

Dimensionality Reduction of Hyperspectral Images

Rishabh Chhabra¹, Shubham Raj² and Shashank Jaiswal³

Abstract—As the use of hyperspectral imaging techniques increases for various applications like astronomy, agriculture, biomedical imaging and mineralogy it becomes extremely important to reduce the complexity of the data obtained. Complex data increases computation time by several manifolds and hence it is difficult to process this data. One way to reduce the complexity of this data is to reduce the dimensionality of the hyperspectral images. This paper focuses on two linear algorithms for the same, name Principal Component Analysis (PCA) and Information Gain (IG).

I. INTRODUCTION

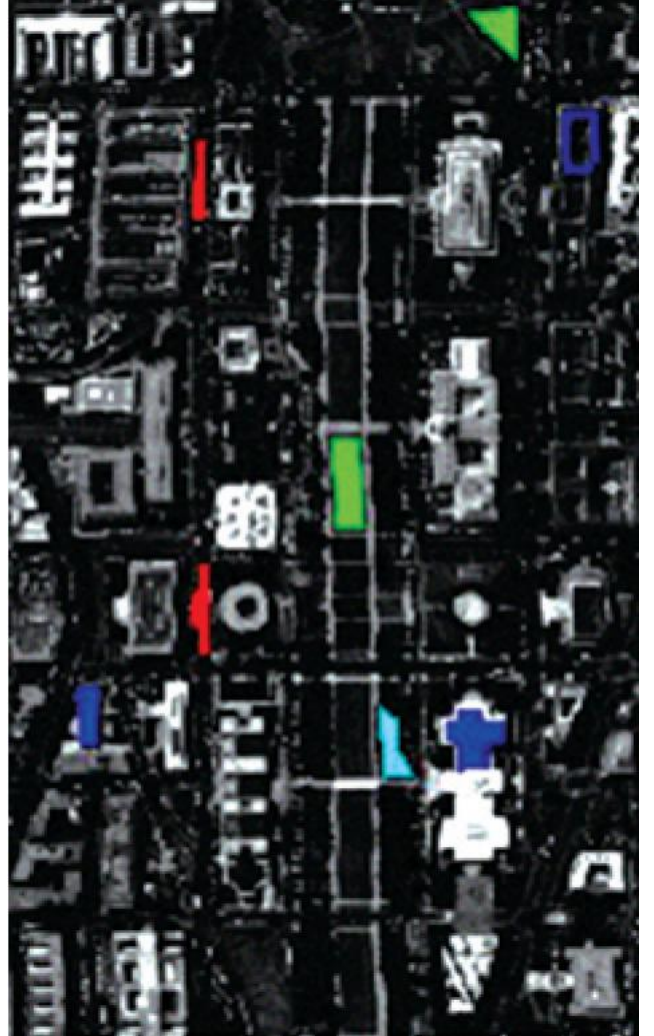
Satellites transacts huge volumes of data between different sources. So reduction of these stack is necessary. Hyperspectral data use a number of dimensions which can be reduced by mapping them into lower order of dimensions without actually losing any considerable amount. Principal Component Analysis (PCA) is perhaps the most popular dimension reduction technique for hyperspectral data. The growth in data volumes due to the large increase in the spectral bands and high computational demands of PCA has prompted the need to develop a fast and efficient algorithm for PCA. In this work, we present on an implementation of Information Gain with PCA dimension reduction of hyper spectral data

¹Rishabh Chhabra, 13116055, Department of Electronics and Communication Engineering, rishabh.chhabra561@gmail.com

²Shubham Raj, 13116064, Department of Electronics and Communication Engineering, shubham.raj.in@gmail.com

³Shashank Jaiswal, 13116059, Department of Electronics and Communication Engineering, shashankjais2008@gmail.com

Fig. 1. Hydice



II. PRINCIPAL COMPONENT ANALYSIS

It is the optimal linear scheme for reducing a set of high dimensional vectors into a set of lower dimensional vectors. The objective of PCA is to perform dimensionality reduction while preserving as much of the randomness in the high-dimensional

space as possible. These lower dimensional vectors form a new coordinate system, called the 'principal components'. Given a set of data on n dimensions, PCA aims to find a linear subspace of dimension d lower than n such that the data points lie mainly on this linear subspace. There are two types of methods for performing PCA, the matrix method, and the data method. In this work, we will focus on the matrix methods.

- 1) Compute covariance matrix $C_X = \frac{1}{n}XX^T$
- 2) We select the matrix P to be a matrix where each row p_i is an eigenvector of $C_X = \frac{1}{n}XX^T$
- 3) If A is a square matrix, a non-zero vector v is an eigenvector of A if there is a scalar λ such that $Av = \lambda v$
- 4) Reduction: there are m eigenvectors, we reduce from m dimensions to k dimensions by choosing k eigenvectors related with k largest eigenvalues λ

$$k = \frac{\lambda_1 + \lambda_2 + \dots + \lambda_k}{\lambda_1 + \lambda_2 + \dots + \lambda_k + \dots + \lambda_m}$$

Stop at PoV ≥ 0.9 . Scree graph plots of PoV vs k , stop at elbow.

III. INFORMATION GAIN

Information Gain is a measure of dependence between the feature and the class label. It is one of the most popular feature selection techniques as it is easy to compute and simple to interpret. The information gain of a given attribute X with respect to the class attribute Y is the reduction in uncertainty about the value of Y when we know the value of X , $IG(X, Y)$. So information gain of a feature or band X and the class labels Y is calculated as

$$IG(X, Y) = H(X) - H(X/Y)$$

Entropy (H) is a measure of the uncertainty associated with a random variable. $H(X)$ and

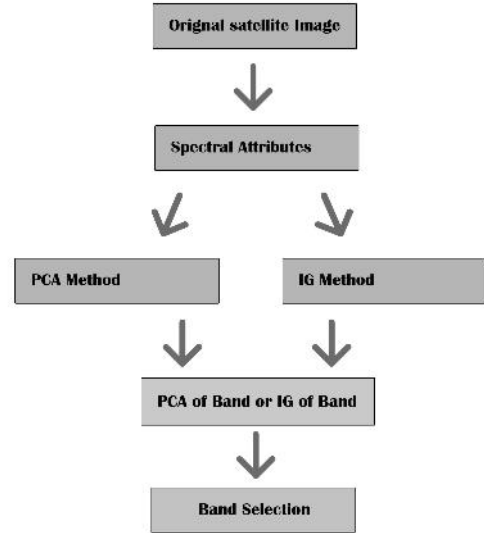
$H(X|Y)$ is the entropy of band X and the entropy of band X after observing Class Y , respectively calculated as—

$$H(X) = - \sum p(x_1) \log_2(p(x_1))$$

$$H(X/Y) = - \sum p(y_1) \sum p(x_1/y_1) \log_2(p(x_1/y_1))$$

The maximum value of information gain is 1. A feature with a high information gain is relevant. Information gain is evaluated independently for each feature and the features with the top- k values are selected as the relevant features. Information Gain does not eliminate redundant features.

Fig. 2. Combining PCA and IG



IV. COMBINING PCA AND IG

Dimension reduction is divided into two main domains, feature selection and feature extraction. Feature selection tries to find a subset of the original higher order variables and feature extraction transforms the data in high dimensional space to a lower dimensional. A method combining the approaches taken by both the techniques is proposed such that

X Band Selected = PCA of Band \cap IG of Band
This method finds a subset of lower dimensions and transforms it into further lower dimensions.

The formulation are as follows:-

For PCA

The vector representation of **image pixel vector** is given by—

$$X_i = [x_1 x_2 x_3 \dots x_n]^T$$

with all the pixel values be taken from corresponding hyperspectral dimension. For hyperspectral image with 'm' rows and 'n' columns of pixel then total will be $M=m*n$. The mean vector of all the image vector is—

$$m = \frac{1}{M} \sum [x_1 x_2 x_3 \dots x_n]^T$$

$$Cov(x) = E(x - E(x))(x - E(x))^T$$

therefore

$$C_x = \frac{1}{M} \sum (x_k - m)(x_k - m)^T$$

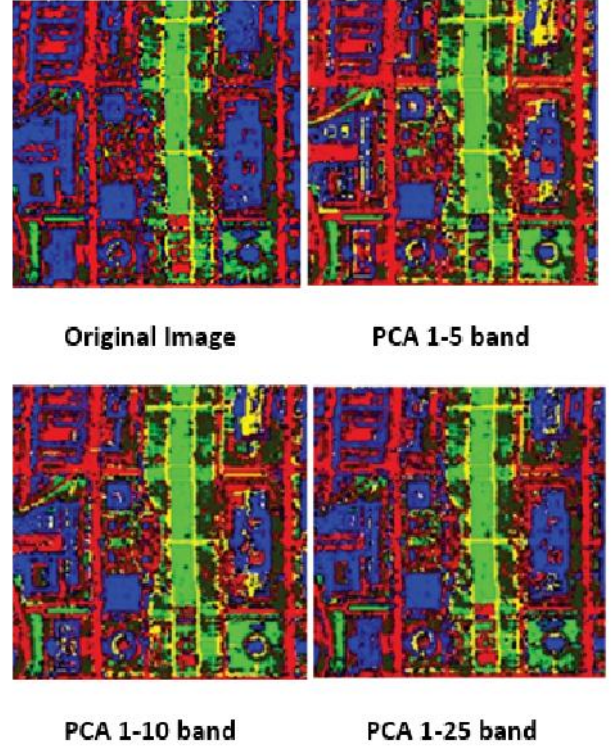
$$z_i = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_K \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1K} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2K} & \dots & a_{2N} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{K1} & a_{K2} & \dots & a_{KK} & \dots & a_{KN} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_K \\ \vdots \\ x_N \end{bmatrix} \quad (i=1,2,\dots,M)$$

where pixel vector z_i will form the first K bands of the PCA images. Such formed PCA bands have the highest contrast or variance in the first band and the lowest contrast or variance in the last band. Therefore, the first K PCA bands often contain the majority of information residing in the original hyperspectral images and can be used for more effective and accurate analyses because the number of image bands and the amount of image noise involved are reduced. After this the new set of k bands are classified through Information gain

removing the band with less information and more noise by setting threshold gain. Hence the number of bands are reduced.

X Band Selected = PCA of Band \cap IG of Band.

Fig. 3. Combining PCA and IG



V. RESULT AND CONCLUSION

The study reveals that the combination of PCA and IG is a useful process for dimensionality reduction when compared to PCA or IG applied alone. Among all the obtained PCA bands, the first few (about 5) bands may contain most of the information contained in the entire hyperspectral image data. After the first 10 PCA bands, all other bands contain mostly noise. Classifications using these first 5 PCA bands yield accurate results and patterns of hyperspectral images. Misclassifications using PCA occur mainly due to non linear nature of data. The use of IG to further classify data significantly reduces the amount of data to be

handled and achieves practically acceptable and accurate classification results that are comparable with those obtained using the entire hyperspectral image data. The complexity of the algorithm used to implement PCA takes up most of the time but the compromise doesn't overweigh the reduction in data to be handled and thus PCA is considered as a beneficial option. Comparing the classification results done by PCA-IG is done using clustering. The clusters formed using the unreduced set of hyperspectral data and the data whose dimensions were reduced using PCA-IG techniques were found to be close. The proximity of the clusters are not as expected when the data has non linear characteristics or when any of the assumptions including those of principal components being orthogonal are not satisfied. Future research efforts will include adding more quantitative and qualitative measures to the PCA-IG technique to further refine the process of dimensionality reduction.

REFERENCES

- [1] Craig Rodarmel and Jie Shan, Surveying and Land Information Systems, Vol. 62, No. 2, 2002.
- [2] J. A. Richards, Remote Sensing Digital Image Analysis: An introduction, Springer-Verlag, Berlin, Heidelberg, 1986.
- [3] Small Multi-Mission Satellite (SMMS) Data Available to: <http://smms.ee.ku.ac.th/index.php>
- [4] Agarwal, A.; El-Ghazawi, T.; El-Askary, H.; Le-Moigne, J.; , "Efficient Hierarchical-PCA Dimension Reduction for Hyperspectral Imagery," Signal Processing and Information Technology, 2007 IEEE International Symposium on , vol., no., pp.353-356, 15-18 Dec. 2007.
- [5] Kaewpijit S., Le-Moige J., El-Ghazawi T., "Hyperspectral Imagery Dimension Reduction Using Principal Component Analysis on the HIVE," In Science Data Processing Workshop, NASA Goddard Space Flight Center, Feb. 2002.
- [6] Padraig Cunningham, Dimension Reduction," Technical Report on Dimension Reduction, University College Dublin, 2007.
- [7] T. M. Cover and J. A. Thomas, "Information Gain," Elements of Information Theory. Wiley, 1991.
- [8] J. E. Jackson, A User Guide to Principal Components, New York: John Wiley and Sons, 1991.
- [9] Richard Kirkby and Eibe Frank, Weka Explorer User Guide, University of Waikato, New Zealand, 2005.
- [10] Frank, A. and Asuncion, A. UCI Machine Learning Repository, [<http://archive.ics.uci.edu/ml/support/Statlog>]. Irvine, CA: University of California, School of Information and Computer Science, 2010.
- [11] Z. Zhao et al., "Advancing Feature Selection Research," Retrieved: Sep 2, 2010
- [12] X. Cheng, Y. R. Chen, Y. Tao, C. Y. Wang, M. S. Kim, A. M. Lefcourt. A novel integrated PCA and FLD method on hyperspectral image feature extraction for cucumber chilling damage inspection. ASAE Transactions, Vol. 47(4), 1313-1320, 2004.