Multiscale Feature Tensor Train Rank Minimization for Multidimensional Image Recovery

Hao Zhang, Xi-Le Zhao, Tai-Xiang Jiang, Michael K. Ng, *Senior Member, IEEE*, and Ting-Zhu Huang

***Abstract*—The general tensor-based methods can recover missing values of multidimensional images by exploiting the low-rankness on the pixel level. However, especially when considerable pixels of an image are missing, the low-rankness is not reliable on the pixel level, resulting in some details losing in their results, which hinders the performance of subsequent image applications (e.g., image recognition and segmentation). In this article, we suggest a novel multiscale feature (MSF) tensorization by exploiting the MSFs of multidimensional images, which not only helps to recover the missing values on a higher level, that is, the feature level but also benefits subsequent image applications. By exploiting the low-rankness of the resulting MSF tensor constructed by the new tensorization, we propose the convex and nonconvex MSF tensor train rank minimization (MSF-TT) to conjointly recover the MSF tensor and the corresponding original tensor in a unified framework. We develop the alternating directional method of multipliers (ADMMs) to solve the convex MSF-TT and the proximal alternating minimization (PAM) to solve the nonconvex MSF-TT. Moreover, we establish the theoretical guarantee of convergence for the PAM algorithm. Numerical examples of real-world multidimensional images show that the proposed MSF-TT outperforms other compared approaches in image recovery and the recovered MSF tensor can benefit the subsequent image recognition.**

***Index Terms*—Feature-level tensor completion, multiscale features (MSFs), tensor train (TT) rank minimization, tensorization.**

Manuscript received 21 October 2020; revised 21 March 2021; accepted 23 August 2021. Date of publication 20 September 2021; date of current version 18 November 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 61876203, Grant 61772003, and Grant 12001446; in part by the Applied Basic Research Project of Sichuan Province under Grant 21YYJC3042; in part by the Key Project of Applied Basic Research in Sichuan Province under Grant 2020YJ0216; in part by the National Key Research and Development Program of China under Grant 2020YFA0714001; in part by the Fundamental Research Funds for the Central Universities under Grant JBK2102001; and in part by HKRGC GRF under Grant 12300218, Grant 12300519, Grant 17201020, and Grant 17300021. This article was recommended by Associate Editor D. Wang. *(Corresponding author: Xi-Le Zhao.)*

Hao Zhang, Xi-Le Zhao, and Ting-Zhu Huang are with the School of Mathematical Sciences/Research Center for Image and Vision Computing, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: aaronzhangfy@163.com; xlzhao122003@163.com; tingzhuhuang@126.com).

Tai-Xiang Jiang is with the FinTech Innovation Center, Financial Intelligence and Financial Engineering Research Key Laboratory of Sichuan Province, School of Economic Information Engineering, Southwestern University of Finance and Economics, Chengdu 611130, China (e-mail: taixiangjiang@gmail.com).

Michael K. Ng is with the Department of Mathematics, University of Hong Kong, Hong Kong (e-mail: mng@maths.hku.hk).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCYB.2021.3108847.

Digital Object Identifier 10.1109/TCYB.2021.3108847

I. Introduction

MULTIDIMENSIONAL images are generally higher-order tensors. For instance, a color image is a 3-D tensor with two indices for spatial variables and one index for the color channel, and a color video is a 4-D tensor added a temporal variable to the color image. However, multidimensional images often suffer from missing values in the acquisition and transmission [1]–[4], which hinders the subsequent image application, such as image recognition [5]; classification [6], [7]; and segmentation [8], [9]. Thus, recovering multidimensional images is to enhance the quality of observed images and enable the recovered results to help subsequent image applications [10], [11]. Tensor completion is the fundamental problem of multidimensional image recovery, becoming an important research topic [12]–[17].

Exploiting the low-rankness of tensors is effective to reconstruct the incomplete tensor data, called low-rank tensor completion (LRTC) [18]–[22]. The LRTC problem can be mathematically formulated as

$$\begin{aligned} \min_{\mathcal{X}} \ & \operatorname{rank}(\mathcal{X}) \\ \text{s.t.} \ & \mathcal{X}_\Omega = \mathcal{O}_\Omega \end{aligned} \tag{1}$$

where $\mathcal{X} \in \mathbb{R}^{n_1 \times \cdots \times n_j}$ is the low-rank tensor with j dimensions, Ω is the index of observed elements, $\mathcal{X}_\Omega$ is the projection operator to keep elements in Ω of $\mathcal{X}$ and make others be zero, and $\mathcal{O}_\Omega$ is the observed tensor with support Ω. Different from the matrix rank, the tensor rank has no unique definition. Based on different tensor decomposition schemes, several definitions of the tensor rank have been developed and considered in LRTC. The CANDECOMP/PARAFAC (CP) scheme [23] decomposes the tensor as the sum of the rank-one tensors. The CP rank [23] is defined as the minimal number of rank-one tensors. However, determining the CP rank of a given tensor is generally an NP-hard problem [24]. The Tucker scheme [25], [26] decomposes the tensor into a core tensor multiplied by matrices along each mode. Liu *et al.* [18] designed a convex surrogate of Tucker rank and developed the corresponding LRTC model as

$$\begin{aligned} \min_{\mathcal{X}} \ & \sum_{k=1}^{j} \alpha_k \left\| \mathbf{X}_{(k)} \right\|_* \\ \text{s.t.} \ & \mathcal{X}_\Omega = \mathcal{O}_\Omega \end{aligned} \tag{2}$$

where $\mathbf{X}_{(k)} \in \mathbb{R}^{n_k \times \prod_{d \neq k} n_d}$ is the mode-k matrization of $\mathcal{X} \in \mathbb{R}^{n_1 \times \cdots \times n_j}$ by the unfolding operator, $\|\cdot\|_*$ is the matrix nuclear norm, that is, the sum of singular values, and α_k is the constant

satisfying $\alpha_k \geq 0$ and $\sum_{k=1}^{j} \alpha_k = 1$. Then, Xu *et al.* [27] suggested a nonconvex Tucker rank optimization by applying the parallel low-rank matrix factorization on each $\mathbf{X}_{(k)}$. Their method, called TMac, is formulated as

$$\begin{aligned} \min_{\mathbf{E}_k, \mathbf{Z}_k, \mathcal{X}} \quad & \sum_{k=1}^{j} \frac{\alpha_k}{2} \left\| \mathbf{E}_k \mathbf{Z}_k - \mathbf{X}_{(k)} \right\|_F^2 \\ \text{s.t.} \quad & \mathcal{X}_\Omega = \mathcal{O}_\Omega \end{aligned} \tag{3}$$

where $\|\cdot\|_F$ is the Frobenius norm, and $\mathbf{E}_k \in \mathbb{R}^{n_k \times r_k}$ and $\mathbf{Z}_k \in \mathbb{R}^{r_k \times \prod_{d \neq k} n_d}$ are low-rank matrices with the rank no more than r_k. The computational cost of TMac is lower than that of (2). However, the matricization of Tucker decomposition produces unbalanced (i.e., not square) matrices [21], [28].

The tensor singular value decomposition (t-SVD) [29]–[32] introduces a tensor tubal rank, which well characterizes the low-rankness of the 3-D tensor. The tubal rank is defined as the number of nonzero tubes in the f-diagonal tensor [33], which comes from the t-SVD operator. Specifically, the t-SVD operator implements the discrete Fourier transform (DFT) along the third dimension and calculates matrix SVDs of frontal slices [34]–[36]. Kilmer and Martin [29] noted that the t-SVD framework can be implemented by an invertible transform, such as discrete cosine transform (DCT). Lu *et al.* [37] and Jiang *et al.* [38] introduced the DCT and framelet transform into the t-SVD framework for LRTC, respectively. However, the tubal rank is generally designed for 3-D data.

In recent years, a new tensor train (TT) decomposition scheme [39] is developed, which can exploit the redundancy of the higher-dimensional tensor. The TT-rank of $\mathcal{X} \in \mathbb{R}^{n_1 \times \cdots \times n_j}$ can be expressed as

$$\operatorname{rank}_{tt}(\mathcal{X}) = \left(\operatorname{rank}\left(\mathbf{X}_{[1]}\right), \ldots, \operatorname{rank}\left(\mathbf{X}_{[j-1]}\right)\right) \tag{4}$$

where $\mathbf{X}_{[k]} \in \mathbb{R}^{(\prod_{d=1}^{k} n_d) \times (\prod_{d=k+1}^{j} n_d)}$ is the mode-k canonical matricization of $\mathcal{X} \in \mathbb{R}^{n_1 \times \cdots \times n_j}$ [21] by the reshaping operator. Compared with Tucker decomposition, the matricization of TT decomposition is more balanced (i.e., square) on the middle mode of the tensor which leads to better results [28]. Based on the mode-k canonical matricization by the reshaping operator, TT-rank is easy to be optimized [21], [40]–[43]. Bengua *et al.* [21] proposed two surrogates for TT-rank-based LRTC: one is the convex surrogate for low-TT-rank tensor completion, which can be formulated as

$$\begin{aligned} \min_{\mathcal{X}} \quad & \sum_{k=1}^{j-1} \alpha_k \left\| \mathbf{X}_{[k]} \right\|_* \\ \text{s.t.} \quad & \mathcal{X}_\Omega = \mathcal{O}_\Omega \end{aligned} \tag{5}$$

where $\|\cdot\|_*$ is the matrix nuclear norm and α_k is the constant satisfying $\alpha_k \geq 0$ and $\sum_{k=1}^{j-1} \alpha_k = 1$. The second one is the nonconvex TT-rank minimization via the parallel low-rank matrix factorization (TMac-TT), which factorizes each $\boldsymbol{X}_{[k]}$ into the product of two smaller matrices, that is

$$\begin{aligned} \min_{\mathbf{E}_k, \mathbf{Z}_k, \mathcal{X}} \quad & \sum_{k=1}^{j-1} \frac{\alpha_k}{2} \left\| \mathbf{E}_k \mathbf{Z}_k - \mathbf{X}_{[k]} \right\|_F^2 \\ \text{s.t.} \quad & \mathcal{X}_\Omega = \mathcal{O}_\Omega \end{aligned} \tag{6}$$

where $\mathbf{E}_k \in \mathbb{R}^{(\prod_{d=1}^{k} n_d) \times r_k}$ and $\mathbf{Z}_k \in \mathbb{R}^{r_k \times (\prod_{d=k+1}^{j} n_d)}$ are low-rank matrices with the rank no more than r_k. Many extensions of TT-rank-based methods are also developed, such as the cyclically connected TT decomposition called tensor ring (TR) decomposition [44]–[46].

A. Tensorizations

For LRTC, tensorizations are usually considered to construct a higher-order tensor for better results in image recovery [21], [47], [48]. Ket augmentation (KA) and visual data tensorization (VDT) both use structured blocks addressing procedures to cast an image into a higher-order tensor, which unavoidably lead to artifacts near the boundary of blocks [21], [47]. The multiway delay-embedding tensorization (MDT) uses the duplicated transform to construct a higher-order tensor on the pixel level, which enhances the correlation of pixels and can capture the delay/shift-invariant structure of the tensor [48]–[51]. In summary, these previous tensorizations are on the pixel level.

B. Motivations and Contributions

These general LRTC-based methods mainly recover images on the pixel level. However, especially when considerable pixels of an image are missing, the low-rankness is not reliable on the pixel level [52]. Specifically, their results often miss some details and subsequent image applications are hindered. In the real world, multidimensional images are naturally with sufficient multiscale features (MSFs) information on each dimension, such as Gabor features [53]–[55] of face images, which can be exploited to benefit the image recovery and subsequent image applications (e.g., image recognition and segmentation). Recently, the multiscale image analysis has been successfully applied in image processing [56], [57]. The power of multiscale image analysis comes from the ability to take advantage of image features at multiple scales. The framelet transform, which is a powerful tool for multiscale image analysis, can capture the image features at multiple scales (e.g., coarse- and fine-scale features) [56], [57]. These motivate us to propose a new tensorization by exploiting the MSFs of multidimensional images based on the framelet transform, which benefits the image recovery and subsequent image applications, and then suggest a new LRTC framework on the feature level.

In this article, based on the framelet transform, we suggest a novel MSF tensorization by exploiting the MSFs of multidimensional images, which not only helps to recover images on the feature level but also benefits subsequent image applications. More concretely, by applying the filter h_r^l (the level parameter l controls the scale of the feature) along different modes of $\mathcal{X} \in \mathbb{R}^{n_1 \times \cdots \times n_j}$, we can obtain the feature tensor at level l. We can construct the MSF tensor $\mathcal{X}_{\mathbf{W}} \in \mathbb{R}^{n_1 \times \cdots \times n_j \times f_1 \cdots f_j \times L}$, by stacking these feature tensors of multiple scales (see Fig. 1). In Table I, we summarize the difference between the KA [21], MDT [49], and the proposed MSF tensorization. In summary, we can conclude that previous tensorizations [21], [49] are on the pixel level and our MSF tensorization is on the feature level.

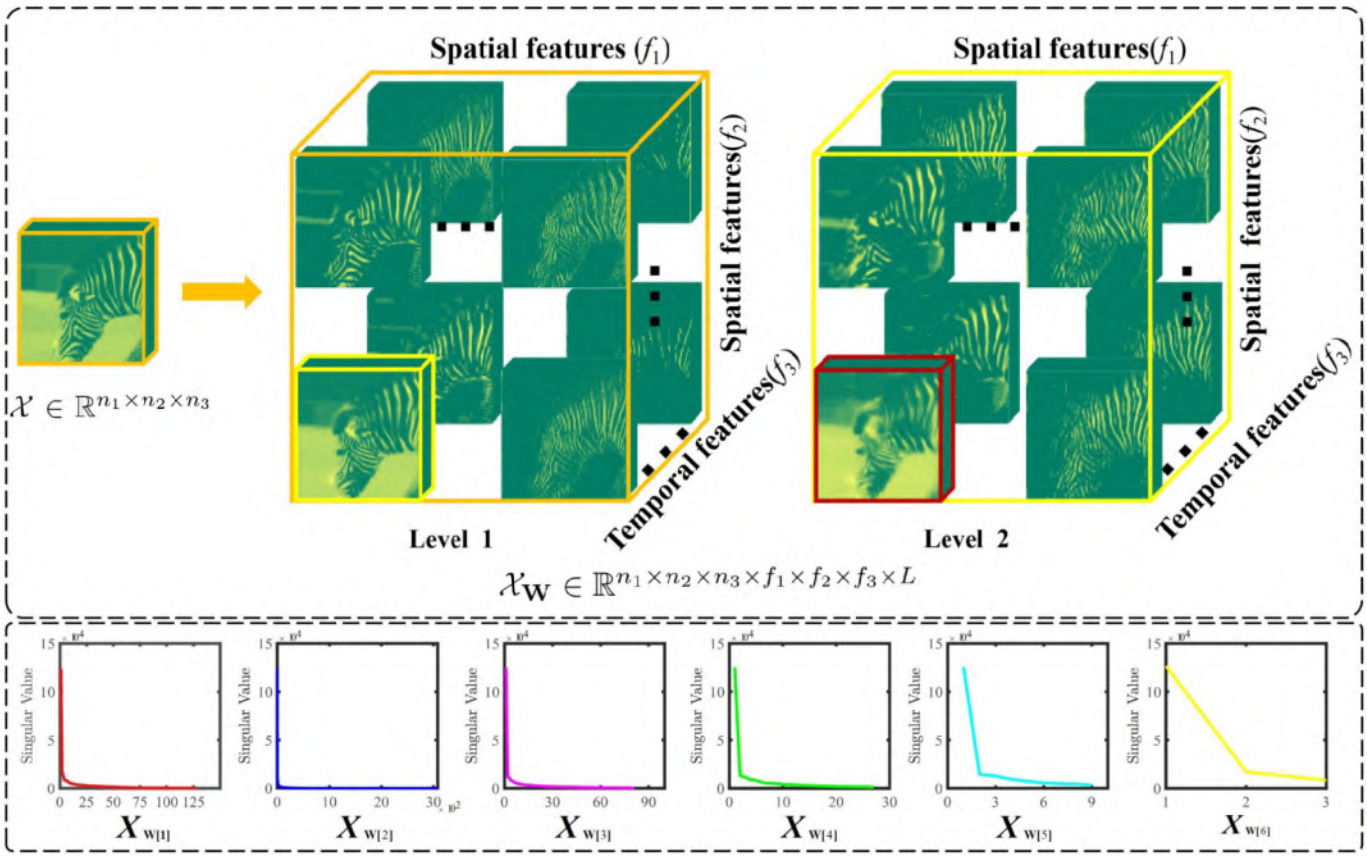


Fig. 1. MSF tensorization, where f_1 and f_2 are dimensions of the spatial features, f_3 is the dimension of the temporal features, and L is the dimension of the level. The curves show the singular values of matrices by mode-k canonical matricization of the multiscale feature tensor $\mathcal{X}_{\mathbf{W}}$.

TABLE I
COMPARISON BETWEEN DIFFERENT TENSORIZATIONS OF TENSOR $\mathcal{X} \in \mathbb{R}^{n \times n \times n}$ FOR TT DECOMPOSITION

Tensorization	Low-rankness		Number of Elements	Computational Cost	Storage Cost	Transform
	Pixel Level	Feature Level				
KA [21]	✓		$\Pi_{d=1}^{j} I_d = n^3$	$O(\sum_{k=1}^{j-1} r_k n^3)$	$O(\Pi_{d=2}^{j-2} r_{d-1} I_d r_d + r_0 I_1 + r_j I_j)$	—
MDT [49]	✓		$\Pi_{d=1}^{j} I_d = \tau^3 (n-\tau+1)^3$	$O(\sum_{k=1}^{j-1} r_k \tau^3 (n-\tau+1)^3)$	$O(\Pi_{d=2}^{j-2} r_{d-1} I_d r_d + r_0 I_1 + r_j I_j)$	duplicated transform
MSF		✓	$\Pi_{d=1}^{j} I_d = n^3 f^3 l$	$O(\sum_{k=1}^{n-1} r_k n^3 f^3 l)$	$O(\Pi_{d=2}^{j-2} r_{d-1} I_d r_d + r_0 I_1 + r_j I_j)$	framelet transform

Since TT decomposition is especially suitable for high-dimensional tensors [21], [28], we consider TT decomposition for the resulting high-dimensional MSF tensor $\mathcal{X}_{\mathbf{W}}$. The curves of Fig. 1 also illustrate that the resulting MSF tensor $\mathcal{X}_{\mathbf{W}}$ is numerically low-TT-rank, which is attributed to the redundancy of the MSFs. Thus, by the resulting MSF tensor, we further propose the new LRTC framework on the feature level called the MSF TT-rank minimization (MSF-TT) to conjointly recover the feature tensor and the corresponding original tensor (image). MSF-TT is formulated as

$$\begin{aligned} \min_{\mathcal{X}} \quad & \text{rank}_{tt}(\mathcal{X}_{\mathbf{W}}) \\ \text{s.t.} \quad & \mathcal{X}_{\Omega} = \mathcal{O}_{\Omega} \end{aligned} \tag{7}$$

where the optimized $\mathcal{X}_{\mathbf{W}} \in \mathbb{R}^{n_1 \times \cdots \times n_j \times f_1 \times \cdots \times f_j \times L}$ is the MSF tensor with $2j+1$ dimensions, f_j is the dimension of the features extracted from n_j, and L is the dimension of the level. The proposed MSF-TT framework enjoys two advantages.

1) MSF-TT can exploit the MSFs to recover the image for more information preservation.
2) The produced MSF tensor can be used to benefit subsequent image applications.

The contributions of our work are summarized as follows.

1) We suggest a novel MSF tensorization by exploiting the MSFs of multidimensional images, which not only helps to recover images on the feature level but also benefits subsequent image applications. Based on the

TABLE II
BASIC DEFINITIONS

Notations	Explanation
$y, \mathbf{y}, \mathbf{Y}, \mathcal{Y}$	scalar, vector, matrix, tensor
$y_{i_1 \dots i_k \dots i_j}$	the $\{i_1, \dots, i_k, \dots, i_j\}$-th elements of $\mathcal{Y}$
$\|\mathcal{Y}\|_F$	Frobenius norm: $\|\mathcal{Y}\|_F = \sqrt{\sum_{i_1}\sum_{i_2}\cdots\sum_{i_j} y^2_{i_1 i_2 \dots i_j}}$
$\|\mathbf{Y}\|_*$	nuclear norm: the sum of singular values of $\mathbf{Y}$
$\mathbf{Y}_{(k)}$	mode-k matricization of $\mathcal{Y}$ by unfolding operator
$\mathbf{Y}_{[k]}$	mode-k canonical matricization of $\mathcal{Y}$ by reshaping operator
$\times_k$	mode-k product

MSF tensorization, we propose a new LRTC framework on the feature level called MSF-TT.

2) We propose the specific convex and nonconvex MSF-TT model to conjointly recover the feature tensor and the corresponding original tensor. Then, we develop the alternating directional method of multipliers (ADMM) to solve the convex MSF-TT and the proximal alternating minimization (PAM) to solve the nonconvex MSF-TT. Moreover, we establish the theoretical guarantee of convergence for the PAM algorithm, under mild assumptions.

This article is arranged as follows. Section II gives notations and preliminaries for developing the MSF tensorization and the convex and nonconvex MSF-TT. Section III proposes the MSF tensorization, the convex and nonconvex MSF-TT models, and the solving algorithms of models. Section IV reports experimental examples and gives some discussions. Section V concludes this article.

II. NOTATIONS AND PRELIMINARIES

In this section, we introduce some notations and preliminaries for developing the MSF tensorization and the convex and nonconvex MSF-TT.

A. Basic Definitions

In Table II, we give some basic definitions. We denote the mode-k matricization of $\mathcal{Y}$ by the unfolding operator as $\mathbf{Y}_{(k)} = \text{unfold}_k(\mathcal{Y})$ of the size $n_k \times \prod_{d\neq k} n_d$ and its inverse operator as $\mathcal{Y} = \text{fold}_k(\mathbf{Y}_{(k)})$. The unfolding operator can be implemented via the following MATLAB command:

$$\mathbf{Y}_{(k)} = \text{reshape}(\text{shiftdim}(\mathcal{Y}, k-1), \text{size}(\mathcal{Y}), []). \tag{8}$$

We denote the mode-k canonical matricization of $\mathcal{Y}$ by the reshaping operator as $\mathbf{Y}_{[k]} = \text{reshape}_k(\mathcal{Y})$ of the size $(\prod_{d=1}^{k} n_d) \times (\prod_{d=k+1}^{j} n_d)$ and its inverse operator as $\mathcal{Y} = \text{unreshape}_k(\mathbf{Y}_{[k]})$. The reshape operator can be implemented by the following MATLAB command:

$$\mathbf{Y}_{[k]} = \text{reshape}\left(\mathcal{Y}, \prod_{d=1}^{k} n_d, \prod_{d=k+1}^{j} n_d\right). \tag{9}$$

B. Framelet Transform

Here, we introduce the framelet transform for developing the MSF tensorization. In image processing, the wavelet system can be discretely represented by the wavelet transform matrix $\mathbf{W}$ of the size $m \times n$ [56]. If the following equivalent holds, that is:

$$\mathbf{x} = \sum_{\mathbf{y}_i \in \mathbf{W}} \langle \mathbf{x}, \mathbf{y}_i \rangle \mathbf{y}_i \quad \forall\ \mathbf{x} \in \mathbb{R}^n \tag{10}$$

where $\mathbf{y}_i$ is the ith row of $\mathbf{W}$, then this is a wavelet tight frame system [56]. The wavelet tight frame system can be represented by the framelet transform matrix $\mathbf{W} \in \mathbb{R}^{wn\times n}$ constructed from a given filter bank $\{h_r^l\}$ with dilations and shifts, where r is from 0 to $f-1$, l is from 1 to L, and $w = (f-1)L+1$: see more details on [56, pp. 2–5]. The $\mathbf{W}$ and its transpose $\mathbf{W}^{\mathrm{T}}$ are also called the decomposition operator and the reconstruction operator [56]. The framelet transform matrix $\mathbf{W}$ satisfies $\mathbf{W}^{\mathrm{T}}\mathbf{W} = \mathbf{I}$ by the unitary extension principle [56].

III. MAIN RESULTS

In this section, we suggest the MSF tensorization, specific convex and nonconvex MSF-TT, and their solving algorithms.

MSF Tensorization: Based on the framelet transform, we propose the MSF tensorization by exploiting the MSFs of multidimensional images. We denote the framelet transform along mode-k of $\mathcal{X} \in \mathbb{R}^{n_1\times\cdots\times n_j}$ by $\mathbf{W}_k \in \mathbb{R}^{w_k n_k \times n_k}$. We have that $\mathcal{X} \times_k \mathbf{W}_k = \text{fold}_k(\mathbf{W}_k \mathbf{X}_{(k)})$, which extracts the MSFs of $\mathcal{X}$ on the kth dimension. Defining $\mathbf{W}_j \in \mathbb{R}^{w_j n_j \times n_j}$ and $\mathbf{W}_k \in \mathbb{R}^{f_k n_k \times n_k}$ with $k = 1, \dots, j-1$, we can construct $\mathcal{X} \times_1 \mathbf{W}_1 \cdots \times_j \mathbf{W}_j \in \mathbb{R}^{f_1 n_1 \times \cdots \times f_{j-1} n_{j-1} \times w_j n_j}$. Then, we incorporate the first $L-1$ coarse scale features and stack all features to construct the higher-dimensional MSF tensor $\mathcal{X}_{\mathbf{W}} \in \mathbb{R}^{n_1\times\cdots\times n_j \times f_1 \cdots f_j \times L}$. The procedure can be denoted as a operator $\mathfrak{R}$. The MSF tensorization can be mathematically formulated as

$$\mathcal{X}_{\mathbf{W}} = \mathfrak{R}\big(\mathcal{X} \times_1 \mathbf{W}_1 \cdots \times_j \mathbf{W}_j\big) \tag{11}$$

and the reconstruction of the original tensor is

$$\mathcal{X} = \mathfrak{R}^{-1}(\mathcal{X}_{\mathbf{W}}) \times_1 \mathbf{W}_1^T \cdots \times_j \mathbf{W}_j^T \tag{12}$$

where $\mathfrak{R}^{-1}$ is the inverse operator of $\mathfrak{R}$. Notably, the MSF tensor has $2j+1$ dimensions being higher than the original tensor. The TT-rank of $\mathcal{X}_{\mathbf{W}} \in \mathbb{R}^{n_1\times\cdots\times n_j \times f_1 \cdots f_j \times L}$ is

$$\text{rank}_{tt}(\mathcal{X}_{\mathbf{W}}) = \big(\text{rank}\big(\mathbf{X}_{\mathbf{W}[1]}\big), \dots, \text{rank}\big(\mathbf{X}_{\mathbf{W}[2j]}\big)\big). \tag{13}$$

Convex MSF-TT: We study the nuclear norm surrogate for developing the convex MSF-TT, that is

$$\begin{aligned} \min_{\mathcal{X}} \ & \sum_{k=1}^{2j} \alpha_k \big\|\mathbf{X}_{\mathbf{W}[k]}\big\|_* \\ \text{s.t.} \ & \mathcal{X}_\Omega = \mathcal{O}_\Omega \end{aligned} \tag{14}$$

where $\mathcal{X}$ is the recovered image, $\mathbf{X}_{\mathbf{W}[k]}$ is generated by the mode-k reshaping operator on the feature tensor $\mathcal{X}_{\mathbf{W}}$, $\mathcal{X}_{\mathbf{W}} = \mathfrak{R}(\mathcal{X} \times_1 \mathbf{W}_1 \cdots \times_j \mathbf{W}_j)$ by (11), $\|\cdot\|_*$ is the nuclear norm, α_k is the constant satisfying $\alpha_k \geq 0$ and $\sum_{k=1}^{2j} \alpha_k = 1$, Ω is the index of observed elements, $\mathcal{O}_\Omega$ is the observed tensor with

support Ω, and $\mathcal{X}_\Omega$ is the projection operator to keep elements in Ω of $\mathcal{X}$ and make others be zero.

Nonconvex MSF-TT: We study the parallel matrix factorization for developing the nonconvex MSF-TT, that is

$$\begin{aligned}\min_{\mathbf{E}_k,\mathbf{Z}_k,\mathcal{X}} \ & \sum_{k=1}^{2j}\frac{\alpha_k}{2}\left\|\mathbf{E}_k\mathbf{Z}_k-\mathbf{X}_{\mathbf{W}[k]}\right\|_F^2\\ \text{s.t.}\ \ & \mathcal{X}_\Omega=\mathcal{O}_\Omega\end{aligned}\tag{15}$$

where $\mathbf{E}_k\in\mathbb{R}^{(\prod_{d=1}^{k}n_d)\times r_k}$, $\mathbf{Z}_k\in\mathbb{R}^{r_k\times(\prod_{d=k+1}^{2j}n_d)}$, r_k is the bound of the rank($\mathbf{E}_k$) and rank($\mathbf{Z}_k$), and $\|\cdot\|_F$ is the Frobenius norm.

To solve the convex and nonconvex MSF-TT, we first introduce indicator functions [58] of $\mathcal{X}$, $\mathbf{E}_k$, and $\mathbf{Z}_k$. For closed subsets $\{\mathcal{X}:\mathcal{X}_\Omega=\mathcal{O}_\Omega\}$, $\{\mathbf{E}_k:\text{rank}(\mathbf{E}_k)\le r_k\}$, and $\{\mathbf{Z}_k:\text{rank}(\mathbf{Z}_k)\le r_k\}$, we denote indicator functions as $\delta_{\mathcal{D}}(\mathcal{X})$, $\delta_R(\mathbf{E}_k)$, and $\delta_R(\mathbf{Z}_k)$, that is

$$\begin{aligned}\delta_D(\mathcal{X})&=\begin{cases}0, & \text{if } \mathcal{X}\in\mathcal{D}=\{\mathcal{X}:\mathcal{X}_\Omega=\mathcal{O}_\Omega\}\\ +\infty, & \text{otherwise}\end{cases}\\ \delta_R(\mathbf{E}_k)&=\begin{cases}0, & \text{if } \mathbf{E}_k\in\mathcal{R}=\{\mathbf{E}_k:\text{rank}(\mathbf{E}_k)\le r_k\}\\ +\infty, & \text{otherwise}\end{cases}\\ \delta_R(\mathbf{Z}_k)&=\begin{cases}0, & \text{if } \mathbf{Z}_k\in\mathcal{R}=\{\mathbf{Z}_k:\text{rank}(\mathbf{Z}_k)\le r_k\}\\ +\infty, & \text{otherwise.}\end{cases}\end{aligned}\tag{16}$$

Then, we reformulate (14) and (15) as the following unconstrained minimizations, that is:

$$\min_{\mathcal{X}}\sum_{k=1}^{2j}\alpha_k\left\|\mathbf{X}_{\mathbf{W}[k]}\right\|_*+\delta_D(\mathcal{X})\tag{17}$$

and

$$\begin{aligned}\min_{\mathbf{E}_k,\mathbf{Z}_k,\mathcal{X}} \ & \sum_{k=1}^{2j}\frac{\alpha_k}{2}\left\|\mathbf{E}_k\mathbf{Z}_k-\mathbf{X}_{\mathbf{W}[k]}\right\|_F^2\\ & +\delta_D(\mathcal{X})+\delta_R(\mathbf{E}_k)+\delta_R(\mathbf{Z}_k).\end{aligned}\tag{18}$$

By minimizing (17) and (18), we have $\delta_{\mathcal{D}}(\mathcal{X})$, $\delta_R(\mathbf{E}_k)$, and $\delta_R(\mathbf{Z}_k)$ to be zero. In (17), if $\delta_{\mathcal{D}}(\mathcal{X})$ is zero, $\mathcal{X}$ belongs to the constrained set $\mathcal{D}$. Therefore, the unconstrained minimization (17) is equal to the convex MSF-TT (14). In (18), if $\delta_{\mathcal{D}}(\mathcal{X})$, $\delta_R(\mathbf{E}_k)$, and $\delta_R(\mathbf{Z}_k)$ are zero, $\mathcal{X}$ belongs to the constrained set $\mathcal{D}$, and $\mathbf{E}_k$ and $\mathbf{Z}_k$ belong to the constrained set $\mathcal{R}$. Thus, the unconstrained minimization (18) is equal to the nonconvex MSF-TT (15).

A. ADMM for the Convex MSF-TT

Since the minimization (17) is convex, we develop the ADMM to solve it. We introduce auxiliary variables $\mathcal{M}_k$ into (17), that is

$$\begin{aligned}\min_{\mathcal{X}} \ & \sum_{k=1}^{2j}\alpha_k\left\|\mathbf{M}_{k[k]}\right\|_*+\delta_D(\mathcal{X})\\ \text{s.t.}\ \ & \mathcal{X}_{\mathbf{W}}=\mathcal{M}_k.\end{aligned}\tag{19}$$

Then, the related augmented Lagrangian function can be formulated as

$$\begin{aligned}\min_{\mathcal{X}} \ & \sum_{k=1}^{2j}\alpha_k\left\|\mathbf{M}_{k[k]}\right\|_*+\sum_{k=1}^{2j}\frac{\beta_k}{2}\|\mathcal{M}_k-\mathcal{X}_{\mathbf{W}}\|_F^2\\ & +\sum_{k=1}^{2j}\langle\mathcal{M}_k-\mathcal{X}_{\mathbf{W}},\mathcal{B}_k\rangle+\delta_D(\mathcal{X})\end{aligned}\tag{20}$$

where $\mathcal{B}_k$ is the Lagrangian multiplier and β_k is the penalty parameter for the violation of the linear constraints. Since each variable $\mathcal{M}_k$ is independent, the solving procedure can be divided into two blocks.

1) *Solving $\{\mathcal{M}_k\}$:* For $\mathcal{X}$, $\|\mathcal{X}\|_F^2=\|\mathbf{X}_{[k]}\|_F^2$. Thus, the problem of $\{\mathcal{M}_k\}$ can be rewritten as the following $2j$ subproblems, that is:

$$\begin{aligned}\mathcal{M}_k^{t+1}=\arg\min_{\mathcal{M}_k}\ & \alpha_k\left\|\mathbf{M}_{k[k]}\right\|_*\\ & +\frac{\beta_k}{2}\left\|\mathbf{M}_{k[k]}-\mathbf{X}^t_{\mathbf{W}[k]}+\frac{\mathbf{B}^t_{k[k]}}{\beta_k}\right\|_F^2.\end{aligned}\tag{21}$$

We solve (21) by the singular value thresholding (SVT) operator, that is

$$\mathcal{M}_k^{t+1}=\text{unreshape}_k\left(\text{SVT}_{\frac{\alpha_k}{\beta_k}}\left(\mathbf{X}^t_{\mathbf{W}[k]}-\frac{\mathbf{B}^t_{k[k]}}{\beta_k}\right)\right).\tag{22}$$

2) *Solving $\mathcal{X}$:* For $\mathcal{X}$, $\|\mathcal{X}\|_F^2=\|\mathbf{X}_{(k)}\|_F^2$. By the unitary extension principle ($\mathbf{W}_k^{\mathrm{T}}\mathbf{W}_k=\mathbf{I}$), we have that

$$\begin{aligned}\mathcal{X}^{t+1}&=\arg\min_{\mathcal{X}\in\mathcal{D}}\sum_{k=1}^{2j}\frac{\beta_k}{2}\left\|\mathcal{M}_k^{t+1}-\mathcal{X}_{\mathbf{W}}+\frac{\mathcal{B}^t_k}{\beta_k}\right\|_F^2\\ &=\arg\min_{\mathcal{X}\in\mathcal{D}}\sum_{k=1}^{2j}\frac{\beta_k}{2}\\ &\quad\times\left\|\mathcal{M}_k^{t+1}-\mathfrak{R}\big(\mathcal{X}\times_1\mathbf{W}_1\cdots\times_j\mathbf{W}_j\big)+\frac{\mathcal{B}^t_k}{\beta_k}\right\|_F^2\\ &=\arg\min_{\mathcal{X}\in\mathcal{D}}\sum_{k=1}^{2j}\frac{\beta_k}{2}\left\|\mathfrak{R}^{-1}\left(\mathcal{M}_k^{t+1}+\frac{\mathcal{B}^t_k}{\beta_k}\right)\right.\\ &\qquad\left.\times_1\mathbf{W}_1^T\cdots\times_j\mathbf{W}_j^T-\mathcal{X}\right\|_F^2.\end{aligned}\tag{23}$$

It is a least-square problem. Then, we have

$$\begin{aligned}\mathcal{X}^{t+1}&=P_{\mathcal{D}}\left(\frac{\sum_{k=1}^{2j}\beta_k\mathfrak{R}^{-1}\left(\mathcal{M}_k^{t+1}+\frac{\mathcal{B}^t_k}{\beta_k}\right)\times_1\mathbf{W}_1^T\cdots\times_j\mathbf{W}_j^T}{\sum_{k=1}^{2j}\beta_k}\right)\\ \mathcal{X}^{t+1}_{\mathbf{W}}&=\mathfrak{R}\Big(\mathcal{X}^{t+1}\times_1\mathbf{W}_1\cdots\times_j\mathbf{W}_j\Big)\end{aligned}\tag{24}$$

where $P_{\mathcal{D}}$ is the projection onto $\mathcal{D}=\{\mathcal{X}:\mathcal{X}_\Omega=\mathcal{O}_\Omega\}$. The multipliers $\{\mathcal{B}_k\}$ can be updated as

$$\mathcal{B}_k^{t+1}=\mathcal{B}_k^t+\beta_k\Big(\mathcal{M}_k^{t+1}-\mathcal{X}^{t+1}_{\mathbf{W}}\Big).\tag{25}$$

We summarize the ADMM for the convex MSF-TT in Algorithm 1.

Algorithm 1 ADMM for the Convex MSF-TT

Input The observed $\mathcal{Y}$, Ω, α_k, β_k, and ϵ.
Output $\mathcal{X}$, $\mathcal{X}_{\mathbf{W}}$.
Initialize $\mathcal{X}^0 = \mathcal{Y}$, $\mathcal{M}_k^0 = \mathcal{B}_k^0 = \mathbf{0}$, $t = 0$, and t_{max}.
While not converged and $t \le t_{max}$ **do**
Update $\mathcal{M}_k^{t+1}$ via (22).
Update $\mathcal{X}^{t+1}$ and $\mathcal{X}_{\mathbf{W}}^{t+1}$ via (24).
Update $\mathcal{B}_k^{t+1}$ via (25).
Let $t = t + 1$.
Check the convergence condition:
$\|\mathcal{X}^{t+1} - \mathcal{X}^t\|_F^2 / \|\mathcal{X}^t\|_F^2 \le \epsilon$ or $t > t_{max}$.
End while

Computational Complexity: The computing burden for the convex MSF-TT mainly comes from updates of $\mathcal{M}_k^{t+1}$ and $\mathcal{X}^{t+1}$. For a j-dimensional tensor $\mathcal{X} \in \mathbb{R}^{n\times\cdots\times n}$, the corresponding $2j+1$-dimensional feature tensor $\mathcal{X}_{\mathbf{W}} \in \mathbb{R}^{I\times\cdots\times I}$, and $\mathbf{W}_k \in \mathbb{R}^{w_k n\times n}$, $O(\min(I^k, I^{2j+1-k})I^{2j+1})$ is the computational complexity for updating each $\mathcal{M}_k$, and the update of $\mathcal{X}$ mainly needs the reconstruction of MSF tensorization, that is, $O(w_j^j n^{j+1})$. Thus, the computational complexity of the convex MSF-TT is $O(\sum_{k=1}^{2j} \min(I^k, I^{2j+1-k})I^{2j+1} + w_j^j n^{j+1})$.

Convergence Analysis: The optimization (19) is convex and well structured. All the variables of (19) can be divided into two independent blocks, including block $\mathcal{X}$ and block $\{\mathcal{M}_k\}$. It is clear that (19) fits the ADMM framework [59]. Therefore, the convergence of Algorithm 1 is theoretically guaranteed.

B. PAM for the Nonconvex MSF-TT

Since the minimization (18) is strongly nonconvex, we develop the PAM to solve it and establish the theoretical guarantee of convergence. We denote (18) as $f(\mathbf{E}, \mathbf{Z}, \mathcal{X})$, where $\mathbf{E}$ and $\mathbf{Z}$ are block-diagonal matrices formed by $\mathbf{E}_k$ and $\mathbf{Z}_k$. The sequence $\{\mathbf{E}, \mathbf{Z}, \mathcal{X}\}$ can be updated as

$$\begin{aligned}
\mathbf{E}^{t+1} &= \arg\min_{\mathbf{E}_k\in\mathcal{R}} \left\{ M_1(\mathbf{E}|\mathbf{E}^t) = f(\mathbf{E}, \mathbf{Z}^t, \mathcal{X}^t) + \frac{\rho}{2}\|\mathbf{E} - \mathbf{E}^t\|_F^2 \right\} \\
\mathbf{Z}^{t+1} &= \arg\min_{\mathbf{Z}_k\in\mathcal{R}} \left\{ M_2(\mathbf{Z}|\mathbf{Z}^t) = f(\mathbf{E}^{t+1}, \mathbf{Z}, \mathcal{X}^t) \right. \\
&\qquad \left. + \frac{\rho}{2}\|\mathbf{Z} - \mathbf{Z}^t\|_F^2 \right\} \\
\mathcal{X}^{t+1} &= \arg\min_{\mathcal{X}\in D} \left\{ M_3(\mathcal{X}|\mathcal{X}^t) = f(\mathbf{E}^{t+1}, \mathbf{Z}^{t+1}, \mathcal{X}) \right. \\
&\qquad \left. + \frac{\rho}{2}\|\mathcal{X} - \mathcal{X}^t\|_F^2 \right\}. \qquad (26)
\end{aligned}$$

1) *Solving $\mathbf{E}$ and $\mathbf{Z}$:* We can see that $\mathbf{E}$ and $\mathbf{Z}$ can be decomposed into $2j$ independent problems. Thus, $\mathbf{E}$ and $\mathbf{Z}$ can be updated as

$$\begin{aligned}
\mathbf{E}_k^{t+1} &= \arg\min_{\mathbf{E}_k\in\mathcal{R}} \frac{\alpha_k}{2}\left\|\mathbf{E}_k\mathbf{Z}_k^t - \mathbf{X}_{\mathbf{W}[k]}^t\right\|_F^2 + \frac{\rho}{2}\|\mathbf{E}_k - \mathbf{E}_k^t\|_F^2 \\
&= \left(\alpha_k \mathbf{X}_{\mathbf{W}[k]}^t (\mathbf{Z}_k^t)^T + \rho\mathbf{E}_k^t\right)\left(\alpha_k \mathbf{Z}_k^t (\mathbf{Z}_k^t)^T + \rho\mathbf{I}\right)^{-1} \\
\mathbf{Z}_k^{t+1} &= \arg\min_{\mathbf{Z}_k\in\mathcal{R}} \frac{\alpha_k}{2}\left\|\mathbf{E}_k^{t+1}\mathbf{Z}_k - \mathbf{X}_{\mathbf{W}[k]}^t\right\|_F^2 + \frac{\rho}{2}\|\mathbf{Z}_k - \mathbf{Z}_k^t\|_F^2 \\
&= \left(\alpha_k \left(\mathbf{E}_k^{t+1}\right)^T \mathbf{E}_k^{t+1} + \rho\mathbf{I}\right)^{-1} \\
&\quad \times \left(\alpha_k \left(\mathbf{E}_k^{t+1}\right)^T \mathbf{X}_{\mathbf{W}[k]}^t + \rho\mathbf{Z}_k^t\right). \qquad (27)
\end{aligned}$$

Algorithm 2 PAM for the Nonconvex MSF-TT

Input The observed $\mathcal{Y}$, Ω, α_k, r_k and ϵ.
Output $\mathcal{X}$, $\mathcal{X}_{\mathbf{W}}$.
Initialize $\mathcal{X}^0 = \mathcal{Y}$, $\mathbf{E}_k^0$, $\mathbf{Z}_k^0$, $t = 0$, and t_{max}.
While not converged and $t \le t_{max}$ **do**
Update $\mathbf{E}_k^{t+1}$ and $\mathbf{Z}_k^{t+1}$ via (27).
Update $\mathcal{X}^{t+1}$ and $\mathcal{X}_{\mathbf{W}}^{t+1}$ via (29).
Let $t = t + 1$.
Check the convergence condition:
$\|\mathcal{X}^{t+1} - \mathcal{X}^t\|_F^2 / \|\mathcal{X}^t\|_F^2 \le \epsilon$ or $t > t_{max}$.
End while

2) *Solving $\mathcal{X}$:* By the unitary extension principle ($\mathbf{W}_k^{\mathrm{T}}\mathbf{W}_k = \mathbf{I}$), we have that

$$\begin{aligned}
\mathcal{X}^{t+1} &= \arg\min_{\mathcal{X}\in\mathcal{D}} \sum_{k=1}^{2j} \frac{\alpha_k}{2}\left\|\mathbf{E}_k^{t+1}\mathbf{Z}_k^{t+1} - \mathbf{X}_{\mathbf{W}[k]}\right\|_F^2 + \frac{\rho}{2}\|\mathcal{X} - \mathcal{X}^t\|_F^2 \\
&= \arg\min_{\mathcal{X}\in\mathcal{D}} \sum_{k=1}^{2j} \frac{\alpha_k}{2}\left\|\mathfrak{R}^{-1}\left(\text{unreshape}_k\left(\mathbf{E}_k^{t+1}\mathbf{Z}_k^{t+1}\right)\right)\right. \\
&\qquad \left. \times_1 \mathbf{W}_1^T \cdots \times_j \mathbf{W}_j^T - \mathcal{X}\right\|_F^2 \\
&\quad + \frac{\rho}{2}\|\mathcal{X} - \mathcal{X}^t\|_F^2. \qquad (28)
\end{aligned}$$

It is a least-square problem. Then

$$\begin{aligned}
\mathcal{X}^{t+1} &= P_{\mathcal{D}}\left(\frac{\sum_{k=1}^{2j}\alpha_k\mathcal{X}_{\mathbf{temp}}^{t+1} + \rho\mathcal{X}^t}{\sum_{k=1}^{2j}\alpha_k + \rho}\right) \\
\mathcal{X}_{\mathbf{W}}^{t+1} &= \mathfrak{R}\left(\mathcal{X}^{t+1} \times_1 \mathbf{W}_1 \cdots \times_j \mathbf{W}_j\right) \qquad (29)
\end{aligned}$$

where $\mathcal{X}_{\mathbf{temp}}^{t+1} = \mathfrak{R}^{-1}(\text{unreshape}_k(\mathbf{E}_k^{t+1}\mathbf{Z}_k^{t+1})) \times_1 \mathbf{W}_1^{\mathrm{T}} \cdots \times_j \mathbf{W}_j^{\mathrm{T}})$. We summarize the PAM for the nonconvex MSF-TT in Algorithm 2.

Computational Complexity: The computing burden for the nonconvex MSF-TT mainly comes from the updates of $\mathbf{E}_k$, $\mathbf{Z}_k$, and $\mathcal{X}$. For a j-dimensional tensor $\mathcal{X} \in \mathbb{R}^{n\times\cdots\times n}$, the corresponding $2j+1$-dimensional feature tensor $\mathcal{X}_{\mathbf{W}} \in \mathbb{R}^{I\times\cdots\times I}$, and $\mathbf{W}_k \in \mathbb{R}^{w_k n\times n}$, the updates of each $\mathbf{E}_k$ and $\mathbf{Z}_k$ need $O()$, and the computational complexity of (29) is $O(w_j^j n^{j+1})$. Thus, solving the nonconvex MSF-TT needs $O(\sum_{k=1}^{2j} r_k I^{2j+1} + w_j^j n^{j+1})$.

Convergence Analysis: Here, we establish the theoretical guarantee of convergence for Algorithm 2, under mild assumptions.

Theorem 1: Assuming that $\mathbf{E}^t$, $\mathbf{Z}^t$, and $\mathcal{X}^t$ are bounded, the sequence $\{\mathbf{E}^t, \mathbf{Z}^t, \mathcal{X}^t\}_{t\in\mathbb{N}}$ generated by Algorithm 2 converges to a critical point of (15).

The detailed proof of Theorem 1 can be found in the Appendix.

IV. Experiments

In this section, we conduct comprehensive experiments to verify the effectiveness of our methods. Color images, multispectral images (MSIs), color videos, and face data were included in experiments, of which the gray values were normalized to [0, 255]. The peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) [60] served as the quantitative evaluation. Generally, better recovery performances are

reflected by higher PSNR and SSIM values. We calculate the mean PSNR (MPSNR) and mean SSIM (MSSIM) by averaging the PSNR values and the SSIM values of all bands for high-dimensional data. For each band, the evaluation metrics (PSNR and SSIM) are computed for all pixels by the following exact equations, that is:

$$\mathrm{PSNR} = 10\log_{10}\frac{N\mathbf{Z}_{\max}^2}{\left\|\mathbf{Z}-\mathbf{Z}^*\right\|_F^2}$$

and

$$\mathrm{SSIM} = \frac{(2\mu_{\mathbf{Z}}\mu_{\mathbf{Z}^*})(2\sigma_{\mathbf{Z}\mathbf{Z}^*}+C_2)}{(\mu_{\mathbf{Z}}^2+\mu_{\mathbf{Z}^*}^2+C_1)(\sigma_{\mathbf{Z}}^2+\sigma_{\mathbf{Z}^*}^2+C_2)}$$

where $\mathbf{Z}^*$ is one band of the original image, $\mathbf{Z}$ is one band of the recovered image, N denotes the total number of pixels in the image, $\mathbf{Z}_{\max}$ is the maximum pixel value of the image, $\mu_{\mathbf{Z}}$ and $\mu_{\mathbf{Z}}^*$ are the mean values of images $\mathbf{Z}$ and $\mathbf{Z}^*$, $\sigma_{\mathbf{Z}}$ and $\sigma_{\mathbf{Z}}^*$ are the standard variances of $\mathbf{Z}$ and $\mathbf{Z}^*$, $\sigma_{\mathbf{Z}\mathbf{Z}^*}$ is the covariance of $\mathbf{Z}$ and $\mathbf{Z}^*$, and $C_1 > 0$ and $C_2 > 0$ are constants.

The current mainstream methods of DCT-TNN [37], MDT-Tucker [48], KA-TT [21], and TRLRF [44] are employed for performance comparison with the proposed methods. DCT-TNN is a tubal rank-based method, which implements the DCT along the third dimension to calculate t-SVD. MDT-Tucker uses the duplicated transform to construct a higher-order tensor and then exploits the low-rankness of the higher-order tensor by Tucker decomposition. KA-TT casts an image into a higher-order tensor and then exploits the low-rankness of the higher-order tensor by TMac-TT. TRLRF is a TR rank-based method, which uses the low-rank matrix factorization to calculate the low-TR-rank tensor. In experiments, we denote the convex MSF-TT by MSF-TT1 and the nonconvex MSF-TT by MSF-TT2 for convenience.

Parameters Setting: Here, we give the parameters setting of the proposed methods, which involves two parts.

Model Part: The first part is the parameters of the multidimensional framelet transform, which influences the feature tensor $\mathcal{X}_{\mathbf{W}}$ of the model. The proposed method is a unified framework, which can flexibly consider different filter banks to generate the feature tensor for image recovery and different image applications. In this work, we mainly consider the framelet transform based on four different filter banks: 1) the Haar wavelet (*Haar*); 2) the piecewise linear B-spline filter bank (*Linear*) [57]; 3) the piecewise cubic filter bank (*Cubic*) [57]; and 4) the Gabor filter bank [53]. The framelet transform settings on each dimension of $\mathcal{X}$ are the same, and the level is limited in $\{1, 2, 3\}$.

Algorithm Part: The second part is the parameters of ADMM and PAM. In ADMM, we control the (α_k/β_k) ranged in $\{10^{-3}, 5\times10^{-3}, 10^{-2}\}$, which governs the solution of the SVT operator (22). In PAM, we control the rank of initial $\mathbf{E}^0$ and $\mathbf{Z}^0$ by rank tolerance (r_{tol}) [40], which is defined as

$$\frac{\sigma_q^{[k]}}{\sigma_1^{[k]}} > r_{\mathrm{tol}} \tag{30}$$

where $\sigma_q^{[k]}$ is the qth singular value of $\mathbf{X}_{\mathbf{W}[k]}$, $q = 1, \ldots, r_k$, and r_{tol} is selected from $\{0.005, 0.007, 0.009, 0.01\}$. The convergence criterion of our proposed algorithms is defined by computing the relative change (RelCha) of the recovered $\mathcal{X}$ between two successive iterations, that is

$$\mathrm{RelCha} = \frac{\left\|\mathcal{X}^{t+1}-\mathcal{X}^t\right\|_F}{\left\|\mathcal{X}^t\right\|_F}. \tag{31}$$

In whole experiments, we terminate the iteration when RelCha is smaller than 10^{-4}.

In all experiments, the optimal parameters of the proposed method and competing methods are chosen to attain the highest PSNR value.

Based on the MATLAB (2017b) programming, all experiments are performed on the Windows 10 system with the CPU Intel Core i7-8700k 3.70 GHz and 16-GB RAM. All the methods are run once for all experiments.

TABLE III
QUANTITATIVE EVALUATION OF COMPETING METHODS ON COLOR IMAGES WITH THE RANDOM MISSING. THE BEST VALUES AND THE SECOND-BEST VALUES ARE HIGHLIGHTED BY BOLD FONTS AND UNDERLINES, RESPECTIVELY

Images	SR / Method	10% PSNR	10% SSIM	20% PSNR	20% SSIM	30% PSNR	30% SSIM
Lena	DCT-TNN	19.20	0.357	23.75	0.663	25.45	0.760
	MDT-Tucker	25.65	0.789	28.22	0.862	30.40	0.898
	KA-TT	23.46	0.637	26.25	0.765	28.21	0.830
	TRLRF	24.64	0.649	27.09	0.775	29.80	0.853
	MSF-TT1	26.09	0.804	28.13	0.834	30.43	0.896
	MSF-TT2	**26.43**	**0.812**	**28.92**	**0.866**	**30.97**	**0.917**
Peppers	DCT-TNN	17.09	0.265	21.13	0.4815	24.12	0.653
	MDT-Tucker	24.87	0.792	26.64	0.852	28.95	0.891
	KA-TT	21.57	0.585	24.12	0.732	29.12	0.864
	TRLRF	22.51	0.604	25.17	0.753	29.10	0.891
	MSF-TT1	23.63	0.781	27.14	0.836	29.01	0.893
	MSF-TT2	**25.25**	**0.804**	**27.79**	**0.868**	**29.51**	**0.901**
Baboon	DCT-TNN	18.07	0.279	20.36	0.443	22.28	0.588
	MDT-Tucker	21.31	0.491	22.32	0.624	23.55	0.704
	KA-TT	19.92	0.331	21.85	0.542	23.62	0.651
	TRLRF	20.95	0.315	21.88	0.561	22.33	0.684
	MSF-TT1	21.37	0.477	22.74	0.603	24.16	0.712
	MSF-TT2	**22.26**	**0.561**	**23.21**	**0.625**	**24.72**	**0.763**
Airplane	DCT-TNN	19.92	0.438	23.18	0.616	25.72	0.744
	MDT-Tucker	22.87	0.671	**25.86**	**0.843**	**27.91**	0.889
	KA-TT	21.37	0.586	24.05	0.784	26.79	0.848
	TRLRF	22.56	0.594	24.47	0.792	26.28	0.861
	MSF-TT1	23.15	0.658	25.04	0.822	27.15	0.879
	MSF-TT2	**24.13**	**0.693**	25.58	0.834	27.64	**0.893**
House	DCT-TNN	20.87	0.412	24.55	0.618	27.43	0.742
	MDT-Tucker	25.03	0.736	26.41	0.787	29.13	0.823
	KA-TT	24.66	0.618	27.86	0.741	29.54	0.827
	TRLRF	25.24	0.628	27.49	0.761	29.88	0.803
	MSF-TT1	26.08	0.712	28.39	0.796	30.34	0.814
	MSF-TT2	**27.06**	**0.768**	**29.78**	**0.802**	**31.14**	**0.847**
Barbara	DCT-TNN	18.01	0.365	21.72	0.582	24.23	0.724
	MDT-Tucker	22.33	0.695	24.24	0.792	25.95	0.838
	KA-TT	22.54	0.540	24.21	0.712	26.87	0.821
	TRLRF	23.02	0.498	26.17	0.723	28.19	0.811
	MSF-TT1	24.01	0.721	26.24	0.797	28.42	0.854
	MSF-TT2	**25.08**	**0.781**	**27.13**	**0.821**	**29.88**	**0.875**

A. Color Images

In this section, we test the proposed methods on color image completion. Color images,[1] including *Lena*, *Peppers*, *Baboon*, *Airplane*, *House*, and *Barbara*, are employed in this experiment. The size of the tested images is $256\times256\times3$. The color images completion includes two different cases: 1) random missing and 2) structural missing.

Random Missing: In the random missing case, we display the challenging sampling rate (SR) cases are 10%, 20%, and 30%, respectively, and the missing pixels are randomly distributed in red, green, and blue channels.

In Table III, we report the quantitative comparisons of competing methods on color images with different SRs. In most

[1] Available at http://sipi.usc.edu/database/database.php.

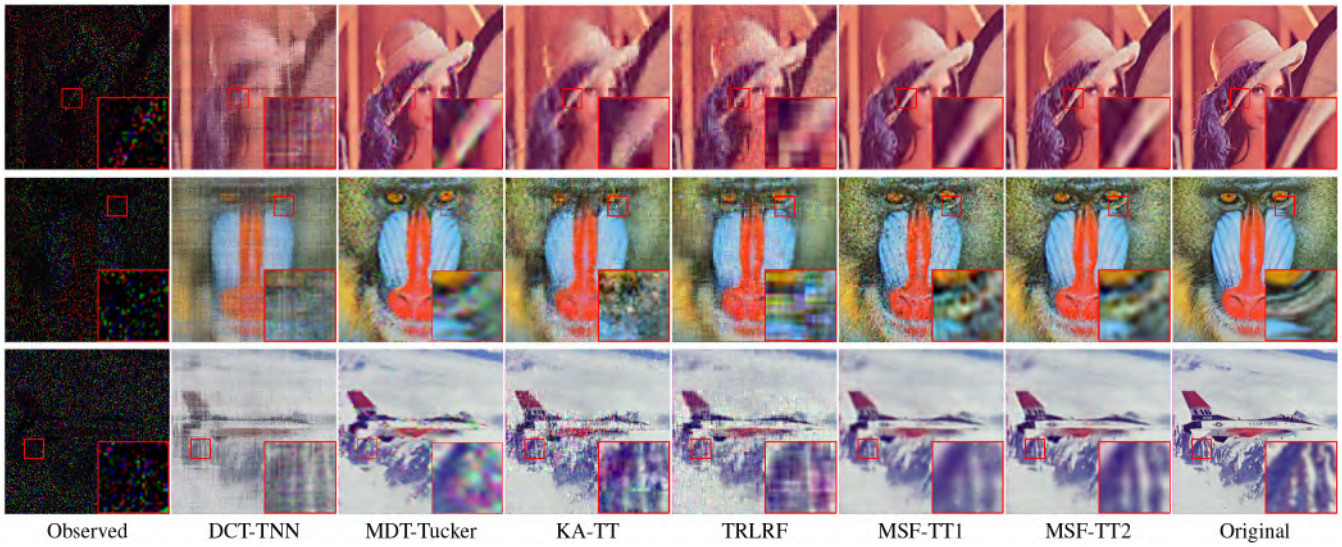


Fig. 2. Recovered images by different methods on color images with SR = 10%. From left to right: the observed images, the recovered images by DCT-TNN, MDT-Tucker, KA-TT, TRLRF, MSF-TT1, and MSF-TT2, and the original images. From top to bottom: *Lena*, *Baboon*, and *Airplane*.

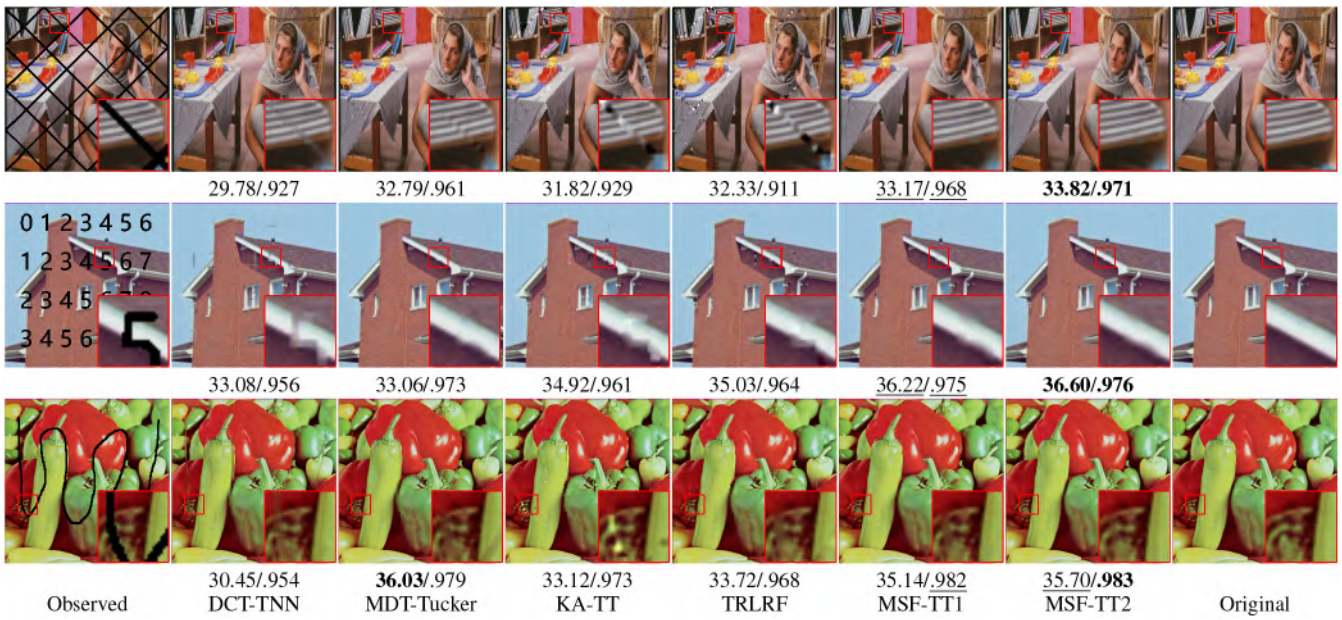


Fig. 3. Recovered images by different methods on color images with different structure missing. From left to right: the observed images, the recovered images by DCT-TNN, MDT-Tucker, KA-TT, TRLRF, MSF-TT1, and MSF-TT2, and the original images. From top to bottom: *Barbara*, *House*, and *Peppers*. PSNR (left) and SSIM (right) are displayed below the image.

cases, our methods significantly outperform other competing methods for the evaluation metrics, for example, on *Barbara* with 10% SR, MSF-TT1 and MSF-TT2 gain around 1.0 and 2.0 dB beyond the third-best method TRLRF in terms of PSNR. Moreover, we can observe that the performance of MSF-TT2 is better than that of MSF-TT1, because the non-convex MSF-TT is more optimal to characterize the MSFs tensor.

Fig. 2 shows the recovered images by different methods on *Lena*, *Baboon*, and *Airplane* with SR = 10%. We can observe that our methods recover the main structures and preserve abundant details well. From the zoomed-in area, we can clearly see that the results of other competing methods lose some details. Notably, there exist apparent block artifacts in the results of KA-TT. The reason is that KA destroys the spatial structure of images. As a comparison, MSF-TT1 and MSF-TT2 both preserve the details. The visual performances of MSF-TT1 and MSF-TT2 are close.

Structural Missing: We test three color images with different structural missing. Specifically, *Barbara*, *House*, and *Peppers* are destroyed by intersecting slices, texts, and random curves, respectively. Fig. 3 shows the visual and quantitative performance by different methods in the structural missing case. In quantitative performance, our methods rank first and second place in terms of the PSNR and SSIM in most cases. In visual quality, MDT-Tucker achieves good visual performance,

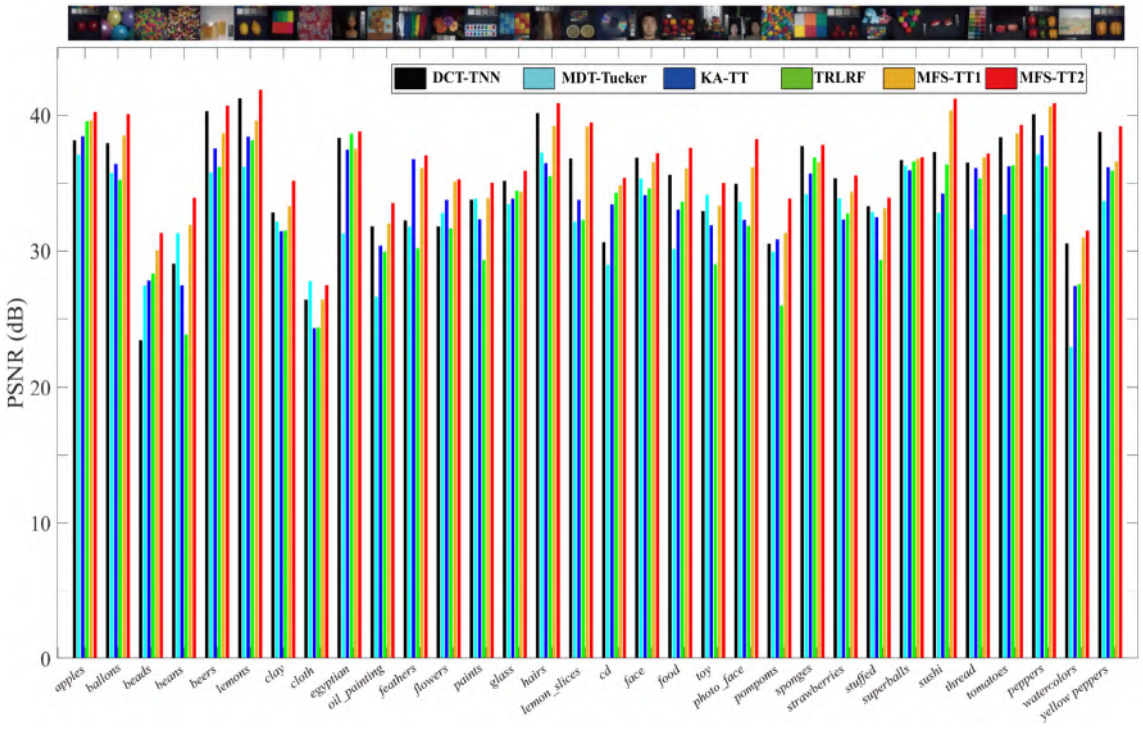


Fig. 4. PSNR values of the recovered images by different methods on MSIs with $SR = 10\%$.

since MDT-Tucker uses duplicated transform to capture some delay/shift-invariant structure of images for missing slices. On recovered *Barbara*, we can see the apparent shadow in the missing area in compared methods. However, our methods produce abundant details in the missing area.

B. MSIs

In this section, we test the proposed methods on MSIs, which have abundant MSFs information in the spectrum. This experiment employs the CAVE MSI data,[2] which contains 32 images. The spatial resolution is resized to 256×256. The SR cases are shown as 10%, 20%, and 30%, respectively. For efficiency, we consider the filter bank as *Linear* and level $= 2$.

Fig. 4 shows the PSNR values of the recovered 32 images by different methods on MSIs with SR $= 10\%$. We can see that MSF-TT2 (red column) achieves the highest values in most cases. In Table IV, we report the quantitative comparisons of competing methods on three MSIs with different SRs. We clearly see that the quantitative performance of our methods is significantly superior to that of the compared methods. For instance, on beads with SR $= 10\%$, the proposed methods achieve around 0.9 in SSIM, which enhances around 0.1 beyond MDT-Tucker, KA-TT, and TRLRF. Fig. 5 displays the spectrum curves of the recovery results by different methods, which illustrates that the proposed methods outperform compared methods in spectral fidelity.

To evaluate the visual quality, we select the 1st, 2nd, and 31st bands to construct the pseudocolor images in Fig. 6. We can observe that DCT-TNN, MDT-Tucker, KA-TT, TRLRF, and the proposed methods obtain similar visual performance and all recover the main structures of images. Zooming-in the results, we see that DCT-TNN and TRLRF cause details losing in some cases, and there exist block artifacts in some images by KA-TT. It is clear that MSF-TT1 and MSF-TT2 both bring significant recovery performance and precisely protect the details, which demonstrates that the proposed methods fully exploit the MSFs information.

[2]Available at http://www.cs.columbia.edu/CAVE/databases/multispectral.

TABLE IV
QUANTITATIVE EVALUATION OF COMPETING METHODS ON MSIS WITH DIFFERENT SRS. THE BEST VALUES AND THE SECOND-BEST VALUES ARE HIGHLIGHTED BY BOLD FONTS AND UNDERLINES, RESPECTIVELY

Images	SR / Method	10% MPSNR	10% MSSIM	20% MPSNR	20% MSSIM	30% MPSNR	30% MSSIM
Beads	DCT-TNN	23.41	0.659	28.72	0.855	32.93	0.933
	MDT-Tucker	27.46	0.839	32.55	0.921	34.23	0.948
	KA-TT	27.82	0.791	32.71	0.902	34.60	0.936
	TRLRF	28.31	0.798	32.80	0.909	34.27	0.925
	MSF-TT1	30.05	0.916	34.57	0.965	37.02	0.966
	MSF-TT2	**31.30**	**0.922**	**35.91**	**0.966**	**37.62**	**0.978**
Feathers	DCT-TNN	32.25	0.902	37.71	0.965	41.65	0.984
	MDT-Tucker	31.81	0.923	37.55	0.976	39.65	0.978
	KA-TT	36.77	0.946	39.52	0.966	41.08	0.975
	TRLRF	30.19	0.792	36.44	0.972	41.64	0.991
	MSF-TT1	36.09	0.955	40.41	0.980	45.11	**0.992**
	MSF-TT2	**37.05**	**0.967**	**42.34**	**0.982**	**45.52**	**0.992**
Flowers	DCT-TNN	31.82	0.906	37.44	0.970	42.24	0.979
	MDT-Tucker	32.79	0.944	37.55	0.975	40.72	0.987
	KA-TT	33.73	0.912	39.44	0.976	41.24	0.983
	TRLRF	31.68	0.871	36.93	0.975	40.54	0.979
	MSF-TT1	35.11	0.958	39.82	0.983	43.99	0.993
	MSF-TT2	**35.27**	**0.953**	**40.57**	**0.987**	**44.13**	**0.994**

C. Color Videos

We test the proposed methods on color videos to evaluate the recovery performance on higher-dimensional tensors.

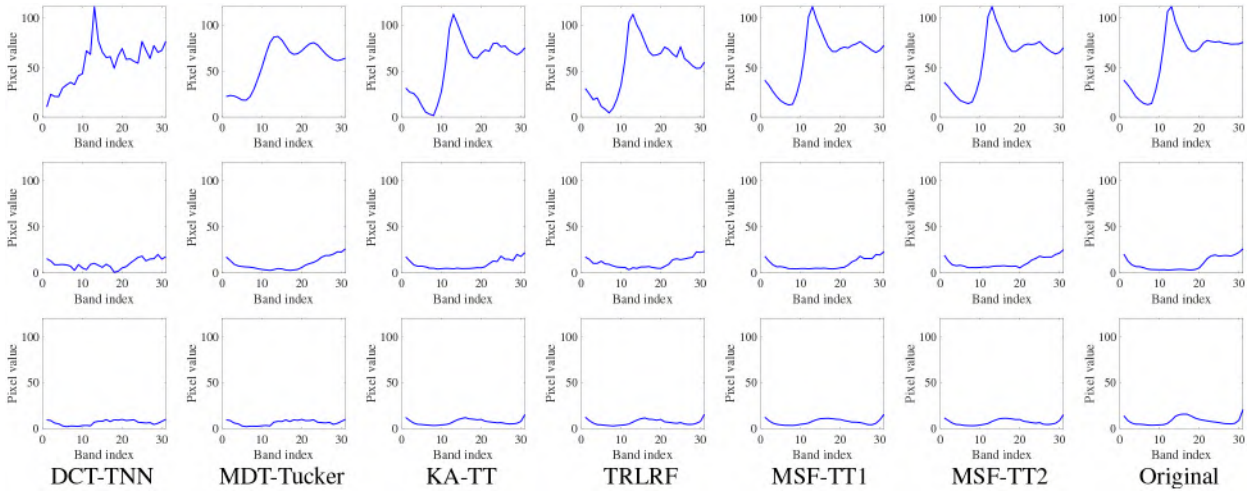


Fig. 5. Spectrum curves of the recovery results by different methods at spatial location (200, 200). From left to right: the results from DCT-TNN, MDT-Tucker, KA-TT, TRLRF, MSF-TT1, MSF-TT2, and the original images. From top to bottom: *Beads*, *Feathers*, and *Flowers* with $SR = 10\%$.

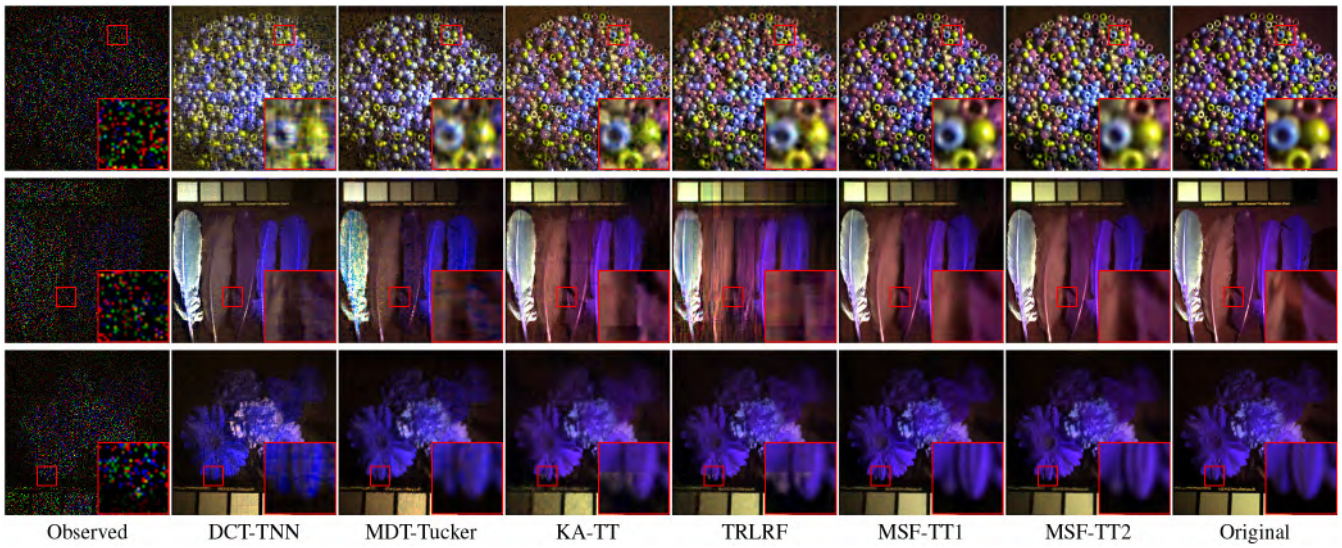


Fig. 6. Recovered images by different methods on MSIs with SR = 10%. From left to right: the observed images, the recovered images by DCT-TNN, MDT-Tucker, KA-TT, TRLRF, MSF-TT1, and MSF-TT2, and the original images. From top to bottom: *Beads*, *Feathers*, and *Flowers*.

Three color videos,[3] including *Football*, *Susie*, and *Bus*, are employed in this experiment. The color videos are higher-dimensional tensors of the size $256 \times 256 \times 3 \times 20$, where the fourth dimension is the temporal variable. We display challenging SR cases as 10%, 20%, and 30%, respectively. Since t-SVD is designed for the 3-D tensor, the temporal mode needs to be merged with the color channel to form a third-order tensor in DCT-TNN. For efficiency, we set the filter bank as *Linear* and level $= 2$.

In Table V, we report the quantitative comparisons of competing methods on color videos with different SRs. It is clear that the proposed methods achieve higher values in quantitative metrics. On *Susie* with SR $= 10\%$, the values of MSSIM are beyond 0.9 by the proposed methods, which means that the recovered video is close to the original one. Fig. 7 shows the PSNR values of each frame of the tested videos. We can see that MSF-TT1 (green lines) and MSF-TT2 (red lines) are undoubtedly superior to compared methods.

To evaluate the visual performance, we show one frame of each video in Fig. 8. The results of our methods are superior to those of the other compared methods. For instance, on *Football*, MSF-TT2 preserves the abundant details from the zoomed-in areas. On *Susie*, the recovered images by our methods are closer to the original images, which agrees with the quantitative performance in Table V. On *Bus*, MSF-TT1 and MST-TT2 are close in visual performance and both outperform other methods' results. These demonstrate that the proposed methods can be applicable for higher-dimensional tensor data.

D. Discussion

In this section, we further discuss some important aspects of the proposed methods.

[3]Available at http://trace.eas.asu.edu/yuv/.

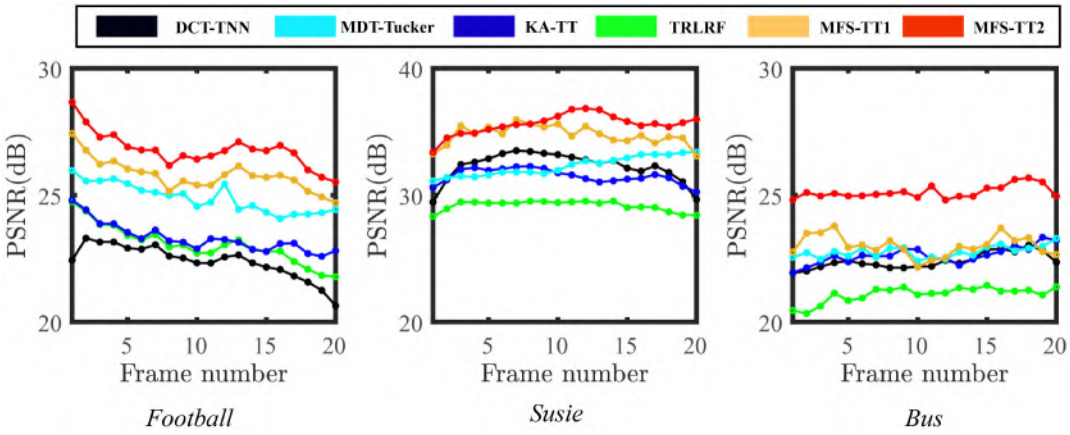


Fig. 7. PSNR values in each frame by different methods on color videos with $SR = 10\%$.

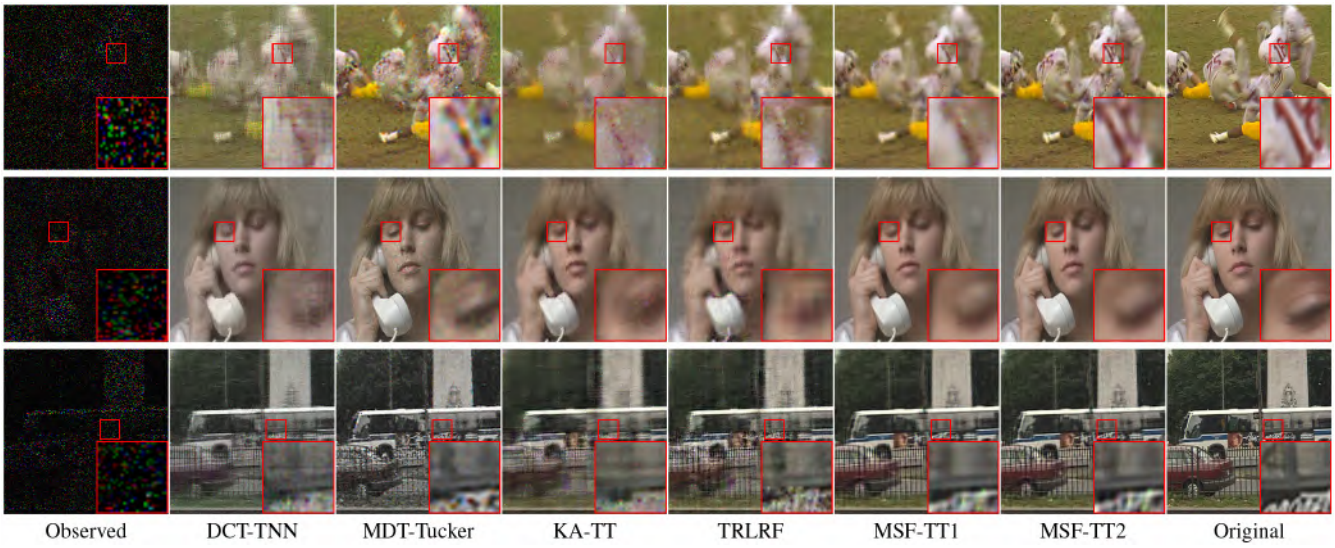


Fig. 8. Recovered images by different methods on color videos with SR = 10%. From left to right: the observed images, the recovered images by DCT-TNN, MDT-Tucker KA-TT, TRLRF, MSF-TT1, and MSF-TT2, and the original images. From top to bottom: *Football*, *Susie*, and *Bus*.

The Subsequent Image Recognition: Here, we demonstrate that the feature tensor not only helps to recover the missing values but also benefits subsequent image applications. We mainly test the performance of the proposed methods on image recognition. In this experiment, we employ a plain classifier for image recognition, that is, nearest neighbor [61], which can be also replaced by any other classifiers, such as the support vector machine (SVM) or the deep neural network (DNN). The *Yale B* face data[4] is employed with the size $48 \times 42 \times 64 \times 38$, where the 3rd dimension has 64 images of one person and the 4th dimension contains 38 different individuals.

Gabor filters use the complex-valued system via local Fourier analysis [53], [55], which has better orientation selectivity compared with the real-valued system (e.g., piecewise linear framelet [62]). Since features of faces (e.g., eyebrows and lips) have natural orientations, Gabor filters can more exactly characterize features of faces. Thus, we set the filter bank as Gabor filters of our methods for face image recovery and use the recovered multiscale Gabor feature tensor to the face image recognition. For the other competing methods, we first extract the Gabor features from their recovered face images and then test by the nearest neighbor. SRs are set as 10%, 20%, and 30%, respectively. In image recognition, we evaluate the performance by the recognition accuracy [0%, 100%]. We set the parameter *rank* (R) as 30 and 60 in the recognition system, which is the dimension of the subspace of the nearest neighbor algorithm.

In Table VI, we report the quantitative evaluation of competing methods on face data with different SRs. In most cases, our methods outperform compared methods in numerical value. In Table VII, we report the recognition accuracy of competing methods on the recovered *Yale B* data with different SRs. It is clear that our methods rank first and second place in terms of accuracy. Notably, with SR = 10%, MSF-TT2 achieves around 3% gain in accuracy beyond the third best-performed method KA-TT. Fig. 9 shows one recovered face by different methods with SR = 10%. Our methods significantly outperform

[4]Available at http://cvc.yale.edu/projects/yalefacesB/yalefacesB.html.

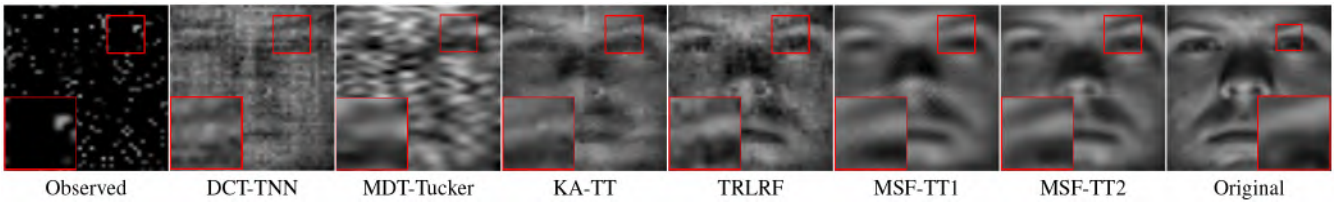


Fig. 9. Recovered images by different methods on face data with $SR = 10\%$. From left to right: the observed images, the recovered images by DCT-TNN, MDT-Tucker, KA-TT, TRLRF, MSF-TT1, and MSF-TT2, and the original images.

TABLE V
QUANTITATIVE EVALUATION OF COMPETING METHODS ON COLOR VIDEOS WITH DIFFERENT SRS. THE BEST VALUES AND THE SECOND-BEST VALUES ARE HIGHLIGHTED BY BOLD FONTS AND UNDERLINES, RESPECTIVELY

Videos	SR / Method	10% MPSNR	10% MSSIM	20% MPSNR	20% MSSIM	30% MPSNR	30% MSSIM
Football	DCT-TNN	22.39	0.472	24.39	0.624	26.51	0.734
	MDT-Tucker	24.90	0.601	27.13	0.768	29.50	0.831
	KA-TT	22.96	0.521	26.05	0.725	27.33	0.807
	TRLRF	23.06	0.536	25.20	0.713	27.10	0.794
	MSF-TT1	25.78	0.655	28.25	0.802	31.24	0.902
	MSF-TT2	**26.38**	**0.703**	**29.73**	**0.861**	**32.88**	**0.930**
Susie	DCT-TNN	32.30	0.871	33.02	0.881	35.58	0.924
	MDT-Tucker	32.21	0.895	37.21	0.943	37.92	0.955
	KA-TT	31.50	0.832	33.64	0.924	36.80	0.956
	TRLRF	29.16	0.786	32.88	0.917	35.99	0.952
	MSF-TT1	34.74	0.924	35.63	0.938	38.67	0.964
	MSF-TT2	**35.49**	**0.935**	**37.32**	**0.953**	**39.87**	**0.971**
Bus	DCT-TNN	22.40	0.557	23.12	0.612	25.46	0.716
	MDT-Tucker	22.77	0.643	23.97	0.722	25.99	0.792
	KA-TT	22.63	0.497	26.12	0.709	27.18	0.814
	TRLRF	21.09	0.564	26.46	0.679	27.09	0.709
	MSF-TT1	23.01	0.668	26.34	0.787	27.97	0.865
	MSF-TT2	**25.12**	**0.748**	**27.64**	**0.804**	**29.07**	**0.915**

TABLE VI
QUANTITATIVE EVALUATION OF COMPETING METHODS ON FACE DATA WITH DIFFERENT SRS. THE BEST VALUES AND THE SECOND-BEST VALUES ARE HIGHLIGHTED BY BOLD FONTS AND UNDERLINES, RESPECTIVELY

Face Data	SR / Method	10% MPSNR	10% MSSIM	20% MPSNR	20% MSSIM	30% MPSNR	30% MSSIM
Yale B	DCT-TNN	19.57	0.501	21.82	0.624	23.81	0.713
	MDT-Tucker	19.25	0.497	22.89	0.612	23.27	0.698
	KA-TT	22.29	0.546	25.32	0.658	27.86	0.681
	TRLRF	21.83	0.525	22.36	0.630	26.39	0.674
	MSF-TT1(Gabor)	22.89	**0.671**	25.64	0.784	27.50	0.831
	MSF-TT2(Gabor)	**23.62**	0.620	**26.60**	**0.806**	**27.93**	**0.868**

TABLE VII
RECOGNITION ACCURACY OF COMPETING METHODS ON THE RECOVERED YALE B DATA WITH DIFFERENT SRS. THE BEST VALUES AND THE SECOND-BEST VALUES ARE HIGHLIGHTED BY BOLD FONTS AND UNDERLINES, RESPECTIVELY

Data	SR / Method	10% R=30	10% R=60	20% R=30	20% R=60	30% R=30	30% R=60
Yale B	DCT-TNN	55.732%	66.283%	60.512%	71.628%	62.788%	73.432%
	MDT-Tucker	54.231%	66.236%	62.241%	71.021%	61.251%	72.151%
	KA-TT	57.957%	66.036%	63.510%	74.095%	64.135%	75.480%
	TRLRF	56.152%	65.296%	59.752%	69.613%	60.896%	71.382%
	MSF-TT1(Gabor)	60.379%	68.298%	63.741%	74.145%	64.186%	75.658%
	MSF-TT2(Gabor)	**61.083%**	**69.367%**	**63.792%**	**74.298%**	**64.267%**	**75.754%**

the other competing methods in the face features protection, which agrees with the recognition performance in Table VII. These show that our methods not only recover images but also effectively help the subsequent image application.

Influence of Filter Bank and Level: We discuss the influence of filter bank and level settings, which determine the feature tensor. We set the filter bank as *Haar*, *Linear*, and *Cubic*, and level as 1, 2 , and 3. Color image *Baboon* with SR = 10% is selected for this discussion, and the size of the image is ranged in $\{64 \times 64 \times 3, 128 \times 128 \times 3, 256 \times 256 \times 3\}$.

Table VIII reports the PSNR, SSIM, and running time (in seconds) on *Baboon* with different sizes. In most cases, the performance of MSF-TT1 and MST-TT2 is about the same, and the results by *Linear* and *Cubic* rank first and second place in terms of the PSNR and SSIM. More levels do not mean better quantitative performance. We can see that the filter bank and level setting also affect the running time, for example, more levels mean more running time in general.

As the size increases, MSF-TT2 has a significant advantage on speed, because the computational complexity of MSF-TT2 is lower than that of MSF-TT1. For the tradeoff between effectiveness and efficiency, we suggest the filter bank as *Linear* and level as 1 or 2.

Influence of the Mode Order: Since the tensor mode order has an effect on the TT-rank minimization, we discuss the influence of mode order of the MSF tensor. In the above numerical experiments, we mainly consider the mode order $(n_1, \ldots, n_j, f_1, \ldots, f_j, L)$ of $\mathcal{X}_{\mathbf{W}}$, which is a natural mode order intuitively from Fig. 1. Now, we mainly consider two classes of mode order: 1) the grouped similar mode $(n_1, \ldots, n_j, f_i, \ldots, f_k, L)$ and 2) the separated similar mode $(n_1, f_i, \ldots, n_j, f_k, L)$. *Beads* is employed with SR = 10%. We set the filter bank as *Linear* and level = 2. The quantitative evaluation of the MSF-TT2 is reported in Table IX for the two classes of mode order. We can see that results of the grouped similar mode generally outperform results of the separated similar mode in most cases.

The Comparison Between Different Tensorizations: The tensorization is usually considered for TT decomposition to construct a higher-order tensor for better results in image recovery [21], [49]. The previous tensorizations are on the pixel level [21], [49], and the proposed tensorization is on the feature level, which can exploit the MSFs of multidimensional images to benefit the image recovery and subsequent image applications. Since TT decomposition is especially suitable for high-order tensors for the computational and storage cost, we consider TT decomposition for the resulting high-order MSF tensor $\mathcal{X}_{\mathbf{W}} \in \mathbb{R}^{n_1 \times \cdots \times n_j \times f_1 \cdots f_j \times L}$. Here, we compare different tensorizations for TT decomposition (TMac-TT). The comparisons (e.g., the computational cost and storage cost) of different tensorizations are summarized in Table I. We can

TABLE VIII
QUANTITATIVE EVALUATION OF COMPETING METHODS ON *Baboon* WITH DIFFERENT SIZES. THE BEST VALUES AND THE SECOND-BEST VALUES ARE HIGHLIGHTED BY BOLD FONTS AND UNDERLINES, RESPECTIVELY

Framelet	Level	Size / Method	$64 \times 64 \times 3$ PSNR	SSIM	Time	$128 \times 128 \times 3$ PSNR	SSIM	Time	$256 \times 256 \times 3$ PSNR	SSIM	Time
Haar	1	MSF-TT1	18.55	0.516	47.44	19.78	0.488	137.29	19.69	0.384	727.69
		MSF-TT2	19.37	0.491	**27.15**	21.26	0.502	**76.56**	20.93	0.427	**379.19**
	2	MSF-TT1	18.07	0.504	193.63	19.79	0.454	208.09	19.94	0.415	1443.17
		MSF-TT2	18.65	0.524	58.78	20.11	0.471	110.79	21.07	0.441	584.12
	3	MSF-TT1	16.86	0.402	133.91	18.95	0.406	447.54	19.41	0.384	1754.34
		MSF-TT2	17.68	17.68	74.72	19.44	0.434	230.97	20.33	0.391	662.57
Linear	1	MSF-TT1	18.71	0.542	139.12	22.29	0.537	237.17	21.37	0.477	1527.58
		MSF-TT2	19.68	0.547	71.17	**22.62**	0.569	113.57	21.84	0.518	863.18
	2	MSF-TT1	19.07	0.538	286.11	18.95	0.471	597.32	21.08	0.512	2139.51
		MSF-TT2	19.62	0.548	103.91	21.22	0.517	345.12	**22.26**	**0.561**	1176.48
	3	MSF-TT1	17.50	0.403	450.31	19.84	0.327	442.35	20.58	0.478	1936.21
		MSF-TT2	17.65	0.412	211.47	20.13	0.409	237.26	21.96	0.516	1343.15
Cubic	1	MSF-TT1	19.42	0.532	688.79	21.23	0.501	1538.12	20.95	0.513	3872.65
		MSF-TT2	**20.97**	**0.587**	406.97	22.50	**0.572**	873.22	22.04	0.557	2187.42
	2	MSF-TT1	18.38	0.515	1340.24	19.95	0.473	1921.71	20.62	0.521	3351.57
		MSF-TT2	19.43	0.484	593.99	21.13	0.521	1012.32	22.05	0.570	2607.17
	3	MSF-TT1	17.52	0.395	1294.05	19.65	0.436	1974.32	19.79	0.486	4532.87
		MSF-TT2	19.08	0.563	449.06	20.16	0.483	1245.14	21.86	0.532	2843.27

TABLE IX
QUANTITATIVE EVALUATION OF MSF-TT2 ON *Beads* WITH DIFFERENT ORDERS. THE BEST VALUES ARE HIGHLIGHTED BY BOLD FONTS

Grouped	*Beads* MPSNR	MSSIM	Separated	*Beads* MPSNR	MSSIM
$(n_1,n_2,n_3,f_1,f_2,f_3,L)$	**25.21**	**0.835**	$(n_1,f_1,n_2,f_2,n_3,f_3,L)$	23.53	0.686
$(n_1,n_2,n_3,f_1,f_3,f_2,L)$	24.68	0.834	$(n_1,f_1,n_2,f_3,n_3,f_2,L)$	22.04	0.700
$(n_1,n_2,n_3,f_2,f_1,f_3,L)$	24.42	0.827	$(n_1,f_2,n_2,f_1,n_3,f_3,L)$	24.53	0.813
$(n_1,n_2,n_3,f_2,f_3,f_1,L)$	24.96	0.830	$(n_1,f_2,n_2,f_3,n_3,f_1,L)$	23.50	0.774
$(n_1,n_2,n_3,f_3,f_1,f_2,L)$	24.84	0.829	$(n_1,f_3,n_2,f_1,n_3,f_2,L)$	24.38	0.792
$(n_1,n_2,n_3,f_3,f_2,f_1,L)$	24.26	0.807	$(n_1,f_3,n_2,f_2,n_3,f_1,L)$	23.79	0.780

TABLE X
QUANTITATIVE EVALUATION OF KA-TT AND MSF-TT ON *Lena*, *Beads*, AND *Football*

Method	*Lena* PSNR	Time	Iteration	*Beads* PSNR	Time	Iteration	*Football* PSNR	Time	Iteration
KA-TT	23.46	17.69	1001	27.82	88.65	819	22.63	147.29	503
MSF-TT	26.43	345.12	310	31.30	473.61	101	25.12	970.61	101

see that MSF increases computational cost and storage cost for TT decomposition compared with KA. Here, we compare the results of KA-TT and MSF-TT. *Lena*, *Beads*, and *Football* with SR $= 10\%$ are selected for this discussion. However, we can see that although MSF-TT needs high calculation time, the recovered results by MSF-TT have significant improvements compared with those of KA from Table X.

Convergence Analysis: Here, we further analyze the convergence of the solving algorithm. Fig. 10 displays RelCha (31) values curves of MSF-TT2 for *Baboon* and *Beads* with $SR = 10\%$. We can see that the RelCha values quickly converge to zero, which numerically validates Theorem 1.

V. CONCLUSION

In this article, to improve the image recovery and help the subsequent image application, we suggested a novel MSF tensorization by extracting the MSFs of multidimensional images. The resulting MSF tensor is high dimensions and has the low-rank structure, which can be better exploited by the TT decomposition. Thus, by the resulting MSF tensor, we proposed convex MSF-TT and nonconvex MSF-TT,

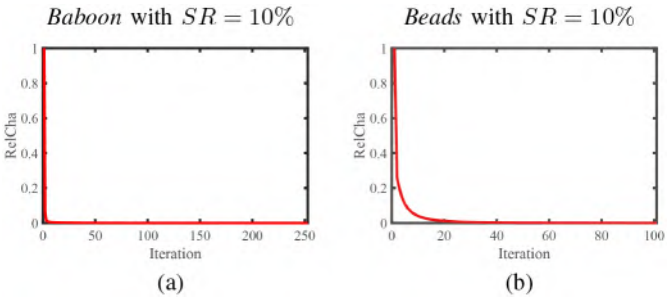


Fig. 10. RelCha values curves with respect to iterations for *Baboon* and *Beads* with $SR = 10\%$.

which can conjointly recover images on the feature level and obtain the feature tensor for subsequent image applications. Then, we developed the ADMM and PAM to solve corresponding models. Moreover, we established the theoretical guarantee of convergence for the PAM algorithm. To verify the effectiveness of the proposed methods, we conducted comprehensive real-data experiments on image recovery and recognition. Examples of real-world imaging data shown that the proposed methods outperformed many compared methods on recovery performance and the recovered feature tensor can help image recognition. In summary, the proposed methods not only improved the recovery quality but also benefited the subsequent image application.

APPENDIX

For the proof of Theorem 1, the following conditions are required [63].

1) $f(\mathbf{E}, \mathbf{Z}, \mathcal{X})$ is a proper lower semicontinuous function.
2) $f(\mathbf{E}, \mathbf{Z}, \mathcal{X})$ has the K-Ł property at each $\{\mathbf{E}^t, \mathbf{Z}^t, \mathcal{X}^t\}_{t\in\mathbb{N}}$.
3) The bounded $\mathbf{E}^t$, $\mathbf{Z}^t$, and $\mathcal{X}^t$ satisfy the sufficient decrease condition and relative error condition.

Proof: The three conditions are dividedly proven as follows.

First, we show that $f(\mathbf{E}, \mathbf{Z}, \mathcal{X})$ is a proper lower semicontinuous function. It is easy to verify that the sum of the Frobenius norm is a C^1 function with locally Lipschitz continuous

gradient. δ_R and δ_D are proper and lower semicontinuous. Thus, $f(\mathbf{E}, \mathbf{Z}, \mathcal{X})$ is a proper lower semicontinuous function.

Second, we show that $f(\mathbf{E}, \mathbf{Z}, \mathcal{X})$ has the K-Ł property at each $\{\mathbf{E}^t, \mathbf{Z}^t, \mathcal{X}^t\}_{t\in\mathbb{N}}$ by proving that $f(\mathbf{E}, \mathbf{Z}, \mathcal{X})$ is a semialgebraic function. Since the low-rank matrices set is semialgebraic, the subsets $\mathcal{R}$ is semialgebraic. Indicator functions of semialgebraic sets are semialgebraic functions [64]. δ_D and the Frobenius norm are semialgebraic functions. Thus, $f(\mathbf{E}, \mathbf{Z}, \mathcal{X})$ is a semialgebraic function. Since the semialgebraic real-valued function is the K-Ł function [64], $f(\mathbf{E}, \mathbf{Z}, \mathcal{X})$ has the K-Ł property at each $\{\mathbf{E}^t, \mathbf{Z}^t, \mathcal{X}^t\}_{t\in\mathbb{N}}$.

Third, we show that the bounded $\mathbf{E}^t$, $\mathbf{Z}^t$, and $\mathcal{X}^t$ satisfy the sufficient decrease condition and relative error condition.

Lemma 1 (Sufficient Decrease): Assume that $f(\mathbf{E}, \mathbf{Z}, \mathcal{X})$ is a $\mathcal{C}^1$ function with locally Lipschitz continuous gradient and $\rho > 0$. Let $\{\mathbf{E}^t, \mathbf{Z}^t, \mathcal{X}^t\}_{t\in\mathbb{N}}$ is generated by Algorithm 2. Then

$$\begin{aligned}
&f\left(\boldsymbol{E}^{t+1}, \mathbf{Z}^t, \mathcal{X}^t\right) + \frac{\rho}{2}\left\|\mathbf{E}^{t+1} - \mathbf{E}^t\right\|_F^2 \le f\left(\mathbf{E}^t, \mathbf{Z}^t, \mathcal{X}^t\right)\\
&f\left(\mathbf{E}^{t+1}, \mathbf{Z}^{t+1}, \mathcal{X}^t\right) + \frac{\rho}{2}\left\|\mathbf{Z}^{t+1} - \mathbf{Z}^t\right\|_F^2 \le f\left(\mathbf{E}^{t+1}, \mathbf{Z}^t, \mathcal{X}^t\right)\\
&f\left(\mathbf{E}^{t+1}, \mathbf{Z}^{t+1}, \mathcal{X}^{t+1}\right) + \frac{\rho}{2}\left\|\mathcal{X}^{t+1} - \mathcal{X}^t\right\|_F^2\\
&\quad\le f\left(\mathbf{E}^{t+1}, \mathbf{Z}^{t+1}, \mathcal{X}^t\right).
\end{aligned} \tag{32}$$

■

Proof of Lemma 1: When $\mathbf{E}^{t+1}$, $\mathbf{Z}^{t+1}$, and $\mathcal{X}^{t+1}$ are optimal solutions of M_1, M_2, and M_3, $\delta_R(\mathbf{E}) = 0$, $\delta_R(\mathbf{Z}) = 0$, and $\delta_D(\mathcal{X}) = 0$. By the definition of M_1, M_2, and M_3, we clearly have that

$$\begin{aligned}
&f\left(\mathbf{E}^{t+1}, \mathbf{Z}^t, \mathcal{X}^t\right) + \frac{\rho}{2}\left\|\mathbf{E}^{t+1} - \mathbf{E}^t\right\|_F^2\\
&\quad= M_1\left(\mathbf{E}^{t+1}|\mathbf{E}^t\right) \le M_1\left(\mathbf{E}^t|\mathbf{E}^t\right) = f\left(\mathbf{E}^t, \mathbf{Z}^t, \mathcal{X}^t\right)\\
&f\left(\mathbf{E}^{t+1}, \mathbf{Z}^{t+1}, \mathcal{X}^t\right) + \frac{\rho}{2}\left\|\mathbf{Z}^{t+1} - \mathbf{Z}^t\right\|_F^2\\
&\quad= M_2\left(\mathbf{Z}^{t+1}|\mathbf{Z}^t\right) \le M_2\left(\mathbf{Z}^t|\mathbf{Z}^t\right) = f\left(\mathbf{E}^{t+1}, \mathbf{Z}^t, \mathcal{X}^t\right)\\
&f\left(\mathbf{E}^{t+1}, \mathbf{Z}^{t+1}, \mathcal{X}^{t+1}\right) + \frac{\rho}{2}\left\|\mathcal{X}^{t+1} - \mathcal{X}^t\right\|_F^2\\
&\quad= M_3\left(\mathcal{X}^{t+1}|\mathcal{X}^t\right) \le M_3\left(\mathcal{X}^t|\mathcal{X}^t\right)\\
&\quad= f\left(\mathbf{E}^{t+1}, \mathbf{Z}^{t+1}, \mathcal{X}^t\right).
\end{aligned} \tag{33}$$

The proof of the sufficient decrease condition is completed.

Lemma 2 (Relative Error): Let $\{\mathbf{E}^t, \mathbf{Z}^t, \mathcal{X}^t\}_{t\in\mathbb{N}}$ is bounded and generated by Algorithm 2, ∇Q is Lipschitz continuous, and $\rho > 0$. Then, there exist V_1^{t+1}, V_2^{t+1}, and V_3^{t+1}, which satisfy

$$\begin{aligned}
&\left\|V_1^{t+1} + \nabla_{\mathbf{E}} Q\left(\mathbf{E}^{t+1}, \mathbf{Z}^t, \mathcal{X}^t\right)\right\|_F^2 \le \rho^2\left\|\mathbf{E}^{t+1} - \mathbf{E}^t\right\|_F^2\\
&\left\|V_2^{t+1} + \nabla_{\mathbf{Z}} Q\left(\mathbf{E}^{t+1}, \mathbf{Z}^{t+1}, \mathcal{X}^t\right)\right\|_F^2 \le \rho^2\left\|\mathbf{Z}^{t+1} - \mathbf{Z}^t\right\|_F^2\\
&\left\|V_3^{t+1} + \nabla_{\mathcal{X}} Q\left(\mathbf{E}^{t+1}, \mathbf{Z}^{t+1}, \mathcal{X}^{t+1}\right)\right\|_F^2 \le \rho^2\left\|\mathcal{X}^{t+1} - \mathcal{X}^t\right\|_F^2
\end{aligned} \tag{34}$$

where $V_1^{t+1} \in \partial\delta_R(\mathbf{E})$, $V_2^{t+1} \in \partial\delta_R(\mathbf{Z})$, and $V_3^{t+1} \in \partial\delta_D(\mathcal{X})$.

Proof of Lemma 2: Q is Lipschitz continuous on any bounded set, and $\mathbf{E}^{t+1}$, $\mathbf{Z}^{t+1}$, and $\mathcal{X}^{t+1}$ are optimal solutions of M_1, M_2, and M_3. Thus, there exists $\rho > 0$ such that for each subproblem of $\mathbf{E}_k$ and $\mathbf{Z}_k$

$$\begin{aligned}
&0 \in \partial\delta_R(\mathbf{E}_k) - \alpha_k\left(\mathbf{E}_k\mathbf{Z}_k^t - \mathbf{X}_{\mathbf{W}[k]}^t\right)\mathbf{Z}_k^{tT} + \rho\left(\mathbf{E}_k - \mathbf{E}_k^t\right)\\
&0 \in \partial\delta_R(\mathbf{Z}_k) - \alpha_k\mathbf{E}^{t+1T}\left(\mathbf{E}_k^{t+1}\mathbf{Z}_k - \mathbf{X}_{\mathbf{W}[k]}^t\right) + \rho\left(\mathbf{Z}_k - \mathbf{Z}_k^t\right).
\end{aligned} \tag{35}$$

The multidimensional framelet transform and the Frobenius norm are linear. Thus

$$0 \in \partial\delta_D(\mathcal{X}) + \nabla Q_{\mathcal{X}} + \rho\left(\mathcal{X} - \mathcal{X}^t\right). \tag{36}$$

Then, we define that

$$\begin{aligned}
V_{1k}^{t+1} &= \alpha_k\left(\mathbf{E}_k^{t+1}\mathbf{Z}_k^t - \mathbf{X}_{\mathbf{W}[k]}^t\right)\mathbf{Z}_k^{tT}\\
&\quad- \rho\left(\mathbf{E}_k^{t+1} - \mathbf{E}_k^t\right) \in \partial\delta_R(\mathbf{E}_k)\\
V_{2k}^{t+1} &= \alpha_k\mathbf{E}^{t+1T}\left(\mathbf{E}_k^{t+1}\mathbf{Z}_k^{t+1} - \mathbf{X}_{\mathbf{W}[k]}^t\right)\\
&\quad- \rho\left(\mathbf{Z}_k^{t+1} - \mathbf{Z}_k^t\right) \in \partial\delta_R(\mathbf{Z}_k)\\
V_3^{t+1} &= -\nabla Q_{\mathcal{X}^{t+1}} - \rho\left(\mathcal{X}^{t+1} - \mathcal{X}^t\right) \in \partial\delta_D(\mathcal{X}).
\end{aligned} \tag{37}$$

It can be seen that

$$\begin{aligned}
&\left\|V_1^{t+1} + \nabla_{\mathbf{E}^{t+1}} Q\right\|_F^2 \le \rho^2\left\|\mathbf{E}^{t+1} - \mathbf{E}^t\right\|_F^2\\
&\left\|V_2^{t+1} + \nabla_{\mathbf{Z}^{t+1}} Q\right\|_F^2 \le \rho^2\left\|\mathbf{Z}^{t+1} - \mathbf{Z}^t\right\|_F^2\\
&\left\|V_3^{t+1} + \nabla_{\mathcal{X}^{t+1}} Q\right\|_F^2 \le \rho^2\left\|\mathcal{X}^{t+1} - \mathcal{X}^t\right\|_F^2.
\end{aligned}$$

The proof of the relative error condition is completed. Thus, the bounded $\mathbf{E}^t$, $\mathbf{Z}^t$, and $\mathcal{X}^t$ satisfy the sufficient decrease condition and relative error condition.

In summary, the bounded sequence $\{\mathbf{E}^t, \mathbf{Z}^t, \mathcal{X}^t\}$ converges to a critical point of $f(\mathbf{E}, \mathbf{Z}, \mathcal{X})$.

REFERENCES

[1] J. Wang, J. Cai, Y. Shi, and B. Yin, "Incoherent dictionary learning for sparse representation based image denoising," *Appl. Comput. Harmon. Anal.*, vol. 41, no. 1, pp. 94–138, 2016.

[2] Y. Chang, L. Yan, X.-L. Zhao, H. Fang, Z. Zhang, and S. Zhong, "Weighted low-rank tensor recovery for hyperspectral image restoration," *IEEE Trans. Cybern.*, vol. 50, no. 11, pp. 4558–4572, Nov. 2020.

[3] Y.-C. Miao, X.-L. Zhao, X. Fu, and J.-L. Wang, "Hyperspectral denoising using unsupervised disentangled spatio−spectral deep priors," *IEEE Trans. Geosci. Remote Sens.*, early access, Aug. 27, 2020, doi: 10.1109/TGRS.2021.3106380.

[4] T.-X. Jiang, L. Zhuang, T.-Z. Huang, X.-L. Zhao, and J. M. Bioucas-Dias, "Adaptive hyperspectral mixed noise removal," *IEEE Trans. Geosci. Remote Sens.*, early access, Jun. 15, 2021, doi: 10.1109/TGRS.2021.3085779.

[5] L. Lin, K. Wang, D. Meng, W. Zuo, and L. Zhang, "Active self-paced learning for cost-effective and progressive face identification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 1, pp. 7–19, Jan. 2018.

[6] H. Zhang, J. Li, Y. Huang, and L. Zhang, "A nonlocal weighted joint sparse representation classification method for hyperspectral imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2056–2065, Jun. 2014.

[7] H. Zhang, H. Zhai, L. Zhang, and P. Li, "Spectral-spatial sparse subspace clustering for hyperspectral remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3672–3684, Jun. 2016.

[8] Z. Zhang, E. Pasolli, M. M. Crawford, and J. C. Tilton, "An active learning framework for hyperspectral image classification using hierarchical segmentation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 2, pp. 640–654, Feb. 2016.

[9] M. Lan, Y. Zhang, L. Zhang, and B. Du, "Global context based automatic road segmentation via dilated convolutional neural network," *Inf. Sci.*, vol. 515, pp. 156–171, Oct. 2020.

[10] P. Zhou, C. Lu, J. Feng, Z. Lin, and S. Yan, "Tensor low-rank representation for data recovery and clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 5, pp. 1718–1732, May 2021, doi: 10.1109/TPAMI.2019.2954874.

[11] Q. Yuan, L. Zhang, and H. Shen, "Hyperspectral image denoising employing a spectral–spatial adaptive total variation model," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 10, pp. 3660–3677, Oct. 2012.

[12] P. Zhou, C. Lu, Z. Lin, and C. Zhang, "Tensor factorization for low-rank tensor completion," *IEEE Trans. Image Process.*, vol. 27, no. 3, pp. 1152–1163, Mar. 2018.

[13] J. Peng, Q. Xie, Q. Zhao, Y. Wang, L. Yee, and D. Meng, "Enhanced 3DTV regularization and its applications on hsi denoising and compressed sensing," *IEEE Trans. Image Process.*, vol. 29, pp. 7889–7903, 2020.

[14] X.-L. Zhao, W.-H. Xu, T.-X. Jiang, Y. Wang, and M. K. Ng, "Deep plug-and-play prior for low-rank tensor completion," *Neurocomputing*, vol. 400, pp. 137–149, Aug. 2020.

[15] J. K. Choi, B. Dong, and X. Zhang, "An edge driven wavelet frame model for image restoration," *Appl. Comput. Harmon. Anal.*, vol. 48, no. 3, pp. 993–1029, 2020.

[16] N. Wang, S. Ma, J. Li, Y. Zhang, and L. Zhang, "Multistage attention network for image inpainting," *Pattern Recognit*, vol. 106, pp. 1–12, Oct. 2020.

[17] J. L. Wang, T. Z. Huang, X. L. Zhao, T. X. Jiang, and M. K. Ng, "Multi-dimensional visual data completion via low-rank tensor representation under coupled transform," *IEEE Trans. Image Process.*, vol. 30, pp. 3581–3596, 2021.

[18] J. Liu, M. Przemyslaw, W. Peter, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 208–220, Jan. 2013.

[19] T. Xie, S. Li, L. Fang, and L. Liu, "Tensor completion via nonlocal low-rank regularization," *IEEE Trans. Cybern.*, vol. 49, no. 6, pp. 2344–2354, Jun. 2019.

[20] Y. Chen, W. He, N. Yokoya, and T.-Z. Huang, "Hyperspectral image restoration using weighted group sparsity-regularized low-rank tensor decomposition," *IEEE Trans. Cybern.*, vol. 50, no. 8, pp. 3556–3570, Aug. 2020.

[21] J. A. Bengua, H. N. Phien, H. D. Tuan, and M. N. Do, "Efficient tensor completion for color image and video recovery: Low-rank tensor train," *IEEE Trans. Image Process.*, vol. 26, no. 5, pp. 2466–2479, May 2017.

[22] L. Zhang, L. Song, B. Du, and Y. Zhang, "Nonlocal low-rank tensor completion for visual data," *IEEE Trans. Cybern.*, vol. 51, no. 2, pp. 673–685, Feb. 2021.

[23] X. Liu, S. Bourennane, and C. Fossati, "Denoising of hyperspectral images using the PARAFAC model and statistical performance analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 10, pp. 3717–3724, Oct. 2012.

[24] C. J. Hillar and L. Lim, "Most tensor problems are NP-hard," *J. ACM.*, vol. 60, no. 6, p. 45, 2013.

[25] T.-Y. Ji, N. Yokoya, X. X. Zhu, and T.-Z. Huang, "Nonlocal tensor completion for multitemporal remotely sensed images′ inpainting," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 6, pp. 3047–3061, Jun. 2018.

[26] R. Dian, S. Li, L. Fang, T. Lu, and J. M. Bioucas-Dias, "Nonlocal sparse tensor factorization for semiblind hyperspectral and multispectral image fusion," *IEEE Trans. Cybern.*, vol. 50, no. 10, pp. 4469–4480, Oct. 2020.

[27] Y. Xu, R. Hao, W. Yin, and Z. Su, "Parallel matrix factorization for low-rank tensor completion," *Inverse Probl. Imag.*, vol. 9, no. 2, pp. 601–624, 2015.

[28] C. Mu, B. Huang, J. Wright, and D. Goldfarb, "Square deal: Lower bounds and improved relaxations for tensor recovery," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2014, pp. 73–81.

[29] M. E. Kilmer and C. D. Martin, "Factorization strategies for third-order tensors," *Linear Algebra Appl.*, vol. 435, no. 3, pp. 641–658, 2011.

[30] Y.-Y. Liu, X.-L. Zhao, Y.-B. Zheng, T.-H. Ma, and H. Zhang, "Hyperspectral image restoration by tensor fibered rank constrained optimization and plug-and-play regularization," *IEEE Trans. Geosci. Remote Sens.*, early access, Jan. 5, 2021, doi: 10.1109/TGRS.2020.3045169.

[31] Y.-B. Zheng, T.-Z. Huang, X.-L. Zhao, T.-X. Jiang, T.-H. Ma, and T.-Y. Ji, "Mixed noise removal in hyperspectral image via low-fibered-rank regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 1, pp. 734–749, Jan. 2020.

[32] T.-X. Jiang, X.-L. Zhao, H. Zhang, and M. K. Ng, "Dictionary learning with low-rank coding coefficients for tensor completion," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Aug. 31, 2021, doi: 10.1109/TNNLS.2021.3104837.

[33] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, and S. Yan, "Tensor robust principal component analysis with a new tensor nuclear norm," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 4, pp. 925–938, Apr. 2020.

[34] Z. Zhang and S. Aeron, "Exact tensor completion using T-SVD," *IEEE Trans. Signal Process.*, vol. 65, no. 6, pp. 1511–1526, Mar. 2017.

[35] Z. Zhang, G. Ely, S. Aeron, N. Hao, and M. Kilmer, "Novel methods for multilinear data completion and de-noising based on tensor-SVD," in *Proc. IEEE Comput. Vis. Pattern Recognit. (CVPR)*, 2014, pp. 3842–3849.

[36] T.-X. Jiang, T.-Z. Huang, X.-L. Zhao, and L.-J. Deng, "Multi-dimensional imaging data recovery via minimizing the partial sum of tubal nuclear norm," *J. Comput. Appl. Math.*, vol. 372, Jul. 2020, Art. no. 112680.

[37] C. Lu, X. Peng, and Y. Wei, "Low-rank tensor completion with a new tensor nuclear norm induced by invertible linear transforms," in *Proc. IEEE Comput. Vis. Pattern Recognit. (CVPR)*, 2019, pp. 5996–6004.

[38] T.-X. Jiang, M. K. Ng, X.-L. Zhao, and T.-Z. Huang, "Framelet representation of tensor nuclear norm for third-order tensor completion," *IEEE Trans. Image Process.*, vol. 29, pp. 7233–7244, 2020.

[39] I. V. Oseledets, "Tensor-train decomposition," *SIAM J. Sci. Comput.*, vol. 33, no. 5, pp. 2295–2317, 2011.

[40] M. Ding, T.-Z. Huang, T.-Y. Ji, X.-L. Zhao, and J.-H. Yang, "Low-rank tensor completion using matrix factorization based on tensor train rank and total variation," *J. Sci. Comput.*, vol. 81, no. 2, pp. 941–964, 2019.

[41] J.-H. Yang, X.-L. Zhao, T.-Y. Ji, T.-H. Ma, and T.-Z. Huang, "Low-rank tensor train for tensor robust principal component analysis," *Appl. Math. Comput.*, vol. 367, Feb. 2020, Art. no. 124783, doi: 10.1016/j.amc.2019.124783.

[42] M. Che and Y. Wei, "Randomized algorithms for the approximations of tucker and the tensor train decompositions," *Adv. Comput. Math.*, vol. 45, no. 1, pp. 395–428, 2019.

[43] M. Ding, T.-Z. Huang, X.-L. Zhao, M. Ng, and T.-H. Ma, "Tensor train rank minimization with nonlocal self-similarity for tensor completion," *Inverse Probl. Imag.*, vol. 15, no. 3, pp. 475–498, 2021, doi: 10.3934/ipi.2021001.

[44] L. Yuan, C. Li, D. Mandic, J. Cao, and Q. Zhao, "Tensor ring decomposition with rank minimization on latent space: An efficient approach for tensor completion," in *Proc. Conf. Artif. Intell. (AAAI)*, vol. 33, 2019, pp. 9151–9158.

[45] W. Wang, V. Aggarwal, and S. Aeron, "Efficient low rank tensor ring completion," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017, pp. 5698–5706.

[46] J. Yu, C. Li, Q. Zhao, and G. Zhou, "Tensor-ring nuclear norm minimization and application for visual data completion," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, 2019, pp. 3142–3146.

[47] L. Yuan, Q. Zhao, L. Gui, and J. Cao, "High-order tensor completion via gradient-based optimization under tensor train format," *Signal Process. Image Commun.*, vol. 73, pp. 53–61, Apr. 2019.

[48] T. Yokota, B. Erem, S. Guler, S. K. Warfield, and H. Hontani, "Missing slice recovery for tensors using a low-rank model in embedded space," in *Proc. IEEE Comput. Vis. Pattern Recognit. (CVPR)*, 2018, pp. 8251–8259.

[49] F. Sedighin, A. Cichocki, T. Yokota, and Q. Shi, "Matrix and tensor completion in multiway delay embedded space using tensor train, with application to signal reconstruction," *IEEE Signal Process. Lett.*, vol. 27, pp. 810–814, 2020.

[50] T. Yokota and H. Hontani, "Tensor completion with shift-invariant cosine bases," in *Proc. IEEE Asia–Pac. Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPAASC)*, 2018, pp. 1325–1333.

[51] T. Yokota, H. Hontani, Q. Zhao, and A. Cichocki, "Manifold modeling in embedded space: An interpretable alternative to deep image prior," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Dec. 4, 2020, doi: 10.1109/TNNLS.2020.3037923.

[52] J. Xue, Y. Zhao, W. Liao, J. C. Chan, and S. G. Kong, "Enhanced sparsity prior model for low-rank tensor completion," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 11, pp. 4567–4581, Nov. 2020, doi: 10.1109/TNNLS.2019.2956153.

[53] H. Ji, Z. Shen, and Y. Zhao, "Digital Gabor filters do generate MRA-based wavelet tight frames," *Appl. Comput. Harmon. Anal.*, vol. 47, no. 1, pp. 87–108, 2017.

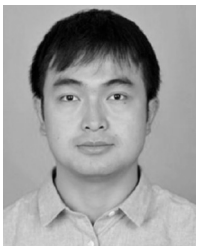
[54] H. Ji, Z. Shen, and Y. Zhao, "Directional frames for image recovery: Multi-scale discrete Gabor frames," *J. Fourier Anal. Appl.*, vol. 23, pp. 729–575, Aug. 2017.
[55] H. Ji, Z. Shen, and Y. Zhao, "Digital Gabor filters with MRA structure," *SIAM J. Multiscale Model Simulat.*, vol. 16, no. 1, pp. 452–476, 2018.
[56] J. Cai, R. H. Chan, and Z. Shen, "A framelet-based image inpainting algorithm," *Appl. Comput. Harmon. Anal.*, vol. 24, no. 2, pp. 131–149, 2008.
[57] J. Cai, H. Ji, C. Liu, and Z. Shen, "Framelet based blind image deblurring from a single image," *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 562–572, Feb. 2012.
[58] H. Attouch, B. Jerome, P. Redont, and A. Soubeyran, "Proximal alternating minimization and projection methods for nonconvex problems. An approach based on the Kurdyka–Lojasiewicz inequality," *Math. Oper. Res.*, vol. 35, no. 2, pp. 438–457, 2010.
[59] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.
[60] A. Horé and D. Ziou, "Image quality metrics: PSNR vs. SSIM," in *Proc. Int. Conf. Pattern Recognit. (ICPR)*, 2010, pp. 23–26.
[61] S. T. Vijay and P. N. Pournami, "Feature based image registration using heuristic nearest neighbour search," in *Proc. Int. Comput. Sci. Eng. Conf. (ICSEC)*, 2018, pp. 1–3.
[62] J. Cai, H. Ji, C. Liu, and Z. Sheng, "Framelet-based blind motion deblurring from a single image," *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 562–572, Feb. 2012.
[63] H. Attouch, J. Bolte, and B. F. Svaiter, "Convergence of descent methods for semi-algebraic and tame problems: Proximal algorithms, forward-backward splitting, and regularized Gauss–Seidel methods," *Math. Program.*, vol. 137, no. 1, pp. 91–129, 2013.
[64] J. Bolte, S. Sabach, and M. Teboulle, "Proximal alternating linearized minimization for nonconvex and nonsmooth problems," *Math. Program.*, vol. 46, no. 1, pp. 459–494, 2014.

Hao Zhang received the M.S. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2021, where he is currently pursuing the Ph.D. degree.

His research interests are modeling and fast algorithms for high-order data recovery based on low rank and sparse prior of tensors.

Xi-Le Zhao received the M.S. and Ph.D. degrees from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2009 and 2012, respectively.

He is currently a Professor with the School of Mathematical Sciences, UESTC. His research interest mainly focuses on model-driven and data-driven methods for image processing problems. His homepage is https://zhaoxile.github.io/.



Tai-Xiang Jiang received the Ph.D. degree in mathematics from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2019.

He was a co-training Ph.D. student with the University of Lisbon, Lisbon, Portugal, supervised by Prof. J. M. Bioucas-Dias from 2017 to 2018. He was the Research Assistant with Hong Kong Baptist University, Hong Kong, supported by Prof. M. K. Ng in 2019. He is currently an Associate Professor with the School of Economic Information Engineering, Southwestern University of Finance and Economics, Chengdu. His research interests include sparse and low-rank modeling and tensor decomposition for multidimensional image processing, especially on the low-level inverse problems for multidimensional images. His homepage is https://taixiangjiang.github.io/.

Michael K. Ng (Senior Member, IEEE) received the Ph.D. degree in mathematics from the Chinese University of Hong Kong, Hong Kong, in 1995.

He is currently the Director of the Research Division for Mathematical and Statistical Science, the Chair Professor with the Department of Mathematics, University of Hong Kong, Hong Kong, and the Chairman of the HKU-TCL Joint Research Center for AI. His research interests include data science, scientific computing, and numerical linear algebra.

Ting-Zhu Huang received the B.S., M.S., and Ph.D. degrees in computational mathematics from the Department of Mathematics, Xi'an Jiaotong University, Xi'an, China, in 1986, 1992, and 2001, respectively.

He is currently a Professor with the School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, China. His current research interests include scientific computation and applications, numerical algorithms for image processing, numerical linear algebra, preconditioning technologies, and matrix analysis with applications.

Prof. Huang is also an Editor of *The Scientific World Journal*, *Advances in Numerical Analysis*, the *Journal of Applied Mathematics*, the *Journal of Pure and Applied Mathematics: Advances in Applied Mathematics*, and the *Journal of Electronic Science and Technology*.