HiNCoT: Hierarchical Nonlinear Continuous Transform-based Tensor Representation for Multi-Dimensional Data Recovery

Tao Yang[1], **Weihao Wu**[1*], **Tingzhu Huang**[1*]

[1]University of Electronic Science and Technology of China, Chengdu, China
taoyangworld@163.com, weihaowu99@163.com, tingzhuhuang@126.com

Abstract

Recently, continuous transform-based tensor representation has emerged as a promising tool for multi-dimensional data recovery. However, the existing continuous transforms are essentially single-layer linear mappings, which limits their ability to capture the complex relationships inherent in multi-dimensional data. To overcome this limitation, we propose a Hierarchical Nonlinear Continuous Transform-based Tensor Representation (HiNCoT) for multi-dimensional data recovery. By leveraging the hierarchical nonlinear continuous transform, HiNCoT constructs the recovered tensor from a latent tensor, which is generated by the deep representation module with a low-rank core tensor as input. Compared with the existing continuous transform-based methods, HiNCoT can more effectively capture the complex nonlinear relationships inherent in multi-dimensional data along the third dimension. To evaluate the effectiveness of the proposed HiNCoT, we suggest an HiNCoT-based multi-dimensional data recovery model. Extensive experiments on diverse degeneration scenarios demonstrate the superiority of our hierarchical nonlinear transform-based method over existing single-layer linear transform-based methods.

Introduction

Multi-dimensional data, such as color images (Jia et al. 2021; Lian, Zagorodnov, and Tan 2005), multispectral images (MSIs) (Anandkumar et al. 2014; Wu et al. 2023), hyperspectral images (HSIs) (Zhao et al. 2025; Li et al. 2020), magnetic resonance imaging (MRI) (Fabian, Heckel, and Soltanolkotabi 2021; Deka, Datta, and Handique 2018), and videos (Lin et al. 2025; Wang et al. 2023b), are widely used in various real-world applications. However, the acquired data are often incomplete due to sensor faults or transmission errors, hindering subsequent analysis and use (Ono 2017).

Transform-based tensor representation is a powerful framework for multi-dimensional data recovery. It applies a specified mode-3 transform to the multi-dimensional data and models the resulting transformed tensor (termed as the latent tensor) for subsequent processing (Wang et al. 2023a; Feng et al. 2023; Qin et al. 2022; Wang et al. 2025; Ji, Zhao, and Sun 2022).

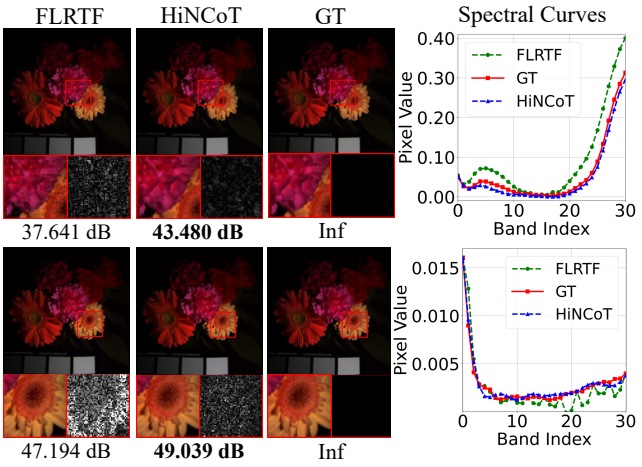

Figure 1: Visual and spectral fidelity comparison of HiNCoT and FLRTF. From left to right: recovered pseudo-color images (R:31, G:15, B:5) with zoom-in views and residuals, followed by spectral curves of the recovered results obtained by two methods. From top to bottom: results for random missing with SR $= 0.1$ on *Flowers* and corresponding spectral curves at spatial location $(120, 120)$; results for slice missing on *Flowers* and corresponding spectral curves at spatial location $(180, 20)$.

Previously, transform-based tensor representation mainly focused on the development of discrete transforms. For example, Braman *et al.* (Braman 2010) used the discrete Fourier transform (DFT) along the third mode and leveraged the singular value decomposition (SVD) to exploit the low-rankness in the transformed domain, Madathil *et al.* (Madathil and George 2018) adopted the real-valued discrete cosine transform (DCT) to replace DFT in order to mitigate the computation burden arising from complex operations. More generally, Lu *et al.* (Lu, Peng, and Wei 2019) used the invertible linear transform and gave the theoretically sampling bound to guarantee the exact recovery, Jiang *et al.* (Jiang et al. 2020) proposed a non-invertible framelet transform in t-SVD framework for tensor completion prob-

*Corresponding authors.

lems. Furthermore, Jiang *et al.* (Jiang et al. 2021) introduced a data-driven transform which has more flexible expressive capabilities compared to pre-defined transforms. Nevertheless, the transforms used in the above methods are single-layer linear transforms, which may not be able to model the complex and diversified topological structures of real-world data. Motivated by these limitations, several recent studies have suggested employing the hierarchical nonlinear transform. For example, Luo *et al.* (Luo et al. 2022) suggested to use a multi-layer perceptron (MLP) as transform to get a better low-rank representation. Li *et al.* (Li et al. 2023) proposed H2TF which use hierarchical nonlinear transform and hierarchical matrix factorization to compactly represent multi-dimensional data. Zhou *et al.* (Zhou et al. 2024) designed a deep transform module to capture the frontal slice relationships of multi-dimensional data and proposed a general tensor representation framework. However, discrete transform-based tensor representation methods struggle to handle slice missing tasks along the third mode of multi-dimensional data (e.g., MSI band recovery and video frame interpolation) without additional regularization. Recently, the continuous transform-based tensor representation has emerged as a promising tool for multi-dimensional data recovery. Luo *et al.* (Luo et al. 2023) proposed a novel tensor function representation for continuously representing multi-dimensional data. Wang *et al.* (Wang and Zhao 2024) proposed the functional transform-based low-rank tensor factorization, where the mode-3 factor matrix is modeled by a functional transform parameterized via an implicit neural representation (INR) (Park et al. 2019; Sitzmann et al. 2020; Mescheder et al. 2019). However, the transforms used in existing methods are typically limited to single-layer linear transforms. Owing to the complex and diversified topological structures inherent in real-world data, the transform from the original tensor to its optimal representation in latent space is likely to be nonlinear and hierarchical (Luo et al. 2022; Li et al. 2022, 2023). Thus, single-layer linear transforms fail to capture the complex nonlinear relationships inherent in real-world data. This naturally raises the question: how can we simultaneously capture nonlinear relationship in real-world data while enabling a continuous representation?

To answer this question, we propose a Hierarchical Continuous Transform-based Tensor Representation (HiNCoT) as a novel representation method for multi-dimensional data in a continuous formulation. Compared to single-layer linear transforms, we design a hierarchical nonlinear continuous transform, which adopts a hierarchical learnable transform coupled with nonlinear activation functions. The transform is formulated using an implicit neural representation, allowing for a continuous representation of multi-dimensional data. The combination of these components allows HiNCoT to simultaneously capture the complex nonlinear relationships and the inherent temporal/spectral smoothness of multi-dimensional data. Meanwhile, we leverage an untrained neural network as the deep representation module to generate the latent tensor by characterizing both intra-slice and inter-slice relationships of the latent tensor. To evaluate the representation ability of the proposed HiNCoT, we consider representative multi-dimensional data recovery tasks and build an unsupervised HiNCoT-based recovery model. We compare the performance of HiNCoT (based on hierarchical nonlinear transform) with FLRTF (Wang and Zhao 2024) (based on single-layer linear transform) under the random missing and slice missing scenarios, as illustrated in Fig. 1. The results indicate that HiNCoT yields more satisfied visual results and higher PSNR values compared with FLRTF, which indicate that our proposed HiNCoT enjoys more powerful representation capability than FLRTF.

Our contributions are summarized as follows:

- To overcome the limitations of single-layer linear transforms, we propose HiNCoT, a hierarchical nonlinear continuous transform-based tensor representation. By synergistically integrating a hierarchical nonlinear continuous transform module with a deep representation module, HiNCoT can effectively capture the complex nonlinear relationships inherent in multi-dimensional data along the third dimension, in contrast to existing methods based on single-layer linear transforms.
- Based on HiNCoT, we propose an unsupervised multi-dimensional data recovery model. Extensive experiments validate that our method outperforms the competing methods across different degradation scenarios, including random, tube, and slice missing.

Notations and Preliminaries

Scalars, vectors, matrices, and tensors are represented as variables x, $\mathbf{x}$, $\mathbf{X}$, and $\mathcal{X}$, respectively. Given a third-order tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, we denote its (i,j,k)-th entry as $\mathcal{X}(i,j,k)$, and its k-th frontal slice as $\mathcal{X}(:,:,k)$. The symbol $\circledast$ denotes the convolution operations. The Frobenius norm of $\mathcal{X}$ is defined as $\|\mathcal{X}\|_F = \sqrt{\sum_{i,j,k} |\mathcal{X}(i,j,k)|^2}$. The Hadamard product of two tensors $\mathcal{X}$ and $\mathcal{Y} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, is denoted as $\mathcal{Z} = \mathcal{X} \odot \mathcal{Y}$, where each element satisfies $\mathcal{Z}(i,j,k) = \mathcal{X}(i,j,k) \cdot \mathcal{Y}(i,j,k)$. Given two third-order tensors, $\mathcal{X} \in \mathbb{R}^{n_1 \times r \times n_3}$ and $\mathcal{Y} \in \mathbb{R}^{r \times n_2 \times n_3}$, their face-wise product (Kilmer et al. 2013) is defined as $\mathcal{Z} = \mathcal{X} \triangle \mathcal{Y} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, where the frontal slice of $\mathcal{Z}$ is given by $\mathcal{Z}(:,:,k) = \mathcal{X}(:,:,k)\mathcal{Y}(:,:,k)$ for all $k \in \{1, \ldots, n_3\}$.

Definition 1 ***(Mode-3 unfolding and folding)*** *(Kolda and Bader 2009)* Given a tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, its mode-3 unfolding is a matrix $\mathbf{X}_{(3)} \in \mathbb{R}^{n_3 \times n_1 n_2}$ defined by $\mathbf{X}_{(3)}(k,l) = \mathcal{X}(i,j,k)$, where $l = (j-1)n_1 + i$. The mode-3 unfolding and folding operators are denoted by $\mathbf{X}_{(3)} = \mathrm{Unfold}_3(\mathcal{X})$ and $\mathcal{X} = \mathrm{Fold}_3(\mathbf{X}_{(3)})$, respectively.

Definition 2 ***(Mode-3 product)*** *(Kolda and Bader 2009)* The mode-3 product of a tensor $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times b}$ and a matrix $\mathbf{H} \in \mathbb{R}^{n_3 \times b}$ is defined as:

$$\mathcal{A} \times_3 \mathbf{H} = \mathrm{Fold}_3(\mathbf{H}\,\mathrm{Unfold}_3(\mathcal{A})). \tag{1}$$

Definition 3 ***(Tensor tubal-rank)*** *(Kilmer et al. 2013)* The tensor tubal-rank of $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is defined as:

$$\mathrm{rank}_t(\mathcal{A}) \triangleq \max_{k=1,2,\ldots,n_3} \mathrm{rank}\left((\mathcal{A} \times_3 \mathbf{F})^{(k)}\right), \tag{2}$$

where $\mathbf{F} \in \mathbb{R}^{n_3 \times n_3}$ is the DFT matrix, and $\mathrm{rank}(\cdot)$ denotes the matrix rank.

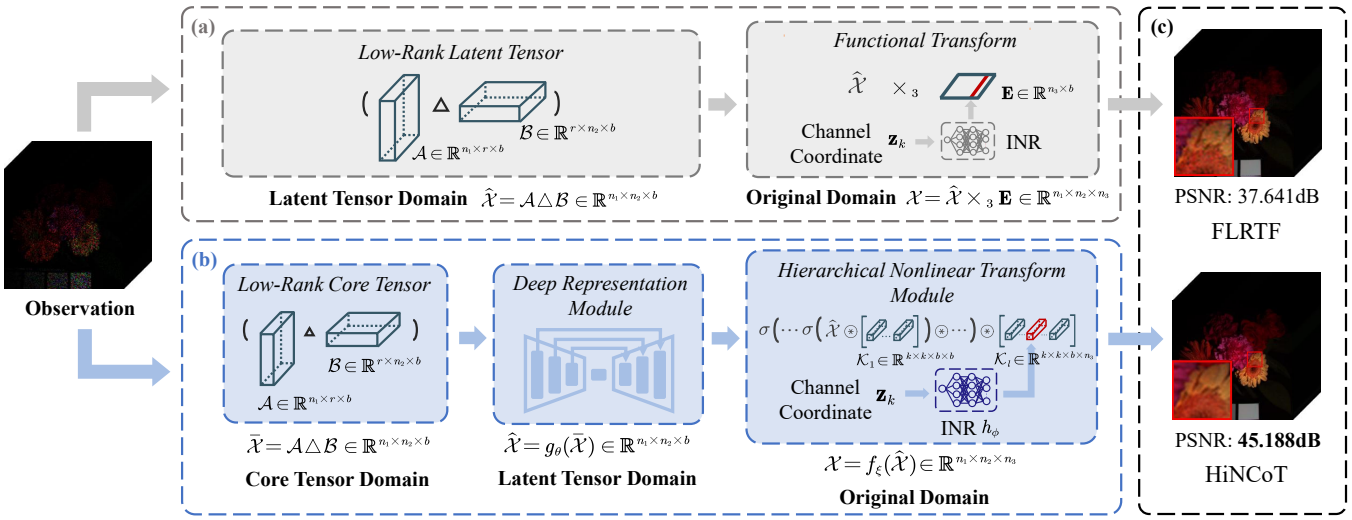

Figure 2: The framework of the proposed HiNCoT. (a) An example of the traditional continuous transform-based tensor representation, e.g., FLRTF (Wang and Zhao 2024). (b) The proposed hierarchical nonlinear continuous transform-based tensor representation, i.e., HiNCoT. (c) Result examples for random missing recovery on *Flowers* dataset of different methods.

Definition 4 ***(T-product)*** *(Kilmer et al. 2013)* The t-product between $\mathcal{A} \in \mathbb{R}^{n_1 \times r \times n_3}$ and $\mathcal{B} \in \mathbb{R}^{r \times n_2 \times n_3}$ is defined as $\mathcal{A} * \mathcal{B} = ((\mathcal{A} \times_3 \mathbf{F}) \triangle (\mathcal{B} \times_3 \mathbf{F})) \times_3 \mathbf{F}^{-1} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, where $\mathbf{F} \in \mathbb{R}^{n_3 \times n_3}$ is the DFT matrix, and $\mathbf{F}^{-1}$ is the inverse DFT matrix.

The Proposed HiNCoT

Continuous Transform-based Tensor Representation

Continuous transform-based tensor representation methods offer a continuous formulation to represent multi-dimensional data, which typically involves two key components: (i) a continuous transform which captures local smoothness along the third dimension, and (ii) the representation of the latent tensor. The core idea of continuous transform-based tensor representation is to model the third-mode transform as a continuous function that maps positional indices (e.g., time, wavelength, or frame) to latent tensor slices, often through parametric models such as basis functions or neural networks. However, existing methods typically adopt a single-layer linear transform (Wang and Zhao 2024), thereby restricting their representational capacity. These limitations motivate the exploration of more expressive transform designs.

The Deep Representation Module

Existing continuous transform-based tensor representation methods usually employ low-rank factorization to characterize the intra-slice structures (i.e., low-rankness) within the frontal slices of the latent tensor, which neglects the complex inter-slice dependencies. In the proposed HiNCoT, we leverage the deep representation module to generate the latent tensor starting from the low-rank core tensor. Specifically, low-rank factorization is first applied to the core tensor domain to capture the intra-slice structures, followed by the use of a deep representation module $g_\theta(\cdot)$ to model the inter-slice relationships, resulting in the latent tensor. The formulation of this process is given as follows:

$$\hat{\mathcal{X}} = g_\theta(\mathcal{A} \triangle \mathcal{B}), \tag{3}$$

where $\hat{\mathcal{X}} \in \mathbb{R}^{n_1 \times n_2 \times b}$ is the latent tensor, $\mathcal{A} \in \mathbb{R}^{n_1 \times r \times b}$ and $\mathcal{B} \in \mathbb{R}^{r \times n_2 \times b}$ are core factor tensors, r represents the tensor tubal rank of the core tensor, b represents the sizes of $\mathcal{A}$ and $\mathcal{B}$ in the third dimension, $g_\theta(\cdot) : \mathbb{R}^{n_1 \times n_2 \times b} \to \mathbb{R}^{n_1 \times n_2 \times b}$ is the deep representation module implemented by an untrained deep neural network, θ denotes the network parameters.

The Hierarchical Nonlinear Continuous Transform Module

In existing continuous transform-based tensor representation methods, the commonly used functional transforms are essentially single-layer linear mappings, which inherently limit their representation capacity. To overcome this limitation, we elaborately design a hierarchical nonlinear continuous transform module $f_\xi(\cdot)$, referred to as the HNC-transform module, which operates along the third dimension of the latent tensor. This module is specifically designed to simultaneously capture the complex nonlinear relationships and the continuity along the third dimension in multi-dimensional data, i.e.,

$$\mathcal{X} = f_\xi(\hat{\mathcal{X}}) = \sigma(\cdots \sigma(\hat{\mathcal{X}} \circledast \mathcal{K}_1) \circledast \cdots \circledast \mathcal{K}_{l-1}) \circledast \mathcal{K}_l, \tag{4}$$

where $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is the multi-dimensional data represented by the proposed HiNCoT, $f_\xi(\cdot) : \mathbb{R}^{n_1 \times n_2 \times b} \to \mathbb{R}^{n_1 \times n_2 \times n_3}$ is the HNC-transform module, ξ denotes the parameters of $f_\xi(\cdot)$, $\{\mathcal{K}_i\}_{i=1}^{l}$ are transform convolution kernels of $f_\xi(\cdot)$ with $\mathcal{K}_i \in \mathbb{R}^{b \times b \times k \times k}$ for $i = 1, 2, \cdots, l-1$ and $\mathcal{K}_l \in \mathbb{R}^{b \times n_3 \times k \times k}$, k is the kernel size, $\sigma(\cdot)$ is a nonlinear activate function.

To represent data in a continuous formulation, we use an INR to parameterize the last transform convolution kernel $\mathcal{K}_l$, i.e.,

$$\mathcal{K}_l(:,i,:,:) = h_\phi(\mathbf{z}_i), \quad i = 1, 2, \cdots, n_3. \tag{5}$$

Here, $h_\phi(\cdot) : D_f \to \mathbb{R}^b$ denotes an INR with parameters ϕ, where $D_f = (0, 1]$ represents the definition domain, $\mathbf{z} = [\frac{1}{n_3}, \frac{2}{n_3}, \cdots, \frac{n_3}{n_3}]$ represents the sampling coordinate vector. In our work, we employ a MLP to parameterize the INR due to its powerful universal approximation capability. We denote the total learnable parameters in the HNC-transform module as $\xi \triangleq \{\mathcal{K}_1, \mathcal{K}_2, \cdots, \mathcal{K}_{l-1}, \phi\}$.

By designing the hierarchical nonlinear continuous transform, HiNCoT can effectively capture complex nonlinear structures while preserving smoothness along the third dimension. This synergy significantly enhances HiNCoT's capability to reconstruct fine spatial details and preserve structural fidelity, especially under severe degradation scenarios.

Hierarchical Nonlinear Continuous Transform-based Tensor Representation

Based on the deep representation module and the HNC-transform module, we propose our hierarchical nonlinear continuous transform-based tensor representation.

Definition 5 ***(Hierarchical nonlinear continuous transform-based tensor representation (HiNCoT))*** Given a tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, its hierarchical nonlinear continuous transform-based tensor representation is defined as:

$$\mathcal{X} = f_\xi \circ g_\theta(\mathcal{A} \triangle \mathcal{B}), \tag{6}$$

where $\mathcal{A} \in \mathbb{R}^{n_1 \times r \times b}$ and $\mathcal{B} \in \mathbb{R}^{r \times n_2 \times b}$ are core factor tensors, $g_\theta(\cdot) : \mathbb{R}^{n_1 \times n_2 \times b} \to \mathbb{R}^{n_1 \times n_2 \times b}$ represents the deep representation module, and $f_\xi(\cdot) : \mathbb{R}^{n_1 \times n_2 \times b} \to \mathbb{R}^{n_1 \times n_2 \times n_3}$ denotes the HNC-transform module, as previously defined.

A general illustration of HiNCoT is provided in Fig. 2. By leveraging a hierarchical nonlinear continuous transform, HiNCoT is capable of simultaneously capturing the complex nonlinear relationships within the multi-dimensional data and preserving smoothness along the third dimension.

Remark: The proposed HiNCoT is designed as a continuous transform-based tensor representation framework that extends the existing method. Specifically, when the number of nonlinear transforms $l = 1$, the kernel size $k = 1$, and the deep representation module $g_\theta(\cdot)$ is set to the identity mapping, HiNCoT degrades to the functional transform-based low-rank tensor factorization (FLRTF) (Wang and Zhao 2024).

HiNCoT based Multi-Dimensional Data Recovery

To examine the representation capability of the proposed HiNCoT, we propose a general HiNCoT-based multi-dimensional data recovery model. Given an observed data $\mathcal{O} \in \mathbb{R}^{n_1 \times n_2 \times \hat{n}_3}$, the model can be formulated as:

$$\min_{\mathcal{A},\mathcal{B},\theta,\xi} \mathcal{L}(\mathcal{O}, \mathcal{X}), \quad \text{s.t.}\, \mathcal{X} = f_\xi \circ g_\theta(\mathcal{A} \triangle \mathcal{B}), \tag{7}$$

where $\mathcal{L}(\cdot,\cdot)$ is a task-specific data fidelity term that ensures consistency with the degradation scenario. Notably, the proposed model (7) is unsupervised and solely requires the observed data without a training dataset, making it adaptable to diverse datasets and degradation scenarios.

In this work, we first validate the effectiveness of the proposed HiNCoT-based multi-dimensional data recovery model on two types of spatial degradation scenarios, i.e., random missing and tube missing. Then we further extend the model to address a spectral degradation scenario, i.e., slice missing.

Random and Tube Missing Recovery In real-world MSIs, incomplete data frequently occurs due to sensor malfunctions. Two common degradation scenarios are random missing and tube missing. In the random missing scenario, each pixel is independently missing with a certain probability, whereas in the tube missing scenario, the entire spectral signature at a specific spatial location is absent. Effectively handling these degradation tasks is essential, as they impair both spatial structures and spectral fidelity, thereby hindering downstream tasks such as material classification and target detection (He et al. 2019). The HiNCoT-based multi-dimensional data recovery model (7) for random and tube missing can be formulated as:

$$\min_{\mathcal{A},\mathcal{B},\theta,\xi} \|\mathcal{O} - \mathcal{M} \odot \mathcal{X}\|_F^2, \quad \text{s.t.}\, \mathcal{X} = f_\xi \circ g_\theta(\mathcal{A} \triangle \mathcal{B}). \tag{8}$$

Here, $\mathcal{O} \in \mathbb{R}^{n_1 \times n_2 \times \hat{n}_3}$ denotes the observed tensor with $\hat{n}_3 = n_3$ in both scenarios, $\mathcal{M} \in \mathbb{R}^{n_1 \times n_2 \times \hat{n}_3}$ is a binary mask tensor that sets the observed location to ones and others to zeros.

Slice Missing Recovery In order to explore the performance upper bound of our method, we consider a more challenging task, i.e., slice missing recovery. Slice missing recovery aims to reconstruct the missing bands by exploiting the spectral correlations and smoothness within the available bands. The recovery model can be formulated as:

$$\min_{\mathcal{A},\mathcal{B},\theta,\xi} \|\mathcal{O} - \mathcal{X}(:,:,\Omega)\|_F^2, \quad \text{s.t.}\, \mathcal{X} = f_\xi \circ g_\theta(\mathcal{A} \triangle \mathcal{B}). \tag{9}$$

Here, $\mathcal{O} \in \mathbb{R}^{n_1 \times n_2 \times \hat{n}_3}$ denotes the observed tensor with $\hat{n}_3 < n_3$, Ω represents the index set of the observed bands, and $|\Omega| = \hat{n}_3$.

In the model (7), the optimization variables include core factor tensors $\mathcal{A}$ and $\mathcal{B}$, as well as the network parameters of $g_\theta(\cdot)$ and $f_\xi(\cdot)$. Since the model (8) and (9) are differentiable with respect to both the factor tensors and network parameters, we adopt the Adam optimizer (Kingma and Ba 2014) to solve the optimization problem. Once the optimal factor tensors and network parameters are obtained, the recovered result $\mathcal{X}$ can be computed via (6). The algorithm of the proposed HiNCoT-based multi-dimensional data recovery model is presented in Algorithm 1.

Experiments

In this section, we conduct comprehensive experiments on three degradation scenarios to verify the effectiveness and superiority of the proposed HiNCoT.

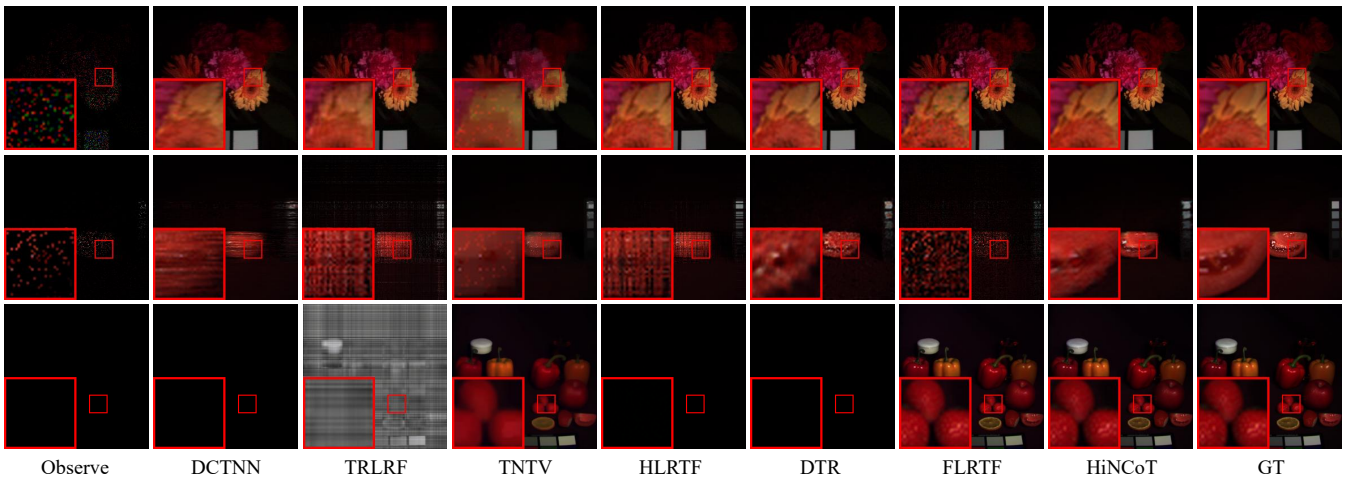

Figure 3: The recovered pseudo-color images by different methods. From top to bottom: random missing (SR $= 0.1$) with *Flowers* dataset (R:30, G:14, B:4), tube missing (SR $= 0.1$) with *Tomatoes* dataset (R:30, G:14, B:4), and slice missing with *Food* dataset (R:29, G:13, B:2).

Algorithm 1: HiNCoT-based multi-dimensional data recovery model

Input: observation $\mathcal{O}$, parameters r and b, maximum iteration number N;
Initialization: core factor tensors $\mathcal{A}$ and $\mathcal{B}$, network parameters θ and ξ, $N = 10000$;

1: **for** $i = 1$ to N **do**
2: Compute the recovered data $\mathcal{X}$ via (6);
3: Compute the model (7) for different tasks;
4: Compute the gradients w.r.t. $\mathcal{A}$, $\mathcal{B}$, θ, ξ;
5: Update $\mathcal{A}$, $\mathcal{B}$, θ, ξ using the Adam optimizer;
6: **end for**

Output: The recovered multi-dimensional data $\mathcal{X}$;

Experimental Settings

Compared Methods We compare HiNCoT with six representative multi-dimensional data recovery methods, including five discrete transform-based methods, i.e., DCTNN (Madathil and George 2018), TRLRF (Yuan et al. 2019), TNTV (Qiu et al. 2021), HLRTF (Luo et al. 2022), and DTR (Zhou et al. 2024), as well as one continuous transform-based method FLRTF (Wang and Zhao 2024), which is the most relevant to our proposed method. All parameters of the compared methods are carefully adjusted according to the suggestions provided by the respective authors in their articles. Experimental results demonstrate that HiNCoT achieves state-of-the-art performance across various degradation scenarios.

Implementation We employ four multispectral images (i.e., *Tomatoes*, *Lemons*, *Flowers*, and *Food*)[1] to test the performance of the proposed HiNCoT. The original size of these datasets is $512 \times 512 \times 31$, and we resize them into $256 \times 256 \times 31$ for our experiments. For random and tube missing scenarios, the sampling rates (SR) are set as $\{0.02, 0.05, 0.10\}$ and $\{0.10, 0.15, 0.20\}$ separately. For slice missing scenario, the sampling bands are set as $\{1, 4, 7, \cdots, 28, 31\}$. The grey values of all datasets are normalized band-by-band into the interval $[0, 1]$. The peak-signal-to-noise-ratio (PSNR) (Huynh-Thu and Ghanbari 2008) and structural similarity (SSIM) (Wang et al. 2004) are selected as the metric indexes to evaluate the quality of the recovered results. All experiments are performed on a PC equipped with one Intel(R) Core(TM) i5-12600KF CPU, one NVIDIA GeForce RTX 4060 Ti GPU and 32GB RAM. HLRTF, DFR, FLRTF and the proposed HiNCoT are implemented by using Python and PyTorch library with GPU calculation, and other comparison methods are implemented by MATLAB (R2022a) with CPU calculation.

Experimental Results

Tables 1–2 summarize the quantitative results of different methods under random missing and tube missing scenarios. Our HiNCoT consistently achieves superior performance in both scenarios. This improvement is attributed to the synergistic effect of the deep representation module and the HNC-transform module, which effectively capture the complex nonlinear relationships inherent in multi-dimensional data. Table 3 reports the experimental results for the slice missing scenario, in which the proposed HiNCoT achieves superior quantitative results than other competitors. Traditional discrete transform-based methods struggle in this scenario due to the discrete nature of their transforms[2], which makes the optimization process intractable. In contrast, the HNC-transform module enables HiNCoT to both capture the local smoothness along the third dimension of data, thus ensuring stable gradient propagation, and effectively characterize

[1] https://cave.cs.columbia.edu/repository/Multispectral

[2] Notably, TNTV can slightly recover the overall structure, primarily due to the incorporation of total variation (TV) regularization.

Sampling rate		0.02		0.05		0.10	
Data	Method	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
MSIs *Tomatoes* (256×256×31)	DCTNN	30.376	0.877	34.409	0.936	39.214	0.975
	TRLRF	30.406	0.853	35.317	0.938	39.395	0.970
	TNTV	26.841	0.843	32.117	0.922	34.303	0.960
	HLRTF	33.300	0.930	37.923	0.970	43.432	0.989
	DTR	32.512	**0.939**	38.137	0.976	41.454	0.986
	FLRTF	29.856	0.811	34.308	0.925	41.040	0.978
	HiNCoT	**34.097**	0.935	**39.029**	**0.979**	**44.742**	**0.993**
MSIs *Lemons* (256×256×31)	DCTNN	31.729	0.868	37.888	0.954	43.284	0.985
	TRLRF	32.276	0.847	36.808	0.928	40.879	0.961
	TNTV	28.619	0.814	34.784	0.933	39.844	0.975
	HLRTF	35.726	0.933	42.516	0.978	47.261	0.992
	DTR	32.029	0.873	40.149	0.969	45.895	0.990
	FLRTF	34.729	0.914	39.453	0.960	46.925	0.990
	HiNCoT	**37.381**	**0.949**	**44.860**	**0.987**	**48.376**	**0.995**
MSIs *Flowers* (256×256×31)	DCTNN	27.209	0.695	32.432	0.857	37.302	0.944
	TRLRF	25.671	0.520	32.015	0.794	36.268	0.897
	TNTV	24.063	0.680	28.856	0.818	32.738	0.902
	HLRTF	28.327	0.726	35.186	0.914	40.401	0.966
	DTR	28.892	0.771	37.216	0.954	42.796	0.981
	FLRTF	28.285	0.690	34.295	0.878	37.641	0.933
	HiNCoT	**32.338**	**0.880**	**38.232**	**0.963**	**45.188**	**0.990**
MSIs *Food* (256×256×31)	DCTNN	27.819	0.753	33.261	0.898	38.027	0.959
	TRLRF	26.252	0.579	32.321	0.833	36.557	0.924
	TNTV	23.438	0.670	29.544	0.864	33.484	0.937
	HLRTF	30.429	0.818	34.761	0.920	40.693	0.974
	DTR	26.415	0.568	34.617	0.928	38.904	0.967
	FLRTF	27.949	0.680	33.251	0.864	38.911	0.951
	HiNCoT	**32.306**	**0.887**	**35.637**	**0.939**	**41.322**	**0.981**

Table 1: The quantitative results of random missing scenarios by different methods.

Sampling rate		0.10		0.15		0.20	
Data	Method	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
MSIs *Tomatoes* (256×256×31)	DCTNN	28.533	0.862	30.268	0.891	31.285	0.910
	TRLRF	16.827	0.044	27.786	0.844	29.746	0.878
	TNTV	30.613	0.939	33.290	0.958	33.957	0.963
	HLRTF	28.613	0.873	30.564	0.906	31.929	0.931
	DTR	31.797	0.930	32.427	0.942	34.225	0.959
	FLRTF	23.329	0.590	28.546	0.760	29.798	0.872
	HiNCoT	**33.053**	**0.943**	**33.824**	**0.954**	**35.064**	**0.962**
MSIs *Lemons* (256×256×31)	DCTNN	27.561	0.776	30.095	0.837	32.380	0.888
	TRLRF	15.874	0.058	19.021	0.193	24.713	0.586
	TNTV	33.692	0.938	37.077	0.964	39.168	0.974
	HLRTF	32.425	0.862	35.685	0.923	37.180	0.944
	DTR	33.102	0.886	34.795	0.906	36.851	0.953
	FLRTF	27.703	0.698	31.830	0.822	31.999	0.830
	HiNCoT	**36.646**	**0.951**	**38.190**	**0.965**	**40.155**	**0.976**
MSIs *Flowers* (256×256×31)	DCTNN	25.260	0.666	26.740	0.721	27.969	0.774
	TRLRF	19.301	0.141	25.086	0.550	26.665	0.635
	TNTV	28.709	0.830	30.852	0.878	31.975	0.902
	HLRTF	25.713	0.685	27.448	0.741	28.803	0.792
	DTR	30.253	0.836	31.295	0.866	33.586	0.920
	FLRTF	23.631	0.559	26.553	0.652	27.159	0.683
	HiNCoT	**30.681**	**0.875**	**32.380**	**0.903**	**33.808**	**0.927**
MSIs *Food* (256×256×31)	DCTNN	25.875	0.741	28.132	0.799	29.432	0.841
	TRLRF	17.433	0.098	24.919	0.542	26.979	0.658
	TNTV	28.527	0.862	30.933	0.913	32.306	0.934
	HLRTF	26.217	0.696	28.489	0.785	30.090	0.835
	DTR	30.270	0.855	31.408	0.881	33.077	0.921
	FLRTF	20.701	0.426	25.178	0.580	26.741	0.650
	HiNCoT	**31.070**	**0.872**	**33.173**	**0.916**	**34.983**	**0.944**

Table 2: The quantitative results of tube missing scenarios by different methods.

complex nonlinear structures, thereby contributing to its superior performance.

To provide a visual comparison of the recovery quality, we present the results of different methods in Fig. 3. As illustrated, HiNCoT effectively reconstructs multi-dimensional data and preserves fine details, thereby delivering the best visual performance across all three types of tasks.

Discussions

The Contribution of HNC-Transform Module

By leveraging the HNC-transform $f_\xi(\cdot)$, the proposed HiNCoT can more effectively capture the complex nonlinear relationships inherent in multi-dimensional data. Thus, $f_\xi(\cdot)$ serves as a pivotal component of the HiNCoT. To evaluate its contribution, we conduct an ablation study with two simplified variants. First, we set the number of transforms $l = 1$, degrading the hierarchical nonlinear transform to a single-layer linear transform. Second, we remove the activation functions in $f_\xi(\cdot)$, resulting in a hierarchical linear transform. We conduct experiments on the *Flowers* dataset under the random missing scenario. As shown in Table 4, both simplified variants exhibit inferior performance compared to the hierarchical nonlinear transform, demonstrating the importance of the hierarchical nonlinear transform in enhancing the model's representation capability.

To better understand the effect of the HNC-transform module, we investigate the impact of the depth of hierarchical nonlinear transform on the model performance. Specifically, we set l from 2 to 4 and evaluate the corresponding

Data	*Tomatoes*		*Lemons*		*Flowers*		*Food*	
Method	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
DCTNN	22.160	0.448	17.118	0.243	16.957	0.379	18.501	0.268
TRLRF	11.803	0.070	9.353	0.087	17.915	0.262	9.662	0.098
TNTV	34.323	0.964	39.608	0.978	33.849	0.924	33.846	0.946
HLRTF	22.206	0.509	17.177	0.333	17.220	0.434	18.552	0.343
DTR	20.924	0.446	17.118	0.244	16.385	0.370	18.501	0.265
FLRTF	49.607	0.990	48.685	**0.992**	47.517	0.991	43.442	0.985
HiNCoT	**50.902**	**0.996**	**49.389**	**0.992**	**48.105**	**0.993**	**45.742**	**0.989**

Table 3: The quantitative results of slice missing scenario by different methods.

results. As shown in Table 4, the proposed HiNCoT attains optimal performance when $l = 3$, followed by a slight degradation as l increases further. This suggests that choosing an appropriate value for l (e.g., $l = 3$) offers a favorable trade-off between model's performance and compactness.

The Synergistic Effect of Different Components

The proposed HiNCoT reconstructs the target tensor by applying the HNC-transform module to the latent tensor, which is generated by the deep representation module from the low-rank core tensor. The synergistic integration of the HNC-transform module, the deep representation module, and the low-rank core tensor endows HiNCoT with powerful representation capability. To investigate the necessity and synergy of the three key components in HiNCoT, we conduct ablation studies by individually removing each one.

Low-rank Core Tensor We compare the proposed HiN-

Sampling rate		0.02		0.05		0.10	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Single, linear	$l = 1$	31.248	0.854	35.734	0.938	42.972	0.984
Hierarchical, linear	$l = 2$	31.505	0.865	37.150	0.952	43.546	0.985
	$l = 3$	32.157	0.860	37.660	0.957	43.059	0.985
	$l = 4$	31.881	0.866	36.847	0.944	42.205	0.981
Hierarchical, nonlinear	$l = 2$	31.810	0.852	37.887	0.955	43.838	0.987
	$l = 3$	**32.338**	**0.880**	**38.232**	**0.963**	**45.188**	**0.990**
	$l = 4$	32.230	**0.880**	37.621	0.956	43.318	0.984

Table 4: The quantitative results with different depth of linear and nonlinear transform in $f_\xi(\cdot)$ under random missing scenarios for *Flowers* dataset

Sampling rate	0.02		0.05		0.10	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
HiNCoT w/o LR	31.628	0.872	37.521	0.952	43.448	0.986
HiNCoT w/o g_θ	29.097	0.763	36.251	0.936	41.480	0.977
HiNCoT w/o f_ξ	31.508	0.851	37.219	0.942	44.403	0.989
HiNCoT	**32.338**	**0.880**	**38.232**	**0.963**	**45.188**	**0.990**

Table 5: The quantitative comparison between HiNCoT and its variants under random missing scenarios for *Flowers* dataset.

CoT, which takes the low-rank core factor tensors $\mathcal{A} \in \mathbb{R}^{n_1 \times r \times b}$ and $\mathcal{B} \in \mathbb{R}^{r \times n_2 \times b}$ as inputs to $g_\theta(\cdot)$, with a variant that omits the low-rank constraint (LR) and directly uses a randomly initialized unconstrained tensor $\mathcal{Z} \in \mathbb{R}^{n_1 \times n_2 \times b}$ as input. As reported in Table 5, using an unconstrained input results in inferior performance, indicating that starting from the low-rank core tensor domain helps preserving the inherent low-rank structures of the core tensor, leading to more accurate reconstruction.

Deep Representation Module The deep representation module $g_\theta(\cdot)$ is an important component of HiNCoT, which represents the latent tensor via an untrained neural network. We ablate $g_\theta(\cdot)$ and directly use the low-rank core tensor as the latent tensor. As shown in Table 5, removing $g_\theta(\cdot)$ leads to a noticeable decline in performance, demonstrating that the deep representation module plays an indispensable role in effectively modeling multi-dimensional data.

HNC-Transform Module The HNC-transform module $f_\xi(\cdot)$ is essential for capturing the complex structures inherent in multi-dimensional data. To evaluate its contribution, we conduct an ablation study by removing $f_\xi(\cdot)$ and directly using the deep representation module to generate the target tensor. As shown in Table 5, removing $f_\xi(\cdot)$ leads to a noticeable performance drop, highlighting the importance of the HNC-transform module in the proposed HiNCoT.

These ablation results collectively demonstrate that the combination of the three components in HiNCoT yields a synergistic effect that is essential for achieving powerful representation capability.

The Influence of Parameters

In HiNCoT, the parameters r and b are two key factors that impact the model's performance and compactness. The parameter r determines the tensor tubal rank of the low-rank core tensor. We vary r from 1 to 30 for the random missing scenario with $\mathrm{SR} = 0.1$ on the *Flowers* dataset. As shown in Fig. 4(a), although the PSNR curve exhibits slight fluctuations, the overall trend increases initially and then declines after reaching the peak value. Therefore, it is essential to select an appropriate value of r based on the specific task.

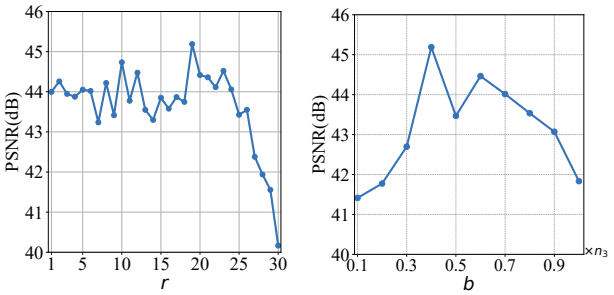

Figure 4: The PSNR curves with respect to the parameter b and r for random missing scenario with $\mathrm{SR} = 0.1$ on *Flowers* dataset.

The parameter b controls the size of the low-rank core tensor in the third dimension. Fig. 4(b) presents the PSNR values for the same setting, where b is set to $k\,n_3$ with $k \in \{0.1, 0.2, \ldots, 1.0\}$. As observed, the PSNR value initially increases with b but drops when b becomes too large. This indicates that selecting a moderate value of b is essential for achieving optimal performance.

Conclusion

In this paper, we proposed a hierarchical nonlinear continuous transform-based tensor representation method, i.e., HiNCoT. The hierarchical nonlinear continuous transform enables HiNCoT to capture the complex nonlinear relationships inherent in multi-dimensional data. Additionally, we leveraged a deep representation module that takes the low-rank core tensor as input to generate the latent tensor, effectively captures both intra-slice structures and inter-slice relationships. Furthermore, HiNCoT provides a continuous representation of multi-dimensional data, enabling it to handle a wide range of degradation tasks in real-world applications. To evaluate the representation capability of HiNCoT, we focused on representative multi-dimensional data recovery tasks, including random, tube, and slice missing scenarios, and proposed an unsupervised HiNCoT-based multi-dimensional data recovery model. Extensive experiments demonstrated that HiNCoT achieved state-of-the-art performance across various degradation scenarios.

Acknowledgements

This work was supported by National Key R&D Program of China (2025YFA1016400), Sichuan Science and Technology Program (No. 2024NSFSC0038, 2024NSFJQ0038) and NSFC (No. 12171072, 12371456).

References

Anandkumar, A.; Ge, R.; Hsu, D. J.; Kakade, S. M.; Telgarsky, M.; et al. 2014. Tensor decompositions for learning latent variable models. *Journal of Machine Learning Research*, 15(1): 2773–2832.

Braman, K. 2010. Third-order tensors as linear operators on a space of matrices. *Linear Algebra and its Applications*, 433(7): 1241–1253.

Deka, B.; Datta, S.; and Handique, S. 2018. Wavelet Tree Support Detection for Compressed Sensing MRI Reconstruction. *IEEE Signal Processing Letters*, 25(5): 730–734.

Fabian, Z.; Heckel, R.; and Soltanolkotabi, M. 2021. Data augmentation for deep learning based accelerated MRI reconstruction with limited data. In *International Conference on Machine Learning*, 3057–3067. PMLR.

Feng, L.; Zhu, C.; Long, Z.; Liu, J.; and Liu, Y. 2023. Multiplex Transformed Tensor Decomposition for Multidimensional Image Recovery. *IEEE Transactions on Image Processing*, 32: 3397–3412.

He, W.; Yao, Q.; Li, C.; Yokoya, N.; and Zhao, Q. 2019. Non-local meets global: An integrated paradigm for hyperspectral denoising. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 6868–6877.

Huynh-Thu, Q.; and Ghanbari, M. 2008. Scope of validity of PSNR in image/video quality assessment. *Electronics letters*, 44(13): 800–801.

Ji, T.-Y.; Zhao, X.-L.; and Sun, D.-L. 2022. Low-Rank Tensor Completion Method for Implicitly Low-Rank Visual Data. *IEEE Signal Processing Letters*, 29: 1162–1166.

Jia, F.; Ma, L.; Yang, Y.; and Zeng, T. 2021. Pixel-Attention CNN With Color Correlation Loss for Color Image Denoising. *IEEE Signal Processing Letters*, 28: 1600–1604.

Jiang, T.-X.; Ng, M. K.; Zhao, X.-L.; and Huang, T.-Z. 2020. Framelet representation of tensor nuclear norm for third-order tensor completion. *IEEE Transactions on Image Processing*, 29: 7233–7244.

Jiang, T.-X.; Zhao, X.-L.; Zhang, H.; and Ng, M. K. 2021. Dictionary learning with low-rank coding coefficients for tensor completion. *IEEE Transactions on Neural Networks and Learning Systems*, 34(2): 932–946.

Kilmer, M. E.; Braman, K.; Hao, N.; and Hoover, R. C. 2013. Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging. *SIAM Journal on Matrix Analysis and Applications*, 34(1): 148–172.

Kingma, D. P.; and Ba, J. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Kolda, T. G.; and Bader, B. W. 2009. Tensor decompositions and applications. *SIAM review*, 51(3): 455–500.

Li, B.-Z.; Zhao, X.-L.; Ji, T.-Y.; Zhang, X.-J.; and Huang, T.-Z. 2022. Nonlinear transform induced tensor nuclear norm for tensor completion. *Journal of Scientific Computing*, 92(3): 83.

Li, H.; Feng, R.; Wang, L.; Zhong, Y.; and Zhang, L. 2020. Superpixel-based reweighted low-rank and total variation sparse unmixing for hyperspectral remote sensing imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 59(1): 629–647.

Li, J.-Y.; Xie, J.-Y.; Luo, Y.-S.; Zhao, X.-L.; and Wang, J.-L. 2023. H2TF for hyperspectral image denoising: Where hierarchical nonlinear transform meets hierarchical matrix factorization. *IEEE Geoscience and Remote Sensing Letters*, 20: 1–5.

Lian, N.-X.; Zagorodnov, V.; and Tan, Y.-P. 2005. Color image denoising using wavelets and minimum cut analysis. *IEEE Signal Processing Letters*, 12(11): 741–744.

Lin, L.; Wang, J.; Wei, G.; Wang, M.; and Zhang, A. 2025. FFSTIE: Video Restoration With Full-Frequency Spatio-Temporal Information Enhancement. *IEEE Signal Processing Letters*, 32: 571–575.

Lu, C.; Peng, X.; and Wei, Y. 2019. 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*, 5996–6004.

Luo, Y.; Zhao, X.; Li, Z.; Ng, M. K.; and Meng, D. 2023. Low-rank tensor function representation for multi-dimensional data recovery. *IEEE transactions on pattern analysis and machine intelligence*, 46(5): 3351–3369.

Luo, Y.; Zhao, X.-L.; Meng, D.; and Jiang, T.-X. 2022. HLRTF: Hierarchical low-rank tensor factorization for inverse problems in multi-dimensional imaging. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 19303–19312.

Madathil, B.; and George, S. N. 2018. DCT based weighted adaptive multi-linear data completion and denoising. *Neurocomputing*, 318: 120–136.

Mescheder, L.; Oechsle, M.; Niemeyer, M.; Nowozin, S.; and Geiger, A. 2019. Occupancy networks: Learning 3d reconstruction in function space. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 4460–4470.

Ono, S. 2017. Primal-Dual Plug-and-Play Image Restoration. *IEEE Signal Processing Letters*, 24(8): 1108–1112.

Park, J. J.; Florence, P.; Straub, J.; Newcombe, R.; and Lovegrove, S. 2019. Deepsdf: Learning continuous signed distance functions for shape representation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 165–174.

Qin, W.; Wang, H.; Zhang, F.; Wang, J.; Luo, X.; and Huang, T. 2022. Low-Rank High-Order Tensor Completion With Applications in Visual Data. *IEEE Transactions on Image Processing*, 31: 2433–2448.

Qiu, D.; Bai, M.; Ng, M. K.; and Zhang, X. 2021. Robust low-rank tensor completion via transformed tensor nuclear norm with total variation regularization. *Neurocomputing*, 435: 197–215.

Sitzmann, V.; Martel, J.; Bergman, A.; Lindell, D.; and Wetzstein, G. 2020. Implicit neural representations with periodic activation functions. *Advances in neural information processing systems*, 33: 7462–7473.

Wang, H.; Peng, J.; Qin, W.; Wang, J.; and Meng, D. 2023a. Guaranteed Tensor Recovery Fused Low-rankness and Smoothness. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(9): 10990–11007.

Wang, J.; and Zhao, X. 2024. Functional Transform-Based Low-Rank Tensor Factorization for Multi-dimensional Data Recovery. In *European Conference on Computer Vision*, 39–56. Springer.

Wang, J.-L.; Huang, T.-Z.; Zhao, X.-L.; and Miao, Y.-C. 2023b. Complex video completion fusing low-rank background and deep foreground priors. *IEEE Signal Processing Letters*, 29: 2737–2741.

Wang, X.-Y.; Li, X.-P.; Sidiropoulos, N. D.; and So, H. C. 2025. Tensor Completion Network for Visual Data. *IEEE Transactions on Signal Processing*.

Wang, Z.; Bovik, A. C.; Sheikh, H. R.; and Simoncelli, E. P. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4): 600–612.

Wu, W.-H.; Huang, T.-Z.; Zhang, H.; Wang, J.-L.; and Zhao, X.-L. 2023. Untrained low-rank neural network prior for multi-dimensional image recovery. *IEEE Signal Processing Letters*, 30: 1647–1651.

Yuan, L.; Li, C.; Mandic, D.; Cao, J.; and Zhao, Q. 2019. Tensor ring decomposition with rank minimization on latent space: An efficient approach for tensor completion. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, 9151–9158.

Zhao, J.-L.; Gao, J.-F.; Fang, S.; Zhang, T.-H.; and Wang, J.-Y. 2025. Robust Tensor Completion via Spatial-Spectral Constrained Deep Low-Rank Tensor Factorization for Hyperspectral Image Recovery. *IEEE Signal Processing Letters*, 32: 551–555.

Zhou, T.-W.; Zhao, X.-L.; Wang, J.-L.; Luo, Y.-S.; Wang, M.; Bai, X.-X.; and Yan, H. 2024. DTR: A Unified Deep Tensor Representation Framework for Multimedia Data Recovery. *arXiv preprint arXiv:2407.05267*.