SPARSE REPRESENTATION BASED BAND SELECTION FOR HYPERSPECTRAL IMAGES

Shuangjiang Li, Hairong Qi

Electrical Engineering and Computer Science
University of Tennessee, Knoxville, TN 37996
Email: {shuangjiang, hqi}@utk.edu

ABSTRACT

Hyperspectral images consist of large number of spectral bands but many of which contain redundant information. Therefore, band selection has been a common practice to reduce the dimensionality of the data space for cutting down the computational cost and alleviating from the Hughes phenomenon. This paper presents a new technique for band selection where a sparse representation of the hyperspectral image data is pursued through an existing algorithm, K-SVD, that decomposes the image data into the multiplication of an overcomplete dictionary (or signature matrix) and the coefficient matrix. The coefficient matrix, that possesses the sparsity property, reveals how importantly each band contributes in forming the hyperspectral data. By calculating the histogram of the coefficient matrix, we select the top K bands that appear more frequently than others to serve the need for dimensionality reduction and at the same time preserving the physical meaning of the selected bands. We refer to the proposed band selection algorithm based on sparse representation as SpaBS. Through experimental evaluation, we first use synthetic data to validate the sparsity property of the coefficient matrix. We then apply SpaBS on real hyperspectral data and use classification accuracy as a metric to evaluate its performance. Compared to other unsupervised band selection algorithms like PCA and ICA, SpaBS presents higher classification accuracy with a stable performance.

***Index Terms*—** Band selection, Sparse representation, Hyperspectral imaging, Image classification

1. INTRODUCTION

Hyperspectral images have been proven beneficial to many different applications, including remote sensing, medical imaging, and quality assurance, to name but a few. Modern remote sensors are producing hyperspectral images which sample hundreds of contiguous narrow spectral bands. However, the challenging problems, like the heavy computational load and the Hughes phenomenon [1], also arise due to the resulting high-dimensional data sets. Band selection is an alternative way to conquer these problems. Unlike feature extraction, band selection tries to identify a subset of original bands through some search strategy and evaluation criterion while keeping the physical meanings of bands unchanged. Therefore, it is generally more preferable.

Fourier, Wavelet, Discrete Cosine transformation, etc. have been widely used in signal and image processing, where by changing the basis of the signal space, these transform domain processing methods can better separate noise from salient features to facilitate feature detection, segmentation, classification and to reduce computational requirements. Inspired by recent work in compressive sensing [2], we propose a technique that transforms the hyperspectral data into a specifically designed basis, maintaining fewest large coefficients and many small or zero coefficients; thus, the hyperspectral data is sparse represented. Sparse representations have increasingly become recognized as providing extremely high performance for applications as diverse as noise reduction, feature extraction, pattern classification and blind source separation. The aim of this transformation is to reveal certain structures of an image and to represent these structures in a compact and sparse representation. Figure 1 uses cuprite image data [3] with totally 188 bands at pixel coordinate (1, 2). The x-axis is the spectral wavelength for totally 188 bands and y-axis is the spectral reflectance in the spatial domain (for Fig. 1(a)) or coefficients in the transformed domain (for Fig. 1(b-d)). We see that Fourier coefficients and Haar wavelet coefficients both enjoy a high level of sparsity.

In this paper, we present a sparse representation based method for band selection, referred to as SpaBS. The algorithm first finds a sparse representation of the hyperspectral image data and then a band ranking criterion using majority voting is applied to the coefficient matrix obtained by setting specific sparsity level. The proposed method avoids transforming the original hyperspectral images to a feature space on which a physical interpretation is not possible. Instead, it tries to gain the large absolute weight coefficients of individual spectral bands and selects the interested bands which contain the maximum information, thereby reducing the dimensionality but retaining most spectral features of hyperspectral images. A novel K-SVD algorithm [4] that generalizing the K-Means clustering processing is adopted in order to achieve sparse representation of hyperspectral data.

The rest of the paper is organized as follows. Section II

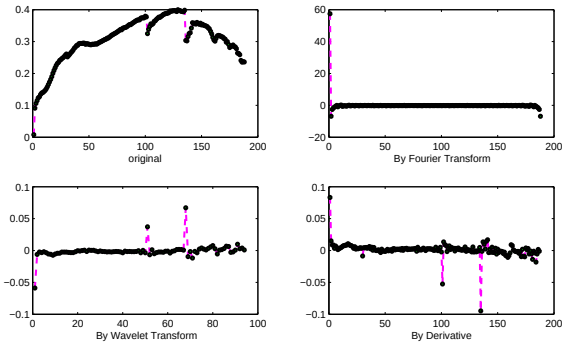

Fig. 1. Transformation of Cuprite image data at pixel coordinate (1,2) (a) Original reflectance (upper-left) (b) reflectance transformed by Fourier transformation (upper-right) (c) reflectance transformed by Haar wavelet transformation (lower-left) (d) reflectance transformed by taking derivative of original image (lower-right).

introduces related work in band selection for hyperspectral image data. Section III reviews the sparse representation of a signal and the K-SVD algorithm. Section IV presents the sparse representation based band selection (SpaBS) method. Various experiments are conducted in Section V to evaluate the performance of SpaBS compared to other selection methods and conventional feature extraction methods. Section VI concludes the paper.

2. RELATED WORK

For many years, the design of efficient and robust feature extraction and feature selection, especially band selection, algorithms has been the most important issue addressed by the remote sensing community. Strong efforts have been devoted to elaborate new band selection algorithms and improve techniques used to reduce dimensionality.

From the selection metric point of view, the band selection algorithms can be categorized into two classes. One class is based on information-theoretic measures, such as entropy and combined entropy, covariance matrix of combined bands [5]. The other class is based on separability of the same classes, such as dispersibility, standard distance between mean values, B distance, and so on [6, 7].

While all these methods try to utilize the existing information in the given spectral bands to find the relation between bands and make them distinguishable, most recently, there has been research focusing on transfer domain processing to attain desirable features. In [8], characterizations in the frequency domain are presented and a very small number of hyperspectral bands are needed to perform classification.

In this paper, we present a sparse representation based band selection (SpaBS) method that maps the bands into a non-physical domain that enjoys the properties of noise reduction, compression yet preserving the most salient information desired. SpaBS uses the Sparse Representation which describes the importance of different spectral bands as the weighted index, and then uses the sparsity level and majority rule as the band selection criteria.

3. SPARSE REPRESENTATION AND K-SVD ALGORITHM

The problem of finding the sparse representation of a signal in a given overcomplete dictionary can be formulated as follows. Given an $L \times K$ matrix $\mathbf{A}$ containing the elements of an overcomplete dictionary in its columns, with $L > K$ and usually $L \gg K$, and a signal $\mathbf{y} \in \mathbb{R}^L$, the problem of sparse representation is to find an $K \times 1$ coefficient vector $\mathbf{x}$, such that $\mathbf{y} = \mathbf{A}\mathbf{x}$ and $||\mathbf{x}||_0$ is minimized, i.e.,

$$\mathbf{x} = \min_{\mathbf{x}'} ||\mathbf{x}'||_0 \qquad s.t. \qquad \mathbf{y} = \mathbf{A}\mathbf{x} \tag{1}$$

where $||\mathbf{x}||_0$ is the l_0 norm and is equivalent to the number of non-zero components in vector $\mathbf{x}$. Finding the solution to Eq. 1 is NP hard due to its nature of combinational optimization. Suboptimal solutions to this problem can be found by iterative methods like the matching pursuit and orthogonal matching pursuit. An approximate solution is obtained by replacing the l_0 norm in Eq. 1 with the l_1 norm, as follows:

$$\mathbf{x} = \min_{\mathbf{x}'} ||\mathbf{x}'||_1 \qquad s.t. \qquad \mathbf{y} = \mathbf{A}\mathbf{x} \tag{2}$$

where $||\mathbf{x}||_1$ is the l_1 norm. In [9], it is proved that if certain conditions on the sparsity is satisfied, i.e., the solution is sparse enough, the solution to Eq. 1 is equivalent to the solution to Eq. 2, which can be efficiently solved by basis pursuit using linear programming. A generalized version of Eq. 2, which allows for certain degree of noise, is to find $\mathbf{x}$ such that the following objective function is minimized:

$$J_1(\mathbf{x}; \lambda) = ||\mathbf{y} - \mathbf{A}\mathbf{x}||_2^2 + \lambda||\mathbf{x}||_1 \tag{3}$$

where the parameter λ is a scalar regularization parameter that balances the tradeoff between reconstruction error and sparsity.

In this paper we adopt the K-SVD algorithm - a generalization of the k-means algorithm, to obtain the sparse representation of the hyperspectral image data, in an unsupervised manner. A detailed description of the algorithm can be found in [4].

4. SPARSITY BASED BAND SELECTION ALGORITHM

In this section, the proposed SpaBS algorithm is introduced. Suppose the spatial dimension of the hyperspectral image data is $M \times N$ and the spectral dimension is L, then we

construct the observation matrix $\mathbf{Y} = [\mathbf{y}_1, \cdots, \mathbf{y}_{MN}]$ where $\mathbf{y}_i, i = 1, \cdots, MN$ is an L-dimensional column vector, representing the spectral reflectance of each pixel. We apply the K-SVD algorithm to calculate a basis (or the dictionary), $\mathbf{A}$ of $L \times L$ dimension, and the coefficient matrix, $\mathbf{X}$ of $L \times MN$ dimension, corresponding to that basis. That is,

$$\mathbf{Y}_{L\times MN} = \mathbf{A}_{L\times L}\mathbf{X}_{L\times MN}.$$

Note that different from the description in Sec. 3 where $\mathbf{A}$ is an overcomplete dictionary, for band selection purpose, we need to set $\mathbf{A}$ as a square matrix of $L \times L$ dimension, thus the coefficient matrix $\mathbf{X}$ would be of the same dimension as the original hyperspectral image, i.e., $L \times MN$. The reason for doing this would become clear through the following description of the analysis algorithm of the coefficient matrix $\mathbf{X}$.

Since $\mathbf{X}$ is derived with the sparsity constraint, most of entries in $\mathbf{X}$ would be equal to or very close to zero. Each column of the coefficient matrix indicates how importantly that each basis contributes in forming each column of the original hyperspectral image. A larger coefficient in $\mathbf{X}$ would mean where the majority information resides. For each column of the coefficient matrix, we select K largest entries out of the L coefficients and calculate the histogram of the corresponding indices of the kept K entries. Here, K is defined by $K = \delta_s \times L$ where δ_s is a pre-defined sparsity level. In our experimental setup, we intentionally increase δ_s slightly to account for minor information.

Algorithm 1 Sparsity based band selection (SpaBS) algorithm

Input:

L-band hyperspectral image $\mathbf{Y}$ in $\mathbb{R}^{L\times MN}$;

Sparsity level δ_s;

Output:

K-band hyperspectral image;

1: Apply K-SVD algorithm on $\mathbf{Y}$ to obtain the coefficient matrix $\mathbf{X}$, $\mathbf{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_{MN}]$;
2: Sort $\mathbf{x}_i$ in the descending order, $i \in [1, \cdots, MN]$;
3: Set K equal to $\delta_s \times L$, select first K entries in each column of $\mathbf{X}$ and use the corresponding indices to form matrix $\mathbf{X}_s$, $\mathbf{X}_s \in \mathbb{R}^{K\times MN}$, $K \ll L$;
4: Calculate the histogram of matrix $\mathbf{X}_s$ that indicates how frequently each band appears in $\mathbf{X}_s$;
5: Select K indices that appear more often according to the histogram and keep the corresponding K spectral bands from the original image data;
6: **return** K-band hyperspectral image data.

5. EXPERIMENTS

5.1. Experimental Results Using Synthetic Data

We first use synthetic data to test the capability of the K-SVD algorithm in generating the relevant bands of which most are sparse. We select 5 signature profiles representing five independent materials from [10]. Each signature profile is of 100 dimension. That is, the signature matrix or the dictionary is of 100×5 dimension. We then generate a sparse coefficient matrix of $5 \times 3,364$ dimension. By multiplying the signature matrix and the coefficient matrix, we construct the synthetic hyperspectral image of 100-band with a $58 \times 58 = 3,364$ spatial resolution. We apply the SpaBS algorithm and set the sparsity level to 5%. Table 1 shows the selected bands and their histogram in descending order. Figure 2 also shows that five bands out of 100 were successfully selected and the coefficients map are exactly the same. We can see that as long as the data enjoys some level of sparsity, the SpaBS algorithm could find a sparse representation of the data and obtain majority of the information at the same time.

Band Number	61	2	1	9	3	4
Histograms	222	178	177	176	166	64

Table 1. Selected bands and its histogram

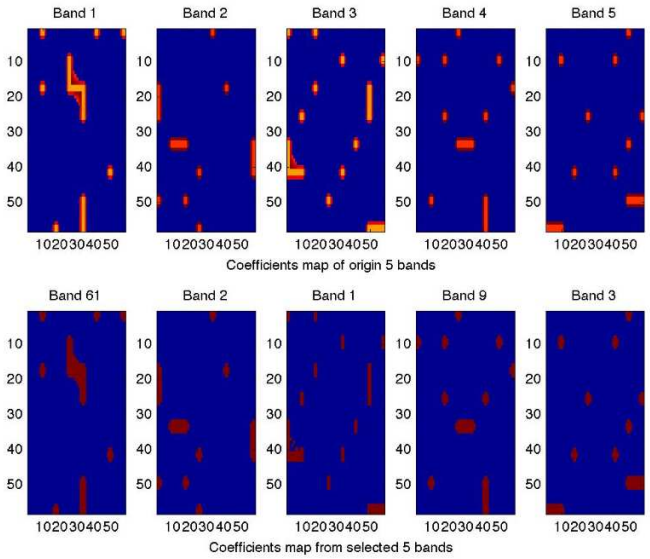

Fig. 2. (Top) The five 58×58 coefficients map to construct the original synthetic data, (bottom) coefficients map corresponding to five selected bands using the SpaBS algorithm.

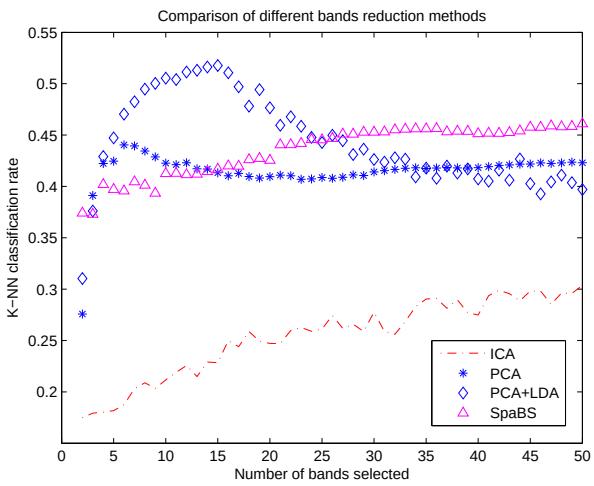

Fig. 3. Classification rate with different sparsity levels.

5.2. Experimental Results Using Real Hyperspectral Scene

The image used here is the 92AV3C image from the NASA Jet Propulsion Laboratory, which consists of 145 by 145 pixels at 224 spectral bands. Four of the 224 spectral bands in the 92AV3C image contain zero values leaving a total of 220 non-zero bands. This image was chosen because it includes the necessary ground truth reference data needed for classification and accuracy assessments.

We apply the k-NN classifier to evaluate the performance of SpaBS using the classification rate. The parameter k[1] is set to 10. We have also tested other selection of k but found no major difference in classification accuracy. A 10-fold cross validation is adopted on the 10,366 pure pixels in one band. The classification results of various dimensionality reduction algorithms with different sparsity levels and compared and the results are shown in Fig. 3.

From Fig. 3, we could see that the SpaBS algorithm outperforms ICA overwhelmingly and catches up with PCA when the number of selected bands increases to around 12. While the classification rate of PCA+LDA is slightly higher than SpaBS when the number of bands is small, the performance is not stable. After increasing the number of bands to 30, SpaBS obtains higher classification rate with stability.

6. CONCLUSION

In this paper, we presented a sparse representation based band selection algorithm. Through experiments conducted on both synthetic and real image data, we observed that sparsity based algorithm can keep important information in the original image to obtain satisfactory classification purpose, while at the same time preserving the physical meaning of the kept information. Through sparse representation, SpaBS provides more stable performance and has the potential of handling noise, missing data, and outliers.

[1] Not to be confused with Capital K, which is number of bands selection in the context.

7. REFERENCES

[1] G. Hughes, "On the mean accuracy of statistical pattern recognizers," *Information Theory, IEEE Transactions on*, vol. 14, no. 1, pp. 55–63, 1968.

[2] Richard G. Baraniuk, "Compressive sensing," *IEEE Signal Processing Magazine*, pp. 118–124, JULY, 2007.

[3] G. Swayze, R. Clark, F. Kruse, S. Sutley, and A. Gallagher, "Ground-truthing AVIRIS mineral mapping at Cuprite, Nevada," in *JPL, Summaries of the Third Annual JPL Airborne Geoscience Workshop.*, 1992, vol. 1.

[4] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation," *Signal Processing, IEEE Transactions on*, vol. 54, no. 11, pp. 4311–4322, 2006.

[5] A. Martınez-Usó, F. Pla, P. Garcıa-Sevilla, and J.M. Sotoca, "Automatic Band Selection in Multispectral Images Using Mutual Information-Based Clustering," .

[6] X. Xu, "A new approach to band clustering and selection for hyperspectral imagery," in *Signal Processing, 2008. ICSP 2008. 9th International Conference on*. IEEE, 2008, pp. 1198–1202.

[7] S. De Backer, P. Kempeneers, W. Debruyn, and P. Scheunders, "A band selection technique for spectral classification," *Geoscience and Remote Sensing Letters, IEEE*, vol. 2, no. 3, pp. 319–323, 2005.

[8] O. Rajadell, P. García-Sevilla, and F. Pla, "Textural features for hyperspectral pixel classification," *Pattern Recognition and Image Analysis*, pp. 208–216, 2009.

[9] D.L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *Information Theory, IEEE Transactions on*, vol. 47, no. 7, pp. 2845–2862, 2002.

[10] C.I. Grove, S.J. Hook, and E.D. Taylor, "Laboratory reflectance of 160 minerals, 0.4 to 2.5 micrometers," *Passadena, National Aeronautics and Space Administration*, p. 355p, 1992.