HYPERSPECTRAL BAND SELECTION USING A COLLABORATIVE SPARSE MODEL

Qian Du [1], *José M. Bioucas-Dias* [2], *Antonio Plaza* [3]

[1] Department of Electrical and Computer Engineering, Mississippi State University, USA
[2] Instituto de Telecomunicações, Instituto Superior Técnico, TULisbon, Portugal
[3] Department of Technology of Computers and Communications, University of Extremadura, Spain
du@ece.msstate.edu

ABSTRACT

In our previous research, we have proposed band-similarity-based unsupervised band selection approaches, which are proven to be very efficient. In this paper, we propose to use a collaborative sparse model for further improvement. Specifically, the pre-selected bands using the fast method, called NFINDR+LP, are further refined using a collaborative sparse model. It not only requires that the linear regression coefficients are sparse, but also requires that the same set of active bands is shared by all the bands to be removed. With the collaborative sparseness constraint being relaxed, the final selected bands can be further improved, that is, the band subset with the same number of bands can provide better classification accuracy. Based on the preliminary result, the proposed sparse model is also capable of finding the minimum number of bands to be selected.

Index Terms — hyperspectral imaging, support vector machine-based classification, band selection.

1. INTRODUCTION

Band selection is a frequently used dimensionality reduction technique for hyperspectral imagery. It selects a subset of original bands without losing their physical meaning. Compared to supervised band selection techniques, unsupervised methods need no prior information about objects or classes [1-2]. In general, they are more practical than supervised methods. However, unsupervised methods may need to analyze the whole dataset, resulting in higher computation complexity than supervised ones, which often consider a limited number of object signatures or class samples.

In this research, we focus on unsupervised band selection. Since the basic idea of unsupervised band selection methods is to find the most distinctive and informative bands, the approaches proposed to search for distinctive spectral signatures as endmembers are natural candidates. The major difference is that the algorithms are applied in the spatial domain for band selection instead of in the spectral domain for endmember extraction. There exist quite a few endmember extraction algorithms. In general, endmember extraction algorithms can be divided into two categories: one extracting distinctive pixels based on similarity measurement, and the one exploiting geometric concept associated with the data sets. The endmember extraction algorithm using unsupervised fully constrained linear unmixing in [3] belongs to the first category, while the well-known N-FINDR algorithm [4] belongs to the second category. In [1], the concept of N-FINDR was applied to band selection and obtained promising results. In [2], we have proposed a band selection algorithm using band linear prediction (LP), which used the similar idea of unconstrained least squares linear unmixing in endmember selection. We have demonstrated that the LP-based method in conjunction with data whitening can outperform other widely used band selection approaches [2]. Thus, we will focus on this method hereafter.

For supervised band selection, the number of bands to be selected is somewhat related to the number of classes [8]. However, for an unsupervised band selection method, such as the LP-based method, this number is difficult to pre-estimated, which usually is changed from time to time. It is preferred to have a smaller band subset, given that data information is well maintained. In this paper, we will propose an approach based on a collaborative sparse model to further refine the LP-selected bands such that the number of bands to be selected can be reduced.

2. PREVIOUS WORK

To select the most distinctive but informative bands, water absorption and low signal-to-noise ratio (SNR) bands need to be pre-removed. This is because these bands are distinctive but not informative. To select the distinctive bands or the most dissimilar bands, a similarity metric needs to be designated. Widely used metrics include distance and correlation, which works on pairs of bands. Here, however, band similarity is evaluated jointly instead of pair-wisely. The proposed band selection methodology, of endmember extraction type, has this property. In addition, due to the large number of original bands, the exhaustive search for optimal band combinations is computationally prohibitive. To save significant computation time, we adopt a sequential forward search (SFS) technique. It begins with the best two-band combination, and it is subsequently augmented to three, four, and so on, until the desired number of bands is selected.

The straightforward criterion that can be employed for similarity comparison is Linear Prediction (LP), which can jointly evaluate the similarity between a single band and multiple bands. Let $\Phi=\{B_0, B_1, B_2\}$ denote a selected set with three bands with N pixels each: B_0 is a column vector of ones and B_1 and B_2 are two initial bands. To find a band that is the most dissimilar to Φ, we project all bands not in Φ into the subspace generated by the elements of Φ and augment Φ with the band having the larger projection error. Band B_0 is included to allow shifts in the subspace. In summary, we have

do

1. U = orth(Φ)
   % compute an orthogonal basis for Φ
2. **for** all bands B not in Φ
   compute W= $(I-UU^T)B$
   % W is the projection error of B onto
   % span{Φ}
3. $\Phi = \Phi \cup \{B_l\}$
   % B_l is the band with larger projection error

while some stop rule is met

The major problem of the LP-based band selection is that computational cost is high if all the pixels are used. The N-FINDR algorithm can be applied for pixel selection, and then the selected pixels are used for band selection [5]. The resulting N-FINDR+LP method provides performance comparable to that of just LP.

3. PROPOSED BAND SELECTION WITH COLLABORATIVE SPARSENESS CONSTRAINT

Due to the use of SFS, the selected band subset from LP or N-FINDR+LP is suboptimal. We also try the more advanced sequential floating forward search (SFFS); however, its performance is not very stable. Here, we propose a new method to further optimize the band subset. The basic idea is to use the fast N-FINDR+LP method to do band pre-selection, and then apply an efficient linear sparse regression (SR) technique to refine the band selection result.

Let $\mathbf{z}$ represent an $N\times1$ unselected band with N pixels, $\mathbf{A}$ be $N\times M$ data matrix including M preselected bands, and $\mathbf{s}$ denote an $M\times1$ coefficient matrix. Assuming that the unselected band is well represented by M preselected bands, then we have

$$\mathbf{z} = \mathbf{A}\mathbf{s}. \tag{1}$$

To minimize the number of selected bands, we formulate the Basis Pursuit Denoising (BPDN) optimization problem

$$\min_{\mathbf{s}} \|\mathbf{s}\|_1 \text{ subject to } \|\mathbf{z}-\mathbf{A}\mathbf{s}\|_2 \le \delta \tag{2}$$

which is known to yield sparse solutions. Problem (2) has the equivalent unconstrained formulation

$$\min_{\mathbf{s}} \|\mathbf{z}-\mathbf{A}\mathbf{s}\|_2 + \lambda\|\mathbf{s}\|_1, \tag{3}$$

where λ is the Lagrange multiplier. For all the L unselected bands represented by an $N\times L$ matrix $\mathbf{Z}$, Eq.(3) can be rewritten as

$$\min_{S} \|\mathbf{Z}-\mathbf{A}\mathbf{S}\|_2 + \lambda\|\mathbf{S}\|_1. \tag{4}$$

where $\mathbf{S}$ is an $M\times L$ coefficient matrix.

The l_1 normin (3) and (4) applies, respectively, to the elements of $\mathbf{s}$ and $\mathbf{S}$ independently. In other words, the sparseness constraint is applied to each element of $\mathbf{S}$. In some applications, it is expected that the same active components from the dictionary are shared in an SR problem, that is, the indexes of the nonzero coefficients in $\mathbf{S}$ are the same for all the samples in the collection. Imposing such dependency gives rise to the so-called collaborative sparse coding problem [6]. With the collaborative sparse model, the entire number of selected bands can be minimized. The collaborative sparseness constrained band selection problem can be formulated as

$$\min_{\mathbf{s}} \|\mathbf{Z}-\mathbf{A}\mathbf{S}\|_2 + \lambda_2 \sum_{j=1}^{M} \|\mathbf{s}_j\|_2, \tag{5}$$

where $\mathbf{s}_j$ represents the jth row of $\mathbf{S}$.. The sparse unmixing algorithm via variable splitting and augmented Lagrangian (SUnSAL) in [7] was used to solve (5). By examining $\|\mathbf{s}_j\|_2$ for $j = 1, \cdots, M$, band selection can be finalized.

4. EXPERIMENT

The dataset used in the experiment is an Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) dataset — Cuprite. As shown in Fig. 1(a), it has been cropped spatially to a size of 350×350 and it is composed of 189 spectral bands after removing water absorption and low SNR bands. Five minerals with known signatures are of interest: alunite (A), buddingtonite (B), calcite (C), kaolinite (K), and muscovite (M). The constrained linear discriminant classifier (CLDA) was adopted for classification, and the classification maps using all the original bands are shown in Fig. 1(b).

The original LP-based band selection method using all the pixels was applied first. The spatial correlation coefficient was calculated between the corresponding classifications maps from using all the original bands and using the selected bands. Let $f_1(x, y)$ and $f_2(x, y)$ be two images with N pixels. The mean and standard deviation of $f_1(x, y)$ are denoted as m_{f_1} and σ_{f_1}, respectively; similarly, m_{f_2} and σ_{f_2} denote the mean and standard deviation of $f_2(x, y)$, respectively. Their spatial correlation coefficient can be computed as $\rho(f_1, f_2) = \dfrac{\frac{1}{N}\sum_{x,y}\left(f_1(x,y) - m_{f_1}\right)\left(f_2(x,y) - m_{f_2}\right)}{\sigma_{f_1}\sigma_{f_2}}$. The averaged correlation coefficient of five classes versus the number of selected bands was plotted in Fig. 2. Then the LP-based band selection was applied to the NFINDR-selected pixels (denoted as N-FINDR+LP). As shown in Fig. 2, the result is comparable.

The 39 N-FINDR+LP-selected bands were further refined using the approach proposed in Section III. 1% pixels were randomly selected for band final selection. As shown in Fig. 2, when the number of selected bands is larger than 28, the performance after final selection is even better than the original LP-based band selection using all the pixels. Fig. 1(c) shows the classification maps using the 28-selected bands, which are quite similar to those using all the bands in Fig. 1(b).

5. CONCLUSION

In this paper, we propose to use a collaborative sparse model for hyperspectral band selection. The pre-selected bands using the fast method (i.e., NFINDR+LP) are further refined using a collaborative sparse model. It not only requires that the linear regression coefficients are sparse, but also requires that the same set of active bands is shared by all the bands to be removed. With the collaborative sparseness constraint being relaxed, the final selected bands can be further improved, that is, the band subset with the same number of bands can provide better classification accuracy. Based on the preliminary result, the proposed method can also be used to find the minimum number of bands to be selected.

REFERENCES

[1] L. Wang, X. Jia, and Y. Zhang, "A novel geometry-based feature-selection technique for hyperspectral imagery," *IEEE Geosciencce Remote Sensing Letters*, vol. 4, no. 1, pp. 171-175, Jan. 2007.

[2] Q. Du and H. Yang, "Similarity-based unsupervised band selection for hyperspectral image analysis," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 4, pp. 564-568, Oct. 2008.

[3] D. C. Heinz, C. I. Chang, "Fully constrained least squares linear Spectral mixture analysis method for material quantification in hyperspectral imagery", *IEEE Trans. Geoscience Remote Sensing*, vol. 39, no. 3, pp. 529-545, Mar. 2001.

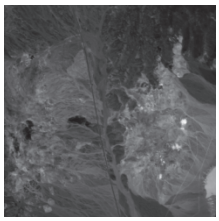
[4] M. E. Winter, "N-FINDR: An algorithm for fast autonomous spectral endmember determination in hyperspectral data," *Proc. SPIE*, vol. 3753, pp. 266-275, 1999.

[5] H. Yang, Q. Du, and G. Chen, "Unsupervised hyperspectral band selection using graphics processing units," *IEEE Journal of Selected Topics in Earth Observation and Remote Sensing*, vol. 4, no. 3, pp. 660-668, Sep. 2011.

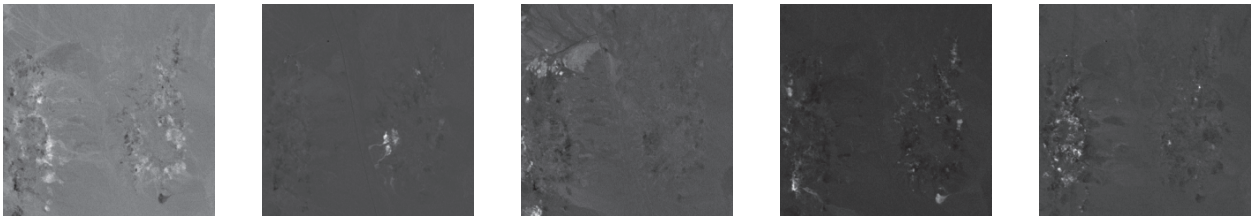
[6] P. Sprechmann, I. Ramirez, G. Sapiro, and Y. C. Eldar, "C-HiLasso: A collaborative hierarchical sparse modeling framework," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4183-4198, Sep. 2011.

[7] J. Bioucas-Dias and M. Figueiredo, Alternating Direction Algorithms for Constrained Sparse Regression: Application to Hyperspectral Unmixing, 2010, Arxiv preprint arXiv:1002.4527.

[8] H. Yang, Q. Du, and G. Chen, "Particle swarm optimization-based hyperspectral dimensionality reduction for urban land cover classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, Apr. 2012.

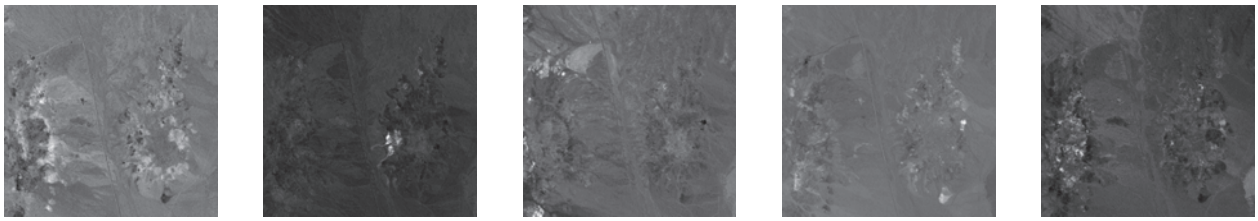


(a) An AVIRIS Cuprite scene



Alunite Buddingtonite Calcite Kaolinite Muscovite

(b) Classification maps using all the bands



Alunite Buddingtonite Calcite Kaolinite Muscovite

(C) Classification maps using 28 selected bands

Fig. 1 The image used in the experiment.

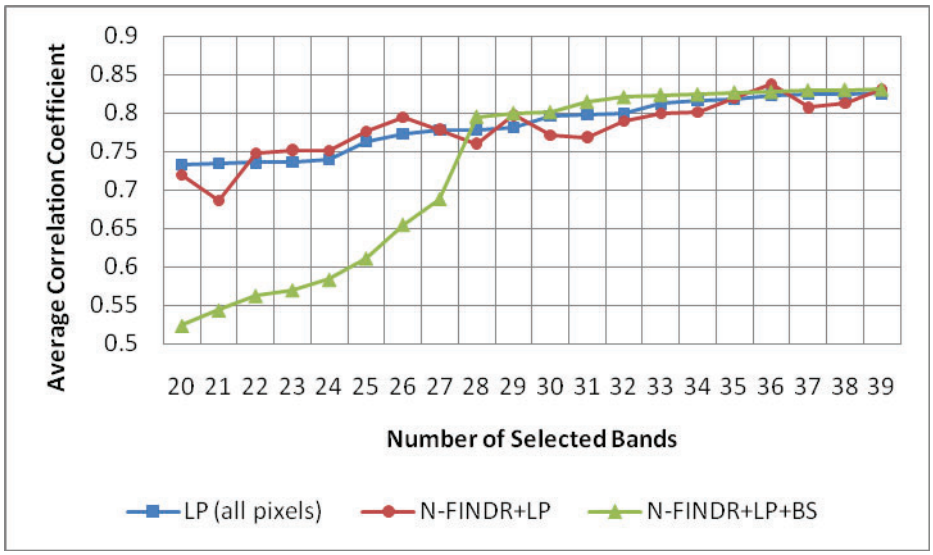


Fig. 2 Band selection performance.