This article has been accepted for publication in IEEE Transactions on Multimedia. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TMM.2022.3216746

Multimodal Core Tensor Factorization and its Applications to Low-Rank Tensor Completion

Haijin Zeng, Jize Xue, Hiêp Q. Luong and Wilfried Philips

***Abstract*—Low-rank tensor completion has been widely used in computer vision and machine learning. This paper develops a novel multimodal core tensor factorization (MCTF) method combined with a tensor low-rankness measure and a better nonconvex relaxation form of this measure (NC-MCTF). The proposed models encode low-rank insights for general tensors provided by Tucker and T-SVD and thus are expected to simultaneously model spectral low-rankness in multiple orientations and accurately restore the data of intrinsic low-rank structure based on few observed entries. Furthermore, we study the MCTF and NC-MCTF regularization minimization problem and design an effective block successive upper-bound minimization (BSUM) algorithm to solve them. Theoretically, we prove that the iterates generated by the proposed models converge to the set of coordinatewise minimizers. This efficient solver can extend MCTF to various tasks such as tensor completion. A series of experiments including hyperspectral image (HSI), video and MRI completion confirm the superior performance of the proposed method.**

***Index Terms*—Tensor, low-rankness, tensor factorization, nonconvex optimization.**

I. Introduction

Low-rankness is a common attribute of many data sources. To date, methods based on low rankness have reported empirical and theoretical success on a large variety of scientific and engineering applications: face modeling [1], gene categorization [2], camera image processing [3], compressive imaging [4], image restoration [5], etc.

A promising method that measures the low-rankness of a matrix should account for the number of nonzero singular values [6]. This low-rankness metric and its relaxations (for example, the L_1 norm and nuclear norm) have been proven useful as regularization terms in applications and have inspired various low-rank models and algorithms to cope with different tasks. On the other hand, a large amount of data generated by modern sensors is naturally represented by high-order tensors, whereas the SVD is restricted to 2D data.

Early high-dimensional data analysis methods reformatted high-dimensional data tensors artificially into 2D matrices and resorted to methods developed for classic two-dimensional analysis methods. However, this flattening strategy and the strict assumptions inherent in two-dimensional analysis do not always match the high-dimensional data well. For example, hyperspectral image (HSI) is the imaging result of different spectral bands from the same spatial scene, which indicates that there is a high correlation in the spectral dimensions [7]; a video contains multiple frames, which contain a high correlation in the temporal dimension, especially for adjacent frames [8]. Therefore, converting these high-dimensional tensor data artificially to 2-D matrices spreads this "local correlation" (e.g., between adjacent frequencies) over large strides in the 2D matrix, complicating the analysis. Thanks to the inherent high-dimensional structure of the data, high-order tensor decomposition allows capturing correlation in a more local fashion along each dimension [9]. In other words, only when analyzing existing inherent multidimensional patterns, we are able to discover the hidden components in the high-dimensional data to model the data more accurately.

The tensor is the generalization of the matrix and vector concept: a vector is a first-order or one-way tensor, and a matrix is a second-order tensor. To measure the low-rankness of tensors, much of the current work involves decomposing the tensor into a combination of several factors to explore its low-rank structure through preliminary tensor decomposition or to unfold the target tensor into matrices according to the modal, and then directly applying the rank of the matrix or the sparsity of the vector to the resulting matrices. Popular decompositions include Tucker [10, 11], Canonical Polyadic (CP) [12], tensor SVD (t-SVD) [13–15], tensor train [16] and tensor ring [17]. There are also some models that further improve the above decomposition or approximation methods. For example, a t-SVD-based nonlocal patches mode is proposed in [18, 19], which achieves robust and promising performance. Ji et al. proposed a smooth matrix factorization-based low-rank tensor completion model [20]. A kernel-based model was proposed in [21] to represent the global correlation prior effectively. This model maps the underlying tensor into the feature space by using a kernel. Some works impose specific prior constraints on the factors obtained by these decompositions [6, 22, 23].

According to the well-established theory of rank function in the matrix case, it seems natural to directly extend matrix completion methods to the tensor completion problem. However, it has been proven that calculating such a tensor rank (whether it is based on Tucker, CP or T-SVD decomposition) is an NP-hard problem [24]. It is difficult to determine or even limit the rank of arbitrary tensors compared to the matrix rank due to tensor low-rankness insight. This should be explained beyond the low-rank properties of all its expanded subspaces. More important, how these subspace low-rank properties are related to the entire tensor structure should also be considered [9]. Some current works directly extend the rank of a matrix to higher-order by simply summing ranks (or its relaxations)

Haijin Zeng, Jize Xue, Hiep Luong and W. Philips are with the Image Processing and Interpretation imec research group at Ghent University, 9000 Ghent, Belgium (e-mail: zeng_navy@163.com, Jize.Xue@UGent.be, Hiep.Luong@UGent.be, Wilfried.Philips@UGent.be)(Corresponding author: Jize Xue.)

This article has been accepted for publication in IEEE Transactions on Multimedia. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TMM.2022.3216746

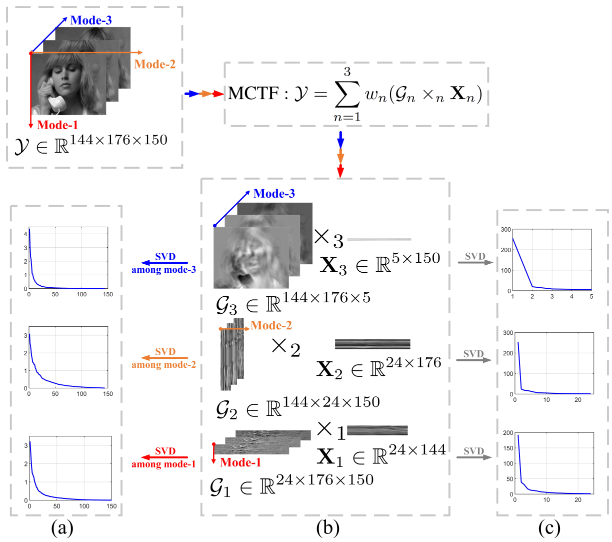


Fig. 1. $\mathcal{Y}$: original video tensor Suzie, there are 150 frames in total, and spatial size of each frame is 144 × 176. (a) shows mode-n singular value curve of $\mathcal{G}_n$. (b) Factor tensor $\mathcal{G}_n$ and factor matrix $\mathbf{X}_n$ obtained by MCTF decomposition along its three modes. (c) Singular value curve of factor matrix $\mathbf{X}_n$.

along all tensor modes [25] or directly consider the global low-rankness of underlying tensors. Different from the matrix scenarios, the simple rank summation term generally lacks a clear physical meaning for tensors [6]. Furthermore, when the sampling rate is very low, it is not sufficient to explore only global low-rankness. As shown in Fig. 1, after exploring the global low-rank prior through tensor decomposition, for the factor obtained by the decomposition, instead of being identical global low-rankness, the low-rankness in different modes or orientations evidently exists and is different. Actually, from Fig. 1-(c), one can see that most singular values of the factor matrices are very close to zero and much smaller than the first several larger singular values. Moreover, instead of being independent, as shown in Fig. 1-(b), there are apparent correlations across different slices of each mode of the factor tensor and matrix.

In this paper, we propose a novel tensor low-rankness measure for the tensor completion problem to effectively model the multimodal low-rankness of high-order tensors. Similar to Tucker, our method is also based upon the tensor and matrix decomposition definitions. However, instead of using the tensor Tucker decomposition directly and requiring the components to be orthogonal, first, a novel tensor decomposition is proposed, in which a high-order low-rank decomposition is introduced into each mode of the underlying tensor. However, the factors are not required to be orthogonal, as shown in Fig. 1(b). Instead of utilizing only one mode low-rankness of the underlying tensor as the tensor nuclear norm based on T-SVD, this decomposition utilizes all mode low-ranknesses of the tensor to give much better performance. Compared with Tucker decomposition, our method does not require the components to be orthogonal; thus, there is no need to use the SVD in our decomposing algorithms, which is computationally much cheaper than Tucker and T-SVD.

TABLE I
NOTATIONAL CONVENTION IN THIS PAPER

Notation	Definition
$\mathbf{x}, \mathbf{y}$	Vectors
$\mathbf{X}, \mathbf{Y}$	Matrices
$\mathcal{X}, \mathcal{Y}$	Tensors
$x_{i_1 \cdots i_N}$	$(i_1, \cdots, i_N)$-th entry of $\mathcal{X}$
$\text{unfold}_n(\mathcal{X}) = \mathbf{X}_{(n)}$	Mode-n unfolding of $\mathcal{X}$
$\text{fold}_n(\text{unfold}_n(\mathcal{X})) = \mathcal{X}$	The inverse of unfold_n
$\text{rank}_n(\mathcal{X}) = \text{rank}(\mathbf{X}_{(n)})$	The n-rank of $\mathcal{X}$

Second, a tensor low-rankness measure based on the proposed decomposition is proposed. This combines both the low-rank prior of the global tensor and the local factors obtained by the proposed tensor decomposition method. Its insight can be easily interpreted as a regularization for the factor tensor and matrix derived from the nonorthogonal multimode low-rank decomposition. Furthermore, an alternative convex relaxation of the proposed low-rankness measure is presented. Such a measure not only unifies the traditional understanding of low-rankness from matrix to tensor but also encodes both sparsity insights delivered by common Tucker, SVD and T-SVD low-rank decompositions for a general tensor.

Third, we apply the proposed low-rank measures to high-dimensional tensor completion tasks, e.g., video, hyperspectral image and MRI completion, and designed a block successive upper-bound minimization (BSUM) method to efficiently solve the resulting models. We also proved that the iterates generated by the proposed model converge to the set of coordinatewise minimizers. The validity of the proposed models is evaluated on a series of experiments including video, MRI and hyperspectral image completion.

II. NOTIONS AND PRELIMINARIES

In this section, we summarized some notations, tensor operations and operators used in this paper.

Table I summarizes the common notations throughout this paper. Following [26, 27], a **fiber** of tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is defined as a vector obtained by fixing all indices of $\mathcal{X}$ except one, and a **slice** of $\mathcal{X}$ is defined as a matrix by fixing all indices of $\mathcal{X}$ except two. $\hat{\mathcal{X}}$ denotes the result of discrete Fourier transformation (FFT) of $\mathcal{X}$ along the 3-rd dimension, and $\mathcal{X}$ can be computed from $\hat{\mathcal{X}}$ via the inverse FFT. Then the multi-rank of $\mathcal{X}$ is defined as the array $\text{rank}(\mathcal{X}) = (\text{rank}(\hat{\mathbf{X}}_{(1)}), \ldots, \text{rank}(\hat{\mathbf{X}}_{(N)}))$, where $\hat{\mathbf{X}}_{(N)}$ denotes the rank of the N-th frontal slice of $\hat{\mathcal{X}}$. Specially, for a 3-way tensor $\mathcal{A} \in \mathbb{C}^{n_1 \times n_2 \times n_3}$, its (i, j, k)-th entry is denoted as a_{ijk} and $\mathcal{A}(i, :, :)$, $\mathcal{A}(:, i, :)$ and $\mathcal{A}(:, :, i)$ represent the i-th horizontal, lateral and frontal slice, respectively.

Based on these common notations of tensor, one can define *inner product, t-product* and *n-mode product*.

Definition 1 (inner product [28]*):* For $\mathcal{X}, \mathcal{Y} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$, their inner product is defined as

$$\langle \mathcal{X}, \mathcal{Y} \rangle = \sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} x_{i_1, \cdots, i_N} y_{i_1, \cdots, i_N}, \tag{1}$$

and $\|\mathcal{X}\|_{\text{F}} = \sqrt{\langle \mathcal{X}, \mathcal{X} \rangle}$ denotes the Frobenius norm of $\mathcal{X}$.

Definition 2 (T-product [29]): Given $\mathcal{X} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ and $\mathcal{Y} \in \mathbb{R}^{d_2 \times d_4 \times d_3}$, their t-product $\mathcal{T} = \mathcal{X} * \mathcal{Y} \in \mathbb{R}^{d_1 \times d_4 \times d_3}$ is a tensor whose $(i,j)^{\text{th}}$ fiber $\mathcal{T}(i,j,:) = \sum_{k=1}^{d_2} \mathcal{X}(i,k,:) \bullet \mathcal{Y}(k,j,:)$, where $\bullet$ is the circular convolution.

Definition 3 (n-mode (matrix) product) [28]: The n-mode product of tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ with a matrix $\mathbf{U} \in \mathbb{R}^{J \times I_n}$, denoted by $\mathcal{X} \times_n \mathbf{U} \in \mathbb{R}^{I_1 \times \cdots \times I_{n-1} \times J \times I_{n+1} \times \cdots \times I_N}$, is defined as

$$(\mathcal{X} \times_n \mathbf{U})_{i_1 \cdots i_{n-1} j i_{n+1} \cdots i_N} = \sum_{i_n=1}^{I_n} x_{i_1 i_2 \cdots i_N} u_{j i_n}.$$

Definition 4 (Conjugate transpose [30]): The conjugate transpose of tensor $\mathcal{A} \in \mathbb{C}^{n_1 \times n_2 \times n_3}$ is the tensor $\mathcal{A}^* \in \mathbb{C}^{n_2 \times n_1 \times n_3}$ obtained by conjugate transposing each of the frontal slices and then reversing the order of transposed frontal slices 2 through n_3.

III. Related Works on Tensor Low-Rankness

The task of tensor completion is to recover the missing values of an incomplete tensor $\mathcal{Y} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$. Since tensor data of high dimensionality usually underlie low-rank [30], the completion of $\mathcal{Y}$ can be modeled as follows:

$$\min_{\mathcal{X}} \operatorname{rank}_t(\mathcal{X}), \quad \text{s.t. } P_\Omega(\mathcal{X} - \mathcal{Y}) = \mathbf{0},$$

where Ω is the given subset of the observed entries, $\operatorname{rank}_t(\mathcal{X})$ denotes the tubal rank of $\mathcal{C}$ and P_Ω is the linear operator that extracts entries in Ω and fills the entries not in Ω with zeros, i.e.,

$$(\mathcal{P}_\Omega(\mathcal{Y}))_{i_1 \cdots i_N} = \begin{cases} y_{i_1, \cdots, i_N}, & (i_1, \cdots, i_N) \in \Omega \\ 0, & \text{otherwise} \end{cases}$$

A tensor is a high-dimensional extension of a two-dimensional matrix, and it can also be reordered into a two-dimensional matrix. Therefore, a natural tensor-filling technique is to unfold the tensor into a matrix and use matrix-based filling methods to achieve its filling and then fold it back to the original high-dimensional tensor. There are two main categories of such methods: low-rank matrix decomposition methods (LRMF) and rank minimization techniques. The principle of low-rank matrix decomposition is to decompose the target matrix into two planar matrices to achieve the inscription of a low-rank prior, while rank minimization achieves this by directly imposing an additional rank constraint on the matrix to be estimated [31, 32].

Although the method of reordering the tensor into a matrix is computationally efficient, this method of dimensionality reduction inevitably destroys the intrinsic structure of the tensor. For example, unfolding hyperspectral or multispectral images along the spectral dimension and those unfolding videos along the temporal dimension will destroy the spatial information of each frequency/time band/frame of these data [6]. Therefore, in the past ten years, much work has focused on completing the task of tensor completion by directly imposing low-rank and sparse constraints on the target tensor. Motivated by the great success of matrix nuclear norms and decomposition, their promotion in the form of tensors has aroused increasing research interest, and there have been many results, e.g., tubal nuclear norm (TNN) [33] and partial sum of the tubal nuclear norm (PSTNN) [34], Tucker rank based on Tucker decomposition, CANDECOMP/PARAFAC (CP) rank based on CP decomposition and framelet-based TNN (FTNN) [35]. Among existing tensor nuclear norms and tensor decomposition, Tucker decomposition, TNN and their extensions have shown superior performance in various applications such as image/video inpainting/denoising [36, 37] and clustering [38].

In Tucker decomposition [10, 36], an N-order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ can be written in the following form:

$$\mathcal{X} = \mathcal{S} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \cdots \times_N \mathbf{U}_N \tag{2}$$

where $\mathcal{S} \in \mathbb{R}^{R_1 \times \cdots \times R_N} (r_i \leq R_i \leq I_i)$ is called the core tensor, $\mathbf{U}_i \in \mathbb{R}^{I_i \times R_i} (1 \leq i \leq N)$ is composed of the R_i orthogonal bases along the i-th mode of $\mathcal{X}$. With this Tucker formula, high-order low-rankness can be quantified as a vector $(r_1, r_2, \cdots r_N)$, i.e., Tucker rank. The degree of freedom of the abovementioned Tucker decomposition is $\prod_{i=1}^N r_i$, which uses the volume of the core tensor to evaluate the low-rankness of the underlying tensor. However, the core tensor obtained by the decomposition of natural data usually has a low-rank structure, which causes this degree of freedom to be further restricted to a smaller number [6]. Therefore, it is difficult for the Tucker rank to take reasonable measures to fully describe the inherent low-rank priors of tensors.

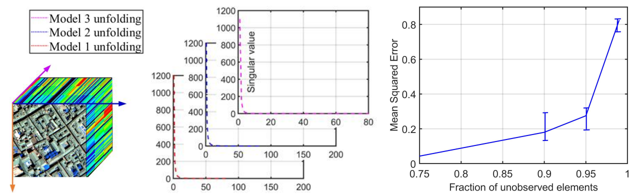
TNN is induced by the T-SVD [13], which attempts to decompose a third-order tensor as a tensor product of three factor tensors. Specifically, let $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$. Then, by using T-SVD, it can be factorized as

$$\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^* \tag{3}$$

where $\mathcal{U} \in \mathbb{R}^{n_1 \times n_1 \times n_3}, \mathcal{V} \in \mathbb{R}^{n_2 \times n_2 \times n_3}$ are orthogonal, and $\mathcal{S} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is a f-diagonal tensor, which is defined as a tensor whose frontal slices is a diagonal matrix. $\mathcal{U} * \mathcal{S}$ and $\mathcal{V}^*$ are the T-product of $\mathcal{U}, \mathcal{S}$ and the conjugate transpose of $\mathcal{V}$, respectively, which are defined in Section II.

The TNN-induced method has reported success in various applications in recent years [39, 40]. It utilizes the low rank of the tensor spectrum, which can well capture the spatiotemporal smoothness. However, by calculating the nuclear norm of the frontal slice after 1-D DFT on the mode-3 fiber, it is sensitive to the orientation of mode and cannot capture the complex intramode and intermode correlation of tensors in multiple directions.

The tensors collected from real scenes often have obvious correlations along each of their modes. Taking HSI as an example, Fig. 2-(b) shows the singular value curve diagram of the three modes of Fig. 2-(a). From the figure, it can be quantitatively observed that only a small fraction of singular values of the three mode unfolding matrices are greater than zero, which means that the three modes along its spectrum and spatial distribution are correlated. This indicates that the tensor along each mode is located on the low-rank subspace, and the entire tensor corresponds to the membership of the subspace along all tensor modes. These facts motivate us to define TNN along different dimensions as the natural intuitive meaning.



(a) (b) (c)

Fig. 2. (a) Real HSI of size 200*200*80; (b) Singular value curves of matrices unfolded along three tensor modes. (c) Tensor completion performance by T-SVD decomposition (TNN) for video "Suzie" dataset. Prediction accuracy severely degenerates when observations are sparse.

On the other hand, in real scenarios, the data representation along a meaningful factor (e.g., $\mathcal{S}, \mathcal{U}, \mathcal{V}$ in (2) or (3)) should always have an evident correlation and thus a low-rank structure [6], Such useful knowledge, however, cannot be well expressed by Tucker or T-SVD decomposition. To ameliorate this issue, we propose a measure for more rationally measuring the low-rankness of the tensor.

IV. MCTF Decomposition-Based Tensor Low-Rankness Measure

Here, we introduce the details of our multimodal core tensor factorization model and then introduce the low-rankness measure metric based on it and a better nonconvex relaxation form of the low-rankness measure.

A. MCTF Decomposition

Tensor low-rankness insight should be interpreted beyond the low-rank property of all its unfolded subspaces and should more importantly consider how such subspace low-rankness are affiliated over the entire tensor structure, especially when the elements in the tensor to be restored are seriously missing. For example, Fig. 2-(c) shows the prediction errors by T-SVD decomposition against the fraction of unobserved elements for a particular HSI dataset, i.e., Pavia City Centre[1]. It can be seen that when fewer data are missing, the error remains a small state, However, when more data are missing, the error starts to increase dramatically.

To reduce the sensitivity to direction and simultaneously capture the complex intra- and intermodal correlations of high-order tensors in multiple directions, improve the limited representation ability and flexibility of the tensor decomposition model in multiorigin correlation modeling, we propose an omnidirectional tensor decomposition strategy called the multimodal core tensor factorization (MCTF) model by employing multilinear techniques.

Definition 2.1 (MCTF). Given an N-way tensor $\mathcal{Y} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, as Tucker proposed, our MCTF decomposition decomposes the input tensor as follows:

$$\begin{aligned}\mathcal{Y} &= w_1(\mathcal{G}_1 \times_1 \mathbf{X}_1) + w_2(\mathcal{G}_2 \times_2 \mathbf{X}_2) + \cdots + w_n(\mathcal{G}_N \times_N \mathbf{X}_N) \\ &= \sum_{n=1}^{N} w_n(\mathcal{G}_n \times_n \mathbf{X}_n),\end{aligned} \tag{4}$$

However, the main difference is that we do not require the components $\mathbf{X}_n$ to be orthogonal, and where $\mathbf{X}_n \in \mathbb{R}^{I_n \times r_n}$ is the n-th ($n = 1, 2, \cdots, N$) factor matrix which reflects the connections (or links) between the latent components and factor matrices, $\mathcal{G}_n \in \mathbb{R}^{I_1 \times \cdots \times I_{n-1} \times r_n \times I_{n+1} \times I_N}$ is a tensor reflecting the joint connections between the latent components in each mode. $w_n (n = 1, 2, \cdots, N)$ are positive weights satisfying $\sum_{n=1}^{N} w_n = 1$. Tucker decomposition imposes the condition of all-orthogonality, instead of diagonality, on tensor $\mathcal{G}_n$, implying that the Tucker is always defined. In fact, $\mathcal{G}$ cannot be diagonal in general, which means that the Tucker test does not necessarily reveal the rank of $\mathcal{Y}$: in the cases where $\mathcal{G}$ is diagonal, and the orthogonality of the matrices of mode-n singular vectors implies that $\mathcal{Y} = \sum_{i_n}^{R_n} g_{i_n i_n \ldots i_n} \mathbf{X}_{i_n}^{(1)} \circ \mathbf{X}_{i_n}^{(2)} \circ \ldots \circ \mathbf{X}_{i_n}^{(N)}$ is a decomposition in a minimal number of rank-1 terms, where $\circ$ is the outer product. On the other hand, the number of nonzero (significant) mode-n singular values corresponds to the mode-n rank (in a numerical sense) of $\mathcal{Y}$ [41].

Actually, the orthogonality constraint in the Tucker decomposition ($\mathcal{Y} = \mathcal{G} \times_n \mathbf{A}$) has two effects. First, it encourages the representations in factor $\mathbf{A}$ to be more distinguishable from each other. $\mathbf{A}$ has column full rank, which means the columns of $\mathbf{A}$ are unrelated to each other. This helps to keep noise out of $\mathbf{A}$. Second, it preserves the distribution of the noise, which enables the denoising of observation $\mathcal{Y}$ by denoising core tensor $\mathcal{G}$. Therefore, the Tucker decomposition is a powerful tool for image denoising, but for the tensor completeness problem that is the focus of this paper, it is virtually noiseless. In addition, the key purpose of employing Tucker decomposition for tensor completion is to represent the high-dimensional low-rankness of the underlying low-rank tensor. However, due to the orthogonality constraint, $\mathbf{A}$ is a column full-rank matrix, and the rank (Tucker) of $\mathcal{G}$ is also bounded with an upper bound and a down bound. Therefore, the efficiency of employing Tucker decomposition to represent low rankness is limited. In addition, for orthogonal Tuker decomposition and nonorthogonal MCTF, we have the following theorem:

Lemma 1. Given a matrix $\mathbf{A}$ with size $m \times n$, $\mathbf{A}$ products (left) a nonsingular matrix $\mathbf{P}$ with size $m \times m$ or products (right) a nonsingular matrix $\mathbf{Q}$ with size $n \times n$ will not change the rank of $\mathbf{A}$.

Proof: due to matrix $\mathbf{P}$ with size $m \times m$ is a nonsingular matrix, i.e., $\mathrm{rank}(\mathbf{P}) = m$, so $\mathrm{rank}(\mathbf{A}) \leqslant \mathrm{rank}(\mathbf{P})$. Let $\mathbf{M} = \mathbf{PA}$. Due to $\mathrm{rank}(\mathbf{PA}) \leqslant \min\{\mathrm{rank}(\mathbf{P}), \mathrm{rank}(\mathbf{A})\}$, we have $\mathrm{rank}(\mathbf{M}) \leqslant \mathrm{rank}(\mathbf{A})$. In addition, according to $\mathbf{A} = \mathbf{P}^{-1}\mathbf{M}$, we have $\mathrm{rank}(\mathbf{A}) \leqslant \mathrm{rank}(\mathbf{M})$. so, $\mathrm{rank}(\mathbf{A}) = \mathrm{rank}(\mathbf{M}) = \mathrm{rank}(\mathbf{PA})$. Similarly, we have $\mathrm{rank}(\mathbf{A}) = \mathrm{rank}(\mathbf{AQ})$.

Theorem 1. Given an N-way tensor $\mathcal{Y}$, $\mathcal{Y} = \mathcal{G} \times_n \mathbf{X}_n$ denotes its Tucker decomposition, where $\mathcal{G} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is the so-called core tensor and $\mathbf{X}_n \in \mathbb{R}^{J_n \times I_n}$ is a semiorthogonal

[1]http://www.ehu.es/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes

This article has been accepted for publication in IEEE Transactions on Multimedia. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TMM.2022.3216746

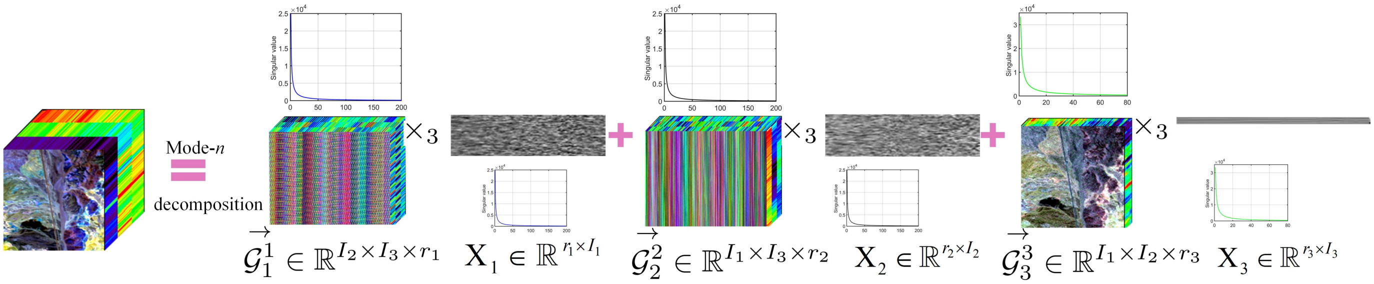


Fig. 3. Illustration of proposed MCTF tensor decomposition for 3D tensors, which encourages simultaneously low rank structure in all orientations.

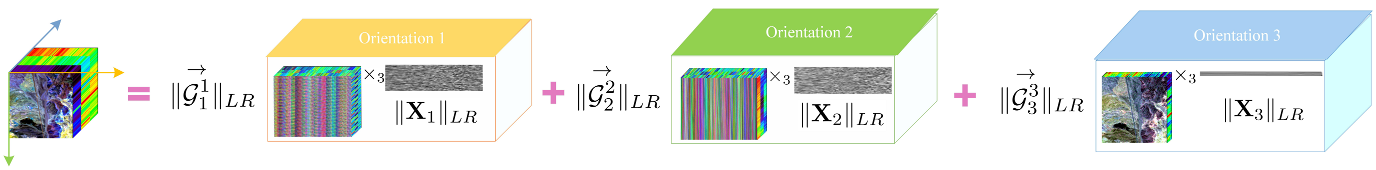


Fig. 4. Illustration of proposed MCTF-based low-rankness measure, which models underlying tensor as mixture of three low-rank combinations of tensor and matrix.

matrix. Then, we have the following conclusions: $\mathrm{rank}(\mathbf{X}_n) = J_n$; if $J_n = I_n, \mathrm{rank}(\mathbf{Y}_{(n)}) = \mathrm{rank}(\mathbf{G}_{(n)})$; if $J_n \neq I_n$, $\mathrm{rank}(\mathbf{Y}_{(n)}) + I_n - J_n \leqslant \mathrm{rank}(\mathbf{G}_{(n)}) \leqslant \mathrm{rank}(\mathbf{Y}_{(n)})$.

Proof: In Tucker decomposition, $\mathbf{X}_n$ is a semiorthogonal matrix and has a column full rank. Therefore, we have $\mathrm{rank}(\mathbf{X}_n) = I_n$ and $\mathcal{G} = \mathcal{Y} \times_n \mathbf{X}_n^*$. By denoting the mode-$n$ product of tensor and matrix as matrix format, $\mathcal{G} = \mathcal{Y} \times_n \mathbf{X}_n^*$ can be rewritten as $\mathbf{G}_{(n)} = \mathbf{X}_n^* \mathbf{Y}_{(n)}$. If $J_n = I_n$, according to Lemma 1, $\mathrm{rank}(\mathbf{Y}_{(n)}) = \mathrm{rank}(\mathbf{G}_{(n)})$. If $J_n \neq I_n$, $\mathrm{rank}(\mathbf{G}_{(n)}) \geqslant \mathrm{rank}(\mathbf{X}_n^*) + \mathrm{rank}(\mathbf{Y}_{(n)}) - J_n$, i.e., $\mathrm{rank}(\mathbf{G}_{(n)}) \geqslant I_n + \mathrm{rank}(\mathbf{Y}_{(n)}) - J_n$. Similarly, for $\mathcal{Y} = \mathcal{G} \times_n \mathbf{X}_n$, we have $\mathrm{rank}(\mathbf{Y}_{(n)}) \geqslant \mathrm{rank}(\mathbf{X}_n) + \mathrm{rank}(\mathbf{G}_{(n)}) - I_n$, i.e., $\mathrm{rank}(\mathbf{Y}_{(n)}) \geqslant \mathrm{rank}(\mathbf{G}_{(n)})$. Therefore, we have $\mathrm{rank}(\mathbf{Y}_{(n)}) + I_n - J_n \leqslant \mathrm{rank}(\mathbf{G}_{(n)}) \leqslant \mathrm{rank}(\mathbf{Y}_{(n)})$.

Theorem 2. Given an N-way tensor $\mathcal{Y} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, $\mathcal{Y} = \mathcal{G} \times_n \mathbf{X}_n$ denotes its MCTF decomposition, where there is no semiorthogonal constraint for $\mathbf{X}_n \in \mathbb{R}^{I_n \times J_n}$. Then, we have the following conclusion: $\mathrm{rank}(\mathbf{X}_n) \leqslant \mathrm{rank}(\mathbf{Y}_{(n)}) - \mathrm{rank}(\mathbf{G}_{(n)}) + I_n$, i.e., there is no lower bound but only an upper bound for $\mathrm{rank}(\mathbf{X}_n)$ and $\mathrm{rank}(\mathbf{G}_{(n)})$.

Therefore, according to Theorem 1 and Theorem 2, if there is a semiorthogonal constraint in the Tucker case, then both the rank of $\mathbf{X}_n$ and $\mathbf{G}_{(n)}$ are fixed, and one cannot further minimize their ranks. If there is no semiorthogonal constraint as in our MCTF, then one can further add low-rank regularization to minimize the rank of $\mathbf{X}_n$ and $\mathbf{G}_{(n)}$.

MCTF encourages a low-rank structure, which means low rankness in the spectral domain of all orientations. It models a data tensor as simultaneously having low tubal rank in all orientations (see Fig. 5 and 4). It differs from TNN, which only considers a low tubal rank of one spectral orientation. For ease of reading, while facilitating the algorithm implementation, we further simplified the form of MCTF. First, we defined a "tensor permutation" operation $\overrightarrow{\mathcal{X}}^k$ to rearrange the dimensions of a tensor. Specifically, for a tensor

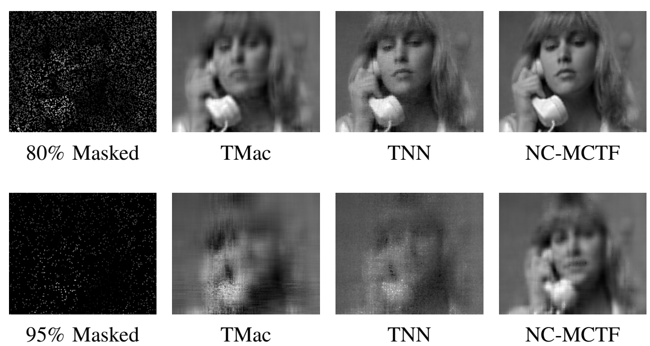


Fig. 5. Two slices of recovered video "Suzie" by TMac, TNN and our NC-MCTF. Sampling rate of first line is 20%, and that of second line is 5%.

$\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, $\overrightarrow{\mathcal{X}}^k$ is defined as the tensor whose ith mode-3 slice is the ith mode-k slice of $\mathcal{X}$., i.e., $\mathcal{X}(i,j,s) = \overrightarrow{\mathcal{X}}^1(j,s,i) = \overrightarrow{\mathcal{X}}^2(s,i,j) = \overrightarrow{\mathcal{X}}^3(i,j,s)$. Then, by using this permutation, the modal-n product of the tensor and the matrix can be uniformly transformed into the modal-3 product, so the MCTF can be rewritten as follows:

$$\begin{aligned}\mathcal{Y} &= w_1(\mathcal{G}_1 \times_1 \mathbf{X}_1) + w_2(\mathcal{G}_2 \times_2 \mathbf{X}_2) + \cdots + w_n(\mathcal{G}_N \times_N \mathbf{X}_N) \\ &= \sum_{n=1}^{N} w_n(\overrightarrow{\mathcal{G}}_n^n \times_3 \mathbf{X}_n).\end{aligned} \tag{5}$$

B. MCTF-Based Tensor Low-Rankness Measure

Existing methods use either factorization or approximation schemes to recover the missing components. However, as the number of missing entries increases, factorization schemes may overfit the model because of incorrectly predefined ranks, while approximation schemes may fail to obtain an easy-to-interpret model factors. Taking the video "Suzie" as an example, as shown in the Fig. 5, when the sampling rate is

This article has been accepted for publication in IEEE Transactions on Multimedia. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TMM.2022.3216746

high, all methods can restore a clear image. However, when the sampling rate is very low, that is, when there are few known entries, both TMac and TNN fail to recover the main information of the image.

Fortunately, there are extra priors that we can utilize, i.e., the model structure is implicitly included in the low-rank factorization model, according to the factor priors, which are usually known a priori in real-world tensor objects. Fig. 3 shows an example of low rank factors. From the figure, one can observe that the factor tensor obtained by the proposed MCTF decomposition is close to low rank along all mode slices. Meanwhile, similar to the factor tensor, the factor matrix is also of low rank. This inspires us to further explore the structure of the factors obtained by the proposed decomposition and then better described the low-rank nature of the original tensor.

To this end, based on the proposed MCTF, we designed the multiple transform domain-based tensor nuclear norm regularization for the factor tensor obtained by the decomposition, together with classic matrix nuclear norm for the factor matrix, to represent the underlying joint manifold drawn from the model factors. We finally propose a low-rankness measure for tensor $\mathcal{Y}$ based on the proposed MCTF, i.e.,

$$\mathcal{S}(\mathcal{Y}) = \sum_{n=1}^{3} (\tau_n \|\mathbf{X}_n\|_* + \lambda_n \|\mathcal{G}_n\|_{\Lambda_n,*}), \tag{6}$$

where $\mathbf{X}_n$ and $\mathcal{G}_n$ are the factor matrix and tensor of $\mathcal{Y}$ with MCTF, respectively; τ_n and λ_n are the parameters to trade off the two terms; $\|\mathbf{X}\|_*$ is the matrix nuclear norm; and $\|\mathcal{X}\|_{\Lambda_n,*}$ denotes the TNN of $\mathcal{X}$ based on transform domain Λ_n. Applying transform domain Λ_n to $\mathcal{X}$ is equivalent to performing the DFT along each mode-n fiber of $\mathcal{X}$ (as shown in the A and A-I column of Fig. 1). This can enhance flexibility for handling different correlations along different modes and reduce the sensitivity to direction. An illustration of the proposed low-rankness measure can be found in Fig. 6. As shown in Fig. 6, the first operation in the proposed (6) is the MCTF decomposition of the underlying tensor, which serves as a complexity measure in the original domain for all orientations. It is based on the matrix decomposition theory to extend the Tucker decomposition and conforms to the internal mechanism of these two decompositions. As shown in Fig. 4, The second term can be interpreted as the number of nonzero singular values of the factor matrix, and the third term models the TNN of $\mathcal{G}_n$ in all orientations, which measures low-rankness in the Fourier domain. They tend to normalize the low-rank attributes across the subspace of each tensor mode. This comprehensive consideration in the proposed measures is conducive to exploring the internal low-rank construction of the factor tensor and the low-rank nature of the tensor quantum space along each mode.

C. MCTF-Based Tensor Nonconvex Low-Rankness Measure

Although the proposed multiple transform domains based on low-rankness measure (6) can provide an efficient numerical solution and report success on low-rank completion experiments (the detailed performance of MCTF can be found in Section VI), we must admit that it also has two shortcomings.

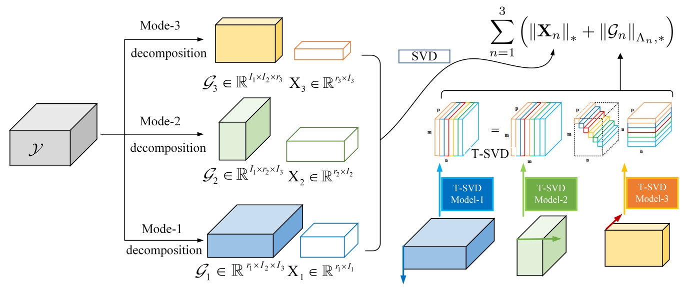


Fig. 6. Visual display of proposed tensor low-rank measure.

First, the TNN in MCTF is essentially the nuclear norm of each spectral slice in the Fourier domain, measuring the L_1 norm of nonzero singular values, which is not an ideal approximation of the tensor tubal rank. Second, both the TNN and NN treat each singular value equally, so the main information may not be well preserved. Larger singular values usually correspond to primary information such as contours, sharp edges and smooth areas, while smaller singular values are mainly composed of noise or outliers [34, 42, 43]. This means that singular values with different numerical values should be treated differently, that is, the punishment for singular values with large numerical values should be reduced, and the punishment for singular values with small numerical values should be increased.

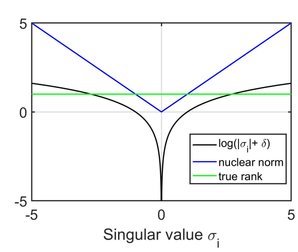


Fig. 7. Different approximations of rank function.

To overcome the above two shortcomings, we designed a novel tensor log-norm and matrix log-norm to perform nonconvex relaxation of TNN and NN to more accurately describe the low-rank structure of the factor tensor and factor matrix:

$$\min \sum_{n=1}^{3} \frac{\alpha_n}{2} \|\mathcal{Y} - \mathcal{G}_n \times_n \mathbf{X}_n\|_F^2 + \tau_n \|\mathbf{X}_n\|_{\log} + \lambda_n \|\mathcal{G}_n\|_{\Lambda_n,\log}, \tag{7}$$

where

$$\begin{aligned} \|\mathcal{G}_n\|_{\Lambda_n,\log} = \|{}_n\overline{\mathcal{G}}_n\|_{\log} &= \| \text{ blockdiag } ({}_n\widehat{\mathcal{G}}_n)\|_{\log} \\ &= \frac{1}{p}\sum_{i=1}^{p} \|{}_n\widehat{\mathcal{G}}_n^{(i)}\|_{\log} \end{aligned} \tag{8}$$

and $\|\mathbf{X}\|_{\log} = \sum_{i=1}^{\min\{m,n\}} (log (|\sigma_i (\mathbf{X})| + \epsilon))$, for $\mathbf{X} \in \mathbb{R}^{m\times n}$, $\sigma_i(\mathbf{X})$ is the i-th singular value of the matrix $\mathbf{X}$, $\epsilon > 0$ is a constant. $\alpha_n, n = 1, 2, 3$, are positive weights satisfying $\sum_{n=1}^{3} \alpha_n = 1$. For $\mathcal{L} \in \mathbb{R}^{m\times n\times p}$, ${}_n\widehat{\mathcal{L}} \in$

$\mathbb{C}^{m\times n\times p}$ denotes the result of Discrete Fourier Transformation (DFT) on $\mathcal{L} \in \mathbb{R}^{m\times n\times p}$ along the n-th dimension, i.e., ${}_n\widehat{\mathcal{L}} = \boldsymbol{F}_n\mathcal{L}$, where $\boldsymbol{F}_n$ is the DFT matrix defined as $\boldsymbol{F}_n = [\boldsymbol{f}_1,\cdots,\boldsymbol{f}_i,\cdots,\boldsymbol{f}_{n_3}] \in \mathbb{R}^{n\times n}$. $\overline{\mathcal{L}}$ denotes the block-diagonal matrix of tensor $\mathcal{L}$ in the Fourier domain. $\overline{\mathcal{L}} =$ blockdiag $(\widehat{\mathcal{L}}) = \mathrm{diag}(\widehat{\mathcal{L}}^{(1)},\widehat{\mathcal{L}}^{(2)},\ddots,\widehat{\mathcal{L}}^{(p)}) \in \mathbb{C}^{mp\times np}$, where $\widehat{\mathcal{L}}^{(i)}$ denotes the i-th frontal slices of $\widehat{\mathcal{L}}$, $i=1,2,\cdots,p$. In addition, a detailed comparison between the true rank, nuclear norm and nonconvex log-norm is shown in Fig. 7. Obviously, the black curve of the log-norm is closer to the green curve of the L_0 norm (rank function) than the blue nuclear norm.

V. MCTF Low-rankness Measure-based Model & its Solving Scheme

Here we introduce the optimization of the proposed two models and analyze their convergence.

A. The MCTF and NC-MCTF minimization models

Before giving the optimization of MCTF and NC-MCTF, we first introduce two lemmas as follows:

Lemma 1 (singular value shrinkage operator, SVT): For $\mathbf{M} \in \mathbb{R}^{n_1\times n_2}$, $\mathbf{M} = \mathbf{P}\mathbf{E}_r\mathbf{Q}^\dagger$ denotes the singular value decomposition (SVD) of matrix $\mathbf{M}$ with rank r, where $\mathbf{E}_r = \mathrm{diag}(\{\sigma_i\}_{1\le i\le r})$, σ_i is the i-th largest singular value of $\mathbf{M}$. Then, the following properties hold,

$$\mathrm{D}_\delta(\mathbf{M}) = \arg\min_{\mathrm{rank}(\mathbf{X})\le r} \delta\|\mathbf{X}\|_* + \frac{1}{2}\|\mathbf{X}-\mathbf{W}\|_{\mathrm{F}}^2,$$

where $\mathrm{D}_\delta(\mathbf{W}) = \mathbf{P}\,\mathrm{diag}\{\max((\sigma_i-\delta),0)\}\,\mathbf{Q}^\dagger$, and $\|.\|_*$ is the matrix nuclear norm.

Lemma 2 (weighted nuclear norm minimization, WNNM) [42]: For any $\gamma>0, \mathbf{Y}\in\mathbb{R}^{m\times n}$ and $0\le d_1\le d_2\le\cdots\le d_r(r=\min(m,n))$, a global optimal solution to the following problem

$$\min_{\mathbf{X}} \sum_{j=1}^{r} \gamma d_j\sigma_j(\mathbf{X}) + \frac{1}{2}\|\mathbf{Y}-\mathbf{X}\|_F^2 \tag{9}$$

is given by the following singular value thresholding

$$\mathbf{X}^* = \mathbf{W}_{\gamma,d}(\mathbf{Y}) = \mathbf{U}\mathbf{S}_{\gamma,d}(\Sigma)\mathbf{V}^T \tag{10}$$

where $\mathbf{Y}=\mathbf{U}\Sigma\mathbf{V}^T$ is the SVD of $\mathbf{Y}$, $\sigma_j(\mathbf{X})$ denotes the j-th singular value of $\mathbf{X}$ and $\mathbf{S}_{\gamma,d}(\Sigma)_{jj} = \max(\Sigma_{jj}-\gamma d, 0)$.

We then analyzing the optimization of the proposed models, the objective function of the proposed MCTF and NC-MCTF are listed as follows:

$$f(\mathbf{X},\mathcal{G},\mathcal{Y}) = \sum_{n=1}^{3}\frac{\alpha_n}{2}\|\mathcal{Y}-\mathcal{G}_n\times_n\mathbf{X}_n\|_{\mathrm{F}}^2 + \tau_n\|\mathbf{X}_n\|_* + \lambda_n\|\mathcal{G}_n\|_{\Lambda_n,*}, \tag{11}$$

$$f(\mathbf{X},\mathcal{G},\mathcal{Y}) = \sum_{n=1}^{3}\frac{\alpha_n}{2}\|\mathcal{Y}-\mathcal{G}_n\times_n\mathbf{X}_n\|_{\mathrm{F}}^2 + \tau_n\|\mathbf{X}_n\|_{\log} + \lambda_n\|\mathcal{G}_n\|_{\Lambda_n,\log}. \tag{12}$$

The minimization of the proposed models are two complicated optimization problems, which are difficult to solve directly. Here, we adopt the block successive upper-bound minimization (BSUM) [44] to solve them.

According to the proximal operator [45], for the k-th iteration, the update can be written as follows:

$$\mathrm{Prox}_f(\mathcal{S},\mathcal{S}^k) = \arg\min_{\mathcal{S}} f(\mathcal{S}) + \frac{\rho}{2}\left\|\mathcal{S}-\mathcal{S}^k\right\|_{\mathrm{F}}^2, \tag{13}$$

where $\rho>0$ is the proximal parameter, $\mathcal{S}=(\mathbf{X},\mathcal{G},\mathcal{Y})$ and $\mathcal{S}^k = (\mathbf{X}^k,\mathcal{G}^k,\mathcal{Y}^k)$.

Let $\mathcal{S}_1^k = (\mathbf{X}^k,\mathcal{G}^k,\mathcal{Y}^k)$, $\mathcal{S}_2^k = (\mathbf{X}^{k+1},\mathcal{G}^k,\mathcal{Y}^k)$, $\mathcal{S}_3^k = (\mathbf{X}^{k+1},\mathcal{G}^{k+1},\mathcal{Y}^k)$. By BSUM, (13) can be rewritten as follows:

$$\begin{aligned}
\mathbf{X}^{k+1} &= \mathrm{Prox}_f(\mathbf{X},\mathcal{S}_1^k) = \arg\min_{\mathbf{X}} f(\mathbf{X},\mathcal{G}^k,\mathcal{Y}^k) + \frac{\rho}{2}\|\mathbf{X}-\mathbf{X}^k\|_{\mathrm{F}}^2,\\
\mathcal{G}^{k+1} &= \mathrm{Prox}_f(\mathcal{G},\mathcal{S}_2^k) = \arg\min_{\mathcal{G}} f(\mathbf{X}^{k+1},\mathcal{G},\mathcal{Y}^k) + \frac{\rho}{2}\|\mathcal{G}-\mathcal{G}^k\|_{\mathrm{F}}^2,\\
\mathcal{Y}^{k+1} &= \mathrm{Prox}_f(\mathcal{Y},\mathcal{S}_3^k) = \arg\min_{\mathcal{Y}} f(\mathbf{X}^{k+1},\mathcal{G}^{k+1},\mathcal{Y}) + \frac{\rho}{2}\|\mathcal{Y}-\mathcal{Y}^k\|_{\mathrm{F}}^2.
\end{aligned} \tag{14}$$

1) Update $\mathbf{X}_n$ with fixing others: By introducing one auxiliary variable $\mathbf{Z}_n$, the $\mathbf{X}_n$-subproblem in (14) can be rewritten as

$$\arg\min_{\mathbf{X}_n,\mathbf{Z}_n} \sum_{n=1}^{3}\Big(\frac{\alpha_n}{2}\|\mathcal{Y}-\mathcal{G}_n\times_n\mathbf{X}_n\|_{\mathrm{F}}^2 + \tau_n\|\mathbf{Z}_n\|_{*\text{ or }\log} + \frac{\rho_n}{2}\|\mathbf{X}_n-\mathbf{X}_n^k\|_{\mathrm{F}}^2\Big), s.t., \mathbf{X}_n = \mathbf{Z}_n. \tag{15}$$

Based on the augmented Lagrange multiplier (ALM) method, the above minimization problem (15) can be transformed into no-contrined problem, and be solved by SVT (V-A) and WNNM operator (10):

$$\mathbf{Z}_n^{k+1} = \mathrm{D}_{\frac{\tau_n}{\rho_n}}(\mathbf{X}_n^k + \Gamma_n^{\mathbf{X}}/\rho_n), n=1,2,\cdots,N; \tag{16}$$

$$\mathbf{Z}_n^{k+1} = W_{\frac{\tau_n}{\rho_n},\epsilon}\left(\mathbf{X}_n^k + \Gamma_n^{\mathbf{X}}/\rho_n\right), n=1,2,\cdots,N. \tag{17}$$

$$\mathbf{X}_n^{k+1} = (\alpha_n\mathbf{G}_n^T\mathbf{G}_n + 2\rho\mathbf{I}_n)^{-1}[\alpha_n\mathbf{G}_n^T\mathbf{Y}_{(n)} + \mu_n(\frac{\mathbf{Z}_n^{k+1} - \Gamma_n^k/\mu_n + \mathbf{X}_n^k}{2}). \tag{18}$$

where (16) for MCTF, (17) for NC-MCTF. Based on the ALM method, the multipliers are updated by the following equations:

$$\Gamma_n^{\mathbf{X}} = \Gamma_n^{\mathbf{X}} + \mathbf{X}_n - \mathbf{Z}_n. \tag{19}$$

2) Update $\mathcal{G}_n$ with fixing others: By introducing an auxiliary variable, the $\mathcal{G}_n$-subproblem can be rewritten as

$$\arg\min_{\mathcal{G}_n} \sum_{n=1}^{3}\Big(\frac{\alpha_n}{2}\|\mathcal{Y}-\mathcal{G}_n\times_n\mathbf{X}_n\|_{\mathrm{F}}^2 + \lambda_n\|\mathcal{J}_n\|_{\Lambda_n,*} + \frac{\rho_n}{2}\|\mathcal{G}_n-\mathcal{G}_n^k\|_{\mathrm{F}}^2\Big), s.t., \mathcal{G}_n = \mathcal{J}_n. \tag{20}$$

This article has been accepted for publication in IEEE Transactions on Multimedia. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TMM.2022.3216746

By using the ALM and SVT operator (V-A), one can also obtain the solutions:

$$ {}_n\widehat{\mathcal{J}}_n^{k+1,(q)} = D_{\frac{1}{\rho_n}}\left({}_n\widehat{\mathcal{U}}_n^{(q)}\right),\ q = 1, 2, \cdots, p. \tag{21} $$

Then, the $(k+1)$-th updating of $\mathcal{J}_n^{k+1}$ can be obtained via inverse Fourier transform

$$ \mathcal{J}_n^{k+1} = \text{ifft}\left({}_n\widehat{\mathcal{J}}_n^{k+1}, [\,], n\right). \tag{22} $$

Similarly, the $\mathcal{J}_n$ related subproblem can be solved by the WNNM operator (10), i.e.,

$$ {}_n\widehat{\mathcal{J}}_n^{k+1,(q)} = W_{\frac{1}{\rho_n},\epsilon}\left({}_n\widehat{\mathcal{U}}_n^{(q)}\right),\ q = 1, 2, \cdots, p. \tag{23} $$

Then, the $(k+1)$-th updating of $\mathcal{J}_n^{k+1}$ can be obtained via inverse Fourier transform

$$ \mathcal{J}_n^{k+1} = \text{ifft}\left({}_n\widehat{\mathcal{J}}_n^{k+1}, [\,], n\right). \tag{24} $$

With other variables fixed, the minimization subproblem for $\mathcal{G}_n$ is also convex and has the following closed-form solution

$$ \begin{aligned} \mathcal{G}_n^{k+1} = \text{fold}\Bigg(\bigg(\mathbf{Y}_{(n)}^k\left(\mathbf{X}_n^{k+1}\right)^T + 2\rho_n\Big(\tfrac{\mathbf{J}_n^{k+1}-\Gamma_n^{\mathcal{G}}/\rho_n+\mathbf{G}_n^k}{2}\Big)\bigg) \\ \left(\mathbf{X}_n^{k+1}\left(\mathbf{X}_n^{k+1}\right)^T + 2\rho_n\mathbf{I}_n\right)^{\dagger}\Bigg), \\ n = 1, 2, \cdots, N. \end{aligned} \tag{25} $$

Finally, the Lagrangian multiplier can be updated by the following equations

$$ \Gamma_n^{\mathcal{G}} = \Gamma_n^{\mathcal{G}} + \mathcal{G}_n - \mathcal{J}_n. \tag{26} $$

3) Update $\mathcal{Y}$ with fixing others: The update of $\mathcal{Y}_{k+1}$ can be written explicitly as

$$ \mathcal{Y}^{k+1} = P_{\Omega^c}\left(\sum_{n=1}^{3}\alpha_n \text{ fold }_n\left(\frac{\mathbf{G}_n^{k+1}\mathbf{X}_n^{k+1} + \rho_n\mathbf{Y}_{(n)}^k}{1+\rho_n}\right)\right) + \mathcal{F} \tag{27} $$

where $\mathcal{F}$ is the observed data; P_Ω is an operator defined in subsection II.

B. Complexity Analysis

The proposed algorithms for the proposed MCTF and NC-MCTF are summarized as Algorithm 1. Further, we discuss the complexity of the proposed algorithms. Complexity Analysis: The cost of computing $\mathbf{X}_n$ is $O\left(I_n r_n^2 + I_n r_n s_n + r_n^2 s_n\right)$; calculating $\mathbf{Z_n}$ has a complexity of $O\left(\Pi_{j\neq n} I_j \times r_n^2\right)$; the complexity of updating $\mathcal{J}_n$ is $O\left(I_n r_n^2\right)$; calculating $\mathcal{G}_n, n = 1, 2, 3$, in both MCTF-based solver and NC-MCTF-based solver, have a complexity of $O(I_1 I_2 I_3(\log(I_1 I_2 I_3) + \sum_{n=1}^{3}\min(I_n, I_{n+1})))$, where we define $I_4 = I_1$; calculating $\mathcal{Y}$ has a complexity of $O(\sum_{n=1}^{N} r_n I_n s_n)$. Then, the total complexity of the proposed algorithms can be obtained by counting the complexity of the above variables. For easily viewing, we list the total complexity of the proposed models as follows:

$$ \begin{aligned} O(3I_n r_n^2 + 3I_n r_n s_n + 3r_n^2 s_n + I_1 I_2 I_3(\log(I_1 I_2 I_3) \\ + \sum_{n=1}^{3}\min(I_n, I_{n+1}))). \end{aligned} \tag{28} $$

Algorithm 1 :Algorithm for the proposed MCTF and NC-MCTF based tensor low-rankness measure.

1: **Input:** The observed tensor $\mathcal{F}$; The set of index of observed entries Ω; The given n-rank, $r = (r_1, r_2, r_3)$; stopping criterion ε.
2: **Outpot:** the completed tensor.
3: Initialize: $\mathbf{X}_n^0 = \mathbf{Z}_n^0 = \mathbf{0}, \mathcal{G}_n^0 = \mathcal{J}_n^0 = \mathbf{0}, \Gamma_n^{\mathbf{X}} = \mathbf{0}, \Gamma_n^{\mathcal{G}} = \mathbf{0}, n = 1, 2, \cdots, N$; $\mu_{\max} = 10^6, \rho = 1.5$, $\mathcal{Y} = \mathcal{P}_\Omega(\mathcal{F})$, and $k = 0$.
4: Repeat until convergence:
5: Update $\mathbf{X}, \mathbf{Z}, \mathcal{G}, \mathcal{J}, \mathcal{Y}, \Gamma^{\mathbf{X}}, \Gamma^{\mathcal{G}}$ via
  1st step: Update $\mathbf{Z}_n$ of MCTF via (16) or $\mathbf{Z}_n$ of NC-MCTF via (17)
  2nd step: Update $\mathbf{X}_n$ via (18)
  3rd step: Update $\mathcal{G}_n$ via (25)
  4th step: Update $\mathcal{J}_n$ of MCTF via (22) or $\mathcal{J}_n$ of NC-MCTF via (24)
  5th step: Update $\mathcal{Y}$ via (27)
  6th step: Update the parameter via (19), (26)
6: Check the convergence condition: $\frac{\left\|\mathcal{Y}^{k+1}-\mathcal{Y}^k\right\|_{\mathrm{F}}}{\left\|\mathcal{Y}^k\right\|_{\mathrm{F}}} < \varepsilon$.

By comparison, the costs of TNN at each iteration is $\mathcal{O}\left(n_1 n_2 n_3 \log n_3 + n_1 n_2 n_3 \min\left(n_1, n_2\right)\right)$. The costs of TMac is $\mathcal{O}\left(\left(r_1 + r_2 + r_3\right) n_1 n_2 n_3\right)$ at each iteration, where r_1, r_2 and r_3 respectively denote the estimated rank of the three unfolded matrices. The costs of PSTNN is $\mathcal{O}\left(n_1 n_2 n_3 \log\left(n_3\right) + n_3 \min\left(n_1 n_2^2, n_2 n_1^2\right)\right)$. The costs of FTNN is $O\left(w n_1 n_2 n_3\left(n_3 + \min\left(n_1, n_2\right)\right)\right)$. The costs of MF-TV is $O\left(I_3 r_3^2 + 2I_3 r_3 s_3 + 3r_3^2 s_3 + r_3 s_3 \log s_3 + \sum_{n\neq 3}\left(2I_n r_n^2 + 3I_n r_n s_n + 2r_n^2 s_n\right)\right)$.

C. Convergence Analysis

For the case of using nuclear norm as the rank approximation, due to nuclear norm being the convex approximation of rank function, the convergence can be established easily. For the case of using the non-convex log norm as the rank approximation, although log-norm is not convex, it is a quasiconvex function. Therefore, the convergence of the resulting algorithm can also be established under the BSUM [44, 45] framework.

Lemma 1 [44, 45]. Given the problem $\arg\min f(x)$, s.t. $x \in \mathcal{X}$, where $\mathcal{X}$ is the feasible set. Assume $h\left(x, x^{k-1}\right)$ is an approximation of $f(x)$ at the $(k-1)$th iteration, which satisfied the following conditions:

$$ \begin{aligned} &1)\quad h_i\left(y_i, y\right) = f(y), \forall y \in \mathcal{X}, \forall i; \\ &2)\quad h_i\left(x_i, y\right) \geq f\left(y_1, \ldots, y_{i-1}, x_i, y_{i+1}, \ldots, y_n\right), \\ &\qquad \forall x_i \in \mathcal{X}_i, \forall y \in \mathcal{X}, \forall i; \\ &3)\quad h_i'\left(x_i, y; d_i\right)\big|_{x_i=y_i} = f'(y; d), v_i = (0, \ldots, d_i \ldots 0) \\ &\qquad \text{s.t. } y_i + d_i \in \mathcal{X}_i, \forall i; \\ &4)\quad h_i\left(x_i, y\right) \text{is continuous in} \left(x_i, y\right), \forall i; \end{aligned} \tag{29} $$

where $h_i\left(x_i, y\right)$ is the sub-problem with respect to the ith block and $f'(y; d)$ is the direction derivative of f at the point y in direction d. Suppose $h_i\left(x_i, y\right)$ is quasiconvex in x_i for $i = 1, 2, \cdots, n$. Furthermore, assume that each sub-problem

$\operatorname{argmin} h_i\left(x_i, x^{k-1}\right)$, s.t. $x \in \mathcal{X}_i$ has a unique solution for any point $x^{k-1} \in \mathcal{X}$. Then, the iterates generated by the BSUM algorithm converge to the set of coordinatewise minimum of f.

Theorem 1. The iterates generated by (13) converge to the set of coordinatewise minimizers.

Proof. According to the notations in (13) and (14), we give the notions for convenience

$$\left\{\begin{array}{l} g(\mathcal{S}, \mathcal{S}^k)=f(\mathcal{S})+\frac{\rho}{2}\left\|\mathcal{S}-\mathcal{S}^k\right\|_{\mathrm{F}}^2, \\ g_1\left(\mathbf{X}, \mathcal{S}_1^k\right)=f\left(\mathbf{X}, \mathbf{A}^k, \mathcal{Y}^k\right)+\frac{\rho}{2}\left\|\mathbf{X}-\mathbf{X}^k\right\|_{\mathrm{F}}^2, \\ g_2\left(\mathbf{A}, \mathcal{S}_2^k\right)=f\left(\mathbf{X}^{k+1}, \mathbf{A}, \mathcal{Y}^k\right)+\frac{\rho}{2}\left\|\mathbf{A}-\mathbf{A}^k\right\|_{\mathrm{F}}^2, \\ g_3\left(\mathcal{Y}, \mathcal{S}_3^k\right)=f\left(\mathbf{X}^{k+1}, \mathbf{A}^{k+1}, \mathcal{Y}\right)+\frac{\rho}{2}\left\|\mathcal{Y}-\mathcal{Y}^k\right\|_{\mathrm{F}}^2 . \end{array}\right. \tag{30}$$

It is easy to verify that $g\left(\mathcal{S}, \mathcal{S}^k\right)$ is an approximation and a global upper bound of $f(\mathcal{S})$ at the k-th iteration, which satisfies the following conditions: (1) $g_i(\mathcal{S}_i, \mathcal{S})=f(\mathcal{S}), \forall \mathcal{S}, i=1,2,3$; (2) $g_i(\bar{\mathcal{S}}_i, \mathcal{S}) \geq f\left(\mathcal{S}_1, \ldots, \bar{\mathcal{S}}_i, \ldots, \mathcal{S}_3\right), \forall \bar{\mathcal{S}}_i, \forall \mathcal{S}, i=1,2,3$; (3) $g_i'(\bar{\mathcal{S}}_i, \mathcal{S}; \mathbf{M}_i)_{\bar{\mathcal{S}}_i=\mathcal{S}_i}=f'(\mathcal{S}; \mathbf{M}^i), \forall \mathbf{M}^i=(0, \ldots, \mathbf{M}_i, \ldots, 0)$; (4) $g_i(\bar{\mathcal{S}}_i, \mathcal{S})$ is continuous in $(\bar{\mathcal{S}}_i, \mathcal{S}), i=1,2,3$; (5) For nuclear norm, nuclear norm is convex, therefore $g_i(i=1,2,3)$ is also convex with respect to $\mathbf{X}, \mathbf{A}$. For the case of non-convex log norm, it is not convex, but it is a quasiconvex, so $g_i(i=1,2,3)$ is quasiconvex with respect to $\mathbf{X}, \mathbf{A}$ and $\mathcal{Y}$ and has a unique solution, where $\mathcal{S}=(\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3)=(\mathbf{X}, \mathbf{A}, \mathcal{Y})$. To sum up, all assumptions in **Lemma 1** are satisfied, and the iterates generated by (13) converge to the set of coordinatewise minimizers.

VI. Numerical Experiments

Three types of public tensor datasets, i.e., video, MRI and hyperspectral image, are selected for verification experiments to evaluate the performance of the proposed model. Five state-of-the-art techniques were proposed between 2013 and 2020, i.e., five tensor completion models related to the proposed models: TMac (2013) [26], MF-TV method (2016) [45], TNN (2016) [33], PSTNN (2020) [34] and FTNN (2020) [35]. These were chosen for comparison.

Two types of standards for evaluation: qualitative visual evaluation of the restored data, five widely used quantitative picture quality indices (PQIs, PSNR [46], SSIM [47], FSIM [48], ERGAS [49] and SAM [50]) were utilized to assess the quality of the restored tensor. All experiments were performed on MATLAB 2018b, Intel Core i7@2.2 GHz and 64.0 GB RAM. For a tensor $\mathcal{Y} \in \mathbb{R}^{I_1 \times \ldots \times I_N}$, The SR can be defined as $\mathrm{SR}=\frac{S_{\text{number}}}{\prod_{n=1}^N I_n}$, where S_{number} denotes the number of sampled entries, and Ω represents the index set. The sampled entries are chosen randomly from a tensor $\mathcal{Y}$ by a uniform distribution.

A. MRI

In this subsection, to further verify the versatility of the proposed models for different datasets, We conducted experiments on cubical MRI data[2] with size $150 \times 150 \times 181$. SRs are set as follows: 5%, 10%, 20% and 30%. Here, we set the rank to (T_1, T_2, T_3), where T_1, T_2, T_3 denote the number of the largest 0.5% singular values of model-1, model-2 and model-3, respectively.

[2]http://brainweb.bic.mni.mcgill.ca/brainweb/selection_normal.html

TABLE II
Averaged PSNR, SSIM, FSIM and ERGAS of recovered results on MRI by MF-TV, TMac, FTNN, PSTNN, TNN, our MCTF and NC-MCTF at different sampling rates. Best values appear in boldface.

				SR =0.05				
method	noisy	MF-TV	TMac	FTNN	PSTNN	TNN	MCTF	NC-MCTF
PSNR	10.258	12.332	20.51	22.540	15.859	18.218	22.951	**23.698**
SSIM	0.228	0.099	0.45	0.508	0.224	0.27	0.528	**0.534**
FSIM	0.473	0.52	0.711	0.732	0.642	0.646	0.771	**0.775**
ERGAS	1030.203	814.747	339.385	268.839	545.77	434.774	277.105	**258.370**
				SR = 0.1				
method	noisy	MF-TV	TMac	FTNN	PSTNN	TNN	MCTF	NC-MCTF
PSNR	10.492	15.406	21.411	27.641	22.061	22.535	29.592	**31.597**
SSIM	0.241	0.25	0.531	0.805	0.482	0.536	0.814	**0.884**
FSIM	0.511	0.587	0.732	0.885	0.764	0.78	0.883	**0.912**
ERGAS	1002.8	584.827	308.655	165.366	275.473	266.753	128.252	**101.607**
				SR = 0.2				
method	noisy	MF-TV	TMac	FTNN	PSTNN	TNN	MCTF	NC-MCTF
PSNR	11.003	27.062	22.33	31.783	29.152	28.571	35.550	**36.471**
SSIM	0.271	0.737	0.586	0.907	0.804	0.802	0.950	**0.960**
FSIM	0.564	0.84	0.754	0.938	0.895	0.891	0.953	**0.960**
ERGAS	945.583	173.636	276.269	100.444	127.133	136.182	64.008	**57.003**
				SR = 0.3				
method	noisy	MF-TV	TMac	FTNN	PSTNN	TNN	MCTF	NC-MCTF
PSNR	11.582	36.355	23.077	34.806	32.608	32.481	37.783	**38.531**
SSIM	0.303	0.954	0.625	0.949	0.895	0.89	0.969	**0.974**
FSIM	0.597	0.962	0.773	0.963	0.939	0.939	0.969	**0.974**
ERGAS	884.608	52.449	252.057	70.542	85.845	87.312	49.323	**45.028**

For quantitative evaluation, Table II lists the PQI of all recovery results in detail, and the best results are marked in boldface. It can be seen from the table that the proposed NC-MCTF obtains the best PQI and the second best is the proposed MCTF, both of which are superior to competing methods of the same type. Fig. 13 shows the detailed PSNR, SSIM and FSIM of all slices of the restored data. The same advantages of our model can also be seen here.

Furthermore, the proposed model is evaluated in terms of visual evaluation. We choose the restoration result of a 0.1 sampling rate as an example. Fig. 8 shows the original MRI data, sampled data and recovery results of different methods. Compared with competing methods, the images recovered by MCTF and NC-MCTF show richer details and clearer structures. In addition, one can see from the figure that the nonconvex metric, as shown in NC-MCTF, produces an impressive improvement over MCTF.

B. Video

In this subsection, two public and classic video datasets, i.e., "Suzie" and "Hall"[3] with a size of $144 \times 176 \times 150$, are selected for comparative experiments to test the performance of our model. SRs are set as 5%, 10% and 20%. The proposed model was comprehensively evaluated from both quantitative and qualitative perspectives. Here, we set the rank to (T_1, T_2, T_3), where T_1, T_2, T_3 denote the number of the largest 0.5% singular values of model-1, model-2 and model-3, respectively.

Quantitative comparison: Table III and Table IV give the detailed PQI of all recovered data at three sampling rates. Boldface indicates the best PQI for each sampling rate. It can be clearly seen from Table III and Table IV that among all of the test methods, the proposed NC-MCTF obtains the

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

This article has been accepted for publication in IEEE Transactions on Multimedia. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TMM.2022.3216746

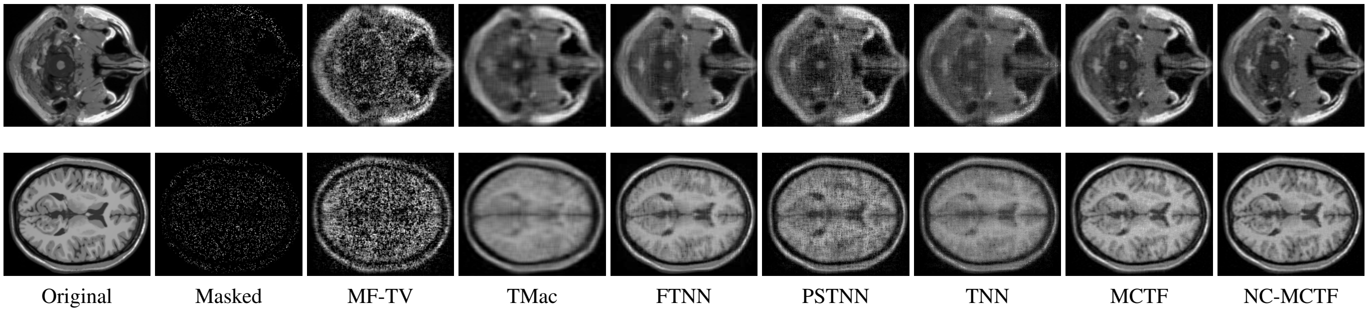


Fig. 8. Slices (7th and 83th) of recovered MRI by MF-TV, TMac, FTNN, PSTNN, TNN, our MCTF and NC-MCTF. Sampling rate is 10%.

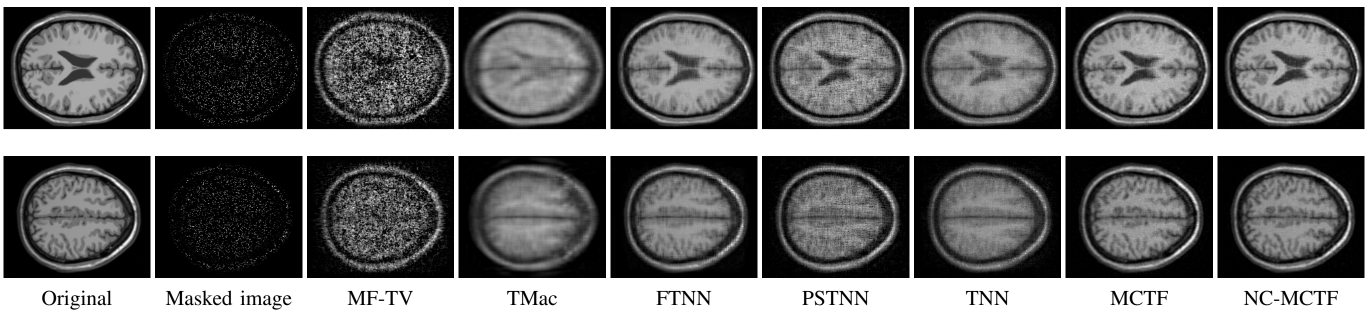


Fig. 9. Slices (95th and 118th) of MRI recovered by MF-TV, TMac, FTNN, PSTNN, TNN, our MCTF and NC-MCTF. Sampling rate is 10%.

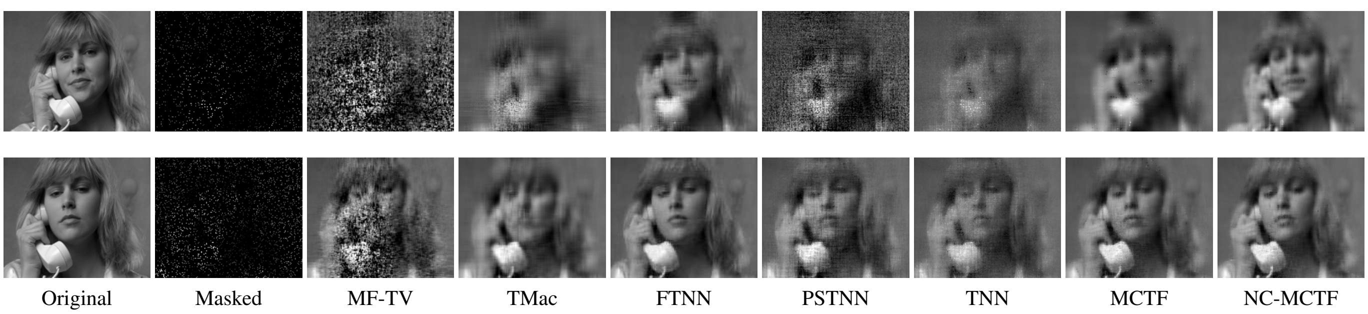


Fig. 10. Slices (94th and 10th) of recovered video for "Suzie" by MF-TV, TMac, FTNN, PSTNN, TNN, our MCTF and NC-MCTF. Sampling rates of first and second rows are 5% and 10%, respectively.

best results, and the evaluation index obtained by the proposed MCTF is superior to competitive methods.

Based on the above quantitative comparison, we conduct a quantitative evaluation of the proposed model in terms of vision. Fig. 10 and Fig. 11 show the partial slice images of restored data under different sampling conditions. The closer the restored result to the original reference image, the better the performance of the corresponding model. It can be seen from the figure that the proposed model achieved a significant advantage in restored images, especially at low sampling rates. When the sampling rate is low, the original image information contained in the input observation image is scarce. If it is desired to restore the data image relatively accurately, as at high sampling rates, it is necessary to impose additional prior constraints on the solution space of the optimization model to increase the accuracy of the obtained solution, as the proposed model does.

TABLE III
AVERAGED PSNR, SSIM, FSIM AND ERGAS OF RECOVERED RESULTS ON VIDEO "SUZIE" BY TMAC, MF-TV, TNN, FTNN, PSTNN AND OUR MCTF AND NC-MCTF AT DIFFERENT SAMPLING RATES. BEST VALUE APPEARS IN BOLDFACE.

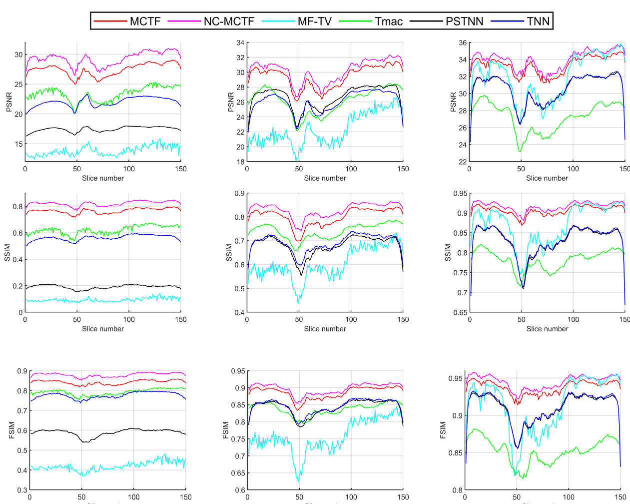
				SR =0.05				
method	noisy	MF-TV	TMac	FTNN	PSTNN	TNN	MCTF	NC-MCTF
PSNR	7.259	13.801	23.385	27.294	17.447	22.005	27.430	**29.312**
SSIM	0.009	0.094	0.622	0.465	0.192	0.563	0.766	**0.822**
FSIM	0.454	0.42	0.792	0.555	0.59	0.776	0.842	**0.880**
ERGAS	1057.282	501.117	167.927	129.27	327.678	194.844	104.955	**84.698**
				SR = 0.1				
method	noisy	MF-TV	TMac	FTNN	PSTNN	TNN	MCTF	NC-MCTF
PSNR	7.493	22.356	26.189	29.484	26.647	26.032	29.414	**30.223**
SSIM	0.014	0.605	0.74	0.585	0.68	0.692	0.801	**0.830**
FSIM	0.426	0.758	0.838	0.670	0.843	0.846	0.886	**0.897**
ERGAS	1029.096	196.059	124.369	95.472	117.104	124.923	84.888	**77.398**
				SR = 0.2				
method	noisy	MF-TV	TMac	FTNN	PSTNN	TNN	MCTF	NC-MCTF
PSNR	8.005	32.064	27.274	32.184	30.566	30.561	33.353	**33.992**
SSIM	0.02	0.872	0.782	0.721	0.829	0.831	0.906	**0.917**
FSIM	0.391	0.916	0.853	0.788	0.91	0.911	0.938	**0.945**
ERGAS	970.285	66.692	109.627	65.322	75.472	75.598	53.121	**49.395**

This article has been accepted for publication in IEEE Transactions on Multimedia. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/TMM.2022.3216746

Fig. 11. Slices (21th and 104th) of recovered video for "Hall" by MF-TV, TMac, FTNN, PSTNN, TNN, our MCTF and NC-MCTF. Sampling rates of first and second rows are 5% and 10%, respectively.

TABLE IV
AVERAGED PSNR, SSIM, FSIM AND ERGAS OF RECOVERED RESULTS ON VIDEO "Hall" BY MF-TV, TMAC, PSTNN, TNN, FTNN, OUR MCTF AND NC-MCTF AT DIFFERENT SAMPLING RATES. BEST VALUES APPEAR IN BOLDFACE.

				SR =0.05				
method	noisy	MF-TV	TMac	FTNN	PSTNN	TNN	MCTF	NC-MCTF
PSNR	4.82	13.539	22.101	**30.022**	16.075	20.78	26.215	27.415
SSIM	0.007	0.412	0.675	0.792	0.36	0.636	0.856	**0.882**
FSIM	0.387	0.612	0.789	0.835	0.672	0.792	0.890	**0.906**
ERGAS	1225.779	452.351	168.866	98.14	335.52	195.315	105.199	**91.728**
				SR = 0.1				
method	noisy	MF-TV	TMac	FTNN	PSTNN	TNN	MCTF	NC-MCTF
PSNR	5.055	24.855	26.936	**32.790**	29.014	28.433	30.731	31.481
SSIM	0.013	0.829	0.854	0.854	0.892	0.905	0.933	**0.942**
FSIM	0.393	0.873	0.888	0.889	0.934	0.936	0.945	**0.952**
ERGAS	1193.075	131.422	97.185	59.375	77.395	82.259	62.923	**57.805**
				SR = 0.2				
method	noisy	MF-TV	TMac	FTNN	PSTNN	TNN	MCTF	NC-MCTF
PSNR	5.567	33.006	27.648	**35.755**	33.629	33.691	33.052	34.097
SSIM	0.025	0.94	0.869	0.902	0.961	0.962	0.956	**0.962**
FSIM	0.403	0.954	0.897	0.927	0.973	**0.974**	0.965	0.970
ERGAS	1124.737	50.971	89.271	44.184	46.123	45.851	48.414	**43.191**

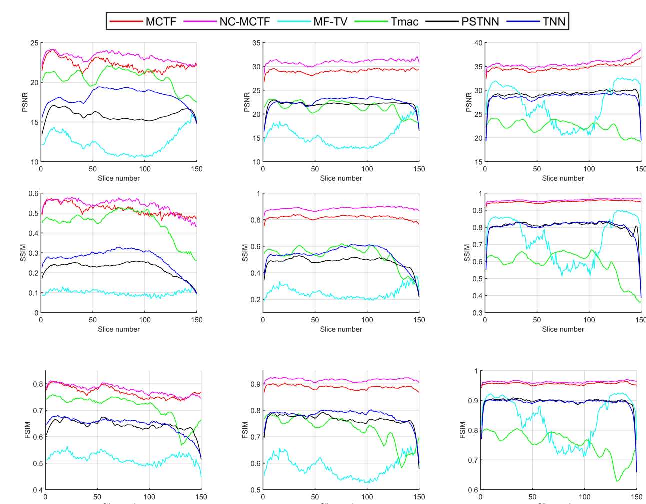


(a) SR = 0.05 (b) SR = 0.1 (c) SR = 0.2

Fig. 12. PSNR, SSIM and FSIM of recovered video "Suzie" by MF-TV, TMac, FTNN, PSTNN, TNN, our MCTF and NC-MCTF for all slices.

TABLE V
PSNR, SSIM, FSIM, ERGAS AND SAM OF RECOVERED RESULTS ON HYPERSPECTRAL IMAGE "Cuprite" BY MF-TV, TMAC, FTNN, PSTNN, TNN, OUR MCTF AND NC-MCTF AT DIFFERENT SAMPLING RATES. BEST VALUES APPEAR IN BOLDFACE.

			SR =0.025				
method	noisy	MF-TV	TMac	PSTNN	TNN	MCTF	NC-MCTF
PSNR	7.666	26.115	21.25	13.387	22.783	31.091	**31.208**
SSIM	0.007	0.539	0.412	0.124	0.554	0.771	**0.774**
FSIM	0.48	0.765	0.755	0.613	0.775	0.842	**0.847**
ERGAS	1043.633	237.074	235.594	539.574	245.333	77.458	**76.503**
SAM	81.221	12.913	7.842	17.98	9.156	2.512	**2.468**
			SR = 0.05				
method	noisy	MF-TV	TMac	PSTNN	TNN	MCTF	NC-MCTF
PSNR	7.779	34.684	28.945	20.621	26.579	34.739	**35.481**
SSIM	0.01	0.845	0.712	0.31	0.663	0.860	**0.879**
FSIM	0.471	0.915	0.846	0.735	0.836	0.907	**0.920**
ERGAS	1030.139	89.372	93.352	234.445	154.292	51.913	**48.063**
SAM	77.268	4.386	3.278	7.886	5.413	1.751	**1.653**
			SR = 0.1				
method	noisy	MF-TV	TMac	PSTNN	TNN	MCTF	NC-MCTF
PSNR	8.013	**40.888**	35.627	35.51	35.015	37.449	37.623
SSIM	0.014	**0.957**	0.885	0.907	0.897	0.912	0.913
FSIM	0.451	**0.978**	0.931	0.951	0.943	0.943	0.943
ERGAS	1002.75	**34.263**	44.518	54.421	57.537	39.232	38.546
SAM	71.695	1.46	1.445	2.072	2.192	1.452	**1.410**



(a) SR = 0.05 (b) SR = 0.1 (c) SR = 0.2

Fig. 13. PSNR, SSIM and FSIM of recovered MRI by MF-TV, TMac, PSTNN, TNN, our MCTF and NC-MCTF for all slices.

C. Hyperspectral Image

In this subsection, we choose two HSI data to apply simulation experiments. The first dataset is five sequential images[4] that were acquired by the Sentinel-2 MSI on 05/09, 15/09, 20/09, 5/10 and 15/10, 2018, in Belgium, with 20-m spatial resolution at 10 $\times$ 10 km. For this dataset, the HSIs are corrupted by various types of missing areas with cloud shapes (see the second row of Fig. 16).

The second dataset is the airborne visible/infrared imaging spectrometer (AVIRIS) copper salt data[5] with size 150 $\times$ 150

[4]https://drive.google.com/file/d/1LlvUKtUWAKoF6R0igbREwvP2 Wfja9U Bv/view

[5]http://aviris.jpl.nasa.gov/html/aviris.freedata.html

TABLE VI
AVERAGED PSNR, SSIM, FSIM AND ERGAS OF RECOVERED RESULTS ON COLOR VIDEO BY SMF-LRTC AND OUR MODEL AT DIFFERENT SAMPLING RATES.

		SR =0.1		
Method	PSNR	SSIM	FSIM	ERGAS
Sampled	7.475	0.025	0.466	1095.701
SMF-LRTC	18.241	0.498	0.733	400.4275
NC-MCTF	29.563	0.860	0.924	93.960
		SR =0.15		
Method	PSNR	SSIM	FSIM	ERGAS
Sampled	7.724	0.033	0.453	1064.745
SMF-LRTC	27.073	0.800	0.895	187.139
NC-MCTF	34.248	0.948	0.970	51.979

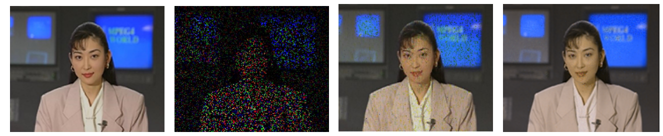
$\times$ 210. The SRs are set as follows: 0.025, 0.05 and 0.1. Here, the missing values are also randomly sampled, and we set the rank to (T_1, T_2, T_3), where T_1, T_2 and T_3 denote the number of the largest 0.5% singular values of model-1, model-2 and model-3, respectively. Because the FTNN did not perform HSI experiments, the original article of the FTNN did not describe the parameter settings of the HSI dataset; therefore, in this subsection, we will not perform comparison experiments on the FTNN.

Table V lists the PQIs of the results restored by all the test models at three different SRs. It can be clearly seen that the two proposed methods obtain the best PQIs among all of the test methods. Fig. 16 and Fig. 17 show the visual results of the ground truth, simulated cloud-covered/missing area, recovery results of TMac, MF-TV, PSTNN, TNN and the proposed NC-MCTF.

D. Color Video Data

Based on the aforementioned 3-D video, MRI and HSI data completion testing, we further consider applying the proposed model to 4-D data. Here, we focus on color videos: Akiyo [6], which has a size of 144 $\times$ 176 $\times$ 50 $\times$ 3. One of the original frames is shown in Fig. 14. The proposed model is compared with a low-rank tensor completion model using smooth matrix factorization (SMF-LRTC) [20] at two sampling rates: 10% and 15%. SMF-LRTC only shows results on 3-D data; therefore, to implement SMF-LRTC, each color frame with size $h \times w \times 3 \times k$ is reshaped to a $(hw) \times 3 \times k$ tensor [43, 51]. Fig. 14 provides a visual comparison of the original frame, the sampled frame, and the results reconstructed by SMF-LRTC and the proposed model. One can observe that SMF-LRTC completes most missing elements but still fails on some local patches, e.g., the hair of the person and the white number on the TV. Table VI lists all of the PQI values. The proposed model outperforms SMF-LRTC with respect to PSNR, SSIM, ERGAS and is slightly higher than SMF-LRTC. It is also worth noting that the advantages of our model are more obvious at a low sample rate, e.g., 0.1, which is consistent with the experimental results on the previous three-dimensional data.

[6]http://trace.eas.asu.edu/yuv/



(a) Original (b) 85% Masked (c) SMF-LRTC (d) NC-MCTF

Fig. 14. One slice (1st) of completed color video "Akiyo" by NC-MCTF and SMF-LRTC. Sampling rate is 15%.

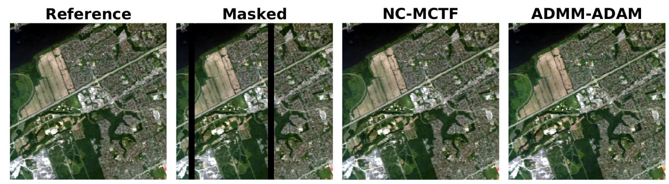


Fig. 15. Visual comparison (18th, 8th, 2nd bands) of ADMM-ADAM [52] and proposed NC-MCTF on Ottawa dataset.

TABLE VII
AVERAGED PSNR, SSIM, UIQI, ERGAS AND SAM OF ALL RECOVERED BANDS ON HYPERSPECTRAL IMAGE "Ottawa" BY ADMM-ADAM AND PROPOSED NC-MCTF WITH MASKED STRIPES.

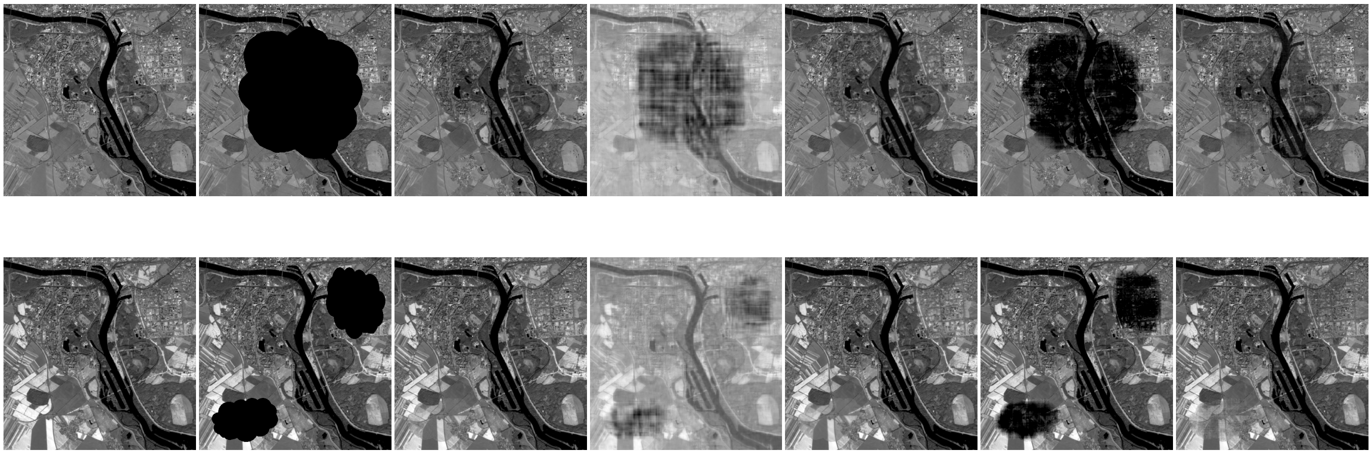
		PQI of the masked 11-60th bands				
Method	Basis	PSNR (↑)	SSIM (↑)	UIQI (↑)	ERGAS (↓)	SAM (↓)
ADMM-ADAM	ADMM+DL	45.503	0.998	0.997	0.288	0.926
NC-MCTF	Tensor theory	47.185	0.942	0.95	3.678	10.081
		PQI of all the 172 bands				
ADMM-ADAM	ADMM+DL	42.248	0.947	0.985	1.552	0.926
NC-MCTF	Tensor theory	60.203	0.985	0.985	1.984	10.081

E. Comparison with Deep Learning

With the successful application of deep learning-based methods in image processing tasks, deep learning has also been introduced to tensor completion. To claim the superiority of the proposed method, an experimental comparison with the latest techniques is necessary. Therefore, in this subsection, we compare the proposed model with the latest tensor completion technique combining the advantages from both convex optimization and deep learning, i.e., ADMM-ADAM [52, 53]. The selected dataset Ottawa with a size of $256 \times 256 \times 172$ is masked with a stripe, the stripe patterns are set as [52], and 50 continuous bands are masked with stripes, where bands 11-13 are completely missing. For a fair comparison, in this subsection, all employed datasets, mask patterns and evaluation metrics are the same as in [52]. As shown in Fig. 15 and Table VII, both the proposed method and ADMM-ADAM successfully reconstruct the masked image.

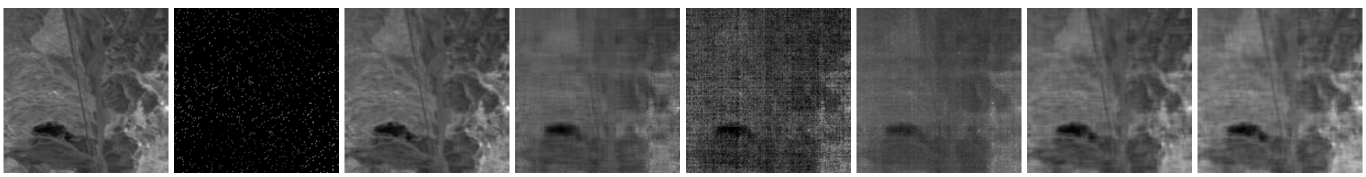
F. Parameter selection and running-time analysis

Since the proposed method consists of two balanced terms, i.e., $\tau_n \|\mathbf{X}_n\|_{* \text{ or } \log} + \lambda_n \|\mathcal{G}_n\|_{\Lambda_n, * \text{ or } \log}$, which need parameters to trade off them, it is necessary to discuss the issue of setting the parameter appropriately. To reduce the workload of adjusting parameters, we fix one of τ_n and λ_n (to enhance the generalization ability of the parameters. Here, we set the same τ_n and λ_n for different n) and then indirectly adjust the ratio of the two, i.e., $C = \frac{\tau_n}{\lambda_n}$. In this subsection, we provide some experiments with real tensor data to analyze this problem. We set the sampling rate to 0.05, 0.2 and 0.3. Under the above three different sampling rates, Fig. 18 visually



Original | Cloud-covered | NC-MCTF | TMac | MF-TV | PSTNN | TNN

Fig. 16. Sentinel-2 MSI multitemporal data sequence cloud removal experiments (20-m spatial resolution; 10 × 10 km; five temporal images) over Mechelen in Belgium.



Original | 95% Masked | MF-TV | TMac | PSTNN | TNN | MCTF | NC-MCTF

Fig. 17. One slice of recovered HSI "Cuprite" by MF-TV, TMac, PSTNN, TNN, our MCTF and NC-MCTF. Sampling rate is 5%.

TABLE VIII
RUNNING TIME (IN MINUTES) OF TMAC, MF-TV, TNN, FTNN, PSTNN, PROPOSED NC-MCTF AND DEEP-LEARNING-BASED ADMM-ADAM ON VIDEO Suzie.

Dataset	Size	ADMM-ADAM	TNN	TMac	FTNN	PSTNN	MF-TV	NC-MCTF
Video-Suzie	$144 \times 176 \times 150$	0.07	1.1	0.3	4.6	1.2	17.1	27.8

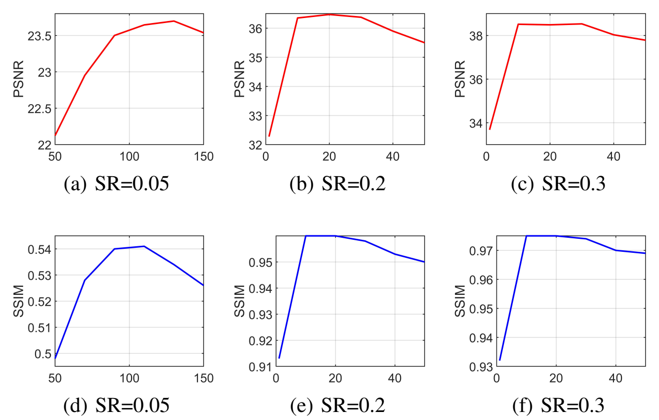


Fig. 18. Performance variation of proposed method in terms of NC-MCTF on different C and different sampling rates.

shows the performance of the proposed method under different settings of C. In addition, for the given rank $r = (r_1, r_2, r_3)$ of the blind mode-n product decomposition, we configure it by using an adaptive rank estimation metric, i.e., approximating the rank of mode-n r_n according the numbers of the largest $0.1\% - 1\%$ singular values [26, 45, 54, 55].

Table VIII lists the running times. For all of these large-scale datasets, all tested methods can complete the reconstruction in less than 30 minutes. Although the proposed NC-MCTF is slower than MF-TV because the modal decomposition and factorial low-rank constraints are parallel for each mode, we can accelerate the proposed algorithm by computing in parallel to compute each mode at the same time.

VII. CONCLUSIONS

In this paper, based on basic tensor decomposition theory, we defined a novel tensor decomposition to explore the multimode low-rank structure of underlying tensors. Then, we studied the structural characteristics of the factors obtained by the decomposition and proposed a novel tensor low-rankness measure. Furthermore, we performed nonconvex relaxation on the submeasure and obtained a better low-rankness measure. We developed an effective method to calculate the optimization problem corresponding to the proposed model and establish the convergence of our algorithm. Numerical experiments verified our theory, and the results of hyperspectral imaging, MRI and video proved that our algorithms can recover a variety of low-rank tensors with significantly fewer samples than the compared methods. It is also worth noting that rank estimation is the key to the LR-based metric but also a challenge. Therefore, in the future, it is necessary to design a fully adaptive rank approximation model with respect to different data to make the LR model effective to industry.

REFERENCES

[1] A. Shashua, "On photometric issues in 3d visual recognition from a single 2d image," *International Journal of Computer Vision*, vol. 21, no. 1-2, pp. 99–122, 1997.

[2] S. Raychaudhuri, J. M. Stuart, and R. B. Altman, "Principal components analysis to summarize microarray experiments: application to sporulation time series," in *Biocomputing 2000*. World Scientific, 1999, pp. 455–466.

[3] F. Heide, M. Steinberger, Y.-T. Tsai, M. Rouf, D. Pajak, D. Reddy, O. Gallo, J. Liu, W. Heidrich, K. Egiazarian *et al.*, "Flexisp: A flexible camera image processing framework," *ACM Transactions on Graphics (TOG)*, vol. 33, no. 6, pp. 1–13, 2014.

[4] A. M. Teodoro, J. M. Bioucas-Dias, and M. A. Figueiredo, "Image restoration and reconstruction using variable splitting and class-adapted image priors," in *2016 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2016, pp. 3518–3522.

[5] H. Zeng, X. Xie, H. Cui, H. Yin, and J. Ning, "Hyperspectral image restoration via global l 1-2 spatial–spectral total variation regularized local low-rank tensor recovery," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 4, pp. 3309–3325, 2020.

[6] Q. Xie, Q. Zhao, D. Meng, and Z. Xu, "Kronecker-basis-representation based tensor sparsity and its applications to tensor recovery," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 8, pp. 1888–1902, 2017.

[7] H.-J. Zeng, X.-Z. Xie, K. Wen-Feng, S. Cui, and J.-F. Ning, "Hyperspectral image denoising via combined non-local self-similarity and local low-rank regularization," *IEEE Access*, vol. 8, pp. 50 190–50 208, 2020.

[8] Y. Wu, H. Tan, Y. Li, F. Li, and H. He, "Robust tensor decomposition based on cauchy distribution and its applications," *Neurocomputing*, vol. 223, pp. 107–117, 2017.

[9] A. Cichocki, D. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, and H. A. Phan, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE Signal Processing Magazine*, vol. 32, no. 2, pp. 145–163, 2015.

[10] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, 1966.

[11] J. Xue, Y. Zhao, W. Liao, J. C.-W. Chan, and S. G. Kong, "Enhanced sparsity prior model for low-rank tensor completion," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 11, pp. 4567–4581, 2019.

[12] R. A. Harshman and M. E. Lundy, "Parafac: Parallel factor analysis," *Computational Statistics & Data Analysis*, vol. 18, no. 1, pp. 39–72, 1994.

[13] M. E. Kilmer, K. Braman, N. Hao, and R. C. Hoover, "Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging," *SIAM Journal on Matrix Analysis and Applications*, vol. 34, no. 1, pp. 148–172, 2013.

[14] Y. Chen, X. Xiao, and Y. Zhou, "Jointly learning kernel representation tensor and affinity matrix for multi-view clustering," *IEEE Trans. Multimedia*, vol. 22, no. 8, pp. 1985–1997, 2019.

[15] H. Zeng, X. Xie, and J. Ning, "Hyperspectral image denoising via global spatial-spectral total variation regularized nonconvex local low-rank tensor approximation," *Signal Processing*, vol. 178, p. 107805, 2021.

[16] I. V. Oseledets, "Tensor-train decomposition," *SIAM Journal on Scientific Computing*, vol. 33, no. 5, pp. 2295–2317, 2011.

[17] Q. Zhao, G. Zhou, S. Xie, L. Zhang, and A. Cichocki, "Tensor ring decomposition," *arXiv preprint arXiv:1606.05535*, 2016.

[18] L. Zhang, L. Song, B. Du, and Y. Zhang, "Nonlocal low-rank tensor completion for visual data," *IEEE Transactions on Cybernetics*, vol. 51, no. 2, pp. 673–685, 2019.

[19] L. Song, B. Du, L. Zhang, L. Zhang, J. Wu, and X. Li, "Nonlocal patch based t-svd for image inpainting: Algorithm and error analysis," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, 2018.

[20] Y.-B. Zheng, T.-Z. Huang, T.-Y. Ji, X.-L. Zhao, T.-X. Jiang, and T.-H. Ma, "Low-rank tensor completion via smooth matrix factorization," *Applied Mathematical Modelling*, vol. 70, pp. 677–695, 2019.

[21] T.-Y. Ji, X.-L. Zhao, and D.-L. Sun, "Low-rank tensor completion method for implicitly low-rank visual data," *IEEE Signal Processing Letters*, vol. 29, pp. 1162–1166, 2022.

[22] Y. Chen, S. Wang, X. Xiao, Y. Liu, Z. Hua, and Y. Zhou, "Self-paced enhanced low-rank tensor kernelized multi-view subspace clustering," *IEEE Trans. Multimedia*, 2021.

[23] J. Xue, Y. Zhao, S. Huang, W. Liao, J. C.-W. Chan, and S. G. Kong, "Multilayer sparsity-based tensor decomposition for low-rank tensor completion," *IEEE Transactions on Neural Networks and Learning Systems*, 2021.

[24] J. Håstad, "Tensor rank is np-complete," *Journal of algorithms (Print)*, vol. 11, no. 4, pp. 644–654, 1990.

[25] J. Liu, P. Musialski, P. Wonka, 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, 2012.

[26] Y. Xu, R. Hao, W. Yin, and Z. Su, "Parallel matrix factorization for low-rank tensor completion," *arXiv preprint arXiv:1312.1254*, 2013.

[27] H. Zeng, X. Xie, and J. Ning, "Tensor completion using enhanced multiple modes low-rank prior and total variation," *arXiv preprint arXiv:2004.08747*, 2020.

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

[29] M. E. Kilmer, K. Braman, N. Hao, and R. C. Hoover, "Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging," *SIAM Journal on Matrix Analysis and Applications*, vol. 34, no. 1, pp. 148–172, 2013.

[30] 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, 2017.

[31] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Journal of the ACM (JACM)*, vol. 58, no. 3, pp. 1–37, 2011.

[32] A. Eriksson and A. Van Den Hengel, "Efficient computation of robust low-rank matrix approximations in the presence of missing data using the l 1 norm," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, 2010, pp. 771–778.

[33] Z. Zhang and S. Aeron, "Exact tensor completion using t-svd," *IEEE Transactions on Signal Processing*, vol. 65, no. 6, pp. 1511–1526, 2016.

[34] 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," *Journal of Computational and Applied Mathematics*, vol. 372, p. 112680, 2020.

[35] 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.

[36] H. Zeng, X. Xie, H. Cui, Y. Zhao, and J. Ning, "Hyperspectral image restoration via cnn denoiser prior regularized low-rank tensor recovery," *Computer Vision and Image Understanding*, vol. 197-198, p. 103004, 2020.

[37] Y. Wang, J. Peng, Q. Zhao, Y. Leung, X.-L. Zhao, and D. Meng, "Hyperspectral image restoration via total variation regularized low-rank tensor decomposition," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 4, pp. 1227–1243, 2017.

[38] Y. Xie, D. Tao, W. Zhang, Y. Liu, L. Zhang, and Y. Qu, "On unifying multi-view self-representations for clustering by tensor multi-rank minimization," *International Journal of Computer Vision*, vol. 126, no. 11, pp. 1157–1179, 2018.

[39] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, and S. Yan, "Tensor robust principal component analysis: Exact recovery of corrupted

low-rank tensors via convex optimization," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 5249–5257.

[40] C. Lu, J. Feng, Z. Lin, and S. Yan, "Exact low tubal rank tensor recovery from gaussian measurements," *arXiv preprint arXiv:1806.02511*, 2018.

[41] L. De Lathauwer and B. De Moor, "From matrix to tensor: Multilinear algebra and signal processing," in *Institute of mathematics and its applications conference series*, vol. 67. Citeseer, 1998, pp. 1–16.

[42] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm minimization with application to image denoising," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 2862–2869.

[43] H. Zeng, Y. Chen, X. Xie, and J. Ning, "Enhanced nonconvex low-rank approximation of tensor multi-modes for tensor completion," *IEEE Transactions on Computational Imaging*, vol. 7, pp. 164–177, 2021.

[44] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," *SIAM Journal on Optimization*, vol. 23, no. 2, pp. 1126–1153, 2013.

[45] T.-Y. Ji, T.-Z. Huang, X.-L. Zhao, T.-H. Ma, and G. Liu, "Tensor completion using total variation and low-rank matrix factorization," *Information Sciences*, vol. 326, pp. 243–257, 2016.

[46] Q. Huynh-Thu and M. Ghanbari, "Scope of validity of psnr in image/video quality assessment," *Electronics Letters*, vol. 44, no. 13, pp. 800–801, 2008.

[47] 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.*, vol. 13, no. 4, pp. 600–612, 2004.

[48] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, 2011.

[49] L. Wald, *Data fusion: definitions and architectures: fusion of images of different spatial resolutions*. Presses des MINES, 2002.

[50] C.-H. Lin, F. Ma, C.-Y. Chi, and C.-H. Hsieh, "A convex optimization-based coupled nonnegative matrix factorization algorithm for hyperspectral and multispectral data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1652–1667, 2017.

[51] 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, 2019.

[52] C.-H. Lin, Y.-C. Lin, and P.-W. Tang, "Admm-adam: A new inverse imaging framework blending the advantages of convex optimization and deep learning," *IEEE Trans. Geosci. Remote Sens.*, 2021.

[53] C.-H. Lin, Y.-C. Lin, P.-W. Tang, and M.-C. Chu, "Deep hyperspectral tensor completion just using small data," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*. IEEE, 2021, pp. 2480–2483.

[54] M.-J. Lai, Y. Xu, and W. Yin, "Improved iteratively reweighted least squares for unconstrained smoothed \ell_q minimization," *SIAM Journal on Numerical Analysis*, vol. 51, no. 2, pp. 927–957, 2013.

[55] Z. Wen, W. Yin, and Y. Zhang, "Solving a low-rank factorization model for matrix completion by a nonlinear successive over-relaxation algorithm," *Mathematical Programming Computation*, vol. 4, no. 4, pp. 333–361, 2012.