CLUSTER CONSTRAINT BASED SPARSE NMF FOR HYPERSPECTRAL IMAGERY UNMIXING

Xinwei Jiang, Lei Ma, Yiping Yang

Institute of Automation, Chinese Academy of Sciences
{xinwei.jiang, lei.ma, yiping.yang}@ia.ac.cn

ABSTRACT

Nonnegative matrix factorization(NMF) has been applied to hyperspectral unmixing in recent years. Different constraints based on geometrical or statistical properties of endmember and abundance are incorporated into NMF model to improve unmixing result. In this paper, a new regularizer based on spectral cluster information is proposed to strengthen the constrained relationship between original image and abundance maps. The new algorithm makes abundances of similar pixels close and abundances of dissimilar pixels be separated completely. Additionally, $L_{1/2}$ sparsity constraint is adopted to make the solutions sparse. Comparative results on real and synthetic hyperspectral datasets prove our proposed method could improve the hyperspectral unmixing accuracy.

Index Terms— Hyperspectral imagery, linear mixing model, nonnegative matrix factorization, spectral cluster

1. INTRODUCTION

Hyperspectral unmixing(HU) is a very important issue in the hyperspectral imagery processing due to the existence of mixed pixels, which is a common phenomenon and limits the application of hyperspectral data like classification, target detection and multisource fusion. The hyperspectral unmixing is expected to find all the endmembers and their abundance maps which show the distribution of endmembers in the scene.

Linear mixing model(LMM) is most widely applied mixing model in current research field due to its simplicity and acceptable approximation for real scenarios [1]. Under the LMM, most classic algorithms are geometrical-based approaches [2–4]. These methods assume that endmembers constitute the data simplex and each endmember is the corresponding vertex. Because of their clear physical concept and light computation burden, geometrical based methods are very often used and treated as baselines. However, when pixels are highly mixed, due to lack of enough geometrical information, these algorithms degrade seriously. In these cases, statistical methods are better alternatives. Under the statistical framework, prior knowledge could be used to better solve inference problem [5]. Bayesian approaches [6–8] are representatives of this kind. Since in most cases, the endmembers and their abundance are unknown, hyperspectral unmixing are formulated as a blind source separation(BSS) problem. so Independent Component Analysis(ICA) is applied to solving the unmixing problem [9–11]. However, the abundance constraints violate the "source independence" assumption.

Recently, many researchers have proposed unmixing algorithms based on NMF because of its nonnegativity of solutions. Using NMF original hyperdata is decomposed into two nonnegative matrices: abundance matrix and endmember matrix. Meanwhile, compared with geometrical based algorithms like PPI [2], NFINDR [3] and VCA [4], unmixing based on NMF doesn't need the assumption of existence of pure pixel and both abundance matrix and endmember matrix are obtained simultaneously. However, due to the nonconvexity of objective function, the solutions are often stuck to local extrema and therefore initialization of solutions are significantly important for good unmixing results. To overcome the nonconvexity and improve solutions, different constraints are incorporated to regularize the solutions based on the properties of spectral and abundance. The representative algorithms are: Miao and Qi [12] proposed minimum volume constrained NMF(MVC-NMF) for endmember extraction. Jia and Qian [13] combines piecewise smoothness and sparseness to improve unmixing result. Liu [14] proposed a new method using abundance separation constraint and smoothness constraint called ASSNMF. Recently, Lu [15]proposed a manifold regularized sparse NMF model(GLNMF) for hyperspectral unmixng that using manifold assumption and spectral sparseness constraint. Based on the LMM, the abundance a_i is the low dimensional representation of pixel x_i in the new space, whose basis are endmembers. Therefore, with manifold assumption we have if two pixels are similar in original space then their abundance will also be similar. As for sparseness constraint, a novel $L_{1/2}$ norm sparse term is used to get sparser solution than L_1 norm sparse regularizer and analytical solution can be obtained.

In our paper, in order to further strengthen the relationship between original pixels and abundance and better capture

The work was supported by the National Natural Science Foundation of China under Grant 61203239.

local geometrical structure of hypercube, a novel algorithm called cluster constraint based sparse NMF(CCSNMF) is proposed for hyperspctral unmixing. CCSNMF applies the idea of distance metric learning to hyperspectral unmixing and incorporates cluster and sparsity constraints into NMF model. Clustering labels of hyperspectral pixels are used as side information. Imposing such constraint not only makes abundance of similar pixels close but makes that of dissimilar pixels well separated. To evaluate the performance of our proposed algorithm, the approach is tested on real and synthetic hyperspectral datasets.

The rest of paper is organized as follows: the section 2 describes our model and its optimization method; section 3 presents the experiment testifying our algorithm on both real and simulated hyperspectral datasets. The last section concludes the whole paper.

2. THE PROPOSED METHOD

With consideration of strengthening the constraint between original space and new low dimension space, we borrow the idea of distance metric learning [16] that if sample x and y belong to "similar" pair, their distance should be close under the new metric and are assigned to the same cluster; on the contrary, if sample x and y belong to "dissimilar" pair, they are pulled far apart and are assigned to different clusters under the new metric. In the hyperspectral unmixing scenario, that is, if two pixels are similar, then their abundance are expected to be similar; if two original pixels are dissimilar, we hope their own abundance can also be separated well. In our work, we first complement cluster to original hyperspectral dataset and then use cluster label as side information to impose similarity and dissimilarity, without having to learn positive semidefinite distance metric matrix. Additionally, a $L_{1/2}$ sparse term is incorporated to get sparse solution.

In this section, we first briefly review linear mixing model and basic NMF model for linear unmixing, then give details of proposed new model and optimization method.

2.1. Linear mixing model

Assume the size of hypercube is $M \times N \times L$, say, there are total $J = M \times N$ pixels and each of them is a L- dimensional vector. In linear mixing model, we assume each spectral vector is linearly composed of endmembers:

$$x_j = \sum_{i=1}^{P} a_i e_i, j = 1, ..., J \tag{1}$$

e_i is called endmember and a_i is the corresponding linear mixing coefficient with abundance nonnegativity constraint (ANC) $a_i \geq 0$ and abundance sum constraint(ASC) $\sum_{i=1}^{P} a_i = 1$, where P is the number of endmembers. With LMM, NMF could naturally be applied to hyperspectral unmixing following the model:

$$\begin{aligned} \min_{A,S} \quad & \frac{1}{2}\|X - AS\|_F^2 \\ s.t. \quad & A \geq 0, S \geq 0 \\ & 1^T s_i = 1^T, i = 1, 2, ..., N \end{aligned}$$

where X, A and S are original hyperspectral dataset, endmember matrix and abundance maps respectively.

2.2. Regularizer based on Distance Metric Learning and Cluster

In distance metric learning, side information like "must-link" and "cannot-link" are needed to impose constraints. In our model, we implement clustering to hyperspectral image and each pixel is assigned a label. If pixel x and y share the same label, they belong to "must-link" set and their abundance are expected to be similar; if not, they belong to "cannot-link" and difference of their abundance should be large. Based on the above analysis, we propose a new constraint term:

$$\begin{aligned} & \frac{1}{2}(\sum_{k=1}^{K} \sum_{i,j \in C_k} \|s_i - s_j\|^2 W_{ij} - \sum_{k=1}^{K} \sum_{\substack{i \in C_k \\ j \in C/C_k}} \|s_i - s_j\|^2 W_{ij}) \\ & = \sum_{i=1}^{N} s_i^2 D_{ii} - \sum_{i,j=1, i \neq j}^{N} s_i^T s_j M_{ij} \\ & = \mathrm{tr}(SLS^T) \end{aligned}$$

where s_i and s_j are the abundance of pixel x_i and x_j, W_{ij} is the similarity between x_i and x_j which is constructed as a k-nearest neighbors graph in a heat kernel method

$$W(i,j) = e^{-\frac{\|x_i - x_j\|^2}{t^2}} \tag{2}$$

Matrix M is defined as

$$M_{ij} = \begin{cases} W_{ij} & x_i \text{ and } x_j \text{ belong to the same cluster} \\ -W_{ij} & \text{otherwise} \end{cases} \tag{3}$$

$L = D - M$, where D is a diagonal matrix with element $D_{ii} = \sum_{j \neq i} M_{ij}$. C_k is the kth cluster and K the number of total clusters.

With such constraint based on cluster label information, we can acquire the constrained relationship between original pixel and their abundance. When x_i and x_j share the same cluster label, if they are similar, their distance are close, so W_{ij} is large and correspondingly the difference of s_i and s_j is likely small to minimize the objective function. If x_i and x_j are in different clusters, when they are dissimilar, W_{ij} is small and the difference of s_i and s_j is inclined to be large.

2.3. Cluster Constrained Sparse NMF

Besides abundance constraint based on spectral cluster, spectral sparseness is also an important property for hyperspectral unmixing. Therefore, $L_{1/2}$ norm regularization is adopted. Compared with well known L_1 norm term, $L_{1/2}$ sparse term has a sparser effect on solutions and analytical solution can be obtained. Mathematically, $L_{1/2}$ norm of a matrix is defined:

$$\|S\|_{\frac{1}{2}} = \sum_{i=1}^{M}\sum_{j=1}^{N} S_{ij}^{1/2} \tag{4}$$

Incorporating regularizer based on distance metric learning and $L_{1/2}$ norm regularization into original NMF, we finally have our unmixing model:

$$\begin{aligned} \min_{A,S} \quad & \frac{1}{2}\|X - AS\|_F^2 + \lambda\|S\|_{\frac{1}{2}} + \frac{1}{2}\mu\,\mathrm{tr}(SLS^T) \\ s.t. \quad & A \geq 0 \\ & S \geq 0 \\ & 1^T s_i = 1^T, i = 1, 2, ..., N \end{aligned}$$

where λ and μ are trade off parameters. To minimize the objective function, we adopt the updating rules proposed in [17]. We construct the Lagrangian function. Let ϕ_{ik} and ψ_{jk} be the lagrange multiplier for nonnegative a_{ik} and s_{jk} respectively and $\Phi = [\phi_{ik}]$, $\Psi = [\psi_{jk}]$.

$$\begin{aligned} L =& \frac{1}{2}\|X - AS\|_F^2 + \lambda\|S\|_{\frac{1}{2}} + \frac{1}{2}\mu\,\mathrm{tr}(SLS^T) + \\ & \mathrm{tr}(\Phi A^T) + \mathrm{tr}(\Psi S^T) \end{aligned}$$

The partial derivatives of L with A and S are:

$$\frac{\partial L}{\partial A} = -XS^T + ASS^T + \Phi \tag{5}$$

$$\frac{\partial L}{\partial S} = -A^T X + A^T AS + \frac{1}{2}\lambda S^{-\frac{1}{2}} + \mu SL + \Psi \tag{6}$$

According to the KKT conditions $\phi_{ik}a_{ik} = 0$ and $\psi_{jk}s_{jk} = 0$, we have

$$\begin{aligned} & (-XS^T + ASS^T).*A = 0 \\ & (-A^T X + A^T AS + \frac{1}{2}\lambda S^{-\frac{1}{2}} + \mu SL).*S = 0 \end{aligned}$$

Using $L = D - W$, the final iterate formulation is

$$A \leftarrow A.*XS^T./ASS^T \tag{7}$$

$$S \leftarrow S.*(A^T X + \mu SW)./\left(A^T AS + \frac{1}{2}\lambda S^{-\frac{1}{2}} + \mu SD\right). \tag{8}$$

The proposed hyperspectral unmixing scheme is summarized in Algorithm 1.

[1] Code can be downloaded from http://www.lx.it.pt/∼bioucas/code.htm.

Algorithm 1 Cluster Constrained Sparse NMF for Hyperspectral Unmixing

Input:
Hyperspectral Dataset X; the number of endmembers P; the number of cluster C; parameter of k-nn K; parameter of constructing similarity graph W k, t; maximum number of Iteration T, trade off parameters λ, μ;

Output:
Endmember Matrix A, Abundance Maps S

```
Implement cluster using k-nn;
Construct similarity matrix W according to Eq. (2)
Construct matrix M according to Eq. (3)
Initialize endmember matrix A and abundance maps S using VCA algorithm[1];
for i=1 to T do
    Update A by Eq. (7);
    Update S by Eq. (8);
end for
return A, S;
```

3. EXPERIMENT

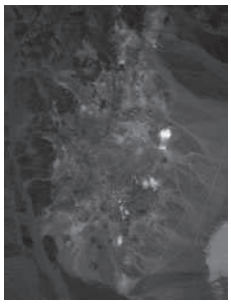
In our experiment, we compare our proposed method CCSNMF and GLNMF [15] on real hyperspectral dataset Cuprite and simulated dataset. In real data experiment, spectral angle distance(SAD) and unmixing error are employed as evaluation metric. In synthetic data experiment, SAD and root mean square error(RMSE) are compared.

3.1. Dataset Description

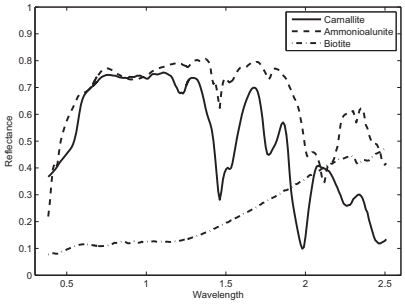
We apply our proposed unmixing algorithm on real hyperspectral dataset obtained by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) in 1997, covering Cuprite in the state of Nevada, US. There are 224 bands in the dataset and a region of size 250×190 is cut from original dataset. Before unmixing, the low quality image band like noisy bands(1-3 and 221-224) and water abortion bands(104-115 and 148-170) are removed, leaving 188 bands. Meanwhile, we use the mineral spectra from the United States Geological Survey(USGS) mineral spectral library as the spectral ground truth. As for synthetic data, three spectral signatures are chosen from USGS library and abundance is generated randomly. Then Gaussian white noise are added to construct the data with different SNR.

3.2. Parameter Selection

In real hyperspectral unmixing, how to determine the number of endmember has always been a difficult problem. But it is not the focus of our paper. Here we set it 12 according to previous research work. In the procedure of cluster, the number of cluster is also set 12 in accordance with number



(a) Sub-scene extracted from Cuprite dataset



(b) Example of endmember spectral

Fig. 1. Real and Synthetic dataset.

of endmembers. When constructing similarity matrix W, the number of nearest neighbors k and heat kernel parameter t are needed and we set 5 and 1 respectively. In synthetic data experiment, the number of endmembers is 3, the number of cluster is also set 3, other parameters are the same as those in real data experiment.

3.3. Experimental Result on Cuprite Dataset

To compare the unmixing result, SAD and unmixing error are used for measuring the accuracy the endmember matrix. The SAD is defined as

$$SAD = \cos^{-1}\left(\frac{a_i^T \cdot \hat{a}_i}{\|a_i\|\|\hat{a}_i\|}\right)$$

where a_i and $\hat{a}_i$ are the estimated ith endmember and the corresponding spectral signature extracted from the USGS library. The unmixing error is defined as

$$error = \|X - AS\|_F$$

where X, A and S are original hyperspectral dataset, endmember matrix and abundance maps respectively.

Table 1. SAD Results On The AVIRIS Cuprite Dataset

	GLNMF	CCSNMF
SAD	0.5316	**0.4684**
error	67.5196	**66.1228**

3.4. Experimental Result on Synthetic Dataset

We conduct experiments on 3 simulated datasets with different SNR to compare our method CCSNMF with GLNMF. Two criteria are used to evaluate the unmixing results. SAD has been defined in the last subsection. The RMSE is defined as

$$RMSE = \sqrt{\frac{1}{P}\sum_{j=1}^{P}(S_j - \hat{S}_j)^2}$$

Table 2. SAD under different SNR on synthetic data

SNR	15dB	20dB	25dB
GLNMF	**0.0807**	0.0676	0.0471
CCSNMF	0.1302	**0.0375**	**0.0196**

Table 3. RMSE under different SNR on synthetic data

SNR	15dB	20dB	25dB
GLNMF	**0.3686**	0.1062	0.0773
CCSNMF	0.3903	**0.0678**	**0.0452**

where S_j and $\hat{S}_j$ are jth abundance and the corresponding reference abundance. P is the number of endmembers.

As shown in Table 1, in real dataset experiment, CCSNMF gets smaller SAD 0.4684 and GLNMF obtains 0.5316. Smaller SAD means better fit of spectral signatures. For unmixing error, CCSNMF yields smaller numerical value, 66.1228, than that of GLNMF, 67.5196, which means better data approximation. In synthetic data experiment, table 2 and 3 show that CCSNMF gets lower metrics both in terms of SAD and RMSE than GLNMF under high SNR.

4. CONCLUSION

In this paper, we have presented a new hyperspectral unmixing algorithm based on spectral cluster information. It incorporates spectral cluster constraint into NMF model and makes a connection between pixels and their abundances, such that abundances of similar pixels are similar and that of dissimilar pixels are more different. Experimental results on real dataset and synthetic data show that our proposed algorithm will improve the accuracy of hyperspectral unmixing.

5. REFERENCES

[1] J Bioucas-Dias, Antonio Plaza, Gustavo Camps-Valls, Paul Scheunders, N Nasrabadi, and JOCELYN Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *Geoscience and Remote Sensing Magazine, IEEE*, vol. 1, no. 2, pp. 6–36, 2013.

[2] Joseph W Boardman et al., "Automating spectral unmixing of aviris data using convex geometry concepts," in *Summaries 4th Annu. JPL Airborne Geoscience Workshop*. JPL Publication 93–26, 1993, vol. 1, pp. 11–14.

[3] Michael E Winter, "N-findr: an algorithm for fast autonomous spectral end-member determination in hyperspectral data," in *SPIE's International Symposium on Optical Science, Engineering, and Instrumentation*. International Society for Optics and Photonics, 1999, pp. 266–275.

[4] José MP Nascimento and José MB Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 43, no. 4, pp. 898–910, 2005.

[5] José M Bioucas-Dias, Antonio Plaza, Nicolas Dobigeon, Mario Parente, Qian Du, Paul Gader, and Jocelyn Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, vol. 5, no. 2, pp. 354–379, 2012.

[6] Saïd Moussaoui, Cédric Carteret, David Brie, and Ali Mohammad-Djafari, "Bayesian analysis of spectral mixture data using markov chain monte carlo methods," *Chemometrics and Intelligent Laboratory Systems*, vol. 81, no. 2, pp. 137–148, 2006.

[7] Nicolas Dobigeon, Saïd Moussaoui, Martial Coulon, J-Y Tourneret, and Alfred O Hero, "Joint bayesian endmember extraction and linear unmixing for hyperspectral imagery," *Signal Processing, IEEE Transactions on*, vol. 57, no. 11, pp. 4355–4368, 2009.

[8] Morten Arngren, Mikkel N Schmidt, and Jan Larsen, "Unmixing of hyperspectral images using bayesian nonnegative matrix factorization with volume prior," *Journal of Signal Processing Systems*, vol. 65, no. 3, pp. 479–496, 2011.

[9] Jessica D Bayliss, J Anthony Gualtieri, and Robert F Cromp, "Analyzing hyperspectral data with independent component analysis," in *26th AIPR Workshop: Exploiting New Image Sources and Sensors*. International Society for Optics and Photonics, 1998, pp. 133–143.

[10] Te-Ming Tu, "Unsupervised signature extraction and separation in hyperspectral images: A noise-adjusted fast independent component analysis approach," *Optical Engineering*, vol. 39, no. 4, pp. 897–906, 2000.

[11] José MP Nascimento and José MB Dias, "Does independent component analysis play a role in unmixing hyperspectral data?," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 43, no. 1, pp. 175–187, 2005.

[12] Lidan Miao and Hairong Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 45, no. 3, pp. 765–777, 2007.

[13] Sen Jia and Yuntao Qian, "Constrained nonnegative matrix factorization for hyperspectral unmixing," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 47, no. 1, pp. 161–173, 2009.

[14] Xuesong Liu, Wei Xia, Bin Wang, and Liming Zhang, "An approach based on constrained nonnegative matrix factorization to unmix hyperspectral data," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 49, no. 2, pp. 757–772, 2011.

[15] Xiaoqiang Lu, Hao Wu, Yuan Yuan, Pingkun Yan, and Xuelong Li, "Manifold regularized sparse nmf for hyperspectral unmixing," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 51, no. 5, pp. 2815–2826, 2013.

[16] Eric P Xing, Andrew Y Ng, Michael I Jordan, and Stuart Russell, "Distance metric learning with application to clustering with side-information," *Advances in neural information processing systems*, pp. 521–528, 2003.

[17] Deng Cai, Xiaofei He, Jiawei Han, and Thomas S Huang, "Graph regularized nonnegative matrix factorization for data representation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 8, pp. 1548–1560, 2011.