Hyperspectral Image Restoration Using Low-Rank Matrix Recovery

Hongyan Zhang, Member, IEEE, Wei He, Liangpei Zhang, Senior Member, IEEE, Huanfeng Shen, Senior Member, IEEE, and Qiangqiang Yuan

Abstract—Hyperspectral images (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. This paper introduces a new HSI restoration method based on low-rank matrix recovery (LRMR), which can simultaneously remove the Gaussian noise, impulse noise, dead lines, and stripes. By lexicographically ordering a patch of the HSI into a 2-D matrix, the low-rank property of the hyperspectral imagery is explored, which suggests that a clean HSI patch can be regarded as a low-rank matrix. We then formulate the HSI restoration problem into an LRMR framework. To further remove the mixed noise, the “Go Decomposition” algorithm is applied to solve the LRMR problem. Several experiments were conducted in both simulated and real data conditions to verify the performance of the proposed LRMR-based HSI restoration method.

Index Terms—Go Decomposition (GoDec), hyperspectral image (HSIs), low rank, restoration.

I. INTRODUCTION

While hyperspectral imaging sensors have experienced significant success, hyperspectral images (HSIs) collected in practice often suffer from various annoying degradations, e.g., noise contamination, stripe corruption, and missing data, due to the sensor, photon effects, and calibration error [1], [2]. As a result, not only the visual appearance but also the applications of these images, which include urban planning, mapping, agriculture, forestry, and so on, are severely influenced. Therefore, as a preprocessing of the image application, HSI restoration is an active and challenging research area.

To date, many different denoising methods have been proposed for the restoration of HSIs. Traditional denoising methods treat each band of the HSI as a gray-level image and restore them separately. More advanced denoising techniques mainly use the spectral information of the HSI and have achieved good results. For example, Letexier and Bourennane [3] considered the HSI data set as a third-order tensor and introduced a generalized multidimensional Wiener filter to denoise the HSIs. Liu et al. [1] developed a powerful multilinear algebra model named parallel factor analysis (PARAFAC) to denoise HSIs, and it has been shown that the PARAFAC model is an efficient denoising method for the removal of Gaussian noise from HSIs.

Recently, researchers have paid most of their attention to the denoising methods that use the spatial and the spectral information together, and these approaches have made significant progress. Othman and Qian [4] proposed a hybrid spatial–spectral derivative-domain wavelet shrinkage model, which benefits from the dissimilarity of the signal regularity in the spatial and spectral dimensions of the HSIs. Karami et al. [5] evaluated a new method for the noise reduction of HSIs, based on a genetic algorithm and kernel hierarchical nonnegative Tucker decomposition. This method also exploits both the spectral and the spatial information in the images. In [6], a multispectral-band conditional random field model was proposed by Zhong and Wang to simultaneously model and use the spatial and spectral dependences in a unified probabilistic framework.

Based on a maximum a posteriori framework, Chen et al. [7] introduced a spatial–spectral domain mixing prior model which takes advantage of the different properties of HSIs in the spatial and spectral domains, and Yuan et al. [8] proposed an HSI denoising algorithm employing a spectral–spatial adaptive total variation (SSAHTV) model. In the SSAHTV algorithm, the spectral noise differences and spatial information differences are both considered in the process of the noise reduction. In [9], the nonlocal similarity and spectral–spatial structure of the hyperspectral imagery were introduced into the sparse representation framework. In addition, principal component analysis, wavelet shrinkage, and anisotropic diffusion have also been adopted for HSI noise removal [10]–[13].

To sum up, most of these restoration methods are based on some specific prior knowledge of the image noise, and they use the spectral or spatial information of the HSIs to recover the current pixel or pixels (image patch). Under the limitation of the prior knowledge, most of the methods mentioned earlier can only remove one or two kinds of noise. However, for real-world HSIs, there usually exists a combination of several different types of noise, e.g., Gaussian noise, impulse noise, dead pixels
or lines, and stripes. To the best of our knowledge, few of the restoration methods mentioned earlier can simultaneously remove all these types of noise.

In recent years, low-rank matrix approximation has emerged as a powerful tool for image analysis, web search, and computer vision [14]–[17]. It describes the problem of finding and exploiting low-dimensional structures in high-dimensional data.Unlike the traditional image restoration methods, the prior knowledge in low-rank matrix approximation-based image restoration is that some components of the clean image are regarded as low rank, and the aim is to remove the various noise types in the contaminated image. Ji et al. [18], [19] introduced this theory to video restoration by formulating the video restoration problem as a low-rank matrix completion/recovery problem. For each reference patch of the selected video frame, a patch matching algorithm is used to search the similar patches in the spatiotemporal domain. These similar patches are lexicographically formulated into a matrix, which can then be decomposed into a low-rank matrix term denoting the clean patches and a sparse matrix term denoting outliers. Li et al. [20] adopted a similar patch matching algorithm for natural images to determine the low-rank parts of a clean image and used spatially adaptive iterative singular value thresholding to recover the clean parts from the noisy image. As to HSI restoration, the main challenge is the determination of the low-rank components and the formation of the low-rank matrix approximation-based image restoration model.

In this paper, we propose an HSI restoration technique on the basis of low-rank matrix recovery (LRMR), which can simultaneously remove Gaussian noise, impulse noise, dead pixels or lines, and stripes. The main contributions of this paper are the exploration of the HSI low-rank property and the application of LRMR to the HSI restoration process. The main ideas of this work can be summarized as follows.

1) HSI data are rearranged as a 2-D matrix, and the low-rank property of the image matrix is explored. Based on the low-rank prior knowledge of the clean HSI and the sparsity property of the non-Gaussian noise [18], [19] (including impulse noise, dead lines, and stripes), we build our HSI restoration model based on LRMR.

2) Both the “Go Decomposition” (GoDec) [21] and the augmented Lagrange multiplier (ALM) numerical optimization algorithms are utilized to solve the proposed restoration model, and the experimental evaluations of the two methods for restoring HSI are provided.

3) The proposed LRMR method provides a new perspective for HSI restoration. The experimental results indicate that it outperforms the other state-of-the-art methods, not only in the case of Gaussian noise only but also in the case of a mixture of several types of noise.

The rest of this paper is organized as follows. In Section II, the LRMR model and the low-rank property of the HSI are introduced. After that, the HSI restoration method based on LRMR is presented in Section III. In Section IV, both the simulated experiment and real data experiments are described and analyzed, followed by the conclusions in Section V.

II. LOW-RANK PROPERTY OF HYPERSONTIAL IMAGERY

A. LRMR Model

The LRMR model was first proposed by Wright et al. [17] and is idealized as a “robust principal component analysis” (RPCA) problem. Assuming that a low-rank matrix \( L \in \mathbb{R}^{m \times n} \) is corrupted by a sparse error matrix \( S \in \mathbb{R}^{m \times n} \), then the observed data matrix \( D \in \mathbb{R}^{m \times n} \) can be decomposed as the sum of a sparse matrix and a low-rank matrix, i.e., \( D = L + S \).

The ideal RPCA problem can be described as follows: Given the observed data matrix \( D \), the low-rank matrix \( L \) and the sparse error matrix \( S \) are unknown, and the goal is to recover \( L \). The formulation of this optimization problem is [17]

\[
\min_{L,S} \quad \text{rank}(L) + \|S\|_0 \quad s.t \quad D = L + S. \tag{1}
\]

Unfortunately, (1) is a highly nonconvex optimization problem, and no efficient solution is known. We can obtain a tractable optimization problem by relaxing (1) and replacing the \( \ell_0 \)-norm with the \( \ell_1 \)-norm and the rank with the nuclear norm [22]–[24], yielding the following convex surrogate:

\[
\min_{L,S} \quad \|L\|_*, + \lambda\|S\|_1 \quad s.t \quad D = L + S \tag{2}
\]

where \( \lambda \) is the regularization parameter used to balance the relative contribution between the nuclear norm and the \( \ell_1 \)-norm. Candes et al. [14] have proven that, when the rank of the matrix \( L \) and the sparsity and distribution of \( S \) obey certain conditions, there is a high probability of recovering the low-rank matrix \( L \) and sparse matrix \( S \).

On the basis of Candes and Plan’s work in [25], Zhou et al. [26] improved the model of RPCA and studied the problem of recovering the low-rank matrix (the principal components) from a high-dimensional data matrix corrupted by both small entrywise noise and gross sparse errors. The revised measurement model assumes that \( D = L + S + N \), where \( D \), \( L \), and \( S \) are as introduced earlier and \( N \) is the noise term, i.e., independent and identically distributed Gaussian noise on each entry of the matrix. The optimization problem of this model is introduced as

\[
\min_{L,S} \quad \|L\|_* + \lambda\|S\|_1 \quad s.t \quad \|D - L - S\|_F \leq \delta \tag{3}
\]

where \( \delta \) is a constant related to the standard deviation of random noise \( N \). Zhou and Tao [21] proposed an equivalent formulation of (3)

\[
\min_{L,S} \quad \|X - L - S\|_F^2 \quad s.t \quad \text{rank}(L) \leq r, \text{card}(S) \leq k \tag{4}
\]

where \( r \) and \( k \), which stand for the upper bound of the rank of \( L \) and the cardinality of \( S \), are set to be the known information. In this paper, a revised model which considers Gaussian noise is utilized.

B. LOW-RANK PROPERTY OF HYPERSONTIAL IMAGERY

As described previously, the baseline of the low-rank-based restoration task is the exploration of the low-rank components.
Differing from the patch matching algorithm used in natural image/video restoration [18]–[20], which exploits the redundancy in the spatial and temporal domains, the high correlation of the spectral information is adopted to formulate the low-rank components here. Fig. 1 shows the matrix representation of a clean HSI patch. Assuming that we have an original HSI cube $u$, denoted by (a) in Fig. 1, the size of $u$ is $M \times N \times B$, in which $M$ represents the width of the image, $N$ stands for the height of the image, and $B$ is the number of spectral bands. Select a subcube $p_{i,j}$ of size $q \times q \times B$ centered at the spatial position $(i, j)$. If we represent the $k$th band of the subcube $p_{i,j}$ as a vector $p_{i,j,k} \in \mathbb{R}^{q^2}$ by lexicographically ordering all its columns, we obtain a 2-D matrix $P_{i,j}$ of size $q^2 \times B$ as follows:

$$ P_{i,j} = (p_{i,j,1}, p_{i,j,2}, \ldots, p_{i,j,B}). \quad (5) $$

Next, we investigate the low-rank property of $P_{i,j}$.

The low-rank property of an HSI can be exploited from the perspective of a linear spectral mixing model [27], [28]. As known, there exist high correlations among the spectral signatures (rows of $P_{i,j}$), as each spectral signature can be represented by a linear combination of a small number of pure spectral endmembers, which is known as the linear spectral mixing model. Supposing the upper bound of the number of pure spectral endmembers for the HSI patch is $r$, then $P_{i,j}$ can be decomposed as $P_{i,j} = MH^T$, where $H \in \mathbb{R}^{B \times r}$ is the matrix whose columns contain pure spectral endmembers and $M \in \mathbb{R}^{q^2 \times r}$ is the abundance matrix. As the upper bound value of the number of endmembers $r$ is usually relatively small, then the rank of $P_{i,j}$ is bounded, i.e., $\text{rank}(P_{i,j}) \leq r$, which suggests the low-rank property of the matrix $P_{i,j}$.

III. LRMR-BASED HSI RESTORATION METHOD

A. HSI Degradation Model

We assume that the HSI is corrupted by four kinds of noise: Gaussian noise, impulse noise, dead pixels or lines, and stripes. The noise degradation model of the HSI can be written as

$$ d = u + s + n \quad (6) $$

where $u$ is the clean HSI, as shown in Fig. 1(a); $d$ is the observed degradation image; $s$ denotes the mixture of impulse noise, dead pixels or lines, and stripes; and $n$ stands for the Gaussian noise. In this degradation model, the images $d$, $u$, $s$, and $n$ are of the same size of $M \times N \times B$.

B. LRMR-Based HSI Restoration Model

In this paper, we divide the HSI into patches and restore each patch sequentially, which can effectively preserve the local details [29], [30]. The size of each patch is $q \times q \times B$. For a patch centered at the pixel $(i, j)$, the image degradation model is expressed as follows:

$$ d_{i,j} = u_{i,j} + s_{i,j} + n_{i,j} \quad (7) $$

where $d_{i,j}$, $u_{i,j}$, $s_{i,j}$, and $n_{i,j}$ are the 3-D subcubes of $d$, $u$, $s$, and $n$, respectively. Rearrange $d_{i,j}$, $u_{i,j}$, $s_{i,j}$, and $n_{i,j}$ into 2-D matrices according to the process shown in Fig. 1, and denote them as $D_{i,j}$, $L_{i,j}$, $S_{i,j}$, and $N_{i,j}$, respectively. The corresponding image degradation model is

$$ D_{i,j} = L_{i,j} + S_{i,j} + N_{i,j} \quad (8) $$

where the matrix $L_{i,j}$ represents the clean HSI patch, the matrix $S_{i,j}$ is denoted as the mixture of impulse noise, dead lines, and stripes, and the noise matrix $N_{i,j}$ is the representation of the Gaussian noise. With the low-rank property of the clean HSI, the matrix $L_{i,j}$ is low rank. Clearly, as the impulse noise, dead pixels or lines, and stripes only corrupt some parts or some bands of the HSI, therefore, the matrix $S_{i,j}$ is sparse. The Gaussian noise matrix $N_{i,j}$ is of a small level and impacts all the pixels.

With the similarity between the HSI degradation model in (8) and the LRMR measurement model, it is natural for us to apply sparse and low-rank matrix decomposition to the noisy matrix $D_{i,j}$ and obtain the clean HSI patch $L_{i,j}$. Thus, the restoration process for the patch centered at $(i, j)$ of the HSI is accomplished. After processing all the patches of the HSI, we synthesize the restored image from these clean patches. In this paper, the image patches are sampled with overlapping regions, and most of the pixels are covered by several restored patches. The value of each pixel in the HSI is then determined by taking the average of the restored patches at this pixel, which will suppress the possible artifacts in the neighborhood of the boundaries of patches.

C. GoDec Algorithm

In recent years, many LRMR optimization algorithms have been proposed. For example, Lin et al. [31] developed and compared two complementary approaches for solving model (2). The first is an accelerated proximal gradient algorithm directly applied to the primal, and the second is a gradient algorithm applied to the dual problem. Yuan and Yang [32] proposed the alternating direction method for solving problem (2), by taking full exploitation to the high-level separable structure of the convex relaxation problem. Recently, Lin et al. [33] applied the method of ALM to solve this convex program. In [26], the authors proved that the methods mentioned earlier also fit the revised model (3).

In this paper, we adopt the GoDec algorithm proposed by Zhou and Tao in [21] to solve the low-rank HSI patch, which can remove the Gaussian noise more effectively. Before solving the optimization problem (4) by the GoDec approach, the upper
bounds of $\text{rank}(L)$ and $\text{card}(S)$, i.e., the values of $r$ and $k$, should be determined. The optimization problem shown in (4) can be converted to solve the following two subproblems alternately until convergence:

$$\begin{align*}
L_t &= \arg\min_{\text{rank}(L) \leq r} \|X - L - S_{t-1}\|_F^2 \\
S_t &= \arg\min_{\text{card}(S) \leq r} \|X - L_t - S\|_F^2.
\end{align*}$$

(9)

(10)

Bilateral random projection (BRP) [34] based low-rank approximation is applied to solve the subproblem (9), as it is near optimal and efficient. Suppose that $X$ is a matrix of size $q^2 \times B$, we have $Y_1 = XA_1$ and $Y_2 = X^TA_2$, wherein $A_1 \in \mathbb{R}^{B \times r}$ and $A_2 \in \mathbb{R}^{q^2 \times r}$ are random matrices, and a fast rank-$r$ BRP approximation of $X$ is denoted by

$$L = Y_1 (A_2^TY_1)^{-1} Y_2^T.$$  

(11)

As for the optimization subproblem (10), $S_t$ is updated via entrywise hard thresholding of $X - L_t$, that is to say

$$S_t = \mathbb{P}_{\Omega}(X - L_t), \Omega : \left|\left(\begin{array}{c}
X - L_t
\end{array}\right)_{i,j}\right| \neq 0$$

and $\geq \left|\left(\begin{array}{c}
X - L_t
\end{array}\right)_{i,j}\right|$, $|\Omega| \leq k$.  

(12)

The GoDec algorithm [21] described earlier is summarized as Algorithm 1.

---

**Algorithm 1**: GoDec Algorithm

**Input**: $X$, $r$, $k$, $\varepsilon$  
**Output**: $L$, $S$

**Initialize**: $L_0 := X$, $S_0 := 0$, $t := 0$

**While** $\|X - L_t - S_t\|_F^2 / \|X\|_F^2 > \varepsilon$

$t := t + 1$

$Y_1 = (X - S_{t-1})A_1$, $A_2 = Y_1$, $Y_2 = (X - S_{t-1})^TA_2$

If $\text{rank}(A_2^TY_1) < r$

then $r := \text{rank}(A_2^TY_1)$, regenerate the matrix $A_1$, and go to the next iteration.

**End if**

$L_t = Y_1 (A_2^TY_1)^{-1} Y_2^T$;  

$S_t = \mathbb{P}_{\Omega}(X - L_t)$, $\Omega$ is the nonzero subset of the first $k$ largest entries of $|X - L_t|$;  

**End while**

---

IV. EXPERIMENTAL RESULTS AND DISCUSSION

To demonstrate the effectiveness of our restoration method for hyperspectral imagery, we perform both simulated and real data experiments, and evaluate the experimental results quantitatively and visually.

A. Simulated Data Experiment

The Hyperspectral Digital Imagery Collection Experiment (HYDICE) image of the Washington DC Mall is used in our simulated experiment. The whole image contains $1208 \times 307$ pixels and 191 spectral channels. Due to the page limitation, a subimage of size $256 \times 256 \times 191$ is used for our experiment, which is presented in Fig. 2. Before the simulated process, the gray values of each band of the HSI are normalized between [0, 1]. After the restoration process, the gray values of each band are stretched to the original level. In our simulated experiment, four kinds of noise are added to the Washington DC image.

1) Zero-mean Gaussian noise is added to all the bands of the HSI. For different bands, the noise intensity is different. The SNR value of each band varies from 10 to 20 dB randomly, and the mean SNR value of all the bands is 14.95 dB.

2) Impulse noise is added to the 11 selected bands from band 20 to band 30. The percentage of impulse noise is 20%.

3) Dead lines are simulated for the four bands from band 70 to band 73. The width of the dead lines is from one line to three lines.

4) Stripes are simulated for the four bands from band 111 to band 114. The width of the stripes is from one line to three lines.

The parameters are set as follows: $q = 20$, which denotes that each hyperspectral patch is $20 \times 20 \times 191$; the step size is set to 4; and the values of $r$ and $k$ are set to 7 and 4000, respectively.

To thoroughly evaluate the performance of our algorithm, we select three different noise reduction methods for comparison, i.e., hard-threshold wavelet denoising [37], the SSAHTV model [8], and video block matching 3-D filtering (VBM3D) [38]. The wavelet method adopts an adaptive noise estimation approach with no parameters. The SSAHTV method utilizes a regularization parameter to balance the fidelity term and the TV regularization term, which is set to 5 in the simulated experiments. For the VBM3D method, the noise variation is set to 15. Before directly applying the VBM3D method to the mixed noisy bands of Gaussian and impulse noise, an internal patch matching procedure is utilized to conduct the preprocessing for supporting impulse noise removal [19].

In this paper, the peak signal-to-noise ratio (PSNR) index and the structural similarity (SSIM) index [35] are adopted to give a quantitative assessment of the restoration results. As introduced in [8], for an HSI, we compute the PSNR and SSIM values...
with each clean band and restored band and then average them. The average values are then referred to as the mean PSNR (MPSNR) and mean SSIM (MSSIM) indices. The definitions of these indices are as follows:

\[
PSNR_i = 10 \times \log_{10} \frac{MN}{\sum_{x=1}^{M} \sum_{y=1}^{N} (\hat{u}_i(x,y) - u_i(x,y))^2}
\]  

\[
SSIM_i = \frac{(2\mu_{u_i}\mu_{\hat{u}_i} + C_1)(2\sigma_{u_i}\sigma_{\hat{u}_i} + C_2)}{\mu_{u_i}^2 + \mu_{\hat{u}_i}^2 + \sigma_{u_i}^2 + \sigma_{\hat{u}_i}^2 + C_1}
\]  

\[
MPSNR = \frac{1}{B} \sum_{i=1}^{B} PSNR_i
\]  

\[
MSSIM = \frac{1}{B} \sum_{i=1}^{B} SSIM_i
\]  

where \(u_i\) and \(\hat{u}_i\) represent the \(i\)th bands of the reference image and the restored image, respectively, \(\mu_{u_i}\) and \(\mu_{\hat{u}_i}\) denote the average values of image \(u_i\) and \(\hat{u}_i\), while \(\sigma_{u_i}\) and \(\sigma_{\hat{u}_i}\) stand for the variances, and \(\sigma_{u_i}\hat{u}_i\) is the covariance between \(u_i\) and \(\hat{u}_i\). Generally speaking, better restoration results are reflected by higher SSIM and PSNR values.

First, some typical bands of the HSI before and after restoration are presented to give the visual effect. Figs. 3 and 4 show bands 4 and 168 before and after restoration, which are only contaminated by Gaussian noise with SNR values of 11.27 and 19.56 dB, respectively. By comparing the restoration results of the four methods, as displayed in Figs. 3 and 4, it can be observed that the proposed LRMR restoration method performs the best, effectively suppressing the Gaussian noise and simultaneously keeping the local details of the original image. While SSAHTV can also remove the Gaussian noise and preserve the edge information, some local details are lost. What is more, when the noise level is high, the performance of SSAHTV is poor. The restoration results of VBM3D are good enough for both the low- and high-noise cases, but there is also some detailed information smoothed and lost. As to the hard-threshold wavelet restoration results, some parts of the images are distorted.

Bands 22 and 27 are corrupted with Gaussian noise and impulse noise, and the restoration results of these two bands are shown in Figs. 5 and 6, respectively. From Figs. 5 and 6, it is shown that our proposed LRMR restoration method can effectively remove the mixed Gaussian noise and impulse noise. VBM3D can also remove this mixed noise, but the details are oversmoothed, while the other two methods perform badly and fail to remove the impulse noise.

Figs. 7 and 8 display the results of removing the hybrid noise of Gaussian noise and dead lines in bands 71 and 72, respectively. Figs. 9 and 10 illustrate the restoration results for the mixed noise of Gaussian noise and stripes in bands 111 and 112, respectively. From Figs. 7–10, it can be clearly observed that the proposed LRMR method suppresses the Gaussian noise and simultaneously removes the dead lines and stripes while preserving the local details of the original image at the same time. The VBM3D method can remove the Gaussian noise effectively but performs poorly with the stripes and dead lines.
Fig. 5. Restoration results in the simulated experiment: (a) Original band 22, (b) noisy band, (c) wavelet, (d) SSAHTV, (e) VBM3D, and (f) LRMR.

Fig. 6. Restoration results in the simulated experiment: (a) Original band 27, (b) noisy band, (c) wavelet, (d) SSAHTV, (e) VBM3D, and (f) LRMR.

Fig. 7. Restoration results in the simulated experiment: (a) Original band 70, (b) noisy band, (c) wavelet, (d) SSAHTV, (e) VBM3D, and (f) LRMR.

Fig. 8. Restoration results in the simulated experiment: (a) Original band 72, (b) noisy band, (c) wavelet, (d) SSAHTV, (e) VBM3D, and (f) LRMR.
Furthermore, the wavelet method also fails to remove the dead lines and stripes and loses a lot of the details.

To further compare the performances of all the restoration algorithms, we show the spectral signatures and profiles of some typical bands before and after restoration. The differences in the spectral signatures between the noise-free and the restoration results are calculated [11]. We also use the mean profiles [36] to show the effectiveness of the four restoration methods for removing stripes and dead lines.

Fig. 11 shows the spectral signatures of pixel (110, 206), which belongs to the roof class, and Fig. 12 shows the spectral signatures of pixel (43, 143), which belongs to the grass class. The differences in the spectral signatures between the noise-free spectrum and the restoration results of pixels (110, 206) and (43, 143) are presented in Figs. 13 and 14, respectively, in which the vertical axis of the figures represents the digital number (DN) values and the horizontal axis shows the spectral-band number. From Figs. 11–14, it can be clearly observed that the proposed LRMR method produces better spectral signatures than the other restoration methods, when compared with the original spectrum.

Figs. 15 and 16 show the horizontal and vertical profiles of band 71 at pixel (179, 96), respectively. In Fig. 15(b), the horizontal profile of the noisy image at pixel (179, 96) is a straight line, which represents a horizontal dead line. Fig. 17 shows the vertical mean profiles of band 70 before and after restoration. The horizontal axis represents the column number, and the vertical axis represents the mean DN value of each
column. As shown in Fig. 17(b), due to the existence of dead lines, there are rapid fluctuations in the curve. From Figs. 15–17, it can be observed that the three profiles produced by the proposed LRMR restoration method are closest to those of the original HSI.

Finally, the PSNR and SSIM values of each band of the experimental results with the different restoration approaches are calculated and displayed in Fig. 18. It can be observed that the PSNR and SSIM values of almost all the bands obtained by our method are higher than those of the other three methods. What is more, for the bands corrupted by Gaussian noise, impulse noise, dead lines, and stripes, our method can still achieve results that are as good as that of the other bands that are contaminated by Gaussian noise only, which indicates the effectiveness of our method for simultaneously removing the
Fig. 18. PSNR and SSIM values of each band of the experimental results with the different restoration methods: (a) PSNR value and (b) SSIM value.

TABLE I

<table>
<thead>
<tr>
<th>Method</th>
<th>Wavelet</th>
<th>SSAHTV</th>
<th>VBM3D</th>
<th>LRMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPSNR</td>
<td>29.23</td>
<td>31.14</td>
<td>35.1259</td>
<td>40.37</td>
</tr>
<tr>
<td>MSSIM</td>
<td>0.8108</td>
<td>0.8566</td>
<td>0.9374</td>
<td>0.9843</td>
</tr>
</tbody>
</table>

Fig. 19. HYDICE urban data set used in the real data experiment 1 (color image composed of bands 2, 103, and 187).

mixed noise. Table I presents the MPSNR and MSSIM values of the four restoration approaches. The quantitative assessment results are consistent with the visual evaluations. It is clear that, in terms of all these indices, our LRMR method outperforms the other three approaches.

B. Real Data Experiments

In this section, two real-world test data sets are used in our experiments. One is an urban area HYDICE data set, and the other is an Earth Observing-1 (EO-1) Hyperion data set.

1) HYDICE Urban Data Set: A HYDICE urban image is used in our first real data experiment. This HSI can be downloaded online at http://www.tec.army.mil/hypercube. The original image is of the size $304 \times 304 \times 210$. As the bands 104–108, 139–151, and 207–210 are seriously polluted by the atmosphere and water absorption and can provide little useful information, we remove them, leaving the remaining test data with a size of $304 \times 304 \times 189$. Fig. 19 shows the color image...
by combining bands 2, 103, and 187, which also shows that the test data are contaminated by stripes and Gaussian noise. In our first real data experiment for LRMR, the parameters are set as follows: $q = 20$, the step size is 4, $r = 4$, and $k = 4000$. For the SSAHTV method, the regularization parameter is set to 5, and the noise variation of the VBM3D method is set to 12.

Figs. 20–23 present bands 2, 103, 187, and 189 of the restored result images. It can be clearly observed that the wavelet method cannot remove the stripes and the VBM3D method also performs poorly. The SSAHTV method can remove the stripes to a certain extent but causes the restoration results to be oversmooth. In Figs. 20–23(d), it can be seen that a lot of the image details are lost. With the proposed LRMR restoration method, the stripes are completely suppressed, and the detailed information is effectively preserved.

Figs. 20–23 show the vertical mean profiles and horizontal mean profiles of band 189 before and after restoration. It can be seen that the curves in Figs. 24(a) and 25(a) show rapid fluctuations, due to the existence of stripes. After the restoration processing, the fluctuations are suppressed to a certain extent. However, the VBM3D and wavelet methods fail to restore the image in some regions, as shown in Fig. 24. Compared to the results of SSAHTV, our method provides smoother curves, which indicates that the stripes have been removed more effectively.
ZHANG et al.: HYPERSPECTRAL IMAGE RESTORATION USING LOW-RANK MATRIX RECOVERY

Fig. 24. Vertical mean profiles of band 189 in the real data experiment 1: (a) Original, (b) wavelet, (c) SSAHTV, (d) VBM3D, and (e) LRMR.

Fig. 25. Horizontal mean profiles of band 189 in the real data experiment 1: (a) Original, (b) wavelet, (c) SSAHTV, (d) VBM3D, and (e) LRMR.

Fig. 26. EO-1 Hyperion data set used in the real data experiment 2: (a) Band 1, (b) band 68, and (c) band 148.

2) EO-1 Hyperion Data Set: Our second real data experiment adopts an EO-1 Hyperion image as the test image. The original image is of size $400 \times 1000 \times 242$ and can be downloaded online at http://datamirror.csdb.cn/admin/dataEO1Main.jsp. After removing the water absorption bands, the final test image is cut to the size of $200 \times 400 \times 166$. In our real data experiment 2, the parameters are set as follows: $q = 20$, the step size is 4, $r = 4$, and $k = 4000$. For the SSAHTV method, the regularization parameter is set to 5, and the noise variation of the VBM3D method is set to 10. Many bands of the test image contain dead lines and stripes, such as the bands 1, 68, and 148 shown in Fig. 26.

Figs. 27–29 present the restoration results of bands 1, 68, and 148 by the different restoration methods. The stripes and dead lines in Figs. 27–29(a) are still obvious, which suggests that the wavelet method fails to suppress the stripes and dead lines. The SSAHTV and VBM3D methods can only remove part of the stripes and dead lines, as presented in Figs. 27–29(b) and (c). What is more, the results produced by SSAHTV are oversmooth, and many details of the original image are lost. The results shown in Figs. 27–29(d) indicate that our method performs the best, completely removing the stripes and dead lines and effectively preserving most of the local details.

Figs. 30 and 31 display the vertical mean profiles of bands 68 and 148 before and after restoration, respectively. The rapid fluctuations in Figs. 29(a) and Figs. 30(a) suggest the existence of stripes and dead lines in bands 68 and 148. From all the restored vertical mean profiles, it can be observed that the LRMR restoration method gives the best restored results with regard to the removal of dead lines and stripes.

C. Discussion

In all of our simulated and real data experiments, the size of the subdata cube is selected as $q = 20$, and the step size is set as 4. In fact, the size of the subdata cube $q$ does not have a large effect on the restoration results of our proposed LRMR method. Table II presents the MPSNR and MSSIM values of the results produced by the proposed LRMR method with different sizes of subcube $q$. It can be observed that the LRMR restoration method is quite robust with regard to the size of the HSI subcube. After the patch size $q$ is determined, we should choose the values of the rank $r$ and the cardinality $k$ before the LRMR processing. As introduced in Section II, the value $r$ stands for the number of the pure spectral endmembers in the subdata cube, and $k$ represents the upper bound of the number of pixels polluted by impulse noise, dead lines, and stripes. Fig. 32 presents the quantitative evaluation results of different rank $r$ and cardinality $k$ in the simulated experiment. It can be clearly seen that the results of the proposed LRMR restoration method are quite robust with regard to the values of $r$ and $k$ when they exceed their corresponding optimal values.

To study the effect of different numerical approaches on the LRMR-based HSI restoration model, the GoDec and ALM numerical methods are adopted and compared in a simulated experiment in terms of PSNR and SSIM quantitative...
evaluation measures. The parameter of the ALM method is set empirically, and the optimal value is 0.05. Fig. 33(a) and (b) presents the PSNR and SSIM values of each band for the GoDec and ALM methods, respectively, and Table III displays the MPSNR and MSSIM values of the two approaches. It is clearly shown that the GoDec method achieves better restoration results, in terms of higher PSNR and SSIM values, which suggests the advantage of the GoDec method for Gaussian noise removal in the HSI restoration problem. In addition, the parameters $r$ and $k$ stand for explicit meanings of the HSI, which allows the users to determine the parameter values more easily.
ZHANG et al.: HYPERSPECTRAL IMAGE RESTORATION USING LOW-RANK MATRIX RECOVERY

Fig. 30. Vertical mean profiles of band 68 in the real data experiment 2: (a) Original, (b) wavelet, (c) SSAHTV, (d) VBM3D, and (e) LRMR.

Fig. 31. Vertical mean profiles of band 148 in the real data experiment 3: (a) Original, (b) wavelet, (c) SSAHTV, (d) VBM3D, and (e) LRMR.

TABLE II
MPSNR AND MSSIM VALUES OF THE LRMR RESULTS WITH DIFFERENT SIZES OF SUBCUBE, WITH THE STEP SIZE BEING 4

<table>
<thead>
<tr>
<th>Subcube Size</th>
<th>MPSNR</th>
<th>MSSIM</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>39.70</td>
<td>0.9725</td>
</tr>
<tr>
<td>12</td>
<td>39.72</td>
<td>0.9742</td>
</tr>
<tr>
<td>16</td>
<td>40.21</td>
<td>0.9792</td>
</tr>
<tr>
<td>20</td>
<td>40.37</td>
<td>0.9843</td>
</tr>
<tr>
<td>24</td>
<td>40.25</td>
<td>0.9823</td>
</tr>
<tr>
<td>64</td>
<td>40.13</td>
<td>0.9831</td>
</tr>
</tbody>
</table>

TABLE III
MPSNR AND MSSIM VALUES OF THE RESTORATION RESULTS IN THE SIMULATED EXPERIMENT VIA DIFFERENT LRMR NUMERICAL APPROACHES

<table>
<thead>
<tr>
<th>Method</th>
<th>MPSNR</th>
<th>MSSIM</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALM</td>
<td>38.467</td>
<td>0.9824</td>
</tr>
<tr>
<td>GoDec</td>
<td>40.37</td>
<td>0.9843</td>
</tr>
</tbody>
</table>

Finally, the computational complexity of our method is discussed. As shown in [21], the GoDec algorithm is an iterative algorithm. For a matrix of size $q^2 \times B$, each iteration step of the GoDec algorithm needs $O(q^2 \times B \times r)$ flops. Then, for each subcube of the HSI, $T \times O(q^2 \times B \times r)$ flops are needed, where $T$ is the number of the iteration steps.

V. CONCLUSION

In this paper, we have proposed an LRMR-based HSI restoration method. By lexicographically ordering the 3-D cube into a 2-D matrix representation, the HSI restoration problem is transformed to an LRMR problem, on the basis of the low-rank property of the clean HSI. The GoDec algorithm is then applied to solve the LRMR-based HSI restoration model. The main advantage of the LRMR restoration algorithm is that all the Gaussian noise, impulse noise, dead lines, and stripes are taken into consideration. One simulated experiment and two real data experiments were conducted. The experimental results confirm that our proposed HSI restoration method can effectively and simultaneously remove the mixed noise of Gaussian noise, impulse noise, dead lines, and stripes. In addition, the proposed method is quite robust and stable with regard to noise types and parameter settings, which improves the potential of the practical application in HSI processing.

Nevertheless, the proposed algorithm still has room for improvement. There is no spatial constraint imposed on the neighboring pixels of the HSI, which may cause an unsatisfactory performance for very large areas of missing pixels. This is a drawback of the method that needs to be overcome in the future. In addition, the adaptive determination of rank $r$ and cardinality $k$ needs to be addressed in our further work.

Fig. 32. Restoration results under different values of the rank $r$ and cardinality $k$.

Fig. 33. PSNR and SSIM values of each band of the experimental results with the different numerical approaches to LRMR: (a) PSNR value and (b) SSIM value.
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 “Go Decomposition” algorithm code.

REFERENCES


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. Since 2010, he has been a Lecturer with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University. His current research interests focus on image reconstruction and remote sensing image processing.

Wei He received the B.S. degree in mathematics and statistics from Wuhan University, Wuhan, China, in 2012, where he is currently working toward the Ph.D. degree in the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing. His research interests include image quality improvement, remote sensing image processing, and low-rank representation.
Liangpei Zhang (M’06–SM’08) received the B.S. degree in physics from Hunan Normal University, Changsha, China, in 1982, the M.S. degree in optics from the Xi’an Institute of Optics and Precision Mechanics of Chinese Academy of Sciences, Xi’an, China, in 1988, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1998.

He is currently with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, as the Head of the Remote Sensing Division. He is also a “Chang-Jiang Scholar” Chair Professor appointed by the Ministry of Education, China. He is currently the Principal Scientist for the 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 200 research papers and is the holder of five patents. His research interests include hyperspectral remote sensing, high-resolution remote sensing, image processing, and artificial intelligence.


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. His research interests include image processing (for quality improvement), remote sensing application, data fusion, and assimilation. He has published more than 60 research papers. He 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).

Qiangqiang Yuan 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 2006 and 2012, respectively.

In July 2012, he joined the School of Geodesy and Geomatics, Wuhan University, where he is currently an Assistant Professor. His research interests focus on image restoration, remote sensing image processing, and application.