DIMENSIONAL REDUCTION OF HYPERSPECTRAL IMAGE DATA USING BAND CLUSTERING AND SELECTION

Researcher:
Muhammad Sohaib
Reg No. 91-FET/MSEE/F07

Supervisor:
Dr. Ihsan ul Haq

Department of Electronic Engineering

Faculty of Engineering and Technology

INTERNATIONAL ISLAMIC UNIVERSITY, ISLAMABAD

1- Remote sensing

2. Image processing - Digital techniques

Dimensional Reduction of Hyperspectral Image Data Using Band Clustering and Selection

By

Muhammad Sohaib
Reg. No. 91-FET/MSEE/F07

A thesis submitted in partial fulfillment of the requirements for the Degree of Master of Science in Electronic Engineering with specialization in image and signal processing at the Faculty of Engineering and Technology
International Islamic University Islamabad.

Supervisor
Dr. Ihsan ul Haq
Assistant Professor
Faculty of Engineering and Technology
International Islamic University
Islamabad .

(August, 2010)

بسم الله الرحمن الرحيم

In the Name of

Allah

The Most Gracious and Merciful

Dedicated to My Loving Parents, respected teachers and dear friends

(Acceptance by the Viva Voice Committee)

Title of Thesis: Dimensional Reduction of Hyperspectral Image Data Using Band Clustering and Selection

Name of Student: Muhammad Sohaib

Registration No: 91-FET/MSEE/F07

Accepted by the Faculty of Engineering and Technology, INTERNATIONAL ISLAMIC UNIVERSITY, ISLAMABAD in partial fulfillment of the requirements for the Master of science degree in Electronic Engineering.

Viva Voce Committee

Supervisor

External Examiner

Intrenal Examinar

Date: August 31, **2010**

Declaration

I certify that except where due to acknowledgments has been made, the work is that of my alone: the work has not been submitted previously, in whole or in the part, to qualify for any other academic award; the content of the thesis is the result of work has been carried out since the official commencement due to approved research program; and any editorial work, paid or unpaid, carried out by a third party is acknowledged.

Muhammad Sohaib

91-FET/ MSEE / F07

ACKNOWLEDGEMENTS

I would first of all pay my thanks to Almighty Allah for His providential guidance, analytical wisdom and vigour to put my best possible effort towards the accomplishment of this thesis.

I express my gratitude to my venerable supervisor **Mr. Ihsan ul Haq** for his vital support and constant encouragement towards the completion of my thesis.

I also express my gratitude to all of my teachers for their kind contribution in my knowledge and experties. I am also indebted to my friend **Mr. Qaisar Mushtaq** for his unconditional support.

I am also thankful to all members of MS/PhD Committee for their kind guidance to ensure the quality of work in my dissertation.

I also express my gratitude to a very kind person **Mr. Ahmad Shuja** (Dean) for his unforgetable support during my stay in this institution.

Muhammad Sohaib

FORWARDING SHEET

The thesis entitled "Dimensional Reduction Using Using Band Clustering and Selection in Hyperspectral Images" submitted by Mr.Muhammad Sohaib in partial fulfillment of Master degree in Electronic Engineering, has been completed under my guidance and supervision. I am satisfied with the quality of student's research work and allow him to submit this thesis for further process of as per IIU rules & regulations.

Date: August 31, 2010.

Supervisor Signature: ____________

Name : DR. IHSAN UL HAQ
Assistant Professor
Department of Electronic Engi
Faculty of Engineering & Te
International Islamic Uni
Islamabad, Paki

ABSTRACT

In this research work an approach for the dimensional reduction of the hyperspectral image data is introduced. For computational and data compression reasons, it is desired to reduce the dimensional of the data set while maintaining good performance in image analysis tasks. Dimensional reduction is a good choice to overcome the challenges of huge data storage, computational load, communication bandwidth and convergence instability in the analysis of hyperspectral image data. The reduction of dimensionality is necessary for high accuracy in unmixing of the pixels, classification and detection.

In our research work the bands are clustered and then selected based on statistical measures of band images. The spread hyperspectral image data is measured in each band and the calculated bands are clustered using the K-means clustering technique. The K-means clustering of bands is performed in such a way that the intra-cluster variance is kept minimum and the inter-cluster variance maximum. The optimal number of band selection is done using the concept of Virtual Dimensionality (VD). The endmember or targets are extracted through Vertex Component Analysis (VCA). The experimental results are compared with other unsupervised band selection techniques to show the effectiveness of the proposed technique.

TABLE OF CONTENTS

Abstract i

Table of Contents ii

List of Tables iv

List of Figures v

1. INTRODUCTION 1

1.1 Introduction to Hyperspectral Imagery................................ 1

1.2 Applications.. 4

1.3 Challenges in Hyperspectral image analysis........................... 5

1.4 Problem Statement 6

1.5 Research objective 7

2. HYPERSPECTRAL IMAGE ANALYSIS.................................. 8

2.1 Introduction... 8

2.2 Dimensional Reduction...... 9

2.2.1 Feature Extraction.. 11

2.2.1.1 Principal Component Analysis 12

2.2.1.2 Independent Component Analysis................................. 15

2.2.1.3 Minimum Noise Fraction (MNF) Transform........... 17

2.2.2 Feature Selection.. 18

2.3 Band Selection... 19

2.3.1 Search-based Methods.. 21

2.3.2 Transform-based Methods... 21

2.3.3 ICA-based Band Selection 21

2.3.4 Information-based Band Selection..22

2.3.4.1 Constrained Band Selection (CBS)........................ 25

2.3.4.2 Maximum Variance Principal Component Analysis..25

2.4 Unmixing of Hyperspectral Data 26

2.4.1 N-FINDER Method..28

2.4.2 Independent Component Analysis..................................... 29

2.4.3 Vertex Component Analysis (VCA)................................. 30

3. PROPOSED METHOD OF DIMENSIONAL REDUCTION......... 31

3.1 Band Clustering and Selection.. 31

3.1.1 K-MEANS Clustering..32

3.2 Proposed Algorithm 34

4. EXPERIMENTAL RESULTS AND DISCUSSION 36

4.1 Hyperspectral Image Data.. 36

4.5 Results and Discussion 38

5. CONCLUSION AND FUTURE WROK.. 44

6. REFERENCES 46

7. RESEARCH CONTRIBUTION.. 56

LIST OF TABLES

Table 4.1 Bands Selection Using Different Techniques

Table 4.2 Spectral Similarity Measurement (Sam) Among Found Endmembers And Ground Truth Endmembers

LIST OF FIGURES

Figure 1: Hyperspectral Image Cube

Figure 2: Correlation Among The Band Image Vectors

Figure 3: Hyperspectral Image Analysis

Figure 4: K-Means Clustering Flow Chart

Figure 5: Flow Chart of the Proposed Algorithm

Figure 6: Spatial Positions Of Four Pure Pixels Which Shows Correspondence To The Minerals: Alunite (A), Buddingtonite (B), Calcite (C), And Kaolinite

Figure 7: USGS spectral signatures of Alunite (A), Buddingtonite (B), Calcite (C), and Kaolinite (K)

Figure 8: Extraction of Endmembers by VCA from Full bands and the selected bands given in Table 4.1

Chapter 1

INTRODUCTION

1.1 Introduction to Hyperspectral Imagery

Hyperspectral sensors- used for hyperspectral imagery collect information as a set of images represented by different bands. This information makes it possible to extract a continuous spectrum for each image cell and unveils materials that cannot be resolved by multispectral sensors [1]. The concept of hyperspectral imagery developed in 1980's, when A.F.H Goetz and his colleagues at NASA's Jet propulsion Laboratory began a revolution in remote sensing. They developed new instruments such as AVIRIS (Airborne Visible Infra Red Imaging Spectrometer) [2]. AVIRIS was able to cover the wavelength region from .4 to 2.5 μm using more than two hundred spectral channels, at nominal spectral resolution of 10 nm. A remotely sensed image is an image in a cubic form with the third dimension specified by spectral wavelengths. The collected image data by hyperspectral remote sensors is simultaneously in hundreds of narrow, adjacent spectral bands over the wavelengths that can range from the near ultraviolet through the thermal infrared at 5nm of fine resolutions. Hyperspectral image cube is shown in fig.1 which shows that every image pixel is a column vector of which each component is denoted by a specific spectral band. Each pixel contains a hyperspectral signature that represents different materials. Remote sensed vision systems for surveillance, object recognition, target identification, estimation of water sedimentation and the creation of maps are of great interest. The opportunity for more detail image analysis can be provided by hyperspectral imagery. As a result of high spectral resolution, hyperspectral systems produce a massive amount of data. These measurements make it possible to derive a continuous spectrum for an image data. Sensors are adjusted and atmospheric and terrain effects are applied, these image spectra can be compared with field or laboratory reflectance spectra to recognize and map surface materials such as particular types of vegetation or diagnostic minerals associated with ore deposits. Hyperspectral data helps the analyst in detection of more materials, objects and regions with enhanced accuracy than previously possible. Hyperspectral images provide a vast amount of information about a scene, but most of that information is redundant as the bands are highly correlated as shown in fig. 2. Due to the reasons of computation and data compression, it is desired to reduce the dimensionality of the data set while maintaining and having no effect on the good performance in image analysis tasks.

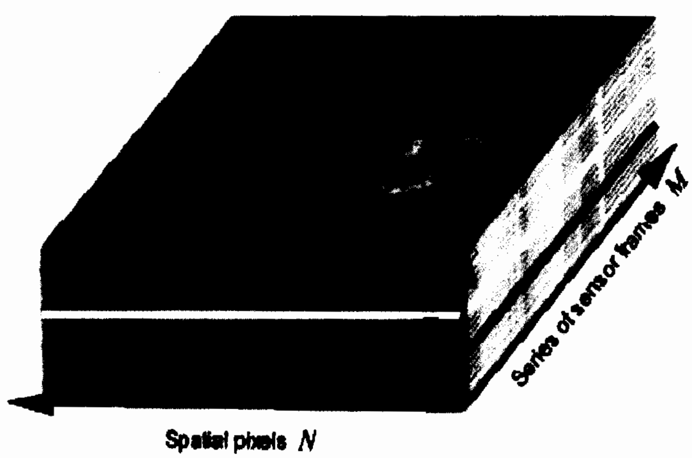


Figure 1: Hyperspectral image cube

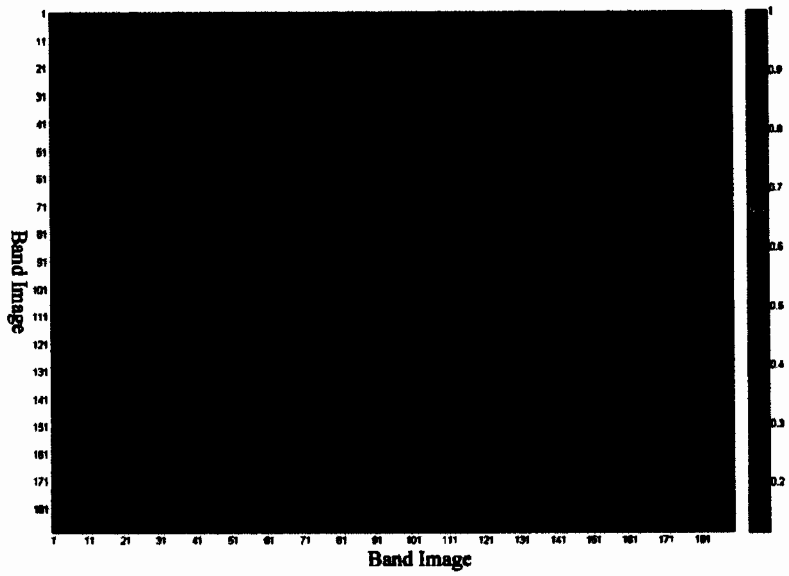


Figure 2: Correlation among the band image vectors

The emergence of image data sets with large numbers of spectral bands has presented image processing and interpretation challenges not experienced with existing multispectral data [3] and hyperspectral data. New techniques for pixel labeling and /or modification of the existing methods may have modification for fully interpreting the information given by hyperspectral image data. More information about materials is provided by hyperspectral imagery than multispectral imagery. As compare to multispectral sensing, hyperspectral sensing can increase the delectability of pixel and sub pixel size targets by exploiting finer detail in the spectral signatures of targets and natural backgrounds.

Hundreds of bands in hyperspectral image data imply high dimensional data which presents several significant challenges to image classification. The performance of many supervised classification methods are strongly affected by the dimension of input space [4]. There is likely to be redundancy between bands and some bands may have less discriminatory information than others. As a result, the imposition of requirements for storage space, computational load and communication bandwidth are against the real time applications.

1.2 Applications

Hyperspectral remote sensing is widely used in real life applications. Mainly used for geology and mining purpose as hyperspectral imaging identifies various minerals for mining and oil industries. It can be used to search for ore and oil [3, 4]. Hyperspectral surveillance is the implementation of hyperspectral scanning technology for surveillance purposes. Hyperspectral imaging is particularly useful in military surveillance because of measures that military entities now take to avoid airborne surveillance. The idea that drives hyperspectral surveillance is that hyperspectral scanning draws information from such a large portion of the light spectrum that any given object should have unique spectral signature in at least a few of the many bands that get scanned. Hyperspectral sensing of minerals is now well developed. Many minerals can be identified from images, and their relation to the presence of valuable minerals such as gold and diamonds is well understood. Currently the move is towards understanding the relation between oil and gas leakages from pipelines and natural wells. In short Hyperspectral images can be used for geology, forestry and agriculture mapping, land cover analysis, and atmospheric analysis, law enforcement, military and defense.

1.3 Challenges in hyperspectral image analysis

Hyperspectral images provide a vast amount of information about a scene. However, much of that information is redundant as the bands are highly correlated. For computational and data compression reasons, it is desired to reduce the dimensional of the data set while maintaining good performance in image analysis tasks. Where more information is carried by hundreds of bands of hyperspectral image data, there are some challenges in analysis of hyperspectral image data. First of all is huge data volume, so there are data storage and transmission problems. Second challenge is redundancy. Information of all the bands is not uncorrelated but some information is shared among two or more than two bands.

Redundancy in data can cause convergence instability. Third challenge is remarkable high processing time either using supervised or unsupervised classification techniques. Fourth one, Hughes phenomenon is observed in hyperspectral image data classification because of limited training data and ratio of the training pixels to the number of band is small. Endmember detection, unmixing and classification accuracy would not always increase with increase of feature used. This is attributed to the fact that more training samples are required to specify the decision boundary for classification for hyperspectral data. It is a very complicated task to analyze information in hyperspectral imagery data for discovering underlying structures. It is difficult to visualize or to classify such a huge amount of data. Analytical technique may be categorized into photo interpretation and machine analysis (classification) for multispectral image data. The former depends upon the use of image enhancement procedures for improving the visual interpretability of image data whereas the latter is based usually on statistical or other forms of numerical algorithms for labeling individual pixels. These traditional analytical techniques face problem with hyperspectral image data. These difficulties are due to enormous data volume, redundancy, the need of calibration and very high dimension of hyperspectral image data.

1.4 Problem Statement

The emergence of image data sets with large numbers of spectral bands has presented image processing and interpretation challenges not experienced with existing multispectral data [108] and hyperspectral data. The curse of dimensionality has been known for more than three decades. There is a need for the development of algorithms for detection, unmixing and classification that utilize the amount of information and separability that hyperspectral image data offers while simultaneously avoiding the difficulties inherent in hyperspectral space. No doubt more information is buried in hundreds of narrow and adjacent bands of hyperspectral data but some information is overlapping among them. So to get rid of this redundancy there are two methods: feature extraction and feature selection [8].

Our concern and focus is enhancing the results for feature extraction through band clustering and selection to reduce the dimensionality of hyperspectral image data and to improve the endmember detections and unmixing accuracy.

1.5 Research Objective

Cost and complexity are the main disadvantages of hyperspectral imagery. Fast processors, sensitive detectors and huge data storage are the requirements of hyperspectral image data analysis. Also one of the hurdles researchers have had is the cost on the transmission of such a huge data.

One way to overcome these difficulties is the reduction of dimensionality of the hyperspectral data. The objective of our research thesis is to reduce the dimensionality of hyperspectral data, reduce the processing time and to improve the detection and classification accuracy. Our concern and focus is enhancing the results for feature extraction through band clustering and selection to reduce the dimensionality of hyperspectral image data and to improve the endmember detections and unmixing accuracy. The reduced dimensional data would be analyzed for unmixing and detection of targets /endmembers.

Chapter 2

Hyperspectral Image Analysis

2.1 Introduction

Hyperspectral image analysis has several steps shown in fig. 3.The first step is the atmospheric correction. The comparison and analysis of the radiance acquired by hyperspectral sensors with a digital spectral library or even with other radiance data sets cannot be done due to illumination and atmospheric effects. The radiance spectra are transformed into reflectance by the atmospheric correction. This operation holds for solar spectrum, path radiance, sensor and sun directions, secondary illumination and shadowing. The second step is dimensional reduction. Since the hyperspectral data is collected in hundreds of bands and are highly correlated. This operation has a great impact since it reduces the amount of data, helps in computational savings in the unmixing step and improves the signal-to-noise ratio (SNR). The third step is spectral unmixing which comprises of two steps: end member determination and inversion. The first step estimates the signatures of the distinct end members present in the scene. The second step estimates the abundance fractions of each end member. The hyperspectral image analysis is shown diagrammatically in the Figure. 3.

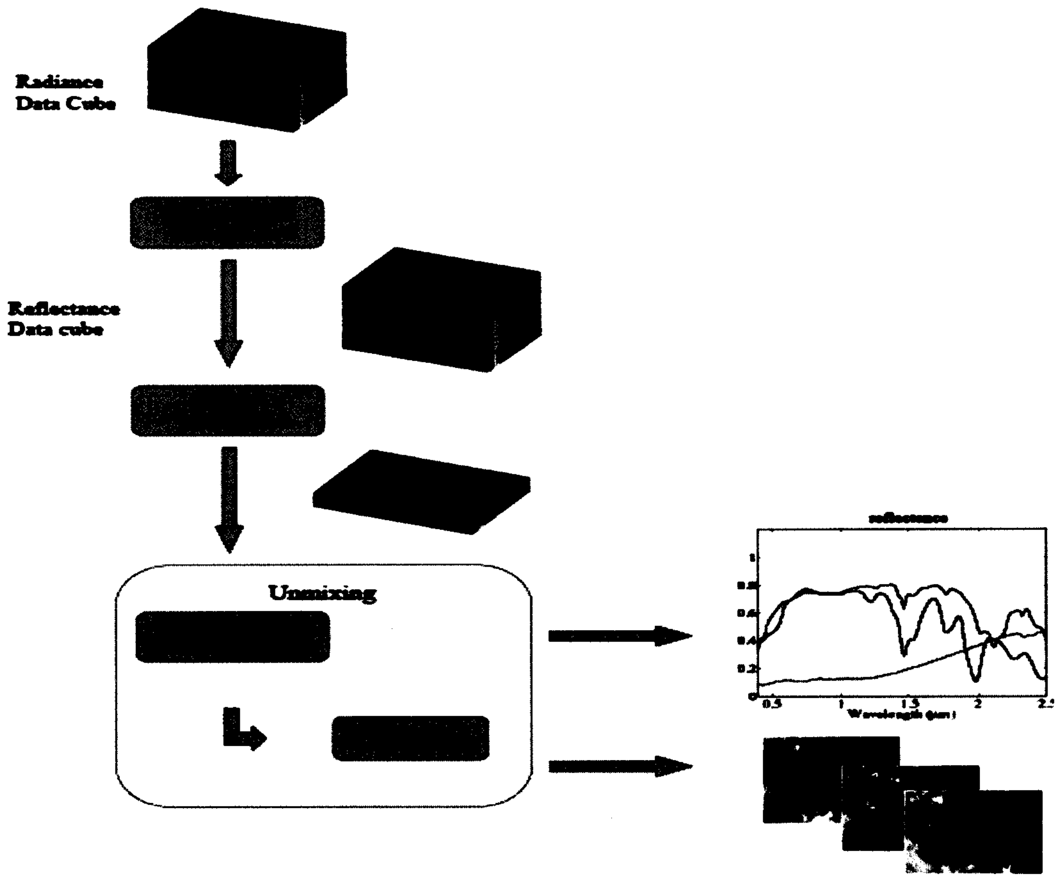


Figure 3: Hyperspectral image analysis

2.2 Dimensional Reduction

Significant amount of information is added to the high dimensional feature space data when the number of spectral bands increases in hyperspectral sensors. It is clear from the previous work that hyperspectral data contains redundant information which is in terms of spectral features [6]. The relevant information contained in hyperspectral data can be represented in lower dimensional subspace for specific applications [6]. Therefore dimensionality reduction of hyperspectral data preserving important information about specific objects of interest is a very important issue for the community of the remote sensing. Each pixel is a vector in hyperspectral image cube data, the dimension of which is equal to the number of bands. These bands are highly correlated due to the contiguousness of bands. Data redundancy occurred due to the high correlation between spectral bands. Hyperspectral imaging data contains high spectral resolution and spatial information, the data of hyperspectral images is huge and it brings high computational burden. The requirements for storage and the time for processing are much higher than those of multispectral imagery. It is very complicated task to analyze the information for discovering the underlying structures. It is difficult to unmix the pixels, the detection and classification of the targets in the presence of the bands which are highly correlated and also for such a huge data. The reduction of dimensionality is necessary for high accuracy in unmixing of the pixels, classification and detection. The performance of the many supervised classification methods are strongly affected by the dimensional reduction [7].

Methods of dimensional reduction can be divided into two categories, feature extraction (Based on transformation) [8] and band or feature selection. In hyperspectral imaging feature or band selection is preferable to feature extraction for dimensional reduction due to two main reasons [9]

- Feature extraction would need the whole (or most) of the original data representation for the extraction of new features, which always forces to obtain and deal with the whole initial representation of the data.
- Since the data are transformed so it is possible that critical information may have been distorted like when dealing with physical measures that are represented in the

hyperspectral image domain, while band selection has the advantage of preserving the relevant original information from the data [10]

The choice between feature extraction and feature selection will mainly depend on the application domain. In feature extraction bands are transformed which provide a better discrimination ability but there may not be a clear physical meaning in the new bands. Therefore the feature selection is very often preferred to feature extraction for dimensional reduction and above mentioned reasons supports this statement [9]. A great research has been done in the field of dimensional reduction of high dimensional data [11, 12, 13, 14, 15, 16, 17]. The techniques used for dimension reduction includes PCA, Isomap [17], multidimensional scaling (MDS) [12], Clustering and feature/ Band selection. [9, 14, 15]

2.2.1 Feature Extraction

In dimensional reduction of hyperspectral image data the original features are not selected. Instead new features are selected in a lower dimension space by applying some methods from statistics. Dimensionality reduction is done by mapping the high correlated dimension space onto uncorrelated low dimension space and due to these projections most desired information is preserved but the physical meaning of each spectral band is changed. Dimensionality reduction can also be done with the help of background knowledge and new features are constructed [31] and in addition to original features these new constructed features are used. Some of the classical techniques include the following

- PCA [32, 10, 33, 18]
- Factor Analysis [34, 35]
- Projection Pursuit [6]
- Wavelets transform [17]
- ICA [36, 37]
- Maximum Noise Transformation [2]
- Decision Boundaries based approach [38]
- Orthogonal Subspace projection approach [39]

2.2.1.1 Principal Component Analysis (PCA)

In hyperspectral Imagery, the amount of data is very large due to which some type of transformation is performed which can preserve the necessary information and reduce the amount of data. There is redundancy, dependency and noise in hyperspectral image data. The transformation methods can be helpful in three ways

- Effective data representation
- Effective feature extraction
- Effective image compression

Orthogonal transforms results in good representation due to the fact that they tend to extract non overlapped information from the data. Linear operations are involved in these transforms. Karhunen-Loeve Transforms (KLT) is the optimal orthogonal transform in the minimum mean square error sense. KLT is also called Principal Component Analysis. A great research has been done on PCA and the literature is well examined in [10, 33, 32, 3]. In PCA data is transformed from high dimensions to low dimension space preserving the maximum information. Due to the conceptual simplicity, analytical properties and relative efficient algorithm for computation [18] PCA is the most widely used technique for the dimensional reduction. PCA is best known as Karhunen-Loeve transform. PCA is a linear transformation on covariance matrix of the multi/ hyperspectral data to rotate the spectral coordinates in to a coordinate space in which the spectral components are uncorrelated. New coordinates are selected in result of PCA transformation for data set in which greatest variance by any projection of the data set resides on the first axis, called the first principal component, the second greatest variance on the second axis and so on. In PCA the calculation of the covariance of the data is involved. Covariance matrix characterizes the scatter of data.

If we have data set$X = (x_1, x_2 \dots x_M)$, then mean of data is defined as

$$\mu = E\{X\} \tag{4.1}$$

Where E is the expectation operator, Covariance matrix of data set is

$$C = E\{(X - \mu)(X - \mu)^T\} \tag{4.2}$$

Each element of covariance is the variance of between xi and xj. This variance shows the spread of component around the mean value, if the two components xi and xj are uncorrelated then the variance is zero. Due to the symmetry property of the covariance matrix, we can calculate an orthogonal basis by finding the eigenvalues and eigenvectors. The solution of the equations is eigenvectors ei and the corresponding eigenvalues λi

$$C_{e_i} = \lambda_i e_i ; \qquad i = 1, 2, \dots, M \tag{4.3}$$

By ordering the eigenvectors in the order of descending eigenvalues, an ordered orthogonal basis with the first eigenvector having the direction of largest variance of data can be created. In this way the directions along which the data set has the most significant amounts of energy can be found. After transformation of a data vector X, we get

$$Y = A(X - \mu) \tag{4.4}$$

This result the points in the orthogonal coordinate system defined by the eigenvectors. The components of Y can be observed as the coordinates in the orthogonal base, since the covariance of matrix Y is diagonal which means that the bands identified by matrix Y are uncorrelated bands. The ordinal data can be reconstruct by

$$X = A^T Y + \mu \tag{4.5}$$

We may represent the data in terms of only a few basis vectors of the orthogonal basis by using all the eigenvectors of the covariance matrix C. if the matrix having the K first eigenvectors as rows is denoted by A_K, taking the transformation of data as

$$Y = A_K(X - \mu) \tag{4.6}$$

This shows that we project the original data vector on the coordinate axes with K dimensions and transforming back the vector by a linear combination of the basis vectors. The mean-square error between the data and represented given number of eigenvectors is minimized. The concentration in a linear subspace may provide a way to compress data without losing much information and can simplify the representation. By selecting the eigenvectors with the largest eigenvalues we can reduce the possibility of losing much information as possible in the mean-square sense. In one way we can choose a fixed number of eigenvectors and their corresponding eigenvalues to represent and abstract the data consistently and a varying amount of energy of the original data is preserved. Alternatively we can choose an approximate amount of energy and the varying amount of eigenvectors and their respective eigenvalues and approximately a consistent amount of information is given in the expense of varying representations with regard to the dimension of the subspace [18].

2.2.1.2 Independent Component Analysis (ICA)

Independent component analysis is a statistical method which is linear transformation of data. In ICA the components are assumed to be mutually statistically independent. The data is transformed into components that are independent from each other as possible so that the bands we get are statistically independent. This way of representing the data captures the essential structure of the data in application including feature extraction and signal separation. The input data is a linear combination of independent and non Gaussian variables including a mixture matrix. ICA is therefore an essential method to extract useful information from data.

Let $X = (x_1, x_2, x_3, \dots, x_M)^T$ be the data set and is considered as a mixture of unknown signals $S = (s_1, s_2, s_3, \dots, s_k)^T$ from independent sources. The ICA model can be presented as

$$X = AS \tag{4.7}$$

Since A is an $M \times K$ matrix called the mixing matrix where$M \geq K$.

The target is to find A and separate the source signal S. The following assumptions are considered in ICA model to solve the problem.

- The principal assumption in ICA is independence. The sources are to be assumed as statistical independent. The joint (PDF) probability density function is the product of two marginal probability density functions that is

$$P(s_1, s_2, s_3, \dots, s_k) = p(s_1)\, p(s_2), p(s_3), \dots, p(s_k)$$

- The second assumption is that one source should have Gaussian distribution at most. Since the higher order cumulates are zero for Gaussian distributions. And therefore the Gaussian sources cannot beat most separated by independence assumption. The original independent components cannot be achieved with more than one Gaussian variable.

- The third assumption is that the mixture matrix is to be squared and invertible which means that the number of independent sources is equal to the number of mixtures. As is in hyperspectral imagery the observed data set should be greater than independent source signals (reflections) but the mixture matrix can be a square matrix.

There are number of ICA algorithms. Many of these algorithms start with one of the following criteria

- Maximization of non-Gaussiannity of the components [40]
- Minimizing mutual information [41]
- Maximum likelihood estimation [42,43]
- Tensorial methods [44]

A relationship can be observed among non-Gaussianity, mutual information and maximum likelihood [45]. Maximization of non-Gaussianity can be done by FastICA algorithms [40, 46].

2.2.1.3 Minimum Noise Fraction (MNF) Transform

Minimum Noise Fraction (MNF) transform is modified version of PCA. Since both depend on the second order moments. The difference between PCA and MNF is that PCA finds out the principal components on the basis of maximum variance of data matrix. But it doesn't work effectively in some cases. There is a linear transformation introduced by Green [47] which transforms the data by arranging the principal components with the signal-to-noise (SNR) decreased. That linear transformation is named as Minimum Noise Fraction Transformation. There are two cascaded Principal Component Analysis (PCA) transformations in MNF transform. In the first step the noise is uncorrelated and resealed and in result the noise in the transformed data has unit variance. This is the noise-whitening of data. In the second step PCA is applied on the noise whited data. Therefore the data is divided in to two parts. The first part, the eigenvalues are large and eigenimages are coherent. The second part has eigenvalues near-unity and noise dominated images. A detailed review is given by Neilsen in [48]. MNF transform is reformulated as the Noise Adjusted Principal Component (NAPC) transforms and is given by Lee at al [49]. The NAPC transform matrix combined these two PC transform matrices and in detail the NAPC transform consists of these stages.

2.2.2 Feature Selection

The method of dimensional reduction by selecting a subset of the original dimensions is called as feature selection. The selected minimum subset (bands, features) has great importance and sufficient for particular application for the improvement of accuracy and reduction of the data size [30]. Feature selection is carried out in the original feature space and the techniques of feature selection do not change the original representation of the variables, but only a subset of the original is selected. There is no transformation taken in selecting the subset o features to reduce the dimensionality, but the concentration is on the selection of features among the original features [22].

Let X be the original feature set, with L dimensions and X' be the subset feature that is $\tilde{X} \subseteq X$ with dimensions $K \leq L$. We suppose that $J(\tilde{X})$ be the objective function that is the feature selection criterion function. The higher value of J is assumed to be a better feature subset. If the probability of error is P_e then $(1 - P_e)$ is one criterion for maximizing J. Probability error as criterion function make feature selection dependent on the size of the data set, classifier/ detection technique and training use as well. Feature selection problem is mathematically an optimizing problem of finding subset features and is given by

$$J(\tilde{X}) = \max_{Y \subseteq X, |\tilde{X}| = K} J(Y) \tag{4.8}$$

Both search algorithm and a criterion function are involved in feature selection techniques [13]. Possible solutions of the feature selection problem which are subsets of features are generated and compared by the search algorithms by applying the criterion function. Several optimal and suboptimal search algorithms have been proposed [50, 51, 52]. The subset that contains a prefixed number of features and is the best in terms of adopted criterion, is identified by the optimal search algorithm where as suboptimal search algorithms select a good subset that contains a prefixed number of features but that is not necessarily the best one. An algorithm named as "A Branch and Bound Algorithm for Feature Selection" is proposed by Nerendra and Fukunaga [50] to find the optimal subset of features much more quickly than exhaustive search. Combinatorial complexity is the reason that optimal search algorithms cannot be used when the number of features exceeds a few tens as in hyperspectral image cube data. In remote sensing a suboptimal search strategy is applied for searching the best subset to minimize the objective function. In literature the information based measure criterion is proposed in [27]. In [53] an unsupervised feature selection algorithm is used together with the maximum information. And compression index as the similarity measure among the feature. Feature selection has a great research work since 70's [54]. Research work has been done in the areas such as statistics [50, 12], data mining [54], pattern recognition [55], machine learning [56] and neural networks [57].

2.3 Band Selection

In remote sensing, the hyperspectral data face many challenges. Some of the challenges include acquisition, transmission analysis, storage process and the extraction of information [58]. There are hundreds of non-overlapping bands along the spectrum covered by the high resolution of hyperspectral sensors. This high resolution results with huge data and high dependency on band images. Hyperspectral data has high correlation, presented in the spatial domain similar to the natural image. There are similar spectral signatures and high correlation in the adjacent locations of hyperspectral images. Due to this high correlation a lot of redundant information is found. Hyperspectral data cannot be analyzed by the traditional techniques of multispectral, as they are not so effective. For this reason it is necessary to reduce dimensionality of hyperspectral image data preserving the important information. Dimension can be reduced by two ways, as discussed above. One is Feature Extraction [23, 59, 60] and the other is Feature/ Band Selection [7, 61]. Feature selection is given the preference on feature extraction as we have discussed above. Bands election is one of the techniques of feature selection in which there is a very low correlation among the selected bands having maximum useful information. There are two main reasons that bands selection is preferred to feature extraction [9]. One is, feature extraction use most of the original data representation for the extraction of new features and the other is, due to the transformation process some critical and crucial information may have been distorted. While on the other hand band selection has a great advantage, that it preserve the important original information of the data [10]. Therefore band selection is considered as an effective means to overcome the problem of dimensionality of remotely sensed image data. Methods for band selection can be categorized in the following four groups [8, 62].

1. Search-based Methods [63, 64, 6]
2. Transform-based Methods [65, 66]
3. ICA-based Band Selection [67]
4. Information-based Methods [68, 62, 10, 20]

2.3.1 Search-based Methods

In hyperspectral data the bands are in order of hundred so the search-based methods are not feasible due to "Combinatorial explosion". Several approaches to optimize algorithms have been applied such as genetic algorithm, hill climbing and greedy for the improvement of search efficiency. Several optimal and suboptimal algorithms are present in literature [50, 51, 52], but since for such number of band combinations, computational cost is still high and the problem of local minima occur [62, 69].

2.3.2 Transform-based Methods

In transform-based methods, there is matrix transformation such as ICA or eigenvector analysis for the projection of data on to lower dimension space [65, 66]. This type of transformation causes loss of the original meaning of spectral data and the interpretation becomes very difficult. Another problem of transform-based method is that it requires the full data cube in original form before the transformation, which is a big disadvantage for real time processing.

2.3.3 ICA-based Band Selection

Pierre Common in 1994 first proposed ICA which has been used in many applications like Blind Source Separation (BSS) [70, 71] recognition etc. Chiang et al and Lennon et al has used ICA in [17] and [36] as a feature extraction method. In this method hyperspectral images are presented in lower dimensional feature space. In [67] ICA-based Band Selection was proposed, in which this method avoids transformation of the original hyperspectral image data to the feature space. Instead of transformation comparison of the average absolute weight coefficients of individual spectral bands is performed and independent bands are selected, which has maximum information. Thereby reducing the dimensionality and preserving most spectral features of hyperspectral image data.

2.3.4 Information-base Band Selection

Information-based Band Selection methods measure the information contains in the individual band. If the information content is related and capable to discriminate, the bands are selected which has higher information. The entropy, the contrast and correlation are the commonly used information metrics [64,20]. As compared to the transformed-base methods, the advantage of the information-based methods is the selection of the subset of hyperspectral data in such a way that the original information retains. There are other band selection techniques which include a trade-off scheme between the resolution and spatial resolution [72], Spectral Angle Mapper maximization [73], high order moments [74], wavelet analysis [75]. Entropy [64, 20] and mutual information [76, 77] has a good role and obviously have a good potential for band selection [69]. In Shannnon's information theory the information content measured by entropy is in terms of uncertainty.

Let X is a random variable and the set δ is the values of X with the probability distributionsPX (δ), then the definition of entropy is

$$H\,(X) \;=\; -\,\Sigma\, PX\,(\delta)\, log\, PX\,(\delta) \qquad (4.9)$$

Methods in [64,20] directly use entropy for band selection, in which the entropy is calculated to estimate the level of the information content in individual band or in wavelength interval. According to entropy values of the spectral bands, the spectral bands are ranked in certain order. Band selection is done by choosing those bands which have higher entropy values. From equation 4.9 entropy is a function of single variable and $H(X)$ is calculated on a single signal, which clearly means that the information measured by entropy has no point of reference. Therefore there is no guarantee of matching the amount of entropy and the information content, which is useful for target classification [62]. Mutual information provides an ideal framework in the sense that it measures the similarity between two random variables. It was introduced for band selection in [76, 77]. MI in information theory is a basic concept of measurement of the statistical dependence between two variables [62]. For two random variables X and Y having the marginal probability distributions $P_X\,(\delta)$ and $P_Y\,(J)$ and $P_{XY}\,(\delta,J)$ is the joint probability distribution. MI is defined as

$$I(X,Y) = \textstyle\sum_{\delta,J} log\, \frac{P_{XY}\,(\delta,J)}{P_X\,(\delta) P_Y\,(J)} \qquad (4.10)$$

The following equations can be derived from equation4.10 between MI and entropy

$$\begin{aligned} I(X,Y) &= H(X) + H(Y) - H(X,Y) \\ &= H(X) - H(X|Y) \\ &= H(Y) - H(Y|X) \end{aligned}$$

Where $H(X)$ and $H(Y)$ are the entropy of X and Y respectively and $H(X,Y)$ is the joint entropy. $H(X/Y)$ is the entropy of X given Y and $H(Y/X)$ is the entropy of Y given X. The dependency between a spectral image and the reference map can be estimated by using MI. It is helpful in investigating how much information a spectral image contains about the reference map [62]. A reference map is always required for MI based band selection. As the reference map is unavailable so it is required to estimate the reference map. An adaptive method is proposed in [62] in which MI is calculated by using the estimated reference map $\check{\mathrm{R}}$. Using the prior knowledge the estimated map can be approximated.

The region of the spectrum or key spectra denoted by Z is assumed to contain the most discriminatory information. If $I_j \in Z$; $1 \leq J \leq M$ is the set of spectral images as important bands then the estimated reference map can be calculated as

$$\hat{R} = \frac{1}{M}\sum_{j=1}^{M} I_j \tag{4.11}$$

Following are the few unsupervised methods of comparison for band selection.

- WaLuMI: Ward's Linkage strategy using Mutual Information [27],
- WaLuDi: Ward's Linkage strategy using Divergence [78,10],
- CBS methods: Constrained Band Selection [9]
- MVPCA: Maximum Variance Principal Component Analysis [77, 78, 10],
- Information Divergence [9].

2.3.4.1 Constrained Band Selection (CBS)

Constrained Band Selection (CBS) was developed in [9] which is a different method from the variance-based or information-based approaches. In this approach there is a linear constrain on a band while minimizing the correlation in hyperspectral image. CBS methods have four solutions for optimizing a problem. Two methods are based on correlation and two on dependency. The Linearly Constrained Minimum Variance (LCMV) and the Constrained Energy minimization (CEM) [79] approaches derives the CBS presented in [9]. Therefore, there are two ways to implement these approaches. The implementation based on CEM has a high computational cost and the implementation based on (LCMV) reduces this complexity substantially. From the experimental results it is clear that the performance of LCMV-CBS and CEM-CBS has a great similarity [9], therefore we use LCMV-CBS for comparison

2.3.4.2 MVPCA: Maximum Variance Principal Component Analysis

For band selection, a joint approach named as band prioritization and band-decorrelation is presented in this section and is used in [77] for classification in hyperspectral image. It is used for comparison of band selection methods in [9]. The band prioritization depends on the eigen analysis in which a matrix is decompose into an eigen form matrix. From eigen form matrix a loading factors matrix could be generated and this loading factor matrix is used to prioritize bands. The priority of each band is determined and according to the associated priorities the bands are ranked and are sorted from high to low variance.

2.4 Unmixing of Hyperspectral Data

The low resolution of hyperspectral sensor causes the existence of distinct substances in a hyperspectral image pixel. Individual spectrum of each material is composed of measured spectral in that pixel [80]. Homogeneous mixture of different material cause is the main cause of mixed pixel and mixed pixel is independent of sensor resolution [80]. The decomposition of the pixel spectrum into a collection of endmember spectral signatures and their respective abundance fractions [73, 80] is called unmixing of hyperspectral image. The unmixing changes the spectral and spatial resolution [81, 82]. The classification of multispectral and hyperspectral analysis is as follows [83, 84].

- Detect known or unknown objects or materials [85, 86]
- Classification [84, 87, 79]
- And estimation of material and respective area occupied with in a pixel [88, 89]

The target spectrum and background spectrum may also be used for endmember detection to observe the spectrum of the mixed pixel [90, 91]. The two main classes of endmember derivation methods are: the algorithm in which the assumption is that end member exist in the image either in pure or mixed pixel and use selection of n-dimensional scatter plot method and convex cone method [92, 93], and the algorithm in which the assumption is that endmembers are derived analytically [81, 82, 86, 94]. To model mixed pixel two models are used:

Linear mixture model [95, 96, 97]

Non linear mixture model [98]

Mostly linear Mixture Model is used widely for source separation [86, 99, 100]. In LMM modeling there is an assumption that the observed pixel spectrum is the linear combination of unique and distinct deterministic spectral signatures i.e. endmembers/targets. The mathematical model of LMM for mixed pixel [101]

$$\mathbf{x} = \sum_{k=1}^{K} a_k s_k + \mathbf{w} = \mathbf{Sa} + \mathbf{w} \tag{4.12}$$

Where $S = [s_1, s_2, s_3, \dots, s_k]$the k endmember spectral or targets and these targets are are assumed to be linearly independent. $a = [a_1, a_2, a_3 \dots a_k]^T$ are abundance fractions for the endmember spectra. w is the additive noise vector and considered as model error. If L is the number of spectral bands then X is a $L \times 1$ column pixel vector and S target signatures matrix of size$L \times K$.

Stochastic mixture model is considered if the end member spectra are random and independently drawn from multivariate normal distributions [85, 102]. The assumption to choose a pixel composition i.e. pure or mixed model selection for spectral variability and selection of mixture procedure leads to different types of target detection algorithms. A binary hypothesis test for detection is formulated with two hypotheses

-Background only (H_0)

-Target and background (H_1)

These two hypotheses have unknown parameters e.g. covariance matrix of the background to be calculated and adaptive detector usually designed using likelihood ratio test approach [103]. Target detection algorithm using the stochastic mixing model, called as finite matched filter is explained in [85] and [102]. Mathematical expression for non linear mixture model is given as

$$\mathbf{x} = f(\mathbf{s}, \mathbf{a}) + \mathbf{w} \tag{4.13}$$

Where f( ) is non linear function for the non linear relation between a and w.

2.4.1 N-FINDER Method

N-FINDER is an algorithm for finding endmember/ targets in mixed pixels of hyperspectral image data; developed by Winter et al [93]. It is assume that pure pixels of endmembers are present in image data and from this assumption, L dimensional spectrum, volume of L dimension is created by a simplex in which the purest pixel locate the vertices. A simplex is the geometric object which spans a given space. N-FINDER is an iterative simplex volume expansion approach and it takes start with a random selection of set of pixels from the scene as initial endmembers. An iterative process continuous and replace the initial endmembers with the pixel tested and calculating the simplex volume. If there is increase in volume, the pixel being tested replaces the initial endmember. The process is repeated again until each pixel is tested as potential endmember. The pixel which are endmember every time at the end of the process are considered as final endmember. Endmember is determined by the N-FINDER from the image regardless of noise and with no prior knowledge. This algorithm doesn't determine a way to find the number of endmember in data. Virtual Dimensional (VD) [79, 88] is used to determine how many endmember exist in the data needed. We have used the new concept of VD [79] for the number of endmembers needs to be calculated by N-FINDER. This algorithm starts with a random set of pixels and this randomness can affect the convergence rate of algorithm and the final results. The Unmixing of image data using N-FINDER is a five step process

- The initialization algorithm (EIA) can be used to select an appropriate set of endmembers for N-FINDER initialization.
- Estimation of total number of endmembers in data (by VD concept)
- The preprocessing in which data redundancy is reduced by MNF-transformation or some other transformation
- Initialization (EIA algorithm)
- Volume finding of simplex with initial endmember considered as vertices and replacing the vertices with testing pixel if the volume of simplex is increased with testing pixel by technique [107] using endmembers.

2.4.2 Independent Component Analysis

ICA is an unsupervised source separation process ICA has applied to linear blind separation problems. The literature for ICA is present in [70, 71, 81, 103, 104]. ICA extracts each source automatically from the observation of linear combination of these sources and assumes the sources to be statistical independent [48]. ICA can work only with sources less than 10 with many observations which is the disadvantage of ICA [48].

2.4.3 Vertex Component Analysis

For multispectral and hyperspectral data as linear mixture model, the vertices of a simplex are the endmembers. To exploit this simple geometry features of data there are several approaches [82, 92, 105, 106]. Minimum Volume Transform (MVT) was proposed by M. D Craig [110], which determines the minimum volume simplex containing the data like N-FINDER method. Method developed by C. Bateson et al [107] was also like MVT but that method considers variability of spectral pixels present in hyperspectral mixtures. VCA was proposed by Jose M. P Nascimento and Jose M. Bioucas Dias [82], is an unsupervised endmember extraction algorithm like N-FINDER, VCA also exploits that the endmember are the vertices of a simplex and works on the assumption that pure pixels of endmember are present in data. Jose M. P Nascimento claimed that the performance of VCA is much better than PPI and better than N-FINDER, but it has a computational complexity between one and two orders of magnitude lower than N-FINDER.

Chapter 3

Proposed Method of Dimensional Reduction

3.1 Band Clustering and Selection

Band Clustering and Selection are two steps used in our work. Clustering of band images keeps the intra-cluster variance minimum and the inter-cluster variance maximum. The method in which dimensionality is reduced by selecting a subset of the original dimensions are known as band/ feature selection. The hyperspectral data is spread in some direction. This data can be measured by using different statistical methods which include MAD (Mean Absolute Deviation), moment, variance, mean, geometric mean and standard deviation. We have used Standard Deviation, MAD and, Variance in our proposed work. Suppose that we have $\{B_l\}_{l=1}^{L}$ band images in our hyperspectral image data cube where L is the total number on bands, if each band image is of size M×N and B_l the mean of the l_{th} band image. The statistics we use for data are given below.

MAD for the l_{th} band is

$$d_l = \frac{1}{MN}\sum_{i=1}^{MN} |b_i - \bar{B}_l| \tag{3.1}$$

Standard Deviation for the l_{th} band image is

$$d_l = \left(\frac{1}{MN}\sum_{i=1}^{MN}(bi - \bar{B}l)^2\right)^{\frac{1}{2}} \tag{3.2}$$

Variance for the l_{th} band image is

$$dl = \frac{1}{MN}\sum_{i=1}^{MN}(bi - \bar{B}l)^2 \tag{3.3}$$

The result from the above statistical methods for L band images is given by

$$\mathbf{d} = \{d_l\}_{l=1}^{L}$$

3.1.1 K-means Clustering

For clustering the bands (band images) K-means clustering technique is used. For K-means clustering city block and Square Euclidean distance metrics are used. K-means clustering is one of the simplest unsupervised algorithms and is well-known for solving the problem of clustering. The flowchart of K-means clustering is shown in figure 4. K-means follows a simple and easy way to classify a given data set through clusters; the number of clusters is fixed and is given a prior. The number of centroids i.e. K are defined for each cluster and which are placed far away from each other as possible. The points which belong to the given data set are taken and are associated to the nearest centroid which results in K number of groups. Again K new centroids are recalculated for new centers of the cluster and a new binding has to be done between the same data set points and the nearest new centroid. A loop is run for the K centroids to change their location step by step until there is no change and the centroids are fixed. The centroids of the clusters are calculated by minimizing the sum of squared errors. The K means algorithm performs three steps until convergence.

1. Determine the centroid coordinate
2. Determine the distance of each object to the centroids
3. Group the object based on minimum distance

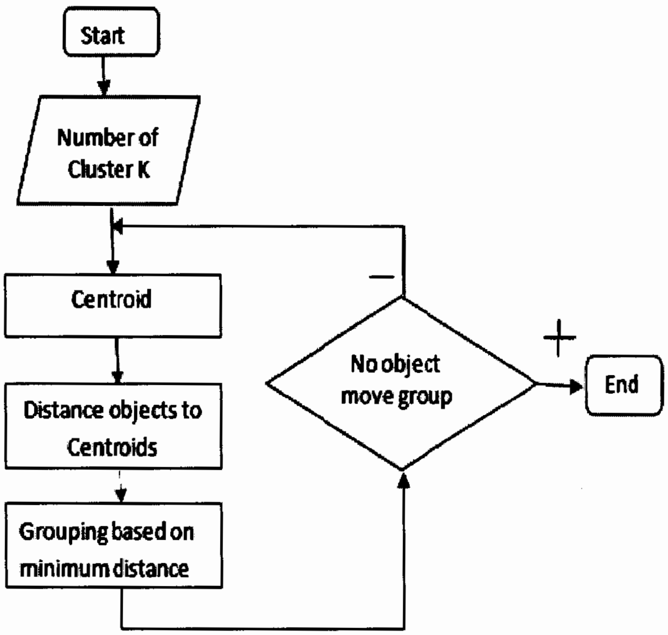


Figure 4: K-Means Clustering Flow Chart

For the observations $X = (x_1, x_2, x_3 \ldots x_n)$, the K-means clustering method divides the n observations into k sets (k < n), $K = \{S_1, S_2, S_3 \ldots S_k\}$, minimizing the sum of squares with-in clusters i.e.

$$\min_{S} \sum_{i=1}^{k} \sum_{x_j \in S_i} \left\| x_j - \mu_i \right\|^2$$

Where μ_i is the mean of points in cluster C_k

K-means computes centroid clusters differently for the different supported distance measures. We have used the following distance metrics in K-means clustering.

- **Square Euclidean**

For an m-by-n data matrix $X = (x_1, x_2, x_3 \dots x_m)$ the distance between the vector x_r and x_s is given as:

$$d_{rs}^2 = (x_r - x_s) D^{-1} (x_r - x_s) \tag{4}$$

Where D is the diagonal matrix

- **City Block metric**

For an m-by-n data matrix $X = (x_1, x_2, x_3 \dots x_m)$ the distance between the vector x_r and x_s is defined as

$$d_{rs} = \sum_{j=1}^{n} |x_{rj} - x_{sj}| \tag{5}$$

3.2 Proposed Algorithm

Following are the steps of the proposed algorithm to summarize the band clustering and selection.

1. Calculate the number of bands i.e. VD.

2. Calculate or measure the data of each band image using Variance (VAR), MEAN Absolute Deviation (MAD) and Standard Deviation (STD).

3. Band clustering using K-means clustering and using distances among the measured values to examine the proximity of band images to each other.

4. According to VD, clusters are created which contain all the measured values.

5. From each cluster, one band having maximum value is picked.

Now the question is how many bands need to be selected preserving the necessary information. This problem can be solved by using the new concept of Virtual Dimensionality (VD) [113] to estimate the minimum number of bands and preserve the maximum useful information. The selected bands are analyzed for the endmember detection. VCA [82] is then used for the unmixing process of the hyperspectral image and the results are compared.

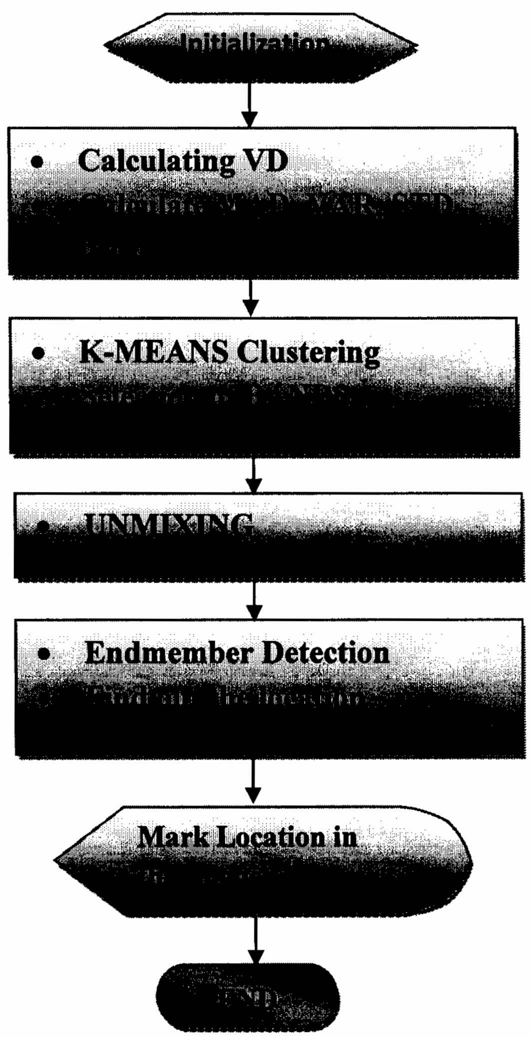


Figure 5: Flow Chart of the Proposed Algorithm

Chapter 4

Experimental Results

4.1 Hyperspectral Image Data

We have used a well known Airborne Visible/ Infrared Imaging Spectrometer [110] for our research work. The Cuprite image is used to compare and evaluate the proposed research work. The image scene is shown in Figure. 6 and it is available at website [111]. It was collected by 224 spectral bands having 10 nm spectral resolutions. These are collected over the Cuprite mining site, Nevada in 1997. Cuprite is a mining area in the south of Nevada with minerals and little vegetation. The geologic summary and mineral map can be found in [112]. Cuprite has been widely used for experiments in remote sensing and has become a standard test site to compare different techniques of hyperspectral image analysis. In our research work, a sub image of size 350×350 with 224 bands of a data set taken on the AVIRIS flight of June 19, 1997. The instrument of AVIRIS covers 0.41 – 2.45 μm regions in 224 bands with a 10 *nm* bandwidth and flying at an altitude of 20 *km*, it has an Instantaneous Field Of View (IFOV) of 20 *m* and views a swath over 10 *km* wide. Prior to the analysis of AVIRIS Cuprite image data, low SNR bands 1 – 3, 105 – 115 and 150 – 170 have been removed and the remaining 189 bands are used for experiments. The ground truth of spatial positions of four pure pixels corresponding to four mineral alunite (A), buddingtonite (B) , calcite (C) and kaolinite

(K) are indicated by "A", "B", "C", and "K" respectively. Endmembers extracted by VCA are verified by using these labels of spatial positions. The USGS signatures of "A", "B", "C" and "K" are also shown in Figure. 7.

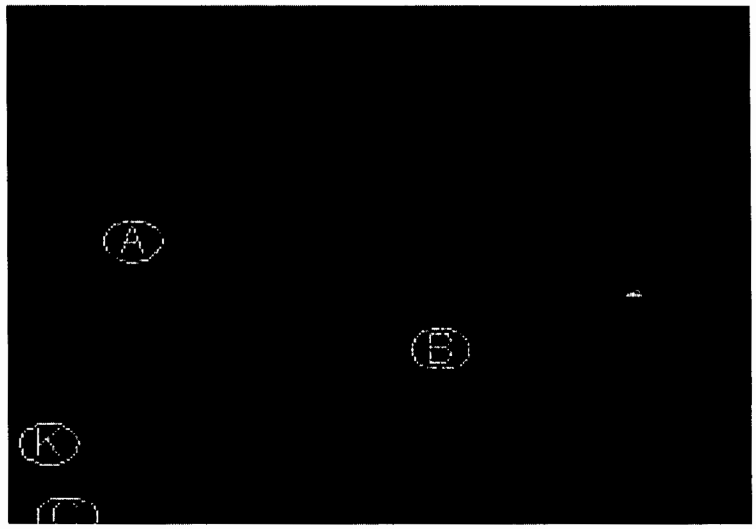


Figure 6: Spatial positions of four pure pixels which shows correspondence to the minerals: Alunite (A), Buddingtonite (B), Calcite (C), and Kaolinite

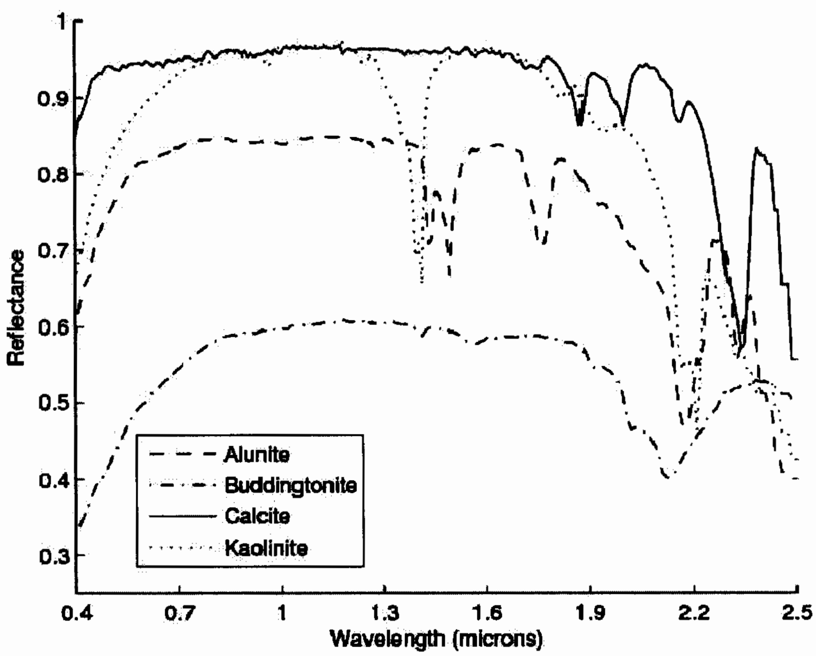


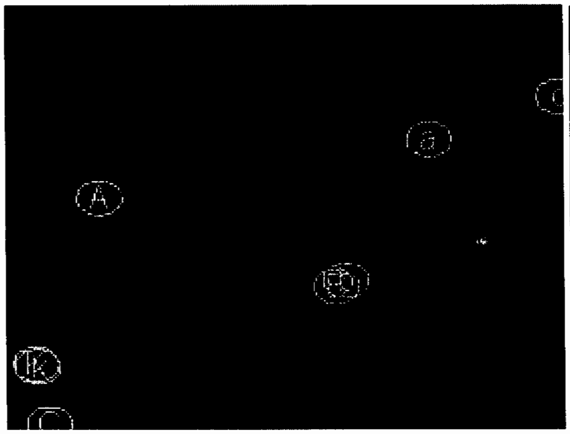
Figure 7: USGS spectral signatures of Alunite (A), Buddingtonite (B), Calcite (C), and Kaolinite (K)

4.2 Results and Discussion

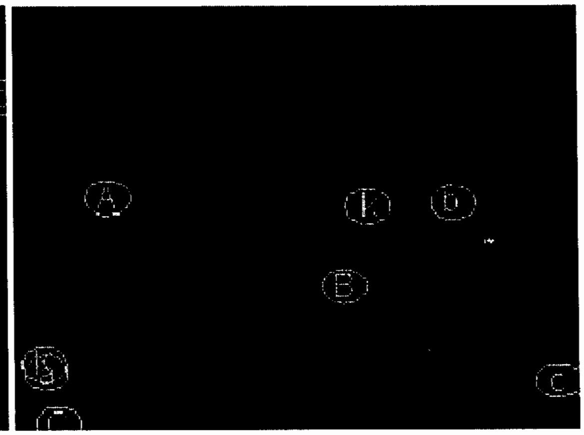
Preserving the maximum information, the number of bands required and estimated VD are 22. In our research work we have tabulated 22 bands. The bands selected by (Linearly Constrained Minimum Variance based Constrained Band Selection) LCMV-CBS, (Minimum Variance Principal Component Analysis) MVPCA and our proposed techniques of clustering according to 22 VD are given in table 4.1. The proposed methods are Mean Absolute Deviation with City Block as distance metric in Clustering abbreviated by MAD-CB, similarly Mean Absolute Deviation with Square Euclidean is MAD-SE. Variance with City Block is VAR-CB and VAR with Square Euclidean is VAR-SE. Standard Deviation with City Block is STD-CB and Standard Deviation with Square Euclidean is STD-SE. Figure. 8 shows the extraction of four end members and also the extracted endmembers by VCA using the 22 selected bands given in table 4.1, the detected endmember/ targets are labeled with "a", "b", "c", "k". the detected endmembers are compared with the ground truth endmember pixels which are labeled as "A", "B","C","K". In addition the measurement of the spectral similarity between the endmember pixels "a","b","c","k" and the ground truth actual endmember pixels "A","B","C","K", we have calculated the Spectral Angle Mapper (SAM), the results of which are tabulated in table 4.2. The location of the "A","B","C","K" and "a","b","c","k" in the image scene are also included in the form of (X,Y) coordinates in brackets. For example the actual location of "A" is (61,161), "B" is (209,234), "C" is (22,298) and "K" is (22,298). Similarly the coordinates for found target/endmembers in brackets shows the location in the image scene.

Table 4.1.Bands selection using different techniques

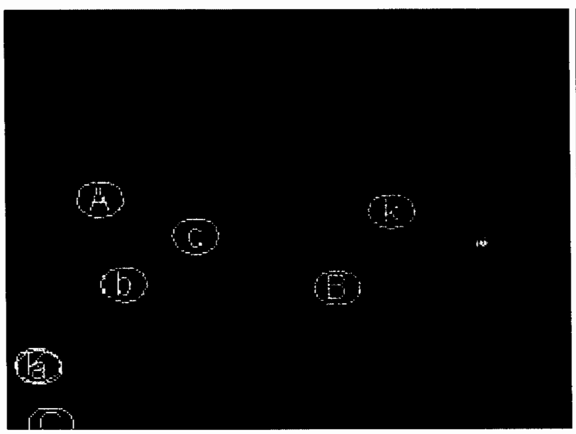
Criteria	Selected bands
MAD-CB	92, 77, 12, 57, 99, 65, 129, 102, 35, 18, 87, 5, 119, 155, 69, 131, 11, 149, 105, 178, 116, 163
VAR-CB	39, 100, 89, 154, 13, 51, 18, 177, 148, 22, 168, 102, 64, 4, 183, 61, 67, 87, 70, 32, 81, 35
STD-CB	5, 85, 13, 170, 51, 89, 151, 33, 73, 177, 18, 81, 64, 147, 67, 155, 168, 87, 9, 176, 23, 61
MAD-SE	53, 19, 102, 87, 17, 170, 168, 35, 155, 69, 183, 39, 178, 149, 112, 4, 24, 48, 94, 107, 65, 164
VAR-SE	82, 107, 39, 26, 163, 128, 57, 51, 21, 177, 5, 147, 89, 99, 14, 151, 78, 34, 72, 18, 87, 67
STD-SE	91, 61, 36, 117, 74, 177, 125, 130, 21, 67, 135, 57, 141, 14, 87, 5, 51, 131, 83, 26, 99, 70
LCMV-CBS CM/BDM	26, 117, 48, 37, 189, 64, 1, 185, 10, 172, 47, 4, 60, 28, 165, 17, 5, 2, 151, 158, 3, 94
LCMV-CBS CC/BDC	185, 37, 2, 3, 5, 64, 8, 9, 6, 7, 10, 165, 4, 11, 12, 14, 151, 13, 28, 15, 16, 153
MVPCA	87, 85, 88, 86, 89, 84, 91, 80, 78, 90, 92, 83, 82, 79, 93, 81, 98, 99, 97, 189, 77, 76



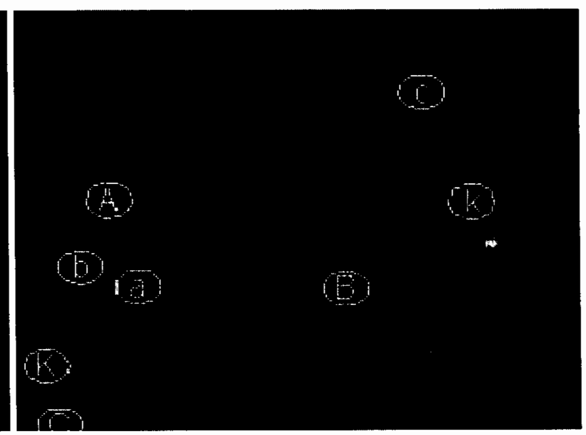
(a)Full Bands



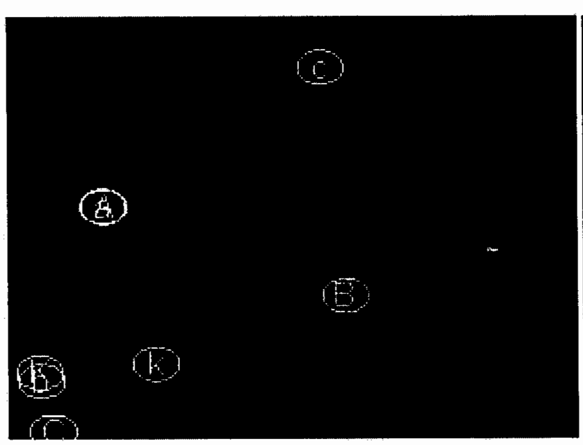
(b) MVPCA



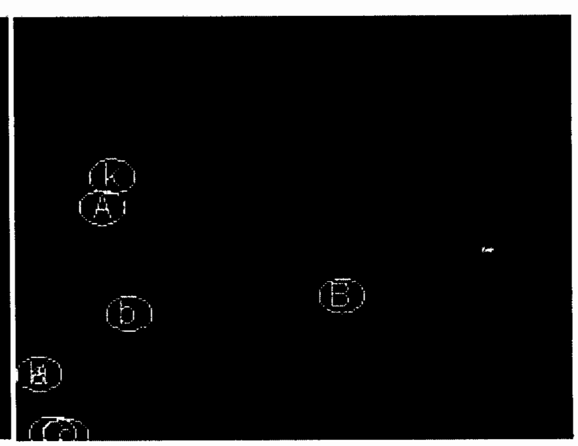
(c) LCMV-CBS BDM



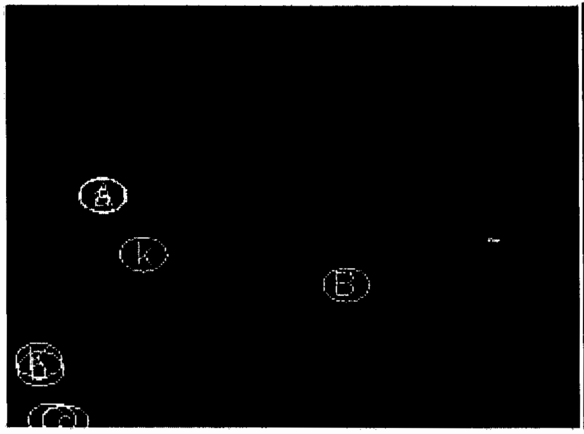
(d) LCMV-CBS BCC/ BDC



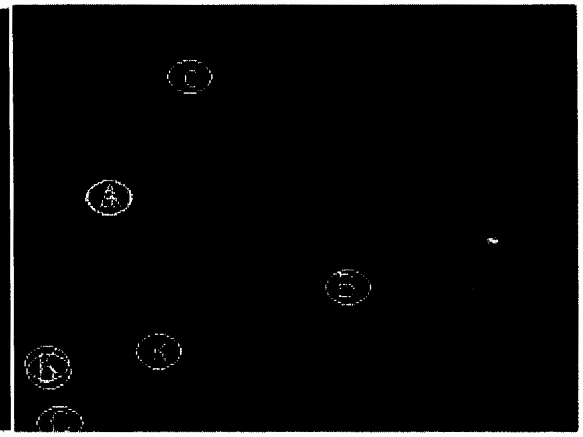
(e) MAD-CB



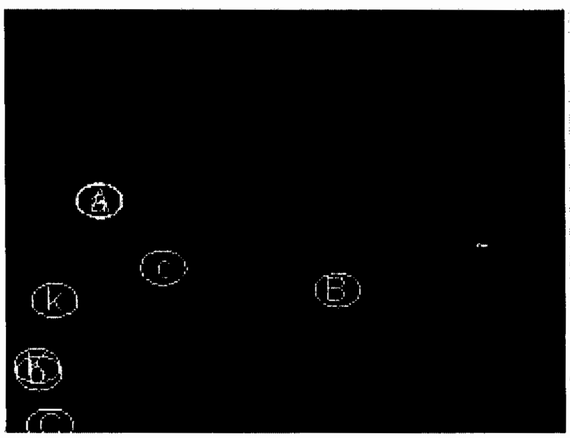
(f) VAR-CB



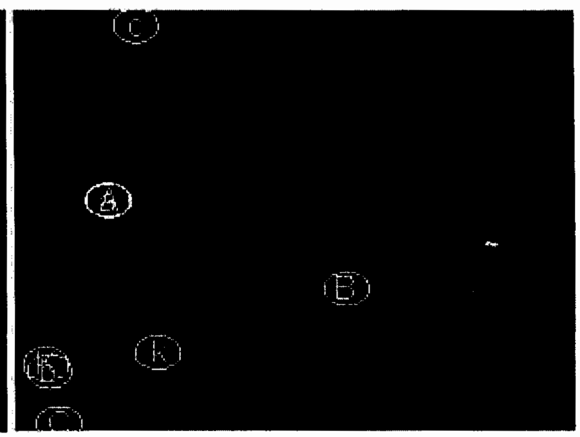
(g) STD-CB



(h) MAD-SE



(i) VAR-SE



(j) STD-SE

Figure 8: Extraction of Endmembers by VCA from Full bands and the selected bands given in Table 4.1

The result obtained from the simulations shows that the performance of the clustering-based band selection techniques, using K-means clustering are better than the band selection using LCMV-CBS and MVPCA and Full bands. The Spectral Similarity is calculated by SAM (Spectral Angle Mapper). The SAM computes the angle between the actual pixel spectrum and the found target spectrum. The smaller the spectral angle the more similar the actual and found target spectra. The values of (SAM) among the same target/ endmembers minerals are highlighted which shows a good similarity. The results of bands selection using Full bands, LCMV-CBS and MVPCA shows that the (SAM) values for some found endmembers are very high as compared to our proposed band selection methods. The best results among all the proposed techniques are STD-SE and MAD-SE due to the good similarity values. Results from the techniques MAD-CB, VAR-CB, STD-CB and VAR-SE are also better than the previously proposed techniques as they too have results in good similarity with the actual endmembers. The detection of the endmember pixel using the selected bands and K-means clustering by VCA gives better results compare to the LCMV-CBS and MVPCA, therefore the detected endmember pixel have high spectral similarities.

Table 4.2 Spectral similarity measurements (SAM) among found endmembers and the ground truth endmembers

	A (61,161)	B (209,234)	C (22,298)	K (22,298)
Full Band				
a (267,113)		0.2146	0.2578	0.1136
b (215,229)	0.1330		0.1089	0.1378
c (349,78)	0.2172	0.1141		0.2408
k (23,300)	0.1043	0.1734	0.2165	
LCMV-CBS BCC/BDC				
a (23,305)		0.1959	0.2354	0.1092
b (277,165)	0.1247		0.1477	0.1403
c (342,312)	0.1680	0.1017		0.1975
k (224,168)	0.0888	0.1834	0.2283	
LCMV-CBS BCM/BDM				
a (23,300)		0.1353	0.1807	0.1075
b (77,231)	0.2046		0.1027	0.2071
c (121,191)	0.1839	0.1027		0.2012
k (243,171)	0.1043	0.1734	0..2165	
MVPCA				
a (80,232)		0.1727	0.2202	0.0913
b (44,216)	0.1763		0.1076	0.1626
c (257,72)	0.1944	0.0748		0.2067
k (288,163)	0.0684	0.1702	0.2097	
MAD-CB				
a (22,298)		0.1413	0.189	0.0969
b (788,248)	0.1643		0.1035	0.1928
c (38,349)	0.1871	0.0839		0.192
k (68,135)	0.0961	0.1733	0.2114	
VAR-CB				
a (60,161)		0.1654	0.2125	0.0933
b (23,305)	0.1329		0.0856	0.1468
c (194,45)	0.2202	0.1002		0.2393
k (93,291)	0.0889	0.1834	0.2283	
STD-CB				
a (61,160)		0.1654	0.2125	0.0933
b (23,305)	0.202		0.0925	0.2179
c (38,349)	0.1871	0.0839		0.192
k (86,309)	0.0889	0.1834	0.2283	
MAD-SE				
a (61,161)		0.1645	0.2115	0.0962
b (23,305)	0.1412		0.0827	0.1541
c (112,62)	0.1962	0.0995		0.209
k (92,288)	0.0889	0.1834	0.2283	
VAR-SE				
a (61,160)		0.1654	0.2125	0.0933
b (23,304)	0.1966		0.0889	0.1996
c (101,216)	0.2021	0.0813		0.1989
k (33,243)	0.104	0.1787	0.2215	
STD-SE				
a (61,160)		0.1654	0.2125	0.0933
b (24,304)	0.1412		0.0827	0.1541
c (79,16)	0.2465	0.1204		0.2528
k (92,288)	0.0938	0.1766	0.2207	

Chapter 5

CONCLUSION AND FUTURE WORK

5.1 Conclusion

The proposed techniques for dimensional reduction and target detection give better results as compare to LCMV-CBS and MVPCA. The results shows that if the dimensions of Hyperspectral data are reduced by clustering the band images using their statistical parameters, then it gives better results for unmixing and detection than other techniques like LCMV-CBS and MVPCA etc. In proposed technique of band clustering and selection using K-means method, a band from each cluster is selected such that intra-cluster variance is maximum and inter-cluster variance is minimum. Furthermore STD-SE is better amongst our proposed technique. The proposed techniques are simple to implement and computes the result very fast. The computation takes seconds for band clustering and selection. Comparing the computation time at any computer with the previously proposed methods of band selection, the results computed on the basis of proposed methods takes very less time. All the endmember/targets are detected well and have high spectral similarities. The proposed method can be used in other applications where data dimensional reduction is a problem. In resultant the proposed clustering techniques are promising techniques for band clustering and band selection.

5.2 Future Work

Bands selection is a one way of feature selection. New techniques may be developed for the bands selection to reduce the dimensionality of hyperspectral image data that can be search based, transform based, ICA based or information based methods. With development of new technique for dimensional reduction, other techniques for unmixing and detection would be implemented for the evaluation of proposed method. Furthermore reduced dimensional data would be analyzed for unmixing and detection of multiple targets. Purpose of development of feature selection technique is not only to reduce dimension but also to improve the detection accuracy not only for single target but also multiple target. For multiple targets, multiple dimensional reduction technique would be developed to get high accuracy through the development of decision fusion techniques.

References

[1].C.-I. Chang, 'Hyperspectral Imaging: Techniques for Spectral Detection and Classification', *Chapter 2, Kluwer Academic/Plenum Publishers, New York* ~2003

[2] A. F. H. Goetz, G. Vane, J. Solomon, and B. Rock, "Imaging spectrometry for earth remote sensing [c]," in *Airborne Imaging Spectrometer Data Analysis Work- shop, JPL Publication 22-29*, vol. 228, pp. 22{29, 1985.

[3] J. Ellis, "Searching for oil seeps and oil-impacted soil with hyperspectral imagery," *Earth Observation Magazine*, January 2001.

[4] R. Smith, "Introduction to hyperspectral imaging with tmips." *MicroImages Tutorial Web site*, July 2006

[6] L. O. Jimenez and D. A. Landgrebe, "Supervised classification in high-dimensional space: geometrical, statistical, and asymptotical properties of multivariate data," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 28, pp. 39{54, Feb 1998.

[7] S. Serpico and L. Bruzzone, "A new search algorithm for feature selection in hyperspectral remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, pp. 1360{1367, Jul 2001

[8] J. Luis and D. Landgrebe, "Supervised classification in high dimensional space: Geometrical, statistical, and asymptotical properties of multivariate data," *IEEE Transaction on System, Man, and Cybernetics*, vol. 28, pp. 39-54, Feb 1998.

[9] C.-I. Chang and S. Wang, "Constrained band selection for hyperspectral imagery," *IEEE Transactions on Geosciences And Remote Sensing*, vol. 44, no. 6, pp. 1575-1585, 2006

[10] A. Martinez-Usomartinez-Uso, F. Pla, J. M. Sotoca, and P. Garcia-Sevilla, "Clustering-Based Hyperspectral Band Selection Using Information Measures [J]," *IEEE Transactions on Geosciences and Remote Sensing*, vol. 45, pp.158 -4171, Dec. 2007

[11] D. A. Richards, "Thematic mapping from multi temporal image data using the principal component transformation," *Remote Sensing of Environment*, vol. 16, pp. 36{46, 1984.29

[12] I. Borg and J. Lingoes, *Multidimensional Similarity Structure Analysis*. Springer, 1 ed., July 1987.

[13] K. Fukunaga, *Introduction to Statistical Pattern Recognition*. San Diego: Academic Press, first ed., 1990

[14] A. Miller, *Subset Selection in Regression*. Washington D.C.: Chapman and Hall, 1990.

[15] L. O. Jimenez and D. A. Landgrebe, "Hyperspectral data analysis and supervised feature reduction via projection pursuit," *IEEE Transaction on Geosciences and Remote Sensing*, vol. 37(6), 1999

[16] T.-M. Tu and P.-Y. Chen, "Determination of data dimensionality in hyperspectral Imagery {a noise-adjusted transformed derschgorin disk approach," *Intelligent Data Analysis*, vol. 4, no. 5, pp. 433-444, 2000

[17] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction." *Science*, vol. 290, pp. 2319-2323,December 2000.

[18] I. T. Jolliffe, *Principal Component Analysis*. Springer, October 2002.

[19] D. Gillis, J. H. Bowles, and M. E. Winter, "*Dimensionality reduction in hyperspectral imagery*," vol. 5093, pp. 45-56, SPIE, 2003

[20] P. Bajcsy and P. Groves, "Methodology for hyperspectral band selection [j]," *Photogrammetric Engineering and Remote Sensing Journal*, vol. 70, pp. 793{802, July 2004

[21] L. Alparone, F. Argenti, M. Dionisio, and L. S. [C], "Dimensionality reduction of hyperspectral imagery based on spectral analysis of homogeneous segments: distortion measurements and classi_cation scores," vol. 5238, pp. 226-233, SPIE, 2004.

[22] D. Mladenic, "Feature selection for dimensionality reduction," pp. 84-102, 2006.

[23] S. Wang and C.-I. Chang, "Linearly constrained band selection for hyperspectral imagery," *in Algorithms and Technologies for Multispectral, Hyperspectral, and 30 Ultraspectral Imagery XII. Edited by Shen, Sylvia S.; Lewis, Paul E. Proceedings of the SPIE, Volume 6233, pp. 62332B (2006)., vol. 6233 of Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference,* June 2006.

[24] R. Archibald and G. Fann, \Feature selection and classi_cation of hyperspectral images with support vector machines [j]," *IEEE Geosciences and Remote Sensing Letters* vol. 4, pp. 674-677, October 2007.

[25] Q. Du and N. H. Younan, *Knowledge-Based Intelligent Information and Engineering Systems*, vol. 5179/2008 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg, September 2008.

[26] T. Xiang and S. Gong, "Spectral clustering with eigenvector selection," vol. 41, pp. 1012-1029, March 2008

[27] A. Martinez-Uso, F. Pla, J. M. Sotoca, and P. Garcia-Sevilla, "Clustering-based multispectral band selection using mutual information [c]," *in International Conference on Patern Recognition* (ICPR'06), vol. 2, (Hong-Kong), pp. 760-763, 2006

[28] B. Guo, R. I. Damper, S. R. Gunn, and J. D. B. Nelson, "A fast separability based feature-selection method for high-dimensional remotely sensed image classification, " *Pattern Recognition*. vol. 41, no. 5, pp. 1670-1679, 2008

[29] J. G. Dy and C. E. Brodley, "Feature selection for unsupervised learning," *Journal of Machine Learning Research*, vol. 5, pp. 845{889, August 2004.

[30] M. Dash and H. Liu, "Feature selection for classification," *Intelligent Data Analysis*, vol. 1, pp. 131-156, 1997

[31] S. Muggleton and L. De Raedt, "Inductive logic programming: Theory and methods," *Journal of Logic Programming*, vol. 19/20, pp. 629-679, 1994.

[32] X. Jia and D. A. Richards, "Segmented principal components transformation for efficient hyperspectral remote sensing image display and classification," *IEEE Transaction on Geosciences and Remote Sensing*, vol. 37, pp. 538-542, 1999.

[33] D. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis: An Introduction*. Berlin, Germany: Springer-Verlag, 3rd ed., 1999.

[34] I. Guyon and A. Elissee_, "An introduction to variable and feature selection," Journal of Machine Learning Research, vol. 3, pp. 1157-1182, March 2003.31

[35] S. K. T. El-ghazawi and J. L. Moigne, "Parallel and adaptive reduction of hyperspectral data to intrinsic dimensionality," (Newport Beach, CA), pp. 102-109, 2001. Proceeding, 3rd IEEE International Conference on Cluster Computing CLUSTER

[36] M. Lennon, G. Mercier, M. C. Mouchot, and L. Hubert-Moy, \Independent component analysis as a tool for the dimensionality reduction and the representation of hyperspectral images," *SPIE Remote Sensing*, vol. 4541, pp. 2893-2895, Sept 19-21 2001.

[37] A. Hyvarinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. Wiley-Interscience, May 2001.

[38] T. Lee and D. Landgrebe, \Feature extraction based on decision boundaries," *IEEE Transaction of Pattern Analysis and Machine Intelligence*, vol. 15, pp. 388-400, April 1993.

[39] C. H. Joseph and C.-I. Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach," *IEEE Transaction on Geoscience anf Remote Sensing*, vol. 32, pp. 779-785, July 1994.

[40] A. Hyvarinen, "Fast and robust _xed-point algorithms for independent component

Analysis," *IEEE Transactions on Neural Networks*, vol. 10, no. 3, pp. :626-634, 1999.

[41] C. H. Zheng, D. S. Huang, Z. L. Sun, M. R. Lyu, and T. M. Lok, "Nonnegative independent component analysis based on minimizing mutual information technique," *Neurocomputing*, vol. 69, pp. 878-883, 2006

[42] E. Oja, "Nonlinear PCA criterion and maximum likelihood in independent component analysis," *Proceedings of International Workshop on Independent Component Analysis and Blind Signal Separation* (ICA'99), pp. 143-148, 1990

[43] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Computation*, vol. 7, pp. 1129-1159, 1995.

[44] J. F. Cardoso, "Source separation using higher order moments," *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP'89), pp. 2109-2112, 1989.

[45] J. F. Cardoso, "Infomax and maximum likelihood for source separation," *IEEE Letters on Signal Processing*, vol. 4, pp. 112-114, 1997

[46] A. Hyvarinen, "One-unit contrast functions for independent component analysis: A statistical analysis," *in IEEE Proceedings on Neural Networks for Signal Processing VII*, (Amelia Island, Florida)), pp. 388-397, 1997.

[47] A. Green, M. Berman, P. Switzer, and M. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal,"

IEEE Transactions on Geoscience and Remote Sensing, vol. 26, pp. 65-74, Jan 1988

[48] A. A. Nielsen, "Spectral mixture analysis: Linear and semi-parametric full and iterated partial unmixing in multi- and hyperspectral image data," *International Journal of Computer Vision*, vol. 42, no. 1-2, pp. 17-37, 2001.

[49] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," *Advances in neural information processing*, vol. 13, 2001.

[50] P. Narendra and K. Fukunaga, "A branch and bound algorithm for feature subset selection," *Computers, IEEE Transactions on*, vol. C-26, pp. 917-922, Sept. 1977.

[51] J. Kittler, "Feature set search algorithms," *Pattern Recognition and Signal Processing*, pp. 41-60, 1978.

[52] P. Pudil, J. Novovicova, and J. Kittler, "Floating search methods in feature selection," *Pattern Recognition Letter* vol. 15, no. 11, pp. 1119-1125, 1994

[53] P. Mitra, C. Murthy, and S. Pal, "Unsupervised feature selection using feature similarity," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, pp. 301-312, Mar 2002

[54] M. syan Chen, J. Hun, P. S. Yu, I. T. J, and W. R. Ctr, "Data mining: An overview from database perspective," *IEEE Transactions on Knowledge and Data Engineering*, vol. 8, pp. 866{883, 1996.

[55] A. Jain and D. Zongker, "Feature selection: evaluation, application, and small sample performance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 153-158, Feb 1997.

[56] M. A. Hall, *Correlation-Based Feature Selection for Machine Learning*. PhD thesis, Department of Computer Science, University of Waikato, Hamilton, New Zealand, April 1999.

[57] R. Setiono and H. Liu, "Neural-network feature selector," *IEEE Transactions on Neural Networks*, vol. 8, pp. 654-662, 1997

[58] G. Motta, F. Rizzo, and J. A. Storer, eds., *Hyperspectral data compression* Springer Science, New York, e-book ed., 2006

[59] L. O. Jimenez-Rodriguez, E. Arzuaga-Cruz, and M. Velez-Reyes, "Unsupervised linear feature-extraction methods and their effects in the classification of high-dimensional data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, pp. 469-483, Feb. 2007.

[60] S. Kumar, J. Ghosh, and M. Crawford, "Best-bases feature extraction algorithms for classification of hyperspectral data," *in IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, pp. 1368-1379, July 2001

[61] L. Bruzzone, F. Roli, and S. Serpico, "An extension of the je_reys-matusita distance to multiclass cases for feature selection," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 33, pp. 1318-1321, Nov 1995

[62] B. D. Baofeng Guo, Steve Gunn and J. Nelson, "Adaptive band selection for hyperspectral image fusion using mutual information," *in IEEE 7th International Conference on Information Fusion (FUSION)*, vol. 8, pp. 7803-9286, May 2005.

[63] G. Petrie, P. Heasler, and T. Warner, "Optimal band selection strategies for hyperspectral data sets," *Geoscience and Remote Sensing Symposium Proceedings, 1998. IGARSS '98. 1998 IEEE International*, vol. 3, pp. 1582-1584, July 1998.

[64] P. Groves and P. Bajcsy, "Methodology for hyperspectral band and classification model selection," *IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, pp. 120-128, 2003.

[65] T.-L. S. C-I. Chang, D. Qian and M. Althouse, "A joint band prioritization and band-decorrelation approach to band selection for hyperspectral image classification" *IEEE transactions on geoscience and remote sensing,*" vol. 37, no. 6, pp. 2631-2641, 1999.

[66] M. Velez-Reyes and L. Jimenez, "Subset selection analysis for the reduction of hyperspectral imagery," *IEEE International Geoscience and Remote Sensing Symposium Proceedings*, vol. 3, pp. 1577-1581, 1998

[67] H. Du, H. Qi, X. Wang, R. Ramanath, and W. E. Snyder, "Band selection using independent component analysis for hyperspectral image processing," *in Applied Imagery Pattern Recognition Workshop*, (Los Alamitos, CA, USA), pp. 93-98, IEEE Computer Society, 2003.

[68] Y. Zhang, D. D. Mita, and Z. Junping, "Adaptive subspace decomposition for hyperspectral data dimensionality reduction," *ICIP*, vol. 99, pp. 326-329, 1999

[69] B. Guo, S. R. Gunn, R. I. Damper, and J. D. B. Nelson, "Adaptive band selection for hyperspectral image classification using mutual information,"

[70] J. H. C. Jutten, "Blind separation of sources, part I: an adaptive algorithm based on neuromimetic architecture," *Signal Processing*.

[71] P. Comon, "Independent component analysis, a new concept," 1994

[72] J. Price, "Spectral band selection for visible-near infrared remote sensing: spectralspatial resolution tradeo_s," *IEEE Transactions on Geoscience and Remote Sens- ing*, vol. 35, no. 5, pp. 1277-1285, 1997.

[73] N. Keshava, "Best bands selection for detection in hyperspectral processing," *in IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol.5, pp. 3149-3152, IEEE, IEEE, 2001

[74] D. Qian, "Band selection and its impact on target detection and classification in hyperspectral image analysis," IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data, pp. 374-377, 2003

[75] S. Kaewpijit, J. L. Moigne, and T. El-Ghazawi, "Automatic reduction of hyperspectral imagery using wavelet spectral analysis," *IEEE Transactions on Geo- science and Remote Sensing*, vol. 41, no. 4, pp. 863-871, 2003.

[76] C. Conesea and F. Masellia, "Selection of optimum bands from tm scenes through mutual information analysis," *ISPRS Journal on Photogrammetry and Remote Sensing*, vol. 48

[77] C.-I. Chang, Q. Du, T.-L. Sun, and M. L. G. Althouse, "A joint band prioritization and band-decorrelation approach to band selection for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, pp. 2631-2641, Nov. 1999.

[78] A. M. Us_o, F. Pla, J. M. Sotoca, and P. G.-S. [C] "Comparison of unsupervised band selection methods for hyperspectral imaging," *in IbPRIA* (1), pp. 30-38, 2007.

[79] C.-I. Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Clas- si_cation.* 2003.

[80] N. Keshava and J. Mustard, "Spectral unmixing," *IEEE Signal Processing Maga- zine*, vol. 19, no. 1, pp. 44-57, 2002.

[81] J. M. P. Nascimento and J. M. B. Diasb, "Independent component analysis applied to unmixing hyperspectral data," *in Image and Signal Processing for Remote Sensing* (L. Bruzzone, ed.), vol. 5238, pp. 306-315, SPIE, SPIE, Bellingham, WA, 2004.

[82] J. M. P. Nascimento and J. M. B. Dias, "Vertex component analysis: a fast algorithm to unmix hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 43, pp. 898-910, April 2005.

[83] A. R. Gillespie, M. O. Smith, J. B. Adams, S. C. Willis, A. F. Fisher, and D. E. Sabol, “Interpretation of residual images: Spectral mixture analysis of aviris images, Owens valley, California," vol. 2 of *AVIRIS Workshop*, JPL, 1990.

[84] M. O. Smith, J. B. Adams, and D. E. Sabol, *Spectral mixture analysis-New strategies for the analysis of multispectral data*. Brussels and Luxemburg, Belgium: Kluwer Academic Publishers, 1994.

[85] A. Stocker and A. Schaum, “Application of stochastic mixing models to hyperspectral detection problems," vol. 3071, *SPI*E, April 1997

[86] C. S. Dimitris Manolakis and G. Shaw, “Hyperspectral subpixel target detection using the linear mixing model," *IEEE Transaction on Geoscience and Remote Sensing*, vol. 39, July 2001

[87] C. Chang, X. Zhao, M. L. G. Althouse, and J. J. Pan, “Least squares subspace projection approach to mixed pixel classi_ation for hyperspectral images," *IEEE Transection on Geoscience Remote Sensing*, vol. 36, pp. 898-912, May 1998

[88] C. Chein. I., “Estimation of number of spectrally distnct signal sources in hyperspectral imagery," *IEEE Transection on Geoscience and Remote Sensing*, vol. 42, March 2004

[89] A. Huck and M. Guillaume, “Independent component analysis-based estimation of anomaly abundances in hyperspectral images," pp. 168-177, 2007

[90] D. Manolakis, “Taxonomy of detection algorithms for hyperspectral imaging applications," *IEEE Signal Processing Magazine*, pp. 29-43, January 2002

[91] D. Manolakis, “Taxonomy of detection algorithms for hyperspectral imaging applications," *Optical Engineering,* vol. 44, no. 6, p. 066403, 2005

[92] A. Ifarraguerri and C.-I. Chang, “Hyperspectral image segmentation with convex cones," *Proceedings of ISSSR*, 1997

[93] Winter and M. E., “Fast autonomous spectral end-member determination in hyperspectral data," *in Proceedings of the Thirteenth International Conference on Applied Geologic Remote Sensing*, vol. 11, (Vancouver, BC, Canada), pp. 337-344, 1999.

[94] H. H. Harmon, *Modern Factor Analysis*. Chicago: University of Chicago Press, 1967.

[95] R. B. Singer and T. B. McCord, \Mars: large scale mixing of bright and dark surface materials and implications for analysis of spectral reactance," in 10th conference on Lunar Planet Science, pp. 1835{1848, 1979.

[96] W. Wei and T. Adali, "Detection using correlation bound in a linear mixure model," *Signal Processing*, vol. 87, pp. 1118-1127, 2007.

[97] M. P. Jos_e and M. B.-D. Jos_e, "Blind hyperspectral unmixing," *in Image and Signal Processing for Remote Sensing XIII (Lorenzo, ed.),* vol. 6748, SIE, 2007

[98] B. Hapke, "Bidirectional reactance spectroscopy. I-theory," *Geophysical Research*, vol. 86, pp. 4571{4586, June 1986

[99] S. Kay, *Fundamentals of Statistical Signal Processing* Englewood Cliffs, NJ: Prentice Hall, 1998.

[100] J. B. Adams and M. O. Smith, "Spectral mixture modeling: a new analysis of rock and soil types at the vicking lander site [j]," *Geophysical Research*, vol. 91, no. B8, pp. 8098-8112, 1986

[101] J. C. Harsanyi and C. I. Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach," *IEEE Transaction Geoscience and Remote Sensing*, vol. 32, pp. 779-785, July.

[102] A. Schaum and A. Stocker, "Spectrally-selective target detection," *in SPIE Proceeding on ISSSR*, vol. 3071, 1997

[103] M. Lennon, G. Mercier, M. C. Mouchot, and L. Hubert-moy, "Spectral unmixing of hyperspectral images with the independent component analysis and wavelet packets," *in International Geoscience and Remote Sensing Symposium*, 2001.

[104] M. Prasad, A. Sowmya, and I. Koch, "Efficient feature selection based on independent component analysis," *in Intelligent Sensors, Sensor Networks and Information Processing Conference,* 2004. Proceedings of the 2004, pp. 427-432, 2004

[105] J. Boardman, "Automating spectral unmixing of aviris data using convex geometry concepts," *Summaries 4th Annual JPL Airborne Geoscience Workshop*, vol. 1, pp. 11-4, 1993.

[106] M. D. Craig, "Minimum-volume transforms for remotely sensed data," *IEEE Transaction on Geoscience and Remote Sensing*, vol. 32, no. 1.

[107] C. Bateson, G. Asner, and C.Wessman, "Endmember bundles: A new approach to incorporating endmember variability into spectral mixture analysis," *IEEE Transaction on Geoscience and Remote Sensing*, vol. 38, pp. 1083-1094, March 2000.

[108] X. Jia and D. A. Richards, *Remote Sensing Digital Image Analysis*: An Introduction. Berlin, Germany: Springer-Verlag, 4th ed., 2006.

[109 Ihsan ul Haq and Xiaojian Xu "*A New Approach to Band Clustering and Selection for Hyperspectral Imagery*" IEEE ICSP2008 Proceedings

[110] G. Vane, R. Green, T. Chrien, H. Enmark, E. Hansen, and W. Porter, "The airborne visi-ble/infrared imaging spectrometer (aviris)," *Remote Sensing of the Environment*, no. 44, pp. 127–143, 1993.

[111] [Online], http://speclab.cr.usgs.gov/cuprite.html.

[112] G. Swayze, S. S. R.N. Clark, and A. Gallagher, "Ground-truthing aviris mineral mapping at cuprite, nevada," *in Third Annual JPL Airborne Geosciences Workshop*, pp. 47–49, 1992

[113] C.-I Chang *Hyperspectral Imaging: Techniques for Spectral Detection and classification.* New York: Plenum, 2003

Research Contribution

Muhammad Sohaib, Ihsan ul Haq and Qaiser Mushtaq, "*Dimensional Reduction Of Hyperspectral Image Data Using Band Clustering And Selection Based On Statistical Characteristics Of Band Images* " 2010 The 2nd International Conference on Intelligence and Information Technology (ICIIT2010) co-sponsored by IEEE officially. University of Central Punjab, Lahore, Pakistan. October 28-30, 2010