOS and Hyperion Images

Reef contains a mixture of many components such as dead and living coral, sand, and sea grasses. This makes it perfect for spectral unmixing analysis.

The sea grass regions are mainly composed of *Thalasia testudinum*. These regions are the ones that have the faster change over time.

The area covered by mangrove changes slowly over time, and they are easy to identify with classification algorithms because of their spectral separability from the other classes.

Sand is found underwater and over the reef flat. Sand over the reef flat is combined with dead coral deposited by the waves during storms. For this experiment both sand regions are considered as one class.

Regions more than five meters deep are considered as deep water. The water column over the other classes is considered as noise for the classification algorithms.

Estimated abundances were calculated for these five endmembers as explained in section 4.1.4 to form the ground truth image.

5.2 End-to-End Processing Experiments

The end-to-end processing for the classification experiments consists of pre-processing, dimension reduction and classification. This is represented in figure 5–5. A set of different images to be classified can be obtained by varying the variability

percentage and cutoff frequency of the PCA and OS filter used as pre-processing respectively.

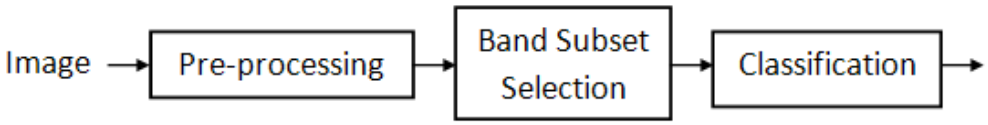

Figure 5–5: End-to-End Classification Testing Diagram

The number of bands of the Hyperion image was reduced to 10 bands using band subset selection. Figure 5–6 shows the five endmembers that will be used later for the unmixing algorithms along with the ten dashed vertical lines indicating the center wavelengths of the bands kept after the dimension reduction. These center wavelengths are tabulated in Table 5–1.

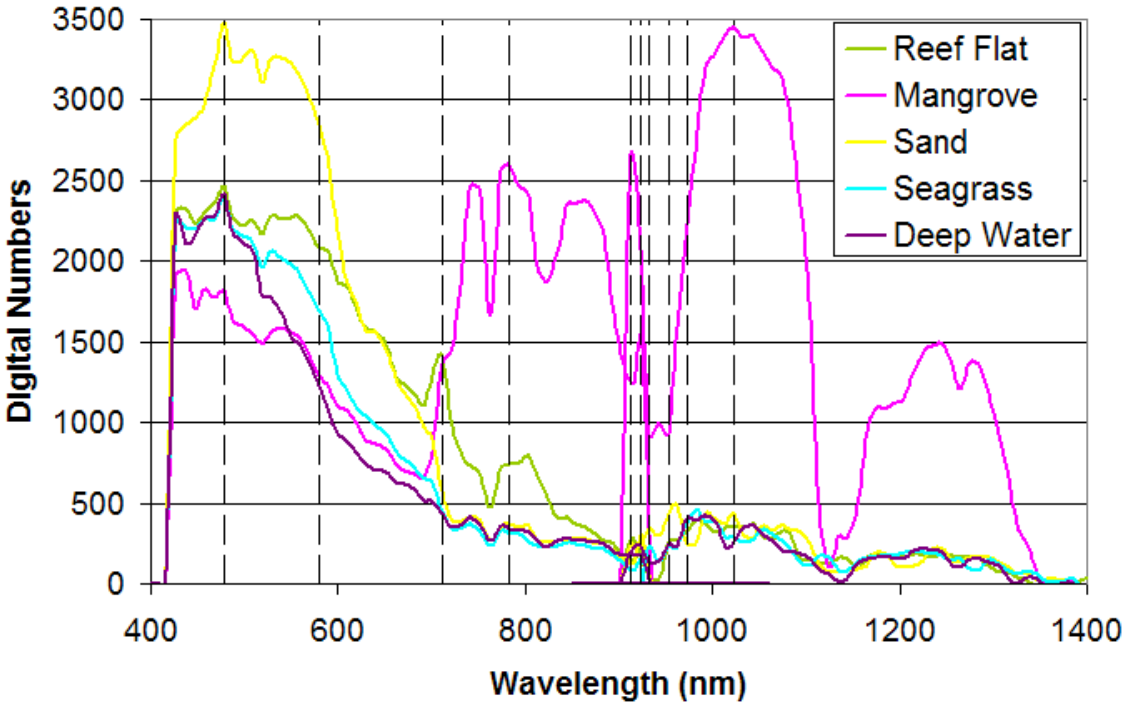

Figure 5–6: Endmembers and bands selected.

Classification Results

The same ground truth composed of abundance maps can be used to test classification accuracy. The endmember associated with the highest abundance value for one pixel will be the selected class for that pixel. The same can be done to all

Table 5–1: Center Wavelengths of bands after the band subset selection.

Band Number	Wavelength (nm)
13	477.69
23	579.45
36	711.72
43	782.95
77	912.45
78	922.54
79	932.64
81	952.82
83	972.99
88	1023.4

the pixels to form a ground truth for the low resolution image, in this case, for the Hyperion image.

Table 5–2 shows the average PCC for the supervised classification of the Enrique reef Hyperion image using different resolution enhancement thresholds. The left column shows the results for the classification algorithms developed at UPRM. These are compared with the PCC for state-of-the-art classifiers.

Table 5–2: Average correct classification for supervised algorithms.

UPRM Classifiers	**PCC (%)**	**State-of-the-art Classifiers**	**PCC (%)**
FML	11.5443	ML	9.4027
FED	86.6026	ED	86.6026
		FLD	88.1527
		MD	69.8048
		AD	86.5530
		SVM	87.3240
		NN	85.3706

Table 5–3 shows the PCC for unsupervised classification algorithms. FLD obtained the highest accuracy as in the previous experiments. FML and FED algorithms obtained higher accuracies than their non fuzzy counterparts.

Confusion matrices provide more information to analyze the results of a classification than the PCC alone. The accuracies for the individual classes can be

Table 5–3: Average correct classification for unsupervised algorithms.

UPRM Classifiers	**PCC (%)**	**State-of-the-art Classifiers**	**PCC (%)**
FML	74.53963	ML	74.43473
FED	74.33275	ED	73.91608
		FLD	81.2267
		MD	64.13459
		AD	81.10433

calculated from the confusion matrix in two ways[21]: the user's accuracy and the producer's accuracy.

$$User's\ Accuracy_i = \frac{n_{ii}}{\sum_j (n_{ij})};\ i = 1, \ldots, 5;\ j = 1, \ldots, 5; \tag{5.1}$$

where n are the elements of the confusion matrix, the first subscript of n indicates the row number and the second subscript the column number. The user's accuracies are calculated for each individual row i.

Using the same nomenclature than above, the producer's accuracies are calculated for each individual column j as:

$$Producer's\ Accuracy_j = \frac{n_{jj}}{\sum_i (n_{ij})};\ i = 1, \ldots, 5;\ j = 1, \ldots, 5. \tag{5.2}$$

The confusion matrices with the user's and producer's accuracies for FLD and ED classifiers are shown in Tables 5–4 and 5–5 respectively.

5.3 Spectral Unmixing Results

The ground truth for abundance estimation obtained above can be used to obtain the correct unmixing index (CUI) as explained in Section 4.2.4.

Table 5–4: Confusion Matrix for the ED algorithm.

			Actual:			**User's**
	Coral	**Mangrove**	**Sand**	**Seagrass**	**Water**	**Accuracy**
Predicted:						
Coral	94	5	16	22	0	68.6%
Mangrove	4	5	0	0	0	55.5%
Sand	0	0	56	7	0	88.9%
Seagrass	60	0	22	115	20	52.9%
Water	7	5	17	23	952	94.8%
Producer's Accuracy	56.9%	33.3%	50.5%	68.9%	97.9%	

Table 5–5: Confusion matrix for the FLD algorithm.

			Actual:			**User's**
	Coral	**Mangrove**	**Sand**	**Seagrass**	**Water**	**Accuracy**
Predicted:						
Coral	133	4	5	1	0	93.0%
Mangrove	2	11	0	0	0	84.6%
Sand	0	0	63	23	0	73.3%
Seagrass	19	0	28	125	5	70.6%
Water	11	0	15	18	967	95.6%
Producer's Accuracy	80.6%	73.3%	56.8%	74.9%	99.5%	

Abundance Estimation Results

Table 5–6 shows the CUI for non-negative sum to one abundance estimation and least squares abundance estimation. The first algorithm was developed at UPRM and both are implemented in the HIAT.

The accuracy for both unmixing methods show which are the best variability percentages to use for the PCA filter and cutoff frequencies for the OS filter.

The accuracy for each individual endmember can be analyzed using the CUI for one pixel element at a time. Results for the images and algorithms used are tabulated in Table 5–7. These measures are useful to analyze the effect of the endmembers chosen on the final accuracy of the image unmixing.

Table 5–6: Correct Unmixing Indices for the Enrique Reef dataset.

Image	**Non negative sum to one CUI**	**Non negative Least Squares CUI**
No Resolution Enhancement	0.8057	0.7897
PCA Filter - 95% Variability	0.8135	0.8026
PCA Filter - 96% Variability	**0.8559**	**0.8557**
PCA Filter - 97% Variability	0.8443	0.8328
PCA Filter - 98% Variability	0.8035	0.7938
PCA Filter - 99% Variability	0.8246	0.7844
OS Filter - 0.1 rad/sam	0.8576	0.8438
OS Filter - 0.2 rad/sam	**0.8590**	0.8524
OS Filter - 0.3 rad/sam	0.8572	**0.8530**
OS Filter - 0.4 rad/sam	0.8517	0.8477
OS Filter - 0.5 rad/sam	0.8454	0.8456
OS Filter - 0.6 rad/sam	0.8542	0.8155
OS Filter - 0.7 rad/sam	0.8377	0.7976
OS Filter - 0.8 rad/sam	0.8063	0.7889
OS Filter - 0.9 rad/sam	0.8103	0.8076

Abundance estimates for reef flat and mangrove are the most accurate with CUI above 0.95. Nevertheless, good results are achieved for all of them. The following pages give details of the unmixing results for each endmember.

Figures 5–7 and 5–8 show the ground truth and estimates for sea grass abundances respectively.

Notice that the abundance map result using the unmixing algorithm detects a sea grass region on the top left corner of the image and the ground truth does not. The reason is that the original spatial crop of the high resolution image in figure 5–1 did not have pixels for that region, and it will not have pixels for that region after the coregistration neither. Care must be taken about this effect of image

Table 5–7: Average CUI for abundances of each class.

Endmember Category	**Non negative Sum to One**	**Non negative Least Squares**
Reef flat	0.952 ± 0.003	0.950 ± 0.004
Mangrove	0.970 ± 0.009	0.975 ± 0.009
Sand	0.848 ± 0.034	0.870 ± 0.031
Sea Grass	0.841 ± 0.007	0.883 ± 0.023
Deep Water	0.816 ± 0.041	0.756 ± 0.051

coregistration. Those extra pixels should not be included when calculating the PCC and CUI values.

The time between the acquisition of the high and low resolution images also affects the abundance estimates due to changes in spatial coverage of the endmembers. It is preferred that both low and high resolution images for the ground truth be taken about the same time.

Figures 5–9 and 5–10 shows the ground truth and abundance map for sand respectively. Notice that the blue region at the bottom of the unmixing result is lighter than the bottom region of the ground truth. A possible reason could be the negative effect of waves that comes from the bottom right to the upper left part of the image. The upper region is deep water because the reef diminished the waves before affecting that region.

The waves effect results in a false detection of more abundance for other endmembers in what should be deep water. In this case, it falsely detects a greater abundance of sand than what should be in that region.

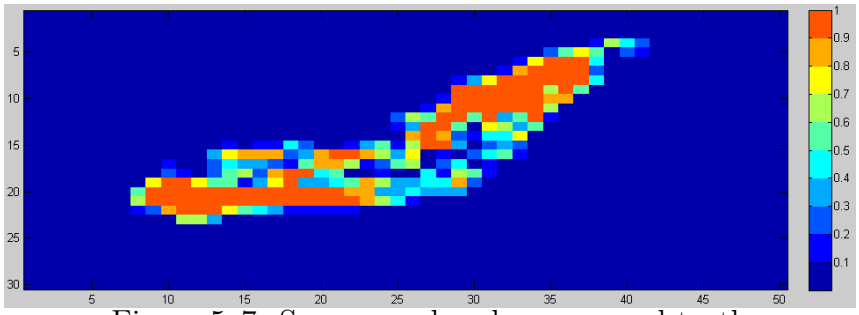

Figure 5–7: Sea grass abundance ground truth

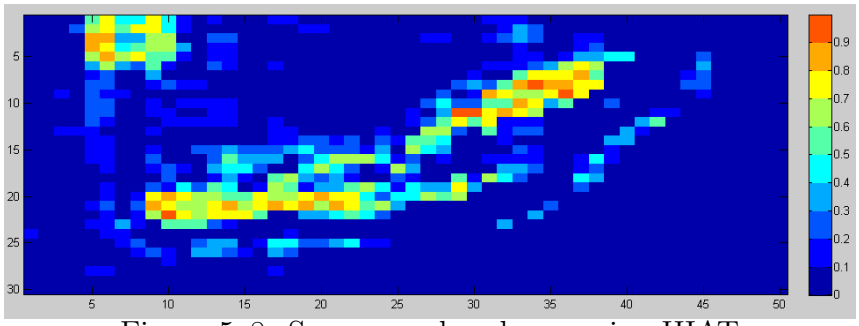

Figure 5–8: Sea grass abundance using HIAT

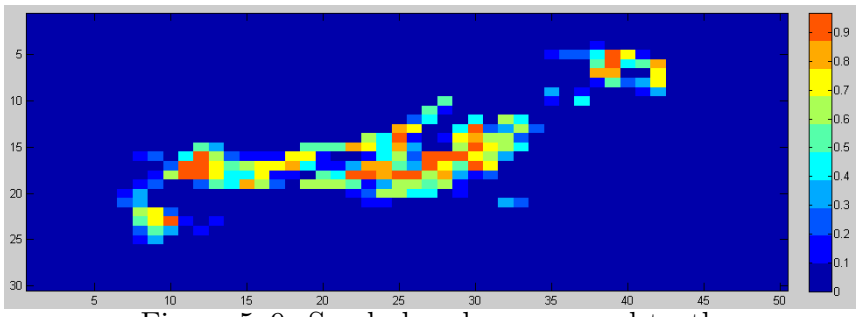

Figure 5–9: Sand abundance ground truth

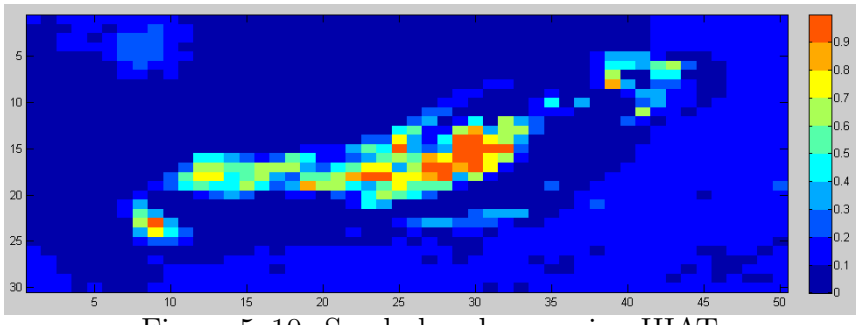

Figure 5–10: Sand abundance using HIAT

Figure 5–11 shows the ground truth for deep water abundance maps. Only for this image the orange regions are deep water and the blue regions are covered by other endmembers.

The blue area is less than 5 meters deep, while the orange area is more than 5 meters deep. The shallow area on the top left corner of the unmixing result, which is not found in the ground truth, results from the coregistration effect explained before.

The waves effect is present in the unmixing result by decreasing the abundance values for deep water. Abundance values on the bottom right of the image were reduced below 90%.

Figures 5–13 and 5–14 shows the ground truth and abundance map for reef flat respectively.

The reef flat contains a mixture of many materials smaller than the resolution of the IKONOS image. The classmap of the IKONOS image does not provide the spatial resolution required for the estimation of abundances in this region. Pixels that have above 90% of reef flat abundance in the ground truth cover a wider area than the unmixing algorithm because the unmixing algorithm detects more abundance for other endmembers in this region. The ground truth has less abundance from other endmembers in the reef flat region due to the limitation in spatial resolution.

Figures 5–15 and 5–16 shows the ground truth and abundance map for mangrove, respectively.

Abundance estimation for mangrove obtained the highest accuracy due to several reasons. The mangrove radiance captured was the least affected by water, because it is the only class that is completely above the water. The spatial coverage of mangrove for this scene was big enough to be well defined by the high resolution image pixels. The spatial coverage of mangrove changes slowly over time. The most important reason is that the spectral separability for mangrove is higher than

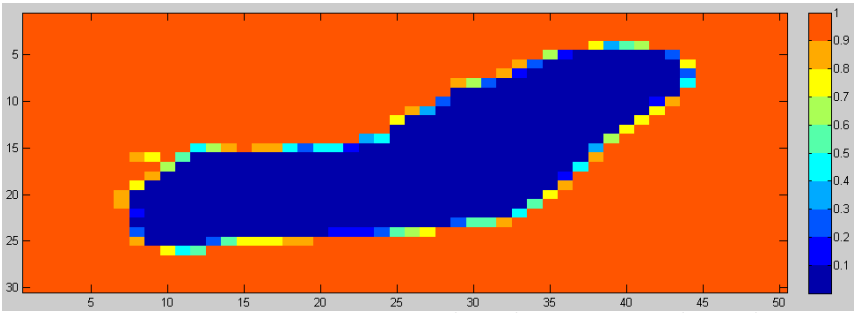

Figure 5–11: Deep water abundance ground truth

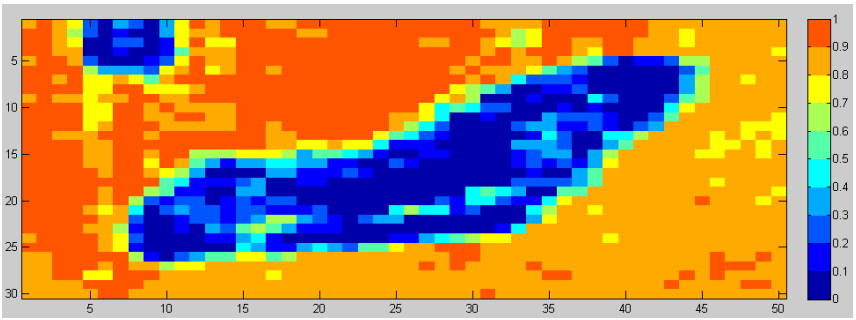

Figure 5–12: Deep water abundance using HIAT

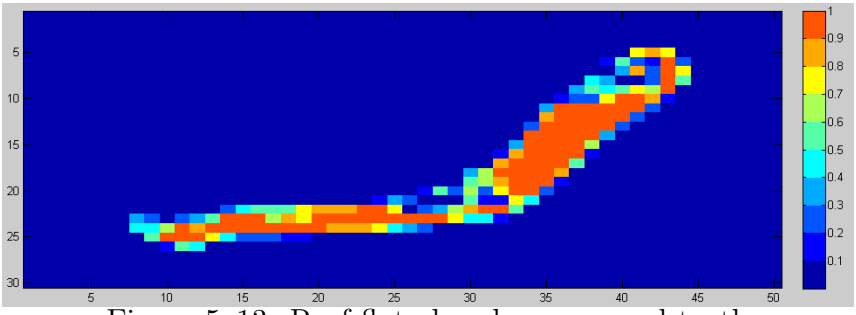

Figure 5–13: Reef flat abundance ground truth

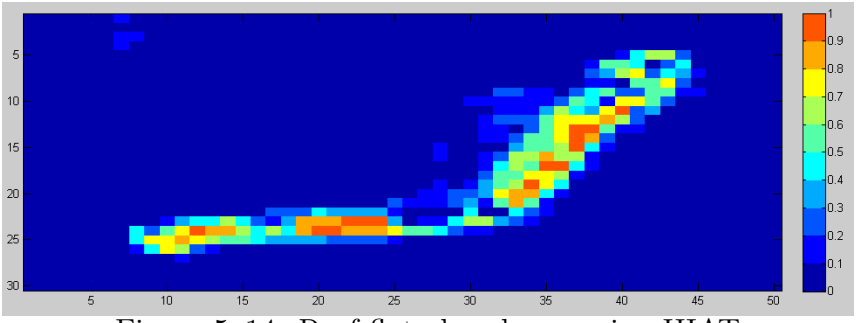

Figure 5–14: Reef flat abundance using HIAT

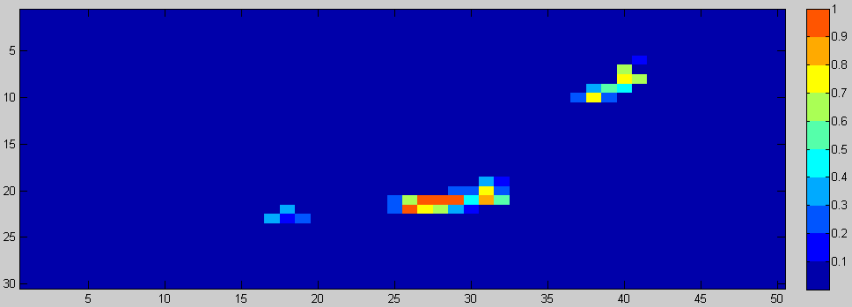

Figure 5–15: Mangrove abundance ground truth

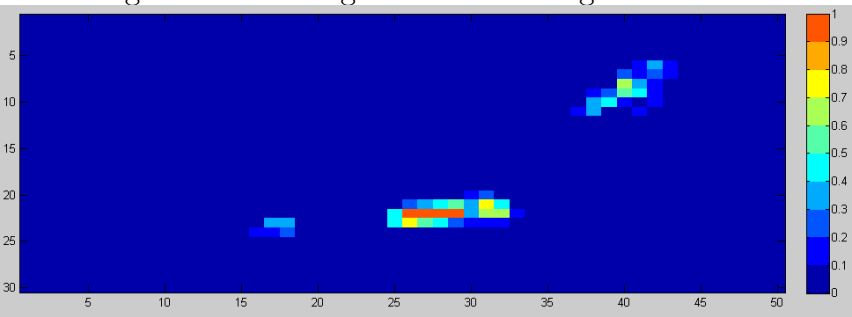

Figure 5–16: Mangrove abundance using HIAT

among the other endmembers. Table 5–8 shows the cosine angle distances between endmembers and that the highest separation occurs when comparing mangrove with the other endmembers.

Table 5–8: Cosine angle distance between endmembers.

	Reef flat	Mangrove	Sand	Sea Grass	Deep Water
Reef flat	1	**0.59**	0.98	0.98	0.96
Mangrove	**0.59**	1	**0.54**	**0.59**	**0.62**
Sand	0.98	**0.54**	1	0.99	0.97
Sea Grass	0.98	**0.59**	0.99	1	0.99
Deep Water	0.96	**0.62**	0.97	0.99	1

5.4 Image Complexity Experiments

Image complexity analysis uses image features to classify an image according to its expected accuracy for some algorithm. Research in the area of an image complexity metric fully independent of the processing involved is not very advanced.

This work uses DIRSIG to generate synthetic remote sensing images with sensor characteristics that imitate those of the Airborne Visible Infrared Imaging Spectrometer (AVIRIS) sensor. AVIRIS is a hyperspectral sensor mounted in an airplane

platform [22]. It consists of 224 spectral bands ranging from 400 to 2500nm. Images were generated which simulate a forest scene and a urban scene, each one with dimensions equal to 128×128 pixels by 224 bands. Ground truth for both images are generated also.

The materials contained in the forest scene are grass, dirt, deciduous and water. The urban scene contains grass, water, dirt, asphalt, roof, shingle and deciduous.

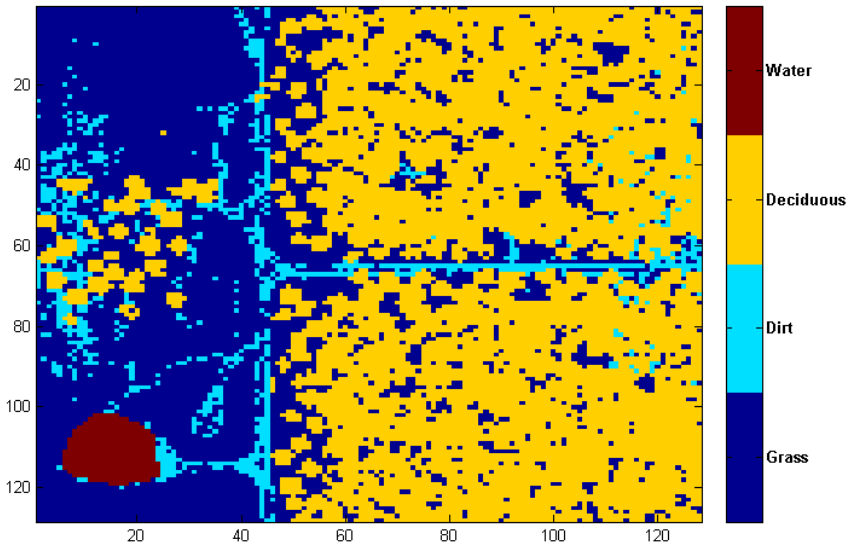

Figure 5–17: Forest Ground Truth Image

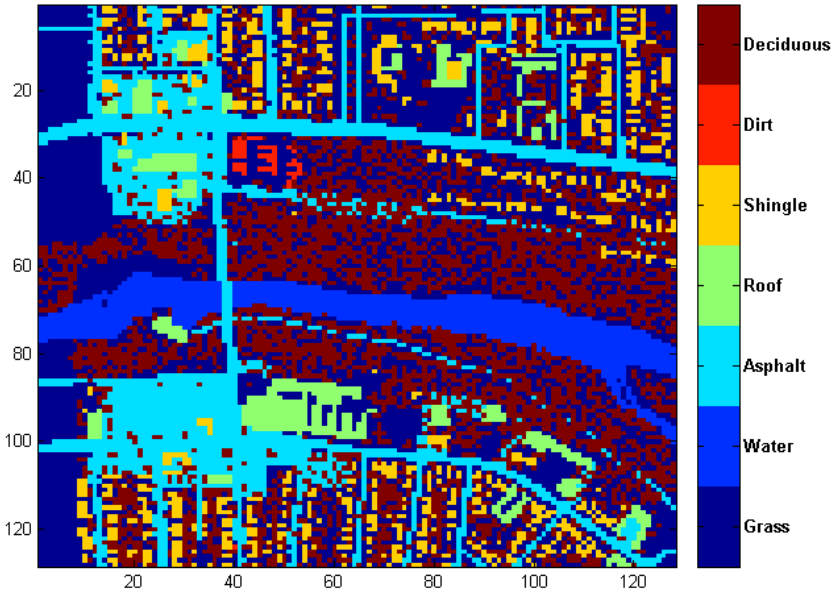

Figure 5–18: Urban Ground Truth Image

Noise was added to both scenes to obtain a dataset of images with increasing complexity. The noise added for each band of the images was proportional to the standard deviation of each band. This ensures noise values in a scale proportional to the signal values which is the norm for remote sensing images [1].

The datasets were classified using the seven classifiers listed in Table 3–1. A major problem with the classifications of the forest scene were the shadows of deciduous trees which were classified as water. In spite of this, the forest scene was used because the classification accuracy of the dataset decreases by increasing noise variance.

PCC values were obtained for both sets of images using their corresponding ground truth images. The image complexity features from Table 3–3 were used to obtain the complexity measures for the forest and urban datasets. The correlation coefficients of the complexity measures with the percentage of correct classification are listed in Table 5–9. These correlation coefficients use PCC values for the two scenes using Fisher's linear discriminant algorithm.

The correlation coefficient provides a measure of the linear relationship between two variables. The relationship between the complexity features and the percentage of correct classification is not necessarily to be linear, but previous works [11] have used correlation coefficients as an indicator of the usefulness of an image feature as a complexity feature.

A linear combination of the complexity features can be used to obtain a single feature that correlates better with the PCC values than each feature alone. Let A be an $N \times m$ matrix composed of m image features organized columnwise for a dataset composed of N images, let $\underline{b}$ be the $100 - \underline{PCC}$ vector for a specific classification experiment, and let $\underline{\omega}$ be a vector with dimension $m \times 1$ composed of the weights for each complexity feature. Then the linear combination for the estimation of the

Table 5–9: Correlation coefficients between image complexity features and PCC values.

Complexity Feature	**Correlation Coefficients for the Urban Scene**	**Correlation Coefficients for the Forest Scene**
Hurst Parameter	-0.9676	-0.9609
Entropy	-0.9930	0.3380
Edge Detection using:		
Sobel	0.9011	0.9057
Prewitt	0.9103	0.9164
Roberts	0.8666	0.8838
Laplacian of Gaussian	-0.9601	-0.9616
Zero-Cross	-0.9601	-0.9616
Canny	-0.9180	-0.9228

accuracy error values $100 - \underline{PCC}$ is

$$A \cdot \underline{\omega} = \underline{b} \tag{5.3}$$

Given A, $\underline{b}$ and $N > m$, $\underline{\omega}$ can be calculated using least squares. After the weights $\underline{\omega}$ have been calculated they can be multiplied to the same features of a different set of images to obtain a single complexity feature.

Table 5–10 shows the correlation coefficients for the linear combination of complexity features with the same Fisher's discriminant analysis results as Table 5–9. The linear combination of complexity features increased the correlation coefficients for different combinations of image features.

Table 5–10: Correlation coefficients and MSE of the lineal combination of complexity features used to estimate the PCC of the FLD classification of the Urban dataset.

Complexity Features	**Correlation Coefficients**	**Mean Square Error**
Entropy and LOG Edge	-0.9118	37.4725
Hurst and LOG Edge	-0.9401	25.8363
Hurst and Entropy	-0.9908	3.9411
Using the three Features	-0.9979	0.8935

Table 5–10 indicates that the LOG edge detection feature decreased the correlation coefficients of the Hurst parameter and entropy features used alone. The last two rows indicate that the LOG edge detection feature decreased the MSE and

increased the correlation coefficient of the other two features combined. The best accuracy estimate was achieved using the three features combined.

The same experiment was repeated for the SVM classifier. The SVM and FLD classification results for the urban scene are plotted in Figure 5–19 against the complexity index that uses the three features discussed previously. Plots of different combinations of two complexity features are shown in Figures 5–20, 5–21, and 5–22. The same training samples were used for both classifications of the dataset. The weights $\underline{\omega}$ were calculated for each individual image using a method called leave-one-out [23] for not to include the same image accuracy to be estimated into the training of the weights. The estimated accuracy curves were almost the same using leave-one-out and not using it.

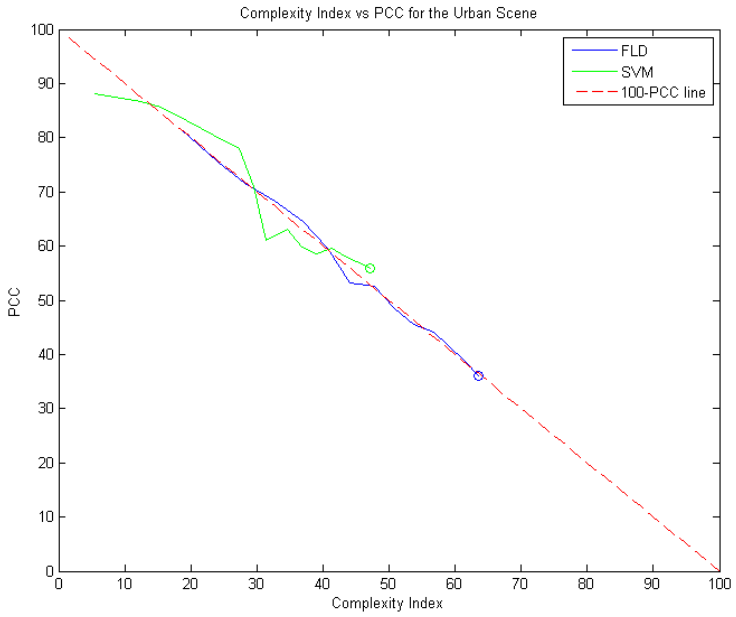

Figure 5–19: Combined complexity feature vs PCC for the urban scene.

The two circles at the end of the curves in Figure 5–19 correspond to the same image. Note that the complete SVM curve is above the FLD curve but displaced to the left, trying to adjust to the $100 - \underline{PCC}$ line. The figure shows that using different classifiers provide different complexity measures for the same image. The complexity metric for each image in the dataset should be unique.

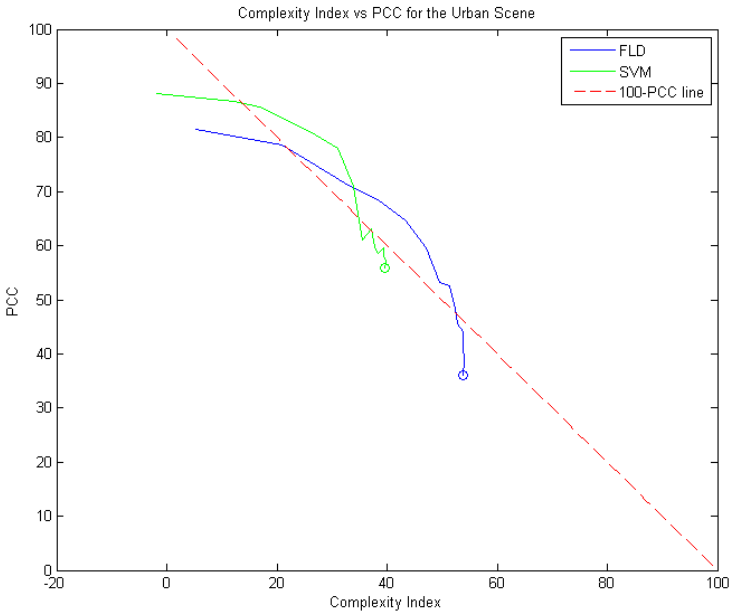

Figure 5–20: Entropy and LOG Edge vs FLD PCC

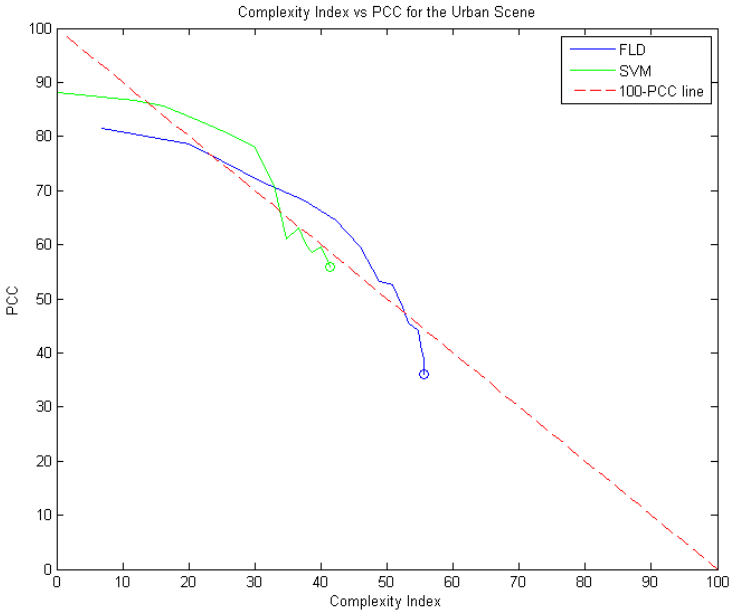

Figure 5–21: Hurst and LOG Edge vs FLD PCC

The weights obtained for the features of a dataset were multiplied by the complexity features of other datasets and the results are not reliable as accuracy estimates. Sometimes the complexity results were negative which is not defined for image complexity. A better approach to analyze the complexity features is by plotting them along with the PCC values for different classifiers.

For this analysis, a third dataset of real images of a tank filled with water was acquired using an SOC700 hyperspectral camera. This camera provides images with 120 bands from 430nm to 900nm and acquires 640 pixels per line by 640 lines. The tank contained some coins, a spoon and a wood stick. Different classification complexities were obtained by adding titanium dioxide to the water. Figure 5–23 shows a true color composite of the tank scene without titanium dioxide and figure 5–24 with titanium dioxide. Figure 5–25 shows the ground truth for the images.

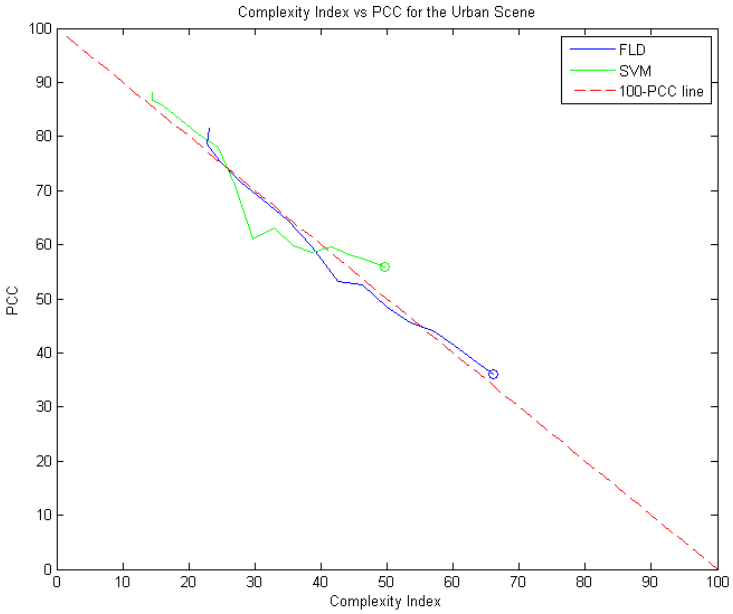

Figure 5–22: Hurst and Entropy features vs FLD PCC

Figure 5–23: Tank image without titanium dioxide.

Figure 5–24: Tank image with titanium dioxide.

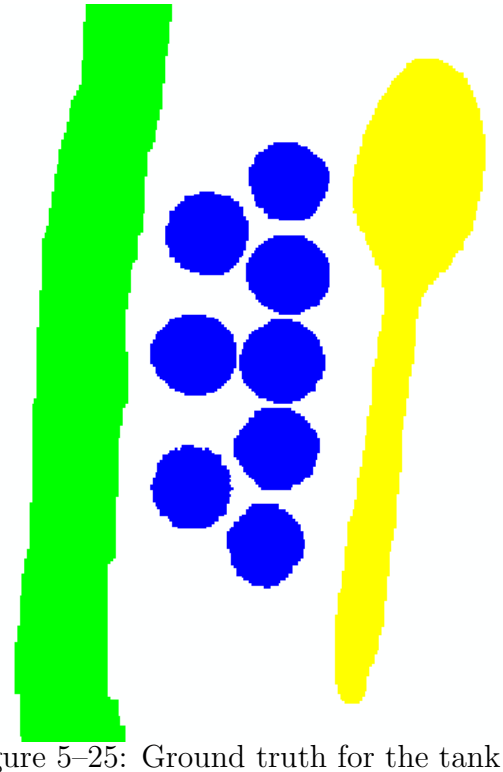

Figure 5–25: Ground truth for the tank images.

Figure 5–26 shows the plot of the LOG edge detection features versus the FLD percentage of correct classification for the Tank, Urban, and Forest image datasets. PCC for the three individual datasets decreases as the edge feature increased, but between different datasets the same cannot be said. This reflects the need for using various complexity features.

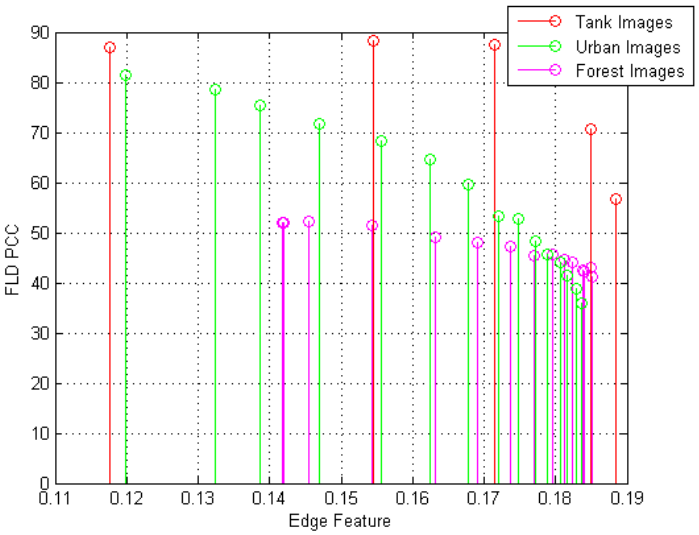

Figure 5–26: LOG edge feature vs FLD PCC.

Figure 5–27 shows the plot of the H parameter and LOG edge detection features versus the FLD percentage of correct classification for the Tank, Urban, and Forest image datasets. The Hurst parameter has a good correlation coefficient with PCC values for the individual datasets but not between them as shown in Figure 5–28. The same can be said about the entropy and edge detection features as shown in Figures 5–29, 5–30, and 5–31

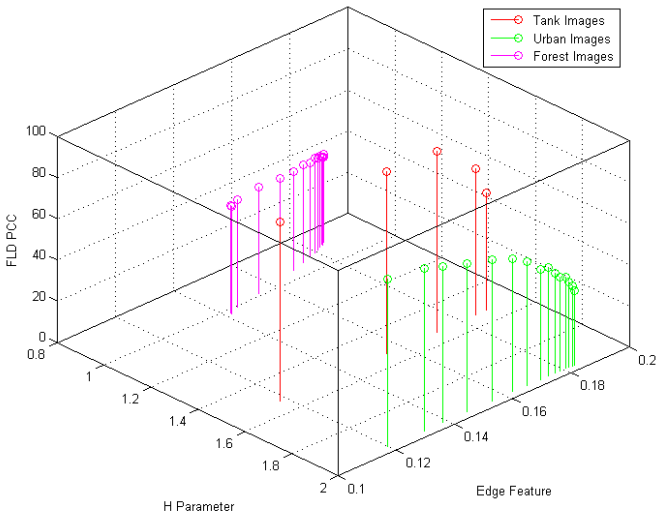

Figure 5–27: H parameter and LOG edge detection vs FLD PCC.

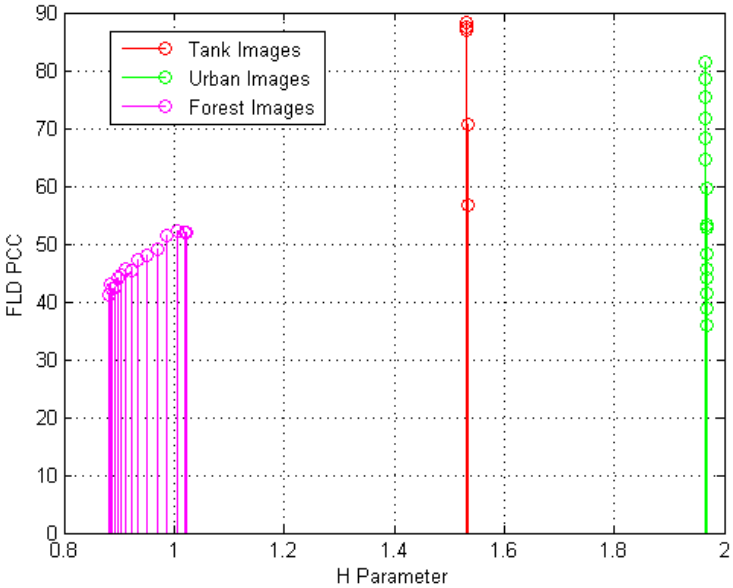

Figure 5–28: Hurst parameter vs FLD PCC.

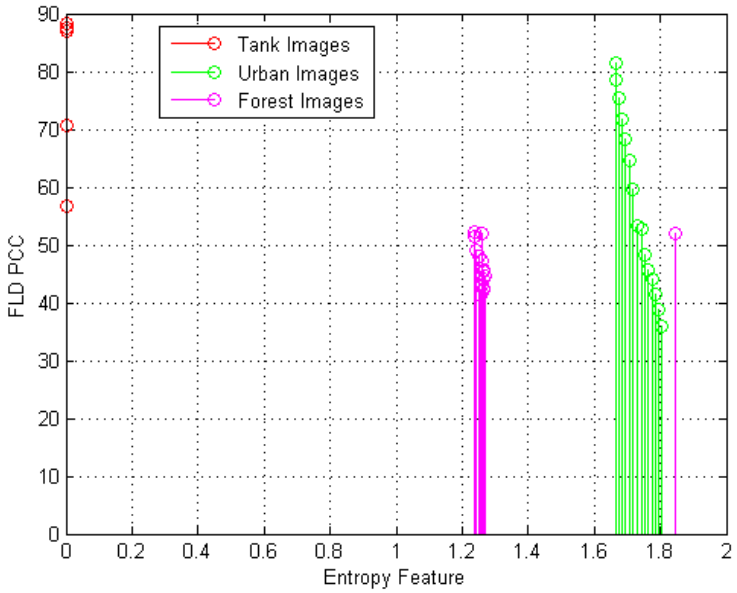

Figure 5–29: Image entropy feature vs FLD PCC.

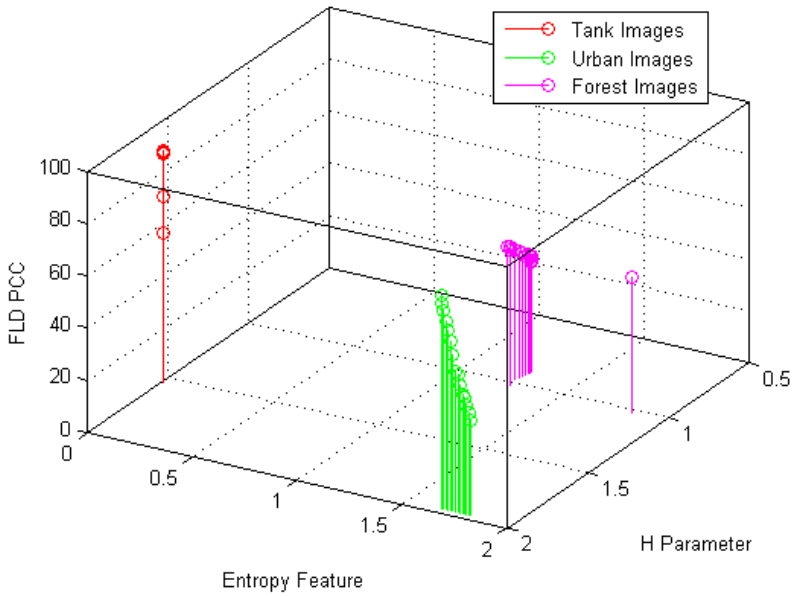

Figure 5–30: Entropy and H parameter vs FLD PCC.

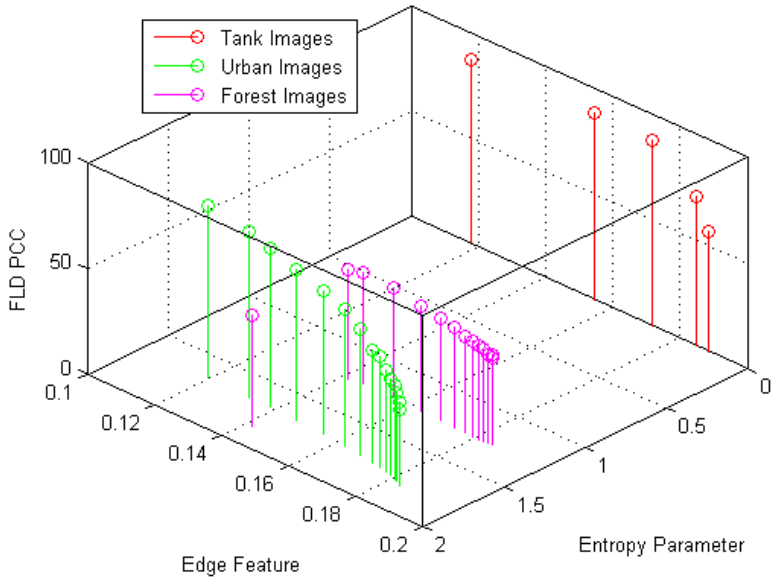

Figure 5–31: LOG edge and entropy features vs FLD PCC.

CHAPTER 6
CONCLUSIONS AND FUTURE WORK

The methodology proposed in Chapter 4 was useful to obtain ground truths for abundance estimation algorithms. This is a meaningful contribution because actually there are no other methodologies in literature for the ground truthing of real images for abundance estimation. The correct unmixing index proposed in section 4.2.4 was useful to quantize the accuracy results of the unmixing algorithms.

The results of supervised and unsupervised classifications were compared with a ground truth generated by the proposed methodology. The algorithms developed at the University of Puerto Rico proved to have accuracies comparable to state-of-the-art algorithms. Pre-processing increased the accuracy of the classifications for both supervised and unsupervised classifications.

The complexity features used correlate well with the percentage of correct classification of the datasets of images. The correlation coefficients indicated which complexity features were better on predicting the classification accuracy. The linear combination of complexity features improved the correlation between the complexity features and the percentages of correct classification. Nevertheless, the linear combination of features did not provide a unique complexity index for the images. The weights used to predict classification accuracy, calculated for a image dataset, were not useful for another datasets. Individual image complexity features are more illustrative providing information about the accuracy of the classifiers for different types of images.

Future Work

Water-column correction can be included as a pre-processing step to increase classification accuracy for underwater classes.

The procedure to obtain the ground truth for the unmixing algorithms use square regions of the high resolution image that matches the area covered for each low resolution pixel. The area on the ground that a pixel represents is not absolutely square. The area on the ground is a parameter of the remote sensing images called the ground instantaneous field of view (GIFOV). A future work would be to obtain the ground truth for the low resolution image by using pixels of the high resolution image enclosed in regions with shapes and dimensions similar to the GIFOV of the low resolution image.

APPENDIX A
HURST PARAMETER COMPUTATION

Appendix B contains the MATLAB code developed to compute the Hurst parameter of hyperspectral images using the variance time plot (VTP) method.

```
% Calculates the Hurst Parameter for the urban images.
N = 15;
HUrb = zeros(N,1);
for zx = 1:N
disp(zx);
switch zx
% Loading images sequentially.
% The file name 'impnoise' means image plus noise.
    case 1, load impnoise0.mat
        % Initialization of temporary variables
        spixels = size(pixels);
        TempX = zeros(spixels(2),1);
        VarX = zeros(spixels(2),1);
        Hb = zeros(spixels(1),1);
        log_m = log10(1:spixels(2));
    case 2, load impnoise25.mat
    case 3, load impnoise50.mat
    case 4, load impnoise75.mat
    case 5, load impnoise100.mat
```

```
    case 6, load impnoise125.mat
    case 7, load impnoise150.mat
    case 8, load impnoise175.mat
    case 9, load impnoise200.mat
    case 10, load impnoise225.mat
    case 11, load impnoise250.mat
    case 12, load impnoise275.mat
    case 13, load impnoise300.mat
    case 14, load impnoise325.mat
    case 15, load impnoise350.mat
end
for b = 1:spixels(1)
% Calculate the Hurst parameter for each individual image band.
    band = squeeze(pixels(b,:));
    for m = 1:spixels(2)
        for k = 1: floor(spixels/m)
            TempX(k) = sum( band(k*m - m + 1 : k*m) );
        end
        VarX(m) = var(TempX/m);
    end
    log_var_X = log10(VarX);
    p = polyfit(log_m', log_var_X, 1);
    Hb(b) = 1 - p(1)/2;
end
HUrb(zx) = mean(Hb);
end
save HUrb.mat HUrb
```

REFERENCE LIST

[1] A. Plaza; P. Martinez; R. Perez; J. Plaza. A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data. *Geoscience and Remote Sensing, IEEE Transactions on*, 42(3):650–663, 2004.

[2] A. Arzuaga-Cruz; LO. Jimenez-Rodriguez; M. Velez-Reyes; D. Kaeli; E. Rodriguez-Diaz; HT. Velazquez-Santana; A. Castrodad-Carrau; LE. Santos-Campis; C. Santiago. A matlab toolbox for hyperspectral image analysis. *IEEE International Geoscience and Remote Sensing Symposium*, 2004.

[3] S. Morillo-Contreras; M. Velez-Reyes; S. Hunt. Effects of resolution enhancement and atmospheric correction pre-processing in classification of hyperspectral imagery. *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery X, Proceedings of SPIE*, 5425, 2004.

[4] S. Hunt; H. Sierra. Spectral oversampling in hyperspectral imagery. *Proceedings of SPIE In Algorithms for Multispectral, Hyperspectral and Ultraspectral Imagery IX*, 5093, 2003.

[5] N. Keshava; P. Boettcher. Relationships between physical phenomena, distance metrics, and best-bands selection in hyperspectral processing. *Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VII*, 4381, 2001.

[6] M. Velez-Reyes; D. Linares; LO. Jimenez-Rodriguez. Two-stage band selection algorithm for hyperspectral imagery. *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery VIII, Proceedings of SPIE*, 4725, 2002.

[7] M. Velez-Reyes; D. Linares. Comparison of principal-component-based band selection methods for hyperspectral imagery. *Image and Signal Processing for*

Remote Sensing VII, Sebastiano Bruno Serpico, Editor, Proceedings of SPIE, 4541, 2002.

[8] M. Velez-Reyes; L.O. Jimenez; D. Linares; H. Velazquez. Comparison of matrix factorization algorithms for band selection in hyperspectral imagery. *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery VI, Proceedings of SPIE*, 4049, April 2000.

[9] N. Keshava; J. Mustard. Spectral unmixing. *IEEE Signal Processing Magazine*, January 2002.

[10] R.A. Peters; R.N. Strickland. Image complexity metrics for automatic target recognizers. *Automatic Target Recognizer System and Technology Conference, Naval Surface Warfare Center, Silver Spring, MD*, Oct 1990.

[11] K. Namuduri; K. Bouyoucef; L. Kaplan. Image metrics for clutter characterization. *IEEE Proceedings. International Conference on Image Processing*, 2000.

[12] O. Fadiran; L. Kaplan. Clutter complexity analysis of hyper-spectral bands. *Proceedings of the IEEE System Theory Symposium*, pages 531–535, 2004.

[13] M. Chacon; A. Corral. Image complexity measure: a human criterion free approach. *North American Fuzzy Information Processing Society, 2005 Annual Meeting*, 2005.

[14] S. Morillo-Contreras; M. Velez-Reyes; S.D. Hunt. A comparison of noise reduction methods for image enhancement in classification of hyperspectral imagery. *Proceedings of SPIE*, 5806:361, 2005.

[15] Y. Masalmah; M. Velez-Reyes; S. Rosario-Torres. An algorithm for unsupervised unmixing of hyperspectral imagery using positive matrix factorization. *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI, Proceedings of SPIE*, 5806:703–710, 2005.

[16] Y. Masalmah; M. Velez-Reyes. A comparison of algorithms to compute the positive matrix factorization and their application to unsupervised unmixing.

Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XII, Proceedings of SPIE, 6233, 2006.

[17] H.S. Bowen. Absolute radiometric calibration of the ikonos sensor using radiometrically characterized stellar sources. *Proceedings of the Pecora 15*, pages 10–14, 2002.

[18] J. Pearlman; S. Carman; C. Segal; P. Jarecke. Overview of the hyperion imaging spectrometer for the nasa eo-1 mission. *Geoscience and Remote Sensing Symposium, 2001. IGARSS'01. IEEE 2001 International*, 7, 2001.

[19] S.D. Brown. Dirsig user documentation: Release 3.4. 2001.

[20] L. Alparone; S. Baronti; A. Garzelli; F. Nencini. A global quality measurement of pan-sharpened multispectral imagery. *Geoscience and Remote Sensing Letters, IEEE*, 1(4):313–317, 2004.

[21] G.M. Foody. Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80(1):185–201, 2002.

[22] R.O. Green; M.L. Eastwood; C.M. Sarture; T.G. Chrien. Imaging spectroscopy and the airborne visible/infrared imaging spectrometer(aviris). *Remote Sensing of Environment*, 65(3):227–248, 1998.

[23] R.O. Duda; P.E. Hart; D.G. Stork. *Pattern Classification*. Wiley-Interscience, 2000.