Pattern Recognition 122 (2022) 108280

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/patcog

Hyperspectral super-resolution via coupled tensor ring factorization

Wei He[a], Yong Chen[b,*], Naoto Yokoya[a,c], Chao Li[a], Qibin Zhao[a]

[a] RIKEN Center for Advanced Intelligence Project, RIKEN, Tokyo 103-0027, Japan
[b] School of Computer and Information Engineering, Jiangxi Normal University, Nanchang 330022, China
[c] Department of Complexity of Science and Engineering, The University of Tokyo, Tokyo 113-8654, Japan

ARTICLE INFO

Article history:
Received 24 February 2020
Revised 15 March 2021
Accepted 24 August 2021
Available online 26 August 2021

Keywords:
Coupled tensor ring decomposition
Super-resolution
Hyperspectral
Multispectral

ABSTRACT

Hyperspectral super-resolution (HSR) fuses a low-resolution hyperspectral image (HSI) and a high-resolution multispectral image (MSI) to obtain a high-resolution HSI (HR-HSI). In this paper, we propose a new model called coupled tensor ring factorization (CTRF) for HSR. The proposed CTRF approach simultaneously learns the tensor ring core tensors of the HR-HSI from a pair of HSI and MSI. The CTRF model can separately exploit the low-rank property of each class (Section 3.3), which has not been explored in previous coupled tensor models. Meanwhile, the model inherits the simple representation of coupled matrix/canonical polyadic factorization and flexible low-rank exploration of coupled Tucker factorization. We further introduce spectral nuclear norm regularization to explore the global spectral low-rank property. The experiments demonstrated the advantage of the proposed nuclear norm regularized CTRF model compared to previous matrix/tensor and deep learning methods.

1. Introduction

Hyperspectral images (HSIs) have been widely used in remote sensing and computer vision applications because they can provide rich and varied spectral information [1–5]. However, owing to the limitations of imaging techniques used [6,7], there is a trade-off between spatial and spectral resolution [8,9]. It is difficult to obtain high-resolution HSIs (HR-HSI) using hyperspectral sensors. In addition, the multispectral imaging sensors limit the spectral resolution to obtain high-resolution multispectral images (MSIs) [7,10]. Therefore, *hyperspectral super-resolution* (HSR) or *hyperspectral and multispectral image fusion*, which fuses low-resolution HSIs and high-resolution MSIs to generate HR-HSIs [7,11], is important in real-world applications.

HSR has been studied extensively since decades [12]. Initially, HSR attempts to fuse an HSI and a panchromatic image [13], including multiresolution analysis (MRA) [12] and sparse representation-based methods [14]. However, these methods were limited to enhancing spatial details in practice [7]. Subsequently, a Bayesian framework was introduced to obtain a HR-HSI from the HSI and MSI, *i.e.,* maximum a posteriori (MAP) [15,16] and Bayesian sparse representation [17]. Very recently, deep learning has also been introduced to fuse the HSI and MSI, and achieving remarkable results [18–22]. In [18], an unsupervised sparse dirichlet-net was introduced for HSR. Following this approach, [23] proposed an unsupervised convolutional neural network (CNN) method with automatical selection of the camera spectral response of the MSI. In [24], an unsupervised adaptation learning method for HSR was introduced; a general image prior to the synthetic data was learned, and then the learned image prior was adapted to the specific HR-HSI under unknown degeneration. The authors of [19] realized a supervised learning method considering the spectral mode of an HR-HSI and introduced an MSI/HSI fusion net. In [25], a progressive zero-centric residual network was proposed for HSR; the network employed spectral-spatial separable convolution with dense connections. In [22], the HSR was reduced to image denoising and a CNN image denoiser was introduced to perform the task.

Unsupervised matrix/tensor-related methods have also been investigated. Over the past few years, numerous state-of-the-art HSR methods have been developed using matrix factorization models [10,11,16,17,26–34]. The original HR-HSI is reshaped into the matrix, and the matrix is decomposed into basis and coefficient matrices. Using the coupled matrix factorization model, the spectral basis from the factorization of the HSI and the spatial coefficients from the MSI can be learned simultaneously. Generally, reshaping a 3-D HR-HSI to a 2-D matrix using an unfolding operator results in a loss of spatial correlation. Although several strategies, such as non-negative regularization of basis and coefficients [27,34] and total variation regularization of coefficient matrix [29], have been introduced to improve the HSR quality, the spatial-spectral corre-

* Corresponding author
E-mail addresses: wei.he@riken.jp (W. He), chenyong1872008@163.com (Y. Chen), naoto.yokoya@riken.jp (N. Yokoya), chao.li@riken.jp (C. Li), qibin.Zhao@riken.jp (Q. Zhao).

https://doi.org/10.1016/j.patcog.2021.108280
0031-3203/

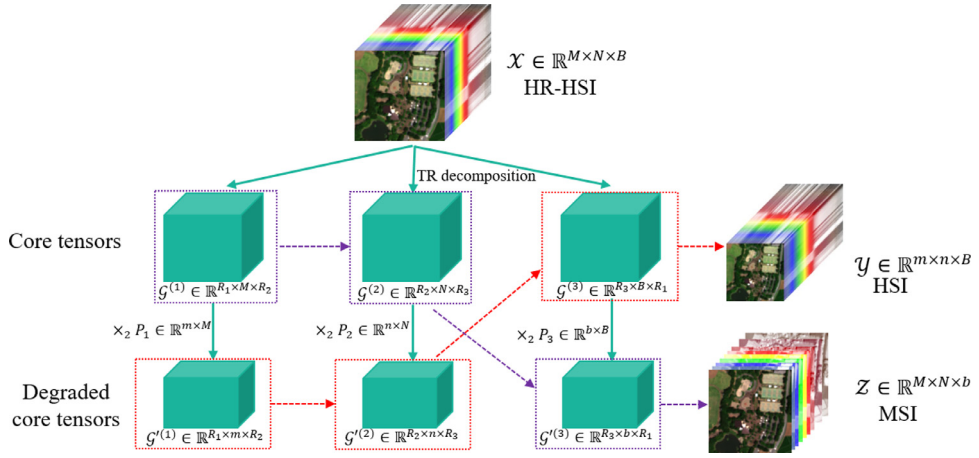


Fig. 1. Tensor ring degradation model from HR-HSI to HSI and MSI, respectively.

lations have not been fully exploited. To extend the coupled matrix factorization model, a coupled tensor factorization model was recently developed for HSR. Well-known methods include the coupled canonical polyadic (CP) factorization model [6] and the coupled Tucker factorization model [35,36]. As presented in [6], the CP factorization model assumes that low-rank properties of different dimensions are the same; however, this is not true in HR-HSI [7,37]. The Tucker decomposition model introduces a 3-D core tensor, which is independent of the entire spatial and spectral degradation process. The existence of this core tensor increases the computation and estimation complexities. Several strategies, including the L_1 norm [35] and L_2 norm [36] regularizations of the core tensor, and non-local processing [8,38,39], have been proposed to improve accuracy. However, this improvement is limited. In summary, developing efficient and effective coupled tensor factorization models for HSR remains a challenge. In this study, we propose a coupled tensor ring factorization (CTRF) model for HSR. Tensor ring (TR) factorization [40] attempts to decompose a high-order tensor into a series of 3-D tensors, called TR cores. As illustrated in [40,41], TR factorization inherits the simple representation of matrix/CP factorization and the flexible low-rank exploration of Tucker factorization. Furthermore, the TR factorization model can exploit the low-rank properties of different classes (discussed further in Section 3.3), which has never been explored in previous coupled tensor factorization models. The core idea of our CTRF model is as follows. First, based on TR factorization, we developed a degradation model from the HR-HSI to the HSI and MSI, as shown in Fig. 1. The HR-HSI is degraded to an HSI by downsampling the spatial core tensors. Meanwhile, the HR-HSI is degraded to the MSI by downsampling the spectral core tensor. Our CTRF model reconstructs the HR-HSI from the coupled input of the HSI and MSI. Subsequently, we build the objective function of the CTRF model and iteratively update the two high-spatial-resolution core tensors and a high-spectral-resolution core tensor from the input HSI and MSI. Finally, the HR-HSI is reconstructed from the optimized high-resolution core tensors via TR. An illustration of this is shown in Fig. 2. The TR factorization model ignores the global low-rank property of the original HR-HSI along the spectral dimension, which has proven to be important in HSR [7,42]. Inspired by [43], we also discovered that the spectral low-rank property of the HR-HSI is bounded by the rank of the third TR core along mode-2 unfolding (Theorem 2). This motivated us to introduce a nuclear norm regularized CTRF (NCTRF) model. The main contributions of this study are summarized as follows:

- Based on TR factorization, we developed a degradation model from the HR-HSI to the MSI and HSI. We proposed a CTRF model for HSR tasks. The nuclear norm regularization of the third TR core with mode-2 unfolding was introduced to further exploit the global spectral low-rank property of the HR-HSI.
- We analyzed the superiority of the CTRF model for HSR and developed an efficient alternating iteration method for the proposed model. The experiments demonstrated the advantage of the CTRF model compared to the previous matrix/tensor and deep learning methods.

Notations.

We mainly adopt the notations from [44] in this paper. Tensors of order $N \geq 3$ are denoted by boldface Euler script letters, e.g., $\mathcal{H} \in \mathbb{R}^{I_1\times I_2\times\cdots\times I_N}$. Scalars are denoted by normal lowercase letters or uppercase letters, e.g., $h, H \in \mathbb{R}$. $H(i_1,\cdots,i_N)$ denotes the element of tensor $\mathcal{H}$ in position $(i_1,\cdots,i_N)$. Vectors are denoted by boldface lowercase letters, e.g., $\mathbf{h} \in \mathbb{R}^I$. Matrices are denoted by boldface capital letters, e.g., $\mathbf{H} \in \mathbb{R}^{I\times J}$. Moreover, we employ two types of tensor unfolding (matricization) operations in this paper. The first mode-n unfolding [44] of tensor $\mathcal{H} \in \mathbb{R}^{I_1\times I_2\times\cdots\times I_N}$ is denoted by $\mathbf{H}_{(n)} \in \mathbb{R}^{I_n\times I_1\cdots I_{n-1}I_{n+1}\cdots I_N}$. The second mode-$n$ unfolding of tensor $\mathcal{H}$ which is often used in TR operations [40] is denoted by $\mathbf{H}_{<n>} \in \mathbb{R}^{I_n\times I_{n+1}\cdots I_N I_1\cdots I_{n-1}}$. $\mathbf{H}[I_1\cdots I_i, I_{i+1}\cdots I_N]$ is the matricization by regarding the first i dimensions as row and the last $N-i$ dimensions as column.

We define the folding operation for the first mode-n unfolding as $\text{fold}_n(\cdot)$, i.e., for a tensor $\mathcal{H}$, we have $\text{fold}_n(\mathbf{H}_{(n)}) = \mathcal{H}$. In addition, the inner product of two tensors $\mathcal{H}$, $\mathcal{W}$ with the same size $\mathbb{R}^{I_1\times I_2\times\cdots\times I_N}$ is defined as $\langle\mathcal{H},\mathcal{W}\rangle = \sum_{i_1}\sum_{i_2}\cdots\sum_{i_N} h_{i_1 i_2\ldots i_N} w_{i_1 i_2\ldots i_N}$. Furthermore, the Frobenius norm of $\mathcal{H}$ is defined by $\|\mathcal{H}\|_F = \sqrt{\langle\mathcal{H},\mathcal{H}\rangle}$. In this paper, we adopt $\mathcal{X} \in \mathbb{R}^{M\times N\times B}$, $\mathcal{Y} \in \mathbb{R}^{m\times n\times B}$, $\mathcal{Z} \in \mathbb{R}^{M\times N\times b}$ to represent the registered HR-HSI, HSI and MSI, respectively. Here, $M > n$ and $N > n$ represent the spatial size and $B > b$ stands for the spectral size. The objective of super-resolution is to estimate the HR-HSI by the fusion of HSI and MSI.

2. Related work

In this section, we briefly discuss the related coupled matrix/tensor decomposition HSR methods, and the development of TR factorization.

2.1. Coupled matrix factorization

Various studies have proposed coupled matrix factorization based methods [10,11,16,17,26,28–32,34]. Specifically, HR-HSI is characterized by strong spectral correlation among different pixels

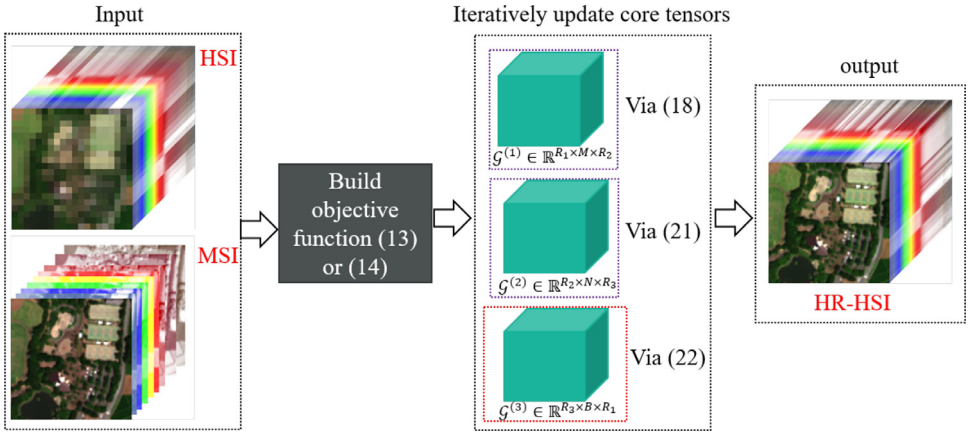


Fig. 2. Flowchart of our proposed CTRF model. The inputs are the MSI and HSI, and the output is the HR-HSI.

and can be expressed by the following matrix factorization model

$$\mathbf{X}_{(3)} = \mathbf{CD}, \tag{1}$$

where $\mathbf{X}_{(3)} \in \mathbb{R}^{B\times MN}$ represents the unfolding HR-HSI tensor $\mathcal{X}$ along the spectral dimension, and $\mathbf{C} \in \mathbb{R}^{B\times L}$ and $\mathbf{D} \in \mathbb{R}^{L\times MN}$ represent the spectral basis and the corresponding coefficient matrix, respectively. In the factorization model (1), the basis matrix can be allocated with different physical meanings with different regularizations, that is, a low dimensional subspace with low-rank regularization [11,16,29], and a spectral endmember matrix with nonnegative regularization [10,32,34]. To explore the relationship between the HR-HSI, HSI and MSI, the researchers usually assume that there exist two linear degradation matrices $\mathbf{P}_0 \in \mathbb{R}^{mn\times MN}$ and $\mathbf{P}_3 \in \mathbb{R}^{b\times B}$ such that

$$\mathbf{Y}_{(3)} = \mathbf{C}(\mathbf{DP}_0^\top), \quad \mathbf{Z}_{(3)} = (\mathbf{P}_3\mathbf{C})\mathbf{D}, \tag{2}$$

where $\mathbf{P}_0$ is the point spread function (PSF) [7]. The basic idea of the coupled matrix factorization model for the HR-HSI method is to jointly factorize $\mathbf{D}$ and $\mathbf{C}$ from HSI and MSI, and reconstruct the HR-HSI via equation (1).

However, the coupled matrix factorization based methods need to reshape the 2-D spatial image into a 1-D vector, ignoring the spatial correlation of the image. Consequently, the total variation was introduced to enhance the spatial smoothness of the corresponding coefficient matrix $\mathbf{D}$ [29]. Unfortunately, the related total variation regularized methods introduced spatial oversmoothing, and the accuracy of the super-resolution result was limited [6].

2.2. Coupled CP factorization

The CP decomposition [44,45] is used to decompose tensor $\mathcal{X}$ into the sum of multiple rank-1 tensors, which can be expressed as

$$\mathcal{X} = \sum_{f=1}^{F} \mathbf{a}_f \circ \mathbf{b}_f \circ \mathbf{c}_f, \tag{3}$$

where $\circ$ denotes the outer product, and F is the *rank* of CP decomposition. $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_f]$, $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_f]$, $\mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_f]$ are called the low-rank latent factors and CP decomposition can be represented as $\mathcal{X} = \lfloor \mathbf{A}, \mathbf{B}, \mathbf{C} \rfloor$. The coupled CP factorization model [6] assumes that there exist three linear degradation matrices $\mathbf{P}_1 \in \mathbb{R}^{m\times M}$, $\mathbf{P}_2 \in \mathbb{R}^{n\times N}$, and $\mathbf{P}_3 \in \mathbb{R}^{b\times B}$ such that

$$\mathcal{Y} = \lfloor \mathbf{P}_1\mathbf{A}, \mathbf{P}_2\mathbf{B}, \mathbf{C} \rfloor, \mathcal{Z} = \lfloor \mathbf{A}, \mathbf{B}, \mathbf{P}_3\mathbf{C} \rfloor \tag{4}$$

The advantage of the coupled CP factorization model is that it preserves the HR-HSI spatial structure. However, the HR-HSI has a stronger spectral low-rank property compared to the spatial perspective. In the coupled CP model, each latent factor among $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ is treated equally with a much larger *rank* F, resulting in insufficient low-rank spectral exploration.

2.3. Coupled tucker factorization

The Tucker decomposition [44,46] decomposes the tensor $\mathcal{X}$ as

$$\mathcal{X} = \mathcal{O} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C}, \tag{5}$$

where $\mathcal{O} \in \mathbb{R}^{R_1\times R_2\times R_3}$ denotes the core tensor, and $\mathbf{A} \in \mathbb{R}^{M\times R_1}$, $\mathbf{B} \in \mathbb{R}^{N\times R_2}$, and $\mathbf{C} \in \mathbb{R}^{B\times R_3}$ are the factors related to different dimensions. We adopted $\times_k$ to represent the mode-k tensor-matrix product [44].

The coupled Tucker factorization [35,36] also assumes that the degradation matrices $\mathbf{P}_1 \in \mathbb{R}^{m\times M}$, $\mathbf{P}_2 \in \mathbb{R}^{n\times N}$, and $\mathbf{P}_3 \in \mathbb{R}^{b\times B}$ such that

$$\begin{aligned} \mathcal{Y} &= \mathcal{O} \times_1 (\mathbf{P}_1\mathbf{A}) \times_2 (\mathbf{P}_2\mathbf{B}) \times_3 \mathbf{C}, \\ \mathcal{Z} &= \mathcal{O} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 (\mathbf{P}_3\mathbf{C}). \end{aligned} \tag{6}$$

In the coupled Tucker factorization model, the factor size can be chosen based on the low-rank property of each dimension. However, it introduces a core tensor $\mathcal{O}$ which is independent of the whole spatial and spectral degradation matrices. L_1 norm [35] and L_2 norm [36] regularizers have been introduced to facilitate the identification of $\mathcal{O}$; unfortunately, the related regularizations introduce additional turning parameters, and the estimation accuracy of core tensor from the HSI and MSI remain problematic [35].

2.4. Review of tensor ring factorization

Based on tensor train factorization [47], TR factorization was first proposed by [40] and subsequently well studied from the algorithm, application and theoretical analysis perspectives. The authors of [40,48] introduced sequential singular value decomposition, alternating least squares (ALS) with adaptive rank, and blockwise ALS methods for the optimization of TR factorization. Owing to its well-balanced structure compared to tensor train factorization, TR factorization has also been directly applied to image completion [41]. To further improve image completion performance, TR with nuclear norm regularization [43] and total variation [49] have been developed. In [50], provable TR completion method was introduced and theoretically verified. In addition to image completion, TR factorization has also been developed for applications such as hyperspectral image denoising [51], neural network compression [52], generative model [53] and so on. More recently, a similar method of low rank coupled TR completion was proposed in

[54]. A comparison between our work and that of [54] highlights a number of differences as follows. First, the methods are designed for different tasks. Our proposed method is used for HSI fusion, while the method in [54] is used for matrix completion. Second, the conditions of the two methods were different. The authors of [54] assumed that the two coupled tensors shared partial core tensors; however, in our model, all three core tensors of the coupled HSI and MSI are related via spatial and spectral downsampling operators.

3. Coupled tensor ring factorization

In this section, we propose a CTRF model for the HSR. The proposed CTRF model has a more flexible *rank* selection strategy than CP decomposition, and a simpler representation compared to Tucker decomposition.

3.1. Tensor ring decomposition

TR decomposition represents tensor $\mathcal{H}$ by circular multilinear products over a sequence of third-order core tensors $\mathcal{G} := \{\mathcal{G}^{(1)}, \dots, \mathcal{G}^{(N)}\}$, where $\mathcal{G}^{(n)} \in \mathbb{R}^{R_n \times I_n \times R_{n+1}}$, $n = 1, 2, \cdots, N$, $R_1 = R_{N+1}$ [40]. Here, $R = [R_1, \cdots, R_N]$ is the TR rank. Each element of tensor $\mathcal{H}$ can be rewritten as

$$H(i_1, \cdots, i_N) = \mathrm{Tr}(\mathbf{G}^{(1)}(i_1) \cdots \mathbf{G}^{(N)}(i_N)), \tag{7}$$

where $\mathbf{G}^{(n)}(i_n)$ is the i_n-th lateral slice matrix of the core tensor $\mathcal{G}^{(n)}$. In this study, we adopt the notation $\mathcal{H} = \Phi(\mathcal{G})$ to represent TR decomposition. Next, we introduce two important properties of TR decomposition.

Definition 1. (Multilinear product [48].) Suppose $\mathcal{G}^{(n)}$ and $\mathcal{G}^{(n+1)}$ are the two nearby cores of TR decomposition. The multilinear product of the two cores is $\mathcal{G}^{(n,n+1)} \in \mathbb{R}^{R_n \times I_n I_{n+1} \times R_{n+2}}$ and denoted by

$$\mathcal{G}^{(n,n+1)}((j_l - 1)I_n + i_k) = \mathcal{G}^{(n)}(i_k)\mathcal{G}^{(n+1)}(j_l), \tag{8}$$

for $i_k = 1, \cdots, I_n$, $j_l = 1, \cdots, I_{n+1}$.

From Definition 1, if the tensor $\mathcal{H}$ can be decomposed via (7), we can immediately obtain the following

$$\mathbf{H}[I_1 \cdots I_n, I_{n+1} \cdots I_N] = \mathbf{G}_{(2)}^{(1,\cdots,n)} \times (\mathbf{G}_{<2>}^{(n+1,\cdots,N)})^{\top}, \tag{9}$$

where $\mathbf{G}_{(2)}^{(1,\cdots,n)}$ denotes the multilinear products of the first n cores and the unfolding along mode-2, and $\mathbf{G}_{<2>}^{(n+1,\cdots,N)}$ represents the multilinear products of the last $N-n$ cores and the second unfolding along mode-2. That is to say, the rank value $R_1 \times R_{n+1}$ is bounded by the rank of unfolding matrix $\mathbf{H}[I_1 \cdots I_n, I_{n+1} \cdots I_N]$.

Proposition 1. *(Circular dimensional permutation invariance [40].) If $\overleftarrow{\mathcal{H}_n} \in \mathbb{R}^{I_{n+1} \times \cdots \times I_N \times I_1 \times \cdots \times I_n}$ is denoted as the circularly shifting the dimensions of $\mathcal{H}$ by n, then we have $\overleftarrow{\mathcal{H}_n} = \Phi(\{\mathcal{G}^{(n+1)}, \dots, \mathcal{G}^{(N)}, \mathcal{G}^{(1)}, \dots, \mathcal{G}^{(n)}\})$*

With Proposition 1, we can easily shift the middle cores of the TR to the first position, and utilize Definition 1 to analyze each core separately.

3.2. Coupled tensor ring factorization

Based on TR factorization, we first developed the degradation model from the HR-HSI to the HSI and MSI, as shown in Fig. 1. Analogous to the coupled CP and Tucker factorization, we also assume that the spatial degradation matrices $\mathbf{P}_1 \in \mathbb{R}^{m \times M}$, $\mathbf{P}_2 \in \mathbb{R}^{n \times N}$, and spectral degradation matrix $\mathbf{P}_3 \in \mathbb{R}^{b \times B}$. $\mathbf{P}_1$ and $\mathbf{P}_2$ can be regarded as the separable operators of $\mathbf{P}_0$ in the coupled matrix factorization model, as $\mathbf{P}_0 = \mathbf{P}_1 \otimes \mathbf{P}_2$, with $\otimes$ representing the Kronecker product. As discussed in [55], separable operators have more advantages in terms of optimization and calculation. Thus, the TR decomposition of HR-HSI can be represented as

$$\mathcal{X} = \Phi(\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \mathcal{G}^{(3)}), \tag{10}$$

where $\mathcal{G}^{(1)} \in \mathbb{R}^{R_1 \times M \times R_2}$, $\mathcal{G}_{(2)} \in \mathbb{R}^{R_2 \times N \times R_3}$ and $\mathcal{G}^{(3)} \in \mathbb{R}^{R_3 \times B \times R_1}$ represents the core tensors related to the spatial, spatial and spectral dimensions, respectively. $R = [R_1, R_2, R_3]$ represents the TR *rank* of the HR-HSI. The HSI can be expressed as

$$\mathcal{Y} = \Phi(\mathcal{G}^{(1)} \times_2 \mathbf{P}_1, \mathcal{G}^{(2)} \times_2 \mathbf{P}_2, \mathcal{G}^{(3)}), \tag{11}$$

and the MSI can be formulated as

$$\mathcal{Z} = \Phi(\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \mathcal{G}^{(3)} \times_2 \mathbf{P}_3). \tag{12}$$

By combining (11) and (12), we can obtain the CTRF model as

$$\min_{\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \mathcal{G}^{(3)}} \left\| \mathcal{Y} - \Phi(\mathcal{G}^{(1)} \times_2 \mathbf{P}_1, \mathcal{G}^{(2)} \times_2 \mathbf{P}_2, \mathcal{G}^{(3)}) \right\|_F^2 + \left\| \mathcal{Z} - \Phi(\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \mathcal{G}^{(3)} \times_2 \mathbf{P}_3) \right\|_F^2. \tag{13}$$

Our objective with the CTRF model is to reconstruct the HR-HSI from the input HSI and MSI.

3.3. Motivation of CTRF for HSR

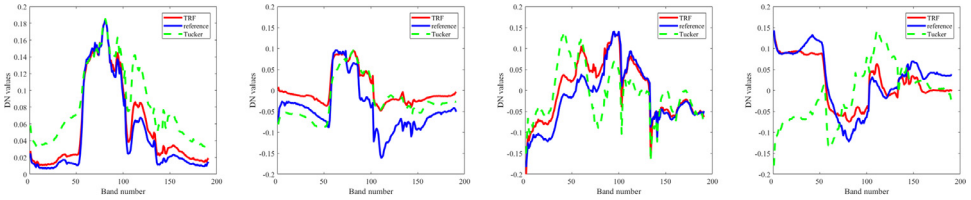
In this section, we investigate (13), and provide insights into why the proposed CTRF model improves on the coupled Tucker decomposition model in HSR. For HSIs, each spectral pixel, denoted as $\mathcal{X}(i, j, :)$ of size $\mathbb{R}^B$, represents the spectrum of one specific class. Typically, one scene of a HSI contains multiple classes, and there are several basis spectral signatures available to construct a low-dimensional subspace for each class [56]. Tucker decomposition usually reshapes HSIs along the spectral dimension, and exploits the spectral low-rank property globally. Conversely, TR factorization can learn the spectral core $\mathcal{G}^{(3)} \in \mathbb{R}^{R_3 \times B \times R_1}$ with R_3 representing the number of classes, and R_1 representing the subspace dimension of each class.

An experiment was constructed to demonstrate the advantages of the CTRF. We selected 50 pixels of class1 (grass) and 50 pixels of class2 (road) from the WDC dataset (Section 4.1), and constructed these 100 pixels to a HR-HSI of size $10 \times 10 \times 191$. TR factorization could obtain a spectral core of size $2 \times 191 \times 2$ from the constructed HR-HSI. Fig. 3 shows the spectral signatures extracted by TR and Tucker decomposition. We applied singular value decomposition (SVD) to each class and extracted the first two principal component vectors as reference signatures. As observed in Fig. 3, the signatures obtained by TR are more similar to those of the references, indicating the advantage of TR factorization for HSI processing.

3.4. NCTRF

As analyzed in Section 3.3, the obtained spectral core $\mathcal{G}^{(3)}$ regards the subspace dimension of each class to be the same. However, in a real case, the subspace dimensions of different classes may differ. If we set a larger subspace dimension R_1, we lose the global low-rank property of the spectral core tensor. However, as reviewed in [7,10,30,57,58], the HSI has a strong spectral correlation, indicating the low-rank property of HSI along the spectral dimension. Inspired by this fact, we propose to regularize the third core tensor $\mathcal{G}^{(3)}$ along mode-2 as low rank, denoted as $rank(\mathcal{G}_{(2)}^{(3)})$. The rank constraint is hard to optimize, and we introduced the nuclear norm $\|\cdot\|_*$, the sum of the singular values of the matrix [59], to regularize the low-rank property. Therefore, the proposed CTRF model with nuclear norm regularization (NCTRF) is formulated as

$$\begin{aligned} \min_{\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \mathcal{G}^{(3)}} & \|\mathcal{Y} - \Phi\left(\mathcal{G}^{(1)} \times_2 \mathbf{P}_1, \mathcal{G}^{(2)} \times_2 \mathbf{P}_2, \mathcal{G}^{(3)}\right)\|_F^2 \\ & + \|\mathcal{Z} - \Phi\left(\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \mathcal{G}^{(3)} \times_2 \mathbf{P}_3\right)\|_F^2 + \lambda \|\mathcal{G}_{(2)}^{(3)}\|_*. \end{aligned} \tag{14}$$



(a) class1-sigature1 (b) class1-sigature2 (c) class2-sigature1 (d) class2-sigature2

Fig. 3. The signature comparison between TR factorization (TRF), Tucker decomposition and the reference. The reference indicates the signature extracted from each class via SVD. Class1-sigature1 indicates the first signature from class 1.

Theorem 2. *Suppose an N-th order tensor* $\mathcal{H} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$*, we have the following property*

$$rank\left(\mathcal{G}^{(n)}_{(2)}\right) \geq rank\left(\mathcal{H}_{(n)}\right), \tag{15}$$

for $n = 1, \cdots N$.

The proof Theorem 2 of can be referred to [43]. By Theorem 2, we can immediately conclude that the spectral rank of HR-HSI $\mathcal{X}$ can be bounded by the rank of core tensor $\mathcal{G}^{(3)}$ along mode-2. The regularization $\|\mathcal{G}^{(3)}_{(2)}\|_*$ can exploit the global low-rank property along spectral dimension.

Remark 3.1. In this study, we only regularize the spectral low-rank property, and ignore the spatial low-rank property [43], because the low-rank property along the spectral dimension is much stronger than that of the spatial dimension [37,42]. The NCTRF model (14) introduces an additional parameter λ. To reduce the complexity of the proposed model, we fixed λ to a constant value in all experiments.

3.5. Optimization

The objective function (14) is non-convex. Fortunately, for each separable variable, the objective function is convex. We first introduce a latent variable $\mathcal{G}^{(0)} = \mathcal{G}^{(3)}$, and convert (14) to the following augmented Lagrangian function

$$\begin{aligned} g(\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \mathcal{G}^{(3)}, \mathcal{G}^{(0)}, \mathcal{L}, \mu) &= \left\|\mathcal{Y} - \mathbf{\Phi}(\mathcal{G}^{(1)} \times_2 \mathbf{P}_1, \mathcal{G}^{(2)} \times_2 \mathbf{P}_2, \mathcal{G}^{(3)})\right\|_F^2 \\ &+ \left\|\mathcal{Z} - \mathbf{\Phi}(\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \mathcal{G}^{(3)} \times_2 \mathbf{P}_3)\right\|_F^2 + \lambda\left\|\mathcal{G}^{(0)}_{(2)}\right\|_* \\ &+ < \mathcal{L}, \mathcal{G}^{(0)} - \mathcal{G}^{(3)} > + \frac{\mu}{2}\left\|\mathcal{G}^{(0)} - \mathcal{G}^{(3)}\right\|_F^2, \end{aligned} \tag{16}$$

where $\mathcal{L}$ represents the Lagrangian multiplier and μ stands for the penalty parameter. Next, we adopt an alternating iterative method [60] to optimize (16). By fixing other variables, and updating one variable for each iteration, the optimization of (16) can be split into four subproblems.

Update $\mathcal{G}^{(1)}$: By fixing other variables, the update of $\mathcal{G}^{(1)}$ can be formulated as

$$\mathcal{G}^{(1)} = \min_{\mathcal{G}^{(1)}} g(\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \mathcal{G}^{(3)}, \mathcal{G}^{(0)}, \mathcal{L}, \mu). \tag{17}$$

By using two kinds of tensor unfolding as in Section 1 and Definition 1, we can convert the optimization of (17) to the following problem

$$\mathcal{G}^{(1)} = \arg\min_{\mathcal{G}^{(1)}} \left\|\mathbf{Y}_{<1>} - \mathbf{P}_1\mathbf{G}^{(1)}_{(2)}\mathbf{A}_1\right\|_F^2 + \left\|\mathbf{Z}_{<1>} - \mathbf{G}^{(1)}_{(2)}\mathbf{B}_1\right\|_F^2, \tag{18}$$

where $\mathbf{Y}_{<1>}$ and $\mathbf{Z}_{<1>}$ stand for the second unfolding of the tensors $\mathcal{Y}, \mathcal{Z}$, respectively, $\mathbf{A}_1 = ((\mathcal{G}^{(2)} \times_2 \mathbf{P}_2) \cdot \mathcal{G}^{(3)})^\top_{<2>}$ and $\mathbf{B}_1 = (\mathcal{G}^{(2)} \cdot (\mathcal{G}^{(3)} \times_2 \mathbf{P}_3))^\top_{<2>}$. Optimization (18) is quadratic and its unique solution is equal to solve the general Sylvester matrix equation [16]

$$\mathbf{P}_1^\top\mathbf{P}_1\mathbf{G}^{(1)}_{(2)}\mathbf{A}_1\mathbf{A}_1^\top + \mathbf{G}^{(1)}_{(2)}\mathbf{B}_1\mathbf{B}_1^\top = \mathbf{P}_1^\top\mathbf{Y}_{<1>}\mathbf{A}_1^\top + \mathbf{Z}_{<1>}\mathbf{B}_1^\top. \tag{19}$$

To avoid the large scale matrix inversion in the closed-form solution of (19), we adopt conjugate gradient (CG) [35] to solve (19).

Update $\mathcal{G}^{(2)}$ and $\mathcal{G}^{(3)}$: We firstly fix other variables and update $\mathcal{G}^{(2)}$. We adopt Proposition 1 to circularly shift the tensors $\mathcal{Y}$ and $\mathcal{Z}$, and convert the optimization of $\mathcal{G}^{(2)}$ to the following optimization:

$$\begin{aligned} &\min_{\mathcal{G}^{(2)}} \left\|\overleftarrow{\mathcal{Y}_1} - \mathbf{\Phi}(\mathcal{G}^{(2)} \times_2 \mathbf{P}_2, \mathcal{G}^{(3)}, \mathcal{G}^{(1)} \times_2 \mathbf{P}_1)\right\|_F^2 \\ &+ \left\|\overleftarrow{\mathcal{Z}_1} - \mathbf{\Phi}(\mathcal{G}^{(2)}, \mathcal{G}^{(3)} \times_2 \mathbf{P}_3, \mathcal{G}^{(1)})\right\|_F^2. \end{aligned} \tag{20}$$

(20) can be converted to the following matrix version optimization problem

$$\mathcal{G}^{(2)} = \arg\min_{\mathcal{G}^{(2)}} \left\|\mathbf{Y}_{<2>} - \mathbf{P}_2\mathbf{G}^{(2)}_{(2)}\mathbf{A}_2\right\|_F^2 + \left\|\mathbf{Z}_{<2>} - \mathbf{G}^{(2)}_{(2)}\mathbf{B}_2\right\|_F^2, \tag{21}$$

where $\mathbf{Y}_{<2>}$ and $\mathbf{Z}_{<2>}$ are equal to the mode-1 unfolding of the tensors $\overleftarrow{\mathcal{Y}_1}, \overleftarrow{\mathcal{Z}_1}$, respectively, $\mathbf{A}_2 = (\mathcal{G}^{(3)} \cdot (\mathcal{G}^{(1)} \times_2 \mathbf{P}_1))^\top_{<2>}$ and $\mathbf{B}_2 = ((\mathcal{G}^{(3)} \times_2 \mathbf{P}_3) \cdot \mathcal{G}^{(1)})^\top_{<2>}$. Thus, it can also be efficiently solved by the CG method. To update $\mathcal{G}^{(3)}$, we also adopt Proposition 1 to circularly shift the tensors $\mathcal{Y}$ and $\mathcal{Z}$, and obtain the following optimization

$$\begin{aligned} &\min_{\mathcal{G}^{(3)}} \left\|\overleftarrow{\mathcal{Y}_2} - \mathbf{\Phi}(\mathcal{G}^{(3)}, \mathcal{G}^{(1)} \times_2 \mathbf{P}_1, \mathcal{G}^{(2)} \times_2 \mathbf{P}_2)\right\|_F^2 \\ &+ \left\|\overleftarrow{\mathcal{Z}_2} - \mathbf{\Phi}(\mathcal{G}^{(3)} \times_2 \mathbf{P}_3, \mathcal{G}^{(1)}, \mathcal{G}^{(2)})\right\|_F^2 \\ &+ < \mathcal{L}, \mathcal{G}^{(0)} - \mathcal{G}^{(3)} > + \frac{\mu}{2}\left\|\mathcal{G}^{(0)} - \mathcal{G}^{(3)}\right\|_F^2, \end{aligned} \tag{22}$$

which can be solved by the CG method.

Update $\mathcal{G}^{(0)}$: By fixing other variables, the optimization of $\mathcal{G}^{(0)}$ can be obtained by solving the following problem

$$\min_{\mathcal{G}^{(0)}} \lambda\left\|\mathbf{G}^{(0)}_{(2)}\right\|_* + \frac{\mu}{2}\left\|\mathbf{G}^{(0)}_{(2)} - \mathbf{G}^{(3)}_{(2)} + \mathbf{L}_{(2)}/\mu\right\|_F^2, \tag{23}$$

which can be solved by the closed-form solution [59]

$$\mathcal{G}^{(0)} = \text{fold}_2(\mathcal{T}_{\frac{\lambda}{\mu}}(\mathbf{G}^{(3)}_{(2)} - \mathbf{L}_{(2)}/\mu)). \tag{24}$$

Here, $\mathcal{T}_{\frac{\lambda}{\mu}}$ is the singular value thresholding (SVT) operator.

Update $\mathcal{L}$ and μ: Finally, we adopt the strategy in [60] to update $\mathcal{L}$ and μ

$$\mathcal{L} = \mathcal{L} + \mu(\mathcal{G}^{(0)} - \mathcal{G}^{(3)}), \quad \mu = min(\mu_1, \rho\mu), \tag{25}$$

where μ_1 and $\rho > 1$ denotes constant values. In summary, the optimization of the proposed NCTRF model is presented in Algorithm 1.

4. Experiments

In this section, we present the experimental results of different methods, followed by the parameter analysis, convergence analysis, computational time, and comparison with deep learning based methods. The experiments were programmed in MATLAB R2018b on a laptop with a Core i7-8750H CPU and 32GB memory.

Algorithm 1 Optimization of NCTRF.

Require: HSI $\mathcal{Y}$, MSI $\mathcal{Z}$, rank R.
1: $\lambda = 0.001, \rho = 1.5, \mu = 10^{-4}, \mu_1 = 10^6$
2: **for** $i = 1, 2, 3, \cdots iter$ **do**
3: Update the TR cores $\mathcal{G}$ using (17), (20), (22), and (23).
4: Update $\mathcal{L}$ and μ via (25).
5: **end for**
6: **return** HR-HSI $\mathcal{X} = \phi(\mathcal{G})$;

Table 1
The size of the image used for HSR experiments.

Image name	HR-HSI	HSI	MSI
WDC	256 × 256×90	64 × 64×90	256 × 256×4
PaC	200 × 200×93	50 × 50×93	200 × 200×4
Indian	144 × 144×100	16 × 16×100	144 × 144×4
CAVE	512 × 512×31	16 × 16×31	512 × 512×3
Real	512 × 512×71	32 × 32×71	512 × 512×3

4.1. Experimental database

To validate the performance of our proposed CTRF method for remote sensing HSR, we conducted the experiments on three HR-HSI datasets, including the Washington DC Mall (WDC)[1], Pavia Center (PaC)[2], and Indian Pines[3]. These three datasets are widely used in simulation-based evaluation of HSR [6,35]. The sizes of the HR-HSI, and simulated HSIs and MSIs are shown in Table 1. We chose the PSF with an average kernel (size 8) to generate the spatial degradation matrices. When we simulated the HSI and MSI from the HR-HSI, Gaussian noise was added to the HSI and MSI, and the signal-to-noise (SNR) ratio changed from 20dB, 30dB to 40dB. All HR-HSIs were scaled to [0, 255].

4.2. Comparison methods and evaluation measures

We chose the following methods for comparison: coupled non-negative matrix factorization (CNMF) [34][4], FUSE [61][5], HySure [29][6], coupled CP factorization (STEREO) [6][7], coupled Tucker factorization (CSTF) [35][8], and non-local sparse tensor factorization (NLSTF) [8]. Related implementation codes were downloaded from the authors' personal website, and the parameters were manually tuned to the best. Our proposed methods are denoted as CTRF and NCTRF, respectively. To evaluate the performance of the proposed method, five quantitative indices were utilized in our study: the peak signal-to-noise ratio (PSNR), root mean square error (RMSE), relative dimensional global error in synthesis (ERGAS) [62], spectral angle mapper (SAM) [63], and structure similarity (SSIM) [64]. Smaller RMSE, ERGAS, and SAM values indicate better super-resolution results. Conversely, the larger PSNR and SSIM values illustrate better quality.

4.3. Experimental results on remote sensing

Quantitative comparison. For each noise level, we calculated the evaluation values of the three datasets and then averaged them, as shown in Table 2. From the table, it can be observed that the proposed CTRF and NCTRF methods displayed more advantages

Table 2
Quantitative comparison of different algorithms under various noise levels. The results are the average of the three datasets, and the best results are in bold. ↑ stands for the larger, the better, and ↓ is the inverse.

SNR = 20					
Method	PSNR ↑	RMSE↓	ERGAS↓	SAM↓	SSIM↑
CNMF	26.73	12.86	4.38	8.53	0.752
FUSE	30.43	8.23	2.94	5.79	0.870
HySure	27.56	11.72	4.51	9.64	0.727
STEREO	30.31	8.20	3.06	7.35	0.846
CSTF	31.48	7.16	2.76	6.17	0.865
NLSTF	25.47	14.36	5.42	12.99	0.688
CTRF	31.94	6.89	2.56	5.76	0.877
NCTRF	**31.98**	**6.84**	**2.55**	**5.74**	**0.877**
SNR = 30					
CNMF	29.15	9.97	3.56	4.85	0.885
FUSE	35.63	4.42	1.67	3.30	0.956
HySure	34.98	4.74	1.84	3.69	0.945
STEREO	35.53	4.68	1.85	4.46	0.942
CSTF	37.93	3.47	1.32	3.21	0.963
NLSTF	34.84	5.03	1.82	4.62	0.940
CTRF	38.15	**3.46**	**1.31**	**3.27**	0.965
NCTRF	**38.16**	**3.46**	**1.31**	**3.27**	**0.966**
SNR = 40					
CNMF	29.31	9.66	3.60	4.95	0.905
FUSE	37.61	3.50	1.40	2.48	0.980
HySure	37.41	3.59	1.43	2.56	0.979
STEREO	38.72	3.44	1.25	3.32	0.968
CSTF	40.69	2.58	1.02	2.45	0.981
NLSTF	**41.99**	**2.34**	**0.85**	**2.22**	**0.985**
CTRF	41.35	2.57	1.00	2.47	**0.985**
NCTRF	41.54	2.54	0.98	2.46	**0.985**

when the SNR was 20dB and 30dB. In the low-noise case, NLSTF achieved the best accuracy. This is because non-local based methods are more effective than global based methods. However, an increase in noise leads to several challenges to the group matching model, resulting in a poorer performance for NLSTF in noisy cases. From the comparison, the CNMF and HySure models performed the worst, demonstrating the advantage of the coupled tensor model. Furthermore, CTRF and NCTRF achieved better results than those of the STEREO and CSTF, demonstrating the advantage of TR decomposition compared to that of Tucker and CP. Finally, the results of NCTRF were slightly better than those of CTRF. This phenomenon demonstrates the advantage of nuclear norm regularization for the third core tensor. As the noise level increases, the gap between CTRF and NCTRF decreases, because we usually choose a smaller TR *rank* $R = [R_1, R_2, R_3]$ in the noisy model. In this case, the efficiency of nuclear norm regularization also decreases, since the value $R_1 \times R_3$ is small enough to explore the global spectral low-rank property of the HR-HSI.

Visual comparison. To further compare the differences between different HSR methods, we chose one band from the WDC dataset and one band from the Indian Pines dataset, to illustrate the related HSI, ground truth HR-HSI, and different HSR results in the case of a noise level of $SNR = 20$ in Fig. 4. We also show the related difference images between the HSR results and the ground truth. From the figure, it can be seen that the proposed CTRF and NCTRF methods achieve the best visual results. The CSTF can also achieve outstanding results. The STEREO and FUSE methods perform better in the case of the WDC dataset, but worse with the Indian Pines dataset. The CNMF, HySure, and NLSTF methods fail to reconstruct the images.

4.4. Comparison with deep learning

Recently, deep learning based methods have also been introduced for the fusion of the HSI and MSI [18–21,23]. We used the

[1] https://engineering.purdue.edu/~biehl/MultiSpec/hyperspectral
[2] http://www.ehu.eus/ccwintco/index.php/
[3] https://engineering.purdue.edu/~biehl/MultiSpec/
[4] http://naotoyokoya.com/Download.html
[5] https://github.com/qw245/BlindFuse
[6] https://github.com/alfaiate
[7] https://sites.google.com/site/harikanats/
[8] https://sites.google.com/view/renweidian/

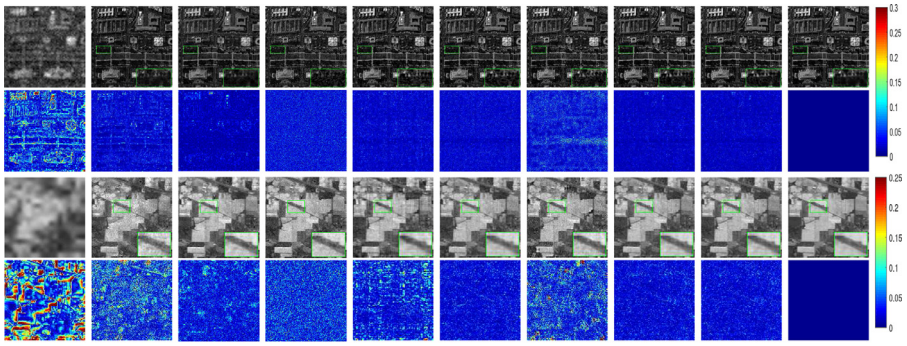


Fig. 4. HSR results of different methods with the WDC and Indian dataset. The noise level corresponds to $SNR = 20$. The first row illustrates the band 13 images from the WDC dataset; the second row shows the related difference images between the HSR results and the ground truth HR-HSI. The third row illustrates the band 49 images from the Indian dataset, and the last row shows the related difference images.

Table 3
Quantitative comparison of different algorithms on CAVE images.

Indices	Method	1	2	3	4	5	6	7	8	9	10	11	12	Averange
PSNR	CNMF	41.71	37.28	29.13	31.17	38.40	37.85	39.60	45.35	40.46	36.86	35.81	37.70	37.61
	FUSE	40.84	36.63	29.84	31.51	38.08	36.06	40.29	40.89	38.91	36.98	34.93	39.24	37.02
	HySure	40.09	33.33	30.41	29.71	37.78	38.25	40.65	38.19	41.09	37.14	36.48	36.75	36.66
	STEREO	37.18	31.41	28.12	28.87	39.92	34.06	40.18	37.16	39.11	31.30	36.53	38.83	35.22
	CSTF	43.86	33.20	35.20	32.19	41.99	40.23	44.30	45.04	45.64	37.57	42.20	44.52	40.50
	NLSTF	43.75	41.14	34.63	33.36	39.37	38.33	43.83	41.90	44.13	37.54	41.41	44.03	40.29
	uSDN	36.33	33.29	29.07	29.14	35.09	36.23	44.29	38.12	36.44	37.06	34.86	39.42	35.78
	UDL	44.69	**41.86**	36.02	**33.99**	41.94	40.24	43.86	46.51	44.68	38.13	**44.24**	**46.21**	41.86
	MHF-net	**46.23**	38.15	**38.08**	33.29	**45.19**	41.16	**46.58**	**46.87**	**47.91**	**43.84**	44.00	44.60	**42.99**
	CTRF	44.40	38.93	37.02	32.00	43.84	**41.70**	45.57	45.46	46.12	40.37	43.50	44.46	41.95
	NCTRF	44.42	38.56	37.11	32.10	44.01	41.64	45.89	45.61	45.96	40.48	43.78	44.68	42.02
RMSE	CNMF	2.95	3.88	9.86	7.75	3.47	3.96	2.94	1.55	2.65	4.36	4.29	3.38	4.25
	FUSE	3.12	4.50	10.80	7.48	3.63	4.94	3.31	2.48	4.08	4.38	5.49	3.01	4.77
	HySure	4.44	7.20	9.39	9.88	3.79	4.38	3.33	3.88	3.72	4.69	4.37	4.13	5.27
	STEREO	6.09	8.58	11.32	11.14	2.92	5.71	3.59	4.37	3.86	7.36	4.07	3.29	6.02
	CSTF	3.62	7.51	6.54	6.97	2.38	3.49	2.27	1.79	1.91	5.08	2.60	1.77	3.83
	NLSTF	3.85	3.66	8.44	6.26	3.21	4.58	2.78	2.89	2.58	4.69	3.32	2.29	4.05
	uSDN	5.11	6.00	9.82	9.24	4.55	4.59	1.69	3.35	3.96	3.86	4.78	2.85	4.98
	UDL	1.94	3.58	5.46	**5.45**	2.34	3.48	1.90	1.40	1.93	3.80	1.92	**1.36**	2.88
	MHF-net	**1.80**	**3.43**	**3.71**	6.79	**1.53**	2.85	**1.65**	**1.37**	**1.54**	**2.85**	**1.88**	2.20	**2.63**
	CTRF	2.27	3.58	4.85	6.74	1.92	2.90	1.83	1.65	1.76	3.35	2.14	1.78	2.90
	NCTRF	2.25	3.86	4.83	6.66	1.90	**2.81**	1.70	1.64	1.79	3.27	2.02	1.71	2.87
SSIM	CNMF	0.975	0.969	0.888	0.962	0.987	0.968	0.986	0.984	0.985	0.974	0.957	0.983	0.968
	FUSE	0.944	0.946	0.812	0.958	0.983	0.947	0.983	0.959	0.971	0.965	0.922	0.971	0.947
	HySure	0.941	0.888	0.834	0.917	0.985	0.954	0.979	0.920	0.934	0.964	0.939	0.968	0.935
	STEREO	0.869	0.792	0.729	0.834	0.971	0.894	0.950	0.907	0.955	0.881	0.920	0.955	0.888
	CSTF	0.955	0.889	0.918	0.934	0.987	0.968	0.984	0.981	0.990	0.955	0.970	0.986	0.960
	NLSTF	0.969	0.970	0.946	0.969	0.986	0.976	0.988	0.983	0.992	0.976	0.961	0.989	0.975
	uSDN	0.915	0.906	0.805	0.824	0.977	0.956	0.964	0.887	0.934	0.980	0.907	0.985	0.920
	UDL	0.971	**0.982**	0.943	0.962	0.984	0.972	0.970	0.974	0.972	0.978	0.980	**0.992**	0.973
	MHF-net	**0.987**	0.979	**0.983**	**0.976**	**0.993**	**0.983**	**0.994**	**0.985**	**0.995**	**0.991**	**0.992**	0.991	**0.988**
	CTRF	0.969	0.955	0.949	0.962	0.991	0.981	0.992	0.983	0.993	0.980	0.984	0.991	0.977
	NCTRF	0.969	0.949	0.949	0.961	0.991	0.981	0.992	0.983	0.991	0.980	0.985	0.991	0.977

unsupervised uSDN [18][9], the unsupervised deep learning (UDL) method introduced in [23], and the supervised MHF-net [19][10] as state-of-the-art deep learning methods for comparison with the proposed methods. With the same setting as in [19], we adopted the CAVE dataset[11] for the experiments, using 20 CAVE images for the training and the remaining 12 CAVE images for the test. The spatial downsampling was the same as that used in [19,23], and the spectral downsampling matrix was the spectral response matrix of the Nikon D700 [23]. The sizes of the simulated CAVE images are listed in Table 1. Table 3 shows comparison between the proposed method and two deep learning based methods. From the table, it can be observed that the proposed NCTRF method achieves better quantitative evaluation results than those of uSDN and UDL method. The supervised MHF-net achieved better results than CTRF and NCTRF. However, our proposed method dose not require additional samples for training compared to the MHF-net. Fig. 5 shows the visual results of different HSR methods, with related difference images between the HSR results and the ground truth HR-HSI CAVE images. From the difference images achieved by the different methods, it can be observed that the proposed method achieved results comparable to those of state-of-the-art unsupervised deep learning methods.

[9] https://github.com/aicip/uSDN
[10] https://github.com/XieQi2015/MHF-net
[11] http://www1.cs.columbia.edu/CAVE/databases/

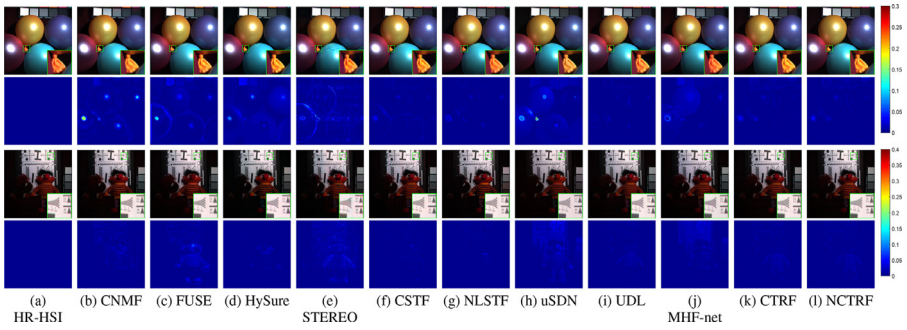


Fig. 5. HSR results of different methods with CAVE dataset at wavelength 460, 540 and 620 nm. The second and forth rows illustrate the difference images between HSR results and the ground-truth HR-HSI.

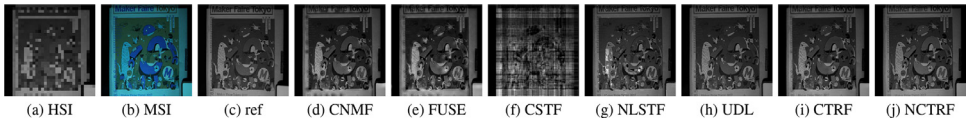


Fig. 6. HSR results of different methods with real images at wavelength 500 nm.

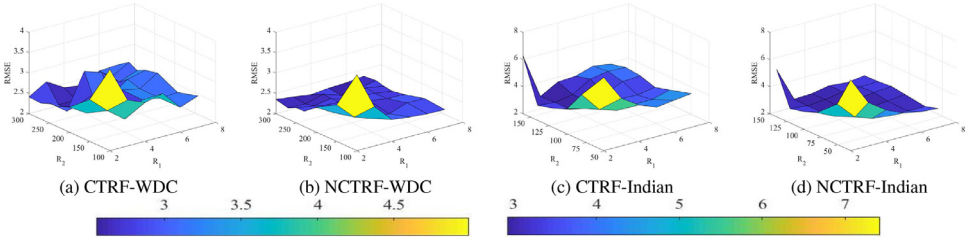


Fig. 7. The changes of RMSE value with different TR *rank* $[R_1, R_2, R_1]$. The noise level is the case of SNR=40.

4.5. Experimental results on real dataset

To further improve the effectiveness of the proposed method, we present HSR results on real images [23]. As illustrated in [23], the HSI was captured using an NH-7 (EBA Japan) hyperspectral camera, and the MSI was captured using a Nikon D5. The sizes of the real HSI and MSI are shown in Table 1. Fig. 6 (a) and (b) show the HSI and MSI images, respectively. Fig. 6 (c) shows a reference image at 500 nm. It can be seen that the UDL method and the proposed CTRF and NCTRF achieve comparable results, indicating the efficiency of the proposed methods to the real applications.

4.6. Discussion

Parameter analysis. TR *rank* $R = [R_1, R_2, R_3]$ is the most important parameter in the proposed CTRF and NCTRF methods. The adaptive selection ofTR *rank* is still a key problem. From Theorem 2, the values $R_1 \times R_2, R_2 \times R_3$, and $R_3 \times R_1$ are bounded by $rank(\mathbf{X}_{<1>}), rank(\mathbf{X}_{<2>})$ and $rank(\mathbf{X}_{<3>})$, respectively. To simplify the complexity of the parameter analysis, we chose $R_1 = R_3$ to maintain the consistency of the spatial TR core tensors. Typically, an HSI $\mathcal{X}$ has a larger spatial rank, and a smaller spectral rank. Therefore, we changed R_1 from the range $[2, 7]$; while changing R_2 from the range $[50, 300]$. Fig. 7 shows the changes in the RMSE values with different TR *ranks* $[R_1, R_2, R_1]$. Typically, when the rank R_1 is larger, the performance of CTRF degrades significantly. This is mainly because the subspace dimensions of some classes may be smaller than R_1. However, with the nuclear norm constraint of the third core tensor, NCTRF can obtain smaller RMSE values, compared to those of CTRF. We choose TR *rank* to be $[3, 150, 3]$, $[4, 200, 4]$, *and*$[5, 250, 5]$ for $SNR = 20, 30, 40$, respectively.

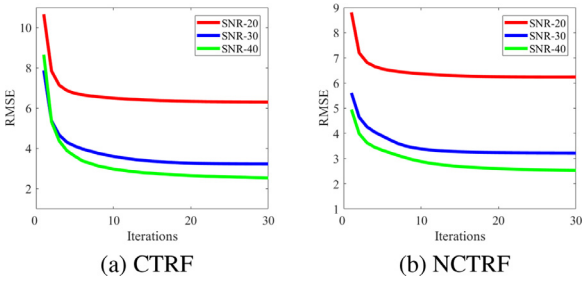


Fig. 8. The changes of RMSE value with the iterations on the WDC dataset.

Table 4
Computational time (s) of different methods with different dataset.

Data	CNMF	FUSE	HySure	STEREO	CSTF	NLSTF	CTRF	NCTRF
WDC	15.8	3.9	50.2	5.3	15.1	20.7	12.0	17.0
Pavia	15.4	3.8	46.1	5.4	14.8	21.9	11.7	17.1
Indian	3.1	1.6	18.6	0.9	3.6	13.2	2.2	3.3

Convergence. We conducted experiments to demonstrate the convergence behavior of the proposed CTRF and NCTRF methods. Fig. 8 shows the changes in RMSE obtained by CTRF and NCTRF, with an increase in the iteration number on the WDC dataset. It can be seen that, with increasing iterations, the RMSE values obtained by CTRF and NCTRF decreased to stable values, indicating the convergence of the proposed methods.

Computational time. Table 4 presents the computational times of the different methods with different datasets. From the table, it can be seen that the proposed methods are competitive when compared to the other traditional methods.

5. Conclusions

In this paper, we propose a CTRF model for HSR. The proposed model offers the advantages of the coupled matrix and Tucker factorization, and is shown to better exploit the low-rank properties of different HSI classes. Furthermore, we propose the NCTRF model by utilizing nuclear norm regularization of the third core tensor to exploit the global spectral low-rank properties of the recovered HR-HSI. An efficient alternating iteration method was developed to optimize the CTRF and NCTRF. Numerous experiments have demonstrated the advantages of the proposed methods compared to other tensor and deep learning methods. In future work, we plan to develop an automatic method to choose the TR *rank*.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported in part by the Japan Society for the Promotion of Science (KAKENHI) under Grant 19K20308, in part by the JST, FOREST, Japan, under Grant JPMJFR206S, and in part by the National Natural Science Foundation of China (NSFC) under Grant 62101222.

References

[1] L. Zhang, Q. Zhang, L. Zhang, D. Tao, X. Huang, B. Du, Ensemble manifold regularized sparse low-rank approximation for multiview feature embedding, Pattern Recognit. 48 (10) (2015) 3102–3112.
[2] H. Ghassemian, A review of remote sensing image fusion methods, Inform. Fusion 32 (2016) 75–89.
[3] S. Li, X. Kang, L. Fang, J. Hu, H. Yin, Pixel-level image fusion: a survey of the state of the art, Inform. Fusion 33 (2017) 100–112.
[4] H. Shen, X. Meng, L. Zhang, An integrated framework for the spatio–temporal–spectral fusion of remote sensing images, IEEE Trans. Geosci. Remote Sens. 54 (12) (2016) 7135–7148.
[5] Hyperspectral document image processing: applications, challenges and future prospects, Pattern Recognit. 90 (2019) 12–22.
[6] C.I. Kanatsoulis, X. Fu, N.D. Sidiropoulos, W.-K. Ma, Hyperspectral super-resolution: a coupled tensor factorization approach, IEEE Trans. Signal Process. 66 (24) (2018) 6503–6517.
[7] N. Yokoya, C. Grohnfeldt, J. Chanussot, Hyperspectral and multispectral data fusion: acomparative review of the recent literature, IEEE Geosci. Remote Sens. Mag. 5 (2) (2017) 29–56.
[8] R. Dian, L. Fang, S. Li, Hyperspectral image super-resolution via non-local sparse tensor factorization, in: CVPR, 2017, pp. 5344–5353.
[9] G.A. Shaw, H.-h.K. Burke, Spectral imaging for remote sensing, Lincoln Laboratory Journal 14 (1) (2003) 3–28.
[10] Q. Wei, J. Bioucas-Dias, N. Dobigeon, J.-Y. Tourneret, M. Chen, S. Godsill, Multiband image fusion based on spectral unmixing, IEEE Trans. Geosci. Remote Sens. 54 (12) (2016) 7236–7249.
[11] R. Kawakami, Y. Matsushita, J. Wright, M. Ben-Ezra, Y.-W. Tai, K. Ikeuchi, High-resolution hyperspectral imaging via matrix factorization, in: CVPR, IEEE, 2011, pp. 2329–2336.
[12] G. Vivone, L. Alparone, J. Chanussot, M. Dalla Mura, A. Garzelli, G.A. Licciardi, R. Restaino, L. Wald, A critical comparison among pansharpening algorithms, IEEE Trans. Geosci. Remote Sens. 53 (5) (2015) 2565–2586.
[13] L. Loncan, L.B. De Almeida, J.M. Bioucas-Dias, X. Briottet, J. Chanussot, N. Dobigeon, S. Fabre, W. Liao, G.A. Licciardi, M. Simoes, et al., Hyperspectral pansharpening: a review, IEEE Geosci. Remote Sens. Mag. 3 (3) (2015) 27–46.
[14] C. Chen, Y. Li, W. Liu, J. Huang, Sirf: simultaneous satellite image registration and fusion in a unified framework, IEEE Trans. Image Process. 24 (11) (2015) 4213–4224.
[15] R.C. Hardie, M.T. Eismann, G.L. Wilson, Map estimation for hyperspectral image resolution enhancement using an auxiliary sensor, IEEE Trans. Image Process. 13 (9) (2004) 1174–1184.
[16] Q. Wei, N. Dobigeon, J.-Y. Tourneret, Fast fusion of multi-band images based on solving a sylvester equation, IEEE Trans. Image Process. 24 (11) (2015) 4109–4121.
[17] N. Akhtar, F. Shafait, A. Mian, Bayesian sparse representation for hyperspectral image super resolution, in: CVPR, 2015, pp. 3631–3640.
[18] Y. Qu, H. Qi, C. Kwan, Unsupervised sparse dirichlet-net for hyperspectral image super-resolution, in: CVPR, 2018, pp. 2511–2520.
[19] Q. Xie, M. Zhou, Q. Zhao, D. Meng, W. Zuo, Z. Xu, Multispectral and hyperspectral image fusion by ms/hs fusion net, CVPR, 2019.
[20] Z. Shi, C. Chen, Z. Xiong, D. Liu, Z.-J. Zha, F. Wu, Deep residual attention network for spectral image super-resolution, in: ECCV, Springer, 2018, pp. 214–229.
[21] S. Nie, L. Gu, Y. Zheng, A. Lam, N. Ono, I. Sato, Deeply learned filter response functions for hyperspectral reconstruction, in: CVPR, 2018, pp. 4767–4776.
[22] R. Dian, S. Li, X. Kang, Regularizing hyperspectral and multispectral image fusion by cnn denoiser, IEEE Trans. Neural Netw. Learn. Syst. DOI:10.1109/TNNLS.2020.2980398 (2020) 1–12.
[23] Y. Fu, T. Zhang, Y. Zheng, D. Zhang, H. Huang, Hyperspectral image super-resolution with optimized rgb guidance, CVPR, 2019.
[24] L. Zhang, J. Nie, W. Wei, Y. Zhang, S. Liao, L. Shao, Unsupervised adaptation learning for hyperspectral imagery super-resolution, in: CVPR, 2020, pp. 3073–3082.
[25] Z. Zhu, J. Hou, J. Chen, H. Zeng, J. Zhou, Hyperspectral image super-resolution via deep progressive zero-centric residual learning, arXiv preprint arXiv:2006.10300 (2020).
[26] R. Dian, S. Li, L. Fang, Learning a low tensor-train rank representation for hyperspectral image super-resolution, IEEE Trans. Neural Netw. Learn. Syst. (2019).
[27] W. Dong, F. Fu, G. Shi, X. Cao, J. Wu, G. Li, X. Li, Hyperspectral image super-resolution via non-negative structured sparse representation, IEEE Trans. Image Process. 25 (5) (2016) 2337–2352.
[28] C. Lanaras, E. Baltsavias, K. Schindler, Hyperspectral super-resolution by coupled spectral unmixing, in: ICCV, 2015, pp. 3586–3594.
[29] M. Simões, J. Bioucas-Dias, L.B. Almeida, J. Chanussot, A convex formulation for hyperspectral image superresolution via subspace-based regularization, IEEE Trans. Geosci. Remote Sens. 53 (6) (2015) 3373–3388.
[30] M.A. Veganzones, M. Simoes, G. Licciardi, N. Yokoya, J.M. Bioucas-Dias, J. Chanussot, Hyperspectral super-resolution of locally low rank images from complementary multisource data, IEEE Trans. Image Process. 25 (1) (2016) 274–288.
[31] Q. Wei, J. Bioucas-Dias, N. Dobigeon, J.-Y. Tourneret, Hyperspectral and multispectral image fusion based on a sparse representation, IEEE Trans. Geosci. Remote Sens. 53 (7) (2015) 3658–3668.
[32] E. Wycoff, T.-H. Chan, K. Jia, W.-K. Ma, Y. Ma, A non-negative sparse promoting algorithm for high resolution hyperspectral imaging, in: ICASSP, IEEE, 2013, pp. 1409–1413.
[33] Y. Xu, Z. Wu, J. Chanussot, P. Comon, Z. Wei, Nonlocal coupled tensor cp decomposition for hyperspectral and multispectral image fusion, IEEE Trans. Geosci. Remote Sens. (2019) 1–15.
[34] N. Yokoya, T. Yairi, A. Iwasaki, Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion, IEEE Trans. Geosci. Remote Sens. 50 (2) (2012) 528–537.
[35] S. Li, R. Dian, L. Fang, J.M. Bioucas-Dias, Fusing hyperspectral and multispectral images via coupled sparse tensor factorization, IEEE Trans. Image Process. 27 (8) (2018) 4118–4130.
[36] H. Li, W. Li, G. Han, F. Liu, Coupled tensor decomposition for hyperspectral pansharpening, IEEE Access 6 (2018) 34206–34213.
[37] Y. Chang, L. Yan, S. Zhong, Hyper-laplacian regularized unidirectional low-rank tensor recovery for multispectral image denoising, in: CVPR, 2017, pp. 4260–4268.
[38] Y. Xu, Z. Wu, J. Chanussot, Z. Wei, Nonlocal patch tensor sparse representation for hyperspectral image super-resolution, IEEE Trans. Image Process. 28 (6) (2019) 3034–3047.
[39] K. Wang, Y. Wang, X.-L. Zhao, J.C.-W. Chan, Z. Xu, D. Meng, Hyperspectral and multispectral image fusion via nonlocal low-rank tensor decomposition and spectral unmixing, IEEE Trans. Geosci. Remote Sens. (2020).
[40] Q. Zhao, G. Zhou, S. Xie, L. Zhang, A. Cichocki, Tensor ring decomposition, arXiv preprint arXiv:1606.05535 (2016).
[41] W. Wang, V. Aggarwal, S. Aeron, Efficient low rank tensor ring completion, in: ICCV, 2017, pp. 5698–5706.
[42] W. He, H. Zhang, L. Zhang, H. Shen, Total-variation-regularized low-rank matrix factorization for hyperspectral image restoration, IEEE Trans. Geosci. Remote Sens. 54 (1) (2016) 178–188.
[43] L. Yuan, C. Li, D. Mandic, J. Cao, Q. Zhao, Tensor ring decomposition with rank minimization on latent space: An efficient approach for tensor completion, in: AAAI, 33, 2019, pp. 9151–9158.
[44] T.G. Kolda, B.W. Bader, Tensor decompositions and applications, SIAM Rev. 51 (3) (2009) 455–500.
[45] N.D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E.E. Papalexakis, C. Faloutsos, Tensor decomposition for signal processing and machine learning, IEEE Trans. Signal Process. 65 (13) (2017) 3551–3582.
[46] L.R. Tucker, Some mathematical notes on three-mode factor analysis, Psychometrika 31 (3) (1966) 279–311.
[47] I. Oseledets, Tensor-train decomposition, SIAM J. Sci. Comput. 33 (5) (2011) 2295–2317.
[48] Q. Zhao, M. Sugiyama, L. Yuan, A. Cichocki, Learning efficient tensor representations with ring-structured networks, in: ICASSP, 2019, pp. 8608–8612.
[49] W. He, N. Yokoya, L. Yuan, Q. Zhao, Remote sensing image reconstruction using tensor ring completion and total variation, IEEE Trans. Geosci. Remote Sens. 57 (11) (2019) 8998–9009.
[50] H. Huang, Y. Liu, J. Liu, C. Zhu, Provable tensor ring completion, Signal Processing 171 (2020) 107486.

[51] Y. Chen, W. He, N. Yokoya, T. Huang, X. Zhao, Nonlocal tensor-ring decomposition for hyperspectral image denoising, IEEE Trans. Geosci. Remote Sens. 58 (2) (2020) 1348–1362.
[52] V. Aggarwal, W. Wang, B. Eriksson, Y. Sun, W. Wang, Wide compression: tensor ring nets, Comput. Vis. Pattern Recognit. (2018) 9329–9338.
[53] M. Kuznetsov, D. Polykovskiy, D. Vetrov, A. Zhebrak, A prior of a googol gaussians: a tensor ring induced prior for generative models, NeurIPS (2019) 4104–4114.
[54] H. Huang, Y. Liu, C. Zhu, Coupled tensor completion via low-rank tensor ring., arXiv: Learning (2020).
[55] Y. Rivenson, A. Stern, Compressed imaging with a separable sensing operator, IEEE Signal Process. Lett. 16 (6) (2009) 449–452.
[56] J.M. Bioucas-Dias, J.M. Nascimento, Hyperspectral subspace identification, IEEE Trans. Geosci. Remote Sens. 46 (8) (2008) 2435–2445.
[57] W. He, Q. Yao, C. Li, N. Yokoya, Q. Zhao, H. Zhang, L. Zhang, Non-local meets global: an integrated paradigm for hyperspectral image restoration, IEEE Trans. Pattern Anal. Mach. Intell. (2020) 1, doi:10.1109/TPAMI.2020.3027563.
[58] W. He, H. Zhang, L. Zhang, Total variation regularized reweighted sparse nonnegative matrix factorization for hyperspectral unmixing, IEEE Trans. Geosci. Remote Sens. 55 (7) (2017) 3909–3921.
[59] J.-F. Cai, E.J. Candès, Z. Shen, A singular value thresholding algorithm for matrix completion, SIAM J. Optim. 20 (4) (2010) 1956–1982.
[60] Z. Lin, R. Liu, Z. Su, Linearized alternating direction method with adaptive penalty for low-rank representation, in: NIPS, 2011, pp. 612–620.
[61] Q. Wei, N. Dobigeon, J. Tourneret, J. Bioucas-Dias, S. Godsill, R-Fuse: robust fast fusion of multiband images based on solving a sylvester equation, IEEE Signal Process. Lett. 23 (11) (2016) 1632–1636.
[62] L. Wald, Quality of high resolution synthesised images: Is there a simple criterion? in: Third conference" Fusion of Earth data: merging point measurements, raster maps and remotely sensed images", SEE/URISCA, 2000, pp. 99–103.
[63] W. He, H. Zhang, H. Shen, L. Zhang, Hyperspectral image denoising using local low-rank matrix recovery and global spatial-spectral total variation, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 11 (3) (2018) 713–729.
[64] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, et al., Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.

Wei He received his B.S. degree in School of Mathematics and Statistics and his Ph.D degree in Surveying, Mapping and Remote Sensing (LIESMARS) from Wuhan University, Wuhan, China, in 2012 and 2017, respectively. He is currently a research scientist at the Geoinformatics Unit, RIKEN Center for Advanced Intelligence Project, Tokyo, Japan, where he was a researcher from 2018 to 2020. His research interests include image quality improvement, remote sensing image processing and low rank representation and deep learning.

Yong Chen received the B.S. degree from the School of Science, East China University of Technology, Nanchang, China, in 2015, and the Ph.D. degree from the School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, China, in 2020. He is currently with the School of Computer and Information Engineering, Jiangxi Normal University, Nanchang. From 2018 to 2019, he was a Research Intern with the Geoinformatics Unit, RIKEN Center for Advanced Intelligence Project, Japan. His research interests include remote sensing image processing and low-rank matrix/tensor representation.

Naoto Yokoya received his M.Eng. and Ph.D. degrees from the Department of Aeronautics and Astronautics, University of Tokyo, in 2010 and 2013, respectively. He is currently a lecturer at the University of Tokyo, Tokyo, Japan, and a unit leader at the RIKEN Center for Advanced Intelligence Project, Tokyo, Japan, where he leads the Geoinformatics Unit. He was an assistant professor at the University of Tokyo from 2013 to 2017. From 2015 to 2017, he was an Alexander von Humboldt Fellow, working at the German Aerospace Center, Oberpfaffenhofen, Germany, and at the Technical University of Munich, Munich, Germany. His research interests include image processing, data fusion, and machine learning for understanding remote sensing images, with applications to disaster management. Dr. Yokoya won the first place in the 2017 IEEE Geoscience and Remote Sensing Society (GRSS) Data Fusion Contest organized by the IEEE Image Analysis and Data Fusion Technical Committee (IADF TC). He was the chair (2019-2021) and a cochair (2017-2019) of the IEEE GRSS IADF TC. He has been an associate editor for IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS) since 2018.

Chao Li received the Ph.D degree in Harbin Engineering University, Harbin, China. He is currently a postdoctoral researcher in Tensor Learning Unit at RIKEN center for Advanced Intelligence Project, Japan. His research interests lies in the area of signal processing, machine learning and applied mathematic. He has collaborated actively with researcher in several other disciplines, particularly neural computing and wireless communication.

Qibin Zhao received the Ph.D. degree in computer science from Shanghai Jiao Tong University, China in 2009. He was a research scientist at RIKEN Brain Science Institute from 2009 to 2017. He is currently a unit leader for tensor learning unit at RIKEN Center for Advanced Intelligence Project (AIP), and is also a visiting professor in Saitama Institute of Technology, and a visiting associate professor in Tokyo University of Agriculture and Technology, Japan. His research interests include machine learning, tensor factorization and tensor networks, computer vision and brain signal processing. He has published more than 100 papers in international journals and conferences and two monographs. He is a senior member of IEEE and serve as an editorial board member for Science China Technological Sciences.