An Efficient $K$-Nearest Neighbors Based Approach for Classifying Land Cover Regions in Hyperspectral Data via Non-Linear Dimensionality Reduction

1K Perumal and 2Dr. R Bhaskaran
1Department of Computer Science, DDE, Madurai Kamaraj University, Madurai-625021, Tamilnadu, India
Email: perumalmk@yahoo.co.in
2School of Mathematics, Madurai Kamaraj University, Madurai -625 021, Tamilnadu, India
Email: raman.bhaskaran@gmail.com

Abstract—In recent times, researchers in the remote sensing community have been greatly interested in utilizing hyperspectral data for in-depth analysis of Earth’s surface. In general, hyperspectral imaging comes with high dimensional data, which necessitates a pressing need for efficient approaches that can effectively process on these high dimensional data. In this paper, we present an efficient approach for the analysis of hyperspectral data by incorporating the concepts of Non-linear manifold learning and $k$-nearest neighbor $(k$-NN). Instead of dealing with the high dimensional feature space directly, the proposed approach employs Non-linear manifold learning that determines a low-dimensional embedding of the original high dimensional data by computing the geometric distances between the samples. Initially, the dimensionality of the hyperspectral data is reduced to a pairwise distance matrix by making use of the Johnson’s shortest path algorithm and Multidimensional scaling (MDS). Subsequently, based on the $k$-nearest neighbors, the classification of the land cover regions in the hyperspectral data is achieved. The proposed $k$-NN based approach is evaluated using the hyperspectral data collected by the NASA’s (National Aeronautics and Space Administration) AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) from Kennedy Space Center, Florida. The classification accuracies of the proposed $k$-NN based approach demonstrate its effectiveness in land cover classification of hyperspectral data.

Index Terms—Remote Sensing, Hyperspectral, Non-linear Dimensionality Reduction (NLDR), Mahalanobis distance, Johnson’s shortest path algorithm, Multidimensional Scaling (MDS), AVIRIS (Airborne Visible/Infrared Imaging Spectrometer), $k$-nearest neighbor ($k$-NN).

I. INTRODUCTION

Data collected from remote sensing serve as a dominant source for information on vegetation parameters that are desirable in all sorts of models meant for describing the processes at the Earth’s surface [1]. In recent times, hyperspectral data added even more power by presenting spectral information about ground scenes on the basis of an enormous number of channels with narrow contiguous spectral bands. Hence, the hyperspectral data can be used to achieve better discrimination of the spectral signatures of land-cover classes that materialize alike when viewed by traditional multispectral sensors [2]. The hyperspectral data are collected using hyperspectral sensors that sample the reflected solar radiation from the Earth’s surface in the portion of the spectrum extending from the visible region through the near-infrared and mid-infrared (wavelengths between 0.3 and 2.5 $\mu$m) in hundreds of narrow (on the order of 10 nm) contiguous bands [4]. These instruments symbolize spectral signatures with much greater detail than traditional multispectral sensors, and thus, can potentially offer improved discrimination of targets. This high spectral resolution yields enormous amounts of data, placing stringent requirements on communications, storage, and processing [3]. For instance, the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) amasses a 512 (along track) $\times$ 614 (across track) $\times$ 224 (bands) $\times$ 12 (bits) data cube in 43 s, corresponding to more than 700 Mb; Hyperion collects 4 Mb in 3 s, corresponding to 366 KByte/km2 [5]. Thus, the application of hyperspectral images brings in new capabilities and with it some difficulties in their processing and analysis.

The determination of land cover types corresponding to the spectral signatures in the hyperspectral image would be a typical application of hyperspectral data, for instance, to examine changes in the ecosystem over large geographic areas [8]. Hyperspectral images, unlike the extensively used multispectral images, can be utilized not only to differentiate distinct categories of land cover, but also the defining components of every land cover category, such as minerals, soil and vegetation type [7]. With all these advantages over multispectral images and with enormous quantities of hyperspectral data available, extracting reliable and accurate class labels for each ‘pixel’ from the hyperspectral images is a non-trivial task, involving either expensive field campaigns or time-consuming
manual interpretation [8]. With no doubts, the bulky amount of data concerned with hyperspectral imagery dramatically increases the processing complexity and time. A unique but principal task for hyperspectral image analysis is to achieve effective reduction in the amount of data involved or selection of the relevant bands associated with a specific application from the entire data set [7]. Some important operations that can be carried out with the information contained in hyperspectral data include characterization, identification, and classification of the land-covers with improved accuracy and robustness. Nevertheless, a number of decisive problems should be considered in classification of hyperspectral data namely, the high number of spectral channels, the spatial variability of the spectral signature, the high cost of true sample labeling and the quality of data [6].

Achieving high classification accuracy and good generalization in hyperspectral image analysis are comparatively simpler than the dimension of the input space, which continues to be a difficult problem, particularly when the number of classes is so large. Moreover, the high dimensionality of the data is challenging for supervised statistical classification techniques that make use of the estimated covariance matrix as the number of known samples is characteristically small relative to the dimension of the data [9]. Earlier studies of supervised methods have revealed that a complex classifier is likely to over train in the aforesaid situations, whereas a weak classifier is often insufficient [10]. With the increase in complexity of the classifiers, the generalization error eventually increases because of over-training [11]. The aforesaid problem can be alleviated using ensemble methods that work by reducing the model variance. Complex classifiers, in addition, do not characteristically perform well when characteristics of the training/test data obtained over the study site evolve in a new area. The above condition is referred as the knowledge transfer problem [12]. In land cover classification, the knowledge transfer problem is considered as a significant problem, as it is frequently difficult to acquire labeled samples from a new area. Changes in spectral signatures can be caused by seasonal changes, unknown land cover types or a different mixture of classes. Hence, it is essential to devise a simple classifier that can acclimatize to such changes and maintain good classification accuracies for the training/testing data. A number of existing classifiers build their models in accordance with the behavior of labeled samples in the reduced or original feature space.

On the contrary, non-linear manifold learning algorithms actually presume that the original high dimensional data lie on a low dimensional manifold defined by local geometric differences between samples. Current research has illustrated the impending of manifold learning algorithms for non-linear dimension reduction and for representation of high dimensional observations through non-linear mapping [13]. Good examples of manifold learning techniques include: Isometric feature mapping (Isomap) [16], Local Linear Embedding (LLE) [15], Laplacian Eigenmaps and Semidefinite Embedding. Even though, these methods were devised to symbolize, high dimensional non-linear phenomena in lower dimensional spaces, the embedded features are eminently helpful for classification of hyperspectral data. Lately, Bachmann et al. [13] and Chen et al. [14] have successfully applied the Isomap method to the hyperspectral data. Yet, the development of a more robust classifier that exploits the advantages of non-linear dimension reduction and computational efficiency is still an area of competitive research.

This research paper proposes an efficient approach based on $k$-nearest neighbors for hyperspectral data analysis using shortest path computation and Multidimensional scaling (MDS). Since, the proposed approach based on $k$-nearest neighbors’ deal with high dimensional data, the primary step is to perform dimensionality reduction of the high dimensional hyperspectral data. Primarily, a novel approach devised for non-linear manifold learning is applied to the high dimensional hyperspectral data, which reduces the dimensionality of the input to a pairwise distance matrix with the aid of the Johnson's shortest path algorithm. MDS is employed to estimate the dimension of the manifold created. Lastly, the $k$-nearest neighbors attained are used to classify the land cover regions in the hyperspectral data based on the distance measures computed during dimensionality reduction.

The rest of the paper is organized as follows. Section II presents a brief review of some recent significant researchers. The proposed non-linear manifold learning approach for dimensionality reduction and $k$-nearest neighbors based approach for hyperspectral data classification are described in section III. Experimental results obtained from hyperspectral data collected from Kennedy Space Center and a formal investigation of the classification accuracies of the proposed approach are presented in Section IV. Section V sums up the paper with the conclusion.

II. REVIEW OF RELATED SIGNIFICANT RESEARCHES

Literature presents with plentiful of researches that perform analysis of hyperspectral data. Of them, a significant number of researches make use of manifold learning for classifying hyperspectral data. A handful of significant works related to the proposed approach are presented below.

Hongjun Su et al. [17] have devised a novel algorithm named OBI, based on a traditional algorithm and fractal dimension, for quicker processing of
hyperspectral remote sensing data. To start with, the fractal dimension was utilized as the criterion to prune the noisy bands, and only those bands with better spatial structure, quality and spectral feature were preserved. Subsequently, the correlation coefficients and covariance amongst all bands were made use of to compute the optimal band index, followed by the selection of the optimum bands. OBI algorithm has proved over other algorithms, on band selection in hyperspectral remote sensing data processing. An algorithm that employs spectral-angle based Support Vector Clustering (SVC) and Principal Component Analysis (PCA) for hyperspectral image analysis was presented by S. Sindhumol and M. Wilsey [18]. In their previous research for hyper-spectral dimensionality reduction based on Principal Component Analysis (PCA), they have not taken into account, the meaning or behavior of the spectrum, and moreover, the results were prejudiced by the majority of the components in the scene. A probable solution to the aforesaid problem is to perform a spectral angle based classification before dimensionality reduction. In their current research, they have proposed a clustering based on support vectors using spectral based kernels that have produced good results in hyperspectral image classification. The algorithm was tested with two hyperspectral image data sets of 210 bands each that were recorded with HYper-spectral Digital Imagery Collection Experiment (HYDICE) air-borne sensors.

Qian Du and Nicolas H. Younan [19] have examined the application of Fisher’s linear discriminant analysis (FLDA) in classifying hyperspectral remote sensing images. The core idea of FLDA is to design an optimal transform so that the classes can be separated well in the low-dimensional space. The difficulty of realistically applying FLDA to hyperspectral images include: the unavailability of sufficient training samples and indefinite information for all the classes present. Hence, the original FLDA is altered to shun the requirements of complete class knowledge, for instance the number of actual classes present. They have also investigated the performance of the class of principal component analysis (PCA) techniques before FLDA and have discovered that the interference and noise adjusted PCA (INAPCA) can offer the improvement in the final classification.

Claudionor Ribeiro da SILVA et al. [20] have presented a method for selecting features from hyperspectral images. From the experiments conducted the feasibility of the use of genetic algorithms for dimensionality reduction of hyperspectral remote sensing data has been proved to improve the accuracy of digital classification. The elitism-based algorithm has attained the best results, by far. The utilization of genetic algorithm adds with an advantage of better flexibility in the search for an optimal solution, by reintroducing a spectral band that is discarded within the evolutionary process, in the optimal solution through reproduction or mutation. Moreover, the genetic algorithm was capable of identifying and discarding noisy bands, on the basis of the fitness criterion computed from the correlation, transformed divergence and optimal number of bands.

Qian Du et al. [21] have investigated the application of independent-component analysis (ICA) to hyperspectral remote sensing image classification. They had focused on the performance of two renowned and commonly utilized ICA algorithms: joint approximate diagonalization of Eigen matrices (JADE) and FastICA; yet their proposed method is also applicable to other ICA algorithms. The chief advantage of utilizing ICA is its capability to perform object classification with unknown spectral signatures in an unknown image scene, i.e., unsupervised classification. Nevertheless, ICA is computationally expensive, which restricts its application to high-dimensional data analysis. So as to, make it applicable or reduce the computation time in hyperspectral image classification, a data-preprocessing procedure has been employed to achieve dimensionality reduction of the data. The experimental results of their proposed approach have illustrated that the chief principal components from the NAPC transform can better preserve the object information in the original data than those from PCA. Consequently, an ICA algorithm could offer better object classification.

Yangchi Chen et al. [22] have examined the concept of L-Isomap and its pros and cons when applied to hyperspectral data. Isomap and L-Isomap were evaluated by conducting experiments on dimensionality reduction and representation of high dimensional observation. Moreover, they have studied on L-Isomap in coincidence with hyperspectral data classification. Their proposed MST-cut landmark selection approach was judged against random selection and k-means cluster centers. Berge et al. [23] have proposed a simple algorithm for reducing the complexity of Gaussian ML-based classifiers for hyperspectral data. The core idea of this research is to determine a sparse approximation to the inverse covariance’s of the component distributions utilized in their classification model. One inspiration for devising this approach was to combat the problems with conventional classifiers because of sample sparsity. The proposed approach reduced the number of parameters when the number of available ground truthed samples is low, whilst sacrificing a little accuracy in modeling the density. The experiments conducted have portrayed that their method performed comparably or better than state of the art conventional classifiers such as SVM, using only a fraction of the full covariance matrices. The performance compared to covariance regularization strategies, in their paper represented by LOOC, seems more than adequate. Their method also performs well in cases where QDA collapses due to sample sparsity and for modeling...
sparseness in covariance matrices is also directly applicable to covariance estimates of component distributions in mixture models.

In general, hyperspectral images contain hundreds of bands leading to covariance matrices having tens of thousands of elements. Of late, the time-series literature has witnessed the usage of general linear regression models in the estimation of inverse covariance matrix. C. Jensen et al. [24] have espoused and applied those ideas to the problems identified in ill-posed hyperspectral image classification. The results of experimentation have shown that at least some of the approaches can give a lower classification error than traditional methods such as the linear discriminant analysis (LDA) and the regularized discriminant analysis (RDA). Moreover, the results have shown in contrast to earlier beliefs that, long-range correlation coefficients are essential to build an effective hyperspectral classifier, and that the high correlations between neighboring bands appear to permit differing sparsity configurations of the covariance matrix to attain similar classification results.

III. PROPOSED APPROACH FOR LAND COVER CLASSIFICATION OF HYPERSPETRAL IMAGES USING NON-LINEAR MANIFOLD LEARNING AND $k$-NEAREST NEIGHBORS

This section details the efficient approach proposed for achieving land cover classification on hyperspectral data. The proposed approach integrates Non-linear manifold learning and the concepts of $k$-NN for effective classification of land cover in hyperspectral data. Generally, effective classification of land cover could be accomplished by using the classical $k$-NN, which classifies the novel observations based on the class label of its $k$-nearest neighbors by making use of a distance measure. But, classical $k$-NN alone does not perform exceptionally well on high dimensional data. As the proposed approach deals with hyperspectral data, it necessitates some add-ons to the classical $k$-NN classifier to exploit the advantages of the classical $k$-NN. Some advantages of the classical $k$-NN classifier that can be exploited include,

1. Easy to implement,
2. Very good classification accuracy on low dimensional problems
3. Provides Non-linear decision boundaries.

The classical $k$-NN classifier has achieved very good classification accuracies on low dimensional problems. So, an effective solution proposed for hyperspectral image classification is to incorporate the concepts of non-linear manifold learning and $k$-NN. Hence, in the proposed approach, we incorporate Non-linear manifold learning as a preprocessing step to hyperspectral image classification. Thereby, the proposed approach for hyperspectral land covers classification is composed of two phases namely,

- Novel manifold learning approach for NLDR
- Land cover classification using the $k$-nearest neighbors

A. Proposed Novel Manifold Learning Approach for NLDR

Dimensionality reduction aims at maintaining only the most significant dimensions, i.e. the dimensions that encompass the most valuable information for the mission at hand, or providing a mapping from the high dimensional space to the low dimensional embedding. Manifold learning is a popular approach for NLDR. Generally, the process of estimating a low-dimensional embedding of high-dimensional space, which underlies the data of interest, is called as manifold learning. Manifold learning algorithms are based on the assumption that most data sets have an artificially high dimensionality; although every data point comprises of possibly thousands of features, it could be illustrated as a function of only a few underlying parameters. Specifically, a non-linear manifold is an abstract mathematical space that is locally Euclidean (i.e., around every point, there is a neighborhood that is topologically the same as described by Euclidean geometry). For any two data points lying on a Non-linear manifold, the “true distance” between them is the geodesic distance on the manifold, i.e. the distance along the surface of the manifold, rather than the straight-line Euclidean distance. Researchers have proposed a number of algorithms for manifold learning, including: Stochastic Neighbor Embedding (SNE), Isomap, Locally Linear Embedding (LLE), Laplacian Eigenmaps, Semidefinite Embedding, and a host of variants of the aforesaid algorithms. Perhaps, most algorithms serve to be the finest and are applied among the multitude of procedures existing for NLDR. In spite of there common usage and optimal behavior, it is still possible to improve the classical manifold learning approaches, so as to achieve better results in its field of application.

One possible extension to the existing NLDR approaches is to tailor them, specifically for handling very large datasets. As the proposed approach for hyperspectral image classification is meant to deal with large datasets, we devise a novel approach for NLDR that aims to discover the significant underlying parameters so as to determine a low-dimensional representation of the data. The proposed approach for non-linear manifold learning is based on shortest path network updating and Multidimensional Scaling (MDS) for NLDR. The steps involved in the proposed approach for NLDR include,

- Initial computation of the pairwise distance matrix among the neighborhood ‘$k$’ using Mahalanobis distance.
Application of Johnson’s shortest path algorithm to compute the geodesic distance between all-pairs of points in the k-nearest neighbor graph constructed.

The NxN distance matrix computed is fed for dimensionality reduction to the classical MDS algorithm.

Given data points \( y_1, y_2, \ldots, y_n \in \mathbb{R}^D \), we assume that the data lies on a \( d \) - dimensional manifold \( M \) embedded within, where \( d < D \). Moreover, we assume the manifold \( M \) is described by a single coordinate chart \( f : M \rightarrow \mathbb{R}^d \). The manifold learning consists of finding \( z_1, z_2, \ldots, z_n \in \mathbb{R}^d \), where \( z_i = f(y_i) \). The processes involved in the proposed non-linear manifold learning approach are listed as follows:

1. **Construct neighborhood graph:** First, we determine a user-defined neighborhood given by ‘\( k' \)’. Subsequently, for every point ‘\( i \)’ in \( \mathbb{R}^D \), we employ the mahalanobis distance measure to estimate the ‘\( k' \)’ neighborhood points on the manifold. Mahalanobis distances computed is given by, \( d'(i,j) \) between pairs of points \( i,j \) in the input space \( Y \). A weighted graph \( G \) is constructed using the neighborhood points with edges \( d'(i,j) \). Mahalanobis distance [28] of each observation, with random vectors \( \tilde{x} \) and \( \tilde{y} \) and the covariance matrix \( S^{-1} \) is given by,

\[
d(\tilde{x}, \tilde{y}) = \sqrt{(\tilde{x} - \tilde{y})^T S^{-1} (\tilde{x} - \tilde{y})}
\]

2. **Shortest path updating using Johnson’s shortest path algorithm:** A shortest path algorithm is generally employed in non-linear manifold learning to compute the geodesic distances among the points that are beyond the neighborhood ‘\( k' \)’. Here, the proposed approach estimates the geodesic distances between all pairs of points in the manifold \( M \), by computing the shortest path distance among points using the neighborhood graph. Johnson’s algorithm is defined as an algorithm that finds the shortest paths between all pairs of vertices by taking advantage of the sparse nature of the graphs, with directed edges and no cycles. It finds a cost \( c(v) \) for each vertex, and as a result on reweighting of the graph, every edge has a non-negative weight. Suppose the graph has a vertex \( s \) that has a path to every other vertex [25]. Johnson’s algorithm computes the shortest paths from \( s \) to every other vertex, using Shimbels’ algorithm (which doesn’t care if the edge weights are negative), and then sets \( c(v) = \text{dist}(s,v) \).

The new weight of every edge is

\[
w'(u \rightarrow v) = \text{dist}(s,u) + w(u \rightarrow v) - \text{dist}(s,v)
\]

**JOHNSONAPSP(V,E,w) :**

create a new vertex \( s \) for every vertex \( v \in V \)

\[
w(s \rightarrow v) \leftarrow 0
\]

\[
w(v \rightarrow s) \leftarrow \infty
\]

\[
dist[s,\cdot] \leftarrow \text{SHIMBEL}(V,E,w,s)
\]

**JOHNSONAPSP(V,E,w) :**

create a new vertex \( s \) for every vertex \( u \in V \)

\[
w'(u \rightarrow v) \leftarrow \text{dist}(s,u) + w(u \rightarrow v) - \text{dist}(s,v)
\]

**apple** if SHIMBEL found a negative cycle

\[
\text{for every vertex } u \in V
\]

\[
w'(u \rightarrow v) \leftarrow \text{dist}(u,v) - \text{dist}(s,u) + \text{dist}(s,v)
\]

The algorithm spends \( \Theta(V) \) time adding the artificial start vertex \( s \), \( \Theta(VE) \) time running SHIMBEL, \( O(E) \) time reweighting the graph, and then \( \Theta(VE + V^2 \log V) \) running \( V \) passes of Dijkstra’s algorithm. Thus, the overall running time is \( \Theta(VE + V^2 \log V) \) compared to \( O(n3) \) for the naive Dijkstra’s algorithm and \( O(n4) \) for the Bellman-Ford-Moore algorithm.

3. **Construct d-dimensional embedding:** The classical MDS is applied to the matrix of graph distances \( D_G = [d_G(i,j)] \), which constructs an embedding \( D_{sp} \) obtained from the distance matrix evaluates the true dimension of the manifold. Given a matrix \( D \rightarrow \mathbb{R}^{n \times n} \) of dissimilarities, MDS constructs a set of points whose interpoint Euclidean distances match those in \( D \) closely and it is used to evaluate the true dimension of the non-linear manifold. Generally, MDS makes use of a stress function to evaluate the true dimension of the manifold. The stress function inversely measures the degree of correspondence between the distances among points implied by MDS map and the matrix input by the user. The general form of the stress function corresponding to MDS is as follows:
The updating distance matrix computed using the Johnson’s shortest path algorithm and the MDS represents the low dimensional manifold corresponding to the high dimensional hyperspectral data. And, the updated distance matrix used for classification preserves the local information on a graph whilst increasing the distance between non-neighbor samples. The input to the classifier devised is the non-linear embedding of the high dimensional data \( D_{stp} \), which would be potentially useful for land cover classification. Given the non-linear manifold \( D_{stp} \), the classifier classifies the rows of the data matrix into groups, based on the grouping of the rows of training. Every training instance is associated with a class label, which defines the grouping. Initially, for every row in \( D_{stp} \), a similarity measure is computed by comparing it with the training instances. The distance metric used in the proposed approach for land cover classification is mahalanobis distance, given by,

\[
\sqrt{\sum \sum (f(x_i) - d_{size8ij})^2}
\]

where, \( f \) is some function of the input data, and \( scale \) refers to a constant scaling factor, used to keep stress values between 0 and 1. When the MDS map perfectly reproduces the input data, \( f(x_i) = d_{ij} \) for all \( i \) and \( j \), so stress is zero. Thus, the smaller the stress, the better the representations of the MDS map.

**B Hyperspectral Land Cover Classification Using The k-Nearest Neighbors**

The proposed approach for the classification of land cover regions in the hyperspectral data makes use of the concepts of the k-NN. The classical k-NN classifier classifies the data based on the class label of its \( k \) nearest neighbors, in the distance sense. In spite of being a competitive algorithm for classification, k-NN, as most classification methods when dealing with high dimensional input data [26, 27] suffers from the curse-of-dimensionality and highly biased estimates. In the proposed approach, the difficulty of high dimensional data classification can be solved, by initially mapping the original data into a lower dimensional space by non-linear manifold learning (which can be viewed as a preprocessing task) and then performing classification on the \( k \)-nearest neighbors. The above process is applicable because, generally, the high dimensional data often represent phenomena that are intrinsically low dimensional.

The updated distance matrix computed using the Johnson’s shortest path algorithm and the MDS represents the low dimensional manifold corresponding to the high dimensional hyperspectral data. And, the updated distance matrix used for classification preserves the local information on a graph whilst increasing the distance between non-neighbor samples. The input to the classifier devised is the non-linear embedding of the high dimensional data \( D_{stp} \), which would be potentially useful for land cover classification. Given the non-linear manifold \( D_{stp} \), the classifier classifies the rows of the data matrix into groups, based on the grouping of the rows of training. Every training instance is associated with a class label, which defines the grouping. Initially, for every row in \( D_{stp} \), a similarity measure is computed by comparing it with the training instances. The distance metric used in the proposed approach for land cover classification is mahalanobis distance, given by,

\[
d(\bar{x}, \bar{y}) = \sqrt{(\bar{x} - \bar{y})^T S^{-1} (\bar{x} - \bar{y})}
\]

where, \( \bar{x} \) and \( \bar{y} \) are random vectors and \( S^{-1} \) is the covariance matrix.

Based on the similarity measure computed and the value of \( k \), the classifier selects \( k \) nearest neighbors associated with the input matrix. Then, based on majority voting of the \( k \)-nearest neighbors the algorithm classifies the unlabeled samples projected in the space of the distance matrix \( D_{stp} \) to their associated groups.

**IV. Experimental Results**

This section presents the results obtained from experimentation on the proposed \( k \)-nearest neighbors based approach for classifying land cover regions in hyperspectral data via non-linear dimensionality reduction. The proposed approach is programmed in MATLAB (Matlab 7.4). The performance (classification accuracy) of the proposed approach for hyperspectral data classification is evaluated by making use of the data collected from Kennedy Space Center (KSC). For clearly depicting the performance, the input KSC data is sampled and the individual samples are further divided into: training instance and test instance. As for any classification technique, the classification accuracy of the proposed approach is determined based on the level of training given. So, we analyze the effectiveness of proposed approach in classifying the hyperspectral data, by means of the classification accuracies obtained at different levels of training.

**A. Data Acquisition**

The experimentation of the proposed approach was performed on the hyperspectral data collected by NASA AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) instrument over the Kennedy Space Center (KSC), Florida. AVIRIS records data in 224 bands of 10 nm width with center wavelengths from 400 - 2500 nm. The KSC data, which has been acquired from an altitude of approximately 20 km, possess a spatial resolution of 18 m. The hyperspectral data analysis was performed 176 bands, after eliminating water absorption and low SNR bands. Selection of the training data was done by making use of the land cover maps derived from color infrared photography provided by the Kennedy Space Center and Landsat Thematic Mapper (TM) imagery. Moreover, the KSC personnel have developed a vegetation classification scheme so as to define functional types that are discernable at the spatial resolution of Landsat and the AVIRIS data. The similarity among the spectral signatures for certain vegetation types makes the discrimination of land cover for this environment very difficult. For classification, 13 classes representing the various land cover types that occur in this environment were defined for the site (see Table I). Here, Classes 4 and 6 represent mixed classes.
B. Classification Results

The KSC data chosen for experimentation is first divided into 10 random samples. Subsequently, for every sample of the KSC data: 75% is for chosen for training and 25% for testing. The results are recorded with distinct levels of training say, 5%, 15%, 30%, 50% and 75%(whole data chosen for training) and testing with 25% of test data. The experimentation is repeated for all ten random samples chosen, and the average classification accuracy of the proposed approach for hyperspectral classification is calculated.

Table II shows the classification accuracies corresponding to the 13 different classes found in KSC data with 75% training and 25% testing. Table III depicts the average classification accuracies obtained with training of 5%, 15%, 30%, 50% and 75% data and testing with 25% data.

V. Conclusion

In this paper, we have proposed an efficient approach for the analysis of hyperspectral data by integrating Non-linear manifold learning and the concepts of $k$-nearest neighbor ($k$-NN). The proposed $k$-NN based approach has employed non-linear manifold learning to determine a low-dimensional embedding of the original high dimensional data by computing the geometric distances between the samples. To start with, the proposed approach has employed the Johnson’s shortest path algorithm and the classical MDS for reducing the dimensionality of the hyperspectral data to a pairwise distance. Then, the classifier is applied to the $k$-nearest neighbors, for the classification of the land cover regions in the hyperspectral data. The proposed $k$-NN based approach has been assessed using the hyperspectral data collected by the NASA’s (National Aeronautics and Space Administration) AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) from Kennedy Space Center (KSC), Florida. The classification accuracies of the proposed $k$-NN based approach have illustrated its efficacy in land cover classification of hyperspectral data.

REFERENCES

[9] D. Landgrebe, "Hyperspectral image data analysis as a high dimensional signal processing problem,” (Invited),

<table>
<thead>
<tr>
<th>Class</th>
<th>No. samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scrub</td>
<td>761 (14.6%)</td>
</tr>
<tr>
<td>Willow swamp</td>
<td>243 (4.66%)</td>
</tr>
<tr>
<td>Cabbage palm hammock</td>
<td>256 (4.92%)</td>
</tr>
<tr>
<td>Cabbage palm/oak hammock</td>
<td>252 (4.84%)</td>
</tr>
<tr>
<td>Slash pine</td>
<td>161 (3.07%)</td>
</tr>
<tr>
<td>Oak/broadleaf hammock</td>
<td>229 (4.18%)</td>
</tr>
<tr>
<td>Hardwood swamp</td>
<td>105 (2.0%)</td>
</tr>
<tr>
<td>Graminoid marsh</td>
<td>431 (8.27%)</td>
</tr>
<tr>
<td>Spartina marsh</td>
<td>520 (9.99%)</td>
</tr>
<tr>
<td>Cattail marsh</td>
<td>404 (7.76%)</td>
</tr>
<tr>
<td>Salt marsh</td>
<td>419 (8.04%)</td>
</tr>
<tr>
<td>Mud flats</td>
<td>503 (9.66%)</td>
</tr>
<tr>
<td>Water</td>
<td>927 (17.8%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Class</th>
<th>Classification Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>98.335</td>
</tr>
<tr>
<td>2</td>
<td>88.525</td>
</tr>
<tr>
<td>3</td>
<td>92.188</td>
</tr>
<tr>
<td>4</td>
<td>66.667</td>
</tr>
<tr>
<td>5</td>
<td>60.976</td>
</tr>
<tr>
<td>6</td>
<td>46.552</td>
</tr>
<tr>
<td>7</td>
<td>85.185</td>
</tr>
<tr>
<td>8</td>
<td>87.963</td>
</tr>
<tr>
<td>9</td>
<td>95.385</td>
</tr>
<tr>
<td>10</td>
<td>97.030</td>
</tr>
<tr>
<td>11</td>
<td>99.048</td>
</tr>
<tr>
<td>12</td>
<td>85.714</td>
</tr>
<tr>
<td>13</td>
<td>98.707</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Training %</th>
<th>Average Classification Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>5%</td>
<td>82.020</td>
</tr>
<tr>
<td>15%</td>
<td>85.692</td>
</tr>
<tr>
<td>30%</td>
<td>87.299</td>
</tr>
<tr>
<td>50%</td>
<td>88.370</td>
</tr>
<tr>
<td>75%</td>
<td>89.671</td>
</tr>
</tbody>
</table>


[28] "Mahalanobis distance" from http://www.aiaccess.net/English/Glossaries/GlosMod/e_gm_mahalanobis.htm