This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Antinoise Hyperspectral Image Fusion by Mining Tensor Low-Multilinear-Rank and Variational Properties

Jie Li, *Member, IEEE*, Xinxin Liu, Qiangqiang Yuan, *Member, IEEE*, Huanfeng Shen, *Senior Member, IEEE*, and Liangpei Zhang, *Fellow, IEEE*

***Abstract*— Enhancing the spatial resolution of hyperspectral (HS) images by fusing with higher spatial resolution multispectral (MS) data is of significance for applications. However, due to the narrow bandwidth, HS images (HSIs) are vulnerable to various types of noise, such as Gaussian noise and stripes, which can severely affect the fusion performance. This paper focuses on antinoise HS and MS image fusion to enhance the spatial details and suppress the noise. By analysis of the intrinsic structure and noise properties, we formulate this problem as the minimization of an objective function. Under the optimization framework, small multilinear ranks in tensor are first used to identify the intrinsic structures of the clean HSI part. Then, considering the high spectral correlation, it is assumed that any bands can be represented by the combination of certain adjacent bands. The difference between one band and its corresponding combination can be used to preserve the spatio-spectral consistency and characterize the distribution of sparse noise (such as stripe noise), based on the variational properties along two directions. The alternating direction method of multipliers (ADMM) is applied to solve and accelerate the model optimization. Experiments with both simulated- and real-data demonstrate the effectiveness of the proposed model and its robustness to the noise, in terms of both qualitative and quantitative perspectives.**

***Index Terms*— Hyperspectral (HS) image, image fusion, multispectral (MS) image, variational optimization.**

Manuscript received October 29, 2018; revised March 27, 2019; accepted May 9, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant 41701400, in part by the Natural Science Foundation of Hubei Province under Grant ZRMS2017000729, in part by the China Postdoctoral Science Foundation under Grant 2018T110803, and in part by the Fundamental Research Funds for the Central Universities under Grant 2042019kf0213 and Grant 531118010209. *(Corresponding author: Qiangqiang Yuan.)*

J. Li is with the School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China (e-mail: aaronleecool@whu.edu.cn).

X. Liu is with the College of Electrical and Information Engineering, Hunan University, Changsha 410008, China (e-mail: liuxinxin@hnu.edu.cn).

Q. Yuan is with the School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China, and also with the Collaborative Innovation Center of Geospatial Technology, Wuhan University, Wuhan 430079, China (e-mail: yqiang86@gmail.com).

H. Shen is with the School of Resource and Environmental Science, Wuhan University, Wuhan 430079, China, and also with the Collaborative Innovation Center of Geospatial Technology, Wuhan University, Wuhan 430079, China (e-mail: shenhf@whu.edu.cn).

L. Zhang is with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China, and also with the Collaborative Innovation Center of Geospatial Technology, Wuhan University, Wuhan 430079, China (e-mail: zlp62@whu.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2019.2916654

I. Introduction

Hyperspectral (HS) imaging is a widely used modality that can simultaneously acquire images of the same scene across a number of different wavelengths. Obtaining dense HS bands is important for remote sensing and computer vision applications [1]–[3] including object segmentation, tracking, and recognition. However, due to the tradeoff between the spatial resolution and spectral resolutions, HS imaging with a high spectral resolution has severe limitations in spatial resolution when compared with regular multispectral (MS) sensors. Namely, it is not easy to simultaneously obtain the high spectral and spatial resolution versions of such images. To meet the high spatial and high spectral requirements of the potential applications, a spatio-spectral fusion method is developed as a solution to yield a high-resolution HS (HRHS) image by merging a low spatial resolution HS image (HSI) into a high spatial resolution MS image [4].

Unfortunately, these observed data coming along with the noise may seriously restrict the fusion accuracy. Especially for HSIs, the narrow bandwidth inevitably causes the various types of noise [5], [6], mainly grouped into random noise and structure (also called fixed-pattern) noise. More specifically, stripe noise, as a typical example of structure noise, appears as a series of striping artifacts in the along-track direction, due to the radiometric miscalibration among detectors, the detectors' response changes influenced by the temperature or the functional failure of individual detector elements. To remove the stripe, researchers have proposed many algorithms [6]–[16], but mainly committed to a single-band destriping task [7]–[12]. For HSI destriping, spectral moment matching [13] is derived based on the abundant spectral autocorrelation. The subspace-based approaches [14], [15] can be used to estimate the striping component and to remove it from the image. Low-rank matrix recovery (LRMR) [6] can also be applied to HSI destriping problem mixed with the Gaussian noise. To treat the multichannel image as a spectral–spatial volume, an anisotropic spectral–spatial total variation (ASSTV) regularization [16] is posed to enhance the smoothness of solution along both the spectral and spatial dimensions. The noise artifacts may result in the obscured details and spectral distortion. The most direct idea to eliminate them in fusion is denoising preprocessing, which yet may smooth details, remain artifacts, and accumulate error. As a result, there is an urgent need to achieve a higher fusion quality with a satisfactory antinoise effect when faced with the problem of noise interference.

To date, many algorithms have been presented for HS and MS image fusion, with the increasing availability of HS systems. Early research mostly extended pansharpening methods [17], [18], which are aimed at fusing panchromatic image and MS image, to HS fusion. These methods mainly focus on the component substitution (CS) and multiresolution analysis (MRA). The CS approaches [19]–[22] mainly rely on the substitution of the spatial component of the HSI using the higher spatial resolution information from the MS image in a transformed space. By dividing the spectrum of the HS data into several regions [19], HS and MS band images are fused in each region using the conventional pan-sharpening techniques [23], such as principal component analysis [21] and Gram–Schmidt (GS) spectral sharpening [22]. The CS method investigates the fusion problem under a low-dimensional spectral subspace and relies upon the correlation between the replaced components. But low correlation easily causes spectral distortion [24]. The MRA approaches [25]–[27] are based on the injection of spatial details, which are obtained through a multiscale decomposition of the high-resolution MS (HRMS) images. In [28], MRA-based pan-sharpening methods were effectively adapted to HS and MS fusion by synthesizing a high-resolution image for each HS band as a linear combination of MS band images via linear regression. However, the performance is highly dependent on the spectral resampling method, which can greatly limit the enhancement of the spatial resolution of the HS band images.

To overcome the above drawback, variational methods have been proposed, among which the Bayesian approach and the matrix factorization methods are widely used methods [17], [29]. The Bayesian fusion methods formulate a posterior distribution with prior knowledge to produce an intuitive interpretation of the fusion process. The ill-posed fusion problem is regularized by adding the appropriate prior distribution for the scene of interest, such as l_2-Gaussian priors [30], total variation prior [31], sparsity prior [32], and low rank [33], [34]. In terms of a Bayesian method based on a maximum *a posteriori* (MAP) estimator developed in [35], a stochastic mixing model and wavelet analysis derived in [36] and [37] can be subsequently combined to capture the spectral characteristic and spatial knowledge. Considering the spatial and spectral characteristics simultaneously, a group spectral embedding-based fusion method [34] was proposed by exploring the multiple manifold structures of the spectral bands and the low-rank structure of the HRHS data. In addition, the optimization of the MAP-based method can be processed in the principal component subspace. This idea of fusing the HS and MS images based on the spectral information of both input images on a subspace has been the main source of inspiration for many HS and MS fusion methods developed subsequently [38]–[40].

In the subspace optimization-based fusion category, the matrix factorization methods are the successful ones. Matrix factorization approaches [41]–[49] intend to exploit the inherent spectral characteristics of the same scene between HS and MS images in a subspace. The observed HSI is usually decomposed into a certain basis (or the spectral signatures of the materials), and the optimized high-resolution coefficient (or abundance) matrices are subsequently combined to obtain a fused image. For example, coupled nonnegative matrix factorization (CNMF) [42] employs unmixing techniques to generate the HS endmember matrix and the high spatial resolution abundance matrix for the low-resolution HS (LRHS) and HRMS image fusion. In addition, the sparse representation-based methods [32], [45] are widely used to produces desired results, by combining the learned dictionary from the HSI and the high-resolution sparse coefficient. To improve the accuracy of matrix factorization, the nonnegative structured sparse representation (NSSR) [46] imposes a structural constraint to ensure the spatial correlation of the coefficients. Moreover, nonparametric Bayesian learning was also employed in [49] to achieve the HRHS image. To better consider and retain a 3-D property of HSIs, a coupled sparse tensor factorization [50] redefines the fusion problem as the estimation of a core tensor and dictionaries of the three modes.

Although the existing variational methods, including the Bayesian approaches and matrix factorization approaches, can obtain promising results under the noise-free conditions, many methods often ignore the real case of mixed noise in HSIs such as Gaussian noise or stripe. Under the influence of noise, the spatial structures and the internal spectral relations of the HSI cannot be well maintained during the process of fusion. And, the structural artifacts induced by stripe noise are even more difficult to be removed with the conventional methods. Furthermore, due to the lack of the consideration of different noise distribution properties, the existing methods [32], [37], [46] are generally designed to remove a certain type of noise, i.e., Gaussian noise, while inadaptable to the structure noise. In order to further improve the fusion quality, it is crucial to design a model which can not only effectively remove the various types of noise, but can also maintain the fused healthy details. Therefore, this paper presents an antinoise HSI fusion method by mining the spatio-spectral properties and characterizing the various noise features. Due to the superiority of the tensor modeling techniques in simultaneously preserving spatial structures and spectral continuity for high-order tensor data [51]–[53], tensor decomposition is employed to separate the noise part, particularly the Gaussian noise, from the intrinsic structures of the clean HSI. Considering only tensor decomposition, which utilizes similarity in space and global correlation in the spectrum, is not enough to eliminate the structure noise because of the similar redundant properties of the structure noise in the HSI. A residual image containing the high-frequency information is used to remove the mixed noise by describing the residual intrinsic spectral and spatial characteristics. From the spectral perspective, an assumption that each HS band can be represented by the adjacent bands based on the high spectral correlation is utilized to design a spectral guidance-based variational (SGV) model. First, the subtraction between each HS band and its estimation based on adjacent bands forms the residual image, which is utilized to restrain the spectral correlation. Second, in the

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

spatial dimension, two differential regularizations in different directions are applied to characterize the residual image by capturing the smoothness along the stripes and the discontinuity across the stripes. Finally, an alternating direction method of multipliers (ADMM) algorithm [54] is used to achieve a robust solution for the HRHS image. The experimental results on several data sets confirm the effectiveness of the proposed method and its obvious antinoising capability for LRHS and HRMS fusion under a noise scenario.

The remainder of this paper is organized as follows. Section II introduces some notations of various parameters and the basic relationship between the two observed images and the desired fusion result. Section III formulates the proposed fusion algorithm and the optimization strategy. The simulated- and real-data experimental results followed by the quantitative and visual assessments are presented in Section IV. Finally, our conclusion is given in Section V.

II. Problem Formulation

A. Notation and Preliminaries

Throughout this paper, we denote scalars, vectors, matrices, and tensors by the nonbold letters, bold lower case letters, bold upper case letters, and calligraphic upper case letters, respectively. It is known that a tensor can be seen as a multi-index numerical array, and its order is defined as the number of its modes or dimensions. We shall provide a brief introduction to tensor algebra in the following.

A tensor of order N is denoted by $\mathcal{X} \in \mathrm{R}^{I_1 \times \ldots \times I_n \times \ldots \times I_N}$ with the N-dimensional data array. Its element is denoted by $x_{i_1 \ldots i_n \ldots i_N}$, where $1 \leq i_n \leq I_n$. The mode-n flattening of the tensor $\mathcal{X}$ is denoted by $\mathbf{X}_{(n)} \in \mathrm{R}^{I_n \times J_n}$, where the tensor element $(i_1 \ldots i_n \ldots i_N)$ maps to the matrix element (i_n, j) with $j = 1 + \sum_{k=1,k\neq n}^{N}(i_k - 1)J_k$, where $J_k = \prod_{l=1,l\neq n}^{k-1} I_l$. The mode-$n$ multiplication of a tensor $\mathcal{X}$ with a matrix $\mathbf{U} \in \mathrm{R}^{J \times I_n}$, denoted by $\mathcal{X} \times_n \mathbf{U}$, is an Nth order tensor, with its elementwise $(\mathcal{X} \times_n \mathbf{U})_{i_1,\ldots,i_{n-1} j i_{n+1},\ldots,i_N} = \sum_{i_n} x_{i_1,i_2,\ldots,i_N} \cdot u_{j,i_n}$. The corresponding Frobenius norm of a tensor is defined as $\|\mathcal{X}\|_F = \sqrt{\langle \mathcal{X}, \mathcal{X} \rangle} = (\sum_{\boldsymbol{i_1,\ldots,i_N}} |\boldsymbol{x_{i_1,\ldots,i_N}}|^{\mathbf{2}})^{1/2}$. As a three-order tensor, the HSI can be translated into three different mode-n matricizations, respectively, representing the properties of two spatial and one spectral characteristics. For convenience, the mode-3 flattening of any HS tensor $\mathbf{X}_{(3)}$ is expressed as $\mathbf{X}$ in the following.

B. Fusion Problem With Noises

HSI $\mathcal{Y} \in \mathbb{R}^{M_h \times N_h \times L_h}$ and MS image $\mathcal{X} \in \mathbb{R}^{M_m \times N_m \times L_m}$ can be thought of as 3-D arrays or tensors, which are often called data cubes. Generally, for convenience, LRHS and HRMS images are represented as the mode-3 multiplication of the image, i.e., $\mathbf{Y} \in \mathbb{R}^{L_h \times M_h N_h}$ and $\mathbf{X} \in \mathbb{R}^{L_m \times M_m N_m}$, meaning lexicographically transforming a 3-D cube into a 2-D matrix representation along the spectral dimension. L_m and L_n are the numbers of spectral bands, and $M_m N_m$ and $M_h N_h$ represent the high-resolution and low-resolution spatial dimension, respectively, for the MS image and HSI. Thus, similar to the LRHS and HRMS images, the high spatial resolution HS cube $\mathcal{Z} \in \mathbb{R}^{M_m \times N_m \times L_h}$ to be estimated is denoted by $\mathbf{Z} \in \mathbb{R}^{L_h \times M_m N_m}$.

The relationship between the observed HSI and the recovered HRHS image can be described as

$$\mathbf{Y} = \mathbf{ZM} + \mathbf{E}_h \tag{1}$$

where $\mathbf{M} \in \mathbb{R}^{M_m N_m \times M_h N_h}$ is the spatial degradation factor (including blurring and down-sampling factors), which indicates spatial degradation and spectral consistency between $\mathbf{Y}$ and $\mathbf{Z}$. $\mathbf{E}_h$ represents the noise and model error. The model can also be called the spatial degradation model.

In real cases, it is known that HSIs are always corrupted by several different types of noise, e.g., Gaussian noise, stripes, and their mixture. The random noise is generally generated due to the limitations of equipment performance like sensor sensitivity, photon effects, and calibration error [55], [56], while the stripes occur mainly because of the inconsistent responses between different detectors [11], [12]. The noises in HSI are, thus, still of acute and can badly influence the fused HSI images if there is no proper way to cope with them.

Following the above description, the noise term $\mathbf{E}_h$ can be divided into two subterms $\mathbf{N}_h$ and $\mathbf{S}_h$, respectively, representing the Gaussian noise term and the sparse noise term like stripes. Thus, the spatial degradation model (1) will be transformed into the following degradation model:

$$\mathbf{Y} = \mathbf{ZM} + \mathbf{N}_h + \mathbf{S}_h. \tag{2}$$

On the other hand, the HRMS image $\mathbf{X}$ can also be measured by left multiplying spectral response transform factor $\mathbf{R}$ to $\mathbf{Z}$

$$\mathbf{X} = \mathbf{RZ} + \mathbf{E}_m \tag{3}$$

where $\mathbf{R} \in \mathbb{R}^{L_m \times L_h}$ is the spectral degradation and describes the spatial consistency between $\mathbf{X}$ and $\mathbf{Z}$. $\mathbf{E}_m$ represents the noise and model error.

III. Proposed MS and HS Image Fusion Algorithm

To recover the desired high spatial resolution HSI through LRHS and HRMS image fusion, the resulting HRHS image is assumed to possess the same spatial resolution as the HRMS image and the same spectral resolution as the input LRHS image. Therefore, based on both the spectral correlation with the LRHS image and the spatial correlation with the HRMS image, the HRHS image can be calculated via the model-based spatio-spectral fusion method in the following minimized cost function $E(\mathbf{Z})$, including the spectral fidelity term, the spatial enhancement term, and the prior term:

$$E(\mathbf{Z}) = f_{\text{spatial}}(\mathbf{Y}, \mathbf{Z}) + f_{\text{spatial}}(\mathbf{X}, \mathbf{Z}) + f_{\text{prior}}(\mathbf{Z}). \tag{4}$$

In general, the prior term is based upon reasonable assumptions or prior knowledge about the recovered HRHS image. After analyzing and mining the image properties, the proposed fusion algorithm exploits the low-multilinear-rank and the variational properties to construct the constraints for the recovered image. Namely, the low-multilinear-rank property of tensor is employed to indicate the high spatio-spectral redundancy, while the variational properties are used to excavate the difference of the desired HRHS image and the noisy image.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

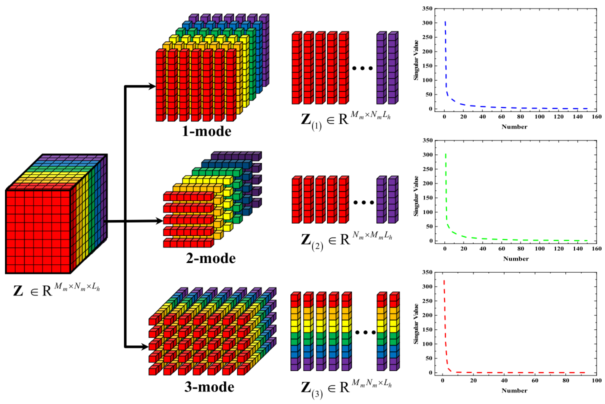


Fig. 1. Low-rank property in the HSI cube.

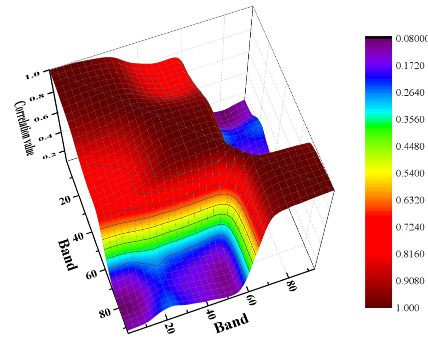


Fig. 2. Correlation surface map of the HSI.

A. Low-Multilinear-Rank Tensor

For the desired HRHS image, each spectral signature (row of the mode-3 matricization $\mathbf{Z}$) can be represented by a linear combination of a small number of pure spectral endmembers. As demonstrated in [57], the number of used endmembers is relatively smaller than the total number of bands, which means that a majority of singular values of $\mathbf{Z}$ are close to zero as shown in the singular value curve of $\mathbf{Z}$ (the so-called $\mathbf{Z}_{(3)}$ in Fig. 1). In addition, similar local patches exist in the space of typical remote sensing, composing of homologous aggregation of microstructures. Therefore, two spatial modes $\mathbf{Z}_{(1)}$ and $\mathbf{Z}_{(2)}$ also contain high correlations which can be reflected as the obvious decaying trends in the curves of the singular values, as shown in Fig. 1. Conversely, the random noise does not have the redundancy. Hence, by characterizing the high spatio-spectral correlation, tensor decomposition can be used to excavate the underlying clean HSI part by subtracting the structures of the heavy noise or artifacts part. For the estimated HSI, the optimization problem that we wish to solve is

$$\begin{aligned} \min_{C,U_1,U_2,U_3} \ & \alpha \|\mathcal{Z} - \mathcal{C} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \mathbf{U}_3\|_{\mathrm{F}}^2, \\ \text{s.t.} \ \ & C \in \mathbb{R}^{r_1 \times r_2 \times r_3}, \quad \mathbf{U}_1 \in \mathbb{R}^{M_m \times \pi}, \ \mathbf{U}_2 \in \mathbb{R}^{N_m \times r_2}, \\ & \mathbf{U}_3 \in \mathbb{R}^{L_h \times r_3} \\ & \text{and } \mathbf{U}_i^T \mathbf{U}_i = \mathbf{I} \quad (i = 1, 2, 3) \end{aligned} \tag{5}$$

where $\mathcal{C} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \mathbf{U}_3$ is the Tucker decomposition with core tensor $\mathcal{C}$ and factor matrices $\mathbf{U}_i$ of rank r_i. The basic idea is to find those components $\mathbf{U}_i$ that best capture the variation in mode n, independent of the other modes. Corresponding to these components, a low-multilinear-ranks constraint $\text{rank}_i(\mathbf{Z}_{(i)}) \leq r_i$ can be used to draw off the components of the dominant singular vectors $\mathbf{Z}_{(i)}$ which describe the clean HSI part.

However, the use of only the tensor decomposition is not enough to eliminate the structure noise, because the structure noise has an obvious pattern in the spatial dimension, i.e., the high redundancy in the HSI. It is, therefore, necessary to characterize the structures of the noise part and add a proper prior to better deal with the mixed noise.

B. Variational Properties via Spectral Guidance

It is well known that the adjacent bands of HSIs are highly correlated and similar according to the correlation in Fig. 2. Furthermore, different bands with high correlation have different gradient strengths, meaning that the high-correlation bands with more detailed information can be better used to enhance the current band. As stated in [58], each image patch can be reconstructed by its nonlocal similar patches, and the spatial noise can be expected to be prominently alleviated by averaging among these similar patches. Similarly, in this work, it is assumed that a band can also be efficiently fitted by its neighbor and similar bands

$$\mathbf{Z}_l = \sum_{j \neq l, j=1}^{L_h} \omega_j \mathbf{Z}_j \tag{6}$$

where $\mathbf{Z}_l$ is the lth band of the HSI, $\mathbf{Z}_j$ is the neighboring bands, and ω_j denotes the weight assigned into the neighboring bands. ω_j can be estimated by the least-squares method [38], [59] and formulated as the weight matrix $\mathbf{W}$. The above assumption provides us with reasonable knowledge to form a gradient relation to ensure the spatial and spectral consistency

$$D\mathbf{Z}_l = D\left(\sum_{j \neq l, j=1}^{L_h} \omega_j \mathbf{Z}_j\right) + \varepsilon \tag{7}$$

where D is the difference operator containing D_1 and D_2, while D_1 and D_2 are the first-order horizontal and vertical difference operators, respectively. With regard to an HSI with stripes, D_1 and D_2 also refer to the across-stripe and along-stripe directions of each band and are used to capture the directional property of the stripe noise. ε is the residual image representing the model error, which is also obtained through the subtraction of the fitted image from the current band l. Each band in the HRHS images is related to the corresponding band in the LRHS images because the LRHS images can be regarded as the spatial degradation of the HRHS images. It is assumed that the similar local geometry in the LRHS images is shared with the HRHS images. Thus, the $\mathbf{W}$ can be calculated through the LRHS image and then transplanted to the HRHS images.

Since the gradient reflects the variation between adjacent pixels, the smoothness can be effectively described by minimizing the local difference along both the vertical and hori-

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

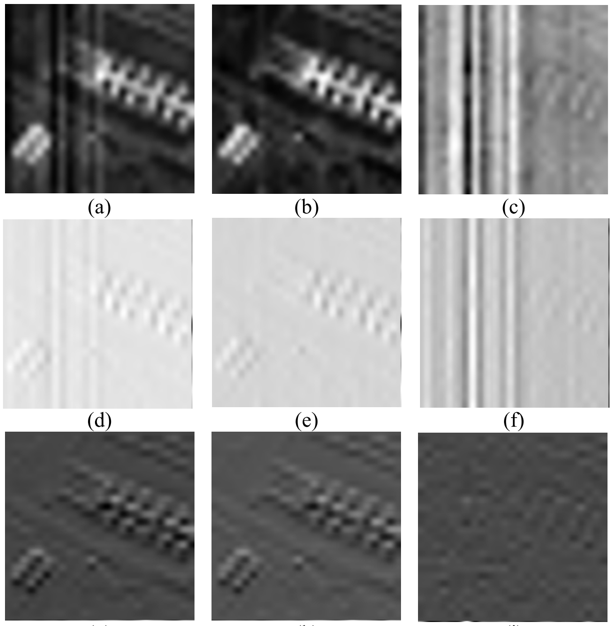


Fig. 3. Variational property of an HSI in both vertical and horizontal modes for the initial 1st band image under a mixed-noise situation. (a)–(c) Current band image $\mathbf{Z}_1$, the fitted image $\sum_{j\neq 1}^{L_h} w_j\mathbf{Z}_j$, and the residual image between them. (d)–(f) Across-stripe (horizontal) differential result of $\mathbf{Z}_1$, the fitted image, and the residual image between the two differential results. (g)–(i) Along-stripe (vertical) differential result of $\mathbf{Z}_1$, the fitted image, and the residual image between the two differential results.

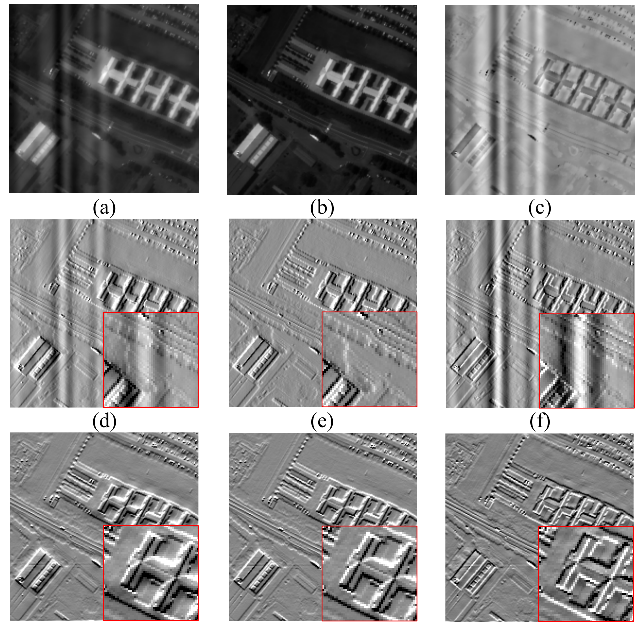


Fig. 4. Variational property of an HSI in both vertical and horizontal modes for the 1st band image after iteration under the mixed-noise situation. (a)–(c) Current band image $\mathbf{Z}_1$, the fitted image $\sum_{j\neq 1}^{L_h} w_j\mathbf{Z}_j$, and the residual image between them. (d)–(f) Across-stripe (horizontal) differential result of $\mathbf{Z}_1$, the fitted image, and the residual image between the two differential results. (g)–(i) Along-stripe (vertical) differential result of $\mathbf{Z}_1$, the fitted image, and the residual image between the two differential results.

zontal directions. For example, the local smooth structure in the along-stripe (vertical) direction is given in Figs. 3(g)–(i) and 4(g)–(i). However, the smooth structure of an HSI in the across-stripe direction will be damaged when the structure noise exists in the space. More specifically, the gradients in the across-stripe (horizontal) direction often have a large absolute value at the stripe locations and a small value at the nonstripe locations. Hence, the saliency characteristics of discontinuity in the across-stripe gradient image can reflect the actual location of a stripe.

Considering the fact that the differences between the along- and across-stripe properties, it is natural to construct two different terms to help suppress the discontinuity of the stripes and maintain the local spatial smoothness. Accordingly, combined with (7), an SGV regularization employing the high-correlation bands is designed to solve the aforementioned mixed noise including both the random and the structure noise

$$\lambda_{\text{utv}}\|D_1(\mathbf{Z}-\mathbf{WZ})\|_1 + \lambda_{\text{tv}}\text{TV}(D(\mathbf{Z}-\mathbf{WZ})) \qquad (8)$$

where λ_{utv} and λ_{tv} are two parameters to balance the constraining degree of the two terms. The first l_1 norm term is the unidirectional total variation used to describe the discontinuity across the stripes, while the second term is the anisotropic total variation-based gradient consistency constraint, representing the distributional properties of the local smoothness of the residual image and the spatial preservation with the sharper fitted image. To facilitate the calculation, we can simplify (8) as $\lambda_1\|D_1(\mathbf{Z}-\mathbf{WZ})\|_1 + \lambda_2\|D_2(\mathbf{Z}-\mathbf{WZ})\|_1$, where λ_1 is the combination of λ_{utv} and λ_{tv}, and λ_2 is equivalent to λ_{tv}.

Three advantages arise when the SGV regularization is implemented. First, on account of the relationship consistency along the spectral bands, the band correlations of the LRHS images can be preserved in the recovered HRHS images, which better characterizes the spectral property. Second, by weighted averaging among similar bands, the spatial noise can be expected to be suppressed, as shown in Figs. 3(b) and 4(b). Furthermore, it can be clearly observed that the along-stripe differential result of $\mathbf{Z}_l$ [Fig. 4(g)] and the fitted image [Fig. 4(h)] have the same gradient direction. Then, the residual image with the opposite gradient then reveals the fact that the differential absolute value $|D\mathbf{Z}_l|$ of $\mathbf{Z}_l$ is smaller than the differential absolute value of its fitted image. In other words, the fitted image can bring more detailed information. Therefore, the gradient consistency constraint between $\mathbf{Z}_l$, and its fitted image can better enhance the spatial structure. Third, due to the sparsity of the residual image after removing the low-frequency information, its differential results in Figs. 3(f) and 4(f) can better point out the characteristics and location of the stripes, in comparison with Figs. 3(d) and 4(d). The sensitivity to the stripes in the residual image, in turn, helps to find and suppress the underlying discontinuous structure of the stripe noise.

C. Proposed Fusion Model

Considering the tensor low-multilinear-rank and variational properties of imagery, we propose a novel antinoise HSI fusion by exploiting low-multilinear-rank tensor constraint and SGV. After combining the proposed regularizers, we can formulate a constrained optimization problem as

$$\min_{\mathbf{Z},\mathbf{S}_h,\mathbf{N}_h} \frac{\alpha}{2}\|\mathbf{X}-\mathbf{R}\mathbf{Z}\|_F^2+\beta\|\mathbf{S}_h\|_1+\gamma\|\mathbf{N}_h\|_F^2 +\lambda_1\|D_1(\mathbf{Z}-\mathbf{W}\mathbf{Z})\|_1+\lambda_2\|D_2(\mathbf{Z}-\mathbf{W}\mathbf{Z})\|_1 \quad \text{s.t. } \mathbf{Y}=\mathbf{Z}\mathbf{M}+\mathbf{N}_h+\mathbf{S}_h,\quad \text{rank}_i(\mathbf{Z}_{(i)})\le r_i \tag{9}$$

where the parameters α, β, and γ are used to control the relative importance of different terms. The first term is the data-fitting term, imposing that the HRHS image $\mathbf{Z}$ should be able to explain the observed HRMS data $\mathbf{X}$ according to the spectral relationship model defined in (3). The constraint term $\mathbf{Y}=\mathbf{Z}\mathbf{M}+\mathbf{N}_h+\mathbf{S}_h$ is the spatial reconstruction constraint with LRHS image $\mathbf{Y}$ and the estimated HRHS image $\mathbf{Z}$ defined in (1) and (2). The l_2 norm regularization $\|\mathbf{N}_h\|_F^2$ can be used to describe the Gaussian distribution of the noise. Meanwhile, the l_1 norm used in the sparsity regularization $\|\mathbf{S}_h\|_1$ and SGV regularization can not only fit the sparsity of the stripes but can also capture the fine details in the residual map. The SGV regularization, in particular, can help to simultaneously characterize the piecewise smooth structure and suppress the discontinuity caused by the noise. The l_1 norm is chosen to ensure that the differential information across the stripe from $D_j(\mathbf{Z})(j=1,2)$ can be more accurately mapped to $D_j(\mathbf{W}\mathbf{Z})(j=1,2)$. As a result, the spectral correlation of the desired image can be well preserved.

D. Optimization Problem

By introducing some auxiliary variables, we rewrite (9) as the following equivalent minimization problem:

$$\begin{aligned}\min_{\mathbf{Z},\mathbf{S}_h,\mathbf{N}_h}\ & \frac{\alpha}{2}\|\mathbf{X}-\mathbf{R}\mathbf{V}\|_F^2+\beta\|\mathbf{S}_h\|_1+\gamma\|\mathbf{N}_h\|_F^2 \\ & +\lambda_1\|\mathbf{Q}_1\|_1+\lambda_2\|\mathbf{Q}_2\|_1 \\ \text{s.t. } & \mathbf{Y}=\mathbf{Z}\mathbf{M}+\mathbf{N}_h+\mathbf{S}_h,\quad \mathbf{H}=\mathbf{Z}, \text{rank}_i(\mathbf{H}_{(i)})\le r_i \\ & \mathbf{V}=\mathbf{Z},\quad \mathbf{E}=\mathbf{V}-\mathbf{W}\mathbf{V}, \mathbf{Q}_1=D_1(\mathbf{E}), \mathbf{Q}_2=D_2(\mathbf{E})\end{aligned} \tag{10}$$

where $\mathbf{Q}_1$, $\mathbf{Q}_2$, $\mathbf{V}$, $\mathbf{H}$, and $\mathbf{E}$ are the auxiliary variables. Then, the constrained optimization problem can be transformed into the following augmented Lagrangian function:

$$\begin{aligned}& L(\mathbf{Z},\mathbf{S}_h,\mathbf{N}_h,\mathbf{H},\mathbf{V},\mathbf{E},\mathbf{Q}_1,\mathbf{Q}_2,\mathbf{P}_1,\mathbf{P}_2,\mathbf{P}_3,\mathbf{P}_4,\mathbf{P}_5,\mathbf{P}_6) \\ &= \frac{\alpha}{2}\|\mathbf{X}-\mathbf{R}\mathbf{V}\|_F^2+\beta\|\mathbf{S}_h\|_1+\gamma\|\mathbf{N}_h\|_F^2 \\ &\quad +\lambda_1\|\mathbf{Q}_1\|_1+\lambda_2\|\mathbf{Q}_2\|_1+\langle\mathbf{P}_1,\mathbf{Y}-\mathbf{Z}\mathbf{M}-\mathbf{N}_h-\mathbf{S}_h\rangle \\ &\quad +\langle\mathbf{P}_2,\mathbf{Z}-\mathbf{V}\rangle+\langle\mathbf{P}_3,\mathbf{V}-\mathbf{W}\mathbf{V}-\mathbf{E}\rangle+\langle\mathbf{P}_4,D_1(\mathbf{E})-\mathbf{Q}_1\rangle \\ &\quad +\langle\mathbf{P}_5,D_2(\mathbf{E})-\mathbf{Q}_2\rangle+\langle\mathbf{P}_6,\mathbf{Z}-\mathbf{H}\rangle \\ &\quad +\frac{\mu}{2}\|\mathbf{Y}-\mathbf{Z}\mathbf{M}-\mathbf{N}_h-\mathbf{S}_h\|_F^2 \\ &\quad +\frac{\mu}{2}\|\mathbf{Z}-\mathbf{V}\|_F^2+\frac{\mu}{2}\|\mathbf{V}-\mathbf{W}\mathbf{V}-\mathbf{E}\|_F^2+\frac{\mu}{2}\|\mathbf{Z}-\mathbf{H}\|_F^2 \\ &\quad +\frac{\mu}{2}\|D_1(\mathbf{E})-\mathbf{Q}_1\|_F^2+\frac{\mu}{2}\|D_2(\mathbf{E})-\mathbf{Q}_2\|_F^2\end{aligned} \tag{11}$$

where $\mathbf{P}_i(i=1,2,3,4,5,6)$ are the Lagrange multipliers, and μ represent the penalty parameter, which determines the step sizes used to update the corresponding Lagrange multipliers. The above function can be solved by the ADMM. Each iteration of the algorithm can be decomposed into seven simpler subproblems, and their variables are updated in an alternating and sequential way.

1) The $\mathbf{Z}$ subproblem is given by all the terms containing $\mathbf{Z}$ from the function

$$\begin{aligned}\min_{\mathbf{Z}}\ & \left\langle\mathbf{P}_1^{(k)},\mathbf{Y}-\mathbf{Z}\mathbf{M}-\mathbf{N}_h^{(k)}-\mathbf{S}_h^{(k)}\right\rangle+\left\langle\mathbf{P}_2^{(k)},\mathbf{Z}-\mathbf{V}^{(k)}\right\rangle \\ & +\langle\mathbf{P}_6,\mathbf{Z}-\mathbf{H}\rangle+\frac{\mu}{2}\left\|\mathbf{Y}-\mathbf{Z}\mathbf{M}-\mathbf{N}_h^{(k)}-\mathbf{S}_h^{(k)}\right\|_F^2 \\ & +\frac{\mu}{2}\|\mathbf{Z}-\mathbf{V}^{(k)}\|_F^2+\frac{\mu}{2}\|\mathbf{Z}-\mathbf{H}^{(k)}\|_F^2.\end{aligned} \tag{12}$$

Equation (12) is a quadratic minimization problem. It has an explicit formula

$$\begin{aligned}\mathbf{Z}(\mathbf{M}\mathbf{M}^T+2\mathbf{I}) &= \left(\mathbf{Y}-\mathbf{N}_h^{(k)}-\mathbf{S}_h^{(k)}\right)\mathbf{M}^T \\ &\quad +\mathbf{V}^{(k)}+\mathbf{H}^{(k)}+\left(\mathbf{P}_1^{(k)}\mathbf{M}^T-\mathbf{P}_2^{(k)}-\mathbf{P}_6^{(k)}\right)/\mu.\end{aligned} \tag{13}$$

The preconditioned conjugate gradients (PCG) method is used to solve (13) and get the mode-3 flattening $\mathbf{Z}$.

2) Solving $\mathbf{H}$ subproblem needs to consider the following problem:

$$\frac{\mu}{2}\|(\mathbf{Z}+\mathbf{P}_6/\mu)-\mathbf{H}\|_F^2 \quad \text{s.t. } \text{rank}_i(\mathbf{H}_{(i)})\le r_i. \tag{14}$$

Due to $\text{rank}_i(\mathbf{H}_{(i)})\le r_i$ is equivalent to a Tucker model for $\mathcal{H}$ with factor matrices $\mathbf{U}_i$ of rank r_i, then, the Tucker decomposition is replaced to optimize the subproblem for $\mathcal{H}$ as follows:

$$\min_{\mathbf{U}_i^T\mathbf{U}_i=\mathbf{I}} \frac{\mu}{2}\|(\mathcal{Z}+\mathcal{P}_6/\mu)-\mathcal{C}\times_1\mathbf{U}_1\times_2\mathbf{U}_2\times_3\mathbf{U}_3\|_F^2. \tag{15}$$

By the classic higher order orthogonal iteration (HOOI) algorithm used in [60], we can easily obtain $\mathcal{C}^{(k+1)}$ and $\mathbf{U}_i^{(k+1)}(i=1,2,3)$. Then, $\mathcal{H}$ can then be updated by $\mathcal{H}^{(k+1)}=\mathcal{C}^{(k+1)}\times_1\mathbf{U}_1^{(k+1)}\times_2\mathbf{U}_2^{(k+1)}\times_3\mathbf{U}_3^{(k+1)}$.

3) The $\mathbf{V}$ subproblem is given by

$$\begin{aligned}\mathbf{V}^{(k+1)} &= \operatorname*{argmin}_{\mathbf{V}} \frac{\alpha}{2}\|\mathbf{X}-\mathbf{R}\mathbf{V}\|_F^2+\left\langle\mathbf{P}_2^{(k)},\mathbf{Z}^{(k+1)}-\mathbf{V}\right\rangle \\ &\quad +\left\langle\mathbf{P}_3^{(k)},\mathbf{V}-\mathbf{W}\mathbf{V}-\mathbf{E}^{(k)}\right\rangle+\frac{\mu}{2}\|\mathbf{Z}^{(k+1)}-\mathbf{V}\|_F^2 \\ &\quad +\frac{\mu}{2}\|\mathbf{V}-\mathbf{W}\mathbf{V}-\mathbf{E}^{(k)}\|_F^2\end{aligned} \tag{16}$$

and hence,

$$\begin{aligned}& [\alpha\mathbf{R}^T\mathbf{R}+\mu\mathbf{I}+\mu(\mathbf{I}-\mathbf{W})^T(\mathbf{I}-\mathbf{W})]\mathbf{V}^{(k+1)} \\ &= \alpha\mathbf{R}^T\mathbf{X}+\mu\mathbf{Z}^{(k+1)}+\mu(\mathbf{I}-\upsilon\mathbf{W})^T\mathbf{E}^{(k)}+\mathbf{P}_2^{(k)}-(\mathbf{I}-\mathbf{W})^T\mathbf{P}_3^{(k)}\end{aligned} \tag{17}$$

which can be simply solved by the PCG method.

4) The $\mathbf{E}$ subproblem is given by

$$\begin{aligned}\mathbf{E}^{(k+1)} &= \operatorname*{argmin}_{\mathbf{E}}\langle\mathbf{P}_3,\mathbf{V}-\mathbf{W}\mathbf{V}-\mathbf{E}\rangle+\left\langle\mathbf{P}_4,D_1(\mathbf{E})-\mathbf{Q}_1^{(k)}\right\rangle \\ &\quad +\left\langle\mathbf{P}_5,D_2(\mathbf{E})-\mathbf{Q}_2^{(k)}\right\rangle+\frac{\mu}{2}\|\mathbf{V}^{(k+1)}-\mathbf{W}\mathbf{V}^{(k+1)}-\mathbf{E}\|_F^2 \\ &\quad +\frac{\mu}{2}\|D_1(\mathbf{E})-\mathbf{Q}_1^{(k)}\|_F^2+\frac{\mu}{2}\|D_2(\mathbf{E})-\mathbf{Q}_2^{(k)}\|_F^2.\end{aligned} \tag{18}$$

This problem can be transformed into the following linear system:

$$\begin{aligned}& (\mu\mathbf{I}+\mu D_1^*D_1+\mu D_2^*D_2)\mathbf{E}^{(k+1)} \\ &= \mu(\mathbf{V}^{(k+1)}-\mathbf{W}\mathbf{V}^{(k+1)})+\mu D_1^*\left(\mathbf{Q}_1^{(k)}\right) \\ &\quad +\mu D_2^*\left(\mathbf{Q}_2^{(k)}\right)+\mathbf{P}_3^{(k)}-D_1^*\left(\mathbf{P}_4^{(k)}\right)-D_2^*\left(\mathbf{P}_5^{(k)}\right)\end{aligned} \tag{19}$$

where $D_j^*(j=1,2)$ indicates the adjoint operator of $D_j(j=1,2)$. Equation (19) can be efficiently solved by fast Fourier transform (FFT).

5) The $\mathbf{Q}_1$ and $\mathbf{Q}_2$ subproblems are given by

$$\mathbf{Q}_1^{(k+1)} = \underset{\mathbf{Q}_1}{\operatorname{argmin}} \langle \mathbf{P}_4^{(k)}, D_1(\mathbf{E}^{(k+1)}) - \mathbf{Q}_1 \rangle + \lambda_1 \|\mathbf{Q}_1\|_1 + \frac{\mu}{2}\|D_1(\mathbf{E}^{(k+1)}) - \mathbf{Q}_1\|_F^2 \tag{20}$$

which can be solved using a soft-threshold shrinkage operator as follows:

$$\mathbf{Q}_1^{(k+1)} = \mathcal{F}_{\frac{\lambda_1}{\mu}}\left(D_1(\mathbf{E}^{(k+1)}) + \frac{\mathbf{P}_4^{(k)}}{\mu}\right) \tag{21}$$

where

$$\mathcal{F}_t(x) = \begin{cases} x-t, & x>t \\ 0, & |x| \le t \\ x+t, & x<-t. \end{cases} \tag{22}$$

Similarly, $\mathbf{Q}_2$ can also be obtained using the above operator.

6) The $\mathbf{S}_h$ subproblem is given by

$$\mathbf{S}_h^{(k+1)} = \underset{\mathbf{S}_h}{\operatorname{argmin}}\, \beta\|\mathbf{S}_h\|_1 + \langle \mathbf{P}_1^{(k)}, \mathbf{Y} - \mathbf{Z}^{(k+1)}\mathbf{M} - \mathbf{N}_h^{(k)} - \mathbf{S}_h \rangle + \frac{\mu}{2}\|\mathbf{Y} - \mathbf{Z}^{(k+1)}\mathbf{M} - \mathbf{N}_h^{(k)} - \mathbf{S}_h\|_F^2. \tag{23}$$

By using the above soft-thresholding operator, the solution of the subproblem can be formulated as

$$\mathbf{S}_h^{(k+1)} = \mathcal{F}_{\frac{\beta}{\mu}}\left(\mathbf{Y} - \mathbf{Z}^{(k+1)}\mathbf{M} - \mathbf{N}_h^{(k)} + \frac{\mathbf{P}_1^{(k)}}{\mu}\right). \tag{24}$$

7) The $\mathbf{N}_h$ subproblem is given by

$$\mathbf{N}_h^{(k+1)} = \underset{\mathbf{N}_h}{\operatorname{argmin}}\, \gamma\|\mathbf{N}_h\|_F^2 + \langle \mathbf{P}_1^{(k)}, \mathbf{Y} - \mathbf{Z}^{(k+1)}\mathbf{M} - \mathbf{N}_h - \mathbf{S}_h^{(k+1)} \rangle + \frac{\mu}{2}\|\mathbf{Y} - \mathbf{Z}^{(k+1)}\mathbf{M} - \mathbf{N}_h - \mathbf{S}_h^{(k+1)}\|_F^2. \tag{25}$$

Then,

$$\mathbf{N}_h^{(k+1)} = \frac{\mu(\mathbf{Y} - \mathbf{Z}^{(k+1)}\mathbf{M} - \mathbf{S}_h^{(k+1)}) + \mathbf{P}_1^{(k)}}{\mu + 2\gamma}. \tag{26}$$

8) Finally, in each iteration, the Lagrange multipliers $\mathbf{P}_i(i=1,2,3,4,5,6)$ are updated as follows:

$$\begin{cases} \mathbf{P}_1^{(k+1)} = \mathbf{P}_1^{(k)} + \mu(\mathbf{Y} - \mathbf{Z}^{(k+1)}\mathbf{M} - \mathbf{N}_h^{(k+1)} - \mathbf{S}_h^{(k+1)}) \\ \mathbf{P}_2^{(k+1)} = \mathbf{P}_2^{(k)} + \mu(\mathbf{Z}^{(k+1)} - \mathbf{V}^{(k+1)}) \\ \mathbf{P}_3^{(k+1)} = \mathbf{P}_3^{(k)} + \mu(\mathbf{V}^{(k+1)} - \mathbf{W}\mathbf{V}^{(k+1)} - \mathbf{E}^{(k+1)}) \\ \mathbf{P}_4^{(k+1)} = \mathbf{P}_4^{(k)} + \mu(D_1(\mathbf{E}^{(k+1)}) - \mathbf{Q}_1^{(k+1)}) \\ \mathbf{P}_5^{(k+1)} = \mathbf{P}_5^{(k)} + \mu(D_2(\mathbf{E}^{(k+1)}) - \mathbf{Q}_2^{(k+1)}) \\ \mathbf{P}_6^{(k+1)} = \mathbf{P}_6^{(k)} + \mu(\mathbf{Z}^{(k+1)} - \mathbf{H}^{(k+1)}). \end{cases} \tag{27}$$

Combining these subproblems from (1)–(8) introduced, we have a one-step iteration for the ADMM. By decomposing the difficult minimization problem into several easy subproblems, the proposed model can be summarized as Algorithm 1.

Algorithm 1 Proposed Fusion Algorithm

Input: HRMS image $\mathbf{X}$, LRHS image $\mathbf{Y}$, parameters α, β, γ, μ, λ_1, λ_2, maximum iterations k_{max}, and stopping criterion τ.

Initialize: $\mathbf{P}_i$ $(i=1,2,3,4,5,6) = \mathbf{0}$,$\mathbf{Q}_1 = \mathbf{0}$,$\mathbf{Q}_2 = \mathbf{0}$,$\mathbf{N}_h = \mathbf{0}$,$\mathbf{S}_h = \mathbf{0}$,$\tau$=$10^{-4}$, $\mathbf{H}$ and $\mathbf{V}$ is the interpolation of $\mathbf{Y}$, $\mathbf{E}$ is initialized by $\mathbf{V} - \mathbf{W}\mathbf{V}$.

While $(\|\mathbf{Z}^{(k)} - \mathbf{Z}^{(k+1)}\|_F^2 / \|\mathbf{Y}\|_F^2 > \tau$ and $k < k_{max})$ **do**

Update $\mathbf{Z}^{(k+1)}$, $\mathbf{H}^{(k+1)}$, $\mathbf{V}^{(k+1)}$, $\mathbf{E}^{(k+1)}$, $\mathbf{Q}_1^{(k+1)}$, $\mathbf{Q}_2^{(k+1)}\mathbf{S}_h^{(k+1)}$, $\mathbf{N}_h^{(k+1)}$by repeating (13), (15), (17), (19), (21), (24), (26)

Update the Lagrange multipliers $\mathbf{P}_i^{(k+1)}$ $(i=1,2,3,4,5,6)$ by (27)

End While

Output: The fused HRHS image $\mathbf{Z}$.

IV. Experiments and Discussion

To highlight our contribution and verify the effectiveness of the proposed method, we compare the results of the related HSI fusion methods to the fusion results of the proposed method, both visually and quantitatively. The five compared methods are HS super-resolution (HySure) [38], CNMF [42], NSSR [46], image fusion based on a sparse representation (BSR) [32], and HS super-resolution using proximal alternating linearized minimization (SupResPALM) [47]. We have conducted experiments on both simulated LRHS images and real-world LRHS images. The mean peak-signal-to-noise ratio (MPSNR) [61], the mean structural similarity (MSSIM) index [62], the mean spectral angle (MSA) mapper [63], the universal image quality index (UIQI) [64], the root-mean-square error (RMSE), the erreur relative globale adimensionnelle de synthèse (ERGAS) [65], and the correlation coefficient (CC) served as evaluation indices for the simulated experiments. Generally speaking, higher MPSNR, MSSIM, UIQI, and CC values reflect the better fusion results, while a lower RMSE, ERGAS, and MSA values means that the fusion results maintain better quality of the fused image, with lower spectral distortion.

The basic parameters of the proposed method are discussed as follows. The maximum iteration number $k_{\max}$ in Algorithm 1 was set as 50. The adjacent spectral band number K was set as 20 for both the simulated- and real-data experiments. Due to different degradation levels of the test images in our experiments, the functioning degree of the corresponding constraint terms in model (10) needs to be adjusted accordingly through the regularization parameters. The coefficient α for the spatial enhancement term can control the degree of the fused detailed information from the MS image. However, the overlarger value of α will damage the spectral fidelity, and thus, the coefficient value was chosen as 1 for all the experiments in this paper. The parameters λ_1 and λ_2 are related to the noise level. Specifically, λ_1 for the across-stripe differential term stands out a superior ability to control the stripe noise levels, and a larger λ_1 can help to eliminate the traces of heavy stripe. When the image is simultaneously

TABLE I

QUANTITATIVE EVALUATION OF THE FUSION RESULTS FOR THE SIMULATED EXPERIMENTS WITH $s = 4$

	Factor	RMSE	UIQI	PSNR	SSIM	SAM	CC	ERGAS
	4 *Case*1	0.015420	0.9897	36.6354	0.9788	2.9998	0.9908	2.0461
HySure	4 *Case*2	0.024627	0.9740	33.5710	0.9384	6.5696	0.9754	2.9198
	4 *Case*3	0.030554	0.9607	30.5449	0.8993	8.4752	0.9617	4.1188
	4 *Case*1	0.010371	0.9956	39.9946	0.9884	2.1815	0.9970	1.3932
CNMF	4 *Case*2	0.064802	0.8822	27.5611	0.8071	17.1633	0.9071	6.7483
	4 *Case*3	0.031551	0.9612	30.2183	0.9129	7.8342	0.9648	4.2804
	4 *Case*1	0.012575	0.9926	39.6127	0.9733	2.7319	0.9929	1.5042
NSSR	4 *Case*2	0.045425	0.9161	28.3574	0.7545	12.4598	0.9179	5.1114
	4 *Case*3	0.040805	0.9304	27.8820	0.8136	11.3535	0.9315	5.4711
	4 *Case*1	0.007946	0.9972	42.8610	0.9903	1.9592	0.9973	1.0341
BSR	4 *Case*2	0.017623	0.9869	36.9323	0.9575	4.8119	0.9872	2.0361
	4 *Case*3	0.024651	0.9741	32.5198	0.9286	6.0807	0.9754	3.2655
	4 *Case*1	0.008256	0.9970	42.1618	0.9911	**1.7187**	0.9970	1.1353
SupResPALM	4 *Case*2	0.020348	0.9825	36.7983	0.9560	5.7875	0.9827	2.1572
	4 *Case*3	0.021323	0.9810	33.7592	0.9505	5.9000	0.9818	2.8632
	4 *Case*1	**0.007398**	**0.9975**	**43.7959**	**0.9918**	1.7823	**0.9977**	**0.9779**
Proposed	4 *Case*2	**0.009250**	**0.9961**	**42.0208**	**0.9900**	**2.2487**	**0.9964**	**1.1983**
	4 *Case*3	**0.010623**	**0.9951**	**40.6601**	**0.9894**	**2.6541**	**0.9956**	**1.3521**

contaminated by the stripes and Gaussian noise, λ_1 should be larger than the value of λ_2, in order to overcome the discontinuity of the stripes. λ_1 and λ_2 were empirically set within the range from 0.0001 to 0.005 in this paper. The proportion of the stripes is in close accordance with the sparsity parameter γ. In addition, the parameter β is used to control the processing strength for the Gaussian noise. The more severe the Gaussian noise level is, the larger should be the parameter β. The parameter β was set within the range [0.5, 1] in our implementations. Notably, the tensor decomposition of rank $r_i (i = 1, 2, 3)$ can extract the clear HSI part, and thus, an image with the high noise intensity needs to select the lower $r_i (i = 1, 2, 3)$ for the three dimensions. To simplify the steps of the parameter adjustment in ADMM algorithm, the coefficient μ was tuned by $\mu^{(k+1)} = 1.2\mu^{(k)}$ with an initial value of 0.001. In all the experiments, the parameters of the other methods were adjusted to the optimum.

The test data sets used in the experiments are described in the following, for both the simulated- and real-data experiments.

1) Data set A was acquired by the Reflective Optics System Imaging Spectrometer (ROSIS) optical sensor over the urban area of the University of Pavia, Pavia, Italy. The reference image scene was cropped to $200 \times 200 \times 93$ after removing 22 water absorption bands. The image spans the 0.43–0.86-μm spectral range and has a spatial resolution of 1.3 m.
2) The data set B used in the real-data experiments was made up of images taken over Suzhou in China, which were obtained by the earth Observing-1 Mission (EO-1) satellite and the Gaofen-1 (GF-1, Gaofen means high resolution in Chinese) satellite. The Hyperion sensor onboard EO-1 is an HS imager with a spatial resolution of 30 m, while the GF-1 provides MS images at resolutions of 16 m, containing four MS bands, spanning the visible to the near-infrared spectral regions from 0.45 to 0.89 μm. To ensure spectral range consistency, the total number of 38 bands and spatial size of 400×362 from Hyperion were used to enhance the spectral information of GF-1. To allow a quantitative evaluation, the gray values of all the HSI bands were normalized to [0,1].

A. Simulated-Data Experiments

In the simulated HSI fusion process, the HSIs from the data set A serving as ground-truth images were used to generate simulated LRHS images and HRMS images. As described in [38], HRMS images $\mathbf{Y}$ were generated by filtering the ground-truth images along the spectral dimension using the reflectance spectral responses like the IKONOS. As for the LRHS images, the original HRHS images were first downsampled to obtain the LRHS images by averaging over disjoint $s \times s$ blocks, where s was the scaling factor of 4 or 8. In the next step, the LRHS images were contaminated with three different cases of additional noise detailed as follows. In case 1, the LRHS image was contaminated by low-intensity Gaussian noise, and the simulation was conducted with 30-dB SNR. The SNR of each band with the noise standard deviation σ_l was defined as follows:

$$\mathrm{SNR}_l = 10\log\left(\frac{\|(\mathbf{ZM})_l\|^2}{\sigma_l^2}\right), \quad l = 1, \ldots, L_h. \tag{28}$$

In case 2, the simulations were generated with 10-dB SNR for each band, meaning that all bands were contaminated by high-intensity Gaussian noise.

TABLE II
QUANTITATIVE EVALUATION OF THE FUSION RESULTS FOR THE SIMULATED EXPERIMENTS WITH $s = 8$

	Factor	RMSE	UIQI	PSNR	SSIM	SAM	CC	ERGAS
HySure	8 *Case*1	0.016477	0.9880	35.8494	0.9744	3.7095	0.9914	1.1600
	8 *Case*2	0.023296	0.9754	33.3873	0.9566	6.1576	0.9799	1.4727
	8 *Case*3	0.034614	0.9491	29.4654	0.9142	9.7372	0.9513	2.4039
CNMF	8 *Case*1	0.012587	0.9937	38.1310	0.9848	2.4431	0.9962	0.8762
	8 *Case*2	0.068036	0.86341	25.5911	0.8115	16.5834	0.8920	3.7817
	8 *Case*3	0.044623	0.9316	27.4963	0.9034	10.6504	0.9424	3.3192
NSSR	8 *Case*1	0.013305	0.9916	39.0229	0.9723	2.8623	0.9922	0.7992
	8 *Case*2	0.047417	0.9096	27.7375	0.7860	12.8284	0.9123	2.7352
	8 *Case*3	0.041987	0.9278	27.6187	0.8339	11.4447	0.9300	2.8744
BSR	8 *Case*1	**0.008808**	**0.9965**	41.9394	0.9896	2.1303	**0.9967**	0.5894
	8 *Case*2	0.016439	0.9885	37.3275	0.9789	4.5130	0.9892	0.9725
	8 *Case*3	0.033085	0.9556	29.8689	0.9253	9.6345	0.9575	2.2936
SupResPALM	8 *Case*1	0.008909	0.9964	41.4379	**0.9905**	**1.8169**	0.9965	0.6084
	8 *Case*2	0.021295	0.9806	35.3212	0.9665	6.0823	0.9815	1.1998
	8 *Case*3	0.026094	0.9728	31.9583	0.9472	6.8736	0.9744	1.8001
Proposed	8 *Case*1	0.008862	0.9964	**42.4135**	0.9896	2.1586	**0.9967**	**0.5785**
	8 *Case*2	**0.010706**	**0.9947**	**41.0899**	**0.9878**	**2.6389**	**0.9954**	**0.6756**
	8 *Case*3	**0.013275**	**0.9929**	**38.6708**	**0.9861**	**3.4728**	**0.9940**	**0.9086**

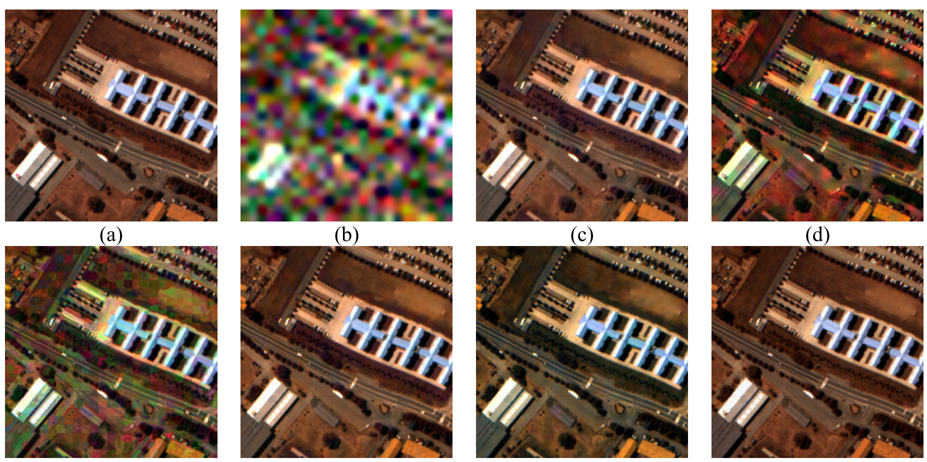


Fig. 5. Results for the data set A in case 2 with $s = 8$. (a) Original HRHS image with bands 59, 30, and 13 (false color). (b) Simulated LRHS image. (c) HySure. (d) CNMF. (e) NSSR. (f) BSR. (g) SupResPALM. (h) Proposed method.

In case 3, the intensity of Gaussian noise was the same as that in case 1. In addition, all bands were also corrupted by the stripes, with 30% of the total row or column number.

In the simulated experiments, the regularization coefficient $[\lambda_1, \lambda_2]$ used in the proposed method was set as [0.0001, 0.0001], [0.0001, 0.0001], and [0.0005, 0.0001], respectively, for case 1 to case 3. To better separate the noise or artifacts part, the ranks $r_i (i = 1, 2, 3)$ of tensor decomposition were selected as [200, 200, 15], [200, 200, 5], and [195, 195, 5] corresponding to different $[\lambda_1, \lambda_2]$ from case 1 to case 3. In addition, the regularization parameter β was empirically set as 0.5 for the low Gaussian noise case and 1 for the high Gaussian noise case, while the parameters γ were empirically set as 1e−4.

To achieve an integrated comparison of the other methods and the proposed method, seven quantitative evaluations, a visual comparison, curves of the spectra, and the overall difference results were used to analyze the spatial and spectral effectiveness of the results of different methods. In addition, the mean cross-stripe profiles were used to reflect the destriping ability. The contrasting evaluation indices for the three cases with various spatial resolution ratios s are listed in Tables I and II, respectively. The best performance for each quality index is marked in bold, and the second-best performance for each quality index is underlined.

From the numerical assessments of all the fused images listed in Tables I and II, the values of the RMSE, UIQI, and PSNR indices show that the proposed method generates the

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

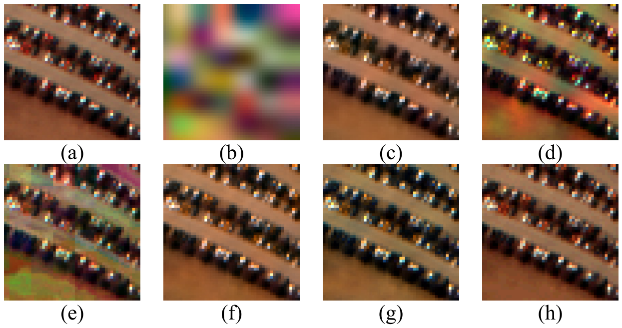


Fig. 6. Magnified results for the data set A in Fig. 5. (a) Original HRHS image with bands 59, 30, and 13 (false color). (b) Simulated LRHS image. (c) HySure. (d) CNMF. (e) NSSR. (f) BSR. (g) SupResPALM. (h) Proposed method.

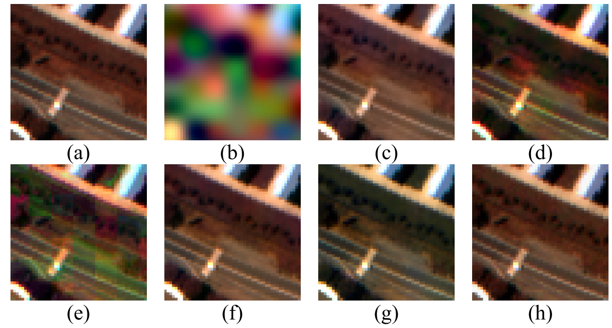


Fig. 7. Magnified results for the data set A in Fig. 5. (a) Original HRHS image with bands 59, 30, and 13 (false color). (b) Simulated LRHS image. (c) HySure. (d) CNMF. (e) NSSR. (f) BSR. (g) SupResPALM. (h) Proposed method.

lowest radiometric distortion in most situations. The SSIM index evaluating the SSIM indicates that the proposed method can capture more details than the other methods. Moreover, in most cases, the proposed method has the lowest SAM and ERGAS values, meaning that the proposed method can preserve more spectral information. The proposed method also achieves the highest CC value, indicating closer similarity with the reference image. Notably, the proposed method shows a stable ability with different noise levels and shows significant superiority in the high-intensity noise and mixed-noise cases. In the low-intensity noise situation, such as case 1, the proposed method obtains better or similar results to BSR and SupResPALM, which also show a good quantitative performance.

Compared with the other algorithms, the proposed method achieves a better visual quality in Figs. 5–7 and 9–11. To verify the fusion effect of the proposed method when confronted with the Gaussian noise, bands 59, 30, and 13 from the HSIs of the compared methods are selected to form a false-color composite for visual comparison in Figs. 5–7. It can be observed that all the fusion methods provide clear and sharp spatial details, compared with the LRHS images. However, there are still some spatial and spectral differences between the reference image and the fusion results of different methods. For example, the results of CNMF and NSSR contain obvious residual noise and spectral distortion. Although fake artifacts are not found in the results of SupResPALM, some areas, especially trees, clearly show spectral distortion. The fusion results of HySure and BSR show good noise suppression and enhancement of the spatial details. However, from the magnified areas in Figs. 6 and 7, some spectral distortion near the trees and streets is obvious. The proposed method obtains the closest results to the original HRHS image by exploring the low-multilinear-rank and the spatio-spectral correlation between bands. Furthermore, the spectral characteristics are better preserved and good spatial structures can be observed.

In order to further compare the performance of the proposed method in spectral preservation, the spectral curves from different objects are plotted in Fig. 8, which shows the spectral signatures of the pixel (101, 100) from data set A in case 2. From Fig. 8, we can see violent fluctuations in the curves of the simulated LRHS image for CNMF and NSSR, which indicate that these two methods fail to remove the spectral noise. Furthermore, HySure, BSR, and SupResPALM cannot approximate the reference spectral curve at the end of the bands very well. For the proposed method, a smoother curve possessing better consistency with the reference curve can be seen in Fig. 8, indicating good recovery of the spectral information.

To further validate the effect of the proposed method when confronted with the mixed noise, Figs. 9–11 show different visual results of different methods with a false-color composite of the bands 84, 48, and 28 from the HSIs. As displayed in Figs. 9–11, although all the results show the larger enhancement of spatial details, the five compared methods retain different degrees of residual stripes, and some of the methods introduce some heavy spectral distortions. The detail region in Fig. 10 reveals that the proposed method has a much better ability to preserve the healthy and sharp information in addition to suppressing the stripe noise successfully.

Using Fig. 12 to further describe the spectral preservation, these results are in good agreement with those visual results in Figs. 9–11. Our method can give a relatively smoother curve than the other methods and maintain a similar wave shape to the reference curve. Due to the less minor uneven on the curve in Fig. 12(h), we give the difference results between all methods and the original HRHS image in Fig. 13 to elaborate the consistency of the spatio-spectral information. The largest minimum, the smallest maximum, and the smallest box region from the 25th to the 75th and the 1st to the 99th percentile reveal that the overall quality of the fusion result of our method is the closest to that of the reference image. Accordingly, it is proved that the proposed fusion algorithm greatly improves the fusion quality under the influence of mixed noise.

In addition, in order to test the abilities of different methods to keep healthy pixels in the fusion process along with destriping, Fig. 14 displays the mean cross-track profiles of band 28 in Case 3. The best mean cross-track profile of the destriped image should be the same as the original reference image. From Fig. 14, we can see that the corresponding profiles of these methods, except for the proposed method, differ a lot from the curve of the reference image. However, the proposed method alleviates the fluctuation and brings the profile into correspondence with the reference image. Especially, the left

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

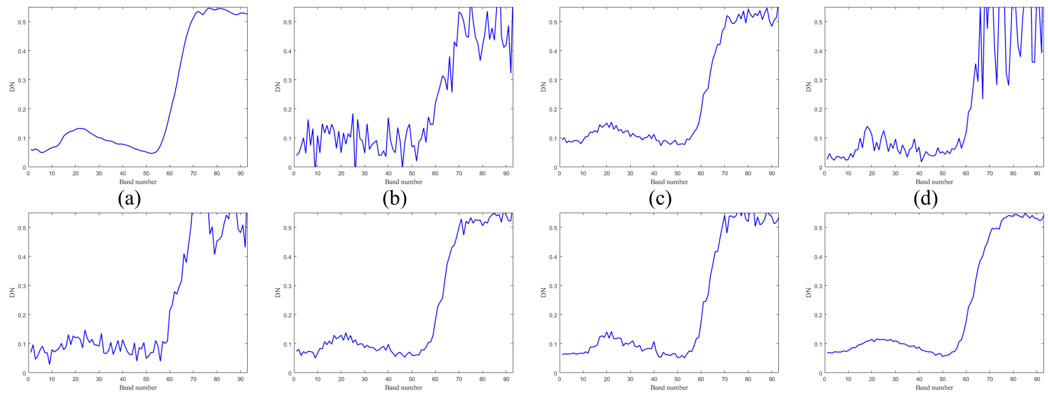


Fig. 8. Spectra of pixel (101, 100) in the fused results with $s = 8$ in case 2. (a) Original HRHS image. (b) Simulated LRHS image. (c) HySure. (d) CNMF. (e) NSSR. (f) BSR. (g) SupResPALM. (h) Proposed method.

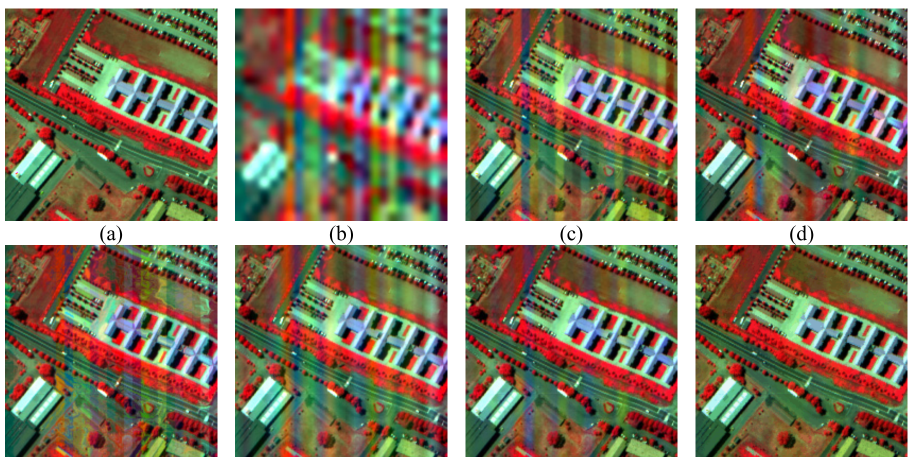


Fig. 9. Results for the data set A in case 3 with $s = 8$. (a) Original HRHS image with bands 84, 48, and 28 (false color). (b) Simulated LRHS image. (c) HySure. (d) CNMF. (e) NSSR. (f) BSR. (g) SupResPALM. (h) Proposed method.

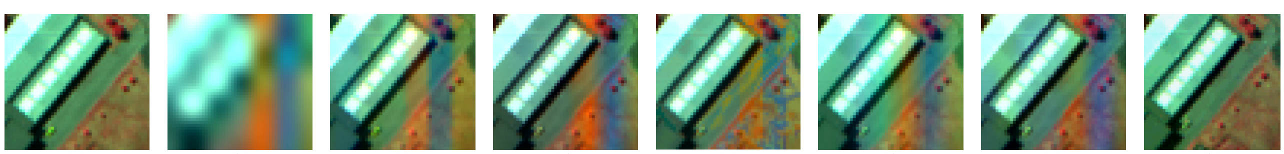


Fig. 10. Magnified results for the data set A in Fig. 9. (a) Original HRHS image with bands 84, 48, and 28 (false color). (b) Simulated LRHS image. (c) HySure. (d) CNMF. (e) NSSR. (f) BSR. (g) SupResPALM. (h) Proposed method.

part of the result, which is a nonstripe region, is the same as the reference image. This confirms that the proposed method is a reliable way to suppress the stripes while preserving healthy detailed information.

B. Real-Data Experiments

In the data set B, most of the bands from the Hyperion were seriously degraded by Gaussian noise, stripes, or the mixed noise. To overcome the influence of the various noises, the parameters κ_1, κ_2, and β were set as 0.005, 0.001, and 1, respectively. In addition, based on our experience in the simulated experiments, the small ranks on the spatial dimensions can be empirically selected as 97.5% of the dimension. Thus, the ranks $r_i (i = 1, 2, 3)$ were set as [390, 350, 5]. To further verify the effectiveness of the proposed method, a visual comparison with the false-color results with the combined bands 38, 19, and 2 bands is given in Figs. 15–17.

In Figs. 15–17, it can be seen that some spectral information is lost in the result of SupResPALM, particularly in the building areas in top-left and top-right corners. The result coincides with the performance of SupResPALM in the simulated case 2. SupResPALM has difficulty solving the problem of heavy Gaussian noise, and the noise results in wrong information, as shown in the top-right corner of Fig. 16(e). In addition, Fig. 16 shows that HySure, CNMF, NSSR, and BSR also

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

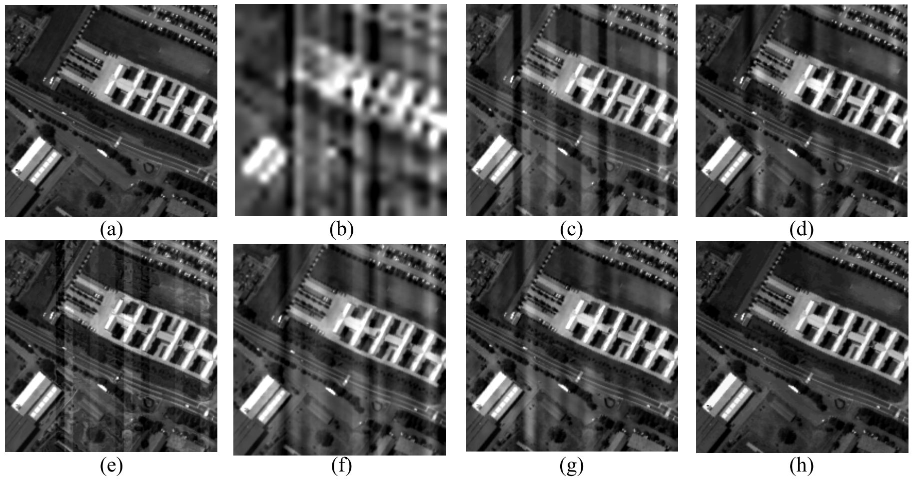


Fig. 11. Results for the data set A in case 3 with $S = 8$. (a) Original HRHS image with band 28. (b) Simulated LRHS image. (c) HySure. (d) CNMF. (e) NSSR. (f) BSR. (g) SupResPALM. (h) Proposed method.

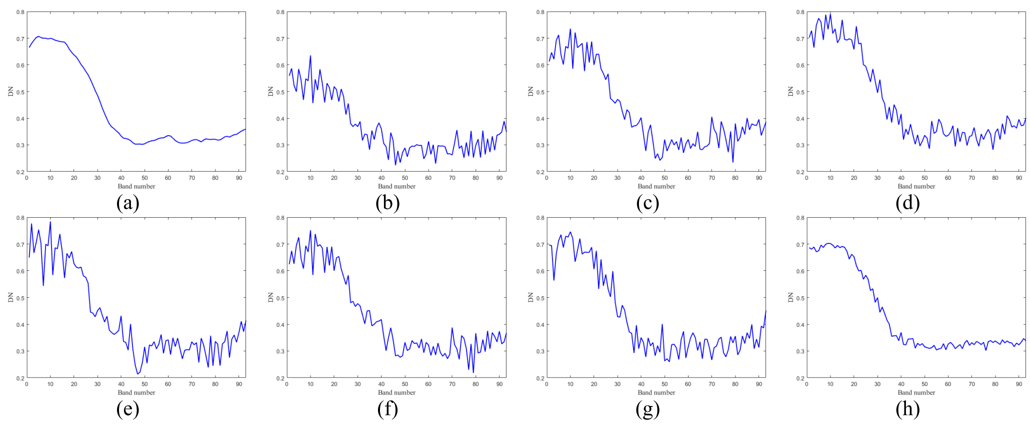


Fig. 12. Spectra of pixel (105, 70) in the fused results for the data set A in case 3 with $S = 8$. (a) Original HRHS image with bands. (b) Simulated LRHS image. (c) HySure. (d) CNMF. (e) NSSR. (f) BSR. (g) SupResPALM. (h) Proposed method.

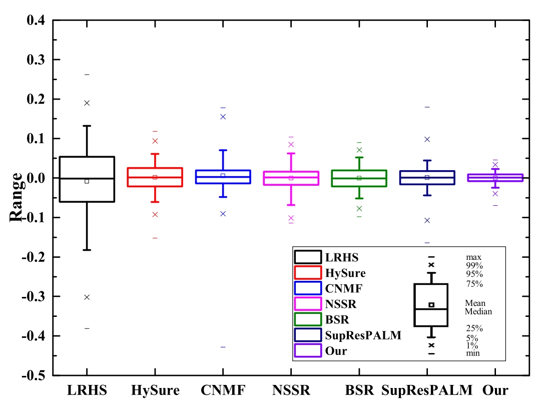


Fig. 13. Box plots of the difference between the compared methods and the original HRHS image for the data set A in case 3. Each box plot represents the difference between a compared method and the original HRHS image.

fail to suppress the spectral distortion induced by the noise of the LRHS image. For example, the color of the building areas shows an abnormal red, due to the impact of the mixed noise. What is worse, these compared methods show obvious residual striping and result in nonsmoothness of the spatial information. Unlike the five existing methods, the fused image produced by the proposed method shows a more natural and sharp visual effect, by effectively removing the mixed noise and fusing the details and structural information. Moreover, Fig. 18 displays the mean cross-track profiles of band 1 in the data set B, where the corresponding profiles of the five existing methods still show too much fluctuation. In contrast, the proposed approach can process stripes well, with the least amount of distortion, and is better able to preserve the fused detailed information.

C. Discussion

1) One-Step Versus Step-by-Step: To further validate the effect of the presented method, we compared a group of fusion approaches with two steps, where the stripes or mixed noise were first removed using a denoising method and then image fusion is performed. These experiments are conducted in two cases. One is the same as Case 3, and ASSTV [16], as one representative method in the multichannel image destriping, is employed to remove the stripes. The other one is set to

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

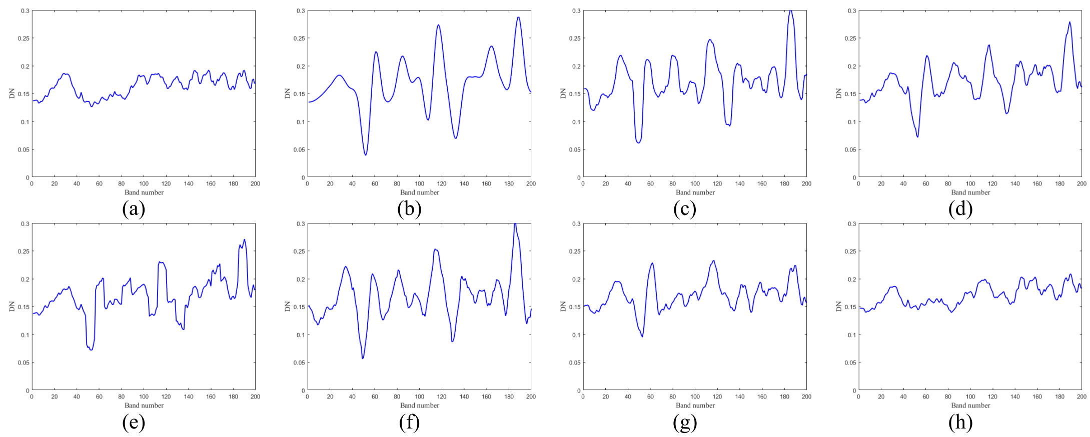


Fig. 14. Mean cross-stripe profiles of band 28 in case 3. (a) Original HRHS image with band 28. (b) Simulated LRHS image. (c) HySure. (d) CNMF. (e) NSSR. (f) BSR. (g) SupResPALM. (h) Proposed method.

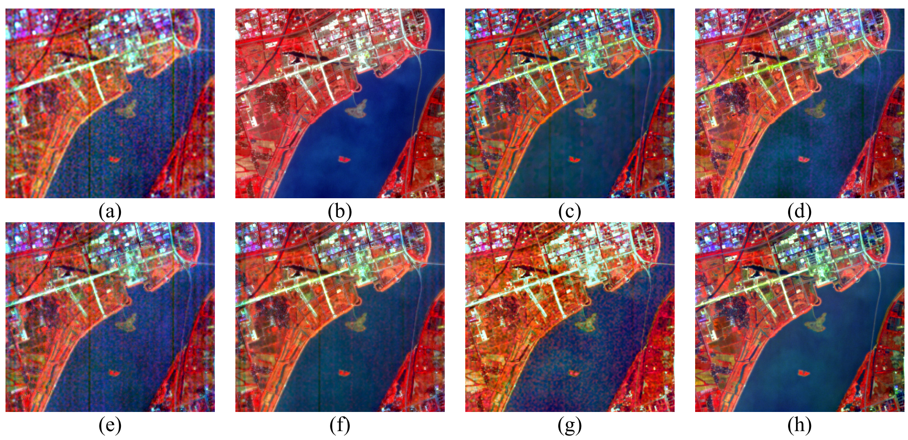


Fig. 15. Results for the data set B. (a) Observed LRHS image with bands 38, 19, and 2 (false color). (b) Observed HRMS image with bands 4, 3, and 2 (false color). (c) HySure. (d) CNMF. (e) NSSR. (f) BSR. (g) SupResPALM. (h) Proposed method.

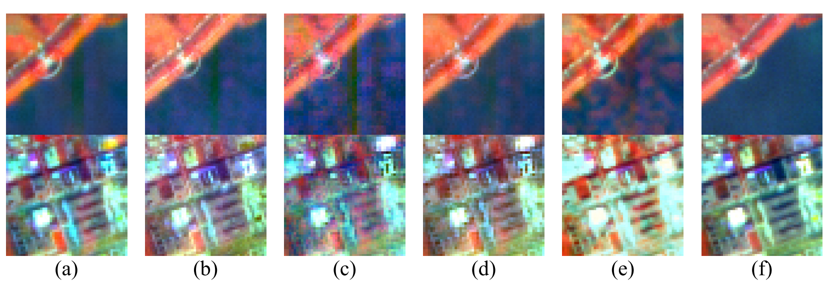


Fig. 16. Magnified results for the data set B in Fig. 15. (a) HySure. (b) CNMF. (c) NSSR. (d) BSR. (e) SupResPALM. (f) Proposed method.

be contaminated by the heavier noise with 10-dB SNR and contain 50% stripes for each band. Due to the high-intensity Gaussian noise in this case, we adopted LRMR [6] which is one of classical and effective methods for the mixed-noises reduction. The best denoising results from ASSTV and LRMR are obtained as the initial values for fusion.

In these experiments, two fusion methods with better performance on quantitative evaluation of the previous simulated experiments are selected to implement the step by step strategy, denoted as "BSR(s)" and "SupResPALM(s)." Contrarily, our method adopts one algorithm to simultaneously achieve the denoising and fusion, and is denoted by "Proposed(o)." The fusion results of the ROSIS image with $s = 4$ are listed in Table III with different methods. All evaluation indices in Table III are almost the best for the proposed method. Furthermore, our method gives a significant superiority with the heavier noise. Nevertheless, the compared methods may lack sufficient capacity in dealing with the denoising results with artifacts or oversmoothing under high-intensity noise condition. The all above results from Tables I–III demonstrate that the proposed method not only can achieve good performance in high SNRs condition but also can better restrain the interference of noise.

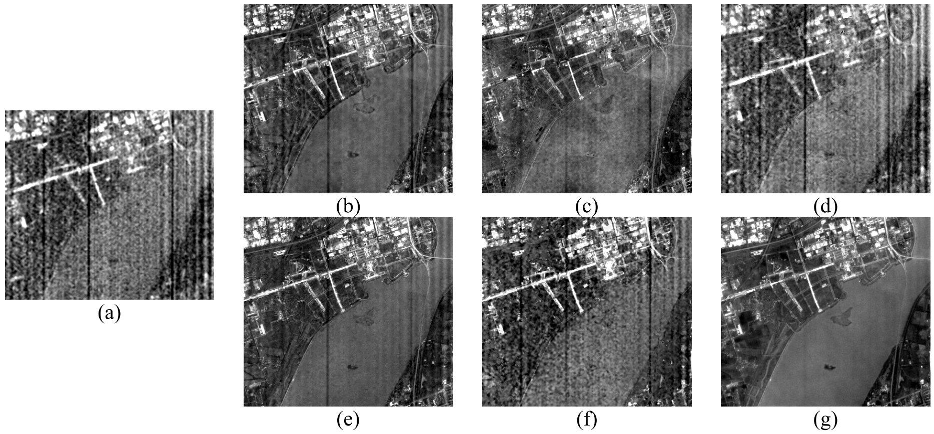


Fig. 17. Results for the data set B with band 1. (a) Observed LRHS image. (b) HySure. (c) CNMF. (d) NSSR. (e) BSR. (f) SupResPALM. (g) Proposed method.

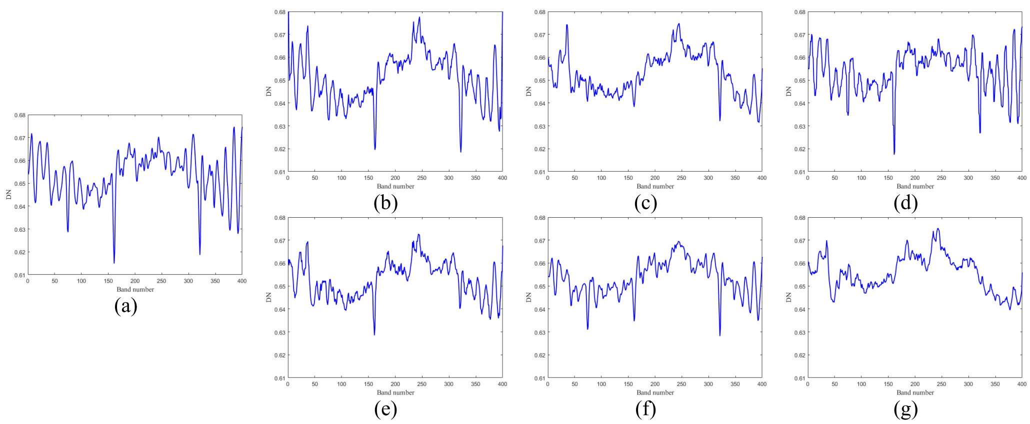


Fig. 18. Mean cross-stripe profiles of band 1 in the data set B. (a) Observed LRHS image. (b) HySure. (c) CNMF. (d) NSSR. (e) BSR. (f) SupResPALM. (g) Proposed method.

TABLE III
QUANTITATIVE EVALUATION OF THE FUSION RESULTS FOR THE SIMULATED EXPERIMENTS WITH $s = 4$

	SNR = 30dB and 30% stripes			SNR = 10dB and 50% stripes		
	BSR(s)	SupResPALM(s)	Proposed (o)	BSR(s)	SupResPALM(s)	Proposed (o)
RMSE	0.01049	0.01154	**0.01062**	0.02126	0.02147	**0.01193**
UIQI	**0.9955**	0.9945	0.9951	0.9817	0.9809	**0.9943**
PSNR	40.3407	39.1969	**40.6601**	33.6851	33.7116	**39.4362**
SSIM	0.9882	0.9849	**0.9894**	0.9592	0.9657	**0.9881**
SAM	2.7082	2.8513	**2.6541**	5.3435	5.2611	**3.0002**
CC	**0.9959**	0.9951	0.9956	0.9834	0.9836	**0.9952**
ERGAS	1.4010	1.5666	**1.3521**	2.9119	3.0784	**1.5711**

TABLE IV
RUN-TIME COMPARISON FOR THE SIMULATED EXPERIMENTS WITH $s = 4$ (SECONDS)

	HySure	CNMF	NSSR	BSR	SupRes PALM	Proposed
Time	52.810	33.159	162.128	516.857	58.159	489.112

2) Run-Time Comparison: To compare the work efficiency of different fusion algorithms, the average running times were recorded for three simulated experiments with $s = 4$ under the same operational environment (Software: Windows 10, MATLAB R2015b; Hardware: 8-GB RAM, i7-6500 CPU), as listed in Table IV. Although BSR and the proposed methods spend a relatively more time to fuse HSI and MS images, they can obtain better performance in Tables I–III and limit time cost within an acceptable range. Hence, to improve the efficiency of the proposed method, the admixture programming with C language, which applied in SupResPALM, can be employed in the future. In addition, the program structure can also be optimized.

V. CONCLUSION

In this paper, we have proposed an antinoise HSI fusion method by combining tensor decomposition and SGV regularization. The tensor decomposition is used to satisfy low multilinear ranks for characterizing the underlying clean HSI part by the subtraction of the structures of the heavy noise and artifacts parts. In order to well remove the stripes in the fusion process, the spectral relationship between bands is explored to form a spectral guidance-based gradient variational constraint. In one respect, the intrinsic gradient complementarity induced by the nonuniform spectral response coefficient is exploited to enhance the details and preserve the spectral property. On the other hand, the differentials of the two directions for the sparse residual image help to suppress the mixed noise. The experiments indicated that the proposed method outperforms the mainstream methods, and the proposed fusion method can produce finer spatial details, while overcoming complex noise, and is better able to preserve the spectral characteristics.

Although the proposed method works well in the HSI fusion, especially in a noisy situation, there are still some limitations for the recovery of a spectral curve when confronted with stripes. Thus, in the future, we will focus on the task of improving the spectral characteristics for HSIs with stripes. The acceleration for fusion should also be considered. Furthermore, for fusion at higher spatial ratios, multisource data and heterogeneous data will be considered in our future work.

REFERENCES

[1] D. Landgrebe, "Hyperspectral image data analysis," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 17–28, Jan. 2002.

[2] A. Mookambiga and V. Gomathi, "Comprehensive review on fusion techniques for spatial information enhancement in hyperspectral imagery," *Multidimensional Syst. Signal Process.*, vol. 27, no. 4, pp. 863–889, Oct. 2016.

[3] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. M. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.

[4] N. Yokoya, C. Grohnfeldt, and J. Chanussot, "Hyperspectral and multispectral data fusion: A comparative review of the recent literature," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 2, pp. 29–56, Jun. 2017.

[5] A. Plaza *et al.*, "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, vol. 113, pp. S110–S122, Sep. 2009.

[6] H. Zhang, W. He, L. Zhang, H. Shen, and Q. Yuan, "Hyperspectral image restoration using low-rank matrix recovery," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 4729–4743, Aug. 2014.

[7] J. Chen, Y. Shao, H. Guo, W. Wang, and B. Zhu, "Destriping CMODIS data by power filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 2119–2124, Sep. 2003.

[8] R. Pande-Chhetri and A. Abd-Elrahman, "De-striping hyperspectral imagery using wavelet transform and adaptive frequency domain filtering," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 5, pp. 620–636, Sep. 2011.

[9] H. Shen, W. Jiang, H. Zhang, and L. Zhang, "A piece-wise approach to removing the nonlinear and irregular stripes in MODIS data," *Int. J. Remote Sens.*, vol. 35, no. 1, pp. 44–53, Nov. 2013.

[10] M. Bouali and S. Ladjal, "Toward optimal destriping of MODIS data using a unidirectional variational model," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 8, pp. 2924–2935, Aug. 2011.

[11] X. Liu, X. Lu, H. Shen, Q. Yuan, Y. Jiao, and L. Zhang, "Stripe noise separation and removal in remote sensing images by consideration of the global sparsity and local variational properties," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 5, pp. 3049–3060, May 2016.

[12] X. Liu, H. Shen, Q. Yuan, X. Lu, and C. Zhou, "A universal destriping framework combining 1-D and 2-D variational optimization methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 808–822, Feb. 2018.

[13] L. Sun, R. Neville, K. Staenz, and H. P. White, "Automatic destriping of Hyperion imagery based on spectral moment matching," *Can. J. Remote Sens.*, vol. 34, no. S1, pp. S68–S81, Jan. 2008.

[14] N. Acito, M. Diani, and G. Corsini, "Subspace-based striping noise reduction in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 4, pp. 1325–1342, Apr. 2011.

[15] Y. Yuan, M. Fu, and X. Lu, "Substance dependence constrained sparse NMF for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 6, pp. 2975–2986, Jun. 2015.

[16] Y. Chang, L. Yan, H. Fang, and C. Luo, "Anisotropic spectral-spatial total variation model for multispectral remote sensing image destriping," *IEEE Trans. Image Process.*, vol. 24, no. 6, pp. 1852–1866, Jun. 2015.

[17] L. Loncan *et al.*, "Hyperspectral pansharpening: A review," *IEEE Trans. Geosci. Remote Sens.*, vol. 3, no. 3, pp. 27–46, Sep. 2015.

[18] X. Meng, H. Shen, H. Li, L. Zhang, and R. Fu, "Review of the pansharpening methods for remote sensing images based on the idea of meta-analysis: Practical discussion and challenges," *Inf. Fusion*, vol. 46, pp. 102–113, Jun. 2018.

[19] Z. Chen, H. Pu, B. Wang, and G.-M. Jiang, "Fusion of hyperspectral and multispectral images: A novel framework based on generalization of pan-sharpening methods," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 8, pp. 1418–1422, Aug. 2014.

[20] C. Grohnfeldt, X. X. Zhu, and R. Bamler, "Splitting the hyperspectral-multispectral image fusion problem into weighted pan-sharpening problems: The spectral grouping concept," in *Proc. 7th Workshop Hyperspectral Image Signal Process., Evol. Remote Sens.*, Jun. 2015, pp. 1–4.

[21] P. S. Chavez, Jr., S. C. Sides, and J. A. Anderson, "Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic," *Photogramm. Eng. Remote Sens.*, vol. 57, no. 3, pp. 295–303, 1991.

[22] B. Aiazzi, S. Baronti, and M. Selva, "Improving component substitution pansharpening through multivariate regression of MS+Pan data," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3230–3239, Oct. 2007.

[23] C. Thomas, T. Ranchin, L. Wald, and J. Chanussot, "Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1301–1312, May 2008.

[24] Q. Yuan, Y. Wei, X. Meng, H. Shen, and L. Zhang, "A multiscale and multidepth convolutional neural network for remote sensing imagery pan-sharpening," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 978–989, Mar. 2018.

[25] R. B. Gomez, A. Jazaeri, and M. Kafatos, "Wavelet-based hyperspectral and multispectral image fusion," *Proc. SPIE*, vol. 4383, pp. 36–42, Jun. 2001.

[26] J. G. Liu, "Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details," *Int. J. Remote Sens.*, vol. 21, no. 18, pp. 3461–3472, Nov. 2000.

[27] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "MTF-tailored multiscale fusion of high-resolution MS and PAN imagery," *Photogramm. Eng. Remote Sens.*, vol. 72, no. 5, pp. 591–596, May 2006.

[28] M. Selva, B. Aiazzi, F. Butera, L. Chiarantini, and S. Baronti, "Hyper-sharpening: A first approach on SIM-GA data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 3008–3024, Jun. 2015.

[29] Y. Wei, Q. Yuan, H. Shen, and L. Zhang, "Boosting the accuracy of multispectral image pansharpening by learning a deep residual network," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 10, pp. 1795–1799, Oct. 2017.

[30] R. C. Hardie, M. T. Eismann, and G. L. Wilson, "MAP estimation for hyperspectral image resolution enhancement using an auxiliary sensor," *IEEE Trans. Image Process.*, vol. 13, no. 9, pp. 1174–1184, Sep. 2004.

[31] X. Bresson and T. F. Chan, "Fast dual minimization of the vectorial total variation norm and applications to color image processing," *Inverse Problems Imag.*, vol. 2, no. 4, pp. 455–484, 2008.

[32] Q. Wei, J. Bioucas-Dias, N. Dobigeon, and J. Y. Tourneret, "Hyperspectral and multispectral image fusion based on a sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3658–3668, Jul. 2015.

[33] K. Rong, L. Jiao, S. Wang, and F. Liu, "Pansharpening based on low-rank and sparse decomposition," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 12, pp. 4793–4805, Dec. 2014.

[34] K. Zhang, M. Wang, and S. Yang, "Multispectral and hyperspectral image fusion based on group spectral embedding and low-rank factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 3, pp. 1363–1371, Mar. 2017.

[35] M. T. Eismann and R. C. Hardie, "Application of the stochastic mixing model to hyperspectral resolution enhancement," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 9, pp. 1924–1933, Sep. 2004.

[36] M. T. Eismann, "Resolution enhancement of hyperspectral imagery using coincident panchromatic imagery and a stochastic mixing model," Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. Dayton, OH, USA, May 2004.

[37] Y. Zhang, S. De Backer, and P. Scheunders, "Noise-resistant wavelet-based Bayesian fusion of multispectral and hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3834–3843, Nov. 2009.

[38] M. Simões, J. B. Dias, L. Almeida, and J. Chanussot, "A convex formulation for hyperspectral image superresolution via subspace subspace-based regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 6, pp. 3373–3388, Jun. 2015.

[39] Y. Zhao, J. Yang, and J. Chan, "Hyperspectral imagery super-resolution by spatial-spectral joint nonlocal similarity," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2671–2679, Dec. 2014.

[40] C. Yi, Y.-Q. Zhao, and J. C.-W. Chan, "Hyperspectral image super-resolution based on spatial and spectral correlation fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 7, pp. 4165–4177, Jul. 2018.

[41] R. Kawakami, Y. Matsushita, J. Wright, M. Ben-Ezra, Y.-W. Tai, and K. Ikeuchi, "High-resolution hyperspectral imaging via matrix factorization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 2329–2336.

[42] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 528–537, Feb. 2012.

[43] N. Yokoya and A. Iwasaki, "Hyperspectral and multispectral data fusion mission on hyperspectral imager suite (HISUI)," in *Proc. IEEE Int. Symp. Geosci. Remote Sens. (IGARSS)*, Melbourne, VIC, Australia, Jul. 2013, pp. 4086–4089.

[44] N. Akhtar, F. Shafait, and A. Mian, "Sparse spatio-spectral representation for hyperspectral image super-resolution," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 63–78.

[45] B. Huang, H. Song, H. Cui, J. Peng, and Z. Xu, "Spatial and spectral image fusion using sparse matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 3, pp. 1693–1704, Mar. 2014.

[46] W. Dong *et al.*, "Hyperspectral image super-resolution via non-negative structured sparse representation," *IEEE Trans. Image Process.*, vol. 25, no. 5, pp. 2337–2352, May 2016.

[47] C. Lanaras, E. Baltsavias, and K. Schindler, "Hyperspectral super-resolution by coupled spectral unmixing," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 3586–3594.

[48] N. Akhtar, F. Shafait, and A. Mian, "Bayesian sparse representation for hyperspectral image super resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 3631–3640.

[49] B. Lin, X. Tao, M. Xu, L. Dong, and J. Lu, "Bayesian hyperspectral and multispectral image fusions via double matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 10, pp. 5666–5678, Oct. 2017.

[50] S. Li, R. Dian, L. Fang, and J. M. Bioucas-Dias, "Fusing hyperspectral and multispectral images via coupled sparse tensor factorization," *IEEE Trans. Image Process.*, vol. 27, no. 8, pp. 4118–4130, Aug. 2018.

[51] Q. Xie *et al.*, "Multispectral images denoising by intrinsic tensor sparsity regularization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 1692–1700.

[52] R. Dian, L. Fang, and S. Li, "Hyperspectral image super-resolution via non-local sparse tensor factorization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 3862–3871.

[53] Y. Wang, J. Peng, Q. Zhao, D. Meng, Y. Leung, and X.-L. Zhao, "Hyperspectral image restoration via total variation regularized low-rank tensor decomposition," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 4, pp. 1227–1243, Apr. 2018.

[54] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method and multipliers," *Found. Trends Mach. Learn.*, vol. 3, pp. 1–122, Jan. 2011.

[55] Y. Chang, L. Yan, H. Fang, S. Zhong, and Z. Zhang, "Weighted low-rank tensor recovery for hyperspectral image restoration," 2017, *arXiv:1709.00192*. [Online]. Available: https://arxiv.org/abs/1709.00192

[56] P. Ghamisi *et al.*, "Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 37–78, Dec. 2017.

[57] R. Feng, Y. Zhong, L. Wang, and W. Lin, "Rolling guidance based scale-aware spatial sparse unmixing for hyperspectral remote sensing imagery," *Remote Sens.*, vol. 9, no. 12, p. 1218, 2017.

[58] A. Buades, B. Coll, and J. M. Morel, "A non-local algorithm for image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 60–65.

[59] Q. Wang, W. Shi, P. M. Atkinson, and Y. Zhao, "Downscaling MODIS images with area-to-point regression Kriging," *Remote Sens. Environ.*, vol. 166, pp. 191–204, Sep. 2015.

[60] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.

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

[62] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[63] R. H. Yuhas, A. F. H. Goetz, and J. W. Boardman, "Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm," in *Proc. Summaries 4th JPL Airborne Earth Sci. Workshop*, 1992, pp. 147–149.

[64] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.

[65] L. Wald, "Quality of high resolution synthesised images: Is there a simple criterion?" in *Proc. Int. Conf. Fusion Earth Data*, Nice, France, Jan. 2000, pp. 99–103.

Jie Li (M'16) received the B.S. degree in sciences and techniques of remote sensing and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2011 and 2016, respectively.

He is currently a Lecturer with the School of Geodesy and Geomatics, Wuhan University. His research interests include the image quality improvement, super-resolution, data fusion, remote sensing image processing, sparse representation, and deep learning.

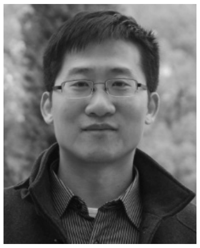
Xinxin Liu received the B.S. degree in geographic information system and the Ph.D. degree in cartography and geographic information system from Wuhan University, Wuhan, China, in 2013 and 2018, respectively.

In 2018, she joined the College of Electrical and Information Engineering, Hunan University, Changsha, China, where she is currently an Assistant Professor. Her research interests include the image quality improvement, remote sensing image processing, and remote sensing mapping and application.

Qiangqiang Yuan (M'13) received the B.S. degree in surveying and mapping engineering and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2006 and 2012, respectively.

In 2012, he joined the School of Geodesy and Geomatics, Wuhan University, where he is currently a Professor. He has authored or coauthored more than 50 research papers, including more than 30 peer-reviewed articles in international journals such as the IEEE TRANSACTIONS ON IMAGE PROCESSING and the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. His research interests include image reconstruction, remote sensing image processing and application, and data fusion.

Dr. Yuan was a recipient of the Top-Ten Academic Star of Wuhan University in 2011 and the Hong Kong Scholar Award from the Society of Hong Kong Scholars and the China National Postdoctoral Council, in 2014. He has frequently served as a referee for more than 20 international journals for remote sensing and image processing.



Huanfeng Shen (M'10–SM'13) received the B.S. degree in surveying and mapping engineering and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2002 and 2007, respectively.

In 2007, he joined the School of Resource and Environmental Sciences, Wuhan University, where he is currently a Luojia Distinguished Professor. He has been supported by several talent programs, such as the Youth Talent Support Program of China in 2015, the China National Science Fund for Excellent Young Scholars in 2014, and the New Century Excellent Talents by the Ministry of Education of China in 2011. He has authored over 100 research papers. His research interests include the image quality improvement, remote sensing mapping and application, data fusion and assimilation, and regional and global environmental changes.

Dr. Shen is currently a member of the Editorial Board of the *Journal of Applied Remote Sensing*.

Liangpei Zhang (M'06–SM'08–F'19) received the B.S. degree in physics from Hunan Normal University, Changsha, China, in 1982, the M.S. degree in optics from the Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, China, in 1988, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1998.

He was the Head of the Remote Sensing Division, State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing (LIESMARS), Wuhan University. He is a "Chang-Jiang Scholar" Chair Professor appointed by the Ministry of Education of China, and he is also a Principal Scientist for the China State Key Basic Research Project (2011–2016) appointed by the Ministry of National Science and Technology of China to lead the Remote Sensing Program in China. He has authored or coauthored more than 600 research papers and 7 books. He holds more than 30 patents. His research interests include hyperspectral remote sensing, high-resolution remote sensing, image processing, and artificial intelligence.

Dr. Zhang is a fellow of IET and an Executive Member (Board of Governor) of the China National Committee of International Geosphere–Biosphere Programme and the China Society of Image and Graphics. He was a recipient of the 2010 Best Paper Boeing Award, the 2013 Best Paper ERDAS Award from the American Society of Photogrammetry and Remote Sensing (ASPRS), and the best reviewer awards from the IEEE Geoscience and Remote Sensing Society (GRSS) for his service to the IEEE JOURNAL OF SELECTED TOPICS IN EARTH OBSERVATIONS AND APPLIED REMOTE SENSING (JSTARS) in 2012 and IEEE GEOSCIENCE AND REMOTE SENSING LETTERS (GRSL) in 2014. His research teams won the top prize of the IEEE GRSS 2018 Data Fusion Contest and top three prizes of the IEEE GRSS 2014 Data Fusion Contest, respectively, and his students have been selected as the winners or finalists of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS) Student Paper Contest in recent years. He is the Founding Chair of the IEEE GRSS Wuhan Chapter. He was the General Chair of the 4th IEEE GRSS Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS) and the Guest Editor of JSTARS. He regularly serves as a Co-Chair of the series SPIE conferences on multispectral image processing and pattern recognition, conference on Asia remote sensing, and many other conferences. He edits several conference proceedings, issues, and geoinformatics symposiums. He also serves as an Associate Editor for the *International Journal of Ambient Computing and Intelligence*, the *International Journal of Image and Graphics*, the *International Journal of Digital Multimedia Broadcasting*, the *Journal of Geo-Spatial Information Science*, and the *Journal of Remote Sensing,* and the Guest Editor of the *Journal of Applied Remote Sensing* and the *Journal of Sensors.* He is currently serving as an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.