

Flexible Generalized Low-Rank Regularizer for Tensor RPCA

Anonymous submission

Abstract

Tensor Robust Principal Component Analysis (TRPCA) has emerged as a powerful technique for tensor recovery. In this paper, we design a novel Flexible Generalized low-rank regularizer termed as FGTNN. Equipped with this, we develop an FGTRPCA framework, which has two desirable properties.

1) generalizability: Many existing TRPCA methods can be viewed as special cases of our framework; **2) flexibility:** Using FGTRPCA as a general platform, we can derive a series of new TRPCA methods by tuning a continuous parameter to improve performance. This endows our model with the capability to adapt to intricate scenarios. Moreover, considering the low-rankness and smoothness priors simultaneously, we use FGTNN to investigate the inherent structural characteristics of gradient tensors and propose another novel regularizer, then leading to the development of the smooth FGTRPCA (SFGTRPCA) model. Besides, we device efficient optimization algorithms based on the Alternating Direction Method of Multipliers (ADMM) framework to implement the proposed models. Experimental results on various denoising and recovery tasks demonstrate the superiority of our models.

Introduction

Tensor data are ubiquitous, many real-world data are usually inherently multidimensional, with information stored in multi-way arrays known as tensors, e.g., images, videos, network flow data, etc. In recent years, significant advancements across various interdisciplinary domains have been made in tensor analysis, such as machine learning (Wen, Chen, and Chen 2024; Phothilimthana et al. 2024), data mining (Zhang et al. 2023a; Huang et al. 2024), and computer vision (Zhao et al. 2024; Liu et al. 2024a). However, due to the inherent limitations of signal acquisition equipment, including sensor sensitivity, photon effects, and calibration errors, tensor data gathered from real-world environments frequently suffer from substantial corruption (Wang et al. 2023a). Consequently, tensor recovery has become a crucial task in tensor analysis.

This paper concentrates on the problem of Tensor Robust Principal Component Analysis (TRPCA) (Huang et al. 2015), which seeks to recover the underlying low-rank tensor \mathcal{L} and sparse tensor \mathcal{E} from their sum \mathcal{M} (see Figure 1 for a visual representation) and solves the following problem

$$\min_{\mathcal{L}, \mathcal{E} \in \mathbb{R}^{d_1 \times d_2 \times d_3}} \text{rank}(\mathcal{L}) + \lambda \|\mathcal{E}\|_1 \text{ s.t. } \mathcal{M} = \mathcal{L} + \mathcal{E}, \quad (1)$$

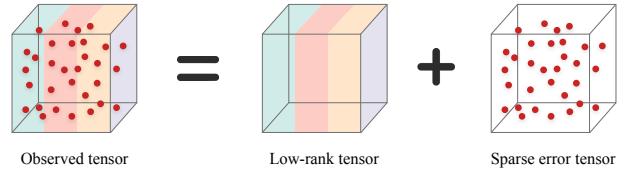


Figure 1: An illustration of TRPCA, which demonstrates the decomposition of observation tensor into low-rank and sparse components.

where $\lambda > 0$ is a regularization parameter, $\text{rank}(\mathcal{L})$ denotes the rank of clean tensor \mathcal{L} and $\|\mathcal{E}\|_1$ is ℓ_1 -norm (sum of the absolute values of all the entries) to measure the sparsity of the noise tensor \mathcal{E} . A key challenge is the definition of tensor rank, which is inherently more complex than matrix rank. Various conventional methods for defining tensor rank originate from distinct tensor decompositions. For instance, inspired by the tensor singular value decomposition (t-SVD), (Kilmer et al. 2013) proposed the tensor tubal rank that can be efficiently computed using the fast Fourier Transform (FFT). Since the non-convexity and discontinuity of the rank function, solving the problem (1) is usually NP-hard. Consequently, (Lu et al. 2020) proposed a novel tensor nuclear norm as a convex approximation to the tensor tubal rank and proposed a new TRPCA method defined as follows

$$\min_{\mathcal{L}, \mathcal{E}} \|\mathcal{L}\|_* + \lambda \|\mathcal{E}\|_1 \text{ s.t. } \mathcal{M} = \mathcal{L} + \mathcal{E}, \quad (2)$$

where $\|\cdot\|_*$ represents the tensor nuclear norm (TNN). Furthermore, recent research by (Kilmer et al. 2021) has demonstrated the optimal representation and compression capabilities of t-SVD, further highlighting the significance of model (2) in capturing the intrinsic low-rank structures of tensors. As a result, the model (2) under t-SVD has garnered considerable interests recently (Hou et al. 2024; Liu et al. 2024c; Qin et al. 2024).

Despite the impressive performance of TRPCA, it still exhibits several limitations. Specifically, when minimizing the TNN, TRPCA employs tensor singular value thresholding to uniformly diminish all singular values. In real-world applications, singular values often carry distinct physical meanings, supported by prior knowledge indicating that larger singular values are generally associated with more signifi-

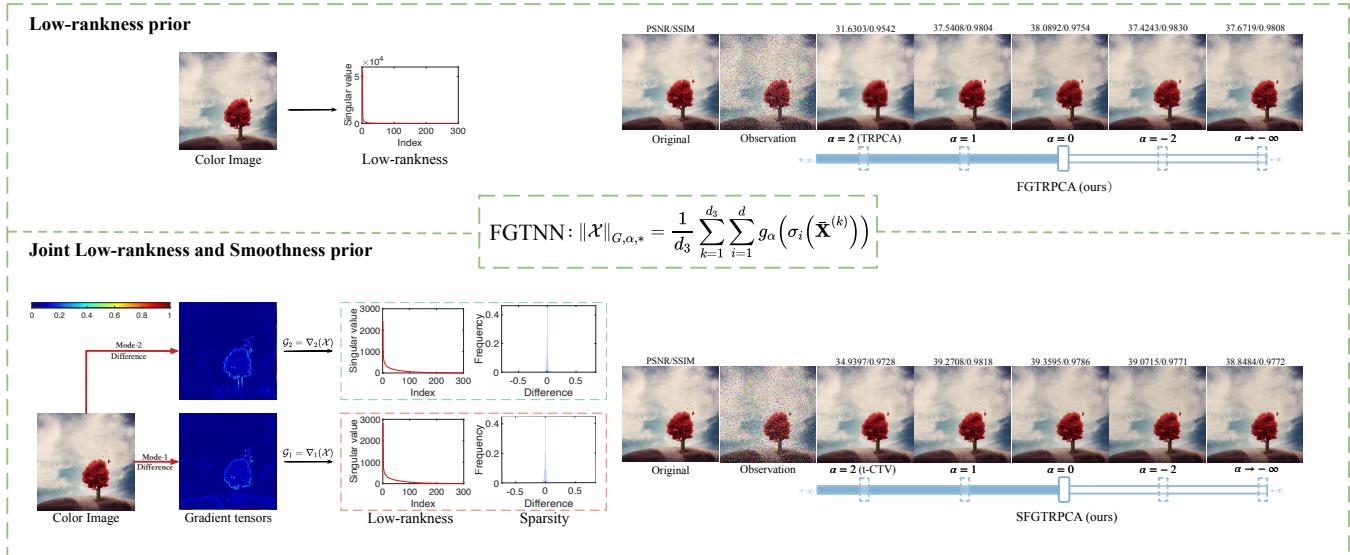


Figure 2: Take a color image sample from ZJU (Hu et al. 2012) dataset as an example. The two frames illustrate the recovery performance of our proposed FGTRPCA and SFGTRPCA with different values of shape parameter α (see Eq. (5)) under different structure priors of color images.

cant information. The uniform shrinkage approach of TRPCA fails to account for these differences among singular values, potentially leading to suboptimal results.

While many existing advanced methodologies (Gao et al. 2020; Jiang et al. 2020; Wang et al. 2023b; Zhang et al. 2023b; Yan and Guo 2024; Liu et al. 2024b) develop various TNN-based low-rank regularizers that penalize large singular values less and small singular values more, thereby efficiently preserving essential information and filtering out irrelevant details, their discrete and fixed models make it unflexible to diverse scenarios. In this paper, we design a novel Flexible Generalized low-rank regularizer (FGTNN) to adaptively assign different penalties to distinct singular values and impose the constraint on the sparse component. We have shown that several existing TRPCA models can be reformulated as special cases of FGTRPCA. Apart from that, we can also derive a wider family of new TRPCA models by tuning a continuous parameter to improve performance. Through this, our model significantly improves flexibility and efficiency in complex situations.

In addition to the low-rankness prior, the smoothness prior, frequently modeled by total variation (TV), is widely utilized in tensor recovery applications (Ko et al. 2020; Qiu et al. 2021). This prior states how similar objects/scenes (with shapes) are adjacently distributed (Peng et al. 2022b). Most previous works encoded the two priors with two independent regularizers and incorporated them into a unified model, which achieved better performance (Peng et al. 2020, 2022a). However, they have two drawbacks: (1) it is challenging to fine-tune the regularization parameter between the two terms; (2) the theoretical guarantee for exact recovery remains unproven for the related methods.

Given the circumstances above, (Wang et al. 2023a) proposed the tensor Correlated Total Variation (t-CTV) norm

which integrates the two priors into a single regularization term, eliminating the need for tuning separate parameters. Moreover, this work offered theoretical guarantees for the precise recovery of analogous tensor methods that concurrently model both priors. Analogously, the integration regularization term was also based on TNN in the gradient domain. Consequently, (Huang et al. 2024) proposed a reweighted regularizer based on ℓ_p norm as a surrogate for t-CTV term. In this paper, we employ FGTNN to explore the inherent structural properties of gradient tensors and introduce a new tensor-correlated Flexible Generalized Joint Prior (t-FGJP) regularizer.

Our principal contributions are outlined as follows:

- We propose a flexible generalized low-rank regularizer (FGTNN) that accounts for the varying importance of different singular values in low-rank tensors and develop a novel FGTRPCA framework. The FGTRPCA, not only regards many existing TRPCA methods as special cases, but also opens a door to design a broad new family of TRPCA methods by tuning a continuous parameter. This enhances the flexibility of our model to counter more intricate scenarios and improve performance.
- Considering the low-rankness and smoothness priors simultaneously, we propose a novel tensor-correlated Flexible Generalized Joint Prior (t-FGJP) regularizer based on FTGNN. The proposed t-FGJP also maintains the flexibility of discriminatively controlling different singular values of the gradient tensors and is applied to the smooth FGTRPCA (SFGTRPCA) model.
- We design ADMM-based (Boyd et al. 2011) algorithmic frameworks tailored for each of the aforementioned models. Our extensive experiments on various tensor denoising and recovery tasks demonstrate the advantages of our models.

Notations and preliminaries

First, we present some key notations and definitions used throughout the paper. We represent scalars, vectors, and matrices using lowercase letters, boldface lowercase letters, and boldface uppercase letters, e.g., x , \mathbf{x} , \mathbf{X} , respectively. Tensors are presented by bold calligraphic letters, e.g., \mathcal{X} . $\mathbf{1}_{d_1 \times d_2}$ and $\mathbf{1}_{d_1 \times d_2 \times d_3}$ represent a matrix of size $d_1 \times d_2$ and a tensor of size $d_1 \times d_2 \times d_3$ with all entries as ones. For a 3-order tensor $\mathcal{X} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$, we denote \mathcal{X}_{ijk} as its (i, j, k) -th entry, $\mathcal{X}(i, :, :)$ as its horizontal slice, $\mathcal{X}(:, j, :)$ as its lateral slice, $\mathcal{X}(:, :, k)$ as its frontal slice, respectively. For convenience, the frontal slice $\mathcal{X}(:, :, k)$ is often denoted as $\mathbf{X}^{(k)}$. The tensor nuclear norm (TNN), tensor ℓ_1 norm (TL1N), tensor Frobenius norm and tensor infinity norm of \mathcal{X} are defined by $\|\mathcal{X}\|_*$, $\|\mathcal{X}\|_1 = |\mathcal{X}_{ijk}|$, $\|\mathcal{X}\|_F = \sqrt{\sum_{ijk} |\mathcal{X}_{ijk}|^2}$ and $\|\mathcal{X}\|_\infty = \max_{ijk} |\mathcal{X}_{ijk}|$, respectively. The transpose of \mathcal{X} is defined as $\mathcal{X}^T \in \mathbb{R}^{d_2 \times d_1 \times d_3}$ (Lu et al. 2020).

Definition 1. (T-SVD) (Kilmer and Martin 2011) For a tensor $\mathcal{X} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$, it can be factorized by t-SVD as

$$\mathcal{X} = \mathcal{U} * \mathcal{S} * \mathcal{V}^T, \quad (3)$$

where $\mathcal{U} \in \mathbb{R}^{d_1 \times d_1 \times d_3}$, $\mathcal{V} \in \mathbb{R}^{d_2 \times d_2 \times d_3}$ are orthogonal tensors, i.e., $\mathcal{U} * \mathcal{U}^T = \mathcal{U}^T * \mathcal{U} = \mathcal{V} * \mathcal{V}^T = \mathcal{V}^T * \mathcal{V} = \mathcal{I}$, and $\mathcal{S} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ is an f-diagonal tensor, i.e., its frontal slices are the diagonal matrices, and “*” is the t-product.

Definition 2. (Tensor Nuclear Norm, TNN) (Lu et al. 2020) For $\mathcal{X} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ and $d = \min(d_1, d_2)$, the Tensor Nuclear Norm of \mathcal{X} is defined as

$$\|\mathcal{X}\|_* = \frac{1}{d_3} \sum_{k=1}^{d_3} \sum_{i=1}^d \sigma_i(\bar{\mathbf{X}}^{(k)}), \quad (4)$$

where $\bar{\mathbf{X}}$ is the result by applying FFT on \mathcal{X} along the third dimension, i.e., $\bar{\mathbf{X}} = \text{fft}(\mathcal{X}, [], 3)$. $\bar{\mathbf{X}}^{(k)}$ is the k -th slice of $\bar{\mathbf{X}}$, $\sigma_i(\bar{\mathbf{X}}^{(k)})$ is the i -th singular value of $\bar{\mathbf{X}}^{(k)}$, and $d = \min(d_1, d_2)$.

Proposed methods

We begin by detailing the motivation for FGTNN and its characteristics, then propose the FGTRPCA framework and implement it using an efficient optimization algorithm. Furthermore, by simultaneously considering low-rankness and smoothness priors, we introduce the t-FGJP regularizer and apply it to the SFGTRPCA model

Flexible Generalized TNN

According to Definition 2, the original TNN uniformly shrinks each singular value of the low-rank tensor \mathcal{L} when minimizing the tensor nuclear norm. Indeed, larger singular values are associated with more critical information within the tensor. Inspired by the common purposes of enhancing TRPCA methods, we introduce a flexible generalized tensor nuclear norm (FGTNN) as a unified framework, which is defined below.

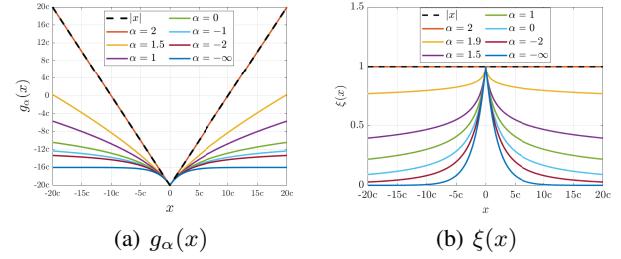


Figure 3: Our proposed $g_\alpha(x)$ and its corresponding weight function $\xi(x)$.

Definition 3. (FGTNN) For a tensor $\mathcal{X} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$, $d = \min(d_1, d_2)$, the Flexible Generalized Tensor Nuclear Norm (FGTNN) is defined as follows

$$\|\mathcal{X}\|_{G,\alpha,*} = \frac{1}{d_3} \sum_{k=1}^{d_3} \sum_{i=1}^d g_\alpha\left(\sigma_i\left(\bar{\mathbf{X}}^{(k)}\right)\right), \quad (5)$$

where $g_\alpha(x)$ is

$$g_\alpha(x) = 2c \cdot \frac{|\alpha - 2|}{\alpha} \left(\left(\frac{|x|/c}{|\alpha - 2|} + 1 \right)^{\alpha/2} - 1 \right), \quad (6)$$

where $\alpha \in \mathbb{R}$ is a continuous parameter that controls the shape of $g_\alpha(x)$ and $c > 0$.

Remark 1. Our proposed FGTNN mainly exhibits two desirable properties. **1) generalizability:** By introducing a continuous parameter α , low-rank regularizers in many existing popular methods such as TRPCA, LRTF, ETR, and DATRPCA can be viewed as special cases of FGTNN with different values of α . (see Table 1 for more details); **2) flexibility:** We can develop plenty of new low-rank regularizers by tuning α and achieve better performance. Compared to the method with fixed-form low-rank regularizer, our model gains flexibility and can adapt to more complex scenarios. Apart from that, $g_\alpha(x)$ in FGTNN controls the penalty strength to singular values. Figure 3(a) intuitively presents the characteristics of $g_\alpha(x)$. We observe that $g_\alpha(x)$ increases slower than $|x|$ for various α , which means less shrunk to large singular values, preserving the critical information within the tensor to a greater extent. More importantly, α is related to the shape of $g_\alpha(x)$. When $\alpha \rightarrow -\infty$, $g_\alpha(x)$ follows an approximately exponential form; When $\alpha = 0$, $g_\alpha(x)$ takes a logarithmic form; When $\alpha = 2$, $g_\alpha(x)$ turns to $|x|$; And in the other case of α , $g_\alpha(x)$ is represented in a approximate power form.

Moreover, we extend FGTNN to the sparse component and define the flexible generalized tensor ℓ_1 norm (FGTL1N).

Definition 4. (FGTL1N) For a tensor $\mathcal{X} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$, the Flexible Generalized Tensor ℓ_1 Norm (FGTL1N) is defined as follows:

$$\|\mathcal{X}\|_{G,\alpha,1} = \sum_{i=1}^{d_1} \sum_{j=1}^{d_2} \sum_{k=1}^{d_3} g_\alpha(|\mathcal{X}_{ijk}|), \quad (7)$$

Author and year	Method	Value of α	$g_\alpha(x)$	Low-rank Regularizer
(Lu et al. 2020)	TRPCA	$\alpha = 2$	$ x $	$\ \mathcal{X}\ _{G,2,*}$
(Chen et al. 2021)	LRTF	$\alpha = 0$	$2c \ln\left(\frac{1}{2} x /c + 1\right)$	$\ \mathcal{X}\ _{G,0,*}$
(Ji and Feng 2023)	ETR	$\alpha = -2$	$2c \frac{2 x /c}{ x /c+4}$	$\ \mathcal{X}\ _{G,-2,*}$
(Wang et al. 2023b)	DATRPCA	$\alpha \rightarrow -\infty$	$2c(1 - \exp(- x /2c))$	$\ \mathcal{X}\ _{G,-\infty,*}$

Table 1: The FGTNN regularizer view for many special cases.

Flexible Generalized TRPCA

By integrating FGTNN and FGTL1N into the TRPCA framework, we formulate the following Flexible Generalized TRPCA (FGTRPCA) model.

$$\min_{\mathcal{L}, \mathcal{E} \in \mathbb{R}^{d_1 \times d_2 \times d_3}} \|\mathcal{L}\|_{G,\alpha,*} + \lambda \|\mathcal{E}\|_{G,\alpha,1} \text{ s.t. } \mathcal{M} = \mathcal{L} + \mathcal{E}. \quad (8)$$

Note that FGTNN includes a series of specific functions that are nonlinear and complex, thus making it hard to obtain the optimal solution of the FGTRPCA model. In this paper, we design an efficient algorithm optimization framework based on the ADMM framework (Boyd et al. 2011) to implement the FGTRPCA model.

Proposition 1. For $g_\alpha(x)$, there exists a convex conjugate function $\phi : \mathbb{R} \rightarrow \mathbb{R}$ which satisfied

$$g_\alpha(x) = \min_{w \in \mathbb{R}_+} (w|x| + \phi(w)), \quad (9)$$

and for fixed x , the minimum is reached at $w = \xi(x)$, which is defined as

$$w = \xi(x) = \begin{cases} 1, & \text{if } \alpha = 2 \\ 2c/(|x| + 2c), & \text{if } \alpha = 0 \\ \exp(-|x|/2c), & \text{if } \alpha = -\infty \\ \left(\frac{|x|/c}{|\alpha-2|} + 1\right)^{\alpha/2-1}, & \text{otherwise.} \end{cases} \quad (10)$$

Remark 2. According to Proposition 1, $g_\alpha(x)$ in Eq. (6) can be optimized by an adaptive alternating weighted minimization scheme. From the perspective of weights, smaller weights represent smaller shrinkages to singular values. As shown in Figure 3(b), TNN assigns equal weight to each singular value, i.e., TNN treats each singular value equally. For our proposed FGTNN, larger singular values will adaptively receive smaller weights, resulting in less shrinkage.

According to Proposition 1, FGTNN can be transformed into

$$\|\mathcal{L}\|_{G,\alpha,*} = \min_{\mathbf{W}} \frac{1}{d_3} \sum_{k=1}^{d_3} \sum_{i=1}^d (W_{ki} \sigma_i(\bar{\mathbf{L}}^{(k)}) + \phi(W_{ki})), \quad (11)$$

where the W_{ki} is the k, i -th element of matrix $\mathbf{W} \in \mathbb{R}^{d_3 \times d}$. The minimum is reached at $W_{ki} = \xi(\sigma_i(\bar{\mathbf{L}}^{(k)}); c)$. Similarly, as for FGTL1N, we have

$$\|\mathcal{E}\|_{G,\alpha,1} = \min_{\mathbf{W}} \sum_{i=1}^{d_1} \sum_{j=1}^{d_2} \sum_{k=1}^{d_3} (\mathcal{W}_{ijk} |\mathcal{E}_{ijk}| + \phi(\mathcal{W}_{ijk})), \quad (12)$$

where the \mathcal{W}_{ijk} is the i, j, k -th element of tensor $\mathbf{W} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$. The minimum is reached at $\mathcal{W}_{ijk} = \xi(|\mathcal{E}_{ijk}|; c)$.

Notably, problem (8) can be reformulated as the weighted tensor nuclear norm minimization problem (11) and the weighted tensor ℓ_1 norm minimization problem (12). The following defines two key concepts: Weighted tensor nuclear norm (WTNN) and weighted tensor ℓ_1 norm (WTLIN).

Definition 5. (Weighted Tensor Nuclear Norm, WTNN) (Wang et al. 2023b) For a tensor $\mathcal{X} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ and a weight matrix $\mathbf{W} \in \mathbb{R}^{d \times d_3}$, $d = \min(d_1, d_2)$, the WTNN of \mathcal{X} is defined as

$$\|\mathcal{X}\|_{\mathbf{W},*} = \frac{1}{d_3} \sum_{k=1}^{d_3} \sum_{i=1}^d W_{ik} \sigma_i(\bar{\mathbf{X}}^{(k)}). \quad (13)$$

Definition 6. (Weighted Tensor ℓ_1 Norm, WTLIN) (Wang et al. 2023b) For a tensor $\mathcal{X} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ and a weight tensor $\mathbf{W} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$, the WTLIN of \mathcal{X} is defined as

$$\|\mathcal{X}\|_{\mathbf{W},1} = \sum_{i=1}^{d_1} \sum_{j=1}^{d_2} \sum_{k=1}^{d_3} |\mathcal{W}_{ijk} \mathcal{X}_{ijk}|. \quad (14)$$

By incorporating Eq. (11) and Eq. (12) into model (8), and according to the definition of WTNN and WTLIN, we have

$$\begin{aligned} \min_{\mathcal{L}, \mathcal{E}, \mathbf{W}, \mathbf{W}} & \|\mathcal{L}\|_{\mathbf{W},*} + \lambda \|\mathcal{E}\|_{\mathbf{W},1} + \Phi_M(\mathbf{W}) + \Phi_T(\mathbf{W}) \\ \text{s.t. } & \mathcal{M} = \mathcal{L} + \mathcal{E}, \end{aligned} \quad (15)$$

where $\Phi_M(\mathbf{W})$ and $\Phi_T(\mathbf{W})$ are defined such that $(\Phi_M(\mathbf{W}))_{ki} = \phi(W_{ki})$ and $(\Phi_T(\mathbf{W}))_{ijk} = \phi(\mathcal{W}_{ijk})$. In the next part, we will present the optimization for implementing FGTRPCA.

Optimization for FGTRPCA

The Lagrangian function of the FGTRPCA model is

$$\begin{aligned} L(\mathcal{L}, \mathcal{E}, \mathbf{W}, \mathbf{W}, \mathcal{Z}, \mu) = & \|\mathcal{L}\|_{\mathbf{W},*} + \lambda \|\mathcal{E}\|_{\mathbf{W},1} + \Phi_M(\mathbf{W}) \\ & + \Phi_T(\mathbf{W}) + \frac{\mu}{2} \left\| \mathcal{L} + \mathcal{E} - \mathcal{M} + \frac{\mathcal{Z}}{\mu} \right\|_F^2 - \frac{\mu}{2} \|\mathcal{Z}/\mu\|_F^2, \end{aligned} \quad (16)$$

where $\mathcal{Z} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ denotes the Lagrangian multiplier and μ is a positive parameter. Each variable can be updated alternately in the scheme of the ADMM framework.

Step1: Update \mathcal{L} by fixing the other variables:

$$\mathcal{L}_{t+1} = \arg \min_{\mathcal{L}} \frac{1}{\mu_t} \|\mathcal{L}\|_{\mathbf{W},*} + \frac{1}{2} \|\mathcal{L} - (\mathcal{M} - \mathcal{E}_t - \mathcal{Z}_t / \mu_t)\|_F^2. \quad (17)$$

The closed-form solution of (17) can be easily obtained with the following proximity operator.

Lemma 1. Given $\mathcal{X} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ with t-SVD $\mathcal{X} = \mathcal{U} * \mathcal{S} * \mathcal{V}^*$ and a weight matrix $\mathbf{W} \in \mathbb{R}^{d \times d_3}$, where \mathbf{w}_k is the k -th column of \mathbf{W} and $d = \min\{d_1, d_2\}$. Considering the following Weighted Tensor Nuclear Norm minimization (WTNNM) problem

$$\text{Prox}_{\|\cdot\|_{\mathbf{W},*}}(\mathcal{X}) = \arg \min_{\mathcal{L}} \frac{1}{2} \|\mathcal{L} - \mathcal{X}\|_F^2 + \|\mathcal{L}\|_{\mathbf{W},*}, \quad (18)$$

where $\|\cdot\|_{\mathbf{W},*}$ denotes the WTNN, and $\text{Prox}_{\|\cdot\|_{\mathbf{W},*}}$ is defined as a proximal operator. For non-descending weights $0 \leq W_{1k} \leq W_{2k} \leq \dots \leq W_{dk}$ ($k = 1, \dots, d_3$), the problem (18) has the global solution which is defined as

$$\mathcal{L}^* = \text{Prox}_{\|\cdot\|_{\mathbf{W},*}}(\mathcal{X}) = \mathcal{U} * \text{ifft}(\mathcal{P}_{\mathbf{W}}(\bar{\mathcal{S}}), [], 3) * \mathcal{V}^*, \quad (19)$$

where $\mathcal{P}_{\mathbf{W}}(\bar{\mathcal{S}})$ is a tensor to meet the conditions of its k -th frontal slice is $\mathcal{P}_{\mathbf{w}_k}(\bar{\mathcal{S}}^{(k)})$ for $k = 1, \dots, d_3$. $\bar{\mathcal{S}}^{(k)}$ is the k -th frontal slice of $\bar{\mathcal{S}}$, and $\mathcal{P}_{\mathbf{w}_k}(\bar{\mathcal{S}}^{(k)})$ denotes a diagonal matrix which can be computed as $(\mathcal{P}_{\mathbf{w}_k}(\bar{\mathcal{S}}^{(k)}))_{ii} = (\bar{\mathcal{S}}_{ii}^{(k)} - w_{ki})_+$, where $(x)_+ = x$ if $x > 0$ and $(x)_+ = 0$ otherwise. w_{ki} is the i -th element of the \mathbf{w}_k .

By recalling the definition of WTNN in Definition 5, we have $\frac{1}{\mu_t} \|\mathcal{L}\|_{\mathbf{W},*} = \|\mathcal{L}\|_{\frac{1}{\mu_t} \mathbf{W},*}$. Based on Lemma 1, the solution of the subproblem (17) can be described as

$$\mathcal{L}_{t+1} = \text{Prox}_{\|\cdot\|_{\frac{1}{\mu_t} \mathbf{W},*}}(\mathcal{M} - \mathcal{E}_t - \mathcal{Z}_t / \mu_t). \quad (20)$$

Step2: Update \mathcal{E} by fixing other variables:

$$\mathcal{E}_{t+1} = \arg \min_{\mathcal{E}} \frac{\lambda}{\mu_t} \|\mathcal{E}\|_{\mathbf{W},1} + \frac{1}{2} \|\mathcal{E} - (\mathcal{M} - \mathcal{L}_{t+1} - \mathcal{Z}_t / \mu_t)\|_F^2. \quad (21)$$

To get the closed-form solution of the above problem, we utilize the tensor soft-thresholding operator (TST) defined below to update \mathcal{E}_{t+1} .

$$\mathcal{E}_{t+1} = \text{TST}(\mathcal{M} - \mathcal{L}_{t+1} - \mathcal{Z}_t / \mu_t, \frac{\lambda}{\mu_t} \mathbf{W}_t), \quad (22)$$

where the ijk -th entry of TST is defined by

$$(\text{TST}(\mathcal{X}, \mathbf{W}))_{ijk} = \text{sign}(\mathcal{X}_{ijk})(|\mathcal{X}_{ijk}| - \mathbf{W}_{ijk})_+. \quad (23)$$

Step3: Update the elements of \mathbf{W} and \mathbf{W} by an adaptive way according to Proposition 1

$$W_{(t+1)ki} = \xi(\sigma_i(\bar{\mathbf{L}}_{t+1}^{(k)}; c), \mathbf{W}_{(t+1)ijk} = \xi(|(\mathcal{E}_{t+1})_{ijk}|; c). \quad (24)$$

Step4: Update the Lagrangian multiplier tensor \mathcal{Z} and the parameter μ by

$$\mathcal{Z}_{t+1} = \mathcal{Z}_t + \mu_t (\mathcal{L}_{t+1} + \mathcal{E}_{t+1} - \mathcal{M}), \quad (25)$$

$$\mu_{t+1} = \rho \mu_t, \quad (26)$$

Algorithm 1: FGTRPCA algorithm

Input: Observation tensor data $\mathcal{M} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$, and the parameter λ .

- 1: Initialize $\mathcal{L}_0 = \mathcal{E}_0 = \mathcal{Z}_0 = 0$, $\mathbf{W}_0 = \mathbf{1}_{d_3 \times d}$, $\mathbf{W}_0 = \mathbf{1}_{d_1 \times d_2 \times d_3}$, $\mu_0 = 10^{-2}$, $\rho = 1.1$, $\epsilon = 10^{-6}$, and $t = 0$.
- 2: **while** not converge **do**
- 3: Update the low-rank tensor \mathcal{L} by Eq. (17).
- 4: Update the sparse tensor \mathcal{E} by Eq. (21).
- 5: Update the weights \mathbf{W} and \mathbf{W} by Eq. (24).
- 6: Update the Lagrangian multiplier \mathcal{Z} by Eq. (25).
- 7: Update the parameter μ by Eq. (26).
- 8: Check the convergence condition in Eq. (27).

9: **end while**

Output: $\mathcal{L} = \mathcal{L}_{t+1}$, $\mathcal{E} = \mathcal{E}_{t+1}$

where $\rho = 1.1$. The convergence conditions are defined as

$$\left\{ \begin{array}{l} \|\mathcal{L}_{t+1} - \mathcal{L}_t\|_\infty \\ \|\mathcal{E}_{t+1} - \mathcal{E}_t\|_\infty \\ \|\mathcal{M} - \mathcal{L}_{t+1} - \mathcal{E}_{t+1}\|_\infty \end{array} \right\} \leq \epsilon. \quad (27)$$

The whole optimization procedure is summarized in Algorithm 1.

Smooth FGTRPCA

Considering a structured tensor that exhibits both low-rankness and smoothness, we devise a novel regularizer that aims to represent both two properties simultaneously on the gradient tensors, instead of employing a combination of two distinct regularizers for encoding the two properties. We first introduce the definition of the gradient tensor and present our proposed tensor-correlated Flexible Generalized Joint Prior (t-FGJP) regularizer.

Definition 7. (*Gradient tensor*) (Wang et al. 2023a) For a tensor $\mathcal{X} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$, its gradient tensor along the k -th mode is defined as

$$\mathcal{G}_k := \nabla_k(\mathcal{X}) = \mathcal{X} \times_k \mathbf{D}_{n_k}, k = 1, 2, 3, \quad (28)$$

where \mathbf{D}_{n_k} is a row circulant matrix of $(-1, 1, 0, \dots, 0)$.

Definition 8. (t-FGJP) For a tensor $\mathcal{X} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$, the proposed t-FGJP norm is defined as

$$\|\mathcal{X}\|_{\text{t-FGJP}} := \frac{1}{\gamma} \sum_{k \in \Gamma} \|\mathcal{G}_k\|_{G,\alpha,*}, \quad (29)$$

where Γ represents a priori set of directions along which \mathcal{X} equips both low-rankness and smoothness priors and $\gamma := \#\{\Gamma\}$ denotes the cardinality of Γ . By incorporating both t-FGJP and FGTL1N into the TRPCA framework, we propose a smooth FGTRPCA (SFGTRPCA) model defined as

$$\min_{\mathcal{L}, \mathcal{E}} \|\mathcal{L}\|_{\text{t-FGJP}} + \lambda \|\mathcal{E}\|_{G,\alpha,1} \text{ s.t. } \mathcal{M} = \mathcal{L} + \mathcal{E}. \quad (30)$$

The SFGTRPCA optimization problem is similar to the FGTRPCA problem. Details of the optimization algorithm and the entire procedure are available in the supplementary material due to space limitations.

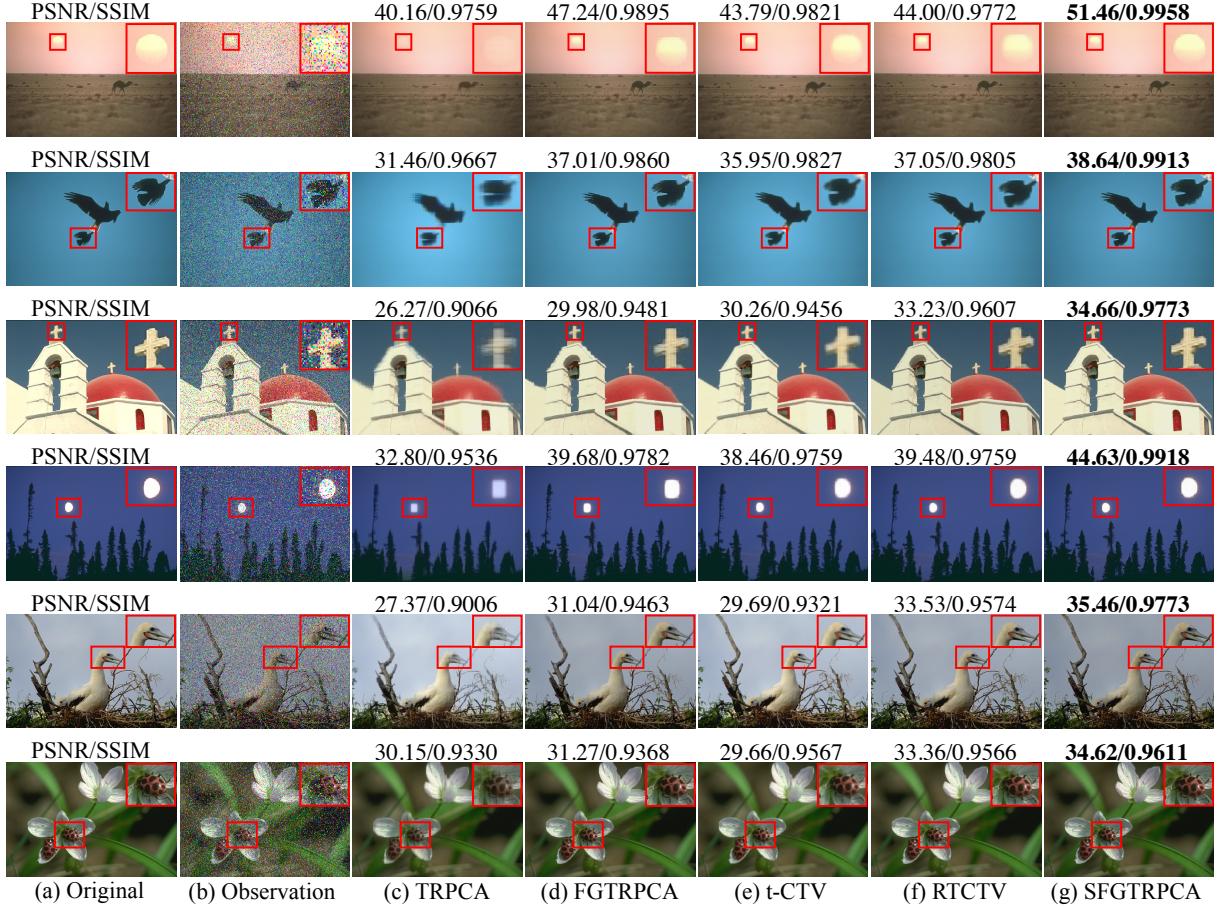


Figure 4: Recovery results on 6 color images from the BSD dataset with 20% noise ratio.

Experiments

In this section, we present several real-world experiments to substantiate the effectiveness of our models. Additional results are provided in the supplementary material.

Settings

Datasets: For comprehensive comparison, we use 4 widely used tensor data types including color images, grayscale videos, hyperspectral images (HSIs), and multispectral images (MSIs). For color images, we choose 3 widely used datasets including Berkeley Segmentation Dataset¹ (BSD) (Martin et al. 2001), Kodak (Kodak 1993) dataset², and ZheJiang University (ZJU) (Hu et al. 2012) dataset³. For grayscale videos, we use 14 grayscale video sequences from the YUV dataset⁴ and select the first 100 frames for each sequence. For HSIs, we select Cuprite⁵, DCMall⁵, Urban⁵, Indian Pines⁵, and Pavia University⁵ (PaviaU) for experiments

and select the first 50 bands from each HSI dataset. For MSIs, we randomly select 10 MSIs from the CAVE dataset (Yasuma et al. 2008).

Baselines: Our baselines are divided into two categories based on different priors. (1) Low-rankness: TRPCA (Lu et al. 2020), ETRPCA (Gao et al. 2020), and PTRPCA (Yan and Guo 2024); (2) Joint Low-rankness & Smoothness: t-CTV (Wang et al. 2023a) and RTCTV (Huang et al. 2024). We utilize the parameters recommended by the authors. For the key parameter α in our models, we search from a candidate set and employ $\alpha = 1$. More detailed parameter settings can be seen in supplementary material.

Evaluation metrics: The peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) are used to evaluate the recovery performance.

Noising Data Construction: For each channel of the color image, each frame of the grayscale video, and each band of HSI and MSI, we add random salt and pepper noise at varying noise ratios of 10%, 20%, and 30%.

¹<https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/>

²<http://r0k.us/graphics/kodak/>

³<https://sites.google.com/site/zjuyaohu/>

⁴<http://trace.eas.asu.edu/yuv/>

⁵<https://lesun.weebly.com/hyperspectral-data-set.html>

	Noise Ratio	10%		20%	
		Methods	PSNR	SSIM	PSNR
Color images	TRPCA	31.20	0.9464	29.55	0.9115
	ETRPCA	33.26	0.9580	31.23	0.9233
	PTRPCA	33.37	0.9622	31.43	0.9350
	FGTRPCA	<u>37.26</u>	<u>0.9796</u>	33.26	0.9415
	t-CTV	32.84	0.9525	31.71	0.9348
	RTCTV	34.96	0.9689	<u>33.46</u>	<u>0.9529</u>
	SFGTRPCA	40.96	0.9907	36.93	0.9782
Grayscale videos	TRPCA	35.19	0.9636	34.16	0.9538
	ETRPCA	38.29	0.9772	36.13	0.9433
	PTRPCA	38.95	0.9807	37.31	0.9669
	FGTRPCA	<u>41.85</u>	<u>0.9858</u>	<u>39.07</u>	<u>0.9743</u>
	t-CTV	37.37	0.9721	36.52	0.9665
	RTCTV	41.11	0.9843	38.62	0.9482
	SFGTRPCA	44.51	0.9911	41.91	0.9846
HSIs	TRPCA	44.18	0.9754	42.48	0.9718
	ETRPCA	44.54	0.9747	43.20	0.9720
	PTRPCA	47.38	0.9815	45.48	0.9772
	FGTRPCA	<u>47.30</u>	<u>0.9858</u>	44.90	0.9787
	t-CTV	45.72	0.9779	44.39	0.9759
	RTCTV	<u>48.29</u>	0.9812	<u>46.76</u>	0.9789
	SFGTRPCA	52.38	0.9888	50.16	0.9856
MSIs	TRPCA	42.07	0.9898	40.41	0.9867
	ETRPCA	45.95	0.9931	44.00	0.9906
	PTRPCA	46.87	0.9939	44.84	0.9920
	FGTRPCA	<u>49.73</u>	<u>0.9960</u>	46.05	0.9921
	t-CTV	46.62	0.9938	45.21	0.9925
	RTCTV	<u>50.19</u>	0.9952	<u>48.69</u>	<u>0.9941</u>
	SFGTRPCA	57.37	0.9977	53.35	0.9955

Table 2: Average PSNR and SSIM results on 4 tensor types with different noise ratios. The best results are marked in bold, and the second-best results are underlined.

Results

Visual Quality. To clearly illustrate the advantages of our methods, Figure 4 presents 6 sample images from the BSD dataset, along with the recovery results under 20% salt and pepper noise. The PSNR and SSIM values are listed above the recovered images to enhance the credibility of the results. The results show that SFGTRPCA constructs more image details and color information (Especially the contour and color of the moon in the 4-th image). Additionally, we have observed that the proposed FGTRPCA and SFGTRPCA methods significantly outperform the baseline methods under corresponding priors. Notably, the average PSNR value of SFGTRPCA surpasses the second-highest (apart from SGTRPCA) method by over 3.1 dB.

Quantitative Quality. Table 2 displays the results of all the competitors on the 4 tensor types with 10% and 20% noise. From the results, we draw the following conclusions:

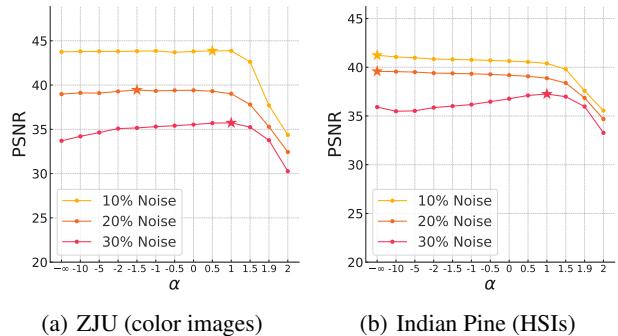


Figure 5: PSNR values of our SFGTRPCA algorithm on different cases. For various values of α , with the point of highest value marked by a pentagram.

- Firstly, our SFGTRPCA method exhibits a substantial performance gain over the comparisons in all cases. Overall, compared to the second-best baseline, SFGTRPCA realizes an average PSNR improvement of more than 4.4 dB. Specifically, in color images and MSIs with 10% noise, SFGTRPCA gains 17.16% and 14.31% PSNR improvements, respectively.
- Secondly, compared to the baselines that only consider the low-rankness prior (TRPCA, ETRPCA, and PTRPCA), our FGTRPCA demonstrates the best performance in most situations, and achieves an average improvement of over 1.7 dB in PSNR compared to the second-best baseline. Notably, our FGTRPCA always leads in SSIM values, which indicates that our FGTRPCA can recover more structural information.
- Thirdly, it is evident that under varying noise levels in different tensor data types, our methods consistently yield competitive scores of evaluation indices. This demonstrates that our methods better leverage the underlying low-rank and sparse structures within the tensor, exhibiting strong recovery ability and robustness.

Parameter Analysis. Obviously, the parameter α plays a crucial role in the performance of our models. This parameter directly influences flexibility and generalizability, making it essential for achieving optimal results in various scenarios. In Figure 5 we show the PSNR of the recovery results for different values of α with various noise ratios on ZJU (color images) and Indian Pines (HSIs) datasets. As seen, by tuning α , our SFGTRPCA model gains significant improvements in various situations. This improvement highlights the advantage of incorporating α as a hyperparameter instead of a fixed formula, and validates the flexibility of our framework.

Conclusion

In this article, we propose a flexible generalized low-rank regularizer termed as FGTNN. Equipped with this, we develop a novel FGTRPCA framework. Many existing popular TRPCA methods can be considered as special cases of FGTRPCA. FGTRPCA offers insights into the connections

between existing TRPCA methods and new approaches of TRPCAs through a continuous parameter. Moreover, considering the low-rankness and smoothness priors simultaneously, we devise a novel regularizer based on FGTNN and put forward a smooth FGTRPCA (SFGTRPCA) model. Compared with existing popular works, our models improve flexibility and generalizability, thus performing better in many tensor denoising and recovery tasks.

References

- Boyd, S.; Parikh, N.; Chu, E.; Peleato, B.; Eckstein, J.; et al. 2011. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning*, 3(1): 1–122.
- Chen, L.; Jiang, X.; Liu, X.; and Zhou, Z. 2021. Logarithmic norm regularized low-rank factorization for matrix and tensor completion. *IEEE Transactions on Image Processing*, 30: 3434–3449.
- Gao, Q.; Zhang, P.; Xia, W.; Xie, D.; Gao, X.; and Tao, D. 2020. Enhanced tensor RPCA and its application. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(6): 2133–2140.
- Hou, J.; Liu, X.; Wang, H.; and Guo, K. 2024. Tensor recovery from binary measurements fused low-rankness and smoothness. *Signal Processing*, 221: 109480.
- Hu, Y.; Zhang, D.; Ye, J.; Li, X.; and He, X. 2012. Fast and accurate matrix completion via truncated nuclear norm regularization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(9): 2117–2130.
- Huang, B.; Mu, C.; Goldfarb, D.; and Wright, J. 2015. Provable models for robust low-rank tensor completion. *Pacific Journal of Optimization*, 11(2): 339–364.
- Huang, K.; Kong, W.; Zhou, M.; Qin, W.; Zhang, F.; and Wang, J. 2024. Enhanced Low-Rank Tensor Recovery Fusing Reweighted Tensor Correlated Total Variation Regularization for Image Denoising. *Journal of Scientific Computing*, 99(3): 69.
- Ji, J.; and Feng, S. 2023. Anchor structure regularization induced multi-view subspace clustering via enhanced tensor rank minimization. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 19343–19352.
- Jiang, T.-X.; Huang, T.-Z.; Zhao, X.-L.; and Deng, L.-J. 2020. Multi-dimensional imaging data recovery via minimizing the partial sum of tubal nuclear norm. *Journal of Computational and Applied Mathematics*, 372: 112680.
- Kilmer, M.; and Martin, C. 2011. Factorization strategies for third-order tensors. *Linear Algebra Appl.*, 435(3): 641–658.
- 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.
- Kilmer, M. E.; Horesh, L.; Avron, H.; and Newman, E. 2021. Tensor-tensor algebra for optimal representation and compression of multiway data. *Proceedings of the National Academy of Sciences*, 118(28): e2015851118.
- Ko, C.-Y.; Batselier, K.; Daniel, L.; Yu, W.; and Wong, N. 2020. Fast and accurate tensor completion with total variation regularized tensor trains. *IEEE Transactions on Image Processing*, 29: 6918–6931.
- Kodak, E. 1993. Kodak lossless true color image suite (PhotoCD PCD0992). URL <http://r0k.us/graphics/kodak>, 6: 2.
- Liu, Y.; Du, B.; Chen, Y.; Zhang, L.; Gong, M.; and Tao, D. 2024a. Convex–Concave Tensor Robust Principal Component Analysis. *International Journal of Computer Vision*, 132(5): 1721–1747.
- Liu, Y.; Wang, Y.; Chen, L.; Wang, L.; and Hu, Y. 2024b. TIHN: Tensor Improved Huber Norm for Low-rank Tensor Recovery. *International Journal of Wavelets, Multiresolution and Information Processing*, 22(05): 2450023.
- Liu, Z.; Han, Z.; Tang, Y.; Zhao, X.-L.; and Wang, Y. 2024c. Low-Tubal-Rank Tensor Recovery via Factorized Gradient Descent. *arXiv preprint arXiv:2401.11940*.
- Lu, C.; Feng, J.; Chen, Y.; Liu, W.; Lin, Z.; and Yan, S. 2020. Tensor Robust Principal Component Analysis with a New Tensor Nuclear Norm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(4): 925–938.
- Martin, D.; Fowlkes, C.; Tal, D.; and Malik, J. 2001. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proceedings eighth IEEE international conference on computer vision. ICCV 2001*, volume 2, 416–423. IEEE.
- Peng, J.; Wang, H.; Cao, X.; Liu, X.; Rui, X.; and Meng, D. 2022a. Fast noise removal in hyperspectral images via representative coefficient total variation. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1–17.
- Peng, J.; Wang, Y.; Zhang, H.; Wang, J.; and Meng, D. 2022b. Exact decomposition of joint low rankness and local smoothness plus sparse matrices. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(5): 5766–5781.
- Peng, J.; Xie, Q.; Zhao, Q.; Wang, Y.; Yee, L.; and Meng, D. 2020. Enhanced 3DTV regularization and its applications on HSI denoising and compressed sensing. *IEEE Transactions on Image Processing*, 29: 7889–7903.
- Phothilimthana, M.; Abu-El-Haija, S.; Cao, K.; Fatemi, B.; Burrows, M.; Mendis, C.; and Perozzi, B. 2024. Tpugraphs: A performance prediction dataset on large tensor computational graphs. *Advances in Neural Information Processing Systems*, 36.
- Qin, W.; Wang, H.; Zhang, F.; Ma, W.; Wang, J.; and Huang, T. 2024. Nonconvex Robust High-Order Tensor Completion Using Randomized Low-Rank Approximation. *IEEE Transactions on Image Processing*, 33: 2835–2850.
- 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.
- 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, Y.; Kou, K. I.; Chen, H.; Tang, Y. Y.; and Li, L. 2023b. Double Auto-Weighted Tensor Robust Principal Component Analysis. *IEEE Transactions on Image Processing*, 32: 5114–5125.
- Wen, T.; Chen, E.; and Chen, Y. 2024. Tensor-view topological graph neural network. In *International Conference on Artificial Intelligence and Statistics*, 4330–4338. PMLR.
- Yan, T.; and Guo, Q. 2024. Tensor robust principal component analysis via dual lp quasi-norm sparse constraints. *Digital Signal Processing*, 150: 104520.
- Yasuma, F.; Mitsunaga, T.; Iso, D.; and Nayar, S. K. 2008. Generalized Assorted Pixel Camera: Post-Capture Control of. *Tech. rep., Department of Computer Science, Columbia University CUCS-061-08*.
- Zhang, C.; Li, H.; Lv, W.; Huang, Z.; Gao, Y.; and Chen, C. 2023a. Enhanced tensor low-rank and sparse representation recovery for incomplete multi-view clustering. In *Proceedings of the AAAI conference on artificial intelligence*, volume 37, 11174–11182.
- Zhang, F.; Wang, H.; Qin, W.; Zhao, X.; and Wang, J. 2023b. Generalized nonconvex regularization for tensor RPCA and its applications in visual inpainting. *Applied Intelligence*, 53(20): 23124–23146.
- Zhao, L.; Xu, X.; Wang, Z.; Zhang, Y.; Zhang, B.; Zheng, W.; Du, D.; Zhou, J.; and Lu, J. 2024. LowRankOcc: Tensor Decomposition and Low-Rank Recovery for Vision-based 3D Semantic Occupancy Prediction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 9806–9815.

Reproducibility Checklist

This paper:

- Includes a conceptual outline and/or pseudocode description of AI methods introduced [yes]
- Clearly delineates statements that are opinions, hypothesis, and speculation from objective facts and results [yes]
- Provides well marked pedagogical references for less-familiar readers to gain background necessary to replicate the paper [yes]

Does this paper make theoretical contributions? (yes)

If yes, please complete the list below.

- All assumptions and restrictions are stated clearly and formally. [yes]
- All novel claims are stated formally (e.g., in theorem statements). [yes]
- Proofs of all novel claims are included. [yes]
- Proof sketches or intuitions are given for complex and/or novel results. [yes]
- Appropriate citations to theoretical tools used are given. [yes]
- All theoretical claims are demonstrated empirically to hold. [yes]
- All experimental code used to eliminate or disprove claims is included. [yes]

Does this paper rely on one or more datasets? [yes]

If yes, please complete the list below

- A motivation is given for why the experiments are conducted on the selected datasets [yes]
- All novel datasets introduced in this paper are included in a data appendix. [NA]
- All novel datasets introduced in this paper will be made publicly available upon publication of the paper with a license that allows free usage for research purposes. [NA]
- All datasets drawn from the existing literature (potentially including authors' own previously published work) are accompanied by appropriate citations. [yes]
- All datasets drawn from the existing literature (potentially including authors' own previously published work) are publicly available. [yes]
- All datasets that are not publicly available are described in detail, with explanation why publicly available alternatives are not scientifically satisfying. [yes]

Does this paper include computational experiments? [yes]

If yes, please complete the list below.

- Any code required for pre-processing data is included in the appendix. [yes]
- All source code required for conducting and analyzing the experiments is included in a code appendix. [yes]
- All source code required for conducting and analyzing the experiments will be made publicly available upon publication of the paper with a license that allows free usage for research purposes. [yes]
- All source code implementing new methods have comments detailing the implementation, with references to the paper where each step comes from [yes]
- If an algorithm depends on randomness, then the method used for setting seeds is described in a way sufficient to allow replication of results. [yes]
- This paper specifies the computing infrastructure used for running experiments (hardware and software), including GPU/CPU models; amount of memory; operating system; names and versions of relevant software libraries and frameworks. [yes]
- This paper formally describes evaluation metrics used and explains the motivation for choosing these metrics. [yes]
- This paper states the number of algorithm runs used to compute each reported result. [yes]
- Analysis of experiments goes beyond single-dimensional summaries of performance (e.g., average; median) to include measures of variation, confidence, or other distributional information. [yes]
- The significance of any improvement or decrease in performance is judged using appropriate statistical tests (e.g., Wilcoxon signed-rank). [yes]
- This paper lists all final (hyper-)parameters used for each model/algorithm in the paper's experiments. [yes]

- This paper states the number and range of values tried per (hyper-) parameter during development of the paper, along with the criterion used for selecting the final parameter setting. [yes]