Hyperspectral image denoising and anomaly detection based on low-rank and sparse representations

Lina Zhuang[a], Lianru Gao[b], Bing Zhang[b,c], and José M. Bioucas-Dias[a]

[a]Instituto de Telecomunicações, Instituto Superior Técnico, Universidade de Lisboa, 1049-001, Lisbon, Portugal.
[b]Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, 100094, Beijing, China.
[c]University of Chinese Academy of Sciences, 100049, Beijing, China.

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. Since HSIs represent natural scenes and their spectral channels are highly correlated, they are characterized by a high level of self-similarity and are well approximated by low-rank representations. These characteristic underlies the state-of-the-art in HSI denoising. However, in presence of rare pixels, the denoising performance of those methods is not optimal and, in addition, it may compromise the future detection of those pixels. To address these hurdles, we introduce **RhyDe** (**R**obust **hy**perspectral **De**noising), a powerful HSI denoiser, which implements explicit low-rank representation, promotes self-similarity, and, by using a form of collaborative sparsity, preserves rare pixels. The denoising and detection effectiveness of the proposed robust HSI denoiser is illustrated using semi-real data.

Keywords: HSI denoising, collaborative sparsity, outlier detection, self-similarity, low-rank representation

1. INTRODUCTION

Hyperspectral images (HSIs) have been widely used in countless applications, (e.g., earth observation, environmental protection and natural disaster monitoring), since they have high spectral resolution (hundreds or thousands of spectral channels), which enables material identification with precision via spectroscopic analysis. However, the measurement noise often precludes the widespread use of HSIs in precise material identification (e.g., precision farming) applications.

Among the recent developments, low-rank and self-similarity based image denoising holds the state-of-the-art in HSI denoising (e.g., NAILRMA [1] and FastHyDe [2]). However, the presence of rare pixels, which are anomalies whose spectral-spatial characteristics are different from majority of pixels (often called background), degrades the denoising performance and may preclude the future detection of those pixels.

The detection of rare pixels, often termed the anomaly detection problem [3], has been object of considerable researcah efforts, part of then devoted to develop accurate models for the background. Representative models for the background are Gaussian densities [4, 5], mixtures of Gaussian densities [6], and (sparse) linear mixing models [7, 8]. The main objective in anomaly detection it to classify a pixel either as background or as anomaly, if it does not fit the background model. Given the usual complexity of the landcover in remote sensing, that classification problem is often quite hard.

The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7-PEOPLE-2013-ITN) under grant agreement n°607290 SpaRTaN. This work was partially supported by the Fundação para a Ciência e Tecnologia, Portuguese Ministry of Science and Higher Education, project UID/EEA/50008/2013.

Further author E-mail information: (Send correspondence to Lina Zhuang)
Lina Zhuang and José M. Bioucas-Dias: {lina.zhuang, bioucas}@lx.it.pt
Lianru Gao and Bing Zhang: {gaolr, zb}@radi.ac.cn

1.1 Contribution

This work aims at endowing our previous FastHyDe denoiser [2] with the ability to preserve rare pixels. It takes full advantage of the structured sparsity linked with the usual highly sparse distribution of rare pixels in the HSI spatial domain. In addition, we also introduce an anomaly detection method, which is a spin-off of the denoiser algorithm and it is competitive with the state-of-the-art anomaly detection methods.

We name the proposed denoising and anomaly detection framework **R**obust **hy**perspectral **De**noising (**RhyDe**).

2. PROBLEM FORMULATION

2.1 Observation Model

Assume that the dataset contains a small number of spectral/spatial outliers in unknown positions. The outliers are usually rare pixels or pixel corruptions due to malfunction of the sensor. In a way similar to robust PCA [9] and to the formulations [1, 2], we adopt the observation model

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

which assumes that the observed data matrix $\mathbf{Y} \in \mathbb{R}^{n_b \times n}$, with n_b rows (spectral bands) and n columns (spatial pixels) is the sum of matrix $\mathbf{X}$, representing the signal component, with matrix $\mathbf{S}$, representing the rare pixels, and $\mathbf{N}$, representing the noise and modeling errors. Given $\mathbf{Y}$, our objective is the estimation of $\mathbf{X}$ and $\mathbf{S}$, by exploiting the fact that $\mathbf{X}$ is low-rank and self-similar and $\mathbf{S}$ is columnwise sparse. This reasoning has connections with that of robust PCA [9]. There are, however, considerable differences: regarding $\mathbf{X}$, we explicitly enforce low-rank and promote self-similarity via suitable patch-based regularization, whereas robust PCA promotes low-rank via nuclear norm regularization and does not use spatial regularization; regarding $\mathbf{S}$, we promote columnwise sparsity via $\ell_{2,1}$ regularization, whereas robust PCA promotes sparsity of any element of $\mathbf{S}$ via ℓ_1 regularization.

An usual assumption in HSIs is that the columns (spectral vectors) of matrix $\mathbf{X}$ live in a low-dimensional subspace that may be estimated from the observed data $\mathbf{Y}$ with good approximation [2, 10]. Thus, we write $\mathbf{X} = \mathbf{E}\mathbf{Z}$ [2], with $\mathbf{E} \in \mathbb{R}^{n_b \times p}$ and $p \ll n_b$, and $\mathbf{E}$ holding an orthogonal basis for the signal subspace. Hence, the observation model (1) may be written as

$$\mathbf{Y} = \mathbf{E}\mathbf{Z} + \mathbf{S} + \mathbf{N}. \tag{2}$$

As stated above, three characteristics of HSIs are exploited in this paper: a) they are well approximated by low dimensional subspaces, 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, such as the BM3D [11] or LRCF [12], and c) anomalies are often spatially sparse.

2.2 Denoising

Based on the above rationale, we propose to estimate the matrix $\mathbf{Z}$ and the sparse matrix $\mathbf{S}$ representing outliers by solving the optimization

$$\{\widehat{\mathbf{Z}}, \widehat{\mathbf{S}}\} \in \arg\min_{\mathbf{Z},\mathbf{S}} \frac{1}{2}||\mathbf{E}\mathbf{Z} + \mathbf{S} - \mathbf{Y}||_F^2 + \lambda_1 \phi(\mathbf{Z}) + \lambda_2 ||\mathbf{S}^T||_{2,1}, \tag{3}$$

where $||\mathbf{X}||_F^2 = \text{trace}(\mathbf{X}\mathbf{X}^T)$ is the Frobenius norm of $\mathbf{X}$. The first term on the right-hand side represents the data fidelity. The second term is the first regularizer expressing prior information tailored to self-similar images [2, 11–13], and the third term is the second regularizer, the mixed $\ell_{2,1}$ norm of $\mathbf{S}^T$ given by $\|\mathbf{S}^T\|_{2,1} = \sum_{i=1}^{n} \|\mathbf{s}_i\|_2$ ($\mathbf{s}_i$ denotes i-th column of $\mathbf{S}$), which promotes column-wise sparsity among the columns of $\mathbf{S}$ (see, e.g. [14]). Finally, $\lambda_1, \lambda_2 \geq 0$ are regularization parameters, setting the relative weight of the respective regularizers. Assuming that ϕ is a convex function, then the optimization (3) is a convex problem.

Let $\mathbf{A} = [\mathbf{Z}^T \mathbf{S}^T]^T$ be a $(p + n_b) \times n$ matrix gathering the $(p \times n)$ eigen-images $\mathbf{Z}$ and the $(n_b \times n)$ outlier matrix $\mathbf{S}$. Problem (3) can be written as

$$\begin{aligned} \widehat{\mathbf{A}} \in \arg\min_{\mathbf{A}} \frac{1}{2}\|\mathbf{Y} - [\mathbf{E}, \mathbf{I}_{n_b}]\mathbf{A}\|_F^2 + \lambda_1 \phi([\mathbf{I}_p, \mathbf{0}_{(p \times n_b)}]\mathbf{A}) \\ + \lambda_2 \|([\mathbf{0}_{(n_b \times p)}, \mathbf{I}_{n_b}]\mathbf{A})^T\|_{2,1}, \end{aligned} \tag{4}$$

where $\mathbf{I}_a$ denotes an identity matrix of size a and $\mathbf{0}_{(a\times b)}$ is zero matrix of size $a \times b$.

We solve the optimization (4) with CSALSA algorithm [15], which is an instance of ADMM [16] conceived to solve convex optimizations with an arbitrary number of terms. CSALSA starts by converting the original optimization into a constrained one, using variable splitting, as follows:

$$\begin{aligned} \min_{\mathbf{A},\mathbf{V}_1,\mathbf{V}_2,\mathbf{V}_3} \quad & \frac{1}{2}\|\mathbf{Y}-\mathbf{V}_1\|_F^2 + \lambda_1\phi(\mathbf{V}_2) + \lambda_2\|\mathbf{V}_3^T\|_{2,1} \\ \text{s.t.} \quad & \mathbf{V}_1 = [\mathbf{E},\mathbf{I}_{n_b}]\mathbf{A} \\ & \mathbf{V}_2 = [\mathbf{I}_p,\mathbf{0}_{(p\times n_b)}]\mathbf{A} \\ & \mathbf{V}_3 = [\mathbf{0}_{(n_b\times p)},\mathbf{I}_{n_b}]\mathbf{A}. \end{aligned} \tag{5}$$

The augmented Lagrangian function of above optimization is

$$\begin{aligned} L(\mathbf{A},&\mathbf{V}_1,\mathbf{V}_2,\mathbf{V}_3,\mathbf{D}_1,\mathbf{D}_2,\mathbf{D}_3) = \\ & \tfrac{1}{2}\|\mathbf{Y}-\mathbf{V}_1\|_F^2 + \lambda_1\phi(\mathbf{V}_2) + \lambda_2\|\mathbf{V}_3^T\|_{2,1} \\ & + \tfrac{\mu_1}{2}\|\mathbf{V}_1 - [\mathbf{E},\mathbf{I}_{n_b}]\mathbf{A} - \mathbf{D}_1\|_F^2 \\ & + \tfrac{\mu_2}{2}\|\mathbf{V}_2 - [\mathbf{I}_p,\mathbf{0}_{(p\times n_b)}]\mathbf{A} - \mathbf{D}_2\|_F^2 \\ & + \tfrac{\mu_3}{2}\|\mathbf{V}_3 - [\mathbf{0}_{(n_b\times p)},\mathbf{I}_{n_b}]\mathbf{A} - \mathbf{D}_3\|_F^2, \end{aligned} \tag{6}$$

where $\mu_1, \mu_2, \mu_3 > 0$ are the CSALSA penalty parameters.

The application of CSALSA to (6) leads to **Algorithm 1**. The optimizations of lines 3 and 4 are quadratic

Algorithm 1 **R**obust **hy**perspectral **De**noising (**RhyDe**)

1: Set $k = 0$, choose $\mu_1, \mu_2, \mu_3 > 0$, $\mathbf{V}_{1,0}$, $\mathbf{V}_{2,0}$, $\mathbf{D}_{1,0}$, $\mathbf{D}_{2,0}$.
2: **repeat**
3: $\quad \mathbf{A}_{k+1} = \arg\min_{\mathbf{A}} \frac{\mu_1}{2}\|\mathbf{V}_{1,k} - [\mathbf{E},\mathbf{I}_{n_b}]\mathbf{A} - \mathbf{D}_{1,k}\|_F^2 + \frac{\mu_2}{2}\|\mathbf{V}_{2,k} - [\mathbf{I}_p,\mathbf{0}_{(p\times n_b)}]\mathbf{A} - \mathbf{D}_{2,k}\|_F^2 + \frac{\mu_3}{2}\|\mathbf{V}_{3,k} - [\mathbf{0}_{(n_b\times p)},\mathbf{I}_{n_b}]\mathbf{A} - \mathbf{D}_{3,k}\|_F^2$
4: $\quad \mathbf{V}_{1,k+1} = \arg\min_{\mathbf{V}_1} \frac{1}{2}\|\mathbf{Y}-\mathbf{V}_1\|_F^2 + \frac{\mu_1}{2}\|\mathbf{V}_1 - [\mathbf{E},\mathbf{I}_{n_b}]\mathbf{A}_{k+1} - \mathbf{D}_{1,k}\|_F^2$
5: $\quad \mathbf{V}_{2,k+1} = \arg\min_{\mathbf{V}_{2,k+1}} \lambda_1\phi(\mathbf{V}_2) + \frac{\mu_2}{2}\|\mathbf{V}_2 - [\mathbf{I}_p,\mathbf{0}_{(p\times n_b)}]\mathbf{A}_{k+1} - \mathbf{D}_{2,k}\|_F^2$
6: $\quad \mathbf{V}_{3,k+1} = \arg\min_{\mathbf{V}_{3,k+1}} \lambda_2\|\mathbf{V}_3^T\|_{2,1} + \frac{\mu_3}{2}\|\mathbf{V}_3 - [\mathbf{0}_{(n_b\times p)},\mathbf{I}_{n_b}]\mathbf{A}_{k+1} - \mathbf{D}_{3,k}\|_F^2$
7: $\quad \mathbf{D}_{1,k+1} = \mathbf{D}_{1,k} - (\mathbf{V}_{1,k+1} - [\mathbf{E},\mathbf{I}_{n_b}]\mathbf{A}_{k+1})$
8: $\quad \mathbf{D}_{2,k+1} = \mathbf{D}_{2,k} - (\mathbf{V}_{2,k+1} - [\mathbf{I}_p,\mathbf{0}_{(p\times n_b)}]\mathbf{A}_{k+1})$
9: $\quad \mathbf{D}_{3,k+1} = \mathbf{D}_{3,k} - (\mathbf{V}_{3,k+1} - [\mathbf{0}_{(n_b\times p)},\mathbf{I}_{n_b}]\mathbf{A}_{k+1})$
10: $\quad k \to k+1$
11: **until** stopping criterion is satisfied.

problems, whose solutions are given by

Line 3:

$$\begin{aligned} \mathbf{A}_{k+1} = \quad & \big(\mu_1[\mathbf{E},\mathbf{I}_{n_b}]^H[\mathbf{E},\mathbf{I}_{n_b}] + \mu_2[\mathbf{I}_p,\mathbf{0}_{p\times n_b}]^H \\ & [\mathbf{I}_p,\mathbf{0}_{p\times n_b}] + \mu_3[\mathbf{0}_{(n_b\times p)},\mathbf{I}_{n_b}]^H[\mathbf{0}_{(n_b\times p)},\mathbf{I}_{n_b}]\big)^{-1} \\ & \big(\mu_1[\mathbf{E},\mathbf{I}_{n_b}]^H(\mathbf{V}_1 - \mathbf{D}_1) + \mu_2[\mathbf{I}_p,\mathbf{0}_{p\times n_b}]^H \\ & (\mathbf{V}_2 - \mathbf{D}_2) + \mu_3[\mathbf{0}_{(n_b\times p)},\mathbf{I}_{n_b}]^H(\mathbf{V}_3 - \mathbf{D}_3)\big), \end{aligned} \tag{7}$$

Line 4:

$$\mathbf{V}_{1,k+1} = (1+\mu_1)^{-1}\big[\mathbf{Y} + \mu_1([\mathbf{E},\mathbf{I}_{n_b}]\mathbf{A}_{k+1} + \mathbf{D}_{1,k})\big]. \tag{8}$$

Lines 5 and 6: these optimizations are the proximity operators (POs) of ϕ applied to $\mathbf{V}'_{2,k} = [\mathbf{I}_p,\mathbf{0}_{(p\times n_b)}]\mathbf{A}_{k+1} + \mathbf{D}_{2,k}$, and of the $\ell_{2,1}$ norm applied to $\mathbf{V}'_{3,k} = [\mathbf{0}_{(n_b\times p)},\mathbf{I}_{n_b}]\mathbf{A}_{k+1} + \mathbf{D}_{3,k}$, respectively. That is

$$\mathbf{V}_{2,k+1} = \Psi_{\lambda_1\phi/\mu_2}(\mathbf{V}'_{2,k}), \tag{9}$$

where

$$\Psi_{\lambda\phi}(\mathbf{U}) = \arg\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{X}-\mathbf{U}\|_F^2 + \lambda\phi(\mathbf{X}), \tag{10}$$

and

$$\mathbf{V}_{3,k+1} = \Psi_{\lambda_2\|\cdot\|_{2,1}/\mu_3}(\mathbf{V}'_{3,k}), \tag{11}$$

where, given $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_u]$,

$$\Psi_{\lambda\|\cdot\|_{2,1}}(\mathbf{U}) = \arg\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{X} - \mathbf{U}\|_F^2 + \lambda\|\mathbf{X}^T\|_{2,1} \tag{12}$$

$$= [\text{soft-vector}(\mathbf{u}_i, \lambda/\mu_3),\ i = 1, \ldots, n], \tag{13}$$

and soft-vector$(\mathbf{x}, \tau)$ is the vector-soft-threshold [17]

$$\mathbf{x} \mapsto \frac{\max(\|\mathbf{x}\| - \tau, \mathbf{0})}{\max(\|\mathbf{x}\| - \tau, \mathbf{0}) + \tau}\mathbf{x}.$$

Actually, $\mathbf{V}'_{3,k}$ and $\mathbf{V}'_{2,k}$ can be regarded, respectively, as an estimate of noisy outlier image, which only contains the information about outliers, and an estimate of noisy eigen-images, respectively, in the k-th iteration. Functions $\Psi_{\lambda_1\phi/\mu_2}(\mathbf{V}'_{2,k})$ and $\Psi_{\lambda_2\|\cdot\|_{2,1}/\mu_3}(\mathbf{V}'_{3,k})$ play the role of denoisers of the respective noisy images.

Here we use the plug-and-play approach in solving the subproblem w.r.t. $\widehat{\mathbf{Z}}$ [18]. By assuming that ϕ is decoupled w.r.t. to the bands of $\mathbf{Z}$, and noting that $\mathbf{E}$ is orthogonal, then the plug-and-play step w.r.t. $\mathbf{Z}$ amounts to apply an available denoiser, such as the BM3D [11] or LRCF [12], to each band of $\mathbf{Z}$. Since those denoisers are not POs, we do not have converge guarantee for the implemented variant of CSALSA. The convergence of the plug-and-play iterative procedures is currently an active area of research [18]. In our case, we have systematically observed converge of RhyDe with the augmented Lagrangian parameters set to $\mu_i = 1$, for $i = 1, 2, 3$.

2.3 Anomaly detection

We propose an anomaly detector derived from the estimate of outlier matrix $\widehat{\mathbf{S}}$ in (3), i.e.,

$$r_i = \|\hat{\mathbf{s}}_i\|_2, \quad i = 1, \ldots, n, \tag{14}$$

where $\hat{\mathbf{s}}_i$ is the i-th column of outlier matrix $\widehat{\mathbf{S}}$. If r_i is larger than a threshold, then the i-th pixel is classified as an anomalous pixel.

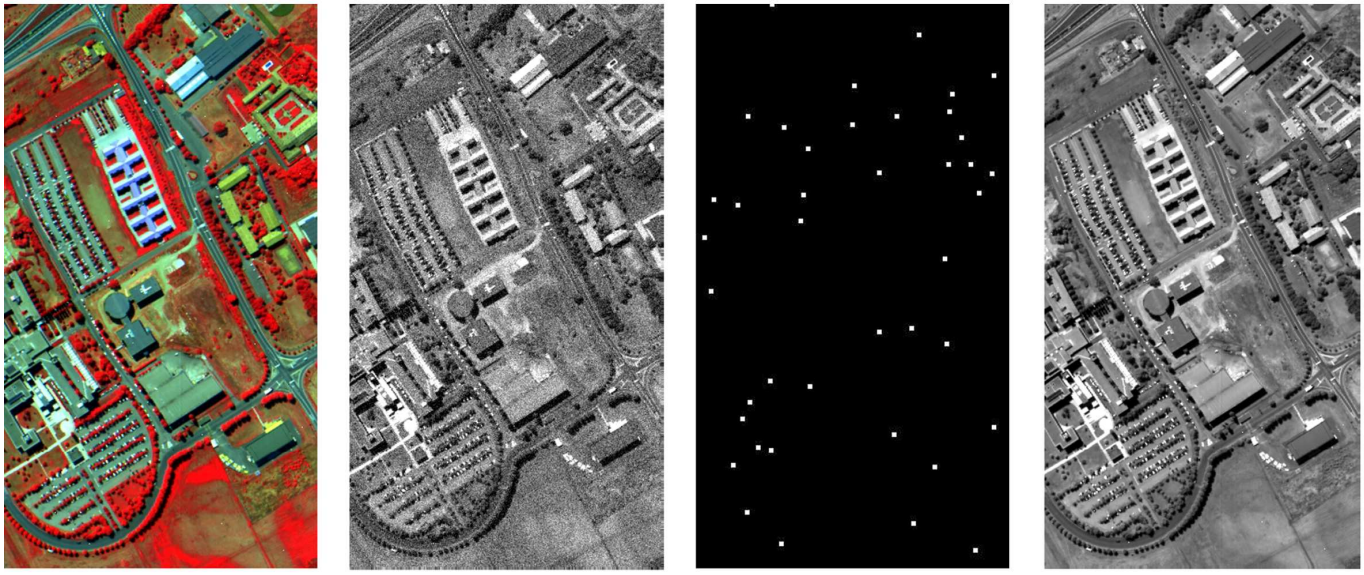
3. EXPERIMENTAL RESULTS

3.1 Dataset

A semi-real hyperspectral dataset (Fig.1-b) was simulated by adding Gaussian noise and anomaly pixels in Pavia University data* as follows: To simulate a clean image, 28 very low signal-to-noise bands were removed. The remaining spectral vectors are then projected onto the signal subspace learned via singular value decomposition (SVD). The projection largely removes the noise. Fig.1-a shows a false color image produced with three bands taken from the clean image.

In order to simulate outliers, we replaced randomly selected 0.02% of the pixels (Fig.1-c) with the spectral signature of the mineral Sillimanite HS186.3B available in the USGS spectral library. By doing this, a clean image containing outliers was obtained. Finally, Gaussian independent with band-dependent variance noise was added to simulated noisy images (Fig. 1-b): $\mathbf{n_i} \sim \mathcal{N}(\mathbf{0}, \mathbf{D})$ where $\mathbf{D}$ is a diagonal matrix with diagonal elements sampled from a Uniform distribution $U(0, u)$ with $u \in \{0.04, 0.02, 0.013, 0.007, 0.004\}$.

* Pavia scenes were provided by Prof. Paolo Gamba from the Telecommunications and Remote Sensing Laboratory, Pavia university (Italy) and can be downloaded from http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes.



(a) Clean image (b) Noisy band (c) Groundth of rare pixels (d) Denoised band by RhyDe

Figure 1. (a) Clean Pavia University scene, (b) A subset of simulated noisy image (Band 61, 20.31 dB) with 0.02% of outliers, (c) Groundtruth of outliers, and (d) A subset of denoised image by RhyDe(Band 61).

3.2 Denoising performance

The denoising performance of RhyDe is compared with those of BM4D [19], NAILRMA [1], and FastHyDe [2]. Since the simulated noise is band variant and all compared methods assume i.i.d noise, the observed data is whitened before the denoisers are applied. For more details see [2].

For quantitative assessment, the signal-to-noise (SNR) index and the structural similarity (SSIM) index [2] of each band are calculated. The corresponding mean SNR (MSNR) and mean SSIM (MSSIM) are reported in Table 1, where we can see that the RhyDe yields uniformly the best performance. The quality of reconstructed spectra may also be inferred from Fig. 2. The results indicate that the strategy in FastHyDe, i.e., combining the prior knowledge self-similarity (also utilized in BM4D) and low-rank (also utilized in NAILRMA) works better than that just using one of them. But the low-rank regularization (NAILRMA and FastHyDe) may lead to loss of rare pixels. In contrast with this scenario, RhyDe is able to preserve anomaly pixels (Fig. 2-right) since it represents rare pixels as a columnwise sparse matrix.

3.3 Anomaly detection performance

The proposed anomaly detector (14) is compared with the state-of-the-art, namely with the global RX [4], local RX [4], OSP global RX [5], OSP local RX [5], NRS [7], and BSJSBD [8]. Receiver operating characteristic (ROC) curve plotting the detection rate against false alarm rate is usually adopted to illustrate the performance of a detector. Here we report the minimum false alarm rates when the detectors reach 100% detection rates in Fig. 3. The detection based on RhyDe yields almost uniformly the best detection results.

4. CONCLUSION

We have proposed RhyDe, a new low-rank based denoising method with preservation of rare pixels. As an extension of FastHyDe, the new method exploits the following characteristics of HSIs: low-rank and self-similarity of clean HSIs, and structured sparsity of outlier matrix. A comparison with the state-of-the-art algorithms was conducted, leading to the conclusion that RhyDe yields better performance for additive noise with preservation of rare pixels. Compared with the state-of-the-art anomaly detector, the derived detector shows superior detection performance.

Table 1. Quantitative assessment of different denoising algorithms applied to semi-real dataset.

Index	Noisy Image	BM4D	NAILRMA	FastHyDe	RhyDe
MSNR (dB)	20.31	28.35	33.25	38.13	**38.58**
MSSIM	0.8295	0.9819	0.9949	0.9982	**0.9983**
Time (Seconds)	-	893	480	**25**	213
MSNR (dB)	25.65	32.37	37.10	42.24	**43.16**
MSSIM	0.9303	0.9932	0.9973	0.9991	**0.9992**
Time (Seconds)	-	881	573	**25**	214
MSNR (dB)	30.84	35.55	40.96	47.58	**49.34**
MSSIM	0.9697	0.9963	0.9988	**0.9997**	**0.9997**
Time (Seconds)	-	884	711	**25**	218
MSNR (dB)	35.40	40.38	44.90	47.36	**51.11**
MSSIM	0.9910	0.9987	0.9995	0.9998	**0.9999**
Time (Seconds)	-	884	718	**26**	220
MSNR (dB)	40.19	43.89	48.19	49.56	**54.27**
MSSIM	0.9964	0.9994	0.9998	**0.9999**	**0.9999**
Time (Seconds)	-	886	675	**26**	222

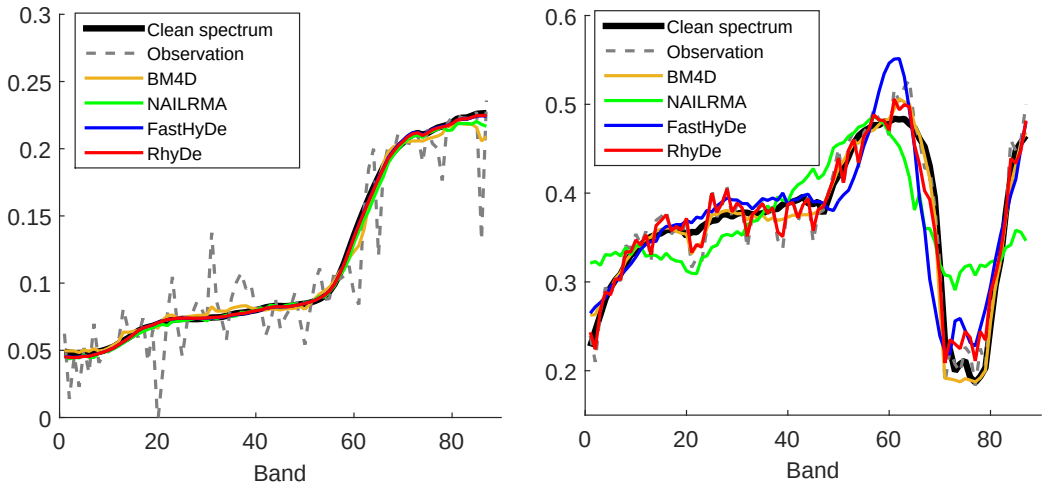


Figure 2. Denoised spectral signatures of a normal pixel (left) and denoised spectral signatures of a anomaly pixel (right) in simulated noisy data (20.31 dB) with 0.02% of outliers. Note that the noise in anomaly pixel is not removed completely by RhyDe since our main objective w.r.t. anomalies is to keep them rather than to denoise them and our output result is $\widehat{\mathbf{Z}} + \widehat{\mathbf{S}}$.

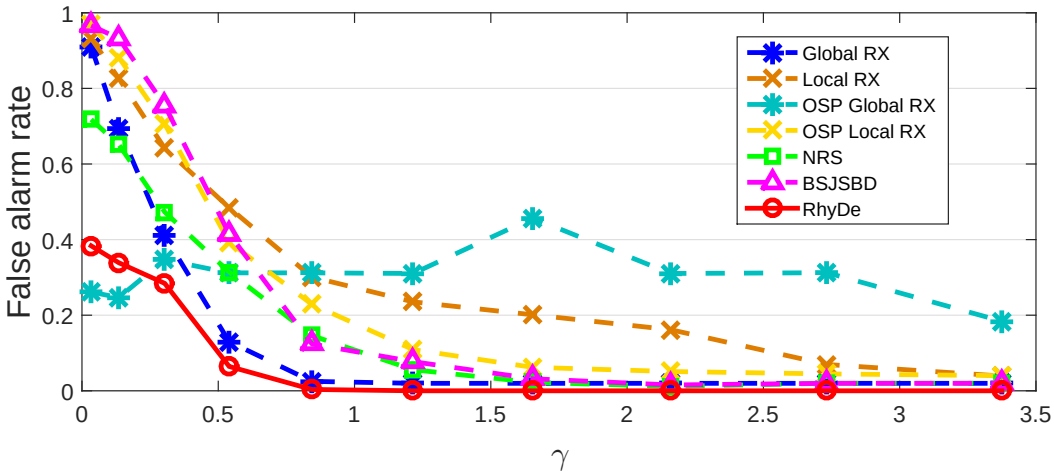


Figure 3. False alarm rate as a function of the relative power of the rare pixels that lies orthogonal complement of the signal subspace, denoted ad γ, for SNR = 20.31dB. As γ decreases, the detection of the outliers becomes more difficult.

References

[1] He, W., Zhang, H., Zhang, L., and Shen, H., "Hyperspectral image denoising via noise-adjusted iterative low-rank matrix approximation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **8**, 3050–3061 (Jun. 2015).

[2] Zhuang, L. and Bioucas-Dias, J., "Fast hyperspectral image denoising based on low rank and sparse representations," in [*IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*], (Jul. 2016).

[3] Matteoli, S., Diani, M., and Corsini, G., "A tutorial overview of anomaly detection in hyperspectral images," *IEEE Aerospace and Electronic Systems Magazine* **25**, 5–28 (Jul. 2010).

[4] Reed, I. S. and Yu, X., "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Transactions on Acoustics, Speech, and Signal Processing* **38**, 1760–1770 (Oct. 1990).

[5] Ma, L. and Tian, J., "Anomaly detection for hyperspectral images based on improved RX algorithm," in [*International Symposium on Multispectral Image Processing and Pattern Recognition*], 67870Q–67870Q, International Society for Optics and Photonics (Nov. 2007).

[6] Stein, D. W. J., Beaven, S. G., Hoff, L. E., Winter, E. M., Schaum, A. P., and Stocker, A. D., "Anomaly detection from hyperspectral imagery," *IEEE Signal Processing Magazine* **19**, 58–69 (Jan. 2002).

[7] Li, W. and Du, Q., "Collaborative representation for hyperspectral anomaly detection," *IEEE Transactions on Geoscience and Remote Sensing* **53**, 1463–1474 (Mar. 2015).

[8] Li, J., Zhang, H., Zhang, L., and Ma, L., "Hyperspectral anomaly detection by the use of background joint sparse representation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **8**, 2523–2533 (Jun. 2015).

[9] Wright, J., Ganesh, A., Rao, S., Peng, Y., and Ma, Y., "Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization," in [*Advances in neural information processing systems*], 2080–2088 (2009).

[10] Bioucas-Dias, J., Plaza, A., Dobigeon, N., Parente, M., Du, Q., Gader, P., and Chanussot, J., "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **5**, 354–379 (Apr. 2012).

[11] Dabov, K., Foi, A., Katkovnik, V., and Egiazarian, K., "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Transactions on Image Processing* **16**, 2080–2095 (Aug. 2007).

[12] Nejati, M., Samavi, S., Soroushmehr, S., and Najarian, K., "Low-rank regularized collaborative filtering for image denoising," in [*IEEE International Conference on Image Processing (ICIP)*], 730–734, IEEE (Sep. 2015).

[13] Buades, A., B-Coll, and Morel, J.-M., "A non-local algorithm for image denoising," in [*IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*], **2**, 60–65, IEEE (Jun. 2005).

[14] Tropp, J. A., Gilbert, A. C., and Strauss, M. J., "Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit," *Signal Processing* **86**, 572–588 (Mar. 2006).

[15] Afonso, M., Bioucas-Dias, J., and Figueiredo, M., "Fast image recovery using variable splitting and constrained optimization," *IEEE Transactions on Image Processing* **19**, 2345–2356 (Sep. 2010).

[16] Eckstein, J. and Bertsekas, "On the DouglasRachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming* **55**, 293–318 (Apr. 1992).

[17] Combettes, P. and Patric, J.-C., "Proximal splitting methods in signal processing," in [*Fixed-point algorithms for inverse problems in science and engineering*], 185–212, Springer (2011).

[18] Venkatakrishnan, S. V., Bouman, C. A., and Wohlberg, B., "Plug-and-play priors for model based reconstruction," in [*IEEE Global Conference on Signal and Information Processing*], 945–948 (Dec. 2013).

[19] Maggioni, M., Katkovnik, V., Egiazarian, K., and Foi, A., "Nonlocal transform-domain filter for volumetric data denoising and reconstruction," *IEEE Transactions on Image Processing* **22**, 119–133 (Jan. 2013).