FAST HYPERSPECTRAL IMAGE DENOISING
BASED ON LOW RANK AND SPARSE REPRESENTATIONS

Lina Zhuang and José M. Bioucas-Dias

Instituto de Telecomunicações, Instituto Superior Técnico,
Universidade de Lisboa, 1900-118, Lisbon, Portugal.

ABSTRACT

The very high spectral resolution of Hyperspectral Images (HSIs) enables the identification of materials with subtle differences and the extraction subpixel information. However, the increasing of spectral resolution often implies an increasing in the noise linked with the image formation process. This degradation mechanism limits the quality of extracted information and its potential applications. This paper presents a new HSI denoising approach developed under the assumption that the clean HSI is low-rank and self-similar. Under these assumptions, the clean HSI admits extremely compact and sparse representations, which are exploited to derive a very fast and competitive denoising algorithm, named Fast Hyperspectral Denoising (FastHyDe), able to cope with Gaussian and Poissonian noise. In a series of experiments, the proposed approach competes with state-of-the-art methods, with much lower computational complexity.

Index Terms— Denoising, hyperspectral image, signal subspace, low rank structure, non-local patch(cube), sparse representation, BM3D, BM4D.

1. INTRODUCTION

Hyperspectral remote sensing images have been widely used in countless applications, (e.g., earth observation, environmental protection and natural disaster monitoring), since they provide remarkably high spectral resolution (hundreds or thousands spectral channels), which enables material identification with precision via spectroscopic analysis. However, the precise identification (e.g., precision farming) requires high quality of hyperspectral image (HSI), whereas the HSIs acquired by imaging instruments are often noisy owing to a number of degradation mechanisms such as electronic noise, Poissonian noise, quatization noise, and atmostheric effects.

Large research efforts have been devoted to image denoising, which is still a challenging problem in image process-ing. Among the recent developments, patch-based image denoising holds the state-of-the-art, e.g., BM3D [1] and BM4D [2, 3]. This line of research make full use of the self-similarity property of real world images, meaning that, given an image, there are similar patches at different locations and scales.

Besides exploiting the spatial information, the high correlation in spectral domain has been widely investigated for hyperspectral data. The works spectral-spatial adaptive hyperspectral TV (SSAHTV)[4] and hybrid spatial-spectral noise reduction (HSSNR) [5] are two paradigmatic examples of this line of research. The usual band dependent noise statistics of HSIs have also been considered, for example in the noise-adjusted PCs transform (NAPCs) method [6] and in the hyperspectral imagery denoising based on oblique subspace projection (DOBSP) [7].

HSIs have very high spectral correlation and spatial self-similarity. As a consequence, this class of images has low-rank structure and admit sparse reprentations on suitable frames. These characteristics of HSIs have received increasing attention, namely in spectral unmixing, classification, and denoising. In the case of denoising, we refer the noise adjusted iterative low-rank matrix approximation (NAILRMA) [8] and the structured tensor TV-based regularization [4].

1.1. Contribution

Most of the published hyperspectral denoising algorithms are time-consuming, namely due to the large sizes of HSIs and due to iterative estimation procedures both in the spatial and the spectral domains. In an effort to mitigate these shortcomings, this paper introduces FastHyDe, a very fast HSI denosing algorithm, which takes full advantage of the HSIs' low rank structure and self-similarity above referred to.

The paper is organized as follows. Section 2 introduces formaly FastHyDe, a denosing approach based on low rank and sparse representations. Section 3 and Section 4 present experimental results including comparisons with the state-of-the-art. Section 5 concludes the paper.

2. PROPOSED FASTHYDE DENOISING

Let $\mathbf{X} := [\mathbf{x}_1, \ldots \mathbf{x}_n] \in \mathbb{R}^{n_b \times n}$ denote a HSI with n spectral vectors (the columns of $\mathbf{X}$) of size n_b (the number of

bands of the sensor). In hyperspectral denoising problems under the additive noise assumption, the observation model may be written as

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

where $\mathbf{Y}, \mathbf{N} \in \mathbb{R}^{n_b \times n}$ represent the observed HSI data and noise, respectively.

We assume that the spectral vectors $\mathbf{x}_i$, for $i = 1, \dots, n$, live in a $k-$dimensional subspace $\mathcal{S}_k$, with and $n_b \gg k$. This is a very good approximation in most real HSIs [9]. Therefore, we may write

$$\mathbf{X} = \mathbf{EZ}, \tag{2}$$

where the columns of $\mathbf{E} = [\mathbf{e}_1, \dots, \mathbf{e}_k] \in \mathbb{R}^{n_b \times k}$ holds a basis for $\mathcal{S}_k$ and matrix $\mathbf{Z} \in \mathbb{R}^{k \times n}$ holds the representation coefficients of $\mathbf{X}$ with respect to (w.r.t.) $\mathbf{E}$. We assume, without loss of generality, that $\mathbf{E}$ is orthogonal, that is $\mathbf{E}^T\mathbf{E} = \mathbf{I}$ with $\mathbf{I}$ representing the identity matrix of appropriate size. Matrix $\mathbf{E}$ may be learned from the data using, e.g., the HySime algorithm [10] or singular value decomposition (SVD) of $\mathbf{Y}$ in the case the noise is independent and identically distributed (i.i.d.). We will herein term the rows of $\mathbf{Z}$ eigen-images.

This paper explores two main characteristics of hyperspectral data: 1) HSIs live in low dimensional subspaces, which opens the door to remove the bulk of the noise using projection-based methods [9]; 2) the eigen-images are self-similar and, therefore, they may be denoised with non-local patch-based methods such as BM3D [1], BM4D [2, 3].

Below we start by considering that the noise is additive and i.i.d. Later, we consider non-i.i.d and non-additive scenarios.

2.1. Additive i.i.d. noise

In the first case, we consider that the noise is additive and i.i.d. over all components of $\mathbf{N}$. Therefore, assuming that the subspace $\mathbf{E}$ has been learned from observed data $\mathbf{Y}$, the eigen-images denoising problem is formulated as

$$
\begin{aligned}
\hat{\mathbf{Z}} &= \arg\min_{\mathbf{Z}} \frac{1}{2} \|\mathbf{EZ} - \mathbf{Y}\|_F^2 + \lambda\phi(\mathbf{Z}) \\
&= \arg\min_{\mathbf{Z}} \frac{1}{2} \|\mathbf{Z} - \mathbf{E}^T\mathbf{Y}\|_F^2 + \lambda\phi(\mathbf{Z}),
\end{aligned} \tag{3}
$$

where $\|\mathbf{X}\|_F^2 = \text{trace}(\mathbf{XX}^T)$ is the Frobenius norm of $\mathbf{X}$. The first term on the right-hand side represents the data fidelity and accounts for the fact that the noise is i.i.d., while the second term is a regularizer expressing prior information tailored to self-similar images. We assume that the function ϕ is decouple w.r.t. the eigen-images, that is

$$\phi(\mathbf{Z}) = \sum_{i=1}^{k} \phi_i(\mathbf{Z}^i) \tag{4}$$

where $\mathbf{Z}^i$ is the i-th eigen-image, i.e., the i-th row of $\mathbf{Z}$. An informal justification for (4) is that the components of $\mathbf{Z}$ tend to be decorrelated. Although decorrelation does not imply

statistical independence, it is a necessary condition for it. In practice, this assumption leads to excellent results as shown in Section 3.

Under the hypothesis (4), the solution of (3) is decoupled w.r.t. $\mathbf{Z}^i$ and may be written as

$$\widehat{\mathbf{Z}} = \psi_{\lambda\phi}(\mathbf{E}^T\mathbf{Y}) = \begin{bmatrix} \psi_{\lambda\phi_1}(\mathbf{e}_1^T\mathbf{Y}) \\ \vdots \\ \psi_{\lambda\phi_k}(\mathbf{e}_k^T\mathbf{Y}) \end{bmatrix}, \tag{5}$$

where

$$\psi_{\lambda\phi_i}(\mathbf{y}) = \arg\min_{\mathbf{w}} \frac{1}{2}\|\mathbf{y} - \mathbf{w}\|_F^2 + \lambda\phi_i(\mathbf{w})$$

is the so-called denoising operator, or Moreau proximity operator of ϕ [11].

The proposed algorithm for HSI denoising, termed *Fast Hyperspectral denoising* (FastHyDe), is summarized in Algorithm 1. Step 2 learns $\mathbf{E}$, e.g., using HySime, step 3 denoises the eigen-images, and step 4 reconstructs the estimate of the original data. In this paper we use the BM3D as denoising operator, as it is the state-of-the-art denoiser for self-similar images [1].

Algorithm 1 FastHyDe denoising algorithm for i.i.d. noise

1: **Input:** Noisy HSI data $\mathbf{Y}$
2: Learn the subspace $\mathbf{E}$ from $\mathbf{Y}$
3: Denoise the eigen-images $\widehat{\mathbf{Z}} = \psi_{\lambda\phi}(\mathbf{E}^T\mathbf{Y})$
4: Reconstruct the denoising HSI: $\widehat{\mathbf{X}} = \mathbf{E}\widehat{\mathbf{Z}}$
5: **Output:** Denoisy HSI $\mathbf{X}$

2.2. non-i.i.d. noise

Consider that in the observation model (1), $\mathbf{N}$ is zero-mean, additive, Gaussian, pixelwise independent with spectral covariance $\mathbf{C}_\lambda = E[\mathbf{n}_i\mathbf{n}_i^T]$, where $\mathbf{n}_i$ is any column of $\mathbf{N}$. Notice that in the i.i.d. case $\mathbf{C}_\lambda = \sigma^2\mathbf{I}$, which is not the case in the non-i.i.d. scenario. We assume that $\mathbf{C}_\lambda$ is positive definite and therefore non-singular. In order to reconvert the non-i.i.d. scenario into the i.i.d. one, we whiten the observed data, that is

$$\widetilde{\mathbf{Y}} := \sqrt{\mathbf{C}_\lambda^{-1}}\mathbf{Y}, \tag{6}$$

where $\sqrt{\mathbf{C}_\lambda^{-1}}$ is a matrix denoting the square root of $\mathbf{C}_\lambda^{-1}$ and $\sqrt{\mathbf{C}_\lambda}$ denotes its inverse. Then the observation model of $\widetilde{\mathbf{Y}}$ becomes

$$\widetilde{\mathbf{Y}} = \sqrt{\mathbf{C}_\lambda^{-1}}\mathbf{X} + \sqrt{\mathbf{C}_\lambda^{-1}}\mathbf{N} = \widetilde{\mathbf{X}} + \widetilde{\mathbf{N}}, \tag{7}$$

whose noise covariance matrix is

$$\widetilde{\mathbf{C}}_\lambda = E[\widetilde{\mathbf{n}}_i\widetilde{\mathbf{n}}_i^T] = \mathbf{I}, \tag{8}$$

where $\widetilde{\mathbf{n}}_i$ is any column of $\widetilde{\mathbf{N}}$

Since the noise in (8) is i.i.d., we may again formulate the denoising problem as

$$\widehat{\widetilde{\mathbf{Z}}} = \arg \min_{\widetilde{\mathbf{Z}}} \frac{1}{2} \|\widetilde{\mathbf{Z}} - \widetilde{\mathbf{E}}^T \widetilde{\mathbf{Y}}\|_F^2 + \lambda \phi(\widetilde{\mathbf{Z}}), \qquad (9)$$

where $\widetilde{\mathbf{E}}$ holds an orthonormal basis learned from $\widetilde{\mathbf{Y}}$. The solution of (9) can be found following the same steps described in Algorithm 1. The clean data of $\widetilde{\mathbf{Y}}$ is estimated as $\widehat{\widetilde{\mathbf{X}}} = \widetilde{\mathbf{E}}\widehat{\widetilde{\mathbf{Z}}}$, followed by recovering original clean data $\mathbf{X}$ from $\widehat{\widetilde{\mathbf{X}}}$ (namely, $\widehat{\mathbf{X}} = \sqrt{\mathbf{C}_\lambda}\widehat{\widetilde{\mathbf{X}}}$).

3. EVALUATION WITH SIMULATED DATA

In order to compare the proposed method with the state-of-the-art denoising algorithms, the experiments were conducted using the same data simulated in the same way as described in [8]. Two datasets, including a subimage of Washington DC Mall dataset (of size $256 \times 256 \times 191$) and a subset of Pavia city center dataset (of size $200 \times 200 \times 80$), are used and each band is normalized to $[0, 1]$ in advance. These subimage of high quality are considered as clean images in this experiment section. Three kinds of additive noise are added to simulate the noisy images as follows:

Case 1 (Gaussian i.i.d. noise): $\mathbf{n}_i \sim \mathcal{N}(\mathbf{0}, \sigma^2\{\mathbf{I}\})$ with $\sigma \in \{0.02, 0.04, 0.06, 0.08, 0.1\}$.

Case 2 (Gaussian non-i.i.d. noise): $\mathbf{n}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{D})$ where $\mathbf{D}$ is a diagonal matrix with diagonal elements sampled from a Uniform distribution $U(0, 1)$.

Case 3 (Poissonian noise): $\mathbf{N} \sim \mathcal{P}(\alpha\mathbf{X})$, where $\mathbf{A} = [a_{ij}]$, with $a_{ij} > 0$, $\mathcal{P}(\mathbf{A})$ stands for a matrix of size($\mathbf{A}$) of independent Poisson random variables whose parameters are given by the corresponding element of $\mathbf{A}$. The parameter α is such that SNR$= \alpha(\sum_{i,j} a_{ij}^2)/(\sum_{ij} a_{ij})$ was set 15dB.

Before being denoised by FastHyDe, the simulated images in Case 2 and 3 were preprocessed in order to to have additive i.i.d. noise. In case 2, we applied the transformation (6). In case 3, we applied the Anscombe transform $\widetilde{\mathbf{Y}} := 2\sqrt{\mathbf{Y} + \frac{3}{8}}$, which converts Poissonion noise into approximately additive noise [12], before the FasHyDe and the remaining algorithms.

The proposed FastHyDe is compared with BM3D [1], applied band by band, BM4D [2], "PCA+BM4D" [13], and NAILRMA [8]. For quantitative assessment, the peak signal-to-noise (PSNR) index and the structural similarity (SSIM) index of each band are calculated. The corresponding mean PSNR (MPSNR) and mean SSIM (MSSIM) in Washington DC Mall are reported in Table 1. Dealing with different kinds of noise, FastHyDe yields almost uniformly the best performance with gains increasing as the noise increases, as it may be concluded from those results. The quality of the FastHyDe reconstruction may also be inferred from Figs. 1-2. The denoising results for Pavia city center dataset exhibit a similar pattern. They are not reported for lack of space.
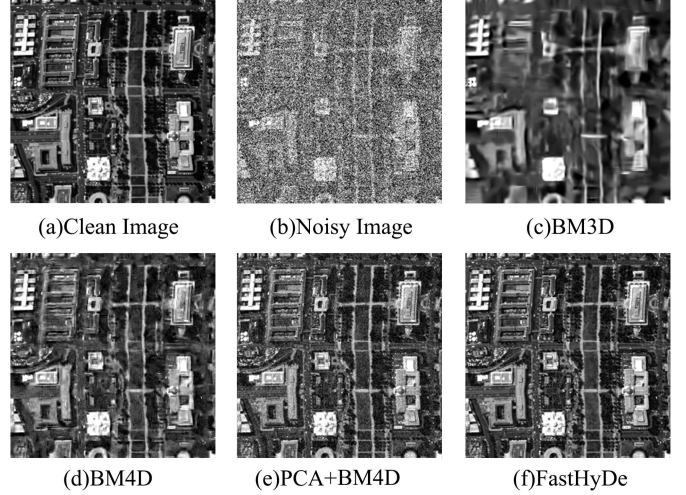

(a)Clean Image   (b)Noisy Image   (c)BM3D

(d)BM4D   (e)PCA+BM4D   (f)FastHyDe

Fig. 1. Denoising band 11 of the Washington DC Mall image in Case 2.

The non-local patch-based BM3D suits very well the eigen-images, which are self-similar. In addition, the fact the that denoising is applied only to the eigen-images, which are much less than the number of bands, significantly reduces the FastHyDe complexity (see Table 2), relative to the competitors. The algorithms were implemented using MATLAB R2010 on a desktop PC equipped with eight Intel Core i7-4970 CPU (at 3.60 GHz) and 16 GB of RAM memory.
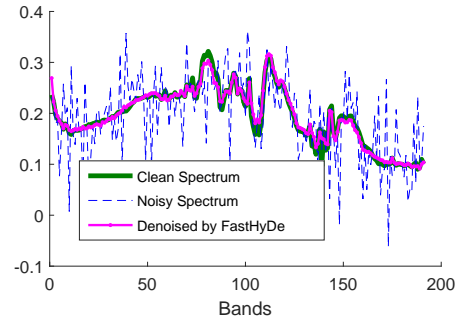

Fig. 2. A denoised spectral signature of the Washington DC Mall image in Case 2.

4. EVALUATION WITH REAL DATA

In this section we apply FastHyde to the AVIRIS (airborne visible/infrared imaging spectrometer) Indian Pines scene. This image, recorded over North- western Indiana in June 1992, has 145×145 pixels, with a spatial resolution of 20 meters per pixel and 220 spectral channels. The image displays strong noise in a number of bands. Fig. 3(a) shows a color image by combining three of these, namely bands 1,

Table 1. Quantitative assessment of different denosing algorithms applied to Wahsington DC Mall.

	σ	Index	Noisy Image	BM3D	BM4D	PCA+BM4D	NAILRMA	FastHyDe
	0.02	MPSNR(dB)	33.98	28.78	42.70	45.14	**46.59**[*]	46.26
		MSSIM	0.9234	0.7981	0.9901	0.9933	0.9936[*]	**0.9957**
	0.04	MPSNR(dB)	27.96	28.31	38.34	40.8758	42.64[*]	**42.73**
		MSSIM	0.7775	0.7786	0.9746	0.9830	0.9864[*]	**0.9913**
Case 1	0.06	MPSNR(dB)	24.44	27.77	35.83	38.31	40.12[*]	**40.56**
		MSSIM	0.6408	0.7546	0.9562	0.9709	0.973[*]	**0.9863**
	0.08	MPSNR(dB)	21.94	27.27	34.09	35.99	38.21[*]	**39.28**
		MSSIM	0.5286	0.7313	0.9365	0.9525	0.9588[*]	**0.9817**
	0.1	MPSNR(dB)	20.00	26.79	32.76	34.72	36.92[*]	**38.05**
		MSSIM	0.4396	0.7090	0.9160	0.9379	0.9425[*]	**0.9762**
Case 2		MPSNR(dB)	28.18	27.95	35.81	37.12	42.05[*]	**45.46**
		MSSIM	0.6971	0.7615	0.9523	0.9500	0.9882[*]	**0.9949**
Case 3		MPSNR(dB)	28.17	29.41	38.86	41.08	-	**41.83**
		MSSIM	0.8073	0.8285	0.9778	0.9850	-	**0.9895**

Note: the numbers with ∗ are cited from [8]

Table 2. Computational time (seconds) of different denoisng algorithms applied to Washington DC Mall.

	σ	BM3D	BM4D	PCA+BM4D	FastHyDe
	0.02	106	647	601	12
	0.04	112	651	608	9
Case 1	0.06	117	669	634	7
	0.08	116	640	609	7
	0.1	120	640	645	6
Case 2		112	644	613	12
Case 3		114	654	628	8

103, and 220 [8]. Indian Pine image was denoised by BM3D, BM4D, 'PCA+BM4D' and FastHyDe, under the assumption of non-i.i.d. noise. The results are exhibited in Fig. 3 (b-e) and corresponding computational times are reported in figure caption. Qualitatively, FastHyDe yields the best result, in the shortest time.
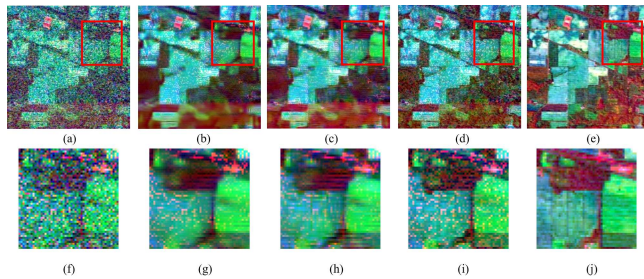

Fig. 3. Denoising results in the real data experiment: (a) original false-color image (R: 1, G: 103, and B: 220); (b) BM3D (40s); (c) BM4D (222s); (d) PCA+BM4D (202s); (e) FastHyDe (9s).

5. CONCLUSIONS

In this paper, we have proposed a new denoising method for HSIs, termed *Fast Hyperspectral denoising* (FastHyDe). The new method exploits two characteristics of HSIs: a) they live in low dimensional subspaces, and b) their images of subspace representation coefficients, herein termed eigen-images, are self-similar and thus suitable to be denoised with non-local patch-based methods. A comparison of FastHyDe with the state-of-the-art algorithms is conducted, leading to the conclusion that FastHyDe yields similar of better performance for additive and Poissonian noise with much lower computational complexity. This characteristics put FastHyDe in a privileged position to be used as an HSI denoiser.

References

[1] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

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

[3] M. Maggioni, G. Boracchi, A. Foi, and K. Egiazarian, "Video denoising, deblocking, and enhancement through separable 4-d nonlocal spatiotemporal transforms," *IEEE Transactions on Image Processing*, vol. 21, no. 9, pp. 3952–3966, Sep. 2012.

[4] Q. Yuan, L. Zhang, and H. Shen, "Hyperspectral image denoising employing a spectral–spatial adaptive total variation model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 10, pp. 3660–3677, Oct. 2012.

[5] H. Othman and S.-E. Qian, "Noise reduction of hyperspectral imagery using hybrid spatial-spectral derivative-domain wavelet shrinkage," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 2, pp. 397–408, Feb. 2006.

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

[7] Q. Wang, L. Zhang, Q. Tong, and F. Zhang, "Hyperspectral imagery denoising based on oblique subspace projection," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 6, pp. 2468–2480, Jun. 2014.

[8] W. He, H. Zhang, L. Zhang, and H. Shen, "Hyperspectral image denoising via noise-adjusted iterative low-rank matrix approximation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 6, pp. 3050–3061, Jun. 2015.

[9] J. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 354–379, Apr. 2012.

[10] J. Bioucas-Dias and J. Nascimento, "Hyperspectral subspace identification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 8, pp. 2435–2445, Aug. 2008.

[11] P. Combettes and V. Wajs, "Signal recovery by proximal forward-backward splitting," *Multiscale Modeling & Simulation*, vol. 4, no. 4, pp. 1168–1200, 2005.

[12] F. Anscombe, "The transformation of poisson, binomial and negative-binomial data," *Biometrika*, pp. 246–254, Dec. 1948.

[13] G. Chen, T. Bui, K. Quach, and S.-E. Qian, "Denoising hyperspectral imagery using principal component analysis and block-matching 4d filtering," *Canadian Journal of Remote Sensing*, vol. 40, no. 1, pp. 60–66, Jan. 2014.