

STRUCTURED DISCRIMINATIVE NONNEGATIVE MATRIX FACTORIZATION FOR HYPERSPPECTRAL UNMIXING

Xue Li¹, Jun Zhou², Lei Tong³, Jianhui Guo¹, Chunxia Zhao¹

¹School of Computer Science and Engineering, Nanjing University of Science and Technology,
Nanjing, Jiangsu, China,

²School of Information and Communication Technology, Griffith University, Nathan, QLD, Australia

³School of Engineering, Griffith University, Nathan, QLD, Australia

ABSTRACT

Hyperspectral unmixing is an important technique for identifying the constituent spectra and estimating their corresponding fractions in an image. Nonnegative Matrix Factorization (NMF) has recently been widely used for hyperspectral unmixing. However, due to the complex distribution of hyperspectral data, most existing NMF algorithms cannot adequately reflect the intrinsic relationship of the data. In this paper, we propose a novel method, Structured Discriminative Nonnegative Matrix Factorization (SDNMF), to preserve the structural information of hyperspectral data. This is achieved by introducing structured discriminative regularization terms to model both local affinity and distant repulsion of observed spectral responses. Moreover, considering that the abundances of most materials are sparse, a sparseness constraint is also introduced into SDNMF. Experimental results on both synthetic and real data have validated the effectiveness of the proposed method which achieves better unmixing performance than several alternative approaches.

Index Terms— Hyperspectral unmixing, nonnegative matrix factorization, sparse, local affinity, distant repulsion

1. INTRODUCTION

Hyperspectral images contain rich spatial and spectral information that are very useful for material identification. Due to the low spatial resolution of hyperspectral remote sensing images, a pixel may contain several disparate substances. Hyperspectral unmixing decomposes a mixed pixel into endmembers which are the spectral responses of a collection of constituent materials, and abundances which are the proportions of endmembers at each location [1]. It is a crucial preprocessing step in many remote sensing applications [2, 3].

Most hyperspectral unmixing approaches are based on Linear Mixture Model (LMM) due to its computational

tractability and flexibility. They are normally performed in a either sequential or simultaneous manner. Several classic methods extract endmember first, then estimate their abundance by fully constrained least squares [4]. These include pixel purity index [5], N-FINDR [6], Vertex Component Analysis (VCA) [7], simplex growing algorithm [8], *et al.* On the contrary, some methods simultaneously estimate the endmembers and abundances. Typical algorithms are independent component analysis [9] and Nonnegative Matrix Factorization (NMF) [10].

NMF is favored in hyperspectral unmixing because it makes the matrix decomposition intuitive and interpretable. Nevertheless, due to the nonconvex objective function, local minima frequently occur and the solution of NMF is not unique. Therefore, a variety of constraints have been introduced to overcome the drawback of standard NMF. Typical examples include $L_{1/2}$ -NMF that uses sparsity constraint [11] and Graph Regularized $L_{1/2}$ -NMF (GLNMF) that uses manifold constraint [12]. Structured constraints have also been introduced to model the complex relationship among data. These methods either combine different manifold structures [13], or explore multilayer factorization techniques [14, 15].

In graph-based methods, the local geometric structure of data is normally modelled based on a local consistency assumption and implemented using the nearest neighbors [16]. There is no constraint on distant data points, which often leads to distorted embedding maps [17, 18]. In order to improve the performance of NMF algorithms, in this paper, we propose a novel method for hyperspectral unmixing, called Structured Discriminative Nonnegative Matrix Factorization (SDNMF). This method preserves the structure of hyperspectral data by incorporating local affinity and distant repulsion constraints. On one hand, local affinity structure forces the local neighborhood relationship be maintained, that is, if x_i and x_j are close in the feature space, their abundance s_i and s_j should also be similar. On the other hand, inspired by the observation that far apart points are generally semantically different, distant repulsion guarantees distant pixels in the feature space

¹This work is supported by the National Science Foundation of China [No.61401214], Natural Science Foundation of Jiangsu Province of China [No.BK20140794] and China Postdoctoral Science Foundation [No.2014M551599].

have dissimilar abundance after unmixing. Additionally, because the sparse contribution of ground substance at various locations in a hyperspectral image, pixels in the same manifold structure are sparsely mixed by a common set of relevant endmembers. Therefore, we also impose a $L_{1/2}$ sparse regularizer on the abundance matrix.

The remainder of this paper is organized as follows. Section 2 gives a brief introduction of the LMM and NMF algorithms. The proposed SDNMF method is presented in Section 3. Experimental results and analysis on both synthetic and real remote sensing data are described in Section 4. Finally, conclusions are drawn in Section 5.

2. RELATED WORK

LMM and NMF are two important models for hyperspectral unmixing. A hyperspectral image with L bands can be described by a matrix $X \in \mathbb{R}^{L \times N}$, where N is the total number of pixels. LMM is expressed as

$$X = AS + E \quad (1)$$

where $A \in \mathbb{R}^{L \times P}$ and $S \in \mathbb{R}^{P \times N}$ refer to an endmember signature matrix and an abundance matrix, respectively. $E \in \mathbb{R}^{L \times N}$ denotes the noises and P is the number of endmembers.

NMF aims at finding two nonnegative matrices A and S simultaneously to approximately represent the nonnegative matrix X . To minimize the reconstruction error, the objective function of NMF can be expressed as

$$\min_{A, S} O(A, S) = \frac{1}{2} \|X - AS\|_F^2, \quad s.t. \quad A \geq 0, S \geq 0 \quad (2)$$

where $\|\cdot\|$ denotes the Frobenius norm. This can be solved by multiplicative update rules [10] as follows

$$A \leftarrow A * XS^T ./ ASS^T; \quad S \leftarrow S * A^T X ./ A^T AS \quad (3)$$

3. APPROACH

In this section, we first explain the framework of structured discrimination, which is helpful to discover the intrinsic geometric relationship of the mixed pixels. Then the objective function of the proposed method and the corresponding iterative updating rules are described.

3.1. Structured Discrimination: Local Affinity and Distant Repulsion

Our method aims to properly preserve the local affinity structure without distorting the distant repulsion property. Therefore, we enforce both local affinity and distant repulsion as explicit constraints on the material abundance.

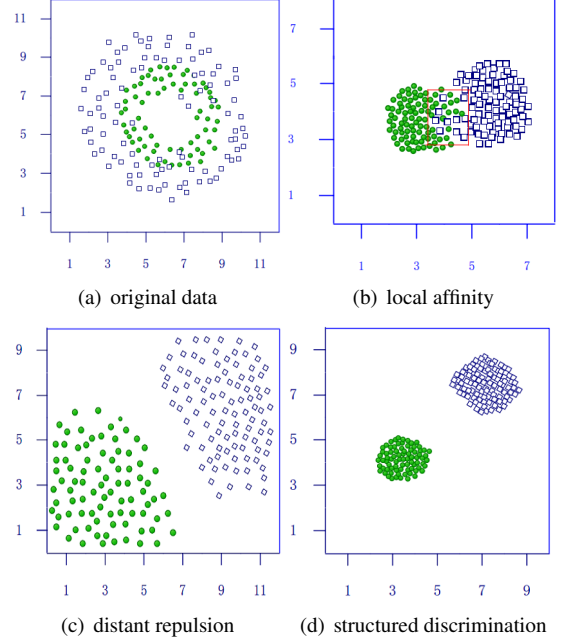


Fig. 1. Illustration of structured discrimination.

Local affinity requires close pixels in the feature space have similar unmixing results. To measure the affinity between x_i and x_j in the original feature space, we adopt a heat kernel [12] to define the affinity matrix, that is, $W_{ij}^l = \exp(-\frac{\|x_i - x_j\|}{2\sigma^2})$, where W_{ij}^l is relatively large if x_i and x_j are similar. In light of its effectiveness, we utilize graph regularization to preserve the local affinity structure

$$\min_s \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N W_{ij}^l \|s_i - s_j\| \quad (4)$$

Through several steps of inference, Eq. (4) can be rewritten as

$$\min_s Tr(SL^l S^T) \quad (5)$$

where $L^l = D^l - W^l$ is the graph Laplacian matrix, D^l is a diagonal matrix whose entries are $D_{ii}^l = \sum_j W_{ij}^l$.

The proposal of distant repulsion property is inspired by elastic embedding [19] which guarantees the mapping of data from its original feature space to a lower dimensional space remains undistorted. In the case of hyperspectral unmixing, distant repulsion enforces the abundance of distant pixels in the original spectral feature space to be dissimilar. We define a repulsive weight matrix $W_{ij}^r = \|x_i - x_j\|$ to measure the repulsion between x_i and x_j . Then distant repulsion constraint can be defined as

$$\min_s \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N W_{ij}^r \exp(-\|s_i - s_j\|^2) \quad (6)$$

The effect of structured discrimination is illustrated in Fig. 1. Fig. 1(a) shows the distribution of two materials, Fig. 1(b) is an example of the distribution of estimated abundance of only considering the local affinity property. This reduces to the aforementioned GLNMF method. It can be clearly seen that the local affinity may result in overlaps (data in the red rectangle) between different materials. Conversely, Fig. 1(c) indicates that only modelling the repulsion correlations without paying attention to the local affinity property leads to loose clusters, *i.e.*, the data is highly dispersed. Fig. 1(d) shows that the performance of structured discrimination leads to non-overlapping between two classes of data while achieving compact distribution.

3.2. Structured Discriminative NMF

Combining both local affinity and distant repulsion constraints defined in Eqs. (5) and (6), and further imposing an $L_{1/2}$ -sparsity constraint, the objective function of SDNMF can be written as follows

$$\min_{A,S} O = \frac{1}{2} \|X - AS\|_F^2 + \frac{\alpha}{2} \text{Tr}(SL^l S^T) + \frac{\beta}{4} \sum_{i,j=1}^N W_{ij}^r \exp(-\|s_i - s_j\|^2) + \gamma \|S\|_{\frac{1}{2}} \quad (7)$$

where the first term is used to measure the reconstruction error of the standard NMF; the second and the third terms are local affinity and distant repulsion constraints, respectively; and the last term is designed to enforce the sparseness of the abundance matrix. α, β, γ are three parameters to balance the weights of the corresponding components.

The objective function of SDNMF in Eq. (7) is not convex when both A and S are to be optimized. Therefore, it is difficult for the optimization algorithm to find the global minima. We introduce Lagrangian multipliers ψ_{im} and ϕ_{mj} for constraints $A_{im} \geq 0$ and $S_{mj} \geq 0$, respectively. Let $\Phi = [\psi_{im}]$ and $\Psi = [\phi_{mj}]$, then the Lagrange \mathcal{L} is

$$\begin{aligned} \mathcal{L} = & \frac{1}{2} \text{Tr}(X^T X) - \text{Tr}(X^T AS) + \frac{1}{2} \text{Tr}(S^T A^T AS) \\ & + \frac{\alpha}{2} \text{Tr}(SL^l S^T) + \frac{\beta}{4} \sum_{i,j=1}^N W_{ij}^r \exp(-\|s_i - s_j\|^2) \\ & + \gamma \|S\|_{\frac{1}{2}} + \text{Tr}(\Psi A^T) + \text{Tr}(\Phi S^T) \end{aligned} \quad (8)$$

The partial derivatives of \mathcal{L} with respect to A and S are

$$\frac{\partial \mathcal{L}}{\partial A} = -XS^T + ASS^T + \Psi \quad (9)$$

$$\frac{\partial \mathcal{L}}{\partial S} = -A^T X + A^T AS + \alpha SL + \frac{\gamma}{2} S^{-\frac{1}{2}} + \Phi \quad (10)$$

where $L = L^l - \lambda \tilde{L}^r$. We define the $\tilde{W}_{ij}^r = W_{ij}^r \exp(-\|v_i - v_j\|^2)$, $W_{ij}^l = W_{ij}^l - \lambda \tilde{W}_{ij}^r$ and their graph Laplacians $\tilde{L}^r =$

$\tilde{D}^r - \tilde{W}^r$ and $L = D - W$ in the usual way, where $\lambda = \beta/\alpha$. According to the Karush-Kuhn-Tucker conditions $\psi_{im} A_{im} = 0$ and $\phi_{mj} S_{mj} = 0$. Introducing $L = L^+ - L^-$, where $L_{ij}^+ = (|L_{ij}| + L_{ij})/2$ and $L_{ij}^- = (|L_{ij}| - L_{ij})/2$, we obtain the following updating rules

$$A \leftarrow A * XS^T ./ ASS^T \quad (11)$$

$$S \leftarrow S * (A^T X + \alpha SL^-) ./ (A^T AS + \alpha SL^+ + \frac{\gamma}{2} S^{-\frac{1}{2}}) \quad (12)$$

4. EXPERIMENTS

In this section, we evaluate the effectiveness of the proposed method. SDNMF is compared with VCA [7], NMF [10], $L_{1/2}$ -NMF [11] and GLNMF [12] on both synthetic and real hyperspectral data. In the experiments, Spectral Angle Distance (SAD) and Root Mean Square Error (RMSE) are adopted to measure the accuracy of the extracted endmembers and their abundances. The detailed definitions of both criteria can be found in [12].

4.1. Experiments on Synthetic Data

We randomly select six spectral signatures from the USGS digital spectral library [20] as candidates to generate the synthetic data. The procedure described in [21] is used to create synthetic images of size 64×64 pixels containing no pure pixels. Then, each image is divided into 8×8 blocks and all pixels in each block are filled up by one type of endmember signature randomly selected from the candidates. To create linear mixture, a spatial low-pass filter of size 9×9 is used to generate the mixed pixels. To remove the probable pure pixels, pixels with abundances greater than 80% are replaced by evenly mixed six endmembers. Finally, zero-mean Gaussian noise is added to the synthetic data to estimate the robustness of the proposed method. The signal-to-noise ratio (SNR) is used to measure the strength of the signal [11].

We first evaluate the performance of different unmixing methods. We set $SNR = 25\text{dB}$ and the total number of endmembers $P = 6$. In practice, local affinity and distant repulsion can be calculated from the k -nearest and k -furthest neighbors of each pixel. In our experiments, the number of nearest and farthest neighbors of each pixel are both set to 5. We run each method for 10 times and then calculate the mean values and standard deviations of SADs and RMSEs, respectively. Fig. 2 shows the experimental results, where the bar and error line stand for the mean SAD and RMSE and their standard deviations, respectively. According to Fig. 2, it can be found that the proposed SDNMF method is superior to other methods. The performances of NMF and $L_{1/2}$ -NMF are limited because they pay no attention to the intrinsic relationship of the hyperspectral data. Furthermore, although both GLNMF and SDNMF consider the structure information of the pixels, the experimental results indicate that SDNMF

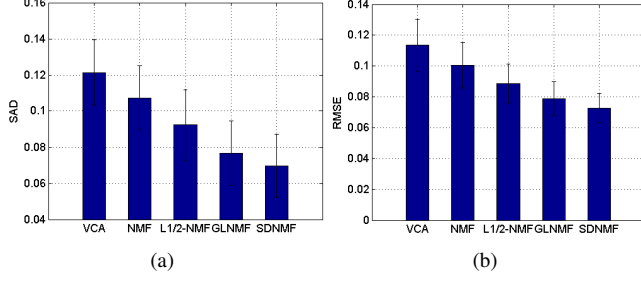


Fig. 2. Performance of different methods. (a) SAD; (b) RMSE.

Table 1. Comparison of methods with different SNRs
(a) SAD

SNR	VCA	NMF	$L_{1/2}$ -NMF	GLNMF	SDNMF
15	0.1523	0.1135	0.1086	0.0758	0.0724
25	0.1255	0.1127	0.0919	0.0739	0.0678
35	0.1243	0.1098	0.0924	0.0663	0.0671
45	0.1224	0.1013	0.0967	0.0753	0.0664
∞	0.1191	0.1067	0.0931	0.0758	0.0644
Mean	0.1287	0.1088	0.0965	0.0734	0.0676

(b) RMSE

SNR	VCA	NMF	$L_{1/2}$ -NMF	GLNMF	SDNMF
15	0.1185	0.1132	0.1049	0.0903	0.0811
25	0.1126	0.1078	0.1012	0.0895	0.0785
35	0.0943	0.0984	0.0954	0.0808	0.0742
45	0.0954	0.0945	0.0927	0.0696	0.0713
∞	0.1007	0.0917	0.0931	0.0705	0.0674
Mean	0.1287	0.1011	0.0975	0.0801	0.0745

consistently outperforms GLNMF. This is mainly due to the fact that when GLNMF encodes the structure information, it only considers the local affinity by the nearest graph regularization, but SDNMF effectively takes both local affinity and distant repulsion constraints into account.

The second experiment evaluates the performance of all five methods under the influence of noise. Table 1 shows the results when SNR varies from 15dB to infinity. It is noticeable that when noise increases, the performances of all methods are worse. Meanwhile, SDNMF achieves the best average performance among all methods. The advantage of SDNMF lies in that it provides a better way to encode the structural information.

4.2. Experiments on Real Hyperspectral Data

In this experiment, SDNMF is applied to an urban hyperspectral data of size of 307×307 captured by Hyperspectral Digital Imagery Collection Experiment (HYDICE). The image contains 210 bands in the 400nm to 2500nm range with spectral resolution of 10nm. After some noisy bands are removed, 162 bands are used for experiments.

In the experiment, four classes of materials are estimat-

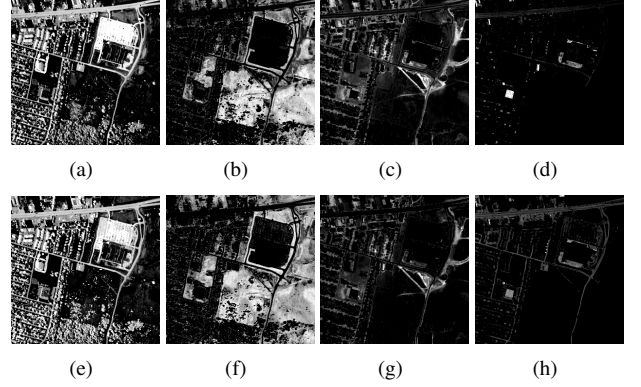


Fig. 3. Abundance maps of different endmembers

ed: Asphalt, Grass, Road and Roof. Figs. 3(a)~(d) show the ground truth reference for the abundance fractions of the endmembers. Figs. 3(e)~(h) give the estimation of endmember abundance maps using our method, in which a dark pixel denotes low abundance of the corresponding endmember and vice versa. It can be seen that the results by our method are consistent with the reference very well. It proves that SDNMF can exhibit more discriminative power by simultaneously taking into account the local affinity and distant repulsion properties, which effectively capture the geometrical structure of the local and distant pixels.

5. CONCLUSION

In this paper, we have presented a new hyperspectral unmixing algorithm called Structured Discriminative Nonnegative Matrix Factorization (SDNMF). With the combination of local affinity, distant repulsion, and sparsity constraints imposed on NMF, SDNMF can discover the intrinsic geometric information of the pixels and use it to discriminate the mixed spectral. SDNMF takes advantage of local affinity property of data to guarantee similar raw data having similar abundance, and simultaneously ensure that dissimilar data have different estimated abundance. As a result, the SDNMF achieves better performance compared with several state-of-the-art methods on both synthetic and real datasets.

6. REFERENCES

- [1] N. Keshava and J.F. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 44–57, 2002.
- [2] J. Solomon and B. Rock, "Imaging spectrometry for earth remote sensing," *Science*, vol. 228, no. 4704, pp. 1147–1152, 1985.

- [3] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. 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] D.C. Heinz and C.I. Chang, "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.
- [5] J.W. Boardman, "Automating spectral unmixing of aviris data using convex geometry concepts," in *Summaries 4th Annu JPL Airborne Geoscience Workshop*, 1993, vol. 1, pp. 11–14.
- [6] M.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*, 1999, pp. 266–275.
- [7] J.M. Nascimento and J.M. 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] C.I. Chang, C. Wu, W. Liu, and Y.C. Ouyang, "A new growing method for simplex-based endmember extraction algorithm," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 10, pp. 2804–2819, 2006.
- [9] J.D. Bayliss, J.A. Gualtieri, and R.F. Crompt, "Analyzing hyperspectral data with independent component analysis," in *26th AIPR Workshop: Exploiting New Image Sources and Sensors*, 1998, pp. 133–143.
- [10] D.D. Lee and H.S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.
- [11] Y. Qian, S. Jia, J. Zhou, and A. Robles-Kelly, "Hyperspectral unmixing via sparsity-constrained nonnegative matrix factorization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4282–4297, 2011.
- [12] X. Lu, H. Wu, Y. Yuan, P. Yan, and X. Li, "Manifold regularized sparse NMF for hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 5, pp. 2815–2826, 2013.
- [13] L. Tong, J. Zhou, X. Bai, and Y. Gao, "Dual graph regularized NMF for hyperspectral unmixing," in *2014 International Conference on Digital Image Computing: Techniques and Applications*, 2014, pp. 1–8.
- [14] R. Rajabi and H. Ghassemian, "Spectral unmixing of hyperspectral imagery using multilayer NMF," *IEEE Transactions on Geoscience and Remote Sensing Letters*, vol. 12, no. 1, pp. 38–42, 2015.
- [15] Z. Shu, J. Zhou, L. Tong, X. Bai, and C. Zhao, "Multilayer manifold and sparsity constrained nonnegative matrix factorization for hyperspectral unmixing," in *2015 IEEE International Conference on Image Processing*, 2015, pp. 2174–2178.
- [16] D. Zhou, O. Bousquet, T.N. Lal, J. Weston, and B. Schölkopf, "Learning with local and global consistency," *Advances in neural information processing systems*, vol. 16, no. 16, pp. 321–328, 2004.
- [17] J. Yang, S. Yan, Y. Fu, X. Li, and T. Huang, "Non-negative graph embedding," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2008, pp. 1–8.
- [18] Y. Peng, S. Wang, S. Wang, and B. Lu, "Structure preserving low-rank representation for semi-supervised face recognition," in *Neural Information Processing*, 2013, pp. 148–155.
- [19] M.A. Carreira-Perpinán, "The elastic embedding algorithm for dimensionality reduction," in *ICML*, 2010, vol. 10, pp. 167–174.
- [20] R.N. Clark, G.A. Swayze, R. Wise, K.E. Livo, T.M. Hoefen, R.F. Kokaly, and S.J. Sutley, "Usgs digital spectral library splib06a: U.s. geological survey, digital data series 231," 2007.
- [21] L. Miao and H. 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.