



Article

Hyperspectral Image Denoising Based on Principal-Third-Order-Moment Analysis

Shouzhi Li ^{1,2,3,*}, Xiurui Geng ^{1,2}, Liangliang Zhu ⁴, Luyan Ji ^{1,2} and Yongchao Zhao ²

¹ Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China; gengxr@sina.com.cn (X.G.); jily@mail.ustc.edu.cn (L.J.)

² Key Laboratory of Technology in Geo-Spatial Information Processing and Application System, Chinese Academy of Sciences, Beijing 100190, China; ofcours_sure@sina.com.cn

³ School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China

⁴ School of Computer Science and Information Engineering, Hefei University of Technology, Hefei 242099, China; llz@hfut.edu.cn

* Correspondence: lishouzhi19@ucas.ac.cn

Abstract: Denoising serves as a critical preprocessing step for the subsequent analysis of the hyperspectral image (HSI). Due to their high computational efficiency, low-rank-based denoising methods that project the noisy HSI into a low-dimensional subspace identified by certain criteria have gained widespread use. However, methods employing second-order statistics as criteria often struggle to retain the signal of the small targets in the denoising results. Other methods utilizing high-order statistics encounter difficulties in effectively suppressing noise. To tackle these challenges, we delve into a novel criterion to determine the projection subspace, and propose an innovative low-rank-based method that successfully preserves the spectral characteristic of small targets while significantly reducing noise. The experimental results on the synthetic and real datasets demonstrate the effectiveness of the proposed method, in terms of both small-target preservation and noise reduction.

Keywords: hyperspectral image; denoising; low-rank approximation



Citation: Li, S.; Geng, X.; Zhu, L.; Ji, L.; Zhao, Y. Hyperspectral Image Denoising Based on Principal-Third-Order-Moment Analysis. *Remote Sens.* **2024**, *16*, 276. <https://doi.org/10.3390/rs16020276>

Academic Editor: Salah Bourennane

Received: 10 November 2023

Revised: 28 December 2023

Accepted: 5 January 2024

Published: 10 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

A hyperspectral image (HSI) is composed of tens or hundreds of spectral reflectance bands, providing abundant spectral information about the objects presenting in the image. Thus, HSIs have been applied in various fields of remote sensing, such as classification [1–3] and target detection [4–6]. In comparison to an RGB-remote-sensing image, an HSI exhibits relatively lower spatial resolution, resulting in many interested man-made small-sized targets in the scene being represented by only a few pixels. The detection of the targets heavily relies on their distinctive spectral characteristic. However, HSIs are often corrupted by noise during acquisition and transmission, which will drown out these spectral characteristics and degrade the accuracy of target detection. Therefore, denoising plays a vital role as a preprocessing step for HSI. The denoising problem is commonly addressed by incorporating the prior knowledge [7] of the HSI signal. Existing denoising algorithms can be broadly categorized into two groups, based on the prior knowledge they employ.

The methods in the first category [8–14] utilize the spatial-smoothness prior of the HSI to build a restoration result. For instance, Zhao [12] proposed the spectral-recognition-spatial-smooth-hyper-spectral filter (SRSSHf), which restores each pixel by taking a weighted average of its neighboring pixels. The weights are determined by the correlation coefficients between the central pixel and its neighboring pixels. This method can effectively preserve the edges in the restoration result. Considering a situation in which the noise level varies with different bands, Yuan [10] proposed the spectral-spatial kernel (SSK), which alleviates the influence of the bands with high noise levels on the calculation of the

weights. Additionally, Buades [15] introduced ‘patch’, which consists of all the pixels in a local region as the fundamental unit, and proposed the non-local-means-(NLM) method, which determines the weights based on the distances between neighboring patches and the central patch. Because the patch provides additional local spatial information, the NLM method achieves improved denoising performance. Further, Maggioni [14] proposed the block-matching-and-4D-filtering-(BM4D) method, which aggregates similar patches and performs collaborative filtering in a transform domain. The above methods only utilize local spatial information for denoising.

In addition, the spatial-smoothness prior also leads to many total-variation-(TV)-regularization-based approaches [16–18] that exploit the smoothness assumption of the entire HSI. Rudin [18] initially introduced the original TV-denoising method, with the aim of finding a restoration image that approximated the noisy image while minimizing the $L1$ norm of its spatial gradient. Aggarwal [16] observed that the spectral dimension also adheres to the smoothness prior. He introduced gradient information along the spectral direction, leading to the development of the spatio-spectral-total-variation-(SSTV) method. Recently, Takemoto [17] proposed the graph-SSTV method, which can better preserve the complex spatial structure in the restoration result. However, by this kind of method, the signal of the small targets with only one or sub-pixel size in the image is often mistakenly considered as noise and is removed in the restoration results.

The methods in the second category are based on the low-rank prior [19–28], and they generally perform notable computational efficiency. Methods of this kind assume that an HSI signal concentrates in a low-dimensional subspace, while the noise exists throughout the space. Consequently, they generally perform denoising by projecting the noisy HSI into the subspace identified by a certain criterion. Many of them employ the variance of the projection components, to evaluate the signal energy present in the corresponding direction. Based on the assumption that noise follows an independent-and-identically-distributed-(i.i.d.) Gaussian distribution, Richards [29] used the first few principal components of the noisy HSI matrix to construct the objective subspace. However, in real-world scenarios, the noise in a natural HSI generally exhibits a complex distribution. Chen [30] observed that the noise distributions of different bands have certain correlations and proposed a non-i.i.d mixture of Gaussian-(NMoG)-noise model to fit the real noise. Xu [31] utilized the bandwise-asymmetric-Laplacian-(BAL) distribution to model the real noise, introducing the BAL-matrix-factorization-(BALMF) method. Additionally, Zhuang [32] and Fu [33] assumed that real noise is a mixture of Gaussian noise and sparse noise and proposed the FastHyMix and RoSEGS methods, respectively. In both methods, the sparse noise is first located and masked, to avoid influencing the estimation of the signal subspace.

Moreover, researchers [26,34] have noticed that HSI signals also exhibit significant high-order statistical features in the signal subspace. They have proposed methods that utilize the high-order statistics of the projection components as criteria. For instance, Chiang [26] extracted the subspace corresponding to the maximum skewness, and the projection results achieved remarkable performance in the target-detection task. Geng [28] developed the principal-skewness-analysis-(PSA) method, which can rapidly extract an arbitrary number of the maximum skewness components. Wang [34] explored the application of the independent-component-analysis-(ICA) method to HSI, utilizing a mixture of skewness and kurtosis as the criterion.

It should be noted that existing low-rank-based denoising methods face a challenge in simultaneously preserving the spectra of small targets and suppressing noise. As small targets constitute only a minor portion of the overall HSI signal, they do not contribute to significant second-order statistical characteristics. Therefore, methods relying on variance as a criterion often struggle to identify the subspace corresponding to the small-target signal. Consequently, the spectral characteristic of the small targets is generally eliminated in the denoised results of the variance-based methods, severely affecting the detection of these small targets. By contrast, high-order-statistic-based methods can effectively capture the target signal, due to their sensitivity to minority components within the signal. However,

the values of skewness and kurtosis are easily influenced by noise, causing these methods to fail to accurately identify the signal subspace in noisy situations (which will be further discussed in the subsequent section). As a result, evident residual noise generally persists in the denoised results of high-order-statistic-based methods.

To overcome these limitations, this paper aims to investigate a new low-rank-based denoising method that can preserve the small target signal while reducing noise. And the pivotal factor hinges on finding a new criterion that is sensitive to the minor portion of the signal and robust to the noise.

This paper is organized as follows. In Section 2, we briefly introduce the PSA algorithm. In Section 3, we propose a novel denoising method. The experimental results are shown in Section 4. Finally, we conclude this paper in Section 5.

2. Background

In this subsection, we provide a brief introduction to the principal-skewness-analysis (PSA) [28] algorithm that is designed to calculate the component with maximum skewness. The commonly used symbols in this paper are summarized at the end of this paper.

Let $\mathbf{X} \in \mathbb{R}^{N \times L}$ represent a noise-free HSI with N pixels and L bands. For simplicity, all HSIs used in this paper have been normalized—that is:

$$\frac{1}{N} \sum_{i=1}^N X_{i,j} = 0, \forall j = 1, 2, \dots, L. \quad (1)$$

Here, $X_{i,j}$ denotes the element located at the i_{th} row and j_{th} column of $\mathbf{X}$. The PSA algorithm initially whitens the data, using the following equation:

$$\mathbf{R} = \mathbf{X}\mathbf{F}, \quad (2)$$

where $\mathbf{R}$ is the whitened data; $\mathbf{F} = \mathbf{E}\mathbf{D}^{-\frac{1}{2}}$ is the whitening operator. Here, $\mathbf{E}$ and $\mathbf{D}$ are, respectively, the eigenvector matrix and the eigenvalue matrix of the covariance matrix of $\mathbf{X}$. Then, the PSA algorithm constructs the co-skewness tensor, calculated as

$$\mathcal{S} = \frac{1}{L} \sum_{i=1}^L \mathbf{R}_i \circ \mathbf{R}_i \circ \mathbf{R}_i. \quad (3)$$

In this equation, $\mathbf{R}_i$ represents the i_{th} column of $\mathbf{R}$, $\circ$ represents the outer product and $\mathcal{S} \in \mathbb{R}^{L \times L \times L}$ is the co-skewness tensor. Subsequently, the maximum skewness direction can be determined by solving the following optimization model:

$$\begin{aligned} \max_{\mathbf{u}} \quad & \mathcal{S} \times_1 \mathbf{u} \times_2 \mathbf{u} \times_3 \mathbf{u} \\ \text{s.t.} \quad & \mathbf{u}^T \mathbf{u} = 1, \end{aligned} \quad (4)$$

where $\mathbf{u} \in \mathbb{R}^{L \times 1}$ represents an arbitrary direction vector. It should be noted that the solution of Equation (4) is the maximum skewness direction of $\mathbf{R}$, and the maximum skewness direction of $\mathbf{X}$ can be obtained by multiplying it by $\mathbf{E}\mathbf{D}^{-\frac{1}{2}}$. And the PSA exhibits effectiveness [28] on extracting the signal subspace of the noise-free HSI.

3. Method

In this section, we initially investigate the issues of the PSA algorithm in the presence of noise. Subsequently, we exploit a new criterion that is robust to Gaussian noise, to determine the projection subspace, and we develop a novel denoising method based on it.

3.1. Maximum Third-Order Moment Criterion

First, we explore the impact of noise on the projection direction searched by the PSA. Actually, the objective of the PSA algorithm is equivalent to finding a unit direction vector $\mathbf{u}$ that maximizes the following function [28]:

$$\text{skew}(\mathbf{v}) = \frac{\mu_3(\mathbf{v})}{(\text{var}(\mathbf{v}))^{\frac{3}{2}}} = \frac{\frac{1}{N} \sum_{i=1}^N v_i^3}{\left(\frac{1}{N} \sum_{i=1}^N v_i^2\right)^{\frac{3}{2}}}, \quad (5)$$

where $\mathbf{v} = \mathbf{X}\mathbf{u}$ represents the projection coefficients along direction $\mathbf{u}$, and where v_i is the i_{th} element of $\mathbf{v}$. The third-order moment and the variance of $\mathbf{v}$ are, respectively, $\mu_3(\mathbf{v})$ and $\text{var}(\mathbf{v})$. As $\mathbf{X}$ has been normalized, the mean of $\mathbf{v}$ is always zero and is eliminated in Equation (5).

In a noise situation, the observed HSI is usually modeled as

$$\mathbf{Y} = \mathbf{X} + \mathbf{N}, \quad (6)$$

where $\mathbf{Y}$ represents the observed noisy HSI and where $\mathbf{N}$ is the zero-mean Gaussian noise with covariance matrix $\mathbf{C} \in \mathbb{R}^{L \times L}$. We denote $\mathbf{w} = \mathbf{Y}\mathbf{u}$ and $\mathbf{n} = \mathbf{N}\mathbf{u}$ as, respectively, the projection coefficients of $\mathbf{Y}$ and $\mathbf{N}$ along direction $\mathbf{u}$. And in the noisy situation, the PSA algorithm aims to maximize the skewness of $\mathbf{w}$:

$$\begin{aligned} \text{skew}(\mathbf{w}) &= \frac{\mu_3(\mathbf{w})}{(\text{var}(\mathbf{w}))^{\frac{3}{2}}} = \frac{\frac{1}{N} \sum_{i=1}^N (v_i + n_i)^3}{\left(\frac{1}{N} \sum_{i=1}^N (v_i + n_i)^2\right)^{\frac{3}{2}}} \\ &= \frac{\frac{1}{N} \sum_{i=1}^N (v_i^3 + 3v_i^2n_i + 3v_in_i^2 + n_i^3)}{\left(\frac{1}{N} \sum_{i=1}^N (v_i^2 + 2v_in_i + n_i^2)\right)^{\frac{3}{2}}}, \end{aligned} \quad (7)$$

where n_i is the i_{th} element of $\mathbf{n}$. As a linear combination of L Gaussian noise, n_i is a random sample from a Gaussian distribution with zero mean and variance σ^2 , where $\sigma^2 = \mathbf{u}^T \mathbf{C} \mathbf{u}$. According to the law of large numbers, as N increases, we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_i v_i^2 n_i = 0 \quad (8)$$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_i v_i n_i^2 = 0 \quad (9)$$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_i v_i n_i = 0 \quad (10)$$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_i n_i^3 = 0 \quad (11)$$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_i n_i^2 = \sigma^2. \quad (12)$$

As a result, Equation (7) can be approximately simplified as

$$\text{skew}(\mathbf{w}) = \frac{\frac{1}{N} \sum_{i=1}^N v_i^3}{\left(\frac{1}{N} \sum_{i=1}^N v_i^2 + \sigma^2\right)^{\frac{3}{2}}}. \quad (13)$$

Comparing Equation (13) to Equation (5), it is evident that the noise introduces an additional term σ^2 in the denominator, resulting in a reduction in skewness along direction $\mathbf{u}$. Additionally, the distribution of the HSI signal in different directions varies, leading to changes in the values of $\sum_{i=1}^N v_i^3$ and $\sum_{i=1}^N v_i^2$, with respect to the projection direction. Consequently, the extent of reduction caused by σ^2 is direction-dependent, suggesting that

the maximum skewness direction of the noisy HSI likely deviates from that of the noise-free HSI. Therefore, the PSA algorithm is unable to accurately identify the signal subspace in the presence of noise.

Fortunately, we find an important conclusion: that the numerator of Equation (13)—namely, the third-order moment—is always equal to that of Equation (5):

$$\mu_3(\mathbf{w}) = \mu_3(\mathbf{v}). \quad (14)$$

Equation (14) indicates that the maximum third-order-moment direction remains unaffected by the Gaussian noise. Furthermore, it has been observed [35] that the third-order-moment value is also sensitive to the minority components in the signal, which implies that the third-order moment retains the property of preserving the small target. Therefore, we consider utilizing the third-order moment as the criterion for determining the projection subspace for denoising.

3.2. Principal Third-Order-Moment Analysis

In this subsection, we develop the principal-third-order-moment-analysis-(PTMA) algorithm, to calculate the maximum third-order-moment directions and apply the PTMA for denoising. The procedure of the PTMA is as follows:

Inspired by the PSA algorithm, we first construct the co-third-order-moment tensor $\mathcal{M}$ as

$$\mathcal{M} = \frac{1}{L} \sum_{i=1}^L \mathbf{Y}_i \circ \mathbf{Y}_i \circ \mathbf{Y}_i, \quad (15)$$

where $\mathbf{Y}_i$ is the i_{th} column of $\mathbf{Y}$. Then, the first third-order-moment direction can be determined by solving the following optimization problem:

$$\begin{aligned} & \underset{\mathbf{u}}{\operatorname{argmax}} \quad \mathcal{M} \times_1 \mathbf{u} \times_2 \mathbf{u} \times_3 \mathbf{u} \\ & \text{s.t.} \quad \mathbf{u}^T \mathbf{u} = 1. \end{aligned} \quad (16)$$

The Lagrangian function of Equation (16) is

$$\mathcal{L}(\mathbf{u}, \lambda) = \mathcal{M} \times_1 \mathbf{u} \times_2 \mathbf{u} \times_3 \mathbf{u} - \lambda(\mathbf{u}^T \mathbf{u} - 1). \quad (17)$$

Subsequently, $\mathbf{u}$ that maximizes Equation (16) satisfies the following equation:

$$\mathcal{M} \times_1 \mathbf{u} \times_3 \mathbf{u} = \lambda \mathbf{u}. \quad (18)$$

We apply the fixed-point-iteration method [36] to solve Equation (18), and a specific iteration step is

$$\begin{cases} \mathbf{u} = \mathcal{M} \times_1 \mathbf{u} \times_3 \mathbf{u} \\ \mathbf{u} = \mathbf{u} / \|\mathbf{u}\|_2 \end{cases}. \quad (19)$$

The convergence result (denoted as $\mathbf{u}_1$) is the first maximum-third-order-moment direction. To obtain the second projection direction while ensuring it does not converge to $\mathbf{u}_1$, we project $\mathbf{Y}$ into the orthogonal complement space of $\mathbf{u}_1$ and reconstruct the co-third-order-moment tensor. These two operations are equivalent to directly multiplying the projection matrix by $\mathcal{M}$, as follows:

$$\mathcal{M} = \mathcal{M} \times_1 \mathbf{P}_{\mathbf{u}_1}^\perp \times_2 \mathbf{P}_{\mathbf{u}_1}^\perp \times_3 \mathbf{P}_{\mathbf{u}_1}^\perp, \quad (20)$$

where $\mathbf{P}_{\mathbf{u}_1}^\perp = \mathbf{I} - \mathbf{u}_1 \mathbf{u}_1^T$ is the projection matrix onto the orthogonal complement subspace of vector $\mathbf{u}_1$, and $\mathbf{I}$ is the $L \times L$ identity matrix. Then, the same iteration process, i.e., Equation (19), is applied to the updated $\mathcal{M}$ to calculate the second maximum-third-order-moment direction $\mathbf{u}_2$, and the following process is conducted in the same manner.

We assume r represents the estimation of the dimension of the signal subspace (many algorithms [25,37,38] have been developed to estimate r). The restoration result is $\mathbf{Y}\mathbf{U}\mathbf{U}^T$, where $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r]$. The pseudocode of the PTMA is presented in Algorithm 1. It should be noted that although the PTMA only takes a minor modification of the PSA algorithm, it achieves a significant theoretical improvement in noise suppression.

Algorithm 1 PTMA algorithm

Require: Centralized Noisy HSI data $\mathbf{Y} \in N \times L$, the estimation of the dimension of the objective subspace r

Ensure: Denoising result $\hat{\mathbf{X}}$

- 1: Calculate the co-skewness tensor $\mathcal{M}$ according to Equation (15)
 - 2: Randomly initialize the transformation matrix $\mathbf{U} \in \mathbb{R}^{L \times r}$
 - 3: **for** $i = 1 : r$ **do**
 - 4: Initialize $\mathbf{u}_i$ with random unit vector
 - 5: **while** stop conditions are not met **do**
 - 6: $\mathbf{u}_i = \mathcal{M} \times_1 \mathbf{u}_i \times_3 \mathbf{u}_i$
 - 7: $\mathbf{u}_i = \mathbf{u}_i / \|\mathbf{u}_i\|_2$
 - 8: **end while**
 - 9: $\mathbf{P}_{\mathbf{u}_i}^\perp = \mathbf{I} - \mathbf{u}_i \mathbf{u}_i^T$
 - 10: $\mathcal{M} = \mathcal{M} \times_1 \mathbf{P}_{\mathbf{u}_i}^\perp \times_2 \mathbf{P}_{\mathbf{u}_i}^\perp \times_3 \mathbf{P}_{\mathbf{u}_i}^\perp$
 - 11: $\mathbf{U}[:, i] = \mathbf{u}_i$
 - 12: **end for**
 - 13: Calculate the denoising result by $\hat{\mathbf{X}} = \mathbf{Y}\mathbf{U}\mathbf{U}^T$
-

4. Experiments

For this section, we utilized simulated and real noisy hyperspectral images (HSIs) to assess the denoising and target-preservation capabilities of the PTMA. For comparison, seven advanced denoising methods were used as competitors: the non-local-means-3-D filter (NLM3D) [13], the block-matching-and-4D filter (BM4D) [14], the low-rank matrix recovery (LRMR) [21], global local factorization (GLF) [39], the spectral-spatial kernel (SSK) [10], the spectral-recognition-spatial-smooth-hyper-spectral filter (SRSSHf) [12] and the PSA.

4.1. Synthetic Experiment

For this subsection, synthetic noisy datasets were generated, based on two public HSI datasets: namely, the subregions of the Washington-DC-Mall data and the University-of-Pavia data. The Washington data were acquired by the airborne-visible/infrared-imaging-spectrometer-(AVIRIS) sensor, with a spectral range of 0.4 to 2.4 μm . And the Pavia data were acquired by the reflective-optics-system-imaging-spectrometer-(ROSIS) sensor, with a spatial resolution of 1.3 m.

To obtain relatively clean images, we first removed the bands severely corrupted by water vapor in the atmosphere. Then, the spectral vectors of each image were projected onto a subspace spanned by the first few principal eigenvectors of each image. And the projection result was considered as the reference noise-free HSI. The processed Washington dataset consisted of 200×200 pixels and 191 bands, whose false-color image composed by bands 60, 27, 14 is shown in Figure 1a. The processed Pavia dataset consisted of 200×200 pixels and 87 bands, whose false-color image composed by bands 60, 40, 20 is shown in Figure 1b.

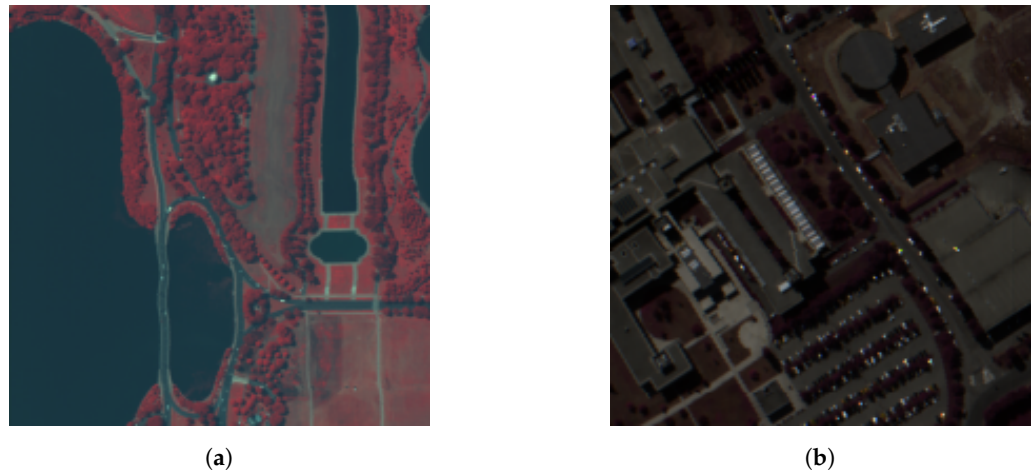


Figure 1. (a) The false-color image composed by bands 60, 27, 14 of the Washington dataset. (b) The false-color image composed by bands 60, 40, 20 of the Pavia dataset.

To generate noisy HSIs, four kinds of noise were added into both clean datasets (both datasets had been normalized). And for each noise type we generated noise at two distinct intensity levels. Specifically, the noise types were as follows.

Case 1 (i.i.d. Gaussian noise): The noise conformed to Gaussian distribution with zero mean. And the standard deviation of the low-intensity situation and the high-intensity situation were set as 0.03 and 0.05, respectively.

Case2 (Non-i.i.d. Gaussian noise): The noise conformed to Gaussian distribution with zero mean and band-varying standard deviation. The standard deviation of the low-intensity situation and the high-intensity situation were sampled from uniform distribution $U(0.01, 0.05)$ and $U(0.01, 0.1)$, respectively.

Case3 (Non-i.i.d. Gaussian noise + Poisson noise): The standard deviation of Gaussian noise in the low- and high-intensity situations were sampled from uniform distribution $U(0.01, 0.02)$ and $U(0.01, 0.07)$, respectively. The Poisson noise was pixel-value correlated, i.e., the output pixel value with Poisson noise was sampled from a Poisson distribution with mean equal to the pixel value. We used a scale factor β to control the intensity of the Poisson noise. For example, if an input pixel value was a , then the corresponding output pixel would be generated from a Poisson distribution with mean of $a \times \beta$ and then divided by β . The scale factor β was set as 7×10^{11} and 5×10^{12} , respectively, for the low-intensity situation and the high-intensity situation.

Case4 (Non-i.i.d. Gaussian noise + Poisson noise + Strip noise): The standard deviations of Gaussian noise in the low- and high-intensity situations were sampled from uniform distribution $U(0.01, 0.02)$ and $U(0.01, 0.03)$, respectively. β was set as 3×10^{11} and 10^{12} . And the number of strips in each band was set as 10 and 30, respectively.

Then, the competitors and our method were employed for denoising. The parameters of the competitors were set as follows. For the synthetic noise, the standard deviation of the noise, which was required as a parameter for BM4D, GLF, the SSK and the SRSHF, was known. For NLM3D, the search-area size, the small-similar-patch size and the big-similar-patch size were set as 7×7 , 3×3 and 5×5 , respectively. For the LRMR, the patch size and the upper bound of the cardinality were set to their default values, 20×20 and 4000, respectively, as recommended in the literature. In addition, GLF and the LRMR, respectively, also required the rank of the signal subspace, which was estimated for each noisy HSI according to the Akaike information criterion.

To quantitatively evaluate the denoising performance of these methods, four metrics were counted, including the mean of the peak signal-to-noise ratio of each band (MPSNR), the mean of the structural similarity index of each band (MSSIM), the mean of the spectral angle mapping (MSAM) and the Erreur Relative Globale Adimensionnelle de Synthèse (ERGAS) [40]. Higher MPSNR, MSSIM values and lower MSAM, ERGAS values generally implied better restoration results.

The MPSNR was calculated as follows:

$$\text{MPSNR} = \frac{1}{L} \sum_{b=1}^L \text{PSNR}_b, \quad (21)$$

$$\text{PSNR}_b = 10 \cdot \log_{10} \left(\frac{\text{MAX}^2}{\text{MSE}_b^2} \right) = 10 \cdot \log_{10} \left(\frac{\text{MAX}^2}{\frac{1}{N} \|\hat{\mathbf{X}}_{:,b} - \mathbf{X}_{:,b}\|_2^2} \right), \quad (22)$$

where PSNR_b denoted the PSNR value of the b_{th} band. MAX was the maximum-possible pixel value of the HSI, which was set as 1 because the HSI has been normalized. $\hat{\mathbf{X}}_{:,b}$ and $\mathbf{X}_{:,b}$ represented the b_{th} band of the denoised HSI and the reference noise-free HSI, respectively.

The MSSIM was calculated as follows:

$$\text{MSSIM} = \frac{1}{L} \sum_{b=1}^L \text{SSIM}_b, \quad (23)$$

$$\text{SSIM}_b = \frac{(2\mu(\hat{\mathbf{X}}_{:,b})\mu(\mathbf{X}_{:,b}) + c_1)(2\sigma(\hat{\mathbf{X}}_{:,b}, \mathbf{X}_{:,b}) + c_2)}{(\mu(\hat{\mathbf{X}}_{:,b})^2 + \mu(\mathbf{X}_{:,b})^2 + c_1)(\sigma(\hat{\mathbf{X}}_{:,b})^2 + \sigma(\mathbf{X}_{:,b})^2 + c_2)}, \quad (24)$$

where SSIM_b was the SSIM value of the b_{th} band, $\mu(\hat{\mathbf{X}}_{:,b})$, $\mu(\mathbf{X}_{:,b})$ represented the mean of $\hat{\mathbf{X}}_{:,b}$ and $\mathbf{X}_{:,b}$, respectively, $\sigma(\hat{\mathbf{X}}_{:,b})$, $\sigma(\mathbf{X}_{:,b})$ were the variance of $\hat{\mathbf{X}}_{:,b}$ and $\mathbf{X}_{:,b}$, respectively, $\sigma(\hat{\mathbf{X}}_{:,b}, \mathbf{X}_{:,b})$ represented their covariance and c_1 and c_2 were constants.

The calculation procedure for the MSAM was

$$\text{SAM} = \frac{1}{N} \sum_{i=1}^N \arccos \left(\frac{\mathbf{X}_i \cdot \hat{\mathbf{X}}_i^T}{\|\mathbf{X}_i\|_2 \cdot \|\hat{\mathbf{X}}_i\|_2} \right), \quad (25)$$

where $\mathbf{X}_i$ and $\hat{\mathbf{X}}_i$ were the i_{th} row of $\mathbf{X}$ and $\hat{\mathbf{X}}$, respectively.

The calculation procedure for the ERGAS was

$$\text{ERGAS} = 100 \sqrt{\frac{1}{L} \sum_{b=1}^L \frac{\|\mathbf{X}_{:,b} - \hat{\mathbf{X}}_{:,b}\|_2}{\mu(\mathbf{X}_{:,b})}}. \quad (26)$$

The metrics of the Washington and Pavia datasets are shown in Tables 1 and 2, respectively. It can be seen that the PTMA achieved the highest MPSNR and MSSIM for most noise types, which indicates that the PTMA is not only applicable to Gaussian noise, but also robust against complex mixed noise. However, considering the ERGAS metric, BM4D obtained the best performance in most situations. It is worth noting that the value of the ERGAS was proportional to the mean of the MSE for each band divided by the average signal value within the respective band and that the PSNR was the inverse of the MSE of the entire HSI. Therefore, the reason that the PTMA achieved higher PSNR and ERGAS values compared to BM4D lay in the fact that for spectral bands with higher signal intensity there was more residual noise in the PTMA than in BM4D, whereas for bands with lower signal intensity, the PTMA demonstrated less residual noise.

We counted the average computation time of the PTMA and the competitors for all the noise cases (Tables 1 and 2). It can be seen that the PTMA was the most computationally efficient algorithm, with a computation time less than 10% of BM4D. In addition, it is worth noting that the PTMA determined the projection direction based on $\mathcal{M} \in \mathbb{R}^{L \times L \times L}$, which meant that the computation complexity of the PTMA was almost uncorrelated to the image size.

Table 1. Performance of the proposed and competitors on Washington data.

	Indexes	Noisy	NLM3D	BM4D	LRMR	GLF	SSK	SRSSHf	PSA	PTMA
Case 1 Low noise level	MPSNR	30.46	42.88	46.37	44.10	37.24	37.04	41.70	45.33	48.45
	MSSIM	0.65	0.97	0.99	0.98	0.93	0.92	0.93	0.94	0.99
	MSAM	10.36	2.39	1.78	2.11	4.14	3.68	3.26	3.58	1.67
	ERGAS	24.68	4.99	3.78	7.49	24.68	7.27	6.33	18.55	4.52
Case 1 High noise level	MPSNR	26.02	39.74	43.62	40.68	33.88	32.93	40.69	40.95	44.80
	MSSIM	0.48	0.95	0.98	0.97	0.87	0.85	0.92	0.92	0.97
	MSAM	16.86	3.28	2.51	3.12	6.24	6.12	3.35	4.75	2.28
	ERGAS	41.31	7.10	4.90	9.17	41.31	14.28	10.72	21.11	8.26
Case 2 Low noise level	MPSNR	34.26	43.47	46.67	44.59	38.68	38.85	42.27	36.64	48.11
	MSSIM	0.73	0.97	0.99	0.98	0.94	0.93	0.92	0.86	0.99
	MSAM	10.23	2.42	1.88	2.33	4.10	3.67	3.53	13.96	1.67
	ERGAS	24.58	4.94	3.85	7.53	24.58	7.03	9.61	55.10	7.36
Case 2 High noise level	MPSNR	32.38	39.23	40.70	43.96	37.65	37.36	42.03	32.92	46.57
	MSSIM	0.82	0.94	0.94	0.98	0.94	0.95	0.94	0.74	0.99
	MSAM	9.13	5.56	5.72	2.63	4.84	4.61	3.78	15.29	2.24
	ERGAS	17.26	6.90	6.76	7.65	17.26	8.19	8.71	57.59	7.60
Case 3 Low noise level	MPSNR	42.80	49.02	51.98	50.21	43.75	44.23	44.03	35.87	52.15
	MSSIM	0.92	0.99	1.00	0.99	0.97	0.98	0.95	0.82	1.00
	MSAM	3.96	1.37	0.91	1.29	2.17	1.73	2.67	14.29	1.27
	ERGAS	9.46	2.75	1.75	5.28	9.46	3.75	8.76	58.55	5.40
Case 3 High noise level	MPSNR	26.67	33.66	35.73	33.08	31.03	31.33	36.59	32.82	36.80
	MSSIM	0.55	0.84	0.84	0.86	0.81	0.82	0.83	0.79	0.87
	MSAM	21.24	11.63	11.41	12.09	13.11	12.57	11.63	16.65	11.34
	ERGAS	35.81	13.00	12.50	14.23	35.81	17.23	12.84	29.83	11.75
Case 4 Low noise level	MPSNR	36.94	43.20	44.24	48.25	41.08	41.47	45.39	33.13	48.95
	MSSIM	0.92	0.97	0.98	0.99	0.96	0.97	0.97	0.75	0.99
	MSAM	5.45	3.18	3.07	1.57	3.15	2.71	2.26	15.34	1.77
	ERGAS	10.49	4.36	3.24	6.48	10.49	4.92	5.09	58.55	7.12
Case 4 High noise level	MPSNR	32.63	39.22	40.47	45.05	37.86	37.69	43.09	32.94	45.24
	MSSIM	0.85	0.94	0.95	0.99	0.94	0.95	0.95	0.77	0.99
	MSAM	8.99	5.59	5.66	2.39	4.78	4.45	3.63	15.37	2.28
	ERGAS	20.08	12.01	11.92	7.55	20.08	8.85	9.09	58.26	7.98
Average computation time (s)		0	155.9	392.6	40.7	23.7	53.6	19.8	29.9	17.1

The bold items correspond to the best denoising performance.

Furthermore, to intuitively evaluate the denoising performance, we illustrate the restoration images of the Washington dataset with Case 1 noise type and high noise level and the Pavia dataset with Case 4 noise type and high noise level in Figures 2 and 3, respectively. For the Washington data, two enlarged local regions are shown on the left of each image, to display the details clearly. The lower region contains abundant edges. In the upper region there is a small ‘white’ vehicle with a single pixel size, labeled by the green circle. It serves as an indicator to observe whether the signal of the small target could be maintained after denoising.

Table 2. Performance of the proposed and competitors on Pavia data.

	Indexes	Noisy	NLM3D	BM4D	LRMR	GLF	SSK	SRSSHf	PSA	PTMA
Case 1 Low noise level	MPSNR	30.46	39.52	42.63	39.83	35.59	33.00	33.45	25.14	43.04
	MSSIM	0.78	0.97	0.99	0.98	0.94	0.89	0.91	0.77	0.98
	MSAM	15.25	4.21	3.37	4.65	6.70	10.84	5.23	7.72	3.07
	ERGAS	58.66	17.91	12.98	13.33	58.66	41.41	54.24	177.50	16.53
Case 1 High noise level	MPSNR	26.02	36.10	39.49	36.22	32.93	28.66	31.80	26.10	39.75
	MSSIM	0.62	0.94	0.97	0.95	0.90	0.76	0.87	0.71	0.98
	MSAM	23.11	5.66	4.50	6.82	9.64	16.74	5.76	7.89	5.68
	ERGAS	98.33	25.16	18.80	21.19	98.33	77.83	74.74	162.03	23.32
Case 2 Low noise level	MPSNR	33.38	40.24	42.59	40.08	35.84	35.14	31.67	24.48	43.65
	MSSIM	0.82	0.97	0.99	0.97	0.94	0.89	0.86	0.70	0.99
	MSAM	15.23	4.26	3.41	5.20	6.70	11.05	5.70	11.14	3.29
	ERGAS	59.97	16.96	12.73	19.42	59.97	49.24	50.08	180.30	17.38
Case 2 High noise level	MPSNR	30.97	36.90	39.03	35.08	33.41	31.75	30.57	24.15	39.98
	MSSIM	0.69	0.94	0.97	0.93	0.91	0.78	0.84	0.66	0.97
	MSAM	24.45	6.10	4.49	9.29	10.20	18.52	6.74	12.01	6.35
	ERGAS	108.00	26.08	19.79	38.07	108.00	94.00	59.38	184.07	25.27
Case 3 Low noise level	MPSNR	43.47	45.46	47.32	45.76	39.23	42.68	34.50	24.28	49.15
	MSSIM	0.96	0.99	0.99	0.99	0.97	0.98	0.93	0.66	1.00
	MSAM	6.27	3.00	2.14	2.35	3.81	4.64	4.55	10.98	1.73
	ERGAS	23.51	7.75	7.36	7.60	23.51	18.20	38.78	202.85	8.56
Case 3 High noise level	MPSNR	25.80	28.41	28.74	28.46	27.85	27.06	27.14	26.95	28.83
	MSSIM	0.66	0.85	0.87	0.87	0.83	0.77	0.78	0.82	0.88
	MSAM	21.60	6.80	5.12	8.97	9.56	15.35	6.94	8.23	4.94
	ERGAS	84.83	22.53	16.44	37.58	84.83	59.47	49.93	26.03	20.50
Case 4 Low noise level	MPSNR	36.52	41.49	41.37	44.11	37.42	37.97	36.68	22.96	44.87
	MSSIM	0.95	0.98	0.99	0.99	0.97	0.97	0.95	0.59	0.99
	MSAM	7.99	3.55	3.29	2.68	4.73	6.09	5.00	13.93	3.09
	ERGAS	25.14	8.81	7.68	8.16	25.14	21.29	30.24	213.85	11.37
Case 4 High noise level	MPSNR	32.69	37.41	37.54	39.95	35.38	34.33	34.60	23.15	43.14
	MSSIM	0.90	0.96	0.97	0.98	0.96	0.94	0.93	0.65	0.99
	MSAM	11.49	5.39	4.90	4.36	6.24	8.72	5.48	13.74	3.37
	ERGAS	35.21	12.86	10.79	12.49	35.21	29.80	45.86	211.69	13.09
Average computation time (s)		0	53.4	169.2	17.6	12.1	40.9	13.5	3.73	2.75

The bold items correspond to the best denoising performance.

From Figure 2, it is evident that the NLM3D, GLF and the SSK caused severe blurring of edges. And the restoration result of the LRMR showed a noticeable bias. Except for the LRMR, the PSA and the PTMA, the small target in the HSIs restored by other methods was not identifiable. In Figure 3, the Pavia data were contaminated by severe strip noise. And in all the results, except for the LRMR and the PTMA, there were still noticeable residual strips, which implies that these methods were not applicable in removing the strip noise. The image information in the PSA result was barely perceptible, which meant that the strip noise severely affected the signal subspace identified by the PSA. By contrast, the PTMA achieved the best performance in removing the strip noise and preserving the image information by applying the third-order moment as the criterion.

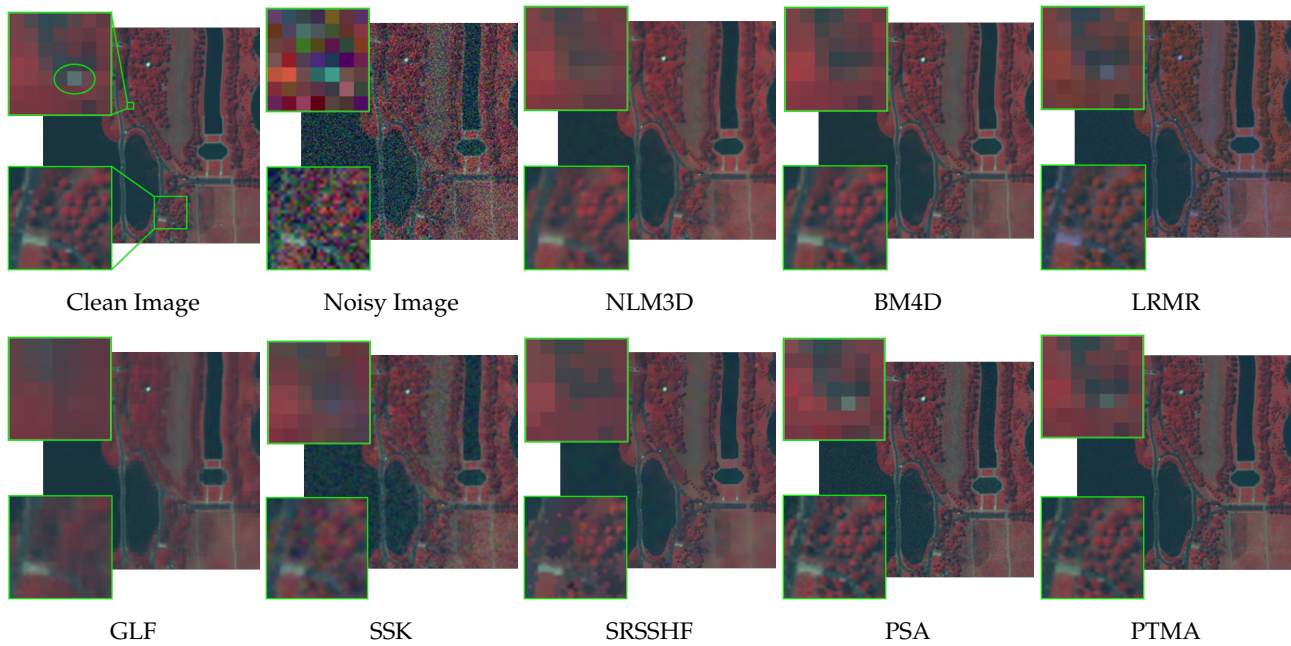


Figure 2. Restoration results of Washington data with Case 1 noise type and high noise level.

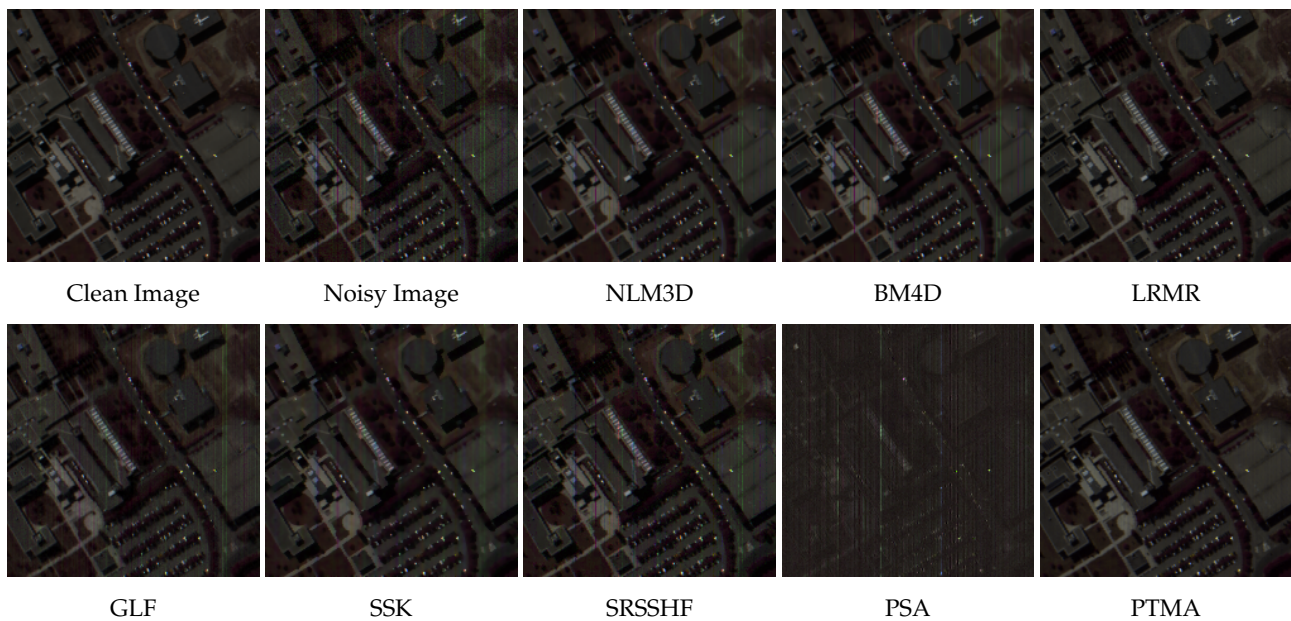


Figure 3. The restoration results of Pavia dataset with Case 4 noise type and high noise level.

Moreover, the comparisons of the denoised spectra to the noise-free spectrum are shown in Figures 4 and 5. For the Washington data, we display the spectra of the small target labeled in Figure 2, to illustrate the ability of these methods to preserve the spectral characteristic of small targets. It can be seen that the spectral shape of the NLM3D, BM4D, GLF and the SSK significantly deviated from the original spectral shape. The LRMR, the PSA and the PTMA could preserve the spectral characteristic of the small target well. But the LRMR and the PSA contained more residual noise than the PTMA. In addition, Figure 5 shows the spectra of the pixel locating at (50, 50) in the Pavia data. The selected pixel corresponded to bare soil, serving as an indicator of whether the spectral characteristics of the background could be maintained by these methods. It can be seen that the denoised spectrum in the PTMA was the closest to the original spectrum. But the spectrum of the PTMA significantly differed from the clean spectrum. Therefore, by replacing the criterion from the skewness to the third-order moment, the PTMA retained the capability

of preserving small targets, while concurrently achieving a significant enhancement in denoising efficacy.

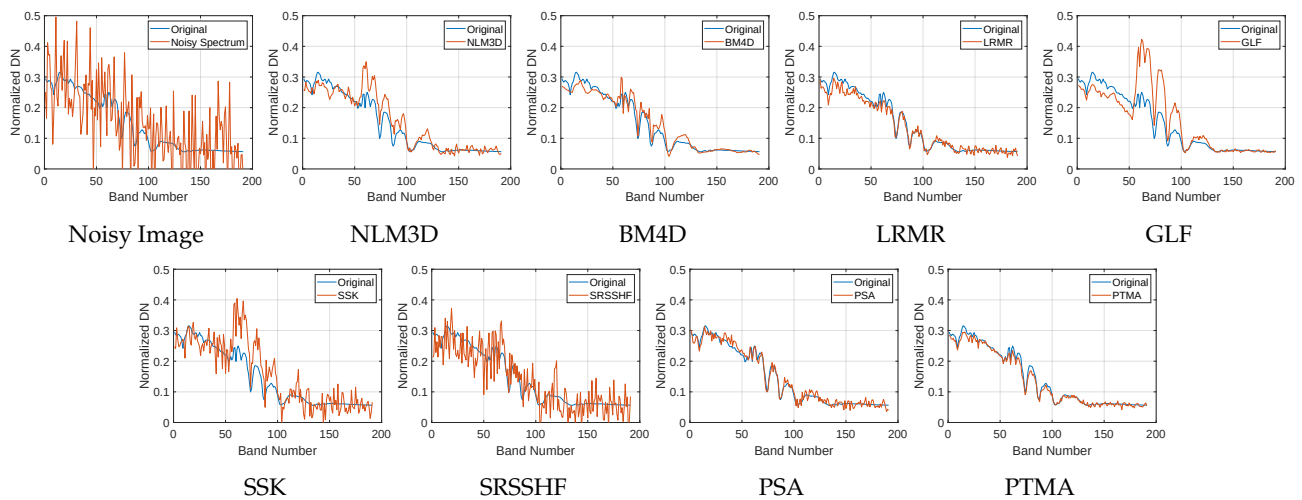


Figure 4. Comparison of the spectra of the small target in the HSIs shown in Figure 2 to the clean spectrum.

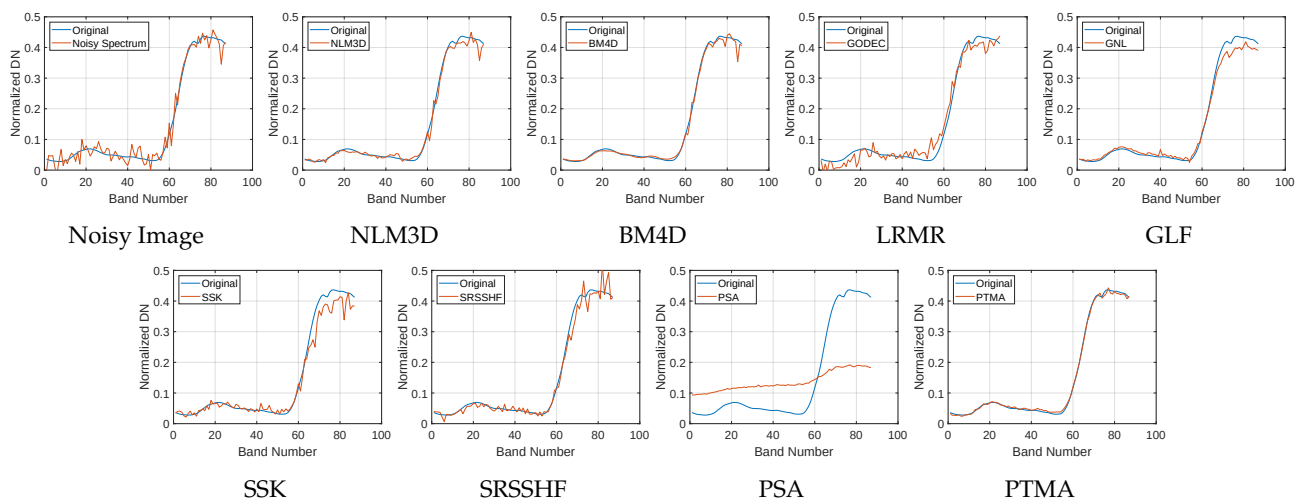


Figure 5. Comparison of the spectra of the small target in the HSIs shown in Figure 3 to the clean spectrum.

4.2. Real Experiment

For this subsection, we utilized two HSIs containing real complex noise as the experiment datasets to evaluate the denoising and target-preservation ability of the PTMA under real-noise conditions. The first noisy HSI was acquired by the airborne-visible/infrared-imaging spectrometer (AVIRIS) over the Indian Pines test site located in northwestern Indiana. It comprised 145×145 pixels and 224 spectral reflectance bands within the wavelength range of 0.4–2.5 μm . Its false-color image, composed of bands 40, 27, 14, is shown in Figure 6a. It can be seen that there was no apparent noise in these three bands. However, the noise level of the Indian Pines dataset varied in different bands. For instance, the image of the second band was affected by intense noise, which is shown in Figure 6b.

The other noisy HSI was acquired by the advanced hyperspectral imager (AHSI) aboard the GaoFen-5 (GF-5) satellite, which is the fifth satellite of a series of the China High-Resolution Earth Observation System. The AHSI is a 330-channel imaging spectrometer with a 30 m spatial resolution covering the 0.4–2.5 μm spectral range. The visible and near-infrared range consists of 150 bands with a spectral resolution of 5 nm, while the shortwave infrared range consists of 180 bands with a spectral resolution of 10 nm. However, due to

the detector's low responsivity, 25 bands (bands number 193–200 and 246–262) had zero signal intensity and, thus, were removed. After preprocessing, the selected HSI consisted of 305 bands and 256×256 pixels, which were acquired over the southwest region of Beijing, China. The false-color image, composed by bands 67, 37, 20, as shown in Figure 6d, exhibited high image quality. However, several bands contained noticeable strip noise, such as band 305, shown in Figure 6e.

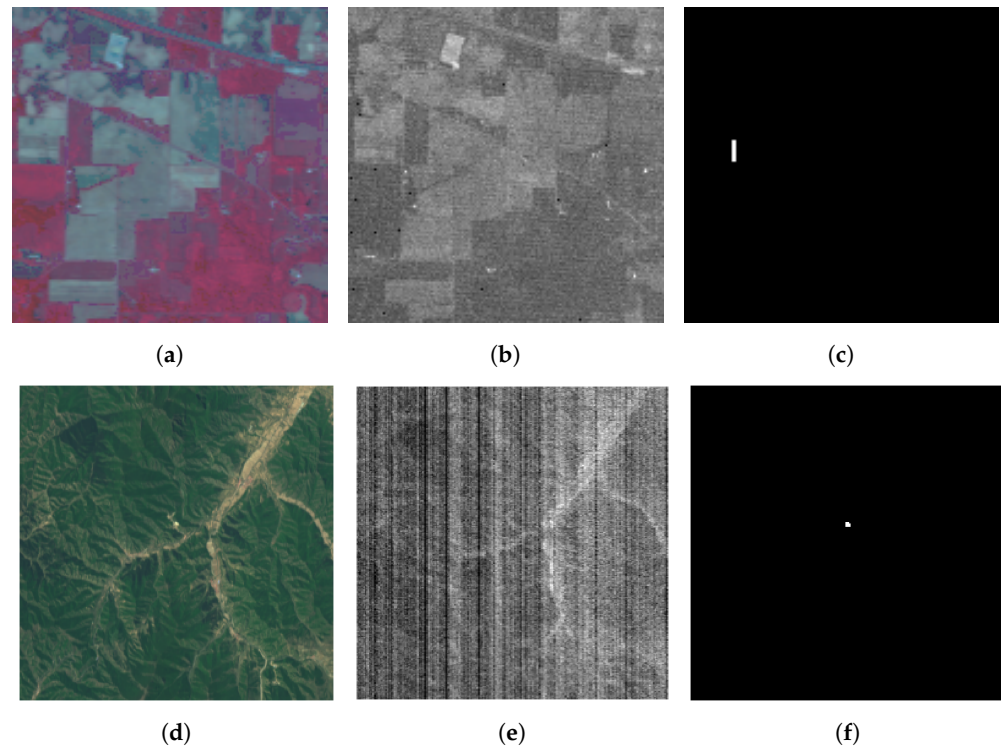


Figure 6. (a) The false-color image of the Indian Pines dataset composed by bands 40, 27, 14. (b) The second band of the Indian Pines dataset. (c) The mask of the target in the Indian Pines dataset. (d) The selected GF-5 data composed by bands 67, 37, 20. (e) The image of the 305th band of the GF-5 data. (f) The mask of the target in the GF-5 dataset.

The PTMA and all the competitors were employed to restore both of these datasets. Through testing, it was determined that the optimal standard deviations for the noise in these two datasets, which were required by certain competitors, should be set to 0.03 and 0.012, respectively. These values were found to yield the best denoising performance. And the remaining parameters remained consistent with those used in the previous subsection.

To intuitively evaluate the denoising performance, the images of the restoration results of the Indian Pines data and the GF-5 data are displayed in Figures 7 and 8. From Figure 7, it is obvious that there was visibly residual noise in the results of NLM3D, BM4D and the SSK. This suggests that NLM3D and BM4D, which achieved good denoising results under the synthetic Gaussian noise, are suitable for a situation where the noise level varies across different bands. The LRMR, GLF and the PTMA all reduced the noise well, while preserving the edges. Comparing the results of the PSA and the PTMA, it is evident that the PTMA achieved much better denoising performance than the PSA, which demonstrates that the third-order moment is also robust to noise with varying variance.

According to Figure 8, it can be observed that the restoration results of NLM3D, the SSK and the SRSSHf were still contaminated by severe stripes, which means that they were not applicable to the strip noise. The result of BM4D was severely blurred. The PSA exhibited some efficacy in suppressing the striping noise, but fell short of complete elimination. The LRMR, GLF and the PTMA achieved relatively good denoising performance. It is worth noting that the number of the residual strips in the HSI denoised by the PTMA was less

than that in the LRMR and GLF results. Consequently, the PTMA demonstrated its ability to accurately identify the signal subspace, even in the presence of stripe noise.

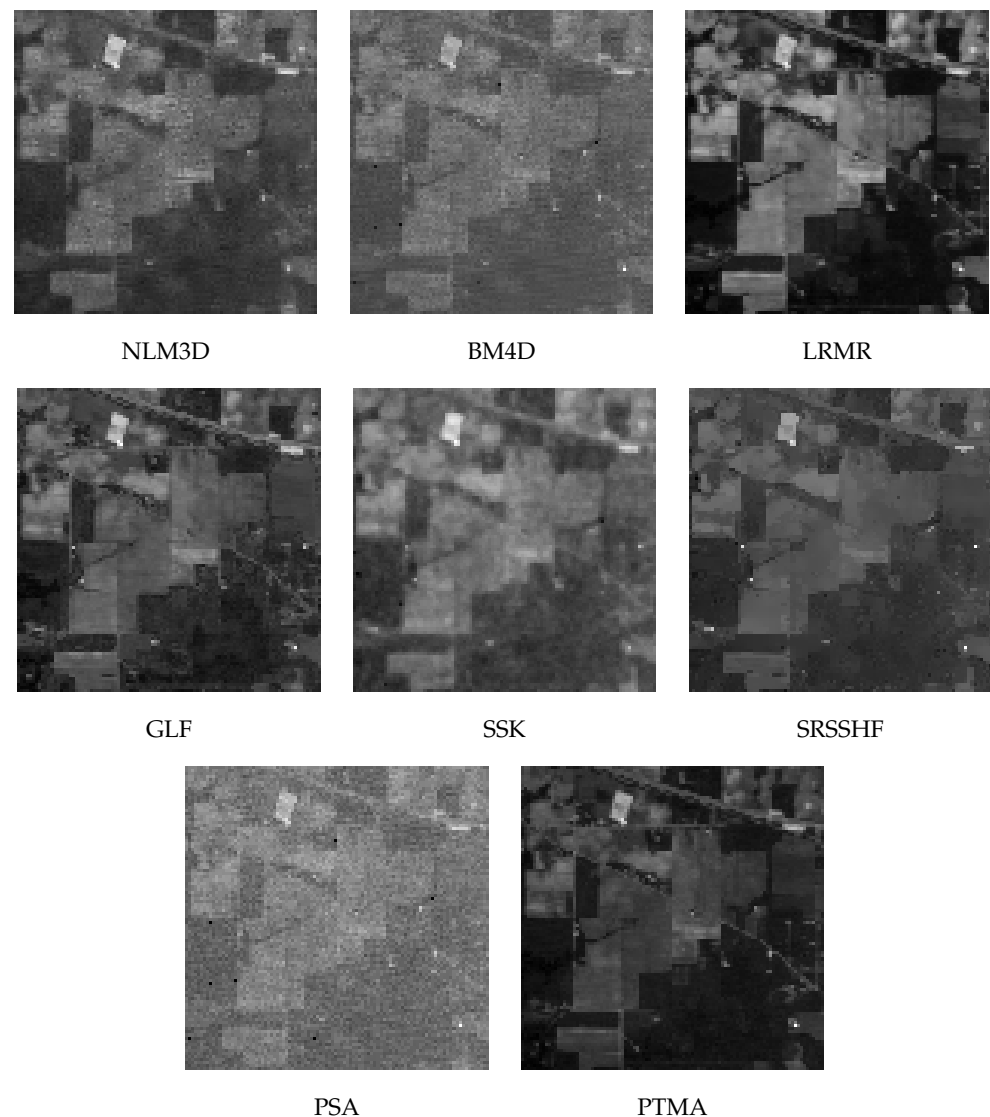


Figure 7. The second band of the restoration results of the Indian Pines data.

To evaluate the target-preservation capabilities of these methods in real noisy conditions, we selected two materials that were present in the datasets and that constituted only a small fraction of the entire image signal, as the small targets to be detected. For the Indian Pines dataset, we selected oats, which consisted of 20 pixels, as the small target. The location of the target is shown in Figure 6c. For the GF-5 dataset, we selected a mine, which consisted of 15 pixels, as the small target. The location of the target is shown in Figure 6f.

Considering whether the spectral characteristic of the small target maintained in the restored HSIs directly affected the target detection accuracy, we used the target detection accuracy to evaluate the target-preservation ability of these methods. The spectral angle mapper (SAM) [41] was employed as the target-detection algorithm. The SAM algorithm required the reference of the target spectrum as the input, so we used the average spectrum of the spectra of all the small-target pixels in the noisy HSI as the reference spectrum, to

minimize the noise in the reference spectrum as much as possible. The output value of the SAM at position i was calculated as

$$\text{SAM}(i) = \arccos\left(\frac{\mathbf{r}_t^T \hat{\mathbf{X}}_i}{\|\mathbf{r}_t\|_2 \|\hat{\mathbf{X}}_i\|_2}\right), \quad (27)$$

in which $\mathbf{r}_t$ and $\hat{\mathbf{X}}_i$ represented the reference spectrum of the target and the i th spectrum in the denoised HSI, respectively. The output range of the SAM was $[0, \pi/2]$, and it was then normalized into $[0, 1]$ to create the target-detection score map.

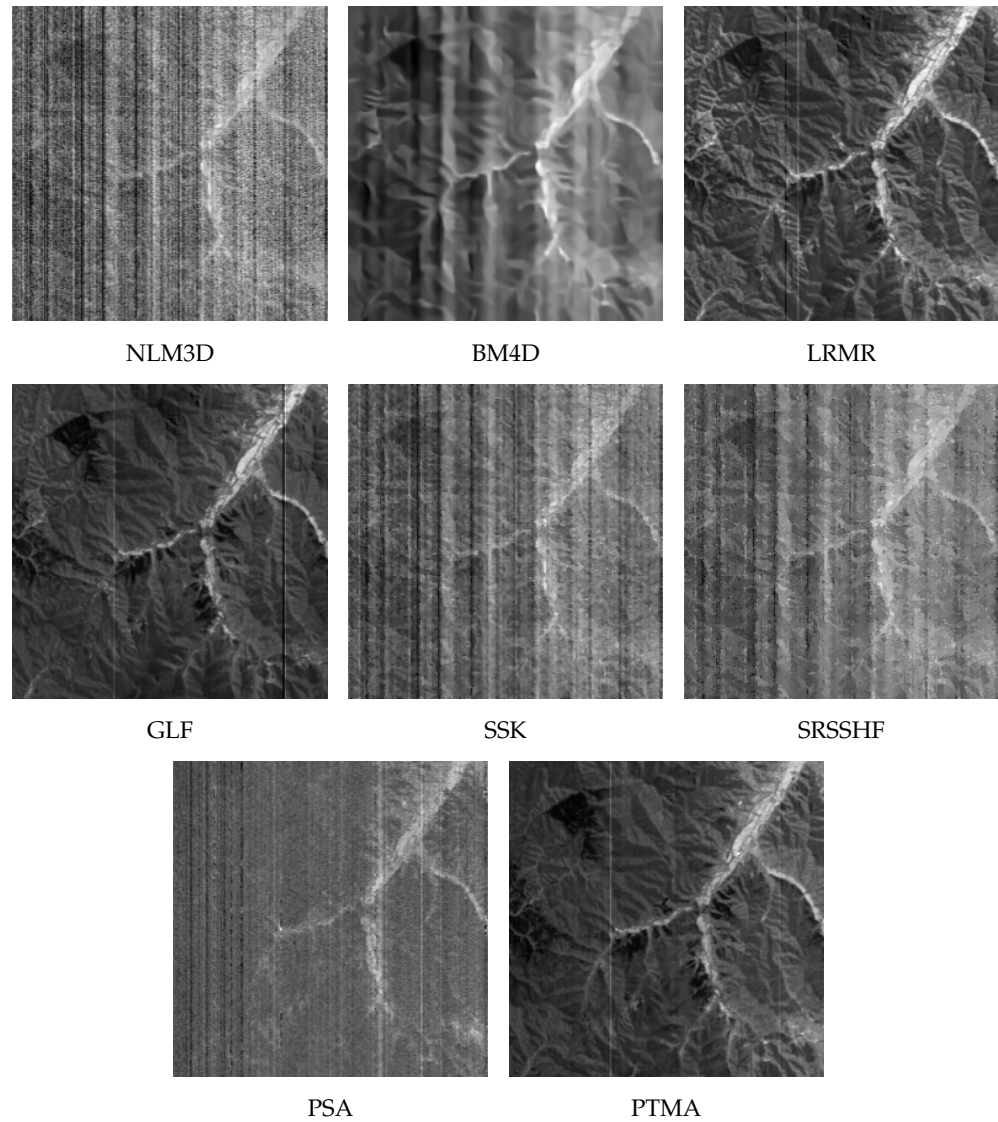


Figure 8. The 305th band in the restoration results of the GF-5 data.

In assessing the target-detection accuracy, we employed the receiver-operating-characteristic-(ROC) curve [42] and the corresponding area under the curve (AUC) as the metric. The ROC curve was the plot of the true positive rate (TPR) against the false positive rate (FPR) at each threshold setting. The TPR and FPR were calculated as

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (28)$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}, \quad (29)$$

where TP meant the true positive rate, FN was the false negative rate, FP was the false positive rate and TN was the true negative rate. An ROC curve that lies closer to the top-left corner indicates better target detection performance, as it maintained a high true positive rate while keeping the false positive rate low. In addition, we calculated the area under the curve (AUC). A higher AUC value indicates better target detection performance.

The ROC and AUC of the detection score map are presented in Figure 9 and Table 3, respectively. It can be seen that the PTMA achieved the highest TPR values at most FPR. In addition, the PTMA obtained the highest AUC value for both datasets. The experiment results affirm that the PTMA is still effective under non-Gaussian noise conditions, in terms of both noise reduction and target preservation.

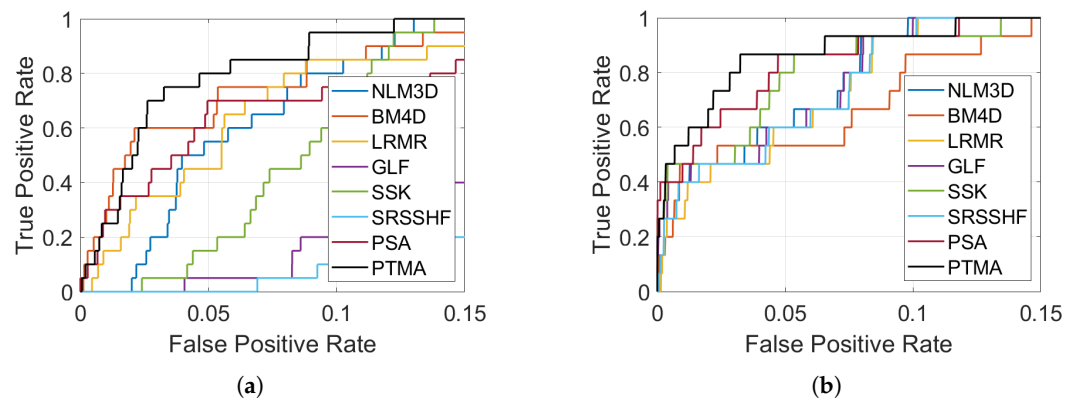


Figure 9. The ROC curves of the restoration results. (a) The Indian Pines dataset. (b) The GF-5 dataset.

Table 3. The AUC of the restoration results of the Indian Pines and GF-5 data.

AUC	NLM3D	BM4D	LRMR	GLF	SSK	SRSSH	PSA	PTMA
Indian Pines	0.940	0.956	0.939	0.823	0.915	0.763	0.936	0.968
GF-5	0.963	0.949	0.958	0.962	0.968	0.959	0.973	0.979

The bold items correspond to the highest target detection accuracy.

5. Conclusions

Existing low-rank-based denoising methods often encounter challenges in simultaneously preserving the signal of small targets and effectively reducing noise. In this paper, we demonstrated that the third-order-moment value remains unaffected by noise and that it exhibits sensitivity to minor components in the signal. Consequently, we adopted the maximum third-order moment as the criterion for identifying the signal subspace. We introduced the principal-third-order-moment-analysis algorithm, capable of rapidly calculating the first few maximum third-order-moment directions. Two simulated and two real HSIs were employed to assess the effectiveness of the proposed method. We compared the proposed denoising method to seven State-of-the-Art competitors. Our extensive experimental results demonstrated the superiority of the proposed method, in terms of noise removal and small-target preservation when compared to the competitors.

Author Contributions: Conceptualization, S.L.; methodology, S.L.; software, S.L.; validation, S.L.; investigation, S.L.; data curation, Y.Z.; writing—original draft preparation, S.L.; writing—review and editing, S.L., X.G., L.Z., L.J.; visualization, S.L.; funding acquisition, Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Key Research and Development Program of China (31400), and the Second Tibetan Plateau Scientific Expedition and Research Program (STEP), Grant No. 2019QZKK0206.

Data Availability Statement: The Washington data, Pavia data and Indian Pines data are all publicly archived dataset, and can be obtained in <http://lesun.weebly.com/hyperspectral-data-set.html>. In addition, the GF-5 data used in this paper can be obtained in <https://github.com/SZ-Li/PTMA-Matlab-Code>.

Conflicts of Interest: The authors declare no conflicts of interest.

Nomenclature

Notation	Description
$\mathbf{X} \in \mathbb{R}^{N \times L}$	Noise-free HSI with N pixels and L bands
$\mathbf{X}_i \in \mathbb{R}^{N \times 1}$	i_{th} column of $\mathbf{X}$, the image vector of the i_{th} band
$\mathbf{X}_{b,:} \in \mathbb{R}^{1 \times L}$	b_{th} row of $\mathbf{X}$, the spectrum vector at position b
$X_{i,j}$	Scalar element located at the i_{th} row and j_{th} column of $\mathbf{X}$
$\mathbf{N}$	Additive-noise matrix
$\mathbf{Y}$	Observed noisy HSI
$\hat{\mathbf{X}}$	Estimated noise-free HSI
$\mathbf{u} \in \mathbb{R}^{L \times 1}$	Unit-projection direction vector
$\mathbf{v}, \mathbf{w}, \mathbf{n} \in \mathbb{R}^{N \times 1}$	Projection coefficients of $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{N}$ along direction $\mathbf{u}$
v_i, w_i, n_i	i_{th} element of $\mathbf{v}$, $\mathbf{w}$ and $\mathbf{n}$

References

1. Sun, G.; Fu, H.; Ren, J.; Zhang, A.; Zabalza, J.; Jia, X.; Zhao, H. SpaSSA: Superpixelwise adaptive SSA for unsupervised spatial-spectral feature extraction in hyperspectral image. *IEEE Trans. Cybern.* **2021**, *52*, 6158–6169. [CrossRef] [PubMed]
2. Huang, S.; Zhang, H.; Pižurica, A. Subspace clustering for hyperspectral images via dictionary learning with adaptive regularization. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 1–17. [CrossRef]
3. Hennessy, A.; Clarke, K.; Lewis, M. Hyperspectral classification of plants: A review of waveband selection generalisability. *Remote Sens.* **2020**, *12*, 113. [CrossRef]
4. Manolakis, D.; Marden, D.; Shaw, G.A. Hyperspectral image processing for automatic target detection applications. *Linc. Lab. J.* **2003**, *14*, 79–116.
5. Sun, X.; Qu, Y.; Gao, L.; Sun, X.; Qi, H.; Zhang, B.; Shen, T. Target detection through tree-structured encoding for hyperspectral images. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 4233–4249. [CrossRef]
6. Zhu, D.; Du, B.; Zhang, L. Target dictionary construction-based sparse representation hyperspectral target detection methods. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 1254–1264. [CrossRef]
7. Zhang, J.; Zhao, D.; Gao, W. Group-based sparse representation for image restoration. *IEEE Trans. Image Process.* **2014**, *23*, 3336–3351. [CrossRef] [PubMed]
8. Tomasi, C.; Manduchi, R. Bilateral filtering for gray and color images. In Proceedings of the Sixth International Conference on Computer Vision (IEEE Cat. No. 98CH36271), Bombay, India, 7 January 1998; pp. 839–846.
9. Peng, H.; Rao, R.; Dianat, S.A. Multispectral image denoising with optimized vector bilateral filter. *IEEE Trans. Image Process.* **2013**, *23*, 264–273. [CrossRef]
10. Yuan, Y.; Zheng, X.; Lu, X. Spectral-spatial kernel regularized for hyperspectral image denoising. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 3815–3832. [CrossRef]
11. Peng, H.; Rao, R. Bilateral kernel parameter optimization by risk minimization. In Proceedings of the 2010 IEEE International Conference on Image Processing, Hong Kong, China, 26–29 September 2010; pp. 3293–3296.
12. Zhao, Y.; Pu, R.; Bell, S.S.; Meyer, C.; Baggett, L.P.; Geng, X. Hyperion image optimization in coastal waters. *IEEE Trans. Geosci. Remote Sens.* **2012**, *51*, 1025–1036. [CrossRef]
13. Manjón, J.V.; Coupé, P.; Martí-Bonmatí, L.; Collins, D.L.; Robles, M. Adaptive non-local means denoising of MR images with spatially varying noise levels. *J. Magn. Reson. Imaging* **2010**, *31*, 192–203. [CrossRef]
14. Maggioni, M.; Katkovnik, V.; Egiazarian, K.; Foi, A. Nonlocal transform-domain filter for volumetric data denoising and reconstruction. *IEEE Trans. Image Process.* **2012**, *22*, 119–133. [CrossRef] [PubMed]
15. Buades, A.; Coll, B.; Morel, J.M. A non-local algorithm for image denoising. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–25 June 2005; Volume 2, pp. 60–65.
16. Aggarwal, H.K.; Majumdar, A. Hyperspectral image denoising using spatio-spectral total variation. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 442–446. [CrossRef]
17. Takemoto, S.; Naganuma, K.; Ono, S. Graph spatio-spectral total variation model for hyperspectral image denoising. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [CrossRef]
18. Rudin, L.I.; Osher, S.; Fatemi, E. Nonlinear total variation based noise removal algorithms. *Phys. Nonlinear Phenom.* **1992**, *60*, 259–268. [CrossRef]

19. He, W.; Zhang, H.; Zhang, L.; Shen, H. Hyperspectral image denoising via noise-adjusted iterative low-rank matrix approximation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 3050–3061. [\[CrossRef\]](#)
20. Green, A.A.; Berman, M.; Switzer, P.; Craig, M.D. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Trans. Geosci. Remote Sens.* **1988**, *26*, 65–74. [\[CrossRef\]](#)
21. Zhang, H.; He, W.; Zhang, L.; Shen, H.; Yuan, Q. Hyperspectral image restoration using low-rank matrix recovery. *IEEE Trans. Geosci. Remote Sens.* **2013**, *52*, 4729–4743. [\[CrossRef\]](#)
22. Bioucas-Dias, J.M.; Nascimento, J.M. Hyperspectral subspace identification. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 2435–2445. [\[CrossRef\]](#)
23. Fan, H.; Li, C.; Guo, Y.; Kuang, G.; Ma, J. Spatial-spectral total variation regularized low-rank tensor decomposition for hyperspectral image denoising. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 6196–6213. [\[CrossRef\]](#)
24. Liu, Y.; Shan, C.; Gao, Q.; Gao, X.; Han, J.; Cui, R. Hyperspectral image denoising via minimizing the partial sum of singular values and superpixel segmentation. *Neurocomputing* **2019**, *330*, 465–482. [\[CrossRef\]](#)
25. Grünwald, P.D. *The Minimum Description Length Principle*; MIT Press: Cambridge, MA, USA, 2007.
26. Chiang, S.S.; Chang, C.I.; Ginsberg, I.W. Unsupervised target detection in hyperspectral images using projection pursuit. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 1380–1391. [\[CrossRef\]](#)
27. Ifarraguerri, A.; Chang, C.I. Unsupervised hyperspectral image analysis with projection pursuit. *IEEE Trans. Geosci. Remote Sens.* **2000**, *38*, 2529–2538.
28. Geng, X.; Ji, L.; Sun, K. Principal skewness analysis: Algorithm and its application for multispectral/hyperspectral images indexing. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 1821–1825. [\[CrossRef\]](#)
29. Richards, J.A.; Richards, J.A. *Remote Sensing Digital Image Analysis*; Springer: Berlin, Germany, 2022; Volume 5.
30. Chen, Y.; Cao, X.; Zhao, Q.; Meng, D.; Xu, Z. Denoising hyperspectral image with non-iid noise structure. *IEEE Trans. Cybern.* **2017**, *48*, 1054–1066. [\[CrossRef\]](#)
31. Xu, S.; Cao, X.; Peng, J.; Ke, Q.; Ma, C.; Meng, D. Hyperspectral Image Denoising by Asymmetric Noise Modeling. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–14. [\[CrossRef\]](#)
32. Zhuang, L.; Ng, M.K. FastHyMix: Fast and parameter-free hyperspectral image mixed noise removal. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *34*, 702–716. [\[CrossRef\]](#)
33. Fu, X.; Guo, Y.; Xu, M.; Jia, S. Hyperspectral Image Denoising Via Robust Subspace Estimation and Group Sparsity Constraint. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5512716. [\[CrossRef\]](#)
34. Wang, J.; Chang, C.I. Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1586–1600. [\[CrossRef\]](#)
35. Tukey, J.W. *Exploratory Data Analysis*; UMI: Boston, MA, USA, 1977; Volume 2.
36. Ortega, J.M.; Rheinboldt, W.C. *Iterative Solution of Nonlinear Equations in Several Variables*; SIAM: Philadelphia, PA, USA, 2000.
37. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [\[CrossRef\]](#)
38. Neath, A.A.; Cavanaugh, J.E. The Bayesian information criterion: background, derivation, and applications. *Wiley Interdiscip. Rev. Comput. Stat.* **2012**, *4*, 199–203. [\[CrossRef\]](#)
39. Zhuang, L.; Fu, X.; Ng, M.K.; Bioucas-Dias, J.M. Hyperspectral image denoising based on global and nonlocal low-rank factorizations. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 10438–10454. [\[CrossRef\]](#)
40. Renza, D.; Martinez, E.; Arquerio, A. A new approach to change detection in multispectral images by means of ERGAS index. *IEEE Geosci. Remote Sens. Lett.* **2012**, *10*, 76–80. [\[CrossRef\]](#)
41. Chang, C.I. An information-theoretic approach to spectral variability, similarity, and discrimination for hyperspectral image analysis. *IEEE Trans. Inf. Theory* **2000**, *46*, 1927–1932. [\[CrossRef\]](#)
42. Fawcett, T. An introduction to ROC analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.