

Adaptive Spatial-Spectral Dictionary Learning for Hyperspectral Image Restoration

Ying Fu¹ · Antony Lam² · Imari Sato³ · Yoichi Sato¹

Received: 15 August 2015 / Accepted: 26 May 2016
© Springer Science+Business Media New York 2016

Abstract Hyperspectral imaging is beneficial in a diverse range of applications from diagnostic medicine, to agriculture, to surveillance to name a few. However, hyperspectral images often suffer from degradation such as noise and low resolution. In this paper, we propose an effective model for hyperspectral image (HSI) restoration, specifically image denoising and super-resolution. Our model considers three underlying characteristics of HSIs: sparsity across the spatial-spectral domain, high correlation across spectra, and non-local self-similarity over space. We first exploit high correlation across spectra and non-local self-similarity over space in the degraded HSI to learn an adaptive spatial-spectral dictionary. Then, we employ the local and non-local sparsity of the HSI under the learned spatial-spectral dictionary to design an HSI restoration model, which can be effectively solved by an iterative numerical algorithm with parameters that are adaptively adjusted for different clusters and different noise levels. In experiments on HSI denoising, we show that the proposed method outperforms many state-of-the-art methods under several comprehensive quantitative assessments. We also show that our method performs well on HSI super-resolution.

Keywords Adaptive spatial-spectral dictionary learning · Hyperspectral image restoration · Self-similarity · High correlation across spectra · Non-local sparse representation

1 Introduction

Restoring a high-quality image from degraded measurements is a typical inverse problem in image processing. It has been extensively studied in the past decades and is widely implemented in our daily life, medical imaging, surveillance, and entertainment, to name a few. Hyperspectral images (HSIs) provide much richer information about scenes than three-band RGB images but often times, they easily suffer from degradation such as noise and low resolution.

Hyperspectral imaging is the process of capturing images of a scene over multiple bands of the electromagnetic spectrum. When captured, a HSI can be thought of as a set of 2D spatially organized “pixels” where each pixel contains an entire spectral distribution over wavelengths. This allows us to see the spectral distribution of any given surface point in a scene and has led to numerous applications in remote sensing (Borengasser et al. 2007; Wang et al. 2009; Teke et al. 2013), classification or detection (Melgani and Bruzzone 2004; Renard and Bourennane 2008; Bourennane et al. 2010; Castrodad et al. 2011; Kang et al. 2014), diagnostic medicine (Stamatas et al. 2003; David and Dicker 2006; Gupta and Ramella-Roman 2008; Hou et al. 2013; Lu and Fei 2014), surveillance (Banerjee et al. 2009), biometrics (Rowe et al. 2005), and more. However, conventional hyperspectral imaging suffers from issues such as limited light in individual bands, which introduces noise into the imaging process. In addition, limited sampling also results in low resolution and lack of details. The degradation of HSIs not only influences the visual appearance of these images but also limits the per-

Communicated by Hiroshi Ishikawa, Takeshi Masuda, Yasuyo Kita and Katsushi Ikeuchi.

✉ Ying Fu
fuying@ut-vision.org

- ¹ Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan
- ² Graduate School of Science and Engineering, Saitama University, 255 Shimo-Okubo, Sakura-ku, Saitama 338-8570, Japan
- ³ National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda, Tokyo 101-8430, Japan

formance of classification and detection methods. Therefore, HSI restoration is a critical step in any HSI based system pipeline.

Recently, sparse representation based models have been widely used in image restoration (Ingrid Daubechies 2004; Bioucas-Dias and Figueiredo 2007; Zibulevsky and Elad 2010; Dabov et al. 2007; Mairal et al. 2009b; Elad and Aharon 2006; Yang et al. 2010; Dong et al. 2011, 2013; Parmar et al. 2008). The success of sparse representation models owes to the development of l_1 -norm optimization techniques (Candes and Tao 2006; Chen et al. 1998; Ingrid Daubechies 2004; Donoho 2004) and the finding that natural images are intrinsically sparse in many domains (Olshausen and Field 1996, 1997). These domains can be represented by dictionaries, which can be constructed from basis functions (e.g. DCT, Wavelets), adaptive learning [e.g. K-SVD (Elad and Aharon 2006)], and stochastic approximations (Mairal et al. 2009a). In addition, due to natural images containing self-repeating patterns, the self-similarity of overlapping patches has been introduced in many papers on topics such as denoising (Buades et al. 2005; Dabov et al. 2007), demosaicing (Buades et al. 2009), super-resolution (Glasner et al. 2009; Yang et al. 2011), and frame rate up-conversion (Shimano et al. 2011).

Dong et al. (2013, 2015) was able to combine both sparse representation and non-local self-similarity to restore grayscale natural images to great effect. However, their grayscale image approach is not ideally suited to HSIs since the relationship between bands is not considered. On the other hand, Lam et al. (2012) showed that spectral domain statistics can help HSI denoising. This suggests there is rich information in the spectral domain that can be utilized. Therefore, the potential use of sparse representation, non-local self-similarity, and spectral domain statistics jointly should be studied.

In this paper, we propose a novel HSI restoration method based on adaptive spatial-spectral dictionary learning as well as local and non-local sparse representations. We first adaptively learn a spatial-spectral dictionary for each overlapping cubic patch of the HSI. By design, each dictionary can sparsely represent its own cubic patch while sharing features with other non-local patches that are similar. Specifically, the dictionary effectively considers the high correlation across the spatial-spectral domain. Then, sparsity and non-local self-similarity constraints are used in the model, which can be solved effectively by an iterative shrinkage algorithm and also allows us to adaptively adjust the regularization parameters for different clusters and different noise levels. Our method outperforms many state-of-the-art HSI denoising methods under several comprehensive quantitative assessments and also achieves high performance in HSI super-resolution.

The remainder of this paper is organized as follows. Section 2 reviews related work. In Sect. 3, our approach to

adaptive spatial-spectral dictionary learning for HSI restoration is presented. In Sect. 4, the numerical algorithm for our model is described in detail. Extensive experimental results are presented in Sect. 5. Finally, conclusions are drawn and future directions are discussed in Sect. 6.

2 Related Work

In the following, we will review the most relevant studies on HSI denoising and super-resolution.

2.1 Hyperspectral Image Denoising

A number of algorithms such as K-SVD (Elad and Aharon 2006), Non-local means (Buades et al. 2005) and BM3D (Dabov et al. 2007) have been developed for 2D image denoising. These methods can be directly applied to HSI denoising by denoising each band image independently. However, considering the bands to be independent limits performance. Qian et al. (2012) and Maggioni et al. (2013) denoised HSIs by considering 3D cubes of the HSI instead of the 2D patches in traditional image restoration but these methods ignore the high correlation across spectra and this limits their performance.

There are several methods based on Wavelets and/or principal component analysis (PCA) for HSI denoising. Atkinson et al. (2003) and Othman and Qian (2006) both proposed wavelet-based HSI denoising algorithms. Guangyi and Q (2009) and Chen and Qian (2011) proposed performing dimensionality reduction and HSI denoising using wavelet shrinkage and PCA. Lam et al. (2012) used PCA for dimensionality reduction in the spectral domain and then performed denoising in local spatial neighborhoods to further improve denoising results.

Several methods used tensor decompositions. Letexier and Bourennane (2008) proposed an HSI denoising method based on a generalized multidimensional Wiener filter by using a third-order tensor. Renard et al. (2008) used a low rank tensor approximation to jointly achieve denoising and spectral dimensionality reduction. Karami et al. (2011) reduced the noise of HSIs by using their Genetic Kernel Tucker Decomposition. Guo et al. (2013) denoised HSIs based on a high-order rank-1 tensor decomposition.

In addition, Murakami et al. (2008) employed spatio-spectral Wiener estimation to denoise multispectral images. Wang et al. (2009, 2010) utilized an alternative hyperspectral anisotropic diffusion scheme to denoise HSIs. Yuan et al. (2012) employed a spatial-spectral adaptive total variation model for their denoising scheme. Zhong and Wang (2013) simultaneously modeled and utilized spatial and spectral dependence in a unified probabilistic framework by multiple-spectral-band conditional random fields. Also, Zhang et al.

(2014) explored the low-rank property of HSIs for image restoration. Many of these methods exploited both spatial and spectral information. However, none have utilized the non-local similarity present in HSIs of natural scenes.

Qian and Ye (2013) introduced the non-local similarity and spatial-spectral structure of HSIs into a sparse representation framework, but their method used a global 3D dictionary built from 3D DCT and 3D OWT, which cannot be adapted to the specific characteristics of a given scene. Peng et al. (2014) proposed an HSI denoising approach based on decomposable non-local tensor dictionary learning, but their method did not explicitly consider the self-similarity within neighboring cubic patches and cannot be directly extend to HSI super-resolution.

2.2 Hyperspectral Image Super-Resolution

Recently, some researchers have begun to employ super-resolution reconstruction approaches to improve the spatial resolution of HSIs. Most of these methods make use of several low resolution HSIs or a low resolution HSI with a high spatial resolution panchromatic image to obtain a single high resolution HSI.

Akgun et al. (2005) fused information from multiple observations and spectral bands to improve spatial resolution and reconstruct the spectrum of the observed scene as a combination of a small number of spectral basis functions. Buttngsrud and Alsberg (2006) proposed a maximum entropy-based HSI super-resolution reconstruction method by using multiple low resolution HSIs. Chan et al. (2010) employed a multi-angular low resolution HSI to reconstruct a high resolution one. Zhang et al. (2012) proposed a maximum a posteriori estimation based multi-frame super-resolution method for HSIs. All these multi-frame based methods need to know or estimate motion parameters between frames.

As an alternative, a number of researchers have tackled the spatial super-resolution restoration of an HSI as an image fusion problem, by use of an auxiliary high resolution image. Eismann and Hardie (2004), Eismann and Hardie (2005), Chen et al. (2014) and Zhao et al. (2011) used a high resolution panchromatic image for HSIs. Ma et al. (2013) and Dong et al. (2016) employed an RGB video/image with high spatial resolution to enhance a single HSI with low spatial resolution.

In our work, we improve the spatial-resolution of low resolution HSIs directly by considering the high correlation across spectra, and do not require any auxiliary high frequency image or multiple observations. Our ability to both denoise and perform super-resolution with good performance demonstrates the extensibility and effectiveness of our method. We now detail how our model is able to solve different restoration problems.

3 Hyperspectral Image Restoration Model

We assume a general model of HSI degradation based on the following equation

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{n} \quad (1)$$

where \mathbf{H} is a degradation matrix, $\mathbf{X} \in \mathbb{R}^{MNB}$ and $\mathbf{Y} \in \mathbb{R}^{M_h N_h B}$ are stacked vector representations of the original clear HSI and the degraded HSI. M stands for the rows of the image, N represents the number of columns, and B is the number of bands. M_h and N_h are the number of rows and columns of the degraded HSI, which is related to the size of degradation matrix \mathbf{H} , and \mathbf{n} is additive noise with the same size as \mathbf{Y} . With different settings of matrix \mathbf{H} , Eq. (1) can represent different image restoration problems, for example, in image denoising, \mathbf{H} is an identity matrix, and in image super-resolution, \mathbf{H} is a composite operator for blurring and down-sampling.

In this section, we will present a novel adaptive spatial-spectral dictionary learning based HSI restoration method, which is regularized using local and non-local sparse constraints. The overall framework of the proposed method is shown in Fig. 1. First, the noisy HSI is used to learn the adaptive spatial-spectral dictionary. Then, the trained dictionary is employed to sparsely represent the HSI using both local details and non-local self-similarities of the image. These two terms are integrated into a unified variational framework for optimization, which is solved by a numerical algorithm. Through the iterations, we alternately update the dictionary and restore the HSI. This process is iterated several times until the optimization converges. In this paper, we apply the proposed method to HSI denoising and super-resolution.

3.1 Adaptive Spatial-Spectral Dictionary Learning

We first show how we build the adaptive spatial-spectral dictionary, which considers both non-local self-similarities and the high correlation across spectra. The HSI is first divided into overlapping cubic patches of size $P \times P \times B$, where $P < M$ and $P < N$. In our method, we consider the two spatial dimensions as a whole. Each cubic patch centered at the spatial location (i, j) can be described as $Z_{i,j} \in \mathbb{R}^{P^2 \times B}$. Each element in $Z_{i,j}$ can be denoted as $Z_{i,j}[n, b]$, where $n = 1, 2, \dots, P^2$ and $b = 1, 2, \dots, B$ index the spatial pixel location and spectral band, respectively.

In studies of the human visual system (Olshausen and Field 1996, 1997), the eye's cell receptive fields code natural images using a small number of structural primitives sparsely chosen from an over-complete code set. This suggests that natural images are intrinsically sparse in some domains, which can be described as a dictionary and constructed by basis functions. Therefore, we can seek a representation

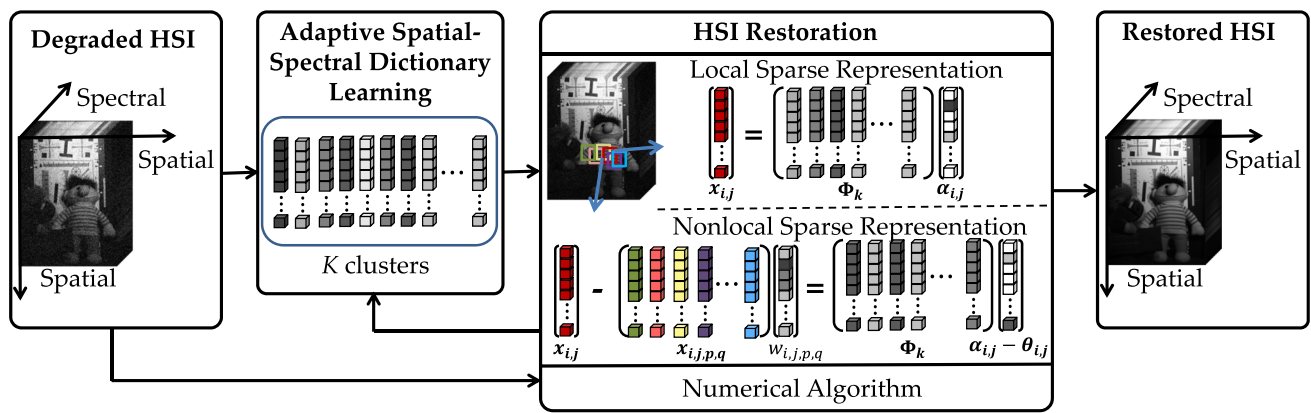


Fig. 1 Overview of the proposed method for HSI restoration. The degraded HSI is first used to learn the adaptive spatial-spectral dictionary. Then, the trained dictionary is employed to sparsely represent the

HSI using both local details and non-local self-similarities of the image. These two terms are integrated into a unified variational framework for optimization, which is solved by a numerical algorithm

that allows us to analyze the high-dimensional $Z_{i,j}$ using a smaller number of components. Then, $Z_{i,j}$ can be represented in terms of a linear superposition of basis functions,

$$Z_{i,j}[n, b] = \sum_m \alpha_m \phi_m[n, b], \tag{2}$$

where $\{\phi_m\}$ is a basis set, and α_m are the corresponding scalar coefficients.

According to previous research (Chakrabarti and Zickler 2011), an HSI’s dictionary can be decomposed into a product of separate spatial and spectral components, that is to say,

$$\phi_m[n, b] = S_d[n]U_r[b], \tag{3}$$

where $\{S_d\}_1^{P^2}$ and $\{U_r\}_1^B$ are orthonormal bases spanning the space of monochrome $P \times P$ spatial patches and the space of spectral distribution from the B channels, respectively. In practice, $\phi_m \in \mathbb{R}^{P^2 \times B}$ is reshaped to a vector $\Phi_m \in \mathbb{R}^{P^2 B}$. This means that spatial basis $\{S_d\}_1^{P^2}$ and spectral basis $\{U_r\}_1^B$ can be separately trained, and then combined to obtain the joint spatial-spectral dictionary.

As to how we specifically represent the bases for the spatial and spectral domains, we first provide details for the spectral domain. It is well known that large sets of spectra consisting of natural scenes or extensive color samples can be represented compactly in just 6 to 8 principal components (Vrhel et al. 1994). So in the spectral domain, we reshape the original HSI into $B \times MN$ and learn the spectral basis $\{U_r\}_1^B$ by PCA (Vrhel et al. 1994; Chakrabarti and Zickler 2011). In addition, the spatial patches can also be represented by PCA basis (Dong et al. 2011; Chakrabarti and Zickler 2011). A simple test on the Columbia Multispectral Image Dataset (Yasuma et al. 2008) reveals that 10,000 random 7×7 spatial patches require only 6 principal components to capture 99% of the variance. We also train the spatial basis $\{S_d\}_1^{P^2}$ using

PCA from monochrome patches pooled across all bands in the spatial domain. The spatial-spectral dictionary $\{\phi_m\}$ can then be constructed by different combinations of S_d and U_r as shown in Eq. (3).

At this point, we could learn a single universal dictionary for all spatial patches like in the cases of the analytically designed wavelet dictionary (Qian and Ye 2013) and the learned K-SVD (Elad and Aharon 2006). Indeed, one dictionary can be used to represent any image patch. However, it would lack sufficient flexibility to sparsely represent a given local patch. In Dong et al. (2013) and Peng et al. (2014), they collected patches that were similar to each other to learn a local dictionary for each cluster instead of a universal dictionary, thus allowing for more sparse representations. Inspired by their methods, instead of learning a single universal dictionary, we first cluster 2D monochrome patches pooled across all bands in the spatial domain by k-means. This clustering process considers the non-local self-similarity of the patches across the full HSI, and we call it the *global non-local self-similarity*. In practice, we cluster spatial image patches by their high frequency patterns (rich textures and edges) as opposed to clustering directly based on the pixel values in the patches to avoid influence from overly similar intensities. The image of high frequency spatial patterns can be obtained by

$$\mathbf{X}_g = \mathbf{X} - \mathbf{G}\mathbf{X}, \tag{4}$$

where matrix \mathbf{G} represents a blurring operator by a Gaussian kernel convolved with the original image. Note that the high frequency image \mathbf{X}_g is only used to cluster similar patches into clusters. We use the corresponding original patch $x_{i,j}$ from \mathbf{X} to learn the PCA basis for each cluster.

Let $\mathbf{S}_k = [S_1^{(k)}, S_2^{(k)}, \dots, S_{P^2}^{(k)}]$ represent the basis in the spatial domain for cluster k ($k = 1, 2, \dots, K$) and $\mathbf{U} =$

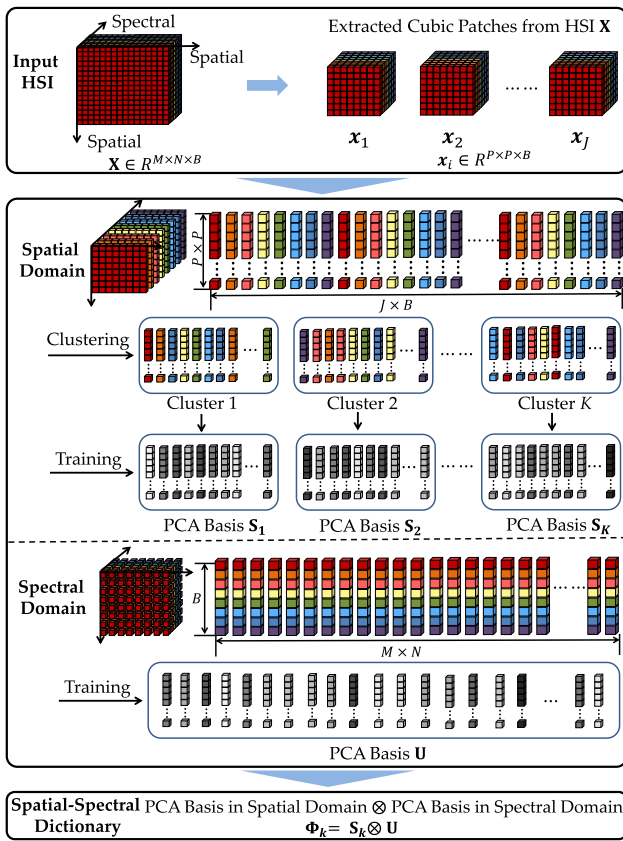


Fig. 2 Illustration of adaptive spatial-spectral dictionary learning. First, the cubic patches are extracted from the HSI in the overlapping pattern. Then, all the band images of the cubic patches are clustered in the spatial domain. The spatial PCA bases for each cluster are then learned independently. In the spectral domain, all spectral distributions from the entire HSI are extracted as vectors and used to train the spectral PCA basis. Finally, the adaptive spatial-spectral dictionary is derived from the spatial and spectral PCA bases in which each cluster has its own spatial-spectral dictionary

$[U_1, U_2, \dots, U_B]$ be the basis in the spectral domain. The combined spatial-spectral basis vectors are then formed by $S_d^{(k)} \otimes U_r$, for all pairs of d and r in cluster k . We obtain

$$\Phi_k = S_k \otimes U, \tag{5}$$

where $\Phi_k = [\Phi_1^{(k)}, \Phi_2^{(k)}, \dots, \Phi_{P^2B}^{(k)}]$ is the spatial-spectral dictionary, which is adaptively learned for each cluster and considers the spatial and spectral dimensions together in an explicit manner. \otimes is the Kronecker product. Figure 2 illustrates the adaptive spatial-spectral dictionary learning process.

Afterward, each cubic patch needs to be represented by a single spatial-spectral dictionary. However, a given cubic patch can be associated with multiple spatial clusters and as a result, multiple dictionaries. This is because the bands of a given cubic patch can be associated with different spatial clusters. Figure 3, shows an example with two band images

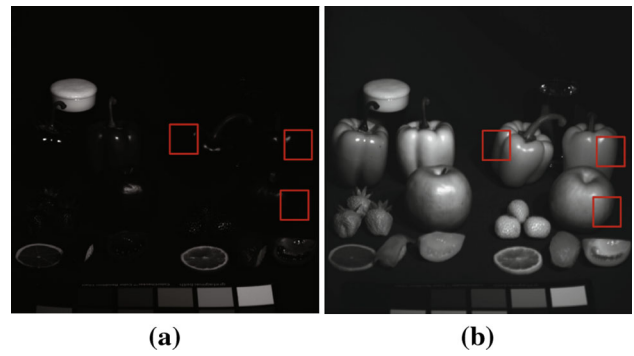


Fig. 3 The two bands from an HSI. The high frequency spatial patterns in the same locations are different for these two bands and will be clustered into different clusters

where the same location exhibits different spatial patterns. To solve this issue, the spatial-spectral dictionary for these cubic patches are obtained by taking the particular spatial dictionary for each band’s spatial patch and computing the Kronecker product with the appropriate row in the spectral dictionary.

3.2 Local and Non-local Sparse Representation Model

Let vector $x_{i,j} \in \mathbb{R}^{P^2B}$ denote a cubic patch extracted from HSI X and centered at the spatial location (i, j) ,

$$x_{i,j} = R_{i,j}X, \tag{6}$$

where $R_{i,j} \in \mathbb{R}^{P^2B \times MNB}$ is the matrix extracting patch $x_{i,j}$ from X . The cubic patches are extracted by shifting over $2D$ spatial locations such that we obtain J cubic patches, where $J = (M - P + 1)(N - P + 1)$. Please note that vector $x_{i,j}$ can also be obtained by reshaping matrix $Z_{i,j}$ as a vector.

As mentioned previously, the learned adaptive spatial-spectral dictionary $\Phi_k \in \mathbb{R}^{B \times P^2 \times m}$ can be used to sparsely code cubic patches $x_{i,j}$, as $\Phi_k \alpha_{i,j}$, when the cubic patch at location (i, j) belongs to the k th cluster C_k . Given the adaptive spatial-spectral dictionary Φ_k , the sparse codes α and corresponding image X can be recovered by

$$\{\hat{\alpha}, \hat{X}\} = \arg \min_{\alpha} \|Y - HX\|_2^2 + \sum_{i,j} \sum_{k=1}^K \sum_{(i,j) \in C_k} (\gamma \|x_{i,j} - \Phi_k \alpha_{i,j}\|_2^2 + \lambda \|\alpha_{i,j}\|_1), \tag{7}$$

where α denotes the concatenation of all $\alpha_{i,j}$, and $\hat{\alpha}$ is the estimation of α . In addition, the cubic patches in an HSI have rich self-similarity with its neighboring patches (Qian et al. 2012), which implies the cubic patch can be effectively described by its neighboring patches. We call this the *neighboring non-local self-similarity* of the HSI to distinguish it

from the *global non-local self-similarity* in Sect. 3.1. In practice, we employ non-local means to describe the *neighboring non-local self-similarity* of the HSI. The non-local means $\hat{\mathbf{x}}_{i,j}$ can be computed by

$$\hat{\mathbf{x}}_{i,j} = \sum_{(p,q) \in \Omega_{i,j}} w_{i,j,p,q} \mathbf{x}_{i,j,p,q}, \tag{8}$$

where $\mathbf{x}_{i,j,p,q}$ denotes a found similar cubic patch within set $\Omega_{i,j}$, which represents a large neighborhood around cubic patch $\mathbf{x}_{i,j}$, and $w_{i,j,p,q}$ is the corresponding weight. Like in [Buades et al. \(2005\)](#), $w_{i,j,p,q}$ can be set as

$$w_{i,j,p,q} = \frac{1}{W} \exp(-\|\mathbf{x}_{i,j} - \mathbf{x}_{i,j,p,q}\|_2^2/d) \tag{9}$$

where W is a normalization factor and d is a pre-determined scalar. Then, we can obtain

$$\begin{aligned} \mathbf{x}_{i,j} - \hat{\mathbf{x}}_{i,j} &= \mathbf{x}_{i,j} - \sum_{(p,q) \in \Omega_{i,j}} w_{i,j,p,q} \mathbf{x}_{i,j,p,q} \\ &= \Phi_k \left(\alpha_{i,j} - \sum_{(p,q) \in \Omega_{i,j}} w_{i,j,p,q} \alpha_{i,j,p,q} \right) \\ &= \Phi_k(\alpha_{i,j} - \theta_{i,j}), \end{aligned} \tag{10}$$

where $\theta_{i,j} = \sum_{(p,q) \in \Omega_{i,j}} w_{i,j,p,q} \alpha_{i,j,p,q}$ is the sparse representation of the non-local means $\hat{\mathbf{x}}_{i,j}$. Since $\hat{\mathbf{x}}_{i,j}$ approximates $\mathbf{x}_{i,j}$, the sparse codes $\theta_{i,j}$ of $\hat{\mathbf{x}}_{i,j}$ should be similar to the sparse codes $\alpha_{i,j}$ of $\mathbf{x}_{i,j}$ ([Qian and Ye 2013](#)). Also, due to the sparsity of $\alpha_{i,j}$, $(\alpha_{i,j} - \theta_{i,j})$ should be sparse enough. Therefore, we use $\|\alpha_{i,j} - \theta_{i,j}\|_1$ as the non-local sparse constraint like [Dong et al. \(2013\)](#). In our method, we add this non-locally sparse constraint into Eq. (7) as a regularization,

$$\begin{aligned} \{\hat{\alpha}, \hat{\mathbf{X}}\} &= \arg \min_{\alpha} \|\mathbf{Y} - \mathbf{H}\mathbf{X}\|_2^2 + \sum_{i,j} \sum_{k=1}^K \sum_{(i,j) \in C_k} \\ &\quad \left(\gamma \|\mathbf{x}_{i,j} - \Phi_k \alpha_{i,j}\|_2^2 + \lambda \|\alpha_{i,j}\|_1 + \eta \|\alpha_{i,j} - \theta_{i,j}\|_1 \right). \end{aligned} \tag{11}$$

We can see that our model in Eq. (11) unifies the local sparsity (i.e. $\|\alpha_{i,j}\|_1$) and neighboring non-local self-similarity induced non-local sparsity (i.e. $\|\alpha_{i,j} - \theta_{i,j}\|_1$) into a variational formulation. The model also clusters the HSI cubic patches across the spatial domain to learn the adaptive spatial-spectral dictionary by considering the global non-local self-similarity over the entire scene (i.e. the summation $\sum_{k=1}^K$ and Φ_k over the K clusters for the cost function). Therefore, our model exploits both local and non-local redundancies across the spatial-spectral domain to restore HSIs. Note that for convenience and without loss of generality, $\sum_{i,j} \sum_{k=1}^K \sum_{(i,j) \in C_k}$ is written as $\sum_{i,j,k}$ in the following sections of this paper.

4 Hyperspectral Image Restoration Algorithm

4.1 Determination of Parameters λ and η

We begin by showing that the Bayesian interpretation for our model (Eq. (11)) provides an explicit way to adaptively set the parameters λ and η for different clusters and noise levels. Under the Bayesian framework, the estimation of sparsity vector α can be cast as a Maximum a Posteriori problem,

$$\begin{aligned} \hat{\alpha} &= \arg \max_{\alpha} P(\alpha|\mathbf{Y}) \\ &= \arg \max_{\alpha} P(\mathbf{Y}|\alpha)P(\alpha) \\ &= \arg \min_{\alpha, \beta} \{-\log P(\mathbf{Y}|\alpha) - \log P(\alpha)\}, \end{aligned} \tag{12}$$

where the two terms correspond to the likelihood and prior terms, respectively. We consider the case where the observation \mathbf{Y} is contaminated with additive Gaussian noise with standard deviation σ_n . Thus, $P(\mathbf{Y}|\alpha)$ satisfies the Gaussian distribution i.e. $\mathbf{Y} = \mathbf{H}\mathbf{X} + \epsilon_1$, $\epsilon_1 \sim \mathcal{N}(0, \sigma_n^2)$ and $\mathbf{x}_{i,j} = \Phi_k \alpha_{i,j} + \epsilon_2$, $\epsilon_2 \sim \mathcal{N}(0, \sigma_n^2/\gamma)$. The likelihood term $P(\mathbf{Y}|\alpha)$ can be characterized as

$$\begin{aligned} P(\mathbf{Y}|\alpha) &= \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left(-\frac{1}{2\sigma_n^2}\|\mathbf{Y} - \mathbf{H}\mathbf{X}\|_2^2\right) \times \\ &\quad \prod_{i,j,k} \frac{1}{\sqrt{2\pi/\gamma}\sigma_n} \exp\left(-\frac{\gamma}{2\sigma_n^2}\|\mathbf{x}_{i,j} - \Phi_k \alpha_{i,j}\|_2^2\right). \end{aligned} \tag{13}$$











Due to the sparsity of $\alpha_{i,j}$ and $(\alpha_{i,j} - \theta_{i,j})$ by l_1 -norm, the Laplace distribution can be used to describe $P(\alpha)$, i.e. $\alpha_{ijl} \sim \text{Laplace}(0, \frac{\delta_{kl}}{\sqrt{2}c_1})$ and $(\alpha_{ijl} - \theta_{ijl}) \sim \text{Laplace}(0, \frac{q_{kl}}{\sqrt{2}c_2})$. Thus, the prior term $P(\alpha)$ in Eq. (12) can be expressed as

$$\begin{aligned} P(\alpha) &= \prod_{i,j,l,k} \frac{c_1}{\sqrt{2}\delta_{kl}} \exp\left(-\frac{\sqrt{2}c_1|\alpha_{ijl}|}{\delta_{kl}}\right) \\ &\quad \frac{c_2}{\sqrt{2}q_{kl}} \exp\left(-\frac{\sqrt{2}c_2|\alpha_{ijl} - \theta_{ijl}|}{q_{kl}}\right), \end{aligned} \tag{14}$$

where α_{ijl} and θ_{ijl} are the l th elements of $\alpha_{i,j}$ and $\theta_{i,j}$, and δ_{kl} and q_{kl} are the standard deviations of α_{ijl} and β_{ijl} when $(i,j) \in C_k$, $l = 1, 2, \dots, P^2B$. c_1 and c_2 are predefined constants. Substituting Eqs. (13) and (14) into Eq. (12), we obtain

$$\begin{aligned} \{\hat{\alpha}, \hat{\mathbf{X}}\} &= \arg \min_{\alpha} \|\mathbf{Y} - \mathbf{H}\mathbf{X}\|_2^2 + \sum_{i,j,k} \left(\gamma \|\mathbf{x}_{i,j} - \Phi_k \alpha_{i,j}\|_2^2 + \right. \\ &\quad \left. c_1 \sum_l \frac{2\sqrt{2}\sigma_n^2}{\delta_{kl}} \|\alpha_{ijl}\|_1 + c_2 \sum_l \frac{2\sqrt{2}\sigma_n^2}{q_{kl}} \|\alpha_{ijl} - \theta_{ijl}\|_1 \right). \end{aligned} \tag{15}$$

Table 1 Denoising results (PSNR(dB)/SSIM/FSIM) of the 10 hyperspectral images from different methods, $\sigma_n = 20$

Imgr.	Metri.	Noisy	BM3D	CSR	ANLM	BM4D	3DCSR	LRTA	SDS	K-SVD	TDL	OurBL	Ours
	PSNR	22.11	36.15	36.12	35.57	39.93	40.01	35.73	36.21	32.51	39.38	39.77	40.81
	SSIM	0.2398	0.9397	0.9437	0.9497	0.9644	0.9655	0.8932	0.8711	0.8002	0.9666	0.9625	0.9721
	FSIM	0.7539	0.9723	0.9722	0.9747	0.9861	0.9868	0.9599	0.9597	0.9603	0.9883	0.9875	0.9905
	PSNR	22.11	30.48	30.35	29.54	34.74	34.64	31.87	34.01	34.70	30.32	35.39	35.95
	SSIM	0.4274	0.8404	0.8359	0.8033	0.9282	0.9255	0.8376	0.9046	0.9323	0.8067	0.9370	0.9408
	FSIM	0.8894	0.9518	0.9476	0.9342	0.9830	0.9804	0.9639	0.9801	0.9844	0.9634	0.9836	0.9871
	PSNR	22.10	38.89	38.82	38.25	41.63	41.60	37.18	36.49	33.98	40.71	41.02	41.97
	SSIM	0.1672	0.9383	0.9411	0.9434	0.9588	0.9596	0.8906	0.8592	0.7829	0.9637	0.9512	0.9611
	FSIM	0.6977	0.9666	0.9649	0.9733	0.9754	0.9751	0.9436	0.9415	0.9473	0.9832	0.9789	0.9837
	PSNR	22.11	37.48	37.73	37.66	40.82	40.97	36.57	37.22	33.87	40.50	40.65	41.21
	SSIM	0.1797	0.9284	0.9337	0.9364	0.9523	0.9534	0.8978	0.8696	0.7873	0.9604	0.9505	0.9557
	FSIM	0.7469	0.9655	0.9660	0.9701	0.9814	0.9819	0.9624	0.9386	0.9567	0.9850	0.9841	0.9878
	PSNR	22.11	36.83	36.82	35.88	39.71	39.69	35.57	34.68	32.58	39.59	39.46	40.26
	SSIM	0.2196	0.9384	0.9407	0.9435	0.9606	0.9588	0.8759	0.8516	0.7947	0.9646	0.9520	0.9605
	FSIM	0.7543	0.9685	0.9682	0.9720	0.9811	0.9832	0.9521	0.9555	0.9589	0.9844	0.9839	0.9869
	PSNR	22.11	39.62	39.66	38.30	41.91	42.05	38.52	37.23	34.28	42.09	42.03	43.20
	SSIM	0.1665	0.9390	0.9450	0.9506	0.9563	0.9601	0.9074	0.8376	0.7847	0.9742	0.9608	0.9690
	FSIM	0.7130	0.9697	0.9702	0.9693	0.9801	0.9860	0.9632	0.9482	0.9486	0.9874	0.9860	0.9891
	PSNR	22.11	33.86	33.82	32.00	36.46	37.01	34.27	34.79	31.77	36.59	37.53	37.93
	SSIM	0.2823	0.8679	0.8665	0.8031	0.9205	0.9299	0.8774	0.8626	0.7652	0.9319	0.9373	0.9421
	FSIM	0.8138	0.9554	0.9540	0.9389	0.9782	0.9823	0.9593	0.9676	0.9522	0.9810	0.9822	0.9852
	PSNR	22.11	36.39	36.30	35.68	38.32	38.21	36.08	33.82	32.66	37.73	38.11	39.07
	SSIM	0.2379	0.9106	0.9090	0.9195	0.9367	0.9318	0.8898	0.8749	0.7949	0.9323	0.9317	0.9334
	FSIM	0.8340	0.9643	0.9641	0.9712	0.9789	0.9790	0.9643	0.9734	0.9616	0.9819	0.9792	0.9848
	PSNR	22.11	37.99	37.99	37.37	40.47	40.55	38.80	36.11	33.77	40.74	40.50	41.56
	SSIM	0.1969	0.9294	0.9332	0.9340	0.9491	0.9566	0.9337	0.8576	0.7878	0.9641	0.9534	0.9626
	FSIM	0.7634	0.9711	0.9718	0.9764	0.9822	0.9844	0.9801	0.9618	0.9612	0.9880	0.9850	0.9884
	PSNR	22.11	33.52	33.41	32.59	36.85	37.55	34.34	34.92	31.47	36.58	37.23	38.07
	SSIM	0.3298	0.9118	0.9100	0.9127	0.9496	0.9535	0.8940	0.9021	0.8164	0.9488	0.9530	0.9605
	FSIM	0.8538	0.9651	0.9643	0.9648	0.9846	0.9850	0.9730	0.9780	0.9675	0.9863	0.9852	0.9890

Comparing Eqs. (15) and (11), we have

$$\lambda_{kl} = c_1 \frac{2\sqrt{2}\sigma_n^2}{\delta_{kl}}, \quad \eta_{kl} = c_2 \frac{2\sqrt{2}\sigma_n^2}{\varrho_{kl}}. \tag{16}$$

In Eq. (16), c_1 and c_2 are set to be the same for different noise or scaling levels. σ_n relies on the noise level of the input HSI. The standard deviations δ_{kl} and ϱ_{kl} for each cluster can be estimated from the sets of α_{ijl} and $(\alpha_{ijl} - \theta_{ijl})$ belonging to cluster C_k , respectively. In practice, we set $\gamma = 0.3$, $c_1 = 0.3$, and $c_2 = 0.7$ for HSI denoising, and $\gamma = 3.2$, $c_1 = 0.2$, and $c_2 = 1.4$ for HSI super-resolution.

4.2 Numerical Algorithm

In Eq. (15), we can see that there are three variables, i.e. \mathbf{X} , α , and θ . α and θ are the concatenation of all $\alpha_{i,j}$ and $\theta_{i,j}$, respectively. To solve Eq. (15), we derive update

rules for \mathbf{X} , α , and θ . We adopt an alternating minimization scheme to reduce the original problem into simpler sub-problems.

Update θ We initialize $\theta_{i,j}$ to 0, i.e. $\theta_{i,j}^{(0)} = 0$, where $a^{(t)}$ represents the t th iteration of variable a . Then, setting $\mathbf{X}^{(0)} = \mathbf{Y}$ for HSI denoising and $\mathbf{X}^{(0)}$ as the bicubic interpolator of degraded HSI \mathbf{Y} for HSI super-resolution, the sparse coding $\alpha_{i,j}^{(0)}$ can be obtained by $\alpha_{i,j}^{(0)} = \Phi_k^T \mathbf{x}_{i,j}^{(0)}$. The similar patches for each local sparse representation $\alpha_{i,j}^{(1)}$ are selected from its neighborhood in terms of the squared difference between them, i.e. $\|\alpha_{i,j} - \alpha_{i,j,p,q}\|^2$. Hence the non-local mean $\theta_{i,j}^{(t)}$ can be updated by

$$\theta_{i,j}^{(t)} = \sum_{(p,q) \in \Omega_{i,j}} w_{i,j,p,q} \alpha_{i,j,p,q}^{(t-1)}. \tag{17}$$

In the t th iteration, the model can be described as

Table 2 Denoising results (PSNR(dB)/SSIM/FSIM) of the 10 hyperspectral images from different methods, $\sigma_n = 40$











Img.	Metr.	Noisy	BM3D	CSR	ANLM	BM4D	3DCSR	LRTA	SDS	K-SVD	TDL	OurBL	Ours
	PSNR	16.09	32.21	32.14	31.51	35.82	35.78	32.28	29.57	29.02	35.30	35.52	36.59
	SSIM	0.1055	0.8778	0.8900	0.8983	0.9105	0.9230	0.8112	0.5617	0.7157	0.9334	0.9172	0.9374
	FSIM	0.5624	0.9382	0.9404	0.9513	0.9670	0.9677	0.9223	0.8737	0.9200	0.9717	0.9695	0.9761
	PSNR	16.09	27.06	26.97	26.99	30.94	30.88	28.32	27.37	25.90	31.05	31.22	32.08
	SSIM	0.1934	0.6968	0.6926	0.6990	0.8483	0.8421	0.7077	0.6649	0.6041	0.8593	0.8638	0.8841
	FSIM	0.7643	0.8833	0.8718	0.8775	0.9553	0.9423	0.9146	0.9406	0.8564	0.9634	0.9537	0.9638
	PSNR	16.10	35.24	35.40	35.74	37.57	37.63	33.65	29.89	31.30	37.10	37.26	38.39
	SSIM	0.0524	0.8781	0.8969	0.9041	0.8959	0.9012	0.8013	0.5236	0.7050	0.9329	0.8983	0.9160
	FSIM	0.4917	0.9343	0.9394	0.9504	0.9499	0.9501	0.9078	0.8399	0.9123	0.9631	0.9593	0.9662
	PSNR	16.10	34.07	34.09	34.86	37.10	37.08	33.41	29.77	31.17	37.05	37.08	38.06
	SSIM	0.0589	0.8714	0.8904	0.9003	0.8961	0.9015	0.8276	0.5186	0.7109	0.9307	0.9025	0.9233
	FSIM	0.5511	0.9306	0.9327	0.9519	0.9590	0.9593	0.9231	0.8671	0.9171	0.9664	0.9639	0.9725
	PSNR	16.09	33.03	33.17	32.62	35.82	35.91	32.16	29.28	29.25	35.96	35.55	36.46
	SSIM	0.0883	0.8801	0.8951	0.9032	0.9106	0.9136	0.7883	0.5616	0.7185	0.9370	0.9032	0.9246
	FSIM	0.5643	0.9358	0.9384	0.9509	0.9580	0.9587	0.9128	0.8718	0.9203	0.9674	0.9643	0.9696
	PSNR	16.08	35.48	35.68	35.62	38.02	38.21	35.16	29.24	31.33	38.57	38.38	39.77
	SSIM	0.0528	0.8735	0.8899	0.8903	0.8858	0.8933	0.8333	0.4646	0.6890	0.9388	0.9108	0.9326
	FSIM	0.5047	0.9345	0.9411	0.9457	0.9567	0.9588	0.9342	0.8421	0.8865	0.9702	0.9704	0.9763
	PSNR	16.10	30.93	31.11	30.48	33.07	33.08	31.40	28.59	30.49	33.30	34.23	34.84
	SSIM	0.1063	0.7760	0.7881	0.7501	0.8417	0.8512	0.8081	0.5780	0.7501	0.8761	0.8837	0.8968
	FSIM	0.6328	0.9158	0.9134	0.9207	0.9527	0.9509	0.9223	0.9023	0.9207	0.9563	0.9600	0.9669
	PSNR	16.09	33.11	33.43	33.59	35.22	35.42	32.62	28.08	29.63	34.97	34.58	35.71
	SSIM	0.0810	0.8607	0.8745	0.8936	0.8961	0.8963	0.8098	0.5547	0.7248	0.8984	0.8760	0.9016
	FSIM	0.8340	0.9351	0.9392	0.9583	0.9569	0.9577	0.9322	0.9152	0.9367	0.9650	0.9553	0.9657
	PSNR	16.09	34.23	34.45	34.01	36.84	37.01	35.30	29.18	30.75	37.12	36.82	30.10
	SSIM	0.0703	0.8668	0.8852	0.8841	0.8873	0.8905	0.8796	0.5117	0.7046	0.9312	0.9065	0.9277
	FSIM	0.5690	0.9394	0.9463	0.9540	0.9607	0.9638	0.9607	0.8783	0.9264	0.9731	0.9679	0.9738
	PSNR	16.09	29.93	29.83	29.57	32.32	32.26	30.87	28.48	28.06	32.89	33.30	34.21
	SSIM	0.1509	0.8416	0.8473	0.8605	0.9087	0.9117	0.8177	0.6230	0.7249	0.9069	0.9040	0.9301
	FSIM	0.7034	0.9265	0.9232	0.9363	0.9658	0.9678	0.9418	0.9253	0.9790	0.9663	0.9618	0.9713

Table 3 Denoising results (PSNR(dB)/SSIM/FSIM) on the Columbia Multispectral Image Dataset from different methods and noise levels

σ_n	Metr.	Noisy	BM3D	CSR	ANLM	BM4D	3DCSR	LRTA	SDS	K-SVD	TDL	OurBL	Ours
10	PSNR	28.13	41.95	42.10	40.77	44.48	44.30	41.11	39.36	36.38	44.28	43.99	44.52
	SSIM	0.4567	0.9671	0.9698	0.9622	0.9791	0.9781	0.9491	0.9492	0.8597	0.9807	0.9770	0.9855
	FSIM	0.8747	0.9820	0.9840	0.9821	0.9876	0.9880	0.9800	0.9753	0.9690	0.9913	0.9898	0.9930
20	PSNR	22.11	38.25	38.30	37.53	40.83	40.89	37.78	35.93	33.49	40.88	41.02	41.85
	SSIM	0.2013	0.9351	0.9382	0.9379	0.9561	0.9578	0.9107	0.8571	0.7911	0.9653	0.9563	0.9678
	FSIM	0.7224	0.9637	0.9639	0.9685	0.9744	0.9755	0.9615	0.9421	0.9480	0.9832	0.9752	0.9886
30	PSNR	18.28	36.03	36.33	35.77	38.58	38.61	35.77	32.24	31.81	38.81	38.42	39.45
	SSIM	0.1141	0.9039	0.9180	0.9169	0.9295	0.9378	0.8774	0.6920	0.7460	0.9500	0.9380	0.9512
	FSIM	0.6087	0.9466	0.9512	0.9564	0.9620	0.9623	0.9457	0.8965	0.9284	0.9742	0.9701	0.9789
40	PSNR	16.09	34.41	34.73	34.51	36.93	37.23	34.32	29.03	30.62	37.32	37.20	38.66
	SSIM	0.0742	0.8743	0.8910	0.8938	0.9001	0.9189	0.8458	0.5159	0.9110	0.9341	0.9112	0.9356
	FSIM	0.5262	0.9307	0.9359	0.9443	0.9496	0.9546	0.9312	0.8444	0.9105	0.9647	0.9534	0.9698
50	PSNR	14.15	33.49	33.59	33.53	35.64	35.86	33.18	26.40	29.76	36.18	35.77	36.96
	SSIM	0.0522	0.8616	0.8780	0.8682	0.8701	0.8901	0.8190	0.3774	0.6828	0.9181	0.9099	0.9222
	FSIM	0.4644	0.9215	0.9235	0.9322	0.9377	0.9476	0.9183	0.7914	0.8948	0.9560	0.9457	0.9611

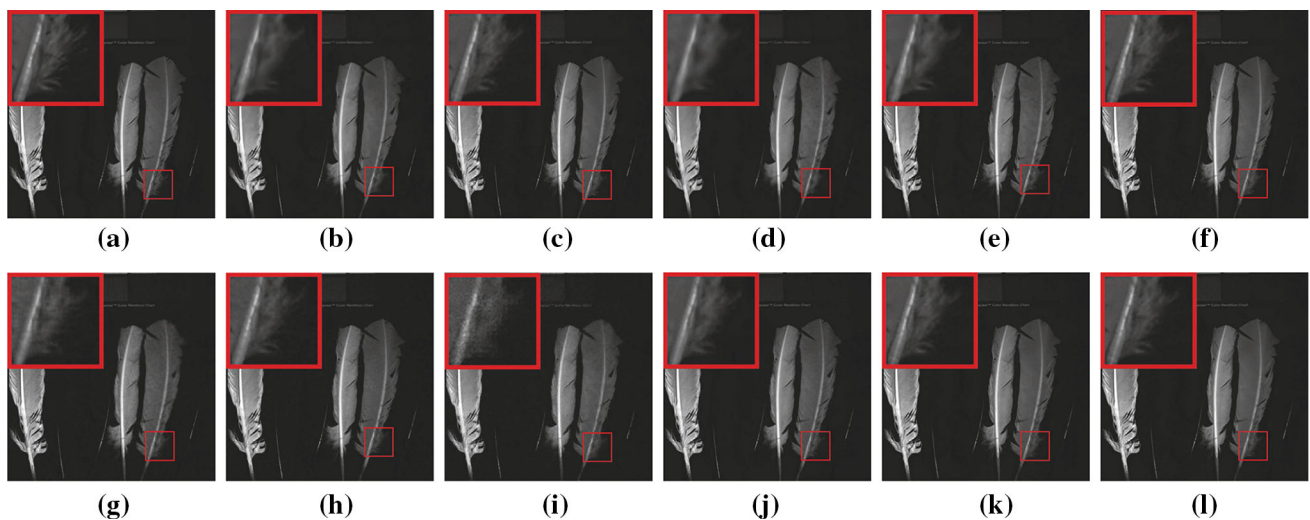


Fig. 4 Visual quality comparison for the HSI feathers under noise level $\sigma_n = 20$. The 620 nm band is shown **a** Original Images **b** BM3D (Dabov et al. 2007) **c** CSR (Dong et al. 2011) **d** ANLM (Manjn et al. 2010) **e**

BM4D (Maggioni et al. 2013) **f** 3DCSR **g** LRRTA (Renard et al. 2008) **h** SDS (Lam et al. 2012) **i** K-SVD (Aharon et al. 2006) **j** DNTDL (Peng et al. 2014) **k** OurBL **l** Ours

$$\{\alpha^{(t)}, \mathbf{X}^{(t)}\} = \arg \min_{\alpha, \mathbf{X}} \|\mathbf{Y} - \mathbf{H}\mathbf{X}\|_2^2 + \sum_{i,j,k} \left(\gamma \|x_{i,j} - \Phi_k \alpha_{i,j}\|_2^2 + \sum_l \lambda_{kl} \|\alpha_{ijl}\|_1 + \sum_l \eta_{kl} \|\alpha_{ijl} - \theta_{ijl}^{(t)}\|_1 \right). \tag{18}$$

Update α For a fixed \mathbf{X} , the optimization of the sparse code α can be described as

$$\alpha^{(t)} = \arg \min_{\alpha} \sum_{i,j} \left(\gamma \|x_{i,j}^{(t-1)} - \Phi_k \alpha_{i,j}\|_2^2 + \sum_l \lambda_{kl} \|\alpha_{ijl}\|_1 + \sum_l \eta_{kl} \|\alpha_{ijl} - \theta_{ijl}^{(t)}\|_1 \right). \tag{19}$$

A bi-variate shrinkage algorithm is employed to solve the function in Eq. (18). In each iteration, we employ an iterative shrinkage operator to update α for fixed θ . Each element in α can be calculated as

$$\alpha_{ijl}^{(t)} = \begin{cases} \mathbf{S}_{\tau_1, \tau_2, \theta_{ijl}}(v_{ijl}^{(t)}), & \theta_{ijl}^{(t)} \geq 0, \\ -\mathbf{S}_{\tau_1, \tau_2, -\theta_{ijl}}(-v_{ijl}^{(t)}), & \theta_{ijl}^{(t)} < 0. \end{cases} \tag{20}$$

where

$$v_{i,j}^{(t)} = \alpha_{i,j}^{(t-1)} - 2c\Phi_k^T (\Phi_k \alpha_{i,j}^{(t-1)} - x_{i,j}^{(t-1)}), \text{ if } (i, j) \in C_k \tag{21}$$

and

$$\tau_1^{(t)} = c\lambda_{kl}^{(t)}/\gamma, \quad \tau_2^{(t)} = c\eta_{kl}^{(t)}/\gamma. \tag{22}$$

where c is an appropriate step size, subscript l denotes the l th element in a vector, and $v_{ijl}^{(t)}$ denote the l th element in the $\mathbf{v}_{i,j}^{(t)}$. when $\theta_{ijl} \geq 0$,

$$\mathbf{S}_{\tau_1, \tau_2, \theta_{ijl}}(v_{ijl}) = \begin{cases} v_{ijl} + \tau_1 + \tau_2, & v_{ijl} < -\tau_1 - \tau_2, \\ 0, & -\tau_1 - \tau_2 \leq v_{ijl} < \tau_1 - \tau_2, \\ v_{ijl} - \tau_1 + \tau_2, & \tau_1 - \tau_2 \leq v_{ijl} < \tau_1 - \tau_2 + \theta_{ijl}, \\ \theta_{ijl}, & \tau_1 - \tau_2 + \theta_{ijl} \leq v_{ijl} < \tau_1 + \tau_2 + \theta_{ijl}, \\ v_{ijl} - \tau_1 - \tau_2, & \tau_1 + \tau_2 + \theta_{ijl} \leq v_{ijl}. \end{cases} \tag{23}$$

Technical details of deriving the bi-variate shrinkage operator $\mathbf{S}_{\tau_1, \tau_2, \theta_{ijl}}$ can be found in Dong et al. (2011). By using this shrinkage operator, the basis vectors in the dictionary are adaptively chosen to represent different image data.

Update \mathbf{X} In this sub-problem, we fix all other variables and optimize \mathbf{X} by

$$\mathbf{X}^{(t)} = \arg \min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{H}\mathbf{X}\| + \gamma \sum_{i,j,k} \|\mathbf{R}_{i,j}\mathbf{X} - \Phi_k \alpha_{i,j}^{(t)}\|_2^2. \tag{24}$$

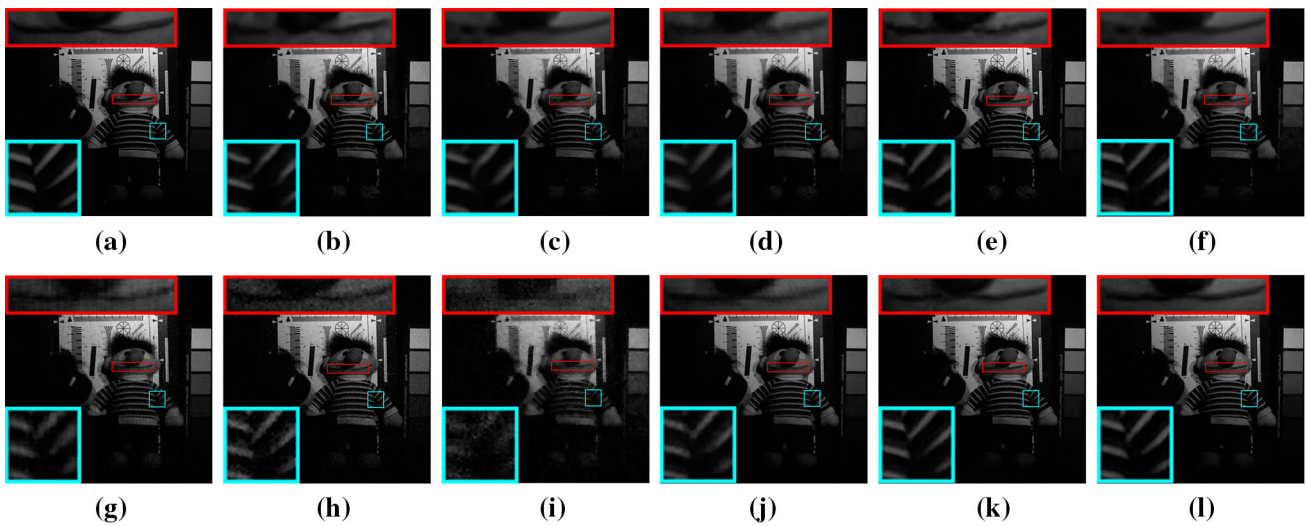


Fig. 5 Visual quality comparison for the HSI *chart* and *stuffed toy* under noise level $\sigma_n = 40$. The 600 nm band is shown **a** Original Images **b** BM3D (Dabov et al. 2007) **c** CSR (Dong et al. 2011) **d** ANLM (Manj

et al. 2010) **e** BM4D (Maggioni et al. 2013) **f** 3DCSR **g** LRTA (Renard et al. 2008) **h** SDS (Lam et al. 2012) **i** K-SVD (Aharon et al. 2006) **j** DNTDL (Peng et al. 2014) **k** OurBL **l** Ours

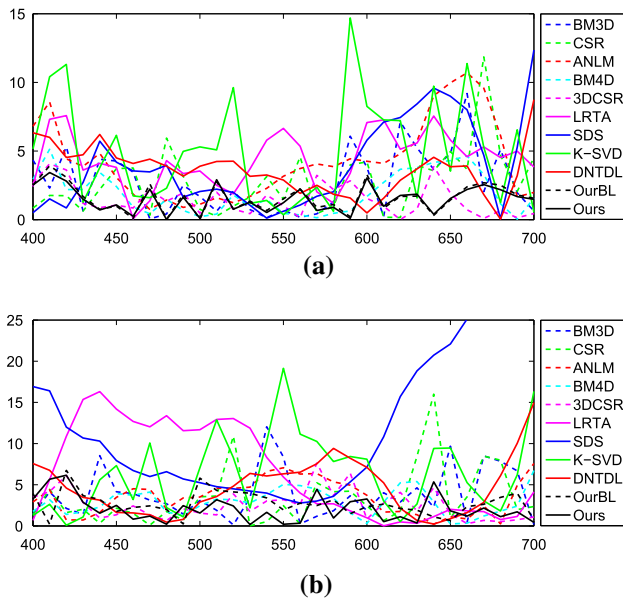


Fig. 6 Absolute difference between the noise-free spectrum and the restoration results of pixel (300, 380) in the *chart* and *stuffed toy* scene under $\sigma_n = 20$ and $\sigma_n = 40$ levels of noise

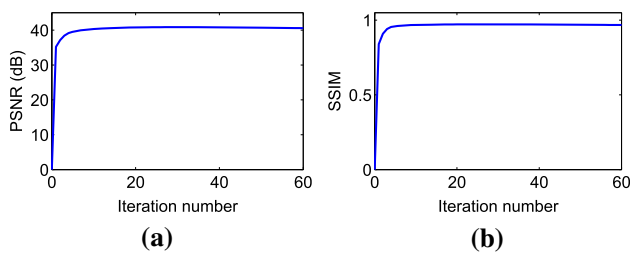


Fig. 7 Verification of the convergence of our method. Progression of the PSNR and SSIM for the *chart* and *stuffed toy* scene under $\sigma_n = 20$ levels of noise

Algorithm 1: Hyperspectral Image Restoration

Input : A degraded hyperspectral image \mathbf{Y}
Output : A restored hyperspectral image \mathbf{X}
Initialization: $\mathbf{X}^{(0)} = \mathbf{Y}$ for HSI denoising, $\mathbf{X}^{(0)} = \text{bicubic}(\mathbf{Y})$ for HSI super-resolution, and $\alpha_{i,j}^{(0)} = \Phi_k^T x_{i,j}^{(0)}$

- 1 **for** $t = 1, 2, \dots, T$ **do**
- 2 Update spatial dictionary $\mathbf{S}_k^{(t)}$ via clustering and PCA;
- 3 Update spectral dictionary $\mathbf{U}^{(t)}$ via PCA;
- 4 Update adaptive spatial-spectral dictionary $\Phi_k^{(t)}$ via Eq. (3);
- 5 Update the $\theta_{i,j}^{(t)}$ via Eqs. (17);
- 6 Compute $v_{i,j}^{(t)}$ via Eq. (21);
- 7 Update the regularization parameters $\lambda_{kl}, \eta_{kl}, \tau_1$, and τ_2 via Eqs. (16) and (22);
- 8 **if** $\text{mod}(t, T_0) = 0$ **then**
- 9 Compute $\alpha_{i,j}^{(t)}$ via shrinkage operator given in Eq. (20);
- 10 Update $\mathbf{X}^{(t)}$ for a given set of sparse codes $\alpha_{i,j}^{(t)}$ by solving Eq. (24);
- 11 **end**
- 12 **end**

Equation (24) is a quadratic minimization problem and thus can be solved in closed-form, i.e.,

$$\mathbf{X}^{(t)} = \left(\mathbf{H}^T \mathbf{H} + \gamma \sum_{i,j,k} \mathbf{R}_{i,j}^T \mathbf{R}_{i,j} \right)^{-1} \left(\mathbf{H}^T \mathbf{Y} + \gamma \sum_{i,j,k} \mathbf{R}_{i,j}^T \Phi_k \alpha_{i,j}^{(t)} \right). \tag{25}$$

In our implementation, we use the conjugate gradient algorithm to compute Eq. (25).

From the above discussion, we can see that non-local self-similarity θ , latent HSI \mathbf{X} , and the sparse coding α are

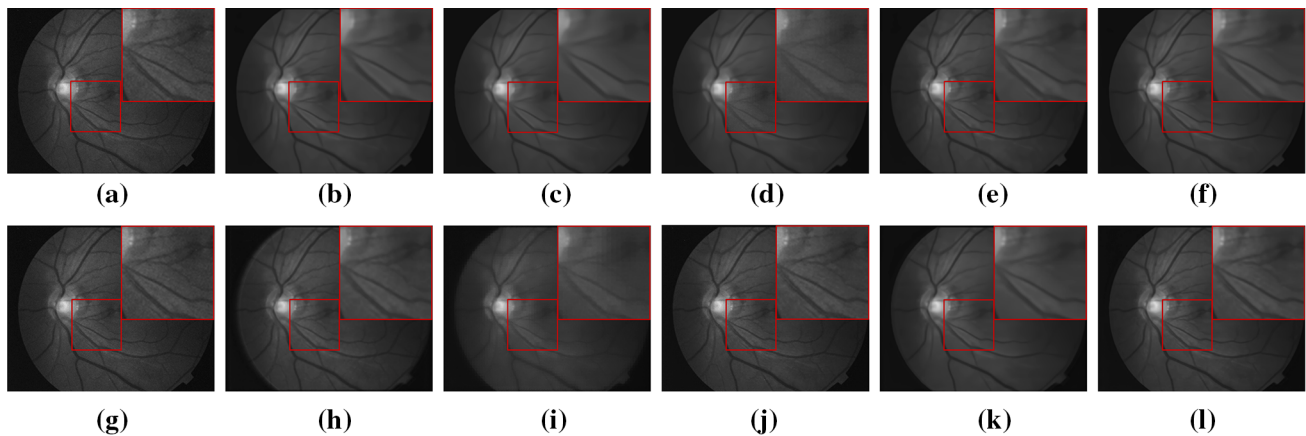


Fig. 8 Visual quality comparison for the real HSI. The 600 nm band is shown **a** Original Images **b** BM3D (Dabov et al. 2007) **c** CSR (Dong et al. 2011) **d** ANLM (Manjn et al. 2010) **e** BM4D (Maggioni et al.

2013) **f** 3DCSR **g** LRTA (Renard et al. 2008) **h** SDS (Lam et al. 2012) **i** K-SVD (Aharon et al. 2006) **j** DNTDL (Peng et al. 2014) **k** OurBL **l** Ours

alternatingly updated in Eq. (11). During the iterations, the accuracy of sparse code $\alpha^{(t)}$ is improved, which in turn, improves the accuracy of non-local self-similarity $\theta^{(t)}$. The clustered similarity patches in the HSI across the spatial domain and the dictionary Φ_k are also updated in terms of the updated $\alpha^{(t)}$. After several iterations, the algorithm converges and the desired sparse code α and latent HSI \mathbf{X} can be obtained. The proposed HSI restoration algorithm is summarized in Algorithm 1.

5 Experimental Results

In this section, extensive experiments are presented to evaluate the performance of the proposed method. We first apply our method to HSI denoising and compare with 8 state-of-the-art methods. Then, our method is also applied to HSI super-resolution. Finally, we show that our method can be widely used for multi-channel images and improve image restoration results.

In our implementation, if not specifically stated, the size of the cubic patches is $7 \times 7 \times B$, the size of the window for searching similar patches is 50×50 , the number of best matched patches is 16, and the number of clusters is 50. We employ PSNR, SSIM (Wang et al. 2004), and FSIM (Zhang et al. 2011) to evaluate the quality of the restored HSI.

5.1 Hyperspectral Image Denoising










We first selected 10 HSIs¹ from the Columbia Multispectral Image Database (Yasuma et al. 2008) to test our method under different noise levels. To demonstrate the effective-

ness of the proposed algorithm, we compared our denoising results with 8 recently developed state-of-the-art denoising methods, including band-wise BM3D (Dabov et al. 2007) (BM3D), band-wise centralized sparse representation (CSR) (Dong et al. 2011), adaptive non-local means denoising for volumetric data (ANLM) (Manjn et al. 2010), non-local transform domain filter for volumetric data (BM4D) (Maggioni et al. 2013), low rank tensor approximation (LRTA) (Renard et al. 2008), spectral domain statistics based method (SDS) (Lam et al. 2012), 3D-cube K-SVD (Aharon et al. 2006) and decomposable non-local tensor dictionary learning (DNTDL) (Peng et al. 2014). All parameters involved in the competing algorithms were optimally assigned or automatically chosen as described in the reference papers. To sufficiently compare with the CSR method, we also extended the CSR method to 3DCSR, where the cubic patches are stacked into long vectors. To show the influence of the number of the clusters on denoising performance, we also implemented our method with a single universal dictionary by setting the number of clusters to one, denoted as OurBL.

In Tables 1 and 2, we provide the PSNR, SSIM, and FSIM results obtained from various methods for each HSI under different noise levels. We can see that the best result for each set is highlighted in bold. The proposed method outperforms other competitive approaches in most cases and achieves state-of-the-art denoising performance in terms of PSNR, SSIM, and FSIM. Comparing our method with OurBL, we see that our method can produce better results, which demonstrates that the proposed adaptive dictionary learning in Sect. 3.1, which captures the characteristics of the scene, can effectively improve the denoising results. We also show more noise levels in terms of the average performance across all images in Table 3.

¹ Due to space limitations, we chose 10 typical and complex scenes to test our method.

Table 4 PSNR (dB), SSIM and FSIM results of the super-resolution hyperspectral image by averaging over all channels under different scaling factors and noise levels

Img.	Metrics	Noiseless				Noise $\sigma_n = 2$			
		Scale=2		Scale=3		Scale=2		Scale=3	
	PSNR	34.23	35.00	31.79	32.55	32.30	33.21	30.38	31.13
	SSIM	0.9657	0.9714	0.9439	0.9515	0.9160	0.9237	0.9058	0.9177
	FSIM	0.9611	0.9684	0.9367	0.9413	0.9214	0.9273	0.9037	0.9123
	PSNR	31.69	32.58	28.02	29.22	29.30	30.25	27.38	28.33
	SSIM	0.9079	0.9299	0.8132	0.8315	0.8427	0.8675	0.7801	0.8012
	FSIM	0.9354	0.9468	0.8705	0.8837	0.8923	0.9043	0.8536	0.8636
	PSNR	43.78	44.38	41.94	42.86	38.64	39.67	38.41	38.28
	SSIM	0.9826	0.9860	0.9767	0.9794	0.9262	0.9333	0.9336	0.9401
	FSIM	0.9841	0.9887	0.9771	0.9821	0.9204	0.9279	0.9369	0.9407
	PSNR	39.95	40.42	37.86	38.43	36.54	37.12	35.99	36.57
	SSIM	0.9750	0.9795	0.9661	0.9698	0.9213	0.9280	0.9297	0.9352
	FSIM	0.9985	0.9990	0.9958	0.9962	0.9779	0.9795	0.9776	0.9796
	PSNR	38.44	39.02	34.85	35.46	35.37	35.94	33.41	34.42
	SSIM	0.9818	0.9858	0.9667	0.9735	0.9287	0.9327	0.9305	0.9377
	FSIM	0.9977	0.9980	0.9886	0.9897	0.9776	0.9792	0.9719	0.9786
	PSNR	45.55	46.24	42.95	43.66	41.25	41.88	40.24	41.09
	SSIM	0.9932	0.9946	0.9891	0.9902	0.9398	0.9410	0.9396	0.9425
	FSIM	0.9909	0.9931	0.9819	0.9852	0.9398	0.9450	0.9337	0.9403
	PSNR	31.53	31.67	30.96	31.27	30.32	30.58	30.23	30.53
	SSIM	0.7624	0.7803	0.7394	0.7586	0.7185	0.7235	0.7131	0.7222
	FSIM	0.8669	0.8893	0.8455	0.8679	0.8673	0.8888	0.8457	0.8663
	PSNR	42.02	43.09	40.80	41.78	38.57	39.65	37.90	38.98
	SSIM	0.9621	0.9715	0.9513	0.9611	0.9208	0.9312	0.9187	0.9292
	FSIM	0.9778	0.9855	0.9690	0.9773	0.9591	0.9677	0.9538	0.9622
	PSNR	37.24	37.90	33.89	34.53	35.03	35.72	33.18	33.86
	SSIM	0.9435	0.9525	0.9092	0.9172	0.8995	0.9112	0.8878	0.8963
	FSIM	0.9623	0.9698	0.9296	0.9352	0.9356	0.9411	0.9208	0.9276
	PSNR	33.74	34.30	30.09	30.91	30.94	31.63	29.17	29.78
	SSIM	0.9403	0.9506	0.8808	0.9020	0.8844	0.8943	0.8508	0.8613
	FSIM	0.9523	0.9605	0.9011	0.9178	0.9164	0.9243	0.8861	0.8989

For each case, the first column shows the results from [Dong et al. \(2013\)](#), and the second column from our method

To visually illustrate the denoising performance of our method, we show two denoised hyperspectral bands from different methods under different levels of noise in Figs. 4 and 5. When the noise level is not very high ($\sigma_n=20$), as shown in Fig. 4, all the competing methods achieve good denoising outputs, but when we enlarge the images and compare the details demarcated in Fig. 4a. We can see that our method recovers nice texture/edge features with rich details. When the noise level is high ($\sigma_n = 40$), as shown in Fig. 5,

LRTA ([Renard et al. 2008](#)), and SDS ([Lam et al. 2012](#)), and 3D-cube K-SVD ([Aharon et al. 2006](#)) tend to generate many visual artifacts. BM3D ([Dabov et al. 2007](#)), CSR ([Dong et al. 2011](#)), ANLM ([Manjñ et al. 2010](#)), BM4D ([Maggioni et al. 2013](#)), and 3DCSR are overly smooth and show many block artifacts. DNTDL ([Peng et al. 2014](#)) and our method work better in this case. Comparing these two methods in detail, the denoised image (Fig. 5l) by our method has much less artifacts than other methods, and is visually more appealing.

The restoration errors across spectra are shown in Fig. 6. The absolute differences in the spectral distributions between the noise-free spectrum and the restoration results of all competing methods at pixel (300, 380) in the *chart and stuffed toy* scene (the same scene in Fig. 5) under $\sigma_n = 20$ and $\sigma_n = 40$ are shown. It is easy to see that our method obtains the best approximation of the true spectral distributions of the original HSI, which is in accordance with our quantitative evaluation.

In addition, we also take the case of HSI denoising when $\sigma_n = 20$ for the *chart and stuffed toy* scene as an example to show the convergence of our method. Figure 7 plots the evaluations of PSNR and SSIM versus iteration numbers for the tested HSI. We can see that both the PSNR and SSIM curves increase monotonically and demonstrate the convergence of our method.

To further demonstrate the denoising performance of our method, we also tested our method on a real HSI, which is an eye fundus that was captured from 500 to 720 nm at 10 nm intervals. This HSI contains 23 bands of size 256×320 . We illustrate the experimental results of an example image located at 600 nm in Fig. 8. To clearly show the experimental results, a part of the image is enlarged for all restored images from all competing methods. We can see that our method effectively removes the noise and properly preserves the underlying structure of the image, while most of the other competing methods make the image blurry or overly sharp as compared with the original image.

5.2 Hyperspectral Image Super-Resolution

We designed a general model for HSI restoration so that it could be applied to more than just denoising. We now apply our model to HSI super-resolution. Most recent HSI super-resolution methods require extra images such as multiple frames (Akgun et al. 2005; Buttingsrud and Alsberg 2006; Chan et al. 2010; Zhang et al. 2012), a high-resolution mono-band image (Eismann and Hardie 2004, 2005; Chen et al. 2014; Zhao et al. 2011), or an RGB image (Ma et al. 2013). In our method, we achieve HSI super-resolution on low resolution HSI without the need for additional information. We do not compare against previous methods, since the inputs are vastly different and comparisons would be unfair. In our experiments, we compare our method with Dong et al. (2013), which used similar constraints to our method, but was designed for grayscale images and did not consider the correlation across spectra.

Since our method is based on adaptive spatial-spectral dictionary learning, it can be easily incorporated into most models that utilize sparse representations with different constraints. Here, we mainly show the extensibility of our model to different HSI restoration problems and also the importance of considering correlation in the spectral domain.

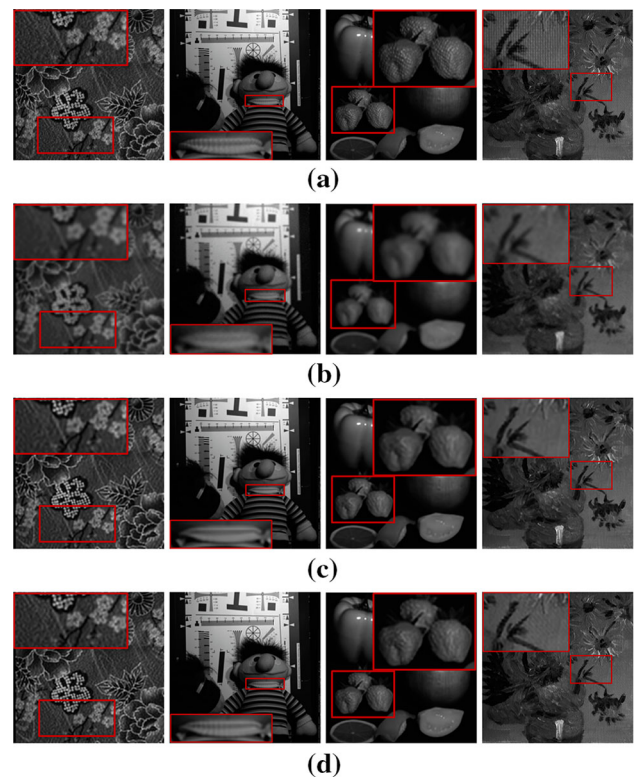


Fig. 9 Reconstructed super-resolution images for the low resolution images. Four scenes are shown at 620 nm. From *top to bottom* original high resolution image; low resolution images; reconstructed super-resolution images by (Dong et al. 2013) for each channel separately; reconstructed super-resolution images by our method. From *left to right*: scaling factor 2 and $\sigma_n = 0$; scaling factor 2 and $\sigma_n = 2$; scaling factor 3 and $\sigma_n = 0$; scaling factor 3 and $\sigma_n = 2$ **a** Original high resolution images **b** Low resolution images **c** Reconstructed super-resolution images by Dong et al. (2013) **d** Reconstructed super-resolution images by our method

We first simulated low resolution images (e.g. Fig. 9b) by blurring high resolution images with a 7×7 Gaussian kernel with standard deviation 1.6, and then downsampled the blurred images by a scaling factor of 2 or 3 in both the horizontal and vertical directions. To make the HSI super-resolution problem more challenging, additive Gaussian noise of standard deviation 2 was also added to the low resolution images.

Table 4 shows that PSNR (dB), SSIM, and FSIM results of the super-resolution hyperspectral image by averaging over all channels under different scaling factors and noise levels. For each case, the first column shows the results from Dong et al. (2013), and the second column from our method. We can see that our method outperforms the image restoration method in Dong et al. (2013) in PSNR, SSIM, and FSIM. It demonstrates that our method, which is based on adaptive spatial-spectral dictionary learning effectively improves the quality of restored results by considering the high correlation across spectra. Visual comparisons between Dong et al.

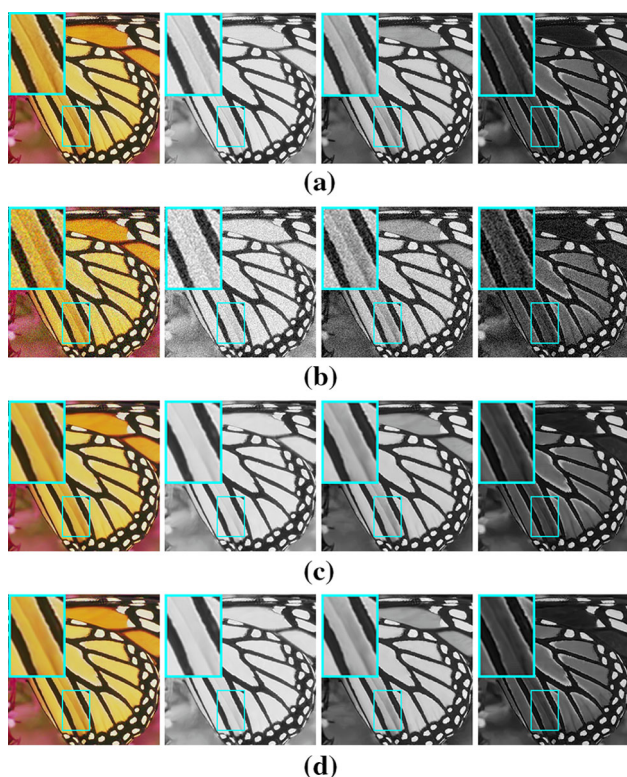


Fig. 10 Effects of high correlation across bands on the *Butterfly* scene. **a** Original image, from left to right RGB image, R channel, G channel and B channel. **b** The corresponding noise images ($\sigma_n = 20$) for each channel. **c** The denoised images for each channel separately. R channel: PSNR = 29.94, SSIM = 0.9368, FSIM = 0.9396. G channel: PSNR = 29.91, SSIM = 0.9315, FSIM = 0.9317. B channel: PSNR = 29.60, SSIM = 0.8942, FSIM = 0.9246. **d** The denoised images from our method that considers the high correlation across spectra. R channel: PSNR = 31.74, SSIM = 0.9528, FSIM = 0.9516. G channel: PSNR = 31.82, SSIM = 0.9472, FSIM = 0.9418. B channel: PSNR = 31.12, SSIM = 0.9241, FSIM = 0.9446

(2013) and our method are shown in Fig. 9. Our method reconstructs better visual high resolution results, especially for the textured/edge parts.

5.3 Applicability to Multi-channel Image Restoration

We have so far applied our method to HSI denoising and super-resolution. One of the benefits of our method is that it considers correlation in the spectral domain. In this section, we show that multi-band images such as RGB images also benefit from our method.

Most image restoration methods for RGB images only pay attention to the intensity channel (the gray image) or restore each channel separately. We apply our method to all channels together, each channel separately, and compare their denoising and super-resolution results. In addition, we also apply our method in YUV space for RGB image super-resolution,

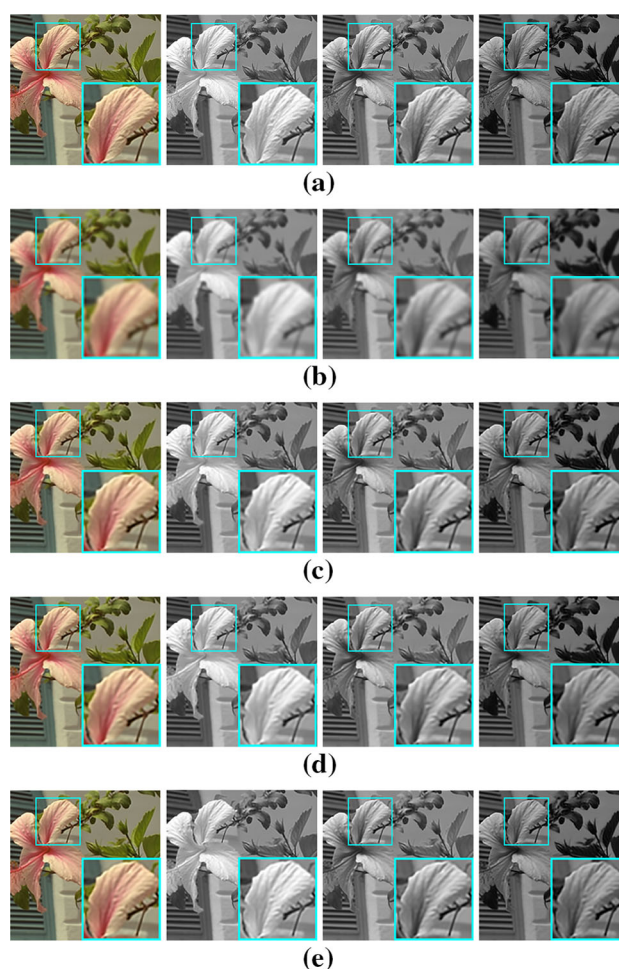











Fig. 11 Effects of high correlation across bands on the *Flower* scene. **a** Original image, from left to right RGB image, R channel, G channel and B channel. **b** The corresponding low-resolution images (scale = 2, $\sigma_n = 0$) for each channel. **c** The super-resolution images on the YUV space by apply the super-resolution method to the Y component and using bicubic interpolation for UV components. R channel: PSNR = 29.90, SSIM = 0.8901, FSIM = 0.9235. G channel: PSNR = 30.03, SSIM = 0.8946, FSIM = 0.9236. B channel: PSNR = 29.71, SSIM = 0.8863, FSIM = 0.9231. **d** The super-resolution images for each channel separately. R channel: PSNR = 30.48, SSIM = 0.9021, FSIM = 0.9303. G channel: PSNR = 30.45, SSIM = 0.9043, FSIM = 0.9306. B channel: PSNR = 30.60, SSIM = 0.9023, FSIM = 0.9324. **e** The super-resolution images from our method that considers the high correlation across spectra. R channel: PSNR = 31.36, SSIM = 0.9209, FSIM = 0.9431. G channel: PSNR = 31.17, SSIM = 0.9195, FSIM = 0.9412. B channel: PSNR = 31.30, SSIM = 0.9174, FSIM = 0.9436

i.e. performing our method only on the intensity channel Y, and using bicubic interpolation for the U and V channels.

Figures 10 and 11 show the effects of correlation across bands in image denoising and super-resolution. In Fig. 10, the visual quality of the denoising results for the *Butterfly* scene is shown. Each channel in the original image (Fig. 10a) is noised using Gaussian noise with $\sigma_n = 20$ (Fig. 10b). Comparing the denoised results from processing each channel separately (Fig. 10c) with our method (Fig. 10d), it is apparent that both

Table 5 PSNR (dB), SSIM and FSIM results of the restored image under different noise levels and scaling factors

Img.	Metrics	Denoising						Super-resolution, scale = 2					
		$\sigma_n = 20$		$\sigma_n = 30$		$\sigma_n = 40$		Noiseless			Noise, $\sigma_n = 2$		
	PSNR	28.32	30.61	25.98	28.18	24.34	26.52	25.59	26.03	26.64	24.42	24.64	25.35
	SSIM	0.8799	0.9270	0.8182	0.8839	0.7888	0.8449	0.8578	0.8712	0.8897	0.8040	0.8228	0.8345
	FSIM	0.9202	0.9478	0.8842	0.9206	0.8505	0.8975	0.8959	0.9028	0.9144	0.8641	0.8737	0.8856
	PSNR	29.82	31.56	27.80	29.41	26.06	27.47	28.72	29.42	30.02	27.00	27.32	28.01
	SSIM	0.9209	0.9414	0.8914	0.9183	0.8590	0.8855	0.9281	0.9382	0.9466	0.8756	0.9018	0.9133
	FSIM	0.9320	0.9460	0.9102	0.9231	0.8772	0.8855	0.9251	0.9332	0.9398	0.8747	0.8970	0.9120
	PSNR	30.05	32.34	27.89	30.11	26.37	28.47	29.88	30.51	31.28	28.58	28.90	29.65
	SSIM	0.8711	0.9229	0.8141	0.8842	0.7629	0.8454	0.8903	0.9029	0.9193	0.8404	0.8625	0.8804
	FSIM	0.9160	0.9475	0.8831	0.9237	0.8526	0.9007	0.9234	0.9311	0.9426	0.8963	0.9079	0.9193
	PSNR	29.95	30.88	28.83	29.73	28.07	28.94	31.12	31.59	31.85	30.38	30.63	30.95
	SSIM	0.7208	0.7833	0.6682	0.7264	0.6305	0.6839	0.7779	0.7957	0.8098	0.7436	0.7522	0.7579
	FSIM	0.8618	0.9010	0.8324	0.8733	0.8066	0.8457	0.9004	0.9120	0.9226	0.8919	0.8869	0.9013
	PSNR	31.58	33.17	29.69	31.16	28.39	29.64	31.05	31.70	32.20	30.00	30.21	30.83
	SSIM	0.8665	0.9024	0.8187	0.8632	0.7809	0.8258	0.8933	0.8887	0.9002	0.8253	0.8458	0.8604
	FSIM	0.9020	0.9256	0.8684	0.8995	0.8370	0.8717	0.9138	0.9218	0.9317	0.8864	0.8979	0.9131
	PSNR	30.26	32.18	27.97	29.81	26.00	27.86	27.96	29.70	30.27	26.68	26.64	27.55
	SSIM	0.9571	0.9714	0.9342	0.9530	0.9047	0.9271	0.9512	0.9661	0.9708	0.9151	0.9309	0.9444
	FSIM	0.9472	0.9594	0.9283	0.9374	0.8998	0.9060	0.9334	0.9497	0.9555	0.8854	0.9080	0.9141
	PSNR	32.31	33.87	30.38	31.91	28.93	30.47	31.12	31.83	32.42	29.92	30.17	30.98
	SSIM	0.8947	0.9155	0.8660	0.8898	0.8443	0.8705	0.9178	0.9329	0.9411	0.8827	0.8974	0.9164
	FSIM	0.9367	0.9502	0.9198	0.9355	0.9061	0.9239	0.9514	0.9579	0.9640	0.9300	0.9415	0.9483
	PSNR	32.73	33.02	30.65	31.01	29.12	29.63	31.43	34.73	35.08	30.70	32.05	32.58
	SSIM	0.8623	0.8691	0.8129	0.8171	0.7886	0.7898	0.8692	0.9381	0.9445	0.8473	0.8623	0.8701
	FSIM	0.9088	0.9164	0.8777	0.8866	0.8486	0.8632	0.9060	0.9473	0.9528	0.8929	0.9047	0.9101
	PSNR	29.04	31.54	27.39	29.57	26.32	28.27	29.59	29.84	30.61	28.25	28.67	28.43
	SSIM	0.7760	0.8751	0.6939	0.8090	0.6294	0.7523	0.8286	0.8396	0.8706	0.7724	0.7833	0.8178
	FSIM	0.8701	0.9279	0.8283	0.8941	0.7918	0.8679	0.9051	0.9091	0.9269	0.8851	0.8835	0.8988

For the denoising problem, the first column shows the results from denoising each channel separately and the second column from our method. For the super-resolution problem, the first column shows the results by applying the method in YUV space, the second column from restoring each channel separately in RGB space, and the third column from our method

methods generate good results in the smooth regions, but we can see that our method obtains much better results in the textured/edge regions. This demonstrates that the quality of the restored image can be effectively improved by using the correlation across bands.

By using our method on RGB image super-resolution, Fig. 11 further demonstrates that our method (Fig. 11e) obtains much better restoration results in the textured/edge regions, compared with performing super-resolution on intensity channel Y only (Fig. 11c) or each channel separately (Fig. 11d). Thus exploiting the correlation across bands also provides better super-resolution results for RGB images.

PSNR, SSIM, and FSIM results on 9 RGB images are shown in Table 5. From Table 5, we can see that our method outperforms restoring for each channel separately or only by applying it to the intensity channel. Thus exploiting the high correlation across spectra provides better RGB restoration results.

6 Conclusion

In this paper, we presented an effective HSI restoration method based on adaptive spatial-spectral dictionary learning, which considered the underlying characteristics of HSIs: the sparsity across the spatial-spectral domain, the high correlation across spectra, and non-local self-similarity over space. We exploited the high correlation across spectra and non-local self-similarity over space to learn the adaptive spatial-spectral dictionary for each overlapping cubic patch of the HSI. By design, each adaptive spatial-spectral dictionary could sparsely represent its own cubic patch across spatial and spectral domains while sharing features with other similar non-local patches across the full HSI. Then, the local and non-local sparsity of the HSI was used to design a HSI restoration model under the trained dictionary. This model was effectively solved by our numerical algorithm with its regularization parameters

adaptively adjusted for different clusters and different noise levels.

The experimental results on HSI denoising showed that the proposed method outperformed many state-of-the-art methods under several comprehensive quantitative assessments. We also demonstrated our method in super-resolution for HSIs and showed that our method restores high resolution images well, especially for detailed parts. While not the main aim of this paper, we also showed that our approach can be applied to multi-band images such as RGB with better results.

We presented a general model that was applied to HSI denoising and super-resolution. Despite its effectiveness, it still has some limitations. For example, we did not consider the complexity of mixing different kinds of noise for denoising (Zhao et al. 2015). For super-resolution, the option to use extra information from multiple frames (Zhang et al. 2012), or a high-resolution mono-band image (Chen et al. 2014) could also be beneficial. In addition, since the spatial-spectral basis can be effectively used for HSI reconstruction from raw data, e.g. extended Bayer filter (Monno et al. 2013), it is worth investigating how we can extend our adaptive spatial-spectral dictionary learning to HSI reconstruction from raw data. In the future, we will investigate these directions and design possible extensions to our model that are specifically adapted for each case to achieve improved empirical results.

References

- Aharon, M., Elad, M., & Bruckstein, A. (2006). K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Image Processing*, 15(11), 4311–4322.
- Akgun, T., Altunbasak, Y., & Mersereau, R. (2005). Super-resolution reconstruction of hyperspectral images. *IEEE Transactions on Image Processing*, 14(11), 1860–1875.
- Atkinson, I., Kamalabadi, F. & Jones, D. (2003). Wavelet-based hyperspectral image estimation. In *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2, 743–745.
- Banerjee, A., Burlina, P. & Broadwater, J. (2009). Hyperspectral video for illumination-invariant tracking. Evolution in Remote Sensing (WHISPERS). In *Workshop on Hyperspectral Image and Signal Processing* (pp. 1–4).
- Bioucas-Dias, J., & Figueiredo, M. (2007). A new TwIST: Two-step iterative shrinkage/thresholding algorithms for image restoration. *IEEE Transactions on Image Processing*, 16(12), 2992–3004.
- Borengasser, M., Hungate, W. S., & Watkins, R. (2007). *Hyperspectral remote sensing: Principles and applications*. Boca Raton: CRC Press.
- Bourennane, S., Fossati, C., & Cailly, A. (2010). Improvement of classification for hyperspectral images based on tensor modeling. *IEEE Geoscience and Remote Sensing Letters*, 7(4), 801–805.
- Buades, A., Coll, B. & Morel, J. M. (2005). A non-local algorithm for image denoising. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2) (pp 60–65).
- Buades, A., Coll, B., Morel, J. M., & Sbert, C. (2009). Self-similarity driven color demosaicking. *IEEE Transactions on Image Processing*, 18(6), 1192–1202.
- Buttingsrud, B. & Alsberg, B. (2006). Superresolution of hyperspectral images. *Chemometrics and Intelligent Laboratory Systems* (pp 62–68).
- Candes, E., & Tao, T. (2006). Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE Transactions on Information Theory*, 52(12), 5406–5425.
- Castrodad, A., Xing, Z., Greer, J., Bosch, E., Carin, L., & Sapiro, G. (2011). Learning discriminative sparse representations for modeling, source separation, and mapping of hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 49(11), 4263–4281.
- Chakrabarti, A. & Zickler, T. (2011). Statistics of real-world hyperspectral images. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp 193–200).
- Chan, J., Ma, J., Kempeneers, P., & Canters, F. (2010). Superresolution enhancement of hyperspectral CHRIS/proba images with a thin-plate spline nonrigid transform model. *IEEE Transaction on Geoscience and Remote Sensing*, 48(6), 2569–2579.
- Chen, C., Yeqing, L., Wei, L., Junzhou, H. (2014). Image fusion with local spectral consistency and dynamic gradient sparsity. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Chen, G., & Qian, S. E. (2011). Denoising of hyperspectral imagery using principal component analysis and wavelet shrinkage. *IEEE Transactions on Geoscience and Remote Sensing*, 49(3), 973–980.
- Chen, S., Donoho, D., & Saunders, M. (1998). Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20(1), 33–61.
- Dabov, K., Foi, A., Katkovnik, V., & Egiazarian, K. (2007). Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8), 2080–2095.
- David, T., & Dicker, J. L. (2006). Differentiation of normal skin and melanoma using high resolution hyperspectral imaging. *Cancer Biology and Therapy*, 5(8), 1033–1038.
- Dong, W., Li, X., Zhang, L. & Shi, G. (2011). Sparsity-based image denoising via dictionary learning and structural clustering. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp 457–464).
- Dong, W., Zhang, L., Shi, G., & Li, X. (2013). Nonlocally centralized sparse representation for image restoration. *IEEE Transactions on Image Processing*, 22(4), 1620–1630.
- Dong, W., Shi, G., Li, X., & Ma, Y. (2015). Image restoration via simultaneous sparse coding: Where structured sparsity meets gaussian scale mixture. *International Journal of Computer Vision*, 114(2), 217–232.
- Dong, W., Fu, F., Shi, G., Cao, X., Wu, J., Li, G., et al. (2016). Hyperspectral image super-resolution via non-negative structured sparse representation. *IEEE Transactions on Image Processing*, 25(5), 2337–2352.
- Donoho, D. L. (2004). For most large underdetermined systems of linear equations the minimal l1-norm solution is also the sparsest solution. *Communications on Pure and Applied Mathematics*, 59, 797–829.
- Eismann, M., & Hardie, R. (2004). Application of the stochastic mixing model to hyperspectral resolution enhancement. *IEEE Transactions on Geoscience and Remote Sensing*, 42(9), 1924–1933.
- Eismann, M., & Hardie, R. (2005). Hyperspectral resolution enhancement using high-resolution multispectral imagery with arbitrary response functions. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3), 455–465.
- Elad, M., & Aharon, M. (2006). Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing*, 15(12), 3736–3745.
- Glasner, D., Bagon, S. & Irani, M. (2009). Super-resolution from a single image. In *Proceedings of International Conference on Computer Vision (ICCV)* (pp 349–356).

- Guangyi, Chen, & Q, S. E. (2009). Denoising and dimensionality reduction of hyperspectral imagery using wavelet packets, neighbour shrinking and principal component analysis. *International Journal of Remote Sensing*, 30(18), 4889–4895.
- Guo, X., Huang, X., Zhang, L., & Zhang, L. (2013). Hyperspectral image noise reduction based on rank-1 tensor decomposition. *ISPRS Journal of Photogrammetry and Remote Sensing*, 83, 50–63.
- Gupta, N., & Ramella-Roman, J.C. (2008). Detection of blood oxygen level by noninvasive passive spectral imaging of skin. In *Proceedings of SPIE*.
- Hou, B., Zhang, X., Ye, Q., & Zheng, Y. (2013). A novel method for hyperspectral image classification based on laplacian eigenmap pixels distribution-flow. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(3), 1602–1618.
- Ingrid Daubechies, M. D. (2004). An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, 57(11), 1413–1457.
- Kang, X., Li, S., & Benediktsson, J. (2014). Feature extraction of hyperspectral images with image fusion and recursive filtering. *IEEE Transactions on Geoscience and Remote Sensing*, 52(6), 3742–3752.
- Karami, A., Yazdi, M., & Zolghadre Asli, A. (2011). Noise reduction of hyperspectral images using kernel non-negative tucker decomposition. *IEEE Journal of Selected Topics in Signal Processing*, 5(3), 487–493.
- Lam, A., Sato, I., Sato, Y. (2012). Denoising hyperspectral images using spectral domain statistics. In *Proceedings of International Conference on Pattern Recognition (ICPR)* (pp. 477–480).
- Letexier, D., & Bourennane, S. (2008). Noise removal from hyperspectral images by multidimensional filtering. *IEEE Transactions on Geoscience and Remote Sensing*, 46(7), 2061–2069.
- Lu, G., & Fei, B. (2014). Medical hyperspectral imaging: A review. *Journal of Biomedical Optics*, 19(1), 010901.
- Ma, C., Cao, X., Tong, X., Dai, Q., & Lin, S. (2013). Acquisition of high spatial and spectral resolution video with a hybrid camera system. *International Journal of Computer Vision*, 110, 141–155.
- Maggioni, M., Katkovnik, V., Egiazarian, K., & Foi, A. (2013). Non-local transform-domain filter for volumetric data denoising and reconstruction. *IEEE Transactions on Image Processing*, 22(1), 119–133.
- Mairal, J., Bach, F., Ponce, J. & Sapiro, G. (2009). Online dictionary learning for sparse coding. In *Proceedings of International Conference on Machine Learning (ICML)* (pp 689–696).
- Mairal, J., Bach, F., Ponce, J., Sapiro, G. & Zisserman, A. (2009). Non-local sparse models for image restoration. In *Proceedings of International Conference on Computer Vision (ICCV)* (pp 2272–2279).
- Manjn, J. V., Coup, P., Mart-Bonmat, L., Collins, D. L., & Robles, M. (2010). Adaptive non-local means denoising of MR images with spatially varying noise levels. *Journal of Magnetic Resonance Imaging*, 31(1), 192–203.
- Melgani, F., & Bruzzone, L. (2004). Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, 42(8), 1778–1790.
- Monno, Y., Tanaka, M., Okutomi, M. (2013). Direct spatio-spectral datacube reconstruction from raw data using a spatially adaptive spatio-spectral basis. In *Proceedings of SPIE 8660, Digital Photography IX* vol 8660 (pp. 866,003–866,003–8)
- Murakami, Y., Fukura, K., Yamaguchi, M., & Ohyama, N. (2008). Color reproduction from low-SNR multispectral images using spatio-spectral Wiener estimation. *Optics Express*, 16(6), 4106.
- Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583), 607–609.
- Olshausen, B. A., & Field, D. J. (1997). Sparse coding with an overcomplete basis set: A strategy employed by v1? *Vision Research*, 37(23), 3311–3325.
- Othman, H., & Qian, S. E. (2006). Noise reduction of hyperspectral imagery using hybrid spatial-spectral derivative-domain wavelet shrinkage. *IEEE Transactions on Geoscience and Remote Sensing*, 44(2), 397–408.
- Parmar, M., Lansel, S. & Wandell, B. (2008). Spatio-spectral reconstruction of the multispectral datacube using sparse recovery. In *Proceedings of IEEE International Conference on Image Processing (ICIP)* (pp 473–476).
- Peng, Y., Meng, D., Xu, Z., Gao, C., Yang, Y. & Zhang, B. (2014). Decomposable nonlocal tensor dictionary learning for multispectral image denoising. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Qian, Y., & Ye, M. (2013). Hyperspectral imagery restoration using non-local spectral-spatial structured sparse representation with noise estimation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(2), 499–515.
- Qian, Y., Shen, Y., Ye, M. & Wang, Q. (2012). 3-D Nonlocal means filter with noise estimation for hyperspectral imagery denoising. In *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 1345–1348).
- Renard, N., & Bourennane, S. (2008). Improvement of target detection methods by multiway filtering. *IEEE Transactions on Geoscience and Remote Sensing*, 46(8), 2407–2417.
- Renard, N., Bourennane, S., & Blanc-Talon, J. (2008). Denoising and dimensionality reduction using multilinear tools for hyperspectral images. *IEEE Geoscience and Remote Sensing Letters*, 5(2), 138–142.
- Rowe, R., Nixon, K. & Corcoran, S. (2005). Multispectral fingerprint biometrics. In *Proceedings of IEEE Information Assurance Workshop (IAW)* (pp 14–20).
- Shimano, M., Okabe, T., Sato, I. & Sato, Y. (2011). Video temporal super-resolution based on self-similarity. In *Proceedings of Asian Conference on Computer Vision (ACCV)* (pp 93–106).
- Stamatas, G.N., Balas, C.J., Kollias, N. (2003). Hyperspectral image acquisition and analysis of skin. In *Proceedings of SPIE* (pp. 77–82).
- Teke, M., Deveci, H., Haliloglu, O., Gurbuz, S. & Sakarya, U. (2013). A short survey of hyperspectral remote sensing applications in agriculture. In *International Conference on Recent Advances in Space Technologies (RAST)* (pp 171–176).
- Vrhel, M. J., Gershon, R., & Iwan, L. S. (1994). Measurement and analysis of object reflectance spectra. *Color Research and Application*, 19(1), 4–9.
- Wang, Y. & Niu, R. (2009). Hyperspectral urban remote sensing image smoothing and enhancement using forward-and-backward diffusion. In *Joint Urban Remote Sensing Event* (pp. 1–5).
- Wang, Y., Niu, R., & Yu, X. (2010). Anisotropic diffusion for hyperspectral imagery enhancement. *IEEE Sensors Journal*, 10(3), 469–477.
- Wang, Z., Bovik, A., Sheikh, H., & Simoncelli, E. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transaction on Image Processing*, 13(4), 600–612.
- Yang, C. Y., Huang, J. B. & Yang, M. H. (2011). Exploiting self-similarities for single frame super-resolution. In *Proceedings of Asian Conference on Computer Vision (ACCV)* (pp 497–510).
- Yang, J., Wright, J., Huang, T., & Ma, Y. (2010). Image super-resolution via sparse representation. *IEEE Transactions on Image Processing*, 19(11), 2861–2873.
- Yasuma, F., Mitsunaga, T., Iso, D. & Nayar, S. (2008). Generalized assorted pixel camera: Post-capture control of resolution, dynamic range and spectrum. Tech. rep.
- Yuan, Q., Zhang, L., & Shen, H. (2012). Hyperspectral image denoising employing a spectral-spatial adaptive total variation model. *IEEE*

Transactions on Geoscience and Remote Sensing, 50(10), 3660–3677.

- Zhang, H., Zhang, L., & Shen, H. (2012). A super-resolution reconstruction algorithm for hyperspectral images. *Signal Processing*, 92(9), 2082–2096.
- Zhang, H., He, W., Zhang, L., Shen, H., & Yuan, Q. (2014). Hyperspectral image restoration using low-rank matrix recovery. *IEEE Transactions on Geoscience and Remote Sensing*, 52(8), 4729–4743.
- Zhang, L., Zhang, D., Mou, X., & Zhang, D. (2011). FSIM: A feature similarity index for image quality assessment. *IEEE Transactions on Image Processing*, 20(8), 2378–2386.
- Zhao, Q., Meng, D., Kong, X., Xie, Q., Cao, W., Wang, Y., & Xu, Z. (2015). A novel sparsity measure for tensor recovery. In *Proceedings of International Conference on Computer Vision (ICCV)* (pp 271–279).
- Zhao, Y., Yang, J., Zhang, Q., Song, L., Cheng, Y., & Pan, Q. (2011). Hyperspectral imagery super-resolution by sparse representation and spectral regularization. *EURASIP Journal on Advances in Signal Processing*, 1, 1–10.
- Zhong, P., & Wang, R. (2013). Multiple-spectral-band CRFs for denoising junk bands of hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 51(4), 2260–2275.
- Zibulevsky, M., & Elad, M. (2010). L1–l2 optimization in signal and image processing. *IEEE Transactions on Signal Processing Magazine*, 27(3), 76–88.



Ying Fu received the B.S. degree in Electronic Engineering from Xidian University in 2009, the M.S. degree in Automation from Tsinghua University in 2012, and the Ph.D. degree in Information Science and Technology from the University of Tokyo in 2015. She is currently a project research associate in the Institute of Industrial Science, the University of Tokyo. Her research interests include physics-based vision, image processing, and computational photography.



Antony Lam received the B.S. degree in Computer Science at the California State Polytechnic University, Pomona in 2004 and the Ph.D. in Computer Science at the University of California, Riverside in 2010. After working at the National Institute of Informatics, Japan, he joined the Graduate School of Science and Engineering at Saitama University as an assistant professor in 2014. His research interests are mainly in computer vision with emphasis on the areas of physics-

based vision and pattern recognition.



Imari Sato received the B.S. degree in policy management from Keio University in 1994. After studying at the Robotics Institute of Carnegie Mellon University as a visiting scholar, she received the M.S. and Ph.D. degrees in interdisciplinary Information Studies from the University of Tokyo in 2002 and 2005, respectively. In 2005, she joined the National Institute of Informatics, where she is currently a professor. Her primary research interests are in the fields of computer vision

(physics-based vision, image-based modeling) and Computer Graphics (image-based rendering, augmented reality). She has received various research awards, including IPSJ Nagao Special Researcher award (2010), The Young Scientists' Prize from The Commendation for Science and Technology by the Minister of Education, Culture, Sports, Science and Technology (2009), and Microsoft Research Japan New Faculty award (2011).



Yoichi Sato is a professor at Institute of Industrial Science, the University of Tokyo. He received his B.S. degree from the University of Tokyo in 1990, and his M.S. and Ph.D. degrees in robotics from School of Computer Science, Carnegie Mellon University in 1993 and 1997 respectively. His research interests include physics-based vision, reflectance analysis, image-based modeling and rendering, and gaze and gesture analysis. He served/is serving

in several conference organization and journal editorial roles including *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *International Journal of Computer Vision*, *Computer Vision and Image Understanding*, *IET Computer Vision*, *IPSJ Journal of Computer Vision and Applications*, *ECCV2012 Program Co-Chair* and *MVA2013 General Chair*.