Signal Processing 111 (2015) 230–248

Contents lists available at ScienceDirect

Signal Processing

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

Hyperspectral image recovery employing a multidimensional nonlocal total variation model

Jie Li [a], Qiangqiang Yuan [b,d,*], Huanfeng Shen [c,d], Liangpei Zhang [a,d]

[a] The State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, PR China
[b] School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, PR China
[c] School of Resource and Environmental Science, Wuhan University, Wuhan 430079, PR China
[d] Collaborative Innovation Center of Geospatial Technology, Wuhan University, Wuhan 430079, PR China

ARTICLE INFO

Article history:
Received 10 July 2014
Received in revised form
29 October 2014
Accepted 18 December 2014
Available online 24 December 2014

Keywords:
Hyperspectral image
Noise reduction
Nonlocal total variation

ABSTRACT

Hyperspectral images (HSIs) have a high spectral resolution and ground-object recognition ability, but inevitably suffer from various factors in the imaging procedure, such as atmospheric effects, secondary illumination, and the physical limitations, which have a direct bearing on the visual quality of the images and the accuracy of the subsequent processing. HSI restoration is therefore a crucial task for improving the precision of the subsequent products. Currently, patch-based schemes have offered promising results for the preservation of detailed information and the removal of additive noise. In HSIs, the information in the spectral dimension is more redundant than the information in the spatial dimension. We therefore propose a multidimensional hyperspectral nonlocal model, in which both the correlation of the spectral bands and the similarity of the spatial structure are considered. In the model, a multidimensional nonlocal total variation constraint is applied to preserve edge sharpness. Experiments with both synthetic and real hyperspectral data illustrate that the proposed method can obtain promising results in HSI restoration.

1. Introduction

Hyperspectral images (HSIs) simultaneously provide spatial and spectral information to identify specific materials in a scene. Unfortunately, during the acquisition procedure of HSIs, atmospheric effects, secondary illumination, and the physical limitations of the sensors (such as artifacts, sensor noise, and dead pixels) degrade the quality of the images [1]. These disturbance factors influence the visual effect of the HSIs and limit the precision of the subsequent applications, such as land-surface classification, object identification, and change detection. To achieve a more accurate estimation, it is important to overcome these limitations and improve the quality of the HSIs.

HSI restoration aims at generating a high-quality image from its degraded version. To date, various HSI restoration techniques have been proposed. We review the existing popular HSI restoration methods in the following. One type of methods is based on the strategy of transform domain [2–6]. With these methods, the input hyperspectral signals are converted into signals in another space, such as the wavelet domain, in which the noise is easily separated from the signal using the compactness of the true signal. The traditional wavelet denoising techniques apply a 2D wavelet transform on each band separately, and thus discard the spectral correlation information. To improve its performance with HSIs, the wavelet transform has been combined with other spectral band decorrelation methods, such as

---

* Corresponding author.
E-mail address: yqiang86@gmail.com (Q. Yuan).

http://dx.doi.org/10.1016/j.sigpro.2014.12.023
0165-1684/

discrete Fourier transform [2] and PCA [3,4]. To exploit the inter-band correlation and spatial information, other advanced HSI noise reduction techniques can be accomplished through wavelet thresholding [5] in the Bayesian estimation framework, and in combination with different prior models [7–10]. However, the biggest drawback of the wavelet-based methods is that they often generate ringing artifacts, shown as additional edges or structures [11].

To preserve the spectral feature, multidimensional filter methods [12–14] have been developed to consider the HSI as a multidimensional data cube in the spatial domain, to simultaneously process the spatial and spectral information. These methods include the multidimensional Wiener filter [12], genetic kernel tucker decomposition [15], and adaptive 3D filtering [16]. However, the classical multidimensional analysis methods can have great difficulty in distinguishing the signal and noise subspaces, and thus may introduce some artifacts, and they also tend to oversmooth the image and lose many textural details [13].

Together with the progress made in remote sensing, to better preserve the textural details and overcome the artifacts, regularization-based approaches [17–23] have emerged in recent years to enhance both the spatial structure and spectral feature. These approaches recover the original image by adding a reasonable assumption or prior knowledge into the observation model. The different priors can be applied to meet different goals, such as preserving edges, protecting textural details, and avoiding artifacts and noise. Yuan et al. [18] employed a spectral–spatial adaptive total variation (TV) model to adaptively denoise image in both the spatial and spectral dimension. Chen and Hu [19] proposed a spatial–spectral domain mixing prior, in which an edge-preserving prior is used to preserve the geometrical structure in the spatial domain, and adaptive spectral weights for the different materials are constructed in the spectral domain. Qian [20] used variance-stabilizing transformation to simplify the mixed-noise into Gaussian noise, and then introduced a structured sparsity-based model to remove the noise.

In regularization-based algorithms, the HSI recovery is cast as the inverse problem of recovering the original high-quality image. A robust estimation for the solution is obtained relying on some strong image priors, and various regularization functions have been proposed to further stabilize the inversion of this ill-posed problem, such as Tikhonov regularization [24], Gaussian Markov random fields regularization [25], Huber-MRF regularization [26], TV regularization [18], nonlocal-based regularization [27,28], and sparse regularization [21,22]. Among these models, the nonlocal-based model [29] is a very popular and powerful tool, which has been widely used in various applications, such as denoising [30], super-resolution reconstruction [31], inpainting [32], and shadow removal [33], because of its good performance in edge and texture preservation.

For HSIs, the simplest way to apply a nonlocal-based regularization is in a band-by-band manner. However, the spectral dependency and inter-channel relationship of the hyperspectral signals will not be fully made use of. Furthermore, owing to the relatively low spatial resolution of HSI [34], the similarity between patches from only a single band is insufficient. At the same time, as the noise-intensity in each band is usually different, the denoising strength should be adaptively adjusted with the noise-intensity in each band. Therefore, we propose a spectrally adaptive multidimensional nonlocal total variation (SAMNLTV) model by exploiting the high correlation of bands to better restore a low-quality HSI. The main ideas and contributions of the proposed approach can be summarized as follows:

(1) A multidimensional nonlocal TV regularization is proposed to acquire more redundancy from the highly correlated bands. Since the intensity of the signal is contiguous in the highly correlated or neighboring bands, they are selected to provide more similar patches in the scheme.

(2) A spectrally adaptive method is proposed for the multidimensional nonlocal TV model. To suppress the different intensities of noise in the different bands, a wavelet method is applied to roughly estimate the strength of noise in the different bands. By making use of the noise strength, an adaptive regularization parameter selection strategy is proposed to improve the restoration results.

(3) A split Bregman iteration algorithm is used to optimize the proposed HSI restoration model. From the experimental results with both simulated and real data, it is illustrated that the proposed model produces good image restoration results.

The rest of this paper is organized as follows. In Section 2, the proposed multidimensional nonlocal total variation model is formulated. Section 3 contains the experimental results and discussion, and Section 4 is the conclusion.

2. The multidimensional nonlocal total variation model

Assuming that we have a HSI $\mathbf{U} \in R^{M_1 M_2 \times B}$ corrupted by an additive noise $\mathbf{V} \in R^{M_1 M_2 \times B}$. Mathematically, this is denoted as $\mathbf{U} \in R^{M_1 M_2 \times B}$, where the matrix representation of the original HSI is of a size of $M_1 \times M_2 \times B$, in which M_1 represents the number of samples in a line, M_2 stands for the number of lines in the image, and B denotes the number of bands. The degradation model for each band can then be defined as

$$\boldsymbol{f}_b = \boldsymbol{u}_b + \boldsymbol{v}_b \tag{1}$$

where $\mathbf{u}_b \in R^{M_1 M_2}$ denotes the vector representation of one band with a size of $M_1 \times M_2$. $\mathbf{f}_b \in R^{M_1 M_2}$ denotes one band of the degraded image $\boldsymbol{F} \in R^{M_1 M_2 \times B}$, and the additive noise is $\mathbf{v}_b \in R^{M_1 M_2}$, which is added to the bands $\mathbf{u}_b$.

Applying the maximum a posteriori probability (MAP) estimator, the HSI restoration model can be represented as the following regularized least squares problem [18]:

$$\widehat{\boldsymbol{U}} = \arg\min_{\mathbf{U}} \left\{ \sum_{b=1}^{B} \|\boldsymbol{u}_b - \mathbf{f}_b\|_2^2 + \lambda \Phi(\boldsymbol{U}) \right\} \tag{2}$$

In the cost function, the first term is called the fidelity term, which denotes the fidelity between the observed noisy data and the original clear data, while the second term $\Phi(\boldsymbol{U})$ is an additional regularization function. λ is the regularization parameter used to balance the tradeoff between the fidelity term and the regularization term.

2.1. Related nonlocal framework

The nonlocal methods in image restoration are generalized from the Yaroslavsky filter [35], and the nonlocal means (NLM) [34] filter was firstly proposed to explore the nonlocal similarity in the image. In comparison with the pixel-based filter, it allows a more robust similarity, and the matching patterns are not restricted to be local.

Given the noisy observation $\mathbf{f}_b$, the pixel value $\boldsymbol{u}_b(i)$ in the clear band $\mathbf{u}_b$ can be estimated by calculating a weighted average of the neighboring pixels $\mathbf{f}_b(j)$, following this formula:

$$\boldsymbol{u}_b(i)=\sum\nolimits_{i\neq j,\forall i\in N_i}\boldsymbol{w}(i,j)\boldsymbol{f}_b(j)\quad\text{with}\quad 0\leq\boldsymbol{w}(i,j)\leq 1$$
$$\text{and}\quad\sum\nolimits_{i\neq j,\forall i\in N_i}\boldsymbol{w}(i,j)=1\tag{3}$$

where N_i is the size of searching window. The weights $\mathbf{w}(i,j)$ can be calculated as

$$\boldsymbol{w}(i,j)=\frac{1}{C(i)}\exp\left\{-G_a*\left(\|\boldsymbol{f}_b(i+\cdot)-\boldsymbol{f}_b(j+\cdot)\|^2\right)/2h_b^2\right\}\tag{4}$$

$$C(i)=\sum\nolimits_{\forall i\in N_i}\exp\left\{-G_a*\left(\|\boldsymbol{f}_b(i+\cdot)-\boldsymbol{f}_b(j+\cdot)\|^2\right)/2h_b^2\right\}\tag{5}$$

where G_a is the Gaussian kernel with standard deviation a, $C(i)$ is the normalizing constant, and h_b is a filtering parameter of each band b. The weight is related to the similarity between the neighboring patches $\boldsymbol{f}_b(i+\cdot)$ and $\boldsymbol{f}_b(j+\cdot)$ with defined radius n, which is centered at pixel i and j.

With the nonlocal filter idea, several nonlocal-based regularization models have been developed, in which the most typical one is the nonlocal total variation (NLTV) model [36]. For a given point i, j represents the point in the searching neighborhood window around i, and the weight $\mathbf{w}(i,j)$ is assumed to be symmetric. We define

$$\text{Nonlocal gradient:}\quad\nabla_w\boldsymbol{u}_b(i,j)=(\boldsymbol{u}_b(j)-\boldsymbol{u}_b(i))\sqrt{\boldsymbol{w}(i,j)}\tag{6}$$

The nonlocal total variation:

$$\begin{aligned}\Phi_{NLTV}(\boldsymbol{u}_b)&=\sum_{i=1}^{M_1M_2}\left|\nabla_w\boldsymbol{u}_b(i)\right|\\&=\sum_{i=1}^{M_1M_2}\sqrt{\sum_{j\in N_i}^{N_i}(\boldsymbol{u}_b(i)-\boldsymbol{u}_b(j))^2\boldsymbol{w}(i,j)}\end{aligned}\tag{7}$$

The gradient measure of NLTV between point i and j is computed with $\sqrt{\boldsymbol{w}(i,j)}$ based on patches. With consideration of the geometrical configuration, the weight measure based on patch distances is robust to noise.

However, recovering hyperspectral data band by band in (7) cannot fully exploit the high correlation between spectral bands, and generates spectral distortion in the recovered HSI. As a result, inspired by color TV regularization [37], Cheng et al. [38] presented a multichannel NLTV (McNLTV) regularization by coupling the channels in a multichannel image.

$$\begin{aligned}\Phi_{McNLTV}(\boldsymbol{U})&=\sum_{i=1}^{M_1M_2}\sqrt{\sum_{b=1}^{B}\left|\nabla_w\boldsymbol{u}_b(i)\right|^2}\\&=\sum_{i=1}^{M_1M_2}\sqrt{\sum_{b=1}^{B}\sum_{j\in N_i}^{N_i}(\boldsymbol{u}_b(i)-\boldsymbol{u}_b(j))^2\boldsymbol{w}(i,j)}\end{aligned}\tag{8}$$

Although the multichannel NLTV preserves the edge information by joining all the nonlocal gradients along the spectral dimension, it is only applied in image inpainting and does not focus on information redundancy in the spectral dimension. In this paper, we compute the nonlocal gradient simultaneously along the spatial and the spectral dimensions, and therefore form a three-dimensional nonlocal gradient cube to maintain the spectral characteristics.

2.2. The multidimensional nonlocal total variation model

Nonlocal regularization is mainly concerned with the similarity of patches. More similar information helps to better improve the processing result [39]. Finding similar patches from the noisy image itself is popular because patches tend to recur within the image. However, the low repeatability of the texture results in a limited performance for this approach with regard to images of low spatial resolution. External clean natural image patches have also been proposed as another source of similar patches [40,41]. Targeted external image databases were effectively applied to denoise images by means of group sparsity and localized priors in [40]. Recently, external databases were also demonstrated to be able to recover patches with edges and texture better than internal databases [41]. HSIs usually have a low spatial resolution and high spectral resolution. Although patches from only one band may be insufficient as a result of the low spatial resolution, the highly correlated bands manifest similar image patterns. Meanwhile, the highly correlated bands with low noise-intensity can give clean patches to preserve the details of the image, in the same way as external databases. Here, we assume that the patches from highly correlated bands offer extra redundant information, instead of only finding the patches from one band. To verify the assumption, the distribution of similar patches in the current band to be restored, and its highly correlated bands, is plotted in Fig. 1. Here, it is clear that the highly correlated bands are able to obtain the same number of similar patches as the current band, and even produce more similar patches than the current band. It is confirmed that local patches from the highly correlated bands have a similar structure.

Therefore, differing from single-band NLTV, we integrate the similarity of the spectral band into the nonlocal TV model. The weights associated with the multiple bands are then computed with the distance between the patch in the current band and the patch in the structurally similar band, as written in (9):

$$\mathbf{w}(i,b;j,d)=\begin{cases}\frac{1}{C(i,b)}\exp\left\{-\frac{G_a*\left(\|\mathbf{f}_b(i+\cdot)-\mathbf{f}_b(j+\cdot)\|^2\right)}{2h_b^2}\right\}, & \text{if}\quad b=d,\quad j\in N_i^b\\ \frac{1}{C(i,b)}\exp\left\{-\frac{G_a*\left(\|\mathbf{f}_b(i+\cdot)-\mathbf{f}_d(j+\cdot)\|^2\right)}{2h_b^2}\right\}, & \text{otherwise},\quad j\in N_i^d\end{cases}\tag{9}$$

$$C(i,b)=\sum_{j\in\Omega_i}^{\Omega_i}\exp\left\{-\frac{G_a*\left(\|\mathbf{f}_b(i+\cdot)-\mathbf{f}_d(j+\cdot)\|^2\right)}{2h_b^2}\right\},\quad d\in NS_b\tag{10}$$

where $\mathbf{w}(i,b;j,d)$ represents the Euclidean distance weight between two patches from the selected b and d bands.

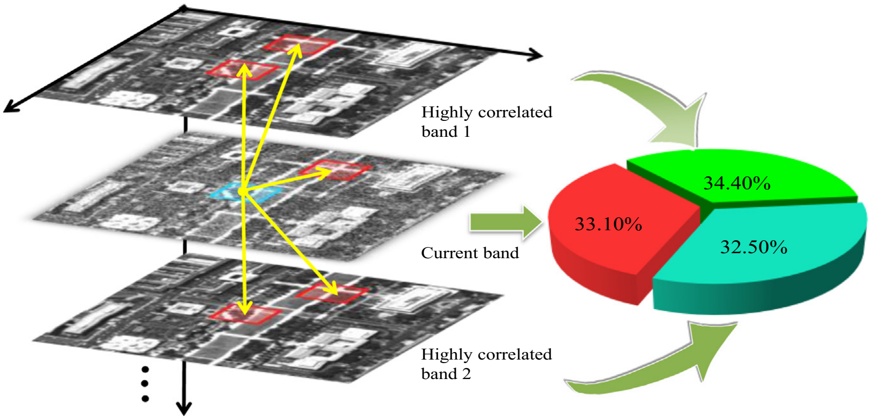

Fig. 1. Percentages of similar patches from three bands.

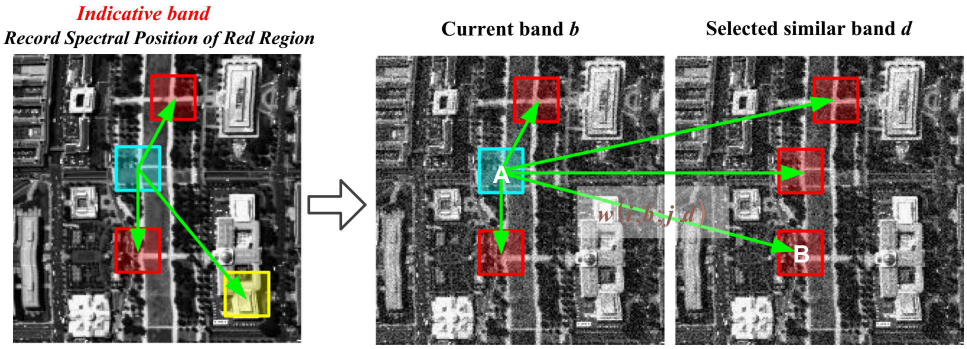

Fig. 2. The searching scheme of the proposed multidimensional nonlocal TV model. “A” denotes the current patch $\boldsymbol{f}_b(i+\cdot)$, and “B” denotes the searched patch $\mathbf{f}_d(j+\cdot)$. For given patches $\mathbf{f}_b(i+\cdot)$ in band b, every similar pixel neighborhood $\mathbf{f}_d(j+\cdot)$ in the structurally similar band is given a weight $\mathbf{w}(i,b;j,d)$.

i denotes the spatial coordinate of band b, and j denotes the spatial coordinate of band d. NS_b denotes the set of all the selected bands for the current bth band. The set contains the current bth band and these similar bands. We define $\boldsymbol{f}_b(i+\cdot)$ as the neighboring patches centered at pixel i with a size of $n\times n$ in the bth band. The searched patch $\mathbf{f}_d(j+\cdot)$ stands the patch at pixel j with a size of $n\times n$ in the dth band. N_i^d is the searching neighborhood window in the dth band, and Ω_i is the union of all $\left\{N_i^d\right\}_{d\in NS_b}$ sets. The normalization factor $C(i,b)$ is the sum of the weights of current band b. Furthermore, the bandwidth h_b controls the softness of the results. Usually, the bandwidth h_b is chosen as the standard deviation of noise added [34]. For the bands with high-intensity noise, h_b is set to be a larger value. Conversely, h_b is set to be a smaller value for low-intensity noise.

Owing to numerous and continuous spectral bands, HSI is firstly divided into many groups. For a clear HSI, these bands with high value of correlation coefficient are assigned to a group. However, in the real noising situation, the correlation coefficients are usually affected by noise. Some bands which should be cross-correlation show the low correlation after corrupted by noise. For this case, we selected the adjacent bands around it within 10 bands range as the highly correlated bands.

For those bands in the same group, we suppose that the patches from the same spatial coordinates (i,j) of different spectral bands should be similar in spatial structure. With less impact of noise, the low-noise band can obtain the optimal weight value and position of the similar patches. Therefore, we can find a batch of more accurate spatial coordinates of similar patches for each group by introducing a low-noise “indicative band”. For each group, the “indicative band” is the lowest-noise band in the current group by the estimated noise-intensity, and is used to choose and record the accurate position of the similar patches. The recorded coordinates are used to extract the similar patches from the selected bands. The process of searching the similar patches in the multidimensional nonlocal TV is described in Fig. 2.

Using the multidimensional nonlocal algorithm to estimate the value of $\mathbf{u}_b(i)$, the function is expressed as

$$\boldsymbol{u}_b(i)=\sum\nolimits_{\forall j\in\Omega_i}\mathbf{w}(i,b;j,d)\boldsymbol{f}_d(j)\,,\quad \Omega_i=\cup\left\{N_i^d\right\}_{d\in NS_b} \tag{11}$$

where $\boldsymbol{f}_d(j)$ is the value of pixel j in the dth band of the corrupted (noisy or contaminated) HSI.

Combining the search mechanism of the multidimensional nonlocal model with the computation of the gradient, the multidimensional nonlocal gradient is defined as

$$\nabla_{Mw}\boldsymbol{u}_b(i,b;j,d)=(\boldsymbol{u}_d(j)-\boldsymbol{u}_b(i))\sqrt{\mathbf{w}(i,b;j,d)},\quad d\in NS_b \tag{12}$$

where $\nabla_{Mw}\boldsymbol{u}_b(i,b;j,d)$ is defined as the vector of all the partial derivatives about i in the current bth band, $\boldsymbol{u}_b(i)$ and

$\boldsymbol{u}_d(j)$ represent the pixel in the current band and the selected band, respectively. When $d=b$, the selected band is equal to the current band, and the nonlocal gradient is analogous to (6), in which the nonlocal gradient is only computed in the spatial domain. When $d \neq b$, the nonlocal gradient is computed with the current bth band and the selected dth band, in both the spatial and spectral dimensions.

For HSIs, the multidimensional nonlocal gradient consists of the spatial dimension and the spectral dimension. The gradient of the spectral dimension preserves the continuity of signals in the spectrum, and the more similar structural patches can be obtained in the spectral domain because of the high spectral resolution.

Using the multidimensional nonlocal gradient, the multidimensional nonlocal TV (MNLTV) is shown as

$$\begin{aligned}\Phi(\boldsymbol{U}) &= \sum_{b=1}^{B}\sum_{i=1}^{M_1M_2}\left|\nabla_{Mw}\boldsymbol{u}_b(i)\right| \\ &= \sum_{b=1}^{B}\sum_{i=1}^{M_1M_2}\sqrt{\sum\nolimits_{d\in NS}\sum\nolimits_{j\in N_i^d}(\boldsymbol{u}_b(i)-\boldsymbol{u}_d(j))^2\boldsymbol{w}(i,b;j,d)}\end{aligned} \tag{13}$$

Then, in conjunction with the proposed MNLTV, we present the following energy minimization problem to address the MNLTV-based model for HSIs:

$$\widehat{\boldsymbol{U}} = \arg\min_{\mathbf{U}}\left\{\sum_{b=1}^{B}\|\boldsymbol{u}_b-\boldsymbol{f}_b\|_2^2+\sum_{b=1}^{B}\lambda_b\sum_{i=1}^{M_1M_2}\left|\nabla_{Mw}\boldsymbol{u}_b(i)\right|\right\} \tag{14}$$

where the parameter $\lambda=[\lambda_1,\ldots,\lambda_b,\ldots,\lambda_B]$ controls the tradeoff between the nonlocal regularization term and the other terms in the objective function, which also means that in different bands, the regularization strength should be different.

2.3. Spectral adaptation for multidimensional nonlocal TV

As noise usually appears in different distributions in different bands, each band should be treated with a different degree of penalty; that is to say, the regularization parameter λ_b should be adaptively computed in different spectral bands with different noise-intensities. For example, for the high-noise bands, a high value of parameter λ_b should be used to reduce the noise. Conversely, a low value of parameter λ_b is needed for the low-noise bands. To reduce the cost of manual selection, it is crucial to define an adaptive parameter for the different noise strengths.

Firstly, to adjust the regularization parameter, the coefficient of the highest frequency wavelet sub-band $\{HH_s\}_{s=1}^{J-1}$ can be used to estimate the intensity of the noise [42].

$$\widehat{\sigma} = \frac{\mathrm{Median}\left(\{\boldsymbol{F}_{HH_S}\}_{s=1}^{J-1}\right)}{0.6745},\quad \{\boldsymbol{F}_{HH_S}\}_{s=1}^{J-1}\in \mathrm{subband}\{HH_s\}_{s=1}^{J-1} \tag{15}$$

where $\{\boldsymbol{F}_{HH_S}\}_{s=1}^{J-1}$ denotes the diagonal detail coefficient matrix.

Next, with the estimated noise standard deviation $\widehat{\sigma}_b$ for each band, we apply them to set the different regularization parameters λ_b, with $\lambda_b \in \lambda$. To treat the different noise level bands with different denoising strengths, we refer to the idea used in the SSAHTV model [18], and propose a new spectrally adaptive strategy for the hyperspectral multidimensional nonlocal TV model in the following:

$$\widehat{\boldsymbol{U}} = \arg\min_{\mathbf{U}}\left\{\sum_{b=1}^{B}\|\boldsymbol{u}_b-\boldsymbol{f}_b\|_2^2+\sum_{b=1}^{B}\lambda_b\sum_{i=1}^{M_1M_2}\left|\nabla_{Mw}\boldsymbol{u}_b(i)\right|\right\}, \tag{16}$$

$$\lambda_b = \frac{\tau\left|\nabla_{Mw}\boldsymbol{u}_b\right|}{\sqrt{\sum_{b=1}^{B}\left|\nabla_{Mw}\boldsymbol{u}_b\right|^2}} \tag{17}$$

where τ is a constant parameter. In (16), the estimated noise standard deviation σ_b is related to $\left|\nabla_{Mw}\boldsymbol{u}_b\right|$, because $h_b=2\sigma_b$.

According to Yuan et al. [18], the value of the TV model is related to the intensity of the noise and can adaptively control the strength of the denoising. Therefore, we believe that the nonlocal TV can also be regarded as an index that reflects the noise-intensity. We can further keep the anomaly of the weights caused by structural differences to a minimum with the band grouping. This is because the highly correlated bands which are grouped into the same cluster have closer gradient information and minimal structural differences. On the other hand, the nonlocal gradient in the spectral dimension can reflect the difference of the spectral noise. Compared to the simple TV model, nonlocal TV is a more robust index for reflecting the intensity of the noise, because of the more stable weight and the more comprehensive gradient direction. To prove the effectiveness of the spectral adaptation in (17), the results of the tuning parameters from SAMNLTV and SSAHTV are discussed in Section 3.

2.4. Model optimization

The SAMNLTV-based recovery model for HSIs is a reweighted L^1 minimization problem, which can be effectively solved by Bregmanized operator splitting (BOS) [43]. With auxiliary variable $\mathbf{Q}$ added into (16), and $\mathbf{u}_b$ regarded as the column vector of $\mathbf{U}$, the algorithm reformulates the problem as

$$\widehat{\boldsymbol{U}} = \arg\min_{\mathbf{U}}\left\{\|\boldsymbol{U}-\boldsymbol{F}\|_F^2+\sum_{b=1}^{B}\lambda_b\sum_{i=1}^{M_1M_2}\left|\mathbf{Q}_{bi}\right|\right\},\quad \text{s.t.}\quad \mathbf{Q}_b=\nabla_{Mw}\boldsymbol{u}_b \tag{18}$$

The constrained problem in (18) can be changed into an unconstrained problem with the Bregman iteration method, as follows:

$$\widehat{\boldsymbol{U}} = \arg\min_{\mathbf{U}}\left\{\|\boldsymbol{U}-\boldsymbol{F}\|_F^2+\lambda\left|\mathbf{Q}\right|+\mu\|\mathbf{Q}-\nabla_{Mw}\boldsymbol{U}-\boldsymbol{P}\|_F^2\right\} \tag{19}$$

where $\mathbf{P}$ is also an auxiliary variable. The problem in (19) can be solved by performing an alternating minimization processing:

$$\boldsymbol{U}^{k+1} = \arg\min_{\mathbf{U}}\left\{\|\boldsymbol{U}-\boldsymbol{F}\|_F^2+\mu\|\mathbf{Q}^k-\nabla_{Mw}\boldsymbol{U}-\boldsymbol{P}^k\|_F^2\right\} \tag{20}$$

$$\mathbf{Q}^{k+1} = \arg\min_{\mathbf{Q}}\left\{\lambda\left|\mathbf{Q}\right|+\mu\|\mathbf{Q}-\nabla_{Mw}\boldsymbol{U}^{k+1}-\boldsymbol{P}^k\|_F^2\right\} \tag{21}$$

To solve the subproblem equation in (20) for $\mathbf{U}^{k+1}$, the equation can be solved as

$$\left(\boldsymbol{U}^{k+1}-\boldsymbol{F}\right)-\mu div_{Mw}\left(\nabla_{Mw}\boldsymbol{U}^{k+1}+\boldsymbol{P}^{k}-\boldsymbol{Q}^{k}\right)=0 \tag{22}$$

For the diagonal linear system (22), one of the most efficient methods is the Gauss–Seidel method:

$$\boldsymbol{U}^{k+1}=\left(1-\mu\Delta_{Mw}\right)^{-1}\left[\boldsymbol{F}+\mu div_{Mw}\left(\boldsymbol{P}^{k}-\boldsymbol{Q}^{k}\right)\right] \tag{23}$$

The $\mathbf{Q}^{k+1}$ subproblem in (21) can be obtained by applying the shrinkage operator for the vector field at each point $(b,\ i)$, representing the point i in band b.

$$\mathbf{Q}_{bi}^{k+1}=\text{shrink}\left(\left(\nabla_{Mw}\mathbf{U}^{k+1}+\mathbf{P}^{k}\right)_{bi},\ \frac{\lambda}{\mu}\right) \tag{24}$$

where $\text{shrink}(p,\lambda/\mu)=(p/|p|)\max\{0,|p|-\lambda/\mu\}$.

Finally, for the parameter $\mathbf{P}$, it should be updated in each iteration in the following way:

$$\boldsymbol{P}^{k+1}=\boldsymbol{P}^{k}+\nabla_{Mw}\boldsymbol{U}^{k+1}-\boldsymbol{Q}^{k+1} \tag{25}$$

We outline the optimization procedure below.

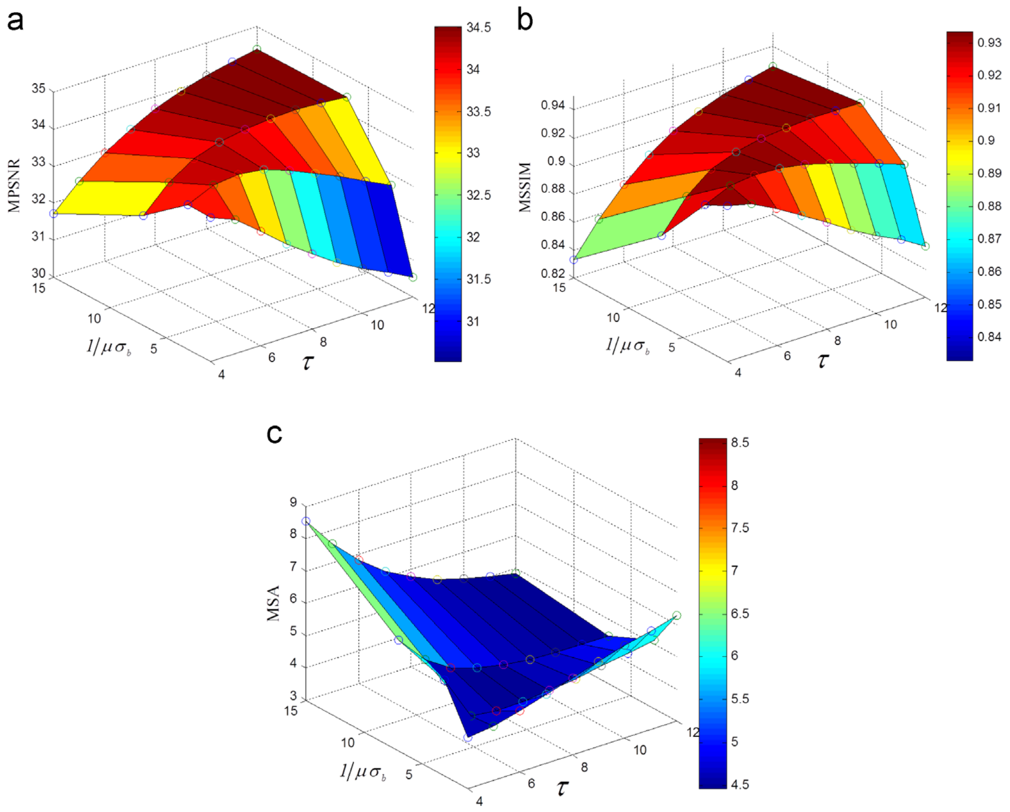

Fig. 3. Results of the proposed method with different values of parameters τ and μ.

Table 1
Quantitative evaluation of the denoising results in the simulated experiment.

Method	Evaluation index	Adaptive Wiener filter	NLM	SANLTV	SSAHTV	**SAMNLTV**
$\sigma=10$	MPSNR	29.401 ± 0.0014	30.265 ± 0.0012	30.912 ± 0.0021	31.614 ± 0.0025	$\mathbf{34.044\pm0.0031}$
	MSSIM	0.846 ± 0.0001	0.879 ± 0.0001	0.888 ± 0.0001	0.906 ± 0.0001	$\mathbf{0.935\pm0.0001}$
	MSA (degree)	5.457 ± 0.0008	5.265 ± 0.0019	4.619 ± 0.0017	4.390 ± 0.0013	$\mathbf{3.999\pm0.0017}$
$\sigma=20$	MPSNR	26.753 ± 0.0016	27.046 ± 0.0057	27.460 ± 0.0033	28.193 ± 0.0025	$\mathbf{29.166\pm0.0039}$
	MSSIM	0.755 ± 0.0001	0.754 ± 0.0002	0.777 ± 0.0002	0.818 ± 0.0001	$\mathbf{0.832\pm0.0001}$
	MSA (degree)	8.363 ± 0.0026	7.163 ± 0.0051	6.738 ± 0.0045	7.223 ± 0.0032	$\mathbf{5.676\pm0.0034}$
$\sigma=rand(25)$	MPSNR	28.725 ± 0.0011	30.064 ± 0.0025	31.812 ± 0.0033	30.775 ± 0.0028	$\mathbf{34.220\pm0.0018}$
	MSSIM	0.816 ± 0.0001	0.846 ± 0.0001	0.865 ± 0.0001	0.879 ± 0.0001	$\mathbf{0.931\pm0.0000}$
	MSA (degree)	6.731 ± 0.0015	6.160 ± 0.0025	5.623 ± 0.0026	5.684 ± 0.0016	$\mathbf{4.611\pm0.0017}$
$\sigma_{spectra}=Gau(150,15)$	MPSNR	30.323 ± 0.0015	47.351 ± 0.0026	50.932 ± 0.0037	33.846 ± 0.003	$\mathbf{51.025\pm0.0037}$
	MSSIM	0.860 ± 0.0001	0.924 ± 0.0001	0.935 ± 0.001	0.933 ± 0.0001	$\mathbf{0.951\pm0.0001}$
	MSA (degree)	5.712 ± 0.0024	5.450 ± 0.004	4.949 ± 0.0045	4.734 ± 0.0031	$\mathbf{4.427\pm0.0057}$

The Optimization Procedure

```
(1) Initialization and parameter estimation:
    (a) By projecting the noisy image into the wavelet domain, the diagonal detail coefficient matrix is used to compute an estimated noise standard deviation $\sigma_b$ for each band.
    (b) The highly correlated and similar bands are grouped together. For each group, the spectrally adaptive parameter $\lambda$ is computed.
    (c) Set $\boldsymbol{U}^0=\boldsymbol{F}$ and $\mathbf{P}=\mathbf{0}$, $\mathbf{Q}=\mathbf{0}$
(2) While $k<nIter$ and $\|\mathbf{U}^{k+1}-\mathbf{U}^k\|_F \geq$ tolerance do
    (a) Solve the $\boldsymbol{U}^{k+1}$ subproblem (20) by the Gauss–Seidel iteration algorithm.
    (b) Solve the $\mathbf{Q}^{k+1}$ subproblem (21) and update the parameter $\boldsymbol{P}^{k+1}$. The pixel $\mathbf{Q}_{bi}^{k+1}$ is obtained by applying the shrinkage operator for the vector field at each point $(b,i)$.
    (c) Update the spectrally adaptive parameter $\lambda^{k+1}$ using the equation in (17).
    End.
```

3. Experimental results and discussion

3.1. Hyperspectral image denoising

To demonstrate the performance of the proposed method, both simulated and real data were tested, which were provided by the Laboratory for Applications of

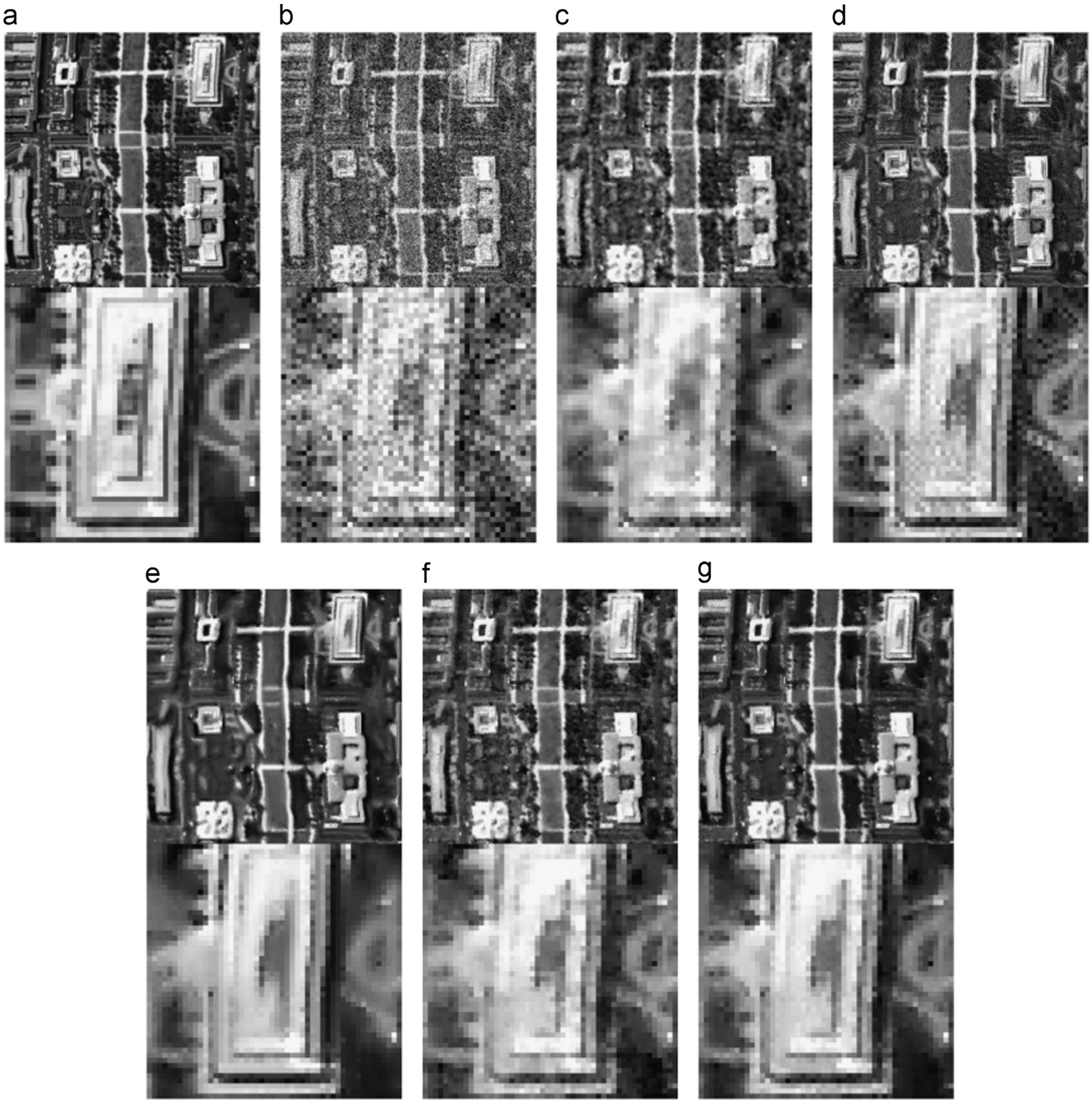

Fig. 4. Denoising results in simulated experiment Case 1 with noise level $\sigma=20$: (a) original band 32; (b) noisy band 32; (c) locally adaptive Wiener filter; (d) NLM; (e) SANLTV; (f) SSAHTV; and (g) the proposed method.

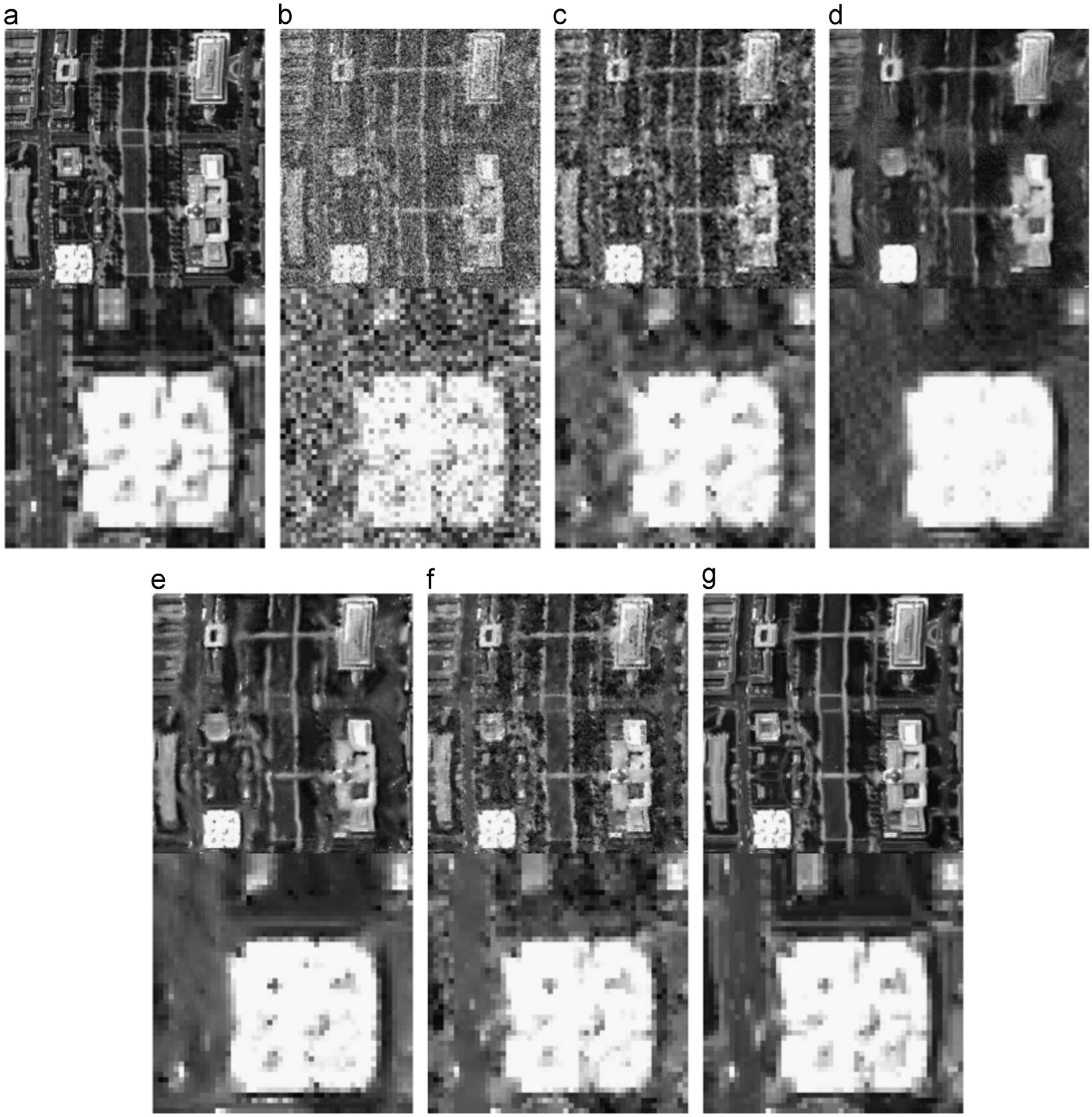

Fig. 5. Denoising results in simulated experiment Case 2 with noise level $\sigma = rand(25)$: (a) original band 1; (b) noisy band 1; (c) locally adaptive Wiener filter; (d) NLM; (e) SANLTV; (f) SSAHTV; and (g) the proposed method.

Remote Sensing, Purdue University. For the experiments, the denoising process involved a search window of 21×21 and a similarity square neighborhood of a size of 9×9 pixels from the noisy image. The denoising results were assessed by the average of the peak signal-to-noise ratio (PSNR) index and the average of the structural similarity index (SSIM) [44], which are abbreviated as MPSNR and MSSIM, respectively. To evaluate the spectral information of the results, the average of the spectral angle (MSA) was also introduced in the synthetic experiment. To obtain the satisfactory denoising results, the filtering parameter h_b was fixed at σ_b. By changing manually the parameter τ and μ, some experiments were used to analyze their variation tendency in Fig. 3. It is observed that the empirical parameter $\tau = 10$ and $\mu = 1/(7\sigma_b)$ can achieve the higher MPSNR and the lower MSA in the proposed method.

For the simulated data experiment, a Hyperspectral Digital Imagery Collection Experiment (HYDICE) airborne hyperspectral dataset from the Washington DC Mall was used, comprising 200 lines and 200 columns. A total of 191 bands of the DC Mall image were utilized to verify the performance of the proposed algorithm. Before the simulated process, the gray

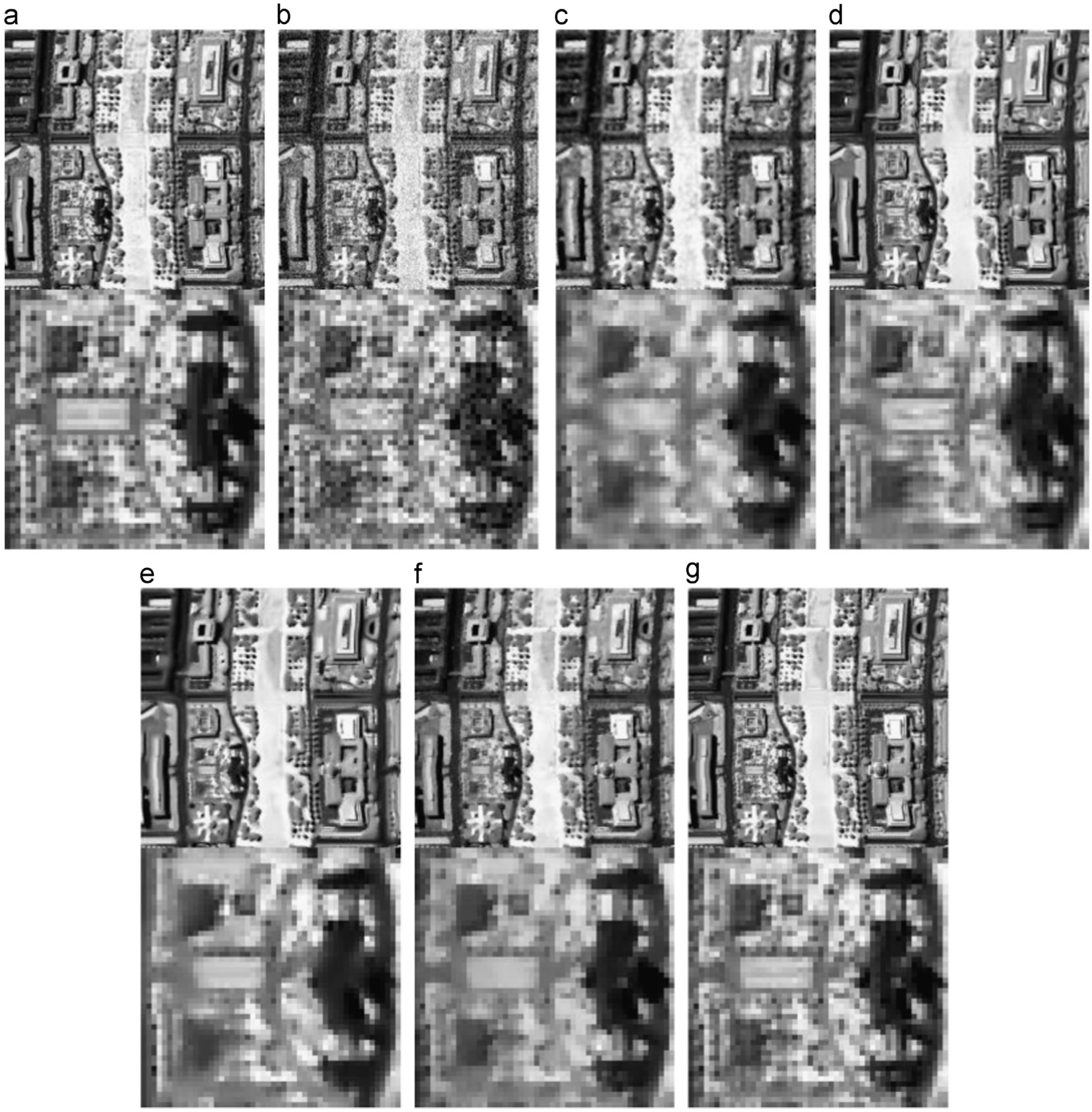

Fig. 6. Denoising results in simulated experiment Case 2 with noise level $\sigma = rand(25)$: (a) original band 93; (b) noisy band 93; (c) locally adaptive Wiener filter; (d) NLM; (e) SANLTV; (f) SSAHTV; and (g) the proposed method.

values of the HSI were normalized between [0, 255] to determine the range of the regularization parameter.

In the simulated process, we simulated the addition of noise in the following three cases: (1) for different bands, the noise-intensity is equal, and the same distribution of zero-mean Gaussian noise is added to all the bands; (2) for different bands, the noise-intensity is different, and different variance zero-mean Gaussian noise is added to different bands, with the standard deviation being randomly selected from 0 to 25. In this case, the noise distribution is represented as $\sigma = rand(25)$; (3) The noise variance σ_b^2 added along the spectral axis like a Gaussian shape centers a middle band $(B/2)$ [45] as

$$\sigma_b^2 = \beta^2 \frac{\exp\left\{-(b-B/2)^2/2\eta^2\right\}}{\sum_{b=1}^{B} \exp\left\{-(b-B/2)^2/2\eta^2\right\}} \tag{26}$$

where the power of the noise is controlled by β and η behaves like the standard deviation for the Gaussian curve. In the simulated experiment, $\beta = 150$, $\eta = 15$ and noise is defined as $\sigma_{spectra} = Gau(\beta, \eta)$.

For each noise-intensity level, the experiments were undertaken 10 times, and the mean and standard deviation of the 10 runs are given in Table 1. To verify the performance

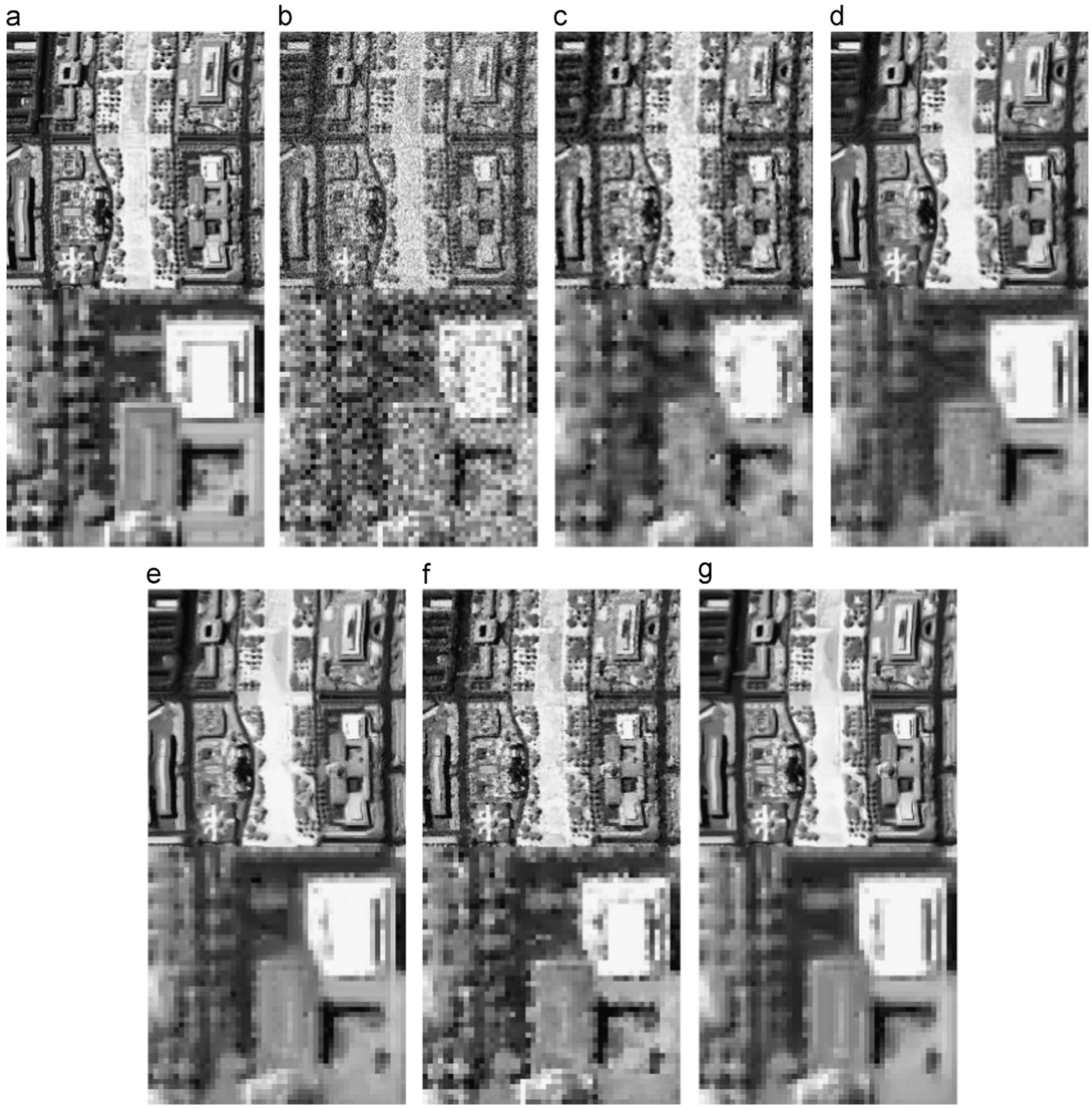

Fig. 7. Denoising results in simulated experiment Case 3: (a) original band 93; (b) noisy band 93; (c) locally adaptive Wiener filter; (d) NLM; (e) SANLTV; (f) SSAHTV; and (g) the proposed method.

of the proposed SAMNLTV model, it was compared with the locally adaptive Wiener filter [46], the NLM filter [29], the spectrally adaptive NLTV model (SANLTV) (in which we added the spectrally adaptive regularization parameter described in Section 2 for the NLTV model), and the SSAHTV model [18].

The contrasting results of the two cases with various noise levels are given in Table 1 by assessing the mean and standard deviation of the MPSNR, MSSIM, and MSA values. To give detailed contrasting results, $\sigma = 20$, $\sigma = rand(25)$ and $\sigma_{spectra} = Gau(150, 15)$ were selected to show the visual effect. Because of the large number of bands in the HSI, only a few bands are presented to give the visual results in each case. Fig. 4 shows the denoising results of the different methods in simulated Case 1, while Figs. 5 and 6 show the denoising results by the different methods in simulated Case 2. Fig. 7 shows the results for the simulated Case 2. In Figs. 8 and 9, the values of PSNR and SSIM from the different bands are presented to assess the denoising result of each band. The partially detail is enlarged to delineate more clearly the difference.

Compared with the other methods in Table 1, the denoising results of SAMNLTV are more robust, achieving the highest values of MPSNR and MSSIM and the lowest values of MSA

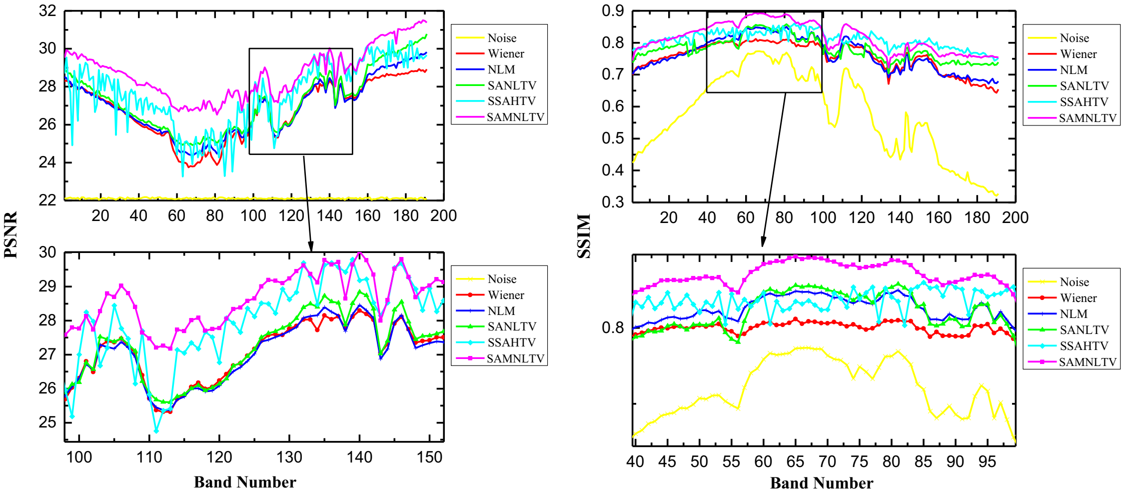

Fig. 8. PSNR and SSIM values of the different denoising approaches in each band of the simulated experiment Case 1 with noise level $\sigma = 20$.

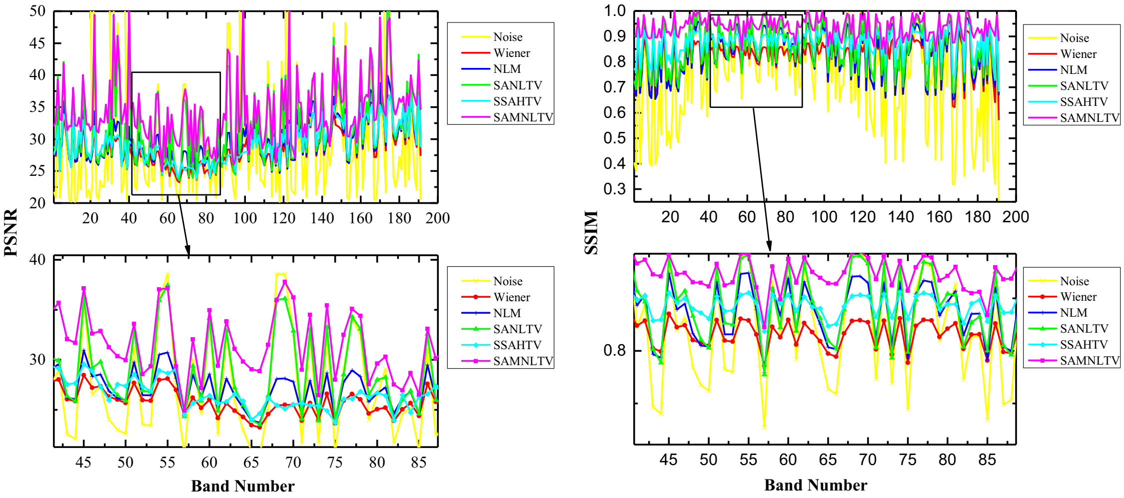

Fig. 9. PSNR and SSIM values of the different denoising approaches in each band of the simulated experiment Case 2 with noise level $\sigma = rand(25)$.

under all the different noise standard deviations. The lower MSA values demonstrate that the proposed method can reflect a better spectral fidelity. The low standard deviations prove that the proposed method can produce more stable results. In the visual comparison of the results with Case 1 from Fig. 4, it can be seen that the SAMNLTV-based method can not only suppress the noise more thoroughly than the other methods, but is also capable of preserving the edge information and the detailed information. In contrast, the result of the locally adaptive Wiener filter is oversmooth, and detailed information is lost. In the NLM-based method, it imports particles into the texture and produces obvious fake artifacts in the smooth regions. Compared with the proposed method, SANLTV generates some discrete error points and loses some detailed information in the texture, although the salient edges are preserved. The result using SSAHTV shows disturbing saw teeth in the edges. The reason for this is that the spatially adaptive based method retains both the spatial information and the residual noise.

In the simulated experiment Cases 2 and 3, different noise levels were added to different bands. In this case, we grouped the data by wavelength range and estimated the noise standard deviation. For example, noisy band 1 with $\sigma = 25$, as shown in Fig. 5, was divided into a group from band 1 to 33. The bands gathered in a group were regarded to be similar and could offered similar patches. In the group, the band with the lowest noise-intensity was regarded as the “indicative band”.

For the high-noise band in the Case 2, Fig. 5 reflects SSAHTV can preserve the edges well but produces some residual noise near the edges. Conversely, the proposed method can thoroughly reduce the noise, and the edges in the SAMNLTV result are not only sharper but are also tidier than SSAHTV, and more detailed information is preserved

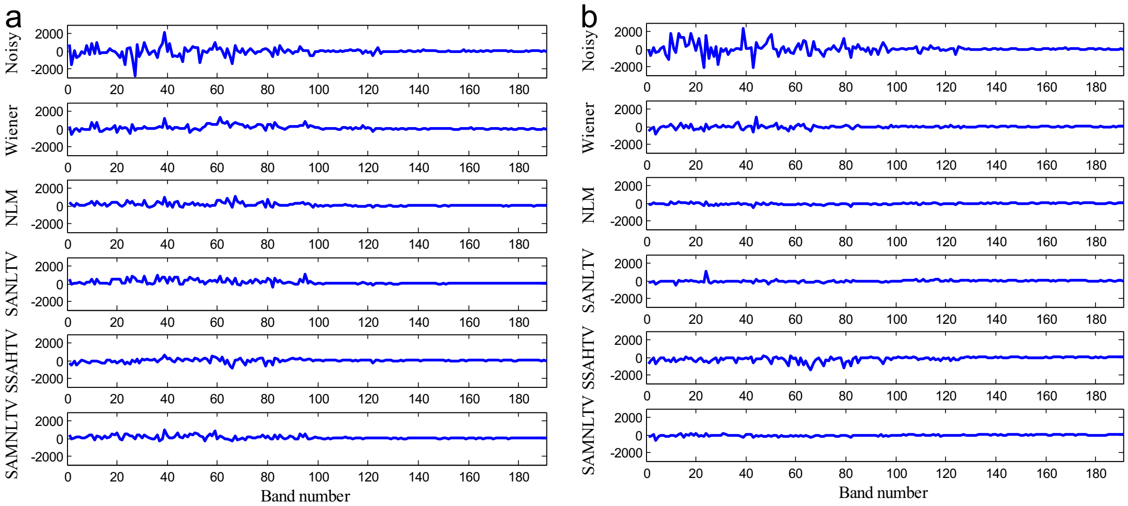

Fig. 10. Difference between the noise-free spectrum and the restoration results of (a) pixel (176, 75), which belongs to the grass class; and (b) pixel (99, 111), which belongs to the road class.

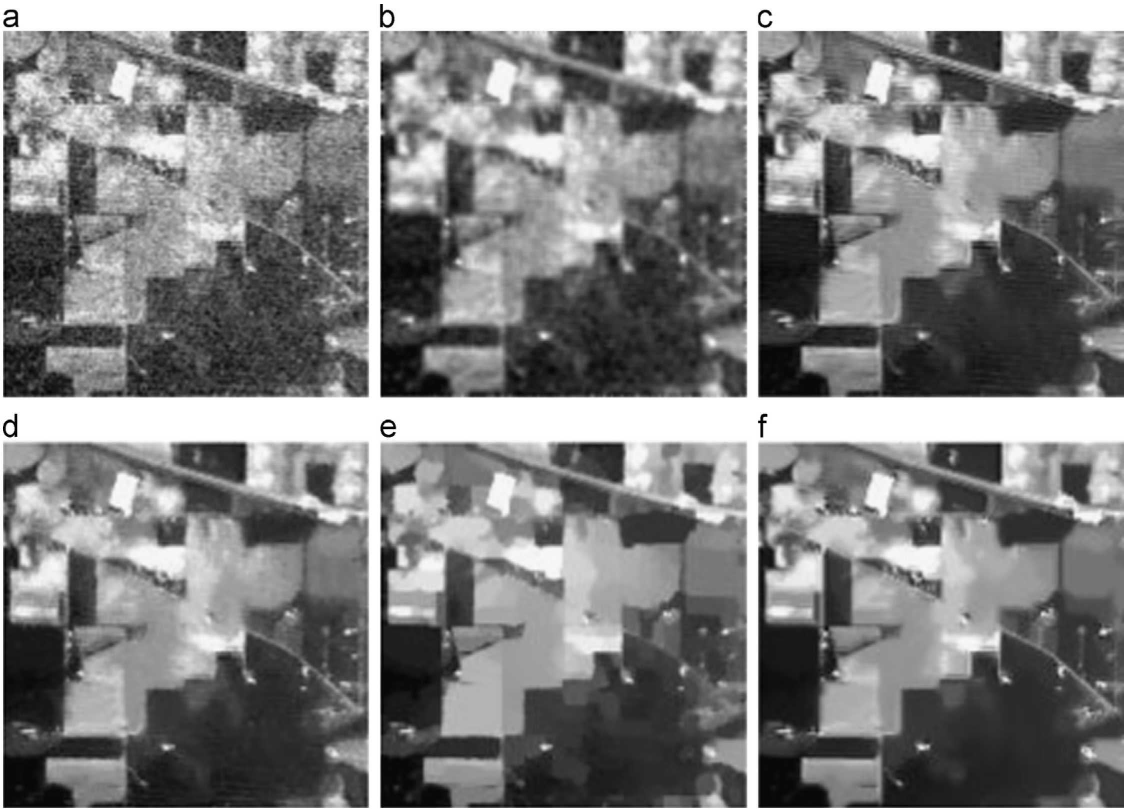

Fig. 11. Denoising results in the real data experiment: (a) noisy band 3; (b) locally adaptive Wiener filter; (c) NLM; (d) SANLTV; (e) SSAHTV; and (f) the proposed method.

than with SANLTV. From Fig. 6 with the low-noise level $\sigma_{93} = 13$, it can be seen that SSAHTV makes the edges oversmooth, while the discrete detail is retained.

For the Case 3, the visual effects in Fig. 7 demonstrate that the proposed method preserves the edges better than the SANLTV and SSAHTV which still leave some residual noises. As shown in Table 1, the proposed method also obtains the best performance for the quantitative evaluation. Although the SANLTV also can gain high values for the spatial information, the MSA is lower than SSAHTV and the proposed method because the SANLTV ignores the spectral information.

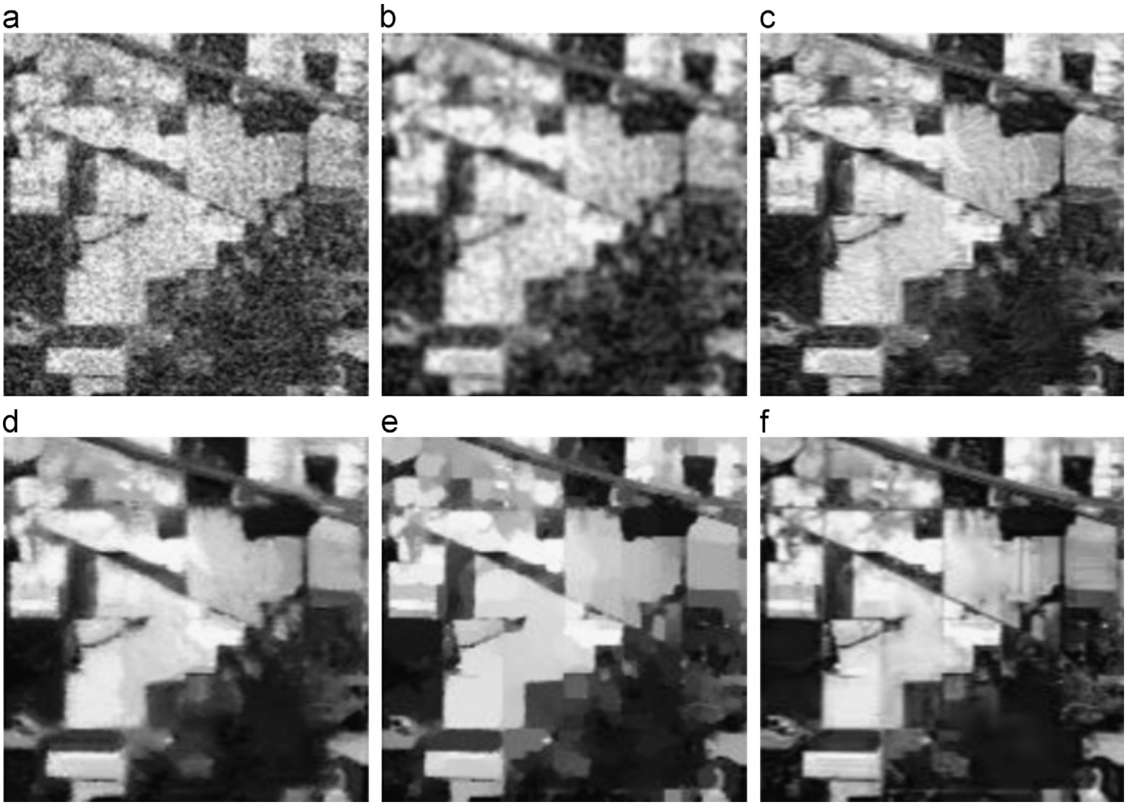

Fig. 12. Denoising results in the real data experiment: (a) noisy band 110; (b) locally adaptive Wiener filter; (c) NLM; (d) SANLTV; (e) SSAHTV; and (f) the proposed method.

The quantitative assessment results are presented in Figs. 8 and 9. Here, it is clear that the proposed method achieves higher PSNR and SSIM values than the other four methods in most of the bands, although there are a few failures in Case 2. The reason for this can be explained as follows. In the same group, the spatial information between bands is similar because of the high band correlation. However, the spectral noise differences can make the weight inaccurate when the weight is computed with patches from different bands. To decrease the error of the weight, the "indicative band" is introduced to choose the appropriate patch. In spite of these failures, the proposed method is still more effective than the other methods, and the advantage is more obvious than with SSAHTV for the low-noise band. From the point of view of the spectral analysis, we compared the spectral difference between the noise-free spectrum and the restoration results of the pixel to assess the hyperspectral denoising algorithm. The differences in the spectral signatures of pixel (176, 75) from the grass class and pixel (99, 111) from the road class are presented in Fig. 10(a) and (b), respectively. In Fig. 10, the vertical axis of the figures represents the digital number (*DN*) values, and the horizontal axis shows the spectral band number. Here, it is clear that the curve of the proposed method is smoother than for the other methods, indicating a lower spectral difference and reduced loss of spectral information.

In the real data experiment, the Airborne Visible Infrared Imaging Spectrometer (AVIRIS) Indian Pines test dataset was used to demonstrate the performance of the proposed method. The data size is 145 × 145 pixels, with 220 bands. Before the denoising processing, the atmospheric and water absorption bands from bands 150 to 163 were removed from the original HSI. Therefore, there were only 206 bands used in this real data experiment. The visual results of bands 3 and 110 are presented in Figs. 11 and 12, respectively, to allow a comparison, and the combination of bands 3, 110, and 196 is shown in Fig. 13. From Figs. 11 and 12, it can be clearly seen that the proposed SAMNLTV algorithm can give better denoising results than the other four methods. The locally adaptive Wiener filter fails to suppress the noise and disturbs the texture of the original image. NLM generates some fake artifacts and fails to remove the striping noise in band 3. Although SANLTV can obtain better denoising results than NLM, the edges appear oversmoothed and a lot of information is lost. With the process of the SSAHTV algorithm, the noise in the smooth areas is suppressed and the edges are sharp; however, considerable and important detailed information is lost. With the proposed SAMNLTV algorithm, it is apparent that the noise can be effectively removed while the edges and small textural features are well maintained.

3.2. *Hyperspectral image inpainting*

To further verify the effectiveness of the proposed approach, a real Hyperion image was also used. The data size is 200 × 200 pixels, with 155 bands. The peculiarity of this data is that it not only includes random noise, but also

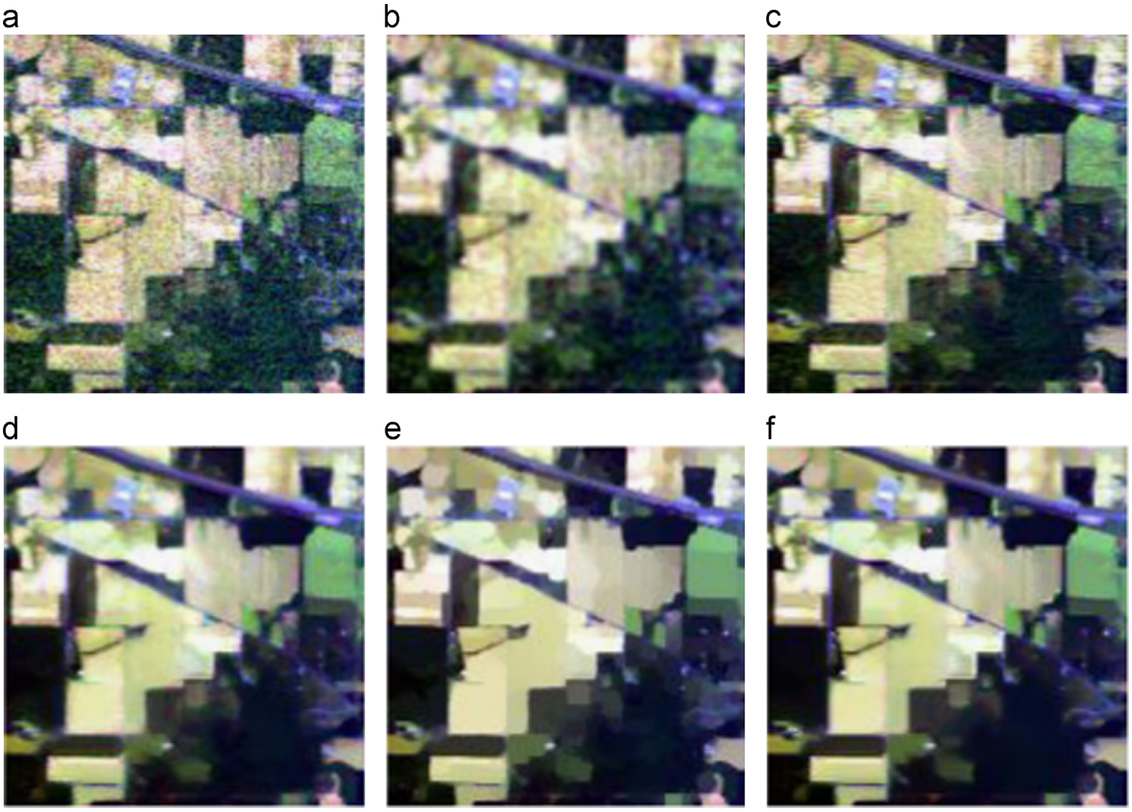

Fig. 13. Denoising results in the real data experiment: (a) noisy bands (196, 110, 3); (b) locally adaptive Wiener filter; (c) NLM; (d) SANLTV; (e) SSAHTV; and (f) the proposed method.

contain some deadlines (Figs. 14(a) and 15(a)), Therefore, we should deal with the denoising and inpainting problem simultaneously.

For this problem, the degradation model (1) can be changed to $\boldsymbol{f}_b = \boldsymbol{A}_b \boldsymbol{u}_b + \boldsymbol{v}_b$. $\boldsymbol{A}_b \in \mathbb{R}^{M_1 M_2 \times M_1 M_2}$ is the set of the diagonal matrix with diagonal elements consisting of 0 and 1, in which 0 represents the missing pixels. To verify that the proposed method can effectively work on both denoising and inpainting problems, we compared the results to the ones obtained by the morphological component analysis (MCA) method [47], the NLTV inpainting algorithm [27], and the multichannel NLTV algorithm [41]. As shown in Figs. 14 and 15, the result using the MCA algorithm has a strong ripple effect. For NLTV and the multichannel NLTV algorithm, the recovery of stripe noise and dead pixels is spatially discontinuous. The result of the NLTV algorithm is not as sharp as the result of the proposed algorithm. By extracting information from the other bands, the proposed method can effectively preserve the edge structures and suppress the irregular particles. The recovery of the spatial information is also continuous, with a more convincing visual result. Fig. 16 shows the spectral signatures of pixel (67, 100). The elliptic region represents a horizontal dead line. Here, it can be observed that the proposed method remedies the loss of spectral information and produces better spectral signatures than the other restoration methods, when compared with the original spectrum.

3.3. Discussion

3.3.1. The effectiveness of the spectrally adaptive idea

To verify the effectiveness of the spectrally adaptive idea in (17), we compared it with the spectrally adaptive idea in reference [18]. The proposed strategy is denoted as mode-1, and the idea in [18] is denoted as mode-2, which has the following expression:

$$\lambda_{b-\mathrm{mod\ e}\ 2} = \frac{\tau |\nabla \boldsymbol{u}_b|}{\sqrt{\sum_{b=1}^{B} |\nabla \boldsymbol{u}_b|^2}} \tag{27}$$

where $|\nabla \boldsymbol{u}_b|$ is the norm of the gradient corresponding to the horizontal and vertical first-order differences.

In the simulation, different noise-intensities were added in different bands. The performance of the λ values from the two different regularization problems is described in Fig. 17(a) and (b). The PSNR values of the three different combinations (mode-1+MNLTV, mode-2+MNLTV, and SSAHTV) are listed in Fig. 17(c).

In Fig. 17(a), the lower wave trough shows that for the low-noise bands, the proposed method can dramatically decrease the value of λ_b from the result in [18], so that the detailed information is also well preserved. In Fig. 17(b), the enlarged rectangular region shows that for the maximum noise level bands, λ_b in the proposed method is larger than in [18]. From Fig. 17(c), it can be observed that the mode-1+MNLTV gives a higher PSNR value for the

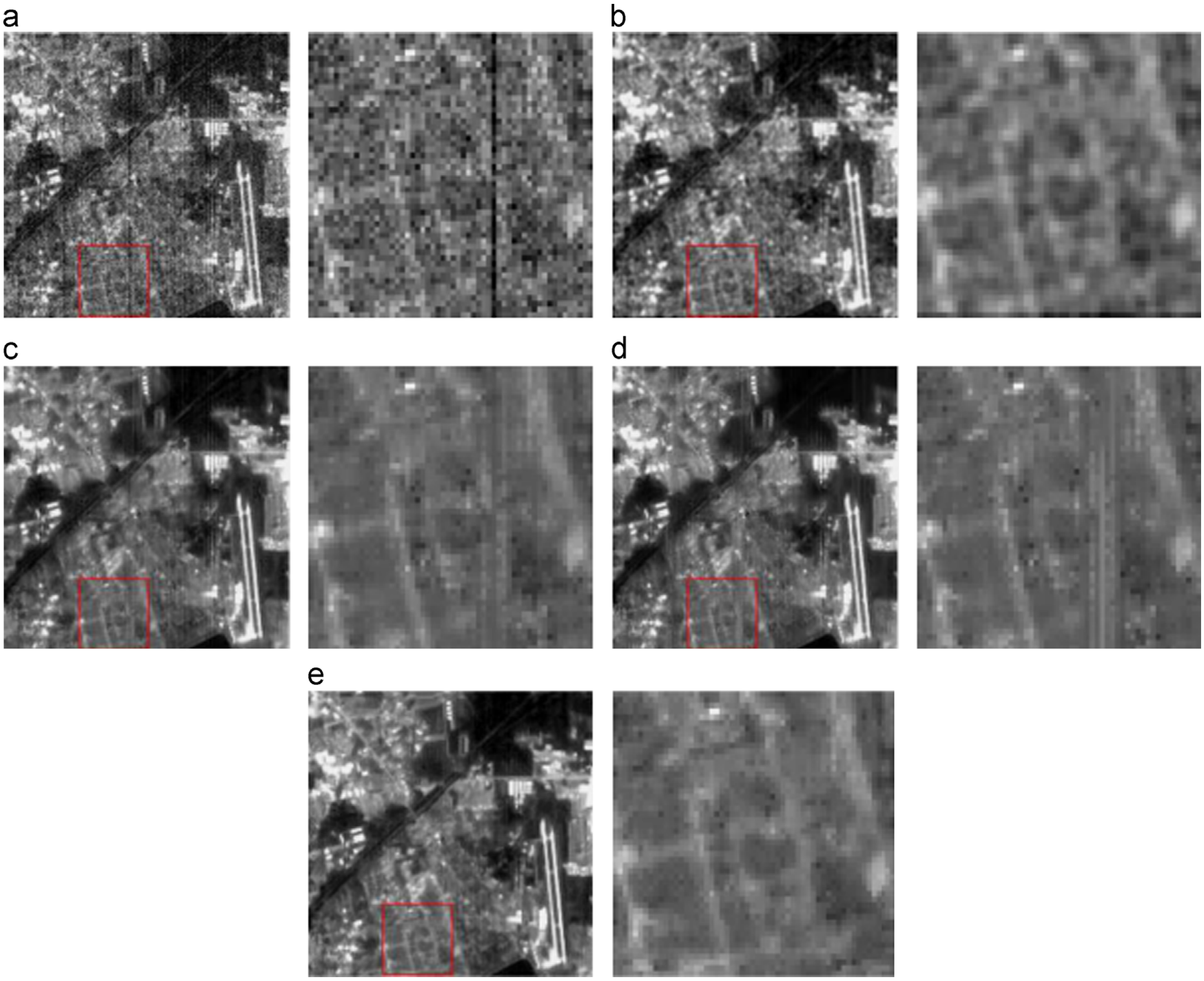

Fig. 14. Inpainting results in the real data experiment: (a) original band 2; (b) the MCA algorithm; (c) NLTV; (d) multichannel NLTV; and (e) the proposed method.

optimum value of the tuning parameters. Therefore, the spectrally adaptive λ_b in the proposed method can satisfactorily reflect the variation of the noise-intensity in the denoising process, and it performs better than the adaptive regularization parameter in [18].

3.3.2. *Sensitivity analysis for "NS" (the number of utilized bands)*

In order to capture more effective information, the structurally similar bands are selected to extract patches which can provide more similar textures. To analyze the impact of "*NS*", different values of "*NS*" were defined in the simulated experiment. From Fig. 18, it is demonstrated that the results of denoising from the multiple bands achieve higher MPSNR and MSSIM values than the results from one band. However, when the number of utilized bands exceeds a certain level, the result of the accuracy assessment gradually declines, and the curve is close to a straight line when the number finally reaches a certain value. The experimental results show that a band number from three to five bands is enough to achieve a satisfactory performance.

3.3.3. *Sensitivity analysis for the "indicative band"*

As described in Fig. 2, to accurately extract the similar patches from the other bands, and to cut down the time of the weight calculation, we use the low-noise band as the "indicative band". To show its effect on the denoising performance, we used the simulated experiment as an example. In the experiment, band 31 from the Washington DC Mall dataset was used to research the relationship between the acquired PSNR and SSIM values and the "indicative band". In Fig. 19, the "indicative band" with different noise standard deviations is provided to show the variation in the results. With band 31 as an explanation of the experimental operation, the similar bands 30 and 32 were used to offer the redundant information, while band 32 was chosen as the "indicative band". The noise standard deviation of band 32 was then changed from 1 to 35. It can be seen that the PSNR and SSIM values using the

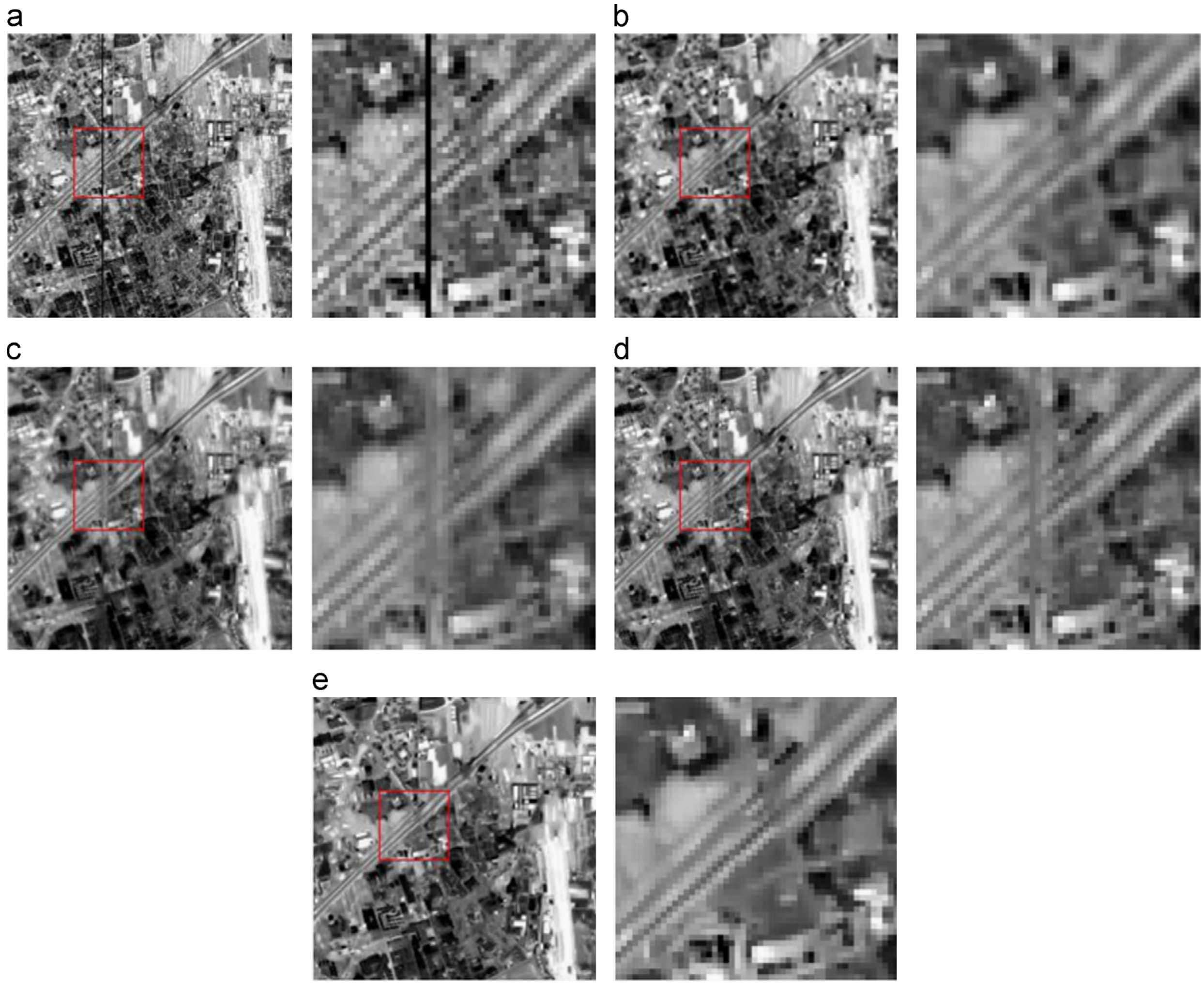

Fig. 15. Inpainting results in the real data experiment: (a) original band 63; (b) the MCA algorithm; (c) NLTV; (d) multichannel NLTV; and (e) the proposed method.

"indicative band" decrease gradually with the increase in the noise-intensity in the "indicative band". This phenomenon can be explained as follows: There is a high correlation between the selected bands. When the "indicative band" has a lower noise-intensity, the difference between the recorded position and the ideal position of the original band is less, so that the result may be closer to the original image; otherwise, using the "indicative band" with a higher noise-intensity, the search for similar patches will encounter a greater deviation as a result of the loss of denoising precision.

4. Conclusion

In this paper, we have established a multidimensional nonlocal image restoration scheme, which can simultaneously consider the similarity from the spatial and spectral dimensions and can adaptively control the denoising strengths for different bands. Meanwhile, a band selection strategy is proposed to improve the value of the weight, in which the highly correlated bands are clustered into a group, and we select the lowest noise-intensity band as the indicative band to obtain an index for the order of the similar patches. Several simulated and real data sets were employed in experiments presented in Section 3 to illustrate that the proposed algorithm can obtain satisfactory results. The proposed method produces denoising results with the noise effectively suppressed and the edges preserved. The adaptive selection of the denoising parameter also satisfactorily balances the denoising strengths of different hyperspectral bands. Furthermore, the proposed method also works well on the mixed-problem with both noise and dead pixels.

Although the proposed model works well in the HSI recovery problem, there are still several research directions that could improve this work. Because of the complexity of noise, the first issue to be considered is to construct a more reliable algorithm to estimate the noise strength for each band. Additionally, the weight between patches will be computed inaccurately when the two compared patches are extracted from different bands with

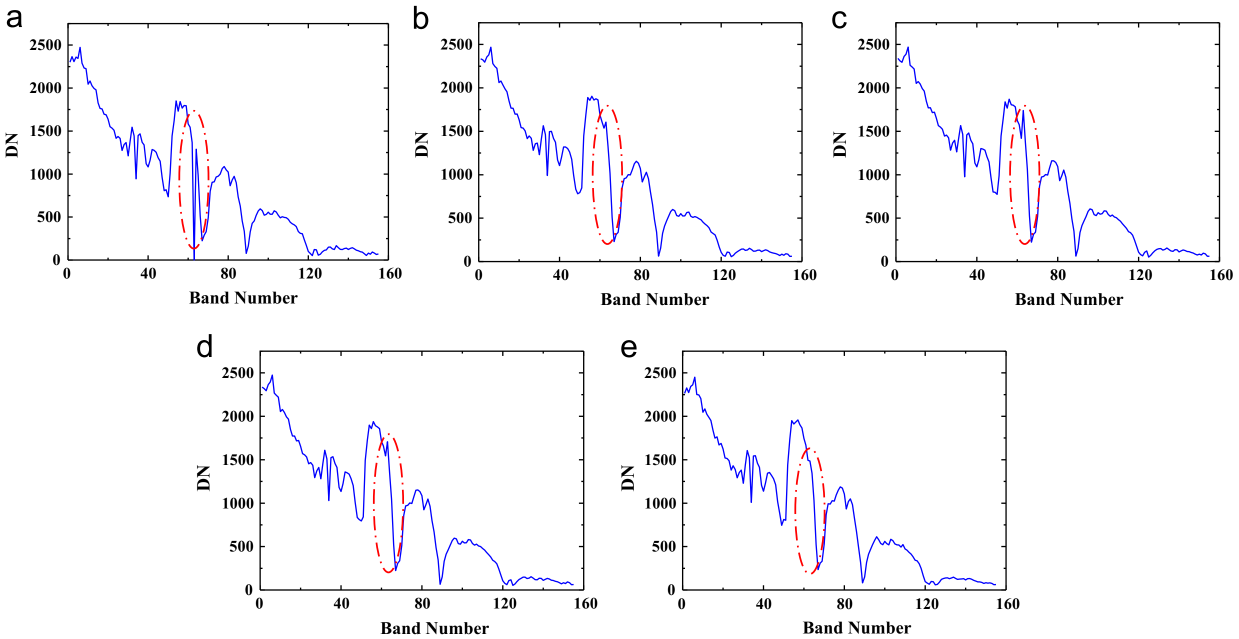

Fig. 16. Spectrum of pixel (67, 100) in the restoration results: (a) original image; (b) the MCA algorithm; (c) NLTV; (d) multichannel NLTV; and (e) the proposed method.

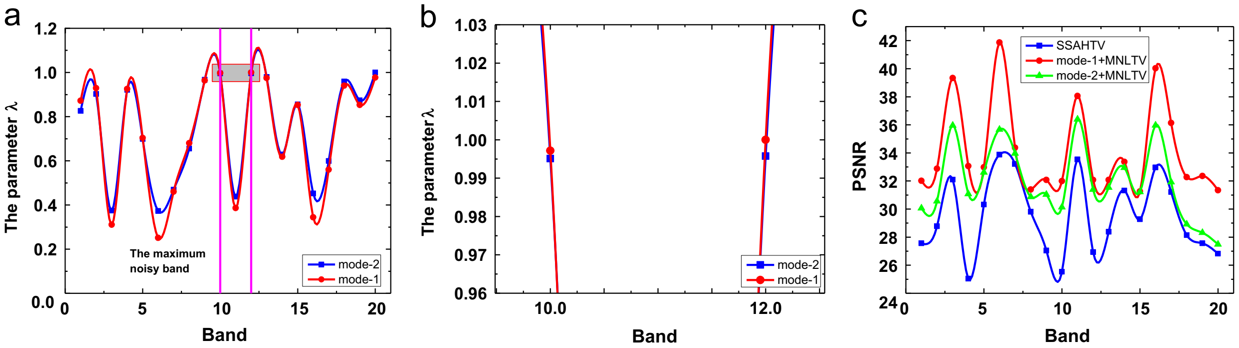

Fig. 17. Analysis of the two tuning parameters λ: (a) and (b) reflect the tendency of the performance of the optimal λ values from SAMNLTV and [18]; and (c) gives the PSNR values from three combinations: mode-1+MNLV, mode-2+MNLTV, and SSAHTV.

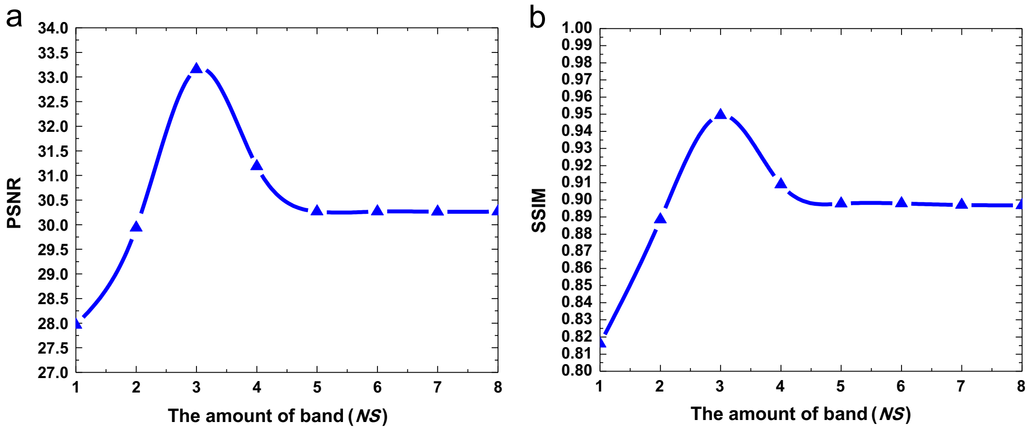

Fig. 18. The impact of "*NS*": (a) the change in the MPSNR value with "*NS*"; and (b) the change in the MSSIM value with "*NS*".

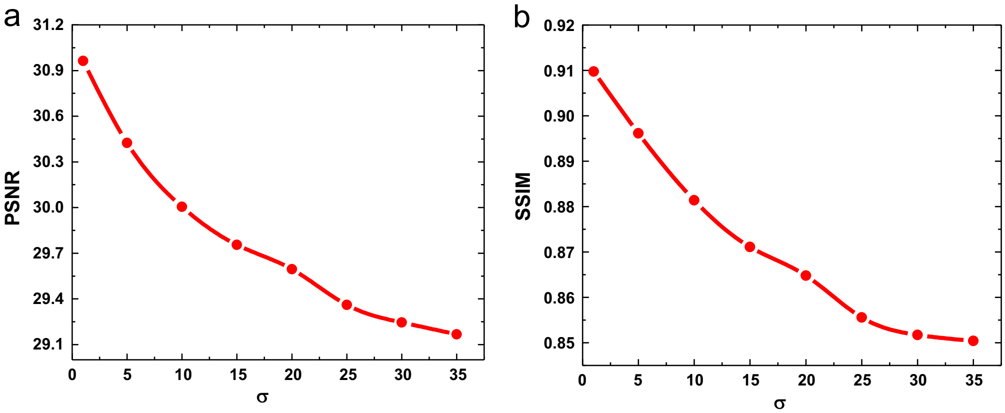

Fig. 19. Sensitivity analysis for the "indicative band" (σ from 1 to 35): (a) the change in the MPSNR value with the "indicative band"; and (b) the change in the MSSIM value with the "indicative band".

different noise intensities. To improve the results, we believe that it will be possible to use an adaptive Gaussian kernel to compute a precise weight between patches. Finally, the patch-based methods need more memory space and computational time. The acceleration of the algorithm of the proposed model will therefore be one of the important subjects in the future.

Acknowledgments

The authors would also like to thank the editors and anonymous reviewers for their careful reading and helpful suggestions. This work is supported by the National Natural Science Foundation of China under (41401383, 91338111 and 41422108), Program for Changjiang Scholars and Innovative Research Team in University (IRT1278).

References

[1] D. Manolakis, D. Mardon, G.A. Shaw, Hyperspectral image processing for automatic target detection applications, Linc. Lab. J. 14 (1) (2003) 79–115.
[2] I. Atkinson, F. Kamalabadi, D.L. Jones, Wavelet-based hyperspectral image estimation, in: Proceedings of the IGARSS, Toulouse, France, July 2003, pp. 743–745.
[3] G. Chen, S. Qian, Simultaneous dimensionality reduction and denoising of hyperspectral imagery using bivariate wavelet shrinking and principal component analysis, Can. J. Remote Sens. 34 (5) (2008) 447–454.
[4] G. Chen, S. Qian, W. Xie, Denoising of hyperspectral imagery using principal component analysis and wavelet shrinkage, IEEE Trans. Geosci. Remote Sens. 49 (3) (2011) 973–980.
[5] P. Scheunders, Wavelet thresholding of multivalued images, IEEE Trans. Image Process. 13 (4) (2004) 475–483.
[6] H. Othman, S. Qian, Noise reduction of hyperspectral imagery using hybrid spatial–spectral derivative-domain wavelet shrinkage, IEEE Trans. Geosci. Remote Sens. 44 (2) (2006) 397–408.
[7] M.S. Crouse, R.D. Nowak, R.G. Baraniuk, Wavelet-based signal processing using hidden Markov models, IEEE Trans. Signal Process. 46 (4) (1998) 886–902.
[8] J. Portilla, V. Strela, M.J. Wainwright, E.P. Simoncelli, Image denoising using scale mixtures of Gaussians in the wavelet domain, IEEE Trans. Image Process. 12 (11) (2003) 1338–1351.
[9] P. Scheunders, S.D. Backer, Wavelet denoising of multicomponent images using Gaussian scale mixture models and a noise-free image as priors, IEEE Trans. Image Process. 16 (7) (2007) 1865–1872.
[10] Y.F. Zhang, S. De Backer, P. Scheunders, Noise-resistant wavelet-based Bayesian fusion of multispectral and hyperspectral images, IEEE Trans. Geosci. Remote Sens. 47 (2009) 3834–3843.
[11] C. Liu, R. Szeliski, S.B. Kang, C.L. Zitnick, W.T. Freeman, Automatic estimation and removal of noise from a single image, IEEE Trans. Pattern Anal. Mach. Intell. 30 (2) (2008) 299–314.
[12] D. Letexier, S. Bourennane, Noise removal from hyperspectral images by multidimensional filtering, IEEE Trans. Geosci. Remote Sens. 46 (7) (2008) 2061–2069.
[13] N. Renard, S. Bourennane, J. Blanc-Talon, Denoising and dimensionality reduction using multilinear tools for hyperspectral images, IEEE Geosci. Remote Sens. Lett. 5 (2) (2008) 138–142.
[14] X. Liu, S. Bourennane, C. Fossati, Denoising of hyperspectral images using the parafac model and statistical performance analysis, IEEE Trans. Geosci. Remote Sens. 50 (10) (2012) 3717–3724.
[15] A. Karami, M. Yazdi, A.Z. Asli, Noise reduction of hyperspectral images using kernel non-negative tucker decomposition, IEEE J. Sel. Top. Signal Process. 5 (3) (2011) 487–493.
[16] S. Bourennane, C. Fossati, A. Cailly, Improvement of target-detection algorithms based on adaptive three-dimensional filtering, IEEE Trans. Geosci. Remote Sens. 49 (4) (2011) 1383–1395.
[17] Y. Wang, R. Niu, X. Yu, Anisotropic diffusion for hyperspectral imagery enhancement, IEEE Sens. J. 10 (3) (2010) 469–477.
[18] Q. Yuan, L. Zhang, H. Shen, Hyperspectral image denoising employing a spectral–spatial adaptive total variation model, IEEE Trans. Geosci. Remote Sens. 50 (10) (2012) 3660–3677.
[19] S. Chen, X. Hu, S. Peng., Hyperspectral imagery denoising using a spatial–spectral domain mixing prior, J. Comput. Sci. Technol. 27 (4) (2012) 851–861.
[20] Y. Qian, M. Ye, Hyperspectral imagery restoration using nonlocal spectral–spatial structured sparse representation with noise estimation, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 6 (2) (2013) 499–515.
[21] B. Rasti, J.R. Sveinsson, M.O. Ulfarsson, J.A. Benediktsson, Hyperspectral image denoising using 3D wavelets, in: Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, IGARSS, 2012, pp. 1349–1352.
[22] Peng Liu, K.B. Eom, Compressive sensing of noisy multispectral images, IEEE Geosci. Remote Sens. Lett. 11 (11) (2014) 1931–1935.
[23] Peng Liu, Fang Huang, Guoqing Li, Zhiwen Liu, Remote-sensing image denoising using partial differential equations and auxiliary images as priors, IEEE Geosci. Remote Sens. Lett. 9 (3) (2012) 358–362.
[24] A. Tikhonov, V. Arsenin, Solutions of Ill-Posed Problems, W.H. Winston & Sons, Washington, DC, 1977.
[25] A. Rangarajan, R. Chellappa, Markov random field models in image processing, in: M. Arbib (Ed.), The Handbook of Brain Theory and Neural Networks, 1995, 564–567.
[26] Y. Chen, Z. Cao, An improved MRF-based change detection approach for multitemporal remote sensing imagery, Signal Process. 24 (1) (2013) 163–175.
[27] X.Q. Zhang, T.F. Chan, Wavelet inpainting by nonlocal total variation, Inverse Probl. Imaging 4 (1) (2010) 191–210.
[28] J. Salmon, Y. Strozecki, Patch reprojections for nonlocal methods, Signal Process. 92 (2) (2012) 447–489.

[29] A. Buades, B. Coll, J.M. Morel, A nonlocal algorithm for image denoising, in: Proceedings of the IEEE CVPR, vol. 2, 2005, pp. 60–65.

[30] B.H. Traceya, E.L. Miller, A constrained optimization approach to combining multiple nonlocal means denoising estimates, Signal Process. 103 (2014) 60–80.

[31] M. Protter, M. Elad, H. Takeda, P. Milanfar, Generalizing the nonlocal-means to super-resolution reconstruction, IEEE Trans. Image Process. 18 (1) (2009) 36–51.

[32] A. Wong, J. Orchard, A nonlocal-means approach to exemplar-based inpainting, in: Proceedings of the IEEE International Conference on Image Processing, 2008, pp. 2600–2603.

[33] H. Li, L. Zhang, H. Shen, An adaptive nonlocal regularized shadow removal method for aerial remote sensing images, IEEE Trans. Geosci. Remote Sens. 52 (1) (2013) 106–120.

[34] A. Buades, B. Coll, J.M. Morel, A review of image denoising algorithms, with a new one, Multiscale Model. Simul. 4 (2) (2005) 490–530.

[35] L. Yaroslavsky, Digital Picture Processing – An Introduction, 2nd ed., Soviet Radio, Moscow, 1985.

[36] G. Gilboa, S. Osher, Nonlocal operators with applications to image processing, Multiscale Model. Simul. 7 (3) (2008) 1005–1028.

[37] P. Blomgren, T.F. Chan, Color TV: total variation methods for restoration of vector-valued images, IEEE Trans. Image Process. 7 (3) (1998) 304–309.

[38] Q. Cheng, H. Shen, L. Zhang, P. Li, Inpainting for remotely sensed images with a multichannel nonlocal total variation model, IEEE Trans. Geosci. Remote Sens. 52 (1) (2014) 175–187.

[39] T. Akgun, Y. Altunbasak, R.M. Mersereau, Super-resolution reconstruction of hyperspectral images, IEEE Trans. Image Process. 14 (11) (2005) 1860–1875.

[40] E. Luo, S.H. Chan, T.Q. Nguyen, Image denoising by targeted external databases, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2014, pp. 3019–3023.

[41] I. Mosseri, M. Zontak, M. Irani, Combining the power of internal and external denoising, in: Proceedings of the ICCP, 2013.

[42] D.L. Donoho, I.M. Johnstone, Ideal spatial adaptation via wavelet shrinkage, Biometrika 81 (1994) 425–455.

[43] X. Zhang, M. Burger, X. Bresson, S. Osher, Bregmanized nonlocal regularization for deconvolution and sparse reconstruction, SIAM J. Imaging Sci. 3 (3) (2009) 253–276.

[44] Z. Wang, A.C. Bovik, H.R. Sheikh, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.

[45] J.M. Bioucas-Dias, J.M.P. Nascimento, Hyperspectral subspace identification, IEEE Trans. Geosci. Remote Sens. 46 (8) (2008) 2435–2445.

[46] Available Online: ⟨http://www.mathworks.cn/help/toolbox/images/ref/wiener2.htm⟩.

[47] M. Elad, J.L. Starck, P. Querre, D.L. Donoho, Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA), Appl. Comput. Harmon. Anal. 19 (3) (2005) 340–358.