Design of KNN Classified Segmentation of Hyper Spectral Imaging

Vinothini .R , Vengateshkumar M.E (Ph.D)
Mohamed Sathak Engineering College

Abstract

Hyperspectral imaging (HSI) collects valuable information for monitoring the surface of Earth , thus addressing important remote sensing applications, including environmental management , agriculture , surveillance , and physics . High spectral dimension of the images gives detailed information about the objects and hence increase the potential to detect and identify the unique trends in the objects with more accuracy. In this Project we have proposed a correlation based approach for band selection and band reduction. So, reducing the number of bands without much compromising the information content has been a challenge in the field of hyper spectral image classification. This paper attempts to address a correlation based approach for band selection and KNN classifier for classification. This approach entails calculation of the correlation among the bands of the hyper spectral image and subsequent selection of those bands having correlation less than a threshold value. The experimental results obtained, have shown that with only a very limited number of bands we can achieve accuracy closer to that of using all the bands.

I INTRODUCTION

Hyperspectral imaging, or imaging spectroscopy, combines the power of digital imaging and spectroscopy. For each pixel in an image, a hyperspectral camera acquires the light intensity (radiance) for a large number (typically a few tens to several hundred) of contiguous spectral bands. Every pixel in the image thus contains a continuous spectrum (in radiance or reflectance) and can be used to characterize the objects in the scene with great precision and detail.

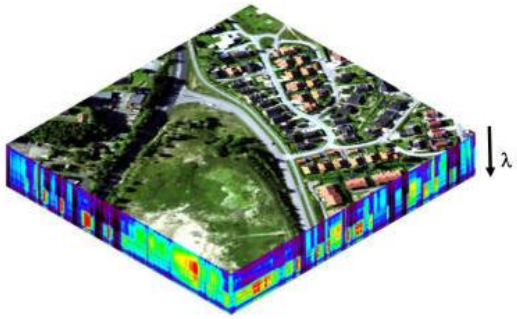


Fig. Illustration of hyperspectral data cube.

Hyperspectral images obviously provide much more detailed information about the scene than a normal color camera, which only acquires three different spectral channels corresponding to the visual primary colors red, green and blue. Hence, hyperspectral imaging leads to a vastly improved ability to classify the objects in the scene based on their spectral properties.

Recent advances in sensor design and processing speed has cleared the path for a wide range of applications employing hyperspectral imaging, ranging from satellite based/airborne remote sensing and military target detection to industrial quality control and lab applications in medicine and biophysics. Due to the rich information content in hyperspectral images, they are uniquely well suited for automated image processing,

whether it is for online industrial monitoring or for remote sensing.

II RELATED WORKS

MiguelA. Et al [1] presented a Hyperspectral Image Segmentation Using a New Spectral Unmixing-Based Binary Partition Tree Representation. The binary partition tree (BPT) is a hierarchical region-based representation of an image in a tree structure. The BPT allows users to explore the image at different segmentation scales. Often, the tree is pruned to get a more compact representation and so the remaining nodes conform an optimal partition for some given task. Here, we propose a novel BPT construction approach and pruning strategy for hyperspectral images based on spectral unmixing concepts. Linear spectral unmixing consists of finding the spectral signatures of the materials present in the image (endmembers) and their fractional abundances within each pixel. The proposed methodology exploits the local unmixing of the regions to find the partition achieving a global minimum reconstruction error.

BaiXue et al [2] presented a completely different approach from a subpixel target detection view point. It implements four stage processes, a preprocessing stage, which uses band selection (BS) and nonlinear band expansion, referred to as BS-then-nonlinear expansion (BSNE), a detection stage, which implements constrained energy minimization (CEM) to produce subpixel target maps, and an iterative stage, which develops an iterative CEM (ICEM) by applying Gaussian filters to capture spatial information, and then feeding the Gaussian-filtered CEM-detection maps back to BSNE band images to reprocess CEM in an iterative manner. Finally, in the last stage Otsu's method is applied to converting ICEM-detected real-valued maps to discrete values for classification. The entire process is called BSNE-ICEM.

Robert Pike et al [3] developed a classification method that combines both spectral and spatial information for distinguishing cancer from healthy tissue on hyperspectral images in an animal model. Methods: An automated algorithm based on a minimum spanning forest (MSF) and optimal band selection has been proposed to classify healthy and cancerous tissue on hyperspectral images. A support vector machine classifier is trained to create a pixel-wise classification probability map of cancerous and healthy tissue. This map is then used to identify markers that are used to compute mutual information for a range of bands in the hyperspectral image and thus select the optimal bands. An MSF is finally grown to segment the image using spatial and spectral information. Conclusion: The MSF based method with automatically selected bands proved to be accurate in determining the tumor boundary on hyperspectral images. Significance: Hyperspectral imaging combined with the proposed classification technique has the potential to provide a noninvasive tool for cancer detection.

Saurabh Prasad et al[4] challenged the conventional approach to hyperspectral classification that typically builds sparsity-based classifiers directly on spectral reflectance features or features derived directly from the data. We assert that hyperspectral image (HSI) processing can benefit very significantly by decoupling data into geometrically distinct components since the resulting decoupled components are much more suitable for sparse representation-based classifiers. Specifically, we apply morphological separation to decouple data into texture and cartoon-like components, which are sparsely represented using local discrete cosine bases and multiscaleshearlets, respectively. In addition to providing a structured sparse representation, this approach

allows us to build classifiers with invariance properties specific to each geometrically distinct component of the data.

Weisheng Dong et al [5] proposes a new hyperspectral image super-resolution method from a low-resolution (LR) image and a HR reference image of the same scene. The estimation of the HR hyperspectral image is formulated as a joint estimation of the hyperspectral dictionary and the sparse codes based on the prior knowledge of the spatial-spectral sparsity of the hyperspectral image. The hyperspectral dictionary representing prototype reflectance spectra vectors of the scene is first learned from the input LR image. Specifically, an efficient non-negative dictionary learning algorithm using the block-coordinate descent optimization technique is proposed. Then, the sparse codes of the desired HR hyperspectral image with respect to learned hyperspectral basis are estimated from the pair of LR and HR reference images.

Mohamed Amine Bendoumi et al [6] proposed a hyperspectral (HS) image resolution enhancement algorithm based on spectral unmixing for the fusion of the high-spatial-resolution multispectral (MS) image and the low-spatial-resolution HS image (HSI). As a result, a high-spatial-resolution HSI is reconstructed based on the high spectral features of the HSI represented by endmembers and the high spatial features of the MS image represented by abundances. Since the number of endmembers extracted from the MS image cannot exceed the number of bands in least-squares-based spectral unmixing algorithm, large reconstruction errors will occur for the HSI, which degrades the fusion performance of the enhanced HSI. Therefore, in this paper, a novel fusion framework is also proposed by dividing the whole image into several subimages, based on which the performance of the proposed spectral-unmixing-based fusion algorithm can be further improved. Finally, experiments on the Hyperspectral Digital Imagery Collection

Miguel A. Veganzones et al [7] proposed to partition the image into patches and solve the data fusion problem independently for each patch. This way, in each patch the subspace/manifold dimensionality is low enough, such that the problem is not ill-posed anymore. We propose two alternative approaches to define the hyperspectral super-resolution through local dictionary learning using endmember induction algorithms. We also explore two alternatives to define the local regions, using sliding windows and binary partition trees. The effectiveness of the proposed approaches is illustrated with synthetic and semi real data.

Mercedes E. Paoletti et al [8] pursued to improve the spectral-spatial features uncovered by the convolutional filters of the network. Specifically, the proposed residual-based approach gradually increases the feature map dimension at all convolutional layers, grouped in pyramidal bottleneck residual blocks, in order to involve more locations as the network depth increases while balancing the workload among all units, preserving the time complexity per layer. It can be seen as a pyramid, where the deeper the blocks, the more feature maps can be extracted. Therefore, the diversity of high-level spectral-spatial attributes can be gradually increased across layers to enhance the performance of the proposed network with the HSI data. Our experiments, conducted using four well-known HSI data sets and 10 different classification techniques, reveal that our newly developed HSI pyramidal residual model is able to provide competitive advantages (in terms of both classification accuracy and computational time) over the state-of-the-art HSI classification methods.

Kai Zhang et al [9] proposed a new group spectral embedding (GSE)-based LRHS and HRMS image fusion method is proposed by exploring the multiple manifold structures of spectral bands and the low-rank structure of HRHS data. First, a low-rank factorization fusion (LRFF)-based robust recovery model is developed for HRHS images, by regarding HRMS images as the spectral degradation of HRHS images and exploring the group sparse prior of difference images. Then, an assumption that grouped spectral bands share the similar local geometry is cast on LRHS and HRHS images, to formulate a GSE regularizer in the LRFF model.

Raúl Guerra et al [10] introduces a novel algorithm named fast algorithm for linearly unmixinghyperspectral images (FUN), which is capable of fully unmixing a hyperspectral image with at least the same accuracy than state-of-the-art approaches while demanding a much lower computational effort, independent of the characteristics of the image under analysis. The FUN algorithm is based on the concept of orthogonal projections and allows performing the estimation of the number of endmembers and their extraction simultaneously, using the modified Gram-Schmidt method. The operations performed by the FUN algorithm are simple and can be easily parallelized. Moreover, this algorithm is able to calculate the abundances using very similar operations, also based on orthogonal projections, which makes it easier to achieve a hardware implementation to perform the entire unmixing process.

III PROPOSED SYSTEM

We have proposed a correlation based approach for band selection and band reduction. We have utilized the correlation that generally exists among the adjacent bands and have shown that correlation can be an effective measure to reduce the number of bands. In our approach we have designed an algorithm which calculates the pair wise correlation coefficient of the bands. of the image and selects only the bands which are moderately correlated. Highly correlated bands are rejected. A comparison has been shown between the experimental results obtained with reduced set of bands, selected by the given approach and that of using all the bands of the image.

In statistics correlation is a technique to measure the strength of relationship between pair of random variables. The correlation between two random variables X and Y is defined as

$$\rho(X,Y) = Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Variance(X)Variance(Y)}} \quad (1)$$

Where ρ is the Product Moment Correlation Coefficient, simply called as Correlation Coefficient and Cov means covariance. Covariance measures how much each of the variable varies from the mean with respect to each other and variance measures the deviation from the mean for the points in one variable.

If we have series of n values for variables X and Y denoted by xi and yi, where i=1,2,3.....n, then the formula for covariance and variance can be written as

$$Cov(X,Y) = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{n-1}$$
$$Variance(X) = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(x_i-\bar{x})}{n-1} \quad (2)$$

Where x' and y' are the means of the observations of variable X and Y respectively.

Using equation 2 equation 1 can be rewritten as

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2 \sum_{i=1}^{n}(y_i-\bar{y})^2}}$$

Where r_{xy} is the Pearson's product moment correlation.

The correlation coefficient can take values from +1 to -1. Positive value of correlation indicates that as the value of one variable increases, so does the other. Negative value indicates an inverse linear relationship. A correlation of 0.5 represents a moderately strong relationship and 0.8 and above represents a very strong relationship.

The proposed band selection technique utilizes the high correlation that exists among the hyper spectral bands. Highly correlated bands share so much information that one band can well represent the other. So, from each group of highly correlated bands we select only one as a representative for the whole group. After finding out the set of selected bands we create training and testing sample using those bands only. We use the Support Vector Machine for training and testing purpose as they have the ability to efficiently deal with large input dimension and also can deal with noisy samples in a robust way.

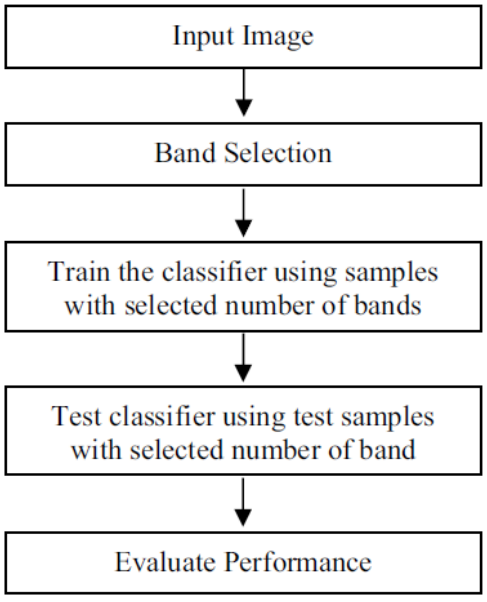


Fig. Proposed Block diagram

A. Algorithm for correlation based band reduction method:

Step1: Take two sets, selected_bands{} and candidate_bands {}. Initialize the set selected_bands {}, to include band 1 of the image and add the rest of the bands to the candidate_bands{}. Also select a threshold value T.
Step2: Select a band B, from the set of candidate_bands{} and find its correlation coefficient with each band belonging to selected_bands{}. If all the correlation coefficents are in the range of +T and -T, then include it in the selected_bands{} else discard the band. Delete B from candidate_bands{}.
Step 3: Repeat step 2 until the set candidate_bands{} ,is not empty. Else go to step 4.
Step 4: Output selected_bands{} as the final set of selected bands.
Step 5: Stop

k-nearest neighbor classifier algorithm

(*k*-NN) is a method for classifying objects based on closest training examples in the feature space. *k*-NN is a type of instance-based learning, or lazy learning where the function is only approximated locally and all computation is deferred until classification. It can also be used for regression.

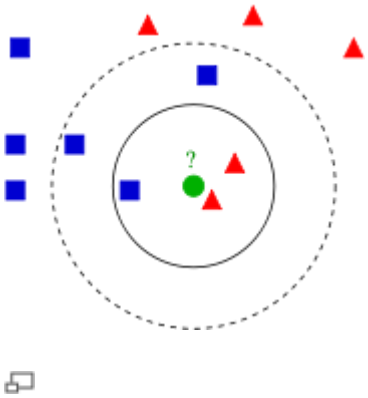
The ***k*-nearest neighbor algorithm** is amongst the simplest of all machine learning algorithms. An object is classified by a majority vote of its neighbors, with the object being assigned the class most common amongst its *k* nearest neighbors. *k* is a positive integer, typically small. If $k = 1$, then the object is simply assigned the class of its nearest neighbor. In binary (two class) classification problems, it is helpful to choose *k* to be an odd

number as this avoids difficulties with tied votes.

The same method can be used for regression, by simply assigning the property value for the object to be the average of the values of its *k* nearest neighbors. It can be useful to weight the contributions of the neighbors, so that the nearer neighbors contribute more to the average than the more distant ones.

The neighbors are taken from a set of objects for which the correct classification (or, in the case of regression, the value of the property) is known. This can be thought of as the training set for the algorithm, though no explicit training step is required. In order to identify neighbors, the objects are represented by position vectors in a multidimensional feature space. It is usual to use the Euclidean distance, though other distance measures, such as the Manhattan distance could in principle be used instead. The *k*-nearest neighbor algorithm is sensitive to the local structure of the data.

A. Algorithm



Example of *k*-NN classification. The test sample (green circle) should be classified either to the first class of blue squares or to the second class of red triangles. If $k = 3$ it is classified to the second class because there are 2 triangles and only 1 square inside the inner circle. If $k = 5$ it is classified to first class (3 squares vs. 2 triangles inside the outer circle).

The training examples are vectors in a multidimensional feature space. The space is partitioned into regions by locations and labels of the training samples. A point in the space is assigned to the class *c* if it is the most frequent class label among the *k* nearest training samples. Usually Euclidean distance is used.

The training phase of the algorithm consists only of storing the feature vectors and class labels of the training samples. In the actual classification phase, the test sample (whose class is not known) is represented as a vector in the feature space. Distances from the new vector to all stored vectors are computed and *k* closest samples are selected. There are a number of ways to classify the new vector to a particular class, one of the most used technique is to predict the new vector to the most common class amongst the K nearest neighbors. A major drawback to use this technique to classify a new vector to a class is that the classes with the more frequent examples tend to dominate the prediction of the new vector, as they tend to come up in the K nearest neighbors when the neighbors are computed due to their large number. One of the ways to overcome this problem is to take into account the distance of each K nearest neighbors with the new vector that is to be classified and predict the class of the new vector based on these distances.

B. Parameter selection

The best choice of *k* depends upon the data; generally, larger values of *k* reduce the effect of noise on the classification, but make boundaries between classes less distinct. A good *k* can be selected by various heuristic techniques, for example, cross-validation. The special case where the class is predicted to be the class of the closest training sample (i.e. when $k = 1$) is called the nearest neighbor algorithm.

The accuracy of the *k*-NN algorithm can be severely degraded by the presence of noisy or irrelevant features, or if the feature scales are not consistent with their importance. Much research effort has been put into selecting or scaling features to improve classification. A particularly popular approach is the use of evolutionary algorithms to optimize feature scaling. Another popular approach is to scale features by the mutual information of the training data with the training classes.

C. Properties

The naive version of the algorithm is easy to implement by computing the distances from the test sample to all stored vectors, but it is computationally intensive, especially when the size of the training set grows. Many optimizations have been proposed over the years; these generally seek to reduce the number of distance evaluations actually performed. Some optimizations involve partitioning the feature space, and only computing distances within specific nearby volumes. Several different types of nearest neighbor finding algorithms include:

- Linear scan
- Kd-trees
- Balltrees
- Metric trees
- Locality sensitive hashing (LSH)
- Agglomerative-Nearest-Neighbor

The nearest neighbor algorithm has some strong consistency results. As the amount of data approaches infinity, the algorithm is guaranteed to yield an error rate no worse than twice the Bayes error rate (the minimum achievable error rate given the distribution of the data). *k*-nearest neighbor is guaranteed to approach the Bayes error rate, for some value of *k* (where *k* increases as a function of the number of data points).

The *k*-NN algorithm can also be adapted for use in estimating continuous variables. One such implementation uses an inverse distance weighted average of the *k*-nearest multivariate neighbors. This algorithm functions as follows:

1. Compute Euclidean or Mahalanobis distance from target plot to those that were sampled.

2. Order samples taking for account calculated distances.

3. Choose heuristically optimal *k* nearest neighbor based on RMSE done by cross validation technique.

4. Calculate an inverse distance weighted average with the *k*-nearest multivariate neighbors

IV EXPERIMENTAL RESULTS

Classification Accuracy of Pavia University Scene

Classifiers	NB Classifier	K-NN Classifier
Classification Accuracy	84	86

Classification accuracy of Naive Bayes classifier is 84 and K-NN classifier is 86 the Pavia university scene, which is shown in Table 7.1. The Classification Accuracy of K-NN and Naïve bayes classifier has been shown in Fig 7.3 and its classified Pavia image has been shown in Fig 7.4.

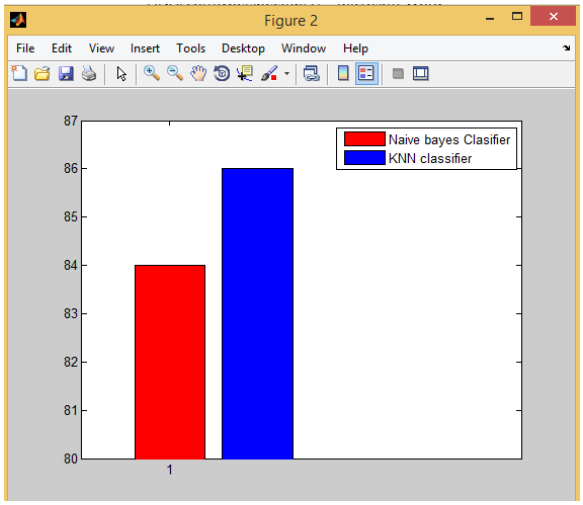


Fig 2 Classification Accuracy of K-NN & NB Classifier

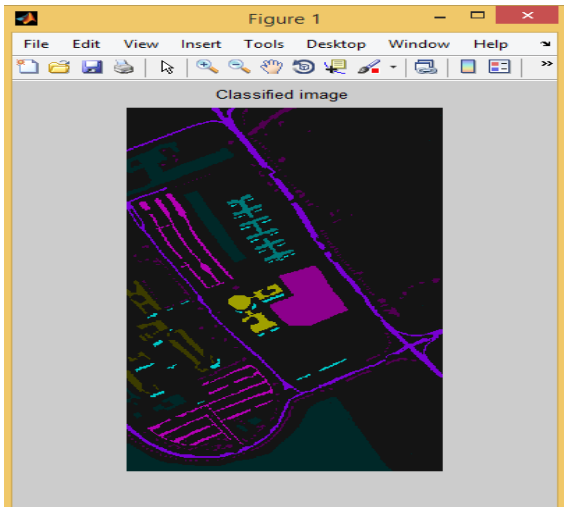


Fig 3 Classified Pavia image

V CONCLUSION

In this project, we presents an approach for band reduction in hyper spectral images based on calculation of correlation coefficient. Here we are taking the advantage of high correlation among the bands of hyper spectral images. We applied the technique over Indian Pine data set and found that with only less than 15% of total number of bands most of the classes can be classified satisfactorily. This reduces the dimension of the image with very little loss of information. For the classes 1, 2 and 8 the result of classification is low even if we use all the bands and degrades significantly the number of bands is reduced. These classes put a challenge for the classification algorithm which we will try to address in our future studies.

REFERENCES

[1]. Veganzones, M. A., Tochon, G., Dalla-Mura, M., Plaza, A. J., &Chanussot, J. (2014). Hyperspectral Image Segmentation Using a New Spectral Unmixing-Based Binary Partition Tree Representation. IEEE Transactions on Image Processing, 23(8), 3574–3589. doi:10.1109/tip.2014.

[2]. Xue, B., Yu, C., Wang, Y., Song, M., Li, S., Wang, L., … Chang, C.-I. (2017). A Subpixel Target Detection Approach to Hyperspectral Image Classification. IEEE Transactions on Geoscience and Remote Sensing, 55(9), 5093–5114. doi:10.1109/tgrs.2017

[3]. Pike, R., Lu, G., Wang, D., Chen, Z. G., &Fei, B. (2016). A Minimum Spanning Forest-Based Method for Noninvasive Cancer Detection WithHyperspectral Imaging. IEEE Transactions on Biomedical Engineering, 63(3), 653–663. doi:10.1109/tbme.2015

[4]. Prasad, S., Labate, D., Cui, M., & Zhang, Y. (2017). Morphologically Decoupled Structured Sparsity for Rotation-Invariant Hyperspectral Image Analysis. IEEE Transactions on Geoscience and Remote Sensing, 55(8), 4355–4366. doi:10.1109/tgrs.2017

[5]. Dong, W., Fu, F., Shi, G., Cao, X., Wu, J., Li, G., & Li, X. (2016). Hyperspectral Image Super-Resolution via Non-Negative Structured Sparse Representation. IEEE Transactions on Image Processing, 25(5), 2337–2352. doi:10.1109/tip.2016.

[6]. Bendoumi, M. A., Mingyi He, &Shaohui Mei. (2014). Hyperspectral Image Resolution Enhancement Using High-Resolution Multispectral Image Based on Spectral Unmixing. IEEE Transactions on Geoscience and Remote Sensing, 52(10), 6574–6583. doi:10.1109/tgrs.2014

[7]. Veganzones, M. A., Simoes, M., Licciardi, G., Yokoya, N., Bioucas-Dias, J. M., &Chanussot, J. (2016). Hyperspectral Super-Resolution of Locally Low Rank Images From Complementary Multisource Data. IEEE Transactions on Image Processing, 25(1), 274–288. doi:10.1109/tip.2015

[8]. Mercedes, "Deep Pyramidal Residual Networks for Spectral-Spatial Hyperspectral Image Classification" IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING,2010

[9]. Zhang, K., Wang, M., & Yang, S. (2017). Multispectral and Hyperspectral Image Fusion Based on Group Spectral Embedding and Low-Rank Factorization. IEEE Transactions on Geoscience and Remote Sensing, 55(3), 1363–1371. doi:10.1109/tgrs.2016

[10]. Guerra, R., Santos, L., Lopez, S., & Sarmiento, R. (2015). A New Fast Algorithm for Linearly UnmixingHyperspectral Images. IEEE Transactions on Geoscience and Remote Sensing, 53(12), 6752–6765. doi:10.1109/tgrs.2015