SCALABLE LOW DIMENSIONAL MANIFOLD MODEL IN THE RECONSTRUCTION OF NOISY AND INCOMPLETE HYPERSPECTRAL IMAGES

Wei Zhu

*Zuoqiang Shi**

Stanley Osher†

Duke University
Mathematics Department

zhu@math.duke.edu

Tsinghua University
Department of Mathematical Sciences
Yau Mathematical Sciences Center

zqshi@tsinghua.edu.cn

UCLA
Department of Mathematics

sjo@math.ucla.edu

ABSTRACT

We present a scalable low dimensional manifold model for the reconstruction of noisy and incomplete hyperspectral images. The model is based on the observation that the spatial-spectral blocks of a hyperspectral image typically lie close to a collection of low dimensional manifolds. To emphasize this, the dimension of the manifold is directly used as a regularizer in a variational functional, which is solved efficiently by alternating direction of minimization and weighted nonlocal Laplacian. Unlike general 3D images, the same similarity matrix can be shared across all spectral bands for a hyperspectral image, therefore the resulting algorithm is much more scalable than that for general 3D data [1]. Numerical experiments on the reconstruction of hyperspectral images from sparse and noisy sampling demonstrate the superiority of our proposed algorithm in terms of both speed and accuracy.

Index Terms— Scalable low dimensional manifold model, hyperspectral image, noisy and incomplete image reconstruction.

1. INTRODUCTION

A hyperspectral image (HSI) is a collection of 2D images of the same spatial location taken at hundreds of different wavelengths [2]. The observed images are typically degraded when such data of high dimensionality are collected. For instance, the images can be very noisy due to limited exposure time, or some of the voxels can be missing due to the malfunctions of the hyperspectral cameras. An important task in HSI analysis is to recover the original image from its noisy incomplete observation. This is an ill-posed inverse problem, and some prior knowledge of the original data must be exploited.

One widely used prior information of HSI is that the 3D data cube has a low-rank structure under the linear mixing model (LMM) [3]. More specifically, the spectral signature of each pixel is assumed to be a linear combination of a few constituent endmembers. Under such an assumption, low-rank

matrix completion and sparse representation techniques have been used for HSI reconstruction [4, 5, 6]. Despite the simplicity of LMM, the linear mixing assumption has been shown to be physically inaccurate in certain situations [7].

Various partial differential equation (PDE) and graph based image processing techniques have also been applied to HSI reconstruction. The total variation (TV) method [8] has been widely used as a regularization in hyperspectral image processing [9, 10, 11, 12]. The nonlocal total variation (NLTV) [13], which computes the gradient in a nonlocal graph-based manner, has also been applied to the analysis of hyperspectral images [14, 15, 16]. However, such methods fail to produce satisfactory results when there is a significant number of missing voxels.

In [17, 18], the authors proposed a low dimensional manifold model (LDMM) for general image processing problems. LDMM is based on the observation that patches of a natural image typically sample a collection of low dimensional manifolds. Therefore the dimension of the patch manifold is directly used as a regularization term in a variational functional. The resulting Euler-Lagrange equation is solved either by the point integral method (PIM) [19], or the weighted nonlocal Laplacian [20]. LDMM achieved excellent results, especially in image inpainting problems from very sparse sampling. LDMM was also extended to 3D scientific data interpolation [1], but such generalization has poor scalability and requires huge memory storage.

In this paper, we exploit the special structure of hyperspectral images and propose a scalable LDMM specifically designed for the reconstruction of HSI from noisy and sparse sampling. The rationale behind the proposed method is that a hyperspectral image is a collection of 2D images of the same spatial location, and hence a single spatial similarity matrix can be shared across all spectral bands. The resulting algorithm is considerably faster than its 3D counterpart: it typically takes less than two minutes given a proper initialization as compared to hours in [1].

2. LDMM FOR HSI RECONSTRUCTION

2.1. Patch Manifold

We first describe the patch manifold of a hyperspectral image. Let $\boldsymbol{u} \in \mathbb{R}^{m \times n \times B}$ be a hyperspectral image, where $m \times n$ and B are the spatial and spectral dimensions of the image. For any $\boldsymbol{x} \in \bar{\Omega} = [m] \times [n]$, where $[m] = \{1, 2, \ldots, m\}$, we define a patch $\mathcal{P}_{\boldsymbol{x}}(\boldsymbol{u})$ as a 3D block of size $s_1 \times s_2 \times B$ of the original data cube $\boldsymbol{u}$, and the pixel $\boldsymbol{x}$ is the top-left corner of the rectangle of size $s_1 \times s_2$. The *patch set* $\mathcal{P}(\boldsymbol{u})$ is defined as the collection of all patches:

$$\mathcal{P}(\boldsymbol{u}) = \{\mathcal{P}_{\boldsymbol{x}}(\boldsymbol{u}) : \boldsymbol{x} \in \bar{\Omega}\} \subset \mathbb{R}^d, \quad d = s_1 \times s_2 \times B. \quad (1)$$

Previous work [1, 17] has shown that the point cloud $\mathcal{P}(\boldsymbol{u})$ is typically close to a collection of low dimensional smooth manifolds $\mathcal{M} = \cup_{l=1}^{L} \mathcal{M}_l$ embedded in $\mathbb{R}^d$. This collection of manifolds is called the *patch manifold* of $\boldsymbol{u}$.

2.2. Scalable LDMM

Our objective is to reconstruct the unknown HSI $\boldsymbol{u}$ from its noisy and incomplete observation $\boldsymbol{b} \in \mathbb{R}^{m \times n \times B}$. Assume that for each spectral band $t \in [B]$, $\boldsymbol{b}$ is only known on a random subset $\Omega^t \subset \bar{\Omega}$, with a sampling rate r (in our experiments $r = 5\%$ or 10%). According to [1, 17], we can use the dimension of the patch manifold as a regularizer to reconstruct $\boldsymbol{u}$ from $\boldsymbol{b}$:

$$\min_{\substack{\boldsymbol{u} \in \mathbb{R}^{m \times n \times B} \\ \mathcal{M} \subset \mathbb{R}^d}} \int_{\mathcal{M}} \dim(\mathcal{M}(\boldsymbol{p})) d\boldsymbol{p} + \lambda \sum_{t=1}^{B} \|\boldsymbol{u}^t - \boldsymbol{b}^t\|_{L^2(\Omega^t)^2}$$

$$\text{subject to:} \quad \mathcal{P}(\boldsymbol{u}) \subset \mathcal{M}, \quad (2)$$

where $\boldsymbol{u}^t$ is the t-th spectral band of the HSI $\boldsymbol{u}$, $\mathcal{M}(\boldsymbol{p})$ denotes the smooth manifold $\mathcal{M}_l$ to which $\boldsymbol{p}$ belongs, and $\int_{\mathcal{M}} \dim(\mathcal{M}(\boldsymbol{p})) d\boldsymbol{p} = \sum_{l=1}^{L} |\mathcal{M}_l| \dim(\mathcal{M}_l)$ is the L^1 norm of the local dimension. Based on Proposition 3.1 in [17], the first term in (2) can be written as the L^2 norm of the coordinate functions $\alpha_i^t : \mathcal{M} \to \mathbb{R}$. More specifically, (2) is equivalent to

$$\min_{\substack{\boldsymbol{u} \in \mathbb{R}^{m \times n \times B} \\ \mathcal{M} \subset \mathbb{R}^d}} \sum_{i=1}^{d_s} \sum_{t=1}^{B} \|\nabla_{\mathcal{M}} \alpha_i^t\|_{L^2(\mathcal{M})}^2 + \lambda \sum_{t=1}^{B} \|\boldsymbol{u}^t - \boldsymbol{b}^t\|_{L^2(\Omega^t)^2}$$

$$\text{subject to:} \quad \mathcal{P}(\boldsymbol{u}) \subset \mathcal{M}, \quad (3)$$

where $d_s = s_1 \times s_2$ is the spatial dimension, α_i^t is the coordinate function that maps every point $\boldsymbol{p} = (p_i^t)_{i,t} \in \mathcal{M}$ into its (i, t)-th coordinate p_i^t. Note that (2) is nonconvex, and we solve it by alternating the direction of minimization with respect to $\boldsymbol{u}$ and $\mathcal{M}$. More specifically, given $\mathcal{M}^{(k)}$ and $\boldsymbol{u}^{(k)}$ at step k satisfying $\mathcal{P}(\boldsymbol{u}^{(k)}) \subset \mathcal{M}^{(k)}$:

- With fixed $\mathcal{M}^{(k)}$, update the data $\boldsymbol{u}^{(k+1)}$ by solving:

$$\min_{\boldsymbol{u}} \sum_{i,t} \|\nabla_{\mathcal{M}^{(k)}} \alpha_i^t\|_{L^2(\mathcal{M}^{(k)})}^2 + \lambda \sum_{t=1}^{B} \|\boldsymbol{u}^t - \boldsymbol{b}^t\|_{L^2(\Omega^t)}^2$$

$$\text{subject to:} \quad \alpha_i^t(\mathcal{P}\boldsymbol{u}^{(k)}(\boldsymbol{x})) = \mathcal{P}_i^t \boldsymbol{u}(\boldsymbol{x}), \quad \boldsymbol{x} \in \bar{\Omega} \quad (4)$$

where $\mathcal{P}_i^t \boldsymbol{u}(\boldsymbol{x})$ is the (i, t)-th element in the patch $\mathcal{P}_{\boldsymbol{x}} \boldsymbol{u}$.

- Update the manifold $\mathcal{M}^{(k+1)}$ as the image under the perturbed coordinate function $\boldsymbol{\alpha}$:

$$\mathcal{M}^{(k+1)} = \boldsymbol{\alpha}(\mathcal{M}^{(k)}) \quad (5)$$

The manifold update (5) is easy to implement, and [18, 1] introduced a way to solve (4) using the weighted nonlocal Laplacian (WNLL) [20]. The idea is to discretize the Dirichlet energy $\|\nabla_{\mathcal{M}^{(k)}} \alpha_i^t\|_{L^2(\mathcal{M}^{(k)})}^2$ as

$$\frac{|\bar{\Omega}|}{|\Omega_i^t|} \sum_{\boldsymbol{x} \in \Omega_i^t} \sum_{\boldsymbol{y} \in \bar{\Omega}} \bar{w}(\boldsymbol{x}, \boldsymbol{y}) \left(\alpha_i^t(\mathcal{P}\boldsymbol{u}^{(k)}(\boldsymbol{x})) - \alpha_i^t(\mathcal{P}\boldsymbol{u}^{(k)}(\boldsymbol{y}))\right)^2$$

$$+ \sum_{\boldsymbol{x} \in \bar{\Omega} \setminus \Omega_i^t} \sum_{\boldsymbol{y} \in \bar{\Omega}} \bar{w}(\boldsymbol{x}, \boldsymbol{y}) \left(\alpha_i^t(\mathcal{P}\boldsymbol{u}^{(k)}(\boldsymbol{x})) - \alpha_i^t(\mathcal{P}\boldsymbol{u}^{(k)}(\boldsymbol{y}))\right)^2,$$

$$(6)$$

where $\Omega_i^t = \{\boldsymbol{x} \in \bar{\Omega} : \mathcal{P}_i^t \boldsymbol{u}^{(k)}(\boldsymbol{x}) \text{ is sampled}\}$ is a spatially translated version of Ω^t, $|\bar{\Omega}|/|\Omega_i^t| = 1/r$ is the inverse of the sampling rate, and $\bar{w}(\boldsymbol{x}, \boldsymbol{y}) = w(\mathcal{P}\boldsymbol{u}^{(k)}(\boldsymbol{x}), \mathcal{P}\boldsymbol{u}^{(k)}(\boldsymbol{y}))$ is the similarity between the patches, with

$$w(\boldsymbol{p}, \boldsymbol{q}) = \exp\left(-\frac{\|\boldsymbol{p} - \boldsymbol{q}\|^2}{\sigma(\boldsymbol{p})\sigma(\boldsymbol{q})}\right), \quad (7)$$

where $\sigma(\boldsymbol{p})$ is the normalizing factor. Combining the WNLL discretization (6) and the constraint in (4), the update of $\boldsymbol{u}$ in (4) can be discretized as

$$\min_{\boldsymbol{u}} \quad \lambda \sum_{t=1}^{B} \|\boldsymbol{u}^t - \boldsymbol{b}^t\|_{L^2(\Omega^t)}^2$$

$$+ \sum_{i,t} \left[\sum_{\boldsymbol{x} \in \bar{\Omega} \setminus \Omega_i^t} \sum_{\boldsymbol{y} \in \bar{\Omega}} \bar{w}(\boldsymbol{x}, \boldsymbol{y}) \left(\mathcal{P}_i^t \boldsymbol{u}(\boldsymbol{x}) - \mathcal{P}_i^t \boldsymbol{u}(\boldsymbol{y})\right)^2 \right.$$

$$\left. + \frac{1}{r} \sum_{\boldsymbol{x} \in \Omega_i^t} \sum_{\boldsymbol{y} \in \bar{\Omega}} \bar{w}(\boldsymbol{x}, \boldsymbol{y}) \left(\mathcal{P}_i^t \boldsymbol{u}(\boldsymbol{x}) - \mathcal{P}_i^t \boldsymbol{u}(\boldsymbol{y})\right)^2 \right]. \quad (8)$$

Remark 1. *Unlike the model in [1], the similarity matrix $\bar{w}$ in (8) is built on 2D coordinates $\boldsymbol{x}, \boldsymbol{y} \in \bar{\Omega}$, which significantly improves the scalability of the model.*

Note that (8) is decoupled with respect to the spectral coordinate t, and for any given $t \in [B]$, we only need to solve

the following problem:

$$\min_{\boldsymbol{u}^t} \quad \lambda \|\boldsymbol{u}^t - \boldsymbol{b}^t\|_{L^2(\Omega^t)}^2$$

$$+ \sum_{i=1}^{d_s} \left[\sum_{\boldsymbol{x} \in \bar{\Omega} \setminus \Omega_i^t} \sum_{\boldsymbol{y} \in \bar{\Omega}} \bar{w}(\boldsymbol{x}, \boldsymbol{y}) \left(\mathcal{P}_i \boldsymbol{u}^t(\boldsymbol{x}) - \mathcal{P}_i \boldsymbol{u}^t(\boldsymbol{y}) \right)^2 \right.$$

$$\left. + \frac{1}{r} \sum_{\boldsymbol{x} \in \Omega_i^t} \sum_{\boldsymbol{y} \in \bar{\Omega}} \bar{w}(\boldsymbol{x}, \boldsymbol{y}) \left(\mathcal{P}_i \boldsymbol{u}^t(\boldsymbol{x}) - \mathcal{P}_i \boldsymbol{u}^t(\boldsymbol{y}) \right)^2 \right]. \quad (9)$$

A standard variational technique shows that (9) is equivalent to the following Euler-Lagrange equation:

$$0 = \mu \sum_{i=1}^{d_s} \mathcal{P}_i^* I_{\Omega_i^t} \left[\sum_{\boldsymbol{y} \in \bar{\Omega}} \bar{w}(\boldsymbol{x}, \boldsymbol{y}) \left(\mathcal{P}_i \boldsymbol{u}^t(\boldsymbol{x}) - \mathcal{P}_i \boldsymbol{u}^t(\boldsymbol{y}) \right) \right]$$

$$+ \sum_{i=1}^{d_s} \mathcal{P}_i^* \left[\sum_{\boldsymbol{y} \in \bar{\Omega}} 2\bar{w}(\boldsymbol{x}, \boldsymbol{y}) \left(\mathcal{P}_i \boldsymbol{u}^t(\boldsymbol{x}) - \mathcal{P}_i \boldsymbol{u}^t(\boldsymbol{y}) \right) \right.$$

$$\left. + \mu \sum_{\boldsymbol{y} \in \Omega_i^t} \bar{w}(\boldsymbol{x}, \boldsymbol{y}) \left(\mathcal{P}_i \boldsymbol{u}^t(\boldsymbol{x}) - \mathcal{P}_i \boldsymbol{u}^t(\boldsymbol{y}) \right) \right]$$

$$+ \lambda I_{\Omega^t} \left(\boldsymbol{u}^t - \boldsymbol{b}^t \right), \quad \forall \boldsymbol{x} \in \bar{\Omega} \quad (10)$$

where $\mu = 1/r - 1$, $\mathcal{P}_i^*$ is the adjoint operator of $\mathcal{P}_i$, I_{Ω^t} is the projection operator that sets $\boldsymbol{u}^t(\boldsymbol{x})$ to zero for $\boldsymbol{x} \notin \Omega^t$. We use the notation $\boldsymbol{x}_{\widehat{j}}$ to denote the j-th component (in the spatial domain) after $\boldsymbol{x}$ in a patch. It is easy to verify that $\mathcal{P}_i \boldsymbol{u}^t(\boldsymbol{x}) = \boldsymbol{u}^t(\boldsymbol{x}_{\widehat{i-1}})$, and $\mathcal{P}_i^* \boldsymbol{u}^t(\boldsymbol{x}) = \boldsymbol{u}^t(\boldsymbol{x}_{\widehat{1-i}})$. Following the analysis similar to [1], we can rewrite (10) as

$$0 = \mu I_{\Omega^t} \left[\sum_{\boldsymbol{y} \in \bar{\Omega}} \sum_{i=1}^{d_s} \bar{w}(\boldsymbol{x}_{\widehat{1-i}}, \boldsymbol{y}_{\widehat{1-i}}) \left(\boldsymbol{u}^t(\boldsymbol{x}) - \boldsymbol{u}^t(\boldsymbol{y}) \right) \right]$$

$$+ \sum_{i=1}^{d_s} \left[\sum_{\boldsymbol{y} \in \bar{\Omega}} 2\bar{w}(\boldsymbol{x}_{\widehat{1-i}}, \boldsymbol{y}_{\widehat{1-i}}) \left(\boldsymbol{u}^t(\boldsymbol{x}) - \boldsymbol{u}^t(\boldsymbol{y}) \right) \right.$$

$$\left. + \mu \sum_{\boldsymbol{y} \in \Omega^t} \bar{w}(\boldsymbol{x}_{\widehat{1-i}}, \boldsymbol{y}_{\widehat{1-i}}) \left(\boldsymbol{u}^t(\boldsymbol{x}) - \boldsymbol{u}^t(\boldsymbol{y}) \right) \right]$$

$$+ \lambda I_{\Omega^t} \left(\boldsymbol{u}^t - \boldsymbol{b}^t \right). \quad \forall \boldsymbol{x} \in \bar{\Omega} \quad (11)$$

After setting $\tilde{w}(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{d_s} \bar{w}(\boldsymbol{x}_{\widehat{1-i}}, \boldsymbol{y}_{\widehat{1-i}})$, (11) is equivalent to

$$0 = 2 \sum_{y \in \bar{\Omega}} \tilde{w}(\boldsymbol{x}, \boldsymbol{y}) \left(\boldsymbol{u}^t(\boldsymbol{x}) - \boldsymbol{u}^t(\boldsymbol{y}) \right) + \lambda I_{\Omega^t} \left(\boldsymbol{u}^t - \boldsymbol{b}^t \right).$$

$$+ \mu I_{\Omega^t} \left[\sum_{\boldsymbol{y} \in \bar{\Omega}} \tilde{w}(\boldsymbol{x}, \boldsymbol{y}) \left(\boldsymbol{u}^t(\boldsymbol{x}) - \boldsymbol{u}^t(\boldsymbol{y}) \right) \right]$$

$$+ \mu \sum_{y \in \Omega^t} \tilde{w}(\boldsymbol{x}, \boldsymbol{y}) \left(\boldsymbol{u}^t(\boldsymbol{x}) - \boldsymbol{u}^t(\boldsymbol{y}) \right), \quad \forall \boldsymbol{x} \in \bar{\Omega} \quad (12)$$

Note that (12) is a linear system for $\boldsymbol{u}^t$ in $\mathbb{R}^{mn}$, but unlike [1], the coefficient matrix is not symmetric because of the projection operator I_{Ω^t}. In our numerical experiments, we always truncate the similarity matrix $\bar{w}(\boldsymbol{x}, \boldsymbol{y})$ to 20 nearest neighbors. Therefore, (12) is a sparse linear system and can be solved by the generalized minimal residual method (GMRES). The proposed algorithm for HSI reconstruction is summarized in Algorithm 1.

Algorithm 1 Scalable LDMM for HSI reconstruction

Input: A noisy and incomplete observation $\boldsymbol{b}$ of an unknown hyperspectral image $\boldsymbol{u} \in \mathbb{R}^{m \times n \times B}$. For every spectral band $t \in [B]$, $\boldsymbol{u}$ is only partially observed on a random subset Ω^t of $\bar{\Omega} = [m] \times [n]$.
Output: Reconstructed HSI $\boldsymbol{u}$.
   Initial guess $\boldsymbol{u}^{(0)}$.
   while not converge **do**
      1. Extract the patch set $\mathcal{P}\boldsymbol{u}^{(k)}$ from $\boldsymbol{u}^{(k)}$.
      2. Compute the similarity matrix on the spatial domain

$$\overline{w}(\boldsymbol{x}, \boldsymbol{y}) = w(\mathcal{P}\boldsymbol{u}^{(k)}(\boldsymbol{x}), \mathcal{P}\boldsymbol{u}^{(k)}(\boldsymbol{y})), \quad \boldsymbol{x}, \boldsymbol{y} \in \bar{\Omega}.$$

      3. Assemble the new similarity matrix

$$\tilde{w}(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{d_s} \bar{w}(\boldsymbol{x}_{\widehat{1-i}}, \boldsymbol{y}_{\widehat{1-i}})$$

      4. For every spectral band t, Update $(\boldsymbol{u}^t)^{(k+1)}$ as the solution of (12) using GMRES.
      5. $k \leftarrow k + 1$.
   end while
   $\boldsymbol{u} = \boldsymbol{u}^{(k)}$.

---

3. NUMERICAL EXPERIMENTS

3.1. Experimental Setup

In this section, we present the numerical results on the following datasets: Pavia University (PU), Pavia Center (PC), Indian Pine (IP), and San Diego Airport (SDA). All images have been cropped in the spatial dimension to 200×200 for easy comparison. The objective of the experiment is to reconstruct the original HSI from 5% random subsample (10% random subsample for noisy data).

Empirically, we found out that it is easier for LDMM to converge if a reasonable initialization is provided. In our experiments, we always use the result of the low-rank matrix completion algorithm APG [21] as an initialization, and run three iterations of manifold update for LDMM. The peak signal-to-noise ratio, PSNR $= 10 \log_{10} (\|\boldsymbol{u}^*\|_\infty / \text{MSE})$, is used to evaluate the reconstruction, where $\boldsymbol{u}^*$ is the ground truth, and MSE is the mean squared error. All experiments were run on a Linux machine with 8 Intel core i7-7820X

3.6 GHz CPUs and 64 GB of RAM. All codes and datasets are available for download at .

	APG		LDMM1		LDMM2	
	PSNR	time	PSNR	time	PSNR	time
IP	26.80	13 s	32.09	8 s	**34.08**	22 s
PC	32.61	17 s	**34.54**	11 s	34.25	31 s
PU	31.51	13 s	33.38	11 s	**33.66**	29 s
SDA	32.43	23 s	40.33	16 s	**44.21**	46 s

Table 1. Reconstruction of the HSIs from their noise-free 5% subsamples. LDMM1 (LDMM2) stands for LDMM with spatial patch size of 1×1 (2×2). The reported time of LDMM does not include that of the AGP initialization.

Original (Band 33)   5% subsample

APG (PSNR = 32.43)   Error

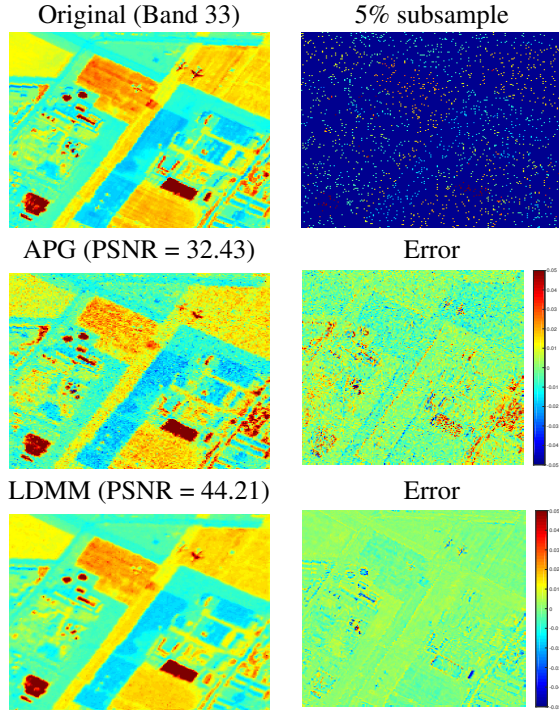LDMM (PSNR = 44.21)   Error

Fig. 1. Reconstruction of SDA from 5% noise-free subsample. Note that the error is displayed with a scale 1/20 of the original data to visually amplify the difference.

3.2. Reconstruction from noise-free subsample

We first present the results of the reconstruction of HSI from 5% noise-free random subsample. Table 1 displays the computational time and accuracy of the low-rank matrix completion (APG) initialization and LDMM with different spatial patch sizes (1×1 and 2×2). It can be observed that LDMM significantly improves the accuracy of APG with comparable extra computational time. Figure 1 provides a visual illustration of the results. Because of the limited space, we only present the reconstruction of SDA on one spectral band.

	APG		LDMM1		LDMM2	
	PSNR	time	PSNR	time	PSNR	time
IP	31.56	18 s	**34.03**	54 s	34.02	56 s
PC	30.22	47 s	30.55	82 s	**31.61**	82 s
PU	29.88	38 s	30.26	77 s	**31.40**	86 s
SDA	33.90	69 s	39.17	186 s	**41.31**	231 s

Table 2. Reconstruction of the noisy HSIs from their 10% subsamples. LDMM1 (LDMM2) stands for LDMM with spatial patch size of 1×1 (2×2). The reported time of LDMM does not include that of the AGP initialization.

Original (Band 38)   Noise added
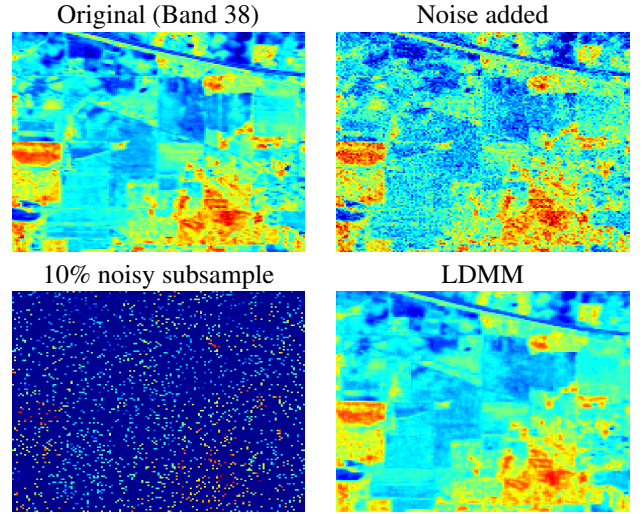
10% noisy subsample   LDMM

Fig. 2. Reconstruction of Indian Pine from 10% noisy subsample.

3.3. Reconstruction from noisy subsample

We then show the results of the reconstruction of HSI from 10% noisy subsample. A gaussian noise with a standard deviation of 0.05 is added to the original image, and then we remove 90% of the voxels from the data cube. The accuracy and computational time is reported in Table 2. Note that LDMM with 2×2 patches typically produce better results than that with 1×1 patches because of the presence of noise. A visual demonstration of the reconstruction is displayed in Figure 2.

4. CONCLUSION

We propose the scalable low dimensional manifold model for the reconstruction of hyperspectral images from noisy and incomplete observations with a significant number of missing voxels. The dimension of the patch manifold is directly used as a regularizer, and the same similarity matrix is shared across all spectral bands, which significantly reduces the computational burden. Numerical experiments show that the proposed algorithm is an accurate and efficient means for HSI reconstruction.

5. REFERENCES

[1] Wei Zhu, Bao Wang, Richard Barnard, Cory D. Hauck, Frank Jenko, and Stanley Osher, "Scientific data interpolation with low dimensional manifold model," *Journal of Computational Physics*, vol. 352, pp. 213 – 245, 2018.

[2] Chein-I Chang, *Hyperspectral imaging: techniques for spectral detection and classification*, vol. 1, Springer Science & Business Media, 2003.

[3] J.M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Qian Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, vol. 5, no. 2, pp. 354–379, 2012.

[4] A. S. Charles, B. A. Olshausen, and C. J. Rozell, "Learning sparse codes for hyperspectral imagery," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 5, pp. 963–978, 2011.

[5] R. Kawakami, Y. Matsushita, J. Wright, M. Ben-Ezra, Y. W. Tai, and K. Ikeuchi, "High-resolution hyperspectral imaging via matrix factorization," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, 2011, pp. 2329–2336.

[6] Zhengming Xing, Mingyuan Zhou, Alexey Castrodad, Guillermo Sapiro, and Lawrence Carin, "Dictionary learning for noisy and incomplete hyperspectral images," *SIAM Journal on Imaging Sciences*, vol. 5, no. 1, pp. 33–56, 2012.

[7] N. Dobigeon, J.-Y. Tourneret, C. Richard, J.C.M. Bermudez, S. McLaughlin, and A.O. Hero, "Nonlinear unmixing of hyperspectral images: Models and algorithms," *Signal Processing Magazine, IEEE*, vol. 31, no. 1, pp. 82–94, 2014.

[8] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, pp. 259–268, 1992.

[9] Q. Yuan, L. Zhang, and H. Shen, "Hyperspectral image denoising employing a spectral-spatial adaptive total variation model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 10, pp. 3660–3677, Oct 2012.

[10] M. D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Total variation spatial regularization for sparse hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 11, pp. 4484–4502, Nov 2012.

[11] W. He, H. Zhang, L. Zhang, and H. Shen, "Total-variation-regularized low-rank matrix factorization for hyperspectral image restoration," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 1, pp. 178–188, Jan 2016.

[12] H. K. Aggarwal and A. Majumdar, "Hyperspectral image denoising using spatio-spectral total variation," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 3, pp. 442–446, March 2016.

[13] Guy Gilboa and Stanley Osher, "Nonlocal operators with applications to image processing," *Multiscale Modeling & Simulation*, vol. 7, no. 3, pp. 1005–1028, 2009.

[14] Huiyi Hu, Justin Sunu, and Andrea L. Bertozzi, *Energy Minimization Methods in Computer Vision and Pattern Recognition: 10th International Conference, EMM-CVPR 2015, Hong Kong, China, January 13-16, 2015. Proceedings*, chapter Multi-class Graph Mumford-Shah Model for Plume Detection Using the MBO scheme, pp. 209–222, Springer International Publishing, Cham, 2015.

[15] W. Zhu, V. Chayes, A. Tiard, S. Sanchez, D. Dahlberg, A. L. Bertozzi, S. Osher, D. Zosso, and D. Kuang, "Unsupervised classification in hyperspectral imagery with nonlocal total variation and primal-dual hybrid gradient algorithm," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 5, pp. 2786–2798, May 2017.

[16] Jie Li, Qiangqiang Yuan, Huanfeng Shen, and Liangpei Zhang, "Hyperspectral image recovery employing a multidimensional nonlocal total variation model," *Signal Processing*, vol. 111, pp. 230 – 248, 2015.

[17] Stanley Osher, Zuoqiang Shi, and Wei Zhu, "Low dimensional manifold model for image processing," *SIAM Journal on Imaging Sciences*, vol. 10, no. 4, pp. 1669–1690, 2017.

[18] Zuoqiang Shi, Stanley Osher, and Wei Zhu, "Generalization of the weighted nonlocal laplacian in low dimensional manifold model," *Journal of Scientific Computing*, Sep 2017.

[19] Zuoqiang Shi and Jian Sun, "Convergence of the point integral method for poisson equation on point cloud," *Res. Math. Sci.*, vol. 4, no. 1, pp. 22.

[20] Zuoqiang Shi, Stanley Osher, and Wei Zhu, "Weighted nonlocal laplacian on interpolation from sparse data," *Journal of Scientific Computing*, vol. 73, no. 2, pp. 1164–1177, Dec 2017.

[21] Kim-Chuan Toh and Sangwoon Yun, "An accelerated proximal gradient algorithm for nuclear norm regularized linear least squares problems," *Pacific Journal of Optimization*, vol. 6, no. 615-640, pp. 15, 2010.