

HYPERSPECTRAL IMAGE DENOISING USING A NONLOCAL SPECTRAL SPATIAL PRINCIPAL COMPONENT ANALYSIS

Duidui Li¹, Linlin Xu^{1,*}, Junhuan Peng¹ Jianguo Ma²

¹Dept. of Land science and technology, China University of Geosciences, Xueyuan Road, Beijing, China
- ldd_cugb@163.com, beyond13031@126.com, pengjunhuan@163.com

²State grid hubei electric power.co.ltd, Hubei, China - majg@hb.sgcc.com.cn

Commission III, WG III/4

KEY WORDS: Hyperspectral Images, Noise Reduction, Nonlocal Similarity, Spectral Spatial Information, Principal Component Analysis

ABSTRACT:

Hyperspectral images (HSIs) denoising is a critical research area in image processing duo to its importance in improving the quality of HSIs, which has a negative impact on object detection and classification and so on. In this paper, we develop a noise reduction method based on principal component analysis (PCA) for hyperspectral imagery, which is dependent on the assumption that the noise can be removed by selecting the leading principal components. The main contribution of paper is to introduce the spectral spatial structure and nonlocal similarity of the HSIs into the PCA denoising model. PCA with spectral spatial structure can exploit spectral correlation and spatial correlation of HSI by using 3D blocks instead of 2D patches. Nonlocal similarity means the similarity between the referenced pixel and other pixels in nonlocal area, where Mahalanobis distance algorithm is used to estimate the spatial spectral similarity by calculating the distance in 3D blocks. The proposed method is tested on both simulated and real hyperspectral images, the results demonstrate that the proposed method is superior to several other popular methods in HSI denoising.

1. INTRODUCTION

Hyperspectral image (HSI) is produced with high spectral resolution, providing contiguous or noncontiguous bands throughout the 400-2500 nm region. HSI is capable of supporting various important application in the field of remote sensing (Wang and Niu, 2009), such as environmental monitoring, discriminating different land cover types, mineral identification and so on. However, the existence of noise changes the spectral curve of HSI, which has a negative impact on various HSI processing tasks, classification, unmixing, subpixel mapping, target detection, and so on. Therefore, how to reduce the noise influence in HSI is an essential step to improve the image quality.

In recent years, image denoising based on principal component analysis (PCA) model has been attracting more attention, and it has been proved that PCA algorithm is very effective and efficient denoising approach because PCA can separate the signal and noise well by converting the data into the PCA domain. Examples include the adaptive PCA denoising scheme (Muresan and Parks, 2003) and SAR image denoising via clustering-Based PCA (Xu et al., 2014). However, in hyperspectral denoising method, PCA denoising algorithm exists only as an auxiliary algorithm. Chen and Qian (Chen and Qian, 2008, Chen and Qian, 2009) proposed to perform dimension reduction and HSI denoising based on wavelet shrink and PCA. Chen *et al.* (Chen and Qian, 2011) proposed a new HSI denoising algorithm, where PCA is first used to decorrelate the data, and then wavelets are used to perform denoising in low energy in low-energy PCA output channels.

Traditionally, HSI denoising techniques are based on band-by-band or pixel-by-pixel processing, i.e., which lead the loss of

correlation between bands and pixels. In recent years, therefore, more and more HSI denoising algorithms tend to exploit the characteristics of high spatial correlation and spectral correlation of hyperspectral data. Yuan *et al.* (Yuan et al., 2012) has proposed spectral-spatial adaptive total-variation (TV) model for HSI denoising, which is capable of accounting for the noise intensity difference between different bands and spatial property differences between different pixels. Karami *et al.* (Karami et al., 2011) proposed genetic kernel Tucker decomposition (GKTD) algorithm for HSI denoising, which exploits both the spectral and spatial information in the image. Linlin xu (Xu et al., 2017) proposed a method using spatial spectral monte carlo sampling approach, which based on posterior probability in a nonparametric manner. Besides in field of HSI noise reduction, spectral-spatial joint structure is also used in HSI data compression (Christophe et al., 2008) and classification (Qian and Ye, 2013). In PCA-based HSI algorithms, although spectral PCA has been proposed for the segmentation of hyperspectral images (Yoshino, 2007), spatial domain information has not been well integrated into the PCA model. Therefore, establishing a PCA denoising model that combine spatial and spectral information is a necessary work.

Most noise reduction methods rely on the local information of signals, the main drawback of which is that the information provided by the neighborhood is too limited to preserve the true structure, details and texture of an image. To deal with this problem, nonlocal algorithm was proposed for image denoising (Buades et al., 2005). Nonlocal approach is based on the assumption that every pixel in image has many similar pixels in the same image. A nonlocal sparse representation based noise reduction algorithm is introduced (Dong et al., 2011), where sparse representation of the similar patches are recovered by a regularized linear regression model with shared constraint of sparsity. For HSI, nonlocality also exists in spectral-spatial space, which sug-

*Corresponding author

gests that the similarity must be considered in every 3D block in HSI. Therefore, the nonlocal similarity is not the distance between two points, but the similarity of two correlation matrix. The Mahalanobis distance (Jr et al., 2010) takes into account the correlation of the dataset and is scale invariant, i.e., not dependent on the scale of measurements. Therefore, Mahalanobis distance can calculate the nonlocal similarity in the spectral-spatial space. As a result, constructing a nonlocal spatial spectral correlation model base on PCA is crucial for hyperspectral denoising.

This paper, therefore, introduces a novel denoising approach base on nonlocal spectral spatial PCA (nonlocal SSPCA) for HSI denoising, which integrate spatial information and spectral information of nonlocal area into the PCA model. The novelty of nonlocal SSPCA lies in the following aspects: 1) separating the signal and noise in HSI only by PCA; 2) exploiting spectral-spatial joint structure of the HSI in PCA; 3) incorporating nonlocality of the HSI into PCA denoising model by calculating the Mahalanobis distance among pixels.

This paper is originated as follows. Section II introduces the proposed denoising framework, as well as the PCA and the nonlocal similarity approaches. Section III introduces the experiments results on both simulated and real hyperspectral images. Section IV concludes the study.

2. METHODOLOGY

2.1 Problem Formulation

The HSI as a collection of all of spatial position and all of the bands is represented by I_{jkb} ($j = 1, 2, 3, \dots, m$), ($k = 1, 2, 3, \dots, n$), ($b = 1, 2, 3, \dots, p$), where j and k determine the location of I_{jkb} in image space, and b represents the band number in spectral domain. $I_{jk\cdot}$ is a $p \times 1$ vector, representing the spectrum curve in the position of (j, k) , and $I_{\cdot\cdot b}$ is a $z \times 1$ ($z = m \times n$) vector, representing the all the pixels in the b th band, where I is a HSI that is contaminated by noise. The noise degradation model of the HSI can be written as (Yuan et al., 2012) (Atkinson et al., 2003)

$$I_{jkb} = X_{jkb} + N_b \quad (1)$$

where X_{jkb} denotes the unobservable noise-free variable and n_b is Gaussian-white noise In proposed PCA denoising model, because I_{jkb} cannot reflect the spatial correlation and spectral correlation of HSI. we use construct Y to entail the spatial spectral information, and use Y as the input to the PCA model for spatial spectral denoising of HSI.

$$Y = [y_1, y_2, \dots, y_z]^T \quad (2)$$

the size of y_i is $t \times 1$ vector. And from (1) we can get

$$I \implies Y = X + N \quad (3)$$

Now the problem is transformed to the estimation of X from the noisy measurement Y .

2.2 HSI denoising in PCA domain

In section 2.1, the problem have been turned into how to denoise the matrix of Y by PCA, The goal of PCA is to find an orthonormal transformation the matrix of P . However, in the process of data centralization, in practice, u is calculated from samples of Y , but not X . But, zero mean noise characteristics dictate that

the mean vector of y_i is the same as x_i , i.e., $E[y] = E[x] = u$. Therefore, we subtract the mean value of u from Y to get the centralized dataset of \bar{Y}

$$\bar{Y} = Y - u = \bar{X} + N \quad (4)$$

the optimal PCA transformation matrix of P can be obtained by computing the covariance matrix of $\Sigma_{\bar{X}}$ of \bar{X} . However, the available dataset \bar{Y} is contaminated by noise so that $\Sigma_{\bar{X}}$ can not be directly computed. Therefore, we need to estimate it by using the linear noise model, which can be expressed as

$$\begin{aligned} \Sigma_{\bar{Y}} &= E[(y - E[y])(y - E[y])^T] \\ &\approx \frac{1}{z} \bar{Y} \bar{Y}^T \\ &= \frac{1}{z} (\bar{X} \bar{X}^T + \bar{X} N^T + N \bar{X}^T + N N^T) \end{aligned} \quad (5)$$

the signal $\Sigma_{\bar{X}}$ and noise N are uncorrelated. Therefore, the $\bar{X} N^T$ and $N \bar{X}^T$ will be nearly zero matrices, we can get

$$\Sigma_{\bar{Y}} \approx \frac{1}{z} (\bar{X} \bar{X}^T + N N^T) = \Sigma_{\bar{X}} + \Sigma_N \quad (6)$$

where $\Sigma_{\bar{X}}$ and Σ_N are the covariance matrices, respectively. Since Gaussian noise is uncorrelated, we can know that Σ_N is a $t \times t$ diagonal matrix with all the diagonal components being σ_N^2 . Therefore, we can get

$$\Sigma_N = \Lambda_N = W_{\bar{X}} (\sigma_N^2 I) W_{\bar{X}}^T = W_{\bar{X}} \Lambda_N W_{\bar{X}}^T \quad (7)$$

Thus we have

$$\begin{aligned} \Sigma_{\bar{Y}} &= \Sigma_{\bar{X}} + \Sigma_N \\ &= W_{\bar{X}} \Lambda_{\bar{X}} W_{\bar{X}}^T + W_{\bar{X}} (\sigma_N^2 I) W_{\bar{X}}^T \\ &= W_{\bar{X}} \Lambda_{\bar{Y}} W_{\bar{X}}^T \end{aligned} \quad (8)$$

where I is identity matrix. We can prove that the PCA transformation matrix of P associated with $\Sigma_{\bar{X}}$ is the same as the PCA transformation matrix associated with $\Sigma_{\bar{Y}}$. Therefore, in PCA denoising model, we can directly get $W_{\bar{X}}$ by decomposing $\Sigma_{\bar{Y}}$ and use $W_{\bar{Y}}$ to replace the $W_{\bar{X}}$. Through the above analysis, we can directly decompose $\Sigma_{\bar{Y}}$ by

$$\begin{aligned} \Sigma_{\bar{Y}} &= \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1t} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{t1} & \sigma_{t2} & \cdots & \sigma_{tt} \end{bmatrix} \\ &= W_{\bar{X}} (\Lambda_{\bar{X}} + \sigma_N^2 I) W_{\bar{X}}^T \\ &= \begin{bmatrix} W_1^T \\ W_2^T \\ \vdots \\ W_t^T \end{bmatrix}^T \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_t \end{bmatrix} \begin{bmatrix} W_1^T \\ W_2^T \\ \vdots \\ W_t^T \end{bmatrix} \end{aligned} \quad (9)$$

where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_t$ is eigenvalue vector of Y , W_r ($r = 1, 2, 3, \dots, t$), $t \times 1$ vectors, denote the sequence of mutually orthogonal PCA bases into which the projection of image stack \bar{Y} produces the PCs with sequentially largest variances represented by λ_r . The orthonormal PCA transformation matrix for \bar{X} is then

$$P = W_{\bar{X}}^T \quad (10)$$

and then we transform our data into a PCA domain by

$$\hat{Y} = \bar{Y}P \quad (11)$$

where \hat{Y} is the de-correlated dataset for signal, most energy of \hat{Y} concentrates on the several most important components, i.e., while the last few components are mostly due to noise. Therefore, by selecting the first K PCs that contains most of the scan signal, the noise can be removed

$$\hat{X} = \hat{Y}(:, 1 : K)P^T(1 : K, :) \quad (12)$$

reformatting \hat{X} results in denoised dataset. Since Y is transformed according to a certain rule based on the original hyperspectral data I , finally we will use \hat{X} to reconstruct our denoised HSI according to the original rules.

2.3 nonlocal spatial spectral PCA denoising

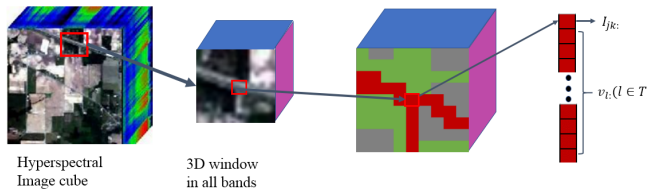


Figure 1. Selecting some similar pixels 3D block around the reference pixels

The core of nonlocal algorithm is to address the non-stationarity of image signals by restructuring the Y . The boundary information and texture information can be obtained by selecting the similar pixels with referenced pixels of I_{jkb} in a nonlocal window. In our proposed method, however, besides the nonlocal similarity in image, many researches have shown that the HSI has strong correlations in both spatial and spectral domains. So, we use 3D blocks instead of 2D patches in the sparse representation (see Fig. 3). The key to the nonlocal algorithm is to select the T pixels with the highest similarity to the central pixel in a $L \times L \times p$ 3D block. Because every pixel with spectral information can be represented $L \times L$ random vectors, the distance of random vectors cannot be calculated by the classical Euclidean distance algorithm. Therefore, the Mahalanobis distance can be considered for calculating the distance of relevant random vectors, which can be represented by

$$S(I_{jk:}, v_{l:}) = \sqrt{(I_{jk:} - v_{l:})^T \Sigma^{-1} (I_{jk:} - v_{l:})} \quad (13)$$

where $v_{l:}$ ($l = L \times L - 1$) represents all pixels other than the center pixel $I_{jk:}$. Since we only select the T pixels that are closest to the center pixel, in the next calculation, we set $l = T$. Therefore, the similar weights can be defined as

$$w(I_{jk:}, v_{l:}) = \frac{1}{P(I_{jk:})} e^{-\frac{S(I_{jk:}, v_{l:})}{h^2}} \quad (14)$$

where satisfy usual conditions $0 \leq w(I_{jk:}, v_{l:}) \leq 1$ and $\sum_T w(I_{jk:}, v_{l:}) = 1$. The parameter h acts as a degree of algorithm. It controls the decay of the exponential function and therefore the decay of weights as a function of the distances $S(I_{jk:}, v_{l:})$. $P(I_{jk:})$ is the normalizing constant

$$P(I_{jk:}) = \sum_{l=1}^T e^{-\frac{S(I_{jk:}, v_{l:})}{h^2}} \quad (15)$$

Through the above analysis, Y can be represented by

$$y_i = (I_{j,k:}^T, v_{1:}^T, v_{2:}^T, \dots, v_{T:}^T)^T \quad (16)$$

$$= (nl_{0:}^T, nl_{1:}^T, \dots, nl_{T:}^T)^T$$

where $nl_{g:}$ ($g = 0, 1, 2, \dots, T$) is $p \times 1$ vector, which is distributed around the referenced pixel. Therefore, in nonlocal SSPCA, we can get a vector of t elements $t = p \times (T + 1)$ in (2), the covariance matrix of y_i can be represented by

$$\Sigma_{\bar{y}} = \begin{bmatrix} \sigma_{nl_{0,1}nl_{0,1}} & \cdots & \sigma_{nl_{0,1}nl_{l,b}} & \cdots & \sigma_{nl_{0,1}nl_{T,p}} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \sigma_{nl_{l,b}nl_{0,1}} & \cdots & \sigma_{nl_{l,b}nl_{l,b}} & \cdots & \sigma_{nl_{l,b}nl_{T,p}} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \sigma_{nl_{T,p}nl_{0,1}} & \cdots & \sigma_{nl_{T,p}nl_{l,b}} & \cdots & \sigma_{nl_{T,p}nl_{T,p}} \end{bmatrix} \quad (17)$$

where $\sigma_{nl_{A,C}nl_{B,D}}$ provide the covariance between different spatial position of A and B and different bands of C and D in 3D blocks. Therefore, nonlocal SSPCA can capture a nonlocal spatial correlation information and spectral correlation information, and utilize these information in PCA denoising model. The PCA denoising process for Y has been introduced in section 2.2, we can get

$$\hat{X}_i = (\hat{I}_{j,k:}^T, \hat{v}_{1:}^T, \hat{v}_{2:}^T, \dots, \hat{v}_{T:}^T)^T \quad (18)$$

where the size of \hat{X} is the same as Y . Therefore we can get noise-free HSI by

$$\bar{I}_{jk:} = \frac{\hat{I}_{jk:} + \sum_{l=1}^T w(I_{jk:}, v_{l:}) \hat{v}_{l:}}{2} \quad (19)$$

2.4 Complete procedure of the propose approach

The detailed procedure of nonlocal SSPCA is given in the following.

1. Select a large 3D block around the referenced pixels of the HSI, in order to obtain region samples at boundary areas, we perform image padding in spatial dimension before obtaining samples.
2. Use Mahalanobis distance algorithm to calculate the distance between all the pixels in 3D blocks and referenced pixel like (13). In this step, spectral information is also used to calculate the distance between two pixels.
3. Select the nearest T pixels in window to contribute to the denoising according to the distance S , and calculating the similarity weight according to the distance S and the number of pixels T using (14) and (15).
4. Use the selected pixels to construct the PCA transformation matrix of Y , where the columns of the Y represent the number of pixels, and the rows represent the nonlocal spatial information and spectral information of the HSI.
5. PCA transformation of Y :
 - 5.1. Data centralization of Y
 - 5.2. Calculate the covariance matrix using (17)
 - 5.3. Factorize $\Sigma_{\hat{Y}} = W \Lambda_{\hat{Y}} W^T$, set the PCA transformation matrix of P

5.4. Transform the dataset to de-correlation PCA domain $\hat{Y} = YP$

5.5. Select the leading K PCs that contains most of the scan signal by (12).

6. Reorganize the HSI using weighted average like (19)

end

Through the above calculation process, we can finally get the denoised HSI.

3. EXPERIMENTS AND DISCUSSION

3.1 Experimental design

In all experiments, several popular denoising techniques in hyperspectral images are compared, such as wavelet-based BayersShrink, wiener filter, and the TV method. The parameters of referenced methods are set by following the suggestion of the respective authors. In proposed methods, a region of 3×3 is used for spatial PCA and SSPCA. In nonlocal SSPCA, the nonlocal area is set to be 19×19 . Approximately 12 pixels are taken to contribute to denoising of referenced pixels. Due to the different methods of constructing the matrix of Y , the value of K is also different. Choosing the appropriate K value to ensure that most of signals are captured is a crucial job in the experiment.

3.2 Test on simulated image

The data is simulated based on the Indian Pines image, which was captured by airborne visible/infrared imaging spectrometer (AVIRIS) over a vegetation area in northwestern Indiana, USA, with a spatial resolution of 20 m, consisting of 145×145 pixels of 16 ground-truth classes and 220 spectral bands. Only the labeling information of Indian Pines images is used in simulation, 10 ground-truth classes with 17 spectral of 224 spectral bands randomly chosen from USGS spectra library. The image is considered as a clean image. And then, we set $SNR=20$.

After testing all the methods on simulated image, in Fig. 2, we show the PSNR and SSIM in each band with line chart. As we can see, for both measure, the line of nonlocal SSPCA is above those of the other methods, indicating that nonlocal SSPCA outperforms the other methods in term of both noise removal and detail preservation. The second best method seems to be SSPCA, and followed closed by spectral PCA. Due to the importance of spectral information in PCA denoising model, spatial PCA achieves the lowest PSNR and SSIM lines among all methods. Wiener, wavelet and TV methods are comparable performance, but wiener have a high SSIM values. The lines of all denoising methods are above the lines of the noisy image.

Fig. 3-4 display the clean, noisy, and denoising images of three bands achieved by different methods. In the picture, we find that nonlocal SSPCA is most similar to the true image, what's more, nonlocal SSPCA not only represses the presence of noise, but also preserve the the edge information and detail information. SSPCA also has a nice denoising effect, but we find that there is ambiguity at the boundaries of the image. TV tends to oversmooth image in denoising processing. There is most of noise still presented in spatial PCA, which fails to reduce the noise.

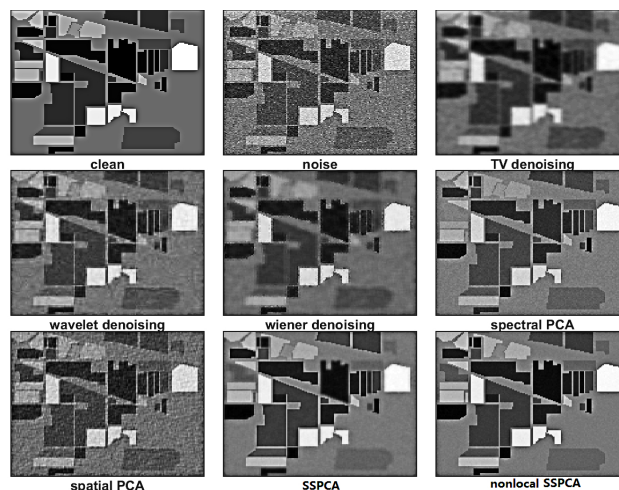


Figure 3. Denoising results achieved by different methods, on band 1 of simulated data. Denoised image by the proposed nonlocal SSPCA method is the most similar one to the true image. The other methods tend to either preserve undesirable artifacts or oversmooth and weak signal.

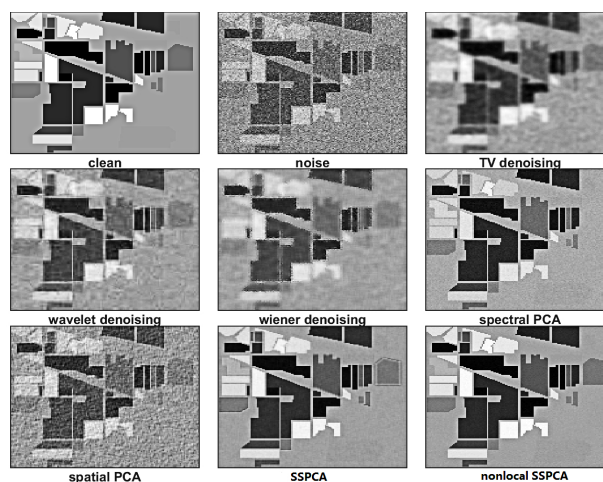


Figure 4. Denoising results achieved by different methods, on band 50 of simulated data.

Fig. 5 show the spectra clean, noisy, and denoising produced by different methods. The similarity of true image reflect the effect of the denoising. The line of nonlocal SSPCA produces the most similar spectra to the true images. SSPCA also have a good denoising effect. Although spectral PCA can obviously restrain the noise, the effect is not very stable with the changes of feature type. Comparing to Wiener and wavelet, denoising performance of TV method is stronger. However, by combining the visual effect of image, we find that wavelet achieves better balance between noise removal and detail preservation. Spatial PCA fails to significantly reduce the noise.

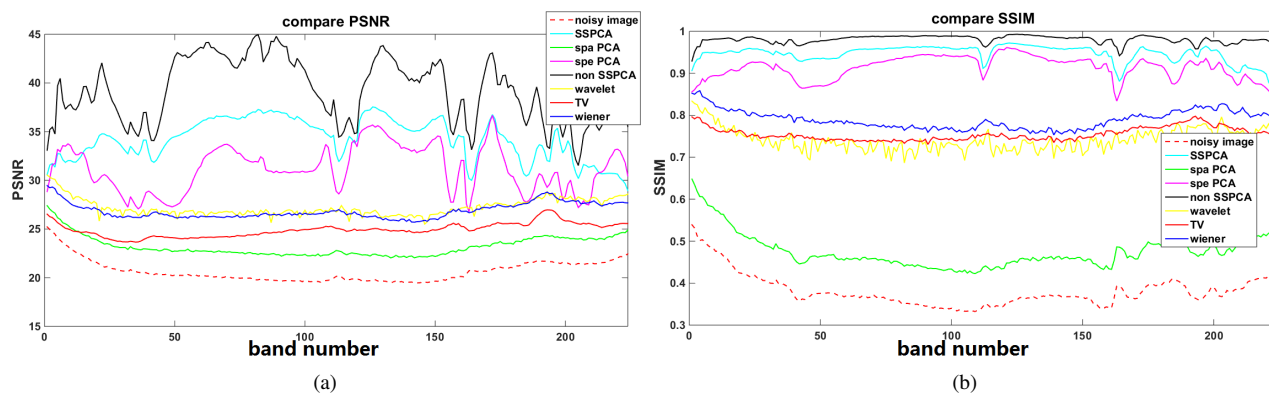


Figure 2. PSNR and SSIM values of the different denoising approaches in each band of the simulated experiment. (a) PSNR. (b) SSIM value. The lines of nonlocal SSPCA above those of the other methods, indicating that nonlocal SSPCA outperforms the other methods in term of both noise removal and detail preservation. The second best method seems to be SSPCA, and followed closed by spectral PCA. Spatial PCA achieves the lowest PSNR and SSIM lines among all methods. Wiener, wavelet and TV methods comparable performance, but wiener have a high SSIM value.

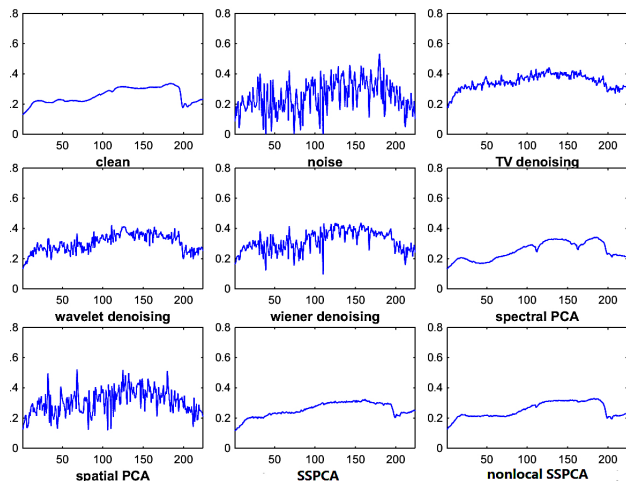


Figure 5. Spectra at pixel (60,60) of the simulated data, before and after denoising by different methods.

3.3 Test on The Real Hyperspectral Data

In this experiment, all methods are tested on the Indian Pines image. The HSI contains 220 spectral bands, with each band contains different intensity of the noise. In order to testing our proposed method, all bands are used in this experiment.

The denoising results of band 200 are shown in Fig. 6. It can be clearly seen that the proposed nonlocal SSPCA method achieves better denoising results than other methods. Because of containing nonlocal spectral information and spectral information in PCA denoising model, the nonlocal SSPCA reduces most random noise efficiently, while also preserves image details, e.g., bright boundary information and point targets very well. Although SSPCA also suppresses most of the noise in the HSI, the denoised image is blurred slightly. Local detailed information, such as edges and texture information of the image are lost in the denoising process. The denoising result using the TV model and wavelet model appear oversmoothed, and most of detailed information is lost. Other methods tend to either keep undesirable artifacts or oversmooth the image.

4. CONCLUSION

In this paper, we have proposed nonlocal SSPCA hyperspectral image denoising algorithm. The PCA was used to separate the noise and signal from contaminated hyperspectral data. Meanwhile, the Mahalanobis distance algorithm was used to estimate the spatial spectral similarity by calculating the distance between the spectrum of referenced pixel and the spectrum of others pixels in nonlocal area, the nonlocal spatial spectral similarity was used to estimated the acceptance probability and captured the image non-stationary information into denoising model. And then, this proposed method was tested on hyperspectral images, Meanwhile in comparison with several other classic denoising method. Through the analysis of both the numerical and visual effects, we have demonstrated that our proposed method have excellent results in denoising. It not only removes a lot of noise, but also preserves detailed information.

5. ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China Grant 41501410 and the State Grid Corporation Science and Technology Project of GCB17201700121.

REFERENCES

Atkinson, I., Kamalabadi, F. and Jones, D. L., 2003. Wavelet-based hyperspectral image estimation. In: *Geoscience and Remote Sensing Symposium, 2003. IGARSS '03. Proceedings. 2003 IEEE International*, pp. 743–745 vol.2.

Buades, A., Coll, B. and Morel, J. M., 2005. A review of image denoising algorithms, with a new one. *Siam Journal on Multi-scale Modeling and Simulation* 4(2), pp. 490–530.

Chen, G. and Qian, S. E., 2008. Simultaneous dimensionality reduction and denoising of hyperspectral imagery using bivariate wavelet shrinking and principal component analysis. *Canadian Journal of Remote Sensing* 34(5), pp. 447–454.

Chen, G. and Qian, S.-E., 2009. Denoising and dimensionality reduction of hyperspectral imagery using wavelet packets, neighbour shrinking and principal component analysis. *International Journal of Remote Sensing* 30(18), pp. 4889–4895.

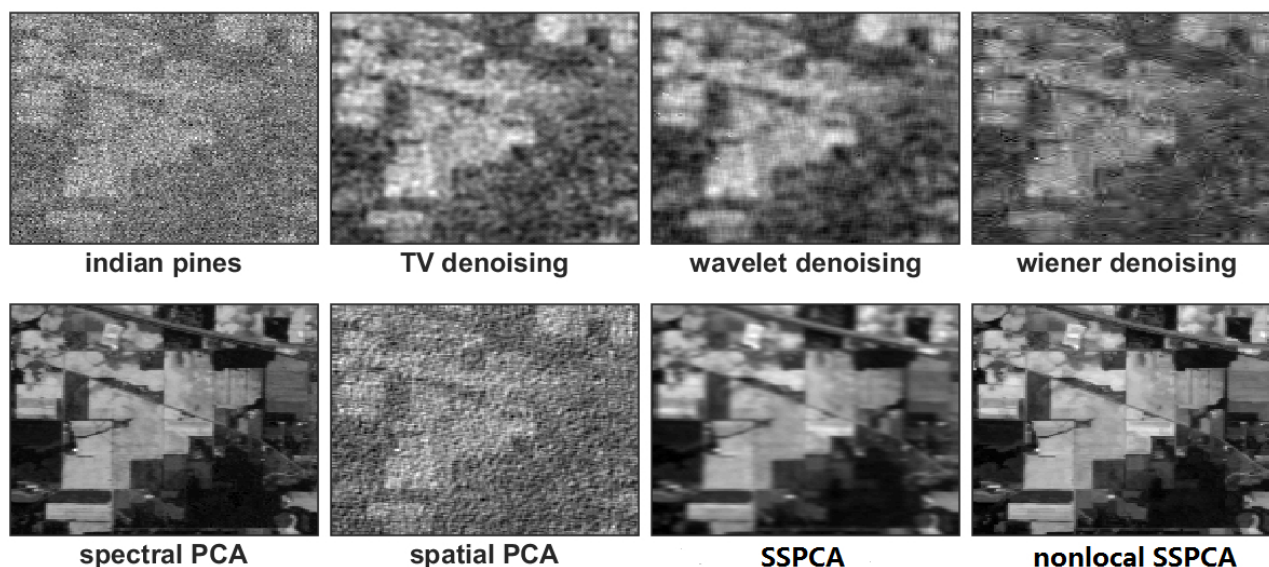


Figure 6. Denoising results achieved by different methods on band 200 of real data. It can be clearly seen that the proposed nonlocal SSPCA method achieved better denoising results than other all methods. Because of containing nonlocal spectral information and spectral information in PCA denoising model, it reduces most random noise efficiently, while also preserving image details, e.g., bright boundary information and point targets very well.

Chen, G. and Qian, S. E., 2011. Denoising of hyperspectral imagery using principal component analysis and wavelet shrinkage. *IEEE Transactions on Geoscience and Remote Sensing* 49(3), pp. 973–980.

Christophe, E., Mailhes, C. and Duhamel, P., 2008. Hyperspectral image compression: Adapting spiht and ezw to anisotropic 3-d wavelet coding. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society* 17(12), pp. 2334.

Dong, W., Li, X., Zhang, L. and Shi, G., 2011. Sparsity-based image denoising via dictionary learning and structural clustering. In: *Computer Vision and Pattern Recognition*, pp. 457–464.

Jr, L. S., Bodanese, B., Zngaro, R. A. and Pacheco, M. T. T., 2010. Discrete wavelet transform for denoising raman spectra of human skin tissues used in a discriminant diagnostic algorithm. *Instrumentation Science and Technology* 38(4), pp. 268–282.

Karami, A., Yazdi, M. and Asli, A. Z., 2011. Noise reduction of hyperspectral images using kernel non-negative tucker decomposition. *IEEE Journal of Selected Topics in Signal Processing* 5(3), pp. 487–493.

Muresan, D. D. and Parks, T. W., 2003. Adaptive principal components and image denoising. In: *International Conference on Image Processing, 2003. ICIP 2003. Proceedings*, pp. I–101–4 vol.1.

Qian, Y. and Ye, M., 2013. Hyperspectral imagery restoration using nonlocal spectral-spatial structured sparse representation with noise estimation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 6(2), pp. 499–515.

Wang, Y. and Niu, R., 2009. Hyperspectral urban remote sensing image smoothing and enhancement using forward-and-backward diffusion. In: *Urban Remote Sensing Event*, pp. 1–5.

Xu, L., Li, F., Wong, A. and Clausi, D. A., 2017. Hyperspectral image denoising using a spatialspectral monte carlo sampling approach. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 8(6), pp. 3025–3038.

Xu, L., Li, J., Shu, Y. and Peng, J., 2014. Sar image denoising via clustering-based principal component analysis. *IEEE Transactions on Geoscience and Remote Sensing* 52(11), pp. 6858–6869.

Yoshino, K., 2007. Spectrally segmented principal component analysis of hyperspectral imagery for mapping invasive plant species. *International Journal of Remote Sensing* 28(5), pp. 1023–1039.

Yuan, Q., Zhang, L. and Shen, H., 2012. Hyperspectral image denoising employing a spectralspatial adaptive total variation model. *IEEE Transactions on Geoscience and Remote Sensing* 50(10), pp. 3660–3677.

Revised October 2018