HYPERSPECTRAL IMAGE SUPER-RESOLUTION BASED ON NON-FACTORIZATION SPARSE REPRESENTATION AND DICTIONARY LEARNING

Xiaolin Han[1], *Jing Yu*[2], *Weidong Sun*[1]

1. State Key Lab. of Intelligence Technology and Systems
Tsinghua National Lab. for Information Science and Technology
Dept. of Electronic Engineering, Tsinghua Univ., Beijing 100084, China
2. Faculty of Information Technology, Beijing Univ. of Technology, Beijing 100124, China

ABSTRACT

Non-negative Matrix Factorization is the most typical model for hyperspectral image super-resolution. However, the non-negative restriction on the coefficients limited the efficiency of dictionary expression. Facing this problem, a new hyperspectral image super-resolution method based on non-factorization sparse representation and dictionary learning (called NFSRDL) is proposed in this paper. Firstly, an efficient spectral dictionary learning method is specifically adopted for the construction of spectral dictionary using some low spatial resolution hyperspectral images in the same or similar areas. Then, the sparse codes of the high-resolution multi-bands image with respect to the learned spectral dictionary are estimated using the alternating direction method of multipliers (ADMM) without non-negative constrains. Experimental results on different datasets demonstrate that, compared with the related state-of-the-art methods, our method can improve PSNR over 1.3282 and SAM over 0.0476 in the same scene, and PSNR over 3.1207 and SAM over 0.4344 in the similar scenes.

Index Terms—*hyperspectral image, super-resolution, non-factorization sparse representation, dictionary learning*

1. INTRODUCTION

High-resolution hyperspectral images have been widely used in various fields, such as environment monitoring, argriculture, military and so on. Nevertheless, the increasing of spectral bands in hyperspectral imaging process, leads to enlarging the size of the photosensitive elements, which results in the limitations on spatial resolution [1]. Normally, the spatial information lost in hyperspectral imaging, can be obtained from high-resolution multi-bands images through some kind of post processing. Therefore, the high-resolution hyperspectral image can be presented by utilizing the high-resolution multi-bands images and the low spatial resolution hyperspectral image.

Recently, hyperspectral image super-resolution methods have been developed rapidly. Specifically, signal processing based methods have been proposed to improve the spatial resolution of hyperspectral image by combining a high spatial resolution panchromatic image with a low spatial hyperspectral image [2]. Furthermore, unmixing based approaches are investigated, which only apply to the images with very similar spectral resolution [3]. Besides, matrix factorization based approaches have played a significant role in hyperspectral image super-resolution. Such as the matrix factorization method based on hyperspectral unmixing, proposed by Kawakami [4], decomposes the image into a basis and a set of sparse coefficients for image super-resolution process. For remote sensing data, a similar method has been proposed by Huang [5], which exploits a downsampling technique in the matrix factorization process. A non-negative sparse approach has also been proposed by Wyckoff [6], which uses the alternating direction method of multipliers (ADMM) for matrix factorization. Yokoya [7] has constructed a coupled matrix factorization method with non-negative constrain for spatial resolution improvement of the hyperspectral images. Akhtar [1] has offered a sparse spatial-spectral representation approach which considers the spectral similarity in the nearby pixels. A non-negative structured sparse representation approach has been proposed by Dong [8], which exploits the strong spectral correlations in similar neighbors.

The non-negative matrix factorization, based on spectral unmixing, is an effective technique for hyperspectral image super-resolution, which decomposes each pixel into endmembers and abundances [4-8]. However, pure spectral signatures are extremely difficult to extract owing to the uncertain numbers of the endmembers. Additionally, the non-negative restriction on the factorization coefficients leads to many errors in hyperspectral representation and high computation complexity [9]. Besides, the performance of non-negative matrix factorization based approach depends on the low spatial resolution hyperspectral images in the same scene, which may be hard to obtain. To deal with the

This work was supported by the National Nature Science Foundation (61501008) and the Natural Science Foundation of Beijing (4172002).

Algorithm 1 The proposed NFSRDL algorithm

Input: Y_R, Y_L, η, λ, μ, K
$D = KSVD(Y_R, K)$
If $\eta = 0$
Compute A via the minimization of Eq.(9)
Else
Compute A via the minimization of Eq.(8)
End
Compute X by Eq.(3)
Output: X

above issues, this paper proposes an approach based on non-factorization sparse representation and dictionary learning (NFSRDL), which is also suitable for the low spatial resolution hyperspectral images in the similar scenes. Firstly, the spectral dictionary is learned by K-singular value decomposition (KSVD) algorithm [10] utilizing the spatial information supplied from the low spatial resolution hyperspectral images in the same or similar scenes, which avoids endmember extraction. Secondly, the sparse codes of the high-resolution multi-bands image with respect to the learned spectral dictionary are estimated using ADMM without non-negative constrains. Finally, the high-resolution hyperspectral image can be obtained with wider applications and better performance.

The remainder is organized as follows. Section 2 formulates the super-resolution framework. Section 3 presents the proposed NFSRDL method. The experiment results and discussions on simulated data are provided in Section 4, followed by the conclusions in Section 5.

2. PROBLEM FORMULATION

The high-resolution hyperspectral image $X \in R^{B\times N}$ is obtained from the high-resolution multi-bands image $Y_L \in R^{b\times N}$ and the low spatial resolution hyperspectral image $Y_R \in R^{B\times n}$, where $N(n)$ and $b(B)$ $(B \gg b, N \gg n)$ represent the number of pixels and the spectral dimensions in Y_L and Y_R. The observed image Y_L and Y_R can be expressed as linear combinations of the high spatial resolution hyperspectral image X:

$$Y_R = XH + E \tag{1}$$

$$Y_L = LX \tag{2}$$

where $L \in R^{b\times B}$ denotes the spectral response function of the multispectral image Y_L; $H \in R^{N\times n}$ denotes the operator of blurring and down sampling as well as $E_R = 0$ when the low spatial resolution multispectral image is obtained in the same scene with Y_L; otherwise, in the similar scenes, $H \in R^{N\times n}$ denotes the matrix transformation process which transforms the X into the similar scene. In addition, the $E \in R^{B\times n}$ denotes the error matrix in transformation, which is close to zero matrix in very similar scenes.

As for sparse representation, the high spatial resolution hyperspectral image X can be represented by the spectral dictionary and sparse coefficients:

$$X = DA \tag{3}$$

where $D \in R^{B\times K}$ ($K \gg B$) is a spectral dictionary, $A \in R^{K\times N}$ is a set of sparse coefficients. Distinguished with non-negative matrix factorization, each column of D in this paper represents a reflectance of the mixed materials instead of the pure signature. Therefore, each column in A is estimated independently without non-negative constraints.

3. PROPOSED SMSR APPROACH

The high-resolution hyperspectral image X can be obtained by spectral dictionary and sparse coefficients according to Eq.(3). The spectral dictionary D can be estimated by Eq.(1) which can be rewritten as:

$$Y_R = D\widehat{A} + E \tag{4}$$

where $\widehat{A} = AH \in R^{B\times n}$. As both D and A are unknown in Eq. (4), the spectral dictionary D can be learned by solving the following sparse decomposition problem:

$$(D, \widehat{A}) = argmin\|Y_R - D\widehat{A}\|_F^2 + \lambda\|\widehat{A}\|_0 \tag{5}$$

Many existing dictionary learning methods, such as KSVD algorithm [10] and online dictionary leaning (ODL) algorithm [11] can be used to solve the above problem. In this paper, the spectral dictionary D is learned by KSVD algorithm. Once the spectral dictionary D is obtained, the sparse coefficients A can be computed from Eq.(2) which can be rewritten in Eq.(6), where the transformed spectral dictionary $\widehat{D}$ equals to LD.

$$Y_L = \widehat{D}A \tag{6}$$

Form the equation above, each column of the sparse coefficient matrix A can be estimated independently, by solving a convex formulation with $l_1 - norm$ constraint below.

$$A = argmin\|Y_L - \widehat{D}A\|_F^2 + \eta\|Y_R - DAH\|_F^2 + \lambda\|A\|_1 \tag{7}$$

When the parameter η equals to zero, it means that the scene of the Y_R is not the same but similar to Y_L, otherwise $\eta \neq 0$. ADMM [12] is exploited to solve the above convex problem in Eq.(7) for fast convergence. Specifically, the augmented Lagrangian function is formulated as bellow:

$$\begin{aligned} L(A, Z, S, V_1, V_2) = {} & \|Y_L - \widehat{D}S\|_F^2 + \eta\|Y_R - ZH\|_F^2 \\ & + \lambda\|A\|_1 + \mu\|S - A + \frac{V_2}{2\mu}\|_F^2 \\ & + \mu\|DS - Z + \frac{V_1}{2\mu}\|_F^2 \end{aligned} \tag{8}$$

$$L(A, S, V_1) = \|Y_L - \widehat{D}S\|_F^2 + \lambda\|A\|_1 + \mu\|S - A + \frac{V_1}{2\mu}\|_F^2 \tag{9}$$

where V_1 and V_2 are Lagrangian multipliers ($\mu > 0$). When the low spatial resolution hyperspectral image Y_R is obtained in the same scene with Y_L ($\eta \neq 0$), minimizing the Eq.(8) leads to the accurate sparse coefficients matrix A for the high-resolution hyperspectral image. When $\eta = 0$, the minimization of Eq.(9) has a better performance on the high-

resolution hyperspectral image estimation. The overall flow of our NFSRDL method is summarized in **Algorithm 1**.

4. EXPERIMENT RESULTS

To evaluate the outperformance of our NFSRDL method, extensive experiments are conducted in comparison with the spatial spectral fusion model (SASFM) method [5], the Generalization of Simultaneous Orthogonal Matching Pursuit (G-SOPM+) method [1] and the Non-negative Structured Sparse Representation (NSSR) method [8] on different image datasets which are normalized in the range of 8-bit grayscale. Furthermore, four objective quality metrics are used to evaluate the various methods, such as mean square error (MSE), peak-signal-to-noise ratio (PSNR), spectral angle mapper (SAM), and universal image quality index (UIQI) [13]. All methods are conducted using MATLAB R2012b on a computer with a 3.60-GHz CPU and 8-GB RAM.

4.1. Experiments on the same scene

In this part of experiments, hyperspectral image acquired on June 19, 1997 (f970619t01p02_r02) by the AVIRS airborne system with dimensions of $512 \times 614 \times 224$, is used. Considering the water absorptions and time costs, only $300 \times 300 \times 93$ of the image is selected in our experiments, which is used as high spatial resolution hyperspectral image. The spectral response function is IKONOS-like spectral responses which covers visible and near infrared spectrum. Therefore, high spatial resolution multi-bands image can be simulated using the above spectral response function. The low spatial resolution hyperspectral image with the dimension of $50 \times 50 \times 93$ is simulated by the 5×5 Gaussian blurring kernel with $\sigma = 1$. On this occasion, the Y_I and Y_R are in the same scene, which means $\eta \neq 0$ in **Algorithm 1**. In addition, the number of columns in spectral dictionary D is chosen to 100; the iteration numbers T is set to 1, and the parameters in augmented Lagrangian function are $\lambda = 10^{-6}$, $\mu = 10^{-3}$, $\eta = 10$.

The performance of our proposed method is shown in Table 1, compared with SASFM, G-SOMP+ and NSSR methods. It can be seen that, our proposed method outperforms all the other methods both in spatial resolution and spectral preservation. Moreover, our proposed method, which is similar to SASFM method, runs faster than the GSOMP and NSSR methods. Fig. 1 shows the visual results of the compared methods on the hyperspectral image at band 10 with one region detailed inside the red square. As can be seen in Fig.1, our proposed method preserves more details of the image, which is obvious in the region with green liner.

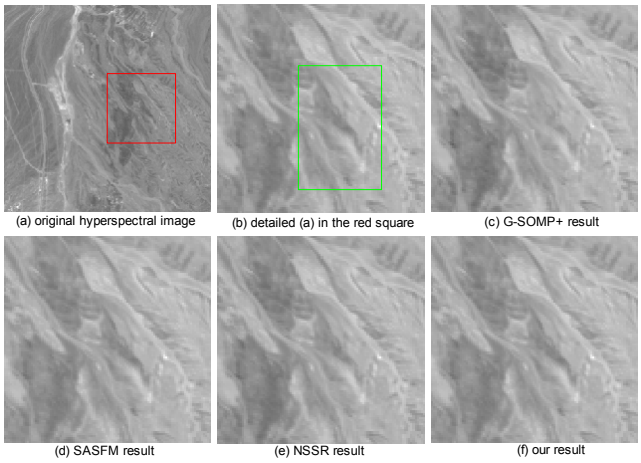

Fig.1 Super-resolution results on the same scene of band 10 in the AVIRS data.

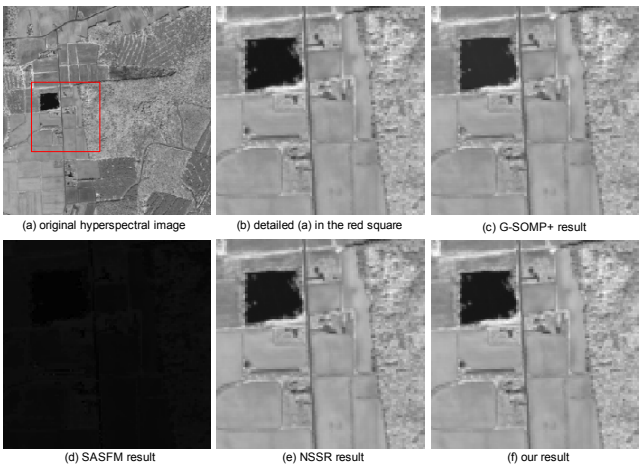

Fig.2 Super-resolution results on the similar scenes of band 40 in the AVIRS datasets.

4.2. Experiments on the similar scenes

In this part of experiments, hyperspectral dataset obtained on July 05, 1996 (f960705t01p02_r05) by the AVIRS, consisting of 7 hyperspectral images in similar scenes with dimensions of $512 \times 614 \times 224$, is used. Removing the interfered bands and taking time costs into account, just $300 \times 300 \times 93$ of the images are reserved, from which the 7th hyperspectral image is used as the high-resolution hyperspectral image. The high-resolution multi-bands image is simulated by the same spectral response function as mentioned above. The low spatial resolution hyperspectral image with dimensions of $50 \times 50 \times 93$ is composed of 2500 spectral vectors from the 1st to the 6th hyperspectral images at random. On this occasion, the Y_I

Table 1 Performance comparisons of different methods on the same scene.

Method	G-SOPM+	SASFM	NSSR	NFSRDL
MSE	1.2862	0.4367	0.5517	**0.3216**
PSNR	47.0377	51.7292	50.7141	**53.0574**
UIQI	0.9918	0.9956	0.9949	**0.9971**
SAM	0.3832	0.3536	0.3675	**0.3060**
TIME/s	82.66	**5.57**	45.15	6.12

Table 2 Performance comparisons of different methods on the similar scenes.

Method	G-SOPM+	SASFM	NSSR	NFSRDL
MSE	3.1043	1861.5003	2.4272	**1.1831**
PSNR	43.2111	15.4322	44.2798	**47.4005**
UIQI	0.9915	0.0380	0.9945	**0.9965**
SAM	1.7956	50.9035	1.6325	**1.1981**
TIME/s	75.58	**6. 33**	418.20	6.41

and Y_R are in the similar scene, where η equals 0 in **Algorithm 1**. For better super-resolution performance, the number of columns in spectral dictionary D is set to 100; the iteration numbers T is set to 1, and the parameters in augmented Lagrangian function are $\lambda = 10^{-4}$, $\mu = 10^{-1}$.

The quantitative and visible results are reported in Table 2 and Fig.2. It can be seen that the G-SOMP+ method performs worst owning to the incorrect estimation of spectral response function. Obviously, our proposed method provides better performance on spectral similarity and spatial edges preserving than the other compared methods in the similar scenes.

5. CONCLUTIONS

In this paper, a non-factorization sparse representation and dictionary learning (NFSRDL) method is proposed for hyperspectral image super-resolution in the same or similar scenes. Spectral dictionary learning takes advantage of spectral information in our proposed method, avoiding the matrix factorization. The sparse codes are estimated without nonnegative constrains, which achieve higher efficiency of spectral dictionary expression. Experiments on different datasets show the superiority of our proposed method both in super-resolution effects and computing time.

6. REFERENCES

[1] N. Akhtar, F. Shafait, and A. Mian, "Sparse spatio-spectral representation for hyperspectral image super-resolution," *IEEE ECCV*, pp. 63-78, 2014.

[2] Luciano Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, & L. M. Bruce, "Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data-fusion contest." *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 10, pp. 3012-3021, 2007.

[3] A. Minghelli-Roman, L. Polidori, S. Mathieu-Blanc, L. Loubersac, & F. Cauneau, "Spatial resolution improvement by merging MERIS-ETM images for coastal water monitoring," *IEEE Geoscience and Remote Sensing Letters*, vol. 3, no. 2, pp. 227-231, 2006.

[4] R. Kawakami, Y. Matsushita, J. Wright, M. Ben-Ezra, Y. W. Tai, & K. Ikeuchi, "High-resolution hyperspectral imaging via matrix factorization," *IEEE CVPR*, pp. 2329-2336, 2011.

[5] B. Huang, H. Song, H. Cui, J. Peng, & Z. Xu "Spatial and spectral image fusion using sparse matrix factorization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 3, pp. 1693-1704, 2014.

[6] E. Wycoff, T. H. Chan, K. Jia, W. K. Ma, & Y. Ma, "A non-negative sparse promoting algorithm for high resolution hyperspectral imaging," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1409-1413, 2013.

[7] N. Yokoya, T. Yairi, A. Iwasaki, "Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 2, pp. 528-537, 2012.

[8] W. Dong, F. Fu, G. Shi, X. Cao, J. Wu, G. Li, & X. Li, "Hyperspectral Image Super-Resolution via Non-Negative Structured Sparse Representation," *IEEE Transactions on Image Processing*, vol. 25, no. 5, pp. 2337-2352, 2016.

[9] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, & 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.

[10] Aharon, Michal, Michael Elad, and Alfred Bruckstein, "K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation," *IEEE Transactions on signal processing*, vol. 54, no. 11, pp. 4311-4322, 2006.

[11] J. Mairal, F. Bach, J. Ponce, &, G. Sapiro, "Online learning for matrix factorization and sparse coding," *Journal of Machine Learning Research*, pp. 19-60, 2010.

[12] S. Boyd, N. Parikh, E. Chu, B. Peleato, & J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1-122, 2011.

[13] Q. Wei, J. Bioucas-Dias, N. Dobigeon, & J. Y. Tourneret, "Hyperspectral and multispectral image fusion based on a sparse representation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 7, pp. 3658-3668, 2015.