

Fusion of low-rankness and smoothness under learnable nonlinear transformation for tensor completion

Yigong Zhang^{a,1}, Zhihui Tu^{a,1}, Jian Lu^{a,b,c,*}, Chen Xu^a, Lixin Shen^d

^a Shenzhen Key Laboratory of Advanced Machine Learning and Applications, School of Mathematical Sciences, Shenzhen University, Shenzhen 518060, China

^b National Center for Applied Mathematics Shenzhen (NCAMS), Shenzhen 518055, China

^c Pazhou Lab, Guangzhou 510335, China

^d Department of Mathematics, Syracuse University, Syracuse, NY 13244, United States of America

ARTICLE INFO

Keywords:

Tensor completion

Learnable nonlinear transform

Fused prior

Tensor correlated total variation

ABSTRACT

The recently proposed tensor correlated total variation (t-CTV) has achieved success in tensor completion. It utilizes the low-rank structure of the gradient tensor under a unified linear transform to jointly encode low-rankness and smoothness priors. However, fixed linear transforms have inherent limitations in fully characterizing gradient tensors in different directions and adapting them to tensors from diverse categories. In this work, we propose the nonlinear tensor correlated total variation (NTCTV) regularization term that leverages the low-rank correlations of the gradient tensor under the learnable nonlinear transformation, providing a more natural approach to fuse the low-rankness and smoothness priors. Specifically, our approach learns the optimal nonlinear implicit low-rank structure of the gradient tensor along different modes separately, and then achieves the expression of fused prior information in a coupled manner. Furthermore, we propose the NTCTV-based tensor completion model and design the proximal alternating minimization (PAM) algorithm to efficiently solve the optimization model. Moreover, we provide a theoretical proof of the global convergence of the algorithm to a critical point. Comprehensive experimental results for hyperspectral images, medical images, multispectral images, and videos demonstrate that the proposed method achieves substantial quantitative and qualitative improvements over many state-of-the-art tensor completion techniques.

1. Introduction

With the exponential increase in data volume and technological advancements, tensors, as multi-dimensional extensions of arrays, naturally represent various types of high-dimensional image data, such as hyperspectral images (HSIs), multispectral images (MSIs), computed tomography (CT), and magnetic resonance images (MRI). Compared to matrix and vector representations, tensor tends to represent native multidimensional structures in data with greater fidelity and accuracy. Thus, there are a wide range of applications involving tensor representations, including anomaly detection [1], computer vision [2], biomedical imaging [3] and recommender systems [4].

Unfortunately, due to defects in the collection equipment and interference from complex real-world situations, the tensor data collected often exhibit overt deteriorations, such as corruptions or absent entries, which significantly degrade the visual quality and affect the subsequent processing tasks. Tensor completion has thus evolved into

one of the elementary issues in tensor research. Mathematically, the problem of recovering the target tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ involves utilizing the potential structural prior information provided by the observation tensor $\mathcal{P}_\Omega(\mathcal{Y}) = \mathcal{P}_\Omega(\mathcal{X})$, where $\mathcal{Y}$ represents the unknown underlying tensor, Ω represents the index of the known elements, and $\mathcal{P}(\cdot)$ denotes the projection operator. This problem, known as tensor completion (TC) [5], falls under the category of typical inverse problems.

To reasonably and accurately estimate the recovery tensor in the context of such an ill-posed problem, extracting and encoding the intrinsic prior structural information inherent in the tensor data into regularization terms are necessary. Low-rankness (denoted by $\mathbf{R}$), which is the most frequently employed global prior, assumes that multi-dimensional tensor data inherently exhibit redundant properties along specific tensor modes. This implies that major information about the tensor is situated within an appropriate low-dimensional subspace, elucidating its information linkage macroscopically along a certain

* Corresponding author at: Shenzhen Key Laboratory of Advanced Machine Learning and Applications, School of Mathematical Sciences, Shenzhen University, Shenzhen 518060, China.

E-mail addresses: zhangyigong2022@email.szu.edu.cn (Y. Zhang), tuzhihui2021@email.szu.edu.cn (Z. Tu), jianlu@szu.edu.cn (J. Lu), chenxu@szu.edu.cn (C. Xu), lshen03@syr.edu (L. Shen).

¹ Yigong Zhang and Zhihui Tu contributed equally to this work.

tensor mode. Thus, this prior gives rise to the classical low-rank tensor completion (LRTC) model:

$$\min_{\mathcal{X}} \mathfrak{R}(\mathcal{X}), \text{ s.t. } \mathcal{P}_{\Omega}(\mathcal{Y}) = \mathcal{P}_{\Omega}(\mathcal{X}), \quad (1)$$

where $\mathfrak{R}(\cdot)$ is a regularization operator representing low-rankness. Unlike the matrix rank definition, the tensor rank definitions are not unique. Over the past few decades, substantial research efforts on tensor rank have been based on different tensor decompositions, such as the classical CANDECOMP/PARAFAC (CP) [6] decomposition and Tucker [7] decomposition. Computing the CP rank, which represents the minimum number of rank one factor tensors in CP decomposition, is an NP-hard task, and its convex relaxation is intractable in practical applications. This renders tensor completion based on CP rank challenging. To avoid this issue, the Tucker rank, which is defined as a vector composed of ranks of each mode- k unfolding matrix of the tensor, and its convex relaxation are used more widely. Liu et al. [5] first proposed the sum of nuclear norms (SNN) [5] as a convex surrogate of the Tucker rank and designed the LRTC method based on the SNN. Xu et al. [8] proposed a new LRTC method that executes low-rank matrix factorization on the all-mode matricizations of the tensor. However, the SNN is limited by the fact that it is not the tightest convex relaxation [9]. In addition, tensor network decomposition frameworks, including tensor train (TT) [10] decomposition, tensor ring (TR) [11] decomposition and tensor fully-connected tensor network (FCTN) [12] decomposition, have been proposed. Each of these decompositions was successful in the realm of high-order LRTC [13–17].

Recently, the tensor singular value decomposition (t-SVD) [18] framework has been proposed and widely applied in LRTC. Based on the tensor-tensor product (t-product), t-SVD establishes a complete tensor decomposition algebraic framework. Under the t-SVD decomposition framework, the tensor tubal rank and its corresponding convex surrogate tensor nuclear norm (TNN) [19] achieved excellent performance. Thus, model (1) based on the t-SVD framework has received considerable attention in characterizing the underlying tensor data that satisfy $\mathbf{R}$ -prior structure. Kernfeld et al. [20] observed that the t-product can be implemented using the discrete Fourier transform (DFT), which has generated substantial interest among researchers in novel transforms founded on the t-product. Jiang et al. [21] proposed a semi-invertible framelet transform that overcame the invertibility constraint. It should be noted that above transforms were predefined. Thus, more researchers have considered data-driven transforms which can better explore the $\mathbf{R}$ structure prior of real data. Luo et al. [22] proposed a self-supervised method to learn transforms adaptively. Li et al. [23] used single layer semi-orthogonal neural network to determine a suitable transform. Despite the favorable results achieved by all the aforementioned purely $\mathbf{R}$ -prior methods, real visual data usually have more intrinsic prior structural information.

Apart from low-rankness prior, the smoothness (denoted by $\mathbf{S}$) prior which is particularly evident in the data structure of visual data, is widely used to smooth adjacent structural information within a tensor. The $\mathbf{S}$ -prior describes the relative local similarity structure within the tensor data, specifically manifested as the adjacent element values along certain tensor modes prone to steady change. It is commonly characterized using total variation (TV), which can be further categorized into two main types: anisotropic TV (TV-1) and isotropic TV (TV-2) [24]. Researchers continually adapt TV based on the nature of the data being analyzed. For HSIs, spectral-spatial TV (SSTV) is often employed to model a smooth structure along its spatial axes [25]. Similarly, Tom and George [26] formulated a temporal-spatial TV for videos.

In practice, to ensure effective tensor completion, the $\mathbf{S}$ -prior, as a type of local prior structure, is commonly employed with the global $\mathbf{R}$ -prior to achieve better performance. Most related studies incorporate the $\mathbf{S}$ -prior into the $\mathbf{R}$ -prior model in an additive manner ($\mathbf{R} + \mathbf{S}$) as follows:

$$\min_{\mathcal{X}} \mathfrak{R}(\mathcal{X}) + \lambda \mathfrak{S}(\mathcal{X}), \text{ s.t. } \mathcal{P}_{\Omega}(\mathcal{Y}) = \mathcal{P}_{\Omega}(\mathcal{X}), \quad (2)$$

where $\mathfrak{S}(\cdot)$ denotes the regularization operator that measures the smoothness prior. By representing the regularization operator $\mathfrak{S}$ in model (2) using various methods, a series of works have achieved promising results for tensor completion. Qiu et al. [27] used the TNN to explore the $\mathbf{R}$ -prior and the simple TV-1 to depict the $\mathbf{S}$ -prior. Yokota et al. [28] achieved tensor completion by incorporating TV-1/TV-2 into a CP-rank model. Other typical works along this line include [29–35].

While the $\mathbf{R} + \mathbf{S}$ -prior model (2) has achieved satisfactory results in various applications, it still has some limitations. On the one hand, the $\mathbf{R}$ -prior and $\mathbf{S}$ -prior are commonly coupled together within real-world natural data, making it challenging to fully represent the data in a straightforward additive manner. On the other hand, determining the valid trade-off parameters between the $\mathbf{R}$ and $\mathbf{S}$ regularizers is non-trivial, since their interaction frequently has a pronounced impact on the performance of (2). Recently, Wang et al. [36] overcame these constraints and improved the fusion prior model by proposing a novel fusion prior regularization operator called the tensor correlated total variation (t-CTV). This operator couples two different prior operators $\mathfrak{R}$ and $\mathfrak{S}$ instead of simply adding them, and provides a theoretical precision recovery guarantee for tensor completion. This implies that the $\mathbf{R} + \mathbf{S}$ -prior is essentially a coupled prior, denoted by $\mathbf{R} \circ \mathbf{S}$ -prior. Thus, the corresponding model can be formulated as follows:

$$\min_{\mathcal{X}} \mathfrak{R} \circ \mathfrak{S}(\mathcal{X}), \text{ s.t. } \mathcal{P}_{\Omega}(\mathcal{Y}) = \mathcal{P}_{\Omega}(\mathcal{X}), \quad (3)$$

where $\mathfrak{R} \circ \mathfrak{S}(\cdot)$ is a fusion prior regularization operator which measures the coupling prior of low-rankness and smoothness. Specifically, t-CTV describes the redundant information inside the tensor by considering the implicit low-rank structure of the directional difference. Suppose that $\mathcal{I}$ is a priori set consisting of directions represented by different modes, with cardinality γ , along which $\mathcal{X}$ has smooth continuity. Then, the model (3) is equivalent to

$$\min_{\mathcal{X}} \sum_{k \in \mathcal{I}} \frac{1}{\gamma} \|\nabla_k(\mathcal{X})\|_{*, \mathcal{L}}, \text{ s.t. } \mathcal{P}_{\Omega}(\mathcal{Y}) = \mathcal{P}_{\Omega}(\mathcal{X}), \quad (4)$$

where $\nabla_k(\mathcal{X}) \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ represents the difference along the k th mode of $\mathcal{X}$, characterized by TNN $\|\cdot\|_{*, \mathcal{L}}$ under t-SVD framework with linear transform $\mathcal{L}$. It is important to note that the gradient tensors in all directions in (4) are projected onto the same linear transform domain, which raises some concerns. First, using an identical transform may not fully facilitate the extraction of the intrinsic-fused-prior-structural information from the original data across different smoothing orientations. Additionally, linear constraints may limit the ability to adequately model the nonlinear traits exhibited in real-world data.

To address these issues, we propose a type of learned nonlinear transform, denoted as $\varphi : \mathbb{R}^{n_1 \times n_2 \times n_3} \rightarrow \mathbb{R}^{r \times n_2 \times n_3}$, for the t-CTV method. This enables joint low-rank and smooth tensor completion without the need to balance the parameters. Specifically, for each directional difference with the fused $\mathbf{R} \circ \mathbf{S}$ -prior, we adaptively determine a semi-orthogonal transform embedded within a nonlinear function. The learned nonlinear transforms can effectively explore the latent structural knowledge present in each prior direction, while utilizing a smaller r to project the observation tensor into a small-scale tensor to reduce computational complexity. Furthermore, processing the fused $\mathbf{R} \circ \mathbf{S}$ -prior in the form of tensor directional difference eliminates the need to select trade-off parameters between the two priors. Utilizing the proposed learnable nonlinear transform, we propose a regularization operator to quantify the fused $\mathbf{R} \circ \mathbf{S}$ -prior, termed nonlinear tensor correlated total variation (NTCTV), which simultaneously captures the $\mathbf{R}$ and $\mathbf{S}$ priors of the underlying tensor under within the nonlinear low-dimensional subspace. Then, we formulate the fused $\mathbf{R} \circ \mathbf{S}$ -prior model based on this operator. Furthermore, we developed solving algorithms based on multi-block proximal alternating minimization (PAM) [37, 38], and analyzed their convergence and effectiveness through theoretical analysis and experiments with real data. The main contributions can be summarized as follows:

- We propose a learnable nonlinear regularizer NTCTV to characterize the fused $\mathbf{R} \circ \mathbf{S}$ -prior with a unique term. The proposed method adaptively learns the optimal nonlinear transform for each gradient tensor with a fused prior.
- Under low-rankness and smoothness priors, we employ the NTCTV for classical TC task based on the PAM algorithm, obtaining the closed-form updating solution for each variable. Additionally, we theoretically analyze the computational complexity and convergence of the proposed algorithm.
- Comprehensive experiments on tensor completion validate the superiority of the proposed NTCTV-based method over state-of-the-art approaches, including those considering $\mathbf{R}$, $\mathbf{R} + \mathbf{S}$ and $\mathbf{R} \circ \mathbf{S}$ -priors. Typically, experimental results show the remarkable performance of our method in HSIs completion, even at high random voxel missing of up to 95%, 98%, and 99%. Additionally, a visual examination of the convergence and the impact of associated parameters are conducted, yielding dependable parameter selection guidelines for varied image categories.

The remainder of this paper is organized as follows. In Section 2, we introduce the t-CTV framework and related definitions. In Section 3, we propose an NTCTV regularizer for tensor completion with a fused prior and establish the corresponding algorithm with a theoretical convergence guarantee and computational complexity. In Section 4, we present TC experiments on HSIs, MSIs, videos, CT, and MRI medical images, respectively. In Section 5, we discuss the proposed algorithm. Section 6 concludes the paper.

2. Notations and preliminaries

In this section, we provide a brief introduction to some basic notations and the t-CTV framework. For more details, please refer to [36, 39, 40].

In this study, we use x , $\mathbf{x}$, $\mathbf{X}$ and $\mathcal{X}$ to denote the scalars, vectors, matrices, and tensors, respectively. For an order-3 tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, its (i, j, k) th element is denoted as $\mathcal{X}(i, j, k)$ or x_{ijk} , and the k th frontal slice of tensor $\mathcal{X}$ is denoted as $\mathcal{X}(:, :, k)$ or $\mathcal{X}^{(k)}$. We then denote $\text{diag}(\mathcal{X})$ with size $n_1 n_3 \times n_2 n_3$, as the block diagonal matrix constructed by all frontal slices. Besides, we denote the ℓ_1 norm as $\|\mathcal{X}\|_1 = \sum_{ijk} |x_{ijk}|$ and the Frobenius norm as $\|\mathcal{X}\|_F = \sqrt{\sum_{ijk} x_{ijk}^2}$. Both the ℓ_1 norm and Frobenius norm can be reduced to the vector or matrix norms. The matrix nuclear norm of $\mathbf{X}$ is denoted as $\|\mathbf{X}\|_* = \sum_i \sigma_i(\mathbf{X})$, where $\sigma_i(\mathbf{X})$ is the i th singular value of $\mathbf{X}$.

Definition 1 (Tensor Mode- k Product [40]). For $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_d}$ and a matrix $\mathbf{M} \in \mathbb{R}^{r_k \times n_k}$, tensor mode- k product is defined as

$$\mathcal{A} := \mathcal{X} \times_k \mathbf{M} = \text{fold}_k(\mathbf{M} \mathcal{X}_{(k)}),$$

where $\mathcal{X}_{(k)}$ is the mode- k matricization of $\mathcal{X}$ and $\text{fold}_k(\cdot)$ is the corresponding inverse operator of matricization that rearranges the elements of a matrix into a tensor.

Definition 2 (Tensor Tubal Rank [39]). For $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ with t-SVD $\mathcal{X} = \mathcal{U} *_{\mathcal{L}} \mathcal{S} *_{\mathcal{L}} \mathcal{V}^T$, its tensor tubal rank is defined as

$$\text{rank}_{\mathcal{L}}(\mathcal{X}) := \#\{i : \mathcal{S}(i, i, :) \neq \mathbf{0}\},$$

where $*_{\mathcal{L}}$ is the t-product under linear transform $\mathcal{L}$ and $\#$ denotes the cardinality of a set.

Definition 3 (TNN [39]). Under the t-SVD framework with a linear transform $\mathcal{L}$, the TNN of order-3 tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is defined as

$$\|\mathcal{X}\|_{\star, \mathcal{L}} := \frac{1}{\ell} \sum_{i=1}^{n_3} \|(\mathcal{L}(\mathcal{X}))^{(i)}\|_*,$$

where $\mathcal{L}(\mathcal{X}) = \mathcal{X} \times_3 \mathbf{L}$, $\mathbf{L}$ is a transform matrix sized $n_3 \times n_3$ with $\mathbf{L}\mathbf{L}^T = \ell \mathbf{I}_{n_3}$.

Definition 4 (Gradient Tensor [36]). For $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_d}$, the gradient tensor along the k th mode is defined as

$$\mathcal{G}_k := \nabla_k \mathcal{X} = \mathcal{X} \times_k \mathbf{D}_{n_k},$$

where ∇_k is the difference operator along the k th mode, $\mathbf{D}_{n_k} \in \mathbb{R}^{n_k \times n_k}$ is a row circulant matrix of $(-1, 1, 0, \dots, 0)$.

Definition 5 (t-CTV [36]). For $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_d}$, denote Γ as a priori set consisting of directions along which $\mathcal{X}$ satisfies the fused $\mathbf{R} \circ \mathbf{S}$ -prior. The t-CTV norm is defined as follows:

$$\|\mathcal{X}\|_{t\text{-CTV}} := \frac{1}{\gamma} \sum_{k \in \Gamma} \|\mathcal{G}_k\|_{\star, \mathcal{L}}, \quad (5)$$

where γ is the cardinality of priori set Γ .

3. Proposed model and optimization algorithm

3.1. Motivations

As shown in (5), for $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ with fused $\mathbf{R} \circ \mathbf{S}$ -prior, the t-CTV exploits the implicit structural information of gradient tensor in the transform domain to achieve the fusion metric of $\mathbf{R}$ and $\mathbf{S}$. However, it utilizes the same predefined linear transform to express low-rank structures of different gradient tensors, which has limited adaptability to tensor data and hinders the expression of nonlinear structural information.

To tackle these problems, we propose a learnable nonlinear transform to adapt to different types of prior structural information. The proposed transform exhibits a composite structure in which a linear matrix is embedded within a nonlinear function. Specifically, for $k \in \Gamma$, the learnable nonlinear transform is defined as:

$$\varphi_k(\mathcal{G}_k) = \phi(\mathcal{G}_k \times_3 \mathbf{T}_k), \quad (6)$$

where $\phi(\cdot)$ denotes the element-wise nonlinear activation functions, $\mathcal{G}_k$ denotes the gradient tensor of $\mathcal{X}$ along the k th mode and $\mathbf{T}_k \in \mathbb{R}^{r \times n_3}$ ($k \in \Gamma$) denotes the learned linear semi-orthogonal matrix satisfying $\mathbf{T}_k \mathbf{T}_k^T = \mathbf{I}_{r \times r}$. In this study, a hyperbolic tangent function ($\tanh$) [41] is employed as the activation transform $\phi(\cdot)$.

From the perspective of $\mathbf{R}$ -prior encoding, the gradient tensor demonstrates a consistent low-rankness with the original tensor, as evidenced by the singular value curve shown in Fig. 1(a). Hence, leveraging the $\mathbf{R}$ -prior of the gradient tensor assists in restoring the original tensor. For this purpose, we employ the proposed data-driven nonlinear transform (6) and define the low-rank metric for the gradient tensor as follows:

$$\|\mathcal{G}_k\|_{\star, \varphi_k} = \sum_{i=1}^r \|\varphi_k(\mathcal{G}_k)^{(i)}\|_*. \quad (7)$$

Furthermore, we construct the nonlinear regularizer corresponding to the fused $\mathbf{R} \circ \mathbf{S}$ -prior, named the *nonlinear tensor correlated total variation* (NTCTV). Fig. 1 depicts the recovery process based on NTCTV.

Definition 6 (Nonlinear Tensor Correlated Total Variation). For $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ with Γ as a priori set consisting of directions along which $\mathcal{X}$ equips fused $\mathbf{R} \circ \mathbf{S}$ -prior, $\mathcal{G}_k$ ($k \in \Gamma$) are correlated gradient tensors. We define the nonlinear tensor correlated total variation norm as

$$\|\mathcal{X}\|_{NTCTV} = \frac{1}{\gamma} \sum_{k \in \Gamma} \|\mathcal{G}_k\|_{\star, \varphi_k} = \frac{1}{\gamma} \sum_{k \in \Gamma} \sum_{i=1}^r \|\varphi_k(\mathcal{G}_k)^{(i)}\|_*, \quad (8)$$

where $\gamma = \#\{k, k \in \Gamma\}$ is the cardinality of the priori set.

From the perspective of $\mathbf{S}$ -prior encoding, the proposed NTCTV encourages sparsity of the gradient tensor data along the smoothness prior modes in the transform domain, which is similar to $\mathbf{S}$ -prior regularizers of TV norm, such as $\|\mathcal{X}\|_{TV_1} := \sum_{k \in \Gamma} \|\mathcal{G}_k\|_1$ and $\|\mathcal{X}\|_{TV_2} :=$

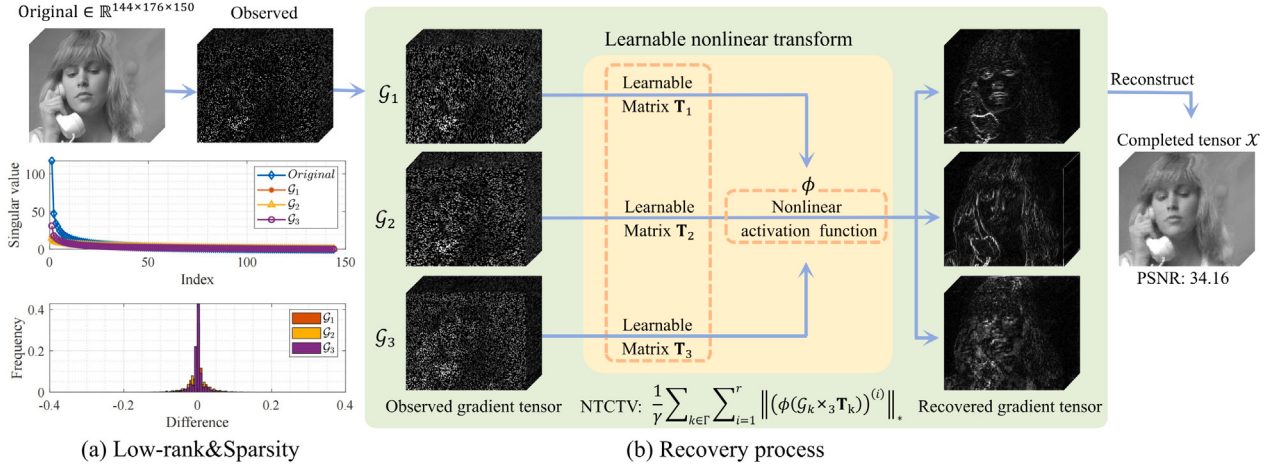


Fig. 1. The pipeline of the proposed NTCTV for tensor inpainting on *Suzie* (SR = 0.15), along with the singular value curves and pixel distribution maps of the gradient tensor.

$\sum_{k \in \Gamma} \|\mathcal{G}_k\|_F$. Moreover, both NTCTV and TV regularization terms characterize the corresponding properties by measuring the energy of the gradient tensor under a specific norm. Proposition 1 demonstrates that the proposed NTCTV achieves a characterization of local smoothness similar to that of the TV norm by exploiting this energy minimization effect in the nonlinear transform domain.

Proposition 1. Suppose that $\phi(\cdot)$ is the nonlinear activation function satisfying $|\phi(x)| \leq |x|$ for $x \in [-1, 1]$, $\mathbf{T}_k \in \mathbb{R}^{n_3 \times n_3}$ ($k \in \Gamma$) represents the learnable matrix corresponding to the gradient tensor of a third-order tensor $\mathcal{X}$ with $\text{rank}_t(\mathcal{X}) = R$. Thus, it holds that

$$\|\mathcal{X}\|_{TV} \lesssim \|\mathcal{X}\|_{NTCTV} \lesssim \sqrt{R} \|\mathcal{X}\|_{TV}.$$

Proof. The proof primarily uses some properties of the t-SVD framework and the inequalities between matrix norms. On the one hand, for any $k \in \Gamma$, according to (7), we have

$$\begin{aligned} \|\mathcal{G}_k\|_{\star, \varphi_k} &= \|bdiag(\varphi_k(\mathcal{G}_k))\|_{\star} \leq \sqrt{n_3 \text{rank}_t(\mathcal{G}_k)} \|bdiag(\varphi_k(\mathcal{G}_k))\|_F \\ &\leq \sqrt{n_3 R} \|bdiag(\mathcal{G}_k \times_3 \mathbf{T}_k)\|_F \\ &= \sqrt{n_3 R} \|\mathbf{T}_k \mathcal{G}_{k(3)}\|_F \\ &= \sqrt{n_3 R} \sqrt{\text{Tr}(\mathcal{G}_{k(3)} \mathbf{I}_{n_3 \times n_3} \mathcal{G}_{k(3)}^T)} \\ &\leq \sqrt{n_3 R} \|\mathcal{G}_k\|_F \leq \sqrt{n_3 R} \sqrt{n_1 n_2 n_3} \|\mathcal{G}_k\|_1, \end{aligned} \quad (9)$$

where the second inequality holds, because $|\phi(x)| \leq |x|$. On the other hand,

$$\begin{aligned} \|\mathcal{G}_k\|_{\star, \varphi_k} &= \|bdiag(\varphi_k(\mathcal{G}_k))\|_{\star} \geq \|bdiag(\varphi_k(\mathcal{G}_k))\|_F = \|bdiag(\phi(\mathcal{G}_k \times_3 \mathbf{T}_k))\|_F \\ &\geq C \|bdiag(\mathcal{G}_k \times_3 \mathbf{T}_k)\|_F = C \|\mathcal{G}_k\|_F \\ &\geq \frac{C}{n_1 n_2 n_3} \|\mathcal{G}_k\|_1, \end{aligned} \quad (10)$$

where the second inequality holds because $\mathcal{G}_k$ and $\mathbf{T}_k$ are bounded. Thus, elements of $\mathcal{G}_k \times_3 \mathbf{T}_k$ are also bounded. We find a constant C which satisfies $\varphi(\mathcal{G}_k)(i, j, k) \geq C(\mathcal{G}_k \times_3 \mathbf{T}_k)(i, j, k)$ for all $(i, j, k) \in N_+^{n_1 \times n_2 \times n_3}$. By combining (9) and (10), we obtain

$$\|\mathcal{G}_k\|_1 \lesssim \|\mathcal{G}_k\|_{\star, \varphi_k} \lesssim \sqrt{R} \|\mathcal{G}_k\|_1, \text{ and } \|\mathcal{G}_k\|_F \lesssim \|\mathcal{G}_k\|_{\star, \varphi_k} \lesssim \sqrt{R} \|\mathcal{G}_k\|_F.$$

The proof is completed by combining the definitions of the TV norm and NTCTV. $\square$

As mentioned previously, the proposed NTCTV can effectively encode both R and S priors. This enables the adaptive exploitation of complementary information from the two priors within a single regularization term, fully harnessing their synergistic effects to enhance the performance in applications.

3.2. Proposed model

Suppose $\mathcal{Y} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is the underlying unknown tensor characterized by the joint R and S -priors. Let $\mathcal{X}$ denote the target tensor. To enhance brevity, we reformulate the NTCTV-based TC model as follows. Please refer to (8) for detailed information.

$$\min_{\mathcal{X}} \|\mathcal{X}\|_{NTCTV}, \text{ s.t. } \mathcal{P}_{\Omega}(\mathcal{Y}) = \mathcal{P}_{\Omega}(\mathcal{X}), \mathbf{T}_k \mathbf{T}_k^T = \mathbf{I}_{r \times r}, \quad (11)$$

where Ω is the index of the observational elements, and $\mathcal{P}_{\Omega}(\cdot)$ is a projection operator which maps the elements in Ω to themselves and projects the remaining elements to zero.

3.3. Optimization algorithm

We employ a PAM-based method to solve (11). To this end, we introduce auxiliary variables $\mathcal{Z}_k$, $\mathcal{N}_k$, $\mathcal{K}$ and indicator functions

$$\delta_{\mathcal{K}, \Omega} = \begin{cases} 0, & \mathcal{P}_{\Omega}(\mathcal{K}) = \mathcal{O} \\ +\infty, & \text{otherwise} \end{cases}, \Psi(\mathbf{T}) = \begin{cases} 0, & \mathbf{T} \mathbf{T}^T = \mathbf{I}_{r \times r} \\ +\infty, & \text{otherwise} \end{cases}. \quad (12)$$

Then, the problem (11) is decoupled using the auxiliary variables and (12), resulting in the following constrained optimization problem:

$$\min_{\mathcal{X}} \sum_{k \in \Gamma} \frac{1}{\gamma} \left(\sum_{i=1}^r \|\mathcal{N}_k^{(i)}\|_{\star} \right) + \Psi(\mathbf{T}_k) + \delta_{\mathcal{K}, \Omega}, \quad (13)$$

$$\text{s.t. } \mathcal{P}_{\Omega}(\mathcal{X}_0) = \mathcal{X} + \mathcal{K}, \nabla_k(\mathcal{X}) = \mathcal{G}_k, \mathcal{G}_k = \mathcal{Z}_k \times_3 \mathbf{T}_k^T, \mathcal{N}_k = \phi(\mathcal{Z}_k).$$

By imposing penalties on the constraint terms in (13), we obtain the following unconstrained problem:

$$\begin{aligned} L(\mathcal{X}, \mathcal{G}_k, \mathcal{N}_k, \mathcal{Z}_k, \mathbf{T}_k, \mathcal{K}) \\ = \sum_{k \in \Gamma} \left(\frac{1}{\gamma} \sum_{i=1}^r \|\mathcal{N}_k^{(i)}\|_{\star} + \frac{\alpha_1}{2} \|\nabla_k(\mathcal{X}) - \mathcal{G}_k\|_F^2 + \frac{\alpha_2}{2} \|\mathcal{G}_k - \mathcal{Z}_k \times_3 \mathbf{T}_k^T\|_F^2 \right. \\ \left. + \frac{\alpha_3}{2} \|\mathcal{Y}_k - \phi(\mathcal{Z}_k)\|_F^2 + \Psi(\mathbf{T}_k) \right) + \frac{\beta}{2} \|\mathcal{P}_{\Omega}(\mathcal{X}_0) - \mathcal{X} - \mathcal{K}\|_F^2 + \delta_{\mathcal{K}, \Omega}, \end{aligned} \quad (14)$$

where α_i ($i = 1, 2, 3$) and β denote the penalty parameters. Utilizing the update framework of the PAM, $\mathcal{X}$, $\mathcal{G}_k$, $\mathcal{N}_k$, $\mathcal{Z}_k$, $\mathbf{T}_k$ and $\mathcal{K}$ are alternately updated as follows:

1) The $\mathcal{X}$ subproblem is

$$\begin{aligned} \mathcal{X}^{p+1} = \arg \min_{\mathcal{X}} \sum_{k \in \Gamma} \frac{\alpha_1}{2} \|\nabla_k(\mathcal{X}) - \mathcal{G}_k^p\|_F^2 \\ + \frac{\beta}{2} \|\mathcal{P}_{\Omega}(\mathcal{X}_0) - \mathcal{X} - \mathcal{K}^p\|_F^2 + \frac{\rho}{2} \|\mathcal{X} - \mathcal{X}^p\|_F^2. \end{aligned} \quad (15)$$

Differentiating (15) with respect to $\mathcal{X}$, we obtain the following linear system:

$$(\beta + \rho)I + \sum_{k \in \Gamma} \alpha_{1,k} \nabla_k^T \nabla_k)(\mathcal{X}) = \beta(\mathcal{P}_\Omega(\mathcal{X}_0) - \mathcal{K}^p) + \rho\mathcal{X}^p + \sum_{k \in \Gamma} \alpha_{1,k} \nabla_k^T(\mathcal{G}_k^p), \quad (16)$$

where $\nabla_k^T(\cdot)$ denotes the transpose operator of $\nabla_k(\cdot)$ and I is the identity tensor. Since the difference operation on tensor is a linear operator within the t-product framework [42], the difference tensor $\mathcal{D}_k$ corresponding to $\nabla_k(\cdot)$ can be represented diagonally by applying FFT. Thus, based on the convolution theorem of Fourier transform, we obtain the closed-form solution of (16) as follows:

$$\mathcal{X} = \mathcal{F}^{-1} \left(\frac{\mathcal{F}(\beta(\mathcal{P}_\Omega(\mathcal{X}_0) - \mathcal{K}^p) + \rho\mathcal{X}^p) + \mathcal{H}}{(\beta + \rho)I + \sum_{k \in \Gamma} \alpha_{1,k} \mathcal{F}(\mathcal{D}_k)^T \odot \mathcal{F}(\mathcal{D}_k)} \right), \quad (17)$$

where $\mathcal{H} = \sum_{k \in \Gamma} \mathcal{F}(\mathcal{D}_k)^T \odot \mathcal{F}(\mathcal{G}_k^p)$, I is a tensor with each element being 1, $\odot$ denotes componentwise multiplication, and the division is also performed componentwise.

2) The $\mathcal{G}_k$ ($k \in \Gamma$) subproblems are

$$\mathcal{G}_k^{p+1} = \arg \min_{\mathcal{G}_k} \frac{\alpha_1}{2} \|\nabla_k(\mathcal{X}^{p+1}) - \mathcal{G}_k\|_F^2 + \frac{\alpha_2}{2} \|\mathcal{G}_k - \mathcal{Z}_k^p \times_3 (\mathbf{T}_k^p)^T\|_F^2 + \frac{\rho}{2} \|\mathcal{G}_k - \mathcal{G}_k^p\|_F^2. \quad (18)$$

For the above least squares problem (18), we can obtain a closed-form solution as follows:

$$\mathcal{G}_k^{p+1} = \frac{\alpha_1 \nabla_k(\mathcal{X}^{p+1}) + \alpha_2 \mathcal{Z}_k^p \times_3 (\mathbf{T}_k^p)^T + \rho \mathcal{G}_k^p}{\alpha_1 + \alpha_2 + \rho}. \quad (19)$$

3) The $\mathcal{N}_k$ ($k \in \Gamma$) subproblems can be reformulated as follows:

$$\mathcal{N}_k^{p+1} = \arg \min_{\mathcal{N}_k} \frac{1}{\gamma} \sum_{i=1}^r \|\mathcal{N}_k^{(i)}\|_* + \frac{\alpha_3 + \rho}{2} \|\mathcal{N}_k - \mathcal{W}_k\|_F^2, \quad (20)$$

where $\mathcal{W}_k = \frac{\alpha_3 \phi(\mathcal{Z}_k^p) + \rho \mathcal{N}_k^p}{\alpha_3 + \rho}$. By employing the singular value thresholding (SVT) operator [43], for each front slice, we can obtain the closed-form solution of (20) as follows:

$$\mathcal{N}_k^{p+1(i)} = \text{SVT}_{1/\gamma(\alpha_3 + \rho)}(\mathcal{W}_k^{(i)}). \quad (21)$$

4) The $\mathcal{Z}_k$ ($k \in \Gamma$) subproblems are

$$\mathcal{Z}_k^{p+1} = \arg \min_{\mathcal{Z}_k} \frac{\alpha_2}{2} \|\mathcal{G}_k^{p+1} - \mathcal{Z}_k \times_3 (\mathbf{T}_k^p)^T\|_F^2 + \frac{\alpha_3}{2} \|\mathcal{N}_k^{p+1} - \phi(\mathcal{Z}_k)\|_F^2 + \frac{\rho}{2} \|\mathcal{Z}_k - \mathcal{Z}_k^p\|_F^2. \quad (22)$$

Furthermore, the subproblems (22) can be equivalently formulated as follows:

$$\mathcal{Z}_{k,(3)}^{p+1} = \arg \min_{\mathcal{Z}_{k,(3)}} \frac{\alpha_2 + \rho}{2} \|\mathcal{Z}_{k,(3)} - \mathbf{W}\|_F^2 + \frac{\alpha_3}{2} \|\phi(\mathcal{Z}_{k,(3)}) - \mathcal{N}_{k,(3)}^{p+1}\|_F^2, \quad (23)$$

where $\mathbf{W} = \frac{\alpha_2 \mathbf{T}_k^p \mathcal{G}_{k,(3)}^{p+1} + \rho \mathcal{Z}_{k,(3)}^p}{\alpha_2 + \rho}$. Then, the problem (23) can be solved using Newton method.

5) The $\mathbf{T}_k$ ($k \in \Gamma$) subproblems can be reformulated as follows:

$$\begin{aligned} \mathbf{T}_k^{p+1} &= \arg \min_{\mathbf{T}_k} \frac{\alpha_2}{2} \|\mathcal{G}_k^{p+1} - \mathcal{Z}_k^{p+1} \times_3 \mathbf{T}_k^T\|_F^2 + \frac{\rho}{2} \|\mathbf{T}_k - \mathbf{T}_k^p\|_F^2 + \Psi(\mathbf{T}_k) \\ &= \arg \min_{\mathbf{T}_k} \frac{\alpha_2}{2} \|\mathcal{G}_{k,(3)}^{p+1} - \mathbf{T}_k^T \mathcal{Z}_{k,(3)}^{p+1}\|_F^2 + \frac{\rho}{2} \|\mathbf{T}_k - \mathbf{T}_k^p\|_F^2 + \Psi(\mathbf{T}_k) \\ &= \arg \min_{\mathbf{T}_k} \frac{\alpha_2}{2} \text{Tr}[(\mathcal{G}_{k,(3)}^{p+1} - \mathbf{T}_k^T \mathcal{Z}_{k,(3)}^{p+1})^T (\mathcal{G}_{k,(3)}^{p+1} - \mathbf{T}_k^T \mathcal{Z}_{k,(3)}^{p+1})] \\ &\quad + \text{Tr}[(\mathbf{T}_k - \mathbf{T}_k^p)^T (\mathbf{T}_k - \mathbf{T}_k^p)] + \Psi(\mathbf{T}_k) \\ &= \arg \max_{\mathbf{T}_k} \text{Tr}[(\alpha_2 \mathcal{G}_{k,(3)}^{p+1} (\mathcal{Z}_{k,(3)}^{p+1})^T + \rho (\mathbf{T}_k^p)^T \mathbf{T}_k) - \Psi(\mathbf{T}_k)], \end{aligned} \quad (24)$$

where $\text{Tr}(\cdot)$ denotes the matrix trace. Let the SVD of $[\alpha_2 \mathcal{G}_{k,(3)}^{p+1} (\mathcal{Z}_{k,(3)}^{p+1})^T + \rho (\mathbf{T}_k^p)^T]$ is $\tilde{\mathbf{U}}_k \tilde{\mathbf{S}}_k \tilde{\mathbf{V}}_k^T$. Thus, we have $\text{Tr}(\tilde{\mathbf{U}}_k \tilde{\mathbf{S}}_k \tilde{\mathbf{V}}_k^T \mathbf{T}_k) = \text{Tr}(\tilde{\mathbf{S}}_k \tilde{\mathbf{U}}_k \tilde{\mathbf{V}}_k^T \mathbf{T}_k)$. Since $\tilde{\mathbf{S}}_k$ is a diagonal matrix with positive elements, maximizing the objective in (24) involves ensuring that the diagonal elements of $\tilde{\mathbf{U}}_k \tilde{\mathbf{V}}_k^T \mathbf{T}_k$ are positive and maximum, with the constraint that $\mathbf{T}_k$ is an orthogonal matrix. By applying the Cauchy-Schwartz inequality, we obtain $\mathbf{T}_k = \tilde{\mathbf{V}}_k \tilde{\mathbf{U}}_k^T$. Thus, the closed-form solution of (24) is

$$\mathbf{T}_k^{p+1} = \tilde{\mathbf{V}}_k \tilde{\mathbf{U}}_k^T, \quad (25)$$

Algorithm 1 PAM-based solver for NTCTV model (11)

Input: observed tensor $\mathcal{P}_\Omega(\mathcal{Y})$; priori set Γ ; proximal operator ρ ; penalty parameters β, α_i ($i = 1, 2, 3$); the row number r of $\mathbf{T}_k$.

1: **Initialization:** $\mathcal{X}^0, \mathcal{G}_k^0, \mathcal{N}_k^0, \mathcal{Z}_k^0, \mathbf{T}_k^0, \mathcal{K}^0$.

2: **while** not convergent **do**

3: Update $\mathcal{X}^{p+1}$ via (17);

4: Update $\mathcal{G}_k^{p+1}$ via (19), $k \in \Gamma$;

5: Update $\mathcal{N}_k^{p+1}$ via (21), $k \in \Gamma$;

6: Update $\mathcal{Z}_k^{p+1}$ via (23), $k \in \Gamma$;

7: Update $\mathbf{T}_k^{p+1}$ via (25), $k \in \Gamma$;

8: Update $\mathcal{K}^{p+1}$ via (26);

9: Check the convergence condition: $\|\mathcal{X}^{p+1} - \mathcal{X}^p\|_F / \|\mathcal{X}^p\|_F \leq 10^{-6}$;

10: **end while**

Output: completed tensor $\hat{\mathcal{X}} = \mathcal{X}^{p+1}$.

6) The $\mathcal{K}$ subproblem is the least squares problem with $\mathcal{P}_\Omega(\mathcal{K}) = \mathcal{O}$. The closed-form solution can be directly obtained as follows:

$$\mathcal{K}^{p+1} = \frac{\beta(\mathcal{P}_\Omega(\mathcal{X}_0) - \mathcal{X}^{p+1}) + \rho \mathcal{K}^p}{\beta + \rho}, \quad \mathcal{P}_\Omega(\mathcal{K}^{p+1}) = \mathcal{O}. \quad (26)$$

3.4. Computational complexity analysis

For Algorithm 1, we consider the input tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and the learnable semi-orthogonal matrix $\mathbf{T}_k \in \mathbb{R}^{r \times n_3}$ ($k \in \Gamma$). As shown in Algorithm 1, the computational complexity can be divided into six parts in each iteration process, that is, step 3 ~ 8. First, updating $\mathcal{X}$ in step 3 primarily involves FFT, which costs $O(n_1 n_2 n_3 \log(n_1 n_2 n_3))$. Second, updating $\mathcal{G}_k$ and $\mathcal{K}$ in steps 4 and step 8 is related to matrix multiplication, which costs $O(r n_1 n_2 n_3)$ and $O(n_1 n_2 n_3)$, respectively. Third, updating $\mathcal{N}_k$ in step 5 cost $O(r n_{(1)} n_{(2)}^2)$, which corresponds to the matrix SVT for each front slice. Fourth, updating $\mathcal{Z}_k$ in step 6 costs $O(n_1 n_2 r)$, primarily because of element-wise computations. Fifth, updating $\mathbf{T}_k$ in step 7 involve matrix multiplication and matrix SVD decomposition, which cost $O(r n_1 n_2 n_3)$ and $O(r n_3^2)$, respectively. The per-iteration computational complexity of Algorithm 1 is determined by the sum of the computational complexities of the individual steps, which can be expressed as $O(n_1 n_2 n_3 \log(n_1 n_2 n_3) + r n_1 n_2 n_3 + r n_{(1)} n_{(2)}^2 + r n_3^2)$.

3.5. Convergence analysis

This subsection presents the theoretical convergence of Algorithm 1 within the PAM algorithm framework.

Theorem 1 (Convergence). Assuming that the $\phi(\cdot)$ is a real analytic function and is continuous on its domain with Lipschitz continuous on any bounded set, the sequence $\{\mathcal{X}^p, \mathcal{G}_k^p, \mathcal{N}_k^p, \mathcal{Z}_k^p, \mathbf{T}_k^p, \mathcal{K}^p\}$ generated by Algorithm 1 converges to a critical point of the optimization problem in (11).

The proof of Theorem 1 can be found in the supplementary material, providing theoretical assurance of the effectiveness of the iterative sequence generated by the designed PAM-based algorithm.

4. Numerical experiments

In this section, we employ the NTCTV-TC method for completing various types of tensor data. These include HSIs, MSIs, videos, MRI, and CT medical images. Two missing scenarios are considered is designing the experiment: random voxel missing and structured missing. Our method is implemented on MATLAB (2021b) with Intel(R) Core(TM) i7-9700 3.00 GHz CPU and 64 GB memory.

Table 1

Quantitative comparison of all methods on HSIs completion results. The best and the sub-optimal values are highlighted in **boldface** and underline respectively. (/s: second).

Data	SR	0.01			0.02			0.05			Average Time/s
	Metric	PSNR	SSIM	FSIM	PSNR	SSIM	FSIM	PSNR	SSIM	FSIM	
HSI Pavia ($200 \times 200 \times 80$)	Observed	13.378	0.008	0.410	13.423	0.012	0.427	13.557	0.025	0.469	–
	TNN	17.727	0.217	0.624	21.744	0.451	0.721	27.693	0.810	0.887	0.324
	SPC+TV	19.737	0.274	0.582	21.251	0.399	0.658	26.573	0.773	0.862	0.750
	TNN+TV	18.416	0.272	0.621	22.347	0.510	0.736	27.501	0.806	0.884	0.570
	fMDT-Tucker	22.473	0.510	0.729	24.469	0.637	0.792	27.075	<u>0.791</u>	0.871	0.190
	MMES	20.051	0.251	0.391	25.901	0.681	0.788	28.702	0.838	0.888	3.715
	NTTNN	24.866	0.673	<u>0.826</u>	27.949	0.831	<u>0.904</u>	<u>34.049</u>	<u>0.952</u>	<u>0.969</u>	0.112
	TCTV	<u>25.735</u>	<u>0.713</u>	0.820	<u>28.287</u>	<u>0.831</u>	0.887	32.567	0.927	0.952	1.112
	NTCTV	26.578	0.755	0.854	30.852	0.903	0.938	38.127	0.980	0.988	0.347
HSI DC ($256 \times 256 \times 191$)	Observed	12.903	0.010	0.383	12.947	0.015	0.411	13.082	0.029	0.465	–
	TNN	18.183	0.322	0.670	22.784	0.597	0.800	29.649	0.882	0.936	1.648
	SPC+TV	19.473	0.341	0.602	22.000	0.539	0.735	27.911	0.845	0.909	2.892
	TNN+TV	18.471	0.354	0.656	22.891	0.603	0.791	28.901	0.867	0.927	2.712
	fMDT-Tucker	21.509	0.433	0.715	22.981	0.551	0.772	24.663	0.693	0.839	0.859
	MMES	20.182	0.292	0.474	24.717	0.635	0.790	29.512	0.877	0.932	16.143
	NTTNN	<u>26.663</u>	<u>0.805</u>	<u>0.893</u>	<u>31.089</u>	<u>0.919</u>	<u>0.953</u>	<u>38.861</u>	<u>0.983</u>	<u>0.990</u>	0.287
	TCTV	25.619	0.736	0.851	28.478	0.857	0.918	33.642	0.950	0.971	5.664
	NTCTV	29.416	0.885	0.933	34.530	0.962	0.976	40.719	0.989	0.992	1.378
HSI Cuprite ($350 \times 350 \times 188$)	Observed	11.045	0.007	0.547	11.089	0.011	0.528	11.224	0.020	0.525	–
	TNN	19.795	0.432	0.754	28.259	0.708	0.839	35.248	0.899	0.946	2.720
	SPC+TV	24.845	0.619	0.710	28.103	0.712	0.793	33.399	0.869	0.912	4.919
	TNN+TV	20.104	0.504	0.752	28.931	0.743	0.844	34.959	0.897	0.942	4.449
	fMDT-Tucker	25.227	0.622	0.779	28.305	0.708	0.829	32.141	0.821	0.897	1.432
	MMES	21.566	0.584	0.599	28.125	0.669	0.680	33.611	0.831	0.884	32.675
	NTTNN	<u>33.054</u>	<u>0.865</u>	<u>0.928</u>	<u>36.765</u>	<u>0.932</u>	<u>0.961</u>	<u>42.125</u>	<u>0.971</u>	<u>0.984</u>	0.516
	TCTV	32.694	0.831	0.891	35.406	0.897	0.940	39.961	0.952	0.976	9.203
	NTCTV	36.391	0.919	0.954	39.691	0.955	0.976	43.363	0.974	0.987	2.331

Comparison algorithm. Since our proposed method is based on the fused $R \circ S$ -prior, we make comparisons with state-of-the-art tensor completion methods that consider R , $R + S$, or $R \circ S$ priors, as well as methods capable of handling structured missing data. These methods include TNN [44], SPC+TV [28], TNN+TV [27], fMDT-Tucker [45], MMES [46], NTTNN [23] and TCTV [36]. Unless specified explicitly, the parameters of the comparison algorithm are selected according to reference recommendations or are manually fine-tuned for optimal performance.

Evaluation index. The recovery performance is assessed through numerical comparison using three metrics: the peak signal-to-noise ratio (PSNR) [47], structural similarity (SSIM) [47], and feature similarity (FSIM) [48]. All three metrics tend to perform better with a larger value. In addition, we assess the computational complexity of the Algorithm 1 by measuring the average time of a single iteration. A lower average time implies lower computational complexity. In all experiments, we define the relative error $\|\mathcal{X}^{p+1} - \mathcal{X}^p\|_F / \|\mathcal{X}^p\|_F \leq 10^{-6}$ as the stopping criterion.

4.1. Hyperspectral image completion

In this subsection, we conduct simulation experiments using three commonly used HSIs datasets (source: <https://rslab.ut.ac.ir/data>). Sub-images of the following sizes are extracted from each dataset: Pavia City ($200 \times 200 \times 80$), DC Mall ($256 \times 256 \times 191$), Cuprite ($350 \times 350 \times 188$). The simulation data used in all tests are randomly sampled from voxel of 3D tensor at sampling rates (SRs) of 1%, 2%, and 5%.

Parameter configuration. As shown in Algorithm 1, the proposed NTCTV method requires tuning several parameters, including penalty parameters represented by α_i ($i = 1, 2, 3$) and β , the proximal parameter ρ , and the row number r of $\mathbf{T}_k$ ($k \in I$), for an observed tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$. The number of rows r in the semi-orthogonal matrix $\mathbf{T}_k$ embedded in the NTCTV regularization term determines the number of frontal slices of the essential tensor obtained by shrinking the gradient tensor. As demonstrated by a computational complexity analysis of the algorithm described in Section 3.4, an appropriately small r can effectively reduce computational costs by utilizing most of the inherent

tensor information to achieve reasonable resource utilization. Additionally, we set the penalty parameters α_i ($i = 1, 2, 3$) to the same value α in all experiments, because all α parameters serve the same purpose of eliminating the constraints introduced by the auxiliary variables. Given that HSIs typically exhibit highly low-rank redundancy, we select the appropriate values of r from the set $\{3, 4, \dots, 10\}$. Furthermore, we set the hyperparameters ρ , α and β to 0.001, 30, and 150, respectively in all experiments in this subsection (see Section 5 for the detailed reasons).

Performance analysis. For a comparison of quality metrics, Table 1 lists PSNR, SSIM, FSIM and average running time of a single iteration of the completed HSIs with different methods, where the highest values of PSNR, SSIM and FSIM are highlighted in boldface and the sub-optimal values are underlined. The results show that the NTCTV method consistently significantly outperforms other methods in terms of evaluation metrics at all SRs and has certain advantages in terms of the average time of a single iteration. Specifically, even at an extremely low SR of 1%, the NTCTV approach achieves 0.8 ~ 3.3-dB improvement in PSNR compared with the next top-performing technique. HSIs typically have strong global low-rankness and local smoothness. This excellent performance can be attributed to the NTCTV method being able to effectively measure the fused $R \circ S$ -prior within HSIs and fully leverage the redundancy it provides to achieve image completion.

To compare the visual quality, Fig. 2 presents the recovered visual results of Pavia city with SR = 5%, DC Mall with SR = 2% and Cuprite with SR = 1%. To facilitate a closer examination of restoration quality, two identical regions from each sub-image are enlarged and displayed in blue and green boxes, respectively. As evident in Fig. 2, the NTCTV approach demonstrates clear improvement in reconstructing textural details and overall fidelity compared to alternative methods. Specifically, the zoomed patches highlight NTCTV's capabilities in retain fine-scale image features and spectral characteristics. The results validate the advantages of NTCTV's joint modeling framework and combined regularization priors for enhancing the quality and visual perception of the completed hyperspectral data.

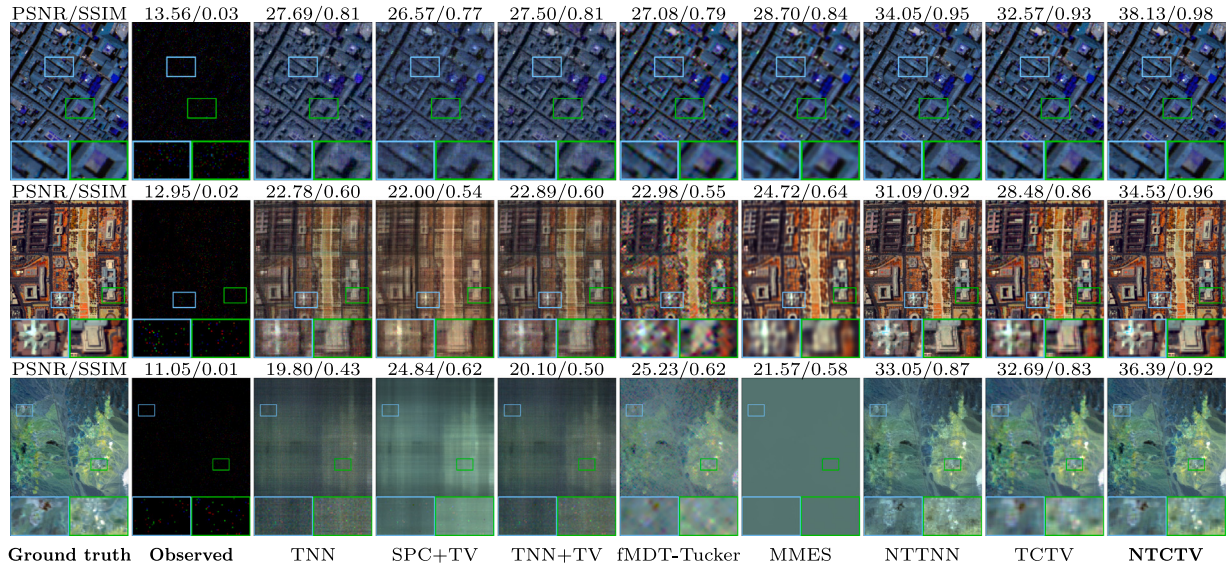


Fig. 2. Pseudo-color mapping of HSIs restoration results obtained from various methods. Top: Pavia (R: 10, G: 30, B: 70) with SR = 5%. Middle: DCmall (R: 60, G: 120, B: 180) with SR = 2%. Bottom: Cuprite (R: 30, G: 90, B: 150) with SR = 1%.

Table 2

Quantitative comparison of all methods on medical images completion results. The best and the sub-optimal values are highlighted in **boldface** and underline respectively. (/s: second).

Data	SR	0.1			0.15			0.2			Average Time/s
	Metric	PSNR	SSIM	FSIM	PSNR	SSIM	FSIM	PSNR	SSIM	FSIM	
MRI Brain (181 × 181 × 103)	Observed	9.943	0.192	0.488	10.191	0.206	0.519	10.453	0.222	0.543	–
	TNN	25.508	0.665	0.833	27.496	0.752	0.872	29.208	0.813	0.901	0.362
	SPC+TV	22.789	0.543	0.777	24.500	0.630	0.814	25.763	0.687	0.838	0.815
	TNN+TV	27.035	0.816	0.873	29.258	0.878	0.912	31.017	0.913	0.935	0.726
	fMDT-Tucker	29.206	0.901	0.907	31.315	0.935	0.934	32.934	0.952	0.950	0.287
	MMES	28.004	0.808	0.890	30.252	0.855	0.921	31.699	0.877	0.938	4.849
	NTTNN	26.558	0.734	0.856	28.787	0.817	0.897	30.635	0.870	0.924	0.338
	TCTV	<u>30.217</u>	0.881	<u>0.913</u>	<u>32.359</u>	0.918	<u>0.939</u>	<u>34.044</u>	0.940	<u>0.955</u>	1.355
	NTCTV	31.388	0.906	0.929	33.485	0.938	0.951	35.097	0.955	0.965	1.416
CT Lumbar (181 × 181 × 103)	Observed	10.129	0.442	0.544	10.378	0.454	0.571	10.640	0.467	0.591	–
	TNN	23.255	0.676	0.780	24.937	0.745	0.823	26.323	0.793	0.854	0.371
	SPC+TV	19.881	0.441	0.696	21.239	0.504	0.729	21.975	0.525	0.747	0.799
	TNN+TV	24.564	0.812	0.844	26.563	0.871	0.885	28.067	0.904	0.911	0.697
	fMDT-Tucker	25.228	<u>0.851</u>	0.850	26.493	0.885	0.880	27.552	0.907	0.901	0.282
	MMES	24.570	0.575	0.851	25.685	0.576	0.875	26.842	0.607	0.899	4.610
	NTTNN	24.728	0.733	0.830	26.627	0.823	0.871	28.206	0.870	0.900	0.336
	TCTV	26.906	0.850	0.872	28.706	0.889	0.902	30.162	0.914	0.922	1.362
	NTCTV	27.679	0.875	0.889	29.398	0.908	0.916	30.857	0.931	0.935	1.392
CT Sacral (181 × 181 × 103)	Observed	11.752	0.553	0.637	11.998	0.564	0.670	12.260	0.575	0.694	–
	TNN	26.182	0.839	0.878	27.828	0.881	0.906	29.193	0.907	0.925	0.359
	SPC+TV	23.739	0.601	0.809	25.254	0.623	0.825	26.417	0.629	0.836	0.787
	TNN+TV	27.048	0.888	0.908	28.871	0.924	0.934	30.328	0.944	0.949	0.691
	fMDT-Tucker	27.041	0.874	0.882	28.308	0.903	0.904	29.345	0.921	0.920	0.288
	MMES	24.791	0.453	0.881	25.962	0.488	0.915	28.139	0.572	0.939	4.739
	NTTNN	27.317	0.858	0.907	29.238	0.916	0.933	30.743	0.940	0.949	0.336
	TCTV	<u>28.932</u>	<u>0.909</u>	<u>0.921</u>	<u>30.649</u>	<u>0.937</u>	<u>0.942</u>	<u>32.035</u>	<u>0.952</u>	<u>0.955</u>	1.321
	NTCTV	29.704	0.931	0.938	31.372	0.951	0.954	32.733	0.963	0.965	1.371

4.2. Medical imaging completion

In this subsection, we further construct experiments targeting CT and MRI medical images (source: <https://www.cancerimagingarchive.net/>), including MRI Brain, CT Lumbar, and CT Sacral. All medical images are resized to $181 \times 181 \times 103$ and tested using random voxel sampling at SRs of 10%, 15% and 20%.

Parameter configuration. Similar to the parameter settings for HSIs completion, we select the proximal parameter ρ and the penalty parameters α and β as 0.001, 30, and 150, respectively. Given that the testing medical images exhibit weaker global low-rankness coupling in

smoothness compared to HSIs, a larger r is necessary to retain the required information content for accurate recovery. Therefore, to balance the improved reconstruction performance with lower computational costs, we select suitable r values from the candidate set $\{40, 50, 60\}$ after experimenting with the medical image data (see Section 5.2 for detailed discussion). The choice of r affects both the accuracy and efficiency of the overall approach.

Performance analysis. Table 2 presents the PSNR, SSIM, FSIM and time complexity metrics for the compared methods on the three medical images across different SRs. The proposed NTCTV approach achieves optimal performance in terms of the evaluation metrics, as indicated by the bolded values. Specifically, NTCTV demonstrates PSNR gains

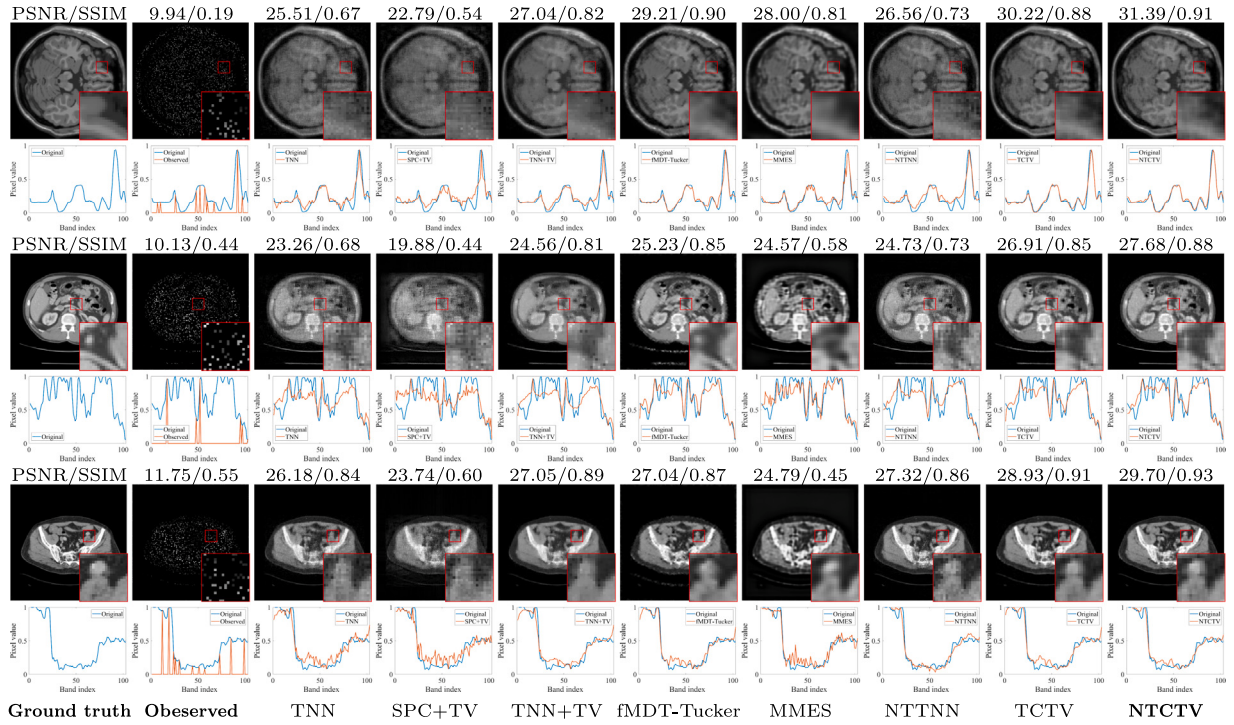


Fig. 3. Grayscale representation and spectral profile curves of medical images restoration results with $SR = 10\%$. Top: Brain. Middle: Lumbar. Bottom: Sacral. The displayed images depict the 30th band.

of $0.7 \sim 1.1$ -dB compared to the next-highest method, underlining its effectiveness for medical image completion. It should be noted that the sub-optimal methods in Tables 1 and 2 differ, and the time complexity is relatively increased because of the weaker low-rankness of the medical images versus HSIs. However, the proposed NTCTV approach still achieves superior performance by fully utilizing the synergistic effects between the global low-rankness and local smoothing priors through nonlinear modeling. The learnable transforms in the fused regularization term enable more effective exploitation of the correlations between the priors. The results validate the benefits of capturing nonlinear relationships within the fused $R \circ S$ -prior.

Fig. 3 shows the visual comparison results and spectrum profiles of three medical images restored by all comparison methods under $SR = 10\%$. The same area in each subfigure is magnified and highlighted with a red box to facilitate a clearer comparison. Qualitatively, the proposed NTCTV approach demonstrates improved preservation of textural structural details compared with competing methods, as evidenced by the magnified images. Rather than exhibiting excessive smoothing, NTCTV retains the necessary information for high-fidelity medical image restoration. Furthermore, the spectral curves obtained using the NTCTV match the profile changes corresponding to the original data more closely.

4.3. Multispectral image completion

In this subsection, we implement all competing methods on three MSIs from the CAVE database (source: <https://www.cs.columbia.edu/CAVE/databases/multispectral/>): Beads, Face and Toys. Each tensor in the dataset is resized to dimensions of $256 \times 256 \times 31$ and tested using random voxel sampling at SRs of 5%, 10% and 15%.

Parameter configuration. Given the similar data structure of MSIs and HSIs, we utilize the same penalty parameters and proximal operator settings as in Section 4.1 for all experiments in this subsection. However, MSIs exhibit lower spectral resolution than HSIs, resulting in less redundant fused prior information. Based on this difference, we select

suitable values for r from the larger candidate set $\{10, 15, 20\}$ to balance the reconstruction accuracy and computational efficiency.

Performance analysis. In terms of the results of the visual metrics, Table 3 lists the detailed MSIs recovery results under varying SRs based on the PSNR, SSIM, FSIM and average single iteration time (in seconds) for all the comparison methods. The optimal values are highlighted by boldface and the second-highest values are underlined. Notably, the visual evaluation metrics of NTCTV demonstrate significant improvements across the different SRs. Specifically, even in the extreme case with a SR of 5%, NTCTV achieved considerable PSNR gains of approximately $1.7 \sim 4.7$ -dB compared with the sub-optimal algorithms for the three multispectral images. This result can be attributed to the learnable nonlinear transform in the regularization term of the fused $R \circ S$ -prior, which can fully characterize the intrinsic connections between different priors. This enables complementary information to be leveraged synergistically to aid model optimization.

In terms of the visual effects of MSIs restoration, Fig. 4 displays the pseudo-color images and associated spectral profiles for all methods at $SR = 5\%$. To facilitate a more distinct visual comparison, the same region across each subfigure is magnified within a red bounding box. On the one hand, it is evident from the locally magnified portion of the subgraph that NTCTV is superior in restoring local texture details while concurrently preserving global high fidelity. On the other hand, the spectrum curve of the recovery result corresponding to the NTCTV method not only exhibits significantly better correspondence with the actual data, but also approximates the curvature variations of the spectral curve of the actual data. This aligns with the notable strength of NTCTV in addressing local texture details. Collectively, these localized and spectral observations demonstrate the superiority of NTCTV in restoring MSIs with enhanced local details and global accuracy.

4.4. Video completion

In this subsection, we conduct the experiments on three video data named News, Suzie, and Carphone from the YUV database (source: <http://trace.eas.asu.edu/yuv/>). All video data are of size 144×176

Table 3

Quantitative comparison of all methods on MSIs completion results. The best and the sub-optimal values are highlighted in **boldface** and underline respectively. (/s: second).

Data	SR	0.05			0.1			0.15			Average Time/s
	Metric	PSNR	SSIM	FSIM	PSNR	SSIM	FSIM	PSNR	SSIM	FSIM	
MSI Beads ($256 \times 256 \times 31$)	Observed	14.574	0.102	0.480	14.805	0.138	0.558	15.053	0.176	0.616	–
	TNN	19.764	0.424	0.718	23.568	0.668	0.824	26.125	0.781	0.880	0.229
	SPC+TV	18.865	0.371	0.673	22.827	0.627	0.803	26.192	0.776	0.878	0.505
	TNN+TV	20.723	0.554	0.748	24.782	0.789	0.871	27.537	0.877	0.920	0.380
	fMDT-Tucker	24.188	0.769	0.845	26.169	0.847	0.887	27.443	0.882	0.910	0.120
	MMES	<u>25.308</u>	<u>0.798</u>	0.863	27.512	0.868	0.906	28.851	0.902	0.928	1.108
	NTTNN	23.958	0.733	0.850	27.741	0.865	0.918	30.675	0.922	<u>0.950</u>	0.156
	TCTV	25.240	0.783	<u>0.865</u>	<u>28.508</u>	<u>0.881</u>	<u>0.920</u>	<u>30.984</u>	<u>0.925</u>	0.948	0.782
	NTCTV	26.996	0.856	0.903	31.866	0.944	0.959	35.445	0.972	0.979	0.926
MSI Face ($256 \times 256 \times 31$)	Observed	17.328	0.317	0.700	17.561	0.354	0.696	17.810	0.389	0.690	–
	TNN	33.301	0.891	0.916	38.564	0.955	0.961	41.643	0.976	0.978	0.219
	SPC+TV	33.111	0.899	0.924	37.904	0.951	0.962	40.223	0.968	0.974	0.456
	TNN+TV	34.695	0.943	0.947	40.296	0.976	0.975	43.345	0.986	0.986	0.381
	fMDT-Tucker	31.404	0.906	0.901	34.488	0.946	0.931	35.918	0.956	0.941	0.126
	MMES	35.922	0.944	0.944	37.787	0.960	0.958	40.245	0.970	0.971	1.134
	NTTNN	37.144	0.951	<u>0.963</u>	40.880	0.978	0.981	43.319	0.987	0.988	0.152
	TCTV	<u>38.511</u>	<u>0.965</u>	0.961	<u>42.826</u>	<u>0.983</u>	<u>0.981</u>	<u>45.885</u>	<u>0.991</u>	<u>0.990</u>	0.777
	NTCTV	43.292	0.989	0.988	47.439	0.995	0.995	50.021	0.997	0.997	0.885
MSI Toys ($256 \times 256 \times 31$)	Observed	11.028	0.248	0.612	11.262	0.284	0.625	11.511	0.320	0.637	–
	TNN	26.725	0.797	0.847	31.455	0.898	0.917	34.274	0.937	0.947	0.218
	SPC+TV	25.973	0.771	0.842	31.273	0.891	0.920	33.812	0.929	0.948	0.483
	TNN+TV	27.167	0.869	0.897	32.124	0.941	0.949	35.345	0.965	0.968	0.382
	fMDT-Tucker	25.374	0.857	0.843	26.921	0.897	0.871	27.986	0.916	0.888	0.120
	MMES	30.378	0.899	0.920	33.102	0.942	0.953	34.663	0.953	0.963	0.984
	NTTNN	<u>31.503</u>	0.912	<u>0.936</u>	<u>36.199</u>	0.958	<u>0.970</u>	<u>39.328</u>	0.975	<u>0.983</u>	0.152
	TCTV	30.408	<u>0.921</u>	0.918	34.905	<u>0.960</u>	0.957	38.163	<u>0.976</u>	0.975	0.778
	NTCTV	34.954	0.964	0.966	41.425	0.989	0.990	45.613	0.995	0.996	0.926

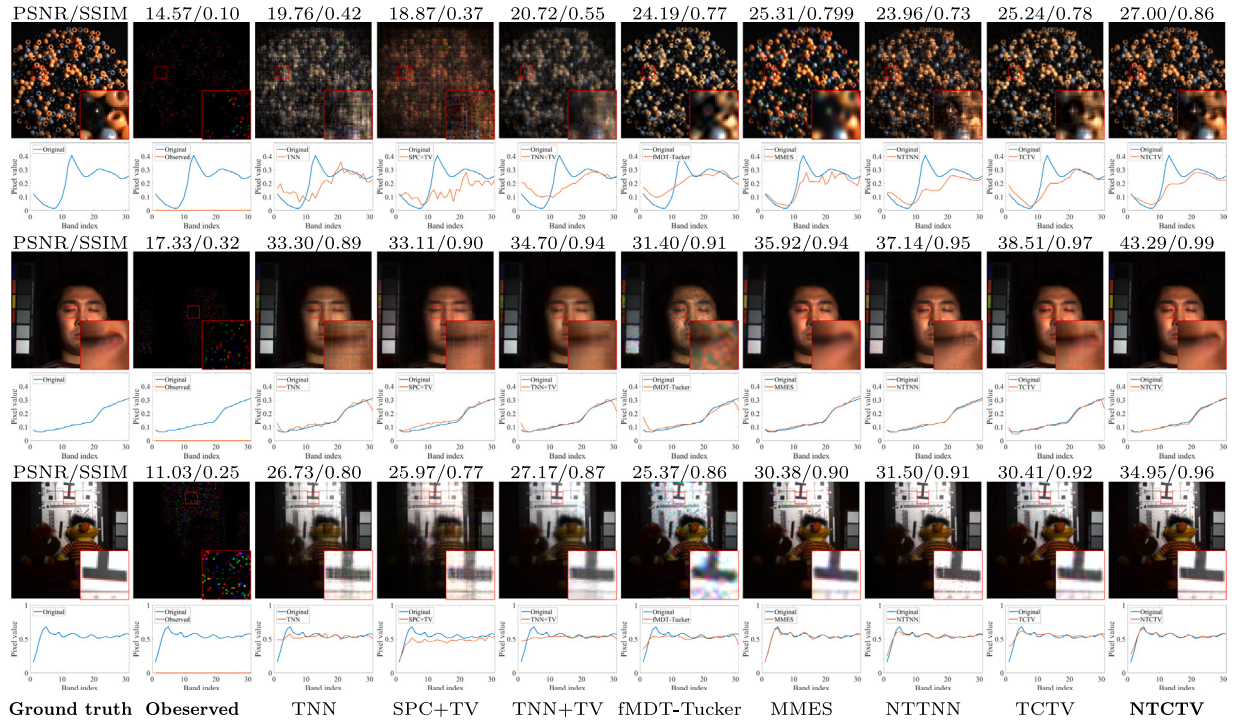


Fig. 4. Pseudo-color mapping and spectral profile curves of MSIs restoration results with SR = 5%. Top: Beads (R: 12, G: 11, B: 10). Middle: Face (R: 30, G: 20, B: 10). Bottom: Toys (R: 30, G: 20, B: 10).

$\times 150$ and tested using random voxel sampling at SRs of 5%, 10% and 15%.

Parameter configuration. Video data are tensors consisting of a sequence of image frames, with each frame containing rich spatial information and exhibiting temporal correlation with adjacent frames. To effectively capture the intrinsic information within the tensor data, we

select the most suitable value of r from the candidate set $\{20, 50, 80\}$. The other hyperparameters, namely, α , β and ρ , are set to the same values as those used in Section 4.1 to ensure consistency across experiments. This selection strategy helps maintain consistency across experiments and indirectly reflects the structural information of various types of data.

Table 4

Quantitative comparison of all methods on video completion results. The best and the sub-optimal values are highlighted in **boldface** and underline respectively. (/s: second).

Data	SR	0.05			0.1			0.15			Average Time/s
	Metric	PSNR	SSIM	FSIM	PSNR	SSIM	FSIM	PSNR	SSIM	FSIM	
Video News (144 × 176 × 150)	Observed	8.799	0.020	0.459	9.033	0.034	0.473	9.281	0.047	0.482	–
	TNN	25.889	0.789	0.891	28.562	0.865	0.930	30.034	0.895	0.946	0.391
	SPC+TV	23.517	0.715	0.845	26.493	0.809	0.897	28.281	0.853	0.920	0.926
	TNN+TV	25.584	0.792	0.888	28.457	0.867	0.929	30.055	0.898	0.946	0.683
	fMDT-Tucker	22.124	0.750	0.830	23.779	0.829	0.874	25.030	0.872	0.899	0.193
	MMES	24.211	0.823	0.883	28.249	0.904	0.933	30.279	0.935	0.954	3.501
	NTTNN	27.604	0.862	0.931	29.847	0.905	0.953	30.917	0.922	0.962	0.168
	TCTV	<u>29.209</u>	<u>0.895</u>	<u>0.937</u>	<u>31.661</u>	<u>0.931</u>	<u>0.959</u>	<u>33.163</u>	<u>0.947</u>	<u>0.969</u>	1.359
	NTCTV	29.892	0.919	0.951	32.735	0.949	0.970	34.024	0.959	0.977	0.954
Video Suzie (144 × 176 × 150)	Observed	7.259	0.009	0.454	7.494	0.014	0.427	7.743	0.017	0.407	–
	TNN	25.805	0.663	0.836	28.407	0.757	0.877	29.770	0.803	0.899	0.387
	SPC+TV	24.483	0.647	0.814	27.390	0.733	0.858	29.074	0.784	0.882	0.936
	TNN+TV	25.951	0.689	0.841	28.407	0.765	0.879	29.827	0.809	0.900	0.691
	fMDT-Tucker	28.370	0.824	0.883	30.794	0.878	0.917	32.276	0.906	0.936	0.208
	MMES	29.785	0.844	0.902	31.270	0.873	0.920	32.325	0.892	0.931	3.470
	NTTNN	27.720	0.739	0.871	30.055	0.819	0.906	31.315	0.854	0.923	0.164
	TCTV	<u>30.381</u>	<u>0.840</u>	<u>0.907</u>	<u>31.994</u>	<u>0.876</u>	<u>0.928</u>	<u>33.147</u>	<u>0.898</u>	<u>0.941</u>	1.396
	NTCTV	30.891	0.860	0.915	32.758	0.897	0.937	34.164	0.920	0.951	1.355
Video Carphone (144 × 176 × 150)	Observed	6.787	0.014	0.473	7.022	0.023	0.452	7.269	0.031	0.436	–
	TNN	25.800	0.737	0.863	28.377	0.816	0.903	29.720	0.852	0.921	0.391
	SPC+TV	24.144	0.699	0.841	27.219	0.792	0.886	28.881	0.835	0.908	0.924
	TNN+TV	25.866	0.756	0.869	28.335	0.823	0.904	29.757	0.857	0.923	0.690
	fMDT-Tucker	24.320	0.826	0.855	25.952	0.884	0.892	26.952	0.910	0.911	0.195
	MMES	27.227	0.853	0.896	30.005	0.907	0.932	31.522	0.929	0.947	3.465
	NTTNN	28.285	0.828	0.910	30.664	0.885	0.938	32.064	0.910	0.951	0.169
	TCTV	<u>29.691</u>	<u>0.879</u>	<u>0.923</u>	<u>31.513</u>	<u>0.910</u>	<u>0.943</u>	<u>32.712</u>	<u>0.927</u>	<u>0.954</u>	1.361
	NTCTV	30.456	0.904	0.935	32.484	0.933	0.955	33.769	0.948	0.965	0.561

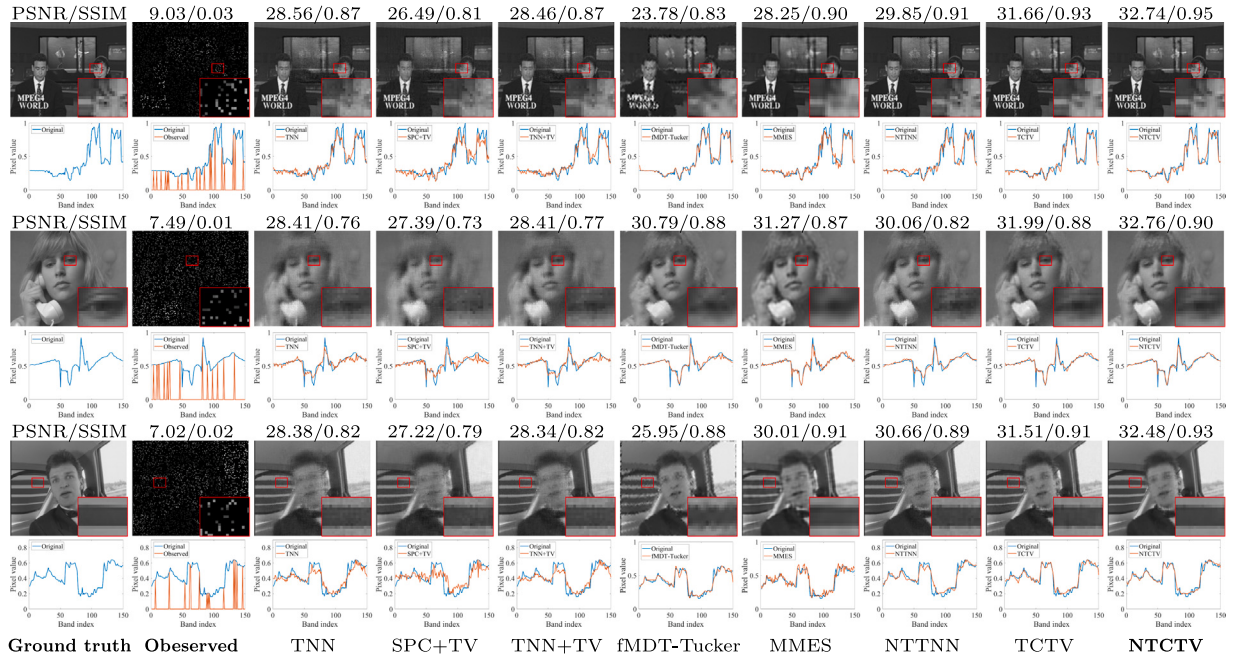


Fig. 5. Grayscale representation and spectral profile curves of videos restoration results with SR = 10%. Top: News. Middle: Suzie. Bottom: Carphone. The displayed images depict the 30th band.

Performance analysis. Table 4 lists the PSNR, SSIM and FSIM values of the recovered tensor data obtained from all comparison methods under different SRs. Besides, the average time (seconds) for each iteration is included. As depicted in the figure, the proposed method exhibits a certain improvement in the visual metrics compared to the other methods. Specifically, NTCTV achieved an improvement of approximately 0.7 ~ 1.1-dB in PSNR. Notably, video data exhibit decreased global low-rank and local smoothness properties versus other visual data types. However, NTCTV can still comprehensively perform image restoration

in a collaborative manner by utilizing the complementary characteristics between the different priors. This outcome effectively validates the superiority of the learnable nonlinear transforms in characterizing the nonlinear relationship between mixed priors.

Fig. 5 presents the grayscale representation and spectral profile curves of the video restoration results with SR = 10%. For enhanced visual discernibility, a localized region, denoted by a red box, is magnified within each subfigure. As shown in the partially enlarged area of the subfigure in Fig. 5, the edge textures and detailed characteristics of the video data are restored more realistically by the proposed NTCTV.

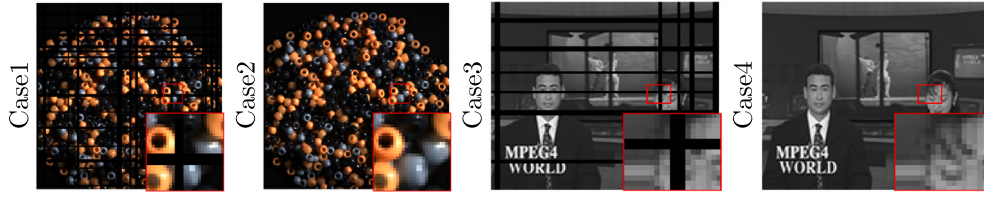


Fig. 6. Examples of different structured missing. Case 1: MSI Beads (sequential rows and columns missing). Case 2: MSI Beads (frontal slices 10~12 missing). Case 3: Video News (sequential rows and columns missing). Case 4: Video News (frontal slices 72~77 missing). Cases 2 and 4 presented are ground truth for visualization purposes.

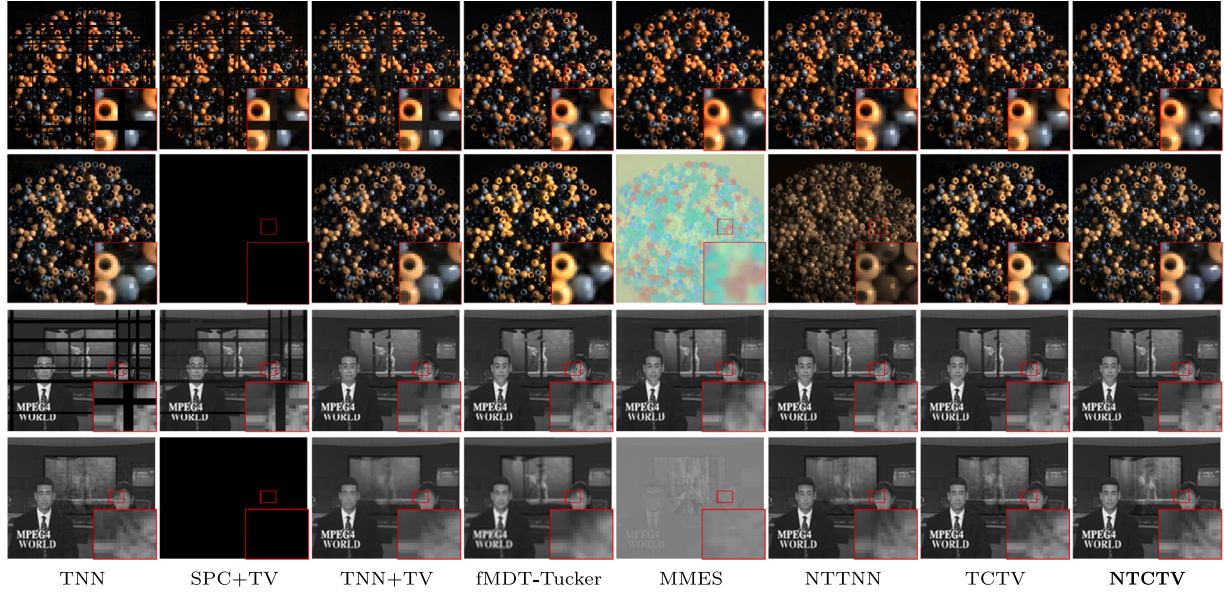


Fig. 7. The visual results of structured missing data completion by all competing methods. From top to bottom: Case 1 (R: 12, G: 11, B: 10), Case 2 (R: 12, G: 11, B: 10), Case 3 (band: 75) and Case 4 (band: 75).

In addition, the spectral profile curve of a single spatial position of NTCTV is more closely with the ground truth. This outcome indicates that the global low-rankness and local smoothness are expressed more realistically.

4.5. Structured missing tensor completion

In this subsection, we further evaluate the performance of NTCTV through a more challenging task of completing tensors with structured missing. Simulation experiments are conducted using MSI Beads and Video News as examples. Two types of structured missing are considered: sequential rows and columns missing, and consecutive frontal slices missing. Refer to Fig. 6 for more detail.

Parameter configuration. To mitigate parameter effects on completing the structured missing tensors, the parameter settings for all methods are aligned with those used in the experiments on random voxel missing.

Performance analysis. Table 5 presents a summary of the quantitative results obtained by all methods for four cases. Specifically, case 2 and 4 show the PSNR, SSIM, and FSIM values of the recovered frontal slices. Our method demonstrates good performance in scenarios involving sequential rows and columns missing. However, its performance deteriorated when dealing with consecutive frontal slices missing. This decline in performance can be attributed to consecutive frontal slice gaps, which pose challenges in learning suitable nonlinear transformations along mode-3 and are inferior to using FFT.

Fig. 7 illustrates visual examples of the recovery results obtained by all methods. By examining the overall image, it is observed that the proposed method lacks saturation in cases 2 and 4. However, when analyzing the magnified patch, the proposed method clearly captured the intricate texture features across all four cases. These results demonstrate the capability of our method to address the more arduous task of completing tensors with structured missing.

5. Discussions

In this section, a sensitivity analysis of the various components of the proposed NTCTV model is presented. This analysis examines the influence of parameters, conducts ablation study, and assesses the numerical convergence of the algorithm.

5.1. Analysis of proximal parameter and regularization parameters

In this subsection, we use the HSI Pavia dataset as an example to discuss the impact of the proximal operator ρ , and penalty parameters α_i ($i = 1, 2, 3$) and β on our proposed model. Because all α_i have the same purpose of eliminating the constraints introduced by auxiliary variables and the constraints related to β are derived directly from the initial constraint terms, we set all terms of α_i to a uniform value α across all experiments. This facilitates a controlled evaluation of the effects of the proximal operator and penalty parameters.

Fig. 8 illustrates the influence of the proximal parameter ρ and the penalty parameters α and β . Specifically, subfigure (a) shows the PSNR

Table 5
Quantitative comparison of all methods on structured missing completion results.

Missing	Case 1			Case 2			Case 3			Case 4		
Method	PSNR	SSIM	FSIM	PSNR	SSIM	FSIM	PSNR	SSIM	FSIM	PSNR	SSIM	FSIM
TNN	19.751	0.682	0.824	29.187	0.882	0.956	15.906	0.597	0.743	<u>29.385</u>	0.902	0.935
SPC+TV	21.042	0.761	0.845	15.200	0.083	0.362	19.138	0.735	0.825	8.802	0.001	0.442
TNN+TV	21.862	0.792	0.848	<u>29.997</u>	0.949	0.967	27.220	0.929	0.932	29.240	0.911	0.929
fMDT-Tucker	26.773	0.892	0.921	29.660	<u>0.949</u>	0.964	29.019	0.943	0.951	27.361	0.892	0.909
MMES	27.512	0.896	0.925	7.078	0.017	0.599	28.258	0.919	0.939	10.112	0.298	0.578
NTTNN	26.668	0.877	0.916	21.242	0.681	0.807	29.021	0.940	0.950	27.854	0.893	0.927
TCTV	27.270	0.893	0.923	30.750	0.925	0.975	29.159	0.948	0.956	31.281	0.930	0.951
NTCTV	27.767	0.906	0.930	29.461	0.922	<u>0.972</u>	29.446	0.948	0.956	28.767	<u>0.914</u>	<u>0.943</u>

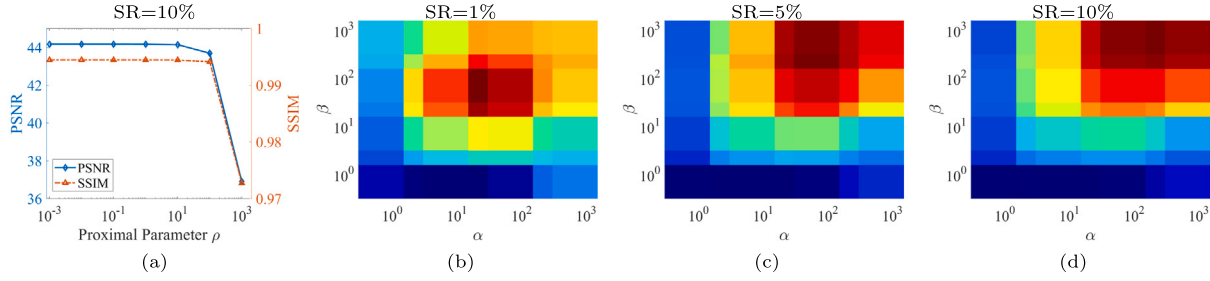


Fig. 8. Visual analysis of the impact of NTCTV parameters. (a) ρ . (b)–(d) α and β .

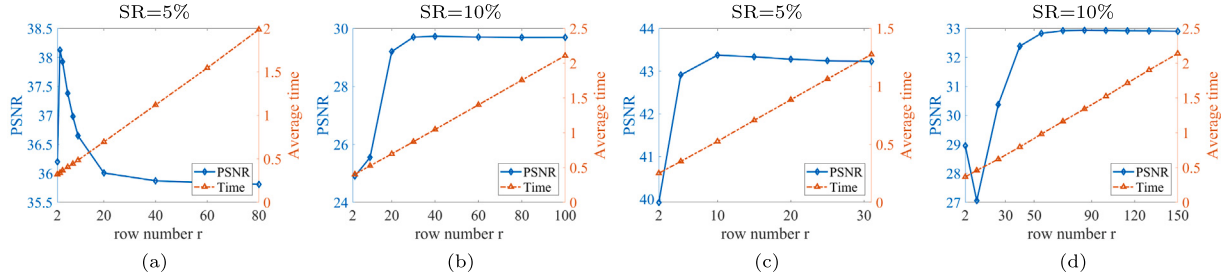


Fig. 9. Visual analysis of the impact of parameter r . (a) HSI Pavia. (b) CT Sacral. (c) MSI Face. (d) Video News.

and SSIM curves resulting from different settings of ρ at SR = 10%. Subfigure (b)–(d) show color maps, projected from the values of α and β onto PSNR, for SRs of 1%, 5% and 10%. The values of α and β are selected from a candidate set $\{0.3, 1.5, 3, 15, \dots, 3 \times 10^2, 15 \times 10^2\}$. Regarding the proximal operator parameter ρ , subgraph (a) demonstrates that it exhibits robustness when selected from $\{10^{-3}, 10^{-2}, \dots, 10^3\}$, and shows a certain anti-interference ability during tensor completion. As for the penalty parameters α and β , subgraphs (b)–(d) indicate some relative sensitivity, but they exhibit similar optimal choices across different SRs. This finding suggests the choice of α and β has a degree of universality for diverse data at varying SRs. Based on the aforementioned analysis and the guidance provided in Fig. 8, we select the ρ , α and β to 0.001, 30 and 150, respectively, in all experiments.

5.2. Analysis of row number r of T_k

In this subsection, we analyze the impact of the row number r of the learned linear semi-orthogonal matrix $T_k (k \in \Gamma)$ using four representative images: HSI Pavia, CT Sacral, MSI Face, Video News. As stated in Remark 1 of [36], the implicit low-rank structure of the gradient tensor in different directions consistently changes with that of the original tensor in the linear transform domain. This indicates that the degree of structural information concentration for the gradient tensor is similar across directions. Thus, we set the number of rows for T_k to a uniform value r .

Fig. 9(a)–(d) depict the PSNR curve and average time (in seconds) curve across varying row numbers r , at particular SRs for each dataset,

respectively. The PSNR curves demonstrate varying sensitivities to r across the four data types. For MSI, medical imaging, and videos, r exhibits robustness when it exceeds a certain threshold, indicating that r is relatively stable in the proposed method for these three types of data. In contrast, lower values of r demonstrate enhanced performance for HSIs processing, suggesting that monotonically increasing r does not necessarily optimize experimental results. This divergence is attributed to the highly low-rank redundancy of HSIs under a fused prior. The average time curves delineate a linear relationship between r and the average time of a single iteration, which aligns with the computational complexity derived theoretically in Section 3.4. Thus, we select an appropriately small r based on the specific experimental data type to achieve a balance between performance and cost in all the experiments.

5.3. Analysis of the effectiveness of nonlinear transform

This subsection analyzes the impact of the nonlinear transform on the proposed NTCTV. Specifically, we compare the performance of NTCTV in restoring MSI Beads without a nonlinear activation function $\phi(\cdot)$ (referred to as NTCTV(linear)) and with different nonlinear activation functions applied, including the sigmoid function [41], softplus function [49] and tanh function.

Table 6 lists the detailed recoveries of the MSI Beads in terms of different nonlinear transforms at different SRs. This indicates that NTCTV(tanh) achieved the best performance across all SRs. In contrast, both NTCTV(sigmoid) and NTCTV(softplus) show inferior performances compared to NTCTV(linear), which differs from the result [23] that all nonlinear activation functions play a promoting role in the original

Table 6

MSI Beads recovery results with different nonlinear activation functions. The optimal and the sub-optimal values are highlighted in **boldface** and underline respectively.

SR	0.05			0.1			0.15		
Metric	PSNR	SSIM	FSIM	PSNR	SSIM	FSIM	PSNR	SSIM	FSIM
Observed	14.574	0.102	0.480	14.805	0.138	0.558	15.053	0.176	0.616
NTCTV(linear)	<u>26.643</u>	<u>0.843</u>	<u>0.896</u>	<u>30.958</u>	<u>0.932</u>	<u>0.952</u>	<u>34.335</u>	<u>0.965</u>	<u>0.975</u>
NTCTV(sigmoid)	24.604	0.774	0.863	28.380	0.898	0.930	31.576	0.948	0.963
NTCTV(softplus)	25.494	0.800	0.877	29.719	0.915	0.942	33.295	0.959	0.971
NTCTV(tanh)	26.996	0.856	0.903	31.866	0.944	0.959	35.445	0.972	0.979

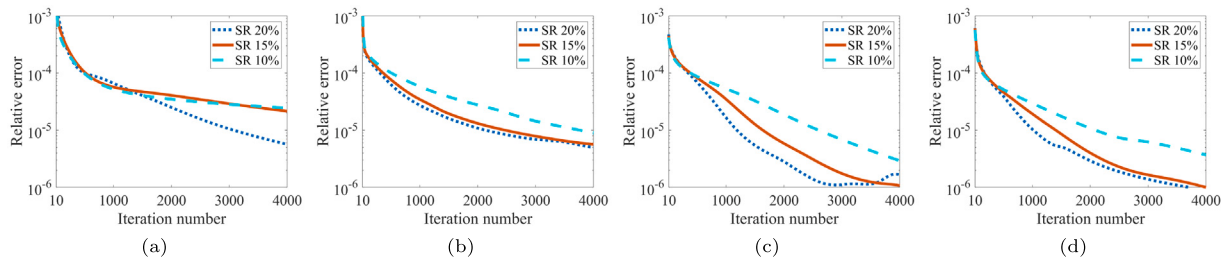


Fig. 10. Curves of relative changes versus iterations. (a) HSI Pavia. (b) CT Sacral. (c) MSI Face. (d) Video News.

TNN framework. This result aligns with Proposition 1 because $\phi(\cdot) = \tanh(\cdot)$ satisfies the condition $|\phi(x)| \leq |x|$. Thus, we employ tanh as the nonlinear activation function.

5.4. Analysis of numerical convergence

This subsection verifies the theoretical convergence of the proposed PAM-based algorithm through numerical convergence. We select HSI Pavia, CT Sacral, MSI Face and Video News datasets as representatives of each image type to test at various SRs. Fig. 10 shows the convergence curves of the proposed PAM-based algorithm, illustrating the relative change over iterations and reflecting the numerical convergence of the algorithm.

6. Conclusions

In this paper, we propose NTCTV, a fusion prior regularization term that incorporates a learnable nonlinear transform. It aims to adaptively explore the structural information present in the fused $R \circ S$ -priors, which are commonly observed in visual tensor data. The proposed nonlinear transform effectively constrains the low-rankness and smoothness of the data by incorporating linear semi-orthogonal transforms specifically designed for the gradient tensor, followed by embedding the nonlinear transform. Additionally, we designed a corresponding optimization model and devised an efficient PAM-based algorithm based on the proposed NTCTV. The theoretical analysis includes the computational complexity and proof of convergence for the proposed algorithm. Furthermore, extensive numerical experimental results confirm that the proposed NTCTV method has excellent performance and achieves tensor recovery even under extremely low SRs.

CRedit authorship contribution statement

Yigong Zhang: Writing – original draft, Methodology, Conceptualization. **Zhihui Tu:** Writing – review & editing, Methodology. **Jian Lu:** Writing – review & editing, Supervision, Funding acquisition. **Chen Xu:** Writing – review & editing, Funding acquisition. **Lixin Shen:** Supervision, Resources.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgments

This work is supported in part by the National Natural Science Foundation of China under grants U21A20455, 12326619, 61972265, and 62372302, the Natural Science Foundation of Guangdong Province of China under grants 2020B1515310008, 2023A1515011691 and 2024A1515011913, the Educational Commission of Guangdong Province of China under grant 2019KZDZX1007, and the National Science Foundation under grants DMS-2208385.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.knsys.2024.111917>.

References

- [1] H. Fanaee-T, J. Gama, Tensor-based anomaly detection: An interdisciplinary survey, *Knowl.-Based Syst.* 98 (2016) 130–147.
- [2] Y. Panagakis, J. Kossaifi, G.G. Chrysos, J. Oldfield, M.A. Nicolaou, A. Anandkumar, S. Zafeiriou, Tensor methods in computer vision and deep learning, *Proc. IEEE* 109 (5) (2021) 863–890.
- [3] S.E. Leh, M.M. Chakravarty, A. Pito, et al., The connectivity of the human pulvinar: a diffusion tensor imaging tractography study, *Int. J. Biomed. Imaging* 2008 (2008).
- [4] E. Frolov, I. Oseledets, Tensor methods and recommender systems, *Wiley Interdiscipl. Rev. Data Min. Knowl. Discovery* 7 (3) (2017) e1201.
- [5] J. Liu, P. Musialski, P. Wonka, J. Ye, Tensor completion for estimating missing values in visual data, *IEEE Trans. Pattern Anal. Mach. Intell.* 35 (1) (2012) 208–220.
- [6] F.L. Hitchcock, The expression of a tensor or a polyadic as a sum of products, *J. Math. Phys.* 6 (1–4) (1927) 164–189.
- [7] L.R. Tucker, Implications of factor analysis of three-way matrices for measurement of change, *Problems Measur. Change* 15 (122–137) (1963) 3.
- [8] Y. Xu, R. Hao, W. Yin, Z. Su, Parallel matrix factorization for low-rank tensor completion, 2013, arXiv preprint [arXiv:1312.1254](https://arxiv.org/abs/1312.1254).

- [9] B. Romera-Paredes, M. Pontil, A new convex relaxation for tensor completion, *Adv. Neural Inf. Process. Syst.* 26 (2013).
- [10] I.V. Oseledets, Tensor-train decomposition, *SIAM J. Sci. Comput.* 33 (5) (2011) 2295–2317.
- [11] Q. Zhao, G. Zhou, S. Xie, L. Zhang, A. Cichocki, Tensor ring decomposition, 2016, arXiv preprint arXiv:1606.05535.
- [12] Y.-B. Zheng, T.-Z. Huang, X.-L. Zhao, Q. Zhao, T.-X. Jiang, Fully-connected tensor network decomposition and its application to higher-order tensor completion, in: *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35, No. 12, 2021, pp. 11071–11078.
- [13] J.A. Bengua, H.N. Phien, H.D. Tuan, M.N. Do, Efficient tensor completion for color image and video recovery: Low-rank tensor train, *IEEE Trans. Image Process.* 26 (5) (2017) 2466–2479.
- [14] W. Wang, V. Aggarwal, S. Aeron, Efficient low rank tensor ring completion, in: *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 5697–5705.
- [15] C.-Y. Lyu, X.-L. Zhao, B.-Z. Li, H. Zhang, T.-Z. Huang, Multi-dimensional image recovery via fully-connected tensor network decomposition under the learnable transforms, *J. Sci. Comput.* 93 (2) (2022) 49.
- [16] J. Miao, K.I. Kou, L. Yang, D. Cheng, Quaternion tensor train rank minimization with sparse regularization in a transformed domain for quaternion tensor completion, *Knowl.-Based Syst.* 284 (2024) 111222.
- [17] Y.-B. Zheng, T.-Z. Huang, X.-L. Zhao, Q. Zhao, Tensor completion via fully-connected tensor network decomposition with regularized factors, *J. Sci. Comput.* 92 (1) (2022) 8.
- [18] M.E. Kilmer, C.D. Martin, Factorization strategies for third-order tensors, *Linear Algebra Appl.* 435 (3) (2011) 641–658.
- [19] O. Semerci, N. Hao, M.E. Kilmer, E.L. Miller, Tensor-based formulation and nuclear norm regularization for multienergy computed tomography, *IEEE Trans. Image Process.* 23 (4) (2014) 1678–1693.
- [20] E. Kernfeld, M. Kilmer, S. Aeron, Tensor–tensor products with invertible linear transforms, *Linear Algebra Appl.* 485 (2015) 545–570.
- [21] T.-X. Jiang, M.K. Ng, X.-L. Zhao, T.-Z. Huang, Framelet representation of tensor nuclear norm for third-order tensor completion, *IEEE Trans. Image Process.* 29 (2020) 7233–7244.
- [22] Y.-S. Luo, X.-L. Zhao, T.-X. Jiang, Y. Chang, M.K. Ng, C. Li, Self-supervised nonlinear transform-based tensor nuclear norm for multi-dimensional image recovery, *IEEE Trans. Image Process.* 31 (2022) 3793–3808.
- [23] B.-Z. Li, X.-L. Zhao, T.-Y. Ji, X.-J. Zhang, T.-Z. Huang, Nonlinear transform induced tensor nuclear norm for tensor completion, *J. Sci. Comput.* 92 (3) (2022) 83.
- [24] L.I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D* 60 (1–4) (1992) 259–268.
- [25] Y. Chang, L. Yan, H. Fang, C. Luo, Anisotropic spectral-spatial total variation model for multispectral remote sensing image destriping, *IEEE Trans. Image Process.* 24 (6) (2015) 1852–1866.
- [26] A.J. Tom, S.N. George, Simultaneous reconstruction and moving object detection from compressive sampled surveillance videos, *IEEE Trans. Image Process.* 29 (2020) 7590–7602.
- [27] D. Qiu, M. Bai, M.K. Ng, X. Zhang, Robust low-rank tensor completion via transformed tensor nuclear norm with total variation regularization, *Neurocomputing* 435 (2021) 197–215.
- [28] T. Yokota, Q. Zhao, A. Cichocki, Smooth PARAFAC decomposition for tensor completion, *IEEE Trans. Signal Process.* 64 (20) (2016) 5423–5436.
- [29] X. Li, Y. Ye, X. Xu, Low-rank tensor completion with total variation for visual data inpainting, in: *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 31, No. 1, 2017.
- [30] Y. Chen, S. Wang, Y. Zhou, Tensor nuclear norm-based low-rank approximation with total variation regularization, *IEEE J. Sel. Top. Sign. Proces.* 12 (6) (2018) 1364–1377.
- [31] J. Lu, J. Tian, Q. Jiang, X. Liu, Z. Hu, Y. Zou, Rician noise removal via weighted nuclear norm penalization, *Appl. Comput. Harmon. Anal.* 53 (2021) 180–198.
- [32] Z. Tu, J. Lu, H. Zhu, H. Pan, W. Hu, Q. Jiang, Z. Lu, A new nonconvex low-rank tensor approximation method with applications to hyperspectral images denoising, *Inverse Problems* 39 (6) (2023) 065003.
- [33] C. Xu, X. Liu, J. Zheng, L. Shen, Q. Jiang, J. Lu, Nonlocal low-rank regularized two-phase approach for mixed noise removal, *Inverse Problems* 37 (8) (2021) 085001.
- [34] J. Lu, L. Shen, C. Xu, Y. Xu, Multiplicative noise removal in imaging: An exp-model and its fixed-point proximity algorithm, *Appl. Comput. Harmon. Anal.* 41 (2) (2016) 518–539.
- [35] H. Zhang, L. Liu, W. He, L. Zhang, Hyperspectral image denoising with total variation regularization and nonlocal low-rank tensor decomposition, *IEEE Trans. Geosci. Remote Sens.* 58 (5) (2019) 3071–3084.
- [36] H. Wang, J. Peng, W. Qin, J. Wang, D. Meng, Guaranteed tensor recovery fused low-rankness and smoothness, *IEEE Trans. Pattern Anal. Mach. Intell.* (2023).
- [37] H. Attouch, J. Bolte, B.F. Svaiter, Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss–Seidel methods, *Math. Program.* 137 (1–2) (2013) 91–129.
- [38] J. Bolte, S. Sabach, M. Teboulle, Proximal alternating linearized minimization for nonconvex and nonsmooth problems, *Math. Program.* 146 (1–2) (2014) 459–494.
- [39] C. Lu, X. Peng, Y. Wei, Low-rank tensor completion with a new tensor nuclear norm induced by invertible linear transforms, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 5996–6004.
- [40] T.G. Kolda, B.W. Bader, Tensor decompositions and applications, *SIAM Rev.* 51 (3) (2009) 455–500.
- [41] S. Haykin, *Neural networks: a comprehensive foundation*, Prentice Hall PTR, 1998.
- [42] Y. Wang, J. Yang, W. Yin, Y. Zhang, A new alternating minimization algorithm for total variation image reconstruction, *SIAM J. Imaging Sci.* 1 (3) (2008) 248–272.
- [43] J.-F. Cai, E.J. Candès, Z. Shen, A singular value thresholding algorithm for matrix completion, *SIAM J. Optim.* 20 (4) (2010) 1956–1982.
- [44] W. Qin, H. Wang, F. Zhang, J. Wang, X. Luo, T. Huang, Low-rank high-order tensor completion with applications in visual data, *IEEE Trans. Image Process.* 31 (2022) 2433–2448.
- [45] R. Yamamoto, H. Hontani, A. Imakura, T. Yokota, Fast algorithm for low-rank tensor completion in delay-embedded space, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 2058–2066.
- [46] T. Yokota, H. Hontani, Q. Zhao, A. Cichocki, Manifold modeling in embedded space: An interpretable alternative to deep image prior, *IEEE Trans. Neural Netw. Learn. Syst.* PP (2020) 1–15, <http://dx.doi.org/10.1109/TNNLS.2020.3037923>.
- [47] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, *IEEE Trans. Image Process.* 13 (4) (2004) 600–612.
- [48] L. Zhang, L. Zhang, X. Mou, D. Zhang, FSIM: A feature similarity index for image quality assessment, *IEEE Trans. Image Process.* 20 (8) (2011) 2378–2386.
- [49] V. Nair, G.E. Hinton, Rectified linear units improve restricted boltzmann machines, in: *Proceedings of the 27th International Conference on Machine Learning*, ICML-10, 2010, pp. 807–814.