doi:10.3772/j.issn.1006-6748.2014.01.009

Denoising of hyperspectral imagery by cubic smoothing spline in the wavelet domain①

Chen Shaolin (陈绍林)②, Hu Xiyuan, Peng Silong, Zhou Zhiqiang
(National ASIC Design and Engineering Center, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, P. R. China)

Abstract

The acquired hyperspectral images (HSIs) are inherently affected by noise with band-varying level, which cannot be removed easily by current approaches. In this study, a new denoising method is proposed for removing such kind of noise by smoothing spectral signals in the transformed multi-scale domain. Specifically, the proposed method includes three procedures: 1) applying a discrete wavelet transform (DWT) to each band; 2) performing cubic spline smoothing on each noisy coefficient vector along the spectral axis; 3) reconstructing each band by an inverse DWT. In order to adapt to the band-varying noise statistics of HSIs, the noise covariance is estimated to control the smoothing degree at different spectral positions. Generalized cross validation (GCV) is employed to choose the smoothing parameter during the optimization. The experimental results on simulated and real HSIs demonstrate that the proposed method can be well adapted to band-varying noise statistics of noisy HSIs and also can well preserve the spectral and spatial features.

Key words: denoising, hyperspectral imagery, cubic spline smoothing, wavelet transform, spectral smoothness

0 Introduction

A hyperspectral image (HSI) is generally comprised of dozens to several hundreds of gray-level images taken in a contiguous spectral range with a high spectral resolution, so it provides rich spectral information utilized to improve the accuracy of hyperspectral applications, like unmixing, terrain classification and object identification. However, HSIs often suffer from various kinds of sensor noises (e. g. photon noise, dark noise and readout noise) and noise from atmospheric effect (e. g. the absorption of water vapor and the scattering effect)[1-8]. Thus, denoising of HSIs is an important task for further applications.

Some researchers apply 2-D filtering methods, for instance wavelet shrinkage technique and anisotropic diffusion, into the spectrum-decorrelated HSIs[4,5,7]. These methods assume noises and signals can be well separated by linear transformations, such as principal component analysis (PCA) and maximum noise fraction (MNF)[9]. But, they only perform well for the HSIs with fixed or low-level band-varying noises. Othman and Qian[6] put forward a hybrid spatial-spectral derivative domain wavelet shrinkage method to improve the HSI's signal-to-noise ratio (SNR). The method assumes that the noise is signal-dependent and it is only effective for HSIs carrying a very high SNR, e. g. 600 :1. The noise in these HSIs is mainly produced by imaging sensors, whereas our target noise in this paper is produced by atmospheric effect, which has a relatively high noise level. Other researchers treat HSIs as 3-D tensors and enhance the HSIs in both the spatial domain and the spectral domain[1,3]. Also, several 1-D smoothing techniques have been widely used to directly smooth the reflectance spectrum separately to reduce the noise, these methods include Savizky-Golay (SG) filtering, average filtering, fast Fourier transform filtering[10,11] and cubic smoothing spline[12]. Since these methods don't use the 2-D spatial information of HSIs, they are only effective for spectra with low level noise.

However, the noise variance of HSI may vary from band to band with a high level. These methods will break down in this kind of noise environment. In this paper, we propose a new denoising method that benefits from the compactness of the wavelet transform and the smoothness of spectra to address this problem. For the HSIs, on one hand, both the high spectral resolu-

① Supported by the National Natural Science Foundation of China (No. 60972126, 60921061) and the State Key Program of National Natural Science of China (No. 61032007).

② To whom correspondence should be addressed. E-mail: shaolin.chen@ia.ac.cn

Received on May 23, 2012

tion and the low spatial resolution lead to strong continuity and smoothness of spectral signal. This phenomenon is especially obvious for the reflectance HSIs. On the other hand, in the spatial domain, all the bands of any HSI have the same geometry structures, and this characteristic will be kept in the transformed multiscale domain of 2-D space. Thus, we smooth noisy coefficients signal in the wavelet domain of spatial domain along the spectral axis to enhance HSIs. We estimate the noise covariance matrix of different bands to adjust the smoothing degree at each spectral position. Thus it can be well adapted to the varying noise level from band to band in HSIs.

The remainder of this paper is organized as follows. In Section 1, we review related work about cubic smoothing spline; Section 2 introduces the proposed denoising framework; Section 3 presents the experimental results and the paper is concluded in the last section.

1 Related work about cubic smoothing spline

Smoothing spline, which has been widely used to reduce experimental noises, is a classic data-driven nonparametric regression method[13,14]. It can keep enough smoothing degree without requiring prior knowledge of regression curves. Considering a 1-D noisy signal y with the following model,

$$y(x_n) = g(x_n) + \epsilon_n, \quad n = 1, \cdots, N \tag{1}$$

where ϵ_n denotes Gaussian noise and $\epsilon_n \sim N(0, \sigma_n^2)$, $g(x)$ is supposed to be a smoothing clean signal, i. e. has continuous derivatives up to some order over the whole domain. $x_1 < \cdots < x_N$ are coordinates of the domain. For the spectral signal, x_n corresponds to the wavelength of the n-th band. The signal g can be estimated by following cost function,

$$\hat{\boldsymbol{g}} = \arg\min_{\boldsymbol{g}} \{ (\boldsymbol{y} - \boldsymbol{g})^{\mathrm{T}} \boldsymbol{D}^{-2} (\boldsymbol{y} - \boldsymbol{g}) + \alpha \int_{x_1}^{x_N} (g^p(t))^2 \, \mathrm{d}t \} \tag{2}$$

where $\boldsymbol{D} = \mathrm{diag}(\sigma_1, \cdots, \sigma_N)$, $\boldsymbol{y} = (y_1, \cdots, y_N)^{\mathrm{T}}$ and $\boldsymbol{g} = (g_1, \cdots, g_N)^{\mathrm{T}}$. The first and second terms correspond to the residual sum-of-squares and the penalty, respectively. The penalty term uses a square integral of the p - th derivative of $\boldsymbol{g}$, which determines the smoothing degree of $\boldsymbol{g}$. It can keep p order continuity of smoothing signal at the inner knots. In this study, we choose $p = 2$, and it leads to a cubic smoothing spline, which can be computed efficiently. Parameter α controls the tradeoff between fitting to data and smoothness of the estimated signal.

Given α, the solution to Eq. (2) with the natural boundary condition is

$$\hat{\boldsymbol{g}} = \boldsymbol{A}(\alpha)\boldsymbol{y} = (\boldsymbol{D}^{-2} + \alpha \boldsymbol{Q}\boldsymbol{R}^{-1}\boldsymbol{Q}^{\mathrm{T}})^{-1}\boldsymbol{D}^{-2}\boldsymbol{y} \tag{3}$$

where $\boldsymbol{Q}$ and $\boldsymbol{R}$ are $N \times (N-2)$ and $(N-2) \times (N-2)$ band matrices, respectively, referring to Ref. [15] for details. $\boldsymbol{A}(\alpha)$ is the influence matrix and will be used for the parameter selection.

2 Cubic smoothing spline in the wavelet domain

For low level noises, smoothing spline[12] and other non-parametric regression methods[10,11] can be used to reduce the spectral noise. But when the noise level is high, this kind of methods cannot work well. In this research, we incorporate the 2-D spatial information of local neighborhood of each pixel during the smoothing by applying the cubic spline smoothing to the coefficient curves in the wavelet domain along spectral dimension, which can lead to a more robust result than directly smoothing spectra in the original domain.

2.1 Observation model with multiscale representation for HSIs

The original HSI consists of $N \in \mathbb{N}^*$ spectral components $s^{(n)}$ of spatial size $L \times L$ with $n \in \{1, \cdots, N\}$. Assume that these components are corrupted by additive Gaussian noise, and in the spatial domain, observation vector $\boldsymbol{r}$ can be expressed as

$$\boldsymbol{r}(\boldsymbol{m}) = \boldsymbol{s}(\boldsymbol{m}) + \boldsymbol{\epsilon}(\boldsymbol{m}), \quad \forall \boldsymbol{m} \in \{1, \cdots, L\}^2 \tag{4}$$

where $\boldsymbol{\epsilon}(\boldsymbol{m}) = (\epsilon^1(\boldsymbol{m}), \cdots, \epsilon^N(\boldsymbol{m}))^{\mathrm{T}}$ and $\boldsymbol{\epsilon}(\boldsymbol{m}) \sim N(0, \boldsymbol{C}^{\epsilon})$.

Noise $\boldsymbol{\epsilon}(\boldsymbol{m})$ is independent of $\boldsymbol{s}(\boldsymbol{m}) = (s^1(\boldsymbol{m}), \cdots, s^N(\boldsymbol{m}))^{\mathrm{T}}$. $\boldsymbol{C}^{\epsilon}$ denotes the covariance matrix of $\boldsymbol{\epsilon}$. We also assume that noise level is different from band to band.

Let $r_j^{(n,o)}(\boldsymbol{k})$ denote the wavelet coefficients of the noisy image at the spatial position $\boldsymbol{k} \in \{1, \cdots, L_j\}^2$, resolution level j, orientation subband o, and image band n. $s_j^{(n,o)}(\boldsymbol{k})$ and $\epsilon_j^{(n,o)}(\boldsymbol{k})$ correspond to wavelet coefficients of the noise-free image and noise, respectively. With the vector representation, the wavelet coefficients of all the N bands at the same spatial position $\boldsymbol{k}$ in a subband of the same orientation o and at given resolution level j are grouped into an N dimensional vector, defined as $\{\boldsymbol{r}, \boldsymbol{s}, \boldsymbol{\epsilon}\}$ with $\boldsymbol{k} \in \{1, \cdots, L_j\}^2$

$$\begin{cases} \boldsymbol{r}_j^{(o)}(\boldsymbol{k}) \triangleq (r_j^{1,o}(\boldsymbol{k}), \cdots, r_j^{N,o}(\boldsymbol{k}))^{\mathrm{T}} \\ \boldsymbol{s}_j^{(o)}(\boldsymbol{k}) \triangleq (s_j^{1,o}(\boldsymbol{k}), \cdots, s_j^{N,o}(\boldsymbol{k}))^{\mathrm{T}} \\ \boldsymbol{\epsilon}_j^{(o)}(\boldsymbol{k}) \triangleq (\epsilon_j^{1,o}(\boldsymbol{k}), \cdots, \epsilon_j^{N,o}(\boldsymbol{k}))^{\mathrm{T}} \end{cases} \tag{5}$$

Therefore, the observation model in the wavelet

domain for all $\boldsymbol{k} \in \{1,\cdots,L_j\}^2$ is given by

$$\boldsymbol{r}_j^{(o)}(\boldsymbol{k}) = \boldsymbol{s}_j^{(o)}(\boldsymbol{k}) + \boldsymbol{\epsilon}_j^{(o)}(\boldsymbol{k}),\ o \in \{1,2,3,a\} \quad (6)$$

Due to the linearity of the orthogonal discrete wavelet transform (DWT), the wavelet coefficients $\boldsymbol{\epsilon}_j^{(o)}(\boldsymbol{k})$ of noise still obey the Gaussian distribution ($N(0, \boldsymbol{C}_j^o)$). It is assumed that the noises of all the points in the same subband are independent and identical in distribution, and thus $\boldsymbol{C}_j^{(\epsilon,o)} = \boldsymbol{C}^{(\epsilon)}$ is got for all the levels $j \in \{1,\cdots,J\}$.

2.2 The cubic smoothing spline in the wavelet domain

In imaging spectroscopy, the high spectral resolution and the strong spectral mixture effect of a low spatial resolution within each pixel together result in the strong continuity and smoothness of spectra. This is more prominent for reflectance spectra, which can be usually obtained by an atmospheric correction from the radiance data. The abruptly large variations (spike-like features) in the spectral signature of a pixel must be caused by noise, therefore they should be suppressed[3]. The spike can be got rid by smoothing along the spectral direction. In addition, HSIs can be treated as a volume of extruded structures delineated by the edges in the original spatial planes[3]. In natural scenes, image edges due to landscape features (the boundary between different landscapes, e.g. buildings, trees, roads) are well correlated along the spectral axis. Hence, their locations do not change from band to band.

Because of these characteristics, the subbands in the wavelet domain for HSI within the same orientation and same resolution, but within different bands should keep very similar geometry structures still, thus, the coefficients vector $\boldsymbol{s}_j^{(o)}(\boldsymbol{k})$ have strong continuity and smoothness too along the spectral axis. Therefore noisy coefficients vector $\boldsymbol{r}_j^{(o)}(\boldsymbol{k})$ can be smoothed by a cubic smoothing spline along spectral dimension to reduce noise of HSIs. According to Eq. (3), it is estimated that clean coefficients vector $\boldsymbol{s}_j^{(o)}(\boldsymbol{k})$ for all $\boldsymbol{k} \in \{1,\cdots,L_j\}^2$ by the following expression

$$\hat{\boldsymbol{s}}_j^{(o)}(\boldsymbol{k}) = (\boldsymbol{C}^{-1} + \alpha \boldsymbol{Q}\boldsymbol{R}^{-1}\boldsymbol{Q}^{\mathrm{T}})^{-1}\boldsymbol{C}^{-1}\boldsymbol{r}_j^{(o)}(\boldsymbol{k}),\ o \in \{1,2,3,a\} \quad (7)$$

where $\boldsymbol{Q}$ and $\boldsymbol{R}$ are the same with counterparts of Eq. (3).

2.3 Proposed denoising framework

The proposed denoising framework (as shown in Fig. 1) is comprised of three steps: firstly, all the bands of a HSI are transformed into the wavelet domain; secondly, perform cubic spline smoothing on each noisy coefficient curve along the spectral axis at each spatial position for all subbands; thirdly, denoised HSI is got by the inverse wavelet transform.

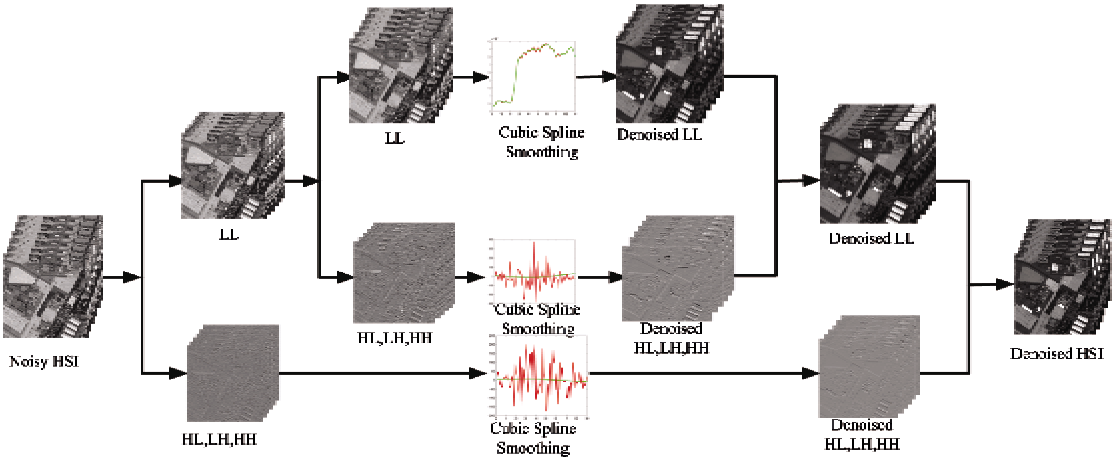


Fig. 1 Denoising framework of the proposed method, only two levels of decomposition are shown, the HH, LH, HL denote the detail subbands and the LL denotes the approximation subband

Since the approximation subband ($o = a$) is obtained by a successive of low pass filtering, each coefficient is the weighted sum of its neighborhood. Applying the smoothing spline to the noisy coefficient curves in the approximation scale is equivalent to incorporating the spatial local information during the denoising. For the orthogonal DWT[16], most of image energies are compressed into the approximation subband. Due to linearity of DWT, the noise intensities are the same with the original domain. Therefore, the noisy coefficient curves in the approximation subband have much higher SNRs than the noisy spectra in the original do-

main. So this can increase the robustness of the curve fitting during denoising.

For the high frequency detail subbands ($o \in \{1, 2, 3\}$), all the coefficients in each subband can be coarsely divided into two parts. The coefficients with large magnitudes correspond to the edges (discontinuity region), and the coefficients with low magnitudes correspond to the continuity region (uniform region), which includes the same substance with very similar spectra. For a clean image, the coefficients of the uniform region in the wavelet domain are zero. For a noisy image in the same region, the magnitudes of coefficients are larger than 0. Thus, the coefficient curves along the spectral dimension in the uniform region are fluctuating direct lines with zero mean. It is easy to fit the noisy coefficient curves to direct lines by any regression methods. It is similar in the discontinuity region, but with a nonzero mean. Therefore, high frequency subbands of DWT are also helpful for denoising a HSI.

The above analysis indicates that the wavelet representation can assist in denoising an HSI through a cubic smoothing spline.

2.4 Optimal parameter selection

This part introduces methods to determine the smoothing parameter α and the noise covariance matrix $\boldsymbol{D}(\boldsymbol{C})$. Parameter α in Eq. (2) controls the tradeoff between the fidelity and the smoothness. And hence, choosing a proper smoothing parameter plays an important role for HSI denoising. Several methods have been proposed to determine the optimal smoothing parameter α[17]. In this investigation, the generalized Cross-Validation (GCV)[17] is utilized to determine the optimal α. GCV can be expressed as the function of the influence matrix $\boldsymbol{A}(\alpha)$,

$$GCV(\alpha) = \frac{\frac{1}{N}\| \boldsymbol{D}^{-1}(\boldsymbol{I} - \boldsymbol{A}(\alpha))\boldsymbol{y} \|^2}{[\frac{1}{N}Trace(\boldsymbol{I} - \boldsymbol{A}(\alpha))]^2} \tag{8}$$

where $Trace(\boldsymbol{I} - \boldsymbol{A}(\alpha))$ denotes the trace of matrix $\boldsymbol{I} - \boldsymbol{A}(\alpha)$, $\boldsymbol{I}$ is an identity matrix with the same size of $\boldsymbol{A}(\alpha)$, N is the length of signal $\boldsymbol{y}$. Optimal α can be estimated through a minimization of $GCV(\alpha)$ within some interval $[a, b]$. For a given α, $GCV(\alpha)$ can be efficiently computed by $p^2 \times N$ operations if $\boldsymbol{D}$ is a diagonal matrix[18]. p denotes the order of derivatives in the penalty term of Eq. (2).

The noise covariance matrix $\boldsymbol{D}(\boldsymbol{C})$ plays the role to finely tune the local smoothing degree of each spectral position. For simplicity, we assume that the noise of different bands is independent. Therefore, we only need to estimate the noise variance of each band. A robust median estimator is used from the finest scale wavelet coefficients[16],

$$\sqrt{\boldsymbol{D}_{n,n}} = \frac{median(|\boldsymbol{s}_1^{n,1}|)}{0.6745} \tag{9}$$

where $\boldsymbol{s}_1^{n,1}$ and $\boldsymbol{D}_{n,n}$ denote the wavelet coefficients in the finest scale and the noise variance of n-th band, respectively.

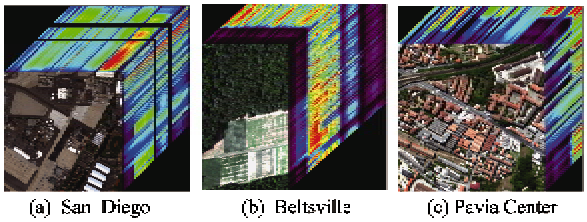
3 Experimental results

In this section, in order to verify the denoising performance of the proposed method on HSIs, an experiment is done on three HSIs (See Fig. 2). For the first two datasets (Fig. 2(a) and (b)), simulated band-varying Gaussian noise is added for quantitative comparisons. Besides, the proposed algorithm is also applied to a real HSI, Pavia Center (Fig. 2(c)). The PCA-based HSI denoising method (PCA-BiShr) presented in Ref. [4], the 2-D (2DBiShr) and 3-D (3DBiShr) bivariate wavelet shrinkage based methods[20,21] are utilized for comparisons. In addition to comparing the visual results of space and spectrum of different techniques, the performance is quantitatively measured by the SNR and root mean square error (RMSE), which are defined as

$$SNR = 10\log10[\frac{\sum_{i,j,l}\hat{\boldsymbol{s}}(i,j,l)^2}{\sum_{i,j,l}(\hat{\boldsymbol{s}}(i,j,l) - \boldsymbol{s}(i,j,l))^2}] \tag{10}$$

$$RMSE = \sqrt{\frac{\sum_{i,j,l}(\hat{\boldsymbol{s}}(i,j,l) - \boldsymbol{s}(i,j,l))^2}{NL^2}} \tag{11}$$

where $\hat{\boldsymbol{s}}$ and $\boldsymbol{s}$ correspond to the restored and the clean HSI, respectively.



(a) San Diego (b) Beltsville (c) Pavia Center

Fig. 2 Hyperspectral datacubes of San Diego (a), Beltsville (b) and Pavia Center (c)

3.1 Experimental datasets

San Diego, Fig. 2(a), is a naval air station in San Diego, California, 3m × 3m for each pixel, about 10nm spectral resolution, 224 bands and collected by the AVIRIS sensor, which has been corrected into reflectance data[22]. Most part of Beltsville (Fig. 2(b))

reflectance dataset is covered by natural vegetation. The spectra of Beltsville have very high smoothness. The spectral resolution is about 5.5nm. A part of 200 $\times$200 $\times$120 is selected for experiments. Pavia center, Fig. 2(c), is captured by the ROSIS-3 senor in 2004 with 102 bands. The spectral resolution is about 4nm. Only a part with the size of 128 $\times$128 $\times$102 is selected for experiments. More than 10 bands (about ranging from 430nm to 480nm) are corrupted by band-varying noise, the noise level decreases as wavelength increases. The spectra of the three datasets all have strong smoothness and continuity because of very high spectral resolutions.

Because the optimal parameter α is not only related to the noise level of the signal, but also related to the signal, we should search an α for each spectrum separately. This leads to a huge computational burden. The whole HSI usually includes many distinct landscapes with very different spectral signatures, which results in a very large searching interval of $[a,b]$, generally $b/a > 10^6$. Finding the optimal α quickly is critical for the practical application of the proposed algorithm. During the experiments, we find that the curve of $GCV(\alpha)$ in the interval $[a,b]$ has an approximate quadratic form, therefore the time for searching an optimal α can be reduced by a quick line searching method. Golden section algorithm[19] is employed to find the optimal α in this work.

3.2 Experiments on simulated dataset

(1) San Diego: The Gaussian noise with band-varying variances is added into 90 continuous bands selected from San Diego dataset, the smallest and largest noise standard variance are 0.5 and 9.3, respectively. Note that the pixel values have been scaled into range [0,255] before adding the noise and then rescaled into the original range. The search region of α is [1.0e $-$5,1.0e +3]. The band 45, which has been heavily corrupted with the largest level noise, is shown in Fig. 3(b). Fig. 3(c)-(e) correspond to the denoising results of BiShr3D, BiShr2D and PCA-BiShr, respectively. Meanwhile, the result of directly applying cubic smoothing spine (SD-CSS) to each spectrum is also given for comparison in Fig. 3(f)[12]. Fig. 3(g) and (h) correspond to the results of the proposed method without (NCRDWT-CSS) and with (RDWT-CSS) estimating the noise variance to regulate local smoothing degree for redundant DWT (à trous wavelet[23]), respectively. Fig. 3 demonstrates that the proposed method performs better than the other methods for the HSI with large band-varying noise. BiShr3D fails to reduce the noise of damaged image, since it is not a local method and estimates the noise variance of 3-D HSI as a whole. BiShr2D denoises each band separately without considering the spectral correlation, thus it also produces poor results.

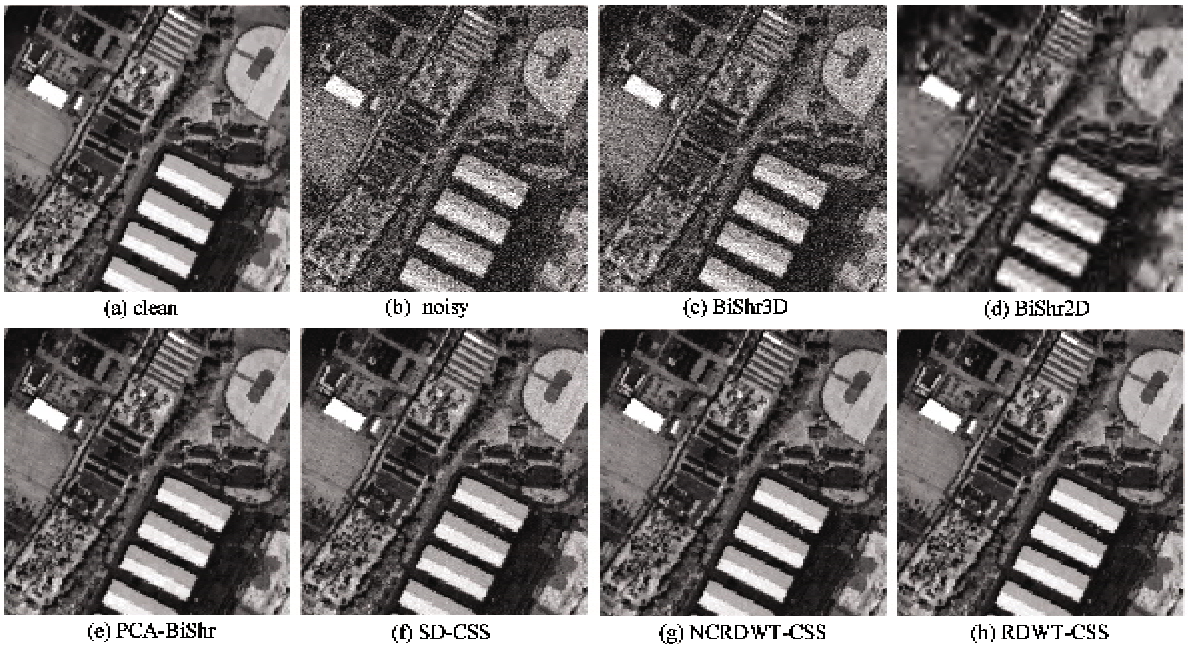


Fig. 3 Denoising results of band 45 for SanDiego

Fig. 4 gives the details of restorations of PCA-BiShr, SD-CSS, NCRDWT-CSS and RDWT-CSS. Better visual result is achieved by the proposed methods. Fig. 5 shows the restored high frequency coefficients along the HL direction at the first scale of band 45 for the redundant DWT. It can be seen that the high frequency coefficients have been well restored by the cubic smoothing spline along the spectral axis direction.

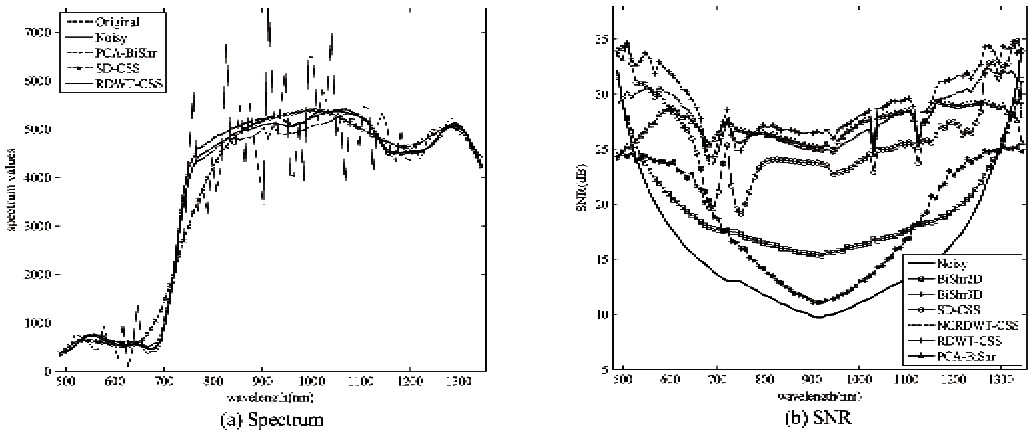
(a) clean (b) PCA-BiShr (c) SD-CSS (d) NCRDWT-CSS (e) RDWT-CSS

Fig. 4 Details comparison of different denoising results with respect to Fig. 3

(a) clean (b) noisy (c) denoised

Fig. 5 The restored high frequency coefficients of HL direction in the first scale of band 45 for the redundant DWT. original (a), noisy (b) and denoised (c) coefficients

Fig. 6(a) illustrates spectra of original, noisy and restorations derived by different methods. The proposed method gives the best restored spectrum (red solid line). The SD-CSS only considers the spectrum continuity without incorporating the spatial information, it can't accurately recover the spectral features with large noise (blue dash star line). The spectrum restored by PCA-BiShr has a steep jump near band 45 (see color image). Thus, RDWT outperforms other methods.



(a) Spectrum (b) SNR

Fig. 6 (a) spectra comparisons of different methods, the parameters used are the same as Fig. 3; (b) the SNR comparisons of each band of different methods

Fig. 6(b) shows the SNRs curves along the wavelength, the SNRs are calculated band by band with Eq. (11). The figure demonstrates that RDWT-CSS gives the best SNR scores and a consistent improvement of all the bands. PCA-BiShr also performs well for the bands with highly noise levels, whereas, it degrades the bands of relatively high qualities (two ends).

The SNRs and RMSEs of different methods are illustrated in Table 1. RDWT-CSS gives the highest SNR and lowest RMSE scores for the whole data. The implementation with noise variance to adjust the local smoothing degree produces a higher SNR than the method without noise variance. Besides, the redundant DWT gives higher SNRs than orthogonal DWT.

(2) Beltsville: The Gaussian noise is added to each band of Beltsville with zero mean and the standard variance is equal to 3% of its maximum amplitude. Thus, the noise levels vary with wavelength. The search region of α is $[1.0e-5, 1.0e+3]$.

Table 1 The SNRs and RMSEs of different denoising methods for simulated San Diego with band-varying noise. (No) Noise Covariance means (no) estimating noise covariance to adjust the local smoothing degree

	Noisy	BiShr3D	BiShr2D	PCA-BiShr	No Noise Covariance			Noise Covariance		
					SD-CSS	DWT-CSS	RDWT-CSS	SD-CSS	DWT-CSS	RDWT-CSS
RMSE	692.17	530.65	386.39	143.47	202.91	152.75	135.60	179.71	135.15	118.37
SNR(dB)	13.18	15.54	17.97	26.63	23.63	26.09	27.14	24.66	27.15	28.30

Fig. 7 and Fig. 8 show the 2-D spatial denoising results and SNR evolution curves of different methods, respectively. The figures demonstrate similar results with the experiments on San Diego are achieved for Beltsville. Some spatial features have been smoothed out by BiShr2D and BiShr3D. Whereas, the proposed multiscale smoothing method (RDWT-CSS) can well restore the spatial features by utilizing spectral smoothness. The RDWT-CSS also performs better than SD-CSS. Fig. 8 and Table 2 illustrate similar results are achieved by RDWT-CSS and PCA-BiShr, since most part of Beltsvillie is covered by natural vegetation with very similar spectra. But RDWT-CSS performs producing much better SNRs than other methods.

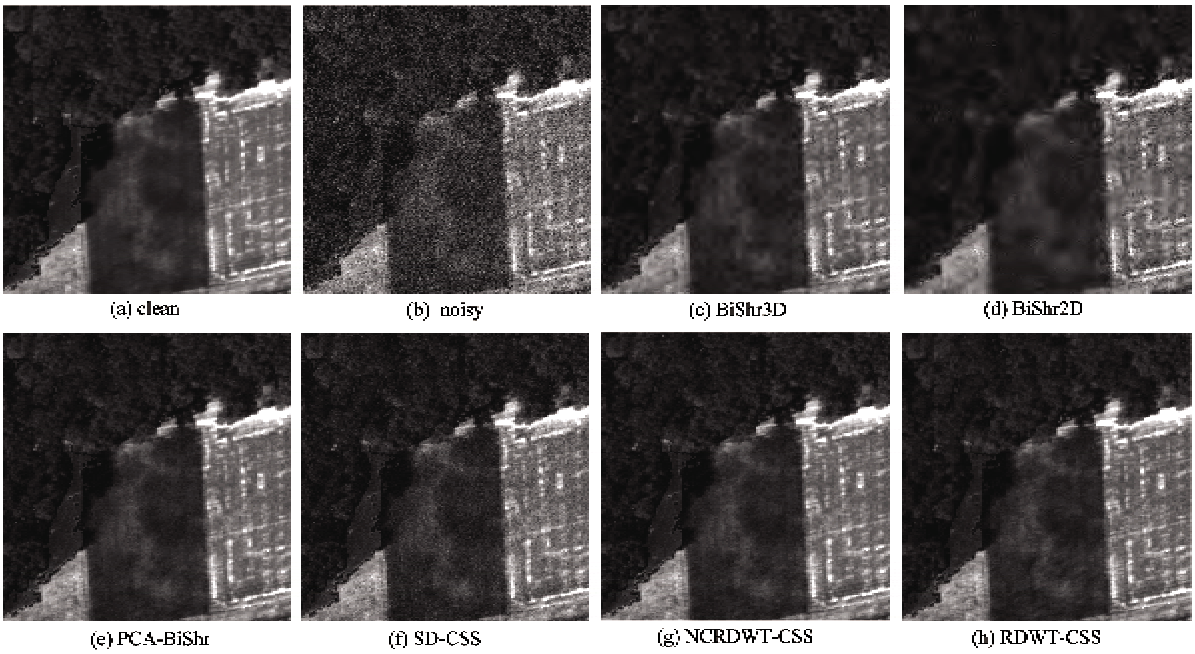


Fig. 7 Simulated Beltsville, denoising results of different methods with band 85. (a) clean; (b) noisy; (c) BiShr3D with 3 scale and 7 ×7 ×7 window size; (d) BiShr2D with 3 scale, haar wavelet and 5 ×5 window size; (e) PCA-BiShr, the first two PCs are kept unchange, haar wavelet, 3 scale, 7 ×7 window size for bivariate wavelet shrinkage for left components; (f) SD-CSS; (g) and (h) correspond to NCRDWT-CSS and RDWT-CSS with 3 scale and haar wavelet, respectively.

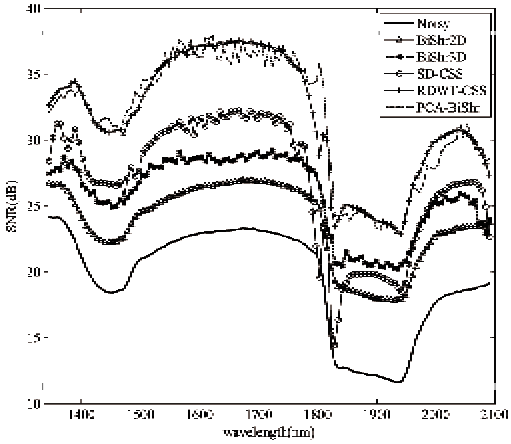


Fig. 8 The SNR comparisons of different methods for Beltsville

3.3 Experiments on pavia center dataset

In this experiment, the proposed method is applied to a real data, Pavia Center, which has often been used for validating the hyperspectral classification algorithm. Because of strong noise in the first 10 bands, researchers always avoid using them for classification[24].

Fig. 9 illustrates the denoising results of different methods. Although the 2DBiShr has removed most of the high frequency noise, it also leads to heavy block effects (Fig. 9(b)). SD-CSS and RDWT-CSS produce slight better results than PCA-BiShr, especially for the band with heavier noise (band 3). Fig. 10 shows smoothed spectra by different methods. The PCA-BiShr

has dramatically changed the spectral values in the region of clean bands, whereas the proposed method can well preserve the original clean spectral signatures of a spectrum while denoising the noisy region.

Table 2 The SNRs and RMSEs of different denoising methods for simulated Beltsville with estimating noise covariance to adjust the local smoothing degree

	Noisy	BiShr3D	BiShr2D	PCA-BiShr	SD-CSS	DWT-CSS	RDWT-CSS
RMSE	129.06	63.64	81.07	30.99	57.36	34.75	30.97
SNR(dB)	21.36	27.46	25.35	33.70	28.36	32.71	33.71

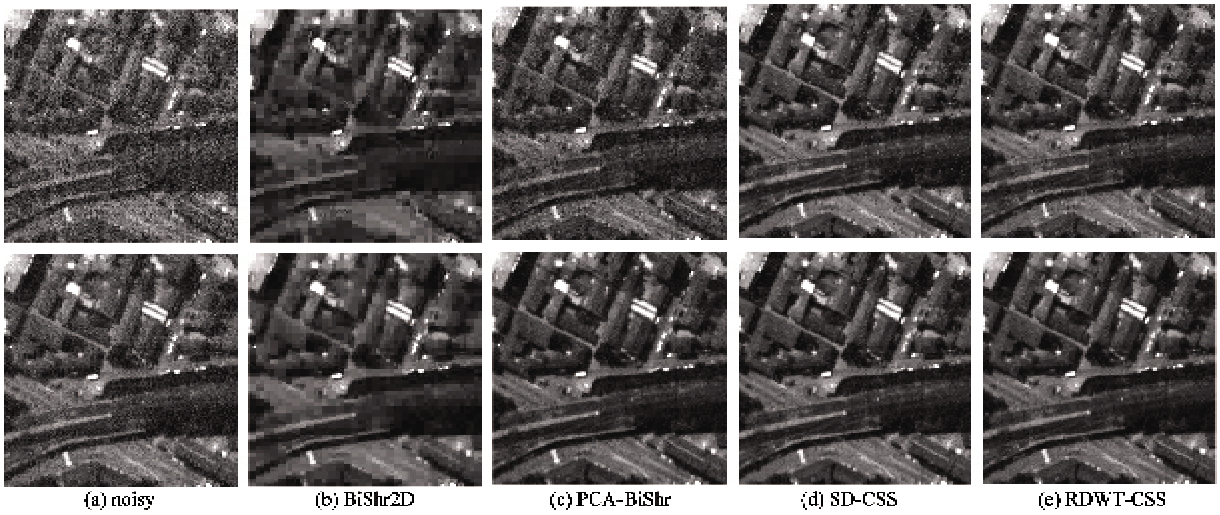


Fig. 9 The denoising results of Pavia Center data by different methods. The first and second rows correspond to the band 3 and band 5, respectively. (a) noisy data; (b) BiShr2D with 3 scale and haar wavelet; (c) PCA-BiShr, the first band principal component is kept unchanged, 3 scale and haar wavelet are used in the bivariate shrinkage, dual complex DWT is used to denoise the spectrum separately; (d) SD-CSS; (e) RDWT-CSS with 3 scale and haar wavelet.

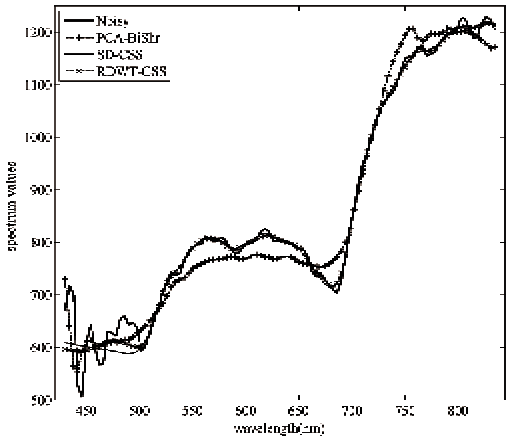


Fig. 10 Original spectrum and smoothed spectra by different methods, the parameters used are the same as Fig. 9

4 Conclusion and discussion

We have presented a multiscale smoothing frame to enhance the noisy HSIs with highly band-varying additive Gaussian noise in this paper. Experimental results illustrate the proposed multiscale method can well recover the spatial and spectral features. It outperforms the wavelet threshold-based methods and directly smooth spectra using the cubic smoothing spline. In addition, the denoising performance can benefit from estimating noise covariance as weights to adjust the smoothing degree of each spectral position. Note that we use the cubic smoothing spline to denoise the spectral signal in this paper, in fact, many other 1-D denoising methods can replace the smoothing spline. The smoothness assumption for the spectral signal is strong to a certain extent for some real HSIs, some sharp spectral signatures may be smoothed out. Therefore, we focus our study on the reflectance HSIs, which have a strong spectral smoothness, especially for the scenes covered by natural vegetation. Next, we will consider some adaptive smoothing methods according to signals. In the high frequency subbands of multiscale decomposition, the energy mainly locates at the strong image edges, the occupied area of which in the whole image may affect the adjusting effects of noise covariance. This needs to be explored thoroughly in the next.

The computational complexity is relatively high, since finding the optimal smoothing parameter. The computational time for DWT-CSS is about 1800 seconds with a Matlab simulation for a HSI with the size of

200 × 200 × 90. The computer is configured with 2.2GHz, Dual Core Process and 2G Memory. The time can be dramatically decreased with a further optimization. It can be also alleviated by a parallel computing, since the proposed method can be paralleled easily.

Reference

[1] Letexier D, Bourennane S. Noise removal from hyperspectral images by multidimensional filtering. *IEEE Trans Geoscience and Remote Sensing*, 2008, 46(7): 2061-2069

[2] Wang Y, Niu R, Yu X. Anisotropic diffusion for hyperspectral imagery enhancement. *IEEE Sensors Journal*, 2010, 10(3): 469-477

[3] Martin-Herrero J. Anisotropic diffusion in the hypercube. *IEEE Trans Geoscience and Remote Sensing*, 2007, 45(5):1386-1398

[4] Chen G, Qian S. Denoising of hyperspectral imagery using principal component analysis and wavelet shrinkage. *IEEE Trans Geoscience and Remote Sensing*, 2011, 49(3):973-980

[5] Lennon M, Mercier G, Hubert-Moy L. Nonlinear filtering of hyperspectral images with anisotropic diffusion. In: Proceedings of the 2002 IEEE International Geoscience and Remote Sensing Symposium, Toronto, Canada, 2002. 2477-2479

[6] Othman H, Qian S. Noise reduction of hyperspectral imagery using hybrid spatial-spectral derivative-domain wavelet shrinkage. *IEEE Trans Geoscience and Remote Sensing*, 2006, 44(2):397-408

[7] Atkinson I, Kamalabadi F, Jones D. Wavelet-based hyperspectral image estimation. In: Proceedings of the IEEE International Geoscience and Remote Sensing Syposium, Toulouse, France, 2003. 743-745

[8] Chang W, Guo L, Liu K, et al. Denoising of hyperspectral data based on wavelet transform and principal component analysis. *Computer Measurement and Control*, 2009 6(17):1070-1076 (in Chinese)

[9] Green A, Berman M, Switzer P, et al. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Trans Geoscience and Remote Sensing*, 1988, 26(1):65-74

[10] Tsai F, Philpot W. Derivative analysis of hyperspectral data. *Remote Sensing of Environment*, 1998, 66(1): 41-51

[11] Miglani A, Ray S, Vashishta D, et al. Comparison of two data smoothing techniques for vegetation spectra derived from EO-1 hyperion. *Journal of the Indian Society of Remote Sensing*, 2011, 39(4):443-453

[12] Lu W, Yu X, Liu J. The application of cubic smooth spline in noise filtering to hyperspectral data. *Journal of Institute of Surveying and Mapping* 2005, 22(1):11-14 (in Chinese)

[13] Craven P, Wahba G. Smoothing noisy data with spline functions. *Numerische Mathematik*, 1978, 31: 377-403

[14] Boor C. A Practical Guide to Splines. New York: Springer, 1978

[15] Fessler J. Nonparametric fixed-interval smoothing with vector splines. *IEEE Trans Signal Processing*, 1991, 39(4):852-859

[16] Donoho D. De-noising by soft-thresholding. *IEEE Trans Information Theory*, 1995, 41(3):613-627

[17] Wahba G. Spline Models for Observational Data. Philadelphia: SIAM, 1990

[18] Hutchinson M, Hoog F. Smoothing noisy data with spline functions. *Numerische Mathematik*, 1985, 47: 99-106

[19] Luenberger D, Ye Y. Linear and Nolinear Programming, 3rd Edition. New York: Springer, 2008

[20] Sendur L, Selesnick I. Bivariate shrinkage with local variance estimation. *IEEE Signal Processing Letters*, 2002, 9(12): 438-441

[21] Chen G, Bui T, Krzyzak A. Denoising of three dimensional data cube using bivariate wavelet shrinking. *International Journal of Pattern Recognition and Artificial Intelligence*, 2011, 25(3):403-413

[22] ENVI: http://www.ittvis.com/productsservices/envi/tutorials.aspx April 2011

[23] Mallat S. A Wavelet Tour of Signal Processing, Second Edition, (*Wavelet Analysis & Its Applications*), 2nd ed. California/ London: Academic Press, 1999

[24] Fauvel M, Benediktsson J, Chanussot J, et al. Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles. *IEEE Trans Geoscience and Remote Sensing*, 2008:46(11) 3804-3814

Chen Shaolin, born in 1982. Now, He is a Ph. D candidate in the Institute of Automation, Chinese Academy of Science. His research interests include hyperspectral images processing and digital image processing.