International Journal of Intelligent Information Technologies
Volume 18 • Issue 2

A Novel Approach for Band Selection Using Virtual Dimensionality Estimate and Principal Component Analysis for Satellite Image Classification

Smriti Sehgal, Amity University, Noida, India*
Laxmi Ahuja, Amity Institute of Information Technology, Amity University, Noida, India
M. Hima Bindu, North Orissa University, India

ABSTRACT

Images, being around us in every aspect of life, have become an emerging field of research. Extensive image analysis has been done on binary as well as coloured images, which has led various researchers to explore images having deep spectral knowledge about a particular area of interest. High resolution images, having more than three spectral bands, capture minute details of an object in various spectral bands resulting in high computational complexity. In this paper, the authors have tried to reduce the complexity of multispectral image by selecting only the relevant bands need to reconstruct an image. Traditional principal component analysis technique is used for band selection of true color bands and classification-assessed results of both the images; original and dimensionality reduced images are compared using partitioning clustering technique. Experimental results show that compressed image after reduction of bands by PCA yields better classification results than the original image.

KEYWORDS

Band Selection, Dimension Reduction, Median Filtering Technique, Multispectral Images, Nonlinear Filters, Principal Component Analysis

INTRODUCTION

In many application areas where visual representation in form of images is involved, the use of spectral information is necessary to perform various tasks such as image enhancement, segmentation, and classification. Spatial and Spectral information is used to analyze the presence of chemical patterns in the high-resolution image to provide a better evaluation of features. The use of spectral correlation among various features forms a basis for the need to select relevant bands from multispectral images having a pool of hundreds of bands. In this paper, the study of high-resolution images such as multispectral and hyperspectral images is illustrated with an utter need to select bands to reduce their dimensionality. The band Selection process is then evaluated with classification technique and is explained in detail in further sections.

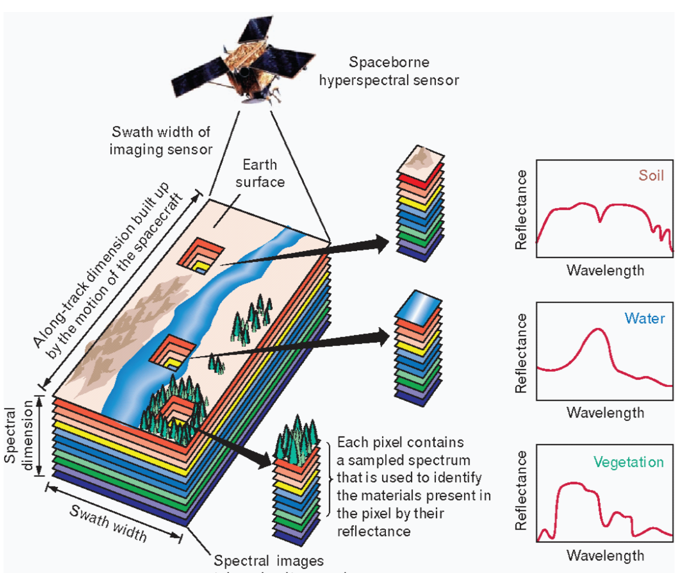
DOI: 10.4018/IJIIT.296272

*Corresponding Author

Taking into consideration of simple image, represented by a 2-d matrix of any object in a scene; Images can be analog or digital (J Kuruvilla et. al, 2016). All aerial photographs are the type of analog images whereas latter images are acquired by electronic sensors of different wavelengths and are stored on the computer system in digital form. Other classifications for the type of image are binary images, colored images, multispectral images, hyperspectral images, and super spectral images. Binary images have the intensity values as 0 or 1, having only 1 band whereas colored or RGB images have intensity values based on the number of bits each pixel holds and has only 3 bands. High-Resolution Images such as multispectral, hyperspectral, and super spectral images are the ones having more than three bands. These images are used to capture minute details of a particular object in a scene across various spectral bands. Here, remote sensing image as a digital image is used in the research work. A Remote Sensing image is an image that represents part of the Earth's surface. Each pixel in an image shows some area on Earth and has an average intensity value that measures solar radiance in a wavelength band reflected from the ground. These images are generally multi-layered images that are constructed by stacking the images taken for the same area by the same or different sensors. Multispectral image (Sotoca, J. M. et. al, 2007) is a special case of a multi-layered image consisting of a few image layers of a particular scene. Each image layer is taken at a particular wavelength depending upon capturing sensor. The image used in this paper is AVIRIS image consisting of several bands in blue, green, red, near-IR, SWIR & thermal bands. Another set of images having hundreds of bands, known as hyperspectral images, is also available which has deep spectral knowledge about an object in the scene, enabling better identification and classification. Currently, this paper is focused on the analysis of high resolution images and their classification (De Backer, S. et. al, 2005).

In figure 1, the concept of spectral imaginary is shown. An airborne sensor simultaneously samples multiple spectral wavebands over a large area in a ground scene. After appropriate processing, the resultant image contains a sample of spectral reflectance measurement which is interpreted to identify the object present. It presents the spectral variation in reflectance for soil, water, and air. Multispectral images are used to measure light in the electromagnetic spectrum, it is different from hyperspectral bands in the way that latter images consist of several bands which increase the complexity of bands. These types of images are used for space imaging and likewise for reporting in painting and for investigation purposes.

Figure 1. Spectral Imaging Concept (Source: https://www.markelowitz.com/Hyperspectral.html)



Due to the advancements in sensor technology, high-end sensors collect data in hundreds of narrow bands. An increase in spectral bands leads to high computational complexity in terms of network transmission and data storage. Large data volume also renders the common image processing techniques and the "curse of dimensionality" problem comes into the picture. With an increase in dimension and fixed training samples and parameters, classification accuracy decreases. High spectral images consume a large amount of space due to the greater number of bands that form a single image. These images store the information of specific wavelengths within the electromagnetic spectrum range. Due to high spectral imaging constraints, denoising of the image is done using various filtering techniques as an important pre-processing step to sharpen the image.

Keeping limited availability of resources in mind, and for its effective utilization, dimension reduction in form of band selection is suggested. Dimension reduction can be achieved using band selection or feature extraction. Band Selection (Nakamura, R. Y. et. al, 2014) (Sun, W. et. al, 2019) is a technique for dimension reduction in which higher dimension data is mapped onto lower dimension data with minimum loss and keeping original features intact. Here, relevant bands according to the problem are chosen and redundant and irrelevant bands are discarded. Feature Extraction techniques (Rodarmel, C., & Shan, J., 2002), also known as transformation-based techniques, combine a subset of original features to generate new features. Various techniques have been used in past for dimension reduction such as PCA, Genetic Algorithm (GA), Discriminant Analysis, Firefly Algorithm, Wavelet-based algorithm (Karami A. et. al, 2012), and many optimization problems (Bajcsy, P. et. al, 2004) (Zhang, B. et. al, 2011). In this paper, band selection in remotely sensed data by the means of PCA is addressed, which converts the number of correlated variables (possible large number of subsets) into several uncorrelated features (small number of subsets) that are called principal components. It is also used for reducing the dimensions of the image that is used to minimize the large sets of variables into smaller variables such that small variables still contain a large amount of information. The principal component analysis can be measured as the rotation of the axes of the original variable coordinate to the new system variable. It does the rotation on the orthogonal axis called as principal axis. This technique relies on the fact that the neighboring bands are correlated in their spectral and textural form in the multispectral image and used to convey similar information about the objects. It employs statistical properties of selecting bands to determine band dependency. The number of principal components to be selected is decided by the Virtual Dimensionality Technique (Ghamary A., et. al, 2016). It provides the number of spectrally distinct bands from hundreds of original bands that give sufficiently large relevant data. Results of PCA are compared with the wavelet-based algorithm, that generates a new reduced combined feature set. It applies Discrete Wavelet Transform (DWT) on the spectral domain of each pixel and removes high-frequency signals that may contain useful information for class separation and identification

In this paper, Satellite Image Classification techniques are applied to original and dimensionality reduced images to measure the accuracy, which is a complex technique grouped into two forms mainly described as Supervised classification technique and Unsupervised classification technique. In the supervised classification technique (Bajcsy, P. et. al, 2004), the algorithm makes use of training examples or training sets. The training sets being used are ground information or ground truth value. The spectral signature is also used to search for the same signatures in the pixel area in the remaining pixel of the image. Expert training help is available in the selection of training sets and biased selection process. Whereas, in an untrained technique, also known as unsupervised techniques (Bajcsy, P. et. al, 2004), where no teacher or knowledge of training information is provided in a selection of training sets or training examples, the natural way of clustering pixels i.e. grey pixels values is chosen. After this, the threshold is defined to limit the class information or sets of classes in the training data. One of the most popular methods, K-mean clustering which comes under untrained classification is used to access both the images.

The remainder of this paper is organized as follows: In Section II, background and related work regarding the satellite image analysis and its processing is discussed. The next section gives a detailed view of the proposed methodology and techniques used. Section IV shows experimental results in image acquisition, filtering for removing noise in image pre-processing, deciding on the number of bands to select, band selection, and classification and its assessment. Finally, the discussion, conclusion, and future aspects are explained in section V.

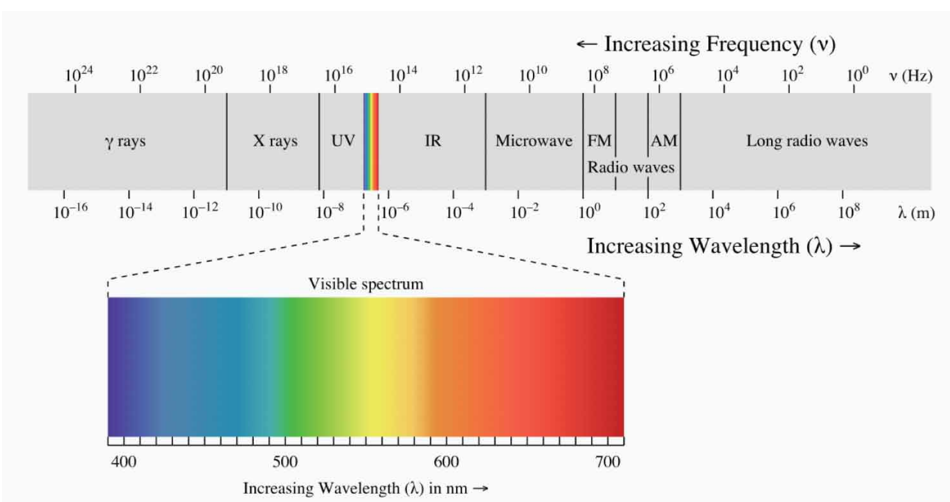
BACKGROUND AND RELATED WORK

The human eye can see colors having a wavelength ranging from 400 nm to 700 nm as shown in figure 2 i.e. colors that come in the visible range of the EM spectrum. A light that is outside the visible range cannot be perceived by the human eye but by other sensors. Capturing an image of the scene from the camera is a physical process. The sensor of the camera uses natural sunlight as the source of energy. Whenever the sunlight falls on any object, the amount of light/energy reflected is recorded and a continuous voltage signal is generated. Sampling and quantization are two processes needed to convert this data into digital form by generating a 2d matrix that is stored on the computer system.

In contrary to ordinary colored images consisting of three-color bands: Red, Green, Blue in the visible region as shown in figure 2, remote sensing images (RSI) are composed of various spectral bands at different wavelengths of EM Spectrum. These images are identified by spectral, spatial, radiometric, and temporal resolution. Spectral Resolution is the capability of the sensor to measure wavelength intervals of electromagnetic radiation. Spatial Resolution is the measure of the smallest object that can be captured by the sensor. Radiometric Resolution differentiates between various energy signals. It determines the number of bits required by each pixel. Temporal Resolution is the amount of time required by the satellite to revisit the same area for recording new energy signals. RSI is stored on the computer system in form of n 2-d matrices, where n is the number of spectral layers. Each spectral layer is known as a band in RSI. The volume of data stored in digital form for RSI is potentially large as the same area is captured at different wavelength bands. Different bands associated with RSI as multispectral and hyperspectral are well explained in (J Kuruvilla et. al, 2016).

Considering the large volume of data in hand and given the task of object identification and classification of the remotely sensed image, deep analysis is required. The curse of dimensionality problem comes into the picture with this. A major challenge is how to reduce dimensions to achieve the goal. Dimension Reduction in hyperspectral imaging is categorized in two forms: transformation-

Figure 2. Electromagnetic Spectrum (Source: wikimedia.org)



based (Koonsanit, K. et.al, 2011) or feature selection approaches (Li, F. et. al, 2016) (De Backer, S. et. al, 2005) (Koonsanit, K. et. al, 2011).

Transformation-based techniques for dimension reduction find the new combination of features from the given set of features but in reduced dimension. Further, approaches in transform-based (Bajcsy, P. et. al, 2004) (Zhang, B. et. al, 2011) are linear or non-linear. Linear methods involve linearly projection of the original data onto a low-dimensional space, whereas latter methods deal with non-linear data only. A few popular linear transformation-based approaches are PCA, Factor Analysis, Wavelet-based, and Linear Discriminant Analysis. The most widely used approach is PCA (Bajcsy, P. et. al, 2004) (X. Wang and M. Leeser, 2017) as it generates principal components independent of each other i.e. there is no correlation among each component. Reducing the number of features also resolves the problem of overfitting, in turn improving algorithm performance. In addition, its focus is on global optimization rather than local achievement.

Principal Component Analysis (Russ, J. C., 2016) is the procedure that is applied to hyperspectral and multispectral remotely sensed data which converts the number of correlated variables into the number of uncorrelated variables i.e. a small number of subsets. It is also referred to as a dimensionality reduction tool that is used to reduce large sets into smaller variables such that it still contains a large amount of relevant information. The first component in this technique contains the variability in data and further, the succeeding component accounts for more variability in data as much as possible. Instead of throwing redundant information, it condensed the information into inter correlated data into a few variables called principal components. The principal component analysis (Bajcsy, P. et. al, 2004) can also be measured as the rotation of the axes of the original variable coordinate to the new system variable. It has rotation on the orthogonal axis called as principal axis. The principal component analysis is based on the fact that the neighboring bands are highly correlated in the multispectral image and are used to convey the same information about the objects. It employs statistical properties of the band selection to examine bands dependency or correlation.

Various names are given to PCA (Koonsanit K. et. al, 2012) such as hoteling transformation or Karhunen Loeve-transformation, but all are based on the same concept of eigenvalue decomposition. Consider the image pixel vector as a vector, x0, a set of N d-dim samples:

$$X_0 = x_1, x_2, x_3 \ldots\ldots x_N$$

The dimension of the image vector is equal to the number of spectral bands. It illustrates that sum of squared distances between x_0 and any x_k must be as small as possible. Mean vector of all image vectors is calculated as:

$$m = \frac{1}{M}\sum_{i=1}^{M}\left[x_1, x_2, x_3 \ldots x_N\right] \tag{1}$$

The covariance matrix of x is calculated as in eq. 2 where E is the expectation operator:

$$\operatorname{cov}(x) = E\left\{\left(x - E\left(x\right)\right)\left(x - E\left(x\right)\right)^T\right\} \tag{2}$$

Eigen Value Decomposition is written as:

$$C_x = ADA^T \tag{3}$$

where D = diagonal matrix consisting of eigenvalues and A = orthogonal matrix having N dim eigenvectors. Linear transformation to be formed for PCA is defined as:

$$y_i = A^T x_i \tag{4}$$

These pixel vector forms combined feature bands of the original image. To compute the approximation of the original image, first, k eigenvectors arranged in descending order are used. The first band is also known as the first principal component holds the highest contrast or variance. It decreases as we create further principal components. Thus, the first k principal components contain the most relevant information about the target object. Another simple yet powerful technique, the Wavelet-Based transformation feature extraction technique is explained in (Karami A. et. al, 2012) for analysis of satellite images. It is a pixel-based transformation by applying DWT on each spectral band and results in the division of sub-images that are un-correlated.

Feature Selection methods such as Backward elimination, Forward Selection, and Random forests (Li, F. et. al, 2016) (De Backer, S. et. al, 2005) (Koonsanit, K. et. al, 2011) only keep the most important features and discard the rest. These methods (Hambal, A. M. et. al, 2017) find some kind of correlation among the bands to find the similarity. These methods can be supervised or unsupervised. When the target object information is known, supervised techniques are applied to find the best class separability. Other various powerful tools for global optimization techniques are nature-inspired algorithms (Markham, B. L. et. al, 2004) in the area of swarm computing. Additionally, in (Goel, L., 2010), the author has compared and contrasted particle swarm optimization with support vector machine-based classification to evaluate its performance. The accuracy of Support Vector Machines (SVM) is higher in PSO (Senthilnath, J. et. al, 2011) (Goel, L., 2010) based selection bands rather than using all the original bands. Likewise, the enhancement for SVM precision can bring out significantly more huge enhancement in classifier combination. The technique in (Li, F. et. al, 2016) applied is supervised feature selection in reducing dimensions to get the relevant feature through the labeled training set.

In hyperspectral imaging, there are large sets of samples that make it difficult to label among a large number of bands, thus the author uses the unsupervised component determination strategy to choose the most proper relevant groups from a lot of test pictures. The information content between phantom groups is dissected identified with the proposed methodology in (Sotoca, J. M. et. al, 2007) that tries to minimize the dependent information and maximize the entropies of the bands been selected. Unsupervised techniques in (Ghamary A. et. al, 2016) are applied whenever the information about an object is unknown. It generally clusters the data into groups having high inter-cluster similarity based upon the similarity index discussed in (Ghamary A. et. al, 2016). In (Yang, H. et. al, 2012), the author explores another nature-inspired algorithm as Glowworm Swarm Optimization (GSO) clustering calculation for various path coverage of remotely sensed images. In unsupervised arrangement strategies, the programmed age of groups to characterize a colossal database isn't abused to their maximum capacity. The proposed procedure in (Yang, H. et. al, 2012) scans for the most ideal number of bunches and its middle utilizing GSO. Another statistical method used for dimension reduction is sparse matrix representation in (De Backer, S. et. al, 2005) in which the orthogonal matching pursuit algorithm is used.

PROPOSED METHODOLOGY

In the proposed work, RSI is analyzed and read into the system using MutliSpec Tool. Noise Removal is done using filtering methods to generate the sharpened image. It removes bands having a low Signal-to-Noise Ratio (SNR) to discard noisy bands. As these images consist of hundreds of spectral bands, all bands cannot be loaded at a time. Increased volume of data increases computational complexity, transmission costs, and storage capacity. To overcome this issue, dimensionality reduction techniques are explored. The number of distinct bands holding relevant information is determined using the Virtual Dimensionality (VD) estimate (Ghamary A., et. al, 2016) and is denoted as there. Next, which bands to select, is decided by a transformation-based algorithm, known as PCA. PCA is well known and explained in detail in literature and has many advantages over other techniques such as it is simple and avoids the problem of overfitting. It generates principal components (PC) by rotating the axis orthogonally to map the maximum number of data points on it. The angle at which PC has the maximum number of points, has the largest variance among variables, and becomes the first PC or first transformed feature. In this paper, PCA works by first constructing a matrix that summarizes how variables are related to each other. Further, this matrix is divided into two components: magnitude and direction. The original multispectral dataset is aligned to these directions and decides which directions are "important". Data along with the directions that are not important are dropped, considering they are irrelevant. It gives n principal components in decreasing order of variance. First t principal components are selected that have the most relevant information about the target and reconstruct an image with new feature bands. This t is the number of bands to be selected and is determined using a VD estimate. With selected bands, a reduced image is constructed and classified to access the selection process. The wavelet-Based technique is used to compare the results generated by PCA. Image with new combined feature set is classified using a simple unsupervised k-means clustering algorithm. The main aim of RSI classification is to classify each pixel into one of the predefined class/groups. The final output class set is denoted as:

$$Y = \left\{c_1, c_2, c_3 \ldots c_m\right\} \tag{5}$$

and input feature band set as:

$$X = \left\{x_1, x_2, x_3 \ldots x_n\right\} \tag{6}$$

Selected feature band set after applying PCA is denoted as:

$$B = \left\{b_1, b_2, b_3 \ldots b_n\right\} \tag{7}$$

If $b_i = 1$, the respective feature band is selected else it is discarded.

k-means is an unsupervised clustering algorithm that classifies each data point in one of the classes based upon the sum of the squared error concept. K-means is chosen for its reasonably accurate estimation of k clusters of each class. Classification accuracy assessment measures like Overall accuracy (OA), Kappa coefficients (K), and Average accuracy (AA) are used to validate the performance of classification with and without dimensionality reduction. These measures are calculated from the error matrix which is generated using the original class label and predicted class label.

Figure 3 shows the flow graph of the proposed technique.

EXPERIMENTAL RESULTS

Image Acquisition

Image, as shown in figure 4, used for experimenting with the algorithm was collected by the Earth Observation Satellite Company (EOSCT), commercial operator of AVIRIS 4 and AVIRIS 5. The

Figure 3. Flow graph of proposed methodology

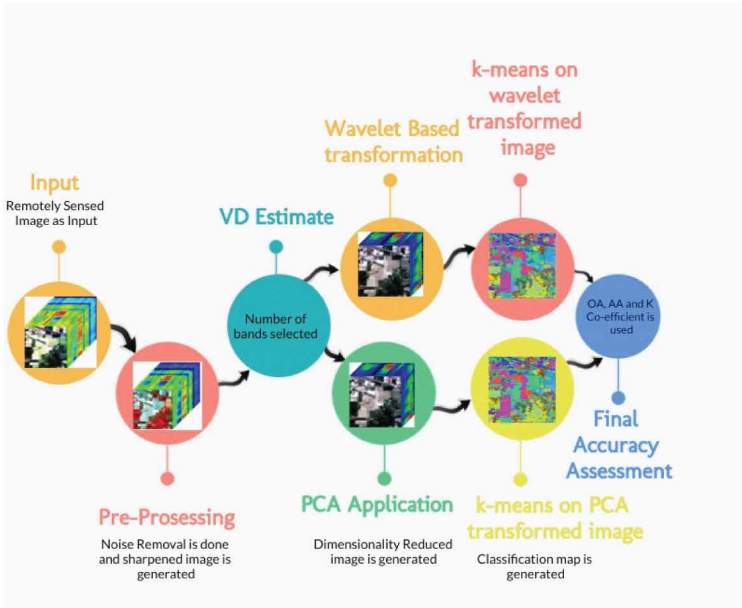
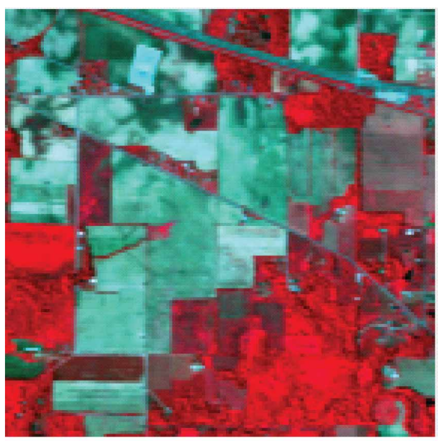


Figure 4. Original image with band 50 as Red, band 27 as Green, and band 17 as Blue



size of the image is 145 x 145 in 220 spectral bands. It was taken on June 12, 1992, AVIRIS image Indian Pine Test Site 3 (2 x 2-mile portion of Northwest Tippecanoe County, Indiana). The repeat cycle of the satellite was recorded as 16 days.

Noise Removal

The image filtering technique is used in this paper to remove noise and clarify the image. In multispectral AVIRIS image, noise is visible as random spots of the wrong color as it tends to increase with temperature and exposure times. The image filtering (Rodarmel, C. et. al, 2002) used here tends to separate and remove noise from image data. The original image of size 145 * 145 is taken and salt and pepper noise is added to it. This type of noise consists of random pixels being set to black or white (the extremes of the data range). Here, medfilt2 did a better job of removing noise, with less blurring of edges.

Now after adding noise to the image, the image is smoothened to sharpen it and improve its quality. Here, the median filtering technique is used to sharpen the image quality as it is easy to implement and is also used for minimizing the intensity variation between the two pixels. In this, pixel values are replaced with median values instead of the mean value. Figure 5 shows the images of before and after noise removal and steps to sharpen the image are stated as follows:

Step 1: Select a two-dimensional window W of size 3*3 pixels. Assume that the pixel being processed is C(x, y).
Step 2: Compute W-med, the median of the pixel values in window W.
Step 3: Replace C(x, y) with W-med.
Step 4: Repeat steps 1 to 3 until all the pixels in the entire image are processed.

Selecting the Number of Bands

Estimating the number of spectrally distinct bands in RSI is done using the concept of VD (Ghamary A. et. al, 2016). It aims at finding distinct bands based on their independence. In this paper, three Neyman Pearson detection theory-based methods namely Harsanyi-Farrand-Chang (HFC), the noise-whitened HFC (NWHFC), and the noise subspace projection (NSP) are used to develop the VD estimate that is summarized in table 1. Here, NSP gave the largest estimate and is used for further steps.

Band Selection

VD estimate using NSP criteria gave near to an optimal number of bands to be chosen. As the probability of false alarm goes to 10-5 and 10-6, the number of bands to be chosen remains unchanged. Taking 16 as the number of bands to be selected, the next question is which 16 bands to select? The answer to this question is solved by PCA (Russ, J. C., 2016) [19].

Figure 5. Noise removal (a) noisy image (b) sharpened image

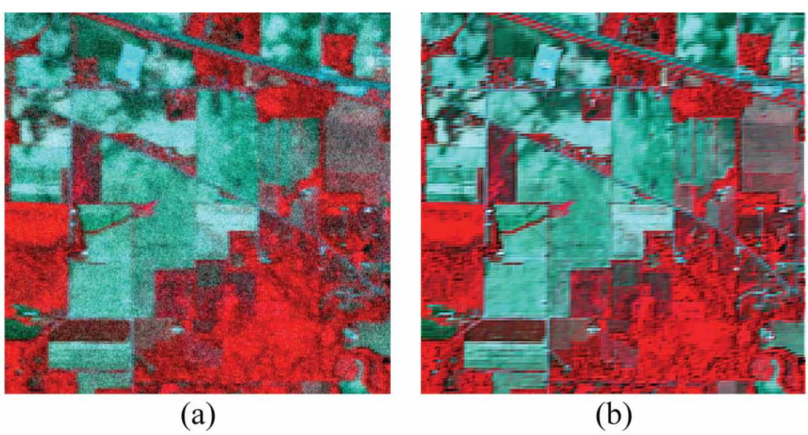


Table 1. VD Estimate for Original Pine Image containing 220 bands

pf (prob of false alarm)	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
VD Estimate	79	60	41	28	16	16

AVIRIS image chosen with all bands is given as an input to the PCA and Wavelet-based feature extraction algorithm. PCA tries to rotate the axis to cover maximum data points i.e. maximum coverage is expected. After generating many principal components by combining original features and arranging them in decreasing order of covariance, the top 16 principal components in form of 16 bands are chosen. Results of PCA are compared with other feature extraction techniques called wavelet-based transformation. In this, DWT is applied to generate sub-images of the original image to reduce redundancy. Table 2 shows the bands selected after applying PCA and Wavelet Transformation.

Classification

Classification accuracies for a given dataset are determined using the clustering technique to evaluate the performance of the model built. K-means clustering is used to classify the original and dimensionality reduced remotely sensed image. Results are shown in figure 6, based on the classification error matrix generated in which the X and Y axis are class labels.

Entry to a particular element is the number of pixels belonging to the Xth class but classified as Yth class. The X-axis denotes original ground classes and Y-axis denotes class predicted. An ideal error matrix for correctly classified classes would be a diagonal matrix. Table 3, Table 4, and Table 5 show the error matrix for the average efficiency of the original image and compressed images using k-means.

Error matrix is generated not using all pixels due to dimensionality constraints and for this, only a few reference pixels from each of the classes are taken and classification accuracy is accessed. OA, AA, and Kappa coefficient are calculated for both the images using error matrix by formulas as shown below and tabulated in table 5:

$$OA = \frac{\left(TP + TN\right)}{\left(TP + TN + FP + FN\right)} \times 100 \tag{8}$$

Table 2. 16 relevant bands selected after PCA and Wavelet Transformation

Method	16 Selected Band numbers (from 220 bands)
PCA	32, 124, 10, 25, 78, 211, 189, 100, 201, 66, 167, 105, 52, 96, 199, 220
Wavelet Based Transformation	10, 25, 31, 50, 64, 77, 95, 100, 105, 122, 167, 190, 195, 200, 211, 220

Figure 6. (a) Original Image using 220 bands; (b) Clustered Image having 16 bands selected using PCA and k-means; (c) Clustered Image having 16 bands using Wavelet Transformation and k-means; and (d) Class Labels

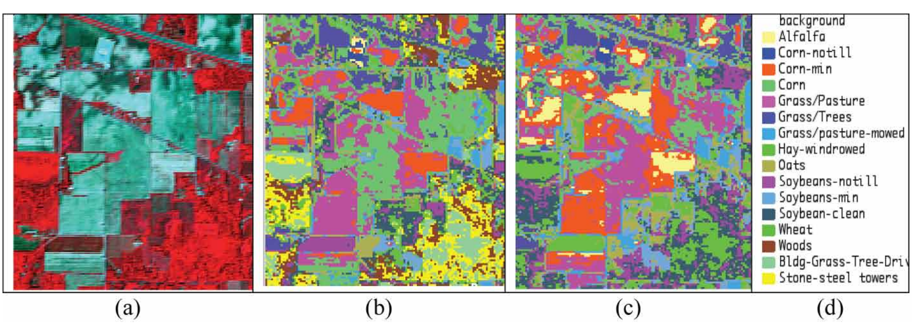


Table 3. Error Matrix for Original Image Classification using k-means

Feature	Alfafa	Corn-notill	Corn-min	Corn	Grass/Pasture	Grass/Trees	Grass/pasture-mowed	Hay	Oats	Soybeans-notill	Soybeans-min	Soybean-clean	Wheat	Woods	Bldg-Grass-Trees-Drives	Stone-steel towers	Total Reference Points
Alfafa	67	5	4	-	-	3	-	-	-	5	-	2	-	-	4	-	90
Corn-notill	2	280	11	9	1	-	-	3	4	-	3	-	-	2	-	9	324
Corn-min	-	5	90	-	-	5	3	8	-	-	-	-	3	4	2	1	121
Corn	9	-	3	62	5	5	-	-	4	-	2	2	6	2	-	-	100
Grass/ Pasture	-	-	3	3	99	5	-	4	-	-	3	3	-	-	-	-	120
Grass/Trees	10	5	-	-	-	201	6	-	14	-	-	13	5	12	7	7	280
Grass/ pasture-mowed	6	8	9	8	-	-	444	9	2	-	-	3	-	-	5	6	500
Hay	-	-	-	2	-	3	1	51	-	-	1	1	-	-	1	1	60
Oats	-	-	-	-	3	-	3	-	12	2	-	-	-	-	-	-	20
Soybeans-notill	5	5	-	2	2	3	-	3	-	160	4	4	2	-	-	-	190
Soybeans-min	-	-	2	-	2	3	4	-	-	-	8	1	-	-	-	-	20
Soybean-clean	4	7	-	-	5	-	3	2	-	1	-	101	8	3	2	-	135
Wheat	-	-	8	2	-	3	-	4	-	-	3	6	22	5	-	2	55
Woods	10	3	-	4	-	3	-	8	-	5	-	3	13	139	10	3	20
Bldg-Grass-Trees-Drives	-	-	-	-	5	5	-	-	-	-	-	3	3	7	67	-	90
Stone-steel towers	3	-	-	-	-	5	-	-	2	4	-	-	-	-	-	21	35

where TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative.

$$AA = \frac{\left(FAR + FRR\right)}{2} \tag{9}$$

where FAR = False Accept Rate and FRR = False Reject Rate and are defined as follows:

$$FAR = \frac{FP}{TP} \times 100 \tag{10}$$

$$FRR = \frac{FN}{TN} \times 100 \tag{11}$$

Table 4. Error Matrix for Reduced Image with 16 bands using PCA and k-means

Feature	Alfafa	Corn-notill	Corn-min	Corn	Grass/Pasture	Grass/Trees	Grass/pasture-mowed	Hay	Oats	Soybeans-notill	Soybeans-min	Soybean-clean	Wheat	Woods	Bldg-Grass-Trees-Drives	Stone-steel towers	Total Reference Points
Alfafa	87	3	1	-	-	1	-	-	-	-	-	-	-	-	-	-	90
Corn-notill	2	300	1	4	1	-	-	3	3	-	1	-	-	2	-	7	324
Corn-min	-	1	110	-	-	3	3	4	-	-	-	-	-	-	-	-	121
Corn	7	-	3	70	1	1	-	-	1	-	2	2	4	1	-	-	100
Grass/ Pasture	-	-	1	1	108	3	-	1	-	-	1	3	-	-	-	-	120
Grass/Trees	5	3	-	-	-	223	3	-	12	-	-	10	3	10	4	7	280
Grass/ pasture-mowed	3	5	5	7	-	-	460	6	2	-	-	2	-	-	5	6	500
Hay	-	-	-	1	-	2	-	55	-	-	1	1	-	-	1	1	60
Oats	-	-	-	-	2	-	2	-	16	-	-	-	-	-	-	-	20
Soybeans-notill	3	2	-	1	2	-	-	-	-	175	3	1	-	-	-	-	190
Soybeans-min	-	-	2	-	2	1	2	-	-	-	12	1	-	-	-	-	20
Soybean-clean	2	5	-	-	2	-	1	1	-	-	-	115	5	1	2	-	135
Wheat	-	-	5	1	-	1	-	1	-	-	1	6	35	3	-	2	55
Woods	5	1	-	1	-	1	-	3	-	1	-	1	5	180	3	-	20
Bldg-Grass-Trees-Drives	-	-	-	-	-	-	-	-	-	-	-	-	-	2	88	-	90
Stone-steel towers	1	-	-	-	-	2	-	-	1	3	-	-	-	-	-	28	35

$$k = \frac{N\sum_{i=1}^{r} x_n - \sum_{i=1}^{r}\left(x_r \cdot r_c\right)}{N^2 - \sum_{i=1}^{r}\left(x_r \cdot r_c\right)} \tag{12}$$

where r = number of rows in error matrix, x_r = total of r^{th} row, xc = total of c^{th} column, N = total number of samples.

As it can be seen from table 6, that when k-means is applied on the three images generated, dimensionality reduced image using PCA gave better results in terms of accuracy as well as Execution Time taken. In addition to improved accuracy percentage, time taken to execute the algorithm on 2.3 GHz Intel Core i5 and 8 GB 2133 MHz LPDDR3, was also monitored and has shown a significant drop.

DISCUSSION AND CONCLUSION

High-Resolution Images store very minute information about an area in contact, as satellite sensors capture data with specific wavelengths across the electromagnetic spectrum. High-resolution images

Table 5. Error Matrix for Reduced Image with 16 bands using Wavelet Transformation and k-means

Feature	Alfafa	Corn-notill	Corn-min	Corn	Grass/Pasture	Grass/Trees	Grass/pasture-mowed	Hay	Oats	Soybeans-notill	Soybeans-min	Soybean-clean	Wheat	Woods	Bldg-Grass-Trees-Drives	Stone-steel towers	Total Reference Points
Alfafa	86	3	2	-	-	1	-	-	-	-	-	-	-	-	-	-	90
Corn-notill	2	300	1	4	1	-	-	3	3	-	1	-	-	2	-	7	324
Corn-min	-	1	111	-	-	2	3	4	-	-	-	-	-	-	-	-	121
Corn	7	-	3	68	1	2	-	-	2	-	2	2	4	1	-	-	100
Grass/ Pasture	-	-	1	1	107	3	-	2	-	-	1	3	-	-	-	-	120
Grass/Trees	5	3	-	-	-	220	3	-	10	-	-	9	3	10	4	7	280
Grass/ pasture-mowed	3	5	5	5	-	-	462	6	2	-	-	2	-	-	5	6	500
Hay	-	-	-	1	-	1	-	54	-	-	1	1	-	-	1	1	60
Oats	-	-	-	-	2	-	1	-	17	-	-	-	-	-	-	-	20
Soybeans-notill	4	4	-	3	2	-	-	-	-	170	3	1	-	-	-	-	190
Soybeans-min	-	-	2	-	2	-	-	-	-	-	15	1	-	-	-	-	20
Soybean-clean	2	5	-	-	2	-	2	2	-	-	-	113	5	1	2	-	135
Wheat	-	-	5	1	-	1	-	1	-	-	2	7	33	3	-	2	55
Woods	5	1	-	1	-	1	-	3	-	1	-	1	-	185	3	-	20
Bldg-Grass-Trees-Drives	-	-	-	1	-	1	-	1	-	-	-	-	-	2	85	-	90
Stone-steel towers	1	-	-	-	-	2	-	-	2	5	-	-	-	-	-	25	35

Table 6. Classification Accuracies for original and compressed image (Indian Pipes)

Image	Accuracy	k-means	Time is taken (ms)
Original image with 220 bands	OA	77%	178
	AA	76.8%	
	K	0.893	
Dimensionality Reduced image with 16 bands using PCA	OA	81.2%	134
	AA	84.5%	
	K	0.921	
Dimensionality Reduced image with 16 bands using Wavelet transformation	OA	77%	151
	AA	74.6%	
	K	0.823	

are categorized as multispectral and hyperspectral images, depending upon the number of spectral bands, image size, image format, and many more. Due to the high requirement of bandwidth for image storage and transmission, dimension reduction was suggested. In this paper, dimension is reduced in the form of band selection using principal component analysis, and its result is compared with Wavelet Transformation Algorithm. Indian Pine Site 3 Image used for experimental analysis was taken by the Earth Observation Satellite Company (EOSCT) at Purdue Foundation having 220 bands. Image is smoothened and sharpened using the noise removal technique in pre-processing step. Original image along with noise and without noise are shown in the experimental results section. The number of bands to be selected is determined using VD estimate and bands are selected using PCA and Wavelet-based Transform Algorithm. Resultant images are recorded, and their classification accuracies generated by k-means, are accessed using OA, AA, Kappa co-efficient, and execution time. Above mentioned classification metrics show that dimensionality reduced image using PCA gives better classification than the original one and the one reduced using Wavelet Transformation. Further in the future, the proposed algorithm can be modified to improve execution time and classification accuracy.

REFERENCES

Asl, G. M., & Mojaradi, B. (2016). Virtual dimensionality estimation in hyperspectral imagery based on unsupervised feature selection. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, 3*(7), 17-23. 10.5194/isprsannals-III-7-17-2016

Bajcsy, P., & Groves, P. (2004). Methodology for hyperspectral band selection. *Photogrammetric Engineering and Remote Sensing*, *70*(7), 793–802. doi:10.14358/PERS.70.7.793

Bajcsy, P., & Groves, P. (2004). Methodology for hyperspectral band selection. *Photogrammetric Engineering and Remote Sensing*, *70*(7), 793–802. doi:10.14358/PERS.70.7.793

De Backer, S., Kempeneers, P., Debruyn, W., & Scheunders, P. (2005). A band selection technique for spectral classification. *IEEE Geoscience and Remote Sensing Letters*, *2*(3), 319–323. doi:10.1109/LGRS.2005.848511

De Backer, S., Kempeneers, P., Debruyn, W., & Scheunders, P. (2005). A band selection technique for spectral classification. *IEEE Geoscience and Remote Sensing Letters*, *2*(3), 319–323. doi:10.1109/LGRS.2005.848511

Goel, L. (2010). *Land Cover Feature Extraction using Hybrid Swarm Intelligence Techniques-A Remote Sensing Perspective*. Academic Press.

Hambal, A. M., Pei, Z., & Ishabailu, F. L. (2017). Image noise reduction and filtering techniques. *International Journal of Scientific Research*, *6*(3), 2033–2038.

Johal, N. K., Singh, S., & Kundra, H. (2010). A hybrid FPAB/BBO algorithm for satellite image classification. *International Journal of Computer Applications*, *6*(5).

Karami, A., Yazdi, M., & Mercier, G. (2012, April). Compression of Hyperspectral Images Using Discerete Wavelet Transform and Tucker Decomposition. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *5*(2), 444–450. doi:10.1109/JSTARS.2012.2189200

Karamizadeh, S., Abdullah, S. M., Manaf, A. A., Zamani, M., & Hooman, A. (2013). An overview of principal component analysis. *Journal of Signal and Information Processing*.

Koonsanit, K., & Jaruskulchai, C. (2011). Band selection for hyperspectral image using principal components analysis and maxima-minima functional. In *Knowledge, Information, and Creativity Support Systems* (pp. 103–112). Springer. doi:10.1007/978-3-642-24788-0_10

Koonsanit, K., Jaruskulchai, C., & Eiumnoh, A. (2012, June). Band Selection for Dimension Reduction in Hyper Spectral Image Using Integrated Information Gain and Principal Components Analysis Technique. *International Journal of Machine Learning and Computing*, *2*(3).

Kuruvilla, J. J., Sukumaran, D., Sankar, A., & Joy, S. P. (2016). A review on image processing and image segmentation. In *Data Mining and Advanced Computing (SAPIENCE), International Conference on* (pp. 198-203). IEEE.

Li, F., & Lu, H. (2016). Hyperspectral images band selection using multi-dictionary based sparse representation. In *IEEE International Geoscience and Remote Sensing Symposium* (pp. 2769-2772). IEEE. doi:10.1109/IGARSS.2016.7729715

Markham, B. L., Storey, J. C., Williams, D. L., & Irons, J. R. (2004). AVIRIS sensor performance: History and current status. *IEEE Transactions on Geoscience and Remote Sensing*, *42*(12), 2691–2694. doi:10.1109/TGRS.2004.840720

Nakamura, R. Y., Fonseca, L. M. G., Dos Santos, J. A., Torres, R. D. S., Yang, X. S., & Papa, J. P. (2014). Nature-inspired framework for hyperspectral band selection. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(4), 2126–2137. doi:10.1109/TGRS.2013.2258351

Opticks. (2020). *OSGeo*. https://www.osgeo.org/projects/opticks/

Rodarmel, C., & Shan, J. (2002). Principal component analysis for hyperspectral image classification. *Surveying and Land Information Science*, *62*(2), 115–122.

Russ, J. C. (2016). *The image processing handbook*. CRC Press. doi:10.1201/b10720

Senthilnath, J., Omkar, S. N., Mani, V., Tejovanth, N., Diwakar, P. G., & Shenoy, A. (2011). Multi-spectral satellite image classification using glowworm swarm optimization. In *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International* (pp. 47-50). IEEE. doi:10.1109/IGARSS.2011.6048894

Sotoca, J. M., Pla, F., & Sanchez, J. S. (2007). Band selection in multispectral images by minimization of dependent information. *IEEE Transactions on Systems, Man, and Cybernetics Part C*, *37*(2), 258–267.

Sun, W., & Du, Q. (2019). Hyperspectral Band Selection: A Review. *IEEE Geoscience and Remote Sensing Magazine*, *7*(2), 118–139. doi:10.1109/MGRS.2019.2911100

Wang, X., & Leeser, M. (2007). K-means Clustering for Multispectral Images Using Floating-Point Divide. *Annual IEEE Symposium on Field-Programmable Custom Computing Machines*, 151-162. doi:10.1109/FCCM.2007.38

Yang, H., Du, Q., & Chen, G. (2012). Particle swarm optimization-based hyperspectral dimensionality reduction for urban land cover classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *5*(2), 544–554. doi:10.1109/JSTARS.2012.2185822

Zhang, B., Sun, X., Gao, L., & Yang, L. (2011). Endmember Extraction of Hyperspectral Remote Sensing Images Based on the Discrete Particle Swarm Optimization Algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(11), 4173–4176. doi:10.1109/TGRS.2011.2131145