

# Survey on Feature Extraction of Hyperspectral Images

Fazeela Hamza<sup>1</sup>, Sreeram S.<sup>2</sup>

P.G. Student, Department of Computer Engineering, MEA Engineering College, Perinthalmanna, Kerala, India<sup>1</sup>

Associate Professor & HOD, Department of Computer Engineering, MEA Engineering College, Perinthalmanna,  
Kerala, India<sup>2</sup>

**ABSTRACT:** Hyperspectral Imaging produces an image where each pixel is having narrow spectral bands with plentiful spectral information. Spectral bands refer to the large number of measured wavelengths bands of Electromagnetic Spectrum. The large number of spectral bands in hyperspectral data increases the computational burden. So it is essential to do dimensionality reduction. Here comes the advantage of using feature extraction. The aim of this paper is to give a review on different feature extraction techniques.

**KEYWORDS:** Hyperspectral images, Electromagnetic Spectrum, Feature Extraction.

## I. INTRODUCTION

With the development in remote sensing technology, several hundred narrow-band images of the same area on the Earth are available. The hyperspectral satellite sensors like Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) provide spectral information covering the visible and the infrared wavelength range. This information is collected over continuous narrow spectral bands by measuring the reflectance response of the surface to different wavelengths. The Hyper Spectral Image (HSI) applications have extensively been researched in the areas of remote sensing, environment monitoring, geological surveying and monitoring due to their distinct advantages in classification and also object identification. Hyperspectral data faces following challenges:

- **High computational cost**  
The computational costs of processing are very high because a large number of image bands (usually 200--250).
- **Storage requirement**  
A single hyperspectral image data set contains a large number of image bands of the scene. The storage of such data needs a few hundreds of megabytes of memory.
- **Redundancy**  
The nearby image bands in the hypercube exhibit a high degree of spatial correlation among them due to the continuous and narrow nature of the hyperspectral sensors. The redundancy of the data needs to be removed for efficient processing and reduction in the computational overheads.

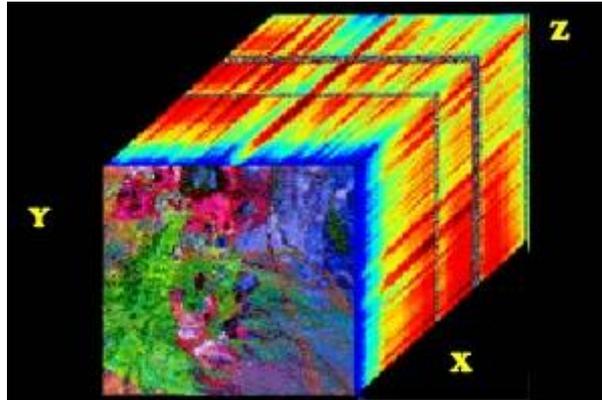


Fig1: Hypercube

In figure1 X and Y, dimensions represents spatial coordinates and Z dimension represent spectral bands. The hyperspectral data present abundant multidimensional information which contains more image bands than those that can be displayed on the standard tristimulus display. Therefore, an efficient and appropriate means of visualization of the hyperspectral data is the need of the hour.

Hyperspectral images are satellite images and are mainly affected by Hughes phenomena . Hughes phenomenon (curse of dimensionality) is a phenomenon that the classification precision increases gradually in the beginning as the number of spectral bands or dimensions increases, but when the band numbers reached at some point, the estimation accuracy begin to decrease dramatically. As a result , when the dimensionality of a data increases, the volume of the space increases so fast that the available data become sparse. The dimensionality of a data must be small for its better classification.

To differentiate different landscapes and to provide spectral information regarding the physical nature of materials, hyperspectral image is used.

## II. DIFFERENT FEATURE EXTRACTION METHODS

**1. Principal Component Analysis (PCA):** In hyperspectral imaging, PCA is based on the fact that neighboring bands are highly correlated which often convey almost the same info about the object. Hence, the analysis is used to transfer higher dimensional hyperspectral data to lower dimensional hyperspectral data removing the correlation among the bands. It employs statistic properties of hyperspectral image [2].

**Advantage:** Most of the information may exclusively be contained in the first few PCA bands. As number of PCA bands increases, the information content of the PCA bands reduces. Remaining bands contain noisy information.

**Disadvantage:** Only the subsets of the original bands are used for the analysis and PCA is performed on these subsets . PCA ca not eliminate anomalies that can be seen in one arbitrary band. PCA transforms an image into a new coordinate system without considering factors such as noise.

**2. Minimum Noise Fraction (MNF) :** MNF is a data transformation technique [3]. In hyperspectral image processing, MNF is commonly used to align the data along axes of decreasing signal to noise ratio (SNR). Since, hyperspectral images tend to contain more noise, MNF is a technique to address the noise present in the image. MNF maximizes the estimated SNR. Few noise filters are used like Adaptive Filters (AF), etc.

**Advantage :** MNF of an image (i.e. higher order bands in hyperspectral image) can be truncated incorporating low SNR that eliminates much of the noise without losing much information

**Disadvantage:** If truncation consists of too many bands, too much noise is left in the image. If truncation consists of too few bands, useful signal may be excluded from the resulting image.

**3. Linear Discriminant Analysis (LDA):** Fisher's linear discriminant analysis is well suited when data is normally distributed. It is a parametric feature extraction method. LDA is based on following three scatter matrices: within-class, between class and mixture (or total) scatter matrices which can be used to formulate criteria of class separability.

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 7, July 2016

Transforms the data from a high dimensional space into low dimensional subspace with maximized class separability. In hyperspectral images, transformation is done using the ratio of number of training samples and spectral features [4] [5].

**Advantage:** This method is distribution-free, fast and easy to apply, gives better within class scatter matrices.

**Disadvantage:** The rank of between-class scatter matrix is the number of classes minus one. Therefore, the maximum number of features that can be extracted remains the same which may not be sufficient for achieving better accuracy in practical application. Cannot perform well for normally distributed data. Sensitive to outliers. In hyperspectral image processing, the ratio of the number of training samples and features is small due to the fact that labeling is expensive and that the feature space has high dimensionality. Hence, it suffers from the singularity problem since within class covariance is singular.

**4. Non-parametric feature extraction:** It is commonly used for high dimensional multiclass pattern recognition problem. Based on non-parametric extension of scatter matrices. To overcome the limitations of LDA non-parametric feature extraction algorithms are developed like nonparametric Discriminant analysis (NDA) based on nonparametric between class scatter matrix, nonparametric weighted feature extraction (NWFE) with fully nonparametric within-class and between-class scatter matrices, nonparametric weighted feature extraction (NWFE), some regularization techniques to solve singularity problem, and nonparametric fuzzy feature extraction (NFFE) which introduces fuzzy membership grades to design its within class and between-class scatter matrices [6] [7].

**Advantage:** Suitable for handling linear problems employment of a suitable regularization Adjustment of the extracted features can be done.

**Disadvantage:** NDA becomes hardy on unequal covariance or complexly distributed data May not be suitable for non-linear problems. To extract the non-linear features kernel version of the algorithms have been used.

**5. Linear Spectral Mixture Analysis (LSMA):** Linear Spectral Mixture Analysis is widely used to address the mixed pixel problem. In hyperspectral image analysis, LSMA is commonly used for subpixel detection and mixed pixel classification. It is based on the assumption that the image pixel is linearly mixed by materials (referred to as targets). It determines the component parts of mixed pixels and identifies the proportion of a pixel which belongs to a particular class or feature based on the spectral characteristics of the end members where end members are the 'pure' spectra corresponding to each of the classes [9].

**Disadvantage:** Sensitive to target knowledge and noise. Abundance fraction of targets present in an image are assumed to be unknown that needs to be estimated. To address these issues some algorithms are modified and developed like Independent component analysis (ICA) based linear spectral random mixture analysis (LSRMA) assuming abundance fraction as random.

**6. Wavelet-based Methods:** Wavelet-based reduction for hyperspectral imagery is based on the fact that the spectral data from original feature space is transformed to a reduced feature space. The main principle of this method is to apply Discrete Wavelet Transform (DWT) to hyperspectral data in spectral domain and at each pixel location. Here, wavelet reduced data represent a spectral distribution similar to the original distribution but in compressed form [8].

**Disadvantage:** This method is Lossy compression method, Removed high frequency signal may contain useful information for class separation and identification.

**7. Independent component analysis (ICA):** Linear transformation method is independent component analysis (ICA), in which the desired representation is the one that minimizes the statistical dependence of the components of the representation [10]. Such a representation seems to capture the essential structure of the data in many applications. As the name implies, the basic goal is to find a transformation in which the components are statistically as independent from each other as possible. The classification accuracy depends on the number of ICs used. The computational time of SVM and ICDA depends on the size of training set. When the training samples used are of medium or larger ones, the computational burden of ICDA method is smaller as compared to SVM.

**Advantage:** Redundancy reduction is achieved, Noise free case which includes Find 'interesting' projections and Special case of projection pursuit.

**8. Weighted sum based image fusion :** This method is fast and effective without multi-scale analysis. The method can be applied for other image fusion applications. In the case of motion detection method, there is no need for the user to assign a reference image for the detection of motion [11]. A good quality fused images could be obtained. This

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 7, July 2016

provides an important role in removing the influence of pixels from the motion objects. It could be applied for infrared and visual image fusion.

**9. Image Fusion with recursive filtering:** Feature extraction is known to be an effective way in both reducing computational complexity and increasing accuracy of Hyperspectral image classification. This feature extraction is a simple yet quite powerful method[1]. First, the hyper spectral image is partitioned into multiple subsets of adjacent hyper spectral bands. Then, the bands in each subset are processed by using image fusion. The fused bands are processed with recursive filtering to get the resulting features for classification. This method give better classification and accurate result. In band partitioning step the adjacent bands are partitioned in to groups. To this group simple averaging fusion method is applied. So from 3d images to 2d image is converted. Next to these fused images transform domain recursive filtering is applied to extract features. Now IFRF (Image Fusion Recursive Filtering) features are extracted and fed in to classifiers. This method efficiently reduces dimensionality and produces output fastly.

### III. CONCLUSION

This paper provides a survey of various dimensionality reduction techniques in terms of feature extraction methods for hyperspectral images. In feature extraction methods, original data are transformed and compact set of original features are generated. Commonly used feature extraction methods used such as principal component analysis, minimum noise fraction transform, Discriminant analysis, non-parametric weighted feature extraction, wavelet transform, and spectral mixture analysis have been discussed for hyperspectral images. By comparing all the feature extraction methods comparatively better method is image fusion with recursive filtering. Also conclude that by making changes in filtering method of image fusion with recursive filtering method further improvement in accuracy can be obtained.

### REFERENCES

- [1] Xudong Kang, Shutao Li, and Jón Atli Benediktsson, "Feature Extraction of Hyperspectral Images With Image Fusion and Recursive Filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3742 – 3752, 2014.
- [2] Craig Rodarmel and Jieshan, "Principal Component Analysis for Hyperspectral Image Classification," *Surveying and Land Information Systems*, vol. 62, no. 2, 2002.
- [3] Rhonda D. Phillips and Layne T. Watson, Christine e. Blinn and Randolph H. Wynne, "An Adaptive Noise Reduction Technique for Improving The Utility of Hyperspectral Data," *Pecora17 – The future of land imaging. Going operational November 18 – 20, 2008.*
- [4] Qian Du, Chein-I Chang, "A Linear Constrained Distance-based Discriminant Analysis for Hyperspectral Image Classification" *Pattern Recognition* 34 (2001).
- [5] Tatyana V. Bandos, Lorenzo Bruzzone and Gustavo Camps-Valls, "Classification of Hyperspectral Images With Regularized Linear Discriminant Analysis," *IEEE transactions on geoscience and remote sensing*, vol. 47, no. 3, March 2009.
- [6] Jinn-Min Yang, Pao-Ta Yu, Bor-Chen Kuo, and Ming-Hsiang Su, "Nonparametric Fuzzy Feature Extraction for Hyperspectral Image Classification," *International Journal of Fuzzy Systems*, vol. 12, no. 3, September 2010.
- [7] Bor-Chen Kuo and David A. Landgrebe, "Hyperspectral Data Classification Using Nonparametric Weighted Feature Extraction".
- [8] B.Salehia and M.J.Valadanoej, "Wavele-based Reduction of Hyperspectral Imagery".
- [9] Chein-I Chang, Shao-Shan Chiang, James A. Smith and Irving W. Ginsberg, "Linear Spectral Random Mixture Analysis for Hyperspectral Imagery," *IEEE transactions on geoscience and remote sensing*, vol. 40, no. 2, February 2002.
- [10] A. Villa, J. A. Benediktsson, J. Chanussot, and C. Jutten, "Hyperspectral image classification with independent component Discriminant analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 12, pp. 4865–4876, Dec. 2011.
- [11] S. Li and X. Kang, "Fast multi-exposure image fusion with median filter and recursive filter," *IEEE Trans. Consum. Electron.*, vol. 58, no. 2, pp. 626–632, May 2012.