Radiometric quality improvement of hyperspectral remote sensing images: a technical tutorial on variational framework

Jie Li,[a] Huanfeng Shen[b,c,]* Huifang Li,[b] Menghui Jiang,[b] and Qiangqiang Yuan[a,c]

[a]Wuhan University, School of Geodesy and Geomatics, Wuhan, China
[b]Wuhan University, School of Resource and Environmental Science, Wuhan, China
[c]Wuhan University, Collaborative Innovation Center of Geospatial Technology, Wuhan, China

Abstract. In hyperspectral remote sensing imagery, the sensor, atmosphere, topography, and other factors often bring about some degradations, such as noise, haze, clouding, and shadowing. Due to inevitable tradeoff between spatial resolution and spectral resolution, low spatial details of hyperspectral images (HSIs) also limit the range of potential applications. Compensating for these degradations through quality improvement is a key preprocessing step in the exploitation of HSIs. A comprehensive analysis of the quality improvement techniques for HSIs is presented. The closely connected techniques, such as denoising, destriping, dehazing, cloud removal, and super-resolution, are linked as a whole by a general reconstruction model in a variational framework. Furthermore, we classify the methods into four categories according to their processing strategies for HSIs, including single-channel prior-based model, cross-channel prior-based model, tensor-based model, and data-driven prior-based model. Then, for several specific tasks, we briefly introduce their architectures of quality improvement, which combine different models and available complementary information from other spectral bands and/or temporal/sensor images. Some experimental results in different tasks are presented to show the effect of variational framework and draw some meaningful conclusions. Finally, some advantages on variational framework are discussed, and several promising directions are provided to serve as guidelines for future work. © 2021 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JRS.15.031502]

1 Introduction

In the field of airborne and satellite remote sensing, hyperspectral imaging (HSI) has matured into one of the most powerful and promising technologies. The continuous spectral bands enable HSI to discriminate different materials on the ground more possibly. The research on HSI processing has been active in the past decades. HSIs have also been more widely used to monitor the earth surface for civilian and military purposes.

However, to obtain a high spectral resolution, the sensors should narrow the bandwidth, as shown in Fig. 1. This approach inevitably reduces the signal-to-noise ratio because less energy can be captured by the sensor. Therefore, in Figs. 2(a)–2(c), degradations, such as random noise, striping, and dead pixels, arise more frequently in HSIs than multispectral (MS) and panchromatic (PAN) images. These noises are produced by atmospheric effects and instrumental failure corruption in the spectral bands with varying degrees. For HSIs, denoising is an essential estimation task of radiometric quality improvement related to the observed image that is degraded by noise sources.

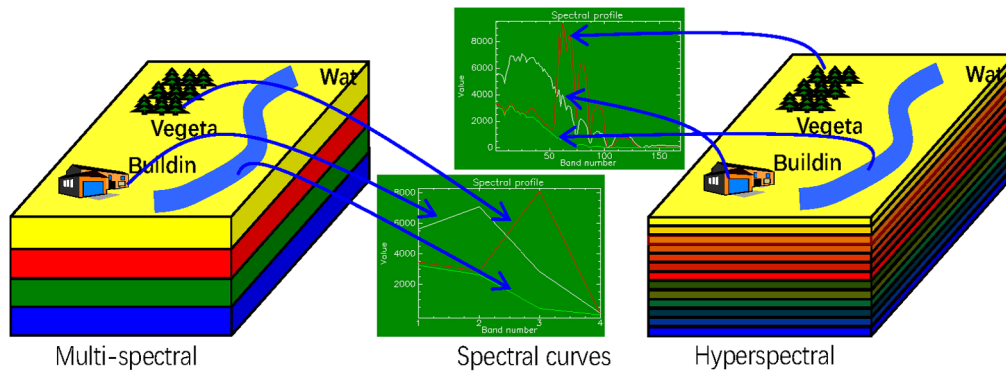*Address all correspondence to Huanfeng Shen, shenhf@whu.edu.cn

Fig. 1 Difference between HSI and MS image.

Moreover, the atmosphere, topography, and other factors often cause additional degradations, such as haze, shadow, and cloud coverage. In Fig. 2(d), haze and thin cloud mainly caused the local unevenness, which shows the increasing intensity due to the complicated atmospheric scattering. Furthermore, in Fig. 2(e), frequent thick clouds and shadow cover the real surface inevitably and causes bad influence on the visual appearance.

From a perspective of satellite remote sensing system limitation, a trade-off remains between the spectral and spatial resolutions, leading to the fact that the HSIs are often not as clear as desired, as shown in Fig. 2(f). The fusion of the complementary information among the multi-source remote sensing observations is a good way to improve the potential applications of remote sensing data.

All these common degradations in HSIs limit the precision of the subsequent processing, such as classification,[1,2] unmixing,[3,4] subpixel mapping,[5–7] and target detection.[8,9] Compensating these degradations through quality improvement is therefore a key preprocessing step in the exploitation of HSIs.[10] With different degradation problems, radiometric quality improvement

(a) Gaussian noise (b) Stripe noise (c) Dead pixels

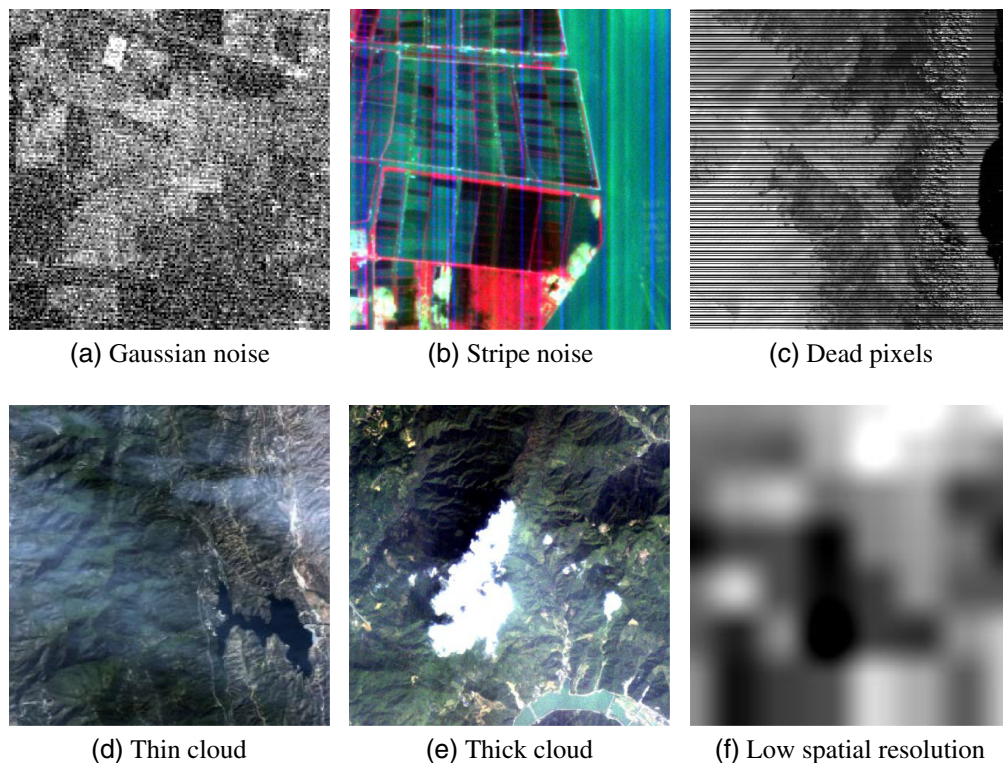(d) Thin cloud (e) Thick cloud (f) Low spatial resolution

Fig. 2 Different degradation problems.

methods for HSIs have been widely researched. These methods can be divided into two classes. The first class is based on physical models, such as the atmosphere correction and reflectance inversion using a radiation transfer equation. The second class is based on statistical image processing techniques, such as denoising, deblurring, destriping, inpainting, haze or cloud removal, super-resolution, and image fusion. To maintain the focus of the paper, we elaborate on the second class. Although different techniques are proposed for a particular degradation in HSIs, the processing techniques for different degradations are in fact closely linked, due to the similar degradation processing and utilization of the HS characteristics (e.g., spatial similarity and the spectral relevance). Thus, the technology barrier among different degradations must really be broken down, and the related techniques should be systematically summarized to benefit the quality improvement of HSIs and to inspire the development of new methods.

In this paper, a comprehensive analysis of the techniques used for the different degradations is presented. We link the closely connected techniques as a whole by providing a general reconstruction model in a variational framework. Most techniques that are based on a Bayesian framework and/or regularization method can be regarded as specific cases within this universal model. Two derived models are given to describe the methods based on spectral transform models, respective for one degraded HS input and multiple inputs with auxiliary complementary data. To embody the differences with the methods used for processing other types of images, the methods are classified according to their special processing strategies for HSIs. According to the access and utilization of prior information, the methods can be divided into single-channel prior-based model, cross-channel-prior based model, tensor-based model, and data-driven prior-based model. Through the application of these different models, researchers can eliminate the influence of denoising, dead pixels, haze, and thin cloud using only useful information from other spectral bands. However, more complementary must be extracted from other temporal/sensor images to solve large missing area and spatial detail loss better.

The remainder of this paper is organized as follows. Section 2 describes the general model for the quality improvement of HSIs in detail. Section 3 gives an introduction to about four categories based on the difference of prior model. Section 4 elaborates the specific applications in the field of HSI quality improvement based on available complementary information. At last, Sec. 5 expounds advantages of variational framework and future development. At last, some concluding remarks are presented in Sec. 6.

2 General Model for HSI

2.1 *Notation and Preliminaries*

Throughout the paper, we denote scalars, vectors, matrices, and tensors by nonbold letters, bold lowercase letters, bold uppercase letters, and calligraphic uppercase letters, respectively. Tensors[10] are multidimensional arrays of numbers that transform linearly under coordinate transformations, which can be represented as $\mathbf{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ with multilinear algebra[11] defined on them. Here, N is the order of the tensor, which is also known as way or mode, and the d'th order of the tensor is of size I_d. An arbitrary element of $\mathbf{X}$ is a scalar denoted by $\mathbf{X}_{i_1,i_2,\cdots,i_N}$, where indices typically range from 1 to the size of their mode, e.g., $1 \leq i_d \leq I_d$ and $1 \leq d \leq N$. The mode-d vectors of $\mathbf{X}$ are defined as the d-dimensional vectors obtained by varying the index i_d while keeping all the other indices fixed. The mode-d matricization, also known as unfolding or flattening, is defined as the reordering of the elements of tensor $\mathbf{X}$ into a matrix $\mathbf{X}_{(d)} \in \mathbb{R}^{I_d \times (I_1 \times I_2 \times \cdots \times I_{d-1} \times I_{d+1} \times \cdots \times I_N)}$ by arranging the mode-d vectors to be the columns of $\mathbf{X}_{(d)}$. The mode-d product of tensor $\mathbf{X}$ by a matrix $\mathbf{U} \in \mathbb{R}^{J_d \times I_d}$, denoted by $\mathbf{X} \times_d \mathbf{U}$, is a tensor with entries $(\mathbf{X} \times_d \mathbf{U})_{i_1,i_2,\cdots,i_{d-1},j_d,i_{d+1},\cdots,i_N} = \sum_{i_d} \mathbf{X}_{i_1,i_2,\cdots,i_N} \cdot \mathbf{U}_{j_d,i_d}$. Some more detailed notations and multilinear rules can be referenced in the literature.[11–14]

2.2 *General Model Description*

The degradations of HSI mainly include noise, strips, dead pixels, clouds, and so on. Many methods can be used for each degradation case. Interestingly, for most techniques, the variational

models (VMs) are mainstream, popular, and promising. Therefore, we concentrate on the variational methods for the quality improvement of HSIs.

In general, the degradation of the HSI can be written as

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

where $\mathbf{Y} \in \mathbb{R}^{n \times B}$ is the observed HSI, and $\mathbf{X} \in \mathbb{R}^{m \times B}$ is the target image. It is noted that $\mathbf{Y}$ and $\mathbf{X}$ are often as matrixes consisting of vectors of all bands, for example, $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \cdots \mathbf{x}_B]$ with B being the number of bands, while they also can be denoted in the form of tensors.[15] The matrices $\mathbf{A}$ and $\mathbf{B}$ denote the degradation operator in the spatial and spectral dimensions, respectively, and $\mathbf{N} \in \mathbb{R}^{n \times B}$ is the additive noise. Our goal for the degradation problems is to recover the unknown image $\mathbf{X}$ based on the observed image $\mathbf{Y}$. The solution of this inverse problem can be summarized and described in the variational framework.

The VMs can be described from the statistical or the algebraic perspective. The standard regularized solution of the inverse problem of quality improvement is the minimum of the function:

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{AX}\|_p^p + \lambda\|\Gamma(\mathbf{X})\|_q^q = \arg \min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{AX}\|_p^p + \lambda p(\mathbf{X}), \tag{2}$$

where the first term is the data fidelity term that provides a measure of the conformance of the present image $\mathbf{X}$ to the observed image $\mathbf{Y}$. The second term is the regularization term that imposes some prior constraint on the solution, where $\|\Gamma(\mathbf{X})\|_q^q$ or $p(\mathbf{X})$ is the constraint function. The parameters p and q are the norm values for the fidelity and regularization term, respectively, and they are often selected as 1, 2, or a decimal value in the interval of 1 and 2. λ is the regularization parameter balancing these two terms. For data fidelity, the l_2 norm ($p = 2$) is widely used in different algorithms because of the simplicity of solution, especially for the noise of Gaussian type.[16–18] For impulse noise and outliers on the images, it has been proven that l_1 fidelity is more effective than l_2 fidelity.[19,20] Compared with l_2 norm, however, the convergence rate of the l_1 norm is often much slower. Some efficient approximation methods[19–21] have been developed for the l_1 optimization. For the consideration of complicated types of noise and model error, an l_1-l_2 hybrid model is also proposed for the fidelity term.[22–24]

To reduce the computational load and avoid the spectral artifacts, HSIs are often processed in a spectral transformation space. Suppose $\mathbf{T}$ is a spectral transformation function, the transformed version of the input HSI is obtained by $\mathbf{Y}_t = \mathbf{TY}$. In the transform field, the image $\mathbf{X}_t$ is then solved as

$$\hat{\mathbf{X}}_t = \arg \min_{\mathbf{X}_t} \|\mathbf{Y}_t - \mathbf{AX}_t\|_p^p + \lambda p(\mathbf{X}_t). \tag{3}$$

Finally, the desired image can be solved by an inverse transform $\hat{\mathbf{X}} = \mathbf{T}^{-1}\hat{\mathbf{X}}_t$.

Except for reducing the dimension of HSI in spectral space, high-dimensional subspace transformation is able to represent and discriminate the characteristics of spatial and/or spectral information. Many types of images captured by different sensors can be sparsely represented using a dictionary of atoms. Hence, sparse representation is devoted to excavating the basic structural unit of an image.[25] Any signals in HSI can be represented as a sparse linear combination with respect to a dictionary, which consists of atoms that represent the structure of the image. Suppose $\mathbf{D}$ is the dictionary, $\mathbf{X}$ can be expressed as $\mathbf{X} = \mathbf{D}\boldsymbol{\alpha}$, where $\boldsymbol{\alpha}$ is the basis coefficient. The optimization problem can be solved by $\hat{\boldsymbol{\alpha}} = \min \|\boldsymbol{\alpha}\|_0$ with some constraint condition, e.g., $\mathbf{Y} = \mathbf{AD}\boldsymbol{\alpha}$. The l_0 minimization is an NP-hard combinatorial search problem that is always difficult to solve. Therefore, it is often translated into a convex optimization under the condition of restricted isometry property.[26] Then, a general regularized solution based on sparse optimization can be obtained:

$$\hat{\boldsymbol{\alpha}} = \arg \min_{\boldsymbol{\alpha}} \|\mathbf{Y} - \mathbf{AD}\boldsymbol{\alpha}\|_p^p + \lambda\|L(\mathbf{D}\boldsymbol{\alpha})\|_q^q + \gamma p(\boldsymbol{\alpha}). \tag{4}$$

Then, the desired image can be obtained with an inverse transform $\hat{\mathbf{X}} = \mathbf{D}\hat{\boldsymbol{\alpha}}$. The subspace transformation matrix $\mathbf{D}$ can be obtained via different approaches, such as PCA, vertex component analysis, or dictionary learning.[27]

Consider improving the image quality with auxiliary data from observation time or sensors, the energy functional model can be given as the following two optimization problems:

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{A}\mathbf{X}\|_p^p + \lambda_1 f(\mathbf{X}, \mathbf{Z}) + \lambda_2 p(\mathbf{X}), \tag{5}$$

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\mathbf{Y} - \mathbf{A}(\mathbf{D}\boldsymbol{\alpha})^{\mathrm{T}}\|_p^p + \lambda_1 f(\boldsymbol{\alpha}, \mathbf{Z}) + \lambda_2 p(\boldsymbol{\alpha}). \tag{6}$$

Here, the second term describes the relationship between a desired image and other complementary observation $\mathbf{Z}$ in Eq. (5), while it denotes the relationship between a coefficient $\boldsymbol{\alpha}$ of images $\mathbf{X}$ in the transform domain and $\mathbf{Z}$. The third term is a regularization term related with $\mathbf{X}$ and $\boldsymbol{\alpha}$. For spatiospectral fusion with two observation images, which is one of the most classic applications, $\mathbf{Z}$ represents the auxiliary high-resolution (HR) image (the corresponding PAN or MS image), and $\mathbf{X}$ is the target high spatial resolution HSI. Then, the second is $\|\mathbf{Z} - \mathbf{X}\mathbf{S}\|_p^p$ and $\|\mathbf{Z} - \mathbf{D}\alpha\mathbf{S}\|_p^p$, which represents data fidelity and has a spectral operator $\mathbf{S}$. λ_1 and λ_2 are constant parameters. In addition, when $\mathbf{S}$ is unknown, Eq. (5) can be replaced by jointly Gaussian.[28,29]

3 Model with Different Priors

HSI is a third-order tensor that incorporates two spatial modes and one spectral mode. Depending on different modes, the previous works on HSIs can be mainly divided into four categories: (A) single-channel prior model to only make use of spatial information on each band, (B) cross-channel prior model to blend spatial and spectral intrinsic structure, (C) tensor-based model to treat HSI or the three-dimensional (3D) patch as a third-order tensor for reserving the original structure, and (D) data-driven prior-based model to reveal the complex nonlinear relation between degraded-clean image. In the A and B categories, the priors/regularization terms act on the HS matrix and are designed the basis of the intrinsic structure of image edges and textures. However, priors in the C and D categories treat an HSI as a 3D array. Especially, the last model is obtained by pretraining on a large number of high-low quality data pairs. This section provides overviews of these technologies, including recent advances.

3.1 Single-Channel Prior-Based Method

As an ill-posed inverse problem, HSI quality improvement method usually employed regularization techniques to solve this problem by adding constraints to the objective function. The main objective of regularization is to incorporate more information about the desired solution to stabilize the problem and find a useful and stable solution. For HSIs, the regularization is usually applied band-by-band to obtain the desired results. For a single-channel image, various popular algorithms[25,30–34] have been developed to solve such problem. In the spatial domain, Tikhonov regularization[30] is first proposed as a basic regularization to enforce a smooth constraint to eliminate noise signals but causes the edges to be blurred. Considering the edge preservation, total variation (TV) regularization[31] allows occasional larger jumps leading to piecewise smoothness instead of overall smoothness and produces powerful results. On the basis of TV, bilateral TV[32] builds different weighting coefficients according to the distance from neighborhood pixels to control the strength of regularization. However, when the noise level is high, these methods, which only utilizing the local correlations, cannot perform very well. Thus, nonlocal TV (NLTV)[33] is designed to make use of the self-similarity of natural images in a nonlocal manner and can better recover the repetitive texture information in particular. As a patch-based approach, since the BM3D[35] framework has a better performance on denoising, Danielyan et al.[36] adapted the BM3D for the inverse problem of image deblurring.

Except to construct *a prior* directly in the space domain, the regularization can be introduced into the transform domain as described in Eq. (4), typically the sparsity penalty term in sparse

representation. As an early machine learning technique, sparse representations have become a trend and are used for restoration problems.[25,37,38] The general idea of the methods is that each patch in the estimated image can be expressed as a linear combination of only few patches from a redundant dictionary, learned using a large group of patches from an image dataset. The representative dictionary learning-based methods include K-clustering with singular value decomposition (K-SVD),[37] learned simultaneous sparse coding,[39] and clustering-based sparse representation.[34] However, the sparse model is still computationally expensive but can only describe the linear relationship between image pairs.

Similar to sparse representation, low-rank-based methods can also describe the redundancy of the signal well and have attracted increasing attention. Typical one is the low-rank matrix factorization with the nuclear norm minimization (NNM).[40] To further improve the flexibility of NNM, Gu et al.[41] proposed a weighted nuclear norm minimization (WNNM) model. Considering the heavy-tailed distributions of both sparse noise/outliers and singular values of matrices, bilinear factor matrix norm minimization model[42] is proposed for corrupted data. Naturally, low-rank constraint can also be integrated with other priors, such as TV prior[43] and nonlocal self-similarity prior.[44]

In fact, due to the extraction of deep structure and the presence of nonlinearities achieved by deep networks, recent progress has been made in incorporating deep priors into general model-based inverse methods, which are detailed in Sec. 3.4.

3.2 Cross-Channel Prior-Based Method

In parallel with these advanced single-channel prior, a number of cross-channel prior approaches have been pioneered for a wide array of HSI problems from image denoising and reconstruction to fusion. The cross-channel prior is usually designed to extract the useful spatial or spectral information from other bands according to the similarity or difference among channels. Due to the importance of preserving spectral information, these cross-channel priors put specific attention on the highly correlated spectral bands to utilize abundant spatial and spectral information in HSI. With the complementary information from other temporal/sensor images, the cross-channel prior can obtain sufficient complementary texture and structure information from these similar bands in other temporal/sensor images to reconstruct the spatial information.

The objective function can be given by integrating the correlation and complementation from other bands:

$$\{\hat{\mathbf{x}}_i\}_{i=1}^{B} = \arg\min_{\{\mathbf{x}_i\}} \sum_{i=1}^{B} \|\mathbf{a}_i\mathbf{x}_i - \mathbf{y}_i\|_2^2 + \lambda p(\mathbf{x}_i, \mathbf{x}_{cc}), \tag{7}$$

where the first item ensures the proximity between the observed image band $\mathbf{y}_i$ and the expected high quality image band $\mathbf{x}_i$, and the vector $\mathbf{a}_i$ represents the degradation process between the degraded-clean band pair; and the second item regularizes the function by imposing constraint from the correlated and complementary spectral bands $\mathbf{x}_{cc}$.

From the well-known TV regularizer[31] to low-rank regularizer,[34] these two-dimensional (2D) priors on a single band can be simply extended into 3D mathematical formulation by excavating the trend of variability along the spectral direction.

3.2.1 Multichannel TV regularizer

Different from the single-channel TV model, the multichannel regularizers in HSIs can account for the spatial and/or spectral variations. Thus, an extended version called spatiospectral TV (SSTV)[45,46] is also widely used. Its anisotropic version is defined as follows:

$$\|\mathbf{X}\|_{\text{sstv}} = \sum_{i,j,k} w_1 |\mathbf{x}_{i,j,k} - \mathbf{x}_{i,j,k-1}| + w_2 |\mathbf{x}_{i,j,k} - \mathbf{x}_{i,j-1,k}| + w_3 |\mathbf{x}_{i,j,k} - \mathbf{x}_{i-1,j,k}|, \tag{8}$$

where $\mathbf{x}_{i,j,k}$ is the (i, j, k)'th entry of $\mathbf{x}$, and $w_n (n = 1, 2, 3)$ is the weight along the j'th mode of $\mathbf{x}$ that controls its strength of regularization. It is worth noting that the above norm defined as

SSTV can fully capture the spatial and spectral differential information of HSI. By characterizing the piecewise smooth structure in both spatial and spectral domains, the Gaussian noise in the spatial and spectral dimensions can be removed.

The authors[47] also proposed an adaptive version for HSI and developed a spectral–spatial adaptive hyperspectral TV (SSAHTV), which have the following formation:

$$\text{HTV}(\mathbf{x}) = \sum_{i=1}^{MN} W_i \sqrt{\sum_{j=1}^{B} (\nabla_{ij}\mathbf{x})^2}, \tag{9}$$

where ∇_{ij} is the linear operator that corresponds to the first-order differences at the i'th pixel in the j'th band. W_i is a weighted parameter to control the regularization strength in the different pixels (see Ref. 47 for more detail about the selection of W_i). In this model, large regularization strength is automatically enforced in the bands with high intensity noise and flat area. Conversely, weak regularization strength is used in low noise intensity bands and edge regions.

3.2.2 Multichannel nonlocal TV regularizer

Inspired by the NLTV,[48] two multichannel NLTV models,[49,50] which introduce the nonlocal gradient on spatiospectral dimension to the inverse problems, have been proposed to suppress the staircase effect from TV, and preserve the fine structures, details, and textures better. One[49] is given as an adaptive hyperspectral NLTV, which has the following formula:

$$R(\mathbf{X}) = \sum_{i \in M \times N} \sqrt{\sum_{b=1}^{B} \sum_{j \in M \times N} (\mathbf{x}_b(i) - \mathbf{x}_b(j))^2 w_b(i,j)}. \tag{10}$$

Let $w_b(i,j)$ be a weight function between two patches centered at locations i and j in b'th band.

The other one[50] is represented as

$$R(\mathbf{X}) = \sum_{b=1}^{B} \sum_{i=1}^{M \times N} |\nabla_w \mathbf{x}_b(i,b)| = \sum_{b=1}^{B} \sum_{i=1}^{M \times N} \sqrt{\sum_{k \in NS} \sum_{j \in N_i^k} (\mathbf{x}_b(i) - \mathbf{x}_k(j))^2 \mathbf{w}(i,b;j,k)}. \tag{11}$$

Compared with Eq. (10), model Eq. (11) integrates the similarity of the spectral band into the NLTV model. It computes the nonlocal gradients for a pixel centered at i in spatial and spectral dimensions with these patches from the current band and the structurally similar band, respectively. NS_b denotes the set of all the selected bands for the current b'th band. The set contains the current b'th band and these similar bands. The regularization Eq. (11) can improve the current bands by exploiting the redundant information from the neighboring bands with higher quality, such as lower noise-intensity or less information missing.

3.2.3 Spatiospectral distributed sparse representation

Applying the sparsity of the image induced by the redundancy of spatial information, this approach has performed well in 2D image denoising.[34,51] With more redundancy in HSIs because of the high spectral correlation caused by the narrow spectral channels, the SSDSR[52] can be designed as

$$\{\hat{\boldsymbol{\alpha}}_c, \hat{\boldsymbol{\alpha}}_b\} = \arg \min_{\mathbf{z}_c, \mathbf{z}_b} \sum_{b=1}^{T} \|(\mathbf{D}_{b,c}\boldsymbol{\alpha}_c + \mathbf{D}_{b,b}\boldsymbol{\alpha}_b) - \mathbf{u}_b\|_2^2 + \mu \left(\sum_{b=1}^{T} \|\boldsymbol{\alpha}_c\|_0 + \|\boldsymbol{\alpha}_b\|_0 \right). \tag{12}$$

To distinguish and utilize the interband correlation and the intraband structure well, the SSDSR decomposes the sparse coefficients into interband vectors $\boldsymbol{\alpha}_c$ and intraband vector α_b. In Eq. (12), $\mathbf{D}_c$ and $\mathbf{D}_b$ are used to learn the common and specific structures for

different bands, as a consequence of improving the results compared with band-by-band sparse representation.

3.2.4 HSI low-rank regularizer

The redundancy in the spatial and spectral information reveals the low-rank property of HSI. From a perspective of spectral dimension, HS pixels are composed of a few pure endmembers, and the number of pure endmembers is much smaller than the HSI dimension. More specifically, supposing the upper bound of the number of pure spectral endmembers for the HSI patch is r, then the rank of $\mathbf{X}_{i,j}$ is bounded,[53] that is, $\text{rank}(\mathbf{X}_{i,j}) \leq r$. In terms of spatial dimension, the nonlocal similarity of patches implicitly shows the low-rank property of spatial domain in HSIs.[53,54] Based on low-rank property, the optimization problem of this model can be represented based on patches from HSI and can restore each patch sequentially:

$$\arg \min_{\mathbf{X},\mathbf{S}} \|\mathbf{X}_{i,j}\|_* + \lambda \|\mathbf{S}_{i,j}\|_1 \quad \text{s.t.} \ \|\mathbf{Y}_{i,j} - \mathbf{X}_{i,j} - \mathbf{S}_{i,j}\|_F \leq \delta, \qquad (13)$$

where $\mathbf{X}_{i,j}$ is the low-rank HS matrix and $\mathbf{S}_{i,j}$ is the sparse error matrix, describing the mixture of sparse noise, such as dead lines and stripes. The constraint $\|\mathbf{Y}_{i,j} - \mathbf{X}_{i,j} - \mathbf{S}_{i,j}\|_F \leq \delta$ is used to measure the relation $\mathbf{Y}_{i,j} = \mathbf{X}_{i,j} + \mathbf{S}_{i,j} + \mathbf{N}_{i,j}$, where δ is a constant related to the standard deviation of random noise or model error $\mathbf{N}_{i,j}$. To obtain a tractable optimization problem, the l_1-norm and the nuclear norm are used to relax the sparsity and rank, respectively. By estimating low-rank matrix and sparse matrix simultaneously, low-rank regularizer can help to restore an HSI corrupted by striping noise and signal-independent noise in many applications.

Naturally, another direction of the cross-channel prior is to consider the spatial and spectral characteristics simultaneously. Spatial and spectral constraints in a 2D or 3D way can also be jointed to utilize the correlation and difference between bands better. For an HSI, a better and robust regularization model is that it cannot only consider the cross-channel correlation among different bands (spectral adaptive) but can also adjust the regularization strength of different pixels in the same band automatically (spatial adaptive).

3.3 Tensor-Based Model

The aforementioned technologies can obtain the promising results and achieve the tremendous progress in various applications. However, several dark clouds exist over the HSI quality improvement, which are urgently needed to be considered. Due to the inherent 3D characteristics of the HSIs, the previous vector-/matrix-based methods have limitations in fully exploiting the spectral–spatial structural correlation in comparison with directly working on a three-order tensor format image.[10–12] The recent works consistently indicate that the tensor-based methods substantially preserve the intrinsic structure correlation with better restoration results. To better utilize low-rankness to one mode of the tensor, some theoretical tensor frameworks have been established to exploit low-dimensional structure in high-dimensional data. For example, the CANDECOMP/PARAFAC (CP) rank[55] is defined as the smallest number of rank one tensor decomposition but is generally an NP-hard problem to compute. Tucker rank[56] obtained by tucker decomposition can represent the low-rank property of each mode-i matricization of tensor.

By applying Tucker decomposition, a desired HRHS image can be represented as a core tensor multiplied by the dictionaries of the width mode $\mathbf{W}$, height mode $\mathbf{H}$, and spectral mode $\mathbf{S}$, specifically $\mathcal{X} = \mathcal{C} \times_1 \mathbf{W} \times_2 \mathbf{H} \times_3 \mathbf{S}$. A low-rank tensor approximation[57] method employs the Tucker factorization method to obtain denoising results well. For a spatially or spectrally downsampled low-resolution (LR) HSI $\mathcal{Y}$ of $\mathcal{X}$, it can be assumed that the point spread function of the HS sensor and the downsampling matrices of the width mode and height modes are separable.[58] The spatial degradation matrices work along the width and height modes, and the spectral downsampling matrix imposes on the spectral mode. Then, the acquired image can be written as

$$\mathcal{Y} = \mathcal{X} \times_1 \mathbf{P}_1 \times_2 \mathbf{P}_2 \times_3 \mathbf{P}_3 = \mathcal{C} \times_1 (\mathbf{P}_1 \mathbf{W}) \times_2 (\mathbf{P}_2 \mathbf{H}) \times_3 (\mathbf{P}_3 \mathbf{S}), \qquad (14)$$

where $\mathbf{P}_1$, $\mathbf{P}_2$, and $\mathbf{P}_3$ denote the possible degradation matrices along the width, height, and spectral modes, respectively, which describe the spatial and spectral responses of the imaging sensors. The tensor $\mathcal{C}$ holds the coefficient of $\mathcal{X}$ over the three dictionaries. The relationship in Eq. (14) can be extended into the data fidelity in the variational framework instead of its matrix form.

As for the regularization term, motivated by the fact that the nuclear norm is the convex envelope of the matrix rank within the unit ball of the spectral norm, the sum of nuclear norms (SNN)[59] on each mode is used as a convex surrogate of the tensor rank. Recently, tensor train (TT) rank[60] and tensor ring (TR) rank[61] have drawn considerable attention because of their computational efficiency and high compression properties. Compared with the Tucker rank, TT rank constitutes the ranks of matrices formed using a well-balanced matricization scheme and has the capacity to capture the global correlation of the tensor entries.[62] TR decomposition model, which represents tensor more flexibly, is regarded as the linear combinations of a group of TT representations.[63] TR representation model can also effectively reveal the characteristics of time-series RS images and different scales in the spatiospectral mode.[64] By constraining the subcomponent from tensor decomposition, we can take the correlation between HSIs over different bands into consideration and attempt to eliminate the information loss generated by tensor flattening.

3.4 Data-Driven Prior-Based Model

The above methods mainly regard the relationship between the observed images and target images as linear simulation. However, the linear model would restrict the recovery quality when observed images are confronted with complex mixed noise, nonuniform blur, uneven haze, or nonoverlap spectral rang among different sensors. Recently, deep network theory can provide a prominent performance to describe the complex nonlinear relationship because of its feature extraction and mapping learning capabilities.[65] While most existing DNN-based methods solve the quality improvement problems by directly mapping low-quality images to desirable high-quality images, the observation models characterizing the image degradation processes have been largely ignored. As a consequence, different from model-based optimization methods that can flexibly handle different image restoration (IR) tasks by exploiting state-of-the-art image prior, these deep network-based methods are usually restricted by specialized tasks. As the data likelihood term has not been explicitly exploited, deep network-based methods[66,67] need to train a different model for various IR tasks separately. Hence, the question is how we should combine DLVM in result to flexibly and effectively solve IR tasks for HSI. To address this issue, a VM combining data-driven prior in the field of remote sensing can be divided into two forms.

The former, which can be called as plug-and-play (PNP) framework, uses a convolutional neural network (CNN) to mine the deep prior of the image and forms an off-the-shelf deep denoiser and then generally plugs the off-the-shelf denoiser to solve subproblem associated the energy function. For example, Zhang et al.[68] trained a set of CNN denoisers and integrated them into model-based IR framework for different IR tasks. Zeng et al.[69] embedded the tensor Tucker decomposition method and a CNN denoiser into the PNP. The tensor Tucker decomposition method can remove sparse noise well and part Gaussian noise by exploring the global spatiospectral correlations. Meanwhile, the CNN denoiser, as a a physical prior, was used to remove the residual noise.

The optimization function can be expressed in general terms as

$$\hat{\mathbf{X}} = \arg \min_{x} f(\mathbf{X}) + \lambda g(\mathbf{X}, \Theta). \qquad (15)$$

In these methods, a deep regularization prior $g(\mathbf{X}, \Theta)$ about $\mathbf{X}$ can be obtained by a pretrained network with the observed degraded image as an input to reveal a prior relationship Θ between degraded–clean image pairs. Mathematically, the solution of deep regularization prior can be served as a preprocessing.[69] It can also be integrated into the subproblem of the optimization algorithm to update image features, such as half-quadratic splitting algorithm and alternating direction method of multipliers (ADMM) algorithm.[68,70]

To reveal the complex nonlinear relationship among different sensors, deep learning regularizers-based method can be generally given as

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{x}} \|\mathbf{Y} - \mathbf{AX}\|_p^p + \lambda \Phi(\mathbf{X}, \theta) + \mu \Psi(\mathbf{X}, \mathbf{Z}, \Theta), \qquad (16)$$

where $\Phi(\mathbf{X}, \theta)$ is the nonlinear function to learn the prior parameters set θ. If given an auxiliary and available data $\mathbf{Z}$, then $\Psi(\mathbf{X}, \mathbf{Z}, \Theta)$ is the nonlinear function that learns the relations between $\mathbf{X}$ and $\mathbf{Z}$ and has a corresponding parameter set Θ. The priors $\Phi(\mathbf{X}, \theta)$ and $\Psi(\mathbf{X}, \mathbf{Z}, \Theta)$ can be plugged in an iterative scheme by decoupling the fidelity term and regularization term.[70] With the predefined nonlinear function learned by deep learning, the model-based optimization methods will be less time-consuming in comparison with sophisticated priors while still retaining flexibility for different tasks.

The latter uses network[71] to connect the model iterative optimization process and the HSI spatiospectral prior. It directly assigns all parameters that would have to be solved in the model into deep learning. In other words, network learning is used to represent and train an unfolding iterative optimization for Eqs. (2)–(4). For example, Yang et al.[72] unfolded the ADMM algorithm to a deep network for fast compressive sensing. Wang et al.[73] also embedded the structure insight of the conjugate gradient algorithm for HSI denoising, which guaranteed the relationship between the desired HSI and the original HSI into a network and formed the data-driven prior, which was also called an optimization-inspired network.

4 Quality Improvement by Available Information

In HSI quality improvement, additional information from other spectral bands/temporal images/sensor images can provide more available spatial and spectral features for solving degradation problems. The complementary information from different sources can be transformed into cross-channel priors, tensor representation, and deep priors to better achieve high quality HSIs. To embody their effectiveness, examples of typical applications are used to show in this section.

4.1 *HSI Restoration with Hybrid Noises*

4.1.1 *Extract complementary information from other spectral bands*

For an HSI, the degradation, such as noise, can not only be considered in the spatial dimension but also in the spectral dimension, as shown in Fig. 3. The regularization item plays a vital role in the variational framework. It gives a prior distribution of the nondegradation image and controls the perturbation of the solution, thereby guaranteeing a stable estimation. Although some methods[74] can solve the degradation problem on spectral dimension to some extent, it still neglects the strong correlation and similarity across different bands. The information from other
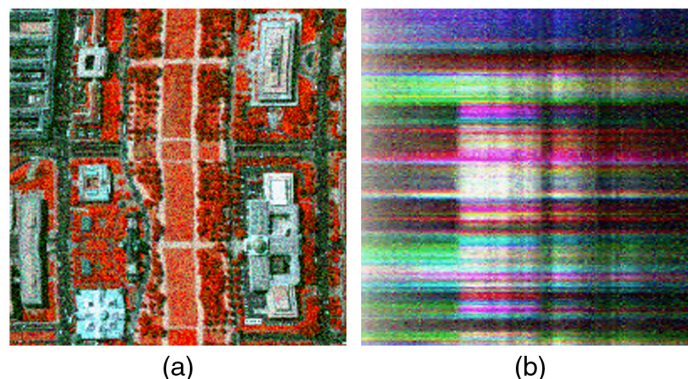

(a)   (b)

Fig. 3 Image degradation in spatial and spectral dimension. (a) Noise degradation in spatial dimension and (b) noise degradation in spectral dimension.

bands can usually constrain and improve the recovery of the present band. Therefore, when no other data with complementary information are available, these mentioned methods in Secs. 3.2–3.4 are potential ways of improving the robustness and precision for various applications.

Moreover, the combination of different regularization terms is an effective and promising way to simultaneously describe the characteristics of different kinds of noises. By applying the sparsity prior to HSI data, Zhao et al.[75] investigated sparse coding to describe the global and local redundancy and correlation (RAC) in the spectral domain and then employed a low-rank constraint to capture the global RAC in the spectral domain. Xie et al.[76] proposed a nonconvex regularized low-rank and sparse matrix decomposition method to simultaneously remove the Gaussian noise, impulse noise, dead lines, and stripes. With regard to the more complicated single super-resolution problem, Guo et al.[77] used the unmixing information and TV minimization to produce a higher resolution HSI. By modeling the sparse prior underlying HSIs, a sparse HSI super-resolution model[78] was proposed. Zhang et al.[79] proposed a maximum *a posteriori*-based HSI super-resolution reconstruction algorithm, in which PCA was employed to reduce computational load while removing noise. Huang et al.[80] presented a super-resolution approach of HSIs by joint low-rank and group-sparse modeling. Their approach can also deal with the situation wherein the system blurring was unknown. Li et al.[81] explored sparse properties in the spectral and spatial domains for HSI super-resolution. An HSI spatial super-resolution[82] was proposed to exploit the nonlocal similar characteristics hidden in several four-dimensional tensors and the local smoothness in the spatial and spectral modes. In general, due to the high spectral–spatial redundancy property for the 3D tensor HSIs, the sparsity or low-rank-based HSI methods[80–82] are the mainstream of artificially designed priors and have achieved state-of-the-art performances.

Given the good nonlinear representation ability of deep learning, a pretrained deep prior embedded in the variational framework is recently a popular trend. However, it is still seldom used for HSI restoration tasks. Nowadays, the main strategy is PNP framework.[69,83] Motivated by Ref. 70 to further reduce their computational complexities, Wang et al.[73] also unfolded the iterative optimization process into a feedforward neural network, whose layers mimic the process flow of the proposed denoising-based IR algorithm. Then, the pretrained deep priors can be jointly optimized with other algorithm parameters.

4.1.2 *Considering separately spatial and spectral degradations using priors*

In addition to directly impose spatiospectral constraints on the original image, the optimization problem can be viewed from two perspectives of the spatial and spectral views by defining the degradation functions and noise according to a specific task, respectively,

$$\mathbf{Y}^{spa} = \mathbf{A}\mathbf{X}^{spa} + \mathbf{N}^{spa}, \tag{17}$$

$$\mathbf{Y}^{spe} = \mathbf{H}\mathbf{X}^{spe} + \mathbf{N}^{spe}. \tag{18}$$

Equations (17) provides the spatial degradation model whereas Eq. (18) describes the contaminated spectral information suffering blur and noise. $\mathbf{X}^{spa}$ and $\mathbf{Y}^{spa}$ are assumed to be HSIs with high- and low-spatial resolutions, respectively, in the spatial views. $\mathbf{X}^{spe}$ and $\mathbf{Y}^{spe}$ are HSIs with high- and low-spectral resolutions, respectively, in the spectral views. $\mathbf{A}$ and $\mathbf{H}$ are the spatial degradation matrix and spectral dimension blurring operator, respectively. $\mathbf{N}^{spa}$ and $\mathbf{N}^{spe}$ are the zero-mean Gaussian distribution noise in spatial and spectral dimensions, respectively. For HSI denoising, Eq. (18) can be simplified as $\mathbf{Y}^{spe} = \mathbf{X}^{spe} + \mathbf{N}^{spe}$.

The simple strategy for regularization is to use single-band prior models on each band separately but ignores the preservation of spectral characteristics. In denoising, the wavelet-based method[74] has been successfully applied in spatial and spectral dimensions. For variational framework, Fig. 4 shows a basic strategy by merging the results from the spatial and spectral domains. Considering the spatial and spectral degradations jointly, the objective function will be first constructed using Eqs. (17) and (18), and then the optimal solution on two different dimensions can be obtained using appropriate optimization algorithms. Then, the weighted correlation is used to combine the spatial and spectral results to achieve a final result.
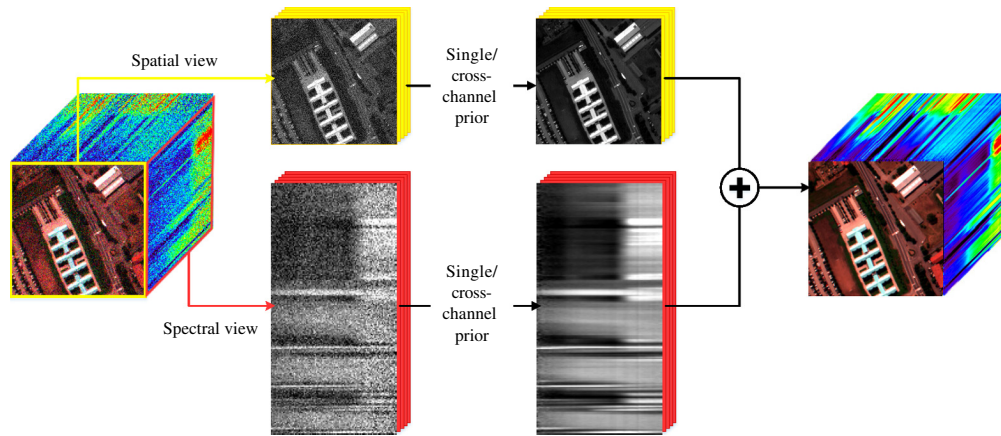
Fig. 4 Processing framework considering degradations in both the spatial and spectral dimensions.

Notably, spatiospectral regularizations in Secs. 3.2–3.4 can also be used in different views and have a potential for better results. An HSI denoising method[84] as a typical example with spatial and spectral view fusion strategy is introduced. The HSI is denoised with SSAHTV model shown in Eq. (9), both from spatial and spectral views. Then, the results for two views are fused band by band in a weighted way $\mathbf{x}_b = (Q_b^{\text{spa}}\mathbf{x}_b^{\text{spa}} + Q_b^{\text{spe}}\mathbf{x}_b^{\text{spe}})/(\mathbf{x}_b^{\text{spa}} + \mathbf{x}_b^{\text{spe}})$. The weights of the different views are adaptively defined using metric Q.[85] Here, $\mathbf{x}_b$, $\mathbf{x}_b^{\text{spa}}$, $\mathbf{x}_b^{\text{spe}}$ are b'th bands of fused HSI, denoised result in the spatial views, and denoised result in the spectral views, respectively. Q_b^{spa} and Q_b^{spe} are the weights of each band computed by metric Q.

4.1.3 *Experimental evaluation*

Comparison between different spatiospectral regularizations. To show the effectiveness of the different regularization models, two cases have been given on the denoising experiments. The adopted data are the Washington DC Mall dataset collected by the Hyperspectral Digital Imagery Collection Experiment (HYDICE) with the cropped size of $200 \times 200 \times 191$. In Fig. 5, the same noise intensity with $\sigma_b = 5$ is added to each band. In Fig. 6, zero-mean Gaussian noise and stripes are simultaneously added to all the bands of HSI. The Gaussian noise standard deviation of each band randomly varies from 0 to 40 dB, while the stripe with one-pixel width covers 30% on each band. The mean peak-signal-to-noise ratio (PSNR), the mean structural similarity (SSIM) index, and the mean spectral angle (MSA) mapper served as evaluation indices for the simulated experiments.

The traditional single-channel methods (TV,[31] NLTV,[33] KSVD,[37] and WNNM[41]) and the cross-channel methods in Sec. 3.2 are conducted on the experiments. The cross-channel methods in Sec. 3.2 can be regarded as variants of single-channel methods. Parameters of all the compared methods are adjusted to the optimal according to their references. The quantitative assessment of the four group results also indicates that the cross-channel regularization models are generally suitable for the recovery of HSIs. The noise is not only well suppressed but also the detail and edge information are also preserved excellently. Although the spurious artifacts almost disappear in cross-channel variation-based methods, the problem of detailed loss still exists in TV-based model. By employing the nonlocal similarity, results are more visually plausible in detail. However, the sparse representation and low-rank decomposition methods can preserve more detailed information for small objects. Especially, HSI low-rank model has an obvious superiority on the reduction of hybrid noise in Fig. 6. By contrast, the sparse representation-based cross-channel regularization exhibited poor performance in removing stripe noise independently because of the limitation of synchronously learned stripe structure in the dictionaries. The comprehensive consideration of the different kinds of noise can help remove the hybrid noises better.
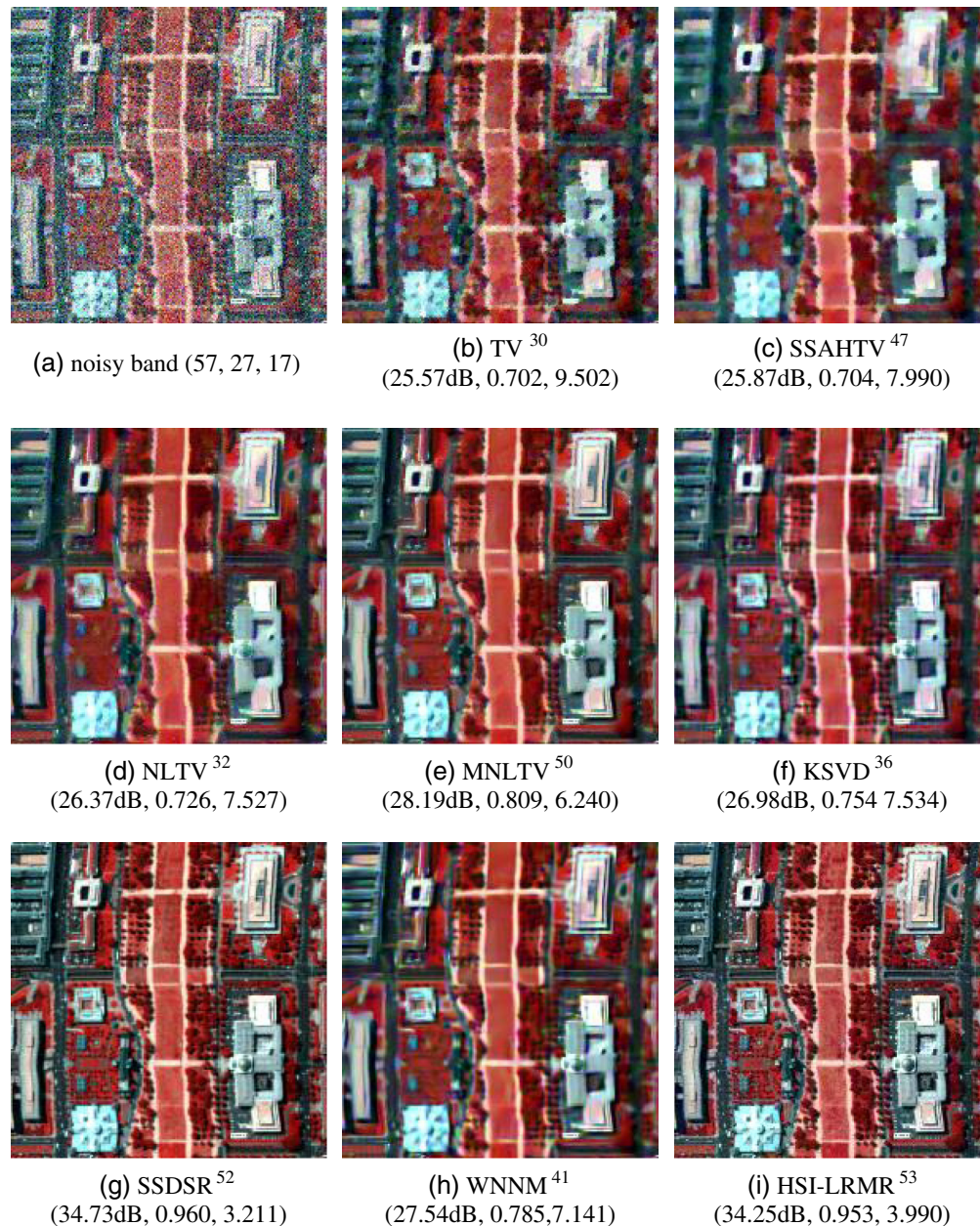
Fig. 5 The visual comparison and quantitative evaluation with PSNR (dB), SSIM, and MSA values of the denoising results in the hyperspectral simulated experiment. (a) Noisy band (57, 27, 17), (b) TV, (c) SSAHTV, (d) NLTV, (e) MNLTV, (f) KSVD, (g) SSDSR, (h) WNNM, and (i) HSI-LRMR.

Effectiveness of fusion strategy. Simulated and real data from the HYDICE are adopted to test the effectiveness of fusion strategy, with the size of $200 \times 200 \times 205$. Washington DC Mall dataset was used in the simulated experiment and Urban dataset was used in the real experiment. The gray values of the HSI were normalized to between 0 and 1. In the simulated experiment, the distribution of zero-mean Gaussian noise was added to each band with a variance of 0.05. In real experiment, Urban dataset was degraded by stripe and Gaussian mixed noise. The SSAHTV was used to remove noise from the spatial view and spectral view. From Eq. (9), it is shown that the optimization of algorithm[84] is only related with regularization parameter λ. In this experiment, λ is selected as the one with the highest Q value.[85] Because the spatial and spectral denoising results can complement each other, the fused denoising result is better than both the sole dimension denoising results. As is shown in Fig. 7(b), although the spatial information is
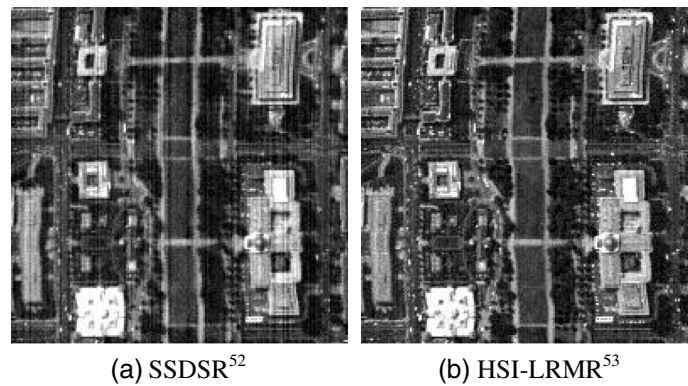
(a) SSDSR[52]   (b) HSI-LRMR[53]

Fig. 6 The comparison of the hybrid noise and stripes reduction results in the simulated experiment. (a) SSDSR[52] result of the stripe band (PSNR = 26.79, SSIM = 0.836, MSA = 7.584), and (b) HSI-LRMR[53] result of the stripe band (PSNR = 33.40, SSIM = 0.957, MSA = 3.858).

preserved well, the noise in the spectral dimension is not completely suppressed, and some noise still remains in the spectral curve. For the spectral view denoising result in Fig. 7(c), the noise in the spectral dimension is suppressed well, the edges are blurred, and the spatial information is not well preserved. For the spatial–spectral view fusion method, not only the spatial information is well preserved but also the noise in the spectral dimension is suppressed better. The result is better than both the individual spatial view and spectral view denoising results.

4.2 Dehazing for Visible Channels

4.2.1 Extract complementary information from other spectral bands

The atmospheric factors not only introduce different types of noises but also produce the haze and thin cloud on the HSIs. Atmospheric attenuation caused by haze and cloud greatly degrades the quality of optical remote sensing images. Fortunately, HS data supply abundant spectral information that covers a region from the visible to infrared spectrum with high spectral resolution. Haze and thin clouds show marked difference in spectrums with large wavelength difference. This characteristic indicates that both adjacent and distant spectra supply complementary information to the concerned spectrum as shown in Fig. 8.

The replacement approaches[86–88] are usually used to remove the haze and thin clouds. They are, however, heavily affected by the correlation of the available fog/cloud-free information. Due to the limitation of majority of MS sensors consisting only of multiple visible bands and one near-infrared (NIR) band, the reflectance of the fog/cloud-free pixel inside an image can only be used to replace and restore the reflectance of another pixel underneath fog/cloud, such as the dark channel prior,[89] and affects the result accuracy. For HSI, since the atmospheric effect is highly wavelength dependent, contamination caused by the atmosphere in the scenes varies in different channels. On the hazy or cloudy days, the observed visible images are vague and lowly contrasted because the visible channels are sensitive to atmospheric conditions. In contrast, the infrared channels with long wavelengths are insensitive to the semitransparent atmosphere. Thus, the scenes in the infrared bands are usually clear and free of haze and cloud. The high correlation in HSI indicates that the infrared channels with cleaner pixels can provide effective prior in variational method to restore the visible images with fog/cloud.

To our knowledge, few variational dehazing algorithms based on complementary spectral information have been developed. However, these methods are oriented for MS image and do not involve HSI. Among them, a cross-channel prior-based method is the mainstream of solving this problem. Its core from a variational gradient-based fusion proposed by Li et al.[90] is to integrate gradient information from referenced clear channels into its highly correlated hazy/cloudy channels to enhance the contrast and remove haze/thin cloud. It has been demonstrated that the short-wave infrared (SWIR) channels at 2.2 μm have a linear correlation with visible channels. Meanwhile, in practice, it is observed that the SWIR channels at 2.2 μm
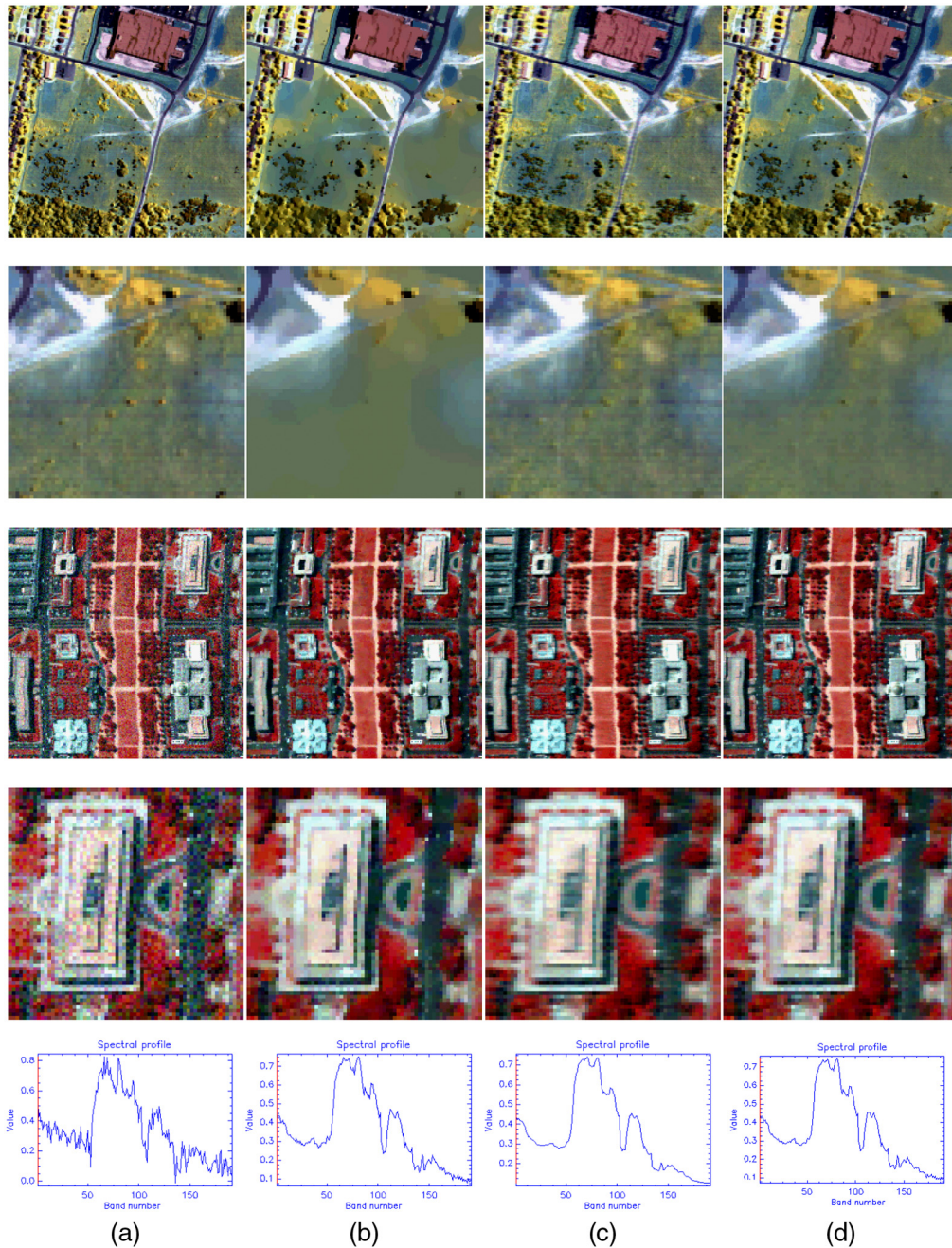
Fig. 7 The denoising results in different views. The first and second rows are the results of Urban dataset, and the third and fifth rows are the results of Washington DC Mall dataset. The fifth row shows the spectral curves. (a) Original noisy image, (b) denoising result in spatial dimension,[47] (c) denoising result in spectral dimension,[47] and (d) denoising results with spatial and spectral view fusion strategy.[84]

exhibit clear land surfaces without contaminations even on hazy days. Thus, the fusion of SWIR and visible band data is explored to remove haze/thin-cloud in visible bands, in which the SWIR channel at 2.2 μm is taken as the referenced channel $\mathbf{x}_{CC}$ to enhance the spatial details in the haze/thin cloud regions. To maintain the spectra invariance after fusion, the constraint of the relationship between the haze/cloud effect and the wavelength, named mean haze projection (MHP),[91] is included in the variational dehazing method. Consequently, by applying Eq. (7), the dehazing model for HS data can be expressed as
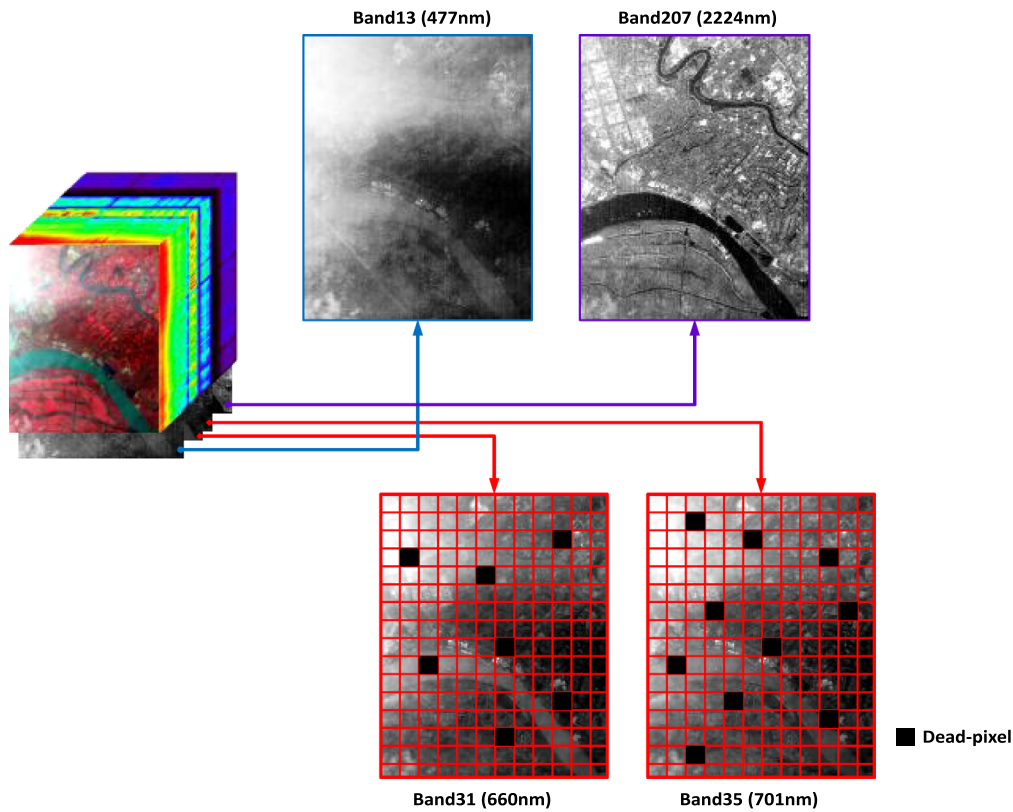
Fig. 8 The complementary information from other spectral bands.

$$\mathbf{x} = \arg\ \min_{\mathbf{x}}\lambda \sum_{b=1}^{B} \|\mathbf{x}_b - \mathbf{y}_b\|_2^2 + \sum_{b=1}^{B}\|\nabla\mathbf{x}_b - \nabla\mathbf{x}_{\mathrm{CC}}\|^2 \quad \text{with the MHP: } \mathbf{x}_{\mathrm{change}} \rightarrow \mathbf{x}_{\mathrm{Achange}},$$

$$(19)$$

where $\mathbf{x}_{\mathrm{change}}$ represents the iteration increment based on the steepest descent numeric solution by Eq. (19), and $\mathbf{x}_{\mathrm{Achange}}$ represents the adjusted $\mathbf{x}_{\mathrm{change}}$ by the MHP. Thus, the iteration equation is expressed as

$$\mathbf{x}_b^{m+1} = \mathbf{x}_b^m - \mathbf{x}_{\mathrm{Achange}}, \qquad (20)$$

where m is the iteration number. More details of this method can be found in Ref. 90.

4.2.2 Experimental evaluation

In this section, variational gradient-based fusion method[90] was extended to deal with HSI. The Hyperion data with 242 bands, covering 400 to 2500 nm spectrum, were used to testify the method. Excluding the infrared and null information bands, the images in the visible spectrum from 467 to 701 nm, covering 24 bands, were taken as the observed contaminated data as shown in Fig. 9(a). The 2193-nm image with clear and sharp gradients was taken as the referenced data $\mathbf{x}_{\mathrm{CC}}$ to remove the haze. The parameter λ was set to 0.01 by considering the visual quality. The result suggests that features of land surfaces, including textures and edges, are enhanced and salient in the fused image as shown in Fig. 9(b). Meanwhile, the spectra of the land surfaces, such as water, building, and vegetation, are reserved in a large degree. The reason is that the radiation increment caused by the thin clouds/haze is directly related to the wavelength, and MHP successfully integrated the relationship between the scattering energy and the wavelength into the variational dehazing algorithm, in which the optimal solution was obtained by simultaneously solving multiple bands. However, the original weak contrast of the homogeneous region
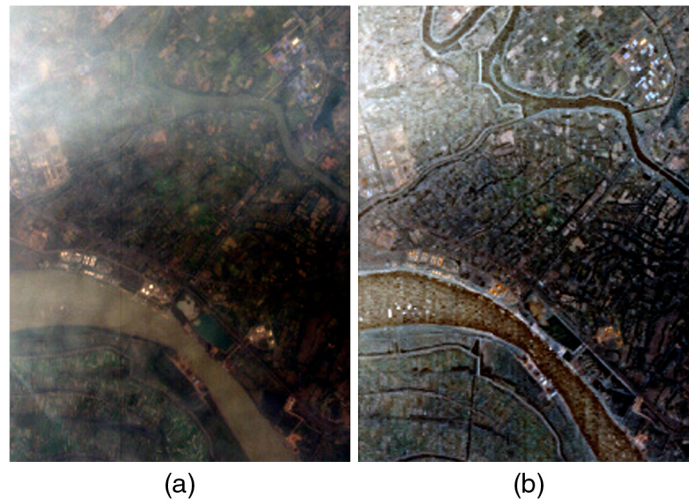
Fig. 9 Experiments on Hyperion data: (a) original hazy visible image. (b) Fused result of the visible and SWIR images.

is also enhanced, while the MHP constraint strengthens the edge and texture features. As shown in Fig. 9, the heterogeneity of the water is also highlighted, resulting in some visual noise.

4.3 Missing Information Reconstruction

Poor weather conditions and/or sensor failure always lead to inevitable information loss for HSIs. Considering the difference in area of missing information, this section illustrates some specific methods to reconstruct the missing information using complementary information and the corresponding VMs. As shown in Fig. 10, taking images from different spectra or periods, it can be easily seen that missing pixels can be surrounded by pixels with complete information after image permutation. Then, missing pixels can be interpolated by employing three different VMs. Currently, existing methods are mostly cross-channel prior-based model, but other models will gradually obtain more attention in the future.

4.3.1 Extract complementary information from other spectral bands for dead-pixel inpainting

Due to long exposure to the harsh environment with intense radiations and space dust collisions, the space-borne sensors are usually subject to partial failures, which results in dead or noisy pixels in the observed imagery. For HSI sensor, the locations of dead or noisy pixels in one CCD are independent of those in the other CCDs. Hence, many redundant spectral information from other bands can be used to reconstruct the missing data in a specific band in Figs. 8 and 10. On the one hand, incomplete (missing information) band can search useful information from spectral bands that are complete or have a different missing location. On the other hand, some remaining information remains in the corrupted band. The basic idea of this class of methods[92–94] is to make use of the other complete spectral bands (one or more) to reconstruct the incomplete band by modeling the relationship between the incomplete and complete bands.

Currently, a few methods[95–97] have been proposed to introduce complementary spectral information into the variational framework. Here, we employ two representative cross-channel regularizers to achieve the HS dead-pixel inpainting. For the dead-pixel inpainting of HSIs, we yield two expressions using Eq. (7).

First, selecting the low-rank regularizer, the optimization model[98] is expressed as

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{X}} \|\mathbf{M}\mathbf{X} - \mathbf{Y}\|_2^2 + \mu \sum_{i=1}^{3} w_i \|\mathbf{X}_{(i)}\|_*, \tag{21}$$
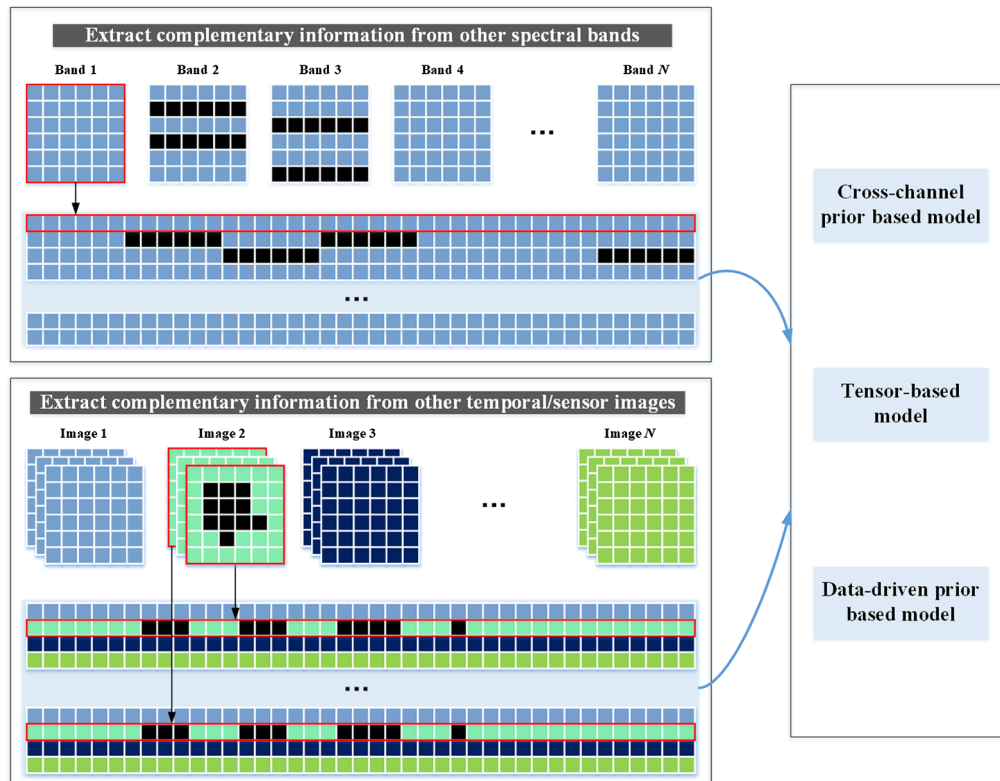
Fig. 10 Missing information reconstruction using other spectral bands or other temporal/sensor images.

where the mask $\mathbf{M}$ comprises 0 and 1, with 0 representing the missing pixels in each band. In Eq. (21), the second term stands for the rank of an HSI $\mathbf{X}$, in which the unfoldings $\mathbf{X}$ with respect to the different dimensions are $\mathbf{X}_{(1)}$, $\mathbf{X}_{(2)}$, and $\mathbf{X}_{(3)}$. w_i is the weight corresponding to the i'th dimension unfolding. Notably, the common way is the equal weights (i.e., 1/3), before the adaptive weighted tensor is derived from the distribution of the singular values of the unfolding matrixes.[98]

Second, considering sparse representation reconstruction, the optimization model[96] is

$$\{\{\hat{\alpha}_k\}_{k=1}, \hat{\mathbf{X}}\} = \arg\min_{\alpha_k, \mathbf{X}} \|\mathbf{MX} - \mathbf{Y}\|_2^2 + \lambda \sum_k \|\mathbf{D}\alpha_k - \mathbf{P}_k\mathbf{X}\|_2^2 + \sum_k \mu_k \|\alpha_k\|_0. \qquad (22)$$

For the framework of sparse representation in Eq. (22), the last two terms are image priors that enable every image patch $\mathbf{P}_k\mathbf{X}$ to have a sparse representation with limited error by introducing the redundant dictionary $\mathbf{D}$, its corresponding coefficients α_k, and the operator $\mathbf{P}_k$ that extracts the k'th patch $\mathbf{x}_k$ of image. Furthermore, the last penalty term is actually used to meet the condition $\|\mathbf{D}\alpha_k - \mathbf{P}_k\mathbf{X}\|_2^2 \le cn_k\sigma^2$ in the corresponding position. Specifically, the second term ensures that the image inpainting effectively extracts available information from other spectral bands by means of spatial–spectral patches. c is a constant that is larger than the maximum eigenvalue of $\mathbf{DD}^T$, n_k counts the existing pixels in this patch, and σ^2 indicates the variance of an additive Gaussian noise. Then, the dictionary should be updated according to the dictionary learning[51] so that it is appropriate for the given image.

4.3.2 *Extract complementary information from other temporal/sensor images large area missing information reconstruction*

In some cases, all bands of the acquired HSI are contaminated the thick clouds. Although we can use IR methods based on single HSI to solve this inverse problem, the result is far from

satisfactory because of the lack of sufficient complementary information. As a result of the scanning deviation of different sensors at different times, data over the same geographical region and acquired at different periods can provide supplementary information. To integrate the complementary information from other temporal/sensor images, the missing information reconstruction can be used to fill the large area missing information caused by thick clouds.

Most of the temporal-based methods[99,100] for reconstructing missing information attempt to build a clear functional relationship (linear or nonlinear) between the corrupted data and the reference data in the temporal domain. In recent years, attempts have been made to establish an unknown relationship using the approach of temporal learning under the perspective of compressed sensing (CS) or sparse representation. The representative methods for multiple bands were proposed by Lorenzi et al.[101] and Li et al.[102] Lorenzi et al. proposed to obtain the CS solution through formulation within a genetic optimization scheme, and Li et al. considered it as a multitemporal dictionary learning issue. Later, to more naturally fill a large area missing information, Li et al.[103] also proposed a patch matching-based multitemporal group sparse representation (PM-MTGSR) by utilizing the local correlations in the temporal domain and the nonlocal correlations in the spatial domain, in which spatial, spectral, and temporal information are joined simultaneously. Different from the classic sparse representation in Eq. (22), group sparse representation is given by joining the similar patches and the target patch

$$\{\hat{\alpha}_{G_k}, \hat{\mathbf{D}}_{G_k}\} = \arg \min_{\alpha_{G_k}, \mathbf{D}_{G_k}} \sum_k \|\mathbf{M}_{G_k} \otimes (\mathbf{D}_{G_k} \alpha_{G_k} - \mathbf{x}_{G_k})\|_2^2 + \sum_k \mu_k \|\alpha_{G_k}\|_0, \qquad (23)$$

where $\mathbf{M}_{G_k}$ is the mask of $\mathbf{x}_{G_k}$ and is extracted from $\mathbf{M}$ of entire image; α_{G_k} is the group sparse coefficient of group $\mathbf{x}_{G_k}$ containing the target patch and its similar patches; $\mathbf{D}_{G_k}$ is the group dictionary; $\otimes$ represents the pointwise product of the two multipliers.

Using image permutation in Fig. 10, the above methods can effectively reconstruct the missing information for MS image. However, hundreds of bands of HSI are difficult to be processed simultaneously. Hence, for each target band image covered by large area clouds, the same band images and highly correlated bands in different periods can be matched. Finally, the target band image and its matched band images are input into VM to achieve the result of target band image after the iterative optimization.

4.3.3 *Experimental evaluation*

Dead-pixel inpainting. The above two types of methods, sparse representation-based[96] and low-rank matrix-based reconstructions,[98] are compared in this section. An experiment was conducted with MODIS reflectance L1B 500-m resolution products, which were directly downloaded from the NASA website.[104] Fifteen of all the 20 detectors in Aqua MODIS band 6 are either nonfunctional or noisy,[97] thereby resulting in large-area dead or noisy pixels [Fig. 11(a)]. At first, we made a mask to label the locations of the missing pixels, then the complementary information from other spectral bands was fully used to recover the missing pixels based on the above two methods. In sparse representation-based reconstruction, the regularization parameters and the dictionary size were set by Ref. 96. In low-rank matrix-based reconstruction, three weights can be adaptively calculated and affect the final reconstruction result, compared with the parameter μ with a value of 1. The comparison of the Aqua MODIS band 6 images before and after inpainting is shown in Figs. 11(a) and 11(b). Since the matrix rank is generally recognized as a rational sparsity measure for matrix, both sparsity and low-rank measure are able to reflect global correlation along the spectral dimension. The result indicates that they are effective for reconstructing the large-scale dead pixels while retaining the edge and texture features well. However, the result using sparsity measure appeared some residual noises while retaining more detailed features. This phenomenon is due to the relatively few redundancies only in one dimension in model Eq. (22) in comparison with the intrinsic characteristics underlying three unfolded dimensions of HSI in model [Eq. (21)].

Large area missing information reconstruction. Owing to the lack of effective variational means of HS cloud removal, considering that the nonlocal self-similarity and highly
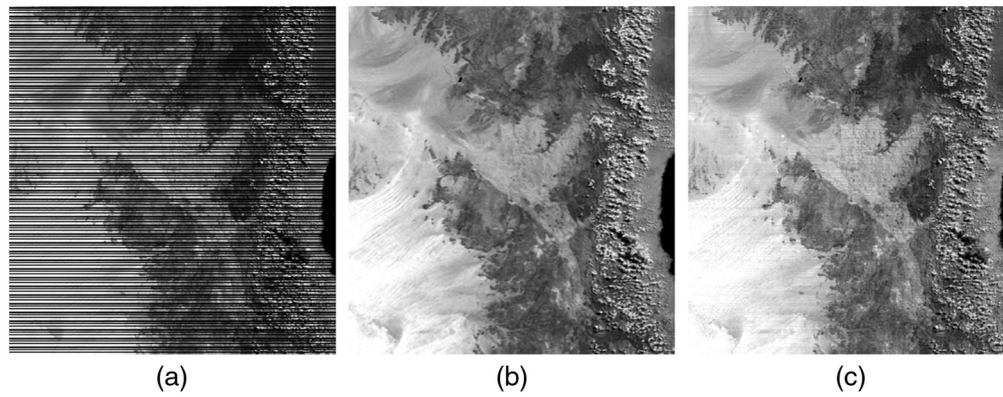
(a)   (b)   (c)

Fig. 11 The Aqua MODIS data recovery (acquired on January 18, 2009), with a size of 400×400. (a) Original Aqua MODIS band 6. (b) Recovery result using low-rank matrix-based reconstruction.[98] (c) Recovery result using sparse representation-based reconstruction.[96]
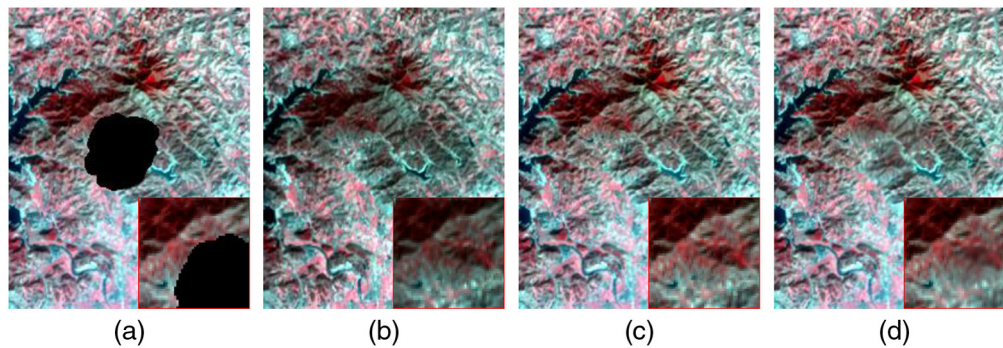

(a)   (b)   (c)   (d)

Fig. 12 Simulated experiment for missing information reconstruction with the band combination of 37, 27 and 17. (a) The Hyperion data in January 2004, with a simulated missing area. (b) The reference Hyperion data in February 2004. (c) The real Hyperion data in January 2004. (d) The reconstruction result in January 2004.

spectral correlation are important for the selection of effectively complementary information, each target band image and its matched band images were put into PM-MTGSR to show the performance of thick cloud removal on Hyperion data with a simulated missing area on all bands. In this experiment, only one band with same spectral range was selected for each target band image. All parameters related to PM-MTGSR were in accord with the values in Ref. 103. The original image is acquired at January 2004, and a reference image acquired at February 2004, is used to provide temporal complementary information. The pair of images is of Hubei Province, China, with the size of $200 \times 250 \times 50$. Due to serious noise interference on Hyperion data, some denoising methods[46,105] were used to remove the mixed noise. The denoising processing may reduce spatial details of image. The reconstruction result is shown in Fig. 12. It can be seen that the reconstruction result is satisfactory due to remarkable spatial and spectral characteristic retention, although radiation resolution differences between two temporal images may lead to the loss of spatial detailed information. If higher resolution data exist, then reconstruction and fusion can be jointly used to further enhance the resolution of HSI.

4.4 HSI and PAN/MSI Fusion

As known, HSI and PAN/MSI fusion technique is used to acquire the high-spatial resolution HSI, which is difficult to acquire directly by sensors because of physical and financial limitations. Compared with PAN sharpening for MSI, extracting complementary information from other sensor images to enhance spatial resolution of HSI is a more sophisticated processing while
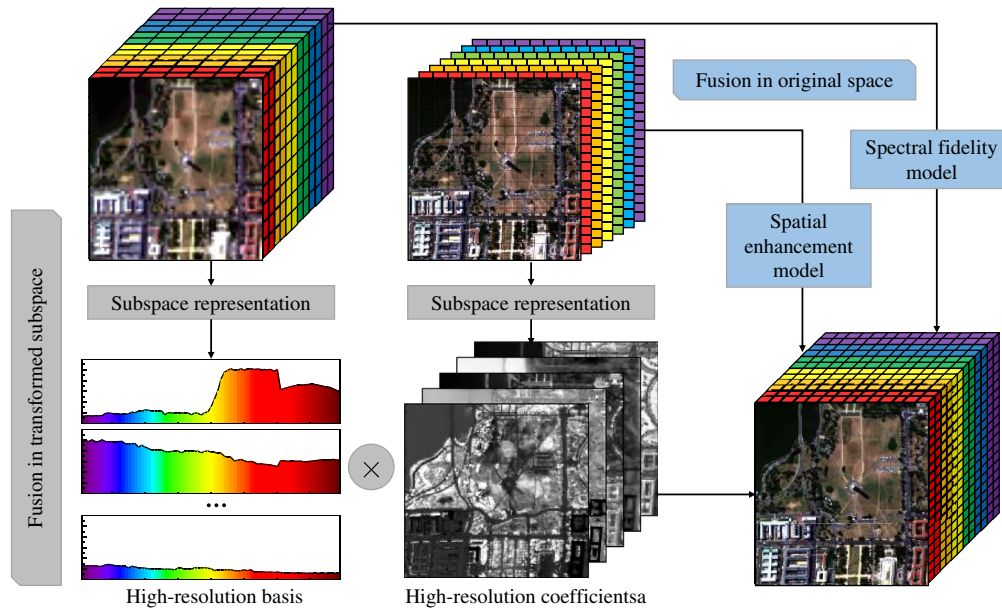
Fig. 13 Two basic variational-based image fusion processing.

keeping spectral features unchanged. For HSI fusion, a registered PAN image or an MS image is often used as the auxiliary data to break through the limitations of the sensor properties on spatial, spectral, and temporal resolutions.

4.4.1 *Space-field method and transform-field method using priors*

The universal framework for fusion has been shown in Eqs. (5) and (6), which can be regarded as Bayesian framework and matrix factorization, respectively. For the Bayesian framework-based method in the original space, methods[106–109] use posterior distribution to estimate the required image according to the image prior. According to the characteristics of circulation and down-sampling matrix in image fusion, the closed-form solution of Sylvester equation can be obtained by integrating prior information of image.[107] Under the Bayesian framework, some spatiospectral joint priors[108] are applied to regularize the solution of the HRHS image. Furthermore, matrix factorization is used to decompose image into some basis (or a set of spectral signatures) and coefficient matrices, by defining as $\mathbf{X} = \mathbf{D}\boldsymbol{\alpha}$. $\boldsymbol{\alpha}$ can be regarded as a component in a transformed subspace. As shown in Fig. 13, the idea of fusing the HS–MS images based on the spectral information of both input images on a subspace has been the main source of inspiration.[110–113] Using geometrical considerations devoted to the hybrid surface,[3] several unmixing-based methods[112,113] have been proposed for HS-MS fusion, which estimate high spatial resolution HSI by sharpening the abundance maps on the endmember feature subspace, resulting in the state-of-the-art fusion performance under the constraints of relative sensor characteristics, such as the coupled non-negative matrix factorization (CNMF).[112]

4.4.2 *Fusion based on tensor-based model*

Currently, the tensor factorization-based HSI fusion methods are few in number. Here, we introduce one primary and state-of-the-art example along this line of research. A coupled sparse tensor factorization (CSTF)[58]-based approach is proposed to fuse HRMS and LRHS images. For the tensor-based model, the optimization problem can be formulated as the following constrained least-squares problem:

$$\mathcal{X} = \arg\ \min \|\mathcal{Y} - \mathcal{C} \times_1 (\mathbf{P}_1\mathbf{W}) \times_2 (\mathbf{P}_2\mathbf{H}) \times_3 (\mathbf{P}_3\mathbf{S})\|_p^p + \|L(\mathcal{C})\|_q^q$$
$$+ \|\mathcal{Z} - \mathcal{C} \times_1 \mathbf{W} \times_2 \mathbf{H} \times_3 (\mathbf{P}_3\mathbf{S})\|_r^r, \tag{24}$$

where the first term is a data-fitting term, imposing that the target HRHS image $\mathcal{X}$ should explain the observed image $\mathcal{Y}$ according to the relationship model defined in Eq. (14). $L(\cdot)$ is the constraint function for prior knowledge of $\mathcal{C}$. $\|\mathcal{Z} - \mathcal{C} \times_1 \mathbf{W} \times_2 \mathbf{H} \times_3 (\mathbf{P}_3\mathbf{S})\|_r^r$ is the function that defines the relationship between $\mathcal{X}$ and other available image $\mathcal{Z}$.

4.4.3 *Fusion based on data-driven prior*

As for data-driven prior, the recent advances[114–117] are gradually found to solve remote sensing image fusion. Dian[114] et al. proposed to incorporate the priors learned by deep CNN-based residual learning into the fusion framework. By two alternations between residual learning and model optimization, the speed up of training is accelerated, and the fusion performance is evidently boosted. Instead of learning regularizer on all pixels of the image, the gradient or spatial detail feature priors learned by deep residual gradient CNN is exploited to construct the spatial fidelity constraint rather than the denoiser prior.[115] By combining DLVM, Shen et al.[115] also proved that in the spatial enhancement terms, the learned gradient consistency prior, which directly represents the spatial information, can obtain better results than the learned image consistency prior. In addition, Xie et al.[117] proposed MHF-net that unfolded the algorithm into an optimization-inspired deep network for MSI/HSI fusion and obtained promising results. In general, such a promising combination actually offers huge flexibility. Hence, the pretrained regularizers exploiting different complementary information can be jointly utilized to solve one specific problem.

Take DLVM that is the first proposed combination method as an example, HSI is divided into different groups according to the spectral range of HSI. We selected one band from each group to compose an LR MSI and then introduced the LR MSI into a pretrained deep residual gradient netwok along with PAN to achieve an initial gradient output, as shown in Fig. 14. The energy function constructed in DLVM can be written as

$$\hat{\boldsymbol{X}} = \arg \min_{\boldsymbol{X}}\|\boldsymbol{Y} - \boldsymbol{A}\boldsymbol{X}\|_F^2 + \lambda \sum_{j=1}^{2} \|\nabla_j \boldsymbol{X} - \boldsymbol{G}_j\|_F^2 + \mu p(\boldsymbol{X}), \qquad (25)$$

where the second term is a spatial enhancer, representing the gradients of the estimated image in the horizontal and vertical directions should be consistent with gradient images $\boldsymbol{G}$ of the HR-MS image obtained through the pretrained network. In DLVM, the regularization in the third term can be a physical prior, such as Laplacian prior adopted in the following experiment, or a pretrained denoiser by CNN. More details of this method can be found in Ref. 115.
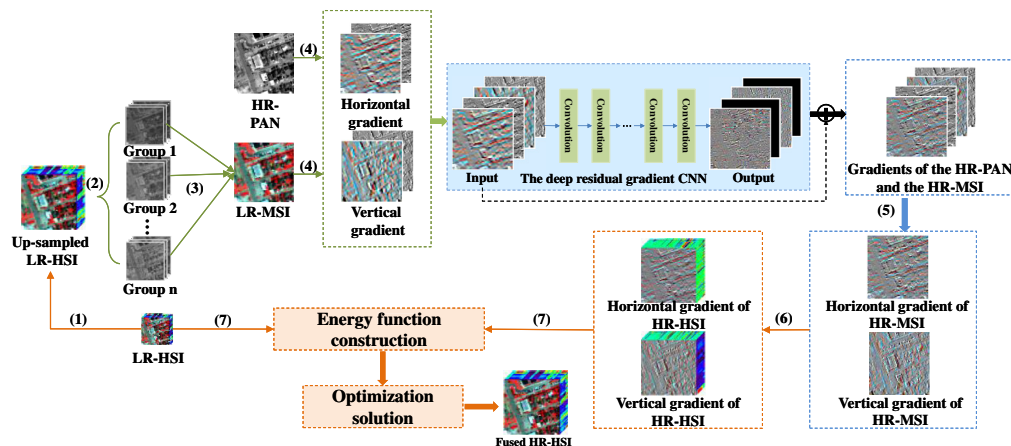

Fig. 14 HSI fusion framework of DLVM with steps (1) to (7).
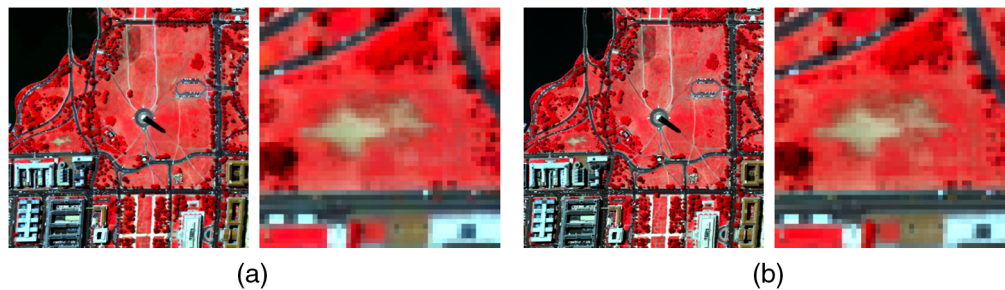
(a)   (b)

Fig. 15 Results (57, 27, and 17 bands) with 4× magnification from two types of fusions in the simulated experiments, derived from the basic function (5) and (6). (a) Fusion result based on the space-field method[109] (PSNR = 29.6632, SSIM = 0.9268, SAM = 6.3686, CC = 0.9908, ERGAS = 2.7727). (b) Fusion result based on transform-field[110] (PSNR = 34.2935, SSIM = 0.9784, SAM = 4.1479, CC = 0.9972, ERGAS = 1.381).

4.4.4 Experimental evaluation

Comparison between space-field method and transform-field method. To further discuss the effectiveness of two kinds of HS-MS fusion algorithms, the methods that are proposed by Shen et al.[109] with simple 3D-Laplace regularizer and Simões et al. (HySure)[110] with TV regularizer, one of the representative and basic algorithms based on Eqs. (5) and (6), have been applied on the HSI in Fig. 15. The parameters of all methods have been tuned to the optimal, according to their references. HySure are available at Ref. 118. In the simulated experiments, the Washington DC Mall dataset with the cropped size of $288 \times 288 \times 79$ is used to obtain the LRHS images of $288 \times 288 \times 5$ size and HRMS images of $72 \times 72 \times 5$. The simulated LRHS image with the spectral range of 450 to 1750 nm was obtained by low-pass filtering and downsampling by a factor of four in the spatial domain. The HRMS image was produced according to the spectral characteristics of Landsat-7 ETM+ bands 1 to 5. As described in Ref. 109, HRMS images were generated by filtering the ground-truth images along the spectral dimension using the reflectance spectral responses, such as the IKONOS. For fusion experiments, PSNR, SSIM, MSA, erreur relative globale adimensionnelle desynthèse (ERGAS), and correlation coefficient (CC) often serve as evaluation indices. The results shown in Fig. 15 suggest that the visual effects of HS-MS fusion in the original space-field and the transform-field are not obviously different. However, after transformation, the transform-field methods can inject high-resolution features and reserve spectral information by effectively removing redundant spatial information. Thus, when the regularizers of two models are simple, the fusion in the transformed subspace may provide higher precision and show better spectral preservation.

Fusion comparison on different number of inputs. In addition, the flexibility of variational methods makes it possible to fuse multiple images that involve more sensors. In this experiment, a method by Shen et al.[109] with simple 3D-Laplace regularizer was chose. An auxiliary PAN image with the size of $288 \times 288 \times 1$ is also created by the spectral characteristics of a Landsat-7 PAN image covering 520 to 900 nm. Different from the experiment in Fig. 15, in the spatial dimension, the simulated MS image was downsampled by a factor of two, and its size was $144 \times 144 \times 5$. From Fig. 16, we can observe that for HSI fusion, more auxiliary images can be used to better improve the final results by introducing more details and adjusting the spectral features in Fig. 16(b). Relatively speaking, the PAN/MS/HS fusion result has slightly better spatial details than the PAN/HS fusion result because of the incorporation of the MS images. From a spectral perspective, the PAN/MS/HS fusion and the MS/HS fusion can effectively recover more spectral characteristics, because the PAN/MS/HS fusion can make full use of the complementary spectral information of the PAN and MS groups.

Effectiveness of tensor-based method. To demonstrate the differences between the fusion results of the matrix-based and tensor-based methods, the experiments were conducted with the University of Pavia image acquired by the reflective optics system imaging spectrometer optical sensor over the urban area of the Pavia University, Italy. The size of the HRHS image is
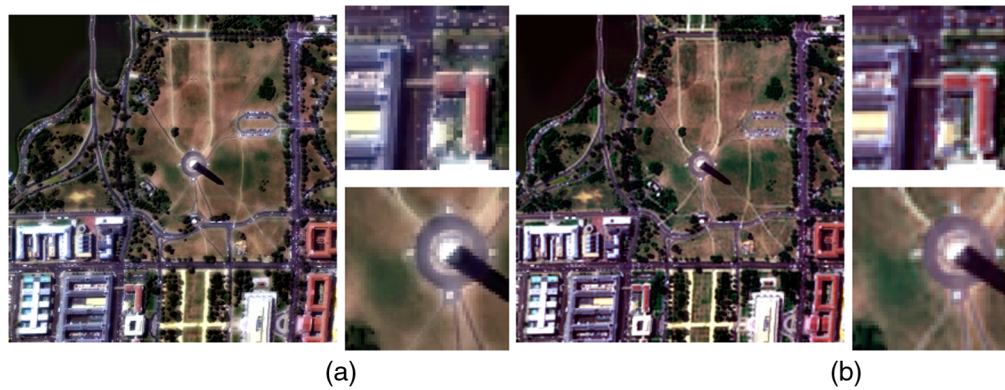
(a)　　　　　　　　　　　　　　　　(b)

Fig. 16 Fusion experiments with 4× magnification using the different number of images. (a) The results (36, 23, and 8 bands) from pan and HS image. (b) The results (36, 23, and 8 bands) from pan, MS, and HS images.

Table 1 Quantitative results of the test methods on the Pavia University.

	PSNR	SSIM	SAM	CC	ERGAS
HySure[110]	38.8075	0.9779	3.019	0.9932	0.8594
CSTF[58]	42.2704	0.9836	2.207	0.9964	0.5941

$256 \times 256 \times 93$. LRHS image is generated by bicubic downsampling HR-HSI, whose size is $32 \times 32 \times 93$. HRMS image is synthetic from HRHS image through the spectral response function of the IKONOS sensor, whose size is $256 \times 256 \times 4$. A state-of-the-art method called CSTF (available at the Github repository: https://github.com/renweidian/CSTF) was used in this experiment, and its parameters were set by Ref. 58. Table 1 shows the average objective results that contain PSNR, SSIM, SAM, CC, and ERGAS. In Table 1 and Fig. 17, the evaluation indices of the CSTF result are better than HySure in terms of both spatial enhancement and spectral fidelity. Consistent with the quantitative results, from the magnified image region, the CSTF method provides clearer and sharper spatial details while HySure produces fewer spectral distortions. The reason is that the tensor-based approach can deeply capture these relationships between HR and LR information from three different dimensions underlying an HSI.

Effectiveness of data-driven prior-based method.　To test the validity of the proposed pretrained priors, HS data downloaded from 2018 IEEE hyperspectral GRSS Data Fusion Contest were used to achieve the PAN/HS fusion. PAN and LRMS images are generated from HRHS image with the cropped size of $200 \times 200 \times 33$. Spectrally, PAN is synthetic from the HRHS image through the spectral response function of the IKONOS sensor, whose size is $200 \times 200 \times 1$. LRHS image is generated by bicubic downsampling HRHS image, whose size
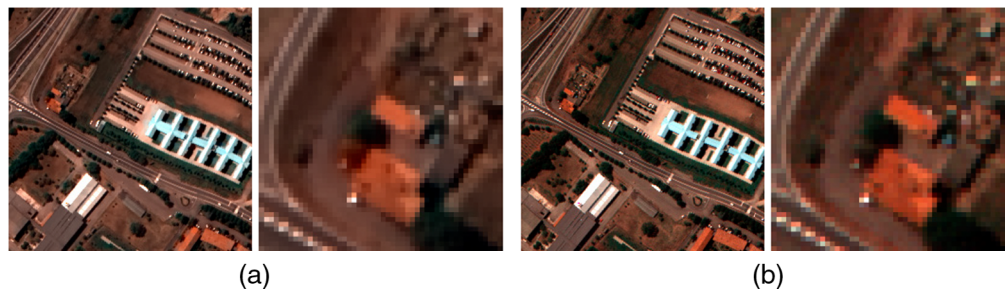

(a)　　　　　　　　　　　　　　　　(b)

Fig. 17 Fusion experiments with 8× magnification on the Pavia University. (a) The matrix-based using HySure.[109] (b) The tensor-based methods using CSTF.[58]

Table 2 Quantitative results of the test methods on 2018 IEEE Contest Dataset.

Method	PSNR	SSIM	SAM	CC	ERGAS
3D-Laplace VM[109]	37.761	0.9556	3.2688	0.9814	2.3858
CNMF[58]	39.481	0.9602	3.2295	0.9834	2.2232
DLVM[115]	43.669	0.9842	1.981	0.9946	1.2505

(a) Original image   (b) 3D-Laplace VM[107]   (c) CNMF[110]   (d) DLVM[113]
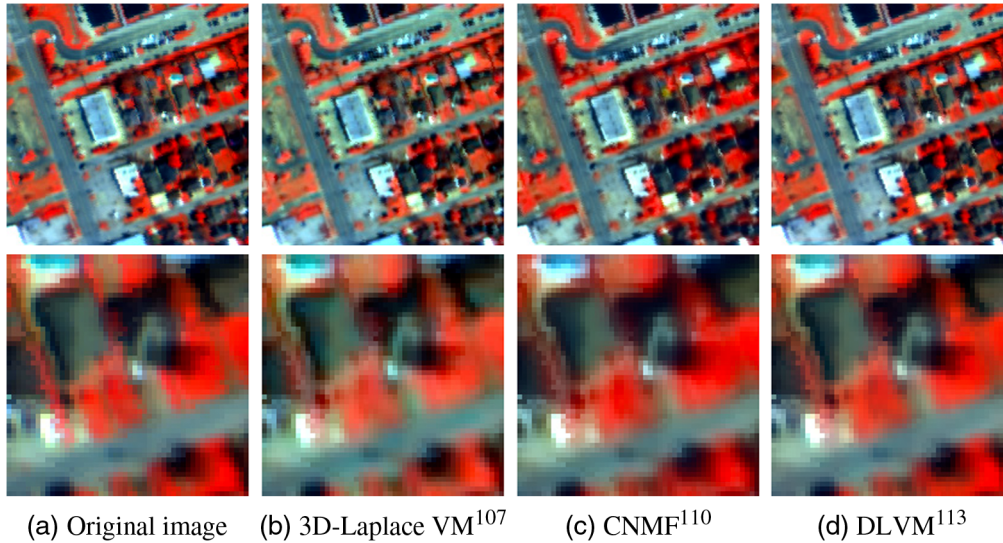
Fig. 18 Experiments with 4× magnification in spatiospectral fusion (displayed as false color with 30, 15, and 3).

is $50 \times 50 \times 33$. To deal with HSI, a state-of-the-art method called DLVM was employed group-by-group in Fig. 14. In this experiment, the regularization parameters for the second and third terms are set to 10 and 1, respectively. These deep learning regularizers must be trained again to adapt to different HSIs. In this experiment, different prior models with 3D-Laplace[109] (cross-channel prior), CNMF[58] (tensor-based model), and DLVM (data-driven prior) were compared to test and verify the fusion effectiveness. Parameters of all the compared methods are adjusted to the optimal. The CNMF code is available at Ref. 119. As shown in Table 2, the combined model DLVM proposed by Shen[115] achieved a satisfying performance. From a visual perspective, DLVM can inject more spatial details and keep better spectral information preservation than the two other fusion cases, especially for the red vegetation area in Fig. 17. However, insufficient samples or poor network structure will greatly affect the accuracy of the results. Therefore, discovering different network architectures is necessary to obtain better and robust performance.

5 Current Advantages and Future Challenges

5.1 Advantages of Variational Framework

Variational framework has proven its power in radiometric quality improvement of HSIs. In this section, we give two main advantages of variational framework.

5.1.1 Variational framework has good compatibility for different problems

On the one hand, different degradations can be integrated into a general reconstruction model based on variational framework. From denoising to reconstruction, variational framework can exploit different degradation operators but same prior information to describe the input–output relationship and the spatiospectral properties. On the other hand, different from deep

learning-based methods that require a retraining for different images, multiple images from more than two sensors can be simultaneously supplied into a VM to achieve the spatio–temporal–spectral complementary information. A clear observation on the HS fusion also proves that more auxiliary images can improve the final results better.

5.1.2 Flexible priors can be put into the variational framework

Variational framework can flexibly embed the most suitable prior according to different degradation problems. These regularizations including cross-channel, tensor, and deep priors are convenient to combine other spectral/temporal/sensor images to obtain available information and optimize the final results by multiple iterations. Furthermore, using the variational framework, some existent methods can be directly introduced as an off-the-shelf prior or denoiser. Take the PNP concept as an example, the energy function with associated different degradation problems can be unrolled by variable splitting technique into some subproblem, and then the prior related subproblem can be replaced by any off-the-shelf denoiser. Remarkably, the denoiser learned by DNN gives rise to promising performance.

5.2 Promising Future Directions for Variational Approach

Despite rapid development and achieved promising progress of variational approach, there are still many open issues for the future work. In the section, we suggest three future interesting topics in HS data quality improvement.

5.2.1 Extension to rarely involved application

The research for haze and missing information problems on HSIs, such as clouds and cloud shadows, is rare. Especially, the simultaneous coverage of clouds on all spectral bands arises the difficulties. Furthermore, for temporal complementary information, it is also necessary to consider the temporal–spatial change and maintain the authenticity of the results. From the perspective of technology, tensor decomposition with excellent preservation for 3D characteristics and data-driven prior with great representation for deep features give a prominent capacity for HSI. However, the combination of tensor and deep learning in a variational framework is still few and requires more research in subsequent HS data processing.

5.2.2 Data-driven prior using insufficient training samples

The technique composed of the degradation model and the parallel deep network is effective for enhancing the spatial resolution and maintaining the spectral feature. However, the insufficiency of and the big difference between HSI training samples contribute greatly to the difficulty in robust learning and credible representation. Although the variational framework based on the degradation model can remedy the deficiency, further research with better optimization should be developed. Unsupervised learning guided by the variational framework will be a valuable direction in the future.

5.2.3 Integrated framework for multiple applications

HSIs are often simultaneously corrupted by multiple degradation factors. As introduced in this review, current methods mainly aim at improving the quality caused by one degradation factor. Although removal of mixed noises, denoising and fusion,[108] and cloud removal and fusion[120] have been proposed in the recent years, they only involve two degradations. Hence, the problems that are mixed by more degradation factors are still crucial and important. In addition, quality improvement can reduce the error of subsequent applications, by eliminating the impact of noises, dead pixels, and partial coverage of clouds. Therefore, an integrated approach containing different degradation removal and their subsequent applications on HSI is a promising future direction, such as denoising and HS mixing,[121] super-resolution and classification,[122] and registration and fusion.[123]

6 Concluding Remarks

Due to its powerful compatibility and flexibility in dealing with different degradations, the variational framework has been a research hotspot. In this paper, we have systematically reviewed the variational framework techniques for HS data, which can link different degradations as a whole by a general reconstruction model. The review starts on generic model of radiometric quality improvement, which can provide a basic architecture for each related degradation tasks. Then, four main models, namely, single-channel prior-based model, cross-channel prior-based model, tensor-based model, and data-driven prior-based model, are also briefly reviewed for HSI quality improvement. Single-channel prior-based model only utilizes information from single band itself. However, cross-channel prior-based model, tensor-based model, and data-driven prior-based model can fully exploit and excavate spatiospectral complementary information from other spectral bands and temporal/sensor images to solve various image degradations. Specifically, available information from spectral bands is sufficient to remove noises, dead pixels, and hazes, while large missing areas reconstruction and image fusion more depend on additional information from other temporal/sensor images. For each specific problem, we introduce corresponding representative methods that utilize spatiospectral priors (cross-channel prior, tensor-based prior, and data-driven prior) and conduct some experiments to demonstrate the effectiveness of variational framework. Finally, to gain a thorough understanding, we summarize the current advantages and propose several promising future directions based on limitations of current variational framework. On the one hand, some applications still need to be deeply explored, such as hazes, clouds, and cloud shadows. To improve the efficiency and accuracy, the integrated approach for multiple applications also requires further research. On the other hand, exploring proper data-driven priors based on an optimization-inspired variational mode for more complex problems also remains a big challenge.

Acknowledgments

References

1. D. Landgrebe, "Hyperspectral image data analysis," *IEEE Signal Process. Mag.* **19**(1), 17–28 (2002).
2. G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.* **43**(6), 1351–1362 (2005).
3. J. M. Bioucas-Dias et al., "Hyperspectral unmixing overview: geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **5**(2), 354–379 (2012).
4. N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.* **19**(1), 44–57 (2002).
5. T. Kasetkasem, M. K. Arora, and P. K. Varshney, "Super-resolution land cover mapping using a Markov random field based approach," *Remote Sens. Environ.* **96**(3-4), 302–314 (2005).
6. J. Verhoeye and R. De Wulf, "Land cover mapping at sub-pixel scales using linear optimization techniques," *Remote Sens. Environ.* **79**(1), 96–104 (2002).
7. P. M. Atkinson, M. E. J. Cutler, and H. Lewis, "Mapping sub-pixel proportional land cover with AVHRR imagery," *Int. J. Remote Sens.* **18**(4), 917–935 (1997).
8. D. W. J. Stein et al., "Anomaly detection from hyperspectral imagery," *IEEE Signal Process. Mag.* **19**(1), 58–69 (2002).
9. D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Process. Mag.* **19**(1), 29–43 (2002).
10. M. A. O. Vasilescu and D. Terzopoulos, "Multilinear (tensor) image synthesis, analysis, and recognition," *IEEE Signal Process. Mag.* **24**(6), 118–123 (2007).

11. L. De Lathauwer, *Signal Processing Based on Multilinear Algebra*, Katholieke Universiteit Leuven (1997).
12. T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.* **51**(3), 455–500 (2009).
13. H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "A survey of multilinear subspace learning for tensor data," *Pattern Recognit.* **44**(7), 1540–1551 (2011).
14. D. Tao et al., "General tensor discriminant analysis and gabor features for gait recognition," *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(10), 1700–1715 (2007).
15. F. Li et al., "Hyperspectral image segmentation, deblurring, and spectral analysis for material identification," *Proc. SPIE* **7701**, 770103 (2010).
16. D. Bertaccini et al., "An adaptive norm algorithm for image restoration," *Lect. Notes Comput. Sci.* **6667**, 194–205 (2011).
17. E. S. Lee and M. G. Kang, "Regularized adaptive high-resolution image reconstruction considering inaccurate subpixel registration," *IEEE Trans. Image Process.* **12**(7), 826–837 (2003).
18. H. Zhang et al., "Image and video restorations via nonlocal kernel regression," *IEEE Trans. Cybern.* **43**(3), 1035–1046 (2013).
19. Y. Dong, M. Hintermüller, and M. Neri, "A primal-dual method for L1 TV image denoising," *SIAM J. Imaging Sci.* **2**(4), 1168–1189 (2009).
20. R. H. Chan et al., "An efficient two-phase L1-TV method for restoring blurred images with impulse noise," *IEEE Trans. Image Process.* **19**(7), 1731–1739 (2010).
21. H. Shen et al., "Super resolution reconstruction algorithm to MODIS remote sensing images," *Comput. J.* **52**(1), 90–100 (2009).
22. H. Fu et al., "Efficient minimization methods of mixed l1-l1 and l2-l1 norms for image restoration," *SIAM J. Sci. Comput.* **27**(6), 1881–1902 (2006).
23. S. Huihui, Z. Lei, and W. Peikang, "An adaptive l1-l2 hybrid error model to super-resolution," in *Proc. IEEE Int. Conf. Image Process.* (2010).
24. H. Shen et al., "Adaptive norm selection for regularized image restoration and super-resolution," *IEEE Trans. Cybern.* **46**(6), 1388–1399 (2016).
25. M. Elad, *Sparse Redundant Representations: From Theory to Applications in Signal Image Processing*, Springer, Berlin (2010).
26. E. Candes, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.* **59**(8), 1207–1223 (2006).
27. C. Jiang et al., "A practical compressed sensing-based pan-sharpening method," *IEEE Geosci. Remote Sens. Lett.* **9**(4), 629–633 (2012).
28. M. T. Eismann and R. C. Hardie, "Hyperspectral resolution enhancement using high-resolution multispectral imagery with arbitrary response functions," *IEEE Trans. Geosci. Remote Sens.* **43**(3), 455–465 (2005).
29. Y. Zhang, A. Duijster, and P. Scheunders, "A Bayesian restoration approach for hyperspectral images," *IEEE Trans. Geosci. Remote Sens.* **50**(9), 3453–3462 (2012).
30. A. N. Tikhonov and V. Y. Arsenin, "Solutions of ill-posed problems," *Math. Comput.* **32**(144), 491–491 (1977).
31. L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D* **60**(1), 259–268 (1992).
32. S. Farsiu et al., "Fast and robust multiframe super resolution," *IEEE Trans. Image Process.* **13**(10), 1327–1344 (2004).
33. Y. Lou et al., "Image recovery via nonlocal operators," *J. Sci. Comput.* **42**(2), 185–197 (2010).
34. W. Dong et al., "Sparsity-based image denoising via dictionary learning and structural clustering," in *Proc. IEEE Comput. Soc. Conf. Comput. Vision Pattern Recognit.*, pp. 457–464 (2011).
35. K. Dabov et al., "Image denoising by sparse 3D transform-domain collaborative filtering," *IEEE Trans. Image Process.* **16**(8), 2080–2095 (2011).
36. A. Danielyan et al., "BM3D frames and variational image deblurring," *IEEE Trans. Image Process.* **21**(4), 1715–1728 (2012).

37. M. Elad and M. Aharon, "Image denoising via sparse redundant representations over learned dictionaries," *IEEE Trans. Image Process.* **15**(12), 3736–3745 (2006).

38. M. Elad and A. Feuer, "Restoration of a single super-resolution image from several blurred, noisy down-sampled measured images," *IEEE Trans. Image Process.* **6**(12), 1646–1658 (1997).

39. J. Mairal et al., "Non-local sparse models for image restoration," in *Proc. IEEE Int. Conf. Comput. Vision*, pp. 2272–2279 (2009).

40. M. Fazel, "Matrix rank minimization with applications," PhD Dissertation, Dept. Elect. Eng., Stanford Univ., Stanford, California (2002).

41. S. Gu et al., "Weighted nuclear norm minimization and its applications to low level vision," *Int. J. Comput. Vis.* **121**(2), 183–208 (2017).

42. F. Shang et al., "Bilinear factor matrix norm minimization for robust PCA: algorithms and applications," *IEEE Trans. Pattern Anal. Mach. Intell.* **40**(9), 2066–2080 (2018).

43. H. Wang et al., "Reweighted low-rank matrix analysis with structural smoothness for image denoising," *IEEE Trans. Image Process.* **27**(4), 1777–1792 (2018).

44. Z. Zha et al., "From rank estimation to rank approximation: rank residual constraint for image restoration," *IEEE Trans. Image Process.* **29**, 3254–3269 (2020).

45. Y. Wang et al., "Hyperspectral image restoration via total variation regularized low-rank tensor decomposition," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **11**(4), 1227–1243 (2018).

46. H. Zhang et al., "Hyperspectral image denoising with total variation regularization and nonlocal low-rank tensor decomposition," *IEEE Trans. Geosci. Remote Sens.* **58**(5), 3071–3084 (2020).

47. Q. Yuan, L. Zhang, and H. Shen, "Hyperspectral image denoising employing a spectral–spatial adaptive total variation model," *IEEE Trans. Geosci. Remote Sens.* **50**(10), 3660–3677 (2012).

48. X. Zhang et al., "Bregmanized nonlocal regularization for deconvolution and sparse reconstruction," *SIAM J. Imaging Sci.* **3**(3), 253–276 (2010).

49. Q. Cheng, H. Shen, and L. Zhang, "Inpainting for remotely sensed images with a multichannel nonlocal total variation model," *IEEE Trans. Geosci. Remote Sens.* **52**(1), 175–187 (2014).

50. J. Li et al., "Hyperspectral image recovery employing a multidimensional nonlocal total variation model," *Signal Process.* **111**, 230–248 (2015).

51. M. Aharon et al., "K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.* **54**(11), 4311–4322 (2006).

52. J. Li et al., "Noise removal from hyperspectral image with joint spectral-spatial distributed sparse representation," *IEEE Trans. Geosci. Remote Sens.* **54**(9), 5425–5439 (2016).

53. H. Zhang et al., "Hyperspectral image restoration using low-rank matrix recovery," *IEEE Trans. Geosci. Remote Sens.* **52**(8), 4729–4743 (2014).

54. Q. Xie et al., "Multispectral images denoising by intrinsic tensor sparsity regularization," in *Proc. IEEE Comput. Soc. Conf. Comput. Vision Pattern Recognit.*, pp. 1692–1700 (2016).

55. M. Zhou et al., "Tensor rank learning in CP decomposition via convolutional neural network," *Signal Process. Image Commun.* **73**, 12–21 (2019).

56. L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika* **31**(3), 279–311 (1966).

57. S. B. N. Renard and J. Blanc-Talon, "Denoising and dimensionality reduction using multilinear tools for hyperspectral images," *IEEE Geosci. Remote Sens. Lett.* **5**(2), 138–142 (2008).

58. S. Li et al., "Fusing hyperspectral and multispectral images via coupled sparse tensor factorization," *IEEE Trans. Image Process.* **27**(8), 4118–4130 (2018).

59. J. Liu et al., "Tensor completion for estimating missing values in visual data," *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(1), 208–220 (2013).

60. I. V. Oseledets, "Tensor-train decomposition," *SIAM J. Sci. Comput.* **33**(5), 2295–2317 (2011).

61. Q. Zhao et al., "Tensor ring decomposition," arXiv:1606.05535 (2016).

62. J. A. Bengua et al., "Efficient tensor completion for color image and video recovery: low-rank tensor train," *IEEE Trans. Image Process.* **26**(5), 2466–2479 (2017).

63. W. He et al., "Remote sensing image reconstruction using tensor ring completion and total variation," *IEEE Trans. Geosci. Remote Sens.* **57**(11), 8998–9009 (2019).

64. Y. Xu et al., "Hyperspectral images super-resolution via learning high-order coupled tensor ring representation," *IEEE Trans. Neural Networks Learn. Syst.* **31**(11), 4747–4760 (2020).

65. X. Zhu et al., "Deep learning in remote sensing: a comprehensive review and list of resources," *IEEE Geosci. Remote Sens. Mag.* **5**(4), 8–36 (2017).

66. C. Dong et al., "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.* **38**(2), 295–307 (2016).

67. L. Xu et al., "Deep convolutional neural network for image deconvolution," in *Proc. 27th Int. Conf. Adv. Neural Inf. Process. Syst.*, Vol. 1, pp. 1790–1798 (2014).

68. K. Zhang et al., "Learning deep CNN denoiser prior for image restoration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 2808–2817 (2017).

69. H. Zeng et al., "Hyperspectral image restoration via CNN denoiser prior regularized low-rank tensor recovery," *Comput. Vis. Image Underst.* **197-198**, 103004 (2020).

70. R. Dian et al., "Regularizing hyperspectral and multispectral image fusion by CNN denoiser," *IEEE Trans. Neural Networks Learn. Syst.* **32**(3), 1124–1135 (2020).

71. W. Dong et al., "Denoising prior driven deep neural network for image restoration," *IEEE Trans. Pattern Anal. Mach. Intell.* **41**(10), 2305–2318 (2019).

72. Y. Yang et al., "Deep ADMM-net for compressive sensing mri," *Proc. 30th Int. Conf. Neural Inf. Process. Syst.*, 10–18 (2016).

73. L. Wang et al., "Hyperspectral image reconstruction using a deep spatial-spectral prior," in *Proc. IEEE Comput. Soc. Conf. Comput. Vision Pattern Recognit.*, pp. 8032–8041 (2019).

74. H. Othman and S. E. Qian, "Noise reduction of hyperspectral imagery using hybrid spatial-spectral derivative-domain wavelet shrinkage," *IEEE Trans. Geosci. Remote Sens.* **44**(2), 397–408 (2006).

75. Y. Zhao and J. Yang, "Hyperspectral image denoising via sparse representation and low-rank constraint," *IEEE Trans. Geosci. Remote Sens.* **53**(1), 296–308 (2019).

76. T. Xie, S. Li, and B. Sun, "Hyperspectral images denoising via nonconvex regularized low-rank and sparse matrix decomposition," *IEEE Trans. Image Process.* **29**, 44–56 (2020).

77. Z. Guo, T. Wittman, and S. Osher, "L1 unmixing and its application to hyperspectral image enhancement," *Proc. SPIE* **7334**, 73341M (2009).

78. Y. Zhao et al., "Hyperspectral imagery super-resolution by sparse representation and spectral regularization," *EURASIP J. Adv. Signal Process.* **2011**, 87 (2011).

79. H. Zhang, L. Zhang, and H. Shen, "A super-resolution reconstruction algorithm for hyperspectral images," *Signal Process.* **92**(9), 2082–2096 (2012).

80. H. Huang, A. G. Christodoulou, and W. Sun, "Super-resolution hyperspectral imaging with unknown blurring by low-rank and group-sparse modeling," in *IEEE Int. Conf. Image Process.*, pp. 2155–2159 (2014).

81. J. Li et al., "Hyperspectral image super-resolution by spectral mixture analysis and spatial–spectral group sparsity," *IEEE Geosci. Remote Sens. Lett.* **13**(9), 1250–1254 (2016).

82. Y. Wang et al., "Hyperspectral image super-resolution via nonlocal low-rank tensor approximation and total variation regularization," *Remote Sens.* **9**(12), 1286 (2017).

83. B. Lin et al., "Hyperspectral image denoising via matrix factorization and deep prior regularization," *IEEE Trans. Image Process.* **29**, 565–578 (2020).

84. Q. Yuan et al., "Hyperspectral image denoising with a spatial-spectral view fusion strategy," *IEEE Trans. Geosci. Remote Sens.* **52**(5), 2314–2325 (2014).

85. X. Zhu and P. Milanfar, "Automatic parameter selection for denoising algorithms using a no-reference measure of image content," *IEEE Trans. Image Process.* **19**(12), 3116–3132 (2010).

86. B. C. Gao et al., "An algorithm using visible and 1.38-μm channels to retrieve cirrus cloud reflectances from aircraft and satellite data," *IEEE Trans. Geosci. Remote Sens.* **40**(8), 1659–1668 (2002).

87. B. C. Gao et al., "Correction of thin cirrus path radiances in the 0.4–1.0 μm spectral region using the sensitive 1.375 μm cirrus detecting channel," *J. Geophys. Res.* **103**(D24), 32169–32176 (1998).

88. B. C. Gao and R. R. Li, "Removal of thin cirrus scattering effects for remote sensing of ocean color from space," *IEEE Geosci. Remote Sens. Lett.* **9**(5), 972–976 (2012).

89. K. He et al., "Single image haze removal using dark channel prior," *IEEE Trans. Pattern Anal. Mach. Intell.* **33**(12), 2341–2353 (2010).

90. H. Li, L. Zhang, and H. Shen, "A variational gradient-based fusion method for visible and SWIR imagery," *Photogramm. Eng. Remote Sens.* **78**(9), 947–958 (2012).

91. A. Karnieli et al., "AFRI—aerosol free vegetation index," *Remote Sens. Environ.* **77**(1), 10–21 (2001).

92. I. Gladkova et al., "Quantitative restoration for MODIS band 6 on Aqua," *IEEE Trans. Geosci. Remote Sens.* **50**(6), 2409–2416 (2012).

93. Z. Xing et al., "Dictionary learning for noisy and incomplete hyperspectral images," *SIAM J. Imaging Sci.* **5**(1), 33–56 (2012).

94. Z. Mingyuan et al., "Nonparametric bayesian dictionary learning for analysis of noisy and incomplete images," *IEEE Trans. Image Process.* **21**(1), 130–144 (2012).

95. X. Li et al., "Dead pixel completion of aqua MODIS band 6 using a robust M-estimator multiregression," *IEEE Geosci. Remote Sens. Lett.* **11**(4), 768–772 (2014).

96. X. Li et al., "Sparse-based reconstruction of missing information in remote sensing images from spectral/temporal complementary information," *ISPRS J. Photogramm. Remote Sens.* **106**, 1–15 (2015).

97. H. Shen et al., "Missing information reconstruction of remote sensing data: a technical review," *IEEE Geosci. Remote Sens. Mag.* **3**(3), 61–85 (2015).

98. M. Ng et al., "An adaptive weighted tensor completion method for the recovery of remote sensing images with missing data," *IEEE Trans. Geosci. Remote Sens.* **55**(6), 3367–3381 (2017).

99. E. Helmer and B. Ruefenacht, "Cloud-free satellite image mosaics with regression trees and histogram matching," *Photogramm. Eng. Remote Sens.* **71**(9), 1079–1089 (2005).

100. C. Zeng, H. Shen, and L. Zhang, "Recovering missing pixels for Landsat ETM + SLC-off imagery using multi-temporal regression analysis and a regularization method," *Remote Sens. Environ.* **131**, 182–194 (2013).

101. L. Lorenzi, F. Melgani, and G. Mercier, "Missing-area reconstruction in multispectral images under a compressive sensing perspective," *IEEE Trans. Geosci. Remote Sens.* **51**(7), 3998–4008 (2013).

102. X. Li et al., "Recovering quantitative remote sensing products contaminated by thick clouds and shadows using multitemporal dictionary learning," *IEEE Trans. Geosci. Remote Sens.* **52**(11), 7086–7098 (2014).

103. X. Li et al., "Patch matching-based multitemporal group sparse representation for the missing information reconstruction of remote-sensing images," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **9**(8), 3629–3941 (2016).

104. http://ladsweb.nascom.nasa.gov/data/search.html.

105. X. Liu et al., "Stripe noise separation and removal in remote sensing images by consideration of the global sparsity and local variational properties," *IEEE Trans. Geosci. Remote Sens.* **54**(5), 3049–3060 (2016).

106. R. Hardie et al., "MAP estimation for hyperspectral image resolution enhancement using an auxiliary sensor," *IEEE Trans. Image Process.* **13**(9), 1174–1184 (2004).

107. Q. Wei et al., "Fast fusion of multi-band images based on solving a Sylvester equation," *IEEE Trans. Image Process.* **24**(11), 4109–4121 (2015).

108. J. Li et al., "Antinoise hyperspectral image fusion by mining tensor low-multilinear-rank and variational properties," *IEEE Trans. Geosci. Remote Sens.* **57**(10), 7832–7848 (2019).

109. H. Shen, X. Meng, and L. Zhang, "An integrated framework for the spatio-temporal-spectral fusion of remote sensing images," *IEEE Trans. Geosci. Remote Sens.* **54**(12), 7135–7148 (2016).

110. M. Simoes et al., "A convex formulation for hyperspectral image superresolution via subspace-based regularization," *IEEE Trans. Geosci. Remote Sens.* **53**(6), 3373–3388 (2015).

111. Q. Wei et al., "Hyperspectral and multispectral image fusion based on a sparse representation," *IEEE Trans. Geosci. Remote Sens.* **53**(7), 3658–3668 (2015).
112. N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled non-negative matrix factorization unmixing for hyperspectral and multispectral data fusion," *IEEE Trans. Geosci. Remote Sens.* **50**(2), 528–537 (2012).
113. C. Lanaras, E. Baltsavias, and K. Schindler, "Hyperspectral superresolution by coupled spectral unmixing," in *IEEE Int. Conf. Comput. Vision*, pp. 3586–3594 (2015).
114. R. Dian et al., "Deep hyperspectral image sharpening," *IEEE Trans. Neural Networks Learn. Syst.* **29**(11), 5345–5355 (2018).
115. W. Xie et al., "Hyperspectral pansharpening with deep priors," *IEEE Trans. Neural Networks Learn. Syst.* **31**(5), 1529–1543 (2020).
116. H. Shen et al., "Spatial–spectral fusion by combining deep learning and variational model," *IEEE Trans. Geosci. Remote Sens.* **57**(8), 6169–6181 (2019).
117. Q. Xie et al., "Multispectral and hyperspectral image fusion by MS/HS fusion net," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, pp. 1585–1594 (2019).
118. http://www.openremotesensing.net/index.php/codes/11-pan-sharpening.
119. http://naotoyokoya.com/Download.html.
120. X. Meng et al., "Pansharpening for cloud-contaminated very high-resolution remote sensing images," *IEEE Trans. Geosci. Remote Sens.* **57**(5), 2840–2854 (2019).
121. J. Yang et al., "Coupled sparse denoising and unmixing with low-rank constraint for hyperspectral image," *IEEE Trans. Geosci. Remote Sens.* **54**(3), 1818–1833 (2016).
122. L. Zhou et al., "Separability and compactness network for image recognition and super-resolution," *IEEE Trans. Neural Networks Learn. Syst.* **30**(11), 3275–3286 (2019).
123. Y. Zhou, A. Rangarajan, and P. D. Gader, "An integrated approach to registration and fusion of hyperspectral and multispectral images," *IEEE Trans. Geosci. Remote Sens.* **58**(5), 3020–3033 (2020).

Jie Li received his BS degree in sciences and techniques of remote sensing and his PhD in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2011 and 2016. He is currently an associate professor with the School of Geodesy and Geomatics, Wuhan University. His research interests include image quality improvement, image super-resolution reconstruction, data fusion, remote sensing image processing, sparse representation, and deep learning. He has authored over 30 research papers in international journals and two books.

Huanfeng Shen received his BS degree in surveying and mapping engineering and his PhD 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 (SRES), Wuhan University, where he is currently a Luojia distinguished professor and associate dean of SRES. His research interests include remote sensing image processing, multisource data fusion, and intelligent environmental sensing. He is the PI of two projects supported by the National Key Research and Development Program of China and six projects supported by the National Natural Science Foundation of China. He has authored over 100 research papers in peer-reviewed international journals. He is a senior member of IEEE, council member of China Association of Remote Sensing Application, education committee member of Chinese Society for Geodesy Photogrammetry and Cartography, and theory committee member of Chinese Society for Geospatial Information Society. He is currently a member of the editorial boards of *Journal of Applied Remote Sensing* and *Geography and Geo-Information Science*.

Huifang Li received her PhD in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2013. She is currently an associate professor with the School of Resources and Environmental Science, Wuhan University. She focuses on the study of radiometric correction of remote sensing images, including cloud correction, shadow correction, and urban thermal environment analysis and alleviation.

Menghui Jiang received her BS degree in geographical science from Wuhan University, Wuhan, China, in 2017. She is currently pursuing her PhD in SRES, Wuhan University, Wuhan, China.

Her research interests include image data fusion, quality improvement, remote sensing image processing, and deep learning.

Qiangqiang Yuan received his BS degree in surveying and mapping engineering and his PhD 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 published more than 70 research papers, including more than 50 peer-reviewed articles in international journals such as the *IEEE Transactions Image Processing* and the *IEEE Transactions on Geoscience and Remote Sensing*. His current research interests include image reconstruction, remote sensing image processing and application, and data fusion. He was the recipient of the Youth Talent Support Program of China in 2019 and the Top-Ten Academic Star of Wuhan University in 2011. In 2014, he received the Hong Kong Scholar Award from the Society of Hong Kong Scholars and the China National Postdoctoral Council. He is an associate editor of *IEEE Access* and has frequently served as a referee for more than 40 international journals for remote sensing and image processing.