

Classification and Processing of Geospatial Hyperspectral data

Vanshika Gupta

Department of Civil Engineering, National Institute of Technology Karnataka,
Surathkal 575025

Dr. Dericks P. Shukla

School of Engineering, Indian Institute of Technology, Mandi, Himachal Pradesh
175005

Table of Contents

Abstract

Abbreviations

1 INTRODUCTION

 1.1 Background/Rationale

 1.2 Objectives of the Research

 1.3 Scope

2 LITERATURE REVIEW

 2.1 Information

3 METHODOLOGY

 3.1 Experimental data used

 3.2 Methods

4 RESULTS AND DISCUSSION

 4.1 ENVI processing results

 4.2 Similarity Matrix

 4.3 Determining value of k

 4.4 Classification Results

5 CONCLUSION AND RECOMTMENDATIONS

 5.1 Conclusion

ACKNOWLEDGEMENTS

References

Source

The Academies' Summer Research Fellowship Programme 2018

Classification and Processing of Geospatial Hyperspectral data

Vanshika Gupta

Department of Civil Engineering, National Institute of Technology Karnataka, Surathkal 575025

Dr. Dericks P. Shukla

School of Engineering, Indian Institute of Technology, Mandi, Himachal Pradesh 175005

Abstract

Mapping of snow and glaciers has been carried out globally using various remote sensors and methods such as optical, multispectral and hyperspectral remote sensing techniques (Yasser Hassebo, 2012, etc). Apart from snow mapping, monitoring the snow characteristics is very important and the output is being used for many applications. A few snow cover characteristics such as snow grain size, albedo, specific surface area (SSA) and snow contamination by soot/dust are retrieved using inversion techniques (Jeff Dozier, 1989) from multispectral as well as hyperspectral satellite data. Multispectral remote sensing refers to collection of reflected, emitted, or backscattered energy from an object or area of interest in multiple bands (3-10) of electromagnetic spectrum whereas Hyperspectral remote sensing, also known as imaging spectroscopy, involves capturing data in hundreds of contiguous narrow bands of the electromagnetic spectrum. It allows whole spectral curves to be recorded with individual absorption features, therefore providing information related to surface material that can be exploited to characterize, quantify and perform automated detection of the targets of interest. Unfortunately, in the Indian Himalayan glaciers, not many studies have been carried out using hyperspectral satellite data. Existing studies on the dynamics of Himalayan glaciers (eg. W. F. Manley, 2015, M. Stoffel, 2012) are not yet at par with the remote sensing capabilities of the current generation of satellites and sensors. The major objective of the project is to study the spectral characteristics of snow from the data obtained from hyperspectral AVIRIS sensor, whose applicability is not yet explored for study of snow/glacier characteristics in the Indian context. The process involves initially applying corrections to the data such as atmospheric correction and bad bands removal. Further, we will performing spectral clustering and band classification using existing algorithms (Bo Du, 2015, Jef Caers, 2010) based on different similarity models to reduce data redundancy between channels. The next step is to identify features i.e. mapping of snow using ENVI software with the help of

spectral indices for snow (eg. NDSI) and if possible, to work on developing methods and models for estimating snow/ice parameters.

Keywords: Hyperspectral Remote Sensing, Spectral Characteristics, Himalyan Glaciers

Abbreviations

PCA	Principal Component Analysis
SAM	Spectral Angle Mapper
WSS	Within-cluster-Sum-of-Sqaures
DC	Distribution Compactness
KNN	K-nearest Neighbors
MKNN	Mutual K-nearest Neighbor
FLAASH	Fast Line-of-sight Atmospheric Analysis of Hypercubes
QUAC	quick atmospheric correction

1 INTRODUCTION

1.1 Background/Rationale

Hyperspectral remote sensing, sometimes called as Imaging Spectroscopy, is used to capture data in hundreds of contiguous narrow bands of the electromagnetic spectrum. Having a higher level of spectral detail in hyperspectral images gives better capability to see the unseen. This allows whole spectral curve to be recorded with individual absorption features, thus providing information related to surface material. This information can be exploited to characterize, quantify and perform automated detection of the targets of interest with much better accuracy than multi-spectral or RGB image (Dutra and Mascarenhas 1984). But its high dimensionality adds to a level of complexity as well as demands high storage. Also one of the major drawbacks of hyper-spectral imagery is its high redundancy due to data recorded at bandwidth as small as 5nm. This means that the image contains multiple bands carrying almost the same information thus resulting in high computational cost and requiring huge amount of data storage for feature extraction or any further processing of the image. It is therefore important to reduce the dimensionality of our data and make it computationally optimized without losing considerable information. Hyperspectral and multispectral images have many real world applications. For example, hyperspectral imagery has been used to map

invasive species and help in mineral exploration and feature characteristic identification. We also use it in the field of agriculture, ecology, oil and gas, oceanography and atmospheric studies.

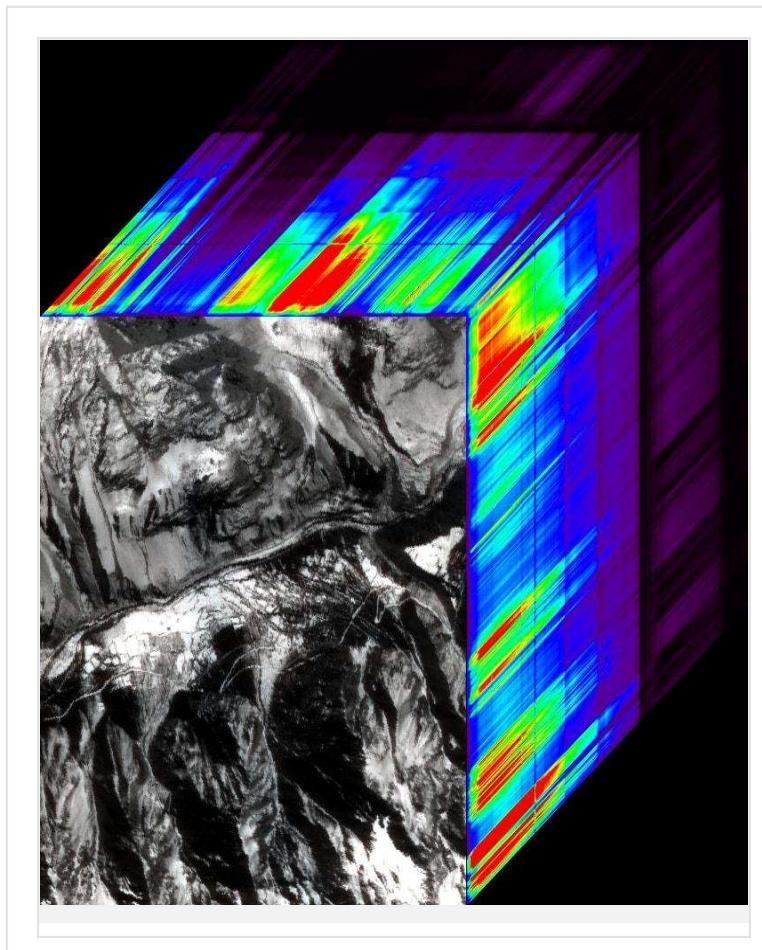


Fig 1 Hyperspectral 3-D cube of NASA AVIRIS-Ng satellite data of Patseo region in the Western Himalayas

1.2 Objectives of the Research

1.2.1 Overall objective

Our objective for the project is to classify the image and extract useful information from the given hyperspectral data. In this, we aim to firstly process the provided dataset by applying corrections. For minimising the redundancy in the information carried by the bands due to its high spectral resolution, we use spectral clustering techniques and at last, classification of the

processed image to identify features of interest. This also involves identifying the pure pixels in the image. The overall objectives can be summarised as below:

- Radiometric Correction for reducing the effects caused by the various governing factors such as atmospheric absorption, water retrieval, etc that can reduce the efficiency of our further processing.
- Spectral Clustering using various clustering techniques to find out the optimum number of bands suitable for representing the image without redundancy in information.
- Band Selection after performing spectral clustering to specify which bands to choose. This is done by using clustering algorithms like k-means and choosing the closest bands to the centroid. We choose 10-20 bands from hundreds of bands present in the image.
- Image Classification using classifiers such as support vector machine and spectral angle mapper and obtaining results for the amount of information losses due to dimensionality reduction. Also, to obtain the accuracy of the classifier using the original image containing all bands to compare with the image obtained after band selection.

1.3 Scope

There are many fields of research which specifically relate to each step involved in this process. This is a topic of ongoing research and is still developing to inculcate and produce better methods. Also, there is much scope in this as it combines civil engineering and machine learning techniques. It is also used in the field of agriculture, ecology, oil and gas, oceanography and atmospheric studies.

2 LITERATURE REVIEW

2.1 Information

Snow mapping from geospatial data has been carried out using various techniques by several people and is a source of extracting valuable information regarding identification of resources and features of interest. Softwares like ENVI and ArcGIS have been used, however, open-source softwares like QGIS may also be used. Snow mapping involves many image processing techniques for optimisation of the involved process. Processing techniques like Atmospheric Correction using FLAASH(Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes).and QUAC(QUick Atmospheric Correction), PCA (principal component analysis), PPI (pixel purity index), Minimum Noise Fraction transform, Match Filtering, Band Math, Spectral indices, etc. (<https://www.harrisgeospatial.com/docs>) have been used previously to extract and comprehend the information contained in the hundreds of bands in the image.

Methods for reducing the data redundancy due to these bands has been carried out using various techniques that fall under spectral clustering. It relates to methods derived from machine learning. Various methods and algorithms have been proposed For e.g. Elbow plot, Davies Bouldin Index, Rand Index, Jaccard Index, Naeini et al 2014, Huband et al 2005 etc. However, we make use of internal measures like Silhouette plot Rousseeuw 1987 , Calinski-Harabasz index, Calinski and Harabasz 1974 etc. to get the optimum number of clusters. The accuracy of these indices depends on the data characteristics. After evaluating the optimal number of bands for our dataset, the final bands need to be chosen. For this purpose, several clustering algorithms have been developed for e.g K-means, Fuzzy clustering Techniques, Hierarchical clustering, etc. K-means is the most simplified and generic algorithm but has several drawbacks. These drawbacks can be overcome by other clustering techniques depending on the data characteristics and requirement. Classification for feature extraction is the sole aim and the main application of all the processing done on the image and improving the classification accuracy is the main objective of any processing performed technique. Several classifiers have been developed previously for e.g. Support Vector Machine (SVM), Spectral Angle Mapper (SAM), ISODATA classification, etc. The most popular among them is the SVM and we will be using it to classify our image after band selection. Each step involved in this process is in itself a research problem which researchers have been trying to solve by developing the best possible methods. However, we will be using some selective methods based on their previous performances and credibility.

3 METHODOLOGY

3.1 Experimental data used

The test data sets used are the openly available standard hyperspectral datasets. The first is the Indian Pines Site 3. This scene was gathered by AVIRIS sensor over the Indian Pines test site in North-western Indiana which consists of 145×145 pixels and 224 spectral reflectance bands in the wavelength range $0.4\text{--}2.5 \times 10^{-6}$ meters. This scene is a subset of a larger one. The Indian Pines scene contains two-thirds agriculture, and one-third forest or other natural perennial vegetation. There are two major dual lane highways, a rail line, as well as some low density housing, other built structures, and smaller roads. Since the scene is taken in June some of the crops present, corn, soybeans, are in early stages of growth with less than 5% coverage. The ground truth available is designated into sixteen classes and is not all mutually exclusive. (http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes#Indian_Pines)

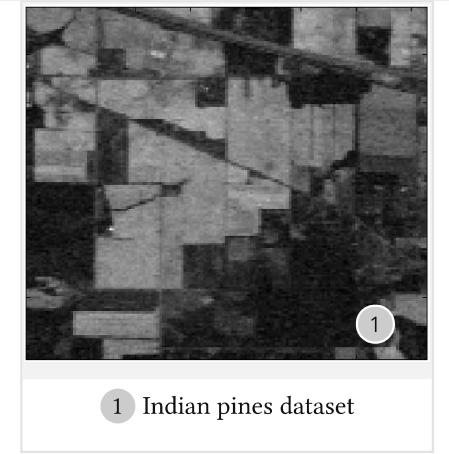


Fig 2 Indian Pines Grayscale image

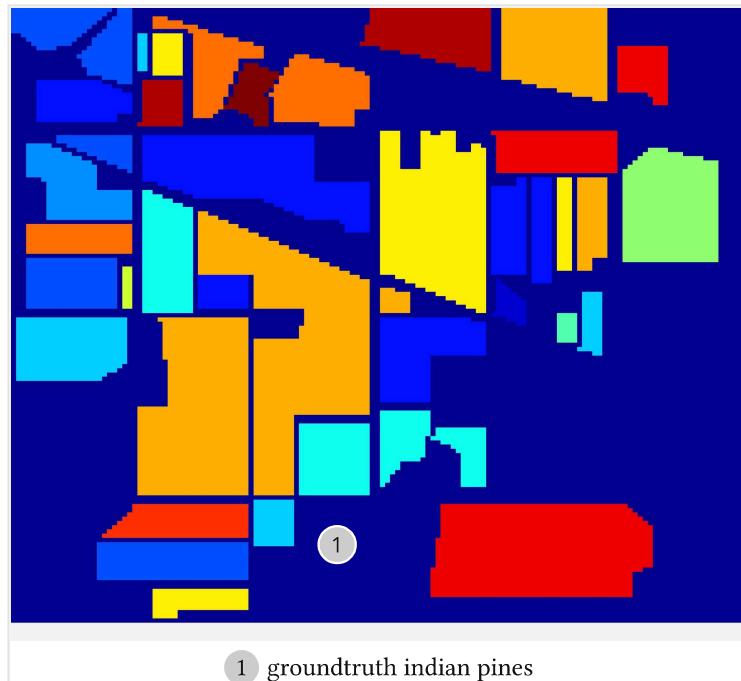


Fig 3 Indian pines groundtruth

The second is the Pavia University dataset. These are two scenes acquired by the ROSIS sensor during a flight campaign over Pavia, nothern Italy. The number of spectral bands is 103. Pavia University is 610*610 pixels, but some of the samples in both images contain no information

and have to be discarded before the analysis. The geometric resolution is 1.3 meters. Both image groundtruths differentiate 9 classes each. The discarded samples in the figures can be seen as abroad black strips. Pavia scenes were provided by Prof Paolo Gamba from the Telcommunications and Remote Sensing Laboratory, Pavia University (Italy).(http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes#Indian_Pines)



Fig 4 Pavia University

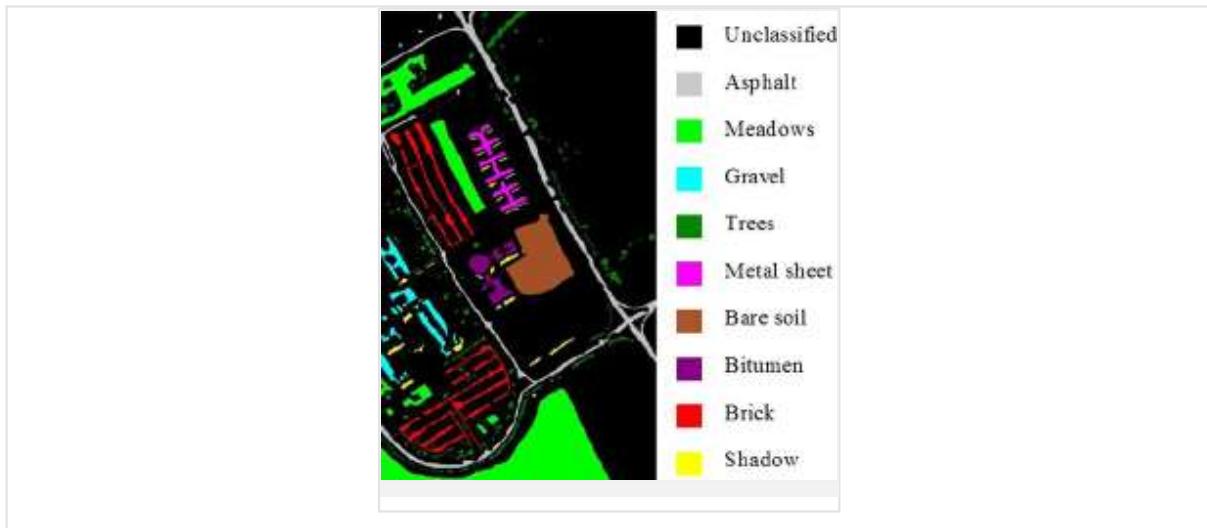


Fig 5 Enter your caption here

For processing on ENVI, we have used a spatial subset of the AVIRIS-NG sensor acquired dataset from NASA of the himalyan region extending from 32.701102° N to 32.719238° S and from 77.155547° W to 77.182222 E°

3.2 Methods

3.2.1 Corrections using ENVI

We apply corrections on the image using ENVI proprietary software

Fast Line-of-sight Atmospheric Analysis of Hypercubes(FLAASH)

FLAASH is a first-principles atmospheric correction tool that corrects wavelengths in the visible regions through near-infrared and shortwave infrared regions, up to 3 μm. Unlike many other atmospheric correction programs that interpolate radiation transfer properties from a pre-calculated database of modeling results, FLAASH uses the MODTRAN radiation transfer code. FLAASH has the following features:

- Correction for the mixing of pixel due to scattering of surface-reflected radiance.
- An option to compute a scene-average visibility (aerosol/haze amount).

- Cirrus and opaque cloud classification map
- Adjustable spectral polishing for artifact suppression.

QUick Atmospheric Correction (QUAC)

QUAC is an atmospheric correction method for multispectral and hyperspectral imagery that works with the visible and near-infrared through shortwave infrared (VNIR-SWIR) wavelength range. It determines atmospheric correction parameters directly from the observed pixel spectra in a scene, without ancillary information. QUAC is based on the empirical finding that the average reflectance of diverse material spectra (excluding highly structured materials such as vegetation, shallow water, and mud) is not dependent on each scene. So processing is much faster compared to first-principles methods. QUAC also allows for any view angle or solar elevation angle. QUAC contains the following enhancements to improve the accuracy of atmospheric correction

- Selects endmembers based on a small subset of available bands for most sensors. When a sensor spans both the visible and NIR-SWIR spectral regions, the algorithm excludes bands in the visible region.
- Constrains the gain curve to be constant for wavelengths below 650 nm.
- Suppresses the effects of dense vegetation.
- Removes cloud endmembers for hyperspectral sensors with 940 to 1020 nm water absorption bands.

Pixel Purity Index (PPI)

PPI is used to find the most spectrally pure (extreme) pixels in multispectral and hyperspectral images. These correspond to mixing endmembers. The PPI is computed by repeatedly projecting n-D scatter plots on a random unit vector. ENVI records the extreme pixels in each projection (those pixels that fall onto the ends of the unit vector) and it notes the total number of times each pixel is marked as extreme(pure). A Pixel Purity Image is created where each pixel value corresponds to the number of times that pixel was recorded as extreme.

Principal Component Analysis (PCA)

PCA is used to produce uncorrelated output bands, to segregate noise components, and to reduce the dimensionality of data sets. Because multispectral data bands are often highly correlated, the principal components (PC) transformation is used to produce uncorrelated output bands. This is done by finding a new set of orthogonal axes that have their origin at the data mean and that are rotated so the data variance is maximized.

PC bands are linear combinations of the original spectral bands and are uncorrelated. You can calculate the same number of output PC bands as input spectral bands. The first PC band contains the largest percentage of data variance and the second PC band contains the second largest data variance, and so on. The last PC bands appear noisy because they contain very little variance, much of which is due to noise in the original spectral data. principal components bands produce more colorful color composite images than spectral color composite images because the data is uncorrelated.

Minimum Noise Fraction (MNF) Transform

MNF transform is used to determine the inherent dimensionality of image data, to segregate noise in the data, and to reduce the computational requirements for subsequent processing. The MNF transform as modified from [Green et al 1988](#) and implemented in ENVI, is a linear transformation that consists of the following separate principal components analysis rotations:

- The first rotation uses the principal components of the noise covariance matrix to decorrelate and rescale the noise in the data (a process known as *noise whitening*), resulting in transformed data in which the noise has unit variance and no band-to-band correlations.
- The second rotation uses the principal components derived from the original image data after they have been noise-whitened by the first rotation and rescaled by the noise standard deviation. Since further spectral processing will occur, the inherent dimensionality of the data is determined by examining the final eigenvalues and the associated images. You can divide the data space into two parts: one part associated with large eigenvalues and coherent eigenimages, and a complementary part with near-unity eigenvalues and noise-dominated images. Using only the coherent portions separates the noise from the data, thus improving spectral processing results.

Bad Band Removal

Bad bands are those bands in the image which contain very little or no information. The sensor could not take information in these wavelength ranges and hence contains null or NaN data values. It is necessary to remove such bands from the image before any further processing of the image as it may reduce the efficiency of the process and add unnecessary complexities. The wavelength range of the bad bands is dependent on the sensor used. One of the methods for detecting bad bands is through visual inspection. It is a time taking process as it involves analysing the grayscale image of all the bands one by one for an image with no information i.e the grayscale image will be black. However, it can also be done using MATLAB code.

All the above listed methods are used for removing bands containing noise, applying different types of corrections and removing bands of no absorption. After all this has been done, we further need to reduce the data redundancy of the image by selecting only those few bands

which are most necessary to represent the image in an informative way without losing considerable information. This is where the concept of Band Selection or Spectral Clustering comes into picture. Now our task is to find the optimum number of bands k .

3.2.2 Similarity Measures

We use the Euclidean distance Ryotaro Kamimura, 2003 and Tanimoto coefficient, as the similarity measure to form the similarity matrix.

$$d_{ij} = \sqrt{S_i - S_j} \quad (1)$$

where S_i denotes the band in an image.

$$T_{ij} = \frac{S_i \cdot S_j}{\|S_i\|^2 + \|S_j\|^2 - S_i \cdot S_j} \quad (2)$$

3.2.3 Similarity Graphs

A nice way of representing the data is in form of the similarity graph $G = (V, E)$. The vertices v_i in this graph represent the data points x_i . Two vertices are connected if the similarity S_{ij} and the edge is weighted by s_{ij} . The problem of clustering can now be reformulated using the similarity graph: we want to find a partition of the graph such that the edges between different groups have a very low weight (which means that points in different clusters are dissimilar from each other) and the edges within a group have high weight (which means that points within the same cluster are similar to each other). It can also be called as a weighted adjacency matrix. Ulrike von Luxburg, 2007

Let $G = (V, E)$ be an undirected graph with vertex set $V = \{v_1, \dots, v_n\}$. In the following we assume that the graph G is weighted, that is each edge between two vertices v_i and v_j carries a non-negative weight. The weighted adjacency matrix of the graph is the matrix $W = (w_{ij})$ $i, j = 1, \dots, n$. If $w_{ij} = 0$ this means that the vertices v_i and v_j are not connected. As G is undirected we require $w_{ij} = w_{ji}$. Ulrike von Luxburg, 2007

3.2.4 Types of Similarity Graphs

The similarity graphs can be constructed by any of the following methods Ulrike von Luxburg, 2007

Epsilon-neighborhood graph: We connect all points whose pairwise distances are smaller than epsilon. This neighborhood graph is usually considered as an unweighted graph.

K-nearest neighborhood Graph: Here the goal is to connect vertex v_i with vertex v_j if v_j is among the k nearest neighbors of v_i . However this leads to a directed graph. To make it undirected, we put a mutual or normalized condition on the graph.

Fully Connected Graph: In this graph, all vertices having positive weight are connected, with edges weighted by W_{ij} . This type of graph lacks the sparse structure we require. The most commonly used function is the Gaussian similarity function

$$W(i, j) = e^{-\frac{d_{ij}^2}{2\sigma}} \quad (3)$$

3.2.5 Graph Laplacian

Different graph Laplacians are exist and have different properties. We will carefully distinguish between different variants of graph Laplacians. Bojan Mohar, 1997

Unnormalized Laplacian W. Freeman, 1998

$$L = D - W \quad (4)$$

where

$$d_i = \sum_{j=1}^N W_{ij} \quad (5)$$

Normalized Laplacian J. Malik, 2000

$$L_{sym} = D^{-1/2} L D^{-1/2} \quad (6)$$

3.2.6 Spectral Clustering Algorithm

The Ng–Jordan–Weiss(NJW) algorithm for spectral clustering is used as follows

1. Construct a similarity graph. Let W be its weighted adjacency matrix..
2. Compute the unnormalized Laplacian L .
3. Compute the first k eigenvectors v_1, \dots, v_k of the generalized eigenproblem $Lv = Dv$.
4. Form u by normalizing each row of v .
5. Cluster the points (y_i) $i=1,\dots,n$ in R_k with the k -means algorithm into clusters C_1, \dots, C_k .

3.2.7 k -means

k -means is one of the simplest unsupervised learning algorithms that solves the well known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed apriori. The main idea is to define k centers, one for each cluster. These centers should be placed in a cunning way because different location causes different result. So, the better choice is to place them as much as possible far away from each other. The next step is to take each point belonging to a given data set and associate it to the nearest center. When no point is pending, the first step is completed and an early group age is done. At this point we need to re-calculate k new centroids as barycenter of the clusters resulting from the previous step. After we have these k new centroids, a new binding has to be done between the same data set points and the nearest newcenter. A loop has been generated. As a result of this loop we may notice that the k centers change their location step by step until no more changes are done or in other words centers do not move any more. Finally, this algorithm aims at minimizing an objective function know as squared error function given by:

$$J(V) = \sum_{i=1}^c \sum_{j=1}^{c_i} (\|x_i - v_j\|)^2$$

Fig 6 Error Function

3.2.8 Finding the optimum number of clusters k

Distribution Compactness (DC) Algorithm

The DC algorithm is a measure of compactness. Yenming Mark Lai, 2015 Compactness is the measure of how close the objects within the same cluster are to each other. It uses a nonparametric estimate of the probability density function of dataset and determines k by analyzing the DC of sparse representations of band vectors in the kernel space. The DC measurement is defined as follows :

$$DC = \sum_{i=1}^N \lambda_i \{1_N^T v_i\}^2 \quad (7)$$

Elbow Method

Given that we have formed i clusters, we find the sum of distance of the centroid and the data points in a cluster corresponding to i clusters and then plot it for varying i . The value of i may be taken as $\{1, \dots, 30\} \in N$. The point where an elbow occurs depicts the approximate range of k . This is regarded as the elbow method. However, this method is not very accurate and works mostly for data groups distant far apart or relatively simple datasets.

Calinski-Harabasz Index

This is based on maximising the Variance Criteria Ratio which is defined as Calinski and Harabasz, 1974

$$CH = \frac{BSS}{k-1} / \frac{WSS}{N-k} \quad (8)$$

This index is plotted for varying i from 1 to 30 and the point where the global maxima occurs is the best value for k .

Davies-Bouldin Index

The davies-bouldin index is defined as: Donald W. Bouldin, 1979

$$DB = \frac{1}{N_c} \sum_{i=1}^{N_c} R_i \quad (9)$$

where R_i is the measure of scatter of the datapoints. This index is plotted for $i=1$ to 30 and the global minima is located. The value of i for which the minima occurs is the value of the optimum number of clusters k .

Dunn's Index

It is an internal evaluation scheme like the above mentioned indices, where the result is based on the clustered data itself. Yuelong Zhu, 2014 As do all other such indices, the aim is to identify sets of clusters that are compact, with a small variance between members of the cluster, and well separated, where the means of different clusters are sufficiently far apart, as compared to the within cluster variance. For a given k , a higher Dunn index indicates better clustering

$$DI = \frac{\min_{1 \leq i < j \leq k} d(C_i, C_j)}{\max_{1 \leq c \leq k} ||N_c||} \quad (10)$$

Silhouette Index

The technique provides a succinct graphical representation of how well each object lies within its cluster. The silhouette value is a measure of how similar an object is to its own cluster (cohesion) compared to other clusters (separation). The silhouette ranges from -1 to +1, where a high value indicates that the object is well matched to its own cluster and poorly matched to neighboring clusters. If most objects have a high value, then the clustering configuration is appropriate. If many points have a low or negative value, then the clustering configuration may have too many or too few clusters. Rousseeuw 1987

$$SI_i = \frac{b_i - a_i}{\max(a_i, b_i)} \quad (11)$$

where a_i is the average distance from the i^{th} point to the other points in its cluster, and b_i is the minimum (over all clusters) of the average distance from the i^{th} point to the points in another cluster.

When we have obtained the value of k using these methods, we will apply the Spectral clustering algorithm described previously and find out the k bands closest to the k centroids respectively, obtained after applying k -means clustering on the normalized $N \times k$ eigenvector,

treating each row as a data point in k-dimensional space. After spectrally subsetting the hyperspectral image for the obtained bands, we need to define the classification accuracy of our selection. For this, we use different classifiers as described below.

3.2.9 Classification Techniques

Support Vector Machine

Support Vector Machine is a supervised machine-learning algorithm that is used for both classification and regression. In this method, we perform classification by finding the right hyper-plane that best differentiates two classes. It is mostly used for binary classifications. Support vectors are simply the coordinates of individual observations. Choosing the right hyperactive plane is also very important in Support Vector Machine. Let us be given a pair of (x,y) coordinates, outputs if it's either *red* or *blue*. A support vector machine takes these data points and outputs the hyperplane (which in two dimensions it's simply a line) that best separates the tags. This line is called the *decision boundary*. Anything that falls to one side of it will be classified as blue, and the rest as red.

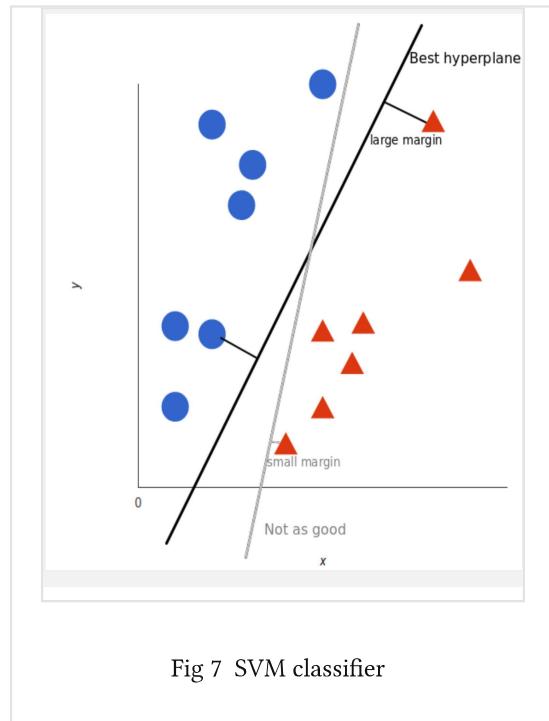
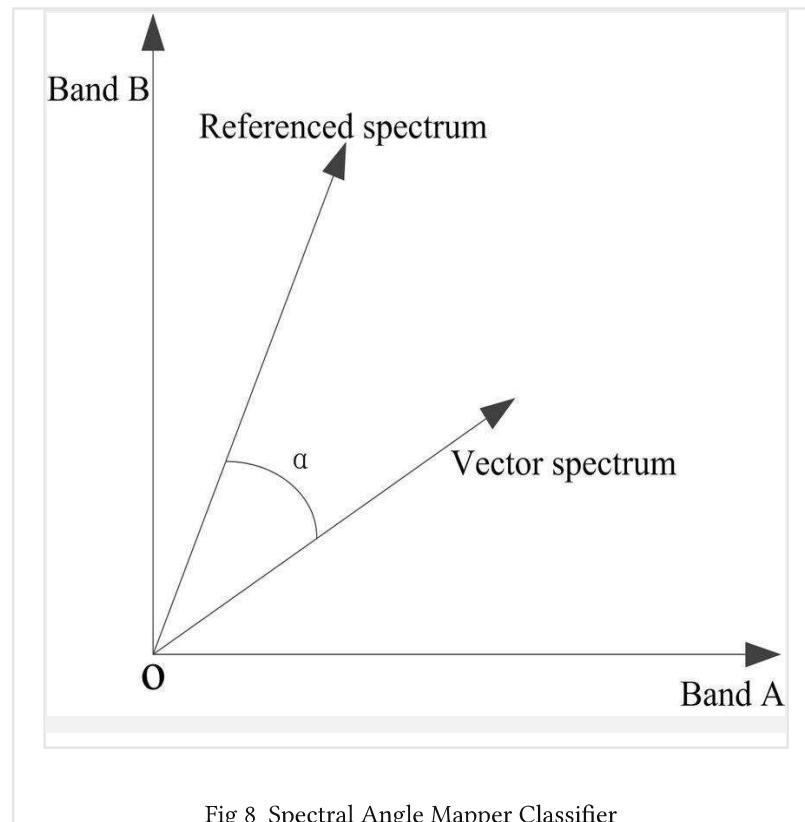


Fig 7 SVM classifier

For SVM, the best hyperplane is the one that maximizes the margins from both tags. In other words: the hyperplane whose distance to the nearest element of each tag is the largest.

Spectral Angle Mapper (SAM)

Spectral Angle Mapper (SAM) is a physically-based spectral classification that uses an n -D angle to match pixels to reference spectra. The algorithm determines the spectral similarity between two spectra by calculating the angle between the spectra and treating them as vectors in a space with dimensionality equal to the number of bands. SAM compares the angle between the endmember spectrum vector and each pixel vector in n -D space. Smaller angles represent closer matches to the reference spectrum. Pixels further away than the specified maximum angle threshold in radians are not classified.



4 RESULTS AND DISCUSSION

4.1 ENVI processing results



Fig 9 MNF result(Band1)

The MNF result shows the noise amount. The bands are arranged in the decreasing order of their eigenvalues and we selected first 20 bands which contain most of the useful information. One can improve the spectral processing results by keeping only the coherent portions/bands. The PCA is also a band reduction technique that is used to select only uncorrelated bands and reduce spectral dimensionality. It is basically an inbuilt band reduction technique in ENVI, however in this we need to choose how many bands are to be selected. We specify the number of output bands and the algorithms searches for those number of most uncorrelated bands from the original image using the correlation matrix.



Fig 10 PCA output band 1

The PPI result shows an image consisting of one band where the intensity of a pixel is proportional to the number of times it was rescored as pure. It is used for identification of pure pixels.



Fig 11 PPI result

4.2 Similarity Matrix

The similarity matrix is constructed using the similarity metrics described in the methodology. Each band is considered as a data point of dimension $D \times 1$ where D is the number of pixels contained in a band, between which the similarity has to be computed. We form a $N \times N$ similarity matrix where N is the number of bands where $(i,j)^{\text{th}}$ cell depicts the measure of similarity between bands i and j . Therefore it is a symmetric matrix.

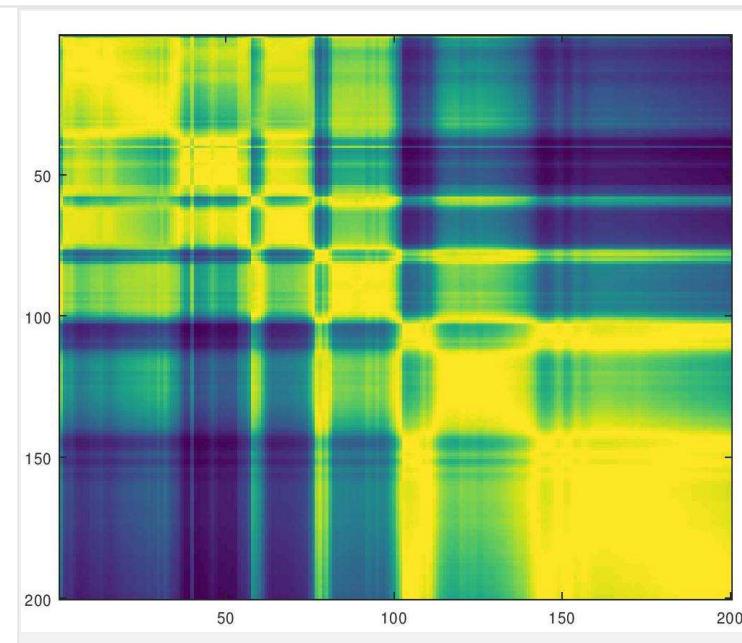


Fig 12 Similarity matrix for Tanimoto similarity on Indian Pines dataset

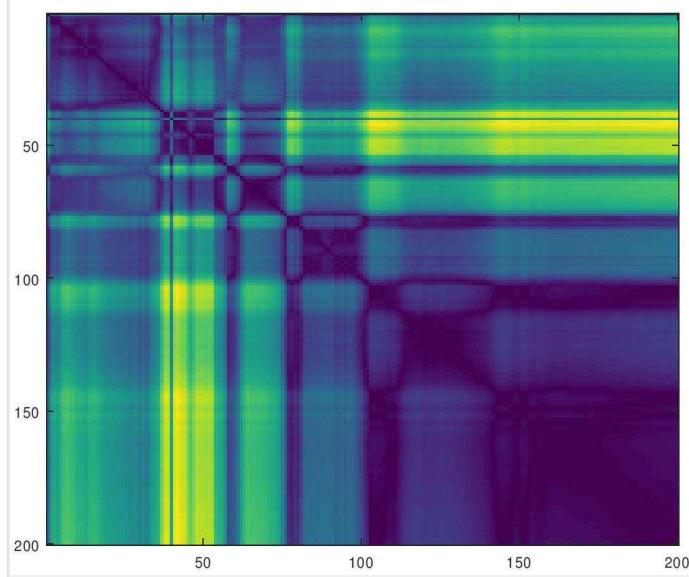


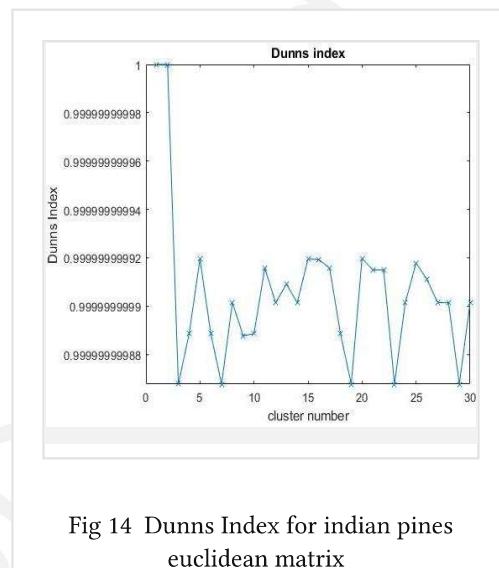
Fig 13 Euclidean matrix for Indian Pines

The matrices shown above are formed using the euclidean and the tanimoto metric and represented in the form an image matrix of dimension $N \times N$ where N represents the number of bands. The indian pines data set consists of 200 bands.

After converting the matrix to similarity graph using KNN algorithm, we get a sparse symmetric adjacency matrix. The symmetric Laplacian matrix is formed which has a block diagonal structure. Further, eigenvalue decomposition is performed on the Laplacian matrix and the corresponding eigenvectors and eigenvalues are computed.

4.3 Determining value of k

To determine value of optimum number of clusters k, we apply the methods listed in the methodology and hence, obtain the following results.



The dunn's index shows a global maxima at k=15. The Davies-Bouldin index shows a global minima at k=23 however, it also shows a local initial minima at k=15.

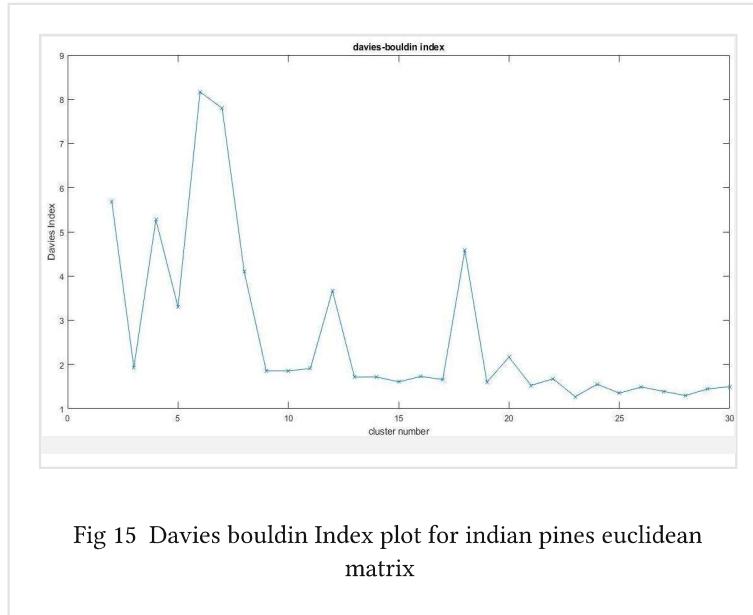


Fig 15 Davies bouldin Index plot for indian pines euclidean matrix

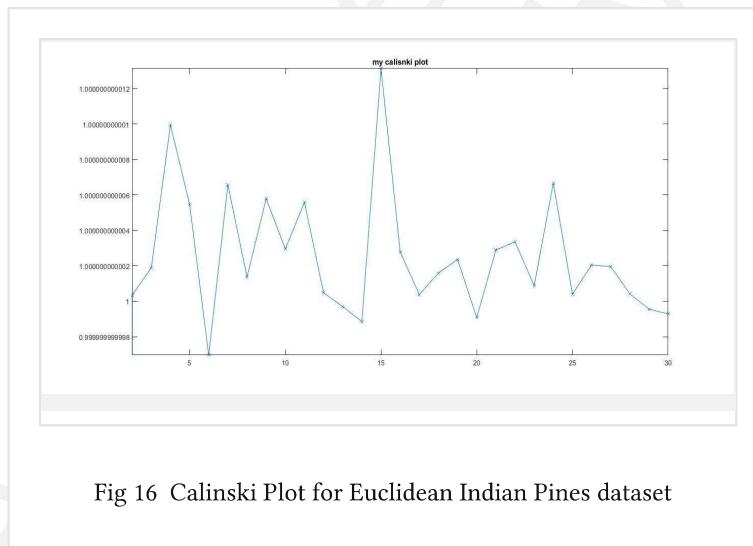


Fig 16 Calinski Plot for Euclidean Indian Pines dataset

The Calinski plot as shown in Fig 13 depicts a clear global maxima at k=15. The different values that we obtain from these three indices are tested for the true value using the silhouette plot for each of the indicated number of clusters. The value of k for which we get the best silhouette plot is hence chosen as the exact value for the number of clusters.

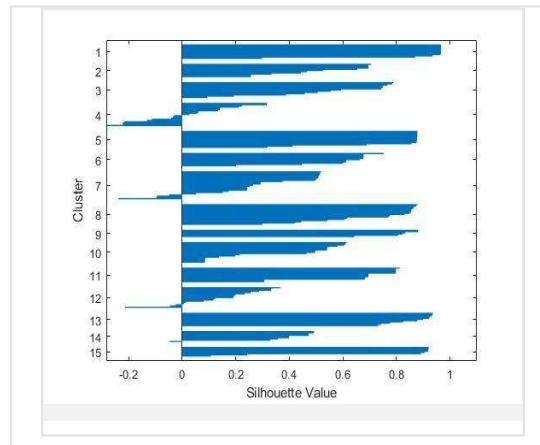


Fig 17 Silhouette Index for Indian Pines
Euclidean Matrix for k=15

The results shown above are the various indices used to indicate the optimum value of k. The results indicate 15 as the value of k. Similar process can be performed for different similarity measures on different datasets. Some of the results are summarised in the given table.

4.4 Classification Results

The classification when done using support vector machine classifier in ENVI gave the following results for the spectral subset of image using tanimoto measure of similarity.

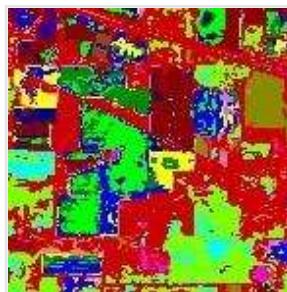


Fig 18 Classified image of Indian Pines using Euclidean Matrix.

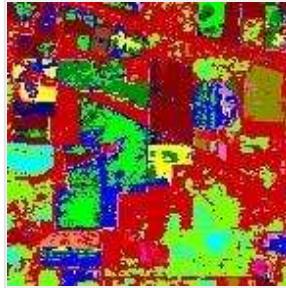


Fig 19 Classified image using Tanimoto coefficient for indian pines

Table 1 Comparative Results for number of clusters on Indian Pines data set

Result Name	Euclidean Matrix	Tanimoto Matrix
Calinski Index	15	17
Dunn's Index	15	16
Silhouette Index	15	17
Davies-Bouldin Index	23	15
Classification Accuracy	75.56%	75.21%

5 CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

Band Selection using Spectral Clustering is an important part of hyperspectral image processing. It allows more information to be stored in the same space, which is much needed when each image file size varies in Gigabytes. We conclude through this paper that accuracy of classification or any other processing changes by a very minimal amount even after selecting only 15 or 17 bands out of the 200 bands. How to select bands and which bands to select is the main process. Here, we reviewed various processing methods developed to make this process more and more efficient based on the idea of a 'good' clustering. Depending on the data set characteristics, some indices and methods have poor performance while some give more accurate results.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my guide Dr. Dericks Praise Shukla for guiding and encouraging me during the course of my fellowship in Indian Institute of Technology, Mandi while working on the project on "Classification and Processing of Geospatial data". The dataset used was provided under the project ID:"IITM/SAC-ISRO/DPS/164" titled "Snow mapping and it's parameter estimation from geospatial(AVIRIS-NG) and field data", funded by SAC-ISRO to Dr. Dericks P. Shukla as Project Investigator. I would like to thank Mr. Sharad Kumar Gupta for helping me carrying out this project smoothly and mentoring me throughout.I would like to sincerely thank the coordinator of Summer Research Fellowship 2018, Mr. CS Ravi Kumar for giving me this opportunity.

References

1. LUCIANO V. DUTRA, NELSON D. A. MASCARENHAS, 1984, Some experiments with spatial feature extraction methods in multispectral classification, *International Journal of Remote Sensing*, Vol. 5, no. 2, pp. 303-313
2. Amin Alizadeh Naeini, Mohammad Saadatseresht, Saeid Homayouni, 2014, Automatic Estimation of Number of Clusters in Hyperspectral Imagery, *Photogrammetric Engineering & Remote Sensing*, Vol. 80, no. 7, pp. 619-626
3. Jacalyn M. Huband, James C. Bezdek, Richard J. Hathaway, 2005, bigVAT: Visual assessment of cluster tendency for large data sets, *Pattern Recognition*, Vol. 38, no. 11, pp. 1875-1886
4. Peter J. Rousseeuw, 1987, Silhouettes: A graphical aid to the interpretation and validation of cluster analysis, *Journal of Computational and Applied Mathematics*, Vol. 20, pp. 53-65
5. T. Calinski, J. Harabasz, 1974. A Dendrite Method for Cluster Analysis, *Communications in Statistics - Simulation and Computation*, Vol. 3, no. 1, pp. 1-27
6. A.A. Green, M. Berman, P. Switzer, M.D. Craig, 1988. 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, no. 1, pp. 65-74
7. Ryotaro Kamimura, 2003, Information-Theoretic Competitive Learning with Inverse

- Euclidean Distance Output Units, Neural Processing Letters, vol. 18, no. 3, pp. 163-204
8. Ulrike von Luxburg, 2007, A tutorial on spectral clustering, Statistics and Computing, vol. 17, no. 4, pp. 395-416
9. Bojan Mohar, 1997, Some applications of Laplace eigenvalues of graphs, Graph Symmetry, pp. 225-275
10. P. Perona, W. Freeman, 1998, A factorization approach to grouping, Computer Vision – ECCV'98, Lecture Notes in Computer Science, pp. 655-670
11. Jianbo Shi, J. Malik, 2000, Normalized cuts and image segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 8, pp. 888-905
12. Weiwei Sun, Liangpei Zhang, Bo Du, Weiyue Li, Yenming Mark Lai, 2015, Band Selection Using Improved Sparse Subspace Clustering for Hyperspectral Imagery Classification, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 8, no. 6, pp. 2784-2797
13. David L. Davies, Donald W. Bouldin, 1979, A Cluster Separation Measure, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-1, no. 2, pp. 224-227
14. Shijin Li, Jianbin Qiu, Xinxin Yang, Huan Liu, Dingsheng Wan, Yuelong Zhu, 2014, A novel approach to hyperspectral band selection based on spectral shape similarity analysis and fast branch and bound search, Engineering Applications of Artificial Intelligence, vol. 27, pp. 241-250

Source

1. Fig 8: https://www.researchgate.net/figure/The-diagram-of-Spectral-Angle-Mapper-SAM_fig1_311359199