TENSOR NON-LOCAL LOW-RANK REGULARIZATION FOR RECOVERING COMPRESSED HYPERSPECTRAL IMAGES

Yongqiang Zhao, Jize Xue and Jinglei Hao*

School of Automation, Northwestern Polytechnical University, Xi'an, 710072, China

ABSTRACT

Sparsity-based methods have been widely used in hyperspectral imagery compression recovery (HSI-CR). However, most of the available HSI-CR methods work on vector space by vectorizing hyperspectral cubes in spatial and spectral domain, which will destroy spatial and spectral correlation and result in spatial and spectral information distortion in the recovery. At the same time, vectorization also make HSI's intrinsic structure sparsity cannot be utilized adequately. In this paper, a tensor non-local low-rank regularization (TNLR) approach is proposed to exploit essential structured sparsity and explore its advantages for CR of hyperspectral imagery. Specifically, a tensor nuclear norm penalty function is utilized as tensor low-rank regularization term to describe the spatial-and-spectral correlation hidden in HSI. To further improve the computational efficiency of the proposed algorithm, a fast implementation algorithm is developed by using the alternative direction multiplier method (ADMM) technique. Experimental results are shown that the proposed TNLR-CR algorithm can significantly outperform existing state-of-the-art CR techniques for hyperspectral image recovery.

Index Terms— Hyperspectral image, compression recovery, tensor low-rank approximation, non-local self-similarity, structured sparsity, alternative direction multiplier method.

1. INTRODUCTION

Hyperspectral imagery (HSI) is widely used in terrain classification, and military surveillance due to high spatial-spectral resolution [1, 2]. However, the HSI acquisition is confronted with the challenges of storage and transmission on limited resource platform at the cost of extremely large data size. Various image compression methods have been proposed to solve this problem. To guarantee the compressive data have no effects on HSI subsequent processing, an accurate HSI-CS algorithm is crucial.

Regularization methods are effective to deal with the reconstruction for compressed HSI [3,4]. One of their most important concerns are to impose a proper sparsity prior for HSI. Many regularizers (e.g., l_0 norm and l_1 norm etc.) usually characterize sparsity on vector independently without considering its structure, which is crucial to improve the reconstruction accuracy of standard sparse learning [5]. Actually, signal reconstruction relies on exploiting structure presenting in the original signal by promoting its sparsity [6]. Many successful compressed image reconstruction schemes rely on the low-rank regularization [7], which exploits structure sparsity of most natural images widely [8]. Besides, low-rank regularization methods applied to HSI-CR have also met with success [3], [4], [9]. The concept of sparsity has evolved into various sophisticated forms including model-based or Bayesian, non-local sparsity [10, 11], and structured/group sparsity [12, 13], where exploiting higher-order dependency among sparse coefficients is shown to be beneficial to compressed image recovery.

In this paper, an HSI with $d_W \times d_H$ spatial resolution and d_S spectral bands can be expressed as a 3^{rd}-order $\mathcal{X} \in \mathbb{R}^{d_W \times d_H \times d_S}$ with two spatial modes and one spectral mode and hidden spatial-and-spectral structures can be discovered by using tensor modelling for the enhancement of its compressed recovery. Specifically, two intrinsic characteristics underlying HSI are fully considered, i.e., the global correlation along spectrum (GCS) and nonlocal self-similarity across space (NSS). To exploit the nonlocal sparsity of hyperspectral images, the HSI-CR recovery is regularized by patch grouping and tensor low-rank approximation. Consequentially, for each exemplar HSI patch we group a set of similar HSI 3D fullband patches (FBP) to form 3^{rd}-order tensor $\mathcal{Y}_p$. Since each patch group contains similar structures, whose rank of recovering tensor $\mathcal{X}_p$ is low and can be implied to be a useful image prior. To efficiently solve the problem of rank minimization, the convex tensor nuclear norm is utilized to characterize the tensor non-local low-rank property. As such, the HSI-CR task resorts to solving an optimization problem, which can be efficiently solved by ADMM.

2. HSI-CR VIA TENSOR NON-LOCAL LOW-RANK REGULARIZATION

For HSI, the compressive measurement operator can be denoted by a matrix Φ with M rows and N columns, where

This work is supported by the National Natural Science Foundation of China (61371152, 61374162), the National Natural Science Foundation of China and South Korean National Research Foundation Joint Funded Cooperation Program (61511140292), the Fundamental Research Funds for the Central Universities (3102015ZY045)

$M \leq N$ and $N = d_W \times d_H \times d_S$. Different compressive measurement operators specify different structures of the matrix Φ. The compressive measurements y can be obtained from the model mathematically:

$$y{=}\Phi x, \tag{1}$$

where y is a vector of length M, x is reconstructive HSI. In this work, the measurement matrix Φ the so called restricted isometry property (RIP) [14]. To recover x from the measurement y, prior knowledge of x is needed. Standard CR method recovers the unknown signal by pursuing the sparsest signal x that satisfies $y{=}\Phi x$, namely

$$x = \arg\min_x \|x\|_0, \ \mathbf{s.t.}\ y{=}\Phi x \tag{2}$$

where $\|\cdot\|_0$ is a pseudo norm counting the number of nonzero entries of x . In theory a K-sparse signal can be perfectly recovered from as low as $M = 2K$ measurements [14]. However, since $\|\cdot\|_0$ norm minimization is a difficult combinatorial optimization problem, solving Eq.(2) directly is both NP-hard and unstable with the presence of noise. For this reason, it has been proposed to replace the nonconvex l_0 norm by its convex l_1 counterpart - i.e.,

$$x = \arg\min_x \|x\|_1, \ \mathbf{s.t.}\ y{=}\Phi x \tag{3}$$

The above l_0-minimization problem can be efficiently solved by various methods, such as iterative shrinkage algorithm [15], Bregman Split algorithm [16] and alternative direction multiplier method (ADMM) [17].

To perfectly promote HSI-CR, more prior information of HSI have been introduced to regular the undetermined problem. Structured sparsity is particularly important to the modeling of HSI that often exhibits rich self-repetitive structures in spatial and spectral domain. Essentially, exploiting the so-called nonlocal self-similarity has led to the well-known nonlocal low-rank methods [7] and block-matching 3D denoising [18]. Both [7] and [18] show state-of-the-art image restoration results. However, their effectiveness in HSI-CR applications has not been documented in the open literature to the best of our knowledge. In this paper, we will present a unified variational framework for CR exploiting the nonlocal structured sparsity via tensor low-rank approximation.

2.1. Tensor non-local low-rank regularization

The proposed regularization model consists of two components: patch grouping for characterizing non-local self-similarity of HSI and tensor low-rank approximation for sparsity enforcement.

Under the assumption that these HSI patches have similar structures. By sweeping across the HSI with overlaps, lots of 3D FBP groups $\{\mathcal{P}_{i,j}\}_{1\leq i\leq d_W-d_w+1,1\leq i\leq d_H-d_h+1} \in \mathrm{R}^{d_w\times d_h\times d_S}(d_w < d_W, d_h < d_H)$ are constructed for the HSI. For each exemplar patch $\mathcal{P}_{i,j}$ (located at spatial position (i,j)), we search for similar patches within a local window (e.g., 70×70). Specifically, we group the similar patches by k-nearest neighbor (k-NN) method into a set of 3^{rd}-order tensors $\{\mathcal{Y}_p\}_{p=1}^P$, where $P = (d_W{-}d_w{+}1)(d_H{-}d_h{+}1)$ denotes the patch number. Each FBP group constructed in this way contains local spatial while global spectral dimensionality, which can easily help us consider the two important properties of a HSI: the nonlocal similarity between spatial patches and the global correlation across all bands. In substance, the formed 3D FBP a low-rank property. In practice, $\{\mathcal{Y}_p\}_{p=1}^P$ may be corrupted by some noise, which could lead to the deviation from the desirable low-rank constraint. One possible solution is to model the data tensor $\mathcal{Y}_p$ as: $\mathcal{Y}_p = \mathcal{X}_p + \mathcal{W}_p$, where $\mathcal{X}_p$ and $\mathcal{W}_p$ denote the low-rank tensor and the Gaussian noise tensor with the variance σ_w^2, respectively. Then the low-rank tensor can be recovered by solving the following optimization problem:

$$\mathcal{X}_p = \arg\min_{\mathcal{X}_p} rank(\mathcal{X}_p), \ \ \mathbf{s.t.}\ \|\mathcal{Y}_p - \mathcal{X}_p\|_F^2 \leq \sigma_w^2 \tag{4}$$

where $\|\mathcal{X}\|_F$ denotes the tensor $\mathcal{X}$ Frobenious norm and $rank(\mathcal{X}_p)$ is the rank function of tensor $\mathcal{X}_p$.

As discussed in the introduction, an HSI possesses the spatial-spectra correlation property. The nonlocal self-similarity indicates that one 3D patch often possesses many similar structure 3D patches across the whole HSI volume. Each 3^{rd}-order tensor $\mathcal{X}_p$ can be approximated as Tucker low-rank tensor decompositions:

$$\mathcal{X}_p \approx \mathcal{G}_p{\times_1}\mathrm{U}_{1p}{\times_2}\mathrm{U}_{2p}{\times_3}\mathrm{U}_{3p} \tag{5}$$

where $\mathcal{G}_p$ is the core tensor and matrices $\mathrm{U}_{1p}, \mathrm{U}_{2p}$, and U_{3p} respectively correspond to factor matrices of the height, width, and spectral mode. Compared to matrix modeling technique, the advantage of tensor modeling technique is that it can not only characterize the spectral correlation but also the spatial correlation in the hyperspectral image.

In this paper, we consider tensor low-rank constraint form based tensor nuclear norm. Following the work [19], low-rank property of a 3^{rd}-order tensor can be measured by following formulas:

$$rank(\mathcal{X}_p) = \|\mathcal{X}_p\|_* = \sum\nolimits_i^3 \alpha_i \left\|\mathcal{X}_{p_{(i)}}\right\|_* \tag{6}$$

where $\alpha_i > 0$ and satisfies $\sum_i^3 \alpha_i = 1$, $\|\mathcal{X}_p\|_*$ denotes the nuclear norm of tensor $\mathcal{X}_p$, and $\mathcal{X}_{p_{(i)}}$ is the i^{th} unfolded matrix of tensor $\mathcal{X}_p$.

2.2. Proposed HSI-CR model

Based on the previous discussions, a regularization model is proposed for solving the HSI-CR problem:

$$\min_{\mathcal{X}_p} \sum\nolimits_{p=1}^P \|\mathcal{Y}_p - \mathcal{X}_p\|_F^2 + \lambda\|\mathcal{X}_p\|_*, \ \mathbf{s.t.}\, y = \Phi x \tag{7}$$

where the scalar λ is regularization parameter.

2.3. ADMM algorithm

Based on the model of HSI-CR, the ADMM strategy [17] can be applied for solving large scale optimization problems.

Firstly, the auxiliary tensors $\{\mathcal{M}_p\}_{p=1}^P$ are introduced and then we can equivalently reformulate (7) as follows:

$$\begin{aligned}\min_{\mathcal{X},\mathcal{M}_p,\mathcal{G}_p,\mathrm{U}_{ip}} &\sum_{p=1}^P \frac{1}{2}\left\|\mathcal{Y}_p-\mathcal{G}_p\times_1\mathrm{U}_{1p}\times_2\mathrm{U}_{2p}\times_3\mathrm{U}_{3p}\right\|_F^2\\ &+\lambda\|\mathcal{M}_p\|_*,\ {\mathrm{U}_{ip}}^T\mathrm{U}_{ip}=I,(i=1,2,3),\\ \text{s.t. }&\mathcal{G}_p\times_1\mathrm{U}_{1p}\times_2\mathrm{U}_{2p}\times_3\mathrm{U}_{3p}-\mathcal{M}_p=0,y=\Phi x\end{aligned} \tag{8}$$

Then its augmented Lagrangian function is in this:

$$\begin{aligned}&L(\mathcal{X},\mathcal{M}_p,\mathcal{G}_p,\mathrm{U}_{1p},\mathrm{U}_{2p},\mathrm{U}_{3p})=\\ &\sum_{p=1}^P\frac{1}{2}\left\|\mathcal{Y}_p-\mathcal{G}_p\times_1\mathrm{U}_{1p}\times_2\mathrm{U}_{2p}\times_3\mathrm{U}_{3p}\right\|_F^2+\lambda\|\mathcal{M}_p\|_*\\ &+\ \langle\mathcal{G}_p\times_1\mathrm{U}_{1p}\times_2\mathrm{U}_{2p}\times_3\mathrm{U}_{3p}-\mathcal{M}_p,\mathcal{P}_p\rangle+\langle\Lambda,y-\Phi x\rangle\\ &+\frac{\mu}{2}\left\|\mathcal{G}_p\times_1\mathrm{U}_{1p}\times_2\mathrm{U}_{2p}\times_3\mathrm{U}_{3p}-\mathcal{M}_p\right\|_F^2+\frac{\beta}{2}\left\|y-\Phi x\right\|_F^2\end{aligned} \tag{9}$$

where $\mathcal{P}_p s$ and Λ are the Lagrange multipliers, μ and β are positive scalars. The Eq.(9) can be decomposed into three sub-problems and iteratively update each variable through fixing the other ones.

(a) $\langle\mathcal{G}_p,\mathrm{U}_{1p},\mathrm{U}_{2p},\mathrm{U}_{3p}\rangle$ sub-problem:

$$\begin{aligned}\min_{\mathcal{G}_p,\mathrm{U}_{ip}}&\sum_{p=1}^P\frac{1}{2}\left\|\mathcal{Y}_p-\mathcal{G}_p\times_1\mathrm{U}_{1p}\times_2\mathrm{U}_{2p}\times_3\mathrm{U}_{3p}\right\|_F^2\\ &+\langle\mathcal{G}_p\times_1\mathrm{U}_{1p}\times_2\mathrm{U}_{2p}\times_3\mathrm{U}_{3p}-\mathcal{M}_p,\mathcal{P}_p\rangle\\ &+\frac{\mu}{2}\left\|\mathcal{G}_p\times_1\mathrm{U}_{1p}\times_2\mathrm{U}_{2p}\times_3\mathrm{U}_{3p}-\mathcal{M}_p\right\|_F^2\end{aligned} \tag{10}$$

which is equivalent to the following sub-problem:

$$\min_{\mathcal{G}_p,\mathrm{U}_{ip}}\sum_{p=1}^P\frac{1}{2}\left\|\mathcal{G}\times_1\mathrm{U}_{1p}\times_2\mathrm{U}_{2p}\times_3\mathrm{U}_{3p}-\mathcal{O}\right\|_F^2 \tag{11}$$

where $\mathcal{O}=\frac{\mathcal{Y}_p-(\mu\mathcal{M}_p-\mathcal{P}_p)}{1+3\mu}$ and ${\mathrm{U}_{ip}}^T\mathrm{U}_{ip}=I$. This sub-problem can be easily solved using the proposed method [20].

(b) $\mathcal{M}_p$ sub-problem:

$$\begin{aligned}\min_{\mathcal{M}_p}&\sum_{p=1}^P\lambda\|\mathcal{M}_p\|_*\\ &+\langle\mathcal{G}_p\times_1\mathrm{U}_{1p}\times_2\mathrm{U}_{2p}\times_3\mathrm{U}_{3p}-\mathcal{M}_p,\mathcal{P}_p\rangle\\ &+\frac{\mu}{2}\left\|\mathcal{G}_p\times_1\mathrm{U}_{1p}\times_2\mathrm{U}_{2p}\times_3\mathrm{U}_{3p}-\mathcal{M}_p\right\|_F^2\end{aligned} \tag{12}$$

Briefly, the sub-problem can be reformulated as

$$\min_{\mathcal{M}_p}\sum_{p=1}^P\frac{\lambda}{\mu}\|\mathcal{M}_p\|_*+\frac{1}{2}\left\|\mathcal{B}_p+\mathcal{P}_p-\mathcal{M}_p\right\|_F^2 \tag{13}$$

where $\mathcal{B}_p=\mathcal{G}_p\times_1\mathrm{U}_{1p}\times_2\mathrm{U}_{2p}\times_3\mathrm{U}_{3p}$. By combinatorially considering (6) and (13), this sub-problem can be rewritten as:

$$\min_{\mathcal{M}_p}\sum_{p=1}\sum_{i=1}^3\frac{\lambda}{\mu}\left\|\mathcal{M}_{p(i)}\right\|_*+\frac{1}{2}\left\|\mathcal{Q}_{p(i)}-\mathcal{M}_{p(i)}\right\|_F^2 \tag{14}$$

where $\mathcal{Q}_{p(i)}=\mathcal{B}_{p(i)}+\mathcal{P}_{p(i)}$. According to [19], its close-form solution is expressed as

$$\mathcal{M}_p=fold_p[S_{\lambda/\mu}(\mathcal{Q}_{p(i)})] \tag{15}$$

where $fold[\cdot]$ denotes folding a tensor into a matrix. For a given matrix X, the singular value shrinkage operator $S_\tau(X)$ is defined as $S_\tau(X):=U_X D_\tau(\Sigma_X)V_X^T$, and where $X=U_X\sigma_X V_X^T$ is the singular value decomposition of X and $D_\tau(A)=\mathrm{sgn}(A_{ij})(|A_{ij}|-\tau)_+$.

(c) $\mathcal{X}$ sub-problem:

$$\begin{aligned}\min_{\mathcal{X}}&\sum_{p=1}^P\frac{1}{2}\left\|\mathcal{Y}_p-\mathcal{G}_p\times_1\mathrm{U}_{1p}\times_2\mathrm{U}_{2p}\times_3\mathrm{U}_{3p}\right\|_F^2\\ &+\langle\Lambda,y-\Phi x\rangle+\frac{\beta}{2}\left\|y-\Phi x\right\|_F^2\end{aligned} \tag{16}$$

It is easy to observe that optimizing L with respect to x can be treated as solving the following linear system:

$$\begin{aligned}x+\beta\Phi^*(\Phi x)=&\Phi^*(\beta y-\Lambda)\\ &+vce(\mathcal{Y}-\mathcal{G}\times_1\mathrm{U}_1\times_2\mathrm{U}_2\times_3\mathrm{U}_3)\end{aligned} \tag{17}$$

where $vec(\cdot)$ indicates the vectorization of the matrix or tensor, and Φ^* indicates the adjoint of Φ . Obviously, this linear system can be solved by off-the-shelf gradient techniques.

(d) According to the ADMM, the multipliers associated with L are updated via the following formulas:

$$\begin{aligned}&\mathcal{P}_p\leftarrow\mathcal{P}_p-\rho\mu(\mathcal{Y}_p-\mathcal{B}_p)(p=1,\ldots,P)\\ &\Lambda\leftarrow\Lambda-\rho\beta(y-\Phi x)\end{aligned} \tag{18}$$

where ρ is a parameter associated with convergence rate with the value, e.g., 1.1, and the penalty parameters μ and follow an adaptive updating scheme. Take β as an example, β is initialized by a small value $\frac{1e-7}{mean(abs(y))}$ and then updated by the scheme ($c=1.05$):

$$\beta\leftarrow c.\beta \tag{19}$$

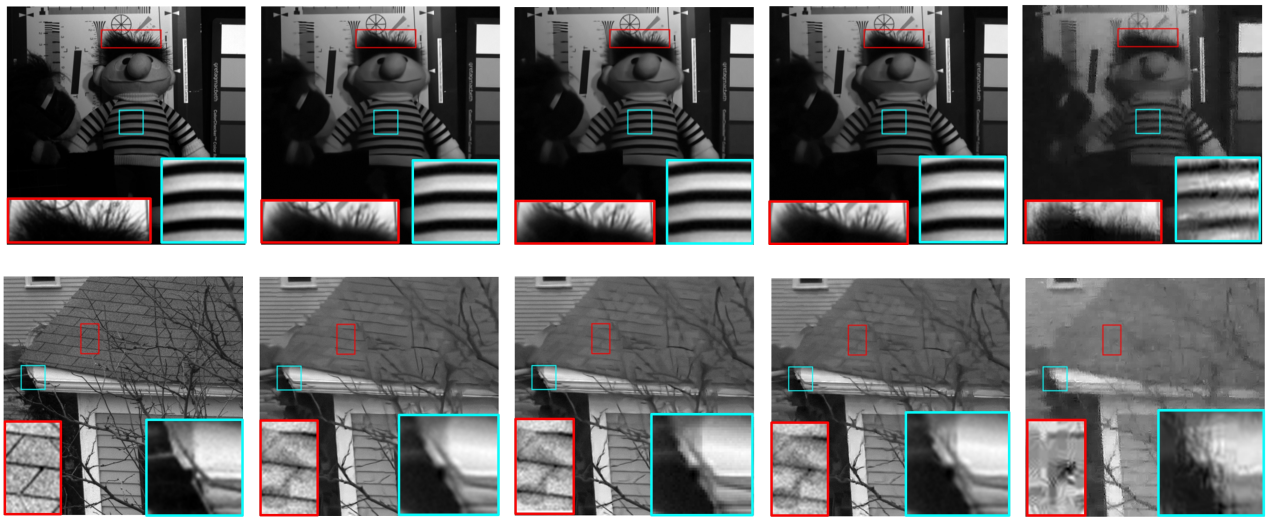
For convenience, the proposed HSI-CR algorithm can be summarized as Algorithm 1, and we denote the proposed method as TNLR.

3. EXPERIMENT

Two different public datasets of hyperspectral images, i.e., the CAVE datasets[1] and the Harvard datasets[2] are used to assess the performance of our proposed method. And, we select crop 500×500 pixels in each band as the test data, then normalized the clipped images, keeping their values in the interval [0-255]. For visual comparison, we compare the proposed TNLR method with three state-of-the-art methods, including Orthogonal Matching Pursuit (OMP) [21], LASSO [22] and K-SVD [23]. Parts of the reconstructed spectral images under sampling rate $\rho=0.05$ are shown in Fig. 1. It can be observed that the recovery-based OMP and LASSO result in some details lose, the K-SVDs reconstruction results will produce artifacts. Proposed TNLR method provides better reconstruction visual effects as shown in Fig. 1 (b), and outperforms the other ones. Further, PSNR results are employed to evaluate the quality of the proposed HSI-CR strategy shown in Fig. 2. Bisides, the mean spectral angle mapper (MSAM) is calculated the average angle between the spectral vectors of the original and reconstructed images as shown in Fig. 3. The HSI-CR method-based TNLR can obtain higher PSNR and

[1] http://www1.cs.columbia.edu/CAVE/databases/multispectral/.

[2] http://vision.seas.harvard.edu/hyperspec/.



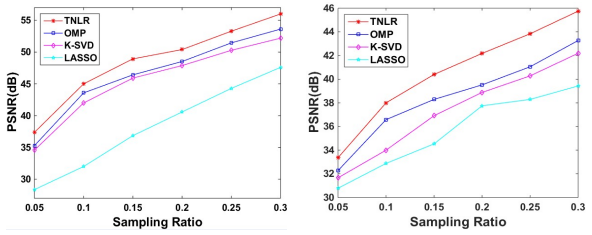
(a) Original Image (b) TNLR (c) OMP (d) K-SVD (e) LASSO

Fig. 1. Visual comparison with sampling rate $\rho = 0.05$. (a)Original image; (b)TNLR; (c)OMP; (d)K-SVD (e) LASSO. The first row shows the reconstructed images for 640nm band in the CAVE; the last row shows the reconstructed images for 540nm band in the Harvard.

Algorithm1: HSI-CR based TNLR

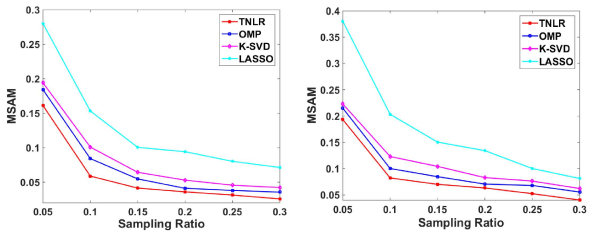
1:Initialization: Initializing an HSI $x^{(0)}$ using a standard CR method(e.g. DCT based CR recovery).
2:For $l = 1 : L$ **do**
3: Construct tensor groups $\{\mathcal{Y}_p\}_{p=1}^{P}$ from $x^{(0)}$:grouping a set of similar 3-D patches into a 3^{rd}-order tensor search the each exemplar patch via k-NN;
4: **For** $p = 1 : P$ **do**
5: Solve the problem (9) by ADMM;
6: Updating $\mathcal{G}_p, \mathrm{U}_{ip}, \mathrm{U}_{2p}, \mathrm{U}_{3p}$ by Eq.(11);
7: Updating $\mathcal{M}_p$ via Eq.(15);
8: **End for**
9: Updating $x^{(l)}$ by Eq.(17);
10: Updating the multipliers by Eq.(18);
11: Updating the parameters by Eq.(19);
12:End for
Out: Compression recovery HSI $x^{(L)}$.

reduce more spectral distortion compared to the others. Because the others only utilize all the spatial information during compression recovery, while TNLR considers both the spatial and spectral structure information.



(a) The CAVE datasets (b) The Harvard datasets

Fig. 2. PSNR value comparison when sampling ratios vary from 0.05 to 0.3 for different datasets.



(a) The CAVE datasets (b) The Harvard datasets

Fig. 3. MSAM value comparison when sampling ratios vary from 0.05 to 0.3 for different datasets.

4. CONCLUSION

In this paper, a novel method for hyperspectral image compressed recovery is proposed by using tensor non-local low-rank regularization. The proposed method considers intrinsic structured sparsity, which the nonlocal similarity between spatial patches and the global correlation across all bands are considered fully. Each patch group contains similar structures, its low-rank property can be regarded as a valuable prior imposed on a tensor. Experimental results reveal that the proposed HSI-CR method outperforms the traditional image compression methods.

5. REFERENCES

[1] J. Han, D. Zhang, G. Cheng, L. Guo, and J. Ren, "Object detection in optical remote sensing images based on weakly supervised learning and high-level feature learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 6, pp. 3325–3337, 2015.

[2] Y. Q. Zhao, L. Zhang, and S. G. Kong, "Band-subset-based clustering and fusion for hyperspectral imagery classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 2, pp. 747–756, 2011.

[3] M. Golbabaee and P. Vandergheynst, "Hyperspectral image compressed sensing via low-rank and joint-sparse matrix recovery," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, 2012, pp. 2741–2744.

[4] B. Huang, K. Xu, J. Wan, and X. Liu, "Distributed compressive sensing of hyperspectral images using low rank and structure similarity property," *Sensing and Imaging*, vol. 16, no. 1, pp. 1–8, 2015.

[5] J. Huang, T. Zhang, and D. Metaxas, "Learning with structured sparsity," *Journal of Machine Learning Research*, vol. 12, no. 7, pp. 417–424, 2009.

[6] W. Dong, X. Wu, and G. Shi, "Sparsity fine tuning in wavelet domain with application to compressive image reconstruction," *IEEE Transactions on Image Processing*, vol. 23, no. 12, pp. 5249–5262, 2014.

[7] W. Dong, G. Shi, X. Li, Y. Ma, and F. Huang, "Compressive sensing via nonlocal low-rank regularization." *IEEE Transactions on Image Processing*, vol. 23, no. 8, pp. 3618–3632, 2014.

[8] W. Dong, X. Li, L. Zhang, and G. Shi, "Sparsity-based image denoising via dictionary learning and structural clustering," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2011, pp. 457–464.

[9] Y. Q. Zhao and J. Yang, "Hyperspectral image denoising via sparse representation and low-rank constraint," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 1, pp. 296–308, 2014.

[10] W. Dong, G. Shi, X. Li, L. Zhang, and X. Wu, "Image reconstruction with locally adaptive sparsity and nonlocal robust regularization," *Signal Processing Image Communication*, vol. 27, no. 10, pp. 1109–1122, 2012.

[11] W. Dong, G. Shi, and X. Li, "Nonlocal image restoration with bilateral variance estimation: a low-rank approach." *IEEE Transactions on Image Processing*, vol. 22, no. 2, pp. 700–711, 2013.

[12] Y. Li, W. Dai, and H. Xiong, "Subspace learning with structured sparsity for compressive video sampling," in *Proceedings of IEEE Data Compression Conference*, 2015, pp. 456–456.

[13] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," in *Proceedings of IEEE International Conference on Computer Vision*, 2010, pp. 2272–2279.

[14] D. L. Donoho, "Compressed sensing," *IEEE Transactions on information theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[15] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413–1457, 2004.

[16] X. Zhang, M. Burger, X. Bresson, and S. Osher, "Bregmanized nolocal regularization for deconvolution and sparse reconstruction." *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 253–276, 2010.

[17] W. Deng and W. Yin, "On the global and linear convergence of the generalized alternating direction method of multipliers," *Journal of Scientific Computing*, vol. 66, no. 3, pp. 889–916, 2016.

[18] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, 2007.

[19] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, pp. 208–220, 2013.

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

[21] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[22] B. Efron, T. Hastie, I. Johnstone, R. Tibshirani *et al.*, "Least angle regression," *The Annals of Statistics*, vol. 32, no. 2, pp. 407–499, 2004.

[23] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.