Hyperspectral Image Denoising via Noise-Adjusted Iterative Low-Rank Matrix Approximation

Wei He, *Student Member, IEEE*, Hongyan Zhang, *Member, IEEE*, Liangpei Zhang, *Senior Member, IEEE*, and Huanfeng Shen, *Senior Member, IEEE*

***Abstract*—Due to the low-dimensional property of clean hyperspectral images (HSIs), many low-rank-based methods have been proposed to denoise HSIs. However, in an HSI, the noise intensity in different bands is often different, and most of the existing methods do not take this fact into consideration. In this paper, a noise-adjusted iterative low-rank matrix approximation (NAILRMA) method is proposed for HSI denoising. Based on the low-rank property of HSIs, the patchwise low-rank matrix approximation (LRMA) is established. To further separate the noise from the signal subspaces, an iterative regularization framework is proposed. Considering that the noise intensity in different bands is different, an adaptive iteration factor selection based on the noise variance of each HSI band is adopted. This noise-adjusted iteration strategy can effectively preserve the high-SNR bands and denoise the low-SNR bands. The randomized singular value decomposition (RSVD) method is then utilized to solve the NAILRMA optimization problem. A number of experiments were conducted in both simulated and real data conditions to illustrate the performance of the proposed NAILRMA method for HSI denoising.**

***Index Terms*—Denoising, hyperspectral image (HSI), low-rank matrix approximation (LRMA), noise-adjusted iteration, randomized singular value decomposition (RSVD).**

I. Introduction

With the wealth of available spectral information, hyperspectral imagery has drawn a lot of attention from various application fields, which include urban planning, mapping, agriculture, forestry, and monitoring [1], [2]. However, hyperspectral images (HSIs) obtained by multidetectors are often corrupted by different types of noise, which severely degrades the quality of the imagery and limits the precision of the subsequent processing, including classification [3]–[5], segmentation [6], unmixing [7], [8], and target detection [9], [10]. It is therefore important to reduce the noise in hyperspectral imagery.

Manuscript received September 23, 2014; revised November 25, 2014; accepted January 27, 2015. Date of publication March 17, 2015; date of current version July 30, 2015. This work was supported in part by the National Basic Research Program of China (973 Program) under Grant 2011CB707105, in part by the 863 Program under Grant 2013AA12A301, in part by the National Natural Science Foundation of China under Grant 61201342 and Grant 41431175, and in part by Program for Changjiang Scholars and Innovative Research Team in University (IRT1278). *(Corresponding author: Hongyan Zhang.)*

W. He, H. Zhang, and L. Zhang are with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Collaborative Innovation Center for Geospatial Technology, Wuhan University, Wuhan 430079, China (e-mail: zhanghongyan@whu.edu.cn).

H. Shen is with the School of Resource and Environmental Science, Wuhan University, Wuhan 430079, China.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2015.2398433

To date, many different denoising methods have been proposed for hyperspectral imagery. HSI data contain hundreds of spectral channels, and each channel can be regarded as a gray-level image. In this regard, a number of gray-level image denoising methods, such as the nonlocal-based algorithm [11], K-SVD [12], and block-matching 3-D filtering (BM3D) [13], can be adopted to denoise the HSI data band by band. However, this kind of processing method ignores the correlations between different spectral bands, and often results in a relatively low-quality result. Alternatively, the most recent hyperspectral denoising methods focus on improving the HSI quality by exploiting both the spatial and spectral information. One of the most famous methods is principal component analysis (PCA), which uses orthogonal transformation to convert the hyperspectral imagery into a set of linearly uncorrelated variables called principal components (PCs). It assumes that the high-dimensional hyperspectral data underlie a low-dimensional intrinsic space, indicating that the first few PCs contain most of the information, and the rest of the PCs are considered to be noise. Thus, the hyperspectral data are denoised via inverse transformation of the first few PCs. However, it was shown in [14] and [54] that the variance in HSIs does not necessarily reflect real SNR, due to the unequal noise variances found in the different bands. Furthermore, it is also possible that some useful information will be included in the remaining PCs. To address these problems, a number of improved methods have been put forward, including noise-adjusted PCs (NAPCs) transform [14], interference and NAPCs analysis (INAPCA) [15], PCA with wavelet denoising [16], and multivariate multiresolution PCA [17]. In [18], a robust PCA (RPCA) model was adopted to simultaneously remove Gaussian noise, impulse noise, and stripes.

By treating the data as a multidimensional data cube, multidimensional analysis methods have also been adopted for HSI denoising. With the data modeled as a 3-D tensor in [19], multidimensional Wiener filtering (MWF) was adopted to detect the signal subspace. To reduce the nonwhite noise, Liu *et al.* [20] proposed a two-stage process composed of a noise prewhitening procedure and an MWF process. To overcome the problems of uniqueness of decomposition and the estimation of multiple ranks for the Tucker3 (three-mode factor analysis) model, Liu *et al.* [21] exploited a powerful multilinear algebra model named parallel factor analysis (PARAFAC), in which the number of estimated ranks is reduced to one. By treating the data as a 3-D cube, many other state-of-the-art denoising methods, such as VBM3D [22] and BM4D [23], can also be used for HSI processing.

The partial differential equation (PDE)-based methods have also been widely used for denoising in HSI processing. In [24], an anisotropic diffusion model was proposed to process HSIs, and in [25]–[27], total variation (TV)-based methods were adopted to denoise the images. Wavelet analysis is another powerful tool for HSI processing. On the basis that only a few of the wavelet coefficients are large (considered as signals) and the majority are small (considered as noise), a hybrid spatial–spectral derivative-domain wavelet shrinkage noise reduction (HSSNR) approach was proposed in [28]. More often, wavelets are combined with other methods in HSI denoising, e.g., PCA [16], multiway Wiener filtering [29], or first-order roughness penalty (FORP) [30].

For an HSI, adjacent bands will typically exhibit strong correlations. Additionally, nearby pixels in the HSI are also typically highly correlated, exhibiting the low-rank structure of hyperspectral imagery. In [18], we investigated the low-rank property of hyperspectral imagery from the perspective of the linear spectral mixing model. Based on the low-rank property of hyperspectral imagery, the low-rank matrix approximation (LRMA) method has been applied to HSI denoising [18], [31]–[33], [52]. However, in HSIs, as the noise intensity in different bands is different, the above low-rank-based methods may get stuck when obtaining optimal results. To handle this problem, in this paper, we propose a noise-adjusted iterative LRMA (NAILRMA) method to denoise hyperspectral imagery, which is an extension of our previous conference work [53] with third prize in the Student Paper Contest of IGARSS 2014. The main ideas and contributions of the proposed method are summarized as follows.

1) Based on the low-rank property of a clean HSI, NAILRMA denoising framework is proposed to further remove the noise in the high-rank part. As the different bands in HSIs have different noise variances, we update the input image band by band, which is related to the noise level of each band. Along with the iteration, the noise is gradually removed and the signal is enhanced. The proposed noise-adjusted iteration framework is also extended to multiple-type mixed-noise removal.
2) Randomized singular value decomposition (RSVD) [34] is introduced and extended to solve the HSI LRMA problem. Benefiting from probabilistic analysis, the complexity of RSVD is much lower than the traditional low-rank factorizations, including singular value decomposition (SVD) [35] and PCA [36].
3) The proposed noise-adjusted iteration framework provides a new perspective to treat the different noise intensities of each band for HSI denoising. The experimental results confirm that the proposed method clearly improves the denoising results, in comparison with some of the aforementioned techniques, both in quantitative evaluations and the visual effect.

The rest of this paper is organized as follows. After introducing LRMA denoising for HSI, the proposed noise-adjusted iteration framework is described in Section II. An extension of NAILRMA to mixed-noise removal is subsequently discussed in Section III. In Section IV, both simulated experiments and a real data experiment are described and analyzed, followed by the conclusion in Section V.

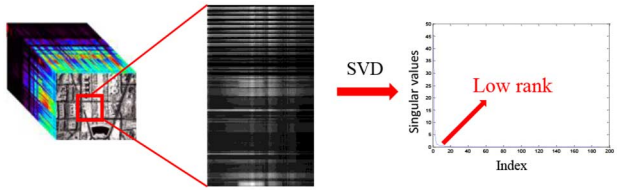


Fig. 1. Singular values of the Casorati matrix from a $20 \times 20 \times 191$ patch extracted from the Washington DC Mall image.

II. Proposed NAILRMA

Observed HSI data, denoted by $\mathbf{Y} \in \mathbb{R}^{MN \times p}$, can be modeled as

$$\mathbf{Y} = \mathbf{X} + \mathbf{N} \tag{1}$$

where $\mathbf{Y} = [\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_p]$ is the Casorati matrix (a matrix whose columns comprise vectorized bands of the HSI) of the degraded HSI $u \in \mathbb{R}^{M \times N \times p}$, $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_p]$ is the Casorati matrix of the clean image $f \in \mathbb{R}^{M \times N \times p}$, and $\mathbf{N} = [\mathbf{N}_1, \mathbf{N}_2, \ldots, \mathbf{N}_p]$ represents the Casorati matrix of the dense noise $\mathcal{N} \in \mathbb{R}^{M \times N \times p}$, which exists in all pixels, and is assumed to obey a Gaussian distribution in this paper. In this model, M represents the width of the HSI data, N stands for the height of the image, and p is the number of spectral bands. $\mathbf{Y}_i$, $\mathbf{X}_i$, and $\mathbf{N}_i \in \mathbb{R}^{MN}$ are the vectors obtained by lexicographically ordering all the pixels of the ith bands of the observed image, clean image, and noise, respectively. The purpose of HSI denoising is to estimate the clean image $\mathbf{X}$ from the noisy image $\mathbf{Y}$. Here, we use $\mathbf{W} = diag(\sigma_1^2, \sigma_2^2, \ldots, \sigma_p^2)$ to denote the noise covariance matrix, and σ_i is the noise standard deviation of band i.

A. Patchwise LRMA Denoising

Hyperspectral data are highly structured, i.e., the HSI lies in a low-dimensional space. From the perspective of the linear spectral mixing model, each spectral signature (row of $\mathbf{X}$) can be represented by a linear combination of a small number of pure spectral endmembers. These considerations inspire us to utilize the LRMA method to solve the denoising problem

$$\min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{X}\|_F^2, \quad s.t \quad rank(\mathbf{X}) \leq r. \tag{2}$$

Consider the matrix $\mathbf{Y}$ of size $MN \times p$; however, the spatial dimensionality greatly exceeds the spectral dimensionality (i.e., $MN \gg p$), and $\mathbf{Y}$ is typically a very thin matrix. In this case, the denoised image by LRMA may result in blurring and loss of details. As a result, we analyze the HSI in a spatial patchwise manner, rather than globally [18], [39].

Let R_b be a binary operator that extracts mn rows from a matrix corresponding to an $m \times n$ spatial patch, specified by the index b within each HSI band, and R_b^* is the inverse. The low-rank property of each patch is shown in Fig. 1. The patchwise LRMA (PLRMA) denoising process can be defined as

$$\text{PLRMA}_r(u) = \left(\sum\nolimits_{b \in \Omega} R_b^* \text{LRMA}_r(R_b u)\right) ./ \left(\sum\nolimits_{b \in \Omega} R_b^* R_b\right) \tag{3}$$

where $\mathrm{LRMA_r}$ is the rank-r approximation [the solution of (2)] of each patch, r is the upper bound of all the Casorati matrices (which is discussed in detail in Section II-B), and Ω denotes the set of overlapping patches that tiles the image domain. In other words, (3) performs LRMA on a family of matrices extracted from the HSI data u, and accumulates a weighted sum of the results. When $m = M, n = N$, and $|\Omega| = 1$, (3) reduces to global LRMA denoising.

B. Noise-Adjusted Iteration Framework

The LRMA solution presented in (2) is more suitable for the case where the noise $\mathbf{N}$ is distributed independently and identically. Unfortunately, in HSIs, the noise intensity in different bands is different, i.e., the noise variance in different columns of $\mathbf{N}$ is different. To thoroughly separate the signal from the noise, instead of whitening the image, an iterative regularization technique combined with LRMA is proposed here to denoise hyperspectral imagery. Iterative regularization is widely used in inverse problems [42], [50], and performance improvements have been reported for both TV [43] and bilateral filtering models [45].

The basic idea of iterative regularization is to add the output denoised image of each iteration back into the input noisy image as the input of the next iteration, i.e.,

$$u^{k+1} = (1-\delta)f^k + \delta u^k \tag{4}$$

where k denotes the iteration number, δ is a relaxation parameter, which is between [0,1], u^k is the input image of the kth iteration, and f^k represents the output image after the PLRMA process on u^k. In the iterative process, if the noise is uniformly distributed in all the bands, the performance improvement is outstanding. However, in most HSIs, different bands have different noise levels. Based on this fact, we extend the idea of iterative regularization, with the selection of δ being related to each band, denoted as

$$u_i^{k+1} = (1-\delta_i)f_i^k + \delta_i u_i^k, \quad i = 1, 2, \ldots, p. \tag{5}$$

We also propose a noise-adjusted determination method for δ_i, based on the noise variance of each band

$$\delta_i = e^{-c(\mathbf{W}(i,i))}, \quad i = 1, 2, \ldots, p. \tag{6}$$

In (5) and (6), u_i^k is the ith band of the kth iteration input image, f_i^k is the ith band of the kth iteration output image, $\mathbf{W}(i,i)$ is the noise variance of the ith band of the original noisy image, and c is a decay parameter, which needs to be predefined.

An intuitive explanation of the decision-making strategy in (5) and (6) is as follows. As the iteration starts, only strong signals (with large singular values) can survive the LRMA processing and contribute to the initial noise-free HSI $\hat{f}$. However, a partially recovered signal will be fed back to the input signal through (5) to lower the estimation of the noise. In return, weaker signals can be identified and added to the signal estimate. In addition, when the noise level is lower in the ith band of the input image, less recovered signal of the ith band will be added back to the input signal in the ith band, and vice versa. This strategy can help to protect the weaker signals in the high-SNR bands, which will be fed back to the performance improvement of the low-SNR bands. As the iteration progresses, it is usually observed that the estimated noise variance monotonically decreases. Meanwhile, the HSI structures are progressively recovered until convergence. The NAILRMA algorithm for HSI denoising is summarized as Algorithm 1.

Algorithm 1. HSI denoising via NAILRMA

Input: Noisy HSI u, noise variance matrix $\mathbf{W}$, upper bound rank r
Output: Denoised HSI $\hat{f}$
Initialization: $u^0 = f^0 = u$, $k = 0$
While $\left\|f^{k+1}(:) - f^k(:)\right\|_2 / \left\|f^k(:)\right\|_2 > \varepsilon$ and $k < iter$, do
  Update u^{k+1} via (5) and (6);
  Update $f^{k+1} = \mathrm{P\,LRMA}_r(u^{k+1})$ via (3);
  Update $k = k + 1$;
End While

C. RSVD Algorithm for LRMA

In the proposed NAILRMA, the major optimization task is to solve (2). This problem can be efficiently solved via SVD, and enjoys a number of optimality properties when the noise is independently and identically distributed Gaussian [35]. Recently, Cai *et al.* [38] proposed a soft-thresholding-based method named singular value thresholding (SVT) to solve problem (2). SVT converts (2) to a slack problem

$$\min \|\mathbf{Y} - \mathbf{X}\|_F^2 + \lambda \|\mathbf{X}\|_* \tag{7}$$

in which $\|\mathbf{X}\|_*$ denotes the sum of all the singular values of $\mathbf{X}$. Candes *et al.* [39] offered a principled and automated way of selecting the regularization parameter in (7) via Stein's unbiased risk estimate (SURE).

Although SVT with SURE can give a naive estimation of a clean image, it still faces some problems in practical denoising. First, the main step of SVT (SVD computation) is time-consuming, containing $O(MNp^2)$ floating-point operations (flops). However, in HSI applications, the datasets can easily reach the million-pixel level, which renders this operation impossible on typical desktop computers. Secondly, this method is effective only in the case of independently and identically distributed Gaussian noise. Unfortunately, in most practical cases, each band of the HSI will have different noise levels. As a result, the input noisy image $\mathbf{Y}$ (vectorized noisy HSI u) should be prewhitened: $\tilde{\mathbf{Y}} = \mathbf{Y}\mathbf{W}^{-1/2}$. However, this method also changes the magnitude value of each band.

Another solution to problem (2) is to apply probabilistic methods which give closely approximated singular vectors and singular values [34], [40], [41], where the complexity is at a much lower level. The proposed RSVD algorithm in [34] explores approximate matrix factorizations by the use of random projections, separating the process into two stages. In the first stage, random sampling is used to obtain a reduced

matrix whose range approximates the range of $\mathbf{Y}$; in the second stage, the reduced matrix is factorized. This process is based on the two facts below [34].

Fact A: For a rank r matrix $\mathbf{Y} \in \mathbb{R}^{MN\times p}$, the range of $\mathbf{Y}$ (the subspace composed by the columns of $\mathbf{Y}$) can be denoted by $\mathbf{\Omega} = \mathbf{YG}$, where $\mathbf{G} \in \mathbb{R}^{p\times r}$ is a Gaussian random matrix.

Fact B: If $\mathbf{Q} \in \mathbb{R}^{MN\times r}$ is a matrix with orthonormal columns which span the range of $\mathbf{Y}$, then

$$\|\mathbf{Y} - \mathbf{QQ}^*\mathbf{Y}\| \approx \min_{rank(\mathbf{X})\le r} \|\mathbf{Y} - \mathbf{X}\|. \tag{8}$$

The RSVD algorithm can be summarized as Algorithm 2.

Algorithm 2. RSVD algorithm

Input: $MN \times p$ matrix $\mathbf{Y}$, and desired rank r
Output: $\mathbf{X}$ (the rank-r approximation of $\mathbf{Y}$)
1 Generate a $p \times r$ matrix $\mathbf{G}$ with independent mean-zero, unit-variance Gaussian entries
2 Form the projection of $\mathbf{Y}$: $\mathbf{H} = \mathbf{YG}$
3 Construct $\mathbf{Q}$ via QR factorization of $\mathbf{H}$: $\mathbf{H} = \mathbf{QR}$, where $\mathbf{Q}$ is an orthogonal basis of the range of $\mathbf{H}$
4 Reconstruct the rank-r approximate matrix $\mathbf{X} = \mathbf{QQ}^*\mathbf{Y}$

From Algorithm 2, it can be seen that the computation of RSVD contains $O(MNp\log(r) + (MN + p)r^2)$ flops, which is more efficient than SVD.

In Algorithm 2, if the singular values of $\mathbf{Y}$ decay gradually, the accuracy of the estimates may be lost. A modified RSVD was introduced in [34], computing $\mathbf{H} = (\mathbf{YY}^*)^q\mathbf{YG}$ instead of $\mathbf{H} = \mathbf{YG}$, since $(\mathbf{YY}^*)^q\mathbf{YG}$ has the same singular vectors as $\mathbf{YG}$, while its singular values decay more rapidly.

D. Adaptive Determination of the Parameters

In the real world, we cannot acquire the information about the noise variance matrix $\mathbf{W}$ and the upper bound rank r of all the patches. In this section, we introduce a noise estimation method. Based on the estimated noise, an upper bound rank estimation technique is also proposed.

Noise estimation is a classical problem for HSIs, and many different techniques have been devoted to it. In this paper, we adopt the multiple regression theory-based approach which was also adopted in [46] and [47]. The main reason for the good performance of multiple regression theory is the strong correlations between neighboring spectral bands.

Let $\mathbf{Y} = [\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_p]$, where $\mathbf{Y}_i$ represents the ith column of noisy image matrix $\mathbf{Y}$, and $\mathbf{Y}_{\partial_i} = [\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_{i-1}, \mathbf{Y}_{i+1}, \ldots, \mathbf{Y}_p]$, which is a $MN \times (p-1)$ matrix. We also assume that $\mathbf{Y}_i$ is explained by a linear combination of the remaining $(p-1)$ bands, denoted as

$$\mathbf{Y}_i = \mathbf{Y}_{\partial_i}\beta_i + \xi_i \tag{9}$$

where $\mathbf{Y}_{\partial_i}$ is the explanatory data matrix, β_i is the regression vector of size $(p-1)\times 1$, and ξ_i is the noise vector of size $MN \times 1$. For each $i \in \{1, \ldots, p\}$, the least squares estimator of the regression vector is given by

$$\hat{\beta} = (\mathbf{Y}_{\partial_i}^T\mathbf{Y}_{\partial_i})^{-1}\mathbf{Y}_{\partial_i}^T\mathbf{Y}_i. \tag{10}$$

As a result, the noise is estimated by

$$\hat{\xi}_i = \mathbf{Y}_i - \mathbf{Y}_{\partial_i}\hat{\beta}. \tag{11}$$

We then obtain the noise matrix $\mathbf{N} = [\hat{\xi}_1, \ldots, \hat{\xi}_p]$ and the correlation matrix by $\mathbf{W} = [\hat{\xi}_1, \ldots, \hat{\xi}_p]^T[\hat{\xi}_1, \ldots, \hat{\xi}_p]/(MN)$. To test the efficiency of this noise estimation method, we randomly added noise to the Washington DC Mall data (the gray values of each HSI band are normalized to [0,1]). Fig. 2 shows the input noise variance of all the bands and the estimated noise variance, which proves the validity of this noise estimation method.

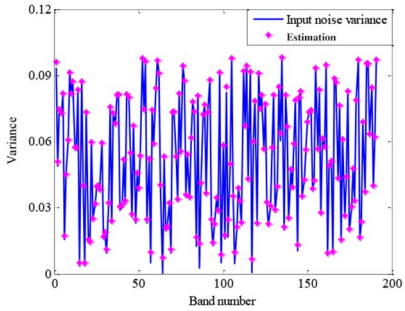


Fig. 2. Input noise variance and the estimated noise variance of each band of the Washington DC Mall data.

In the linear spectral mixing model, the rank of the input matrix denotes the dimension of the signal subspace. Therefore, most of the subspace estimation methods [47], [48] can be applied to the rank estimation in global LRMA denoising. In PLRMA processing, the dimension estimation of HSI can also provide an upper bound for all the patch matrices, in spite of the fact that the ranks of most of the patch matrices are lower than the underlying dimension of the global HSI. In this part, we provide a simple upper bound rank estimation method based on SVD.

For the input noisy HSI data $\mathbf{Y} \in \mathbb{R}^{MN\times p}$, the noise matrix is $\mathbf{N} = [\hat{\xi}_1, \ldots, \hat{\xi}_p]$, where $\hat{\xi}_i$ is estimated via (11). We apply SVD to the matrices $\mathbf{Y}$ and $\mathbf{N}$, respectively, and obtain the singular values of $\mathbf{Y}$and $\mathbf{N}$, denoted by $[\sigma_1, \ldots, \sigma_p]$, $[S_1, \ldots, S_p]$. We can now find the parameter value r, such that $\sigma_r \ge S_1$ and $\sigma_{r+1} < S_1$. The upper bound of the rank is set to r, meaning that the first r-dimension subspaces contain more signals. Fig. 3 shows the first 20 singular values of noisy image $\mathbf{Y}$ and noise $\mathbf{N}$, with the simulated case shown in Fig. 2. The estimated upper bound rank is 7, and the next step of the denoising is to separate the noise from the seven-dimension signal subspaces for each patch.

Up to now, only three parameters need to be specified in the NAILRMA algorithm: the patch size, the step size, and the decay parameter c in (6). The patch size is set to $20 \times 20 \times p$, in order to keep pace with the low-rank matrix recovery (LRMR) method in [18]. The step size is set as 8×8, i.e., we only select patches every eight pixels along both the horizontal and vertical

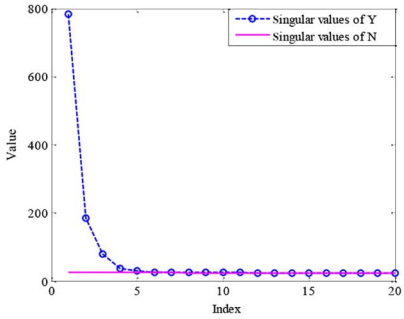


Fig. 3. First 20 singular values of noisy image $\mathbf{Y}$ and noise $\mathbf{N}$. The input noise variance is shown in Fig. 2.

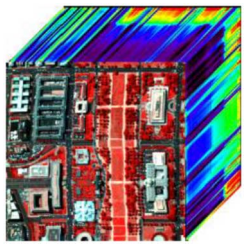


Fig. 4. HYDICE Washington DC Mall dataset used in the simulated experiment (R: 60, G: 27, and B: 17).

directions, to speed up the computation. From the most intuitive point of view, it seems necessary to re-estimate the noise of u^k in the kth iteration, in order to determine the combination of the input and output images. However, this is time-consuming work and, furthermore, the noise estimation may go wrong after so many iterations, resulting in a poor performance. On the other hand, in each patchwise RSVD processing, the noise attenuation rate of each band is constant due to the unified upper bound rank estimation. Thus, the noise variance rate of each different band does not change. We simply set the decay parameter c as 5, ignoring the absolute value but centralizing on the relative value of the iteration parameter δ ($\delta(i) = e^{-c(\mathbf{W}(i,i))}$), as introduced in (6). We also give the reason for this decision in the experimental part.

III. Extension to Mixed-Noise Removal

As introduced in [18], HSIs are often degraded by a mixture of various kinds of noise in the acquisition process, which can include Gaussian noise, impulse noise, dead lines, stripes, and so on. In this case, the noise degradation model of the HSI can be written as

$$\mathbf{Y} = \mathbf{X} + \mathbf{S} + \mathbf{N} \tag{12}$$

where $\mathbf{S}$ is the Casorati matrix of the sparse noise $s \in \mathbb{R}^{M\times N\times p}$, which is the mixture of impulse noise, dead pixels or lines, and stripes. $\mathbf{Y}$, $\mathbf{X}$, and $\mathbf{N}$ are the matrices introduced before. The low-rank approximation denoising is then extended to the well-known RPCA model [37], [44]

$$\min_{\mathbf{X},\mathbf{S}} \|\mathbf{Y}-\mathbf{X}-\mathbf{S}\|_F^2 \text{ s.t. } rank(\mathbf{X}) \le r, \quad card(\mathbf{S}) \le b. \tag{13}$$

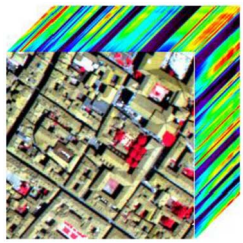


Fig. 5. Pavia city center dataset used in the simulated experiment (R: 80, G: 34, and B: 9).

Here, b stands for the upper bound cardinality of the sparse noise. The LRMR [18] algorithm, which utilizes the RPCA model (13) to denoise the HSI patch by patch, as in (3), is one of the start-of-the-art HSI mixed-noise removal methods. To further improve the dense noise removal performance of the LRMR algorithm, the proposed noise-adjusted iteration framework, as presented in Algorithm 1, is also adopted, combined with LRMR (denoted as NAILRMR), to denoise the HSI. The only modification is to utilize the RPCA model (13) instead of (2) in step 6) of Algorithm 1. Clearly, the NAILRMA method is a special case of NAILRMR, with consideration of the presence of only dense noise.

The major task is now the optimization of (13). To date, many different methods have been proposed for the problem of (13). In this paper, we extended the RSVD algorithm to solve this RPCA model, which is summarized as Algorithm 3.

Algorithm 3. Extended RSVD for the RPCA model

1 **Input:** $MN \times p$ matrix $\mathbf{Y}$, desired rank r, upper bound cardinality b, and stopping criterion ε
2 **Output:** $\mathbf{X}$
3 **Initialize:** $\mathbf{X}_0 := \mathbf{Y}, \mathbf{S}_0 := 0, k := 0, stopcriterion = \varepsilon + 1$
4 **While** $stopcriterion > \varepsilon$ **do**
5 $k = k + 1$;
6 Update $\mathbf{X}_k$: the rank-r approximation of $(\mathbf{Y} - \mathbf{S}_{k-1})$ via Algorithm 2
7 Update $\mathbf{S}_k := \mathrm{P}_\Omega(\mathbf{Y} - \mathbf{X}_k)$, where Ω is the nonzero subset of the first b largest entries of $|\mathbf{Y} - \mathbf{X}_k|$;
8 Update $stopcriterion := \|\mathbf{Y} - \mathbf{X}_k - \mathbf{S}_k\|_F^2 \big/ \|\mathbf{Y}\|_F^2$
9 **End While**

IV. Experimental Results and Discussions

Both simulated and real image data experiments are undertaken to demonstrate the effectiveness of the NAILRMA method for HSI denoising. To thoroughly evaluate the performance of the proposed algorithm, we select five different noise removal methods for comparison, i.e., the SSAHTV model [26], the VBM3D method [22], the BM4D method [23], SURE-SVT [39], and LRMR [18]. Before the denoising processing, the gray values of each HSI band are normalized to [0,1]. The SSAHTV method utilizes a regularization parameter λ to balance the fidelity term and the TV regularization term, and

TABLE I
REGULARIZATION PARAMETER SELECTION IN SSAHTV AND LRMR

		Case 1					Case 2
		0.02	0.04	0.06	0.08	0.1	
SSAHTV	Washington DC Mall	$\lambda=8$	$\lambda=3$	$\lambda=2$	$\lambda=2$	$\lambda=1$	$\lambda=3$
	Pavia city center	$\lambda=8$	$\lambda=5$	$\lambda=3$	$\lambda=2$	$\lambda=2$	$\lambda=3$
LRMR	Washington DC Mall	$r=11$	$r=9$	$r=7$	$r=7$	$r=7$	$r=7$
	Pavia city center	$r=7$	$r=6$	$r=5$	$r=5$	$r=4$	$r=5$

TABLE II
QUANTITATIVE EVALUATION OF THE DIFFERENT DENOISING ALGORITHMS WITH WASHINGTON DC MALL IMAGE IN CASE 1

Noise variance	Evaluation index	SSAHTV	VBM3D	BM4D	SURE-SVT	LRMR	NAILRMA
0.02	MPSNR (dB)	36.74	42.77	42.93	43.79	44.79	**46.59**
	MSSIM	0.9628	0.9902	0.9880	0.9937	0.9936	**0.9954**
0.04	MPSNR (dB)	32.48	38.38	38.85	39.76	40.48	**42.64**
	MSSIM	0.9094	0.9745	0.9728	0.9860	0.9864	**0.9896**
0.06	MPSNR (dB)	30.33	35.79	36.00	37.21	37.81	**40.12**
	MSSIM	0.8584	0.9549	0.9543	0.9771	0.9730	**0.9826**
0.08	MPSNR (dB)	28.53	33.77	34.39	35.40	35.80	**38.21**
	MSSIM	0.7999	0.9280	0.9328	0.9672	0.9588	**0.9742**
0.1	MPSNR (dB)	27.65	32.12	32.92	33.99	34.22	**36.92**
	MSSIM	0.7512	0.9167	0.9121	0.9563	0.9425	**0.9667**

TABLE III
QUANTITATIVE EVALUATION OF THE DIFFERENT DENOISING ALGORITHMS WITH THE PAVIA CITY CENTER IMAGE IN CASE 1

Noise variance	Evaluation index	SSAHTV	VBM3D	BM4D	SURE-SVT	LRMR	NAILRMA
0.02	MPSNR (dB)	36.79	43.41	43.55	42.87	44.84	**46.53**
	MSSIM	0.9663	0.9923	0.9924	0.9935	0.9944	**0.9949**
0.04	MPSNR (dB)	33.14	38.92	39.23	38.75	40.30	**41.40**
	MSSIM	0.9285	0.9793	0.9803	0.9847	0.9850	**0.9866**
0.06	MPSNR (dB)	30.91	36.34	36.74	36.06	37.29	**38.74**
	MSSIM	0.8844	0.9633	0.9662	0.9735	0.9723	**0.9763**
0.08	MPSNR (dB)	29.30	34.38	35.00	34.13	35.24	**36.80**
	MSSIM	0.8363	0.9431	0.9504	0.9610	0.9571	**0.9647**
0.1	MPSNR (dB)	28.36	32.71	33.68	32.61	33.80	**35.27**
	MSSIM	0.8079	0.9167	0.9342	0.9474	0.9395	**0.9534**

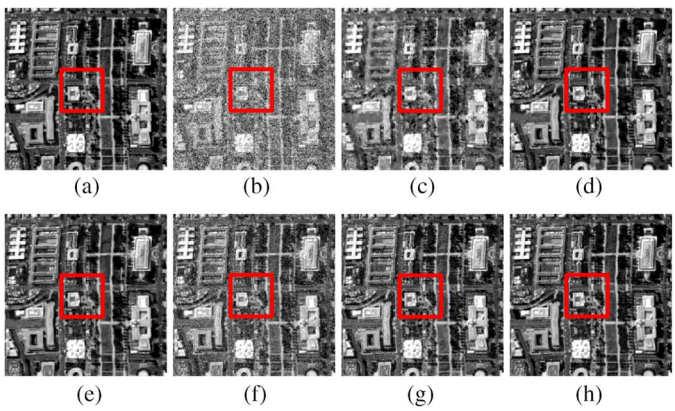


Fig. 6. Denoising results in the simulated experiments in Case 2: (a) original band 11 of the Washington DC Mall image; (b) noisy band; (c) SSAHTV; (d) VBM3D; (e) BM4D; (f) SURE-SVT; (g) LRMR; and (h) NAILRMA.

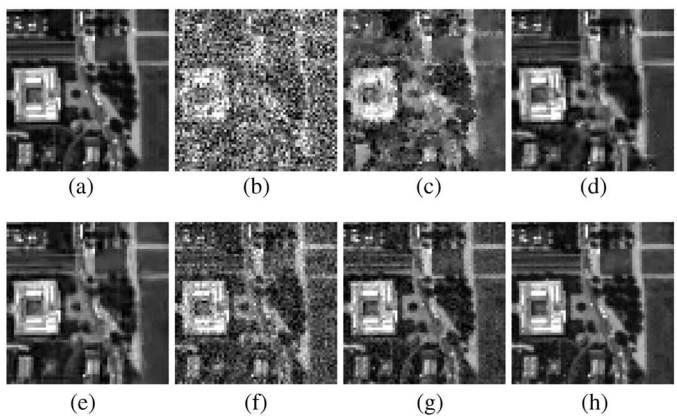


Fig. 7. Magnified results of Fig. 6: (a) original band 11 of the Washington DC Mall image; (b) noisy band; (c) SSAHTV; (d) VBM3D; (e) BM4D; (f) SURE-SVT; (g) LRMR; and (h) NAILRMA.

the values are manually adjusted in each experiment. For the VBM3D method, the noise variation is estimated via the technique introduced in Section II-D. BM4D is a parameter-free method, and for SURE-SVT, the patch size, step size, and noise variation estimation method are the same as NAILRMA. The upper bound rank r is also set the same in LRMR and the proposed method.

A. *Simulated Data Experiments for Dense Noise Removal*

Two HSI datasets, the Washington DC Mall dataset and the Pavia city center dataset, are used in the simulated experiments. The Washington DC Mall dataset (available: https://engineering.purdue.edu/~biehl/MultiSpec/hyperspectral.html) was collected by the Hyperspectral Digital Imagery

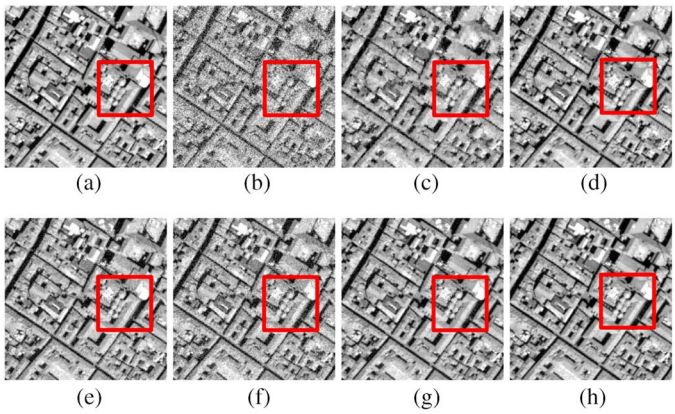


Fig. 8. Denoising results in the simulated experiments in Case 2: (a) original band 59 of the Pavia city center image; (b) noisy band; (c) SSAHTV; (d) VBM3D; (e) BM4D; (f) SURE-SVT; (g) LRMR; and (h) NAILRMA.

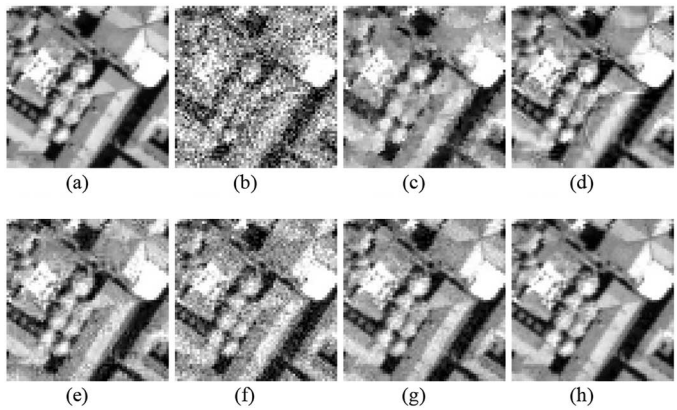


Fig. 9. Magnified results of Fig. 8: (a) original band 59 of the Pavia city center image; (b) noisy band; (c) SSAHTV; (d) VBM3D; (e) BM4D; (f) SURE-SVT; (g) LRMR; and (h) NAILRMA.

Collection Experiment (HYDICE), and the whole image contains 1208×307 pixels and 191 spectral bands. In the experiments, only a subimage of size $256 \times 256 \times 191$ is used, which is presented in Fig. 4. The Pavia city center dataset (http://www.ehu.es/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes) was collected by the Reflective Optics System Imaging Spectrometer (ROSIS-03). As some spectral bands of the Pavia city center dataset are heavily contaminated by noise, they cannot be used as the reference for denoising. Therefore, the first 22 bands (contain all the noisy bands) of this data are removed, and the size of subimage is selected as $200 \times 200 \times 80$, which is presented in Fig. 5. In the simulated experiments, the peak signal-to-noise ratio (PSNR) index and the structural similarity index (SSIM) [49] are adopted to give a quantitative assessment of the results of the simulated image experiments. For the HSI, we computed the PSNR and SSIM values between each noise-free band and denoised band, and then averaged them. These metrics are denoted as MPSNR (mean PSNR) and MSSIM (mean SSIM index).

For the simulated process, we first add noise as in the following two cases.

Case 1: Zero-mean Gaussian noise is added to all the bands of the Washington DC Mall and Pavia city center data. The noise variance values are 0.02, 0.04, 0.06, 0.08, and 0.1, respectively.

Case 2: Different variance zero-mean Gaussian noise is added to each band of the two HSI datasets. The variance values are randomly selected from 0 to 0.1, and the mean variance values are 0.053 and 0.051, respectively.

The selection of the regularization parameters λ of the SSAHTV method and r (r is estimated via the method presented in Section II-D) of the LRMR method, in both Cases 1 and 2, are shown in Table I. The denoised results of the different algorithms for the Washington DC Mall and Pavia city center images in Case 1 are presented in Tables II and III, respectively. The best results for each quality index are labeled in bold, and the second-best results for each quality index are underlined. From Tables II and III, it can be seen that the proposed method provides the highest values in both MPSNR and MSSIM, confirming the advantage of the proposed method over the other methods.

For Case 2, some typical bands of the HSI before and after denoising are presented to give the visual effect. Figs. 6 and 8 show band 11 of the Washington DC Mall image and band 59 of the Pavia city center image, before and after denoising, which are contaminated by Gaussian noise, with noise variance values of 0.088 and 0.092, respectively. Figs. 7 and 9 present magnified parts of the results in Figs. 6 and 8, respectively. By comparing the denoising results of the five methods, it can be clearly seen that the proposed method performs the best and effectively suppresses the Gaussian noise. As displayed in Figs. 6–9, we can also see that SSAHTV fails to restore the image and most of the details are lost. In fact, the SSAHTV method is more suitable for low-level noise removal. As the noise intensity grows, the performance of SSAHTV drops. VBM3D and BM4D can acquire relatively better results, but artifacts are introduced into the denoised results. SURE-SVT cannot effectively remove noise from the image. LRMR can obtain comparable results to NAILRMA, but the removal of Gaussian noise is not complete.

Table IV presents the MPSNR and MSSIM values of all the denoising methods for both datasets in Case 2, in which the proposed NAILRMA obtains the best denoising performance. From Tables II to IV, it is clear that the proposed method is more competitive in Case 2 (in which the different bands have different noise variances), because of the introduction of the adaptive iterative regularization. Figs. 10 and 11 show the PSNR and SSIM values of each band of the Washington DC Mall and Pavia city center images in the Case 2 experiments. The MPSNR and MSSIM values are higher than those of the other five methods in both experimental datasets, indicating that the NAILRMA method outperforms the other methods in HSI noise removal.

To further compare the performances of all the denoising algorithms, we calculated the mean spectral angle distance (MSAD) of all the spectral signatures between the noise-free and the denoised HSIs. The degree of MSAD is defined as

$$\text{MSAD} = \frac{1}{MN}\sum_{i=1}^{MN}\frac{180}{\pi} \times \arccos\frac{(\mathbf{X}^i)^T \cdot (\hat{\mathbf{X}}^i)}{||\mathbf{X}^i|| \cdot ||\hat{\mathbf{X}}^i||} \tag{14}$$

TABLE IV
QUANTITATIVE EVALUATION OF THE DIFFERENT DENOISING ALGORITHMS IN CASE II

Data	Evaluation index	SSAHTV	VBM3D	BM4D	SURE-SVT	LRMR	NAILRMA
Washington DC Mall	MPSNR(dB)	30.59	35.35	35.98	35.61	37.64	**42.05**
	MSSIM	0.8593	0.9459	0.9553	0.9405	0.9698	**0.9882**
Pavia city center	MPSNR(dB)	31.30	36.12	36.79	34.79	37.49	**41.40**
	MSSIM	0.8885	0.9576	0.9683	0.9381	0.9730	**0.9880**

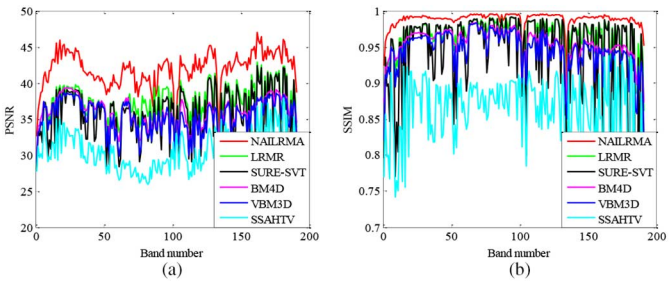


Fig. 10. PSNR and SSIM values of each band of the experimental results of the Washington DC Mall image in Case 2: (a) PSNR values and (b) SSIM values.

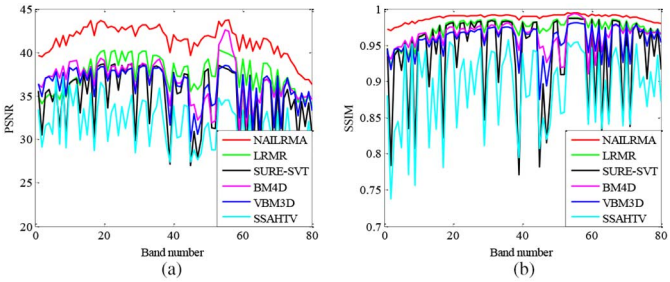


Fig. 11. PSNR and SSIM values of each band of the experimental results of the Pavia city center image in Case 2: (a) PSNR values and (b) SSIM values.

where $\mathbf{X}^i$ and $\hat{\mathbf{X}}^i$ are the ith spectral signatures of noise-free and the denoised HSIs, respectively. Table V gives the MSAD values in degree of all noisy and denoised HSIs. Here, it can be clearly seen that the proposed NAILRMA method produces better spectral signatures than the other denoising methods in both HSIs.

B. Simulated Data Experiment for Mixed-Noise Removal

For this simulated process, we add noise as follows.

Case 3: Dense Gaussian noise is first added to the Washington DC Mall data, as in Case 2. Impulse noise of 10% is then added to all the bands. Finally, dead lines are simulated for the four bands from band 60 to band 63.

In this case, we compare the denoised results of the proposed NAILRMR with LRMR. For a fair comparison, the upper bound rank r is set to 7, the cardinality b is set to 6000, and the step size is set as 8×8 in both the NAILRMR and LRMR methods.

Fig. 12 shows the false-color composite of the Washington DC Mall image in Case 3, before and after denoising, and Fig. 13 presents magnified results from Fig. 12. It is clear that both LRMR and NAILRMR can effectively remove sparse noise, and NAILRMR performs better in the dense noise removal. Fig. 14 shows the PSNR and SSIM values of each band of the Washington DC Mall image after LRMR and NAILRMR denoising, which indicates the good performance of the noise-adjusted iteration framework in dense noise removal.

C. Real Data Experiment

The Indian Pines dataset was acquired by the NASA AVIRIS (airborne visible/infrared imaging spectrometer) instrument over the Indian Pines test site in Northwestern Indiana in 1992. The data size is 145×145 pixels and 220 bands. The dataset was provided by Professor David Landgrebe and can be downloaded from https://engineering.purdue.edu/~biehl/MultiSpec/hyperspectral.html. Fig. 15 shows the color image by combining bands 3, 147, and 219, and the 10 land-cover classes used in the classification. In this real data experiment, the regularization parameter λ of the SSAHTV method is set to 3, and the noise variance of each band and the upper bound rank r are estimated via the methods introduced in Section II-D.

Fig. 16 presents the false-color image composed of bands 1, 103, and 220, before and after denoising, and Fig. 17 presents magnified results from Fig. 16. The Indian Pines image is mainly corrupted by dense noise, and sparse noise is also attached to some bands of the data. Here, it can be clearly observed that NAILRMR achieves the best results, effectively suppressing the noise and simultaneously keeping the local details of the original image. The result of NAILRMA is a little worse than NAILRMR. This is mainly because the NAILRMA method fails to remove the sparse noise which exists in some bands of the original Indian Pines data. LRMR can more or less obtain a comparable result to NAILRMA, but some dense noise still exists in the denoised result. For SSAHTV, the edges are enhanced, but the result is over-smooth and some details are lost. The VBM3D and BM4D methods fail to restore the image and introduce some artifacts. SURE-SVT can remove the noise and preserve the details, but not completely.

To further verify the effectiveness of the proposed denoising method, the classification results of the HSI before and after denoising are given for comparison purposes. Support vector machine (SVM) [51], with cross validation (CV) for selecting the tuning parameters, is utilized to conduct the supervised classification task for the HSI. The main idea of SVM is to project the data to a higher-dimensional space and to use a hyperplane to acquire a better separation. The classification process is repeated 10 times (training samples are chosen randomly 10 times), and the mean overall accuracy (OA) and kappa coefficient are chosen as the evaluation criteria.

Fig. 18 shows the classification results by the use of SVM on the whole Indian Pines dataset. The classification accuracy evaluation results by the use of the OA and kappa coefficient, along with the classification accuracies for each class,

TABLE V
MSAD VALUES OF THE NOISY AND DENOISED HSI DATA IN CASE II

Data	Noisy	SSAHTV	VBM3D	BM4D	SURE-SVT	LRMR	NAILRMA
Washington DC Mall	22.8116	8.5707	6.5860	6.1913	8.7799	6.3830	**3.8220**
Pavia city center	21.4757	9.3221	6.3710	6.4215	9.5098	6.1203	**3.4970**

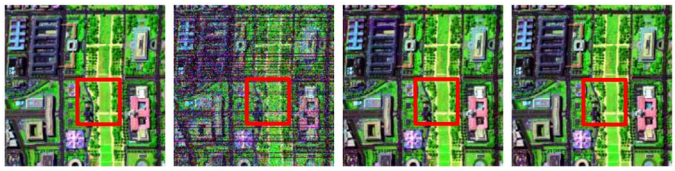


Fig. 12. Denoising results of the Washington DC Mall image in Case 3: (a) original false-color image (R: 180, G: 61, and B: 9); (b) noisy band; (c) LRMR; and (d) NAILRMR.

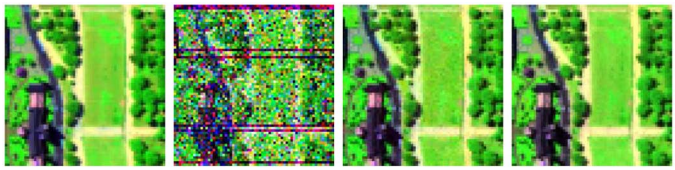


Fig. 13. Magnified results of Fig. 12: (a) original false-color image (R: 180, G: 61, and B: 9); (b) noisy band; (c) LRMR; and (d) NAILRMR.

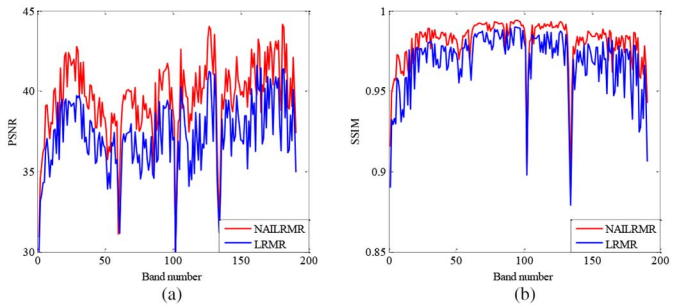


Fig. 14. PSNR and SSIM values of each band of the experimental results in Case 3: (a) PSNR values and (b) SSIM values.

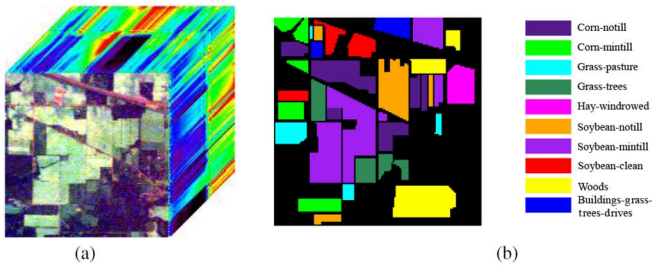


Fig. 15. AVIRIS Indian Pines dataset used in the real data experiment. (a) The HSI 3-D cube and (b) the training and test samples used in the classification.

are presented in Table VI. Here, it can be clearly seen that the classification result has been significantly improved after the denoising process. In the original noisy classification result, the classification appears fragmentary, because of the effect of the strong noise existing in most of the bands. However, in the denoised image classification results, the fragmentary

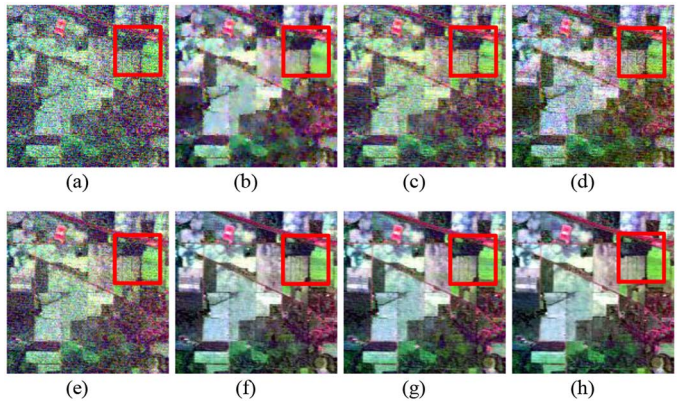


Fig. 16. Denoising results in the real data experiment: (a) original false-color image (R: 1, G: 103, and B: 220); (b) SSAHTV; (c) VBM3D; (d) BM4D; (e) SURE-SVT; (f) LRMR; (g) NAILRMA; and (h) NAILRMR.

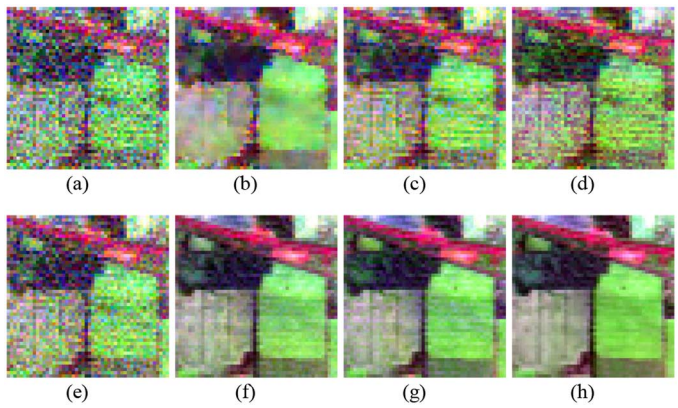


Fig. 17. Magnified results of Fig. 16: (a) original false-color image (R: 1, G: 103, and B: 220); (b) SSAHTV; (c) VBM3D; (d) BM4D; (e) SURE-SVT; (f) LRMR; (g) NAILRMA; and (h) NAILRMR.

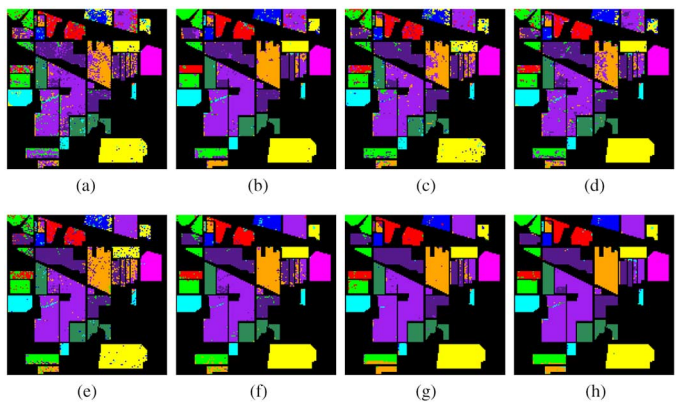


Fig. 18. Classification results of the Indian Pines data before and after denoising: (a) classification result of the original image; (b) classification result after SSAHTV denoising; (c) VBM3D; (d) BM4D; (e) SURE-SVT; (f) LRMR; (g) NAILRMA; and (h) NAILRMR.

TABLE VI
CLASSIFICATION ACCURACY EVALUATION RESULTS OF THE ORIGINAL AND DENOISED INDIAN PINES DATA

Class name	Train./ test	Origin	SSAHTV	VBM3D	BM4D	SURE-SVT	LRMR	NAILRMA	NAILRMR
Corn-notill	143/1285	0.7505	0.9128	0.8777	0.8587	0.8774	0.8875	0.9268	**0.9492**
Corn-mintill	83/747	0.61245	0.8436	0.7740	0.7689	0.8308	0.9229	**0.9428**	0.9426
Grass-pasture	48/435	0.8418	0.9549	0.8903	0.9434	0.9053	**0.9639**	0.9563	0.9632
Grass-trees	73/657	0.9154	0.9871	0.9335	0.975	0.9537	0.9941	0.9907	**0.9953**
Hay-windrowed	48/430	0.9923	0.9988	0.9949	0.9974	0.9960	0.9974	0.9981	**0.9995**
Soybean-notill	97/875	0.6398	0.8582	0.7584	0.7454	0.771429	0.8511	0.9285	**0.9381**
Soybean-mintill	246/2209	0.7531	0.9011	0.8762	0.8537	0.9019	0.9518	0.9699	**0.9820**
Soybean-clean	59/534	0.6112	0.8779	0.7629	0.8223	0.7903	0.9238	0.9507	**0.9624**
Woods	127/1138	0.9100	0.9639	0.9199	0.9645	0.9385	0.9960	0.9939	**0.9962**
Buildings-grass-trees-drives	39/347	0.5568	0.8470	0.6473	0.8346	0.6536	0.9735	0.9761	**0.9767**
OA	–	0.7618	0.9123	0.8562	0.8689	0.8757	0.9406	0.9615	**0.9707**
Kappa	–	0.7398	0.9009	0.8394	0.8536	0.8604	0.9323	0.9558	**0.9663**

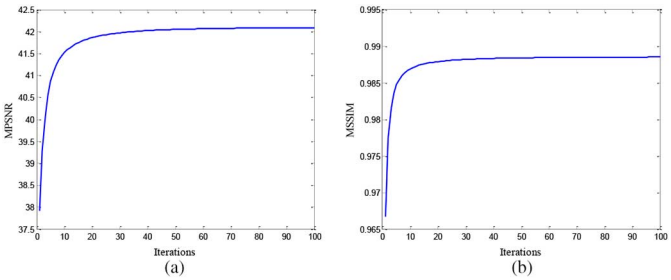


Fig. 19. MPSNR and MSSIM values of each step in the first 100 iterations in the simulated data experiment in Case 2 for the Washington DC Mall data: (a) MPSNR values of each step and (b) MSSIM values.

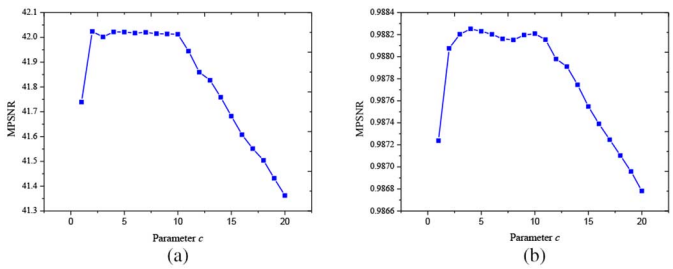


Fig. 20. Sensitivity analysis of parameter c (c from 1 to 20): (a) change in the MPSNR value and (b) change in the MSSIM value.

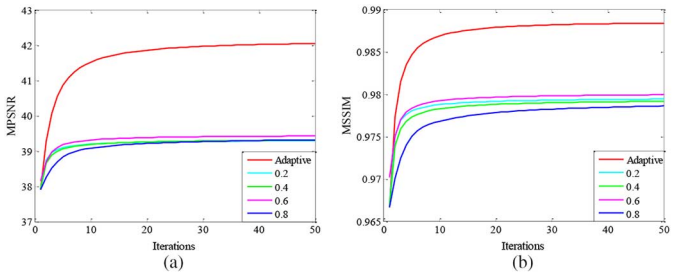


Fig. 21. Comparison of the noise-adjusted update mode (5) and update mode (4) with δ varied between 0.2, 0.4, 0.6, and 0.8: (a) MPSNR values of each iteration in Case 2 for the Washington DC Mall data and (b) MSSIM values.

phenomenon is greatly reduced. Among all the classification results of the seven denoising methods, NAILRMR achieves the highest OA and kappa coefficient values, indicating the best performance in the noise removal. For SSAHTV, reducing the value of parameter λ can improve the classification result, but this is accompanied with a smoother denoised image, and more details are lost.

D. Discussions

As the major concern of this paper is centralized on the performance improvement of dense noise removal, we only discuss and analysis the NAILRMA method here. However, NAILRMR has the same property as NAILRMA, except for the sparse noise removal.

1) Convergence of NAILRMA: To verify the convergence of the iteration, we give the MPSNR and MSSIM values of each iterative step in the first 100 iterations in the simulated experiment in Case 2 for the Washington DC Mall data in Fig. 19. From Fig. 19, we can see that as the iteration progresses, the MPSNR and MSSIM values of the output image monotonically increase, indicating the performance improvement of NAILRMA. In Fig. 19, we can also see that the MPSNR and MSSIM values increase rapidly in the first few iterations, and then grow slowly in the subsequent iterations. Therefore, in all of the simulated and real data experiments, the stopping criterion is set as $\varepsilon = 1e-3$, and the maximum iteration number is 50.

2) Sensitivity Analysis of Parameter c: In (5) and (6), the selection of the iteration factor δ is related to the noise variance of each band and the decay parameter c. To show the effect of parameter c on the final denoising performance, we undertake a sensitivity analysis in the simulated experimental results in Case 2 for the Washington DC Mall data, as presented in Fig. 20. To confirm the superiority of the noise-adjusted iterative update method, as introduced in (5), we also undertake a comparison of NAILRMA with (5) and (4) as the update mode.

Fig. 20 presents the MPSNR and MSSIM values with the change of parameter c from 1 to 20. When c falls within the range of 2–10, the NAILRMA method can acquire stable MPSNR and MSSIM values. Therefore, we set c as 5 in all the experiments. Fig. 21 gives the output MPSNR and MSSIM values of each iteration in NAILRMA with the two different

update modes, as in (5) and (4). For update mode (4), the iteration factor δ is the same for all the HSI bands, and we choose δ as 0.2, 0.4, 0.6, and 0.8 for comparison. From Fig. 21, we can see that the gaps in the output MPSNR and MSSIM values are small after 50 iterations when δ is varied between 0.2, 0.4, 0.6, and 0.8. Overall, it can be clearly observed that the noise-adjusted iteration update mode has a clear advantage over update mode (4).

V. Conclusion

In this paper, we have proposed an NAILRMA method for HSI denoising. Patchwise randomized singular value decomposition is first applied to each Casorati matrix (obtained by lexicographically ordering the 3-D patch into a 2-D matrix) to denoise the HSI. An iterative regularization technique is subsequently adopted, based on the patchwise LRMA, to further separate the signal and noise. As for the HSI, different bands have different noise levels. A noise-adjusted update mode is then proposed to update the input image of the iteration. That is, in each iteration step, the input image is composed of the previous iteration of the input and output images, and the input image of the previous iteration accounts for the main part in the low-noise case, and vice versa. As the iteration progresses, the noise in each band is suppressed. Meanwhile, the image structures are progressively recovered until convergence. This noise-adjusted iteration framework is also extended to mixed-noise removal. Simulated and real HSI experiment results confirm the effectiveness of the proposed method.

Acknowledgment

The authors would like to thank Prof. D. Landgrebe at Purdue University for providing the free downloads of the HYDICE image of the Washington DC Mall, and Prof. D. Tao and T. Zhou for the GoDec algorithm code.

References

[1] P. Zhong and R. Wang, "Multiple-spectral-band CRFs for denoising junk bands of hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 4, pp. 2260–2275, Apr. 2013.

[2] Y. Qian and M. Ye, "Hyperspectral imagery restoration using nonlocal spectral-spatial structured sparse representation with noise estimation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 499–515, Apr. 2013.

[3] J. Li, H. Zhang, Y. Huang, and L. Zhang, "Hyperspectral image classification by nonlocal joint collaborative representation with a locally adaptive dictionary," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3707–3719, Jun. 2014.

[4] A. Plaza, P. Martinez, J. Plaza, and R. Perez, "Dimensionality reduction and classification of hyperspectral image data using sequences of extended morphological transformations," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 2, pp. 466–479, Mar. 2005.

[5] H. Zhang, J. Li, Y. Huang, and L. Zhang, "A nonlocal weighted joint sparse representation classification method for hyperspectral imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2056–2065, Jun. 2014.

[6] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Semisupervised hyperspectral image segmentation using multinomial logistic regression with active learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 11, pp. 4085–4098, Nov. 2010.

[7] J. M. Bioucas-Dias *et al.*, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.

[8] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2014–2039, Jun. 2011.

[9] D. W. Stein *et al.*, "Anomaly detection from hyperspectral imagery," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 58–69, Jan. 2002.

[10] L. Zhang, L. Zhang, D. Tao, and X. Huang, "Sparse transfer manifold embedding for hyperspectral target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 2, pp. 1030–1043, Feb. 2014.

[11] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR'05)*, 2005, pp. 60–65.

[12] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, Dec. 2006.

[13] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[14] A. A. Green, M. Berman, P. Switzer, and M. D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Trans. Geosci. Remote Sens.*, vol. 26, no. 1, pp. 65–74, Jan. 1988.

[15] C.-I. Chang and Q. Du, "Interference and noise-adjusted principal components analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 5, pp. 2387–2396, Sep. 1999.

[16] G. Chen and S. Qian, "Denoising of hyperspectral imagery using principal component analysis and wavelet shrinkage," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 3, pp. 973–980, Mar. 2011.

[17] F. Bollenbeck, A. Backhaus, and U. Seiffert, "A multivariate wavelet-PCA denoising-filter for hyperspectral images," in *Proc. 3rd Workshop Hyperspectral Image Signal Process. Evol. Remote Sens. (WHISPERS)*, 2011, pp. 1–4.

[18] H. Zhang, W. He, L. Zhang, H. Shen, and Q. Yuan, "Hyperspectral image restoration using low-rank matrix recovery," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 4729–4743, Aug. 2014.

[19] D. Letexier and S. Bourennane, "Noise removal from hyperspectral images by multidimensional filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 7, pp. 2061–2069, Jul. 2008.

[20] X. Liu, S. Bourennane, and C. Fossati, "Nonwhite noise reduction in hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 3, pp. 368–372, May 2012.

[21] X. Liu, S. Bourennane, and C. Fossati, "Denoising of hyperspectral images using the PARAFAC model and statistical performance analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 10, pp. 3717–3724, Oct. 2012.

[22] K. Dabov, A. Foi, and K. Egiazarian, "Video denoising by sparse 3D transform-domain collaborative filtering," presented at the European Signal Processing Conf., Poznan, Poland, Sep. 2007, pp. 145–149.

[23] M. Maggioni, V. Katkovnik, K. Egiazarian, and A. Foi, "Nonlocal transform-domain filter for volumetric data denoising and reconstruction," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 119–133, Jan. 2013.

[24] J. Martín-Herrero, "Anisotropic diffusion in the hypercube," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1386–1398, May 2007.

[25] H. Zhang, "Hyperspectral image denoising with cubic total variation model," *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. I-7, pp. 95–98, 2012.

[26] Q. Yuan, L. Zhang, and H. Shen, "Hyperspectral image denoising employing a spectral-spatial adaptive total variation model," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 10, pp. 3660–3677, Oct. 2012.

[27] Q. Yuan, L. Zhang, and H. Shen, "Hyperspectral image denoising with a spatial-spectral view fusion strategy," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2314–2325, May 2014.

[28] H. Othman and S. Qian, "Noise reduction of hyperspectral imagery using hybrid spatial-spectral derivative-domain wavelet shrinkage," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 2, pp. 397–408, Feb. 2006.

[29] T. Lin and S. Bourennane, "Hyperspectral image processing by jointly filtering wavelet component tensor," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 6, pp. 3529–3541, Jun. 2013.

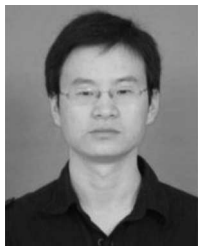
[30] B. Rasti, J. R. Sveinsson, M. O. Ulfarsson, and J. A. Benediktsson, "Hyperspectral image denoising using first order spectral roughness penalty in wavelet domain," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2458–2467, Jun. 2014.

[31] G. Ely, S. Aeron, and E. L. Miller, "Exploiting structural complexity for robust and rapid hyperspectral imaging," arXiv preprint arXiv:1305.2170, 2013.

[32] B. Rasti, J. R. Sveinsson, and M. O. Ulfarsson, "Wavelet-based sparse reduced-rank regression for hyperspectral image restoration," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 10, pp. 6688–6698, Oct. 2014.
[33] T. Hu, H. Zhang, H. Shen, and L. Zhang, "Robust registration by rank minimization for multiangle hyper/multispectral remotely sensed imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2443–2457, Jun. 2014.
[34] N. Halko, P.-G. Martinsson, and J. A. Tropp, "Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions," *SIAM Rev.*, vol. 53, no. 2, pp. 217–288, 2011.
[35] M. Fazel, "Matrix rank minimization with applications," Ph.D. dissertation, Elec. Eng. Dept, Stanford Univ., Stanford, CA, USA, Mar. 2002.
[36] I. Jolliffe, *Principal Component Analysis*. Hoboken, NJ, USA: Wiley, 2005.
[37] E. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *J. ACM*, vol. 58, no. 3, pp. 1–37, May 2011.
[38] J. Cai, E. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956–1982, Jan. 2010.
[39] E. Candes, C. A. Sing-Long, and J. D. Trzasko, "Unbiased risk estimates for singular value thresholding and spectral estimators," arXiv preprint arXiv:1210.4139, 2012.
[40] R. Witten and E. Candès, "Randomized algorithms for low-rank factorizations: Sharp performance bounds," *Algorithmica*, pp. 1–18, 2013.
[41] T. Zhou and D. Tao, "Bilateral random projections," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2012, pp. 1286–1290.
[42] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, Mar. 2009.
[43] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, "An iterative regularization method for total variation-based image restoration," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 460–489, Jun. 2005.
[44] T. Zhou and D. Tao, "Godec: Randomized low-rank & sparse matrix decomposition in noisy case," in *Proc. 28th Int. Conf. Mach. Learn. (ICML)*, 2011, pp. 33–40.
[45] W. Dong, G. Shi, and X. Li, "Nonlocal image restoration with bilateral variance estimation: A low-rank approach," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 700–711, Feb. 2013.
[46] L. Gao, Q. Du, B. Zhang, W. Yang, and Y. Wu, "A comparative study on linear regression-based noise estimation for hyperspectral imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 488–498, Feb. 2013.
[47] J. M. Bioucas-Dias and J. M. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2435–2445, Aug. 2008.
[48] N. Dobigeon, S. Moussaoui, M. Coulon, J.-Y. Tourneret, and A. O. Hero, "Joint Bayesian endmember extraction and linear unmixing for hyperspectral imagery," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4355–4368, Nov. 2009.
[49] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[50] M. Guo, H. Zhang, J. Li, L. Zhang, and H. Shen, "An online coupled dictionary learning approach for remote sensing image fusion," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1284–1294, Apr. 2014.
[51] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778–1790, Aug. 2004.
[52] Q. Li, H. Li, Z. Lu, Q. Lu, and W. Li, "Denoising of hyperspectral images employing two-phase matrix decomposition," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 9, pp. 3742–3754, Sep. 2014.
[53] W. He, H. Zhang, L. Zhang, and H. Shen, "A noise-adjusted iterative randomized singular value decomposition method for hyperspectral image denoising," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Quebec, Canada, 2014, pp. 1536–1539.
[54] Q. Wang, L. Zhang, Q. Tong, and F. Zhang, "Hyperspectral imagery denoising based on oblique subspace projection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 2468–2480, Jun. 2014.

Wei He (S'14) received the B.S. degree in mathematics and statistics from Wuhan University, Wuhan, China, in 2012. He is currently pursuing the Ph.D. degree at the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing (LIESMARS), Wuhan University.

His research interests include image quality improvement, remote sensing image processing, and low-rank representation.



Hongyan Zhang (M'13) received the B.S. degree in geographic information system and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2005 and 2010, respectively.

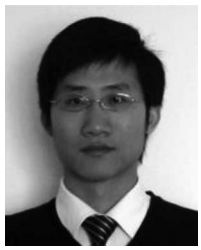
He is currently an Associate Professor with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University. His research interests include image reconstruction, hyperspectral image processing, sparse representation, and low-rank methods for sensing image imagery.

Dr. Zhang is a Reviewer of about 10 international academic journals, and has published more than 30 research papers.

Liangpei Zhang (M'06–SM'08) received the B.S. degree in physics from Hunan Normal University, Changsha, China, the M.S. degree in optics from Chinese Academy of Sciences, Xian, China, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1982, 1988, and 1998, respectively.

He is currently the Head of the Remote Sensing Division, State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University. He is also a Chang-Jiang Scholar Chair Professor appointed by the Ministry of Education of China. He is currently a Principal Scientist with China State Key Basic Research Project (2011–2016) appointed by the Ministry of National Science and Technology of China to lead the remote sensing program in China. He has more than 360 research papers. He is the holder of 15 patents. His research interests include hyperspectral remote sensing, high-resolution remote sensing, image processing, and artificial intelligence.

Dr. Zhang is a Fellow of the Institution of Engineering and Technology, an Executive Member (Board of Governor) of the China National Committee of the International Geosphere–Biosphere Programme, and an Executive Member of the China Society of Image and Graphics. He regularly serves as a Co-Chair of the series SPIE Conferences on Multispectral Image Processing and Pattern Recognition, Conference on Asia Remote Sensing, and many other conferences. He has authored several conference proceedings, issues, and geoinformatics symposiums. He also serves as an Associate Editor of the *International Journal of Ambient Computing and Intelligence*, *International Journal of Image and Graphics*, *International Journal of Digital Multimedia Broadcasting*, *Journal of Geo-spatial Information Science*, *Journal of Remote Sensing*, and IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.



Huanfeng Shen (M'11–SM'13) received the B.S. degree in surveying and mapping engineering, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2002 and 2007, respectively.

In July 2007, he joined the School of Resource and Environmental Science, Wuhan University, where he is currently a Full Professor. He has authored more than 60 research papers. His research interests include image processing (for quality improvement), remote sensing application, data fusion, and assimilation.

Dr. Shen has been supported by several talent programs, including the New Century Excellent Talents by the Ministry of Education of China (2011) and the Hubei Science Fund for Distinguished Young Scholars (2011).