

# MULTIPLE GRAPH REGULARIZED NMF FOR HYPERSPECTRAL UNMIXING

Lei Tong<sup>1</sup>, Jun Zhou<sup>2</sup>, Yuntao Qian<sup>3</sup> and Yongsheng Gao<sup>1</sup>

<sup>1</sup>School of Engineering, Griffith University, Nathan, Australia.

<sup>2</sup>School of Information and Communication Technology, Griffith University, Nathan, Australia

<sup>3</sup> College of Computer Science, Zhejiang University, Hangzhou, P.R. China.

## ABSTRACT

Hyperspectral unmixing is an important technique for estimating fraction of different land covers from remote sensing imagery. In recent years, nonnegative matrix factorization (NMF) methods with various constraints have been introduced into hyperspectral unmixing. Among these methods, graph based constraint has been proved to be useful in capturing the latent manifold structure of the hyperspectral data in the feature domain. However, due to the complexity of the data, only using single graph can not adequately reflect the intrinsic property of the data. In this paper, we propose a multiple graph regularized NMF method for hyperspectral unmixing, which approximates the manifold and consistency of data by a linear combination of several graphs constructed in different scales. Results on both synthetic and real data have validated the effectiveness of the proposed method, and shown that it has outperformed several state-of-the-arts hyperspectral unmixing methods.

**Index Terms**— Hyperspectral Unmixing, Nonnegative Matrix Factorization, Multiple Graph.

## 1. INTRODUCTION

Hyperspectral remote sensing imagery, which contains both spatial and spectral information captured by imaging sensors, has been widely used for land cover detection [1]. Because of low spatial resolution of image, pixels in hyperspectral imagery are composed of mixed spectral responses from multiple ground targets. Therefore, hyperspectral unmixing has become an important technique which decomposes mixed pixels to a collection of endmembers and their corresponding proportions, i.e., abundances [2].

Amongst various mixture models, linear mixture model treats each pixel as a linear combination of endmembers and assumes no interference between endmembers [3, 4]. Methods based on linear mixture model can be divided into three categories: simplex geometric based, statistics based, and sparse representation methods. There are several classic simplex geometric methods, such as N-FINDR [5], pixel purity index [6], and vertex component analysis (VCA) [7]. A typical statistics based method is independent components analysis [8]. Sparse representation based approaches have also been proposed [9, 10, 11]. They select endmembers from an existing library of standard spectral reflectance and estimate the abundance at the same time.

More recently, nonnegative matrix factorization (NMF) [12] has been proposed as a statistics based unmixing method. It treats the unmixing as a blind source separation problem, and decomposes image data into endmember matrix and abundance matrix simultaneously.

However, due to non-convexity of its object function, the outcome of NMF algorithm usually falls into local minima. One solution is to add constraints to NMF according to different considerations. Therefore, several constraints have been combined to NMF methods, such as minimum volume constraint [13], dissimilarity constraint [14], sparse constraint [15] and graph constraint [16].

The graph constraint method proposed by [16] is built up on the top of  $L_{1/2}$ -NMF method and based on the manifold theory. It seeks the nearest neighbour of pixels in the image data to build a graph that is used to guide the unmixing process. This means that data with close distance in the high dimension hyperspectral image shall have similar unmixing outcomes. However, this method imposes only one manifold constraint to the unmixing model. Selecting an appropriate graph and its parameter is quite important for its performance. Discrete grid search is usually used to select the graph model and its parameters, which is time consuming. A sub-optimal graph model may not well reflect the intrinsic property of data in the feature space. In this paper, we solve this problem by using a linear combination of graphs and propose a multiple graph regularized NMF method for hyperspectral unmixing. Each graph is built from different scales in terms of different nearest neighbours. The weights of the graph are optimized to control the contribution of different manifolds. Therefore, our method is a generalised version of GNMF which can not only use different graphs but also estimate the parameters automatically in the optimization process.

## 2. RELATED WORK

### 2.1. Linear Mixture Model

In hyperspectral image, each pixel is considered as a pure endmember or the mixture of several endmembers. Let the number of wavelength-indexed bands in an image be  $H$  and the number of endmembers be  $K$ . A pixel  $y$  in a hyperspectral image is an  $H \times 1$  column vector whose entries correspond to the measured data in different bands. Let  $M$  be an  $H \times K$  matrix  $(m_1, \dots, m_j, \dots, m_K)$ , where  $m_j$  is an  $H \times 1$  column vector representing the spectral signature of the  $j^{\text{th}}$  endmember. Then  $y$  can be approximated by a linear combination of endmembers

$$y = Mr + e \quad (1)$$

where  $r$  is a  $K \times 1$  column vector for endmember abundances, and  $e$  is the additive Gaussian white noise. It is natural to extend the above pixel-level mixing model to the whole image. Let  $N$  be the number of pixels contained in a hyperspectral image  $Y$ , the linear

model becomes

$$Y = MR + E \quad (2)$$

where  $Y \in \mathbb{R}_+^{H \times N}$ ,  $R \in \mathbb{R}_+^{K \times N}$ , and  $E \in \mathbb{R}^{H \times N}$  represent the hyperspectral image, the abundance matrix, and the additive noise, respectively. The goal of unmixing is to estimate  $M$  and  $R$  from  $Y$ .

## 2.2. Graph Regularized NMF

The objective function of the graph regularized sparse NMF is defined as

$$obj(M, R) = \frac{1}{2} \|Y - MR\|_F^2 + \lambda \|R\|_{1/2} + \frac{\mu}{2} Tr(RLR^T) \quad (3)$$

where  $\mu \in \mathbb{R}_+$  is a scalar which weights the contribution of manifold constraint.  $Tr(\cdot)$  represents the trace of a matrix.  $\lambda$  is a parameter to adjust the sparseness.  $L$  is a Laplacian matrix used as the manifold constraint. It can be computed as the difference of the degree matrix  $D$  and the adjacency matrix  $W$  of the graph

$$L = D - W \quad (4)$$

Given two adjacent pixels  $y_i$  and  $y_j$ , the entry of the adjacency matrix  $W$  is calculated using a heat kernel

$$W_{ij} = \exp - \frac{\|y_i - y_j\|^2}{\sigma} \quad (5)$$

The diagonal matrix  $D$  is then computed as follows

$$D_{ii} = \sum_{i=1} W_{ij} \quad (6)$$

## 3. APPROACH

In this section, we firstly introduce the multiple graph regularized NMF method in detail. Then some implementation issues are discussed.

### 3.1. Graph Construction

We could generate different graphs using  $k$  nearest neighbours with different graph models and parameters. The most commonly used models are heat kernel and 0 – 1 weighting. The heat kernel is defined as follows

$$W_{ij} = \begin{cases} \exp - \frac{\|y_i - y_j\|^2}{\sigma}, & y_j \in N(y_i) \text{ or } y_i \in N(y_j) \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

If  $y_i$  and  $y_j$  are similar, the value of  $W_{ij}$  is relatively large.

Alternatively, 0 – 1 weighting gives a binarized weight definition

$$W_{ij} = \begin{cases} 1, & y_j \in N(y_i) \text{ or } y_i \in N(y_j) \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

In practices, different  $k$  nearest neighbours of these two models are generated to represent different scales of manifold constraint.

### 3.2. Multiple Graph Regularized NMF

A number of graphs, which is converted to the representation of Laplacian matrices, are used in our approach to regularize the NMF method. We adopt a linear combination method to optimize the contribution for each matrix [17].

Suppose that we have a number of Laplacian matrices  $L_1, \dots, L_S$ , whose combination is defined as follows

$$L = \sum_{s=1}^S \alpha_s L_s \quad \text{s.t.} \quad \sum_{s=1}^S \alpha_s = 1, \alpha_s \geq 0 \quad (9)$$

where  $\alpha_s$  is the weight of  $s^{th}$  graph. To avoid negative contribution, we constrain  $\sum_{s=1}^S \alpha_s = 1, \alpha_s \geq 0$ .

With multiple graphs generated before, the objective function of the multiple graph regularized sparse NMF can be defined as

$$obj(M, R) = \frac{1}{2} \|Y - MR\|_F^2 + \lambda \|R\|_{1/2} + \frac{\mu}{2} \sum_{s=1}^S \alpha_s Tr(RL_s R^T) + \beta \|\alpha\|^2 \quad (10)$$

As can be seen, the first two terms are the same as the formulation in the  $L_{1/2}$ -NMF method, which force sparsity of the abundance matrix. The third term is the graph regularization term. The last term is to avoid parameter  $\alpha$  overfitting to one graph, while  $\alpha$  is a parameter to balance the contribution of different graphs.  $\mu \in \mathbb{R}_+$  and  $\beta$  are also used here to control the contribution of terms.

In order to get the optimal  $M$  and  $R$ , we take a two-step strategy to iteratively optimize Equ. (10) because of its complexity. In each iteration, only  $(M, R)$  or  $\alpha$  is updated while the other is fixed.

#### 3.2.1. Minimizing $(M, R)$

When parameter  $\alpha_s$ ,  $s = 1, \dots, S$  is fixed, the objective function becomes

$$obj(M, R) = \frac{1}{2} \|Y - MR\|_F^2 + \lambda \|R\|_{1/2} + \frac{\mu}{2} \alpha Tr(RLR^T) \quad (11)$$

After computing the partial derivative of this objective function with respect to  $M$  and  $R$ , and doing some computations, the iterative updating rule of the graph regularized NMF becomes

$$M \leftarrow M * YR^T ./ MRR^T \quad (12)$$

$$R \leftarrow R * (M^T Y + \mu \alpha R W) ./ (M^T M R + \frac{\lambda}{2} R^{-\frac{1}{2}} + \mu \alpha R D) \quad (13)$$

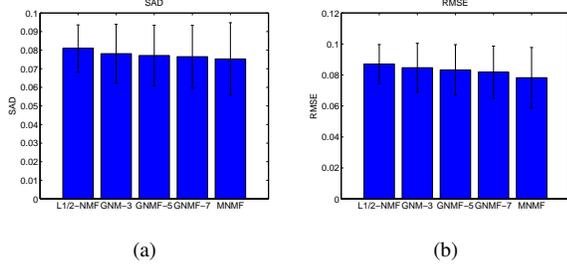
Note here that  $M$  and  $R$  are updated iteratively. This is similar to the GNMF method.

#### 3.2.2. Minimizing $\alpha$

Now, we fix  $(M, R)$  and remove other terms. The objective function becomes

$$\sum_{r=1}^R \alpha_r Tr(WL_r W^T) + \beta \|\alpha\|^2 \quad \text{s.t.} \quad \sum_{r=1}^R \alpha_r = 1, \alpha_r \geq 0 \quad (14)$$

This is a constrained convex optimization problem. We solve this problem by the convex optimization approach in [18].



**Fig. 1.** Performance of different algorithms. (a) Mean SAD; (b) Mean RMSE.

### 3.3. Implementation Issues

#### 3.3.1. Parameters

As mentioned in Section 2, the abundance matrix should meet the sum-to-one constraint. In the implementation, the method in [19] is adopted to apply this constraint. The data matrix  $Y$  and the signature matrix  $M$  are replaced by  $Y_f$  and  $M_f$  which are augmented matrices by a row of constants defined as follows

$$Y_f = \begin{pmatrix} Y \\ \delta \mathbf{1}_N^T \end{pmatrix} M_f = \begin{pmatrix} M \\ \delta \mathbf{1}_N^T \end{pmatrix} \quad (15)$$

where  $\delta$  is a parameter to balance the impact of the additivity constraint of the abundance matrix. A large  $\delta$  forces the sum of abundances at each pixel close to unity. In the iteration, these augmented matrices replace  $Y$  and  $M$  in the update rule of  $R$  given in Equation 13. The value of the parameter  $\lambda$  is dependent on the sparsity of the material abundances. Here we adopt [20] method to compute  $\lambda$

The procedure of our multiple graph regularized NMF method is described below

---

#### Algorithm 1: Unmixing procedure

---

**Data:** Hyperspectral image  $Y$

1. Initialize endmember matrix  $M$  and abundance matrix  $R$ .
2. Construct multiple graphs.
3. Augment  $Y$  and  $M$  to  $Y_f$  and  $M_f$ .

**while not converge do**

1. Update  $R$  by equation (13).
2. Update  $M$  by equation (12).
3. Update  $\alpha$  by equation (14).

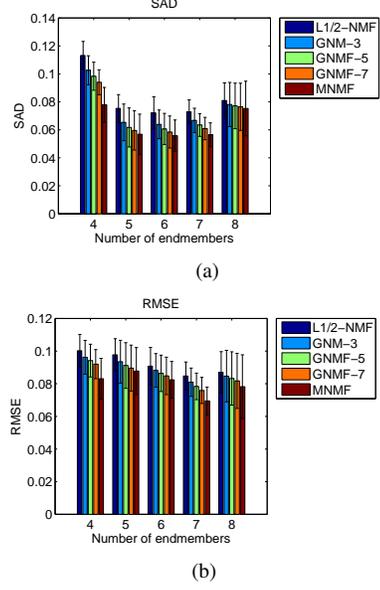
**end**

Output  $M$  and  $R$

---

## 4. EXPERIMENTS

A series of experiments are developed to evaluate the performance of the proposed multiple graph regularized NMF method. Two evaluation criteria are used: spectral angle distance (SAD) and the root mean squared error (RMSE). The SAD is used to compare the similarity of the  $p^{th}$  endmember signature  $M_p$  and its estimate  $\widehat{M}_p$ . The root mean square error (RMSE) is used to evaluate the the estimation of abundance matrix. The definition of both criteria can be found in [15].



**Fig. 2.** The result of increasing number of endmembers. (a) Mean SAD; (b) Mean RMSE.

### 4.1. Synthetic Data

We first evaluate and analyze the proposed method on synthetic data. Six spectral signatures are chosen from the USGS digital spectral library [21] to generate the synthetic data. Similar to [13], there are 5 steps to generate this synthetic data. 1. Generate an image which contains  $64 \times 64$  pixels and divided into  $8 \times 8$  regions. 2. In each region the same type of ground cover was initialized with one of the endmembers chosen randomly. 3. A low pass filter was applied to generate mixed data. 4. For the pixel whose abundance is larger than  $\theta = 0.8$ , the abundance of this pixel is replaced with a mixture of all endmembers with equal abundances. 5. Zero-mean Gaussian noise is added to synthetic data.

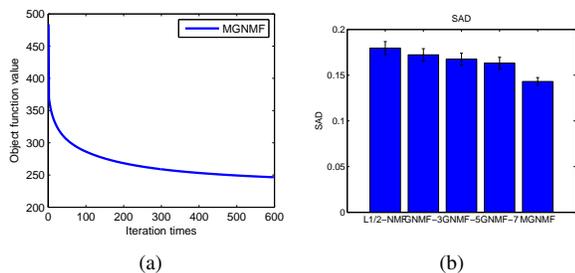
The signal to noise ratio (SNR) is defined as

$$SNR = 10 \log_{10} \frac{E[y^T y]}{E[e^T e]} \quad (16)$$

where the  $y$  and  $e$  are the observation and noise of a pixel.  $E[\cdot]$  denotes the expectation operator.

We plot the convergence curve in Fig. 3(a) to show the convergence process of the proposed algorithm. We can see that at beginning, the object function value of our method is quite large. When iteration processes, the function value of MGNMF method becomes smaller. In the end, we can see that the convergence curve of MGNMF method become stable.

We also evaluate the performance of different unmixing methods:  $L_{1/2}$ -NMF [15], GNMF [16] and our MGNMF method. In the experiment, we choose  $3NN$ ,  $5NN$  and  $7NN$  as the nearest neighbours of GNMF and use heat kernel to generate graphs. Once the nearest neighbours are selected, the way each graph is generated is the same as GNMF. In this experiment, the parameters setting in  $L_{1/2}$ -NMF, GNMF and our method are the same. We set  $SNR = 20$  and total number of endmembers  $P = 6$  in this experiment. We run each method for 10 times and then calculate the mean value and standard deviation of SAD and RMSE. Fig. 1 shows



**Fig. 3.** (a)Convergence curve of the MGNMF. (b)Results of SAD on endmember estimation on the Urban data.



**Fig. 4.** Results of abundance estimation on the Urban HYDICE Dataset.

the experimental results. The bar and error line stand for the mean SAD and RMSE and their standard deviations, respectively. It can be seen that our method has clearly outperformed all other methods in endmember estimation, and is better than  $L_{1/2}$ -NMF and GNMFM methods in abundance estimation. We also conduct the experiment with different endmembers. Fig. 2 shows that MGNMF method is also better than others. The advantage of our method over GNMFM lies in that it uses different scales of graphs. Therefore, MGNMF method can better use manifold information and get better performance.

#### 4.2. Remote Sensing Data

We also conducted experiment on the widely used urban image obtained by the Urban Hyperspectral Digital Imagery Collection Experiment (HYDICE) sensor [15]. This image is composed of 210 spectral channels with spectral resolution of 10nm acquired in the 400nm and 2500nm region. After low SNR bands had been removed (channels 1–4, 76, 87, 101–111, 136–153, and 198–210), 162 bands remained for the experiments, i.e.,  $H = 162$ .

In the experiment, we considered 4 types of endmembers: Asphalt, Tree, Grass and Roof. Fig. 3 shows the mean SAD of endmember estimation with different methods. The results show that MNMF can help to get better endmember estimation. Fig 4 shows the abundance maps of endmembers.

### 5. CONCLUSION

In this paper, we have introduced a multiple graph regularized NMF for hyperspectral unmixing. Built on top of the graph regularized sparse NMF approach, this method provides new constraints on the NMF model by the construction of multiple graphs. A linear combination of multiple graph is proposed to adopt information from different graphs. An iterative optimization process is developed to generate a unified solution for both endmember and abundance estimation. This method is validated on both synthetic and real data,

and have achieved better performance when compared against the  $L_{1/2}$ -NMF and graph regularized NMF method. In the future, we will explore other effective way to combine multiple feature of hyperspectral images to get better unmixing results.

### 6. REFERENCE

- [1] Alexander FH Goetz, Gregg Vane, Jerry E Solomon, and Barrett N Rock, "Imaging spectrometry for earth remote sensing," *Science*, vol. 228, no. 4704, pp. 1147–1153, 1985.
- [2] Nirmal Keshava and John F Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 44–57, 2002.
- [3] 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," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 354–379, 2012.
- [4] Nirmal Keshava, "A survey of spectral unmixing algorithms," *Lincoln Laboratory Journal*, vol. 14, no. 1, pp. 55–78, 2003.
- [5] Michael E Winter, "N-FINDR: an algorithm for fast autonomous spectral endmember determination in hyperspectral data," in *SPIE International Symposium on Optical Science, Engineering, and Instrumentation*. International Society for Optics and Photonics, 1999, pp. 266–275.
- [6] Joseph W Boardman, "Automating spectral unmixing of aviris data using convex geometry concepts," in *JPL Airborne Geoscience Workshop Summaries 4th Annu. JPL Publication 93–26*, 1993, vol. 1, pp. 11–14.
- [7] José MP Nascimento and José MB Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 4, pp. 898–910, 2005.
- [8] Aapo Hyvärinen and Erkki Oja, "Independent component analysis: algorithms and applications," *Neural networks*, vol. 13, no. 4, pp. 411–430, 2000.
- [9] M.-D. Iordache, J.M. Bioucas-Dias, and A Plaza, "Sparse unmixing of hyperspectral data," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 49, no. 6, pp. 2014–2039, June 2011.
- [10] M. Iordache, J. Dias, and A. Plaza, "Collaborative sparse regression for hyperspectral unmixing," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 52, no. 1, pp. 341–354, January 2013.
- [11] M-D Iordache, José M Bioucas-Dias, and Antonio Plaza, "Total variation spatial regularization for sparse hyperspectral unmixing," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 50, no. 11, pp. 4484–4502, 2012.
- [12] V Paul Pauca, Jon Piper, and Robert J Plemmons, "Nonnegative matrix factorization for spectral data analysis," *Linear algebra and its applications*, vol. 416, no. 1, pp. 29–47, 2006.
- [13] Lidian Miao and Hairong Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 3, pp. 765–777, 2007.
- [14] Nan Wang, Bo Du, and Liangpei Zhang, "An endmember dissimilarity constrained non-negative matrix factorization method for hyperspectral unmixing," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 2, pp. 554–569, April 2013.
- [15] Yuntao Qian, Sen Jia, Jun Zhou, and A. Robles-Kelly, "Hyperspectral unmixing via  $L_{1/2}$  sparsity-constrained nonnegative matrix factorization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4282–4297, 2011.
- [16] Xiaoqiang Lu, Hao Wu, Yuan Yuan, Pingkun Yan, and Xuelong Li, "Manifold regularized sparse NMF for hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 5, pp. 2815–2826, 2013.
- [17] Jim Jing-Yan Wang, Halima Bensmail, and Xin Gao, "Multiple graph regularized nonnegative matrix factorization," *Pattern Recognition*, vol. 46, no. 10, pp. 2840–2847, 2013.
- [18] Michael Grant and Stephen Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*, V. Blondel, S. Boyd, and H. Kimura, Eds., Lecture Notes in Control and Information Sciences, pp. 95–110. Springer-Verlag Limited, 2008.
- [19] Daniel C Heinz et al., "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 3, pp. 529–545, 2001.
- [20] Patrik O Hoyer, "Non-negative matrix factorization with sparseness constraints," *The Journal of Machine Learning Research*, vol. 5, pp. 1457–1469, 2004.
- [21] R. Clark, G. Swayze, R. Wise, E. Livo, T. Hoefen, R. Kokaly, and S. Sutley, "USGS digital spectral library splib06a: U.S. Geological Survey, Digital Data Series 231," USGS, 2007.