Multi-Dimensional Visual Data Completion via Low-Rank Tensor Representation Under Coupled Transform

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

***Abstract*— This paper addresses the tensor completion problem, which aims to recover missing information of multi-dimensional images. How to represent a low-rank structure embedded in the underlying data is the key issue in tensor completion. In this work, we suggest a novel low-rank tensor representation based on coupled transform, which fully exploits the spatial multi-scale nature and redundancy in spatial and spectral/temporal dimensions, leading to a better low tensor multi-rank approximation. More precisely, this representation is achieved by using two-dimensional framelet transform for the two spatial dimensions, one/two-dimensional Fourier transform for the temporal/spectral dimension, and then Karhunen–Loéve transform (via singular value decomposition) for the transformed tensor. Based on this low-rank tensor representation, we formulate a novel low-rank tensor completion model for recovering missing information in multi-dimensional visual data, which leads to a convex optimization problem. To tackle the proposed model, we develop the alternating directional method of multipliers (ADMM) algorithm tailored for the structured optimization problem. Numerical examples on color images, multispectral images, and videos illustrate that the proposed method outperforms many state-of-the-art methods in qualitative and quantitative aspects.**

***Index Terms*— 2D framelet transform, multi-scale representation, tensor nuclear norm, tensor completion.**

Manuscript received February 29, 2020; revised November 29, 2020; accepted February 18, 2021. Date of publication March 8, 2021; date of current version March 12, 2021. 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 the HKRGC under Grant GRF 12200317, Grant 12300218, Grant 12300519, and Grant 17201020. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Chandra Sekhar Seelamantula. *(Corresponding authors: Ting-Zhu Huang; Xi-Le Zhao.)*

Jian-Li Wang, Ting-Zhu Huang, and Xi-Le Zhao are with the School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: wangjianli_123@163.com; tingzhuhuang@126.com; xlzhao122003@163.com).

Tai-Xiang Jiang is with the Financial Intelligence and Financial Engineering Research Key Laboratory of Sichuan Province, FinTech Innovation Center, School of Economic Information Engineering, Southwestern University of Finance and Economics, Chengdu 610074, China (e-mail: taixiangjiang@gmail.com).

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

Digital Object Identifier 10.1109/TIP.2021.3062995

I. Introduction

As THE high-dimensional extension of vector/matrix [1]–[3], tensor provides a more diverse and flexible representation for multi-dimensional visual data, which usually contains two spatial dimensions and another temporal (or spectral) dimension, such as color images [4], videos [5], [6], hyperspectral images [7]–[13], seismic data [14], etc. Unfortunately, due to hardware restrictions and various degradations, the obtained data are usually incomplete, which significantly degrades the visual quality and limits the subsequent processing tasks. The problem of recovering the missing information in multi-dimensional visual data from its small known observations is called tensor completion (TC) [15]–[18], which is a typical inverse problem in image processing. An effective restoration process generally relies on prior knowledge about the desired solution.

Low-rankness is a powerful tool to describe the internal redundancy of tensor, and how to exploit the embedded low-rank structure has been widely studied in TC [19]. Mathematically, a unified low-rank tensor completion (LRTC) model can be written as

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

where $\mathcal{X} \in \mathbb{R}^{n_1 \times \cdots \times n_k}$ is the underlying tensor, $\mathcal{O} \in \mathbb{R}^{n_1 \times \cdots \times n_k}$ is the observed data, Ω is the index set of the observed data, and $\mathcal{P}_{\Omega}(\cdot)$ is the sampling operator that remains the elements in Ω while making the others to be zeros. In LRTC, a central issue is the definition of the tensor rank, i.e., how to characterize the correlations along different dimensions within a data. However, it is a difficult task due to the complex algebraic structure of tensors, and there is no universally accepted solution to the definition of tensor rank [20]. Over the past decades, many research efforts have been devoted to this topic, such as CANDECOMP/PARAFAC (CP) rank [21], [22], Tucker rank [23], [24], and tubal rank [25], [26], all of which capture the tensor redundancy from their perspectives.

The CP decomposition [22] decomposes a tensor into a sum of rank-one factors constructed by the vector outer product, and the CP-rank is defined as the smallest number of these factors among all such decompositions. However, calculating the CP-rank of a known tensor is an NP-hard problem [27], [28],

and there is generally no accurate algorithm for estimating CP-rank. The Tucker decomposition factorizes a tensor into a product of a core tensor and factor matrices, and the Tucker rank (also called as "n-rank") [23], [24], [29] is a vector composed of the ranks of unfolding matrices of the target tensor along different dimensions [30]. However, Tucker rank suffers from a limitation that the global correlation within the tensor is destroyed by its unfolding scheme.

Recently, a new tensor decomposition method called the tensor singular value decomposition (t-SVD) [25], [26] was proposed for third-order tensors, which decomposes a three-dimensional tensor into the product of two orthogonal tensors and one f-diagonal tensor (see Section II for details). t-SVD has the advantage of characterizing the correlation along the third dimension via constructing group-rings along the tensor fibers, i.e., executing the 1D fast Discrete Fourier Transform (DFT) along the third mode. Based on the t-SVD framework, a new tensor tubal rank [31], [32] is defined as the number of non-zero tubes of the f-diagonal tensor in t-SVD. However, direct minimizing tubal rank is an NP-hard problem. As the convex surrogate of tensor tubal rank, tensor nuclear norm (TNN) [25], [33], [34] is considered for tubal rank minimization with a strong theoretical guarantee for LRTC.

Later works [35] have been proposed to improve the performance of TNN. Considering that TNN is a biased approximation [36]–[38] to the tubal rank, some works replace the nuclear norm with non-convex surrogates [36]–[38] to obtain a better low-rank approximation, including weighted tensor nuclear norm [39], Laplace function [40], and the partial sum of tubal nuclear norm [37]. Other works focus on a more effective characterization of the data correlation by finding a more suitable transform than DFT, such as invertible linear transform [41], [42], multiple linear transform [43], [44], and framelet transform [45].

Recently, deep learning (DL)-based completion methods [46]–[48] have been rapidly developed to learn deep image priors from a large number of example images and have shown promising performance due to its high capacity. However, most of these methods are designed for particular tasks, e.g., inpainting (tube-wise sampling) [46]–[48], and their performance is essentially dependent upon the diversity and volume of training datasets. Therefore, the lack of generalization hinders its direct application to the general tensor completion problem for diverse samplings and data compared with regularized LRTC methods.

The above-mentioned LRTC methods only consider the global data correlation (i.e., the low-rankness) but ignore the other significant properties, such as the spatial multi-scale nature and redundancy in the spatial and spectral/temporal dimensions. As a result, those methods can only recover the global cartoon but fail to repair specific textures, especially when the sampling rate is low. To fully exploit the intrinsic structure within multi-dimensional visual data, we propose a novel low-rank tensor representation under coupled transform, which can characterize the correlations along different dimensions in a unified framework. The main idea is to explore suitable transforms to decorrelate the spatial and temporal (or spectral) dimensions, achieving an enhanced low-rank representation of the underlying data. More precisely, the proposed representation involves three layers: $\mathcal{B} = \mathcal{B}_3 \circ \mathcal{B}_2 \circ \mathcal{B}_1$.

- In the first layer $\mathcal{B}_1$, we use a two-dimensional framelet transform with abundant bases to describe the local spatial correlation, which has two advantages. First, with the help of the filters to construct the framelet system, the abundant spatial structures can be elaborately expressed with respect to different frequencies and spatial directions. Second, the multi-level nature of the framelet decomposition helps to capture the multi-scale features of the tensor, as shown in Fig. 1 (a).
- In the second layer $\mathcal{B}_2$, we use a Fourier transform to characterize the global correlation of the temporal (or spectral) dimension. This layer characterizes the temporal (or spectral) global correlation by constructing group-rings along the corresponding dimensions and reveals the intrinsic low-rank property of the multi-scale representation coefficients achieved by the first layer, in which the spatial features of different directions and levels are gathered together with strong correlations.
- In the third layer $\mathcal{B}_3$, we use a Karhunen–Loéve (KL) transform (via singular value decomposition), regarded as a special transform, to characterize the global spatial correlation (i.e., low-rankness) of the representation coefficients by the first two layers, which incorporates both the global correlation and local geometric details in a unified framework. The sparsity of the result after KL transform is indeed the rank of the spatial slice of the framelet and Fourier transformed result.

Fig. 1 gives an illustration of the proposed representation on the color image "Butterfly", in which Fig. 1 (a) shows the third-layer structure of the proposed representation; Fig. 1 (b) shows the comparison of the accumulation energy ratio (AccEgy $= \sum_{i=1}^{k} \sigma_i^2 / \sum_j \sigma_j^2$, where σ_i is the i-th singular value) of the proposed representation and t-SVD. Form Fig. 1 (b), we can observe that the proposed representation always achieve higher structural similarity (SSIM) values and lower relative squared error (RSE) than t-SVD on the same AccEgy, which implies that our representation coefficients exhibit higher sparsity, i.e., the representation is more efficient in terms of data compression.

Based on the above sparsity of the representation coefficients, we formulate a novel model for LRTC under coupled transform, termed as CT-LRTC, which is expressed as follows:

$$\begin{aligned} &\min_{\mathcal{X}} \ \|\mathcal{B}(\mathcal{X})\|_1 \\ &\text{s.t.} \ \ \mathcal{P}_\Omega(\mathcal{X}) = \mathcal{P}_\Omega(\mathcal{O}), \end{aligned}$$

where $\mathcal{B} = \mathcal{B}_3 \circ \mathcal{B}_2 \circ \mathcal{B}_1$ is coupled transforms which involves three layer, i.e., the framelet transform, Fourier transform, and KL transform.

Our work also provides a unified framework that can accommodate the previous TNN-based methods, as well as a deconstruction perspective to view their inner mechanisms. In our framework, they can be viewed as a compound of transforms, such as FFT+KL (TNN) [33] and DCT+KL [41] (see Section II-B for more details).

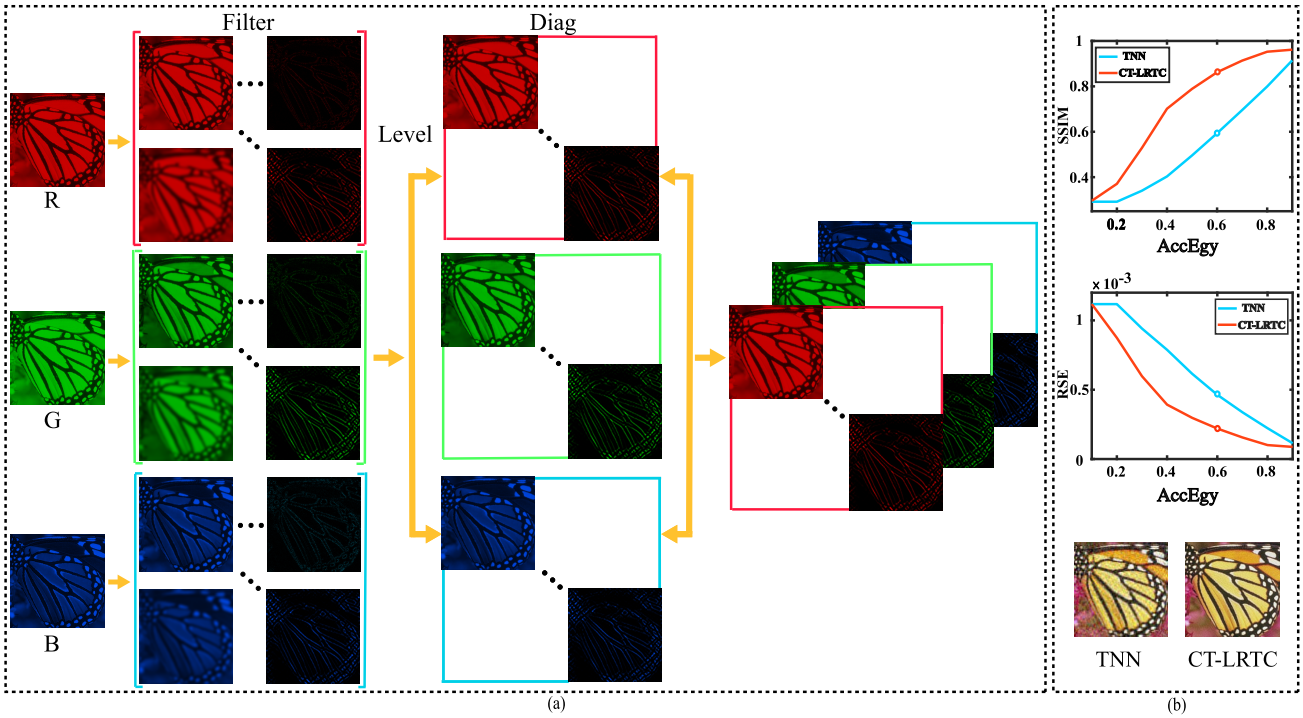


Fig. 1. An illustration of the proposed CT-LRTC on the color image "Butterfly". (a) Shows the three-layer structure of the proposed representation. (b) Compares the accuracy of the proposed representation and t-SVD with respect to the change of AccEgy and shows the image constructed by AccEgy $= 0.6$.

The main contributions of this work are summarized as follows.

- We propose a novel low-rank tensor representation under coupled transform for multi-dimensional visual data, which provides an enhanced low-rank approximation and a novel perspective to exploit the implicit low-rank structure, i.e., being not directly low-rank in the original domain but low-rank in the transformed domain with well-chosen transformation.
- We formulate a novel low-rank tensor completion model for recovering missing information in multi-dimensional visual data based on the proposed low-rank tensor representation, and develop an efficient alternating directional method of multipliers (ADMM) [49] solving algorithm.
- Numerical examples on color images, multispectral images, and videos illustrate the superiority and effectiveness of the proposed CT-LRTC compared with many existing methods.

The rest of this paper is organized as follows. Section II gives some preliminaries of tensor used throughout this paper. Section III presents the proposed CT-LRTC and the corresponding ADMM solver in detail. Section IV displays experimental results to verify the effectiveness of the proposed CT-LRTC. Finally, Section V concludes our work.

II. Preliminaries

In this section, we give some preliminary knowledge of tensor and t-SVD used throughout this paper.

A. Tensor Notations

For a better reading, we summarize some basic notations of tensors in Table I. Below we introduce some necessary definitions; see [41], [50] for more details.

TABLE I
NOTATIONS ABOUT A THIRD-ORDER TENSOR

Notation	Description
x, $\mathbf{x}$, $\mathbf{X}$, $\mathcal{X}$	Scalar, vector, matrix, tensor.
$\mathcal{I}$	The identity tensor, whose first frontal slice is the identity matrix, and the other frontal slices are zero matrices.
$\mathbf{X}^{(i)}$	The i-th frontal slice of $\mathcal{X}$.
$\text{unfold}_n(\mathcal{X})/\mathbf{X}_{(n)}$	The mode-n unfolding of $\mathcal{X}$.
$\text{fold}_n(\mathbf{X})$	The folding operation of $\mathbf{X}$ along the mode-n.
$\bar{\mathcal{X}}$	The DFT result of $\mathcal{X}$ along its third dimension e.g., $\bar{\mathcal{X}} = \text{fft}(\mathcal{X}, [], 3)$.
$\|\mathcal{X}\|_F$	The Frobenius norm of $\mathcal{X}$, which is the square root of the squared sum of each element of $\mathcal{X}$, i.e., $\|\mathcal{X}\|_F = \sqrt{\sum_{ijk} x_{ijk}^2}$.
$\|\mathcal{X}\|_{\text{TNN}}$	The TNN of $\mathcal{X}$, which is the sum of the nuclear norm of all frontal slices of $\bar{\mathcal{X}}$, i.e., $\|\mathcal{X}\|_{\text{TNN}} = \sum_{i=1}^{n_3} \|\bar{X}^{(i)}\|_*$.
$\mathcal{X}^{\mathbf{H}}$	The conjugate transpose of $\mathcal{X}$, which is obtained by conjugate transposing each frontal slice and then reversing the order of the transposed frontal slices 2 through n_3, i.e., $(\mathbf{X}^{(1)})^{\mathbf{H}} = (\mathbf{X}^{\mathbf{H}})^{(1)}$ and $(\mathbf{X}^{(n_3+2-i)})^{\mathbf{H}} = (\mathbf{X}^{\mathbf{H}})^{(i)}$ $(i = 2, \cdots, n_3)$.

B. The t-SVD Framework

In this subsection, we briefly introduce the t-SVD framework since the proposed CT-LRTC is somewhat related to TNN. Here are some related definitions.

Definition 2.1 (t-Product [41]): The tensor-tensor product of $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and $\mathcal{Y} \in \mathbb{R}^{n_2 \times n_4 \times n_3}$ is defined as

$$\mathcal{X} * \mathcal{Y} = \text{Fold}(\text{bcirc}(\mathcal{X}) \cdot \text{Unfold}(\mathcal{Y})) \in \mathbb{R}^{n_1 \times n_4 \times n_3},$$

where $\text{bcirc}(\mathcal{X}) \in \mathbb{R}^{n_1 n_3 \times n_2 n_3}$ is the following block circulant matrix

$$\text{bcirc}(\mathcal{X}) := \begin{bmatrix} X^{(1)} & X^{(n_3)} & \dots & X^{(2)} \\ X^{(2)} & X^{(1)} & \dots & X^{(3)} \\ \vdots & \vdots & \ddots & \vdots \\ X^{(n_3)} & X^{(n_3-1)} & \dots & X^{(1)} \end{bmatrix},$$

$\text{Unfold}(\mathcal{Y})$ is a matrix of size $n_2\ n_3 \times n_4$:

$$\text{Unfold}(\mathcal{Y}) = \begin{bmatrix} \mathcal{Y}^{(1)} \\ \mathcal{Y}^{(2)} \\ \vdots \\ \mathcal{Y}^{(n_3)} \end{bmatrix}, \quad \text{Fold}(\text{Unfold}(\mathcal{Y})) = \mathcal{Y}.$$

When $n_3 = 1$, the t-product degenerates to a matrix-matrix product.

Definition 2.2 (Orthogonal Tensor [50]): A tensor $\mathcal{Q} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is called an orthogonal tensor, if it satisfies $\mathcal{Q}^{\mathbf{H}} * \mathcal{Q} = \mathcal{Q} * \mathcal{Q}^{\mathbf{H}} = \mathcal{I}$.

Definition 2.3 (f-Diagonal Tensor [50]): A tensor is called an f-diagonal tensor, if each of its frontal slices is a diagonal matrix.

Definition 2.4 (t-SVD [41]): Every tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ can be factorized as

$$\mathcal{X} = \mathcal{U} * \mathcal{S} * \mathcal{V}^{\mathbf{H}},$$

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

Definition 2.5 (Tensor Multi Rank [33]): The multi rank of $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is a vector $r \in \mathbb{R}^{n_3}$, whose i-th element equals to the rank of i-th frontal slice of $\bar{\mathcal{X}}$, i.e., $r_i = \text{rank}(\bar{X}^{(i)})$.

Definition 2.6 (Tensor Tubal Rank [33]): The tubal rank of $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is defined as the number of non-vanishing tubes of $\mathcal{S}$ in $\mathcal{X} = \mathcal{U} * \mathcal{S} * \mathcal{V}^{\mathbf{H}}$. That is

$$\text{rank}_{tubal}(\mathcal{X}) = \sharp\{i : \mathcal{S}(i, i, :) \neq 0\}.$$

III. Low-Rank Tensor Completion Under Coupled Transform

This section is divided into three parts. Subsection III-A gives the proposed CT-LRTC for tensor completion. Subsection III-B provides justification for several key ideas of coupled transform. Subsection III-C provides the corresponding ADMM [49] algorithm with guaranteed convergence in details.

A. The Proposed CT-LRTC

Based on the above sparsity of the representation coefficients, we formulate a novel model for LRTC under coupled transforms, termed as CT-LRTC, which is as follows:

$$\begin{aligned} \min_{\mathcal{X}} \ & \|\mathcal{B}(\mathcal{X})\|_1 \\ \text{s.t.} \ & \mathcal{P}_\Omega(\mathcal{X}) = \mathcal{P}_\Omega(\mathcal{O}), \end{aligned}$$

where $\mathcal{X} \in \mathbb{R}^{n_1 \times \cdots \times n_k}$ is the underlying tensor, $\mathcal{O} \in \mathbb{R}^{n_1 \times \cdots \times n_k}$ is the observed data, Ω is the index set of the observed data, $\mathcal{P}_\Omega(\cdot)$ is the sampling operator that remains the elements in Ω while making the others to be zeros, and $\mathcal{B} = \mathcal{B}_3 \circ \mathcal{B}_2 \circ \mathcal{B}_1$ is a coupled transform which involves three layer, i.e., the framelet transform, Fourier transform, and KL transform.

Algorithm 1 Coupled Transform for a Third-Order Tensor

Input: $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$.
Step 1: $\mathcal{W}(\mathcal{X}) \leftarrow \text{fold}_4((\mathbf{W}\mathbf{X}_{(3)}^\top)^\top)$;
$\mathcal{B}_1(\mathcal{X}) \leftarrow \mathcal{L}(\mathcal{W}(\mathcal{X}))$;
Step 2: $\mathcal{B}_2(\mathcal{B}_1(\mathcal{X})) :\leftarrow \text{fft}(\mathcal{B}_1(\mathcal{X}), [], 3)$;
Step 3: $\mathcal{S} \leftarrow \mathcal{B}_2(\mathcal{B}_1(\mathcal{X}))$;
for $i = 1$ to n_3
$[\mathbf{U}, \mathbf{S}, \mathbf{V}] = \text{svd}\,(\mathcal{B}_2(\mathcal{B}_1(\mathcal{X}))(:, :, i))$;
$\mathcal{S}(:, :, j) = \mathbf{S}$;
end
Output: $\mathcal{B}(\mathcal{X}) \leftarrow \mathcal{S}$.

In the first layer $\mathcal{B}_1$, the framelet transform is conducted on the spatial slices, which can be expressed as a composition of two operators, i.e., $\mathcal{B}_1(\mathcal{X}) = \mathcal{L}(\mathcal{W}(\mathcal{X}))$. Here, $\mathcal{W} : \mathbb{R}^{n_1 \times n_2 \times n_3 \times \cdots \times n_k} \rightarrow \mathbb{R}^{m^2 l \times n_1 \times n_2 \times n_3 \times \cdots \times n_k}$ is a framelet transform operator, i.e., $\mathcal{W}(\mathcal{X}) = \text{fold}_4((\mathbf{W}\mathbf{X}_{(3)}^\top)^\top)$, where $\mathbf{W} \in \mathbb{R}^{m^2 l n_1 n_2 \times n_1 n_2}$ is a framelet transform matrix with m filters and l levels, and $\mathcal{L} : \mathbb{R}^{m^2 l \times n_1 \times n_2 \times n_3 \times \cdots \times n_k} \rightarrow \mathbb{R}^{m^2 l n_1 \times m^2 l n_2 \times n_3 \times \cdots \times n_k}$ is a reshape operator, i.e., rearranges $\mathcal{W}(\mathcal{X})$ in a diagonal manner which regards $\mathcal{W}(\mathcal{X})$ as $m^2 l$ k-order tensors. In fact, $\mathcal{B}_1 : \mathbb{R}^{n_1 \times n_2 \times n_3 \times \cdots \times n_k} \rightarrow \mathbb{R}^{m^2 l n_1 \times m^2 l n_2 \times n_3 \times \cdots \times n_k}$ is a linear operator. Here, the two-dimensional framelet transform $\mathbf{W}$ is equivalent to a collection of over-complete basis over $L^2(R)$ which satisfies $\mathbf{W}^\top \mathbf{W} = \mathbf{I}$ [51]–[53], where $\mathbf{W}^\top$ and $\mathbf{I}$ are the inverse framelet transform and equivalent transform, respectively. The detailed description of the generation process of $\mathbf{W}$ can be found in [54], [55].

In the second layer $\mathcal{B}_2$, the Fourier transform is implemented along the spectral or temporal mode through Matlab's fft operation, which is a linear operator $\mathcal{B}_2 : \mathbb{R}^{n_1 \times n_2 \times n_3 \times \cdots \times n_k} \rightarrow \mathbb{C}^{n_1 \times n_2 \times n_3 \times \cdots \times n_k}$.

In the third layer $\mathcal{B}_3$, the KL transform projects the data to the bases, which is the principal directions of the data itself $\mathcal{B}_3 : \mathbb{C}^{n_1 \times n_2 \times n_3 \times \cdots \times n_k} \rightarrow \mathbb{R}^{n_1 \times n_2 \times n_3 \times \cdots \times n_k}$. As a contrast, Fourier transform and framelet transform only project the data to some fixed bases.

The compound of the KL transform, framelet transform, and Fourier transform can be viewed as a couple of the predefined transform and data-adaptive transform. Moreover, the coupled transform $\mathcal{B}$ can be efficiently obtained by calculating a series of operations in Matlab; see **Algorithm 1** for more details.

We discuss the connection between the proposed work and previous TNN-based methods. On the one hand, we provide a unified framework that can accommodate the previous TNN-based methods, as well as a deconstruction perspective to view their inner mechanisms. In our framework, they can be viewed as a sparse approximation under a compound of transforms, such as FFT+KL (TNN) [33] and DCT+KL [41]. On the other hand, our method can be seen as an improvement of the existing work, namely TNN in the framelet domain. Thus, for a third-order tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, the proposed

CT-LRTC is equivalent to

$$\begin{aligned} \min_{\mathcal{X}} \quad & \|\mathcal{B}_1(\mathcal{X})\|_{\text{TNN}} \\ \text{s.t.} \quad & \mathcal{P}_\Omega(\mathcal{X}) = \mathcal{P}_\Omega(\mathcal{O}), \end{aligned} \tag{1}$$

where $\mathcal{B}_1$ is a linear operator which is a two-dimensional framelet transform, it is computed within a local spatial neighborhood of the original tensor. The following theorem shows that the proposed model is a convex optimization problem.

Theorem 1: The proposed CT-LRTC (1) is a convex optimization problem.

Proof. The key to proving the convexity of the proposed CT-LRTC (1) is that $\{\mathcal{X}|\mathcal{P}_\Omega(\mathcal{X}) = \mathcal{P}_\Omega(\mathcal{O})\}$ is a convex set and $\|\mathcal{B}_1(\mathcal{X})\|_{\text{TNN}}$ is a convex function.

First, it is easy to check that set $\{\mathcal{X}|\mathcal{P}_\Omega(\mathcal{X}) = \mathcal{P}_\Omega(\mathcal{O})\}$ is convex by the definition.

Second, we prove that the function $\|\mathcal{B}_1(\mathcal{X})\|_{\text{TNN}} : \mathcal{X} \in \mathbb{R}^{n_1\times n_2\times n_3} \to \mathbb{R}$ is convex. In fact, for $\forall\ \mathcal{X}_1, \mathcal{X}_2 \in \mathbb{R}^{n_1\times n_2\times n_3}$, and $0 \le \theta \le 1$, we have

$$\begin{aligned} &\|\mathcal{B}_1(\theta\mathcal{X}_1 + (1-\theta)\mathcal{X}_2)\|_{\text{TNN}} \\ &\quad = \|\theta\mathcal{B}_1(\mathcal{X}_1) + (1-\theta)\mathcal{B}_1(\mathcal{X}_2)\|_{\text{TNN}} \\ &\qquad \times\ (\mathcal{B}_1 \text{ is a linear operator}) \\ &\quad \le \theta\|\mathcal{B}_1(\mathcal{X}_1)\|_{\text{TNN}} + (1-\theta)\|\mathcal{B}_1(\mathcal{X}_2)\|_{\text{TNN}} \\ &\qquad \times (\text{TNN [41] is a convex function}). \end{aligned}$$

Thus, the function $\|\mathcal{B}_1(\mathcal{X})\|_{\text{TNN}}$ is convex by definition.

Combining the above facts, the proposed CT-LRTC (1) is a convex optimization problem.

B. The Coupled Transform

Below, we provide explanations for several key ideas of the proposal.

1) Role of Different Transforms: The framelet transform captures local spatial correlation, since it is computed within a local spatial neighborhood of the original tensor. The Fourier transform is believed to exploit the global temporal (or spectral) correlation, since its computation refers to the whole temporal (or spectral) vectors. The KL transform reflects the global spatial correlation (i.e., low-rankness), since it is computed within a spatial slice of the framelet and Fourier transformed result. The sparsity of the result after KL transform is indeed the rank of the spatial slice of the framelet and Fourier transformed result.

These three transforms are inherently related, and they have collaborated. As we can see, the framelet transform is applied on the spatial slices while the Fourier transform is conducted along the third mode (temporal or spectral mode). The carrying out of framelet and Fourier transforms relatively independent but spatial and temporal (or spectral) are coupled. Meanwhile, the KL transform is based on the bases related to the data distribution, being different from the fixed bases provided by framelet and Fourier transforms.

2) Spatial and Spectral Decorrelation: First, the framelet transform has been widely used for image processing (Please see [51], [54] for example) with its abundant bases which are greatly suitable for the spatial geometric structures, textures, and details existed in a majority of natural images. Second, along the temporal (or spectral) dimension, this framelet transformed result is further decomposed into low-frequency components and high-frequency components with respect to the Fourier bases. Thus, after the framelet and Fourier transforms, the multi-dimensional images are thoroughly decomposed with respect to the semi-orthogonal framelet bases along the spatial dimensions and the orthogonal Fourier bases along the temporal (or spectral) dimension. The introduced framelet transform faithfully exploits the spatial multi-scale nature widely existed in imaging data, especially for implicit spatial low-rank images, which results in more sparse coefficients after KL transform (i.e., more low-rankness) as compared with coupled Fourier and KL transforms (i.e., TNN).

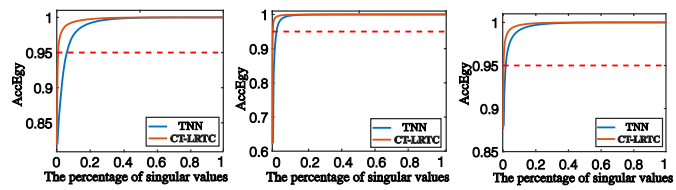


Fig. 2. The AccEgy with respect to the percentage of singular values of the original tensor and the framelet transformed tensor. From left to right: the color image *Butterfly* of size $256 \times 256 \times 3$, the multispectral image (MSI) *Cd* of size $512 \times 512 \times 31$, and the grayscale video *Bus* of size $288 \times 352 \times 30$, respectively.

3) Enhanced Low-Rank Representation: The proposed low-rank tensor representation based a coupled transform faithfully exploits the spatial multi-scale nature widely existed in imaging data, and this is expected to help generating a better low tensor multi-rank approximation. Fig. 2 plots the AccEgy with respect to the percentage of singular values of the original tensor and the framelet transformed tensor on different types of multi-dimensional data. An immediate observation is that the framelet together with Fourier transform significantly reform singular values for different multi-dimensional data and the energy of the singular value of the transformed tensor is more concentrated after the framelet transform. More distinctly, as pointed out by the auxiliary dashed lines, after the framelet and Fourier transforms, the transformed data could occupy 95% of the whole energy with a smaller proportion of singular values, compared with the case in which only the Fourier transform is involved (i.e., TNN). Therefore, we can use a better tensor low multi-rank approximation to achieve the same AccEgy. The above numerical experiments justify our contribution.

C. ADMM-Based Optimization Algorithm

We develop an ADMM [49] algorithm to solve the proposed CT-LRTC (1). By introducing an auxiliary variable $\mathcal{Y} = \mathcal{B}_1(\mathcal{X})$, we rewrite the original problem (1) as the equivalent constrained optimization problem

$$\begin{aligned} \min_{\mathcal{X},\mathcal{Y}} \quad & l_{\mathbb{S}}(\mathcal{X}) + \|\mathcal{Y}\|_{\text{TNN}} \\ \text{s.t.} \quad & \mathcal{Y} = \mathcal{B}_1(\mathcal{X}), \end{aligned} \tag{2}$$

where $l_{\mathbb{S}}(\mathcal{X})$ is an indicator function defined as

$$l_{\mathbb{S}}(\mathcal{X}) = \begin{cases} 0, & \mathcal{X} \in \Phi, \\ \infty, & \text{otherwise}, \end{cases}$$

where $\Phi := \{\mathcal{X} \in \mathbb{R}^{n_1\times n_2\times n_3},\ \mathcal{P}_\Omega(\mathcal{X}) = \mathcal{P}_\Omega(\mathcal{O})\}$. The corresponding augmented Lagrangian function is

$$L_\beta(\mathcal{X},\mathcal{Y},\Lambda)=l_{\mathbb{S}}(\mathcal{X})+\|\mathcal{Y}\|_{\text{TNN}}+\frac{\beta}{2}\|\mathcal{B}_1(\mathcal{X})-\mathcal{Y}+\frac{\Lambda}{\beta}\|_F^2, \quad (3)$$

where $\Lambda \in \mathbb{R}^{m^2ln_1\times m^2ln_2\times n_3}$ is the Lagrangian multiplier and β is a positive penalty parameter. In the end, ADMM iterates as follows:

$$\begin{cases}\mathcal{Y}^{k+1} = \arg\min\limits_{\mathcal{Y}}\ L_\beta(\mathcal{X}^k,\mathcal{Y},\Lambda^k),\\ \mathcal{X}^{k+1} = \arg\min\limits_{\mathcal{X}}\ L_\beta(\mathcal{X},\mathcal{Y}^{k+1},\Lambda^k),\\ \Lambda^{k+1} = \Lambda^k+\beta(\mathcal{B}_1(\mathcal{X}^{k+1})-\mathcal{Y}^{k+1}).\end{cases}$$

Below, we give the details of updating each minimizing subproblem.

Step 1, the $\mathcal{Y}$-subproblem at the k-th iteration is

$$\mathcal{Y}^{k+1} = \arg\min_{\mathcal{Y}}\ \|\mathcal{Y}\|_{\text{TNN}}+\frac{\beta}{2}\|\mathcal{B}_1(\mathcal{X}^k)-\mathcal{Y}+\frac{\Lambda^k}{\beta}\|_F^2. \quad (4)$$

A closed-form solution of (4) can be obtained by a tensor singular value thresholding (t-SVT) operator [56], [57], i.e.,

$$\mathcal{Y}^{k+1} = \text{t-SVT}_{\frac{1}{\beta}}(\mathcal{B}_1(\mathcal{X}^k)+\frac{\Lambda^k}{\beta}), \quad (5)$$

and its computational complexity is $\mathcal{O}(m^2ln_1n_2n_3\log(n_3)+m^2ln_3\min(n_1n_2^2,n_2n_1^2))$.

Step 2, the $\mathcal{X}$-subproblem at the k-th step is

$$\begin{aligned}\mathcal{X}^{k+1} &= \arg\min_{\mathcal{X}}\ l_{\mathbb{S}}(\mathcal{X})+\frac{\beta}{2}\|\mathcal{B}_1(\mathcal{X})-\mathcal{Y}^{k+1}+\frac{\Lambda^k}{\beta}\|_F^2\\ &= \arg\min_{\mathcal{X}}\ l_{\mathbb{S}}(\mathcal{X})\\ &\quad+\frac{\beta}{2}\|\mathbf{W}\mathbf{X}_{(3)}^\top-(\mathcal{L}^{-1}(\mathcal{Y}^{k+1}-\frac{\Lambda^k}{\beta})_{(4)})^\top\|_F^2. \quad (6)\end{aligned}$$

To solve (6), we introduce the following Theorem.

Theorem 2: Let $\mathbf{E}$ be a semi-orthogonal matrix, i.e., $\mathbf{E}^\top\mathbf{E}=\mathbf{I}$, where $\mathbf{I}$ is the identity matrix, then

$$\mathbf{E}^\top\mathbf{Y} = \arg\min_{\mathbf{Z}}\|\mathbf{E}\mathbf{Z}-\mathbf{Y}\|_F^2.$$

Proof. The derivation process is as follows

$$\begin{aligned}&\arg\min_{\mathbf{Z}}\|\mathbf{E}\mathbf{Z}-\mathbf{Y}\|_F^2\\ &= \arg\min_{\mathbf{Z}}\|\mathbf{E}\mathbf{Z}\|_F^2-2<\mathbf{E}\mathbf{Z},\mathbf{Y}>+\|\mathbf{Y}\|_F^2\\ &= \arg\min_{\mathbf{Z}}\|\mathbf{Z}\|_F^2-2<\mathbf{Z},\mathbf{E}^\top\mathbf{Y}>+\|\mathbf{E}^\top\mathbf{Y}\|_F^2\\ &= \arg\min_{\mathbf{Z}}\left\|\mathbf{Z}-\mathbf{E}^\top\mathbf{Y}\right\|_F^2.\end{aligned}$$

Thus, $\mathbf{E}^\top\mathbf{Y} = \arg\min_{\mathbf{Z}}\left\|\mathbf{Z}-\mathbf{E}^\top\mathbf{Y}\right\|_F^2 = \arg\min_{\mathbf{Z}}\|\mathbf{E}\mathbf{Z}-\mathbf{Y}\|_F^2$.

By using ***Theorem 2***, the corresponding (6) equals to

$$\begin{aligned}\mathcal{X}^{k+1} &= \arg\min_{\mathcal{X}}\ l_{\mathbb{S}}(\mathcal{X})\\ &\quad+\frac{\beta}{2}\|\mathbf{X}_{(3)}-(\mathcal{L}^{-1}(\mathcal{Y}^{k+1}-\frac{\Lambda^k}{\beta})_{(4)})\mathbf{W}\|_F^2.\end{aligned}$$

Algorithm 2 Low-Rank Tensor Completion Under Coupled Transform

Input: The observed data $\mathcal{O}\in\mathbb{R}^{n_1\times n_2\times n_3}$; the index set Ω; and the parameter β;

Initialization: $\mathcal{X}^{(0)}=\mathcal{O}$; $\mathcal{Y}^{(0)}=\Lambda^{(0)}=\text{zeros}(m^2ln_1\times m^2ln_2\times n_3)$; $\mathbf{W}$; and $k=0$.

while $\frac{\|\mathcal{X}^{k+1}-\mathcal{X}^k\|_F}{\|\mathcal{X}^k\|_F}>10^{-3}$ and $k\leqslant 200$ **do**
  Update $\mathcal{Y}^{(k+1)}$ by solving (5);
  Update $\mathcal{X}^{(k+1)}$ by solving (7);
  Update $\Lambda^{(k+1)}$ by solving (8);
end while

Output: The completed tensor $\mathcal{X}$.

By minimizing the $\mathcal{X}$-subproblem, we have $l_{\mathbb{S}}(\mathcal{X})=0$. Thus, the solution of $\mathcal{X}$-subproblem can be obtained by

$$\begin{cases}\mathcal{P}_\Omega(\mathcal{X}^{k+1})=\mathcal{P}_\Omega(\mathcal{O}),\\ \mathcal{P}_{\Omega^C}(\mathcal{X}^{k+1})=\mathcal{P}_{\Omega^C}(\text{fold}_3((\mathcal{L}^{-1}(\mathcal{Y}^{k+1}-\frac{\Lambda^k}{\beta})_{(4)})\mathbf{W})),\end{cases} \quad (7)$$

where Ω^C denotes the complement of set Ω, and its computational complexity is $\mathcal{O}(m^2ln_1n_2n_3)$.

Step 3, updating the multiplier Λ by

$$\Lambda^{k+1}=\Lambda^k+\beta(\mathcal{B}_1(\mathcal{X}^{k+1})-\mathcal{Y}^{k+1}). \quad (8)$$

The overall ADMM iteration strategy is summarized in **Algorithm 2**. The total computational complexity at each iteration is $\mathcal{O}(m^2ln_1n_2n_3\log(n_3)+m^2ln_3\min(n_1n_2^2,n_2n_1^2))$. The following theorem shows the convergence of the proposed algorithm.

Theorem 3: The sequences generated from the proposed ADMM algorithm converges to the minimizer of the convex optimization problem (1).

Proof. Since $\mathcal{B}_1$ is a linear operator, $\mathcal{B}_1(\mathcal{X})$ can be written as a matrix-vector product, i.e., $\mathbf{B}_1\mathbf{x}$, where $\mathbf{x}$ denotes the vectorization of $\mathcal{X}$. The linear constraints can be reformulated as the following matrix-vector product:

$$\mathbf{I}\mathbf{y}+(-\mathbf{B}_1)\mathbf{x}=0,$$

where $\mathbf{I}$ and $\mathbf{y}$ denote the identity matrix and the vectorization of $\mathcal{Y}$, respectively. We separate all the variables into two groups, $\mathcal{X}$ and $\mathcal{Y}$, decompose the objective function as $f+g$ with $f=l_{\mathbb{S}}(\mathcal{X})$ and $g=\|\mathcal{Y}\|_{\text{TNN}}$, both of which are convex functions. Therefore, the problem (2) fits the framework of ADMM [49], and the convergence of the algorithm is theoretically guaranteed, i.e., the sequence $\{\mathcal{X}^k,\mathcal{Y}^k\}$ generated from the proposed ADMM algorithm is convergent.

IV. Experimental Results

In this section, we test the performance of the proposed CT-LRTC and compare it with state-of-the-art tensor completion methods. Subsection IV-A provides some experimental settings. Subsection IV-B tests the proposed CT-LRTC in synthetic data completion. Subsections IV-C–IV-E present experimental results on multi-dimensional visual data, including color images, multispectral images, grayscale videos, and color videos. Finally, some discussions are given in Subsection IV-F.

TABLE II
THE COMPUTATIONAL COMPLEXITY OF DIFFERENT ALGORITHMS ON AN $n_1 \times n_2 \times n_3$ TENSOR. HERE, R IS THE RANK OF THE TR FACTORIZATION

Method	The computational complexity
TNN	$\mathcal{O}\left(n_1 n_2 n_3 \log(n_3) + n_3 \min\left(n_1 n_2^2, n_2 n_1^2\right)\right)$
TNN-DCT	$\mathcal{O}\left(n_1 n_2 n_3 \log(n_3) + n_3 \min\left(n_1 n_2^2, n_2 n_1^2\right)\right)$
t-TNN	$\mathcal{O}\left(n_1 n_2 n_3 \log(n_3) + n_3 \min\left(n_1 n_2^2, n_2 n_1^2\right)\right)$
PSTNN	$\mathcal{O}\left(n_1 n_2 n_3 \log(n_3) + n_3 \min\left(n_1 n_2^2, n_2 n_1^2\right)\right)$
TRLRF	$\mathcal{O}(3R^2 n_1^2 n_3 + 3R^6)$
CT-LRTC	$\mathcal{O}(m^2 l n_1 n_2 n_3 \log(n_3) + m^2 l n_3 \min(n_1 n_2^2, n_2 n_1^2)$

A. Experimental Setting

1) Parameter Setting: The proposed method involves the following parameters: m and l in the framelet transform matrix $\mathbf{W}$, and the positive penalty parameter β in the augmented Lagrangian function (3). The parameter m controls the type of the selected filter, including the Haar wavelet, piece-wise linear B-spline, and piece-wise cubic B-spline, which are briefly referred to as "Haar", "Linear", and "Cubic", severally. The parameter l controls the level of framelet transform, and we set it to the range of [1, 3] with increments 1. The effects of m and l are discussed in subsection IV-F. The parameter β controls the speed of convergence, and we fine-tuned it in the range of [0.01, 0.1] with increments 0.01 to obtain the highest peak signal-to-noise ratio (PSNR) [58] value.

2) Compared Methods: We compare the proposed method with several existing approaches, including TNN [33], TNN-DCT [41], t-TNN [59], PSTNN [37], and TRLRF [60]. For a fair comparison, parameters in the compared method are manually adjusted according to the authors' suggestions to obtain the highest PSNR value. Table II compares the computational complexity of different methods.

3) Data Generation and Experimental Environment: In order to test the generalization, we consider four types of multi-dimensional visual data: color images, multispectral images, grayscale videos, and color videos. All data are normalized to [0, 1]. We also consider two types of sampling: the element-wise sampling and the tube-wise sampling as shown in Fig. 3. In the element-wise sampling, we generate incomplete data by sampling elements of data randomly, with different sampling rates (SRs) $\in \{5\%, 10\%, 20\%\}$. As for tube-wise sampling, we generate incomplete data by deleting elements of the original data with three different masks (text, graffiti, and grid mask). All tests are implemented under Windows 10 and MATLAB R2018a running on a desktop with an Intel(R) Core (TM) i9-9900K CPU at 3.60GHz and 32GB RAM.

4) Evaluation Indices: We consider the following indices for quantitative evaluation: PSNR and the structural similarity index [58] (SSIM). PSNR and SSIM are defined as

$$\text{PSNR} = 20\ \log_{10} \frac{\sqrt{n_1 \times n_2}}{\|\mathbf{X} - \mathbf{X}^*\|_F}$$

and

$$\text{SSIM} = \frac{(2\mu_{\mathbf{X}} \cdot \mu_{\mathbf{X}^*} + c_1)(2\sigma_{\mathbf{X}\mathbf{X}^*} + c_2)}{(\mu_{\mathbf{X}}^2 + \mu_{\mathbf{X}^*}^2 + c_1)(\sigma_{\mathbf{X}}^2 + \sigma_{\mathbf{X}^*}^2 + c_2)}.$$

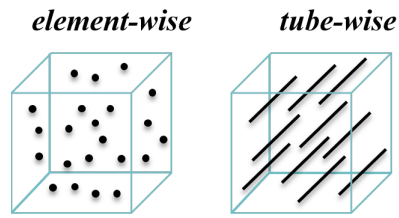


Fig. 3. An illustration of the element-wise sampling and tube-wise sampling.

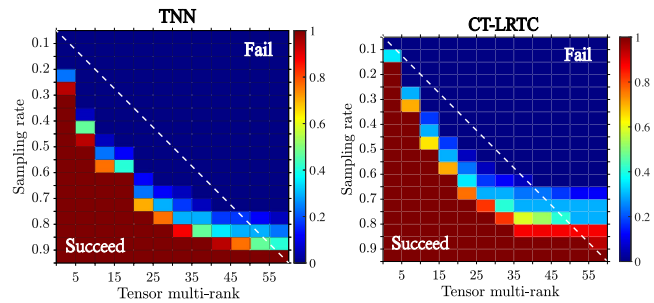


Fig. 4. The success rates for synthetic data with varying multi-ranks and SRs. The left and right are the result of the TNN-based and the proposed CT-LRTC method, respectively. The white dashed lines are placed on the diagonal line for easier comparison.

Fig. 5. Ten benchmark color images used in color image completion. From left to right, and from top to bottom: *Barbara, Airplane, Peppers, Fruits, Baboon, Cherry, Parrot, Vegetable, Flower,* and *Sailboat*.

Here, $\mathbf{X}^* \in \mathbb{R}^{n_1 \times n_2}$ is the recovered image, $\mathbf{X}$ is the ground truth image, $\mu_{\mathbf{X}}$ and $\mu_{\mathbf{X}^*}$, $\sigma_{\mathbf{X}}$ and $\sigma_{\mathbf{X}^*}$ are the mean values and standard variances of $\mathbf{X}$ and $\mathbf{X}^*$, respectively, $\sigma_{\mathbf{X}\mathbf{X}^*}$ is the covariance of $\mathbf{X}$ and $\mathbf{X}^*$, and $c_1, c_2 > 0$ are constants. For a third-order tensor, the PSNR and SSIM values are calculated by averaging the PSNR and SSIM values of all frontal slices.

B. Synthetic Data Completion

To verify our motivation, we conducted the experiments with respect to the tensor multi-rank and the sampling rate (SR) by the TNN-based tensor completion method and the proposed CT-LRTC. We conduct 50 independent experiments on random color image patch of size $64 \times 64 \times 3$ and calculate the success rate, where one test is successful if the relative square error of the recovered tensor $\widehat{\mathcal{X}}$ and the ground-truth tensor $\mathcal{X}$, i.e., $\frac{\|\mathcal{X} - \widehat{\mathcal{X}}\|_F^2}{\|\mathcal{X}\|_F^2}$ is less than 10^{-3}. Fig. 4 shows the success rates for various multi-ranks and different SRs. We can observe that the performance of the proposed CT-LRTC shows larger reddish-brown areas than the TNN-based method, implying higher success rates.

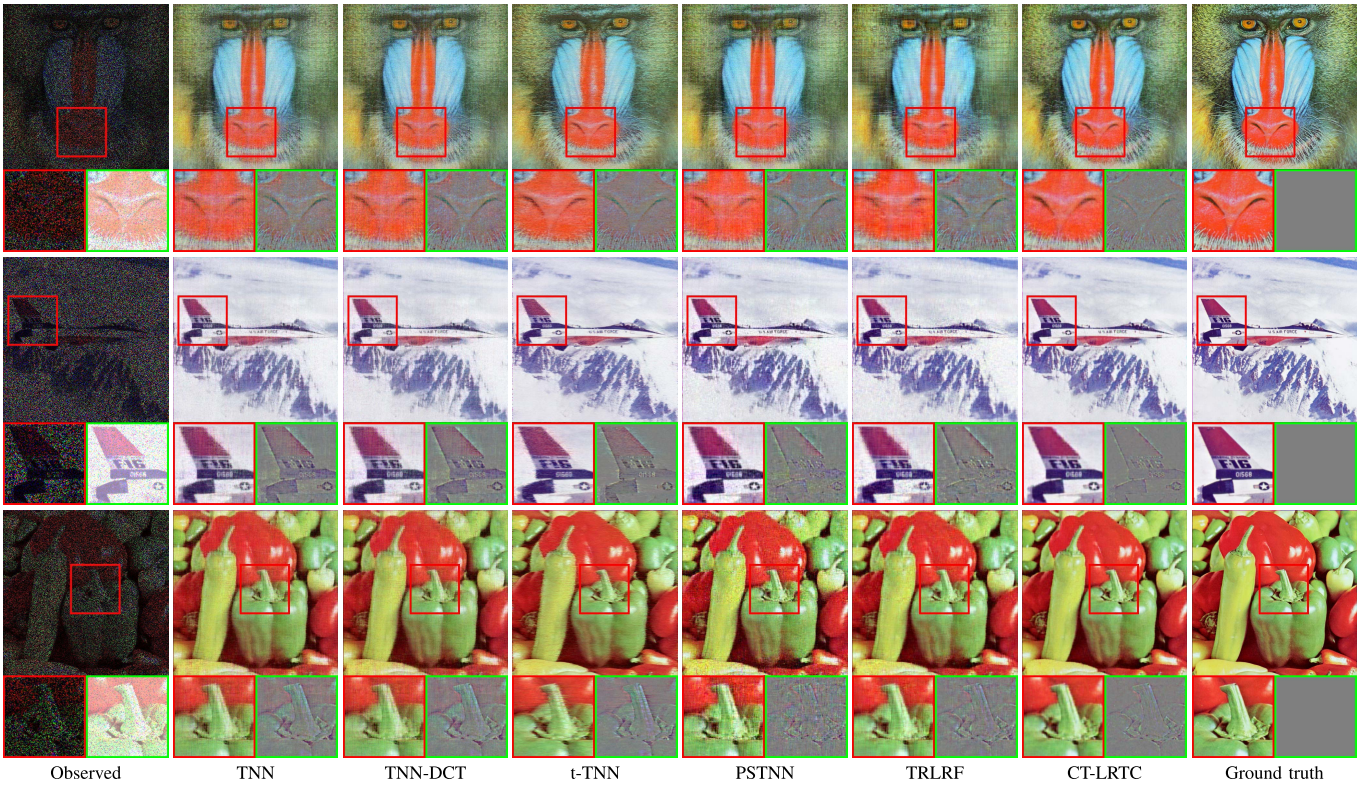


Fig. 6. The recovered results by different methods on color images under SR = 20%. From top to bottom: *Barbara, Airplane,* and *Peppers*, respectively. For better viewing, under each image, we display the magnified map of a patch and the corresponding error map (difference from the ground truth). Error maps with less color information indicate better restoration performance.

C. Color Image Completion

Ten benchmark color images[1] are shown in Fig. 5 for color image completion. We conduct two types of experiments: the first five benchmark images of size $512 \times 512 \times 3$ with element-wise sampling, and the last five benchmark images of size $500 \times 500 \times 3$ with tube-wise sampling (i.e., color image inpainting). Here, we test three different masks: text, graffiti, and grid mask.

1) The Element-Wise Sampling: Fig. 6 shows the recovered results by different methods on color images under SR = 20%. As one can see, the recovered results obtained by TNN only recover the coarse structure, which produce evident blur and artifacts. We can observe that other TNN-based methods (i.e., TNN-DCT, t-TNN, and PSTNN) also suffer from such problems. The recovered results obtained by TRLRF recover some details and do not alleviate the blurriness compared to TNN-based methods. The underlying reason is that only the global low-rankness prior is not sufficient to recover the potential image. In comparison, the proposed CT-LRTC provides the most visually pleasing results with clear and sharp spatial details, due to CT-LRTC can capture the intrinsic structure of the multi-dimensional visual data represented in a multi-scale manner.

Table III presents the PSNR/SSIM values and average CPU time (in minutes) of the recovered results by different methods on color images under different SRs. We can observe that the proposed CT-LRTC consistently outperforms the compared ones in terms of both PSNR and SSIM values and the CPU time is comparable with the other compared methods. More precisely, CT-LRTC outperforms the second-best methods about 3 dB in PSNR and 0.1 in SSIM. Such an improvement is mainly attributed to that the use of multi-scale representation by CT-LRTC to implicitly utilize the spatial information of images, thereby promoting the improvement of PSNR and SSIM values.

2) The Tube-Wise Sampling: Fig. 7 shows the recovered results by different methods for color image inpainting. Notice that the magnified map of a patch and the corresponding error map (difference from the ground truth) are displayed below each image. The corresponding PSNR and SSIM values are displayed above the image. Clearly, CT-LRTC obtains the most visually satisfying results among the competing methods, and the PSNR and SSIM results are higher than the corresponding compared methods in all cases. In particular, the advantages of the proposed CT-LRTC are particularly evident in the text and mesh masking problem.

D. Multispectral Image Completion

In this subsection, we perform LRTC on five multispectral images (MSIs) from the CAVE dataset,[2] i.e., *Clay, Balloons, Feathers, Cd,* and *Beads*. All MSIs are of size $512 \times 512 \times 31$ with SR = 5%, 10%, and 20%.

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

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

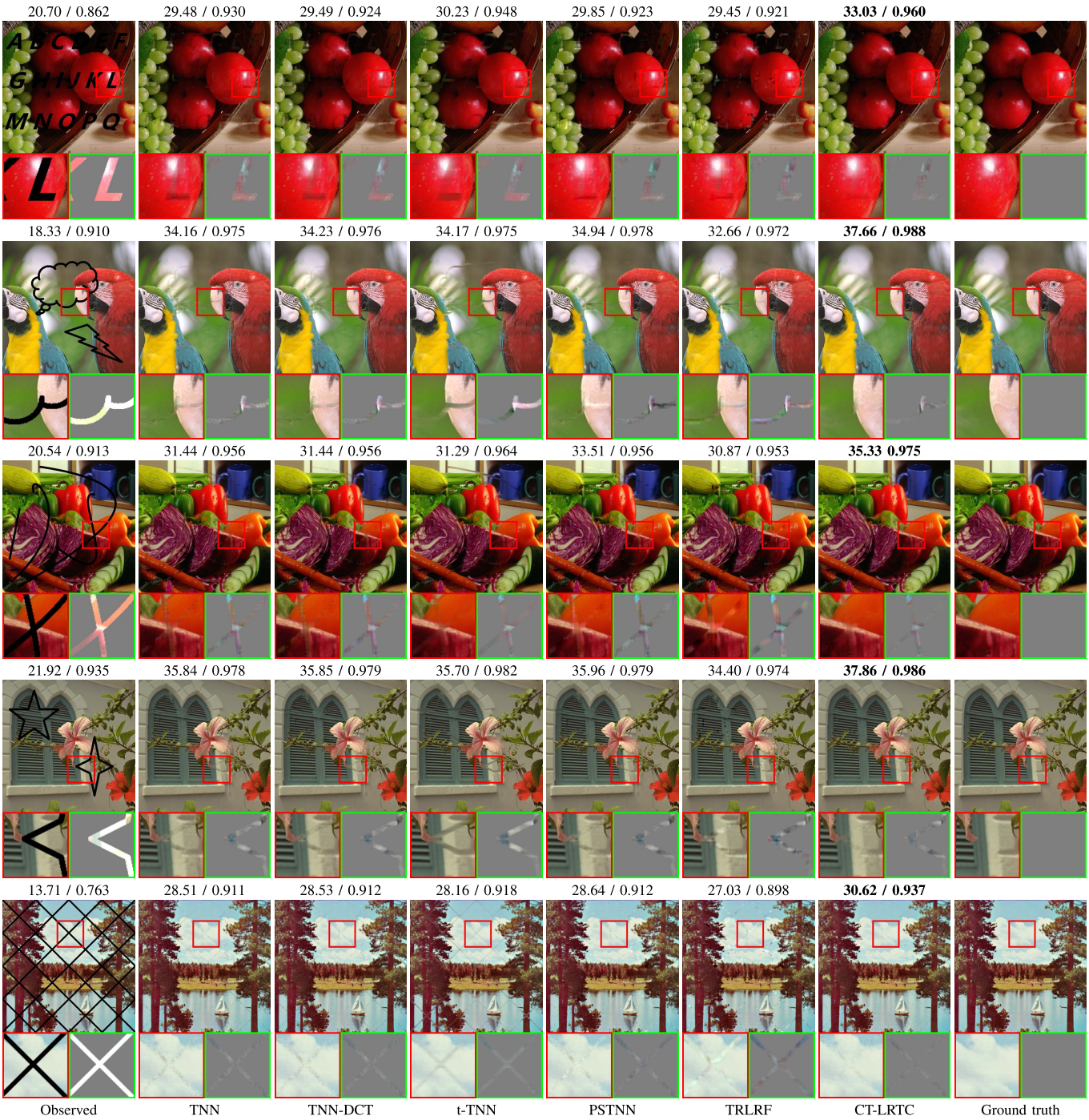


Fig. 7. The recovered results by different methods on the five masked image. The corresponding PSNR and SSIM values are displayed above the image. For better visual comparison, under each image, we display the magnified map of a patch and the corresponding error map (difference from the ground truth). Error maps with less color information indicate better restoration performance.

Fig. 8 shows the pseudo-color image (composed of the 1st, 2nd, and 31st bands) of the recovered results by different methods on MSIs under SR $= 10\%$, where the pixel intensity is readjusted for better viewing. We observe that TNN, TNN-DCT, PSTNN, and TRLRF cannot recover all the elements; t-TNN performs slightly better while the proposed CT-LRTC fills almost all elements. Regarding visual quality, the proposed CT-LRTC produces the closet result to the ground truth, which is superior to other compared methods.

Table IV lists the quantitative indexes (PSNR/SSIM) and average CPU time of the recovered results by different methods on MSIs under different SRs. The proposed CT-LRTC obtains the highest PSNR and SSIM values in almost all cases under different MRs. In addition, Fig. 9 displays the PSNR values of each frontal slice of multispectral image *Clay*. As observed, in all frontal slices, the PSNR values of the proposed CT-LRTC are much higher than those of the other compared methods.

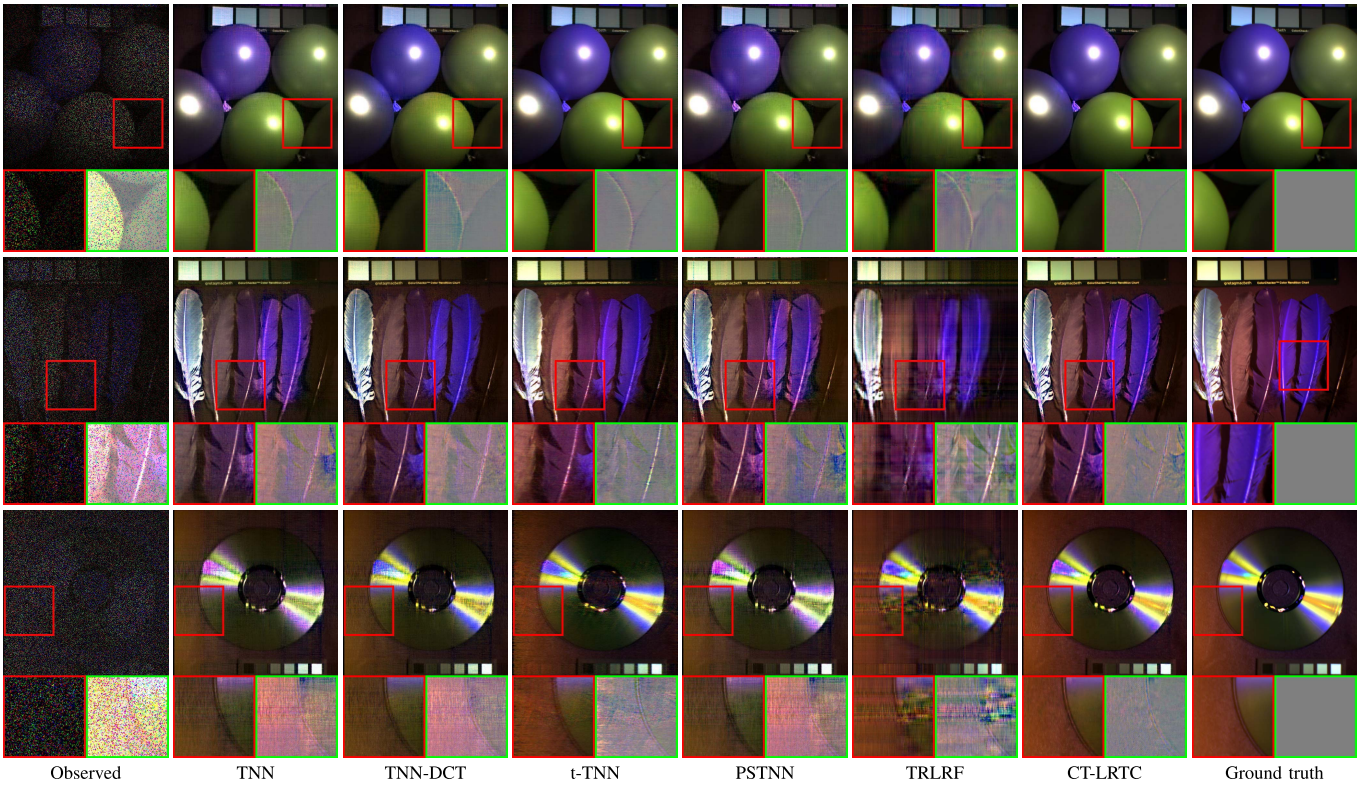


Fig. 8. The pseudo-color image (composed of the 1st, 2ed, and 31st bands) of the recovered results by different methods on MSIs under SR = 10%. From top to bottom: *Balloons, Feathers,* and *Cd*, respectively. The intensity of the pixels in the shown image is readjusted for better display.

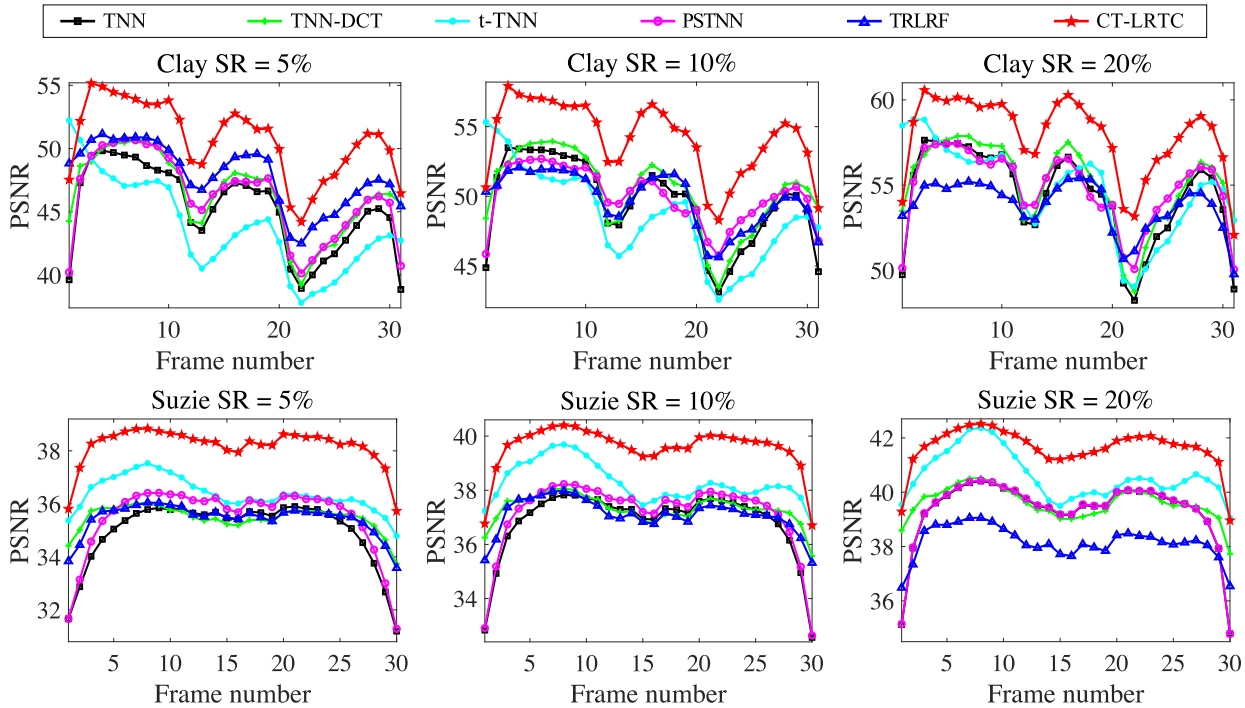


Fig. 9. The PSNR values of all frontal slices obtained by different methods on the multispectral image *Clay* and the grayscale video *Suzie* under different SRs.

E. Video Completion

In this subsection, we test four grayscale videos with differenr SRs: *Birdhouse* of size $700 \times 860 \times 30$, *Suzie* of size $480 \times 700 \times 30$, *Bus* of size $288 \times 352 \times 30$, and *Hall* of size $144 \times 176 \times 100$. And we also test three color videos with differenr SRs: *Bus* of size $288 \times 352 \times 3 \times 30$, *Container* of size $144 \times 176 \times 3 \times 30$, and *Coastguard* of size $144 \times 176 \times 3 \times 30$. These videos are available online http://trace.eas.asu.edu/yuv/.

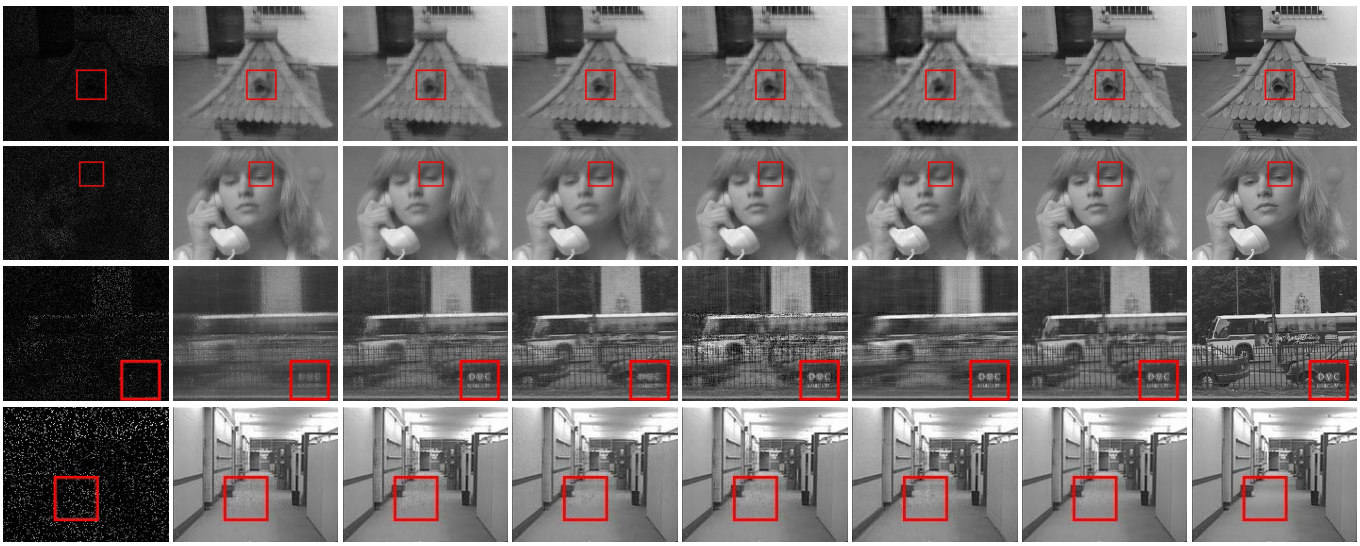


Fig. 10. The 15-th frame of the recovered results by different methods on grayscale videos under SR = 10%. From top to bottom: *Birdhouse, Suzie, Bus,* and *Hall*, respectively.

Fig. 11. The 15-th frame (color figure) of the recovered results by different methods on color videos under SR = 5%. From top to bottom: *Bus*, *Container*, and *Coastguard*, respectively. For better viewing, under each image, we display the magnified map of a patch and the corresponding error map (difference from the ground truth). Error maps with less color information indicate better restoration performance.

1) Grayscale Video Completion: Fig. 10 shows the 15-th frame of the recovered examples by different methods on grayscale videos under SR = 10%, and the corresponding PSNR, SSIM values, and average CPU time are summarized in Table V. Clearly, CT-LRTC obtains the most visually pleasant results among the compared methods. In the red regions of Fig. 10, we can see that the recovered results by the proposed CT-LRTC are more clear and sharper in spatial texture compared with the other methods. Table V shows that the proposed CT-LRTC obtains the highest quantitative indexes in all testing videos. Fig. 9 displays the PSNR values of each frontal slice of *Suzie*, and the quantitative indexes are consistent with the above results.

2) Color Video Completion: Compared with the other TNN-based methods, the proposed CT-LRTC can also be used for color videos completion by simply changing $\mathcal{B}_2$ into a two-dimensional Fourier transform. In order to facilitate comparison, for the other TNN-based methods, we use a reshaping operation to convert the color video (fourth-order) into third-order data, and then input it into the model for solution.

TABLE III

THE PSNR/SSIM VALUES AND AVERAGE CPU TIME (IN MINUTES) OF THE RECOVERED RESULTS BY DIFFERENT METHODS ON COLOR IMAGES UNDER DIFFERENT SRS. THE **BEST** RESULTS ARE HIGHLIGHTED IN **BOLDER** FONTS

	SR	5%		10%		20%		Time
	Method	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	
Barbara	TNN	16.22	0.234	21.55	0.474	22.54	0.563	0.47
	TNN-DCT	16.16	0.219	19.34	0.374	22.89	0.583	0.36
	t-TNN	17.19	0.286	19.39	0.426	22.19	0.607	**0.27**
	PSTNN	16.97	0.256	21.65	0.475	22.99	0.586	1.88
	TRLRF	17.10	0.244	19.69	0.374	23.20	0.586	6.58
	CT-LRTC	**22.81**	**0.573**	**24.96**	**0.650**	**27.72**	**0.747**	4.30
Airplane	TNN	18.07	0.331	21.01	0.464	24.56	0.645	0.27
	TNN-DCT	17.75	0.311	21.09	0.475	24.66	0.648	0.97
	t-TNN	20.41	0.559	22.94	0.677	25.83	0.798	**0.57**
	PSTNN	18.53	0.350	21.75	0.471	24.58	0.641	1.75
	TRLRF	18.41	0.360	20.65	0.450	25.27	0.672	4.62
	CT-LRTC	**23.41**	**0.695**	**25.99**	**0.772**	**29.11**	**0.856**	10.37
Peppers	TNN	14.23	0.128	18.63	0.279	22.59	0.466	**0.58**
	TNN-DCT	14.75	0.151	18.71	0.286	22.53	0.467	0.60
	t-TNN	18.00	0.188	20.43	0.444	23.49	0.612	0.59
	PSTNN	14.12	0.097	19.46	0.294	20.80	0.339	2.16
	TRLRF	16.36	0.205	19.70	0.296	23.26	0.486	5.50
	CT-LRTC	**22.81**	**0.573**	**24.96**	**0.650**	**27.72**	**0.747**	13.13
Fruits	TNN	18.32	0.203	20.36	0.365	23.90	0.543	0.60
	TNN-DCT	17.61	0.249	20.59	0.377	24.18	0.559	**0.47**
	t-TNN	19.19	0.411	21.49	0.547	24.75	0.710	0.60
	PSTNN	18.03	0.220	20.91	0.383	23.83	0.495	1.76
	TRLRF	17.84	0.255	20.46	0.335	24.80	0.590	6.56
	CT-LRTC	**23.85**	**0.646**	**25.50**	**0.704**	**28.68**	**0.805**	7.46
Baboon	TNN	14.69	0.148	17.26	0.247	19.13	0.411	0.68
	TNN-DCT	15.38	0.150	17.41	0.253	19.29	0.418	**0.62**
	t-TNN	16.85	0.240	18.37	0.369	20.52	0.553	0.51
	PSTNN	15.93	0.142	17.47	0.254	19.31	0.418	1.62
	TRLRF	16.54	0.181	17.76	0.272	19.22	0.407	2.69
	CT-LRTC	**18.66**	**0.294**	**19.80**	**0.412**	**21.34**	**0.568**	9.09

TABLE IV

THE PSNR/SSIM VALUES AND AVERAGE CPU TIME (IN MINUTES) OF THE RECOVERED RESULTS BY DIFFERENT METHODS ON MSIS UNDER DIFFERENT SRS. THE **BEST** RESULTS ARE HIGHLIGHTED IN **BOLDER** FONTS

video	SR	5%		10%		20%		Time
	Method	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	
Clay	TNN	39.20	0.961	43.53	0.983	48.34	0.993	8.18
	TNN-DCT	40.36	0.970	44.29	0.985	49.09	0.994	**4.93**
	t-TNN	37.80	0.939	42.58	0.978	48.73	0.994	5.12
	PSTNN	40.19	0.965	43.97	0.982	48.77	0.992	32.33
	TRLRF	41.86	0.962	43.75	0.976	47.79	0.988	58.78
	CT-LRTC	**44.71**	**0.988**	**48.16**	**0.993**	**51.92**	**0.996**	68.49
Balloons	TNN	34.99	0.924	39.61	0.967	44.84	0.987	10.43
	TNN-DCT	36.05	0.939	40.87	0.974	46.16	0.990	5.59
	t-TNN	38.37	0.966	43.21	0.986	48.30	0.995	**4.79**
	PSTNN	36.16	0.929	40.94	0.969	45.63	0.988	37.84
	TRLRF	36.43	0.932	40.13	0.963	42.35	0.976	60.84
	CT-LRTC	**41.47**	**0.985**	**45.07**	**0.992**	**48.99**	**0.996**	100.01
Feathers	TNN	29.28	0.812	33.10	0.897	37.69	0.954	6.41
	TNN-DCT	30.33	0.850	34.51	0.926	39.52	0.971	**4.48**
	t-TNN	30.94	0.855	35.86	0.942	**42.04**	0.982	4.72
	PSTNN	29.98	0.808	34.66	0.903	38.74	0.956	33.30
	TRLRF	31.33	0.826	33.70	0.895	35.45	0.916	54.58
	CT-LRTC	**33.77**	**0.938**	**37.14**	**0.966**	41.16	**0.983**	99.34
Cd	TNN	28.39	0.859	32.22	0.911	36.30	0.953	10.63
	TNN-DCT	29.19	0.881	32.65	0.924	36.62	0.960	5.52
	t-TNN	31.58	0.899	34.50	0.938	38.61	0.972	**4.95**
	PSTNN	29.72	0.847	32.87	0.911	37.28	0.948	49.57
	TRLRF	30.66	0.870	33.96	0.902	35.71	0.920	52.34
	CT-LRTC	**35.56**	**0.954**	**37.90**	**0.970**	**41.06**	**0.983**	58.32
Beads	TNN	21.93	0.650	25.59	0.688	30.09	0.845	5.86
	TNN-DCT	21.96	0.498	25.90	0.699	30.73	0.865	5.36
	t-TNN	25.41	0.745	29.70	**0.887**	**35.21**	**0.965**	**4.38**
	PSTNN	22.42	0.467	26.10	0.669	31.30	0.837	31.90
	TRLRF	23.64	0.583	25.63	0.698	26.69	0.747	55.64
	CT-LRTC	**26.90**	**0.797**	**29.88**	0.886	33.63	0.947	61.82

Fig. 11 shows the 15-th frame of the recovered examples by different methods on color videos under SR = 10%, and the corresponding PSNR, SSIM values, and average CPU time are summarized in Table VI. It is obvious that CT-LRTC visually outperforms the other compared methods in keeping details of the recovered images and achieves the highest quantitative results under different SRs.

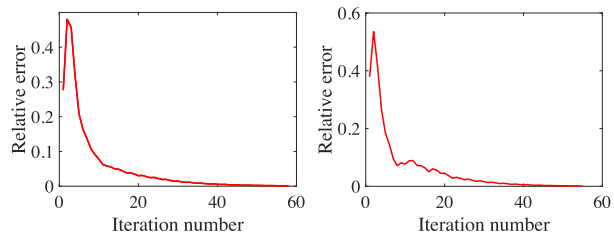


Fig. 12. The relative error curves of the ADMM algorithm on multispectral image *Clay* with SR = 10% (left) and SR = 20% (right), respectively.

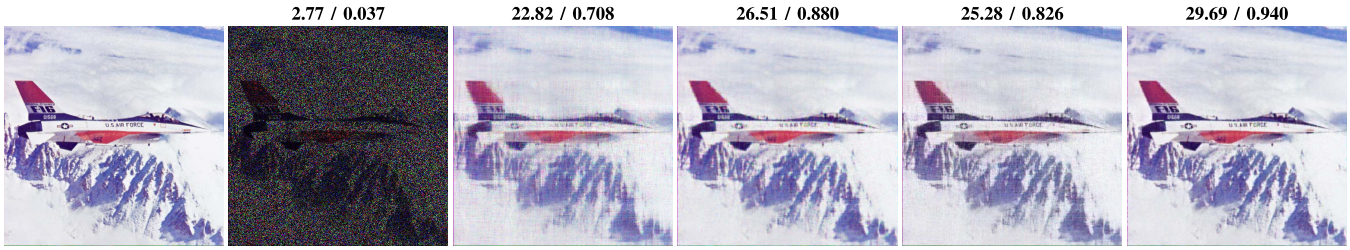
F. Discussion

In this subsection, we give some discussions.

1) A Rigorous Comparison: In order to further verify the effectiveness of the proposed method, we add a rigorous comparison with the compared methods about missing percentage from 10% to 90% with an interval of 20%. Table VII lists the quantitative indexes (PSNR/SSIM) by different methods on the color image *Flower*, the multispectral image *Cd*, and the grayscale video *Hall* under different SRs. We can observe that the proposed CT-LRTC obtains visually satisfying results and consistently outperforms the other compared methods in terms of both PSNR and SSIM values, especially when the sampling is extremely low. These observations correspond exactly to the relationships we mentioned earlier.

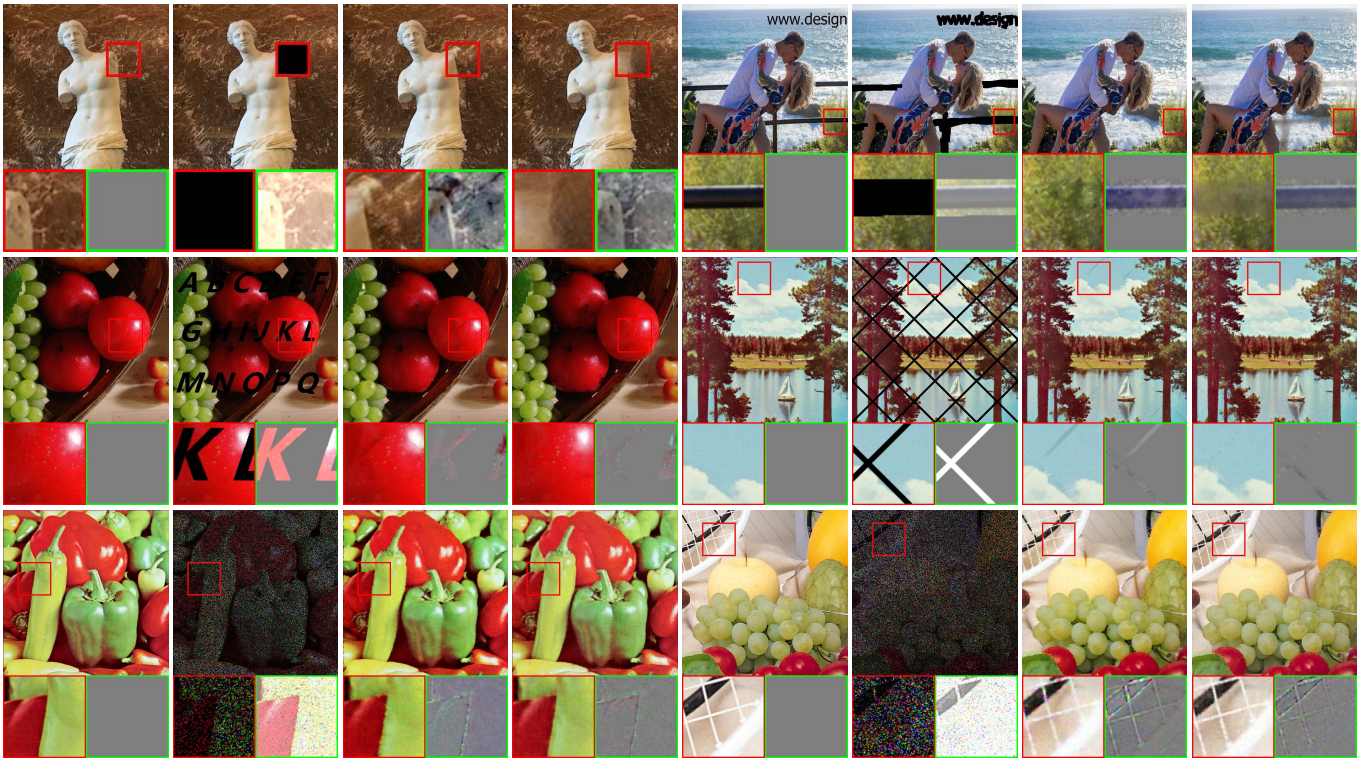
2) Framelet Setting: We discuss the effects of two important parameters of the framelet transform, which are the parameter m that controls the selection of filter and the parameter l that controls the decomposition level. Table VIII shows the quantitative indexes of the proposed CT-LRTC with different framelet settings on the color image *peppers* under SR = 5%. From Table VIII, we can see that CT-LRTC is not very sensitive to changes in the framelet transform parameters, but they directly affect the computational complexity of the proposed algorithm.

3) Mumerical Convergence of the Algorithm: In Fig. 12, we illustrate the convergence behavior of the proposed ADMM solver. Here, the relative error in each iteration is defined as $\|\mathcal{X}^{k+1} - \mathcal{X}^k\|_F / \|\mathcal{X}^k\|_F$, where $\mathcal{X}^{k+1}$ and $\mathcal{X}^k$ are the two successive reconstructed tensors. Fig. 12 suggests a strong convergence behavior of the proposed ADMM solver; it can reach a relative decrease of 10^{-2} in 30 iterations, and the relative error is decreased with the increase of iterations so that higher accuracy can be obtained.



(a) Ground truth (b) Observed (c) KL (d) Framelet + KL (e) FFT + KL (f) Framelet + FFT + KL (CT-LRTC)

Fig. 13. The recovered results by different methods on color image *Airplane* of size $512 \times 512 \times 3$ under SR $= 20\%$. (a) The ground truth; (b) The observed image; (c-f) The recovered results by CT-LRTC without framelet and Fourier transforms, CT-LRTC without Fourier transform, CT-LRTC without framelet transform, and the proposed CT-LRTC, respectively. The corresponding PSNR and SSIM values are displayed above the image.



Ground truth Observed CA CT-LRTC Ground truth Observed CA CT-LRTC

Fig. 14. The recovered results by different methods for diverse sampling and data. From top to bottom: the color images *Sculpture* and *Beach* from the testing dataset with tube-wise sampling; the color images *Cherry* and *Sailboat* not from the testing dataset with tube-wise sampling; and the color images *Peppers* and *Fruits* not from the testing dataset with element-wise sampling. For better visualization, under each image, we show enlargements of a demarcated patch and the corresponding error map (difference from the ground-truth). Error maps with less color information indicate better restoration performance.

4) Effects of the Three Transforms: Numerically, to further clarify that the inner-correlation between these three transforms are mutually collaborative, we have added experiments to test the performance of different combinations of these three transforms in Fig. 13. From Fig. 13, we can observe that the recovered results with KL and Framelet + KL are color distorted, which implies the necessity of Fourier transform in the spectral (i.e., color) fidelity. The recovered results with KL and FFT+KL lose some image details and textures, especially in the tail, which implies the necessity of Framelet transform in the spatial fidelity. Although the coupled transform may adversely affect real-time performance, the proposed CT-LRTC has the best visual performance in the spatial and spectral fidelity, which validates our motivation.

5) Compared With Deep Learning: To further illustrate that the proposed CT-LRTC has good generalization compared with some DL-based completion methods, some experiments are given. Here, we choose "Generative Image Inpainting with Contextual Attention" (termed as CA) [46] as the compared method, whose main idea is to utilize the surrounding image features as references during network training to make better predictions. The model is a feed-forward, fully convolutional neural network which is trained on the color image dataset with tube-wise sampling. We consider two color images from the testing dataset with tube-wise sampling, two color images not from the testing dataset with tube-wise sampling, and two color images not from the testing dataset with element-wise sampling.

TABLE V

THE PSNR/SSIM VALUES AND AVERAGE CPU TIME (IN MINUTES) OF THE RECOVERED RESULTS BY DIFFERENT METHODS ON GRAYSCALE VIDEOS UNDER DIFFERENT SRS. THE **BEST** RESULTS ARE HIGHLIGHTED IN **BOLDER** FONTS

video	SR	5%		10%		20%		Time
	Method	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	
Birdhouse	TNN	23.74	0.467	24.79	0.526	26.97	0.625	7.86
	TNN-DCT	23.25	0.475	24.85	0.530	27.02	0.629	5.22
	t-TNN	24.53	0.563	26.08	0.633	28.09	0.733	**5.00**
	PSTNN	23.21	0.413	25.07	0.523	27.17	0.630	58.01
	TRLRF	23.43	0.492	24.74	0.594	25.70	0.659	64.72
	CT-LRTC	**26.85**	**0.673**	**27.93**	**0.703**	**29.78**	**0.770**	58.77
Suzie	TNN	28.94	0.728	30.80	0.775	33.28	0.841	5.41
	TNN-DCT	29.44	0.741	31.28	0.788	33.57	0.848	4.20
	t-TNN	30.35	0.771	32.26	0.823	34.58	0.880	**4.14**
	PSTNN	29.40	0.725	31.10	0.779	33.30	0.841	20.01
	TRLRF	29.40	0.742	31.07	0.789	32.14	0.830	51.49
	CT-LRTC	**32.14**	**0.826**	**33.54**	**0.855**	**35.60**	**0.896**	38.72
Bus	TNN	19.70	0.383	21.23	0.486	23.21	0.635	2.25
	TNN-DCT	19.84	0.388	21.14	0.485	23.07	0.622	**0.99**
	t-TNN	20.87	0.488	22.45	0.606	24.85	0.748	1.14
	PSTNN	19.81	0.387	21.86	0.549	24.54	0.813	8.21
	TRLRF	19.47	0.396	20.59	0.484	22.51	0.612	11.18
	CT-LRTC	**21.84**	**0.557**	**23.06**	**0.636**	**25.10**	**0.752**	7.03
Hall	TNN	24.44	0.636	30.91	0.935	33.81	0.959	2.67
	TNN-DCT	24.65	0.648	30.85	0.933	33.70	0.957	**1.38**
	t-TNN	25.08	0.652	29.95	0.907	33.44	0.949	2.16
	PSTNN	25.31	0.615	31.30	0.937	34.09	0.960	7.62
	TRLRF	25.60	0.631	30.55	0.909	32.49	0.936	16.83
	CT-LRTC	**27.27**	**0.791**	**32.74**	**0.958**	**35.33**	**0.973**	22.10

TABLE VI

THE PSNR/SSIM VALUES AND AVERAGE CPU TIME (IN MINUTES) OF THE RECOVERED RESULTS BY DIFFERENT METHODS ON COLOR VIDEOS UNDER DIFFERENT SRS. THE **BEST** RESULTS ARE HIGHLIGHTED IN **BOLDER** FONTS

video	SR	5%		10%		20%		Time
	Method	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	
Bus	TNN	19.73	0.448	21.36	0.566	23.79	0.712	4.95
	TNN-DCT	19.76	0.445	22.27	0.556	23.56	0.698	**4.08**
	t-TNN	20.75	0.549	22.91	0.691	26.52	0.844	4.17
	PSTNN	16.38	0.343	21.27	0.569	24.11	0.723	27.01
	TRLRF	19.16	0.413	20.57	0.532	21.66	0.619	62.71
	CT-LRTC	**22.31**	**0.646**	**24.35**	**0.758**	**27.79**	**0.879**	48.20
Container	TNN	29.35	0.902	33.61	0.952	38.36	0.977	1.89
	TNN-DCT	28.46	0.884	33.42	0.948	39.28	0.980	**0.96**
	t-TNN	27.14	0.849	33.41	0.947	40.12	0.983	1.47
	PSTNN	30.45	0.911	34.44	0.954	38.90	0.977	6.60
	TRLRF	27.62	0.855	32.03	0.916	34.37	0.941	12.34
	CT-LRTC	**30.99**	**0.932**	**35.40**	**0.967**	**40.31**	**0.985**	10.09
Coastguard	TNN	24.83	0.733	27.05	0.825	30.18	0.905	1.65
	TNN-DCT	24.66	0.720	26.90	0.815	30.05	0.899	**0.87**
	t-TNN	25.01	0.767	28.43	0.877	**33.34**	**0.953**	1.29
	PSTNN	24.25	0.730	27.62	0.840	30.59	0.910	6.31
	TRLRF	23.49	0.661	25.86	0.773	27.83	0.841	9.57
	CT-LRTC	**26.55**	**0.811**	**29.08**	**0.886**	32.92	0.948	10.28

Fig. 14 shows the recovered results by different methods for diverse samplings and data. We can observe that the performance of CA is better than the proposed CT-LRTC on the testing image with the tube-wise sampling. As we expected, the performance of the proposed CT-LRTC is superior to CA on the image not from testing dataset with diverse samplings, which demonstrates the generalization ability of our method.

TABLE VII

THE PSNR/SSIM VALUES OF THE RECOVERED RESULTS BY DIFFERENT METHODS UNDER DIFFERENT SRS. THE **BEST** RESULTS ARE HIGHLIGHTED IN **BOLDER** FONTS

Dataset	SR	10%		30%		50%		70%		90%	
	Method	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Flower	TNN	22.05	0.522	28.33	0.810	33.96	0.933	40.98	0.984	51.23	**0.998**
	TNN-DCT	22.11	0.526	28.38	0.813	34.03	0.934	41.06	0.984	51.42	**0.998**
	t-TNN	23.74	0.690	29.65	0.892	34.94	0.964	41.01	0.990	49.60	**0.998**
	PSTNN	22.55	0.538	28.59	0.801	34.87	0.937	41.83	0.985	**51.62**	**0.998**
	TRLRF	22.54	0.536	27.94	0.775	31.58	0.878	34.48	0.929	39.52	0.975
	CT-LRTC	**26.73**	**0.801**	**32.35**	**0.932**	**35.57**	**0.976**	**43.72**	**0.992**	50.40	**0.998**
Cd	TNN	32.22	0.911	39.22	0.922	43.90	0.989	48.66	0.996	55.65	**0.999**
	TNN-DCT	32.65	0.924	39.53	0.976	44.20	0.991	49.00	0.997	56.04	**0.999**
	t-TNN	34.50	0.938	41.60	0.985	46.27	**0.995**	50.77	**0.998**	57.50	**0.999**
	PSTNN	32.87	0.911	40.23	0.972	44.50	0.989	49.06	0.996	55.85	**0.999**
	TRLRF	33.96	0.902	36.55	0.942	38.36	0.958	40.84	0.973	45.74	0.990
	CT-LRTC	**37.90**	**0.970**	**43.26**	**0.989**	**47.09**	**0.995**	**51.20**	**0.998**	**57.71**	**0.999**
Hall	TNN	30.91	0.935	35.90	0.971	39.54	0.985	43.65	0.993	50.33	0.998
	TNN-DCT	30.85	0.933	35.77	0.970	39.33	0.984	43.46	0.993	50.19	0.998
	t-TNN	29.95	0.907	35.85	0.966	39.76	0.984	44.02	0.993	**50.73**	0.998
	PSTNN	31.30	0.937	36.07	0.972	39.62	0.985	43.69	0.993	50.33	0.998
	TRLRF	30.55	0.909	33.13	0.944	35.02	0.961	37.33	0.976	42.15	0.992
	CT-LRTC	**32.74**	**0.958**	**37.57**	**0.982**	**40.96**	**0.991**	**44.87**	**0.996**	50.13	**0.999**

TABLE VIII

THE PSNR/SSIM/TIME (IN SECONDS) VALUES OF THE PROPOSED CT-LRTC ON THE COLOR IMAGE *Peppers* UNDER DIFFERENT FRAMELET SETTINGS WITH SR = 5%. THE **BEST** RESULTS ARE HIGHLIGHTED IN **BOLDER** FONTS

Filters	Index	Level = 1	Level = 2	Level = 3	Level = 4	Level = 5
Haar	PSNR	20.40	22.73	23.33	23.44	23.51
	SSIM	0.579	0.714	0.748	0.756	0.757
	Time(s)	**166**	189	231	288	357
Linear	PSNR	21.89	23.12	23.24	23.17	23.14
	SSIM	0.665	0.748	0.758	0.757	0.756
	Time(s)	245	313	481	635	804
Cubic	PSNR	22.89	23.15	**23.66**	22.68	22.64
	SSIM	0.725	0.763	**0.767**	0.750	0.748
	Time(s)	625	823	932	1024	1206

V. CONCLUSION

In this paper, we proposed a novel low-rank tensor representation under the coupled transforms, which can characterize both the global correlation and local geometric details in a unified framework, and obtain a better low multi-rank approximation. The proposed low-rank tensor representation can be formulated via the two-dimensional framelet transform, Fourier transform, and Karhunen–Loéve transform (via singular value decomposition). Further, we formulated a novel model for LRTC (named as CT-LRTC) by promoting the proposed low-rank tensor representation. Then, we developed an efficient ADMM algorithm to optimize the proposed CT-LRTC, in which every variable has a closed-form solution in each iteration. Finally, experimental examples of real-world imaging data illustrated that the proposed CT-LRTC outperforms many existing approaches in both qualitative and quantitative aspects. In future work, we will work hard to speed up CT-LRTC, establish the theoretical justification, and expand the low-rank tensor representation under coupled transforms to other applications, such as denoising [61], [62] and subspace clustering [63], [64].

REFERENCES

[1] L. Zhuang, C.-H. Lin, M. A. T. Figueiredo, and J. M. Bioucas-Dias, "Regularization parameter selection in minimum volume hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 9858–9877, Dec. 2019.

[2] L. Zhuang and J. M. Bioucas-Dias, "Fast hyperspectral image denoising and inpainting based on low-rank and sparse representations," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 730–742, Mar. 2018.

[3] X. Fu, K. Huang, N. D. Sidiropoulos, and W.-K. Ma, "Nonnegative matrix factorization for signal and data analytics: Identifiability, algorithms, and applications," *IEEE Signal Process. Mag.*, vol. 36, no. 2, pp. 59–80, Mar. 2019.

[4] L. Li, F. Jiang, and R. Shen, "Total variation regularized reweighted low-rank tensor completion for color image inpainting," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2018, pp. 2152–2156.

[5] W. Hu, Y. Yang, W. Zhang, and Y. Xie, "Moving object detection using tensor-based low-rank and saliently fused-sparse decomposition," *IEEE Trans. Image Process.*, vol. 26, no. 2, pp. 724–737, Feb. 2017.

[6] Y. T. Wang, X. L. Zhao, T. X. Jiang, L. J. Deng, Y. Chang, and T. Z. Huang, "Rain streaks removal for single image via Kernel-guided convolutional neural network," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Aug. 21, 2020, doi: 10.1109/TNNLS.2020.3015897.

[7] X.-L. Zhao, F. Wang, T.-Z. Huang, M. K. Ng, and R. J. Plemmons, "Deblurring and sparse unmixing for hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 4045–4058, Jul. 2013.

[8] X. Fu, K. Huang, B. Yang, W.-K. Ma, and N. D. Sidiropoulos, "Robust volume minimization-based matrix factorization for remote sensing and document clustering," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6254–6268, Dec. 2016.

[9] W. He, L. Yuan, and N. Yokoya, "Total-variation-regularized tensor ring completion for remote sensing image reconstruction," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2019, pp. 8603–8607.

[10] W. He, Q. Yao, C. Li, N. Yokoya, and Q. Zhao, "Non-local meets global: An integrated paradigm for hyperspectral denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2019, pp. 1–12.

[11] H. Xu, H. Zhang, W. He, and L. Zhang, "Superpixel-based spatial-spectral dimension reduction for hyperspectral imagery classification," *Neurocomputing*, vol. 360, pp. 138–150, Sep. 2019.

[12] X. Li, L. Zhang, B. Du, and L. Zhang, "On gleaning knowledge from cross domains by sparse subspace correlation analysis for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 6, pp. 3204–3220, Jun. 2019.

[13] 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.

[14] M. Hingee, H. Tkalčić, A. Fichtner, and M. Sambridge, "Seismic moment tensor inversion using a 3-D structural model: Applications for the Australian region," *Geophys. J. Int.*, vol. 184, no. 2, pp. 949–964, Feb. 2011.

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

[16] 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.

[17] 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.

[18] 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, Nov. 2019.

[19] 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, Jan. 2013.

[20] 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, Feb. 2017.

[21] Q. Zhao, L. Zhang, and A. Cichocki, "Bayesian CP factorization of incomplete tensors with automatic rank determination," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 9, pp. 1751–1763, Sep. 2015.

[22] A.-H. Phan, P. Tichavský, and A. Cichocki, "Partitioned hierarchical alternating least squares algorithm for CP tensor decomposition," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Mar. 2017, pp. 2542–2546.

[23] 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 Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 8251–8259.

[24] M. Ashraphijuo, V. Aggarwal, and X. Wang, "A characterization of sampling patterns for low-Tucker-rank tensor completion problem," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2017, pp. 531–535.

[25] 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 Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 3842–3849.

[26] 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.

[27] J. Håstad, "Tensor rank is NP-complete," *J. Algorithms*, vol. 11, no. 4, pp. 644–654, Dec. 1990.

[28] Z. Long, Y. Liu, L. Chen, and C. Zhu, "Low rank tensor completion for multiway visual data," *Signal Process.*, vol. 155, pp. 301–316, Feb. 2019.

[29] R. Grotheer, S. Li, A. Ma, D. Needell and J. Qin, "Stochastic iterative hard thresholding for low-Tucker-rank tensor recovery," in *Proc. Inf. Theory Appl. Workshop*, Feb. 2020, pp. 1–5.

[30] M. Ashraphijuo, V. Aggarwal, and X. Wang, "Deterministic and probabilistic conditions for finite completability of low-Tucker-rank tensor," *IEEE Trans. Inf. Theory*, vol. 65, no. 9, pp. 5380–5400, Sep. 2019.

[31] Y. Wang *et al.*, "Rubik: Knowledge guided tensor factorization and completion for health data analytics," in *Proc. KDD*, 2015, pp. 1265–1274.

[32] Z. Kong and X. Yang, "Color image and multispectral image denoising using block diagonal representation," *IEEE Trans. Image Process.*, vol. 28, no. 9, pp. 4247–4259, Sep. 2019.

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

[34] S. Xue, W. Qiu, F. Liu, and X. Jin, "Low-rank tensor completion by truncated nuclear norm regularization," in *Proc. Int. Conf. Pattern Recognit.*, Aug. 2018, pp. 2600–2605.

[35] Y.-B. Zheng, T.-Z. Huang, X.-L. Zhao, T.-X. Jiang, T.-Y. Ji, and T.-H. Ma, "Tensor N-tubal rank and its convex relaxation for low-rank tensor recovery," *Inf. Sci.*, vol. 532, pp. 170–189, Sep. 2020.

[36] S. Cai, Q. Luo, M. Yang, W. Li, and M. Xiao, "Tensor robust principal component analysis via non-convex low rank approximation," *Appl. Sci.*, vol. 9, no. 7, p. 1411, Apr. 2019.

[37] 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.

[38] T.-Y. Ji, T.-Z. Huang, X.-L. Zhao, T.-H. Ma, and L.-J. Deng, "A non-convex tensor rank approximation for tensor completion," *Appl. Math. Model.*, vol. 48, pp. 410–422, Aug. 2017.

[39] Y. Mu, P. Wang, L. Lu, X. Zhang, and L. Qi, "Weighted tensor nuclear norm minimization for tensor completion using tensor-SVD," *Pattern Recognit. Lett.*, vol. 130, pp. 4–11, Feb. 2020.

[40] W.-H. Xu *et al.*, "Laplace function based nonconvex surrogate for low-rank tensor completion," *Signal Process., Image Commun.*, vol. 73, pp. 62–69, Apr. 2019.

[41] 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 Conf. Comput. Vis. Pattern Recognit.*, Jun. 2019, pp. 5996–6004.

[42] E. Kernfeld, M. Kilmer, and S. Aeron, "Tensor–tensor products with invertible linear transforms," *Linear Algebra Appl.*, vol. 485, pp. 545–570, Nov. 2015.

[43] C. Li, W. He, L. Yuan, Z. Sun, and Q. Zhao, "Guaranteed matrix completion under multiple linear transformations," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2019, pp. 11136–11145.

[44] Y. Wang, X. Liao, and S. Lin, "Rescaled boosting in classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 9, pp. 2598–2610, Sep. 2019.

[45] 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.

[46] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang, "Generative image inpainting with contextual attention," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 5505–5514.

[47] M.-C. Sagong, Y.-G. Shin, S.-W. Kim, S. Park, and S.-J. Ko, "PEPSI : Fast image inpainting with parallel decoding network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2019, pp. 11352–11360.

[48] C.-T. Li, W.-C. Siu, Z.-S. Liu, L.-W. Wang, and D. P.-K. Lun, "DeepGIN: Deep generative inpainting network for extreme image inpainting," pp. 1–13, 2020, *arXiv:2008.07173*. [Online]. Available: http://arxiv.org/abs/2008.07173

[49] 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, 2010.

[50] 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 J. Matrix Anal. Appl.*, vol. 34, no. 1, pp. 148–172, Jan. 2013.

[51] J.-F. Cai, H. Ji, C. Liu, and Z. Shen, "Framelet-based blind motion deblurring from a single image," *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 562–572, Feb. 2012.

[52] A. Ron and Z. Shen, "Affine systems in $l_2(\mathbb{R}^d)$: The analysis of the analysis operator," *J. Funct. Anal.*, vol. 148, no. 2, pp. 408–447, 1997.

[53] J.-F. Cai, J. K. Choi, and K. Wei, "Data driven tight frame for compressed sensing MRI reconstruction via off-the-grid regularization," *SIAM J. Imag. Sci.*, vol. 13, no. 3, pp. 1271–1301, 2020.

[54] G. Zhang, Y. Xu, and F. Fang, "Framelet-based sparse unmixing of hyperspectral images," *IEEE Trans. Image Process.*, vol. 25, no. 4, pp. 1516–1529, Apr. 2016.

[55] A. Chai and Z. Shen, "Deconvolution: A wavelet frame approach," *Numerische Math.*, vol. 106, no. 4, pp. 529–587, May 2007.

[56] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956–1982, Jan. 2010.

[57] 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.

[58] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[59] W. Hu, D. Tao, W. Zhang, Y. Xie, and Y. Yang, "The twist tensor nuclear norm for video completion," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 12, pp. 2961–2973, Dec. 2017.

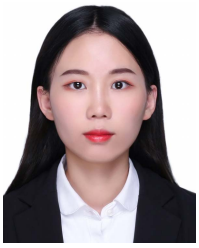
[60] L. Yuan, C. Li, D. Mandic, J. Cao, and Q. Zhao, "Tensor ring decomposition with rank minimization on latent space: An efficient approach for tensor completion," in in *Proc. AAAI Conf. Artif. Intell.*, vol. 33, 2019, pp. 9151–9158.

[61] J.-L. Wang, T.-Z. Huang, T.-H. Ma, X.-L. Zhao, and Y. Chen, "A sheared low-rank model for oblique stripe removal," *Appl. Math. Comput.*, vol. 360, pp. 167–180, Nov. 2019.

[62] J.-L. Wang, T.-Z. Huang, X.-L. Zhao, J. Huang, T.-H. Ma, and Y.-B. Zheng, "Reweighted block sparsity regularization for remote sensing images destriping," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 12, pp. 4951–4963, Dec. 2019.

[63] X. Peng, C. Lu, Z. Yi, and H. Tang, "Connections between nuclear-norm and Frobenius-norm-based representations," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 1, pp. 218–224, Jan. 2018.

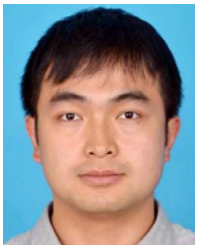
[64] X. Peng, Z. Yu, Z. Yi, and H. Tang, "Constructing the L_2-graph for robust subspace learning and subspace clustering," *IEEE Trans. Cybern.*, vol. 47, no. 4, pp. 1053–1066, Apr. 2017.



Jian-Li Wang received the B.S. degree in mathematics and applied mathematics from Neijiang Normal University, Neijiang, China, in 2017. She is currently pursuing the Ph.D. degree with the School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, China. Her current research interests include sparse and low-rank modeling for image processing problems.

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. 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. He 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*, China.

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 interests include the models and algorithms of high-dimensional image processing problems.



Tai-Xiang Jiang received the Ph.D. degree in mathematics from the University of Electronic Science and Technology of China (UESTC) in 2019. He was a co-training Ph.D. student with the University of Lisbon supervised by Prof. J. M. Bioucas-Dias from 2017 to 2018. He was the Research Assistant with the Hong Kong Baptist University 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. His research interests include sparse and low-rank modeling and tensor decomposition for multi-dimensional image processing, especially on the low-level inverse problems for multi-dimensional images.

Michael K. Ng (Senior Member, IEEE) is currently the Director of the Research Division for Mathematical and Statistical Science, the Chair Professor with the Department of Mathematics, The University of 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.