A Tensor Subspace Representation-Based Method for Hyperspectral Image Denoising

Jie Lin[1]

Ting-Zhu Huang[1], Xi-Le Zhao[1], Tai-Xiang Jiang[2], Lina Zhuang[3]

[1]University of Electronic Science and Technology of China

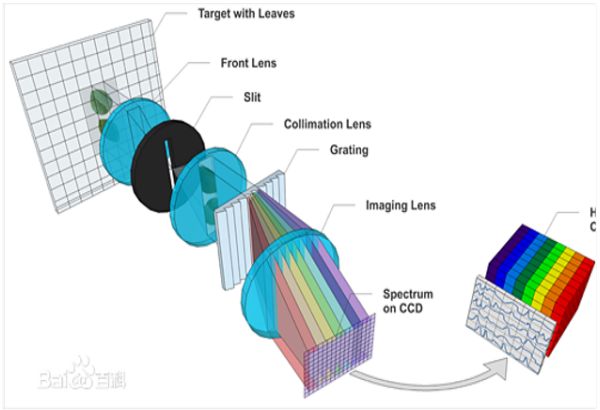
[2]Southwestern University of Finance and Economics

[3]The University of Hong Kong

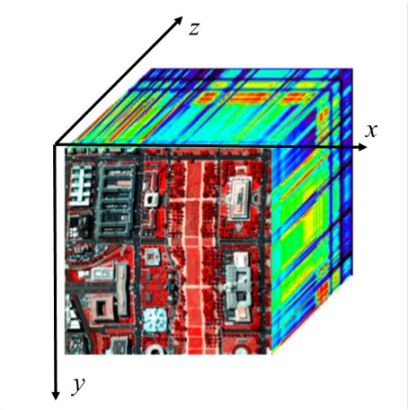
Outline

Outline

Hyperspectral image (HSI) tensor data

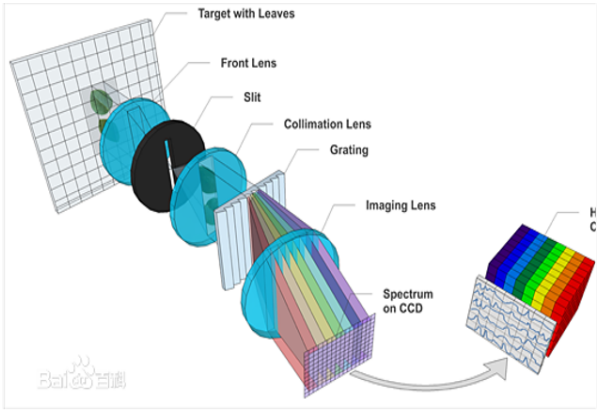


Imaging process

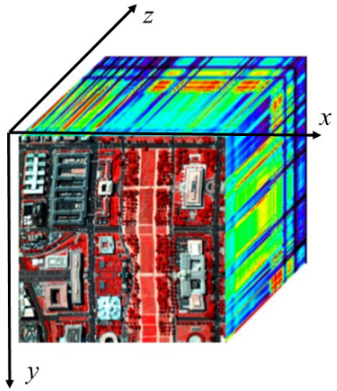


HSI tensor

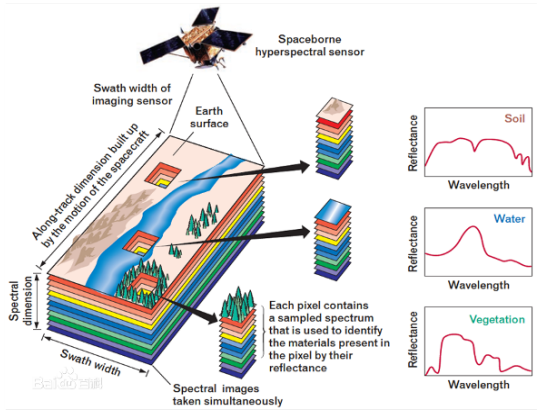
Hyperspectral image (HSI) tensor data



Imaging process



HSI tensor

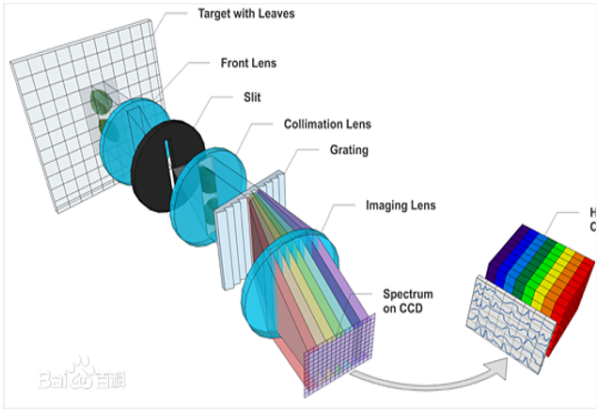


Application

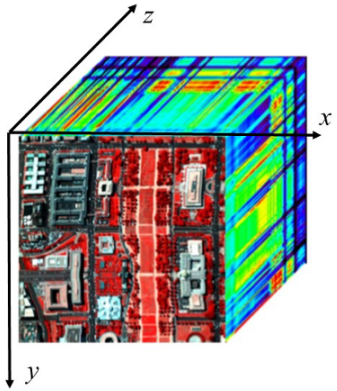
Mineral exploration

Urban planning

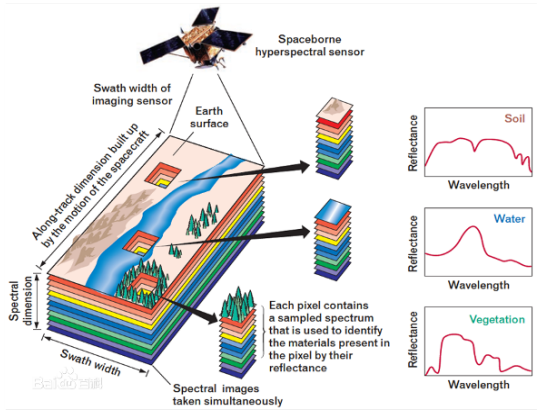
Hyperspectral image (HSI) tensor data



Imaging process



HSI tensor



Application

Mineral exploration

Urban planning

- Noise: especially Gaussian nosie

Tensor-Singular Value Decomposition (T-SVD)

Tensor-tensor product (t-prod) [Kilmer et al.. LAA, 2011]

The tensor-tensor-product (t-prod) $\mathcal{C} = \mathcal{A} * \mathcal{B}$ of $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and $\mathcal{B} \in \mathbb{R}^{n_2 \times n_4 \times n_3}$ is a tensor of size $n_1 \times n_4 \times n_3$, where the (i, j)-th tube $\mathbf{c}_{ij:}$ is given by

$$\mathbf{c}_{ij:} = \mathcal{C}(i, j, :) = \sum_{k=1}^{n_2} \mathcal{A}(i, k, :) * \mathcal{B}(k, j, :)$$

where $*$ denotes the circular convolution between two tubes of same size.

Tensor-Singular Value Decomposition (T-SVD)

t-SVD [Kilmer et al.. LAA, 2011]

For $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, the t-SVD of $\mathcal{A}$ is given by

$$\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^{\mathrm{H}} \tag{1}$$

where $\mathcal{U} \in \mathbb{R}^{n_1 \times n_1 \times n_3}$ and $\mathcal{V} \in \mathbb{R}^{n_2 \times n_2 \times n_3}$ are orthogonal tensors, and $\mathcal{S} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is an f-diagonal tensor.

The t-SVD is illustrated in Figure 1.

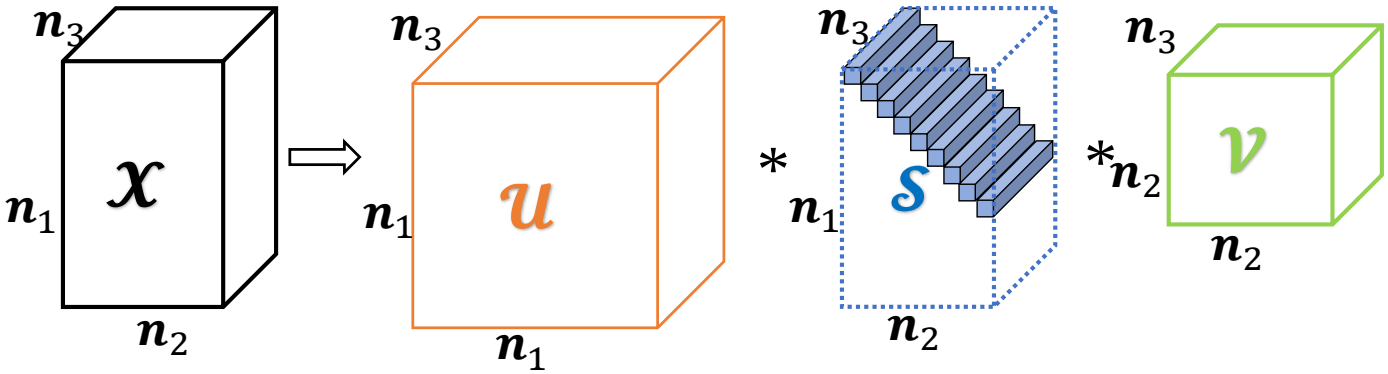
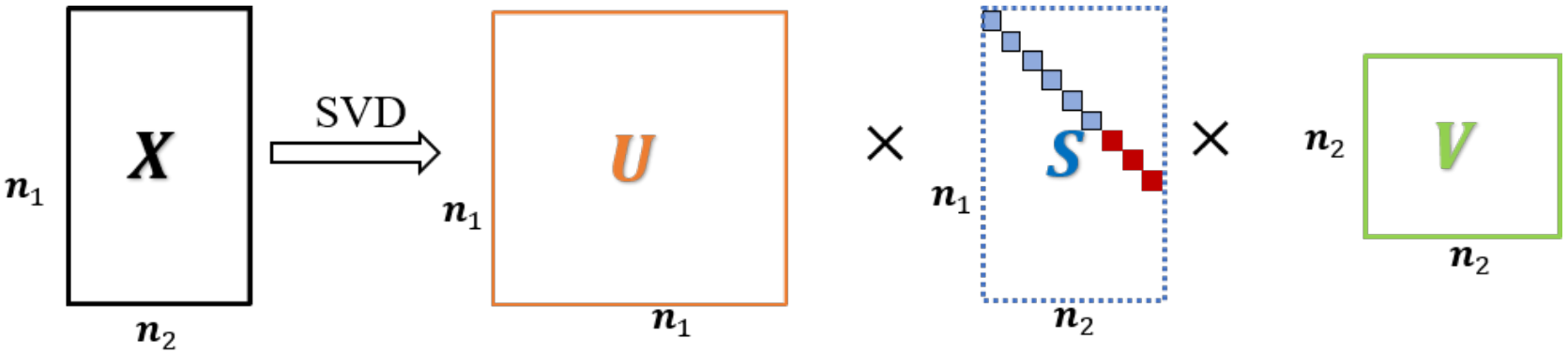


Figure 1: The t-SVD of an $n_1 \times n_2 \times n_3$ tensor.

Outline

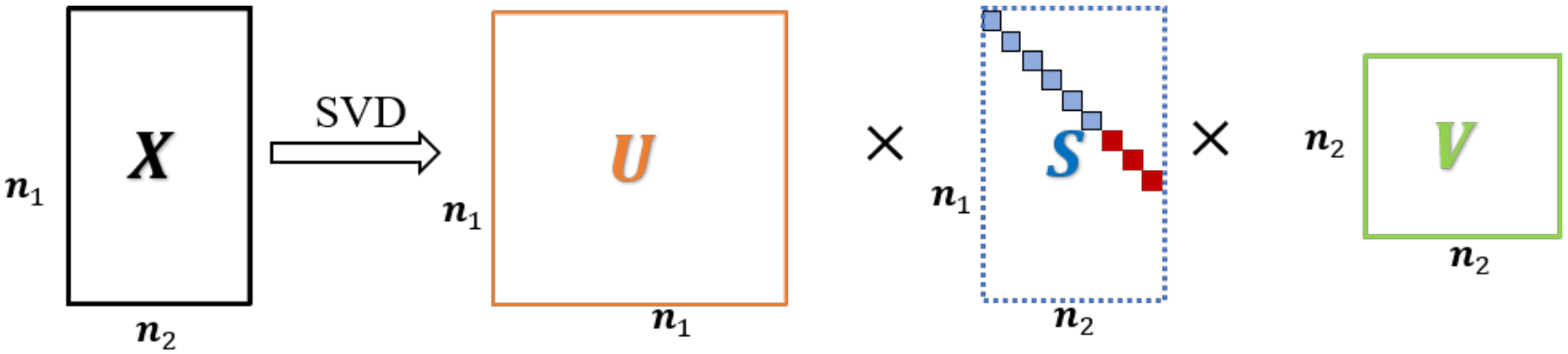
Matrix Subspace

- In linear algebra, the rank of a matrix **X** is the dimension of the vector space generated (or spanned) by its columns. [Bourbaki, Algebra, ch. II, Sec. 10.12, p. 359]



Matrix Subspace

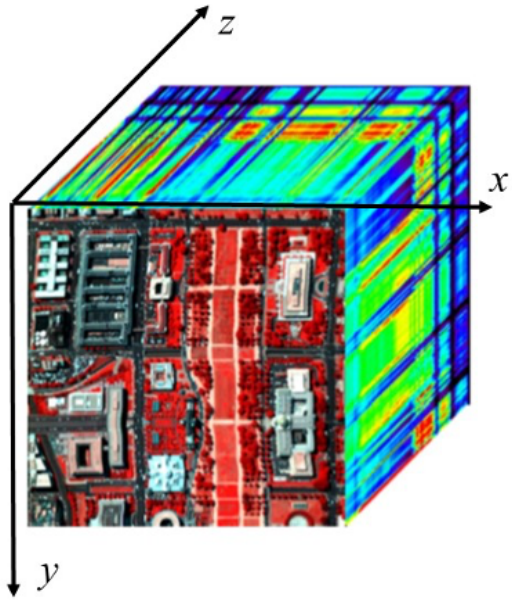
- In linear algebra, the rank of a matrix **X** is the dimension of the vector space generated (or spanned) by its columns. [Bourbaki, Algebra, ch. II, Sec. 10.12, p. 359]



- Subspace perspective: linear representation
- Related work:

  FastHyDe [Zhuang et al. IEEE J-STARS 2018], NGmeet [He et al. CVPR 2019]

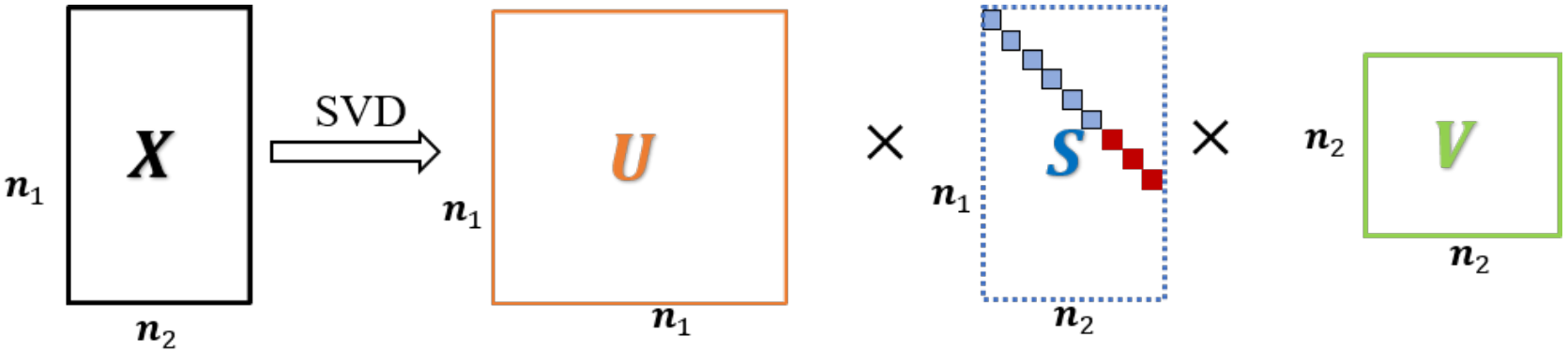
Tensor Subspace



HSI tensor

Subspace for tensors?

Inspiration



Tensor-linear (T-linear) Representation and Tensor Subspace

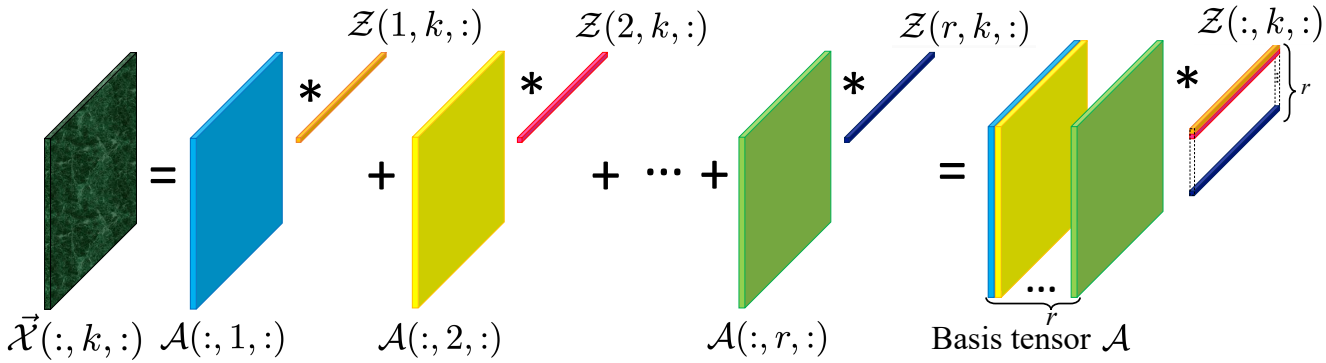


Figure 2: The kth lateral slice of $\vec{\mathcal{X}}$, i.e., the kth band of HSI, is represented by a t-linear combination of the r bases and coefficients.

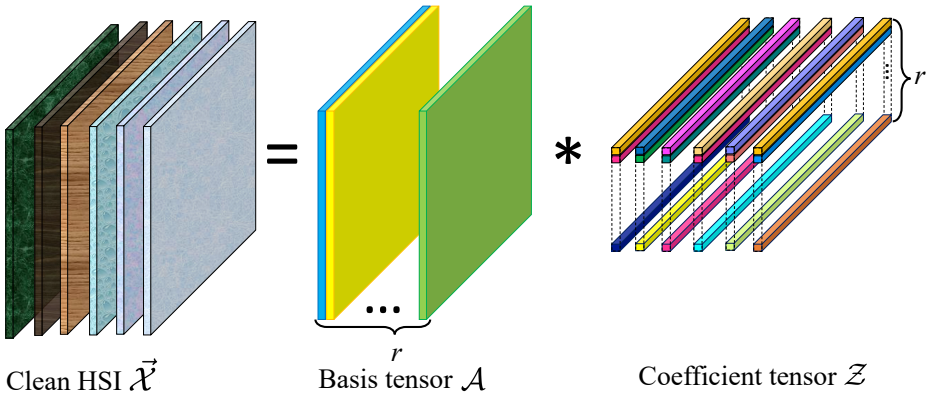


Figure 3: The clean HSI is represented by the basis tensor and coefficient tensor.

Advantage of Tensor Subspace

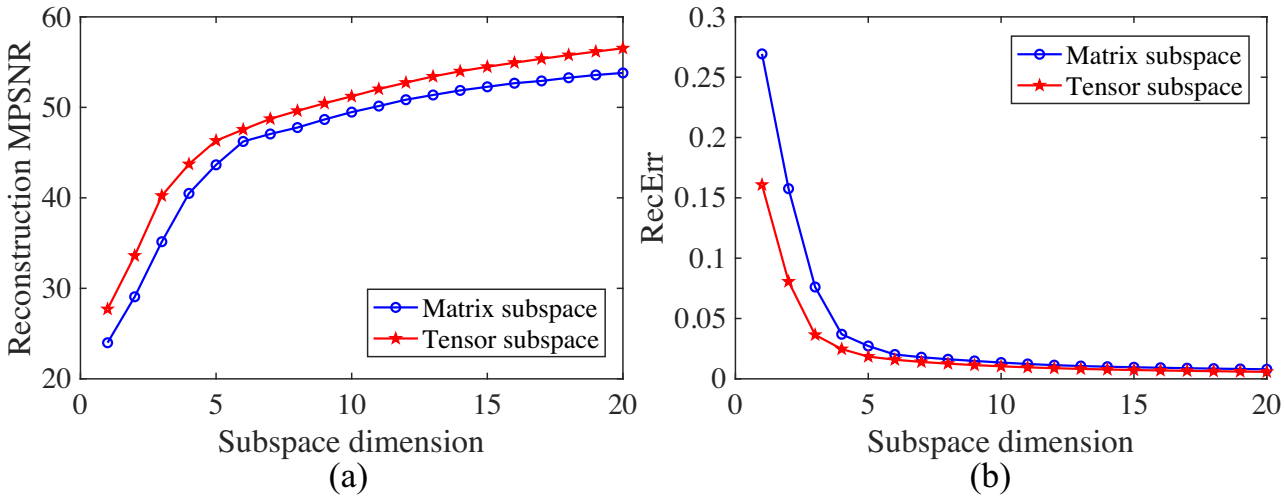


Figure 4: Presentation capability comparison of the matrix and tensor subspace. (a) Reconstruction MPSNR with respect to the subspace dimension; (b) Reconstruction error with respect to the subspace dimension.

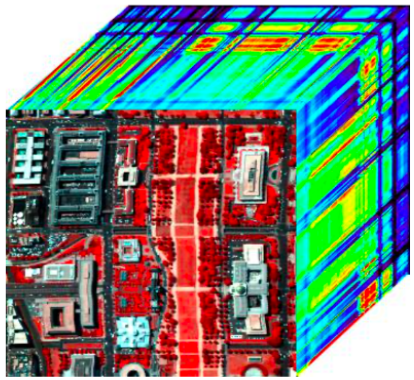
Outline

Denoising Model

Degradation model:

$$\mathcal{Y} = \mathcal{X} + \mathcal{N},$$

Prior knowledge:



Intrinsic property

Global correlation in spectrum and space (GCSS)

Nonlocal self-similarity (NSS)

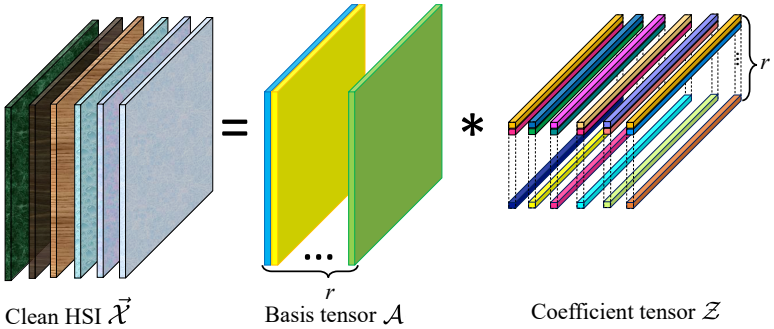
Denoising Model

Regularization Design:

GCSS

↓

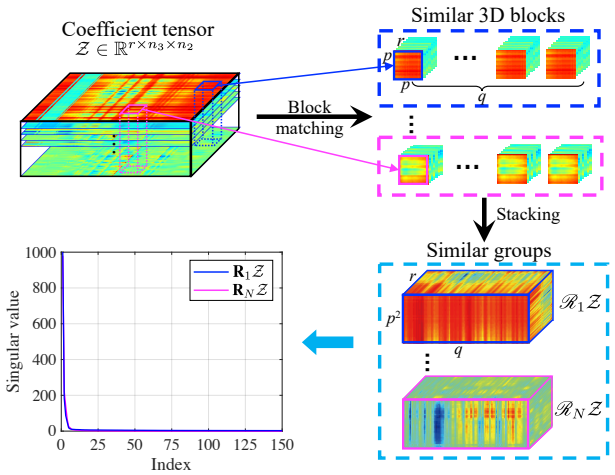
Tensor subspace



NSS

↓

Weighted nuclear norm



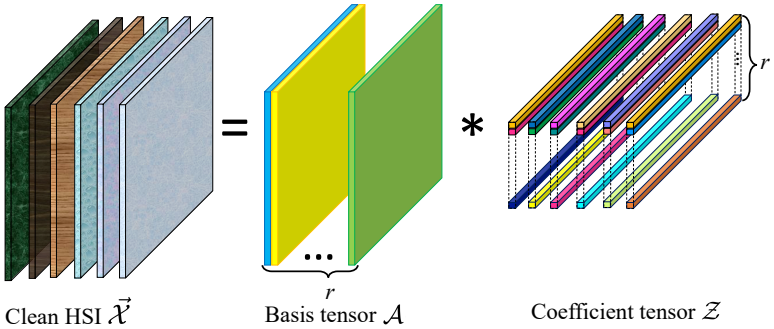
Denoising Model

Regularization Design:

GCSS

↓

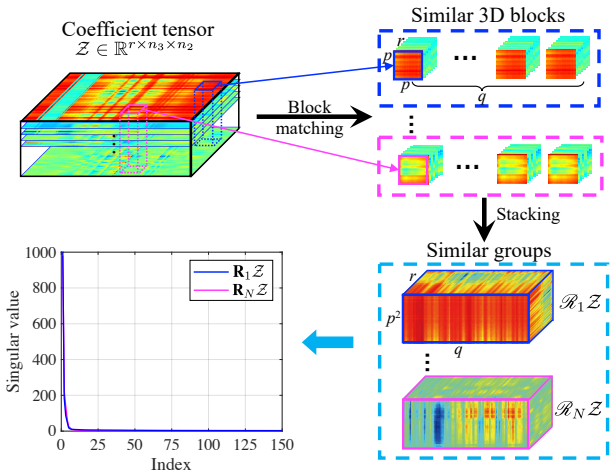
Tensor subspace



NSS

↓

Weighted nuclear norm



Main model:

$$\min_{\mathcal{A},\mathcal{Z}} \frac{1}{2}\|\vec{\mathcal{Y}} - \mathcal{A} * \mathcal{Z}\|_F^2 + \lambda \sum_i \|\mathbf{R}_i\mathcal{Z}\|_{\mathbf{w},*}, \quad (2)$$

$$s.t.\ \mathcal{A}^T * \mathcal{A} = \mathcal{I},$$

Solving Algorithm

Main model:

$$\begin{aligned} &\min_{\mathcal{A},\mathcal{Z}} \frac{1}{2}\|\vec{\mathcal{Y}} - \mathcal{A} * \mathcal{Z}\|_F^2 + \lambda \sum\nolimits_i \|\mathbf{R}_i \mathcal{Z}\|_{\mathbf{w},*}, \\ &s.t.\ \mathcal{A}^T * \mathcal{A} = \mathcal{I}, \end{aligned} \tag{3}$$

Proximal alternating minimization (PAM) algorithm:

$$\begin{cases} \mathcal{Z}^{k+1} = \arg\min_{\mathcal{Z}} \frac{1}{2}\|\vec{\mathcal{Y}} - \mathcal{A}^k * \mathcal{Z}\|_F^2 + \lambda \sum\nolimits_i \|\mathbf{R}_i \mathcal{Z}\|_{\mathbf{w},*} + \frac{\rho}{2}\|\mathcal{Z} - \mathcal{Z}^k\|_F^2 \\ \mathcal{A}^{k+1} = \arg\min_{\mathcal{A}^T * \mathcal{A} = \mathcal{I}} \frac{1}{2}\|\vec{\mathcal{Y}} - \mathcal{A} * \mathcal{Z}^{k+1}\|_F^2 + \frac{\rho}{2}\|\mathcal{A} - \mathcal{A}^k\|_F^2, \end{cases} \tag{4}$$

$\mathcal{Z}$ subproblem

$$\begin{aligned}\mathcal{Z}^{k+1} &= \arg\min_{\mathcal{Z}} \frac{1}{2}\|(\mathcal{A}^k)^T * \vec{\mathcal{Y}} - \mathcal{Z}\|_F^2 + \lambda \sum\nolimits_i \|\mathbf{R}_i \mathcal{Z}\|_{\mathbf{w},*} + \frac{\rho}{2}\|\mathcal{Z} - \mathcal{Z}^k\|_F^2 \\ &= \arg\min_{\mathcal{Z}} \frac{1}{2}\left\|\frac{(\mathcal{A}^k)^T * \vec{\mathcal{Y}} + \rho\mathcal{Z}^k}{1+\rho} - \mathcal{Z}\right\|_F^2 + \frac{\lambda}{1+\rho}\sum\nolimits_i \|\mathbf{R}_i \mathcal{Z}\|_{\mathbf{w},*}.\end{aligned} \tag{5}$$

$\mathcal{Z}$ subproblem

$$\begin{aligned}\mathcal{Z}^{k+1} &= \arg\min_{\mathcal{Z}} \frac{1}{2}\|(\mathcal{A}^k)^T * \vec{\mathcal{Y}} - \mathcal{Z}\|_F^2 + \lambda \sum\nolimits_i \|\mathbf{R}_i\mathcal{Z}\|_{\mathbf{w},*} + \frac{\rho}{2}\|\mathcal{Z} - \mathcal{Z}^k\|_F^2 \\ &= \arg\min_{\mathcal{Z}} \frac{1}{2}\left\|\frac{(\mathcal{A}^k)^T * \vec{\mathcal{Y}} + \rho\mathcal{Z}^k}{1+\rho} - \mathcal{Z}\right\|_F^2 + \frac{\lambda}{1+\rho}\sum\nolimits_i \|\mathbf{R}_i\mathcal{Z}\|_{\mathbf{w},*}.\end{aligned} \tag{5}$$

We apply the alternating direction method of multipliers (ADMM) to solve problem (5), which can be guaranteed to converge globally [Wang et al. J. Sci. Comput, 2019]. Introducing auxiliary variables $\mathcal{M}_i = \mathcal{Z}(i = 1, 2, \cdots, N)$ to (5), we have

$$\min_{\mathcal{M}_i,\mathcal{Z}} \frac{1}{2}\|\tilde{\mathcal{Z}} - \mathcal{Z}\|_F^2 + \frac{\lambda}{1+\rho}\sum\nolimits_i \|\mathbf{M}_i^{(2)}\|_{\mathbf{w},*}, \ s.t. \ \mathbf{M}_i^{(2)} = \mathbf{R}_i\mathcal{Z}. \tag{6}$$

$\mathcal{Z}$ subproblem

The augmented Lagrangian function of (6) is defined as

$$\begin{aligned} L(\mathcal{M}_i,\mathcal{Z},\mathcal{T}_i) &= \frac{1}{2}\|\tilde{\mathcal{Z}}-\mathcal{Z}\|_F^2 \\ &+ \frac{\lambda}{1+\rho}\sum\nolimits_i \|\mathbf{M}_i^{(2)}\|_{\mathbf{w},*} + \frac{\gamma}{2}\sum\nolimits_i \left\|\mathcal{M}_i - R_i\mathcal{Z} + \frac{\mathcal{T}_i}{\gamma}\right\|_F^2, \end{aligned} \tag{7}$$

Updating:

$$\begin{cases} (\mathbf{M}_i^{(2)})^{p+1} = \min\limits_{\mathbf{M}_i^{(2)}} \dfrac{\gamma(1+\rho)}{2\lambda}\left\|\mathbf{M}_i^{(2)} - \left(\mathbf{R}_i\mathcal{Z}^p - \dfrac{(\mathbf{T}_i^{(2)})^p}{\gamma}\right)\right\|_F^2 + \|\mathbf{M}_i^{(2)}\|_{\mathbf{w},*}, \\ \mathcal{Z}^{p+1} = \min\limits_{\mathcal{Z}} \dfrac{1}{2}\|\tilde{\mathcal{Z}}-\mathcal{Z}\|_F^2 + \dfrac{\gamma}{2}\sum\nolimits_i \left\|R_i\mathcal{Z} - \left(\mathcal{M}_i^{p+1} + \dfrac{\mathcal{T}_i^p}{\gamma}\right)\right\|_F^2, \\ \mathcal{T}_i^{p+1} = \mathcal{T}_i^p + \gamma(\mathcal{M}_i^{p+1} - R_i\mathcal{Z}^{p+1}). \end{cases} \tag{8}$$

Algorithm 1

Algorithm 1 ADMM algorithm for solving (3)

Input: The penalty parameter γ.

1: Initialize: $\mathcal{Z}^0 = \tilde{\mathcal{Z}}$, $\mathcal{T}_i^0 = 0$.
2: **for** $p = 0 : P$ **do**
3: Update $\mathcal{M}_i^{p+1}$ by (10), $i = 1, 2, \cdots, N$;
4: Update $\mathcal{Z}^{p+1}$ by (12);
5: Update $\mathcal{T}_i^{p+1}$ by (8), $i = 1, 2, \cdots, N$.
6: **end for**

Output: The solution $\mathcal{Z}^{p+1}$.

$\mathcal{A}$ subproblem

$$\mathcal{A}^{k+1} = \arg\min_{\mathcal{A}^T * \mathcal{A} = \mathcal{I}} \frac{1}{2}\|\vec{\mathcal{Y}} - \mathcal{A} * \mathcal{Z}^{k+1}\|_F^2 + \frac{\rho}{2}\|\mathcal{A} - \mathcal{A}^k\|_F^2, \tag{9}$$

$\mathcal{A}$ subproblem

$$\mathcal{A}^{k+1} = \arg\min_{\mathcal{A}^T * \mathcal{A} = \mathcal{I}} \frac{1}{2}\|\vec{\mathcal{Y}} - \mathcal{A} * \mathcal{Z}^{k+1}\|_F^2 + \frac{\rho}{2}\|\mathcal{A} - \mathcal{A}^k\|_F^2, \tag{9}$$

To solve (9), we give the following theorem.

Theorem

For any $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, the following problem:

$$\min_{\mathcal{A}} \frac{1}{2}\|\mathcal{Y} - \mathcal{A} * \mathcal{Z}\|_F^2 + \frac{\rho}{2}\|\mathcal{A} - \mathcal{B}\|_F^2, \ s.t. \mathcal{A}^T * \mathcal{A} = \mathcal{I}, \tag{10}$$

has the closed-form solution $\mathcal{A}^* = \mathcal{V} * \mathcal{U}^T$, where $\mathcal{U}$ and $\mathcal{V}$ are from the t-SVD of $\mathcal{Z} * \mathcal{Y}^T + \rho\mathcal{B}^T = \mathcal{U} * \mathcal{S} * \mathcal{V}^T$.

Algorithm 2

Algorithm 2 PAM Algorithm for TenSRDe

Input: Noisy HSI $\mathcal{Y}$, regularization parameter λ, proximal parameter ρ, iterative regularization parameter θ, and increment δ.

Initialize: $\mathcal{A}^0$, $\mathcal{Z}^0 = (\mathcal{A}^0)^T * \vec{\mathcal{Y}}$, $\mathcal{X}^0 = \mathcal{O}$, and $\epsilon = 10^{-3}$.
for $t = 0 : T-1$ **do**
    Compute $r^{t+1} = r^t + \delta$;
    Compute $\mathcal{Y}^{t+1} = \theta\mathcal{X}^t + (1-\theta)\mathcal{Y}$.
    while not converged **do**
        Update $\mathcal{Z}^{k+1}$ by (3);
        Update $\mathcal{A}^{k+1}$ by (18);
        Check the convergence condition:

$$\|\mathcal{A}^{k+1} * \mathcal{Z}^{k+1} - \mathcal{A}^k * \mathcal{Z}^k\|_F^2 / \|\mathcal{A}^k * \mathcal{Z}^k\|_F^2 \leq \epsilon.$$

    end while
    Let $\mathcal{X}^{t+1} = \text{inv-Permute}(\mathcal{A}^{k+1} * \mathcal{Z}^{k+1})$.
end for

Output: Estimated HSI $\mathcal{X}$.

Outline

Compared Methods and Datasets

Compared methods:

- BM3D [Dabov et al. IEEE TIP 2007]
- BM4D [Maggioni et al. IEEE TIP 2013]
- KBR [Xie et al. IEEE TPAMI 2018]
- TDL [Peng et al. CVPR 2014]
- LRTFL_0 [Xiong et al. IEEE TGRS 2019]
- FastHyDe [Zhuang et al. IEEE J-STARS 2018]
- SNLRSF [Cao et al. IEEE J-STARS 2019]
- NGmeet [He et al. CVPR 2019]

Compared Methods and Datasets

Compared methods:

- BM3D [Dabov et al. IEEE TIP 2007]
- BM4D [Maggioni et al. IEEE TIP 2013]
- KBR [Xie et al. IEEE TPAMI 2018]
- TDL [Peng et al. CVPR 2014]
- LRTFL_0 [Xiong et al. IEEE TGRS 2019]
- FastHyDe [Zhuang et al. IEEE J-STARS 2018]
- SNLRSF [Cao et al. IEEE J-STARS 2019]
- NGmeet [He et al. CVPR 2019]

Datasets:

- Pavia Centre (PaC)
- Washington DC (WDC)
- Indian Pines
- Urban

Metrics: PSNR, SSIM, and ERGAS

Simulated Experiments

Table 1: Quantitative Evaluation and Running Time (in Minutes) of Different Methods on Three Datasets with Simulated Noise.

Dataset	σ	Index	Noisy	BM3D	BM4D	KBR	TDL	LRTF$L0$	FastHyDe	SNLRSF	NGmeet	TenSRDe
PaC	0.02	MPSNR	33.983	36.744	43.862	45.994	46.367	45.010	47.078	47.761	48.251	**49.247**
		MSSIM	0.9326	0.9670	0.9930	0.9954	0.9959	0.9948	0.9966	0.9971	0.9973	**0.9978**
		ERGAS	74.069	53.384	23.703	18.909	18.211	21.220	16.753	15.589	14.677	**13.469**
		Time	—	0.5245	1.6773	7.6816	0.1454	15.127	**0.0873**	6.3234	0.5243	0.3777
	0.10	MPSNR	19.998	28.100	33.655	35.500	35.642	35.702	37.046	37.483	38.210	**38.645**
		MSSIM	0.4362	0.7984	0.9334	0.9585	0.9586	0.9580	0.9701	0.9733	0.9775	**0.9783**
		ERGAS	370.51	144.13	76.337	62.159	61.375	61.528	52.530	49.573	45.543	**44.932**
		Time	—	0.6334	1.6863	6.5056	0.1443	5.8963	**0.0877**	6.1883	0.5176	0.2777
WDC	0.02	MPSNR	33.977	36.315	43.604	47.556	47.592	46.742	49.687	49.041	49.787	**51.623**
		MSSIM	0.9366	0.9667	0.9934	0.9971	0.9973	0.9964	0.9981	0.9980	0.9982	**0.9988**
		ERGAS	75.448	56.575	24.684	16.245	16.733	17.797	12.891	14.826	12.645	**10.368**
		Time	—	2.1015	7.6493	24.075	0.5627	127.33	**0.1565**	10.543	0.9347	1.5100
	0.10	MPSNR	20.001	27.311	32.935	35.949	35.913	36.764	38.702	39.203	39.113	**40.263**
		MSSIM	0.4741	0.7788	0.9279	0.9679	0.9671	0.9703	0.9803	0.9835	0.9829	**0.9859**
		ERGAS	377.06	158.62	83.534	59.399	61.032	54.871	44.322	41.511	42.031	**37.456**
		Time	—	2.4635	7.5322	23.983	0.5167	36.033	**0.1617**	10.682	0.9404	1.2337

Simulated Experiments

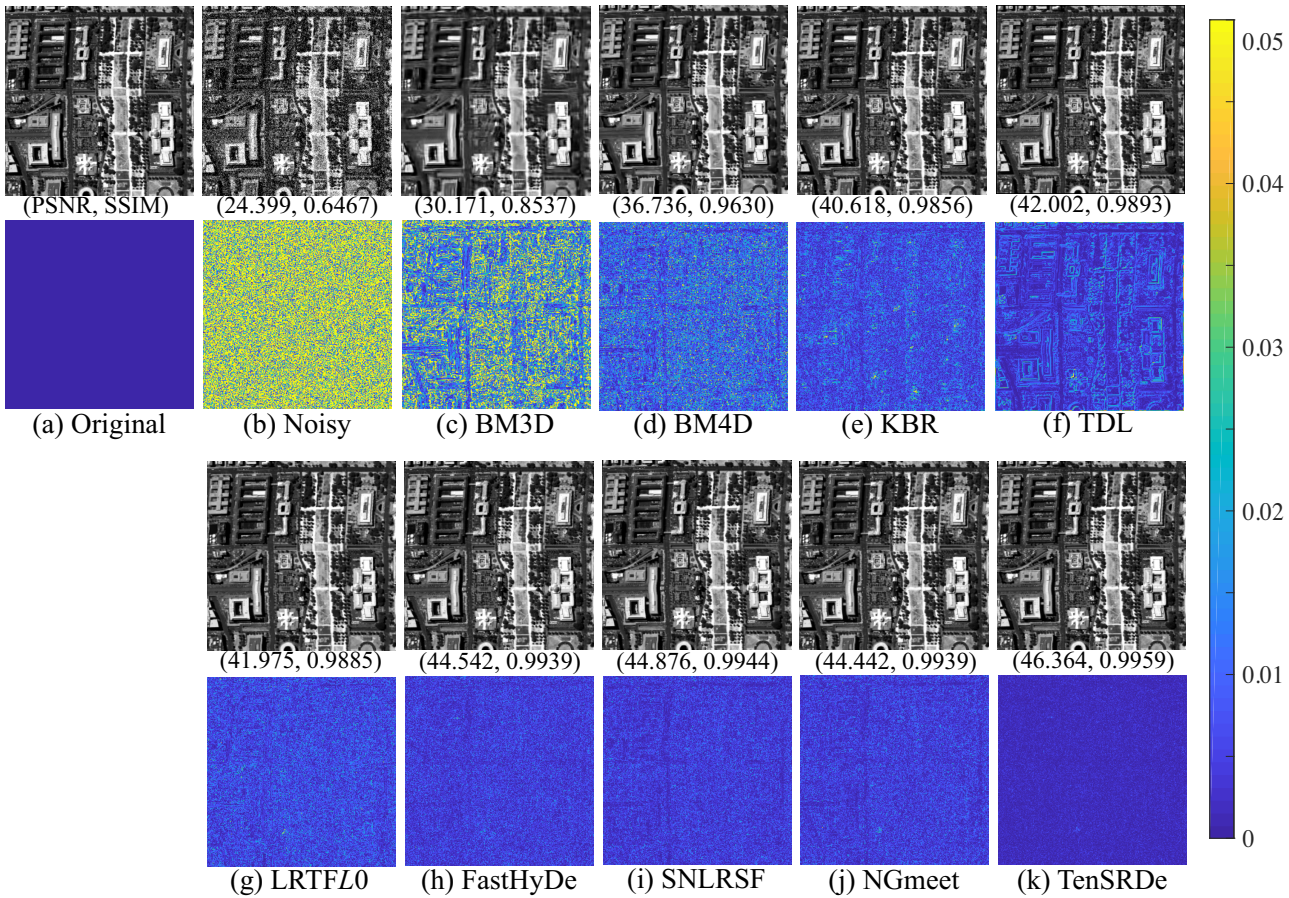


Figure 5: (a) Original image and residual image of the WDC dataset in band 163; (b) Noisy image and residual image under Gaussian noise level $\sigma = 0.06$; (c)-(k) Denoised images and residual images (the difference between each denoised band with the original band) of different methods.

Real Experiments

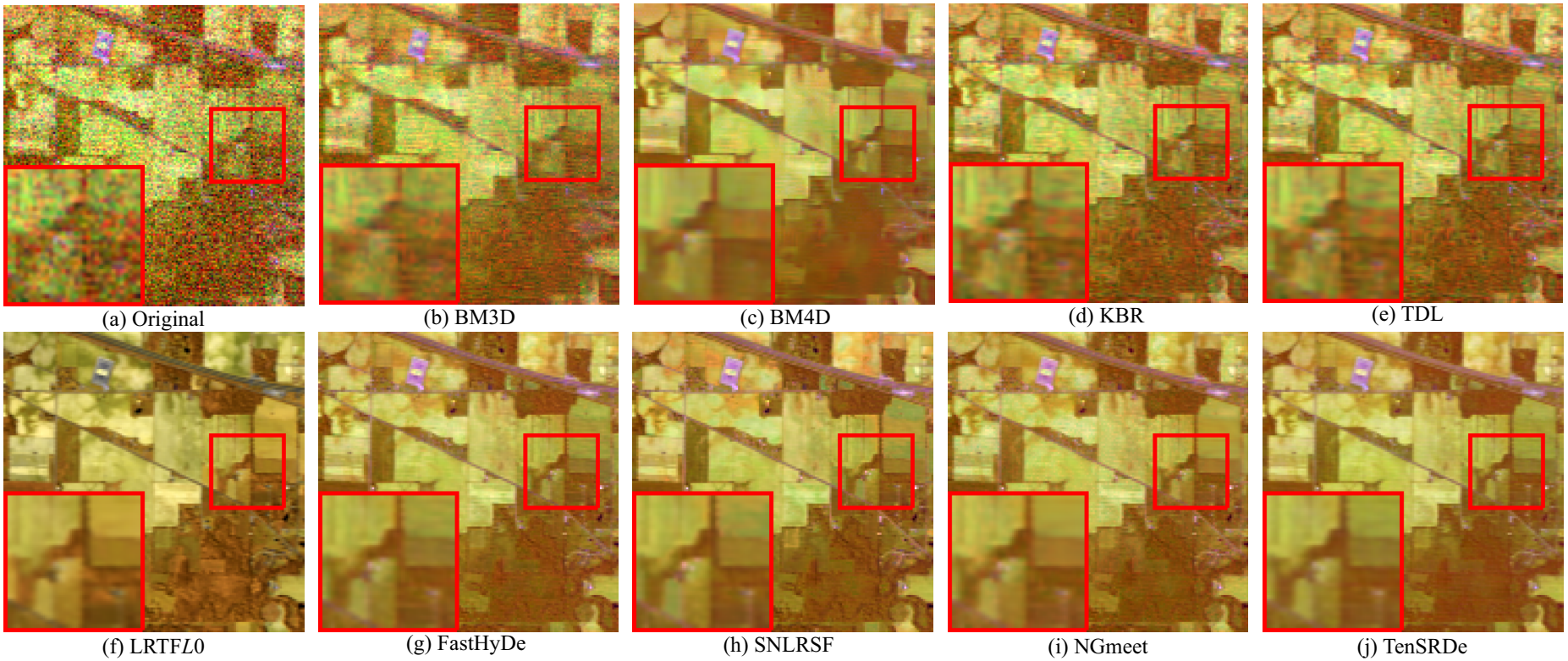


Figure 6: (a) Pseudo-color image (R: 200, G: 144, B: 3) of the original Indian Pines dataset; (b)-(j) Pseudo-color image of the denoising results by different methods.

Conclusions

- We design a novel basis tensor and develop a tensor subspace representation.
- We propose a TenSR-based HSI denoising model for Gaussian noise removal.
- We develop an efficient PAM algorithm to solve the proposed non-convex model.

Thank you!