Fast Guaranteed Robust Local-Smooth Principal Component Separation

Mingdi Hu[1]*, Hailin Wang[2]*, Shuaijiang Li[1], Kexin Shi[2], Jiangjun Peng[3,4]†

[1] School of Communications and Information Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710121, China

[2] School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, China

[3] School of Mathematics and Statistics, Northwestern Polytechnical University, Xi'an 710129, China

[4] Shenzhen Research Institute of Northwestern Polytechnical University, Nanshan, Shenzhen 518057, China

humingdi2005@xupt.edu.cn, wanghailin97@163.com, lishuaijiang@stu.xupt.edu.cn, kexinshi@stu.xjtu.edu.cn, pengjj@nwpu.edu.cn

Abstract

Leveraging intrinsic data priors is critical for effective data recovery. However, existing approaches often struggle to achieve theoretical guarantees, strong performance, and computational efficiency simultaneously. In this paper, we introduce a novel *Representative Coefficient Correlated Total Variation (RCCTV)* regularizer that captures the recently observed low-rank and local smoothness properties of the representative coefficient tensor derived from a low-rank decomposition. RCCTV regularizer offers three key advantages: (1) it operates on a compact representative coefficient image significantly smaller than the original data, enabling highly efficient optimization; (2) it jointly enforces low-rankness and spatial smoothness through a single regularizer, eliminating the need for trade-off parameters; and (3) when integrated into a robust PCA framework (i.e., *RCCTV-RPCA* model), it admits provable exact recovery under mild conditions. To solve the resulting model, we develop an efficient ADMM-based algorithm accelerated via fast Fourier transform. Extensive experiments on both synthetic and real-world datasets demonstrate that the *RCCTV-RPCA* model achieves state-of-the-art accuracy while running significantly faster. Our code and Supplementary Material are available at https://github.com/mendy-2013/RCCTV.

Introduction

Robust principal component analysis (RPCA) (Candès et al. 2011) is a foundational model in data recovery that aims to separate a low-rank matrix $\mathbf{X_0}$ and a sparse error matrix $\mathbf{S_0}$ from a fully observed but corrupted matrix $\mathbf{Y} = \mathbf{X_0} + \mathbf{S_0}$. This decomposition is obtained by solving the following convex optimization problem:

$$\min_{\mathbf{X}} \ \|\mathbf{X}\|_* + \beta\|\mathbf{X}\|_{\mathrm{TV}}, \tag{1}$$

where $\|\mathbf{X}\|_*$ denotes the nuclear norm (NN, sum of singular values), and $\|\mathbf{S}\|_1$ represents the entry-wise ℓ_1-norm (sum of absolute values of all entries). The parameter $\lambda > 0$, typically set as $1/\sqrt{\max(n_1, n_2)}$ for an $n_1 \times n_2$ matrix, balances the trade-off between the low-rank and sparse components. Under certain incoherence conditions (Candès and Recht 2009; Chen 2015), RPCA guarantees exact recovery with high probability (Candès et al. 2011; Lu et al. 2019).

Regularizer	Prior	Efficiency	Guarantees
$\|\mathbf{X}\|_*$	L of $\mathbf{X}$	$\mathcal{O}(MNB^2)$	✓
$\|\mathbf{X}\|_{\mathrm{TV}}$	LS of $\mathbf{X}$	$\mathcal{O}(MNB\log(MN))$	×
$\|\mathbf{X}\|_* + \beta\|\mathbf{X}\|_{\mathrm{TV}}$	LS, L of $\mathbf{X}$	$\mathcal{O}(MNB^2)$	×
$\|\mathbf{X}\|_{\mathrm{CTV}}$	LS, L of $\mathbf{X}$	$\mathcal{O}(MNB^2)$	✓
$\|\mathbf{X}\|_{\mathrm{RCCTV}}$	LS, L of $\mathbf{U}$	$\mathcal{O}(MNR^2)$	✓

Table 1: Comparison of the properties of different regularization terms applied to data $\mathbf{X} \in \mathbb{R}^{MN\times B}$ with rank $R(R \ll B)$, where $\mathbf{U} \in \mathbb{R}^{MN\times R}$ denotes the coefficient matrix representing $\mathbf{X}$. Here, "L" refers to low-rankness and "LS" to local smoothness.

The simplicity and theoretical elegance of RPCA have led to its widespread adoption in numerous machine learning and computer vision tasks, including image denoising (Peng et al. 2022b), background subtraction in videos (Wang et al. 2023), etc (Bouwmans et al. 2018). However, RPCA only captures the global low-rank structure of data; it often falls short when applied to structured visual data such as hyperspectral images (HSI) (Wang et al. 2017; Peng et al. 2020), multispectral images (MSI) (Peng et al. 2025b), and surveillance videos (Cao, Yang, and Guo 2015; Cao et al. 2016). These data exhibit not only global low-rank structure but also strong local smoothness in the spatial domain.

Local smoothness refers to the property that neighboring pixels in images often share similar intensity values, owing to continuity in materials and scene content. This spatial regularity is especially prominent in remote sensing and video data, where abrupt changes are rare. As illustrated in Figure 1, this smoothness manifests as sparsity in the gradient domain, motivating the use of total variation (TV) regularization (Rudin, Osher, and Fatemi 1992; Peng et al. 2024) to capture local structure. To jointly exploit both low-rank and local smoothness, a natural extension is to combine NN and TV regularizer for the principal component data $\mathbf{X}$, i.e.,

$$\|\mathbf{X}\|_* + \beta\|\mathbf{X}\|_{\mathrm{TV}}, \tag{2}$$

*These authors contributed equally.

†Corresponding author.

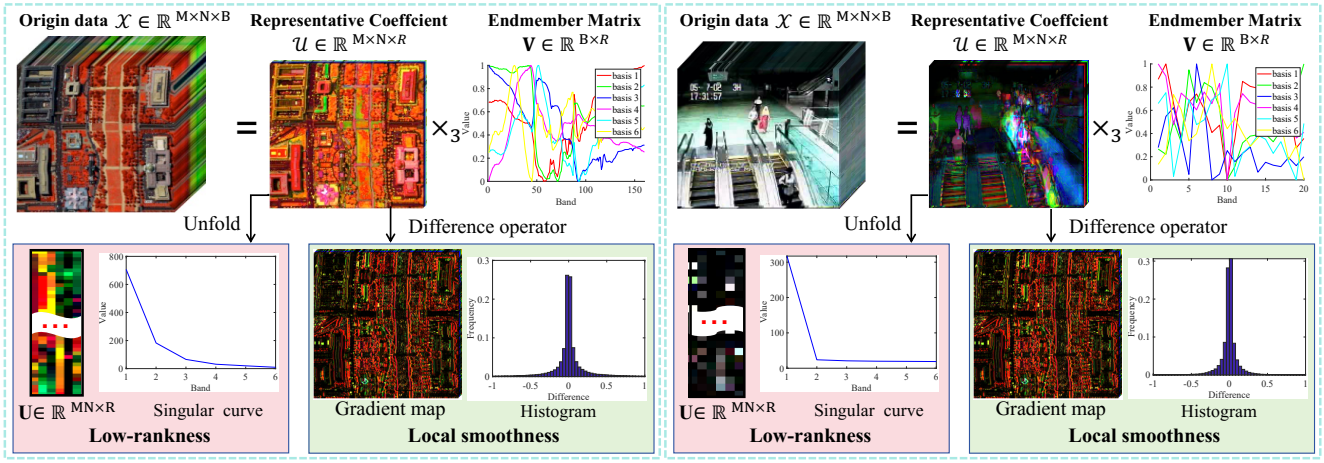


Figure 1: Illustration of subspace inheritance of low-rank and local smoothness priors from the original data. A hyperspectral image and a surveillance video are shown as representative examples.

where $\beta > 0$ controls the trade-off between the two priors. Several works have adopted this formulation and reported empirical improvements (He et al. 2015; Wang et al. 2017; Peng et al. 2018). However, the fusion model (2) suffers from two significant challenges: First, selecting the parameter β is nontrivial and data-dependent. Second, theoretical analysis becomes intractable due to the absence of a recovery theory for TV regularization.

To overcome these limitations, the correlated total variation (CTV) regularizer has been recently proposed (Peng et al. 2022b; Wang et al. 2023; Peng et al. 2024), defined as:

$$\|\mathbf{X}\|_{\mathrm{CTV}} := \|\nabla(\mathbf{X})\|_*, \tag{3}$$

where $\nabla(\cdot)$ is the spatial gradient (or difference) operator. The CTV regularizer elegantly integrates both local smoothness and global low-rank priors into a single term by applying the nuclear norm to the gradient map. This eliminates the need for a trade-off parameter while enabling theoretical recovery analysis. As shown in Table 1, models employing CTV enjoy both practical and theoretical benefits.

Despite these advances, existing models still suffer from scalability issues. Both NN, TV, and CTV regularizers operate on the original high-dimensional data, which becomes computationally burdensome for large-scale visual data such as HSI/MSI and videos. This raises a fundamental question: *Can we design a computationally efficient model that preserves local smoothness and low-rank priors, while also offering theoretical guarantees for recovery?*

In this work, we address this question by examining the structure of the low-rank subspace. Large-scale data of interest such as HSI/MSI and videos can typically be modeled as a low-rank tensor $\mathcal{X} \in \mathbb{R}^{M\times N\times B}$, which admits the factorization $\mathcal{X} = \mathcal{U} \times_3 \mathbf{V}$, where $\mathcal{U} \in \mathbb{R}^{M\times N\times R}$ $(R \ll B)$ contains the representative coefficients, and $\mathbf{V} \in \mathbb{R}^{B\times R}$ is the end-member matrix. The mode-3 product $\times_3$ projects the tensor onto a lower-dimensional subspace (Kolda and Bader 2009). Interestingly, we find that in real data, the representative tensor $\mathcal{U}$ inherits the local smoothness of $\mathcal{X}$—i.e., its spatial gradients remain sparse—while the unfolded matrix $\mathbf{U} = \mathrm{unfold}_3(\mathcal{U}) \in \mathbb{R}^{MN\times R}$ retains low-rank structure. See Figure 1 for illustrations. This observation inspires us to define a regularizer on the much smaller matrix $\mathbf{U}$, significantly improving computational efficiency:

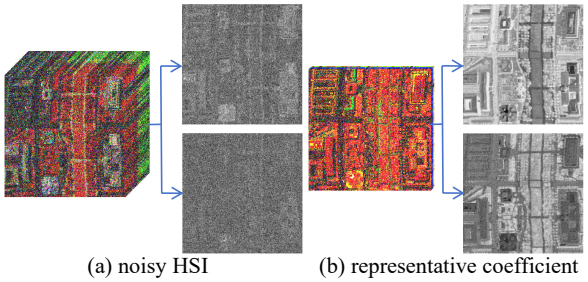
$$\|\mathbf{X}\|_{\mathrm{RCCTV}} := \|\nabla(\mathbf{U})\|_*, \quad \text{s.t.} \mathbf{X} = \mathbf{U}\mathbf{V}^\top, \tag{4}$$

where $\mathbf{V}$ is a orthogonal matrix that satisfies $\mathbf{V}^\top\mathbf{V} = \mathbf{I}$, and $\mathbf{X} = \mathrm{unfold}_3(\mathcal{X}) \in \mathbb{R}^{MN\times B}$. We refer to Eq. (4) as the RCCTV regularizer, where the nuclear norm is applied to the spatial gradients of the representative coefficient matrix $\mathbb{U}$. Because $\mathbb{U}$ contains only R slices, this formulation significantly reduces computational cost. Defining the regularizer in the low-rank subspace further preserves essential structural priors while facilitating theoretical analysis. As summarized in Table 1, RCCTV is the only approach that simultaneously achieves unified prior modeling, computational efficiency, and provable guarantees.

In summary, our main contributions are:

- **Regularization Design**: We introduce the RCCTV regularizer, which simultaneously captures low-rank and local smoothness properties on a compact subspace representation. This avoids trade-off parameters and enhances computational efficiency.
- **Theoretical Analysis**: We formulate the RCCTV-RPCA model and establish exact recovery guarantees under mild conditions. To our knowledge, this is the first subspace-based model that unifies spatial smoothness and low-rankness with rigorous theory.
- **Algorithm and Implementation**: We develop an efficient optimization algorithm based on the Alternating Direction Method of Multipliers (ADMM), accelerated using the fast Fourier transform. Extensive experiments on synthetic and real-world datasets demonstrate state-of-the-art performance in accuracy and speed.

Notations. For a given dataset $\mathcal{X} \in \mathbb{R}^{M\times N\times B}$ that exhibits joint low-rankness and local smoothness, we let M, N, and



(a) noisy HSI (b) representative coefficient

Figure 2: Illustration of the effect of Gaussian noise ($\sigma = 0.4$) on an HSI and its representative coefficient images.

B denote the sizes of its three modes, respectively. We define the mode-3 unfolding of $\mathcal{X}$ as $\mathbf{X} \in \mathbb{R}^{MN\times B}$, which satisfies $\mathbf{X} = \text{unfold}(\mathcal{X})$ and $\mathcal{X} = \text{fold}(\mathbf{X})$. For a matrix $\mathbf{X} \in \mathbb{R}^{MN\times B}$ of rank R, it admits the low-rank decomposition $\mathbf{X} = \mathbf{U}\mathbf{V}^\top$, where $\mathbf{U} \in \mathbb{R}^{MN\times R}$ and $\mathbf{V} \in \mathbb{R}^{B\times R}$ denote the representative coefficient matrix and the endmember matrix, respectively. We further define the difference operators on $\mathbf{U}$ as $\nabla_i, i = 1, 2$, given by

$$\nabla_i(\mathbf{U}) = \text{unfold}\left(\mathbf{D}_i\big(\text{fold}(\mathbf{U})\big)\right) = \mathbf{G}_i, \quad i = 1, 2, \quad (5)$$

where $\mathbf{D}_i, i = 1, 2$ denotes the spatial difference operator applied along the horizontal and vertical directions of the representative coefficient tensor $\mathcal{U} = \text{fold}(\mathbf{U}) \in \mathbb{R}^{M\times N\times R}$.

Representative Coefficient Correlated Total Variation Regularizer

Next, we introduce the *Representative Coefficient Correlated Total Variation (RCCTV)* Regularizer from three perspectives: the model, the theory, and the algorithm.

RCCTV-RPCA Model

To illustrate the motivation behind the proposed RCCTV regularizer, we consider two representative examples: a hyperspectral image (HSI) and a surveillance video. As depicted in Figure 1, we observe three key properties:

1. A low-rank tensor $\mathcal{X} \in \mathbb{R}^{M\times N\times B}$ of rank R along the third mode has a representative coefficient tensor with only R slices. When the data is strongly low-rank (i.e., $R/B \ll 1$), operating on the representative coefficients enables significantly faster optimization.
2. The representative coefficient tensor obtained via low-rank decomposition exhibits clear local smoothness. This is evidenced by the sparsity in the spatial gradient maps of $\mathcal{U}$ shown in Figure 1.
3. The unfolded coefficient matrix $\mathbf{U}$ retains approximate low-rank structure, as indicated by a rapidly decaying singular value spectrum, again seen in Figure 1.

Furthermore, as shown in Figure 2, adding Gaussian noise to the original data severely distorts its structure. However, after applying the low-rank decomposition, the resulting coefficient tensor $\mathcal{U}$ largely preserves the underlying spatial patterns of the clean data. This robustness arises because the decomposition extracts the principal components of the signal, which are inherently less sensitive to noise. Consequently, the representative coefficient tensor often maintains greater structural integrity than the noisy data itself.

Motivated by these observations and inspired by the CTV regularizer, we introduce a new variant applied to the representative coefficients. This formulation captures both low-rankness and local smoothness within a lower-dimensional subspace, leading to the *RCCTV* regularizer, defined as:

$$\|\mathbf{X}\|_{\text{RCCTV}} := \sum_{i=1}^{2} \|\nabla_i(\mathbf{U})\|_*, \quad \text{s.t.}\mathbf{X} = \mathbf{U}\mathbf{V}^\top, \quad (6)$$

where $\mathbf{U} \in \mathbb{R}^{MN\times R}$ is the coefficient matrix, $\mathbf{V} \in \mathbb{R}^{B\times R}$ is the basis matrix with $\mathbf{V}^\top\mathbf{V} = \mathbf{I}$, and R is a predefined rank. Note that Eqs. (4) and (6) are equivalent forms of the RCCTV regularizer, with Eq. (6) explicitly indicating the directional gradient terms. *We now explain why RCCTV regularizer effectively captures both the low-rank structure of the original data and the local smoothness property of its subspace factor.*

1. The original tensor $\mathcal{X}$ exhibits strong low-rankness along the third mode. In the RCCTV regularizer, this structure is explicitly modeled by fixing the rank R in the low-rank decomposition.
2. As shown in Figure 1, the coefficient matrix $\mathbf{U}$ also displays approximate low-rank structure, but this property is weaker than that of the original tensor. Enforcing an even lower rank may cause information loss. Therefore, applying the nuclear norm to $\mathbf{U}$ provides a more flexible and adaptive way to encourage low-rankness.
3. By norm compatibility, we have:

$$\|\nabla_i(\mathbf{U})\|_{\text{F}} \le \|\nabla_i(\mathbf{U})\|_* \le \|\nabla_i(\mathbf{U})\|_1, \quad (7)$$

where the Frobenius and ℓ_1 norms correspond to isotropic and anisotropic TV regularizations, respectively (He et al. 2015; Peng et al. 2022a). This inequality indicates that minimizing the nuclear norm of the gradients (i.e., CTV) also promotes local smoothness. Consequently, the RCCTV regularizer naturally inherits and enforces spatial continuity in the coefficient domain (Peng et al. 2022b).

Following the classical RPCA framework (Candès et al. 2011), we incorporate the RCCTV regularizer into the objective to formulate a new model for recovering locally smooth principal components:

$$\min_{\mathbf{X},\mathbf{S}} \ \|\mathbf{X}\|_{\text{RCCTV}} + 2\lambda\|\mathbf{S}\|_1 \quad s,t.\mathbf{Y} = \mathbf{X} + \mathbf{S}, \quad (8)$$

where $\mathbf{Y}$ denotes the observed corrupted data, and $\mathbf{S}$ represents sparse noise or outliers. By substituting Eq. (6) into Eq. (8), we obtain the equivalent formulation:

$$\begin{aligned} \min_{\mathbf{X}=\mathbf{U}\mathbf{V}^\top,\mathbf{S}} \quad & \sum_{i=1}^{2} \|\nabla_i(\mathbf{U})\|_* + 2\lambda\|\mathbf{S}\|_1, \\ \text{s.t.} \quad & \mathbf{Y} = \mathbf{U}\mathbf{V}^\top + \mathbf{S}, \mathbf{V}^\top\mathbf{V} = \mathbf{I}. \end{aligned} \quad (9)$$

By expressing $\mathbf{X}$ in its low-rank factorized form $\mathbf{U}\mathbf{V}^\top$, Eq. (9) provides an equivalent formulation that operates in a

compact subspace. Here, the nuclear norm of the gradients $\nabla_i(\mathbf{U})$ enforces local smoothness in the coefficient domain, and the orthogonality constraint $\mathbf{V}^\top\mathbf{V} = \mathbf{I}$ ensures identifiability of the low-rank factors. This formulation enables efficient and theoretically grounded recovery of low-rank, locally smooth components.

Theorem Guarantee

Next, we establish the exact recovery guarantee for the proposed RCCTV-RPCA model (9). As with other RPCA-based formulations, a key requirement for successful recovery is the *incoherence condition* on the low-rank component. This condition ensures that the left and right singular vectors of the underlying low-rank matrix are not excessively concentrated or aligned with the standard basis, which could otherwise impede the separation from sparse corruptions (Candès and Recht 2009; Chen 2015; Zhang et al. 2020).

Since the nuclear norm is defined on the gradient map of the subspace factor (i.e., representative coefficient matrix $\mathbf{U}$), we impose the incoherence condition directly on $\nabla_i(\mathbf{U})$ rather than on the gradient of $\nabla_i(\mathbf{X})$. Specifically, let each gradient map $\nabla_i(\mathbf{U})$ $(i = 1, 2)$ admit the singular value decomposition $\mathbf{L}_i\mathbf{\Sigma}_i\mathbf{R}_i^\top$, where $\mathbf{L}_i \in \mathbb{R}^{n_1\times r}$, $\mathbf{R}_i \in \mathbb{R}^{n_2\times r}$. We then make the following subspace incoherence assumption with constant μ for the RCCTV-RPCA model (8):

Assumption 1 (Subspace Incoherence) *For a low-rank matrix $\nabla_i(\mathbf{U}_0) \in \mathbb{R}^{n_1\times n_2}$ with rank r, it satisfies the incoherence condition with parameter μ, i.e.,*

$$\begin{aligned}\max_k\|\mathbf{L}_i^\top\hat{\mathbf{e}}_k\|^2 \le \frac{\mu r}{n_1}, \max_k\|\mathbf{R}_i^\top\hat{\mathbf{e}}_k\|^2 \le \frac{\mu r}{n_2},\\ \|\mathbf{L}_i\mathbf{R}_i^\top\|_\infty \le \sqrt{\frac{\mu r}{n_1 n_2}},\end{aligned} \tag{10}$$

where $\mathbf{L}_i \in \mathbb{R}^{n_1\times r}$ and $\mathbf{R}_i \in \mathbb{R}^{n_2\times r}$ are obtained from the singular vector decomposition of $\nabla_i(\mathbf{U}_0)$ and $\hat{\mathbf{e}}_k$ is the k-th standard basis vector (Candès et al. 2011).

Based on the subspace incoherence assumption (i.e., Assumption 1), we can obtain the following exact recoverability theorem.

Theorem 1 *Suppose that the gradient maps of the representative coefficient matrix of the clean data $\mathbf{X}_0$ (i.e., $\nabla_i(\mathbf{U}) \in \mathbb{R}^{n_1\times n_2}, i = 1, 2$) satisfy Assumption 1, and that the support set Ω of $\mathbf{S}_0$ is uniformly distributed among all sets of cardinality m. Then, there exists a numerical constant $c > 0$ such that, with probability at least $1 - cn_1^{-10}$ (over the choice of the support of $\mathbf{S}_0$), the RCCTV-RPCA model (8) with $\lambda = 1/\sqrt{n_1}$ exactly recovers the underlying components, i.e., the solution $(\hat{\mathbf{X}}, \hat{\mathbf{S}}) = (\mathbf{X}_0, \mathbf{S}_0)$, provided that*

$$\textit{rank}(\mathbf{X}_0) \le \rho_r n_2\mu^{-1}(\log n_1)^{-2}, \textit{ and } m \le \rho_s n_1 n_2, \tag{11}$$

where ρ_r and ρ_s are some positive numerical constants, and m is the number of the support set of $\mathbf{S}_0$.

The proof of the above theorem follows a similar strategy to (Peng et al. 2025a) and is provided in the supplementary material. This result asserts that when the gradient maps $\nabla_i(\mathbf{U})$, $i = 1, 2$, of the representative coefficient matrix $\mathbf{U}$ satisfy the subspace incoherence condition (i.e., Assumption 1), and the support of $\mathbf{S}_0$ is randomly distributed, the RCCTV-RPCA model (9) can exactly recover both the low-rank, locally smooth components and the sparse component with high probability.

Model	Computational Complexity
CTV	$\mathcal{O}(MNB\log(MN) + MNB^2)$
RCCTV	$\mathcal{O}(MNR\log(MN) + MNR^2 + MNB + BR^2)$

Table 2: Comparison of the computational complexity for CTV- and RCCTV-based RPCA model.

Optimization Algorithm

To efficiently solve the RCCTV-RPCA model (9), we adopt the well-established *Alternating Direction Method of Multipliers* (ADMM) framework (Boyd, Parikh, and Chu 2011). This approach enables the original optimization problem to be decomposed into a sequence of simpler subproblems, each of which can be solved efficiently in closed form. We begin by formulating the corresponding augmented Lagrangian function as follows:

$$\begin{aligned}&\mathcal{L}(\mathbf{U}, \mathbf{V}, \mathbf{S}, \{\mathbf{\Gamma}_i\}_{i=1}^3, \{\mathbf{G}_i\}_{i=1}^2)\\ &\quad= \sum_{i=1}^{2}\left(\|\mathbf{G}_i\|_* + \frac{\mu}{2}\left\|\nabla_i(\mathbf{U}) - \mathbf{G}_i + \frac{\mathbf{\Gamma}_i}{\mu}\right\|_F^2\right)\\ &\qquad+ 2\lambda\|\mathbf{S}\|_1 + \frac{\mu}{2}\left\|\mathbf{Y} - \mathbf{U}\mathbf{V}^\top - \mathbf{S} + \frac{\mathbf{\Gamma}_3}{\mu}\right\|_F^2.\end{aligned} \tag{12}$$

where $\mathbf{G}_1$ and $\mathbf{G}_2$ are auxiliary variables introduced to decouple the non-smooth nuclear norm terms, $\mu > 0$ is the penalty parameter, and $\mathbf{\Gamma}_1$, $\mathbf{\Gamma}_2$, and $\mathbf{\Gamma}_3$ are the associated Lagrange multipliers. Each subproblem in this framework admits an efficient closed-form update. For brevity, the full update rules and detailed derivations are provided in the supplementary material.

The detail process of solving Eq. (12) via ADMM summarized in the supplementary material, each iteration of the algorithm involves several efficient operations: updating $\mathbf{U}$ via the Fast Fourier Transform (FFT), updating $\mathbf{V}$ and $\mathbf{G}_i$ $(i = 1, 2)$ via small-scale singular value decomposition (SVD), updating $\mathbf{S}$ using soft-thresholding, and performing matrix multiplications. The computational complexity of each step is as follows: $\mathcal{O}(MNB)$ for soft-thresholding, $\mathcal{O}(BR^2)$ for updating $\mathbf{V}$, $\mathcal{O}(MNR^2)$ for updating $\mathbf{G}_i$, and $\mathcal{O}(RMN\log(MN))$ for updating $\mathbf{U}$ via FFT. Consequently, the total per-iteration complexity is

$$\mathcal{O}(MNB + BR^2 + MNR^2 + RMN\log(MN)), \tag{13}$$

which demonstrates that the algorithm scales efficiently with both the data dimensions and the subspace rank, making it well-suited for large-scale, high-dimensional datasets.

It is worth noting that the CTV-RPCA model can be regarded as a special case of the RCCTV-RPCA framework when $\mathbf{V} = \mathbf{I}$. However, unlike RCCTV-RPCA, CTV-RPCA

does not exploit the low-rank structure of the representative coefficient matrix $\mathbf{U}$. Furthermore, since RCCTV-RPCA operates in a substantially smaller subspace ($R \ll B$), it achieves significantly lower computational complexity. A detailed comparison of computational costs is provided in Table 2, which clearly demonstrates that RCCTV-RPCA is far more efficient than CTV-RPCA, particularly in high-dimensional settings.

Simulation Experiments

This section presents experiments on synthetic data to evaluate the RCCTV-RPCA models performance and theoretical claims, using datasets with both low-rank and local-smooth.

Low-rank and Local-smooth Data Generation

We first generate the tensor $\mathcal{X}_0 = \mathcal{U} \times_3 \mathbf{V}$, where the coefficient tensor $\mathcal{U} \in \mathbb{R}^{M \times N \times R}$, the basis matrix $\mathbf{V} \in \mathbb{R}^{B \times R}$, and $R \ll \min MN, B$. To ensure that $\mathcal{X}_0$ exhibits strong local smoothness, we adopt the same data generation mechanism as in (Peng et al. 2022b):

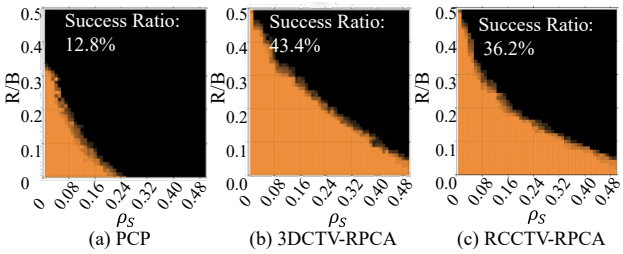
1. Generate the representative coefficient tensor $\mathcal{U}$ by randomly selecting R initial points. Assign each remaining point to its nearest initial point, thereby partitioning the space into R regions, with each region sharing the same representative coefficient vector. Each element is sampled independently.
2. Generate the basis matrix $\mathbf{V}$ by first sampling each vector $\mathbf{V}(i,:)$ independently and then applying smoothing.
3. Normalize the grayscale values of each column of $\mathbf{V}$ to lie within $[0,1]$ and obtain $\mathbf{X}_0$ via $\mathbf{X}_0 = \text{unfold}_3(\mathcal{X}_0) = \text{unfold}_3(\mathcal{U} \times_3 \mathbf{V})$.

Additionally, $\mathbf{S}_0$ is generated by randomly and uniformly selecting a support set Ω of size m, and setting $\mathbf{S}_0 = \mathcal{P}_\Omega(\mathbf{E})$, where $\mathbf{E}$ is a matrix with independent Bernoulli ± 1 elements. Finally, we set $\mathbf{Y} = \mathbf{X}_0 + \mathbf{S}_0$. In all subsequent experiments, we set $M = N = 10$ and $B = 200$.

Phase Transition in Sparsity and Rank

Next, we investigate how the rank of $\mathbf{X}_0$ and the sparsity of $\mathbf{S}_0$ influence the performance of the RCCTV-RPCA, 3DCTV-RPCA, and standard RPCA models. We vary the sparsity ρ_s of $\mathbf{S}_0$ and the rank-to-dimension ratio R/B of $\mathbf{X}_0$ between 0 and 0.6, performing 30 trials for each pair (ρ_s, R) to reduce randomness. A trial is considered a successful recovery if the recovered matrix $\mathbf{X}$ satisfies $\|\mathbf{X}_0 - \mathbf{X}\|_F / \|\mathbf{X}_0\|_F < 0.05$.

Figure 3 shows that both the RCCTV-RPCA and 3DCTV-RPCA models outperform PCP[1] in recovery, achieving success rates of 36.2%, 43.4%, and 12.8%, respectively. Since the generated data is continuous along all three dimensions, it is naturally more suitable for the 3DCTV-RPCA model, which enforces continuity in three dimensions. In contrast, the RCCTV regularizer proposed in this work performs a low-rank decomposition along the third mode to accelerate computation. While the CTV regularizer can only be applied to the representative coefficients, the RCCTV approach does not explicitly account for connectivity along the third dimension, leading to a slightly smaller recoverable region compared to 3DCTV-RPCA. However, because the rank R is much smaller than the spectral dimension B (see Table 2), the RCCTV regularizer achieves significantly higher computational efficiency. Taking into account this acceleration, the phase transition results of RCCTV-RPCA are still considered highly satisfactory.



(a) PCP (b) 3DCTV-RPCA (c) RCCTV-RPCA

Figure 3: Fraction of correct recoveries across 30 trials as a function of the sparsity of $\mathbf{S}_0$ (x-axis) and the rank of $\mathbf{X}_0$ (y-axis). The phase transition diagrams of (a) PCP, (b) 3DCTV-RPCA, and (c) RCCTV-RPCA models are shown. (Yellow and black areas: successful and failed recovery cases).

Real Experiments

To further validate the RCCTV-RPCA model, we evaluate its performance on three tasks: HSI denoising, MSI denoising, and surveillance video background extraction.

Comparison Methods. We compare our approach with several classical and state-of-the-art low-rank and sparse decomposition methods, including 3DCTV-RPCA (Peng et al. 2022b), RPCA (Candès et al. 2011), MPCP (Zhan and Vaswani 2015), PCPF (Chiang, Hsieh, and Dhillon 2016), VBPRCA (Babacan et al. 2012), RegL1 (Zheng et al. 2012), WNNM (Gu et al. 2014), PRMF (Gu et al. 2014), MOG (Meng and De La Torre 2013), LRMR (Zhang et al. 2014), and LRTV (He et al. 2015). Additional implementation details are provided in the supplementary material.

Noise Setting. Denoising experiments were conducted under two types of noise: sparse noise and mixed noise. Sparse noise was simulated by adding varying levels of salt-and-pepper noise, whereas mixed noise was generated by combining sparse noise with zero-mean Gaussian noise of different standard deviations.

HSI and MSI Denoising

The HSI dataset used is *DC Mall* (Zhang et al. 2014), with dimensions $200 \times 200 \times 160$. The MSI dataset used is the *CAVE* dataset (Yasuma et al. 2010)[2], which includes 32 multispectral images, each of size $512 \times 512 \times 31$. Additional details are provided in the supplementary material. For quantitative evaluation, we use peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) (Wang et al. 2004).

[1]Robust Principal Component Analysis (RPCA) seeks principal components, and in (Candès et al. 2011), RPCA is also referred to as PCP.

[2]https://www.cs.columbia.edu/CAVE/databases/multispectral/

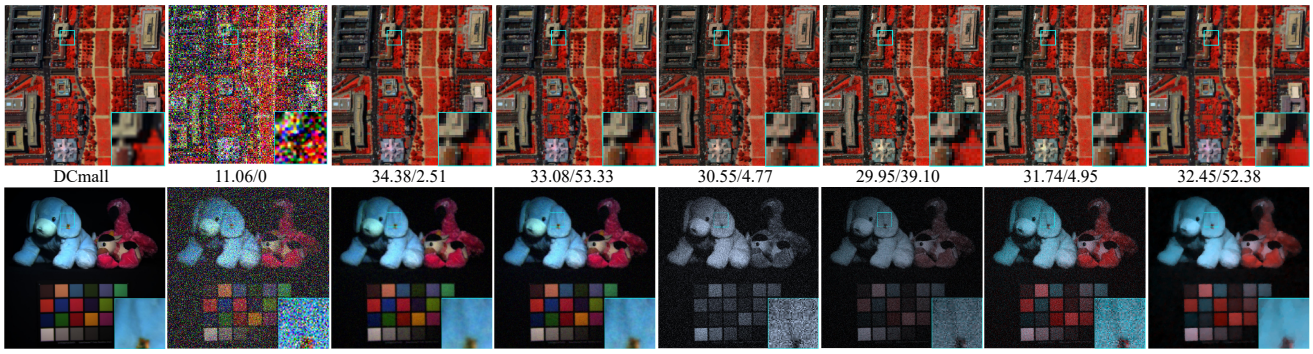


Figure 4: Pseudo-color reconstructions of HSI (bands 60-27-17) and MSI (bands 25-15-2) under Gaussian noise ($\sigma = 0.1$) and sparse noise ($\rho = 0.2$). PSNR (dB) and runtime (s) are reported below each image, which are magnified 3 times for clarity.

Noise	Metric	RCCTV -RPCA	CTV -RPCA	PCP	VBRPCA	RegL1	WNNM	PRMF	MoG	LRTV	LRMR	MPCP	PCPF
Sparse noise													
S = 0.10	psnr	**51.94**	49.42	45.03	43.13	43.64	42.26	44.11	42.95	39.28	43.62	47.10	45.33
	ssim	0.998	**0.999**	0.998	0.997	0.992	0.996	0.997	0.991	0.985	0.994	**0.999**	0.998
S = 0.20	psnr	**49.05**	48.28	43.54	42.40	43.50	41.90	44.15	37.63	39.05	41.73	45.10	43.76
	ssim	**0.999**	**0.999**	0.998	0.998	0.991	0.996	0.997	0.976	0.979	0.996	0.999	0.998
S = 0.30	psnr	**47.57**	46.98	41.31	41.43	38.22	38.43	44.03	37.41	35.61	41.23	42.83	41.32
	ssim	**0.999**	0.999	0.997	0.997	0.980	0.990	0.997	0.975	0.967	0.996	0.998	0.997
S = 0.40	psnr	**45.65**	45.18	38.59	40.12	35.37	36.05	43.54	35.66	35.26	40.51	39.95	38.60
	ssim	**0.998**	**0.998**	0.997	0.996	0.967	0.985	0.997	0.963	0.967	0.995	0.996	0.994
Sparse noise and Gaussian noise													
G=0.05 S=0.10	psnr	**38.97**	37.89	35.30	34.66	37.08	34.92	36.83	38.36	36.24	35.56	34.96	35.27
	ssim	**0.986**	0.983	0.981	0.963	0.973	0.977	0.978	0.970	0.972	0.977	0.981	0.976
G=0.10 S=0.10	psnr	**34.96**	33.81	31.37	30.32	31.31	32.51	31.90	33.63	33.54	32.41	31.94	31.35
	ssim	**0.967**	0.957	0.952	0.915	0.901	0.951	0.921	0.943	0.944	0.946	0.952	0.951
G=0.05 S=0.20	psnr	**35.54**	33.49	31.47	30.21	30.96	31.41	31.92	32.51	31.46	31.09	31.71	31.56
	ssim	**0.968**	0.964	0.952	0.917	0.912	0.951	0.924	0.934	0.946	0.930	0.957	0.953
G=0.10 S=0.20	psnr	**31.75**	30.16	29.27	28.29	28.84	29.58	30.22	29.52	29.14	28.80	29.45	29.41
	ssim	**0.934**	0.926	0.906	0.890	0.879	0.916	0.913	0.904	0.907	0.912	0.912	0.911
Mean Times(s)		7.10	85.87	**4.98**	47.97	19.51	9.56	65.79	1214	84.77	66.86	540.7	14.99

Table 3: Quantitative comparison of all methods on the *DC Mall* dataset under varying sparse (S) and Gaussian (G) noise levels. Each value denotes the average of five runs. The best and second-best results are shown in **bold** and underlined, respectively.

Methods	data										Mean Times(s)
	curt.	airp.	foun.	esca.	shop.	boot.	lobb.	camp.	wate.	Average	
RCCTV-RPCA	0.9126	**0.9168**	**0.9506**	0.9144	**0.9504**	**0.9475**	**0.9274**	**0.8936**	**0.9432**	**0.9280**	0.22
3DCTV-RPCA	0.8922	0.9158	0.9375	**0.9145**	0.9497	0.9300	0.9158	0.8878	0.9399	0.9200	3.01
PCP	0.8721	0.8721	0.9417	0.9049	0.9450	0.9131	0.9169	0.8919	0.8357	0.8990	0.41
RegL1	0.8899	0.8977	0.9194	0.4159	0.9423	0.8819	0.9249	0.8871	0.8920	0.8501	1.33
PRMF	0.8820	0.8894	0.9164	0.9061	0.9413	0.8846	0.9215	0.8865	0.8791	0.9007	4.27
PCPF	0.8720	0.8720	0.9419	0.9048	0.9438	0.9135	0.9150	0.8912	0.8323	0.8984	0.34
GODEC	0.9049	0.9143	0.9419	0.9112	0.9502	0.9252	0.9238	0.8877	0.8899	0.9165	0.14
OMoGMF	**0.9130**	0.9001	0.9099	0.9125	0.9210	0.9336	0.9046	0.8693	0.9370	0.9112	**0.06**

Table 4: AUC results of competing methods on the Li dataset. Values are averaged over annotated frames.

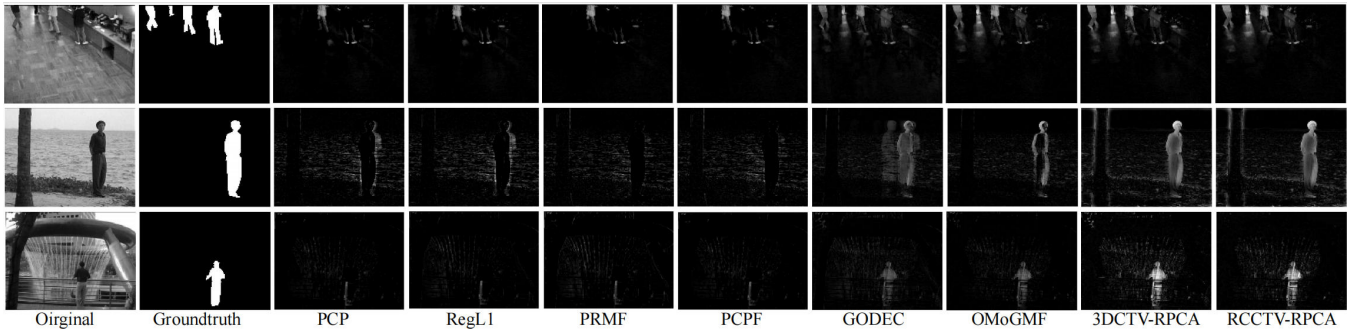


Figure 5: Visual comparison of foreground restoration across methods. Columns show original frames, ground-truth masks, and method outputs. Rows correspond to *Watersurface*, *Shoppingmall*, and *Fountain*.

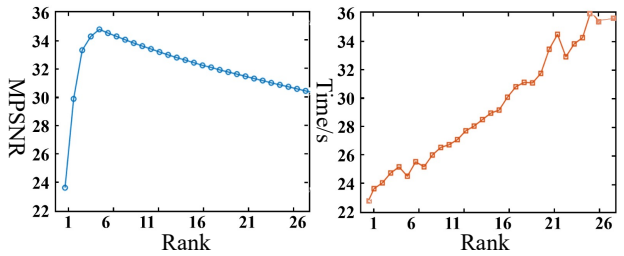


Figure 6: Sensitivity analysis of rank R under Gaussian noise with variance of 0.1 and sparse noise with percentage of 0.2 on *DC Mall* dataset.

Table 3 reports quantitative results on the *DC Mall* dataset, where RCCTV-RPCA model consistently outperforms both baseline and state-of-the-art methods across all metrics. Figure 4 shows pseudo-color visual comparisons under different noise conditions, demonstrating that RCCTV-RPCA better preserves color fidelity and spatial smoothness while effectively handling sparse and mixed noise. PSNR and runtime (in seconds) are listed below each image, highlighting its computational efficiency. By modeling representative coefficients in a low-rank subspace, RCCTV-RPCA achieves high-quality restoration at lower cost than full-tensor methods. Additional results are provided in the supplementary material.

Video Foreground-Background Separation

This task is a classic RPCA application, where the video background is modeled as the low-rank component and the moving foreground as the sparse component. We evaluated performance using the *Li* dataset[3] for evaluation. As reported in Table 4, both RCCTV-RPCA and 3DCTV-RPCA achieve outstanding performance in background-foreground separation. Compared to 3DCTV-RPCA, RCCTV-RPCA offers a clear advantage in runtime, demonstrating the efficiency gains from subspace modeling and operating on the compact representative coefficients.

Figure 5 compares the background–foreground separation results of different methods. Both RCCTV-RPCA and 3DCTV-RPCA effectively extract moving objects while preserving clean and stable backgrounds. RCCTV-RPCA further reduces ghosting artifacts and better maintains temporal consistency across frames, especially under illumination changes and dynamic scene variations. Overall, RCCTV-RPCA achieves superior accuracy, visual quality, and computational efficiency, demonstrating the effectiveness of integrating low-rank subspace modeling with locally smooth representative coefficients.

[3]http://perception.i2r.a-star.edu.sg/bkmodel/bkindex.html

Ablation Experiments

For model (9), performance is mainly governed by two hyperparameters: the rank R and the trade-off parameter λ.

Influence of Rank. The rank R affects both reconstruction accuracy and computational cost. Larger R values capture richer spectral correlations but incur higher complexity. As shown in Fig. 6, RCCTV-RPCA remains stable over a wide range of R under complex noise on the DC Mall dataset, indicating robustness to rank selection. While runtime increases approximately linearly with R, reconstruction performance saturates beyond a moderate rank. In practice, setting R between 6 and 10 achieves a good balance between accuracy and efficiency.

Influence of trade-off parameter λ. The parameter λ balances the low-rank component and sparse residuals. According to RPCA theory, setting $\lambda = 1/\sqrt{n_1}$ (with n_1 denoting the first dimension of the unfolded data) provides a principled trade-off between the nuclear norm and ℓ_1 norm terms. Empirically, adjusting λ based on noise characteristics can further improve performance. Ablation results in the supplementary material show that appropriate tuning of λ consistently enhances PSNR and SSIM. In the absence of prior knowledge, $\lambda = 1/\sqrt{n_1}$ remains a reliable default choice.

Conclusion

This paper introduces a fast RCCTV regularizer that jointly captures low-rank structure and local smoothness of representative coefficients in a low-dimensional subspace. Based on this regularizer, we develop the RCCTV-RPCA model and establish its exact recovery guarantee. Extensive experiments demonstrate the effectiveness and scalability of RCCTV for large-scale structured visual data recovery.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China under Grant 12401674, the Guangdong Basic and Applied Basic Research Foundation under Grant 2025A1515011453, the Fundamental Research Funds for the Central Universities under Grant D5000240095, and the Xi'an Municipal Science and Technology Bureau under Project 2025JH-GXKJRC-0149.

References

Babacan, S. D.; Luessi, M.; Molina, R.; and Katsaggelos, A. K. 2012. Sparse Bayesian methods for low-rank matrix estimation. *IEEE Transactions on Signal Processing*, 60(8): 3964–3977.

Bouwmans, T.; Javed, S.; Zhang, H.; Lin, Z.; and Otazo, R. 2018. On the applications of robust PCA in image and video processing. *Proceedings of the IEEE*, 106(8): 1427–1457.

Boyd, S.; Parikh, N.; and Chu, E. 2011. *Distributed optimization and statistical learning via the alternating direction method of multipliers*. Now Publishers Inc.

Candès, E. J.; Li, X.; Ma, Y.; and Wright, J. 2011. Robust principal component analysis? *Journal of the ACM*, 58(3): 11.

Candès, E. J.; and Recht, B. 2009. Exact matrix completion via convex optimization. *Foundations of Computational Mathematics*, 9(6): 717.

Cao, W.; Wang, Y.; Sun, J.; Meng, D.; Yang, C.; Cichocki, A.; and Xu, Z. 2016. Total variation regularized tensor RPCA for background subtraction from compressive measurements. *IEEE Transactions on Image Processing*, 25(9): 4075–4090.

Cao, X.; Yang, L.; and Guo, X. 2015. Total variation regularized RPCA for irregularly moving object detection under dynamic background. *IEEE Transactions on Cybernetics*, 46(4): 1014–1027.

Chen, Y. 2015. Incoherence-optimal matrix completion. *IEEE Transactions on Information Theory*, 61(5): 2909–2923.

Chiang, K.-Y.; Hsieh, C.-J.; and Dhillon, I. 2016. Robust principal component analysis with side information. In *International Conference on Machine Learning*, 2291–2299.

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

He, W.; Zhang, H.; Zhang, L.; and Shen, H. 2015. Total-variation-regularized low-rank matrix factorization for hyperspectral image restoration. *IEEE Transactions on Geoscience and Remote Sensing*, 54(1): 178–188.

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

Lu, C.; Feng, J.; Chen, Y.; Liu, W.; Lin, Z.; and Yan, S. 2019. Tensor robust principal component analysis with a new tensor nuclear norm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(4): 925–938.

Meng, D.; and De La Torre, F. 2013. Robust matrix factorization with unknown noise. In *Proceedings of the IEEE International Conference on Computer Vision*, 1337–1344.

Peng, J.; Luo, Y.; Cao, X.; Xu, S.; and Meng, D. 2025a. Beyond low-rankness: guaranteed matrix recovery via modified nuclear norm. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence*, 6066–6074.

Peng, J.; Wang, H.; Cao, X.; Jia, X.; Zhang, H.; and Meng, D. 2024. Stable Local-Smooth Principal Component Pursuit. *SIAM Journal on Imaging Sciences*, 17(2): 1182–1205.

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, H.; Cao, X.; and Xu, S. 2025b. Fast guaranteed tensor recovery with adaptive tensor nuclear norm. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence*, 6057–6065.

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.

Peng, J.; Zeng, D.; Ma, J.; Wang, Y.; and Meng, D. 2018. CPCT-LRTDTV: cerebral perfusion CT image restoration via a low rank tensor decomposition with total variation regularization. In *Medical Imaging 2018: Physics of Medical Imaging*, volume 10573, 1057337.

Rudin, L. I.; Osher, S.; and Fatemi, E. 1992. Nonlinear total variation based noise removal algorithms. *Physica D Nonlinear Phenomena*, 60(1–4): 259–268.

Wang, H.; Peng, J.; Qin, W.; Wang, J.; and Meng, D. 2023. Guaranteed tensor recovery fused low-rankness and smoothness. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(9): 10990 – 11007.

Wang, Y.; Peng, J.; Zhao, Q.; Leung, Y.; Zhao, X.-L.; and Meng, D. 2017. Hyperspectral image restoration via total variation regularized low-rank tensor decomposition. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(4): 1227–1243.

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.

Yasuma, F.; Mitsunaga, T.; Iso, D.; and Nayar, S. K. 2010. Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum. *IEEE Transactions on Image Processing*, 19(9): 2241–2253.

Zhan, J.; and Vaswani, N. 2015. Robust PCA with partial subspace knowledge. *IEEE Transactions on Signal Processing*, 63(13): 3332–3347.

Zhang, F.; Wang, J.; Wang, W.; and Xu, C. 2020. Low-tubal-rank plus sparse tensor recovery with prior subspace information. *IEEE transactions on Pattern Analysis and Machine Intelligence*, 43(10): 3492–3507.

Zhang, H.; He, W.; Zhang, L.; Shen, H.; and Yuan, Q. 2014. Hyperspectral Image Restoration Using Low-Rank Matrix Recovery. *IEEE Transactions on Geoscience and Remote Sensing*, 52(8): 4729–4743.

Zheng, Y.; Liu, G.; Sugimoto, S.; Yan, S.; and Okutomi, M. 2012. Practical low-rank matrix approximation under robust l 1-norm. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 1410–1417.