

HYPERSENSITIVELY SENSITIVE CHANGE DETECTION USING TEMPORAL PRINCIPAL COMPONENT ANALYSIS

By

Vanessa Ortiz Rivera

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

MASTER OF SCIENCE

in

ELECTRICAL ENGINEERING

UNIVERSITY OF PUERTO RICO

MAYAGUEZ CAMPUS

2005

Approved by:

Shawn D. Hunt, Ph.D.
Member, Graduate Committee

Date

Badrinath Roysam, Ph.D.
Member, Graduate Committee

Date

Miguel Vélez-Reyes, Ph.D.
President, Graduate Committee

Date

Roy Armstrong, Ph.D.
Representative of Graduate Studies

Date

Isidoro Couvertier, Ph.D.
Chairperson of the Department

Date

Abstract

This work deals with the detection of changes using hyperspectral images. Change detection is the process of automatically identifying and analyzing regions that have undergone spatial or spectral changes from multi temporal images. Detecting and representing change provides valuable information of the possible transformations a given scene has suffered over time. Change detection in sequences of hyperspectral images is complicated by the fact that change can occur in the temporal and/or spectral domains. This work studies the use of Temporal Principal Component Analysis (TPCA) for change detection in multi/hyperspectral images. Additional methods were implemented in order to compare its results with TPCA. These were: Image Differencing and Conventional Principal Component Analysis. Hyperspectral imagery from different sensors showing different scenarios was used to test and validate the methods presented in this study. The algorithms were implemented using Matlab, and its performances are presented in terms of false alarms and missed changes. Overall results showed that the performance of TPCA was the best, obtaining the smallest error percentage.

Resumen

Este trabajo atiende la detección de cambios utilizando imágenes hiperespectrales. La detección de cambios es el proceso automático de identificar y analizar las áreas que han sufrido cambios en imágenes multitemporales, ya sea en el dominio espacial o espectral. El detectar y representar los cambios provee información valiosa sobre las posibles transformaciones que pudo haber sufrido una escena, a través del tiempo. La detección de cambios en secuencia de imágenes hiperespectrales se torna complicada por el hecho de que el cambio puede ocurrir tanto en el dominio espectral como en el temporal. Este trabajo estudia el uso del Análisis Temporal de Componentes Principales para detectar cambios en imágenes hiperespectrales. Dos métodos adicionales fueron implementados para comparar sus resultados con el método presentado en este trabajo. Estos fueron: Diferencia de Imágenes y el Análisis Convencional de Componentes Principales. Imágenes tomadas de distintos sensores hiperespectrales fueron utilizadas para probar y validar la capacidad de los algoritmos estudiados en detectar los cambios. Los algoritmos presentados fueron implementados en el programa MATLAB, y su desempeño se muestra en términos de la cantidad de falsas alarmas y pixeles de cambio no detectados. Los resultados generales demostraron que el método de Análisis Temporal de Componentes Principales fue el mejor obteniendo el menor porcentaje de error.

TABLE OF CONTENTS

1. INTRODUCTION.....	15
1.1 MOTIVATION.....	15
1.2 PROBLEM STATEMENT	16
1.3 OBJECTIVES	16
1.4 LITERATURE REVIEW	17
1.5 CONTRIBUTION	19
1.6 THESIS OVERVIEW	19
2. BACKGROUND	21
2.1 SPECTROSCOPY	21
2.2 IMAGE SPECTROSCOPY.....	23
2.3 REMOTE SENSING.....	23
2.4 SPECTRAL SENSING.....	25
2.5 IMAGE CLASSIFICATION	26
2.5.1 <i>Supervised Classification</i>	27
2.5.2 <i>Unsupervised Classification</i>	28
2.6 DIABETIC RETINOPATHY	29
2.7 PRINCIPAL COMPONENTS ANALYSIS	32
2.8 CHANGE DETECTION.....	33
2.9 CHANGE DETECTION TECHNIQUES.....	36
2.9.1 <i>Image Differencing</i>	36
2.9.2 <i>Change Vector Analysis (CVA)</i>	37
2.9.3 <i>Conventional Principal Component Analysis</i>	38
2.9.4 <i>Post-Classification Comparison</i>	39
2.9.5 <i>Multivariate Alteration Detection (MAD)</i>	40
2.10 SUMMARY	43
3. TEMPORAL PRINCIPAL COMPONENT ANALYSIS (TPCA).....	44
3.1 INTRODUCTION.....	44
3.2 METHODOLOGY FOR TPCA	45
3.3 SUMMARY	48
4. METHODOLOGY.....	49

4.1	METHODOLOGY USED FOR CHANGE DETECTION	49
4.1.1	<i>Image Pre-Processing</i>	50
4.1.1.1	<i>Image Registration</i>	50
4.1.1.2	<i>Atmospheric Correction</i>	51
4.1.2	<i>Change detection</i>	52
4.1.2.1	<i>Image Differencing</i>	52
4.1.2.2	<i>Conventional Principal Component Analysis</i>	54
4.1.2.3	<i>Post-Classification Comparison (HYPERION Images)</i>	54
4.1.3	<i>Decision Rule</i>	55
4.1.3.1	<i>Global Threshold Selection</i>	55
4.1.3.2	<i>Per-band Change detection</i>	56
4.2	HYPERSPECTRAL CAMERA	57
4.3	HYPERION SENSOR	58
4.4	TOPCON TRG 50IA FUNDUS CAMERA	59
4.5	HYPERSPECTRAL DATA	60
4.5.1	<i>Test Images</i>	60
4.5.2	<i>HYPERION Data – Benthic Habitat of Enrique Reef at La Parguera</i>	61
4.5.3	<i>Fundus Camera Data- Images from a Human Retina</i>	63
4.6	MATLAB HYPERSPECTRAL IMAGE ANALYSIS TOOLBOX	63
4.7	SUMMARY	66
5.	EXPERIMENTAL RESULTS.....	67
5.1	TEST DATA – COINS OVER CONCRETE	67
5.1.1	<i>Image Differencing with Global Threshold Selection</i>	67
5.1.2	<i>Image Differencing with Per-band Threshold Selection</i>	68
5.1.3	<i>CPCA with Global Threshold Selection</i>	69
5.1.4	<i>CPCA with Per-band Threshold Selection</i>	72
5.1.5	<i>TPCA with Global Threshold Selection</i>	72
5.1.5	<i>TPCA with Per-band Threshold Selection</i>	73
5.2	TEST DATA – PLASTIC STRAW OVER CONCRETE SIDEWALK	74
5.2.1	<i>Image Differencing with Global Threshold Selection</i>	74
5.2.2	<i>Image Differencing with Per-band Threshold Selection</i>	75
5.2.3	<i>CPCA with Global Threshold Selection</i>	75
5.2.4	<i>CPCA with Per-band Threshold Selection</i>	78
5.2.5	<i>TPCA with Global Threshold Selection</i>	78
5.2.6	<i>TPCA with Per-band Threshold Selection</i>	79
5.3	ALGORITHM PERFORMANCE	79
5.4	APPLICATION OF CHANGE DETECTION TO BENTHIC HABITATS AT ENRIQUE REEF IN LA PARGUERA	83
5.4.1	<i>Image Differencing with Global Threshold Selection</i>	83
5.4.2	<i>Image Differencing with Per-band Threshold Selection</i>	84

5.4.3	<i>CPCA with Global Threshold Selection</i>	84
5.4.4	<i>CPCA with Per-band Threshold Selection</i>	87
5.4.5	<i>TPCA with Global Threshold Selection</i>	87
5.4.6	<i>TPCA with Per-band Threshold Selection</i>	88
5.4.7	<i>Post Classification Comparison</i>	88
5.5	APPLICATION OF CHANGE DETECTION TO THE FUNDUS CAMERA IMAGES	90
5.5.1	<i>Image Difference with Global Threshold Selection</i>	90
5.5.2	<i>Image Difference with Per-band Threshold Selection</i>	91
5.5.3	<i>CPCA with Global Threshold Selection</i>	91
5.5.4	<i>CPCA with Per-band Threshold Selection</i>	93
5.5.5	<i>TPCA with Global Threshold Selection</i>	94
5.5.6	<i>TPCA with Per-band Threshold Selection</i>	94
5.6	SUMMARY	95
6.	CONCLUSIONS AND FUTURE WORK	97
6.1	CONCLUSIONS	97
6.2	FUTURE WORK.....	98
	REFERENCES	100
	APPENDICES	103
A.	PRINCIPAL COMPONENT ANALYSIS	103
A.1	<i>Derivation</i>	103
A.2	<i>Mathematical and Statistical Properties</i>	106
B.	SINGULAR VALUE DECOMPOSITION NOISE FILTERING.....	111
B.1	<i>Signal Model</i>	111
B.2	<i>Computing the filtered image</i>	112
C.	MATLAB CODES:	113

LIST OF FIGURES

	Page
Figure 1 - Spectral Range.....	22
Figure 2: Remote Sensing Scheme (from CCRS website).	25
Figure 3 : Hyperspectral Images (from TNTmips website).....	26
Figure 4: Image Classification (from CCRS website).	26
Figure 5: Supervised Classification (from CCRS website).	27
Figure 6: Unsupervised Classification (from CCRS website).	28
Figure 7: Effects to vision of diabetic retinopathy (from St. Luke’s Eye website).	29
Figure 8: Illustration by Mark Erickson (from http://www.stlukeseye.com).....	30
Figure 9: Effects to vision of macular degeneration (from St. Luke’s Eye website).....	31
Figure 10: Wet and dry AMD illustrations (from St. Luke’s Eye website).....	31
Figure 11: Temporal Principal Component Analysis Procedure.	33
Figure 12: General steps for the estimation of the change mask.	35
Figure 13: Image Differencing Method.	36
Figure 14: Change Vector Analysis.	38
Figure 15: Conventional Principal Component Analysis.....	39
Figure 16: Conventional Principal Component Analysis.....	39
Figure 17: Post-Classification Comparison.	40
Figure 18: Multivariate Alteration Detection.....	42
Figure 19: Principal Component Analysis-General Concept.....	46
Figure 20: Temporal Principal Component Analysis Procedure.	46
Figure 21: Change detection diagram for the Temporal Principal Component Analysis method.....	47
Figure 22: Hyperspectral Cubes conversion to vector form.	52
Figure 23: Training Areas	55
Figure 24: Procedure for the application of a global threshold.....	56
Figure 25: Procedure for the application of per-band change detection approach.	57
Figure 26: SOC-700 Hyperspectral Camera (from Surface Optics web page).....	58
Figure 27: Hyperion Sensor (NASA EO1 Webpage)	58
Figure 28: Topcon TRC 50IA fundus camera (from www.colcon.com).....	59

Figure 29: Test hyperspectral image of sidewalk.	60
Figure 30: Test hyperspectral image with coins.	60
Figure 31: Test hyperspectral image of sidewalk.	61
Figure 32: Test hyperspectral sidewalk image with straw on it.	61
Figure 33: Area of La Parguera, taken by Hyperion (This map is from Lonely Planet web site).	62
Figure 34: Image at time one, taken on August 28, 2002	62
Figure 35: Image at time two, taken on July 15,2003.	62
Figure 36: Images taken with the Topcon TRC 50IA fundus camera.	63
Figure 37: Hyperspectral Image Analysis Toolbox.	64
Figure 38: Image Difference Magnitude of Change	68
Figure 39: Image Difference Change Mask	68
Figure 40: Image Difference Change Mask from individual thresholds.	68
Figure 41: CPCA Principal Components.	70
Figure 42: CPCA Magnitude of Change.	71
Figure 43: CPCA Change Mask, components 2-12.	72
Figure 44: CPCA Change Mask with individual threshold selection.	72
Figure 45: TPCA Magnitude of Change with global threshold.	73
Figure 46: TPCA Change Mask with global threshold selection.	73
Figure 47: TPCA Change Mask with local threshold selection.	74
Figure 48: Image Difference Change Mask from global threshold selection.	74
Figure 49: Image Difference Change Mask from local thresholds selection.	75
Figure 50: CPCA Principal Components.	76
Figure 51: CPCA Change Mask, components 3-8.	77
Figure 52: CPCA Change Mask with individual threshold selection.	78
Figure 53: TPCA Change Mask with global threshold selection.	79
Figure 54: TPCA Change Mask with local threshold selection.	79
Figure 55: (a) Reference Image, (b) Image difference, (c) CPCA and (d) TPCA change masks.	81
Figure 56: (a) Reference Image, (b) Image difference, (c) CPCA and (d) TPCA change masks.	81

Figure 57: (a) Reference Image, (b) Image difference, (c) CPCA and (d) TPCA change masks.....	82
Figure 58: (a) Reference Image, (b) Image difference, (c) CPCA and (d) TPCA change masks.....	82
Figure 59: Image Differencing Magnitude of Change.....	83
Figure 60: Image Differencing Change Mask with global threshold selection.	83
Figure 61: Image Differencing Change Mask with individual threshold selection.	84
Figure 62: CPCA Principal Components of Enrique Reef.....	85
Figure 63: CPCA Magnitude Image PCs 2-8.....	86
Figure 64: CPCA Change Mask obtained with global threshold selection.....	86
Figure 65: CPCA Change Mask obtained with local threshold selection.....	87
Figure 66: TPCA Magnitude Image.....	87
Figure 67: TPCA Change Mask with global threshold selection.....	88
Figure 68: TPCA Change Mask with individual threshold selection.	88
Figure 69: Training samples areas.	89
Figure 70: Classification Maps of the Enrique Reef at years (a) 2002 and (b) 2003.....	89
Figure 71: Post Classification Comparison Change Mask.....	90
Figure 72: Change Mask obtained by Image differencing with global threshold selection.	90
Figure 73: Change Mask obtained by Image differencing with per-band threshold selection.	91
Figure 74: Principal Components and percentages of variability.	92
Figure 75: Change Mask obtained by CPCA with global threshold selection.....	93
Figure 76: Change Mask obtained by CPCA with per-band threshold selection.	93
Figure 77: Change Mask obtained by TPCA with global threshold selection.....	94
Figure 78: Change Mask obtained by TPCA with per-band threshold selection.....	95

LIST OF TABLES

	Page
Table 1 –Regions of the Spectral Range	22
Table 2: Hyperion Sensor Specifications	59
Table 3: Percentages of Variability and variance for each principal component.	71
Table 4: Percentages of Variability and eigenvalues of each principal component.	77
Table 5: Methods Performances using global threshold selection.....	81
Table 6: Methods Performances using per-band threshold selection.....	81
Table 7: Methods Performances using global threshold selection.....	82
Table 8: Methods Performances using per-band threshold selection.....	82
Table 9: Percentages of Variability and eigenvalues of the principal components.	86
Table 10: Overall error for the images of coins over concrete sidewalk.	96
Table 11: Overall error for the images of a drinking straw over concrete sidewalk.....	96

To the Lord, Jesus Christ, without your grace is impossible. To my parents, Ramón and Lucy, to my sister Lumary, and to George, thanks for your love.

ACKNOWLEDGEMENTS

I would like to specially thank my advisor Dr. Miguel Vélez for all his interest and patience during all this year. I also want to thank my graduate committee members Dr. Shawn Hunt and Dr. Badrinath Roysam, for their support and collaboration in this work.

Special thanks go to my Dad, you are my role model, the struggles you've had in life have not stopped you and that is my inspiration.

This work was supported by CenSSIS, the Center for Subsurface Sensing and Imaging Systems, under the Engineering Research Centers Program of the National Science Foundation (Award Number EEC-9986821).

CHAPTER 1

1. Introduction

1.1 Motivation

Timely and accurate change detection of Earth's surface features is extremely important for understanding relationships and interactions between human and natural phenomena in order to promote better decision making (Lu et al. 2004). Detecting and delineating changes in images of the same scene taken at different times has interested researchers for many years owing to a large number of applications in diverse disciplines (Radke et al. 2003). Examples of important applications include remote sensing, medical diagnosis and treatment, civil infrastructure and others. Because of the advantage of repetitive data acquisition, its synoptic view, and digital format suitable for computer processing, remotely sensed data have become the major source of information for different change detection applications during the past decades (Lu et al. 2004). Some aspects that could be monitored using remote sensing technologies for change detection are, for example, land-use and land-cover, forest or vegetation change, deforestation, wetland, landscape or vegetation change, and crop monitoring.

A variety of multispectral change detection techniques have been developed and its selection depends on the application. Examples of these are: Image Differencing, Post-Classification Comparison, Change Vector Analysis, Principal Component Analysis, among others. Most of these methods perform change detection using single bands. This is they perform a band per band analysis of the images, making necessary a suitable selection of bands.

The selection of bands to perform change detection is one of the major challenges the analyzer has to face. Since each band provides a certain amount of change information it becomes difficult to decide which bands are going to be chosen. The motivation of this work responds to the lack of methods that perform change detection by using multiple bands of hyperspectral images at the same time. By doing this it is intended to take advantage of all the possible change information available in the data.

1.2 Problem Statement

Change detection in a sequence of hyperspectral images is complicated by the fact that change can occur in the temporal and/or spectral domains. Since hyperspectral images are composed of a considerable number of bands, it is difficult to determine which features are going to be selected to perform the detection of changes. On the other hand, considering for example that hyperspectral images with N spectral bands yields an N -dimensional change space, change can occur independently in each spectral band. The majority of the existing change detection algorithms used to analyze spectral data are implemented to detect changes using specific bands for specific applications. In this work a change detection technique based on the Temporal Principal Component Analysis is presented. All the bands of each hyperspectral image are used to perform change detection. By using all the bands, we expect to maximize the amount of information that can be extracted by each image. Two threshold techniques (global and local threshold selection), based on simple hypothesis testing were applied in order to produce the binary change/no-change maps for each method.

1.3 Objectives

The objectives of this study are:

- Study and compare existing change detection methods used with multispectral imagery.

- Study change detection using hyperspectral images.
- Study the temporal principal component analysis method and modify it to use in change detection using hyperspectral images.
- Implement the most common change detection algorithms used with multispectral imagery and compare their performance with the technique studied. For this work, the algorithms to be implemented were the following:

- (1) Image Differencing
- (2) Conventional Principal Component Analysis
- (3) Post-Classification Comparison

1.4 Literature Review

Currently, there are many change detection methods that have been implemented and their use depends on the application, which is in this case remote sensing. This section presents the different change detection methods, encountered in the literature review.

- (1) Bruzzone et al. (2002), presented an automatic approach to unsupervised classification of changes in multitemporal remote sensing images. In this method, two bands are subtracted and then a threshold is selected in order to generate a binary change/ no change map. The two bands are the same but correspond to different points in time. The threshold selection is based on the formulation of the unsupervised change detection problem in terms of the Bayesian Approach. The statistical terms associated with the gray levels of changed and unchanged pixels in a difference image are estimated and used to make the decision if whether change has or has not occurred.
- (2) Wiemker et al. (1997) presented a change detection method using principal component analysis. Here a single band from each image is represented in vector form. These two vectors are used to form a bi-temporal feature space, in which

the first vector corresponds to an image at time 1 and the second vector corresponds to the same image at time 2. Principal Component Analysis is performed over this bi-temporal feature space. All the pixels that lie around the first component are considered to be no-change pixels, and the ones that lie around the second component are considered as pixels that have suffered change. This means that the information of the magnitude of change is contained in the second principal component. In order to generate a binary change/no change map, a threshold was obtained by using Bayes decision rule. This permits to observe the change detection problem as an unsupervised classification, where each pixel is assigned to one out of two classes: 'change' or 'no-change'.

- (3) Nielsen et al. (1998) introduced the Multivariate Alteration Detection Transformation, which is based on the canonical correlation analysis. This method basically finds the difference between linear combinations of the spectral bands from the two acquisitions. The differences produced are orthogonal and constructed so they show maximum variance. The threshold for the generation of a change/no change map was selected by using the Bayesian approach.
- (4) Borrego et al. (2001) presented the change vector analysis detection method. The change vector analysis generates two variables, the magnitude of variation and the angle of the change vector. The change vector is obtained by subtracting the images represented in vector form. The magnitude of variation between the two vectors is obtained calculating the Euclidean distance between the differences in positions of the same pixel from different times. The angle of the vector, which indicates the type of change that has occurred, varies according to the number of components used. Each vector is a function of the positive or negative changes through the spectral bands.
- (5) Huang et al. (2000) presented the Post Classification Change Detection method to detect changes over an area of natural grass. In this method both images were classified using supervised classification (Maximum Likelihood Classification).

The classification maps of both images were compared in order to determine if change occurred. After comparing both classification results a change map was constructed.

This study presents how Principal Component Analysis can be used to analyze temporal data for change detection purposes. The method uses the properties of PCA to transform the data in a way change can be detected. A group of multispectral and hyperspectral images were analyzed and the performance of the method was calculated in terms of percentage of false alarms and missed changes by using test images.

1.5 Contribution

From the literature review realized as part of this work, it was encountered that the change detection methods currently used with multispectral imagery receive as input a single band or just a subset, from each image set, to perform the analysis. This demands the expertise of the analyst to choose suitable bands for a given application. In this study change detection was performed, by using all the bands of each image set. By doing this, it was pretended to exploit all the possible change information contained in both spatial and spectral domains. According to the experiments performed in this study, the proposed modified Temporal Principal Component Analysis method improved change detection in comparison with the Image Differencing and Conventional Principal Component Analysis techniques. Overall results also demonstrated that local threshold selection method decreased the overall error.

1.6 Thesis Overview

Chapter Two presents all the general concepts about remote sensing, hyperspectral images and change detection. A description of current change detection algorithms used to analyze multispectral imagery is also provided.

Chapter Three shows how the Principal Component Analysis was used to analyze temporal data.

Chapter Four describes the methodology followed in order to achieve the objectives of this work. It also gives information about the hyperspectral images that were used in this research.

Chapter Five presents the results obtained by each change detection technique implemented in this work. The metrics used to evaluate each algorithm performance is in terms of percentage of false alarms and missed changes.

Chapter Six presents the conclusions and suggests possible future work.

The appendices include the Matlab codes used to perform change detection over all the multi/hyperspectral images and provide a theoretical and mathematical background of Principal Component Analysis and Singular Value Decomposition.

CHAPTER 2

2. Background

This chapter presents the general concepts in hyperspectral imaging, remote sensing and change detection necessary for the understanding of the fundamental aspects of this research work.

2.1 Spectroscopy

Spectroscopy uses the absorption, emission, or scattering of electromagnetic radiation by matter to qualitatively or quantitatively study the matter or to study physical processes. Three basic interactions between light and matter are identified:

- Absorption: It is the transition from a lower level to a higher level with transfer of energy from the radiation field to an absorber, atom, molecule, or solid.
- Emission: It is the transition from a higher level to a lower level with transfer of energy from the emitter to the radiation field. If no radiation is emitted, the transition from higher to lower energy levels is called non radioactive decay.
- Scattering: It refers to the redirection of light due to its interaction with matter. Scattering might or might not occur with a transfer of energy, i.e., the scattered radiation might or might not have a slightly different wavelength compared to the light incident on the sample (Tissue 2000).

The instrument used to record the spectral properties of an object is called a spectrometer. There are four general parameters that describe the capability of a spectrometer:

1. Spectral Range: It is the wavelength range measured by the instrument.
The light spectra is divided in five regions (see Figure 1 and Table 1).

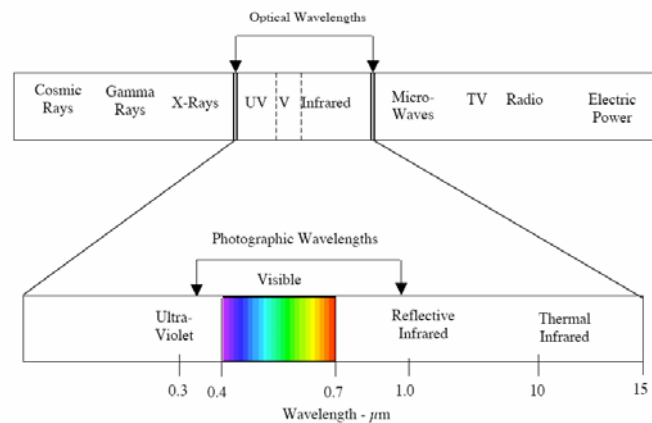


Figure 1 - Spectral Range

Table 1 –Regions of the Spectral Range

Region	Spectral Range
Ultraviolet	0.001 to 0.4 μm
Visible	0.4 to 0.7 μm
Near-Infrared	0.7 to 3.0 μm
Mid-Infrared	3.0 to 30 μm
Far-Infrared	30 μm to 1mm

2. Spectral resolution: this is the width of an individual spectral channel in the spectrometer.

3. Spectral sampling: it is the distance in wavelength between the spectral band pass profiles for each channel in the spectrometer as a function of wavelength.

4. Signal-to-noise ratio: is a measure of the strength of the spectral signal in comparison to that of the background noise, usually calculated as the average signal divided by the root-mean-square noise.

2.2 Image Spectroscopy

In Hyperspectral Imaging or Imaging Spectroscopy, spatial information along with high-spectral resolution of the scene under study is collected. As the object of interest is embedded in a translucent media (i.e. atmosphere, coastal waters or skin), the measured spectral signature is a distorted version of the object signature mixed with clutter. The high spectral resolution may allow discrimination between media and object contributions enabling the retrieval of information about the object of interest.

2.3 Remote Sensing

By definition, “Remote Sensing is the science and art of obtaining information about an object, area, or phenomenon through the analysis of data acquired by a device that is not in contact with the object, area, or phenomenon under investigation”. One way to achieve the retrieval of information in remote sensing applications is by the use of electromagnetic energy sensors that are currently operated from airborne and spaceborne platforms to assist in inventorying, mapping and monitoring earth resources. These sensors acquire data on the way various earth surface features emit and reflect energy, and these data are analyzed to provide information about the resources investigated (Lillesand et al. 1999).

The process of remote sensing involves the interaction between the incident radiation and the objects that are observed. The basic elements involved in the retrieval of the scene spectral properties, are the following, see Figure 2:

- A) Energy Source or Illumination - the first requirement for remote sensing is to have an energy source, which illuminates or provides electromagnetic energy to the target of interest.
- B) Radiation and the Atmosphere - as the energy travels from its source to the target, it will come in contact with and interact with the atmosphere it passes through. This interaction may take place a second time as the energy travels from the target to the sensor.
- C) Interaction with the Target - once the energy makes its way to the target through the atmosphere; it interacts with the target depending on the properties of both the target and the radiation.
- D) Recording of Energy by the Sensor - after the energy has been scattered by, or emitted from the target, it is required to the sensor (remote - not in contact with the target) to collect and record the electromagnetic radiation.
- E) Interpretation and Analysis - the processed image is interpreted, visually and/or digitally or electronically, to extract information about the target which was illuminated.
- G) Application - the final element of the remote sensing process is achieved when the information extracted from the imagery about the target is applied in order to better understand it, reveal some new information, or assist in solving a particular problem (CCRS 2004).

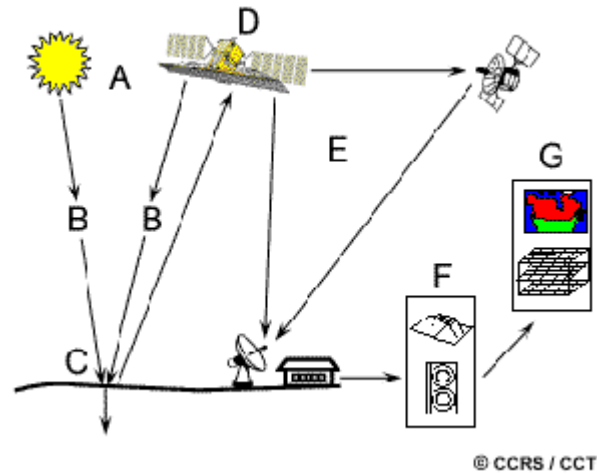


Figure 2: Remote Sensing Scheme (from CCRS website).

The final result of the process described above is an image cube with the third dimension specified by spectral wavelengths. Each image pixels is indeed a column vector, of which each component represents a particular spectral band (CCRS 2004).

2.4 Spectral Sensing

Multispectral sensors are instruments that measure multiple, wide, separated wavelength bands while hyperspectral sensors acquire images in many, very narrow, contiguous spectral bands throughout the visible, near-IR, mid-IR, and thermal-IR portions of the spectrum (Shippert 2002). These sensors take advantage of hundreds of contiguous spectral channels to uncover spectral features that usually cannot be resolved by multispectral sensors (Chang 2003). Although most hyperspectral sensors measure hundreds of wavelength bands, and most multispectral sensors measure only a few, it is not the number of measured wavelengths that defines a sensor as multispectral or hyperspectral, but the narrowness and contiguous nature of the measurements.

Hyperspectral images contain a wealth of data, but interpreting them requires an understanding of exactly what properties of ground materials are intended to be measured, and how they relate to the measurements actually made by the hyperspectral sensor.

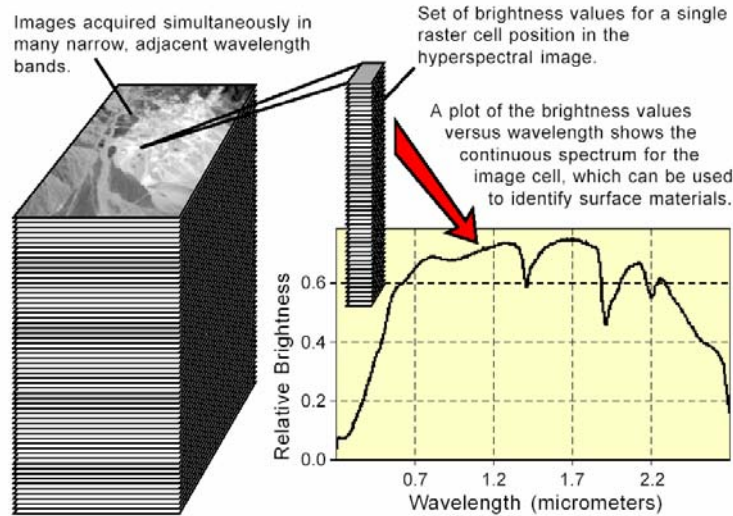


Figure 3 : Hyperspectral Images (from TNTmips website)

2.5 Image Classification

Digital image classification uses the spectral information represented by the digital numbers in one or more spectral bands, and attempts to classify each individual pixel based on this spectral information. This type of classification is termed spectral pattern recognition. In either case, the objective is to assign all pixels in the image to particular classes or themes (e.g. water, coniferous forest, deciduous forest, corn, wheat, etc.). The resulting classified image is comprised of a mosaic of pixels, each of which belongs to a particular theme, and is essentially a thematic map of the original image (see Figure 4, CCRS 2004).

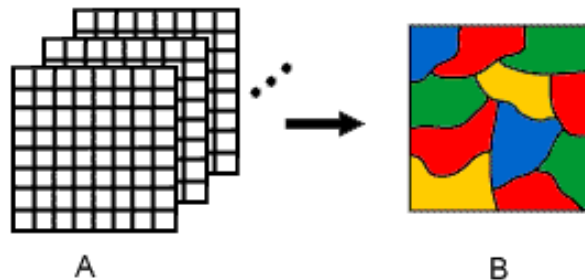


Figure 4: Image Classification (from CCRS website).

Common classification procedures can be broken down into two broad subdivisions based on the method used: supervised classification and unsupervised classification.

2.5.1 Supervised Classification

In a supervised classification, the analyst identifies in the imagery homogeneous representative samples of the different surface cover types of interest. These samples are referred to as training areas (see Figure 5). The selection of appropriate training areas is based on the analyst's familiarity with the geographical area and their knowledge of the actual surface cover types present in the image. Thus, the analyst is supervising the categorization of a set of specific classes. The numerical information in all spectral bands for the pixels comprising these areas is used to train the computer to recognize spectrally similar areas for each class. The computer uses a special program or algorithm, to determine the numerical "signatures" for each training class. Once the computer has determined the signatures for each class, each pixel in the image is compared to these signatures and labeled as the class it most "closely resembles" digitally. Thus, in a supervised classification there is first an identification of the information classes, which are then used to determine the spectral classes that represent those (CCRS 2004).

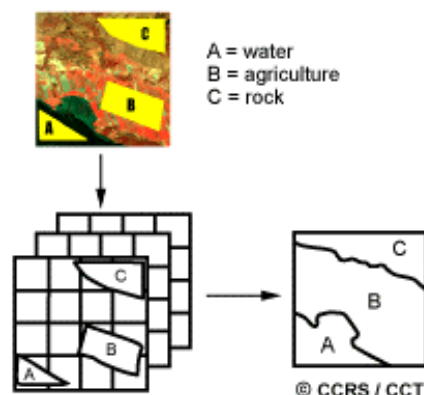


Figure 5: Supervised Classification (from CCRS website).

2.5.2 Unsupervised Classification

Unsupervised classification in essence reverses the supervised classification process. Spectral classes are grouped first, based solely on the numerical information in the data, and are then matched by the analyst to information classes (see Figure 6). Clustering algorithms are used to determine the statistical structures in the data. Usually, the analyst specifies how many groups or clusters are to be looked for in the data. In addition to specifying the desired number of classes, the analyst may also specify parameters related to the separation distance among the clusters and the variation within each cluster. The final result of this iterative clustering process may result in some clusters that the analyst will want to subsequently combine, or clusters that should be broken down further each of these requiring a further application of the clustering algorithm. Thus, unsupervised classification is not completely without human intervention. However, it does not start with a pre-determined set of classes as in a supervised classification (CCRS 2004).

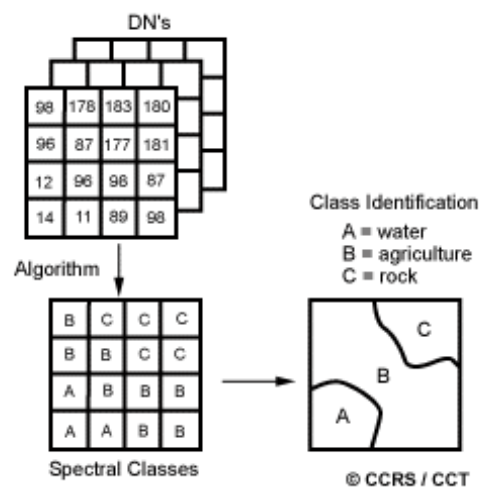


Figure 6: Unsupervised Classification (from CCRS website).

2.6 Diabetic Retinopathy

The most common causes of vision loss among the elderly are age-related macular degeneration, glaucoma, cataract and diabetic retinopathy. The effect of diabetes on the eye is called diabetic retinopathy. Over time, diabetes affects the circulatory system of the retina. The earliest phase of the disease is known as background diabetic retinopathy. In this phase, the arteries in the retina become weakened and leak, forming small, dot-like hemorrhages. These leaking vessels often lead to swelling or edema in the retina and decreased vision. The next stage is known as proliferative diabetic retinopathy. In this stage, circulation problems cause areas of the retina to become oxygen-deprived or ischemic. New, fragile, vessels develop as the circulatory system attempts to maintain adequate oxygen levels within the retina. This is called neovascularization. Unfortunately, these delicate vessels hemorrhage easily. Blood may leak into the retina and vitreous, causing spots or floaters, along with decreased vision (St. Luke's Cataract and Laser Institute).

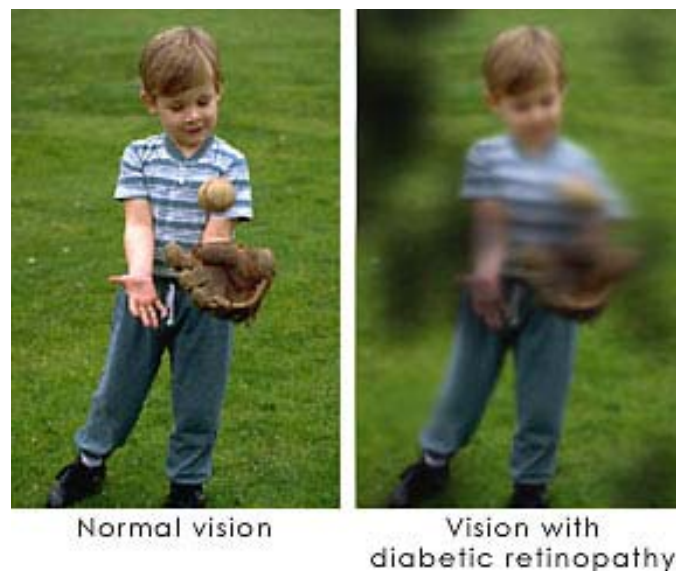


Figure 7: Effects to vision of diabetic retinopathy (from St. Luke's Eye website).

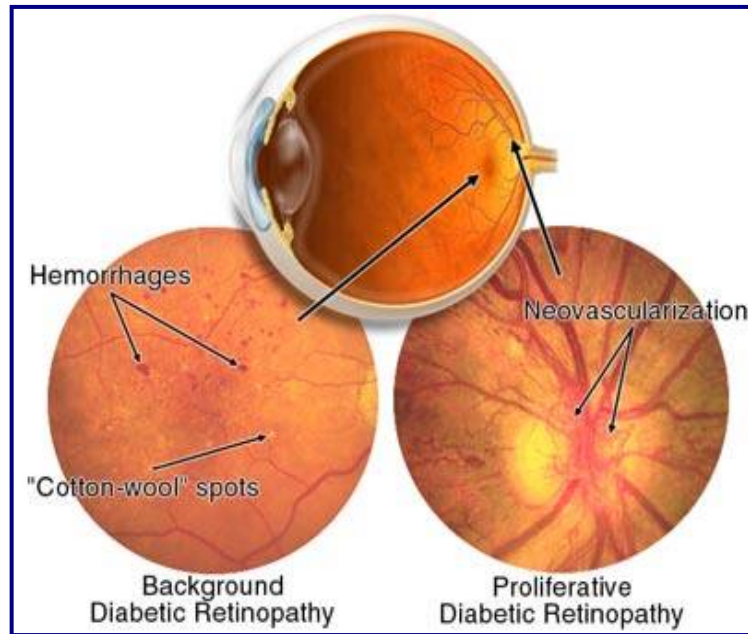


Figure 8: Illustration by Mark Erickson (from <http://www.stlukeseye.com>).

Another common cause of vision loss is age-related macular degeneration (AMD) (Quillen 1999). AMD is a degenerative condition of the macula (the central retina). It is the most common cause of vision loss in the United States in those 50 or older, and its prevalence increases with age. AMD is caused by hardening of the arteries that nourish the retina. This deprives the sensitive retinal tissue of oxygen and nutrients that it needs to function and thrive. As a result, the central vision deteriorates.

AMD can be divided into two categories: nonexudative (or dry and the most common) and exudative (wet) AMD (see Figure 10). Wet AMD occurs when abnormal blood vessels behind the retina start to grow under the macula. These new blood vessels tend to be very fragile and often leak blood and fluid. The blood and fluid raise the macula from its normal place at the back of the eye. Damage to the macula occurs rapidly. With wet AMD, loss of central vision can occur quickly. Wet AMD is considered to be advanced AMD and is more severe than the dry form. Dry AMD occurs when the light-sensitive cells in the macula slowly break down, gradually blurring central vision in the affected eye. As dry AMD gets worse, you may see a blurred spot in the center of your vision. Over time, as less of the macula

functions, central vision in the affected eye can be lost gradually. The most common symptom of dry AMD is slightly blurred vision. You may have difficulty recognizing faces. You may need more light for reading and other tasks. Dry AMD generally affects both eyes, but vision can be lost in one eye while the other eye seems unaffected (Medline Plus website). Within dry AMD there are drusen and geographic atrophy (GA). The latter is responsible for nearly all cases of severe loss of vision associated with dry AMD. Geographic atrophy is characterized by round or oval patches of atrophy of the retina, retinal pigment epithelium and underlying choroid. Over time, the patches may increase in size and number or may coalesce to form larger areas of atrophy.



Figure 9: Effects to vision of macular degeneration (from St. Luke's Eye website).

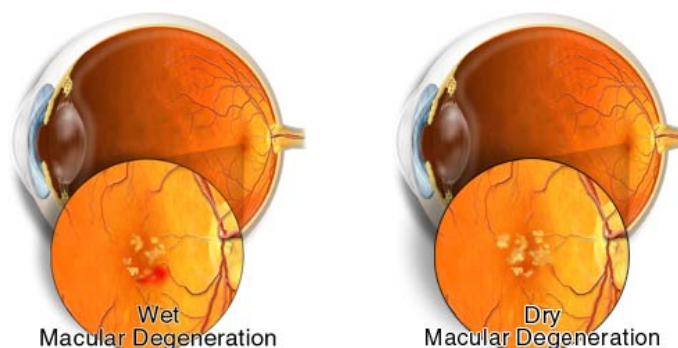


Figure 10: Wet and dry AMD illustrations (from St. Luke's Eye website).

2.7 Principal Components Analysis

Principal Component Analysis (PCA) is a linear transformation that chooses a new coordinate system for the data set such that the greatest variance by any projection of the data set comes to lie on the first axis (then called the first principal component), the second greatest variance on the second axis, and so forth.

Suppose that \mathbf{x} is a vector of p random variables, and that the variances of the p random variables and the structure of the covariances or correlations between the p variables are of interest. Unless p is small, or the structure is very simple, it is not always very helpful to simply look at the p variances and all of the $(1/2)p(p-1)$ correlations or covariances. An alternative approach is to look for a few ($\ll p$) derived variables that preserve most of the variability information.

The first step is to look for a linear function $z_1 = \boldsymbol{\alpha}_1' \mathbf{x}$ of the elements of \mathbf{x} having maximum variance, where $\boldsymbol{\alpha}_1$ is a vector of p constants $\alpha_{11}, \alpha_{12}, \dots, \alpha_{1p}$, and $'$ denotes transpose, so that,

$$z_1 = \boldsymbol{\alpha}_1' \mathbf{x} = \alpha_{11}x_1 + \alpha_{12}x_2 + \alpha_{13}x_3 + \dots + \alpha_{1p}x_p = \sum_{j=1}^p \alpha_{1j}x_j$$

$$\boldsymbol{\alpha}_1 = \arg \{ \max(\text{var}(z_1)) \}$$

The next step is to look for a linear function $\boldsymbol{\alpha}_2' \mathbf{x}$ uncorrelated with $\boldsymbol{\alpha}_1' \mathbf{x}$ having maximum variance, and so on, so that the k th stage a linear function of $\boldsymbol{\alpha}_k' \mathbf{x}$ is found that has maximum variance subject to being uncorrelated with z_1, z_2, \dots, z_{k-1} .

The k th derived variable, $\boldsymbol{\alpha}_k' \mathbf{x}$ is the k th principal component. Up to p principal components could be found, but in most applications the variability in \mathbf{x} is accounted for by few, $m \ll p$ principal components (Jolliffe 2002).

Let $\boldsymbol{\Sigma}$ be the covariance of \mathbf{x} . For $k=1, 2, \dots, p$ the k^{th} PC is given by $z_k = \boldsymbol{\alpha}_k' \mathbf{x}$ where $\boldsymbol{\alpha}_k$ is an eigenvector of $\boldsymbol{\Sigma}$ corresponding to the largest eigenvalue λ_k and $\boldsymbol{\alpha}_k$ is chosen to

have unit length $\mathbf{a}_k' * \mathbf{a}_k = 1$. Notice that $\text{var}(z_k) = \lambda_k$. A quantity of the variability **explained** by the first m principal components is the percentage of variability,

$$\%var = \frac{\sum_{i=1}^m \lambda_i}{\sum_{i=1}^p \lambda_i} \times 100$$

In dimensionality reduction m is chosen so that $\%var \geq 90\%$.

2.8 Change Detection

Change detection is the process of automatically identifying and analyzing regions that have undergone spatial or spectral changes from multi temporal images. Detecting and representing change provides valuable information of the possible transformations a given scene has suffered over time. A typical change detection algorithm takes the images Im_1 and Im_2 as input and generates a binary image \mathbf{B} , called a change mask that identifies changed regions (see Figure 11).

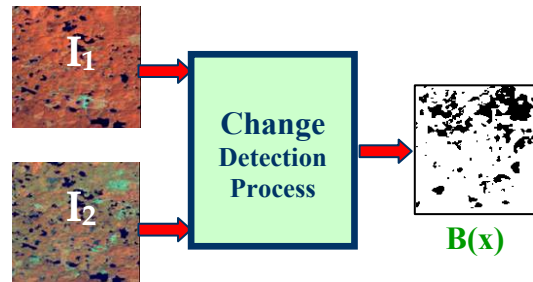


Figure 11: Temporal Principal Component Analysis Procedure.

Where, $B(x) = \begin{cases} 1 & \text{if pixel } \mathbf{x} \text{ corresponds to a significant change from } Im_1(x) \text{ to } Im_2(x) \\ 0 & \text{Otherwise} \end{cases}$

The decision at pixel x generally involves evaluating a cost function and selecting a suitable decision threshold (Radke et al. 2003).

The estimation of the change mask involves a series of steps, presented in Figure 12. These steps involve: image pre-processing, feature extraction, dimension reduction, image comparison, decision rule application, and image post-processing. Nevertheless, these steps are not necessarily performed by every change detection technique; it depends on the method used. Below, a general description of each change detection step is provided:

- (1) Image pre-processing- this step makes the two images comparable in both the spatial and spectral domain. Concerning the spatial domain, the two images should be co-registered so that pixels with the same coordinates in the image may be associated with the same area on the ground. With regard to the spectral domain, changes in the illumination and atmospheric conditions between the two acquisition times may be a potential source of errors and should be taken into account in order to obtain accurate results (Bruzzone et al. 2002). This problem can be addressed performing radiometric calibration of the images. Other pre-processing steps are noise filtration, normalization of an image with respect to the other and topographic correction.
- (2) Feature Extraction - it consists on the extraction of a subset of new features from the original feature set by means of some functional mapping keeping as much information in the data as possible. The features selected depend on the application the change detection algorithm is going to be used for. Examples: Normalized Difference Vegetation Index (NDVI), and Principal Components.
- (3) Dimension reduction – it involves the representation of each image in a new space of lower dimension. In this step the pixel values of each image

can be fitted into polynomial functions or generated by normal distributions. By this, the change detection problem reduces to the comparison of the polynomial coefficients or the normal distributions, to determine if any change occurred.

- (4) Image comparison – The two registered and corrected images or their representation are compared, pixel-by-pixel, in order to generate a further image. A comparison measure used widely because of its simplicity is image differencing. The difference image is computed in such a way that pixels associated with changed areas present gray-level values significantly different from those pixels associated to unchanged areas.
- (5) Decision Rule- it is the most critical step of any change detection method. It involves the selection of a threshold, to decide if a pixel has or has not change from one time to another. The selection of the decision threshold is of major importance as the accuracy of the change mask depends on this choice (Bruzzone et al. 2002) .
- (6) Post-processing- it is needed when change detection results are noisy or inadequately smooth. Change masks could be post processed to adjust the size, shape or boundary of regions representing change (Radke et al. 2003).

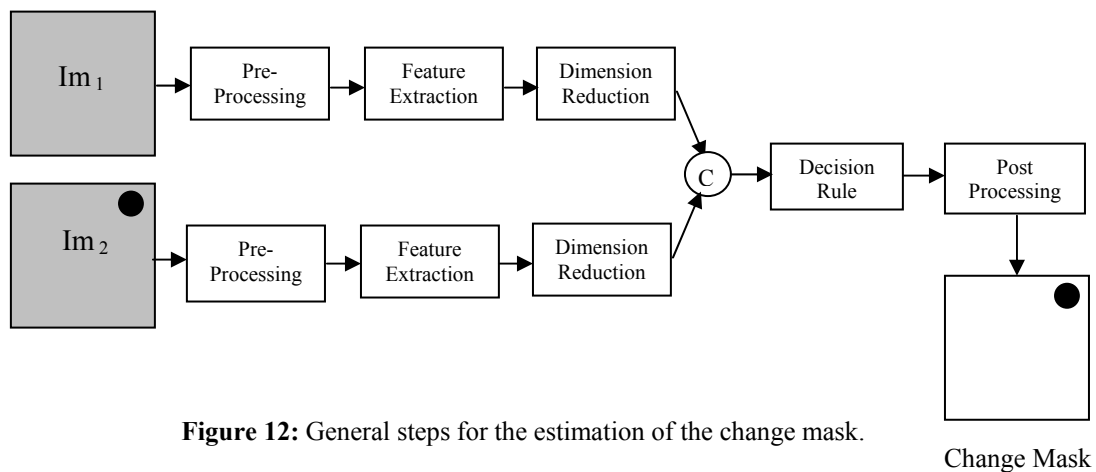


Figure 12: General steps for the estimation of the change mask.

2.9 Change Detection Techniques

Many change detection techniques have been developed and their use depend in the application. As part of this study, the methodology of the following multispectral change detection techniques is described: Image differencing, Conventional Principal Component Analysis (PCA), Change Vector Analysis (CVA), Post Classification Comparison and Multivariate Alteration Detection (MAD). Due to the importance of detecting changes in many fields, new techniques are being developed constantly; nevertheless the ones mentioned above, are the most commonly used with multispectral imagery.

2.9.1 Image Differencing

Image differencing is used widely because of its simplicity to implement and interpret. It consists in the subtraction of the first data image from a second-date image, pixel by pixel. One of the disadvantages that this change detection technique has is that it cannot provide a detailed change matrix, thus it requires a selection of a suitable threshold in order to produce a change/no-change binary map.

Figure 13 presents a detailed diagram describing the change detection process for image differencing.

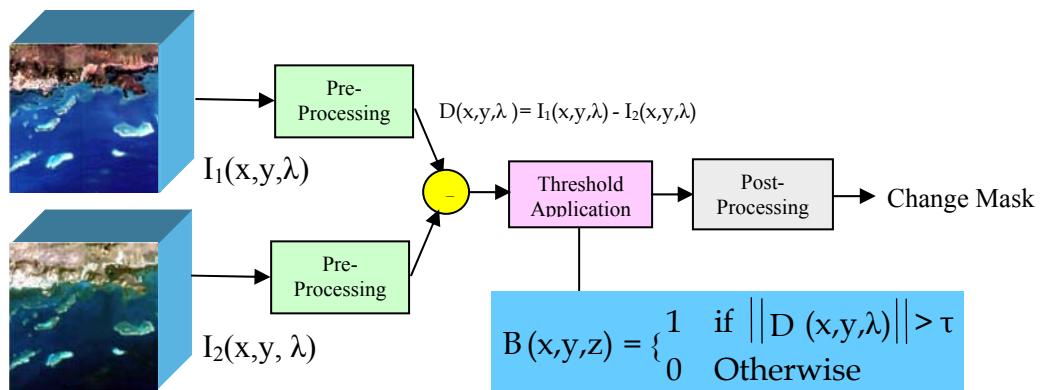


Figure 13: Image Differencing Method.

2.9.2 Change Vector Analysis (CVA)

The Change Vector Analysis (CVA) technique, illustrated in Figure 10, generates two outputs: magnitude and direction of the change vector. Changed areas may then be described in these terms. CVA is a multivariate technique, which accepts as input n-bands, transforms, or spectral features. The first step on this method is to eliminate any redundant information or bands (band transformation) in order to concentrate change analysis in the features of interest.

The change vector can be formed as follows:

$$\Delta I(x, y, \lambda) = I_1(x, y, \lambda) - I_2(x, y, \lambda) = \begin{bmatrix} b_{11}(x, y, \lambda_1) - b_{12}(x, y, \lambda_1) \\ b_{21}(x, y, \lambda_2) - b_{22}(x, y, \lambda_2) \\ \vdots \\ b_{n1}(x, y, \lambda_n) - b_{n2}(x, y, \lambda_n) \end{bmatrix}$$

The next step is to calculate the magnitude of variation among spectral change vectors between the images pairs. The magnitude of the change vector is calculated by the Euclidean distance formula which measures the intensity of change.

$$\|\Delta I(x, y, \lambda)\| = \sqrt{(b_{11}(x, y, \lambda_1) - b_{12}(x, y, \lambda_1))^2 + \cdots (b_{n1}(x, y, \lambda_n) - b_{n2}(x, y, \lambda_n))^2}$$

The angle of the vectors, which indicates the nature of change that occurred, varies according to the number of components used. In other words, each vector is a function of the combination of positive or negative changes through channels or spectral bands.

Figure 14 presents a detailed diagram describing the change detection process for CVA.

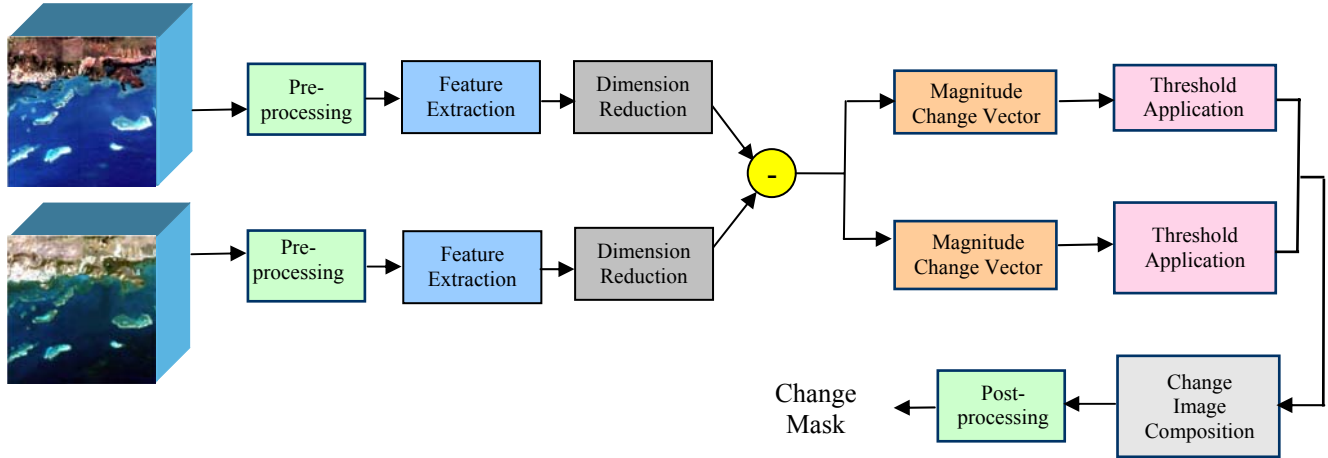


Figure 14: Change Vector Analysis.

2.9.3 Conventional Principal Component Analysis

There are two ways to apply conventional Principal Component Analysis, and they are:

- (1) To put two or more images of different time instants into a single “cube”, then perform the analysis and analyze the minor principal component images for change information (see Figure 15).
- (2) Transform each image cube into its principal components and then apply change vector analysis in the principal component space (see Figure 16).

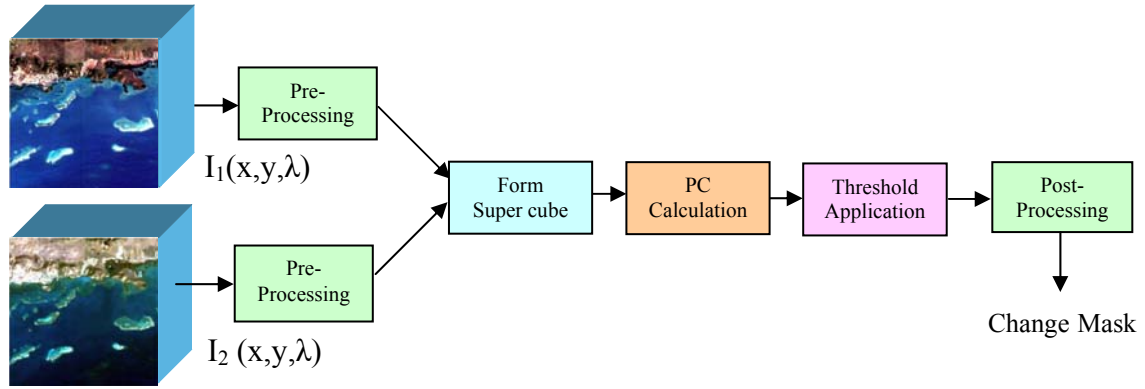


Figure 15: Conventional Principal Component Analysis.

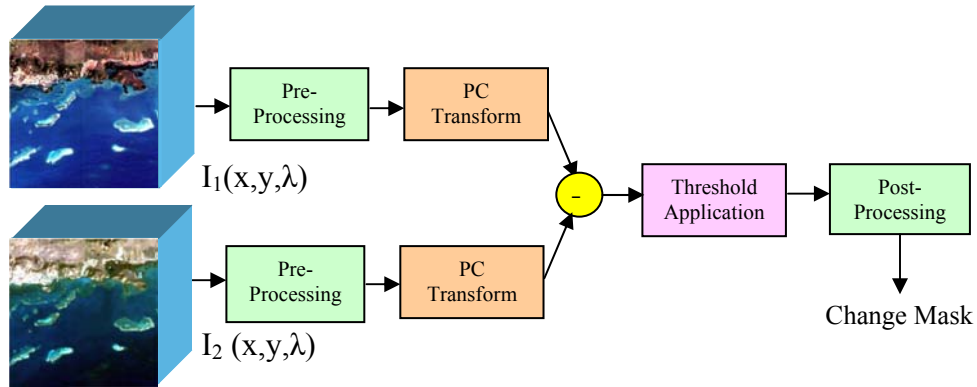


Figure 16: Conventional Principal Component Analysis.

2.9.4 Post-Classification Comparison

This change detection method separately classifies multi-temporal images and then implements comparison of the thematic maps pixel by pixel (see Figure 17). Within the advantages this method offers, it minimizes the impact of atmospheric, sensor and environmental effects between multi-temporal images and also provides a complete matrix of change information. This is there is no need of thresholding. It requires a great amount of time and expertise to create classification products and its final accuracy

depends on the quality of each classified image. If a supervised classification technique is considered, sufficient training sample data for classification is required.

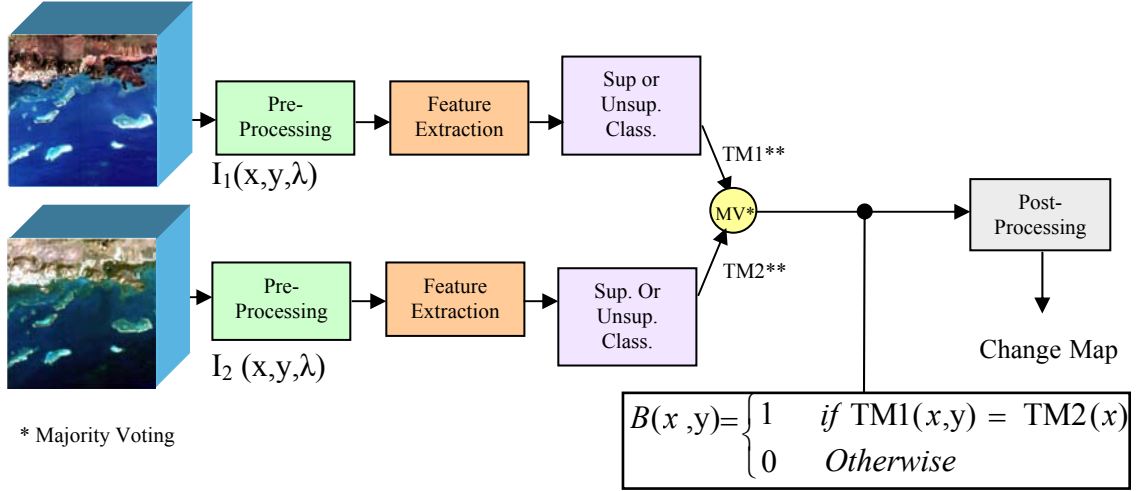


Figure 17: Post-Classification Comparison.

2.9.5 Multivariate Alteration Detection (MAD)

The Multivariate Alteration Detection (MAD) method is based on a classical statistical transformation referred to as canonical correlation analysis and is described in Figure 18. Canonical correlations analysis investigates the relationship between two groups of variables. It finds two sets of linear combinations of the original variables, one for each group. The first two linear combinations are the ones with the largest correlation. This correlation is called the first canonical correlation, and the two linear combinations are called the first canonical variates. The second two linear combinations are the ones with the largest correlation subject to the condition that they are orthogonal to the first canonical variates. This correlation is called the second canonical correlation, and the two linear combinations are called the second canonical variates. Higher-order canonical correlations and canonical variates are defined similarly. Because corresponding pairs of canonical variates are linear combinations of the original variables ordered by correlation

or similarity between pairs, it seems natural to base a change detection scheme on differences between these pairs of variates (Nielsen et al. 1998).

If the multispectral pixel intensities measured at two different times by random vectors X and Y can be represented as follows:

$$X = [X_1 \quad \cdots \quad X_k]^T$$

$$Y = [Y_1 \quad \cdots \quad Y_k]^T$$

Where K is the number of spectral components, then linear combinations are calculated:

$$a^T X = a_1 X_1 + \cdots + a_k X_k$$

$$b^T Y = b_1 Y_1 + \cdots + b_k Y_k$$

Such that the difference of the transformed vectors has maximum variance:

$$\max(\text{var}\{a^T X - b^T Y\})$$

Subjected to the constraints:

$$\text{var}\{a^T X\} = 1$$

$$\text{var}\{b^T Y\} = 1$$

Under these constraints:

$$\text{var}\{a^T X - b^T Y\} = 2(1 - \text{corr}\{a^T X - b^T Y\})$$

Since this method is used for change detection, it is requested that $a^T X$ and $b^T Y$ are positively correlated. Therefore, determining the difference between linear combinations

with maximum variance corresponds to determining linear combinations with minimum (non-negative) correlation,

$$\min(\text{corr}(\{a^T X, b^T Y\}))$$

This complies with canonical correlations analysis (Nielsen et al. 1998).

Then the multivariate alteration detection is defined as:

$$\begin{bmatrix} X \\ Y \end{bmatrix} \rightarrow \begin{bmatrix} a_p^T X_p - b_p^T Y_p \\ \vdots \\ a_1^T X_1 - b_1^T Y_1 \end{bmatrix}$$

where a_i and b_i are the defining coefficients from a standard canonical correlation analysis. The first difference (with the variates $\mathbf{a}_p^T \mathbf{X}$, $\mathbf{b}_p^T \mathbf{Y}$) shows maximum variation. The j th difference (with the variates $\mathbf{a}_{p-j+1}^T \mathbf{X}$, $\mathbf{b}_{p-j+1}^T \mathbf{Y}$) shows maximum variation subject to the constraint of being uncorrelated with the previous ones.

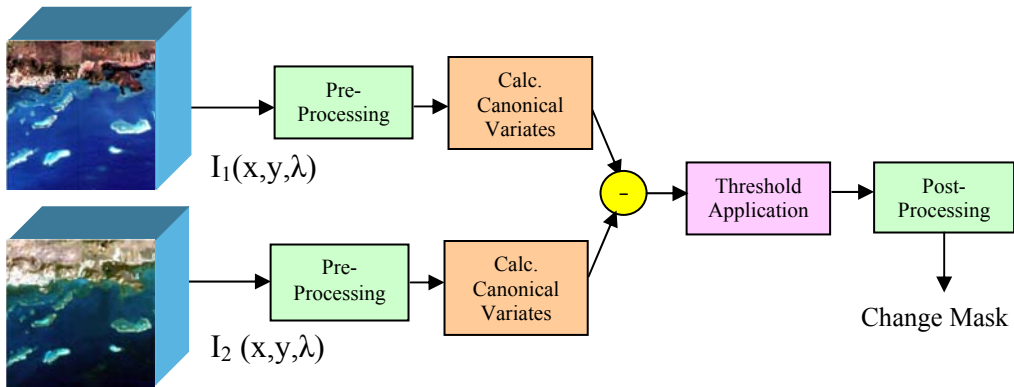


Figure 18: Multivariate Alteration Detection.

2.10 Summary

This chapter presented the basic concepts of image spectroscopy and hyperspectral remote sensing and change detection. The basic steps of the process of change detection as well as the most commonly used methods were described. Basic concepts on supervised and unsupervised image classification were also introduced.

CHAPTER 3

3. Temporal Principal Component Analysis (TPCA)

This chapter presents how Principal Component Analysis was used in this work to analyze temporal data.

3.1 Introduction

The main application of Principal Components Analysis (PCA) is to reduce the dimensionality of a data set that consists of a large number of interrelated variables, while retaining as much as possible of the variability present in the data set. This can be achieved by transforming to a new set of variables, or principal components, which are uncorrelated, and ordered so that the first few retain most of the variability present in the original variables. Applications such as the classification of multi/hyperspectral images benefit from dimensionality reduction of data, since the transformed data set provides sufficient information for the classifier to discriminate between the possible classes present in a given scene. In change detection applications the use of principal components differs from that in classification. As it was mentioned before the principal components are uncorrelated among them. The first components accumulate most of the common variability through the data set, while the latter components include the uncommon features. If PCA is applied to temporal data sets, the latter components will contain the information about features that are not common. This information can be interpreted as the features that are not present in either of the sets, i.e. change information.

3.2 Methodology for TPCA

In this method the images were arranged in the same fashion as in the Image differencing method. But this time a bi-temporal feature space was constructed by placing the two image vectors in the same space. This is,

$$\text{if } X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix},$$

the bitemporal feature space is defined as,

$$Z = \begin{bmatrix} X_1 & Y_1 \\ X_2 & Y_2 \\ \vdots & \vdots \\ X_N & Y_N \end{bmatrix}$$

where N is the number of spectral bands.

By constructing a bitemporal feature space, it is possible to use the general concept of principal components applied to change detection, which interprets that the pixels that have not suffered change will tend to lie around the first principal component and the changed pixels will tend to lie around the second principal component, which is orthogonal to the first. Figure 19 shows the general concept of change detection using principal components.

In this work all the bands are used to construct the bi-temporal feature space. By doing so, it is possible to project in two dimensions all the transformed data obtained from the images. This will lead to an easier interpretation of the results because all the change

information will be expected to be in the second principal component. It also permits to take advantage of all the information in the images. This means that for every band there is going to be a second image component that will offer unique information of change.

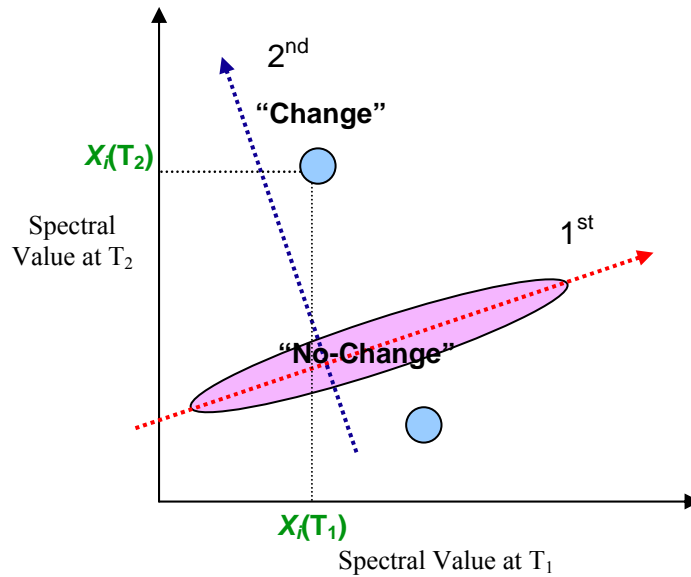


Figure 19: Principal Component Analysis-General Concept

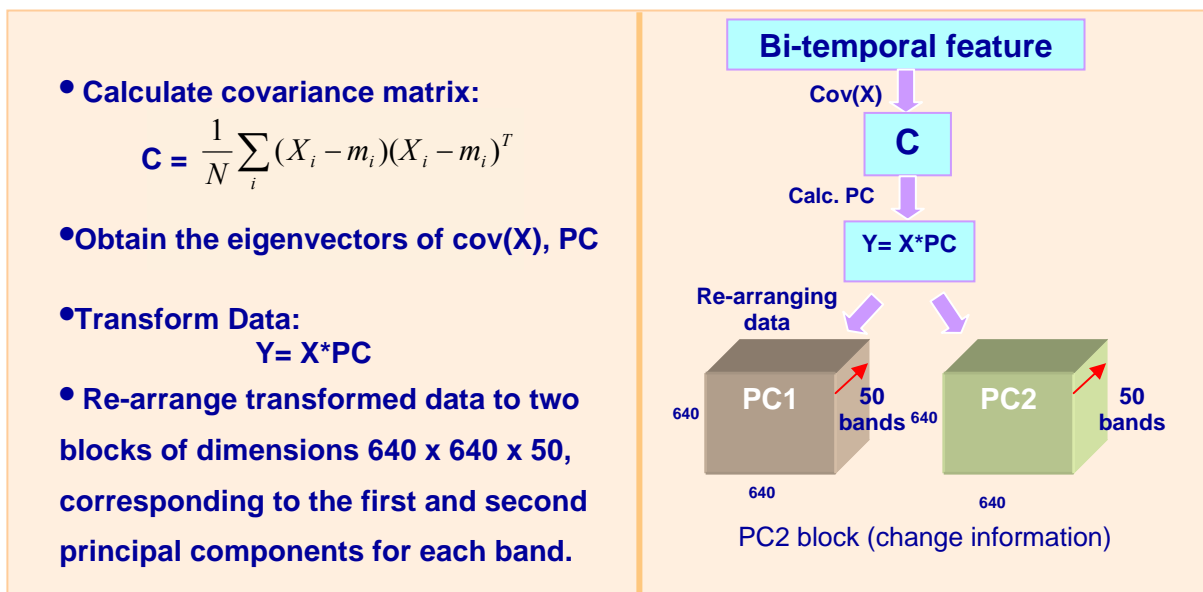


Figure 20: Temporal Principal Component Analysis Procedure.

Figure 20 presents step by step diagram of the modified Temporal Principal Component Analysis method.

Figure 21 corresponds to the change detection diagram for the modified Temporal Principal Component Analysis method using the images from the sensor Hyperion. For the images obtained with the SOC-700 hyperspectral camera, there was no need for the registration or atmospheric correction stage. The images taken with the fundus camera were registered and corrected from illumination effects.

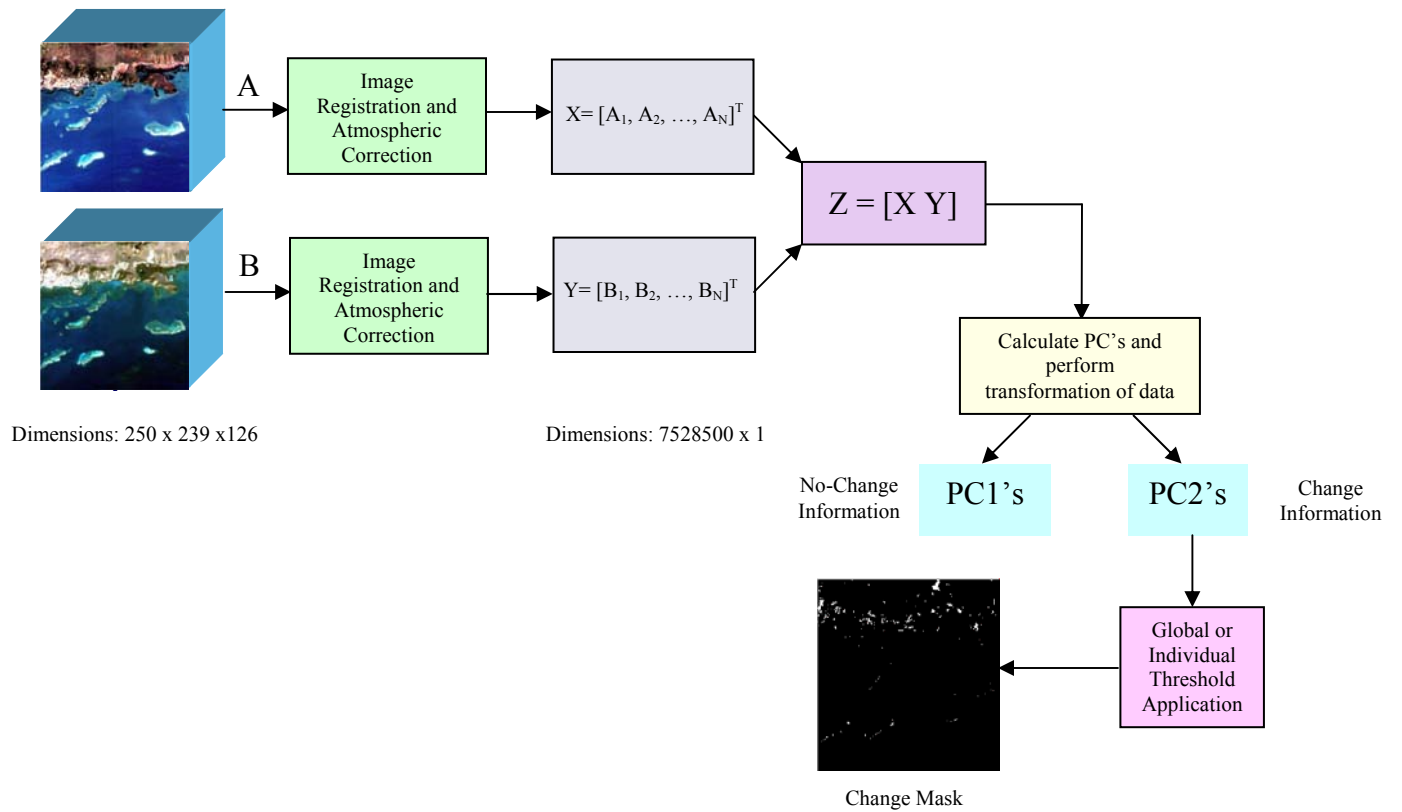


Figure 21: Change detection diagram for the Temporal Principal Component Analysis method.

3.3 Summary

This chapter showed how Principal Components Analysis was used to analyze temporal data. The methodology for the implementation of the change detection method proposed in this study (TPCA) was also presented.

CHAPTER 4

4. Methodology

This chapter presents the procedure that was followed in order to achieve the objectives presented in this work and mentioned in chapter one.

4.1 Methodology used for change detection

The methodology used to compare the performance of the different change detection methods studied in this work is described in this section. For this study, images from a hyperspectral camera, the sensor Hyperion and a Topcon TRG 50IA fundus camera were utilized to test the performance of the algorithms. The images taken with the SOC-700 hyperspectral camera available on campus at the Laboratory for Applied Remote Sensing and Image Processing (LARSIP) correspond to a concrete scene in which changes were induced by placing four coins and a drinking straw. On the other hand, the images from the fundus camera show the effects of diabetes to the human retina and the ones from the spaceborne sensor Hyperion were obtained over Enrique Reef at La Parguera (southwest of Puerto Rico).

As it was described in Chapter 2, the process of change detection involves a series of steps necessary for the estimation of a change mask, which indicates the pixels that have suffered from any kind of transformation. These steps are: image pre-processing, feature extraction, dimension reduction, image comparison, decision rule application, and image post-processing.

4.1.1 Image Pre-Processing

Image preprocessing is one of the steps that some change detection methods has to perform in order to obtain correct change detection results. One of the most common preprocessing procedures performed over all images is registration, which means the alignment of pixels with the same coordinates in the image so it can be associated with the same area on the ground. With regard to the spectral domain, changes in the illumination and atmospheric conditions between the two acquisition times may be a potential source of errors and should be taken into account in order to obtain accurate results.

In this work, the images from the sensor Hyperion were registered and corrected atmospherically. There was no need to register the test images obtained with the SOC-700 hyperspectral camera since registration was taken care of during acquisition. The images from the fundus camera were also registered and corrected from illumination (Narasimha-Iyer et al.).

4.1.1.1 Image Registration

The registration of the images of Enrique Reef in La Parguera was done using routines for image registration from the program ITK (Insight Segmentation and Registration Toolkit). ITK is implemented in C++. ITK is cross-platform, using the CMake build environment to manage the compilation process. In addition, an automated wrapping process generates interfaces between C++ and interpreted programming languages such as Tcl, Java, and Python (using CableSwig). This enables developers to create software using a variety of programming languages (ITK website 2003).

The registration procedure involves the calculation of a transformation matrix from the Hyperion image of the year 2002. This transformation matrix was obtained by using the best bands of that image. After the transformation matrix is produced it is applied to the image of the year 2003 in order to register it. The transformations performed were basically shifts and rotations.

4.1.1.2 Atmospheric Correction

The program used to perform atmospheric correction over the Hyperion images was ACORN (Atmospheric Correction Now) (Gao et al. 2000). ACORN uses look-up-tables calculated with the Modtran4 (Adler-Golder et al. 1998) radiative transfer code to model atmospheric gas absorption as well as molecular and aerosol scattering effects. These modeled atmospheric characteristics are used to convert the calibrated sensor radiance measurements to apparent surface reflectance. The technique uses a fast and accurate look-up-table approach to calculate water vapor amounts on a pixel-by-pixel basis. The user may choose to use the water vapor absorption bands at 940 or 1150nm or both for the water vapor derivations. Additionally the user may input a visibility or ask ACORN to attempt to estimate the visibility from the data. A set of artifact suppression options are also included in the ACORN software.

ACORN is designed to work with all airborne and spaceborne calibrated hyperspectral and multispectral data types including HYPERION, ASTER, LANDSAT, AVIRIS, IKONOS, among others (Gao et al. 2000).

ACORN provides the following specific modes of atmospheric correction:

- Mode 1 Radiative transfer atmospheric correction of calibrated hyperspectral data.
- Mode 1.5 Radiative transfer atmospheric correction of calibrated hyperspectral data with water vapor and liquid water spectral fitting
- Mode 2 Single spectrum enhancement of a hyperspectral atmospheric correction.
- Mode 3 Atmospheric correction using the empirical line method for hyperspectral data.
- Mode 4 User specified convolution of hyperspectral data to multispectral data.
- Mode 5 Radiative transfer atmospheric correction of calibrated multispectral data.
- Mode 5.5 Radiative transfer atmospheric correction of calibrated multispectral data with independently know water vapor image for spatially varying water vapor correction

- Mode 6 Single spectrum enhancement of a multispectral atmospheric correction.
- Mode 7 Atmospheric correction by the empirical line method for multispectral data (Gao et al. 2000).

4.1.2 Change detection

After the pre-processing stage was completed, three change detection algorithms were applied to both images. These were: Image Differencing, Conventional Principal Component Analysis, and Temporal Principal Component Analysis. The latter was described in Chapter 3. Post Classification Comparison was applied only to the HYPERION images of Enrique Reef at La Parguera since it has more classes than the images from the taken with the hyperspectral camera.

4.1.2.1 Image Differencing

In this method, the two images corresponding to different times were placed in vector form, by stacking the columns of each band together. Figure 22 shows the procedure followed to convert the hyperspectral cube into a long vector.

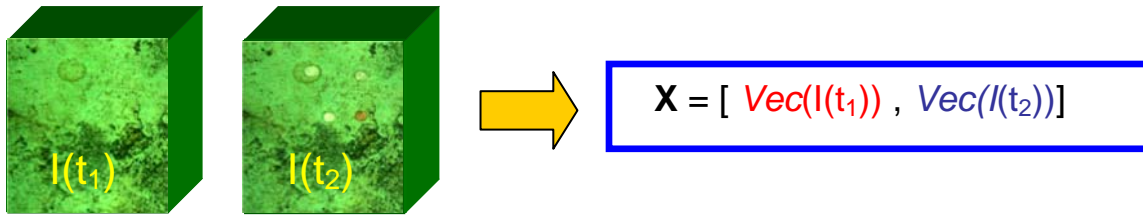


Figure 22: Hyperspectral Cubes conversion to vector form.

Where $\text{Vec}(I)$ converts the image cube into a vector by stacking the columns of each band.

By doing this two vectors were constructed, both corresponding to the images at time one and two respectively. This is, by letting X be a $m \times n$ image at time one and Y , a $m \times n$ image at time two, the new images in vector form are expressed as follows,

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix}$$

where N is the number of spectral bands, and each component X_1, X_2, \dots, X_N , is itself a vector with $m \times n$ rows and one column. Then the image vectors were subtracted in order to obtain a new vector D , containing the information of change. The absolute of the difference was calculated. This is,

$$D = \text{abs}(X - Y) = \text{abs} \left(\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix} - \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix} \right)$$

Since the Image differencing technique does not produce a change matrix, it was necessary to select a threshold in order to generate the binary change/no-change mask. In this work two threshold selection techniques (global and per band threshold selection) were used to generate the change maps. The procedure followed by both methods is described in section 4.1.3.

Image differencing change detection technique is widely used because of its simplicity in the implementation and the interpretation of its results.

4.1.2.2 Conventional Principal Component Analysis

In this approach, hyperspectral images measured at different times are stacked into a supercube and PCA is performed on it. Those components with smaller variances are associated with change.

4.1.2.3 Post-Classification Comparison (HYPERION Images)

As a feature extraction process for classification, we used band subset selection based on Singular Value Decomposition (SVD). The subset band selection algorithm requires as input the number of bands to select, which in turn depends on the dimensionality of the data. The dimensionality of the image input to the band subset selection procedure was determined as the number of principal components needed to explain at least 99.5% of the total variability (Umaña et al, 2003). For this case the dimensionality of the data was 5 and the bands used for classification were: 5, 13, 16, 27 and 56. A supervised classifier (Euclidean Distance) was used to classify each image independently.

After obtaining the corresponding classification maps, they were compared, in order to produce a change map. The training samples were manually selected over the image, by using the classification tool of the Hyperspectral Image Analysis MATLAB Toolbox (Arzuaga et al, 2004). A high spatial-resolution image (IKONOS Sensor) was used to aid in the selection of the training samples. Figure 23 shows the areas from which the samples were selected.

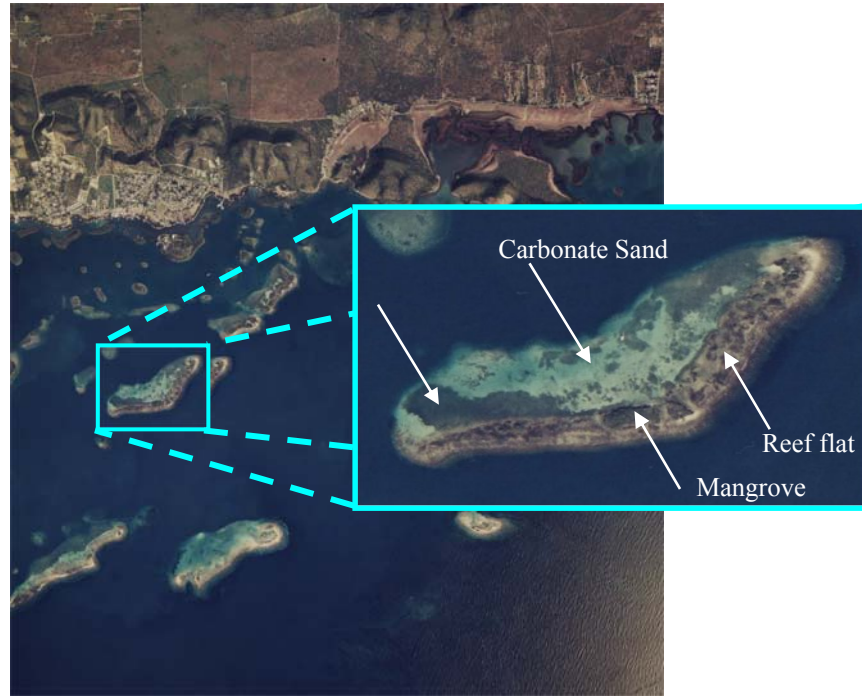


Figure 23: Training Areas

4.1.3 Decision Rule

The decision rule is the most critical step of any change detection method. It involves the selection of a threshold, to decide if a pixel has or has not change from one time to another. In this work, two ways of selecting a threshold were implemented. This section presents the procedures of both threshold selection techniques.

4.1.3.1 Global Threshold Selection

This method calculates a global magnitude image from all the change information bands generated by each change detection technique. Then the standard deviation of the magnitude image is obtained and the threshold is set to three times this value. Figure 24 shows the procedure of the global threshold selection method.

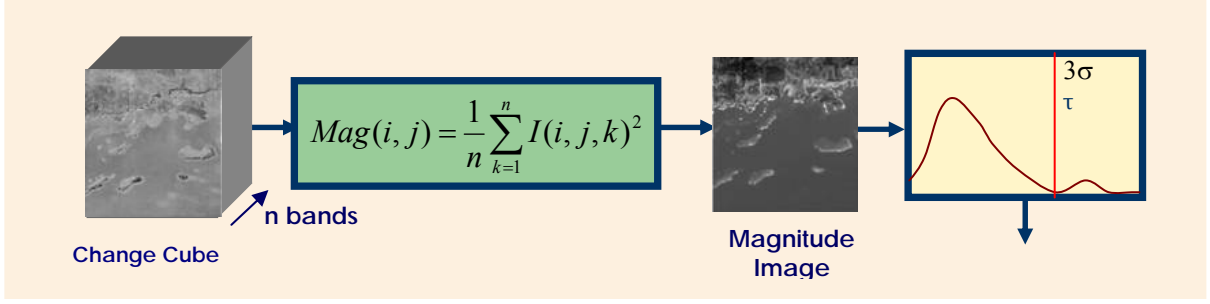


Figure 24: Procedure for the application of a global threshold.

4.1.3.2 Per-band Change detection

This method produces per-band change masks from each one of the change information bands. These masks are produced by applying thresholds to each change band, using the same three-times the standard deviation criterion. Then we take the sum of all the maps and generate a final change mask taking into account the level of noise of the images. Then if the number of bands that labeled a certain pixel as change is greater or equal than:

$$\%Noise\ Level \times n,$$

where n is the number of total bands, the pixel is identified as change in the final mask. By doing this we pretend to minimize the probability that a pixel may be classified as change because of noise.

Figure 25 shows the procedure of the threshold selection technique discussed in this section.

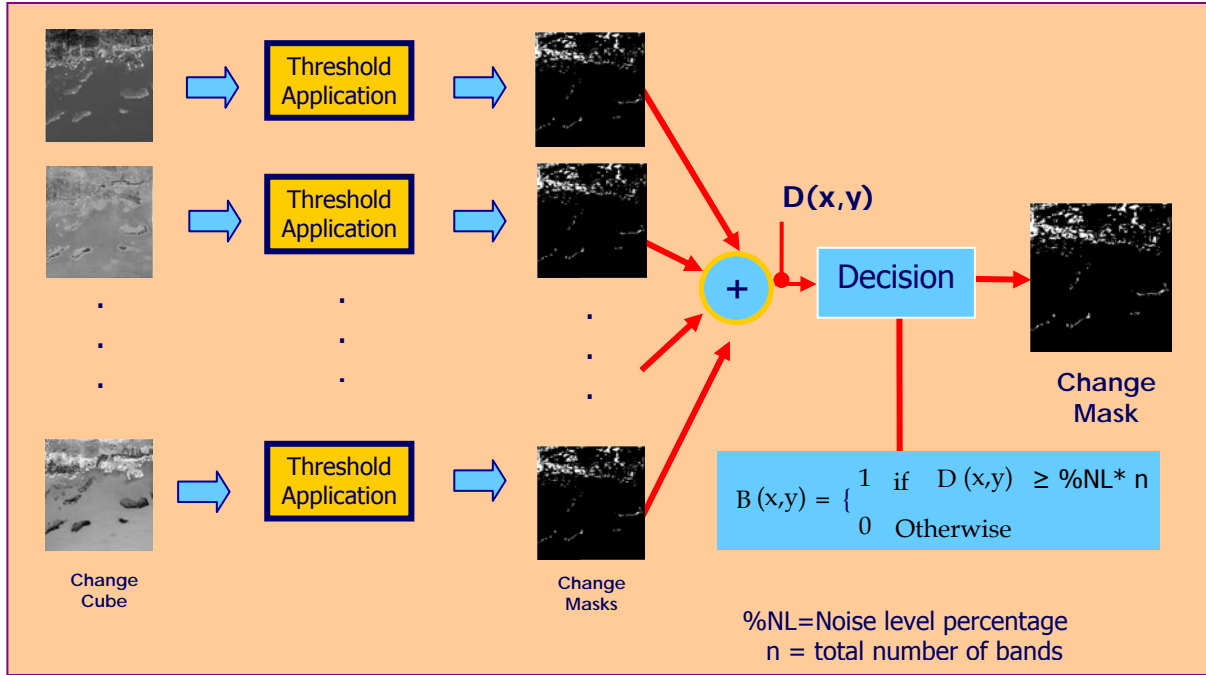


Figure 25: Procedure for the application of per-band change detection approach.

4.2 Hyperspectral Camera

For this study, the images used for the implementation of the different change detection methods were obtained from the sensor Hyperion and the hyperspectral camera, available at the Laboratory of Remote Sensing and Image Processing (LARSIP). The model of the camera is SOC-700 built by the Surface Optics Corporation (see Figure 26). It covers the spectral range from 0.4 to 0.9 μm in 120 bands (640 x 640), with a spectral resolution of 4nm.

The images obtained with the hyperspectral camera were used to test the algorithms under controlled conditions.



Figure 26: SOC-700 Hyperspectral Camera (from Surface Optics web page)

4.3 Hyperion Sensor

The Hyperion sensor onboard the EO-1 satellite is the first hyperspectral sensor on an Earth Observation Satellite. It covers the reflective spectral range from 0.4 to 2.5 μm at 10nm resolution. Such comprehensive spectral resolution permits very detailed land cover classifications or identifications to be performed. The instrument images a 7.5 km by 100 km land area per scene providing detailed spectral mapping across all 220 channels.



Figure 27: Hyperion Sensor (NASA EO1 Webpage)

Table 2 presents the general specifications of the Hyperion sensor.

Table 2: Hyperion Sensor Specifications

Specifications	
IFOV	0.043 mrad
Bands	220
Altitude	705 km
Swath	7.5 km
Range	VNIR: 355 – 1000 nm SWIR: 900 – 2577 nm

4.4 Topcon TRG 50IA Fundus Camera

Figure 28 shows a picture of the instrument (fundus camera) used to take the images from the retina. The model of this camera is TRC 50IA and was built by the company Topcon (Tokyo, Japan). This camera takes digital photos of the retina and stores them in 35mm films. These images aid to monitor and study the transformations the retina may suffer as a consequence of diabetes.



Figure 28: Topcon TRC 50IA fundus camera (from www.colcon.com).

4.5 Hyperspectral Data

4.5.1 Test Images

Two sets of images were used in this work to evaluate the performance of the algorithms, under controlled conditions. These images were taken with the SOC-700 hyperspectral camera. The first set of these images (see Figures 29 to 30) corresponds to a scene of a concrete sidewalk in which changes were induced deliberately by placing four metal coins. In the second set of images (see Figures 31 to 32), a drinking straw was placed over the sidewalk as well. In this case, the changes were caused by the object and its shadow.

The images were originally 120 bands each, and then reduced to 50, when most of the noisy bands were eliminated. This reduction was performed by using Singular Value Decomposition (Velez 1998).

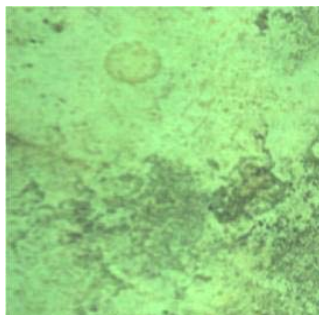


Figure 29: Test hyperspectral image of sidewalk.

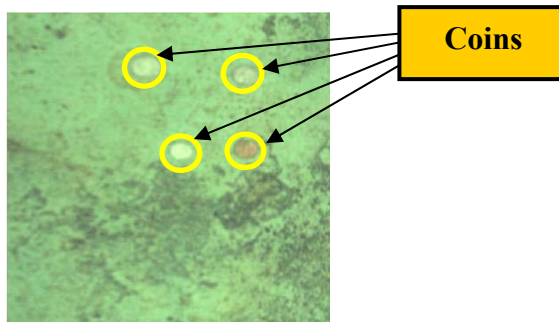


Figure 30: Test hyperspectral image with coins.

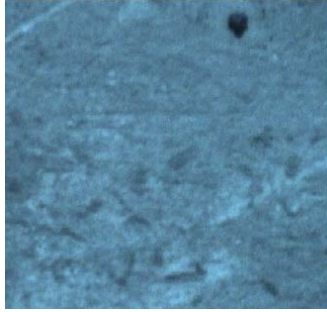


Figure 31: Test hyperspectral image of sidewalk.

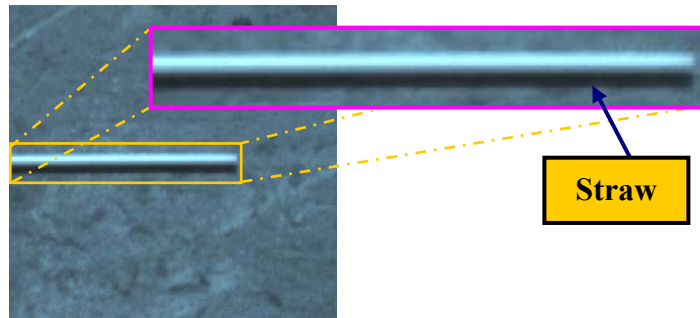


Figure 32: Test hyperspectral sidewalk image with straw on it.

4.5.2 HYPERION Data – Benthic Habitat of Enrique Reef at La Parguera

The images obtained from the Hyperion sensors corresponding to the benthic habitat of Enrique Reef at La Parguera (Lajas, Puerto Rico) were used for change detection. The images have an area of 27 x 46 pixels and the number of bands used in the analysis was 126.

The images were taken in August 28, 2002 and July 15, 2003, respectively. Figure 33 present the location of the municipality of Lajas and shows an approximate location where Hyperion images were taken. Figures 34 and 35 present RGB composites of the HYPERION images.



Figure 33: Area of La Parguera, taken by Hyperion (This map is from Lonely Planet web site).



Figure 34: Image at time one, taken on August 28, 2002

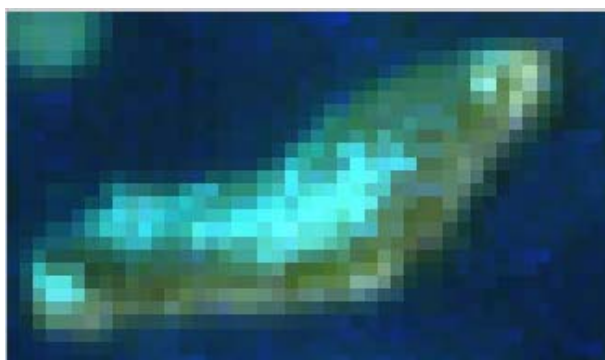


Figure 35: Image at time two, taken on July 15, 2003.

4.5.3 Fundus Camera Data- Images from a Human Retina

Two images corresponding to a human eye were taken using a Topcon TRC 50IA fundus camera in order to detect changes related to the effects of diabetes on the retina. Figure 36 shows the images used for the analysis, in which the most noticeable changes are the enlargement of the geographic atrophy and the presence of small dot hemorrhages. The images have an area of 640 x 700 and are of three bands each.

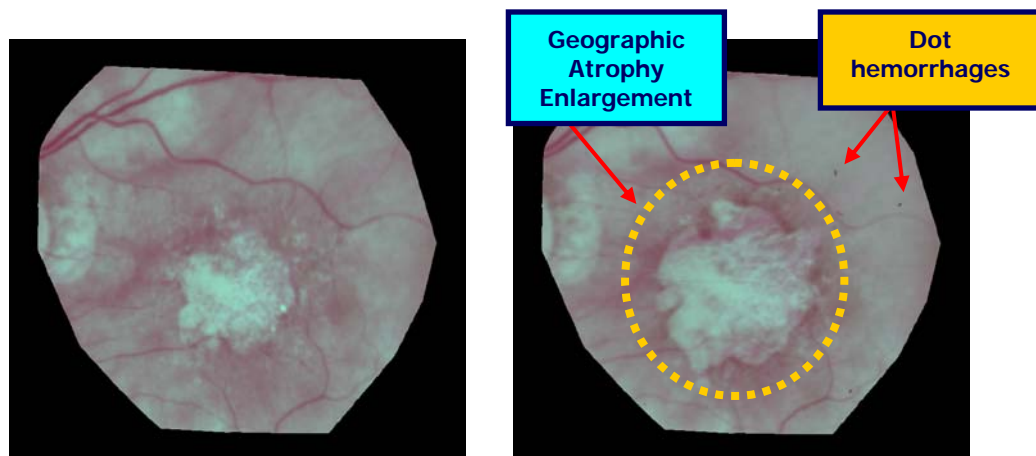


Figure 36: Images taken with the Topcon TRC 50IA fundus camera.

4.6 MATLAB Hyperspectral Image Analysis Toolbox

The hyperspectral image analysis toolbox (see Figure 37) developed at the Laboratory for Applied Remote Sensing and Image Processing (LARSIP) of the University of Puerto Rico at Mayaguez was used to aid in the reduction of bands and the supervised classification of the images for the post-classification comparison method.

The Hyperspectral Image Analysis Toolbox (HIAT) is a collection of algorithms that extend the capability of the MATLAB numerical computing environment for the processing of hyperspectral and multispectral imagery. The purpose of the HIAT Toolbox

is to provide information extraction algorithms to users of hyperspectral and multispectral imagery in environmental and biomedical applications. HIAT has been developed as part of the NSF Center for Subsurface Sensing and Imaging (CenSSIS) Solutionware that seeks to develop a repository of reliable and reusable software tools that can be shared by researchers across research domains. HIAT provides easy access to supervised and unsupervised classification algorithms developed at LARSIP over the last 8 years (Rosario-Torres 2005).

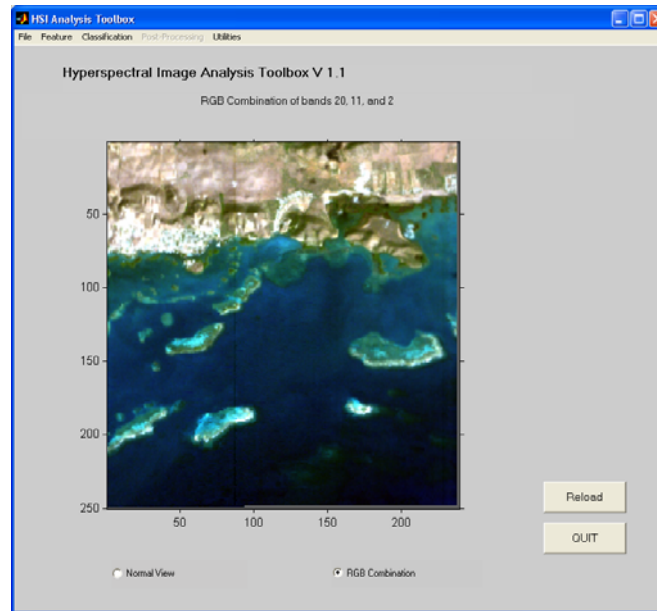


Figure 37: Hyperspectral Image Analysis Toolbox.

Within the features the hyperspectral images analysis toolbox has, there are the following:

- Load and Save options for different image formats such as MATLAB data (.mat), Remote Sensing binary formats (bil, bis, bsq), jpeg and tiff.
- Images can be visualized, as RGB composites or band by band.
- The toolbox provides for the selection of the following available feature extraction/selection techniques:

1. Principal Components Analysis
2. Discriminant Analysis
3. Singular Value Decomposition (SVD)
4. Band Subset Selection
5. Projection Pursuit
6. Optimized Information Divergence Projection Pursuit

- Observation of the spectral responses at each pixel in the images.
- Supervised and unsupervised classification options. The classification techniques available in the toolbox are the following:

1. Euclidean Distance
2. Fisher's Linear Discriminant Analysis
3. Mahalanobis Distance
4. Maximum Likelihood
5. Angle Detection

Training and testing samples for the spectral classes present in the image can be selected by the user in order to perform classification. The unsupervised classification module also allows the user to select the stopping criteria for unsupervised classification. The stopping criteria available in the toolbox are:

1. Bhattacharyya Distance
2. Covariance matrix
3. Pixel quantity of variation
4. Sum of squared error
5. Sum of squared error with covariance matrix information

- Post-processing techniques in order to enhance classification results. These are:

1. Supervised & Unsupervised ECHO classifier

2. Spatial Information Integration based on Markov Random Fields

- After classification is finished, a color map is produced. Each color represents a different class present in the scene.

4.7 Summary

The methodology for the implementation of the change detection algorithms studied in this work was presented in this chapter. Information about the instruments used to obtain the images as well as the tools for their analysis was also described.

CHAPTER 5

5. Experimental Results

In order to evaluate the performance of the different change detection methods studied in this work, several experiments were conducted by using test and validation data. Two sets of images were used as control, to calculate how many missed changes and false alarms produced each change detection method. The images were obtained from the SOC-700 hyperspectral camera, and correspond to concrete sidewalk, where changes were induced by placing four coins and a drinking straw in the scene. Images corresponding to Enrique Reef in La Parguera and a human retina were also analyzed.

5.1 Test Data – Coins over concrete

This section presents test results obtained with hyperspectral images taken with the SOC-700 camera where changes were induced by placing four coins and a straw over a concrete sidewalk.

5.1.1 Image Differencing with Global Threshold Selection

Figure 38 shows the histogram and magnitude of the change vector obtained by subtracting the two images. From the histogram a global threshold was calculated and set to three times the standard deviation of the pixels values data. In this case the mean value and the standard deviation were 0.0248 and 0.0282, respectively. The threshold value used to produce the binary change mask was 0.1094. Figure 39 presents the change mask generated after applying the global threshold.

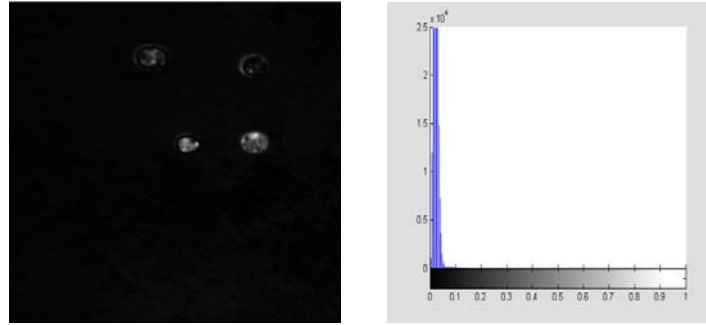


Figure 38: Image Difference Magnitude of Change



Figure 39: Image Difference Change Mask

5.1.2 Image Differencing with Per-band Threshold Selection

Figure 40 shows the change mask obtained by applying per-band thresholds to each of the change information bands. As can be seen the algorithm was capable to detect changes better than in the case of global threshold selection, in which failed to label as change most of the coins.

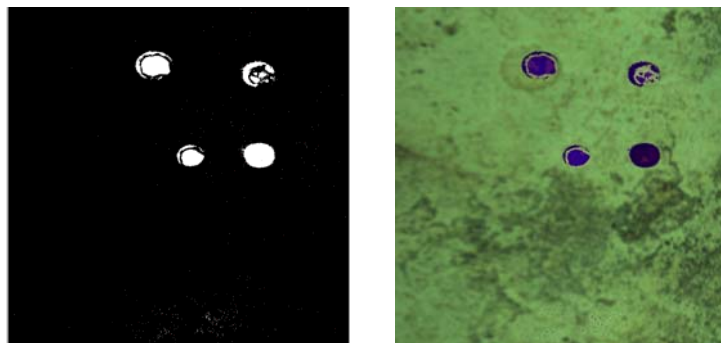
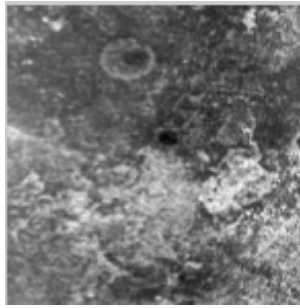


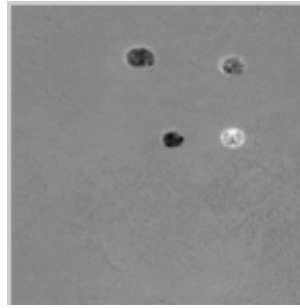
Figure 40: Image Difference Change Mask from individual thresholds.

5.1.3 CPCA with Global Threshold Selection

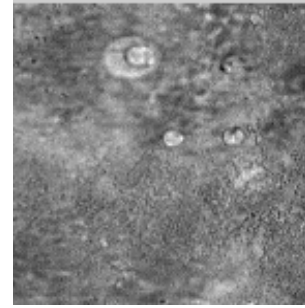
Figure 41, presents the first 12 principal components obtained by performing the analysis over the coins data. We can see that each component offers different information. This is due, as it was mentioned earlier, to their orthogonal property. Table 3 shows the variances and the percentages of variability. The first image component (see Figure 35a) contains the no-change information while the remaining ones are expected to highlight changes. For this study, the components considered to carry change information were the ones outside 90% of the total percentage of variability. Therefore the image components that should be taken into account for change detection in this case were PC2 to PC50. Nevertheless by looking at the principal components from PC8 to PC50 it was clear that they were noisy, thus there was not any substantial change information that could be obtained from them.



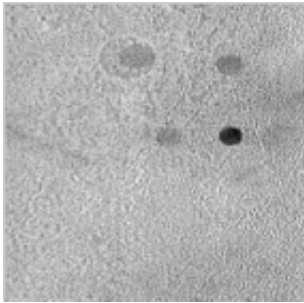
(a) PC1



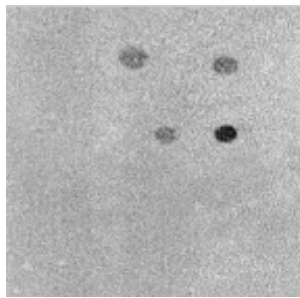
(b) PC2



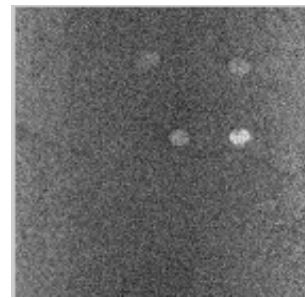
(c) PC3



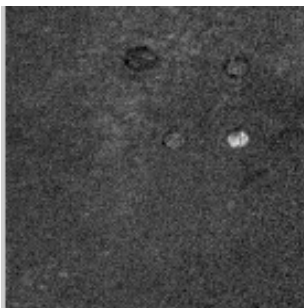
(d) PC4



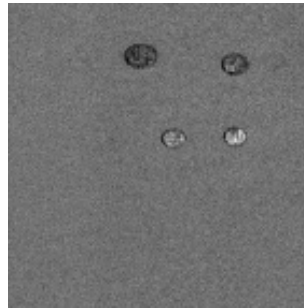
(e) PC5



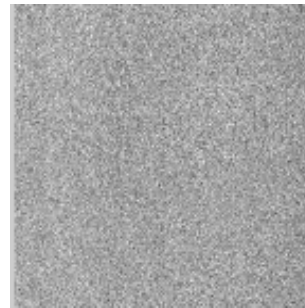
(f) PC6



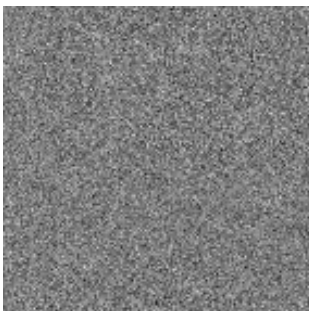
(g) PC7



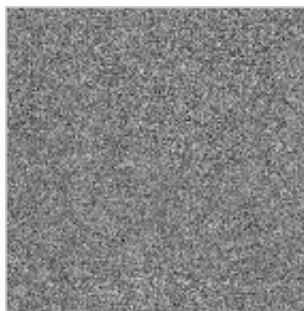
(h) PC8



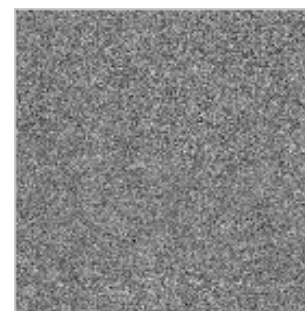
(i) PC9



(j) PC10



(k) PC11



(l) PC12

Figure 41: CPCA Principal Components.

Table 3: Percentages of Variability and variance for each principal component.

Component	% of variability	Eigenvalue
1	95.248	1.6272e+005
2	1.2451	2188.6
3	0.83702	1424.3
4	0.31579	562.58
5	0.10936	192.77
6	0.071902	120.84
7	0.05744	96.648
8	0.052216	87.651
9	0.03582	58.194
10	0.034254	53.689
11	0.034063	53.094
12	0.033845	52.722

Figure 42 presents the histogram and the magnitude of change obtained from the principal components 2 to 8. After the magnitude was obtained, the mean and the standard deviation of the data were calculated and the values were 0.0137 and 0.0126 respectively. The threshold selected then was 0.0515. Figure 43 shows the change mask corresponding to the conventional principal component analysis using global threshold selection to produce a change map. In this case the algorithm was able to detect more changes than Image differencing with global threshold selection.

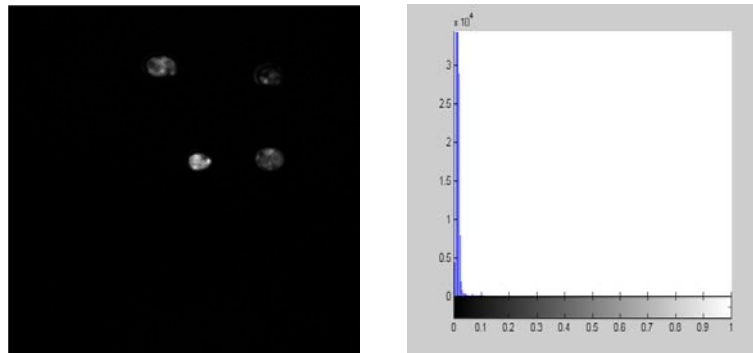


Figure 42: CPCA Magnitude of Change.

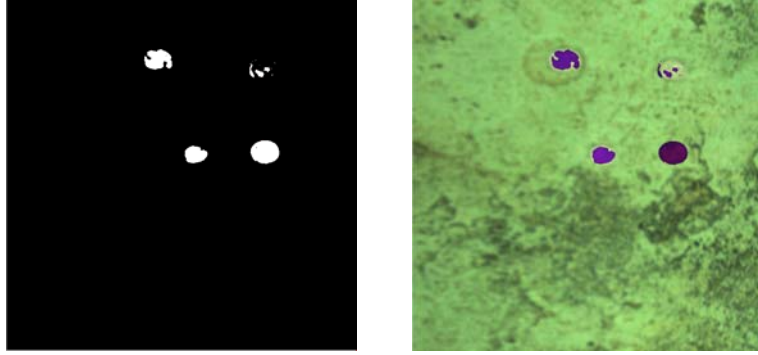


Figure 43: CPCA Change Mask, components 2-12.

5.1.4 CPCA with Per-band Threshold Selection

Figure 44 presents the change mask obtained from the per-band thresholds selection. In this case thresholds were calculated for each principal component separately. As a result seven change maps were generated (from PC2 to PC8). These maps were then compared to generate a final change mask. As can be seen the algorithm was capable to detect changes better than in the case of CPCA with global threshold selection.

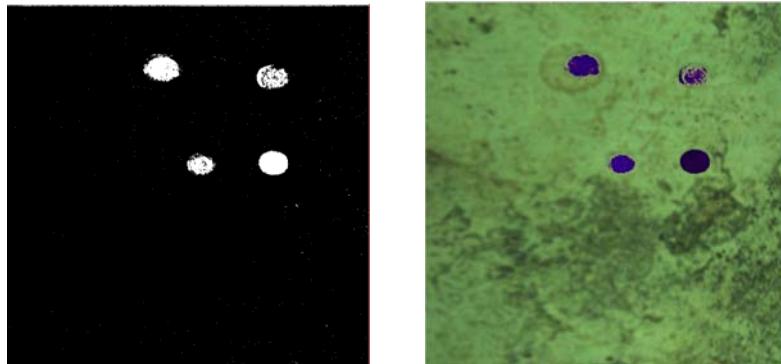


Figure 44: CPCA Change Mask with individual threshold selection.

5.1.5 TPCA with Global Threshold Selection

Figure 45 shows the histogram and the magnitude image of the second principal component cube obtained from the modified Temporal Principal Component Analysis method. The mean and the standard deviation values were 0.0067 and 0.0344 respectively.

The threshold was set to 0.1099. Figure 46 presents the change mask obtained after applying the threshold. In this case the algorithm was able to detect more changes than CPCA and Image differencing with global threshold selection.

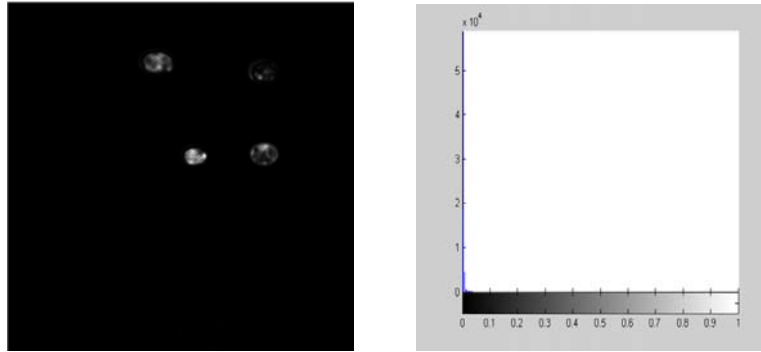


Figure 45: TPCA Magnitude of Change with global threshold.

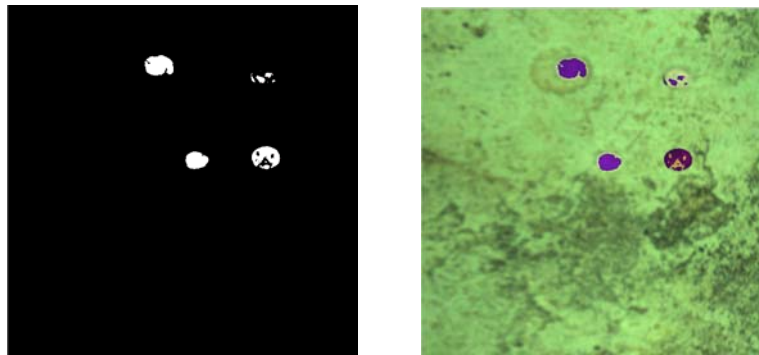


Figure 46: TPCA Change Mask with global threshold selection.

5.1.5 TPCA with Per-band Threshold Selection

Figure 47 shows the change mask obtained from per-band threshold selection in TPCA. It was constructed by using all the second principal image components obtained after the analysis. A local threshold was applied to each component, generating different change maps. These maps were then compared to generate a final change mask. As can be seen the algorithm was capable to detect changes better than in the case of TPCA with global threshold selection.

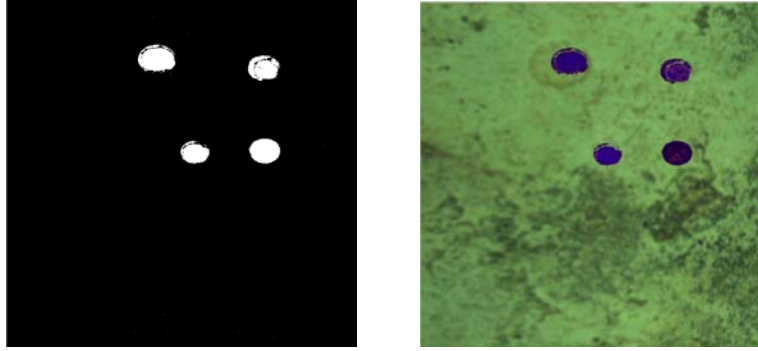


Figure 47: TPCA Change Mask with local threshold selection.

5.2 Test Data – Plastic straw over concrete sidewalk

This section presents the results obtained with the hyperspectral images in which changes were introduced by placing a plastic drinking straw over the concrete surface.

5.2.1 Image Differencing with Global Threshold Selection

Figure 48 corresponds to the change mask produced by the difference of the two image vectors. A histogram of the magnitude was calculated and the global threshold was set to three times the standard deviation of the pixels values data. In this case the mean value and the standard deviation were 0.0081 and 0.0606, respectively. The threshold value used to produce the binary change mask was 0.19.

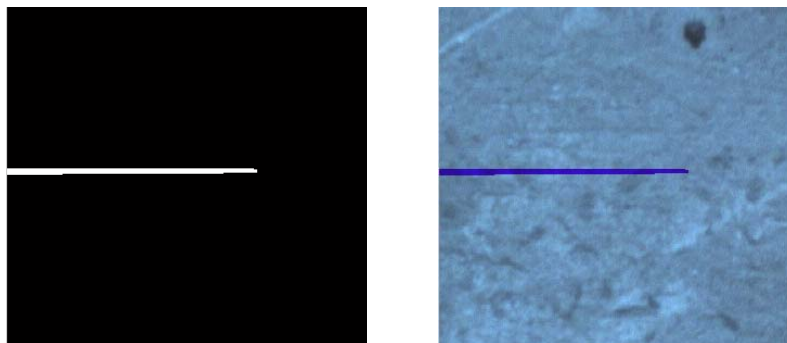


Figure 48: Image Difference Change Mask from global threshold selection.

5.2.2 Image Differencing with Per-band Threshold Selection

Figure 49 shows the change mask obtained by applying per-band thresholds to each of the change information bands. As can be seen the algorithm was capable to detect changes better than in the case of Image differencing with global threshold selection. In this case the changes related to the shadow of the straw were highlighted, unlike the previous case.

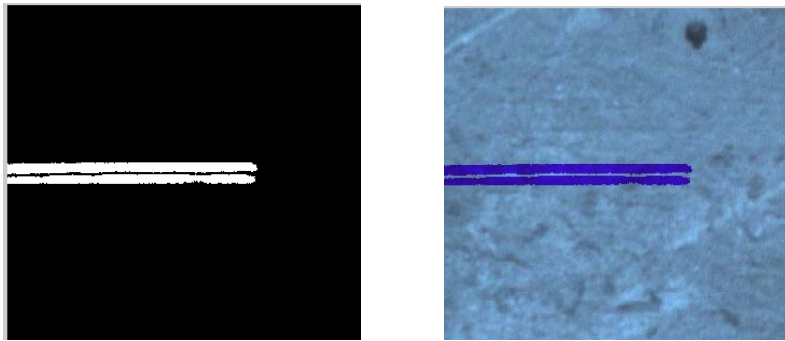
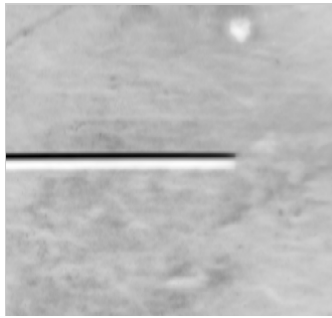


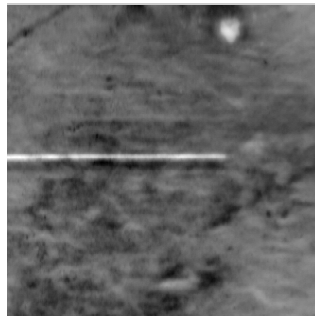
Figure 49: Image Difference Change Mask from local thresholds selection.

5.2.3 CPCA with Global Threshold Selection

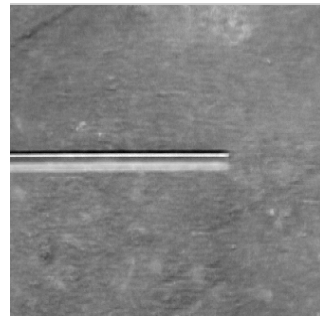
Figure 50, presents the image principal components and Table 4 shows the eigenvalues and the percentages of variability obtained from their variances. The first image component (see Figure 50a) contains the no-change information while the remaining ones are expected to highlight changes. For this study, the components considered to carry change information were the ones outside 90% of the total percentage of variability. Therefore the image components that were taken into account for change detection in this case were PC3 to PC50. Nevertheless by observing the principal components from PC9 to PC50 it was clear that they were noisy, thus there was not any substantial information that could be obtained from them.



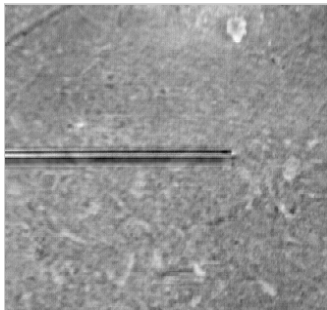
(a) PC1



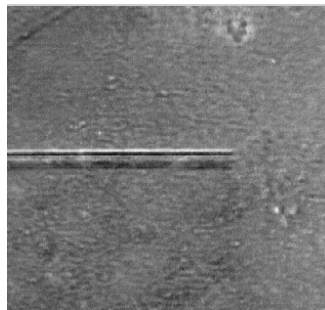
(b) PC2



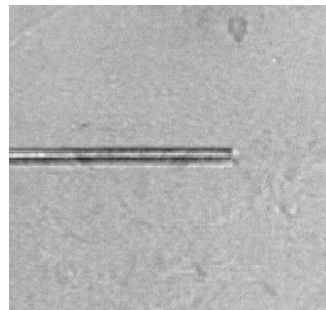
(c) PC3



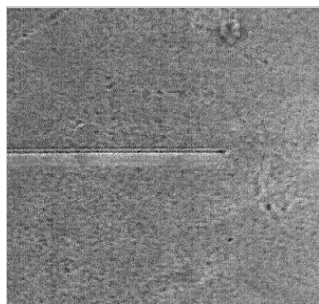
(d) PC4



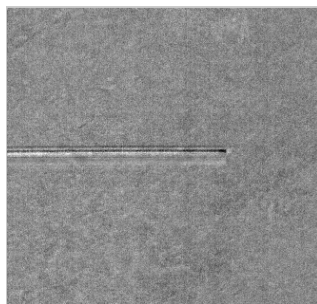
(e) PC5



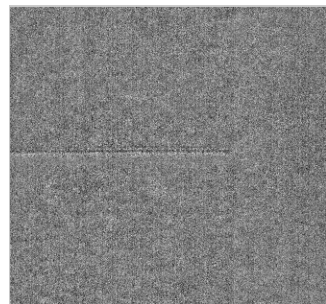
(f) PC6



(g) PC7



(h) PC8



(i) PC9

Figure 50: CPCA Principal Components.

Table 4: Percentages of Variability and eigenvalues of each principal component.

Component	% of variability	Eigenvalue
1	86.316	3.1875e+005
2	11.088	40947
3	0.80552	2974.6
4	0.39854	1471.7
5	0.13422	495.66
6	0.075569	279.06
7	0.047663	176.01
8	0.037209	137.4
9	0.032516	120.08

After the magnitude was obtained, the mean and the standard deviation of the data were calculated and the values were 0.1965 and 0.0838 respectively. The threshold selected then was 0.4479. Figure 51 shows the change mask corresponding to the conventional principal component analysis using global threshold selection to produce a binary change/no-change map. By looking at the change mask it can be noticed that the algorithm was not able to detect the majority of changes produced by the straw and its shadow. However it increased the amount of false alarms by identifying as change some areas of the concrete sidewalk. In this case the algorithm was not able to perform better than Image differencing with global threshold selection.

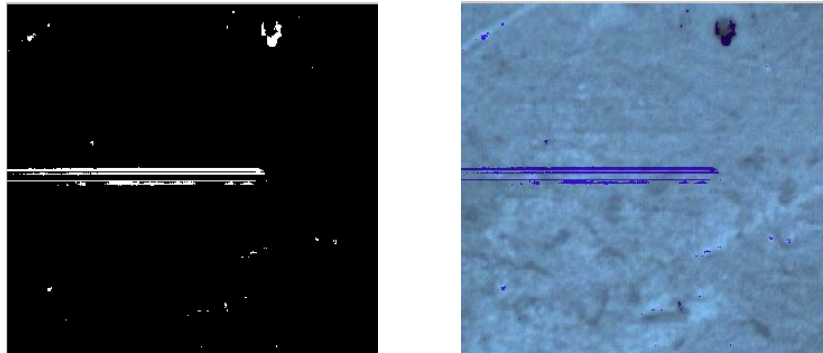


Figure 51: CPCA Change Mask, components 3-8.

5.2.4 CPCA with Per-band Threshold Selection

Figure 52 presents the change mask obtained from the per-band threshold selection. It was constructed by using the same principal components from the global threshold selection. This time a threshold was applied to each component, generating different change masks. These masks were then compared to generate a final change map. The results obtained for this case were better than those from CPCA with global threshold selection. As can be seen the number of missed changes and false alarms decreased, unlike the previous case.

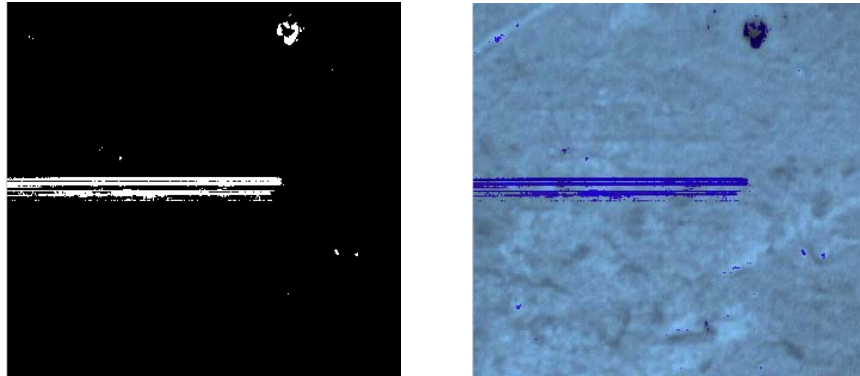


Figure 52: CPCA Change Mask with individual threshold selection.

5.2.5 TPCA with Global Threshold Selection

Figure 53 corresponds to the change mask produced by the TPCA change detection method. A histogram of the magnitude was calculated and the global threshold was set to three times the standard deviation of the pixels values data. The mean and the standard deviation values were 0.0094 and 0.0596 respectively. The threshold was set to 0.1882. In this case the algorithm was able to perform better than Image differencing and CPCA with global threshold, by decreasing the amount of misses.

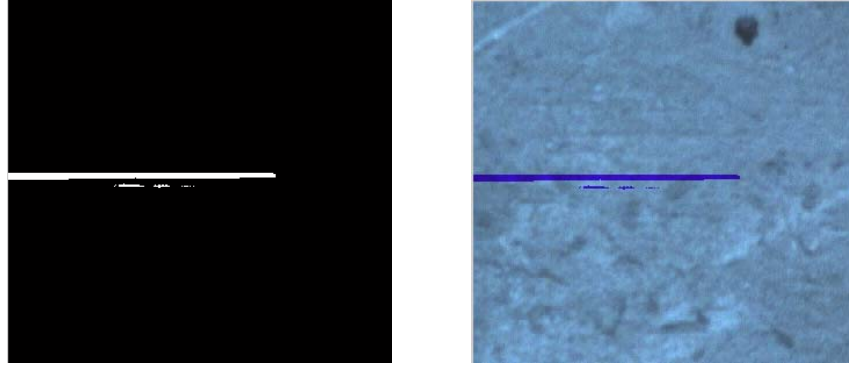


Figure 53: TPCA Change Mask with global threshold selection.

5.2.6 TPCA with Per-band Threshold Selection

Figure 54 presents the change mask obtained from the per-band thresholds selection using Temporal Principal Component Analysis. As can be seen the algorithm was capable to detect changes better than in the case of TPCA with global threshold selection, by detecting most of the shadow produced by the object. In comparison to the previous techniques, TPCA also obtained the best performance.

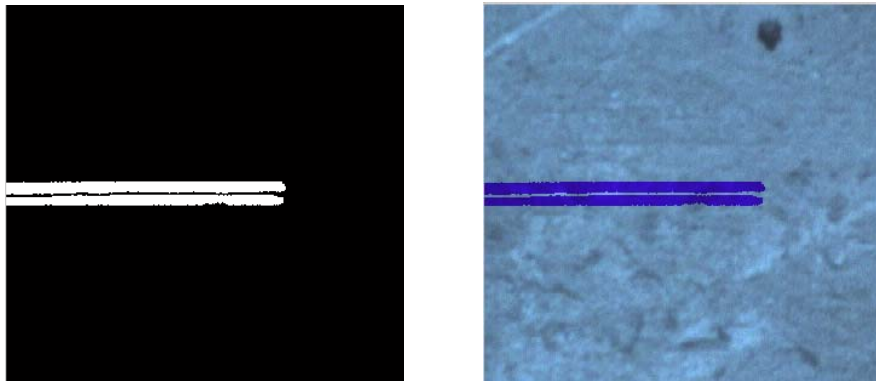


Figure 54: TPCA Change Mask with local threshold selection.

5.3 Algorithm Performance

The performance of the algorithms was tested in terms of missed changes and false alarms. Since the changes in the images were induced deliberately it was possible to

construct a reference mask in order to calculate which pixels that represented changes were classified as background and which pixels that represented background were classified as changes.

Figure 55 shows all the change masks obtained from the different change detection methods, by using global threshold selection. Table 5 shows the percentage of missed changes and false alarms obtained. The procedure was repeated for the per-band threshold selection technique and the results are presented in Table 6. It is possible to notice that the overall error was higher for the methods with the global threshold selection than with per-band selection. Regardless of the way of how thresholds were selected, the method that produced the smallest percentage of error was Temporal Principal Component Analysis. Figures 57 and 58 show the results obtained for the image corresponding to a drinking straw over a concrete sidewalk.

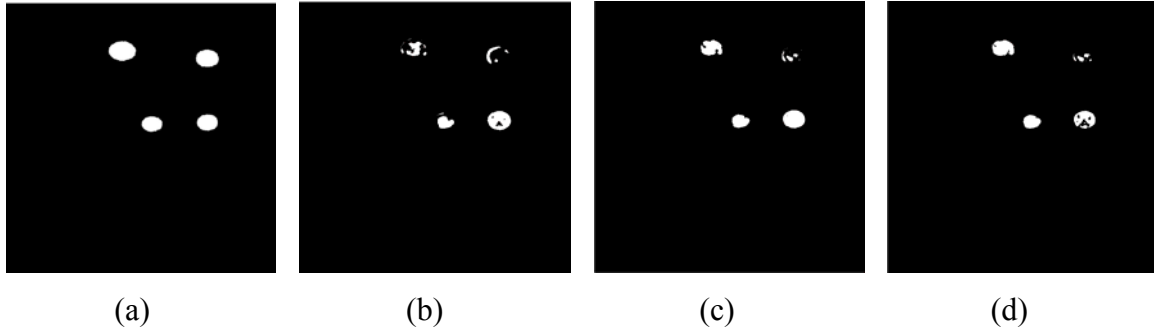


Figure 55: (a) Reference Image, (b) Image difference, (c) CPCA and (d) TPCA change masks.

Table 5: Methods Performances using global threshold selection.

Method	Im-Diff	CPCA	TPCA
Misses	53.4279%	40.2482%	36.9385%
False Alarms	0.1964%	0.0861%	0.0707%
Overall Error	1.0759%	0.7498%	0.6799%

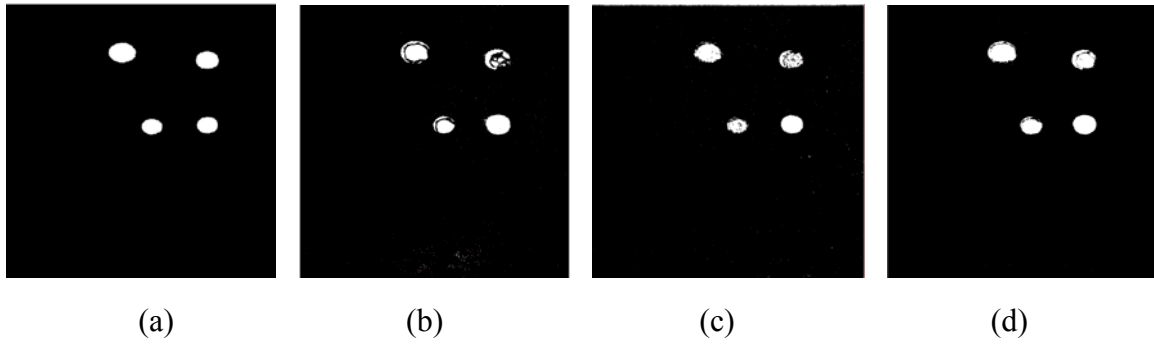


Figure 56: (a) Reference Image, (b) Image difference, (c) CPCA and (d) TPCA change masks.

Table 6: Methods Performances using per-band threshold selection.

Method	Im-Diff	CPCA	TPCA
Misses	25.2955%	31.4569%	5.7033%
False Alarms	0.5926%	0.1053%	0.3969%
Overall Error	1.000%	0.6233%	0.4846%

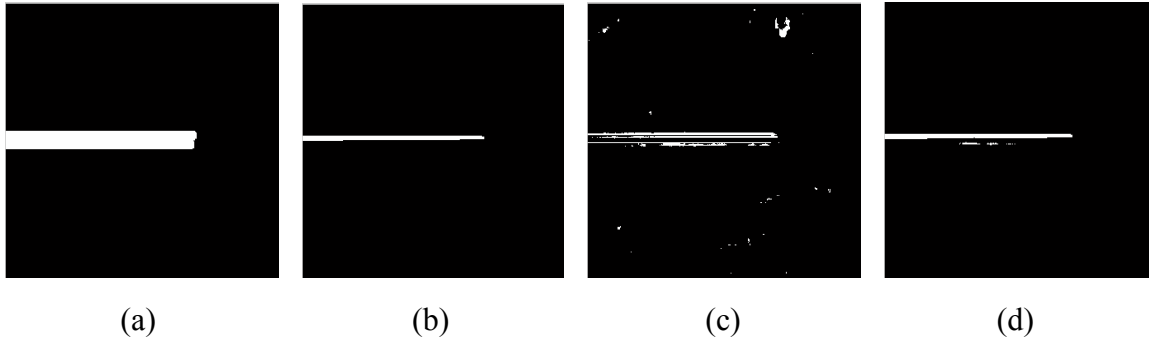


Figure 57: (a) Reference Image, (b) Image difference, (c) CPCA and (d) TPCA change masks.

Table 7: Methods Performances using global threshold selection.

Method	Im-Diff	CPCA	TPCA
Misses	75.8485%	68.8328%	74.3171%
False Alarms	0%	0.2775%	0%
Overall Error	3.5269%	3.4653%	3.4557%

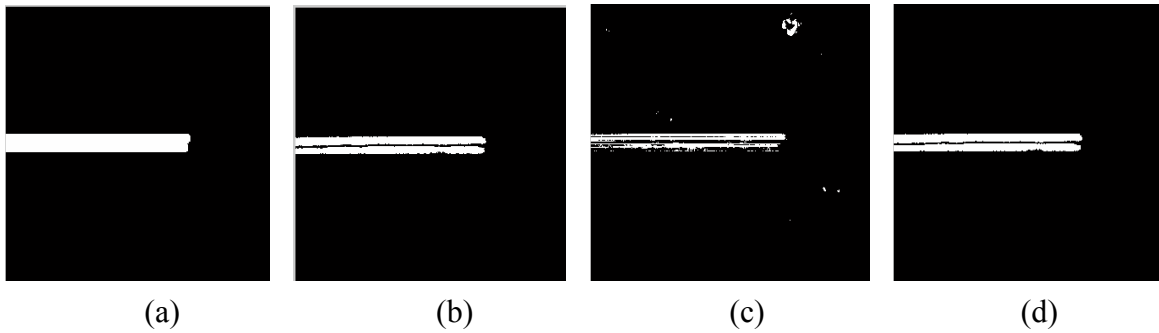


Figure 58: (a) Reference Image, (b) Image difference, (c) CPCA and (d) TPCA change masks.

Table 8: Methods Performances using per-band threshold selection.

Method	Im-Diff	CPCA	TPCA
Misses	19.3295%	48.4478%	17.5290%
False Alarms	0.0040 %	0.2089%	0.0050%
Overall Error	0.9027%	2.4520%	0.8199%

5.4 Application of change detection to Benthic Habitats at Enrique Reef in La Parguera

The 2002 and 2003 HYPERION images of Enrique Reef sensor were used to test the change detection algorithms presented. This section includes the results obtained by each technique.

5.4.1 Image Differencing with Global Threshold Selection

This section presents the results obtained with the Image Differencing method. Figure 59 shows the magnitude of the change vector and its histogram. The mean and the standard deviation obtained for this case were 0.0879 and 0.0747 respectively. The value of the threshold was 0.3121. Figure 60a shows the change mask obtained by applying the global threshold. The change mask over a RGB composite of the image is illustrated in Figure 60b. The changes detected correspond to areas of Mangrove and reef.

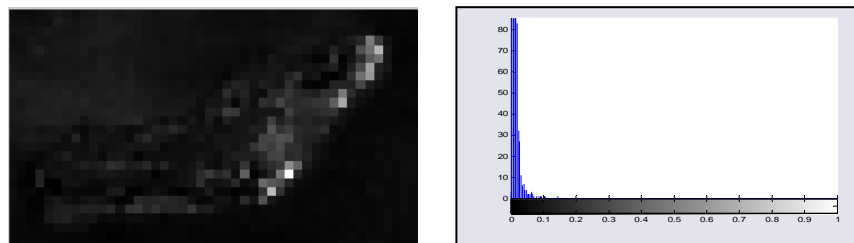


Figure 59: Image Differencing Magnitude of Change.

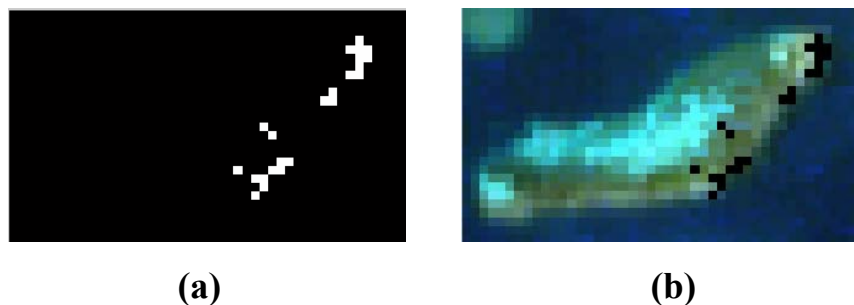


Figure 60: Image Differencing Change Mask with global threshold selection.

5.4.2 Image Differencing with Per-band Threshold Selection

Figure 61 shows the change mask obtained by applying per-band thresholds as well as the change mask over a RGB composite of the image. As can be seen a higher number of changes were detected, compared to Image differencing with global threshold selection. Changes can be observed in the areas of Sea Grass, Reef and Mangrove.

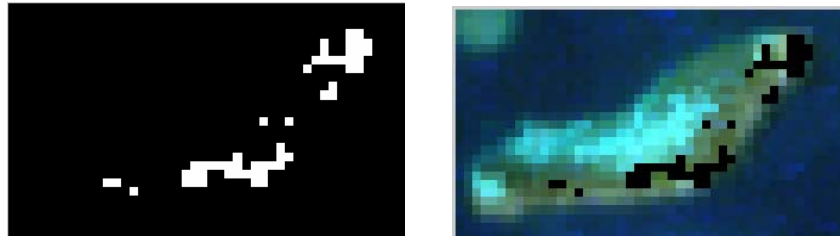
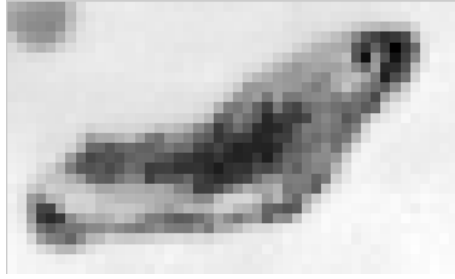


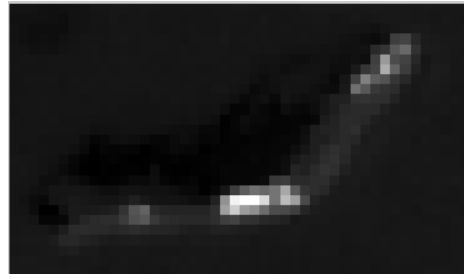
Figure 61: Image Differencing Change Mask with individual threshold selection.

5.4.3 CPCA with Global Threshold Selection

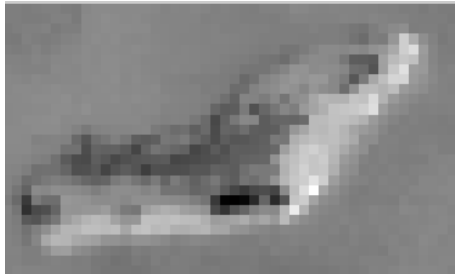
Figure 62, presents the image principal components obtained by performing the analysis over the HYPERION data. Table 9 shows the variances and the percentages of variability obtained from the conventional analysis of principal components. The components selected for change detection were the ones outside a 90% of the total variability. Therefore the image components that should be taken into account for change detection are PC4 to PC50. Nevertheless by observing the principal components from PC9 to PC50 it was clear that they were noisy, thus there was not any substantial information that could be obtained from them.



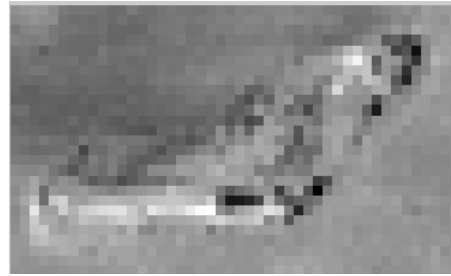
(a) PC1



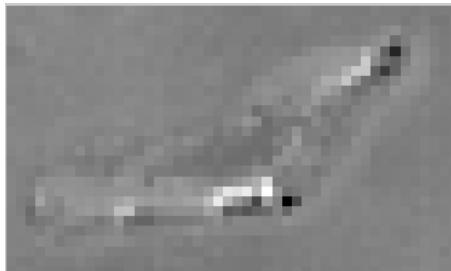
(b) PC2



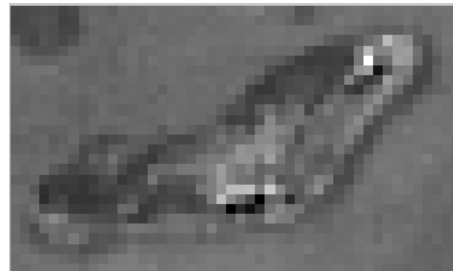
(c) PC3



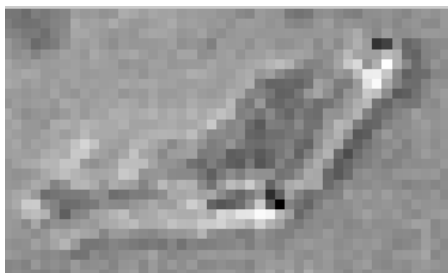
(d) PC4



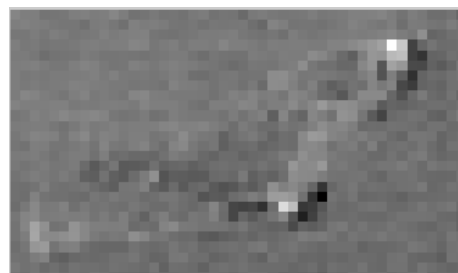
(e) PC5



(f) PC6



(g) PC7



(h) PC8

Figure 62: CPCA components of Enrique Reef.

Table 9: Percentages of Variability and eigenvalues of the principal components.

Component	% of variability	Eigenvalue
1	64.612	1.5148
2	20.173	0.4730
3	6.3838	0.1497
4	2.5968	0.0609
5	1.5583	0.0365
6	0.59832	0.0140
7	0.43891	0.0103
8	0.29206	0.0068

Figure 63 shows the magnitude of change and its histogram obtained from the fusion of the principal components 4 to 8. The mean and standard deviation calculated from the data were 0.0528 and 0.0607. The threshold selected to produce the change mask was 0.2351. Figure 64 shows the change map obtained after threshold application. In this case more changes from the Mangrove area were detected. Just a few pixels from the reef were labeled as change.

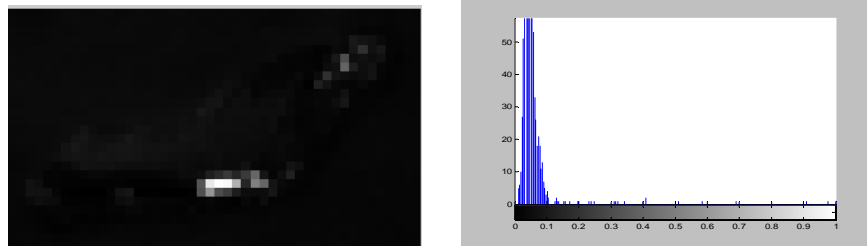


Figure 63: CPCA Magnitude Image PCs 2-8

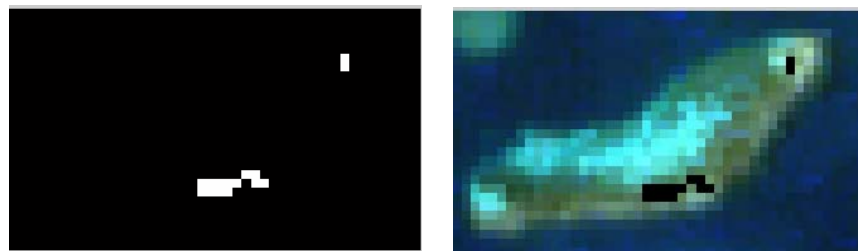


Figure 64: CPCA Change Mask obtained with global threshold selection.

5.4.4 CPCA with Per-band Threshold Selection

Figure 65 presents the change mask obtained from the individual threshold selection. As can be seen a higher number of changes were detected, compared to CPCA with global threshold selection. Changes can be observed in the areas of Sea Grass, Reef and Mangrove.

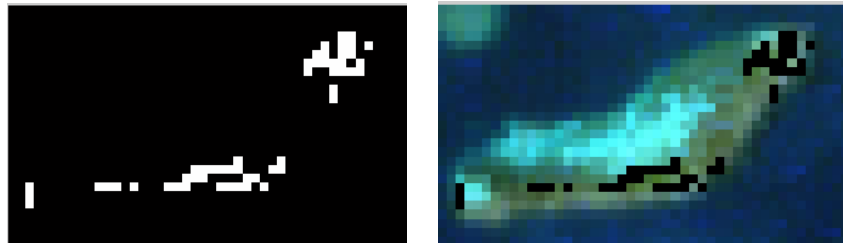


Figure 65: CPCA Change Mask obtained with local threshold selection.

5.4.5 TPCA with Global Threshold Selection

Figure 66 shows the magnitude of change obtained with TPCA. The mean and the standard deviation values were 0.0432 and 0.0802 respectively. The threshold was set to 0.2837. Figure 67 shows the change maps obtained after threshold application. In this case the changes were also from the areas of Mangrove and reef, but it is possible to observe a few pixels from the area of Carbonate sand labeled as change as well.

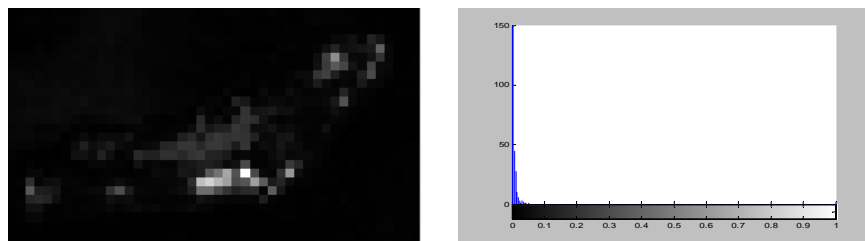


Figure 66: TPCA Magnitude Image

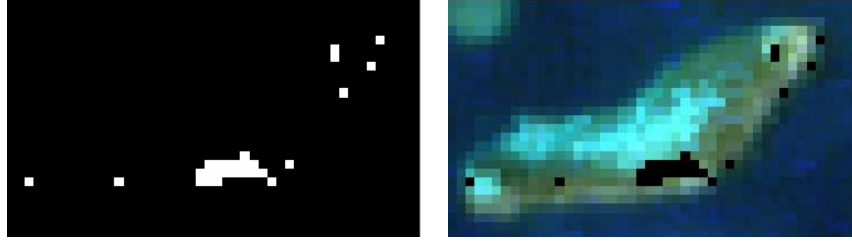


Figure 67: TPCA Change Mask with global threshold selection.

5.4.6 TPCA with Per-band Threshold Selection

Figure 68 presents the change mask obtained from the local thresholds selection using Temporal Principal Component Analysis. The changes highlighted in this case correspond to the area of reef-flat.

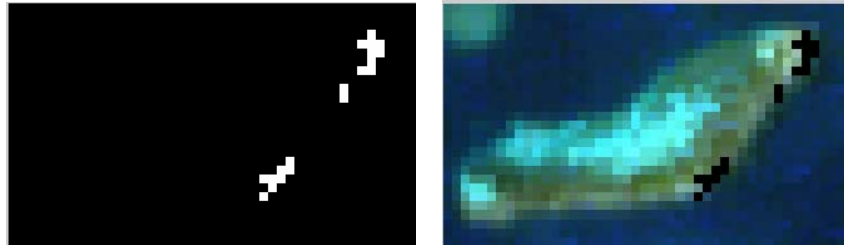


Figure 68: TPCA Change Mask with individual threshold selection.

5.4.7 Post Classification Comparison

The two images from the sensor Hyperion were classified using the Euclidean distance criterion. The supervised classification was performed over each image separately. Then the classification maps were compared to produce the change map. The areas from which samples were taken to train the classifier are presented in Figure 69. Figure 70 presents the classification maps obtained for each image.



Figure 69: Training samples areas.

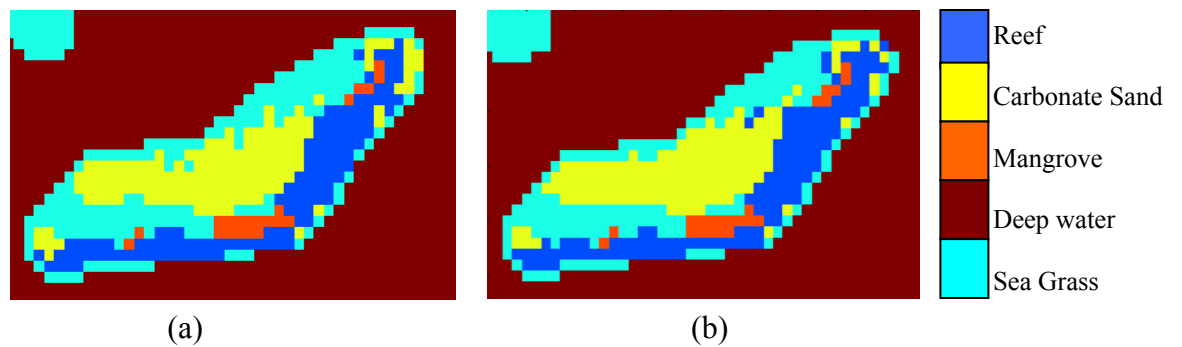


Figure 70: Classification Maps of the Enrique Reef at years (a) 2002 and (b) 2003.

The images were labeled into the following classes:

1. Reef -flat
2. Carbonate Sand
3. Mangrove
4. Deep Water
5. Sea Grass

Figure 71 shows the change mask obtained as a result of the comparison of the two classification maps presented in Figure 70. In this case just a few changes corresponding to the areas of Reef, Carbonate sand and Sea Grass, were detected. However the areas that seem to have suffered the majority of changes were the ones labeled as Reef and Carbonate sand in the image of 2002.

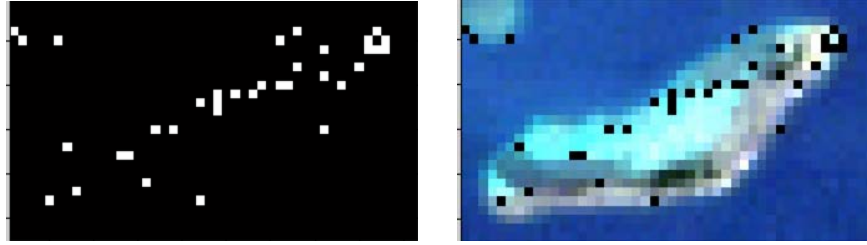


Figure 71: Post Classification Comparison Change Mask

5.5 Application of change detection to the fundus camera images

Two images, corresponding to a human retina were analyzed to detect changes related to diabetes (diabetic retinopathy). These images were taken with a Topcon TRG 501A fundus camera. This section presents the results obtained by each algorithm.

5.5.1 Image Difference with Global Threshold Selection

Figure 72 shows the result obtained with Image differencing with global threshold selection. As can be seen the algorithm was able to detect some changes related to the enlargement of the geographic atrophy as well as the dot hemorrhages.

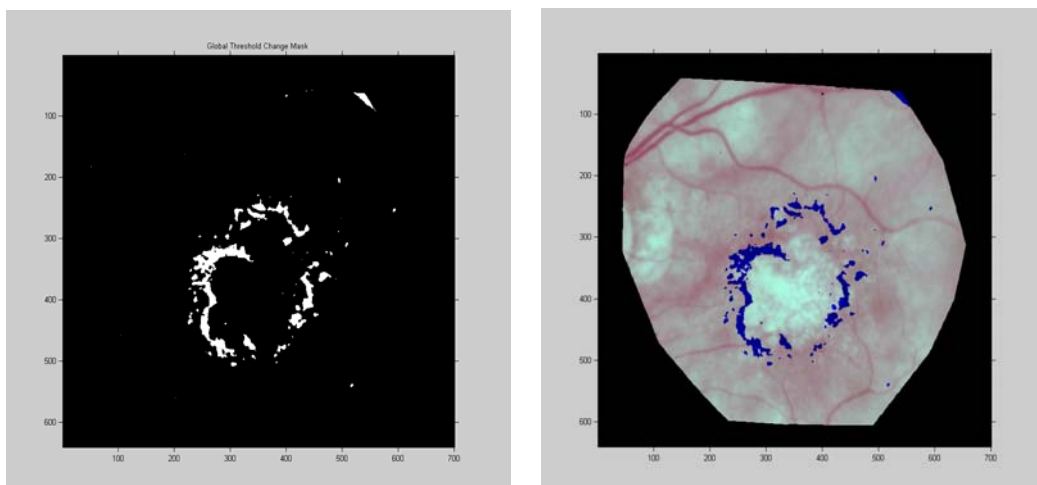


Figure 72: Change Mask obtained by Image differencing with global threshold selection.

5.5.2 Image Difference with Per-band Threshold Selection

Figure 73 shows the result obtained with Image differencing with per-band threshold selection. For this case the algorithm was able to detect more changes related to the enlargement of the geographic atrophy in comparison to the previous method. The dot hemorrhages were detected as well.

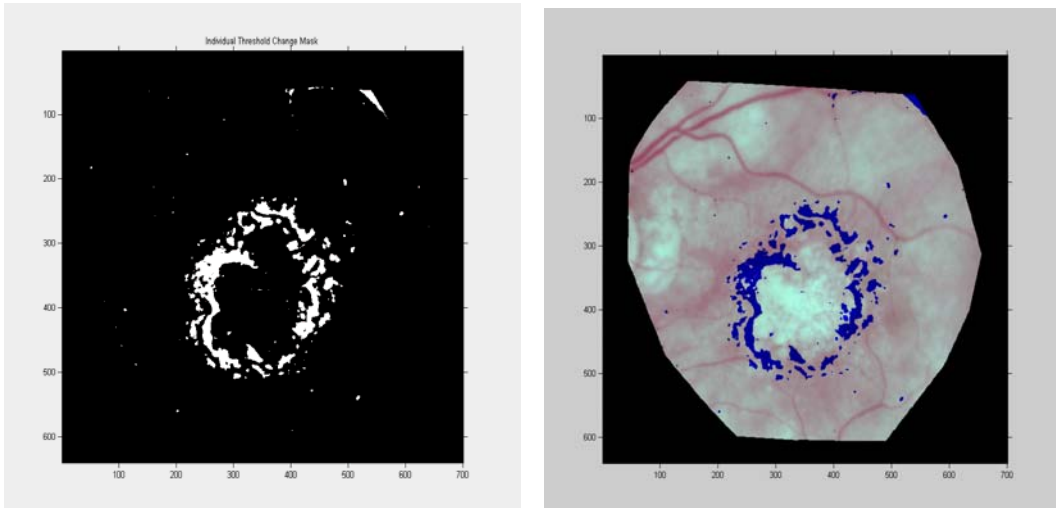


Figure 73: Change Mask obtained by Image differencing with per-band threshold selection.

5.5.3 CPCA with Global Threshold Selection

Figure 74 presents the image principal components obtained by performing the analysis over the fundus camera data. The components selected for change detection were the ones outside a 90% of the total variability. Therefore the image components that should be taken into account for change detection in this case are PC2 to PC6. Nevertheless by observing the principal component PC6 it was clear that it was noisy, thus there was not any substantial information that could be obtained from it.

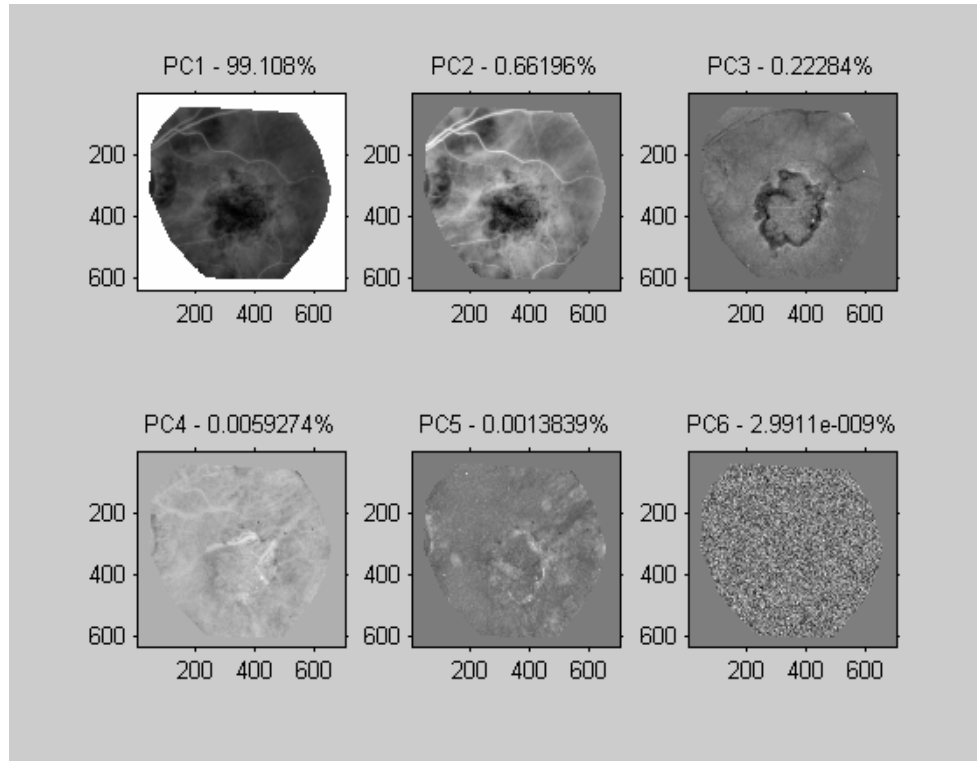


Figure 74: Principal Components and percentages of variability.

Figure 75 presents the change mask generated by the CPCA method with global threshold selection. In this case, the algorithm failed to detect changes related to the enlargement of the geographic atrophy. Instead, it labeled as change areas of the upper blood vessels and the center of the macula.

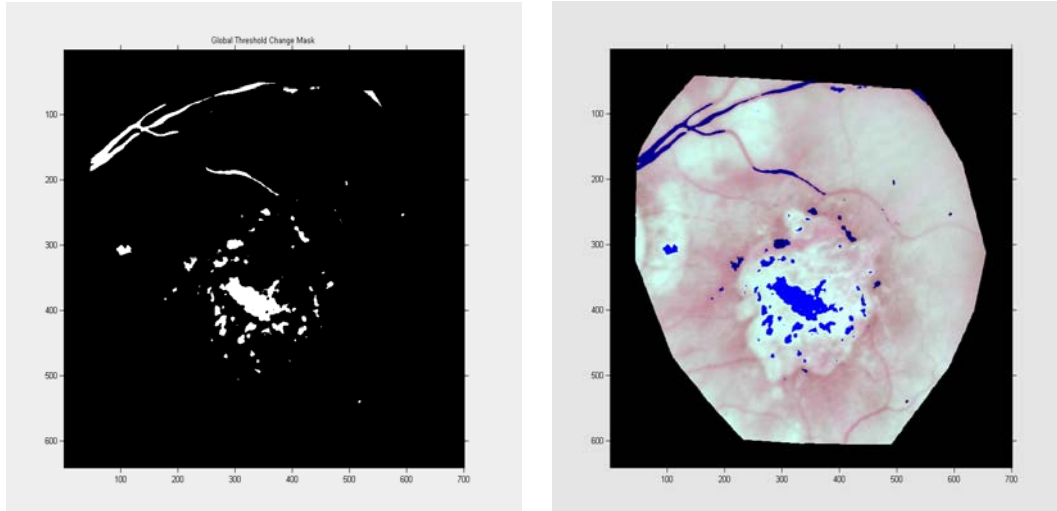


Figure 75: Change Mask obtained by CPCA with global threshold selection

5.5.4 CPCA with Per-band Threshold Selection

Figure 76 presents the change mask generated by the CPCA method with per-band threshold selection. As can be seen the change detection performance of the algorithm improved for this case. Now the false alarms decreased considerably by not identifying the blood vessels and the center of the macula, as change. The changed areas are now related to geographic atrophy.

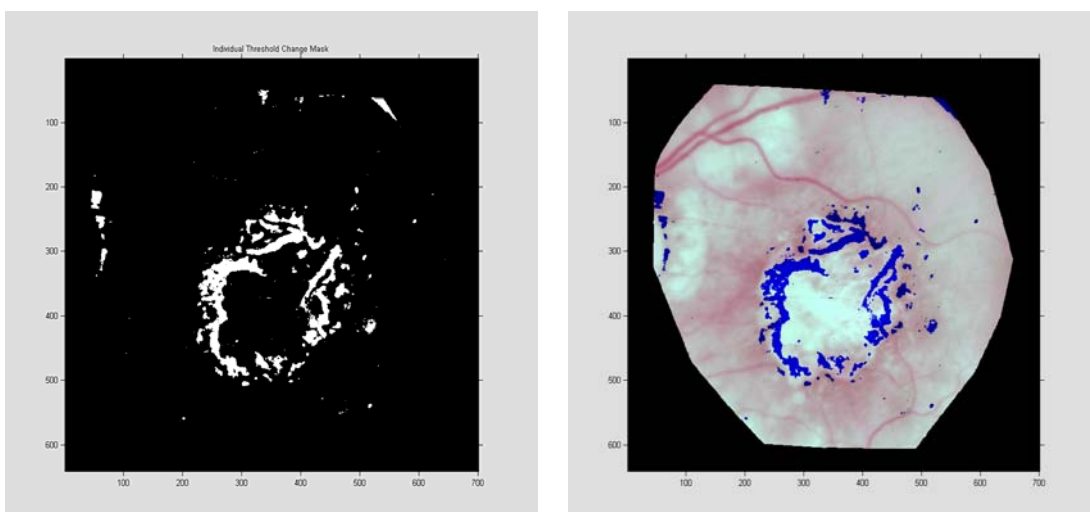


Figure 76: Change Mask obtained by CPCA with per-band threshold selection.

5.5.5 TPCA with Global Threshold Selection

Figure 77 shows the result obtained with Image differencing with global threshold selection. As can be seen the algorithm was able to detect some changes related to the enlargement of the geographic atrophy as well as the dot hemorrhages.

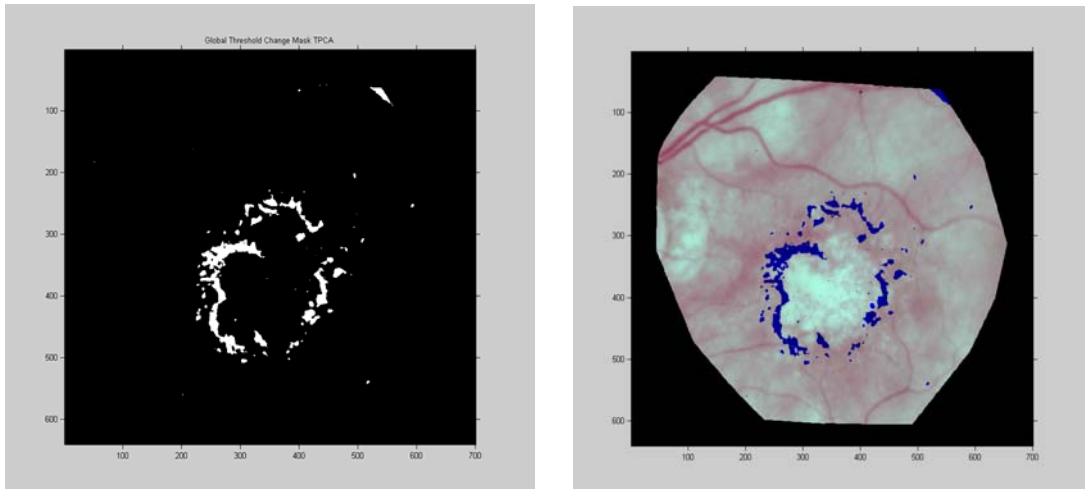


Figure 77: Change Mask obtained by TPCA with global threshold selection.

5.5.6 TPCA with Per-band Threshold Selection

Figure 78 shows the result obtained with TPCA with per-band threshold selection. For this case the algorithm was able to detect more changes related to the enlargement of the geographic atrophy in comparison to the previous method. The dot hemorrhages were detected as well. The results are similar to those obtained with Image differencing.

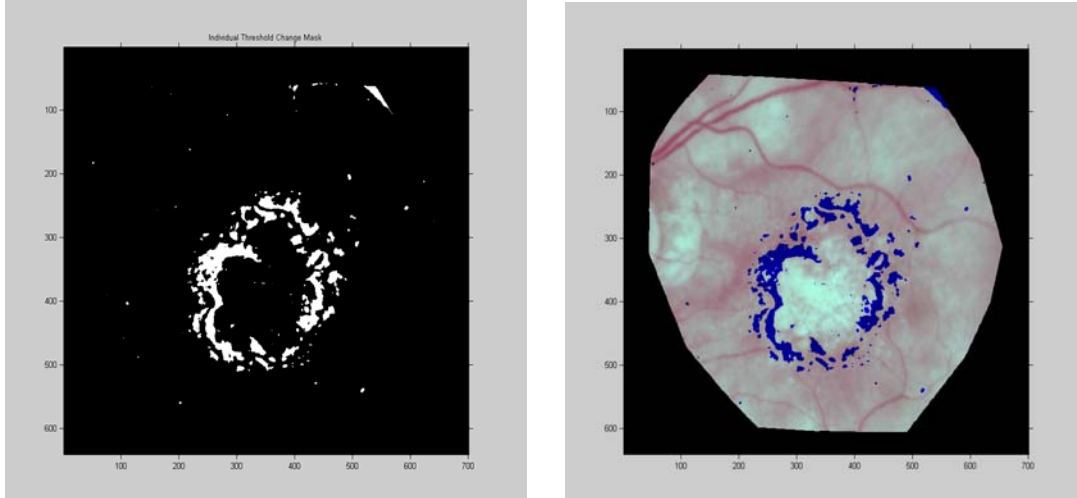


Figure 78: Change Mask obtained by TPCA with per-band threshold selection.

5.6 Summary

In this chapter, the results of the experiments realized as part of this study, were presented. Three change detection techniques were implemented by using hyperspectral data. The performance of each method was measured in terms of missed changes and false alarms. The results obtained from two threshold methods, global and per-band selection, were also included.

As it can be seen in Tables 10 and 11, the Temporal Principal Component Analysis change detection technique decreased the overall errors better than the other two methods, in both global and local thresholds selection.

Table 10: Overall error for the images of coins over concrete sidewalk.

Method	Overall Error % (Global Threshold)	Overall Error % (Local Threshold)
Image Difference	1.0759%	1.0000%
CPCA	0.7498%	0.6233%
TPCA	0.6799%	0.4846%

Table 11: Overall error for the images of a drinking straw over concrete sidewalk.

Method	Overall Error % (Global Threshold)	Overall Error % (Local Threshold)
Image Difference	3.5269%	0.9027%
CPCA	3.4653%	2.4520 %
TPCA	3.4557%	0.8199%

CHAPTER 6

6. Conclusions and Future Work

Three change detection methods have been tested and their performances compared. These methods were: Image Differencing, Conventional Principal Components Analysis (CPCA), and Temporal Principal Component Analysis (TPCA). The latter was modified as a contribution of this work. This chapter presents the conclusions of this work and proposes ideas for future work.

6.1 Conclusions

The experiments performed over the test images shown that the modified Temporal Principal Component Analysis obtained the best results in terms of the amount of missed changes, false alarms and overall errors, with both global and individual threshold selection.

The modified Temporal Principal Component Analysis (TPCA) technique decreased the overall error in change detection by approximately 0.5931% in the case of global threshold selection and 0.5923% for the local threshold selection method. However the overall error was improved by the selection of local thresholds rather than a global one by approximately 1.11%. This is because by calculating a global magnitude image with all the change components, an average of all the information contained by each, was generated. Then just a single threshold is selected to produce a change mask, leading perhaps to errors on the classification of change/no change pixels. The idea of taking a global magnitude came from the synoptic nature hyperspectral images have. This means that each band provides different change information from the same scene, and it was

intended to take advantage of it by placing in a single space or component all the change information encountered in the analysis in order to make it simpler. On the other hand, by using per-band thresholds, different values were applied to produce the binary maps. Since these values are calculated from the statistics of each change information band or component, the final result corresponded to the change information given by each component. A final change mask was then obtained. This final mask is the sum of all the contributions of change information given by each band.

From the images obtained with the sensor HYPERION the most common changes detected belonged to the areas of mangrove and reef. Unfortunately there is no ground truth of the images at both dates, making it difficult to determine which change detection method presented in this study did a best job. On the other hand the results with the fundus camera images, with exception of CPCA with global threshold selection, seemed to detect the most evident changes related to the enlargement of the geographic atrophy and the formation of dot hemorrhages. However it looks like the methods using per-band threshold selection, performed better than the ones with global threshold selection.

Based on our results using test data, it can be concluded that Temporal Principal Components Analysis had a better performance over Image differencing and Conventional Principal Components change detection techniques, by decreasing the overall errors with both global and per-band threshold selection methods.

6.2 Future Work

Test the algorithms with new data. Perform controlled experiments with different types of changes.

Evaluate the effects of different pre-processing methods such as atmospheric or illumination correction and Spectral Resolution Enhancement, in change detection.

Compare multispectral vs hyperspectral change detection.

Combine unmixing of hyperspectral images with change detection and see how the performance of the algorithms is affected.

Improve programming of TPCA and CPCA algorithms before incorporating into the HSI Toolbox.

REFERENCES

Adler-Golden, S., Berk, A., Bernstein, L.S., Richtmeier, S., Acharya, P.K., Matthew, M.W., Anderson, G.P., Allred, C.L., Jeong, L.S., and Chetwynd, J.H. 1998. "FLAASH, A Modtran4 Atmospheric Correction Package for Hyperspectral Data Retrievals and simulations". AVIRIS Earth Science and Applications Workshop Proceedings.

Arzuaga-Cruz, A., Jimenez-Rodriguez, LO., Velez-Reyes, M., Kaeli, D., Rodriguez-Diaz, E Velazquez-Santana, HT., Castrodad-Carrau, A., Santos-Campis, LE., and Santiago, C., 2004. "A MATLAB Toolbox for Hyperspectral Image Analysis". IEEE Internacional Geoscience and Remote Sensing Symposium.

Borrego, R., dos Santos, J.R., Shimabukuro, J.E., Foster, Heinrick, 2001, "A Change vector analysis technique to monitor land-use/land-cover in SW Brazilian Amazon: Acre state", National Institute for Space Research – INPE.

Brunn, A., "Monitoring Mining Induced Plant Alteration and Change Detection in a German Coal Mining area using Airborne Hyperspectral imagery", Technical University of Clausthal, Institute for Geotechnical Engineering and Mine Surveying.

Bruzzone, L., Prieto, D., April 2002, " An Adaptive Semi-parametric and context-based Approach to Unsupervised Change Detection in Multitemporal Remote-Sensing Images", IEEE Transactions in Image Processing, Vol.11, No.4.

CCRS, Canada Center for Remote Sensing, 2004. "Fundamentals of Remote Sensing Tutorial". <http://www.ccrs.nrcan.gc.ca/ccrs/learn/tutorials>

Chang, C.I., 2003. Hyperspectral Imaging, Techniques for Spectral Detection and Classification. Kluwer Academic/Plenum Publishers, New York.

Cho, S.H., 2000 "Digital Change Detection by Post-Classification Comparison of Multitemporal Remotely-Sensed data", Journal of the Korean Society of Remote Sensing, Vol. 16, pp.367-373.

Quillen, D.A., "Common Causes of Vision loss in elderly patients", Published by the American Academy of Family Physicians, Vol. 60, No. 1, July 1999.

Gao, BC., Montes, MJ., Ahmad, Ziauddin and Davis, CO., 2000. Atmospheric correction algorithm for Hyperspectral remote sensing of ocean color from space. Applied Optics, Vol. 39, No. 6.

Gonzalez, R., Woods, R., 2002, "Digital Image Processing", Prentice Hall, 2nd edition.

Hansen, P.C., and Jensen, S.H., 1998. "FIR Filter Representations of Reduced-Rank Noise Reduction". IEEE Trans. On Signal Processing. 46:1737-1741.

Huang, H.H., and Hsiao, C.J., 2000, "Post-Classification and Detection of Simulated Change for Natural Grass", ACRS.

ImSpec LLC, 2002. "ACORNTM User's Guide", ACORN Version 3.12. ImSpec LLC

Insight Segmentation and Registration Toolkit:

<http://www.itk.org/HTML/About.htm>

Jolliffe, I.T., 2002, "Principal Component Analysis", Second Edition, Springer Series in Statistics.

Lillesand, T.M., Kiefer, R.W., 1999, "Remote Sensing and Image Interpretation", Wiley, 4th edition.

Lu, D., Mausel, P., Brondizio, E.S., and Moran, E.F., June 20, 2004, "Change Detection techniques", International Journal of Remote Sensing, Vol. 25, 2365-2407.

Lunetta, R., Elvidge, C., 1999, "Remote Sensing Change Detection: Environmental Monitoring Methods and Applications", Taylor and Francis.

Medline Plus website:

<http://www.nlm.nih.gov/medlineplus/maculardegeneration.html>

Moser, G., Melgani, F., Serpico, S., Caruso, A., 2002, "Partially Supervised Detection of Changes from Remote Sensing Images", IEEE Transactions.

Nielsen, A., Conradsen, K., and Simpson, J., 1998, "Multivariate Alteration Detection and MAF Post Processing in Multispectral, Bitemporal Image Data: New Approaches to Change Detection Studies", Remote Sensing Environment, 64:1-19.

Nikiforof, I., January 1995, "A generalized Change detection problem", IEEE Transactions on Information Theory, Vol. 41, No.1.

Niemeyer, C., 2001, "Object-Oriented Post-Classification of change Images", Forschungszentrum J  lich GmbH, STE, D-52425 J  lich, Germany.

Radke, R., Andra, S., Al-Kohafi, O. and Roysam, B., 2005, "Image Change Detection algorithms: A Systematic Survey". Image Processing, IEEE Transactions, Vol.14, Issue 3, 294-307.

St. Luke's Cataract and Laser Institute:

<http://www.stlukeseye.com/Conditions/DiabeticRetinopathy.asp>

Surface Optics Corporation (SOC), “Hyperspectral Imager and Imaging Spectrometer System”. <http://www.surfaceoptics.com/Products/Hyperspectral.htm>

Tissue, B.M., “Spectroscopy”
<http://www.chem.vt.edu/chem-ed/courses/spec/atomic.html>

Trefethen, L. N., Bau, D, 1997. *Numerical Linear Algebra*. Siam, Philadelphia, PA.

Umaña, A., and Vélez-Reyes, M., 2003. “Determining the dimensionality of hyperspectral imagery for unsupervised band selection”. In Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery IX, Proceedings of SPIE Vol. 5093.

Vélez, M. and Jiménez. L. 1998. “Subset Selection Analysis for the Reduction of Hyperspectral Imagery”. Geoscience and Remote Sensing Symposium Proceedings, 1998. IGARSS '98. IEEE International 3:1577 -1581.

Wehn, N., Goldstein, I., Burke, Bult, T., 1998, “Hyperspectral Mosaicking , Co-Registration and Change detection”, IEEE Transactions.

Wiemker, R., Speck, A., Kulbach, D., Spitzer, H., and Bienlein, J., July 1997, “Unsupervised Robust Change Detection on Multispectral Imagery using spectral and spatial features”, Third International Airborne Remote Sensing Conference and Exhibition, Denmark.

<http://www.geoplace.com/gw/2002/0210/0210hyp.asp>

<http://www.hesc.it/microimages/tutorial/hyprspec-ENG.pdf>

APPENDICES

A. Principal Component Analysis

A.1 Derivation

The vector \mathbf{x} has a known covariance matrix Σ . This is the matrix whose (i,j)th element is the covariance between the ith and jth elements of \mathbf{x} when $i \neq j$, and the variance of the jth element of \mathbf{x} when $i=j$. For $k=1,2, \dots, p$, the kth PC is given by $z_k = \alpha_k' \mathbf{x}$ where α_k is an eigenvector of Σ corresponding to its kth largest eigenvalue λ_k . Furthermore, if α_k is chosen to have unit length ($\alpha_k' \alpha_k = 1$), then $\text{var}(z_k) = \lambda_k$, where $\text{var}(z_k)$ is the variance of z_k .

Considering first $\alpha_1' \mathbf{x}$, the vector α_1 maximizes $\text{var}[\alpha_1' \mathbf{x}] = \alpha_1' \Sigma \alpha_1$. The maximum will not be achieved for finite α_1 ; therefore a normalization constraint must be imposed. The constraint used in the derivation is $\alpha_1' \alpha_1 = 1$, that is the sum of squares of elements of α_1 equals 1. The use of constraints other than $\alpha_1' \alpha_1 = \text{constant}$ in the derivation leads to a more difficult optimization problem, and it will produce a set of derived variables different from the principal components. To maximize $\alpha_1' \Sigma \alpha_1$ subject to the constraint $\alpha_1' \alpha_1 = 1$, the standard approach is to use the technique of Lagrange multipliers. Maximize

$$\alpha_1' \Sigma \alpha_1 - \lambda(\alpha_1' \alpha_1 - 1),$$

where λ is a Lagrange multiplier. Differentiation with respect to α_1 gives

$$\Sigma \alpha_1 - \lambda \alpha_1 = 0,$$

or

$$(\Sigma - \lambda I_p) \alpha_1 = 0,$$

where I_p is the $(p \times p)$ identity matrix. Thus, λ is an eigenvalue of Σ and α_1 is the corresponding eigenvector. To decide which of the p eigenvectors gives $\alpha_1'x$ with maximum variance, note that the quantity to be maximized is

$$\alpha_1' \Sigma \alpha_1 = \alpha_1' \lambda \alpha_1 = \lambda \alpha_1' \alpha_1 = \lambda,$$

so λ must be as large as possible. Thus, α_1 is an eigenvector of Σ , and $\text{var}(\alpha_1'x) = \alpha_1' \Sigma \alpha_1 = \lambda_1$, the largest eigenvalue.

In general, the k th principal component of \mathbf{x} is $\alpha_k'x$ and $\text{var}(\alpha_k'x) = \lambda_k$, where λ_k is the k th largest eigenvalue of Σ , and α_k is the corresponding eigenvector.

The second principal component, $\alpha_2'x$, maximizes $\alpha_2' \Sigma \alpha_2$ subject to being uncorrelated with $\alpha_1'x$, or equivalently subject to $\text{cov}[\alpha_1'x, \alpha_2'x] = 0$, where $\text{cov}(x, y)$ denotes the covariance between the random variables x and y . But

$$\text{cov}[\alpha_1'x, \alpha_2'x] = \alpha_1' \Sigma \alpha_2 = \alpha_2' \Sigma \alpha_1 = \alpha_2' \lambda_1 \alpha_1 = \lambda_1 \alpha_2' \alpha_1 = \lambda_1 \alpha_2' \alpha_1.$$

Thus, any of the equations

$$\alpha_1' \Sigma \alpha_2 = 0, \quad \alpha_2' \Sigma \alpha_1 = 0,$$

$$\alpha_1' \alpha_2 = 0, \quad \alpha_2' \alpha_1 = 0$$

could be used to specify zero correlation between $\alpha_1'x$ and $\alpha_2'x$. By choosing the last of these, and noting that a normalization constraint is again necessary, the quantity to be maximized is

$$\alpha_2'\Sigma\alpha_2 - \lambda(\alpha_2'\alpha_2 - 1) - \phi\alpha_2'\alpha_1,$$

where λ, ϕ are Lagrange multipliers. Differentiation with respect to α_2 gives

$$\Sigma\alpha_2 - \lambda\alpha_2 - \phi\alpha_1 = 0$$

and multiplication of this equation on the left by $\alpha_1'x$ gives

$$\alpha_1'\Sigma\alpha_2 - \lambda\alpha_1'\alpha_2 - \phi\alpha_1'\alpha_1 = 0,$$

which, since the first two terms are zero and $\alpha_1'\alpha_1 = 1$, reduces to $\phi = 0$. Therefore, $\Sigma\alpha_2 - \lambda\alpha_2 = 0$, or equivalently $(\Sigma - \lambda I_p)\alpha_2 = 0$, so λ is once more an eigenvalue of Σ , and α_2 the corresponding eigenvector.

Again, $\lambda = \alpha_2'\Sigma\alpha_2$, so λ is to be as large as possible. Assuming that Σ does not have repeated eigenvalues, λ cannot equal λ_1 . If this happens, then $\alpha_1 = \alpha_2$, violating the constraint $\alpha_1'\alpha_2 = 0$. Hence λ is the second largest eigenvalue of Σ and α_2 is the corresponding eigenvector. It can be shown that for the third, fourth, ..., p th PCs, the vectors of coefficients $\alpha_3, \alpha_4, \dots, \alpha_p$ are the eigenvectors of Σ corresponding to $\lambda_3, \lambda_4, \dots, \lambda_p$, the third and fourth largest, ..., and the smallest eigenvalue, respectively. Furthermore,

$$\text{var}[\alpha_k'x] = \lambda_k \quad \text{for } k = 1, 2, \dots, p.$$

A.2 Mathematical and Statistical Properties

This section presents the mathematical and statistical properties of principal components in both the contexts of population and sample covariance.

Population Principal Components

Notation:

Considering the derivation presented above, and denoting by \mathbf{z} the vector whose k th element is z_k , the k th PC, $k=1, 2, \dots, p$. Then,

$$\mathbf{z} = \mathbf{A}'\mathbf{x}$$

where \mathbf{A} is the orthogonal matrix whose k th column, α_k , is the k th eigenvector of Σ . Thus, the principal components are defined by an orthonormal linear transformation of \mathbf{x} . From the derivation above,

$$\Sigma\mathbf{A} = \mathbf{A}\Lambda, \mathbf{A}'\Sigma\mathbf{A} = \Lambda, \text{ or } \Sigma = \mathbf{A}\Lambda\mathbf{A}'$$

where Λ is the diagonal matrix whose k th diagonal element is λ_k , the k th eigenvalue of Σ , and $\lambda_k = \text{var}(\alpha_k'x) = \text{var}(z_k)$.

The orthonormal linear transformation, $\mathbf{z} = \mathbf{A}'\mathbf{x}$, has a number of properties which are presented below.

Property 1. For any integer q , $1 \leq q \leq p$, consider the orthonormal linear transformation

$$\mathbf{y} = \mathbf{B}'\mathbf{x},$$

where \mathbf{y} is a q -element vector and \mathbf{B}' is a $(q \times p)$ matrix, and let $\Sigma_y = \mathbf{B}'\Sigma\mathbf{B}$ be the variance-covariance matrix for y . Then the trace of Σ_y , denoted $\text{tr}(\Sigma_y)$, is maximized by taking $\mathbf{B}=\mathbf{A}_q$ consists of the first q columns of \mathbf{A} .

Property 2. Considering the orthonormal transformation

$$\mathbf{y} = \mathbf{B}'\mathbf{x},$$

with \mathbf{x} , \mathbf{B} , \mathbf{A} and Σ_y defined as before. Then $\text{tr}(\Sigma_y)$ is minimized by taking $\mathbf{B}=\mathbf{A}_q^*$ where \mathbf{A}_q^* consists of the last q columns of \mathbf{A} .

Property 3. Spectral Decomposition of Σ .

$$\Sigma = \lambda_1\alpha_1\alpha_1' + \lambda_2\alpha_2\alpha_2' + \cdots + \lambda_p\alpha_p\alpha_p'$$

Property 4. Considering the transformation $\mathbf{y} = \mathbf{B}'\mathbf{x}$, if $\det(\Sigma_y)$ denotes the determinant of the covariance matrix \mathbf{y} , then $\det(\Sigma_y)$ is maximized when $\mathbf{B}=\mathbf{A}_q$.

Property 5. Suppose that it is wished to predict each random variable, x_j in \mathbf{x} by a linear function of \mathbf{y} , where $\mathbf{y} = \mathbf{B}'\mathbf{x}$, as before. If σ_j^2 is the residual variance in predicting

x_j from \mathbf{y} , then $\sum_{j=1}^p \sigma_j^2$ is minimized if $\mathbf{B}=\mathbf{A}_q$.

Property 6. For any integer q , $1 \leq q \leq p$, consider the orthonormal linear transformation

$$\mathbf{y} = \mathbf{B}'\mathbf{x},$$

Letting $R_{j,q}^2$ be the squared multiple correlation between x_j and the q variables y_1, y_2, \dots, y_q , defined by the elements of \mathbf{y} . The criterion,

$$\sum_{j=1}^p R_{j:q}^2,$$

is maximized when y_1, y_2, \dots, y_q are the first q correlation matrix principal components. The maximized value of the criterion is equal to the sum of the q largest eigenvalues of the correlation matrix.

Property 7. Consider the family of p -dimensional ellipsoids

$$\mathbf{x}'\Sigma^{-1}\mathbf{x} = \text{constant},$$

The principal components define the principal axes of these ellipsoids.

Property 8. Supposing that $\mathbf{x}_1, \mathbf{x}_2$ are independent random vectors, both having the same probability distribution, and that $\mathbf{x}_1, \mathbf{x}_2$ are both subjected to the same linear transformation

$$\mathbf{y}_i = \mathbf{B}'\mathbf{x}_i, i=1,2$$

If \mathbf{B} is a $(p \times q)$ matrix with orthonormal columns chosen to maximize $E[(\mathbf{y}_1 - \mathbf{y}_2)'(\mathbf{y}_1 - \mathbf{y}_2)]$, then $\mathbf{B} = \mathbf{A}_q$, using the same notation.

Sample Principal Components

Notation:

If there are n independent observations on the p element random vector \mathbf{x} , these n observations are denoted by $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$. \mathbf{S} is defined as the covariance matrix for $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$, and α_k is the corresponding eigenvector for $k = 1, 2, \dots, p$. Then the $(n \times p)$ matrices \mathbf{X} and \mathbf{Z} are defined to have (i,k) th elements equal to the value of the k th element x_{ik} of \mathbf{x}_i , and to z_{ik} , respectively. Then \mathbf{Z} and \mathbf{X} are related by $\mathbf{Z} = \mathbf{XA}$, where \mathbf{A} is the $(p \times p)$

p) orthogonal matrix whose k th columns is α_k . If the mean of each element of \mathbf{x} is known to be zero, then $\mathbf{S} = \frac{1}{n} \mathbf{X}' \mathbf{X}$. If the mean of \mathbf{x} is unknown \mathbf{S} is

$$\frac{1}{n-1} \sum_{i=1}^n (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)$$

where

$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij}, \quad j=1, 2, \dots, p.$$

Therefore the matrix \mathbf{S} can be written as follows,

$$\mathbf{S} = \frac{1}{n-1} \mathbf{X}' \mathbf{X},$$

where \mathbf{X} is an $(n \times p)$ matrix with (i,j) th element $(x_{ij} - \bar{x}_j)$. The notation x_{ij} is used to denote the (i,j) th element of \mathbf{X} , so that x_{ij} is the value of the j th variable measured about its mean \bar{x}_j for the i th observation. Then the matrix of principal component scores is denoted by,

$$\mathbf{Z} = \mathbf{XA}$$

Properties 1, 2, 4, 5

Defining the linear transformation as

$$\mathbf{y}_i = \mathbf{B}'\mathbf{x}_i, \quad \text{for } i=1, 2, \dots, n$$

where \mathbf{B} , as in Properties 1, 2, 4, 5, (of population pc) is a $(p \times q)$ matrix whose columns are orthonormal. Properties 1, 2, 4, 5, still hold but with the sample covariance matrix of the observations \mathbf{y}_i , $i=1, 2, \dots, n$, replacing Σ_y , and with the matrix \mathbf{A} now defined as having k th column \mathbf{a}_k , with $\mathbf{A}_q, \mathbf{A}_q^*$, representing its first and last q columns.

Property 3. *The spectral decomposition property is given by,*

$$\mathbf{S} = l_1 \mathbf{a}_1 \mathbf{a}_1' + l_2 \mathbf{a}_2 \mathbf{a}_2' + \cdots + l_p \mathbf{a}_p \mathbf{a}_p'$$

Property 6. *Supposing now that X consists of n observations on p predictor variable x measured about their sample means, and that the corresponding regression equation is*

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

where \mathbf{y} is the vector of n observations on the dependent variable, again measured about the sample mean. Supposing that X is transformed but h equation $\mathbf{Z} = \mathbf{X}\mathbf{B}$, where \mathbf{B} is a $(p \times p)$ orthogonal matrix. The regression equation can then be rewritten as

$$\mathbf{y} = \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\epsilon},$$

where $\boldsymbol{\gamma} = \mathbf{B}^{-1} \boldsymbol{\beta}$. The usual least squares estimator for $\boldsymbol{\gamma}$ is $\hat{\boldsymbol{\gamma}} = (\mathbf{Z}'\mathbf{Z})^{-1} \mathbf{Z}'\mathbf{y}$. Then the elements of $\hat{\boldsymbol{\gamma}}$ have successively, the smallest possible variances if $\mathbf{B} = \mathbf{A}$, the matrix whose k th column is the k th eigenvector of $\mathbf{X}'\mathbf{X}$, and hence the k th eigenvector of \mathbf{S} . Thus \mathbf{Z} consists of values of the sample principal components for \mathbf{x} .

Property 7. *Suppose that the observations $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ are transformed by $\mathbf{y}_i = \mathbf{B}'\mathbf{x}_i$, $i = 1, 2, \dots, n$, where \mathbf{B} is a $(p \times q)$ matrix with orthonormal columns, so that $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$ are projections of $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ onto a q dimensional subspace. A measure of goodness-of-fit of this q -dimensional subspace to $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ can be defined as the sum of squared perpendicular distances of $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ from the subspace. This measure is minimized when $\mathbf{B} = \mathbf{A}_q$.*

B. Singular Value Decomposition Noise Filtering

B.1 Signal Model

Let us model the Hyperspectral pixel as a combination of a random spectral signature, \mathbf{s} , (pure or mixed) with additive noise, \mathbf{n} , given by:

$$\mathbf{x} = \mathbf{s} + \mathbf{n} .$$

where \mathbf{s} is the spectral signature with covariance Σ_s , \mathbf{n} is white noise with covariance $\mu^2 I$ and \mathbf{s} and \mathbf{n} are uncorrelated. The $n \times n$ covariance matrix for \mathbf{x} , Σ_x , given by

$$\Sigma_x = \Sigma_s + \mu^2 I$$

where Σ_s is of rank $p < n$ which is equal to the dimension of the signal subspace. The eigenvalues of Σ_x , λ_{xi} , are given by:

$$\lambda_{xi} = \lambda_{si} + \mu^2, i = 1, 2, \dots, n$$

where λ_{si} is the i -th eigenvalue of Σ_s and the eigenvalues are ordered in descending order of magnitude. Since Σ_s is rank p , $\lambda_{si} = 0$ for $i > p$, so that:

$$\lambda_{xi} = \lambda_{si} + \mu^2, i = 1, 2, \dots, p \text{ (signal and noise)}$$

and

$$\lambda_{xi} = \mu^2, i = p+1, \dots, n \text{ (only noise contribution)}$$

If we have a good signal to noise ratio, we should expect that $\lambda_{si} \gg \mu^2$ for all i .

Of the measured signal \mathbf{x} , only the first p principal components carry signal information and a filtered \mathbf{x} is constructed using the first p eigenvectors of the Σ_x ,

$$\mathbf{x}_p = \mathbf{V}_p \mathbf{V}_p^T \mathbf{x}$$

Where \mathbf{x}_p is the filtering version of \mathbf{x} , \mathbf{V}_p is the matrix of the first p eigenvectors of the covariance matrix. Notice that \mathbf{x}_p is a projection of \mathbf{x} onto the space spawned by the signal eigenvectors. This reconstruction is optimal in a mean squared sense; and it is used in many multivariable statistical applications.

B.2 *Computing the filtered image*

Let $\mathbf{A} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$ (pixels of the Hyperspectral cube) be a set of independent measurement of the random variable \mathbf{x} (mean subtracted).

The singular Value Decomposition (SVD) of a matrix \mathbf{A} of size $m \times n$ is a decomposition of the form (Trefethen et al, 1997):

$$\mathbf{A} = \mathbf{U} \mathbf{S} \mathbf{V}^T$$

where \mathbf{U} and \mathbf{V} are orthogonal, square matrices of sizes $m \times m$ and $n \times n$ respectively, and \mathbf{S} is a diagonal matrix of size $m \times n$. The columns u_i ($1 \leq i \leq m$) and v_j ($1 \leq j \leq n$) of \mathbf{U} and \mathbf{V} are called the left and right singular vectors, respectively, and the diagonal elements s_{kk} ($1 \leq k \leq \min\{m, n\}$) of \mathbf{S} are called the singular values of \mathbf{A} . The left (u_i) and right (v_j) singular vectors of \mathbf{A} form orthonormal bases, and are related by the following relation:

$$\mathbf{A} v_i = s_i u_i$$

showing that each right singular vector is mapped onto the corresponding left singular vector, and the “magnification factor” is the corresponding singular value.

Since $\Sigma_x = \frac{A^t A}{N}$, we get that $\lambda_{xi} = \frac{\sigma_i^2}{N}$, where σ_i is the i -th singular value of A . So we can use the SVD of A to compute the filtered mode. The desired approximation is the rank p approximation to A . Given by $A_p = U_p \Sigma_p V_p^T$ where U_p is the matrix of the first p left singular vectors, Σ_p is a diagonal matrix with the first p singular values, and V_p is the matrix of the first p right vectors. Notice that the right singular vectors of A are equal to the eigenvalues of Σ_x .

SVD has a variety of applications in scientific computing, signal processing, automatic control, and many other areas. In signal processing, the SVD decomposition, followed by the truncation of the lower singular values has been traditionally used for noise reduction, in voice and image signals (Hansen et al, 98). The central idea can be summarized as follows: let A be the matrix that represents the discretized noisy signal, compute the SVD decomposition of A , and then discard the smaller singular values of A and their corresponding singular vectors. It can be shown that the small singular values of A , mainly represent the noise, and thus the rank- p matrix A_p ($p < n$) represents a filtered signal with less noise (A key problem is to determine the dimension of the signal space p). This application of the SVD is called Truncated SVD (TSVD) or Reduced-Rank Noise Reduction. (Hansen et al.1998).

C. Matlab Codes:

```
Image differencing code

pixels1=reshape(imagen1,640*640,50);
X=pixels1';
Im_T1=X';
clear X

pixels2=reshape(imagen2,640*640,50);
Y=pixels2';
Im_T2=Y';
clear Y

clear pixels1 pixels2

%Constructing Difference Image
```

```

Im_diff=abs(Im_T1-Im_T2);
a=reshape(Im_diff,640,640,49);
b=imrotate(a,-90);
c=flipdim(b,2);%Final Image of difference.

```

Conventional Principal Component Analysis Code

```

X=pixels;
Im_T1=X';
clear X

Y=pixels;
Im_T2=Y';
clear Y
clear pixels

Imagen=[Im_T1 Im_T2];
clear X Y

[pc, eigval, porcentages]=pcacov(cov(Imagen));

T= Imagen*pc; %Transformed Image

clear Imagen

```

Temporal Principal Component Analysis Code

```

pixels1=reshape(imagen1,640*640,49);
X=pixels1';
Im_T1=X';
clear X

pixels2=reshape(imagen2,640*640,49);
Y=pixels2';
Im_T2=Y';
clear Y
clear pixels1 pixels2

% Temporal Principal Component Analysis:

% 1. Reshaping Original data

[m,n]=size(Im_T1(:,1:50));
T1=reshape(Im_T1(:,1:50),m*n,1);
T2=reshape(Im_T2(:,1:50),m*n,1);

clear Im_T1 Im_T2

Im_Mat=[T1 T2];
clear T1 T2

[pc,EigVal,Percentages]=pcacov(cov(Im_Mat));

```



```

%Rotating Data to the PC's

Rot=Im_Mat*pc;
clear Im_Mat
A1=Rot(:,1);
A2=Rot(:,2);

clear Rot

A1r=reshape(A1,640,640,49);
A2r=reshape(A2,640,640,49); % Change Information (Second Component)

```