

Robust Core Tensor Dictionary Learning with Modified Gaussian Mixture Model for Multispectral Image Restoration

Leilei Geng¹, Chaoran Cui¹, Qiang Guo¹, Sijie Niu², Guoqing Zhang³ and Peng Fu^{4,*}

Abstract: The multispectral remote sensing image (MS-RSI) is degraded existing multispectral camera due to various hardware limitations. In this paper, we propose a novel core tensor dictionary learning approach with the robust modified Gaussian mixture model for MS-RSI restoration. First, the multispectral patch is modeled by three-order tensor and high-order singular value decomposition is applied to the tensor. Then the task of MS-RSI restoration is formulated as a minimum sparse core tensor estimation problem. To improve the accuracy of core tensor coding, the core tensor estimation based on the robust modified Gaussian mixture model is introduced into the proposed model by exploiting the sparse distribution prior in image. When applied to MS-RSI restoration, our experimental results have shown that the proposed algorithm can better reconstruct the sharpness of the image textures and can outperform several existing state-of-the-art multispectral image restoration methods in both subjective image quality and visual perception.

Keywords: Multispectral remote sensing, image restoration, modified Gaussian mixture, sparse core tensor, tensor dictionary learning.

1 Introduction

The radiance of a real scene is distributed across a wide range of spectral bands. Traditional multispectral remote sensing image (MS-RSI) is achieved by integrating the product of the intensity at four typical band intervals. Characterized by the distinctive advantages of spatial-spectrum information, MS-RSI has been widely in many fields such as land-cover and resource survey, geological exploration, environment monitoring and natural disaster monitoring [Xu, Chen, Xia et al. (2018); Fu, Xu, Zhang et al. (2019); Canbaz, Gursoy and Gokce (2018); Liu, Yang, Zhao et al. (2017); Xia and Hu (2016)]. In real cases, however, an MS-RSI is always corrupted by some noises that are generally conducted by equipment limitations like sensor sensitivity, photon effects and calibration error [Gong and Chen (2019)]. Moreover, the MS-RSI includes both 2D spatial

¹ School of Computer Science and Technology, Shandong University of Finance and Economics, Jinan, 250014, China.

² School of Computer Science and Engineering, University of Jinan, Jinan, 250024, China.

³ School of Computer Science and Engineering, Nanyang Technological University, 639798, Singapore.

⁴ School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, 210094, China.

* Corresponding Author: Peng Fu. Email: fupeng@njust.edu.cn.

Received: 02 February 2020; Accepted: 30 April 2020.

information and 1D spectral information, it has severe limitations in radiation quality comparing against regular panchromatic image due to the out-of-focus blur of the optical system, the vibration of camera platform, the disturbance of the turbulent atmosphere, the lowpass filtering effect and the integral average effect of the lens. To obtain high-quality MS-RSI, the image restoration process is often necessary when the multispectral imaging system is fixed.

In the MS-RSI restoration field, there are mainly two approaches including the 2D approach and 3D-cube-based approach. As one of the classical problems in computer vision, 2D image restoration has been addressed for more than 50 years and a large amount of researches have been proposed on this problem, such as NLM [Penna and Mascarenhas (2018)], BM3D [Chen, Luo, Tian et al. (2017)], K-SVD [Chen, Liu, Huang et al. (2017)], dictionary learning [Zhang, Porikli, Sun et al. (2020)] and the low-rank matrix approximations. For example, the multilayer dictionary model restrains the noise by learning the image feature with similar spectral characters [Liu, Ma, Yang et al. (2017)]; the Laplacian scale mixture distribution and nonlocal low-rank regularization are used to regularize the image recovery process [Huang, Dong, Xie et al. (2017)]. These 2D methods can be directly applied to MS-RSI restoration by processing the images located at different bands separately. However, the details and textures cannot be reconstructed effectively caused by ignoring the inter-spectral correlation of MS-RSI.

The more reasonable extension is specifically designed for the 3D-cube-based image restoration methods [Sun and Jeon (2018); Chen and Shen (2015)]. Recently, the tensor-based approaches have led to many promising results, which regards MS-RSI as a 3rd-order tensor by applying the tensor factorization techniques [Peng, Meng, Xu et al. (2014); Dong, Huang, Shi et al. (2018)]. More specifically, Peng et al. proposed an effective nonlocal tensor dictionary learning method for multispectral image denoising. To exploit nonlocal spatiotemporal redundancy, each set of grouped similar 3D patches are linearly approximated by low-rank tensor approximation. Different from the rank of matrix, the multi-rank of tensor cannot be estimated accurately. Therefore, the sparse coding approaches have attracted increasingly more attention [Dian, Fang, Li et al. (2017); Yang, Wang, Li et al. (2015)]. Dian et al. proposed a noniterative recovery of hyperspectral images via sparse tensors and non-local spatial self-similarities. However, the existing multispectral tensor representation cannot take full use of the image structures and the spatial-spectral correlation. Our own recent works demonstrates the potential of sparse core tensor dictionary learning for MS-RSI restoration by exploiting self-repeating patterns in abundant similarity of edges and textures [Geng, Nie, Niu et al. (2018)].

To further exploit the sparse distribution prior of the core tensor, we propose a novel core tensor dictionary learning approach with the robust modified Gaussian mixture model (CTDL-MGM). The contributions of this paper are two-fold. First, to improve the accuracy of sparse core tensor coding, the asymmetric and the non-Gaussian distribution of MS-RSI data is modeled by modified Gaussian mixture model with non-zero means and positive scaling variables. Second, an efficient core tensor dictionary learning model using alternating optimization algorithm is proposed.

2 Proposed method

In this section, the tensor dictionary learning is presented firstly. Then we analyze the limitation of the existing approaches. Finally, we construct a novel core tensor dictionary learning approach with the robust modified Gaussian mixture model.

2.1 Tensor dictionary learning for MS-RSI

We aim at recovering a high quality MS-RSI $\mathcal{X} \in \mathbb{R}^{L \times N}$ from its distorted measurement \mathcal{Y} where $N=W \times H$ denotes the number of pixels and L is the number of spectral bands:

$$\mathcal{Y} = \mathcal{H}\mathcal{X} + \mathcal{N} \quad (1)$$

where \mathcal{H} is the blurring operators and \mathcal{N} is the additive noise. Considering both structural correlation in space and global correlation in spectrum, the MS-RSI patch is modeled by 3^{rd} -tensor $\mathcal{Y}_i = \mathcal{R}_i \mathcal{Y}$, (sized by $n \times n \times L$) with the operator matrix \mathcal{R}_i at spatial position i . Due to the redundant information, the spatial textures cannot be characterized with the multispectral tensor \mathcal{Y}_i . Then, we construct the core tensor $\tilde{\mathcal{S}}_i$ by high-order singular value decomposition (HOSVD) as in Eq. (2), which preserves spatial structure information with small amount of computation.

$$\mathcal{Y}_i = \tilde{\mathcal{S}}_i \times_1 U_{i,1} \times_2 U_{i,2} \times_3 U_{i,3} \quad (2)$$

where $U_{i,1}, U_{i,2}, U_{i,3} \in \mathbb{R}^{n \times n}$ are orthogonal matrices; the core tensor $\tilde{\mathcal{S}}_i \in \mathbb{R}^{n \times n \times L}$ is a 3D coefficient array; \times_i denotes the i -th mode product. Thanks for the both spatial-structural similarity and spectral correlation of the \mathcal{Y}_i , most components of $\tilde{\mathcal{S}}_i = \mathcal{Y}_i \times_1 U_{i,1}^\top \times_2 U_{i,2}^\top \times_3 U_{i,3}^\top$ are close to zero and the corresponding 1D representations \tilde{s}_i is highly sparse as in Fig. 1. Due to the orthogonality of the $U_{i,j} (j = 1,2,3)$, the $(U_{i,1}, U_{i,2}, U_{i,3})$ can be regarded as dictionary \mathcal{D} and the distorted \tilde{s}_i can be recovered using sparse model as follows:

$$\hat{s}_i = \arg \min_{s_i} \|\tilde{s}_i - s_i\|_2^2 + \lambda \psi(s_i) \quad (3)$$

where the $\|\tilde{s}_i - s_i\|_2^2$ is the fidelity constraint term and the $\psi(s_i)$ is the sparse constraint term. Popular choices $\psi(\cdot)$ of include the pseudo-norm l_0 and the l_1 norm, which exactly lead to the hard and soft thresholding of core tensor coefficient array \tilde{s}_i respectively.

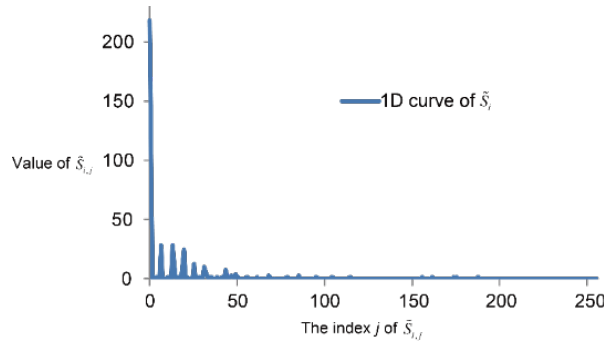


Figure 1: 1D curve of the core tensor

2.2 The limitation of the tensor dictionary learning

For effective MS-MRI restoration, the sparse core tensor $\hat{s}_{i,y}$ obtained from distorted \mathbf{Y} by Eq. (3) is expected to be as close as possible to the $\hat{s}_{i,x}$ obtained from clean \mathbf{X} . However, due to the degradation, $\hat{s}_{i,y}$ will deviate from its true value $\hat{s}_{i,x}$ and would directly decrease the quality of the reconstructed image. To further illustrate these differences, we define the deviation degree of the core tensor $\hat{s}_{i,y}$ as:

$$Dev(\hat{s}_{i,y}, \hat{s}_{i,x}) = \|\hat{s}_{i,y} - \hat{s}_{i,x}\|_F / \|\hat{s}_{i,x}\|_F \quad (4)$$

where $\|\cdot\|_F$ denotes the Frobenius norm. The definition of $Dev(\hat{s}_{i,y}, \hat{s}_{i,x})$ describes the similarity between $\hat{s}_{i,x}$ and $\hat{s}_{i,y}$. To show the $Dev(\hat{s}_{i,y}, \hat{s}_{i,x})$, we performed simulated experiments on distorted MS-RSI, degraded by the Gaussian blurring and Gaussian white noise. For fair comparison, both $\hat{s}_{i,x}$ and $\hat{s}_{i,y}$ were obtained using the same projection matrices $(U_{i,1}, U_{i,2}, U_{i,3})$. As shown in Fig. 2, most components of $Dev(\hat{s}_{i,y}, \hat{s}_{i,x})$ concentrated in 0.1~0.25 and the $\hat{s}_{i,y}$ deviates from its true value obviously. The definition of $Dev(\hat{s}_{i,y}, \hat{s}_{i,x})$ indicates that the accurate $\hat{s}_{i,y}$ cannot be obtained by the tensor dictionary learning model and the reconstructed performance can be further improved by learning the optimal solution of $\hat{s}_{i,y}$. This observation motivates us to model a novel tensor dictionary learning approach, as will be further discussed in Subsection 2.3.

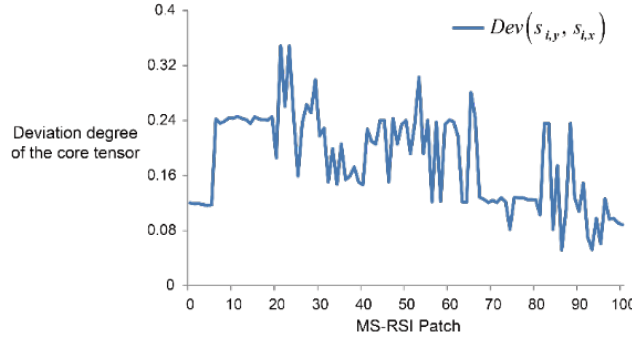


Figure 2: Distributions of $Dev(\hat{s}_{i,y}, \hat{s}_{i,x})$ in degraded MS-RSI

2.3 Core Tensor dictionary learning via modified Gaussian mixture modeling

As we can see from Eq. (3), the regularization function $\psi(\cdot)$ is critical in sparse tensor approximation. In this subsection, we propose a Maximum a Posterior (MAP) method for estimating s from distorted \tilde{s} , which can be formulated as:

$$s = \arg \min_s \{-\log P(\tilde{s}|s) - \log P(s)\} \quad (5)$$

where $\log P(\tilde{s}|s)$ is given by the Gaussian distribution of noise, i.e.:

$$\log P(\tilde{s}|s) \propto -\|\tilde{s} - s\|_2^2 \quad (6)$$

and $P(s)$ is the distribution prior of s . To exploit the nonlocally similar textures, the tensors $\{\mathcal{Y}_i, i = 1, \dots, N\}$ are clustered into K sets and their corresponding core tensor sets

$\{\Omega_k | \Omega_k = \{s_{k,j}\}\}$ are obtained. Then each core tensor set should be characterized by the same prior (i.e., the probability density function with the same parameters). Motivated by the BMM algorithm [Nguyen, Wu, Mukherjee et al. (2014)] and GM-based method [Mallouli (2019)], in this paper, we modify the asymmetric Gaussian mixture model to fit different shapes of observed data and incorporate the spatial information. The density function $f(s_i | \Pi, \Theta)$ for each core tensor s_i is defined as:

$$f(s_i | \Pi, \Theta) = \sum_{k=1}^K \pi_{ik} \cdot \prod_{s_j \in \phi_i} p(s_j | \theta_k)^{w_{ij}} \quad (7)$$

where $\Pi = \{\pi_{ik}\}$ is the set of between-cluster prior distributions which models the prior probability of core tensor s_i in label Ω_k , and satisfies the constraints $0 \leq \pi_{ik} \leq 1$ and $\sum_{k=1}^K \pi_{ik} = 1$. The probability $p(s_j | \theta_k)$ is a statistical distribution, and w_{ij} is the weight factor of j -th tensor \mathcal{Y}_j in the local neighborhood $\phi_i = \{s_1, \dots, s_m\}$ around tensor \mathcal{Y}_i , which will be presented in details in subsection 2.4. Since the observation s_i is generally considered as statistically independent among the core tensor set Ω_k , the joint conditional density over the core tensor set $s = \{s_i\}$ can be calculated using the novel modified Gaussian Mixture (MGM) Model:

$$\Phi(s | \Pi, \Theta) = \prod_{i=1}^N \sum_{k=1}^K \pi_{ik} \cdot \prod_{s_j \in \phi_i} p(s_j | \theta_k)^{w_{ij}} \quad (8)$$

The joint conditional density $\Phi(s | \Pi, \Theta)$ is mainly influenced by the probabilities of core tensors $p(s_j | \theta_k)$ in its neighborhood $\phi_i = \{s_1, \dots, s_m\}$ around tensor \mathcal{Y}_i . The distribution $p(s_j | \theta_k)$ is the important component which can be any kind of probability statistical distribution (e.g., the Laplacian, Generalized Gaussian, and student's t-distribution). However, these traditional distributions are symmetric and unimodal, which does not consider the spatial information and is often shorter than required. It means that these are not flexible enough to fit the shape of the multispectral core tensor s_i with non-Gaussian and asymmetric distribution. So each component of s can be characterized by a multiple-statistical distributions. In this paper, we propose a modified Gaussian mixture distribution with a positive scaling variable to model core tensor s_i , which can be written in the form:

$$p(s_i | \theta_k) = \frac{1}{(2\pi)^{D/2}} \cdot \frac{1}{|\Sigma_k|^{1/2}} \exp \left\{ -\frac{1}{2} (s_i - \mu_k)^T \Sigma_k^{-1} (s_i - \mu_k) \right\} \quad (9)$$

where $\Theta = \{\theta_k | k = (1, \dots, K)\}$ is the set of Gaussian parameters with $\theta_k = \{\mu_k, \Sigma_k\}$. The D -dimensional vector μ_k is the mean value, the $D \times D$ matrix Σ_k is the covariance value, and $|\Sigma_k|$ denotes the determinant of Σ_k . Consequently, we have also adopted this option in this work, which translates the core tensor dictionary learning (described in Eq. (5)) into:

$$(\hat{s}, \theta) = \arg \min_{s, \theta} \| \hat{s} - s \|_2^2 + \alpha \| s \|_1 + \frac{1}{2} \omega \log |\Sigma| + \frac{1}{2} \omega (s - \mu)^T \Sigma^{-1} (s - \mu) \quad (10)$$

where the $\| \hat{s} - s \|_2^2$ is the fidelity constraint term and the $\alpha \| s \|_1$ is the sparse constraint term. As shown in Fig. 2, the core tensor s is naturally sparse due to the both spatial-structural similarity and spectral correlation. Then the sparse constraint term $\alpha \| s \|_1$ can be removed. Therefore, the above core tensor dictionary learning model can be converted into:

$$(\hat{s}, \theta) = \arg \min_{s, \theta} \|\hat{s} - s\|_2^2 + \frac{1}{2} \omega \log |\Sigma| + \frac{1}{2} \omega (s - \mu)^T \Sigma^{-1} (s - \mu) \quad (11)$$

where ω is the weight factor vector of the core tensor s in its local neighborhood, which will be presented in details in subsection 2.4. μ is the mean vector with D dimension. Σ is the covariance matrix with $D \times D$ dimension, and $|\Sigma|$ is the determinant of Σ . We call such novel formulation *Core Tensor Dictionary Learning with the Modified Gaussian mixture* (CTDL-MGM) and the computationally efficient solution will be discussed in Section 3.

2.4 Construction of weight factor

In our proposed model, the conditional probability of the core tensor s_i is influenced by the probabilities of core tensors in its neighborhood ϕ_i . To incorporate the spatial-spectral information and core tensor intensity information, we add weigh factor w_{ij} for distant core tensors in order to distinguish among the contributions of different core tensors, as the weighted parameters decrease with increasing distance.

Most state-of-the-art algorithms with spatial-spectral information, including nonlocal core tensor dictionary learning method [Geng, Nie, Niu et al. (2018)] and clustering methods [Li, Zhang, Lu et al. (2017); Chen, Xiong, Xu et al. (2019)], have yielded effective results, but still have some disadvantages: (1) Although the introduction of local spatial and spectral information makes the corresponding algorithm overcome the impact of noise, they still lack enough robustness to noise with different types or levels, especially in absence of prior knowledge of the noise; (2) One or more controlling parameters are generally introduced to balance the robustness to noise and effectiveness of preserving the details. Unfortunately, the selection of these parameters has to be made by experience or trial and error experiments.

To overcome the above mentioned disadvantages, a new spatial weight factor should be constructed with the following characteristics: controlling the influence of the neighborhood core tensors depending on their distance from each other among the neighborhood to further improve the robustness to noise; utilizing the intensity information of the original image to make the algorithm avoid sensitive to initializations; being free of any balance parameter selection. Let R be the radius of the neighborhood ϕ_i centered as core tensor s_i . There are $(2R + 1)^2$ core tensors in each neighborhood obviously, but not all the core tensors are useful for similarity computation. In other words, different core tensors in the neighborhood should have different weights. Therefore, we construct a novel weight factor defined as:

$$w_{ij} = \frac{\exp\left(-\sum_{s_m \in \phi_i, m \neq j} \frac{\|s_j - s_m\|_2}{N_{\phi_i} \times (N_{\phi_i} - 1)}\right)}{\sum_{s_h \in \phi_i} \exp\left(-\sum_{s_m \in \phi_i, m \neq j} \frac{\|s_h - s_m\|_2}{N_{\phi_i} \times (N_{\phi_i} - 1)}\right)} \quad (12)$$

where N_{ϕ_i} is the number of core tensors in the neighborhood ϕ_i for core tensor s_i . As shown in above equation, it can be seen that the factor w_{ij} is formulated without setting any control or balance parameters to balance the robustness to noise and effectiveness of preserving the details. Moreover, the influence of core tensors within the local window in w_{ij} is flexibly by using their spatial-spectral Euclidean distance from other core tensors in

the neighborhood. Therefore, the proposed weight factor can reflect the damping extent of the neighbors with the spatial-spectral distance from other neighborhood core tensors. Thus more local spatial information can be used. It is worth noting that the shape of the local window used in our experiments is square, but also, windows with other shapes like circle or diamond can be easily adapted to the proposed weight factor.

3 MS-RSI Restoration with CTDL-MGM

In the previous sections, we have seen how to solve CTDL-MGM problem for a single core tensor s (a collection of image patches similar to a chosen exemplar). In this section, we generalize such formulation to whole-image reconstruction and study the applications of CTDL-MGM into image restoration including image denoising and image deblurring. For the degraded image $\mathbf{y} = \mathcal{H}\mathbf{X} + \mathcal{N}$ (as in Eq. (1)), the whole-image \mathbf{X} can be reconstructed by solving the optimization problem:

$$(\mathbf{X}, \{\theta_i\}) = \arg \min_{s, \theta} \|\mathbf{y} - \mathbf{X}\|_F^2 + \sum_i \|\tilde{\mathcal{R}}_i \mathbf{X} - S_i \times_1 U_{i,1} \times_2 U_{i,2} \times_3 U_{i,3}\|_F^2 + \frac{1}{2} \omega \sum_i \log |\Sigma_i| + \frac{1}{2} \omega \sum_i (S_i - \mu_i)^T \Sigma_i^{-1} (S_i - \mu_i) \quad (13)$$

where $\tilde{\mathcal{R}}_i = [\tilde{\mathcal{R}}_{i_0}, \tilde{\mathcal{R}}_{i_1}, \dots, \tilde{\mathcal{R}}_{i_{m-1}}]$ denotes an operator extracting similar 3D tensors at spatial position i ; $\theta_i = \{\mu_i, \Sigma_i\}$ is the parameter in modified Gaussian mixture model. The above global minimization problem can be decomposed into the following two subproblems: minimization of \mathbf{X} for a fixed $\{S_i\}$ and $\{\theta_i\}$; minimization of $\{S_i\}$ and $\{\theta_i\}$ for a fixed \mathbf{X} , which can be efficiently solved by the method of alternating optimization. Specifically, both subproblems admits closed-form solutions when the dictionary $(U_{i,1}, U_{i,2}, U_{i,3})$ is orthogonal.

3.1 Solving for whole image

Let $\hat{\mathcal{X}}_i = \hat{S}_i \times_1 U_{i,1} \times_2 U_{i,2} \times_3 U_{i,3}$ denotes the reconstructed tensor with an initial estimate of S_i . Then, for a fixed $\{\hat{S}_i\}$, the whole image \mathbf{X} can be recovered by solving the following l_2 -minimization problem:

$$\mathbf{X} = \arg \min_{\mathbf{X}} \|\mathbf{y} - \mathbf{X}\|_F^2 + \eta \sum_{i=1}^N \|\tilde{\mathcal{R}}_i \mathbf{X} - \hat{\mathcal{X}}_i\|_F^2 \quad (14)$$

where η is an increasing factor and $\tilde{\mathcal{R}}_i$ is the operator matrix at spatial position i . Then, the reconstructed MS-MRI $\hat{\mathbf{X}}$ is reconstructed by averaging each restored tensor using straightforward least-square solution:

$$\hat{\mathbf{X}} = \left(\lambda \mathcal{H}^T \mathcal{H} + \eta \sum_i \tilde{\mathcal{R}}_i^T \tilde{\mathcal{R}}_i \right)^{-1} \left(\lambda \mathcal{H}^T \mathbf{y} + \eta \sum_i \tilde{\mathcal{R}}_i^T \hat{\mathbf{y}}_i \right) \quad (15)$$

where $\hat{\mathbf{y}}_i$ can be constructed by $\hat{\mathbf{y}}_i = \hat{S}_i \times_1 U_{i,1} \times_2 U_{i,2} \times_3 U_{i,3}$, $\tilde{\mathcal{R}}_i^T \tilde{\mathcal{R}}_i = \sum_{j=0}^{m-1} \tilde{\mathcal{R}}_{i_j}^T \tilde{\mathcal{R}}_{i_j}$ and $\tilde{\mathcal{R}}_i^T \hat{\mathbf{y}}_i = \sum_{j=0}^{m-1} \tilde{\mathcal{R}}_{i_j}^T \hat{\mathbf{y}}_{i_j}$. Note that it is image denoising application where $\mathcal{H} = \mathcal{I}$, and can be computed easily.

3.2 Solving for $\{s_i\}$ and $\{\theta_i\}$

When the image \mathcal{X} is fixed, the subproblem boils down to a sequence of tensor-level CTDL-MGM problems, i.e., for each exemplar:

$$(\hat{s}_i, \theta_i) = \arg \min_{s, \theta} \|\tilde{s}_i - s_i\|_2^2 + \frac{1}{2} \omega \log |\Sigma_i| + \frac{1}{2} \omega (s_i - \mu_i)^T \Sigma_i^{-1} (s_i - \mu_i) \quad (16)$$

where we have used $\tilde{S}_i = \mathcal{X}_i \times_1 U_{i,1}^\top \times_2 U_{i,2}^\top \times_3 U_{i,3}^\top$. This is exactly the problem studied in the previous section. The μ_i and Σ_i can be computed by:

$$\mu_i = \sum_{s_j \in \phi_i} w_{ij} \cdot s_j \quad (17)$$

$$\Sigma_i = \sum_{s_j \in \phi_i} w_{ij} (s_j - \mu_i)(s_j - \mu_i)^T \quad (18)$$

where w_{ij} is the weight factor constructed in Subsection 2.4. Putting things together, a complete image restoration based on CTDL-MGM can be summarized as following section.

3.3 Algorithm of CTDL-MGM

The main procedure of the proposed CTDL-MGM model is summarized in **Algorithm 1**. Compared with the existing approaches, there three main advantages in the CTDL-MGM model as follows: (1) The MS-RSI patch is modeled by 3rd-tensor preserving both structural correlation in space and global correlation in spectrum. (2) The core tensor sparse-based model is constructed to solve the core tensor approximation effectively. (3) Exploiting the sparse distribution prior of the core tensor, the proposed CTDL-MGM can obtain a better estimation of the core tensor.

Algorithm 1 CTDL-MGM model for MS-MRI restoration

- **Initialization:**

Initialize the estimate $\hat{\mathcal{X}}^{(0)} = \mathcal{Y}$, the parameter γ and the number of cluster K .

- **Outer loop:** iterate on $t = 1, 2, \dots, T_{max}$.

1. Obtain the similar tensors set $\{\phi_i\}_i^{K(t)}$ from $\hat{\mathcal{X}}^{(t)}$ via K-Means.

2. Calculate the core tensor vector s_i for each tensor $\hat{\mathcal{X}}_i^{(t)}$ with HOSVD.

3. **Inner loop** (core tensor $\{s_i\}$ and statistical distribution $\{\theta_i\}$ approximation by solving Eq. (12)): iterate on $l = 1, 2, \dots, L_{max}$.

(1) Calculate the weight factor w_{ij} for each core tensor s_i with Eq. (13).

(2) Update both the mean values μ_i with Eq. (18) and the covariance value Σ_i with Eq. (19) for fixed s_i .

(3) Update \hat{s}_i for fixed μ_i and Σ_i .

End for

4. Reconstruct $\hat{\mathcal{X}}^{(t+1)}$ from $\{\hat{s}_i\}$ with Eq. (16).

5. If $\|\hat{\mathcal{X}}^{(t+1)} - \hat{\mathcal{X}}^{(t)}\|_2^2 \leq \varepsilon$, iteration is stopped.

End for

- **Output** $\hat{\mathcal{X}}^{(t+1)}$

4 Experimental results

In this section, we apply the proposed CTDL-MGM algorithm to image restoration,

including image denoising and image deblurring. To verify the performance of our proposed method, we utilized the 16 ZY-3 remote sensing images, each with 4 spectral bands including red, blue, green and near-infrared. The parameters of the proposed method are set as follows: the multispectral tensor size is $6 \times 6 \times 4$ with 2-pixel-width overlap and its corresponding core tensor size is $6 \times 6 \times 4$; the number of cluster is $K=10$; the mean values for each Gaussian component are initialized randomly; the covariance matrices and the prior probabilities are then initialed based on the means. To comprehensively assess the performance, we employ two objective image quality indices: peak signal-to-noise ratio (PSNR) and feature similarity (FSIM).

4.1 Image denoising

In the subsection, the Gaussian noise is added to the MS-RSIs, which comes from many natural sources, such as the spontaneous thermal generation of electrons. We parameterized the Gaussian noise by its standard deviation $\sigma=10, 20, 35, 50$. We have compared CTDL-MGM based image denoising method against three current state-of-the-art methods, including the structured sparse coding with Gaussian scale mixture (SSC-GSM) [Dong, Shi, Ma et al. (2015)], hierarchical dictionary learning (HDL) [Liu, Ma, Yang et al. (2017)], nonlocal tensor dictionary learning (NTDL) [Peng, Meng, Xu et al. (2014)]. Both average PSNR (dB) and average FSIM at different noise levels are reported in Tab. 1. As can be seen from Tab. 1, the proposed CTDL-MGM method consistently outperforms all other leading algorithms. The average PSNR improvements over HDL, SSC-GSM and NTDL methods are larger than about 3 dB, 3 dB and 1 dB, respectively. Besides, the average FSIM improvements 0.01~0.02 on low noise levels ($\sigma=10, 20$) and 0.04~0.05 on high noise levels ($\sigma=35, 50$). On the average, HDL and SSC-GSM are mostly performing methods. However, they fall behind 3rd-tensor dictionary learning methods (including NTDL and CTDL-MGM) by less than 1.5 dB in PSNR and 0.01 in FSIM. By exploiting the sparse distribution prior of the multispectral 3rd-tensor, the proposed CTDL-MGM method has achieved better competitive denoising performance to the NTDL using nonlocal tensor similarities.

Table 1: The PSNR (dB) and FSIM of the denoised MS-RSIs

Noise Level		HDL	SSC-GSM	NTDL	Proposed CTDL-MGM
$\sigma = 10$	PSNR	34.13	34.19	35.94	37.21
	FSIM	0.987	0.986	0.993	0.995
$\sigma = 20$	PSNR	30.24	30.28	31.88	32.85
	FSIM	0.966	0.965	0.982	0.986
$\sigma = 35$	PSNR	27.37	27.31	29.87	30.32
	FSIM	0.935	0.930	0.967	0.970
$\sigma = 50$	PSNR	25.65	25.63	27.85	28.17
	FSIM	0.907	0.900	0.950	0.953

Fig. 3 includes the visual comparison of denoising results for ZY-3 remote sensing images at the moderate noise level ($\sigma=20$). It can be easily observed that the restored images by HDL and SSC-GSM both suffer from noticeable artifacts especially around the

smooth areas close to the bridge; by contrast, the NTDL and CTDL-MGM based on tensor dictionary learning seem to deliver the better visual quality. Obviously, the MS-RSI restored by our method properly removes the noise while finely preserves the structures underlying the image.

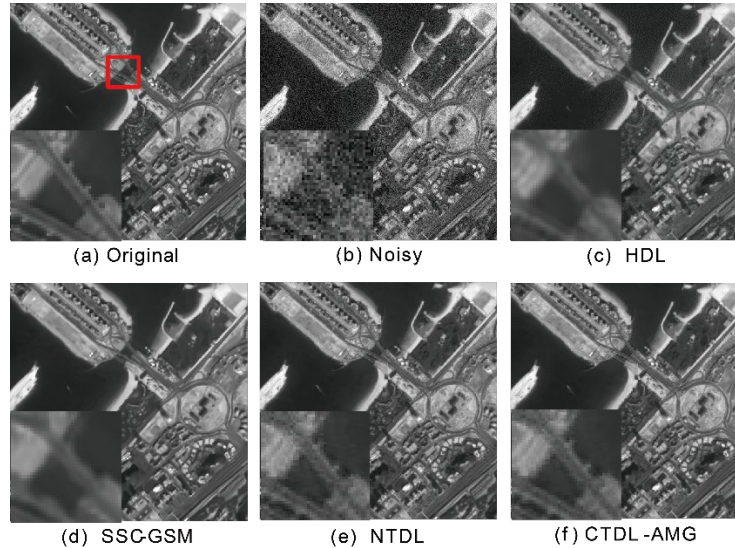


Figure 3: Denoising performance comparison: (a) Original images at red band (image size: 400×400); (b) The image corrupted by Gaussian noise of $\sigma=20$; denoised images by (c) HDL (PSNR=30.13 dB, FSIM=0.967); (d) SSC-GSM (PSNR=30.41 dB, FSIM=0.964); (e) NTDL (PSNR=31.92 dB, FSIM=0.983); (f) the proposed CTDL-MGM (PSNR=32.87 dB, FSIM=0.986)

4.2 Image deblurring

We have also compared CTDL-MGM based image deblurring and three other competing approaches in the literature: structural compact core tensor dictionary learning (SC-CTDL) [Geng, Nie, Niu et al. (2018)], multispectral image deblurring with jitter image trajectory (MID-JIT) [Song, Fang, Pan et al. (2017)], multispectral image deblurring using interchannel correlation (MID-IC) [Chen and Shen (2015)]. The performance of the proposed CTDL-MGM model was verified on a set of 16 MS-RSIs. The original images were blurred by a blur kernel and then added by the additive Gaussian white noise with standard deviation $\sqrt{2}$. In our comparative study, two blur kernels were commonly used: (1) 9×9 uniform blur simulating the relative motion between the subject and multispectral camera; (2) Gaussian blur with standard deviation 1.6 simulating the optics blur.

Uniform blur. We have applied the proposed method and three other competing approaches on 16 blurred scenes by uniform blur. The PSNR and FSIM curves are depicted in Fig. 4. In PSNR results, the proposed CTDL-MGM method significantly outperforms three other methods by 2~3 dB. For the FSIM, our proposed method is better than three other algorithms about 0.04. From Figs. 5(c) to 5(f), we present the deblurring results of a city region at blue band. It can be easily observed that MID-JIT and MID-IC are quit poor

without considering the correlation between spatial and spectral information into account. The SC-CTDL method constructs the central texture based on the details in its neighborhood which are also corrupted. Therefore, the corresponding result is not satisfied enough especially for the region with abundant edges or textures. The proposed method leads to the best visual quality. It cannot only remove the blurring effects and noise, but also reconstruct more and sharper image edges than other methods. The excellent edge preservation owes to exploiting the sparse distribution prior of the core tensor.

Gaussian blur. For 7×7 Gaussian blurring with standard deviation 1.6, the superior performance of our proposed CTDL-MGM is demonstrated as shown in Fig. 6. We can see that the proposed method achieves the best results. Its performance is more outstanding than other three methods with 1~3 dB in PSNR and 0.01~0.06 in FSIM. Moreover, the MID-JIT and MID-IC is comparable in FSIM. To further verify the effectiveness of the proposed method, we present the deblurring island image at green band in Fig. 7. The MID-JIT produces over-smoothed results and eliminates much image details. The MID-IC also fails to reconstruct fine image edges and tends to generate some “ghost” artifacts around the edges. The SC-CTDL results seem the acceptable visual quality. In contrast, the proposed method presents the clearer and more details images when compared with the ground truth.

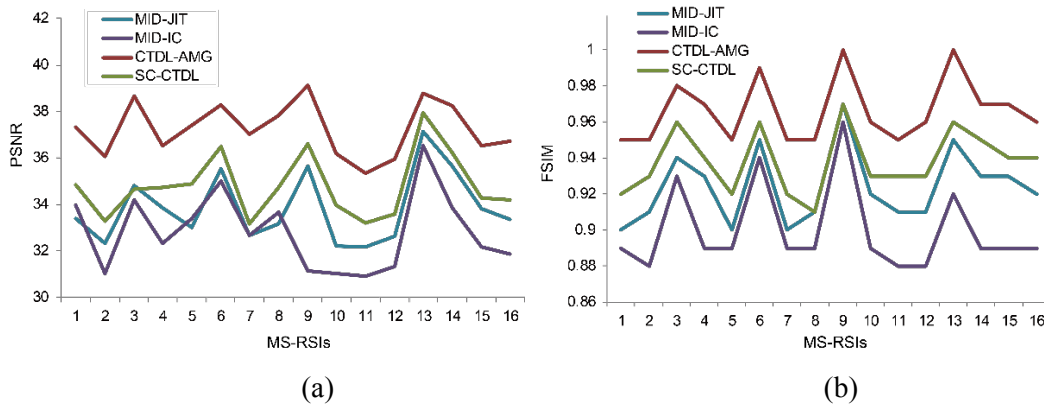


Figure 4: Performance of four image restoration algorithms for the uniform blurring: (a) PSNR curves; (b): FSIM curves

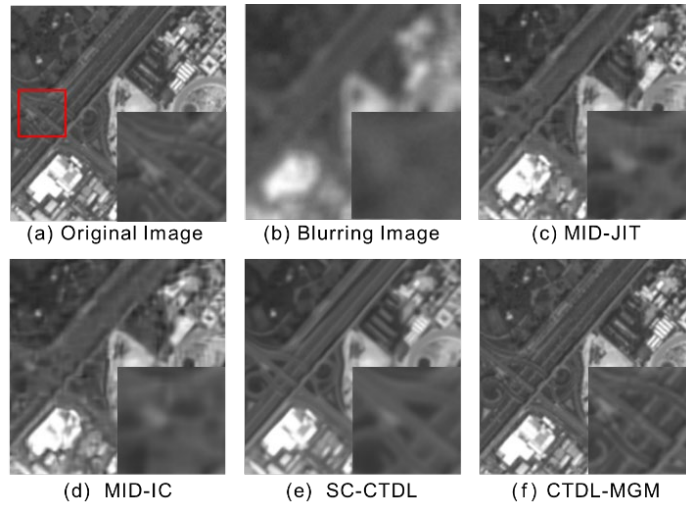


Figure 5: Deblurring performance comparison: (a) Original images at blue band (image size: 144×144); (b) Noisy and blurred image (9×9 uniform blur, $\sigma = \sqrt{2}$); deblurred images by (c) MID-JIT (PSNR=34.81 dB, FSIM=0.941); (d) MID-IC (PSNR=34.20 dB, FSIM=0.934); (e) SC-CTDL (PSNR=34.66 dB, FSIM=0.963); (f) the proposed CTDL-MGM (PSNR=38.78 dB, FSIM=0.986)

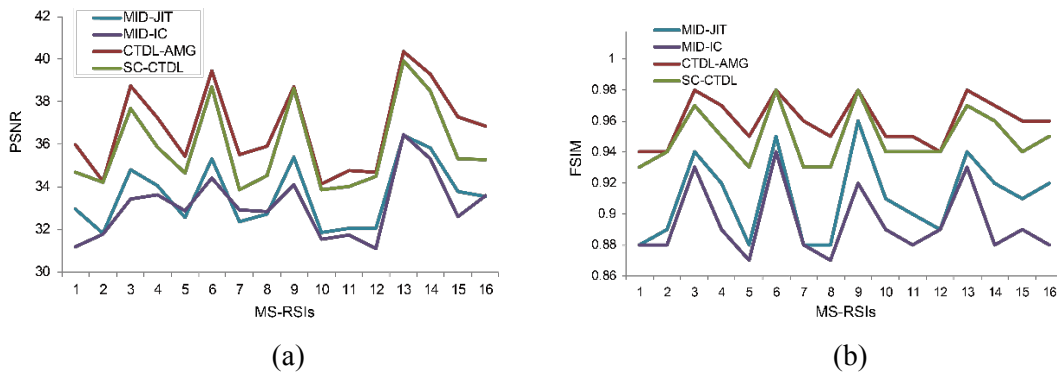


Figure 6: Performance of four image restoration algorithms for the Gaussian blurring: (a) PSNR curves; (b) FSIM curves

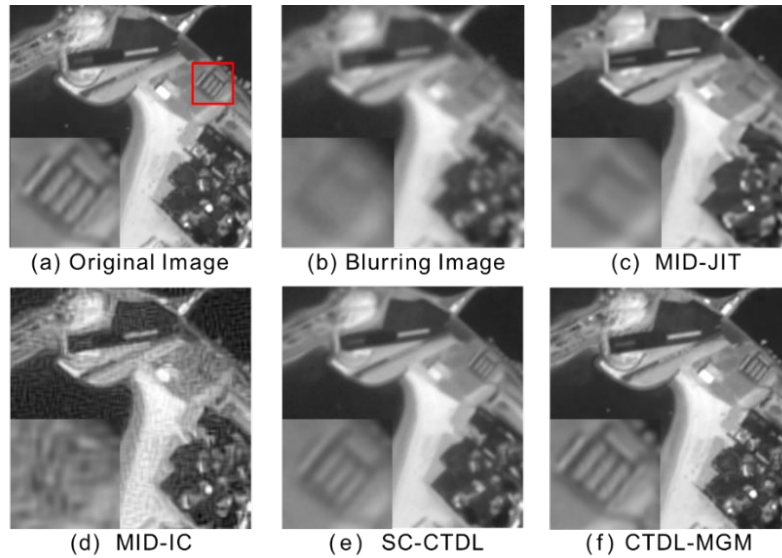


Figure 7: Deblurring performance comparison: (a) Original images at green band (image size: 144×144); (b) Noisy and blurred image (Gaussian blur, $\sigma=1.6$); deblurred images by (c) MID-JIT (PSNR=34.86 dB, FSIM=0.944); (d) MID-IC (PSNR=33.42 dB, FSIM=0.933); (e) SC-CTDL (PSNR=37.69 dB, FSIM=0.972); (f) the proposed CTDL-MGM (PSNR=38.98 dB, FSIM=0.983)

Overall the proposed CTDL-MGM method leads to the best result on both subjective image quality and visual perception. One of the main reasons is that there is essential difference of data structure for MS-RSI patch between the proposed method and other 2D-patch-based methods. In CTDL-MGM method, the MS-RSI patch is conducted by tensor incorporating spatial structure and spectral correlation. Additionally, the proposed method is capable of greatly improving the accuracy of core tensor by exploiting the sparse distribution prior. In contrast, MID-JIT ignored the spectral correlation and restored spectral image separately; MID-IC focused out-of-focus blur but not on other blur types; SC-CTDL emphasized more on exploiting nonlocal similarity and failed to strike a good balance between local variation and nonlocal invariance. Therefore, our proposed method can outperform other three methods.

5 Conclusions

To further improve the accuracy of sparse core tensor and the quality of reconstructed image, in this paper, we presented a novel prior distribution for MS-RSI restoration. In order to incorporate spatial-structural similarity and spectral correlation of MS-RSIs into a unified representation, the MS-RSI patch is modeled by 3^{rd} -tensor. Considering both the sparsity of core tensors and the orthogonality of projection matrices, the core tensors are restored by the sparse core tensor dictionary learning model. For the purpose of improving the accuracy of core tensors, the modified Gaussian mixture is introduced into the core tensor dictionary learning model. To this end, the core tensor dictionary learning approach with the modified Gaussian mixture model (CTDL-MGM) is proposed. An

efficient alternating optimization algorithm was presented for solving the CTDL-MGM minimization problem. Our results on Gaussian noise, uniform blur kernel and Gaussian blur kernel show that the proposed algorithm is capable of producing more accurate image restoration results than several state-of-the-art algorithms.

One limitation of the proposed algorithm is the manual selection of the window size for the tensor. One possible extension of this work is to set the radius based on the amount of noise in the target image. Making our algorithm with adaptive window size selection will be our direction of the future work. Another limitation of the proposed algorithm is how to determine the number of classes/components automatically for different applications. The traversal strategy is generally used to determine the number of clusters/components based on the criteria in mixture models. Thus, addressing such questions is out of the scope of this paper and subjects of future research.

Acknowledgment: We thank anonymous reviewers for their feedback which helped in the improvement and presentation of this article. We also acknowledge Professor Quansen Sun (The Nanjing University of Science and Technology, China) for his feedback on an earlier version of the manuscript.

Funding Statement: This work was supported by the Project of Shandong Province Higher Educational Science and Technology Program [KJ2018BAN047, Geng, L.]; National Natural Science Foundation of China [61801222, Fu, P.]; Fundamental Research Funds for the Central Universities [30919011230, Fu, P.]; Science and Technology Innovation Program for Distributed Young Talents of Shandong Province Higher Education Institutions [2019KJN045, Guo, Q.]; Shandong Provincial Key Laboratory of Network Based Intelligent Computing [<http://nbic.ujn.edu.cn/>].

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- Canbaz, O.; Gursoy, O.; Gokce, A.** (2018): Detecting clay minerals in hydrothermal alteration areas with integration of ASTER image and spectral data in Kösedag-Zara (Sivas), Turkey. *Journal of the Geological Society of India*, vol. 91, no. 4, pp. 483-488.
- Chen, G.; Luo, G.; Tian, L.; Chen, A.** (2017): Noise reduction for images with non-uniform noise using adaptive block matching 3D filtering. *Chinese Journal of Electronics*, vol. 26, no. 6, pp. 1227-1232.
- Chen, J.; Liu, S.; Huang, M.; Gao, J.** (2017): Dictionary learning for MRI denoising based on modified K-SVD. *Journal of Imaging Science & Technology*, vol. 61, no. 3, pp. 1-10.
- Chen, S.; Shen, H.** (2015): Multispectral image out-of-focus deblurring using interchannel correlation. *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 4433-4445.
- Chen, Y.; Xiong, J.; Xu, W.; Zuo, J.** (2019): A novel online incremental and decremental learning algorithm based on variable support vector machine. *Cluster Computing*, vol. 22, no. 8, pp. 7435-7445.

Dian, R.; Fang, L.; Li, S. (2017): Hyperspectral image super-resolution via non-local sparse tensor factorization. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3862-3871.

Dong, W.; Huang, T.; Shi, G.; Ma, Y.; Li, X. (2018): Robust tensor approximation with Laplacian scale mixture modeling for multiframe image and video denoising. *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 6, pp. 1435-1448.

Dong, W.; Shi, G.; Ma, Y.; Li, X. (2015): Image restoration via simultaneous sparse coding: where structured sparsity meets gaussian scale mixture. *International Journal of Computer Vision*, vol. 114, no. 2-3, pp. 217-232.

Fu, P.; Xu, Q.; Zhang, J.; Geng, L. (2019): A noise-resistant superpixel segmentation algorithm for hyperspectral images. *Computers, Materials & Continua*, vol. 59, no. 2, pp. 509-515.

Geng, L.; Nie, X.; Niu, S.; Yin, Y.; Lin, J. (2018): Structural compact core tensor dictionary learning for multispectral remote sensing image deblurring. *IEEE International Conference on Image Processing*, pp. 2865-2869.

Gong, X.; Chen, W. (2019): Multi-spectral image denoising with shared dictionaries and low-rank representation. *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1707-1711.

Huang, T.; Dong, W.; Xie, X.; Shi, G.; Bai, X. (2017): Mixed noise removal via Laplacian scale mixture modeling and nonlocal low-rank approximation. *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3171-3186.

Li, T.; Zhang, L.; Lu, W.; Hou, H.; Liu, X. et al. (2017): Interval kernel fuzzy C-Means clustering of incomplete data. *Neurocomputing*, vol. 237, pp. 316-331.

Liu, S.; Ma, W.; Yang, S.; Chen, P. (2017): Multispectral imagery denoising using hierarchical dictionary learning with spatial-spectral information. *Journal of Xidian University*, vol. 44, no. 4, pp. 174-180.

Liu, S.; Yang, Y.; Zhao, Z.; Yao, Y.; Mao, X. et al. (2017): Application requirement analysis of high spectral and high spatial resolution satellite for environment remote sensing monitoring. *Chinese Society for Optical Engineering Conferences*, vol. 10255, pp. 339-343.

Mallouli, F. (2019): Robust EM algorithm for Iris segmentation based on mixture of Gaussian distribution. *Intelligent Automation and Soft Computing*, vol. 25, no. 2, pp. 243-248.

Nguyen, T. M.; Wu, Q. M. J.; Mukherjee, D.; Zhang, H. (2014): A Bayesian bounded asymmetric mixture model with segmentation application. *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 1, pp. 109-119.

Peng, Y.; Meng, D.; Xu, Z.; Gao, C.; Yang, Y. et al. (2014): Decomposable nonlocal tensor dictionary learning for multispectral image denoising. *IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 2949-2956.

Penna, P. A. A.; Mascarenhas, N. D. A. (2018): (Non-) homomorphic approaches to denoise intensity SAR images with non-local means and stochastic distances. *Computers and Geosciences*, vol. 111, pp. 127-138.

Song, X.; Fang, Z.; Pan, J. (2017): Image deblurring method for satellite platform caused by jitter. *Journal on Communications*, vol. 38, no. Z1, pp. 186-192.

Sun, L.; Jeon, B. (2018): Hyperspectral mixed denoising via subspace low-rank learning and BM4D filtering. *IEEE International Geoscience and Remote Sensing Symposium*, pp. 8034-8037.

Xia, Q.; Hu, Z. (2016): A novel method to monitor coal fires based on multi-spectral Landsat images. *Spectroscopy and Spectral Analysis*, vol. 36, no. 8, pp. 2712-2720.

Xu, Z.; Chen, J.; Xia, J.; Du, P.; Zheng, H. et al. (2018): Multisource earth observation data for land-cover classification using random forest. *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 5, pp. 789-793.

Yang, S.; Wang, M.; Li, P.; Jin, L.; Wu, B. et al. (2015): Compressive hyperspectral imaging via sparse tensor and nonlinear compressed sensing. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 11, pp. 5943-5957.

Zhang, G.; Porikli, F.; Sun, H.; Sun, Q.; Xia, G. et al. (2020): Cost-sensitive joint feature and dictionary learning for face recognition. *Neurocomputing*.