Neurocomputing 273 (2018) 171–177

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Super-resolution of hyperspectral image via superpixel-based sparse representation

Leyuan Fang, Haijie Zhuo, Shutao Li*

College of Electrical and Information Engineering, Hunan University, Changsha 410082, China

ARTICLE INFO

Article history:
Received 8 December 2016
Revised 19 May 2017
Accepted 4 August 2017
Available online 18 August 2017

Communicated by Yue Gao

Keywords:
Hyperspectral image
Super-resolution
Sparse representation
Superpixel

ABSTRACT

In this paper, a novel superpixel-based sparse representation (SSR) model is proposed for hyperspectral image (HSI) super-resolution. Specifically, given a HSI with low spatial resolution and a multispectral image (MSI) with high spatial resolution, the proposed SSR approach first learns a spectral dictionary from HSI and constructs a transformed dictionary corresponding to MSI. Then, the SSR method clusters the MSI into superpixels, whose shape and size can be adaptively adjusted according to the local structures. Since pixels within each superpixel have strong similarities, the SSR method simultaneously decomposes them on the transformed dictionary to generate the corresponding fractional abundance coefficient matrix, which can exploit the similarities within the superpixel to improve the sparse decomposition. Finally, the high resolution hyperspectral image can be reconstructed with the obtained fractional abundance coefficient matrix. Experimental results show that the proposed approach is superior to some well-known HSI super-resolution methods.

1. Introduction

Hyperspectral sensor can acquire images in many contiguous and very narrow spectral bands from visible to infrared spectrum. Recently, hyperspectral image (HSI) has been widely applied in many tasks, such as classification [1,2], detection [3,4], and tracking [5]. However, the real acquired HSI usually has low spatial resolution, which will seriously deteriorate the performance in applications. Super-resolution [6] is a technique which aims to enhance resolution in a software manner, and two recent works in [7,8] have utilized the learning method to greatly enhance the resolution of nature images. Therefore, it is also essential to develop efficient super-resolution methods to improve the spatial resolution of HSI.

Compared with HSI, MSI of the same scene has higher spatial resolution, and therefore the rich spatial information is usually utilized to enhance the spatial resolution of HSI. The objective of HSI super-resolution [9] is to fuse the spectral context of the observed HSI with the spatial information of MSI. For HSI super-resolution, unmixing [10] is one of the effective tool. The unmixing technique can obtain endmember matrix and fractional abundance matrix from the HSI and MSI of the same scene, respectively. By combining the endmember matrix with fractional abundance matrix, the final high resolution HSI will be reconstructed [11-13]. Work in [11] adopts the non-negative matrix factorization (NMF) to separately exploit the spectral information and spatial structures, which provides a good super-resolution performance. In general, the NMF usually have multiple factorizations and will result in unstable performance. To ensure the unique of factorization, work in [12] adopts a non-smooth Gauss–Newton algorithm [14] to obtain the sparsest fractional abundance matrix. Work in [13] considers both non-negativity of endmember and sparsity of abundance, which provides a very promising result. However, such a method requires very high computational cost in the unmixing process.

Recently, sparse representation has demonstrated to be powerful tool for many hyperspectral image applications. For example, in [15], a multitask sparsity model is utilized to effectively select the meaningful bands from the redundant HSI image. In [16,17], similar multitask sparsity model has been adopted to effectively exploit the spatial-spectral information for the HSI classification. In [18-21], the joint sparsity model has been combined with the ensemble learning, multitask learning or nonlocal weighting for the HSI classification, which can achieve very high accuracies. In [22,23], similar cubes of the HSI are first grouped and then the joint sparsity is utilized to detect the target, which has obtained very promising detection results. Very recently, the sparsity model has also been applied to HSI super resolution, which successively learns a spectral dictionary from the observed HSI and extracts the spatial sparse coefficients from the observed MSI image, respectively [9,24,25]. The reconstructed high-resolution HSI would be a combination of the obtained spectral dictionary and spatial

* Corresponding author.
E-mail addresses: fangleyuan@gmail.com (L. Fang), haijiezhuo@hnu.edu.cn (H. Zhuo), shutao_li@hnu.edu.cn, 992359791@qq.com (S. Li).

http://dx.doi.org/10.1016/j.neucom.2017.08.019
0925-2312/

coefficient. In [24,25], a pixel-wise sparsity model is used to achieve the HSI super-resolution. However, the pixel-wise sparse representation ignores the similarity of neighbor pixels and its results are not very satisfactory. To cope with this problem, work in [9] simultaneously encodes the pixels within a local fixed window, which can provide a promising reconstruction result. Note that, MSI usually has very complex structures, and thus different variations might still be existed in the fixed windows. This will result in an inferior reconstruction. Therefore, to sufficiently utilize the spatial context of MSI, the local windows should be adaptively adjusted according to the structures of MSI.

In this paper, a superpixel based sparse representation model (SSR) is proposed for the super-resolution of HSI. Firstly, the SSR learns a spectral dictionary via online dictionary learning (ODL) [26] whose atoms represent spectral signatures of materials in the captured area, and a transformed dictionary representing the spectral characteristics of MSI can be constructed by directly selecting several bands from the learned spectral dictionary. Then, the SRR clusters the MSI into superpixels via an efficient oversegmentation algorithm [27]. Each superpixel is a small region whose shape and size can be adaptively adjusted according to the structures of MSI, and thus the pixels within each superpixel have similar spatial structures. To exploit local spatial similarity of superpixel in MSI, the SSR adopts the joint sparse regularization to simultaneously decompose superpixel on the transformed dictionary to obtain the corresponding coefficient matrix. Finally, the high spatial resolution HSI can be reconstructed by multiplying the spectral dictionary with the obtained coefficient matrix.

The main contribution of this paper can be summarized as follows. 1) Instead of using the fixed size patch, this paper introduces an efficient shape adaptive superpixel strategy to better reflect the complex structures of MSI. 2) This paper proposes a superpixel based joint sparsity model, which utilizes the joint sparsity regularization to exploit the spatial-spectral information of superpixels to enhance the performance of HSI super-resolution.

The rest of the paper is organized as follows. In Section II, the sparse representation based HSI super-resolution is briefly reviewed. The proposed SSR method is introduced in Section III. Experimental results are reported in Section IV. Section V concludes the paper.

2. Sparse representation based HSI super-resolution reconstruction

HSI super-resolution reconstruction aims to generate a high spatial resolution HSI $\mathbf{Z} \in \mathbb{R}^{B\times N}$ using the observed low spatial resolution HSI $\mathbf{X} \in \mathbb{R}^{B\times n}$ and high spatial resolution MSI $\mathbf{Y} \in \mathbb{R}^{b\times N}$ of the same scene, where $N = W \times H$ and $n = w \times h$. W, w, H, and h indicate the spatial dimension, and B, b indicate the spectral dimension. Here, $W \gg w$, $H \gg h$, $B \gg b$, we assume that the low spatial resolution HSI $\mathbf{X}$ and the high spatial resolution MSI $\mathbf{Y}$ can be linearly mapped from the target image $\mathbf{Z}$

$$\mathbf{X} = \mathbf{ZH} \tag{1}$$

$$\mathbf{Y} = \Gamma\mathbf{Z} \tag{2}$$

where $\mathbf{H} \in \mathbb{R}^{N\times n} : \mathbb{R}^{B\times N} \to \mathbb{R}^{B\times n}$ is a spatial dimensionality reduction operator, while $\Gamma \in \mathbb{R}^{b\times B} : \mathbb{R}^{B\times N} \to \mathbb{R}^{b\times N}$ is a spectral dimensionality reduction operator.

The sparse representation [9] assumes that a pixel in $\mathbf{Z} \in \mathbb{R}^{B\times N}$ can be represented as linear combinations of several distinct spectral signatures:

$$\mathbf{z}_i \approx \mathbf{D}\alpha_i \tag{3}$$

where $\mathbf{D}$ is a spectral dictionary and $\boldsymbol{\alpha}_i$ is the sparse coefficient. Each atom in $\mathbf{D}$ represents the spectral vector of materials in the captured area. Since the observed HSI $\mathbf{X}$ is the low spatial resolution version of $\mathbf{Z}$, x_i can be represented as:

$$\mathbf{x}_i = \sum_{j\in W_i} h_j\mathbf{z}_j = \mathbf{D}\sum_{j\in W_i} h_j\alpha_i = \mathbf{D}\beta_i \tag{4}$$

where W_i is a window centered on pixel i and h_j denotes the weighting coefficients. Each pixel in $\mathbf{X}$ can be linearly represented by the pixels of $\mathbf{Z}$ within window W_i. As illustrated in (2), pixel-wise form of MSI $\mathrm{y}_i$can be represented as:

$$\mathbf{y}_i = \Gamma\mathbf{z}_i \approx \Gamma\mathbf{D}\alpha_i \tag{5}$$

As can been seen in (4) and (5), we can easily estimate the spectral dictionary $\mathbf{D}$ and coefficient matrix $\mathbf{A} = [\alpha_1, ..., \alpha_i, ..., \alpha_N]$ from the observed $\mathbf{X}$ and $\mathbf{Y}$. Then, the target image $\mathbf{Z}$ would be reconstructed as:

$$\mathbf{Z} \approx \mathbf{DA} \tag{6}$$

For the whole image, we can rewrite (4) as $\mathbf{X} = \mathbf{DB}$, where $\mathbf{B} = [\beta_1, ..., \beta_i, ..., \beta_n]$. Generally, spectral dictionary $\mathbf{D}$ can be obtained by solving the following optimal problem:

$$\arg\min_{\mathbf{D},\mathbf{B}} \frac{1}{2}\|\mathbf{X} - \mathbf{DB}\|_F^2 + \lambda\|\mathbf{B}\| \tag{7}$$

According to (5), we have $\mathbf{Y} = \Gamma\mathbf{DA}$. Therefore, once the spectral dictionary $\mathbf{D}$ is learned, fractional abundance coefficient matrix $\mathbf{A}$ can be obtained by decomposing the MSI $\mathbf{Y}$ on the transformed dictionary $\Gamma\mathbf{D}$. To pursue a compact representation $\mathbf{A}$, Song et al. introduced a pixel based sparse model [24] to solve the following problem:

$$\arg\min_{\alpha_i} \|\alpha_i\|_0 \text{s.t. } \|\mathbf{y}_i - \mathbf{D}\alpha_i\|_F \le \eta_1 \tag{8}$$

The fractional abundance coefficient matrix $\mathbf{A}$ can be constructed as $\mathbf{A} = [\alpha_1, ..., \alpha_i, ..., \alpha_N]$.

In general, pixel-based sparse model in [24,25] ignores the local spatial similarity of the MSI image and thus the spatial information cannot be sufficiently utilized. To address this issue, Akhtar et al. [9] introduces a patch-based sparse model. Specifically, the MSI is firstly divided into many non-overlapped fixed patches, and the ith patch can be denoted as $\mathbf{Y}_i^P$, which is composed of multiple pixels. Then, a joint sparse regularization is utilized to decompose the patch $\mathbf{Y}_i^P$:

$$\arg\min_{\mathbf{A}_i^P} \left\|\mathbf{A}_i^P\right\|_{row\_0} \text{s.t.} \left\|\mathbf{Y}_i^P - \Gamma\mathbf{DA}_i^P\right\|_F \le \eta_2, \mathbf{A}_i^P \ge 0 \tag{9}$$

where $\mathbf{A}_i^P$ is the sparse coefficients matrix of $\mathbf{Y}_i^P$, η_2 denotes the tolerable error, and ℓ_{row_0} denotes the row sparse, which constrains pixels within $\mathbf{Y}_i^P$ to be decomposed on the same atoms with different weights. The final fractional abundance matrix is $\mathbf{A} = [\mathbf{A}_1^P, ..., \mathbf{A}_i^P, ..., \mathbf{A}_{k_1}^P]$, where k_1 is the total number of patches. After estimating the spectral dictionary and fractional abundance vectors from the observed HSI $\mathbf{X}$ and MSI $\mathbf{Y}$, the target image can be reconstructed by $\mathbf{Z} = \mathbf{DA}$.

3. Proposed superpixel based sparse reconstruction model for HSI super-resolution

Since HSI usually has very complex structures, different variations might be existed within a window of fixed size. Thus, the patch-based joint sparsity model cannot reflect such complex structures of HSI. In this paper, we propose a superpixel based sparse representation (SSR) model for HSI super-resolution reconstruction. The SSR can be divided into the following three parts: 1) Learning spectral dictionary from low resolution HIS to extract spectral information; 2) Applying the joint decomposition on the

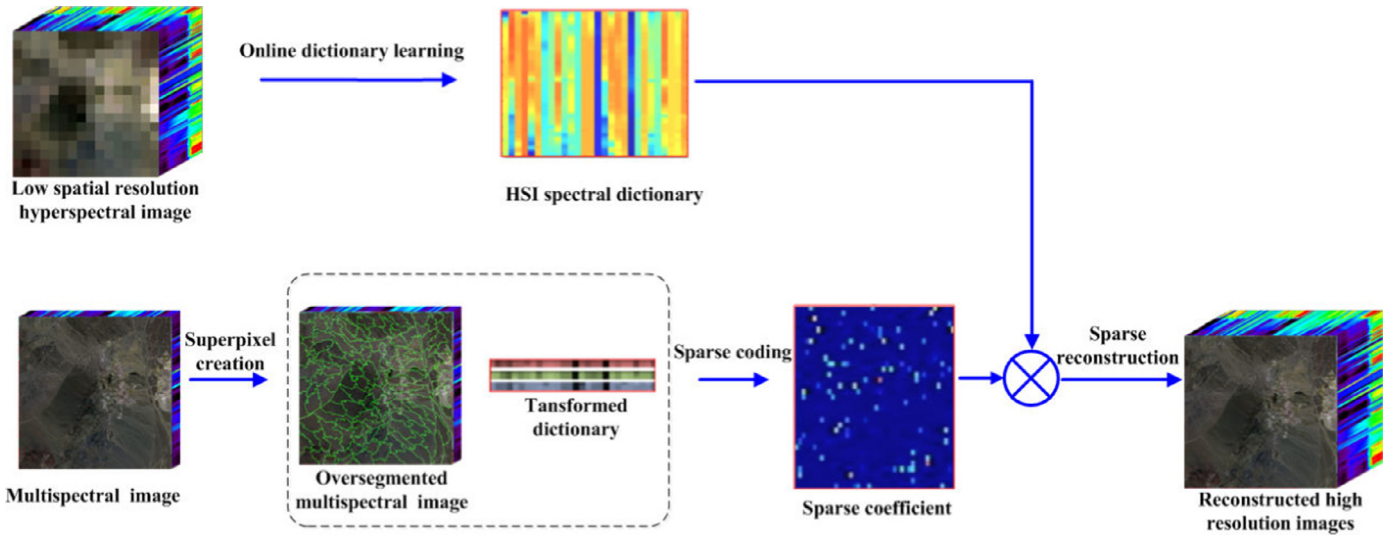

Fig. 1. The whole procedure of the proposed SSR approach.

observed MSI to obtain spatial information; and 3) Fusing the spectral with spatial information to reconstruct the high resolution HSI. More descriptions of the proposed SSR method will be detailed as follows (also illustrated in Fig. 1).

Since dictionary learning has achieved promising result on image reconstruction [28], we first train a spectral dictionary $\mathbf{D}$ from the observed HSI via an online dictionary learning (ODL) algorithm [26]. Specially, given the HSI $\mathbf{X}$, the spectral dictionary can be learned by solving the following optimization problem:

$$\{\mathbf{D},\mathbf{B}\} = \underset{\mathbf{D},\mathbf{B}}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{X}-\mathbf{D}\mathbf{B}\|_F^2 + \lambda\|\mathbf{B}\|_1 s.t. d_k > 0 \tag{10}$$

where $d_k > 0$ means that the spectral vectors of HSI are positive. Once the HSI spectral dictionary $\mathbf{D}$ is obtained, the transformed dictionary corresponding to MSI would be constructed by reducing spectral dimension of HSI spectral dictionary as $\Gamma\mathbf{D}$.

Secondly, the MSI is decomposed on the transformed dictionary to obtain the sparse coefficient carrying the spatial information. To sufficiently utilize the spatial information, the joint decomposition is applied on the superpixels rather than fixed patches. The main steps of the joint decomposition on the superpixel are summarized as follows: 1) superpixels segmentation on MSI; and 2) joint decomposition of superpixels.

To segment the observed MSI more efficiently, we first apply the sparse principal component analysis [29] to get the first principal component. Since the first principal component contain the most important information of the MSI, it is used as the base image to create superpixels segmentation map via entropy rate superpixel (ERS) segmentation [27]. Specially, the base image is first mapped to a graph $G = (V, E)$, where the vertices set V denote the pixels of base image and E is the edge set representing the pairwise similarity among adjacent pixels. Then, the graph image is clustered into K connected subgraphs (corresponding to the superpixels) by choosing a subset A from E. To obtain the homogeneous, compact and balanced superpixels, the entropy rate constraint $\mathcal{H}(A)$ and balance constraint $\mathcal{B}(A)$ are considered and the objective function of superpixels segmentation can be formulated as follows:

$$\max_{A} [\mathcal{H}(A) + \mu\mathcal{B}(A)] \text{ s.t. } A \subseteq E \tag{11}$$

where $\mu \geq 0$ is the weight controlling the contribution of entropy rate term and balance term. The problem can be effectively solved by a greedy algorithm as described in [27]. Once the map is created, the superpixels patch can be extracted from MSI following the guidance of the segmentation map. The procedure of superpixels segmentation map creation is illustrated in Fig. 2.

After obtaining the superpixels, we can jointly decompose superpixels on the transformed dictionary to get the sparse coefficients. In general, pixels within a superpixel $\mathbf{Y}_i^{SP}$ are assumed to have very similar spectral characteristics. The SSR exploits the local similarity by imposing a joint sparse regularization on pixels of each superpixel. Specially, the pixels within each superpixel can be simultaneously decomposed on the same sets of atoms in transformed dictionary $\Gamma\mathbf{D}$ as $\mathbf{Y}_i^{SP} \approx \Gamma\mathbf{D}\mathbf{A}_i^{SP}$, where $\mathbf{A}_i^{SP}$ is the joint coefficients for the $\mathbf{Y}_i^{SP}$. As illustrated in (9), the joint sparse regularization is achieved by placing a ℓ_{row_0} norm on $\mathbf{A}_i^{SP}$. The ℓ_{row_0} norm can restrict the sparse coefficients of pixels in each superpixel to be row sparse, with slight different values reflecting the differences among pixels. Then, the SSR pursuits the $\mathbf{A}_i^{SP}$ by solving the following optimal problem:

$$\arg\min_{\mathbf{A}_i^{SP}} \left\|\mathbf{A}_i^{SP}\right\|_{row\_0} \text{s.t.} \left\|\mathbf{Y}_i^{SP} - \Gamma\mathbf{D}\mathbf{A}_i^{SP}\right\|_F \leq \eta_3, \mathbf{A}_i^{SP} \geq 0. \tag{12}$$

The problem (12) can be efficiently solved by a generalized simultaneous orthogonal match pursuit (G-SOMP+) [9]. The

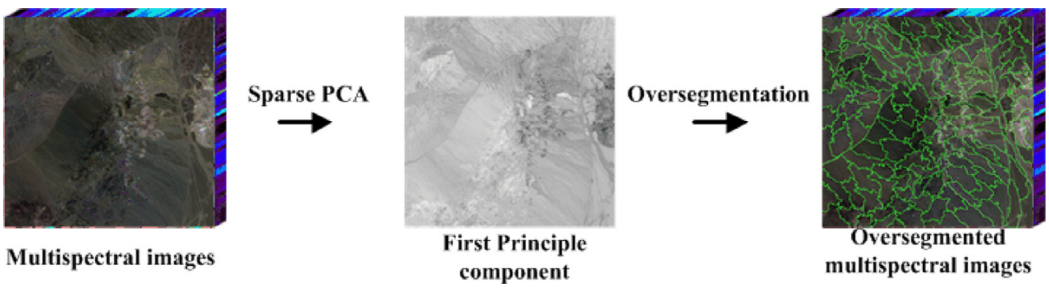

Fig. 2. Procedure of superpixel creation.

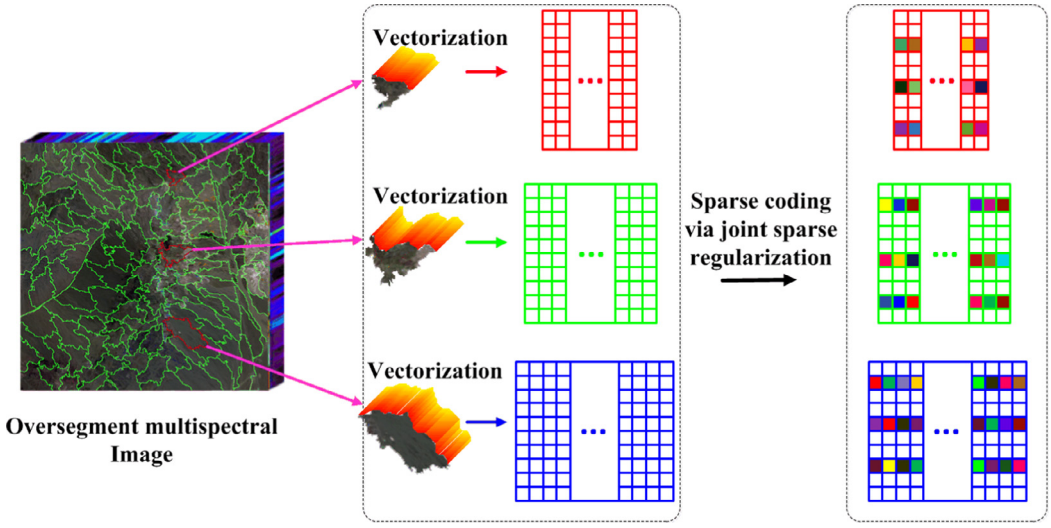

Fig. 3. Schematic diagram of superpixel-based sparse representation.

schematic diagram of the superpixel-based sparse representation is shown in Fig. 3.

Finally, we can reconstruct the high resolution HSI $\mathbf{Z}$ by multiplying the learned spectral dictionary $\mathbf{D}$ with the sparse coefficient$\mathbf{A}$as $\mathbf{Z} = \mathbf{DA}$, where the $\mathbf{A} = [\mathbf{A}_1^{SP}, ..., \mathbf{A}_i^{SP}, ..., \mathbf{A}_K^{SP}]$ is the overall coefficient matrix, and K is the total number of superpixels.

4. Experimental results

4.1. Datasets and quantitative metrics

We test the proposed SSR approach on three popular HSI datasets: Indian Pines [30], Cuprite Mine Nevada [31] and Pavia Center [32]. The Indian Pines image is captured by Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) over the agricultural Indian Pines test site of northwestern Indiana and 20 water absorption bands (104–108, 150–163, and 220) are removed in our experiments. The dataset can be downloaded on the website: http://www.dynamo.ecn.purdue.edu/~biehl/MultiSpec/. We crop the bottom right corner of the data (of a size 512×512) as the ground truth image, since it has rich spatial structure. The electromagnetic wavelength of Cuprite mine Nevada image ranges from 400 nm to 2500 nm with an interval of 10 nm and the spatial resolution is 20 m. This dataset is available on the following website: http://aviris.jpl.nasa.gov/data/free_data.html. In our experiments, as in [9], we crop the top left region (of size 512×512) as the ground truth image after abandoning those bands with low SNR and water absorptions. Pavia Center image is recorded by the Reflective Optics System Imaging Spectrometer (ROSIS-03) sensor over the Center of Pavia area, and its spectrum ranges from 430 nm to 860 nm with an interval of 4 nm. The Pavia Center dataset can be downloaded from: http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes. After abandoning the nosiest band, the bottom right part of this dataset is cropped to the size of $512 \times 512 \times 102$, which is used as the ground truth image in the following experiment.

In our experiment, the ground truth image is downsampled 32 times to obtain the low spatial resolution HSI $\mathbf{X}$. Specifically, each pixel $\mathbf{x}_i \in \mathbf{X}$ is generated by averaging pixels within a 32×32 window of ground truth image centering on location i. For MSI $\mathbf{Y}$, we directly select several bands from the ground truth image. For Indian Pine and Cuprite Mine images, we choose the bands whose center wavelengths are the same as USGS/NASA Landsat 7 images to simulate multispectral image. That is, the bands whose center wavelengths are 480, 560, 660, 830, 1650, and 2220 nm would be chosen. For Pavia Center image, we choose the blue, green, red and near-infrared channel (corresponding to 480 nm, 560 nm, 660 nm, and 830 nm, respectively) of ground truth to simulate multispectral image. Three widely used metrics: root mean square error (RMSE), relative dimensionless global error in synthesis (ERGAS) [33] and spectral angle mapper (SAM) [34] to measure performance of different methods. For the RMSE, ERGAS and SAM metrics, the smaller value corresponds to the better performance.

4.2. Results comparison

We compare the proposed method with two well-known HSI super-resolution approaches: SASFM [25] and G-SOMP+ [9] methods. For the proposed method, the number of atoms number L and superpixels K are two main parameters needed to be tuned. For the atoms number L, we choose it to be 30, 270 and 270 for Cuprite Mine, Indian Pine and Pavia Center images, respectively. For the number of superpixels K, we set it to be 6000 for all the three test images. For the parameters of the two compared methods, their parameters are selected to the default values in [9,25] to reach their best results.

Qualitative and quantitative results are shown in Figs. 4–6, and Table 1, respectively. As can be seen in Fig. 4, the river region in Indian Pine image is not well reconstructed, large error occurred between the SASFM reconstructed image and ground truth. The G-SOMP+ method delivered a better result, but still creating a lot of reconstructed burrs in the river. By contrast, the proposed SSR method can provide a better reconstruction with less burrs. The results on the Cuprite Mine image (see Fig. 5) and the Pavia Center image (see Fig. 6) also show the superiority of the proposed method over the compared methods. This is because the SASFM coded the multispectral image in a pixelwise way, thus ignoring local structure similarity of the image. Although G-SOMP+ [9] method considers local structure similarity within a fixed window, the fixed window may not sufficiently utilize the spatial information. The proposed SSR method can generally better reconstruct the high-resolution HSI and produce less burrs with the adopted superpixel strategy. The experiment results in Table 1 also demonstrate the superiority of our method over other compared methods.

4.3. Effect of the selected atoms number and number of superpixels

Here, we discuss the effects of selected atoms number L and number of superpixels K on the performance of the SSR method.

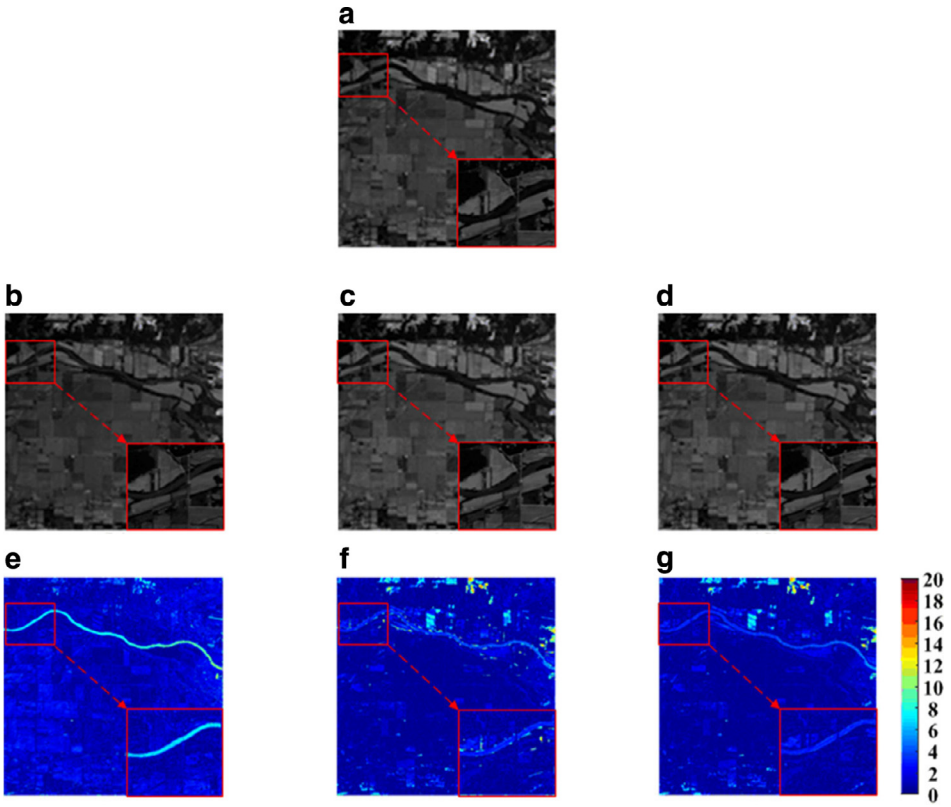

Fig. 4. Comparisons of SASFM [25], G-SOMP+ [9] and proposed SSR methods on the Indian Pines image. (a) Ground truth image, (b) SASFM reconstructed result, (c) G-SOMP+ reconstructed result, (d) SSR reconstructed result, (e) Absolute error between SASFM result and ground truth, (f) Absolute error between G-SOMP+ result and ground truth, (g) Absolute error between SSR result and ground truth.

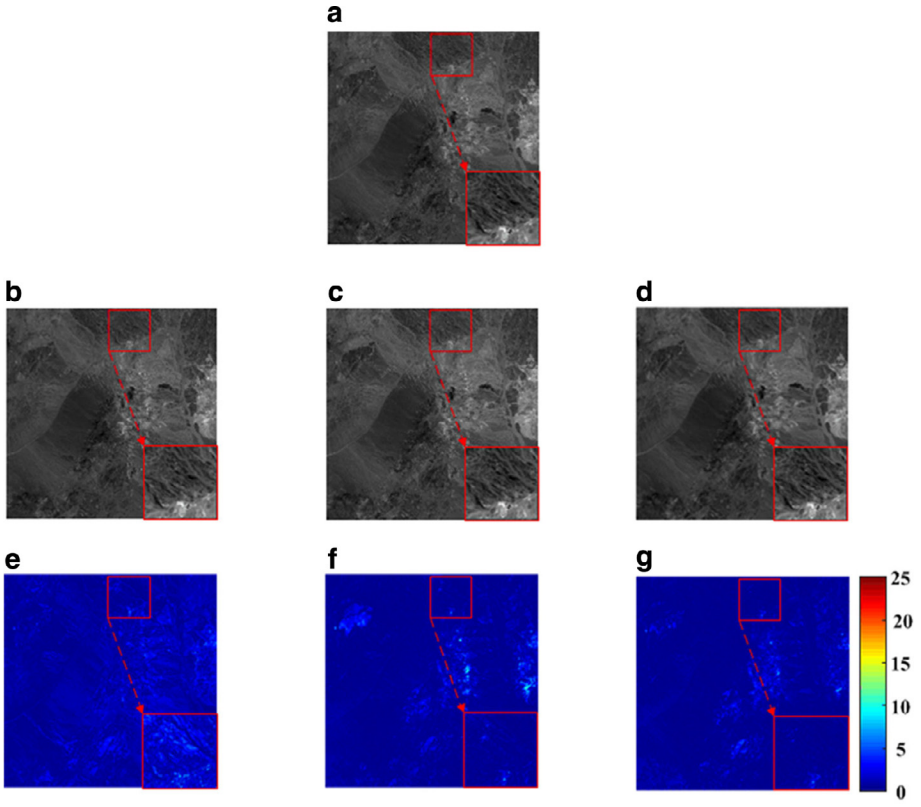

Fig. 5. Comparisons of SASFM [25], G-SOMP+ [9] and proposed SSR methods on Cuprite Mine image. (a) Ground truth image, (b) SASFM reconstructed result, (c) G-SOMP+ reconstructed result, (d) SSR reconstructed result, (e) Absolute error between SASFM result and ground truth, (f) Absolute error between G-SOMP+ result and ground truth, (g) Absolute error between SSR result and ground truth.

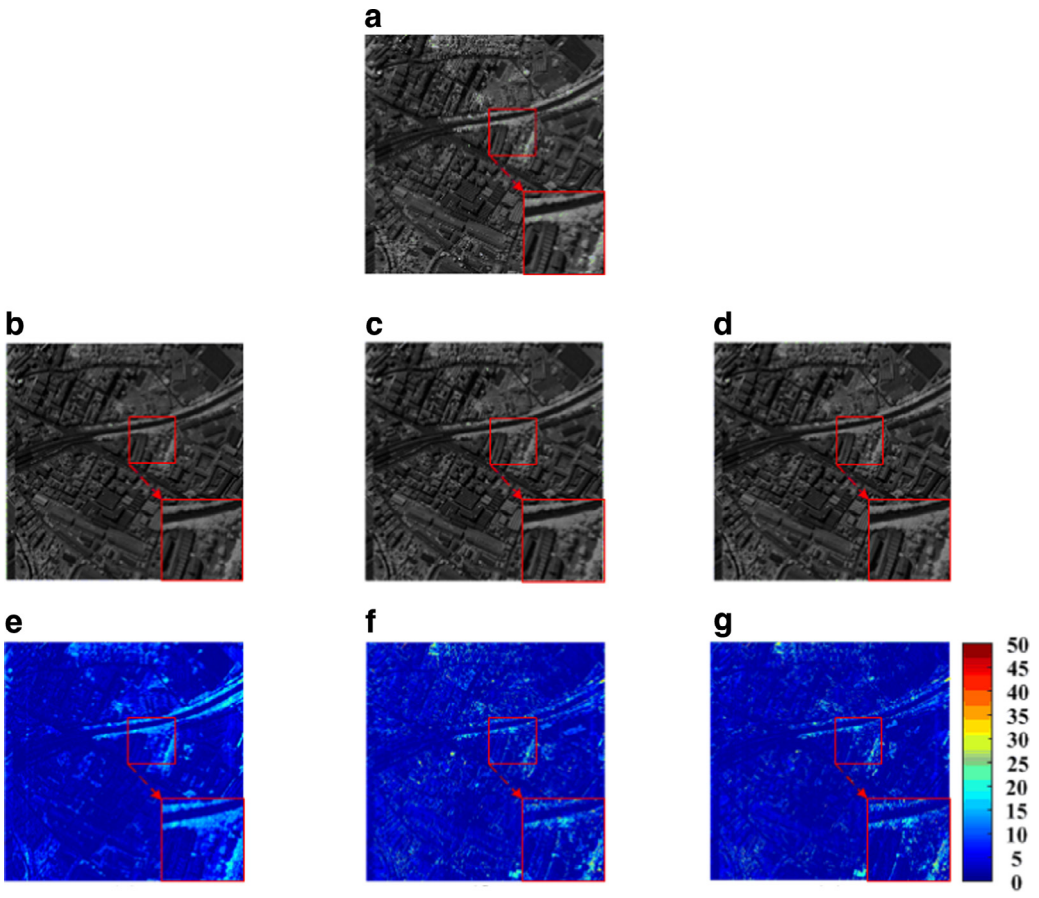

Fig. 6. Comparisons of SASFM [25], G-SOMP+ [9] and proposed SSR methods on Pavia Center image. (a) Ground truth image, (b) SASFM reconstructed result, (c) G-SOMP+ reconstructed result, (d) SSR reconstructed result, (e) Absolute error between SASFM result and ground truth, (f) Absolute error between G-SOMP+ result and ground truth, (g) Absolute error between SSR result and ground truth.

Table 1
Comparisons among different methods: SASFM [25], G-SOMP+ [9] and the proposed SSR. The best result is labeled in bold.

Methods	SASFM [25]	G-SOMP+ [9]	SSR
Indian Pine			
RMSE	1.174	1.125	**1.092**
ERGAS	**0.056**	0.058	**0.056**
SAM	0.871	0.737	**0.716**
Cuprite Mine			
RMSE	1.331	1.016	**0.943**
ERGAS	0.051	0.040	**0.039**
SAM	0.802	0.555	**0.526**
Pavia Center			
RMSE	4.230	3.810	**3.450**
ERGAS	0.301	0.250	**0.235**
SAM	4.038	3.302	**2.981**

Table 2 shows the overall reconstructed performance (in RMSE) under different atoms numbers for all the three datasets. As can be observed, the RMSE value of our proposed method will first decrease and then become stable, as the atom number increases. But, when the atoms number is too large, heavy computational cost will be created on the sparse coding part.

Table 2
The effect of selected atoms number L on the performance of SSR method (in RMSE).

Datasets\L	1	30	60	90	120	150	180	210	240	270	300
Indian Pine	1.22	1.11	1.10	1.10	1.10	1.10	1.09	1.09	1.09	1.09	1.09
Cuprite Mine	0.97	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.93	0.93	0.94
Pavia Center	60.57	3.92	3.77	3.66	3.60	3.67	3.61	3.54	3.49	3.45	4.30

Table 3 shows the influence of the number of superpixels K on the performance of the proposed SSR method (in RMSE). As can be observed, the value of RMSE is relatively high when the number of superpixels is lower than 1000. However, when the number further increases, the value of RMSE becomes comparatively stable. The reason is that when the superpixel number is large, the superpixel size will become small, and thus the pixels within superpixels will be very similar. On the other hand, further reducing superpixel size will also create an additional computational complexity for the sparse coding.

5. Conclusion

In this paper, we proposed a novel superpixel-based sparse representation model (SSR) for hyperspectral image super-resolution

Table 3
The effect of number of superpixels K on the performance of SSR method (in RMSE).

Datasets\K	1	1000	2000	3000	4000	5000	6000	7000	8000	9000	10,000
Indian Pine	1.53	1.09	1.09	1.09	1.09	1.09	1.09	1.09	1.09	1.09	1.09
Cuprite Mine	0.99	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94
Pavia Center	4.20	3.41	3.41	3.41	3.41	3.43	3.45	3.46	3.47	3.48	3.48

reconstruction. With the observed low resolution HSI and MSI in the same scene, the proposed SSR method successively extracts spectral signatures and sparse coefficients via the superpixel based sparse model, and the target HSI is reconstructed by multiplying the spectral signatures by sparse coefficients. Experimental results on three real HSIs demonstrate the superiority of the proposed SSR algorithm over other methods.

In our experiments, the number of superpixels was selected empirically for different images, and so our future work will design an automatic way to select the superpixel number.

Acknowledgments

This work was supported by the National Natural Science Fund of China for Distinguished Young Scholars under Grant 61325007, the National Natural Science Fund of China for International Cooperation and Exchanges under Grant 61520106001, and the National Natural Science Foundation for Young Scientist of China under Grant No. 61501180.

References

[1] L. Fang, S. Li, X. Kang, J.A. Benediktsson, Spectral-spatial hyperspectral image classification via multiscale adaptive sparse representation, IEEE Trans. Geosci. Remote Sens. 52 (12) (2014) 7738–7749.
[2] Y. Gu, H. Liu, Sample-screening MKL method via boosting strategy for hyperspectral image classification, Neurocomputing 173 (1) (2016) 1630–1639.
[3] W. Li, Q. Du, Collaborative representation for hyperspectral anomaly detection, IEEE Trans. Geosci. Remote Sens 53 (3) (2015) 1463–1474.
[4] Y. Gu, Y. Wang, H. Zheng, Y. Hu, Hyperspectral target detection via exploiting spatial-spectral joint sparsity, Neurocomputing 169 (11) (2015) 5–12.
[5] H.V. Nguyen, A. Banerjee, R. Chellappa, Tracking via object reflectance using a hyperspectral video camera, in: IEEE Conference on Computer Vision and Pattern Recognition Workshop, 2010, pp. 44–51.
[6] J. Yang, J. Wright, T.S. Huang, Y. Ma, Image super-resolution via sparse representation, IEEE Trans. Image Process. 19 (11) (2010) 2861–2873.
[7] Q. Wang, Y. Yuan, Learning to resize image, Neurocomputing 131 (5) (2014) 357–367.
[8] Q. Wang, Y. Yuan, High quality image resizing, Neurocomputing 131 (5) (2014) 348–356.
[9] N. Akhtar, F. Shafait, A. Mian, Sparse spatio-spectral representation for hyperspectral image super-resolution, in: European Conference on Computer Vision, 2014, pp. 63–78.
[10] B. Du, S. Wang, N. Wang, L. Zhang, D. Tao, L. Zhang, Hyperspectral signal unmixing based on constrained non-negative matrix factorization approach, Neurocomputing 204 (5) (2016) 153–161.
[11] N. Yokoya, T. Yairi, A. Iwasaki, Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion, IEEE Trans. Geosci. Remote Sens 50 (2) (2012) 528–537.
[12] R. Kawakami, J. Wright, Y.W. Tai, Y. Matsushita, M. Ben-Ezra, K. Ikeuchi, High-resolution hyperspectral imaging via matrix factorization, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011, pp. 2329–2336.
[13] E. Wycoff, T.H. Chan, K. Jia, W.K. Ma, Y. Ma, A non-negative sparse promoting algorithm for high resolution hyperspectral imaging, in: International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2011, pp. 1409–1413.
[14] Q. Geng, H. Wang, J. Wright, On the local correctness of $\ell 1$ minimization for dictionary learning, in: IEEE International Symposium on Information Theory (ISIT), 2014, pp. 3180–3184.
[15] Y. Yuan, G. Zhu, Q. Wang, Hyperspectral band selection by multitask sparsity pursuit, IEEE Trans. Geosci. Remote Sens 53 (2) (2015) 631–644.
[16] Y. Yuan, J. Lin, Q. Wang, Hyperspectral image classification via multitask joint sparse representation and stepwise MRF optimization, IEEE Trans. Cybern 46 (12) (2016) 2966–2977.
[17] L. Fang, C. Wang, S. Li, J.A. Benediktsson, Hyperspectral image classification via multiple-feature-based adaptive sparse representation, IEEE Trans. Instrum. Meas (2017) In Press.
[18] E. Zhang, X. Zhang, L. Jiao, L. Li, B. Hou, Spectral-spatial hyperspectralimage ensemble classification via joint sparse representation, Pattern Recognit 59 (2016) 42–54.
[19] W. Li, Q. Du, Joint within-class collaborative representation for hyperspectral image classification, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens 7 (6) (2014) 2200–2208.
[20] J. Li, H. Zhang, L. Zhang, X. Huang, L. Zhang, Joint collaborative representation with multitask learning for hyperspectral image classification, IEEE Trans. Geosci. Remote Sens 52 (9) (2014) 5923–5936.
[21] H. Zhang, J. Li, Y. Huang, L. Zhang, A nonlocal weighted joint sparse representation classification method for hyperspectral imagery, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens 7 (6) (2014) 2056–2065.
[22] Y. Zhang, B. Do, L. Zhang, T. Liu, Joint sparse representation and multitask learning for hyperspectral target detection, IEEE Trans. Geosci. Remote Sens. 55 (2) (2017) 894–906.
[23] Y. Chen, N.M. Nasrabadi, T.D. Tran, Simultaneous joint sparsity model for target detection in hyperspectral imagery, IEEE Geosci. Remote Sens. Lett 8 (4) (2011) 676–680.
[24] H. Song, B. Huang, K. Zhang, H. Zhang, Spatio-spectral fusion of satellite images based on dictionary-pair learning, Inf. Fusion 18 (7) (2014) 148–160.
[25] B. Huang, H. Song, H. Cui, J. Peng, Z. Xu, Spatial and spectral image fusion using sparse matrix factorization, IEEE Trans. Geosci. Remote Sens 52 (3) (2014) 1693–1704.
[26] J. Mairal, F. Bach, J. Ponce, G. Sapiro, Online dictionary learning for sparse coding, in: International Conference on Machine Learning (ICML), 2009, pp. 689–696.
[27] M.-Y. Liu, O. Tuzel, S. Ramalingam, R. Chellappa, Entropy rate superpixel segmentation, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011, pp. 2097–2104.
[28] Y. Shen, J. Li, Z. Zhu, W. Cao, Y. Song, Image reconstruction algorithm from compressed sensing measurements by dictionary learning, Neurocomputing 151 (3) (2015) 1153–1162.
[29] H. Zou, T. Hastie, R. Tibshirani, Sparse principal component analysis, J. Comput. Graph. Stat 15 (2) (2006) 265–286.
[30] D.A. Landgrebe, L. Biehl, 220 Band Hyperspectral Image: AVIRIS image Indian Pine Test Site 3, Purdue University, West Lafayette, School of Engineering. 1992. http://www.dynamo.ecn.purdue.edu/~biehl/MultiSpec/.
[31] R.O. Green, et al., Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS), Remote Sens. Environ 65 (1998) 227–248.
[32] F. Dell'Acqua, P. Gamba, A. Ferrari, Exploiting spectral and spatial information for classifying hyperspectral data in urban areas, in: IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2003, pp. 464–466.
[33] L. Wald, Quality of high resolution synthesised images: is there a simple criterion? in: International Conference on Fusion Earth Data, 2000, pp. 99–103.
[34] Q. Wei, J. Bioucas-Dias, N. Dobigeon, J.Y. Tourneret, Hyperspectral and multispectral image fusion based on a sparse representation, IEEE Trans. Geosci. Remote Sens 53 (7) (2015) 3658–3668.

Leyuan Fang received the B.S. in 2008 and Ph.D. degree in 2015, all in the College of Electrical and Information Engineering, Hunan University, Changsha, China.

From September 2011 to September 2012, he was a visiting Ph.D. student with the Department of Ophthalmology, Duke University, Durham, NC, USA, supported by the China Scholarship Council. Since Jan. 2017, he has been an associate professor with the College of Electrical and Information Engineering, Hunan University. His research interests include sparse representation and multiresolution analysis in remote sensing and medical image processing.

He has won the Scholarship Award for Excellent Doctoral Student granted by Chinese Ministry of Education in 2011.

Haijie Zhuo received the B.S from Nanchang University, Nanchang China, in 2014. He is currently working toward the M.S degree in the College of Electrical and Information Engineering, Hunan University, Changsha, China.

His research interest includes hyperspectral image super-resolution.

Shutao Li received his B.S., M.S., and Ph.D. degrees in electrical engineering from the Hunan University, in 1995, 1997, and 2001, respectively. He joined the College of Electrical and Information Engineering, Hunan University, in 2001.

He was a Research Associate in the Department of Computer Science, Hong Kong University of Science and Technology, from May 2001 to October 2001. From November 2002 to November 2003, he was a postdoctoral fellow at the Royal Holloway College, University of London, working with Prof. John Shawe-Taylor. During April 2005 to June 2005, he has visited the Department of Computer Science, Hong Kong University of Science and Technology as a visiting professor. Now, he is a full professor with the College of Electrical and Information Engineering, Hunan University. He has authored or coauthored more than 160 refereed papers. His professional interests are compressive sensing, sparse representation, image processing, and pattern recognition.

He is now an Associate Editor of the *IEEE Transactions on Geoscience and Remote sensing (TGRS), IEEE Transactions on Instrumentation and Measurement (TIM)* and is member of the Editorial Board of the *Information Fusion* and the *Sensing and Imaging*. He was a recipient of two 2nd-Grade National Awards at the Science and Technology Progress of China in 2004 and 2006.