

# SUPERPIXEL-BASED NONNEGATIVE TENSOR FACTORIZATION FOR HYPERSPECTRAL UNMIXING

Fengchao Xiong<sup>1</sup> Jingzhou Chen<sup>1</sup> Jun Zhou<sup>2</sup> Yuntao Qian<sup>1</sup>

<sup>1</sup> College of Computer Science, Zhejiang University, Hangzhou, China

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

## ABSTRACT

Hyperspectral unmixing aims at decomposing a hyperspectral image (HSI) into a number of constituted materials and associated proportions. Recently, nonnegative tensor factorization (NTF) based methods have been proved effective and natural for hyperspectral unmixing owing to their virtue of representing an HSI without any information loss. However, these methods take an HSI as a whole, partly ignoring the local information in distinct local regions. In addition, HSIs are high likely to be disturbed by various noise, making the global information unnecessarily reliable. To alleviate these drawbacks, we propose a superpixel-based matrix-vector nonnegative tensor factorization (S-MV-NTF) method for hyperspectral unmixing, where both the global information and local information are taken into consideration. In this method, the HSI is firstly partitioned into numerous superpixels, homogeneous regions with adaptive sizes and compact boundaries, representing the local spatial structure information. Then, such local information is integrated to the tensor factorization to make the pixels lying in the same superpixel share similar abundances. Experimental results on synthetic data and real-world data show that the proposed method dominates the state-of-the-art methods.

**Index Terms**— Hyperspectral unmixing, joint spectral-spatial information, superpixel, nonnegative tensor factorization,

## 1. INTRODUCTION

Hyperspectral imagery (HSI), containing hundreds of continuous narrow spectral bands, has been widely applied in various fields [1]. Many pixels cover more than one kind of materials because of sensor's limited spatial resolution. The existence of these pixels has caused a range of difficulties in real-world applications. Hyperspectral unmixing aims to detect the existence of a collection of constitute materials, i.e., endmembers, and estimate their corresponding fractions, i.e., abundances. It offers an attractive way to tackle this problem.

This work was supported by the National Natural Science Foundation of China 61571393

From the statistic analysis point of view, this problem can be considered as blind source separation problem. Nonnegative matrix factorization (NMF) [2] is effective to solve this problem. Nevertheless, there exist many local optimal solutions because of its non-convex objective function. In order to make the model more powerful, various constraints incorporating auxiliary prior information are added to NMF. Considering hyperspectral images vary smoothly in its spatial domain, Lu *et al.* [3] proposed a manifold regularized sparse NMF (MRS-NMF) to make the pixels in a low-dimensional submanifold behave similarly. In [4], hypergraph structure was employed to model more complicated similarity relationship among the spatial nearby pixels. Tong *et al.* [5] proposed a region-based NMF (R-NMF) method to keep the structure consistency within the regions while discriminating the differences between regions.

Compared with converting a 3D HSI into a 2D matrix form, a third-order tensor is more natural in representing HSI. Recently, Qian *et al.* [6] proposed a matrix-vector nonnegative tensor factorization (MV-NTF) for unmixing. This model factorizes an HSI into  $R$  rank  $(L, L, 1)$  component tensors where each one is the outer product of a low rank matrix and a vector, representing abundance and endmember respectively. Experimental results demonstrate that this method outperforms many NMF based unmixing approaches. However, MV-NTF takes an HSI as a whole so that the detailed spatial structure information can not be fully described. In addition, in real-world applications, the tensor data suffers from some changing cases, such as low-SNR, bare identifiability and ill-condition [7], which makes tensor decomposition fail to be unique.

Embedding local spatial information to tensor factorization is an effective manner to alleviate the drawbacks of NTF. For example, in [8], a 3D total variation and  $L_1$  sparsity regulation were incorporated into low rank Tucker decomposition to remove the sparse noise and Gaussian noise simultaneously while preserving the spatial structure and spectral signatures of HSI. In addition, superpixel, which groups similar local pixels into meaningful homogeneous regions, provides an effective manner to obtain more elaborate spatial information. Owing to its superior advantages in representing spatial structure information while matching the boundary of distinct se-

mantic region, superpixel has been widely applied in hyperspectral denoising [9] and classification [10].

In this paper, in order to alleviate the drawbacks of MV-NTF, a superpixel-based matrix-vector nonnegative tensor factorization (S-MV-NTF) method is proposed for hyperspectral unmixing. As shown in Fig. 1, this approach first uses simple linear iterative clustering (SLIC) [11] to segment the HSI into several superpixels with adaptive sizes and shapes. Then, two similarity graphs are constructed to model the similarity in superpixels. Finally, these two similarity graphs are embedded into the original MV-NTF to make the pixels in the same superpixel share similar abundances. Experimental results on synthetic data and real-world data show that the proposed method outperforms alternative methods.

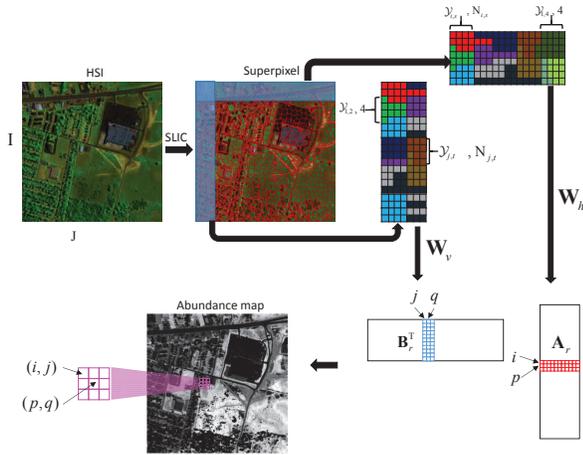


Fig. 1: The framework of S-MV-NTF.

## 2. RELATED WORK

As mentioned in [6], MV-NTF is consistent with linear spectral mixture model (LMM). Mathematically, its objective function can be formulate as:

$$\min_{\mathbf{A}, \mathbf{B}, \mathbf{C}} \frac{1}{2} \|\mathcal{Y} - \sum_{r=1}^R (\mathbf{A}_r \mathbf{B}_r^T) \circ \mathbf{c}_r\|_F^2 + \frac{\delta}{2} \|\mathbf{1}_{\mathbf{I} \times \mathbf{J}} - \mathbf{A} \mathbf{B}^T\|_F^2 \quad (1)$$

where  $\mathcal{Y} \in \mathbb{R}^{I \times J \times K}$  represents the observed 3D HSI with  $I \times J$  pixels and  $K$  bands,  $\mathbf{A}_r \mathbf{B}_r^T$  is the abundance map of  $r$ th endmember, approximately represented by two low-rank matrices  $\mathbf{A}_r$  and  $\mathbf{B}_r$ ,  $\mathbf{c}_r$  is the  $r$ th endmember,  $\mathbf{1}_{\mathbf{I} \times \mathbf{J}}$  is a matrix filled with all ones and  $\delta \in \mathbb{R}^+$  balances the tradeoff between the reconstruction error and sum-to-one constraint.

Though this approach performs unmixing under tensor notation, it takes an HSI as a whole, ignoring the detailed spatial structure information of distinct regions. To facilitate better performance, auxiliary local spatial information

exploited from the HSI should be integrated to tensor factorization. This will be presented in the next section.

## 3. SUPERPIXEL BASED MATRIX-VECTOR NTF FOR HYPERSPECTRAL UNMIXING

In this section, we give the details of the proposed method. It integrates the local information represented by superpixels to the original MV-NTF method to make the pixels in the same superpixel share similar abundances. The corresponding similarity graph construction and update rules are also discussed.

### 3.1. Superpixel guided similarity graph construction

In this paper, we adopt SLIC [11] to generate superpixels because of its virtue of good accuracy, high computational speed and boundary recall properties. Two similarity graphs  $\mathbf{W}_h$  and  $\mathbf{W}_v$ , representing horizontal and vertical spatial relationship between pixels, are built based on these superpixels to make the pixels in the same superpixel behave similarly. Taking the horizontal orientation for example, mean feature vector  $\mathcal{Y}_{i,s}$  of the pixels belonging to the  $s$ -th superpixel in the  $i$ -th row is given by:

$$\mathcal{Y}_{i,s} = \frac{1}{N_{i,s}} \sum_{j=1, l(i,j)=s}^J \mathcal{Y}_{i,j} \quad (2)$$

where  $N_{i,s}$  denotes the number of such pixels. An edge  $e_{ip}$  is built between  $i$ -th row and  $j$ -th row, whose weight is :

$$w(e_{ip}) = \frac{1}{N} \sum_{s \in S} \exp - \frac{\|\mathcal{Y}_{i,s} - \mathcal{Y}_{p,s}\|^2}{\sigma} \quad (3)$$

where  $S$  denotes the superpixels that exist in both  $i$ -th row and  $p$ -th row and  $N$  represents associated total number. In the same way, the vertical similarity graph  $\mathbf{W}_v$  can also be constructed for  $\mathbf{B} = [\mathbf{B}_1 \cdots \mathbf{B}_R]$  to represent the vertical spatial structure information.

### 3.2. Proposed method

With the obtained weight matrix  $\mathbf{W}_h$ , we can enforce the similarity between rows in  $\mathbf{A}_r$ , which is formulated by:

$$\frac{1}{2} \sum_{i=1}^I \sum_{p=1}^I \|\mathbf{A}_r^i - \mathbf{A}_r^p\|^2 \mathbf{W}_h^{ip} = \text{Tr}(\mathbf{L}_h \mathbf{A}_r \mathbf{A}_r^T) \quad (4)$$

where  $\text{Tr}$  denotes the trace of a matrix,  $\mathbf{D}_h$  is a diagonal matrix where  $\mathbf{D}_h^{ii} = \sum_j \mathbf{W}_h^{ij}$  and  $\mathbf{L}_h$  is a laplacian matrix of  $\mathbf{W}_h$  which is represented as  $\mathbf{L}_h = \mathbf{D}_h - \mathbf{W}_h$ . Extending this constraint to  $\mathbf{A} = [\mathbf{A}_1 \cdots \mathbf{A}_R]$ , the regularization term is

$$\frac{1}{2} \sum_{r=1}^R \text{Tr}(\mathbf{L}_h \mathbf{A}_r \mathbf{A}_r^T) = \text{Tr}(\mathbf{L}_h \mathbf{A} \mathbf{A}^T) \quad (5)$$

In the same way, adding the similarity constraint to the rows in  $\mathbf{B}_r$ , we can get:

$$\frac{1}{2} \sum_{r=1}^R \text{Tr}(\mathbf{L}_v \mathbf{B}_r \mathbf{B}_r^T) = \text{Tr}(\mathbf{L}_v \mathbf{B} \mathbf{B}^T) \quad (6)$$

Incorporating these two regulation terms into the MV-NTF model, the goal now becomes minimizing the following cost function:

$$\begin{aligned} \min_{\mathbf{A}, \mathbf{B}, \mathbf{C}} \frac{1}{2} \|\mathcal{Y} - \sum_{i=1}^R \mathbf{A}_r \mathbf{B}_r^T \circ c_r\|_F^2 + \frac{\delta}{2} \|\mathbf{1}_{\mathbf{I} \times \mathbf{J}} - \mathbf{A} \mathbf{B}^T\|_F^2 \\ + \frac{\alpha}{2} \text{Tr}(\mathbf{L}_h \mathbf{A} \mathbf{A}^T) + \frac{\mu}{2} \text{Tr}(\mathbf{L}_v \mathbf{B} \mathbf{B}^T) \end{aligned} \quad (7)$$

where  $\alpha \in \mathbb{R}^+$  and  $\mu \in \mathbb{R}^+$  are two regularization parameters to control two constraints on  $\mathbf{A}$  and  $\mathbf{B}$  respectively. Under the framework of multiplicative update, the associated update rules are given by:

$$\mathbf{A} \leftarrow \mathbf{A} \cdot \left( \mathbf{Y}_{(1)}^T \mathbf{M} + \delta \mathbf{1}_{\mathbf{I} \times \mathbf{J}} \mathbf{B} + \alpha \mathbf{W}_h \mathbf{A} \right) / \left( \mathbf{A} \mathbf{M}^T \mathbf{M} + \delta \mathbf{A} \mathbf{B}^T \mathbf{B} + \alpha \mathbf{D}_h \mathbf{A} \right) \quad (8)$$

where  $\mathbf{M} = \mathbf{B} \odot \mathbf{C}$ .

$$\mathbf{B} \leftarrow \mathbf{B} \cdot \left( \mathbf{Y}_{(2)}^T \mathbf{M} + \delta \mathbf{1}_{\mathbf{I} \times \mathbf{J}} \mathbf{A} + \mu \mathbf{W}_v \mathbf{B} \right) / \left( \mathbf{B} \mathbf{M}^T \mathbf{M} + \delta \mathbf{B} \mathbf{A}^T \mathbf{A} + \mu \mathbf{D}_v \mathbf{B} \right) \quad (9)$$

where  $\mathbf{M} = \mathbf{C} \odot \mathbf{A}$ .

$$\mathbf{C} \leftarrow \mathbf{C} \cdot \mathbf{Y}_{(3)}^T \mathbf{M} / (\mathbf{C} \mathbf{M}^T \mathbf{M}) \quad (10)$$

where  $\mathbf{M} = [(\mathbf{A}_1 \odot \mathbf{B}_1) \mathbf{1}_L \cdots (\mathbf{A}_R \odot \mathbf{B}_R) \mathbf{1}_L]$ .

## 4. EXPERIMENTS

Here, we conduct experiments on synthetic data and real-world data to evaluate the performance of the proposed method. Manifold regularized sparse NMF (MRS-NMF) [3], hypergraph regularized  $L_{1/2}$ -NMF (HGL $_{1/2}$ -NMF) [4], region-based NMF (R-NMF) [5] and matrix-vector NTF (MV-NTF) [6] are selected as alternative methods to compare with proposed method. The spectral angle distance (SAD) and root mean squared error (RMSE) defined in [5] are adopted to evaluate the unmixing performance.

### 4.1. Experiments on Synthetic Data

The synthetic data is generated by using a modified version of the method presented in [5]. Six pure signatures (Carnallite, Ammonio-jarosite, Almandine, Brucite, Axinite and Chlonte) are selected from USGS library to generate endmembers. For abundance generation, we follow the following steps. First, a  $z^2 \times z^2$  synthetic image is dividend into  $z^2$  patches, and each one is a sub-image with  $z \times z$  pixels. Then, each pixel of one

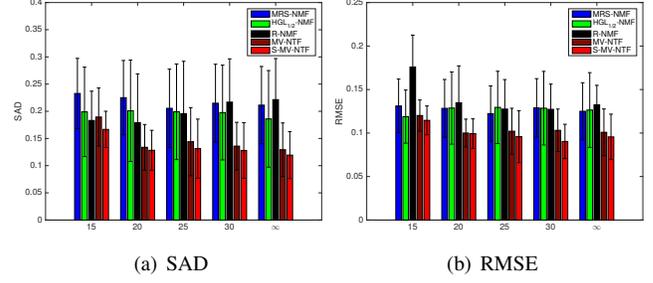


Fig. 2: SAD and RMSE with respect to the noise level.

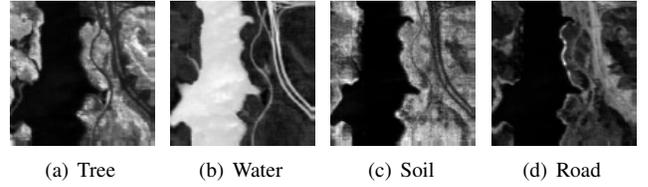


Fig. 3: Abundance maps extracted by S-MV-NTF

region is filled with two kinds of randomly selected endmembers whose proportions are set as  $\beta$  and  $1 - \beta$  respectively. Finally, a  $(z + 1) \times (z + 1)$  lower pass filter is applied to the image to get a highly mixed image. To evaluate the robustness of proposed method under different noise levels, the obtained clean HSI is disturbed by zero-mean white Gaussian noise with pre-specified signal-to-noise ratio (SNR) that is defined as

$$SNR = 10 \log_{10} \frac{E[\mathbf{y}^T \mathbf{y}]}{E[\mathbf{e}^T \mathbf{e}]} \quad (11)$$

where  $\mathbf{y}$  and  $\mathbf{e}$  are the clean signal and the noise at a pixel.  $E[\cdot]$  denotes the expectation operator.

Now, we show the robustness of five methods to different noise levels. The clean HSI that is generated with  $z = 8$ ,  $\beta = 0.8$  is disturbed by noise of various SNRs, 15dB, 20dB, 25dB, 30dB and  $\infty$  (noise-free). Fig. 2 shows the experimental results. The bars and error lines stand for the mean SAD, mean RMSE and their corresponding standard deviations, respectively. It can be seen that the standard deviations of tensor based methods are less than those of the other methods, which shows that tensor based methods are more stable. On the other hand, the proposed method performs the best among all the methods in all the cases, which maybe because super-pixel provides an efficient way for S-MV-NTF to exploit local spatial correlation in an HSI. This contributes S-MV-NTF to a better estimate of abundances and endmembers.

### 4.2. Experiments on Real-world Data

We also conduct experiment on the widely used Jasper Ridge dataset. It is a  $512 \times 614$  image, which was collected by the airborne visible/infrared imaging spectrometer (AVIRIS)

**Table 1:** Means and standard deviations of the SAD on Jasper Ridge Data

Algorithm	MRS-NMF	HGL <sub>1/2</sub> -NMF	R-NMF	MV-NTF	S-MV-NTF
Tree	0.0452 ± 0.02%	0.0755 ± 1.02%	<b>0.0371±0.31%</b>	0.2119 ± 1%	0.2121 ± 0.69%
Water	0.1112±0.48%	<b>0.1±1.11%</b>	0.1596±0.59%	0.2105±3.93%	0.1846±7.64%
Dirt	0.1069±0.70%	0.2221±20.72%	<b>0.0864±0.55%</b>	0.1185±4.16%	0.1118±4.90%
Road	0.8329±7.97%	0.6111±31.54%	0.7190±4.56%	0.1841±3.51%	<b>0.1822±2.04%</b>
Mean	0.2740±2.17%	0.2522±3.30%	0.2505±1.03%	0.1813±1.47%	<b>0.1727±0.12%</b>

sensor over Jasper Ridge in central California, USA. After low SNR and water-vapour absorption bands are removed, a sub-image with  $100 \times 100$  pixels and 198 bands is selected to conduct unmixing in our experiment. In this experiment, we set there are four distinct target of interests, i.e. soil, water, tree, road.

The SAD comparison is shown in Table 1. From the table, we can see that S-MV-NTF performs best among all the methods, especially in the road signature, which may due to the fact that the borders between road and other targets are delicate, and superpixel can adaptively group these pixels into irregular regions so that the spatial structure information can be well incorporated into tensor factorization. A visual comparison is shown in Fig. 3, where Figs. 3(a)-(d) represent tree, water, dirt and road respectively. As can be seen, all the abundance maps are perceptually smooth.

## 5. CONCLUSION

In this paper, an S-MV-NTF method is proposed for hyperspectral unmixing, where the accurate and detailed spatial structure information represented by superpixels is embedded into the original MV-NTF method. In this method, superpixels are firstly generated by SLIC algorithm, then, these superpixels are used to construct similarity graphs. Finally, these graphs are embedded into two factor matrices to enforce the abundances in the same superpixels to be smooth. Experimental results on synthetic data and real-world data demonstrate that our method outperforms other baseline methods. In our future work, other regularizations such as sparsity regularization, and cross-mode similarity regularization will be studied to improve the unmixing performance.

## 6. REFERENCES

- [1] N. Keshava, "A survey of spectral unmixing algorithms," *Lincoln Laboratory Journal*, vol. 14, no. 1, pp. 55–78, 2003.
- [2] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Proc. Advances in Neural Information Processing Systems (NIPS)*, 2001, pp. 556–562.
- [3] X. Lu, H. Wu, Y. Yuan, P. Yan, and X. Li, "Manifold regularized sparse NMF for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2815–2826, 2013.
- [4] W. Wang, Y. Qian, and Y. Y. Tang, "Hypergraph-regularized sparse NMF for hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 2, pp. 681–694, 2016.
- [5] L. Tong, J. Zhou, X. Li, Y. Qian, and Y. Gao, "Region-based structure preserving nonnegative matrix factorization for hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 4, pp. 1575–1588, 2017.
- [6] Y. Qian, F. Xiong, S. Zeng, J. Zhou, and Y. Y. Tang, "Matrix-vector nonnegative tensor factorization for blind unmixing of hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 3, pp. 1776–1792, 2017.
- [7] N. D. Sidiropoulos, L. D. Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3551–3582, 2017.
- [8] Y. Wang, J. Peng, Q. Zhao, Y. Leung, X. L. Zhao, and D. Meng, "Hyperspectral image restoration via total variation regularized low-rank tensor decomposition," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 4, pp. 1227–1243, 2018.
- [9] F. Fan, Y. Ma, C. Li, X. Mei, J. Huang, and J. Ma, "Hyperspectral image denoising with superpixel segmentation and low-rank representation," *Information Sciences*, vol. 397, pp. 48–68, 2017.
- [10] S. Jia, B. Deng, J. Zhu, X. Jia, and Q. Li, "Superpixel-based multitask learning framework for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 2575–2588, 2017.
- [11] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Ssstrunk, "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, 2012.