IOP Conf. Series: Materials Science and Engineering **392** (2018) 062112 doi:10.1088/1757-899X/392/6/062112

A Novel Method for Improving the Accuracy of Hyperspectral Remote Sensing for Detecting Surface Minerals on the Earth

Yang Shao[1], Jinhui Lan[1,*]

[1]School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China

E-mail: lanjh@ustb.edu.cn;

Abstract. Hyperspectral remote sensing technology has made remarkable progress and shows a huge application prospect, such as in mineral distribution detection. Hyperspectral unmixing is a key step in applying hyperspectral remote sensing to detect surface minerals, which extracts endmember spectrum of minerals and detects mineral distribution. Non-negative matrix factorization (NMF) has been introduced into hyperspectral unmixing in the last decade. To reduce the influence of the non-convexity of NMF on spectral unmixing accuracy, the paper proposes a novel hyperspectral unmixing model. The proposed method uses the kernel density function to estimate the intrinsic data structure of hyperspectral images and uses regularization to establish the relation between high-dimensional hyperspectral image and low-dimensional abundance matrix. The proposed method makes the decomposed abundance matrix preserve the hyperspectral data structure, which leads to a more desired spectral unmixing performance. The experimental results on real hyperspectral image prove the superiority of the proposed method in surface mineral detection compared with other typical spectral unmixing methods.

1. Introduction

Hyperspectral remote sensing technology has made remarkable progress. Because of high spectral resolution of the hyperspectral imagery, it can be used as a reference for identifying ground object. Hyperspectral remote sensing technology shows a huge application prospect, such as in mineral distribution detection [1]. Hyperspectral unmixing is a key step in applying hyperspectral remote sensing to detect surface minerals, which extracts endmember spectrum of minerals and detects mineral distribution.

The extraction of endmembers is a key step in spectral unmixing [2]. Many scholars have proposed corresponding methods for endmember extraction, such as N-FINDER [3], VCA [4], SGA [5]. In addition, there are methods that can generate endmember and abundance matrices at the same time, such as non-negative matrix factorization (NMF) [6] and MVES [7]. In the above methods, because NMF can generate endmember matrix and abundance matrix at the same time and is suitable for the extraction of non-pure pixels. However, NMF simply performs mathematical operations and lacks clear geographical significance [8]. To apply NMF to spectral unmixing, many scholars add different constraints to the standard NMF objective function, making the mathematical model more in line with geographical significance and achieving a certain spectral unmixing effect [9-10]. MVC-NMF [9] regards the minimum monomorphic volume formed by the endmember spectrum as a constraint. L1/2-NMF [10] with sparseness constraints is proposed to obtain the best sparse solution.

The above-mentioned NMF algorithms don't explore the global information of hyperspectral image, (i.e., the relation between pixels of the hyperspectral image). The relation between hyperspectral pixels

is referred to data structure of hyperspectral image. The data structure of hyperspectral image can be established using kernel density function. This paper proposes NMF model with new constraint conditions. The new method can preserve the data structure of hyperspectral images to the abundance matrix in low-dimensional space. Because the abundance matrix of the low-dimensional space preserves data structure of the hyperspectral image, the accuracy of the endmember spectral information obtained by the proposed algorithm in the paper is improved.

2.The Proposed Method

The relation between the spectrum of each pixel of the hyperspectral image and other pixels of the same image is called the data structure of hyperspectral image. The data structure of the hyperspectral image can be established by kernel density function. The kernel density function used here is as formula (1).

$$K_\sigma\left(x_i,x_j\right)=\frac{\left(1+\left\|x_i-x_j\right\|^2\right)^{-1}}{\sum_{i\neq j}\left(1+\left\|x_i-x_j\right\|^2\right)^{-1}} \tag{1}$$

x_i , x_j are the i-th and j-th column vectors of the original hyperspectral data X. The data structure of the hyperspectral image is equivalent to the matrix P in formula (2).

$$P=\begin{pmatrix} K_\sigma\left(x_1,x_1\right) & \cdots & K_\sigma\left(x_1,x_n\right) \\ \vdots & \ddots & \vdots \\ K_\sigma\left(x_n,x_1\right) & \cdots & K_\sigma\left(x_n,x_n\right) \end{pmatrix} \tag{2}$$

The formula (3) is used to establish the relation between abundance matrix A and hyperspectral image.

$$\arg.\min\sum_i\sum_j K_\sigma(x_i,x_j)\left\|a_i-a_j\right\|_2^2 \tag{3}$$

Where, a_i , a_j are the column vectors of the abundance matrix A.

We learn from formula (1) that the more similar x_i and x_j are, the bigger $K_\sigma\left(x_i,x_j\right)$ is. It is obvious that the more similar x_i and x_j are, the more similar the abundance a_i of the given pixel x_i is to the abundance a_j of the given pixel x_j. Different weight coefficients $K_\sigma\left(x_i,x_j\right)$ are set according to the difference values between different pixels. The formula (3) is a regularization term that makes the abundance matrix A preserve the data structure of hyperspectral image. We add the formula (3) to the NMF objective function and can get non-negative matrix factorization with data structure invariant constraints algorithm model in Equation (4).

$$\min\left\{\left\|X-EA\right\|_F^2+\gamma\sum_i\sum_j K_\sigma(x_i,x_j)\left\|a_i-a_j\right\|_2^2\right\} \tag{4}$$

In the paper, we abbreviate the non-negative matrix factorization with data structure invariant constraints algorithm model as NMF-DSI.

The NMF-DSI model uses an iterative algorithm to calculate the optimal solution for E and A. The objective function's gradient calculations for E and A are as follows:

$$\nabla_E f(E,A)=2EAA^T-2XA^T \tag{5}$$

$$\nabla_A f(E,A)=2(E^TEA+\gamma*W.*A)-2(E^TX+\gamma*AP) \tag{6}$$

$(.)^T$ represents the transpose matrix of the matrix. W(:,i) in formula (6) is equivalent to $\sum_{j=1}^n K_\sigma\left(x_i,x_j\right)$.

Using the above gradient formulas (5) and (6), we develop the following multiplicative rules.

$$E\Leftarrow E.*(XA^T)./(EAA^T) \tag{7}$$

$$A \Leftarrow A.*(E^T X + \gamma * AP)./(E^T EA + \gamma * W.*A) \tag{8}$$

Where, $.*$ and $./$ represent the multiplication of elements and the division of elements within the matrix, respectively.

In the NMF-DSI algorithm model, we use the VCA algorithm or other endmember extraction algorithm to set the initial value of endmember matrix E. The initial value of the abundance matrix A can be calculated using the least squares method, as follows:

$$A = \max(((E^T E)^{-1} E^T X), 0) \tag{9}$$

The NMF-DSI algorithm model sets two stop criteria in the iterative optimization process. One of the stop criteria is the maximum number of iterations , another stop criteria is that a threshold τ should be specified for the formula (10).

$$\frac{1}{n}\sum_{i=1}^{n}\sqrt{\frac{1}{L}\left\|X - EA\right\|_F^2} \le \tau \tag{10}$$

3. Experimental Results and Analysis

The paper uses real hyperspectral images to verify and analyze the algorithm model. The experiment is to verify the effectiveness of the proposed NMF-DSI algorithm.

Spectral angle distance (SAD) is used to measure spectral unmixing accuracy and is defined in the formula (11). If the calculated SAD value is smaller, then the closer the extracted spectral endmember accuracy is to the true spectral value.

$$SAD_i = \arccos\left(\frac{E_i^T E_i}{\left\|E_i^T\right\|\left\|E_i\right\|}\right) \tag{11}$$

The parameters of the NMF-DSI algorithm model are as follows : (1) the maximum number of iterations is set to 1000 in real hyperspectral data experiments; (2) regularization parameter γ in NMF-DSI algorithm model is experimentally selected from the interval $[0.005n/m^2 \quad 0.1n/m^2]$ for real hyperspectral data experiments; (3) threshold τ should be specified as 0.001.

To evaluate the actual performance of the NMF-DSI algorithm model, this section will analyze the map of the Cuprite mineral area in the western part of Nevada, USA. This image has been obtained by an AVIRIS sensor. Cuprite data is available on the internet. The real spectral library of ground objects in this area is complete and is recorded in the “USGS Spectral Database” [11]. The experiment uses the "USGS spectral database" as a reference standard and uses SAD to measure the accuracy of the extracted endmember spectrum.226 x 176 sub-images have been taken from the raw data for experimentation. Figure 1 shows the 50th band of this image. In experiments, the low SNR band and the water vapor absorption band (1-6, 104-113, 148-167, and 219-224) have been removed. Using the VD method[12] with false alarm probability $P_f = 0.00001$ to estimate the number of endmembers in the selected region is equal to 10 .The endmember spectra extracted by the NMF-DSI algorithm and other algorithms are compared with the USGS library, and spectral similarities are quantified using the SAD standard. As can be seen from Table I, the NMF-DSI algorithm has the largest number of endmember spectral minimum SAD values, and the NMF-DSI algorithm has the smallest endmember spectral SAD mean value.

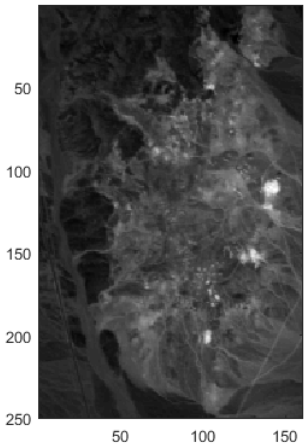
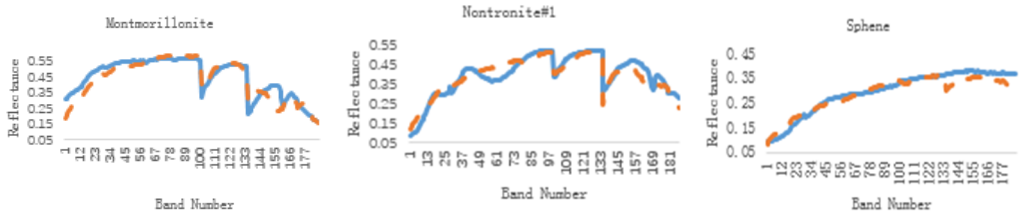


Figure 1. The AVIRIS Cuprite image (band 50).

Figure 2 is a comparison of the endmember spectrum of the NMF-DSI algorithm with the mineral curve of the USGS library. Figure 3 is mineral abundance distribution maps for NMF-DSI, L1/2-NMF, MVC-NMF, and MVES. The color of the distribution map corresponds to the value 0-1. The closer the color is to the blue, the closer the abundance value is to 0. This indicates that there is no such mineral distribution in the area. The closer the color is to the yellow, the closer the abundance value is to 1, indicating that more minerals are distributed in the area. From the mineral distribution map, it can be seen that the mineral distribution map of MVES can hardly find the existence of minerals, and NMF-DSI and L1/2-NMF can obviously find the corresponding mineral distribution. However, the endmember spectral similarity accuracy of L1/2-NMF is worse than that of NMF-DSI. L1/2-NMF only considers the sparsity of abundance. Combined with the mineral abundance maps in Figure 3 and Table I, it can be seen that the NMF-DSI algorithm is superior to other algorithms (L1/2-NMF, MVC-NMF and MVES).

TABLE 1 SAD RESULTS ON THE AVIRIS CUPRITE DATA

	NMF-DSI	L1/2-NMF	MVC-NMF	MVES
Montmorillonite	**0.1043**	0.1095	0.2231	0.1602
Nontronite#1	0.0841	**0.0807**	0.1120	0.1444
Sphene	**0.0707**	0.0741	0.1164	0.2638
Nontronite#2	**0.1212**	0.1253	0.1530	0.1288
Kaolin#1	**0.0985**	0.1068	0.1809	0.1304
Kaolin#2	**0.0748**	0.0785	0.1258	0.1222
Kaolin#3	**0.0904**	0.0956	0.1464	0.2132
Nontronite#3	0.1593	0.1629	0.1639	**0.1565**
Desert_Varnish	0.2106	**0.2088**	0.2370	0.2949
Muscovite	0.0938	0.0927	**0.0874**	0.1127
AVERAGE	**0.1118**	0.1146	0.1596	0.1697



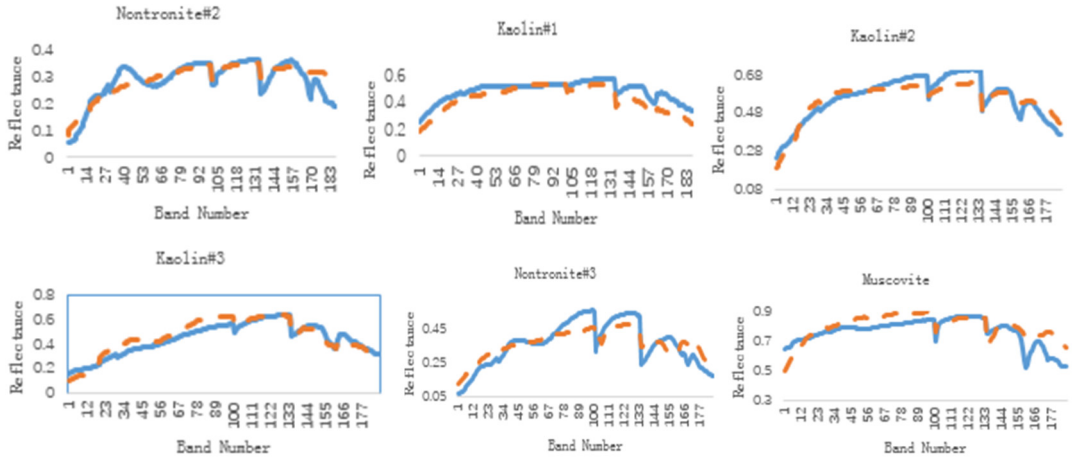


Figure 2. Results on the AVIRIS Cuprite image: Comparison of the USGS library spectra (solid line) with the signatures extracted by NMF- (dotted line).

	MVC-NMF	MVES	L1/2-NMF	NMF-DSI
Montmorillonite				
Nontronite#1				
Sphene				
Nontronite#2				
Kaolin#1				
Kaolin#2				

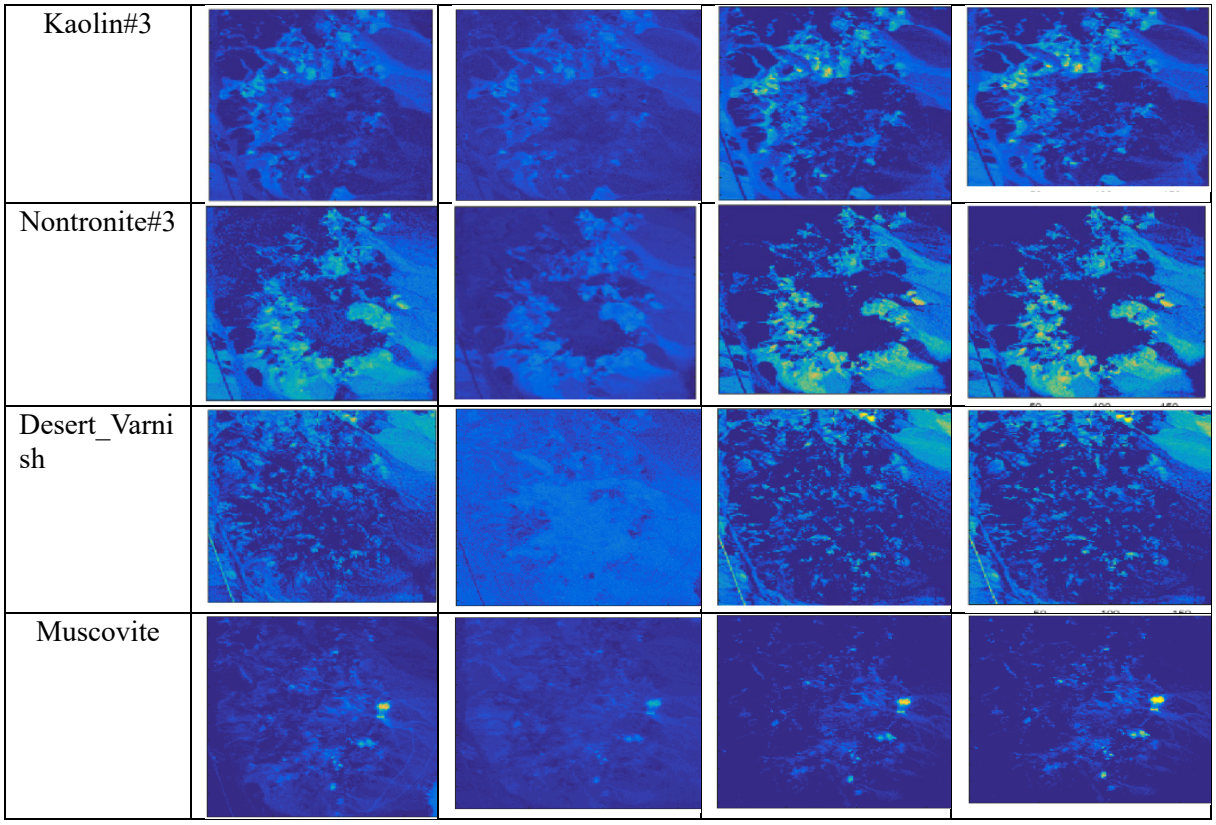


Figure 3. Abundance maps of MVC-NMF, MVES, $L_{1/2}$-NMF, and NMF-DSI for the Cuprite dataset

4. Conclusion

The paper considers the application of global information of hyperspectral imagery (i.e., the relation between pixels of the hyperspectral image) and uses kernel function estimation methods to establish the data structure of the hyperspectral image. The regularization term designed in the paper makes the abundance matrix preserve the data structure of hyperspectral image. Because the abundance matrix preserves data structure of the hyperspectral image, the accuracy of the endmember spectrum obtained by NMF-DSI algorithm is improved. The experimental results of real data prove that NMF-DSI algorithm is superior to the typical algorithms in detecting surface minerals.

Acknowledgments:

The work is supported by the Major Special Project of the China High-Resolution Earth Observation System.

References

[1] Bioucas-Dias J M, Plaza A, Camps-Valls G, et al. Hyperspectral remote sensing data analysis and future challenge. IEEE Geoscience and Remote Sensing Magazine 2013, 1(2): 6-36

[2] LAN Jinhui, ZOU Jinlin, HAO Yanshuang, ZHANG Yuzhen,DONG Mingwei. Research progress on unmixing of hyperspectral remote sensing imagery. Journal of Remote Sensing ,2018,22(01):13-27.

[3] M. E. Winter, “N-FINDR: An algorithm for fast autonomous spectral endmember determination in hyperspectral data,” in Proc. SPIE Conf.Imaging Spectrometry, Pasadena, CA, Oct. 1999, pp. 266–275.

[4] J. Nascimento and J. Bioucas-Dias, “Vertex component analysis: A fast algorithm to unmix hyperspectral data,” IEEE Trans. Geosci. Remote Sens., vol. 43, no. 4, pp. 898–910, Apr. 2005.

[5] C. I. Chang et al., "A new growing method for simplex-based endmember extraction algorithm," IEEE Trans. Geosci. Remote Sens., vol.44, no. 10, pp. 2804–2819, Oct. 2006.
[6] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in Adv. Neural Inform. Process. Syst. Cambridge, MA:MIT Press, 2000, vol. 13, pp. 556–562.
[7] T. H. Chan et al., "A convex analysis based minimum-volume enclosing simplex algorithm for hyperspectral unmixing," IEEE Trans. Signal Process., vol. 57, no. 11, pp. 4418–4432, Nov. 2009.
[8] P.Paura, J.Piper,R. J.Plemmons. Nonnegative matrix factorization for spectral data analysis. Linear Algebra &Its Applications, 2006,416 (1) :29-47.
[9] L.Miao, H.Qi. Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization. IEEE Trans. Geosci. Remote Sens.,2007,45(3) :765-777.
[10] Y Qian, S Jia, J Zhou, A. Robleskelly. Hyperspectral Unmixing Via L1/2 Sparsity-constrained Nonnegtive Matrix Factorization. IEEE Trans. Geosci. Remote Sens., 2014, 49(!1):4282-4297.
[11] CLARK R. N. ,SWAYZ A. G. ,WISEE R., et a1. USGS digital spectral library splih05a[M].[s.n.]: USGS Open File Report, 2003: 3-395.
[12] C. I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," IEEE Trans. Geosci. Remote Sens., vol. 42, no. 3, pp. 608–619, Mar. 2004.