

Tensor Completion via Complementary Global, Local, and Nonlocal Priors

Xi-Le Zhao¹, Member, IEEE, Jing-Hua Yang, Graduate Student Member, IEEE, Tian-Hui Ma²,
Tai-Xiang Jiang³, Member, IEEE, Michael K. Ng⁴, Senior Member, IEEE, and Ting-Zhu Huang⁵

Abstract—Completing missing entries in multidimensional visual data is a typical ill-posed problem that requires appropriate exploitation of prior information of the underlying data. Commonly used priors can be roughly categorized into three classes: global tensor low-rankness, local properties, and nonlocal self-similarity (NSS); most existing works utilize one or two of them to implement completion. Naturally, there arises an interesting question: can one concurrently make use of multiple priors in a unified way, such that they can collaborate with each other to achieve better performance? This work gives a positive answer by formulating a novel tensor completion framework which can simultaneously take advantage of the global-local-nonlocal priors. In the proposed framework, the tensor train (TT) rank is adopted to characterize the global correlation; meanwhile, two Plug-and-Play (PnP) denoisers, including a convolutional neural network (CNN) denoiser and the color block-matching and 3 D filtering (CBM3D) denoiser, are incorporated to preserve local details and exploit NSS, respectively. Then, we design a proximal alternating minimization algorithm to efficiently solve this model under the PnP framework. Under mild conditions, we establish the convergence guarantee of the proposed algorithm. Extensive experiments show that these priors organically benefit from each other to achieve state-of-the-art performance both quantitatively and qualitatively.

Index Terms—Tensor train rank, convolutional neural network, plug-and-play, color block-matching and 3 D filtering, proximal alternating minimization, alternating direction method of multipliers.

I. INTRODUCTION

TENSOR is the multidimensional generalization of vector and matrix, which has recently attracted much attention in various real-world applications. However, in the data acquisition process, observed tensor data often suffers from damaged entries, which severely damage the data quality and hinder subsequent applications. Tensor completion (TC) aims at estimating the missing or damaged entries of the underlying data, which can improve the data quality and is beneficial for subsequent applications, such as classification [1], target detection, and recognition [2]. Therefore, tensor completion is one of the most important problems in computer vision, image processing, and machine learning. TC has received increasing attentions and achieved success in various applications, such as video recovery [3]–[7], hyperspectral image recovery [8]–[12], hyperspectral compression [13], traffic data recovery [14]–[16], seismic data recovery [17], [18], recommender systems [19], high speed video completion [20], and cloud removal [21], [22].

Generally, this ill-posed inverse problem can be tackled via the maximum a posteriori (MAP) estimation with the Bayes' rule [23]. As for TC, since the degradation process is determined, the likelihood term (corresponding to the data fidelity term) can be formulated as a Dirac delta function or a constraint with a projection operator, which enforces the solution being consistent with the observed data in specific locations. Thus, the key issue of TC under MAP framework is to reasonably analysis and effectively utilize the prior knowledge of the underlying data. When dealing with multidimensional visual data, three types of prior knowledge are widely investigated, namely, the global, local, and nonlocal priors. Below we give a brief introduction of each type of priors; more details can be found in Section II.

Although being of a mass volume, real-world multidimensional data, such as high-order web links [24] and seismic data [25], are always inner structured and globally correlated. It also goes for multidimensional visual data. For example, the bands of a hyperspectral image are highly correlated such that its spectral vectors live in a low-dimensional subspace [26]. This low-dimensionality can be mathematically formulated as low-rankness, i.e., representing the high-dimensional data under learned lower-dimensional bases. Although there has many definitions of the tensor rank, four mainstream works are the CANDECOMP/PARAFAC (CP) rank [27], the Tucker rank [28], the tubal rank [29], and the tensor train (TT) rank [30]. Regularizing the global low-rankness

Manuscript received June 17, 2020; revised May 14, 2021 and August 24, 2021; accepted December 15, 2021. Date of publication December 31, 2021; date of current version January 10, 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 61876203, Grant 61772003, Grant 11901450, Grant 12001446, and Grant 12171072; in part by the Key Project of Applied Basic Research in Sichuan Province under Grant 2020YJ0216; in part by the Applied Basic Research Project of Sichuan Province under Grant 2021YJ0107; in part by the Hong Kong Research Grant Council (HKRGC) through the General Research Fund (GRF) under Grant 12300218, Grant 12300519, Grant 17201020, and Grant 17300021; in part by the National Key Research and Development Program of China under Grant 2020YFA0714001; and in part by the Macao Science and Technology Development Fund through Macao Funding Scheme for Key Research and Development Projects under Grant 0025/2019/AKP. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Xiaolin Wu. (Xi-Le Zhao and Jing-Hua Yang are co-first authors.) (Corresponding author: Tai-Xiang Jiang.)

Xi-Le Zhao and Ting-Zhu Huang are with the School of Mathematical Sciences/Research Center for Image and Vision Computing, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, China (e-mail: xlzhao122003@163.com; tingzhu Huang@126.com).

Jing-Hua Yang is with the Faculty of Information Technology, Macau University of Science and Technology, Macau, China (e-mail: yangjinghua110@126.com).

Tian-Hui Ma is with the School of Science, Civil Aviation University of China, Tianjin 300300, China (e-mail: nkmth0307@126.com).

Tai-Xiang Jiang is with the School of Economic Information Engineering, Southwestern University of Finance and Economics, Chengdu, Sichuan 610074, China (e-mail: taixiangjiang@gmail.com; jiangtx@swufe.edu.cn).

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

Digital Object Identifier 10.1109/TIP.2021.3138325

TABLE I
EXAMPLES OF SOME EXISTING TENSOR COMPLETION METHODS
UTILIZING DIFFERENT PRIORS

TC methods	Types of the prior knowledge		
	Global	Local	Nonlocal
HaLRTC [34]	✓		
TSVD [33]	✓		
SiLRTC-TT [55]	✓		
MF-TV [39]	✓	✓	
MF-Framelet [56]	✓	✓	
NLS-LR [57]	✓		✓
NGmeet [58]	✓		✓
SMF-LRTC [41]	✓	✓	
NLRR-TC [54]			✓
KBR [52]			✓
TMac-TT [55]	✓		
TT-TV [59]	✓	✓	
GLON	✓	✓	✓

for TC, referred to as low-rank TC (LRTC), has shown an impressive capability for recovering real-world tensor data [24], [31]–[35].

When it comes to multidimensional visual data, the local property should not be neglected. For example, the successive frames of the videos, the adjacent bands of hyperspectral images, and the neighbor spatial pixels are often homogeneous. Multidimensional visual data invariably consist of spatial slices while the spectral (or temporal) local relation between these spatial slices are different. Therefore, in this paper, we mainly focus on the spatial local property, which widely exists on all kinds of the multidimensional visual data. Traditional visual tensor data recovery methods exploit the spatial local priors by minimizing some hand-craft regularizers, such as the total variation (TV) [36] used in [37] and [38] and the tight wavelet frame (framelet) [39], [40]. More recently, Zhao *et al.* [4] formulated an implicit regularization term by plugging a deep convolutional neural network (CNN) denoiser, i.e., the FFDNet [41], to make use of the learned deep denoising prior. Since the architecture of FFDNet is a cascade of convolution layers, and rectified linear units (ReLU) [42] layers, which are all local operations, it can be viewed as expressing the local prior knowledge.

Another important property for visual data is the nonlocal self-similarity (NSS), which describes the existence of abundant nonlocal similar structures including regular and repetitive patterns [43]. Over the past decade, NSS underlies the state-of-the-art in gray image denoising and has been fully exploited in nonlocal means [44] and block-matching and 3 D filtering (BM3D) [43]. This idea has been extended to the recovery of color images [45]–[47], ultrasound images [48], volumetric images [49], videos [50], and multi-spectral/hyperspectral images [51]–[53].

We summarize the types of prior knowledge used in some representative TC methods in Table I. One can see that existing TC methods, although achieving good performance, utilize only one or two types of the priors mentioned above. Intuitively, these three types of priors are complementary to each other since they depict the characteristics of multidimensional visual data from different views. Naturally, there arises an interesting question: can one jointly make use of multiple priors in an unified way, such that these priors can collaborate with each other to achieve better performance? This paper gives a positive answer by formulating a novel

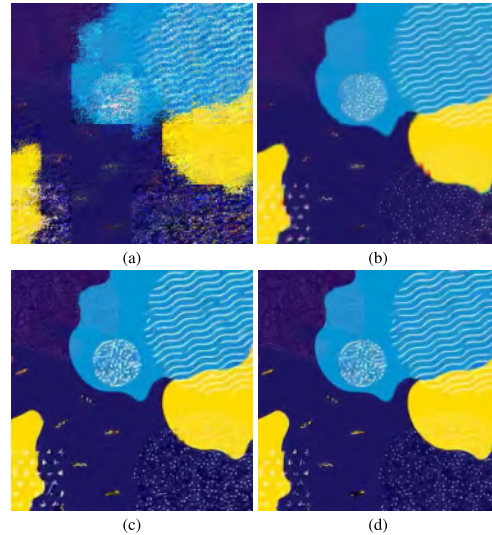


Fig. 1. An illustration of effects of different priors in the color image completion with sampling rate (SR)=0.1. (a)–(d): the recovered results by GLON without FFDNet and CBM3D, GLON without FFDNet, GLON without CBM3D, and the proposed GLON, respectively.

tensor completion framework which can simultaneously and complementarily employ these three types of priors, and we refer it as Global-Local-Nonlocal (GLON).

Considering that TMac-TT (see Sec. II.A) is an effective and efficient regularizer to expressing the global low-rankness, we directly use the TT low-rank constraint, remaining the local and nonlocal regularizers in a general formulation, to characterize the global low-rankness prior. Without loss of generality, the suggested GLON framework is formulated as

$$\begin{aligned} \min_{\mathcal{M}} \quad & \sum_{i=1}^{l-1} \frac{\alpha_i}{2} \|\mathcal{K}(\mathcal{M})_{[i]} - X_i Y_i\|_F^2 + \lambda_1 \Phi_L(\mathcal{M}) + \lambda_2 \Phi_N(\mathcal{M}), \\ \text{s.t.} \quad & U_{\Omega}(\mathcal{M}) = U_{\Omega}(\mathcal{T}), \end{aligned} \quad (1)$$

where $\mathcal{M} \in R^{n_1 \times \dots \times n_l}$ is the underlying completed tensor, $\mathcal{T} \in R^{n_1 \times \dots \times n_l}$ is the observed tensor with missing data, Ω is the index of observed entries, $U_{\Omega}(\cdot)$ is the projection operator on Ω , λ_1 and λ_2 are regularization parameters, the first term is TMac-TT expressing the global prior, and $\Phi_L(\mathcal{M})$ and $\Phi_N(\mathcal{M})$ denote regularization terms expressing the local prior and the nonlocal prior, respectively. In the experimental part, we directly adopt two denoisers, i.e., FFDNet and CBM3D, to respectively enforce the local and nonlocal priors under the Plug-and-Play (PnP) framework. Then, we design an effective proximal alternating minimization (PAM) algorithm for solving the proposed tensor completion model and establish the theoretical convergence guarantee. Extensive numerical experiments illustrate that the proposed method can generate satisfactory results, and exhibits obvious superiority over compared methods both visually and quantitatively.

In Fig. 1, we display a toy example to illustrate that the regularization terms in our model characterize different types of prior knowledge, and they are indeed complementary to each other. It can be found that, when only TMac-TT (GLON without FFDNet and CBM3D) used, we can only see the overall structure from the completion result but failed to distinguish details and textures. When TMac-TT cooperates with CBM3D (GLON without FFDNet), the performance becomes much better since the repetitive textures are well preserved.

However, the result by TMac-TT+CBM3D is a little bit blurry (See the bottom-left corner of Fig.1-(b)). Similarly, GLON without CBM3D (TMac-TT+FFDNet) acts well on the edges but destroys some repetitive patterns (See the wavy lines in the top-right corner of Fig.1-(c)). In comparison, the proposed method, which incorporates three types of the prior knowledge, deals well with the overall structure, repetitive patterns, and sharp edges, and herein achieves the best performance.

In summary, this paper mainly has three contributions.

- Based on MAP principle, we propose a unified low-rank tensor completion framework for multi-dimensional image recovery, in which different types of local and nonlocal priors knowledge can be flexibly and simultaneously utilized. The proposed framework makes the incorporated priors benefit from each other, achieving a comprehensive and multi-perspective characterization of the intrinsic properties of multidimensional images.
- Based on the above framework, we choose two denoisers to respectively express the local and nonlocal priors. More precisely, for the local prior, we plug a CNN regularizer to characterize complex image features by exploiting a large training dataset. Moreover, we use CBM3D to explore the nonlocal redundancy of data, and complement the poor generalization of CNN for different data to achieve satisfactory results.
- We design an effective strategy to solve the rather challenging optimization- and learning-based tensor completion framework and establish the theoretical convergence guarantee. More precisely, we use PAM to decompose the original problem into easier subproblems, which has closed-form solutions or can be efficiently solved by the alternating direction method of multipliers (ADMM) algorithm under the PnP framework.

The remainder of this paper is organized as follows. Section II reviews some related works. Section III introduces some preliminary notation and knowledge. Section IV presents the proposed framework with an efficient solving algorithm. Section V provides experimental results and discusses some details. Section VI summarizes this paper.

II. RELATED WORK

As previously mentioned, the prior knowledge utilized for TC can be categorized into three types: global low-rankness, local property, and nonlocal self-similarity. In this section, we go through the representative works related to these three categories.

A. Global Low-Rankness

The global low-rankness reflects the global correlation by representing the high-dimensional data under learned lower-dimensional bases. There are several representative works characterizing the tensor global low-rankness as follows.

1) *CP Rank*: The CP rank [27] is the smallest number of rank-one tensors that produce the target tensor. Although many methods have proposed to handle CP-rank minimization problem [34], it is a NP-hard problem [59], [60]. So finding the optimal low-CP-rank approximation of the target tensor is still an open problem.

2) *Tucker Rank*: The Tucker rank [28] is defined as the vector whose elements are the ranks of unfolding matrices of the target tensor. Liu *et al.* [33] used the sum of nuclear norms (SNN) as its convex surrogate to solve tensor completion problem. Xu *et al.* [61] proposed a low-rank matrix

factorization method by factorizing each unfolding matrix into the product of two smaller matrices. However, the Tucker rank is developed based on an unbalanced unfolding scheme, such that it only captures the correlation between one single mode of the tensor and all other modes [54].

3) *Tubal Rank*: The tubal rank [62] is defined as the number of non-zero singular tubes obtained by the tensor singular value decomposition of the target tensor. In [32], Zhang *et al.* proposed the tensor nuclear norm (TNN), which is a convex surrogate of the tensor tubal rank, for the LRTC. Jiang *et al.* further consider the semi-invertible transform [63] and the data-adaptive dictionary [64] for better representation of the global low-rankness within this framework. TNN is also applied for the tensor robust principal component analysis by Lu *et al.* [65]. Liu *et al.* [66] considered the low-tubal-rankness to explore the global correlations of data in tensor completion problem.

4) *Tensor Train Rank*: The TT rank [30] is the vector that its elements are the ranks of canonical matrices in the matrix product state representation of the target tensor. Bengua *et al.* [54] factorized each canonical matrix as the product of two low-rank matrices to approximate the TT rank and performed tensor completion using parallel low-rank matrix factorization method (TMac-TT). Compared with Tucker rank, TT rank is obtained by a more balanced matricization scheme, i.e., matricize the tensor along permutations of modes. Therefore, TT rank can capture the correlations between different modes, which is more suitable for higher-orders. The TT rank has shown promising results to capture the low-rank structure of high-order tensor data [54], [58], [67]. In this work, TMac-TT is introduced to characterize the global low-rankness of the multi-dimensional visual data.

B. Spatial Local Property

As mentioned in Sec. I, we mainly consider the spatial local property, which invariably exists in multidimensional visual data, instead of the spectral (or temporal) local relation, which could be inconstant for different types of data. For the 2 D image denoising, TV [36], [68] regularizer is well-known for its remarkable ability of preserving the edges and enhancing the piece-wise local continuity of the visual data. For the multidimensional visual data completion in the tensor form, Li *et al.* [37] and Ji *et al.* [38] utilized TV together with the SNN constraint and low-rank matrix factorization, respectively. Jiang *et al.* [55] further introduced used framelet to enhance the spatial smoothness of completion results by low-Tucker-rank regularization. For various image recovery tasks, Zhang *et al.* [69] learned the deep denoising prior expressed by a CNN. The representation ability of the CNN is remarkably higher than those hand-craft regularizers. In [4], Zhao *et al.* adopted FFDNet to regularize the spatial slices of the tensor visual data and their method showed promising capability of recovering the fine spatial details.

C. Nonlocal Self-Similarity

Nonlocal self-similarity (NSS) describes the existence of abundant nonlocal similar structures, such as regular and repetitive patterns, by stacking together patches of similar textures across a nature image. A representative work is the block-matching and 3 D filtering (BM3D) algorithm proposed by Dabov *et al.* [43], which combines sparsity and NSS for

image denoising. Dong *et al.* [70] exploited NSS to obtain good estimates of the sparse coding coefficients and proposed the nonlocally centralized sparse representation model for image restoration. Gu *et al.* [71] proposed the weighted nuclear norm minimization (WNNM) by combining NSS and low-rank matrix approximation. Recently, NSS-based methods have been successfully applied to multidimensional images recovery, including the LRTC problem. Xie *et al.* [53] integrated the nonlocal spatial-spectral similarity into the LRTC model. Li *et al.* [56] promoted the NSS of the tensor by employing BM3D for tensor completion. He *et al.* [57] designed the spatial non-local self-similarity regularizer with respect to the representation coefficients in the subspace spanned by left singular vectors. These methods are good at completing regular high-frequency textures, but more or less compromise local image details.

III. PRELIMINARIES

In this section, we briefly introduce some preliminary notation and knowledge used in our work.

A. Tensor Basics

The scalars, vectors, matrices, and tensors are denoted as m , \mathbf{m} , M , and \mathcal{M} , respectively. The order (or mode) of a tensor is the number of its dimensions. Given a tensor $\mathcal{M} \in R^{n_1 \times n_2 \times \dots \times n_l}$, the mode- $(1, 2, \dots, i)$ ($i = 1, 2, \dots, l - 1$) canonical matricization is denoted as $M_{[i]} \in R^{\prod_{k=1}^i n_k \times \prod_{k=i+1}^l n_k}$ [30]. It can be implemented via the function “reshape”¹ in Matlab as

$$M_{[i]} = \text{reshape}(\mathcal{M}, [\prod_{k=1}^i n_k, \prod_{k=i+1}^l n_k]).$$

We use “unreshape” denotes the inverse operator of reshape, i.e., $\text{unreshape}_{[i]}(M_{[i]}) := \mathcal{M}$. The TT rank is defined as $\mathbf{r} = (\text{rank}(M_{[1]}), \text{rank}(M_{[2]}), \dots, \text{rank}(M_{[l-1]}))$. Please refer to [30] for the details of the TT-rank.

B. Visual Data Tensorization

Using the visual data tensorization (VDT) [72], TT-based TC methods can efficiently utilize the global correlation of multidimensional visual data in a high-order manner. We introduce VDT for dimension augmentation by rearranging the tensor elements. Given a third-order visual data $\mathcal{Z} \in R^{m \times n \times p}$, the details of VDT process are as follows. Assume that the two spatial dimensions m and n have factorizations $m = \prod_{d=1}^q m_d$ and $n = \prod_{d=1}^q n_d$, we first factorize $m \times n$ to $m_1 \times m_2 \times \dots \times m_q \times n_1 \times n_2 \times \dots \times n_q$, then we permute the order to $m_1 \times n_1 \times m_2 \times n_2 \times \dots \times m_q \times n_q$ and reshape the dimension to the size $m_1 n_1 \times \dots \times m_q n_q$. Finally, we transform the original tensor \mathcal{Z} to a high-order $\tilde{\mathcal{Z}}$ of size $m_1 n_1 \times \dots \times m_q n_q \times p$. After applying the completion algorithm on the higher-order tensor, using the reverse operation of VDT on the result into the original tensor form.

C. Plug-and-Play Framework

The Plug-and-Play (PnP) framework provides a flexible way to embed different types of priors into MAP-based probability models. Given a(n) (implicit) prior $\Phi(x)$, its proximal operator of $\text{prox}_\phi : R^n \rightarrow R^n$ is given by

$$\text{prox}_\phi(y) = \arg \min_x \{\Phi(x) + \frac{\beta}{2} \|x - y\|^2\}. \quad (2)$$

¹<https://www.mathworks.com/help/matlab/ref/reshape.html>

The minimizer of (2) can be obtained through the mapping of the input y , which is equivalent to map the noisy image to the clean image. Thus, (2) can be replaced by a denoiser under the PnP framework. Thereby, we remark that $\Phi(x)$ denotes an implicit regularizer that deliver various image priors, particularly the deep priors.

PnP has been widely used in image processing. Based on the image prior information, some hand-crafted denoisers have been designed, such as TV denoiser, CBM3D denoiser [43], [45], and WNNM denoiser [73], [73]. Recently, based on training on a large number of natural image sets for specific tasks, denoisers [32] based on deep learning have become popular to learn data-driven image priors, such as FFDnet [41], MLP [74], and DCNN [75]. Under the PnP framework, these traditional denoisers and deep learning-based denoisers can be flexibly plugged as part of the modular. The above PnP framework has been successfully applied in various problems [69], [76], [77]. In addition, it has also been pioneered for higher order tensor data in [78].

IV. TENSOR COMPLETION USING GLOBAL-LOCAL-NONLOCAL PRIORS

This section is divided into four parts. Section IV.A establishes a unified low-rank tensor completion framework under the MAP framework. Section IV.B designs a PAM algorithm to optimize the proposed framework. Section IV.C establishes the theoretical convergence guarantee. Section IV.D shows the rank-increasing scheme.

A. Model Formulation

Tensor completion aims at estimating the underlying tensor data $\mathcal{M} \in R^{n_1 \times n_2 \times \dots \times n_l}$ from its partial observation \mathcal{T} with the support Ω . The MAP estimator maximizes the posterior probability $P(\mathcal{M}|\mathcal{T})$ under the Bayes rule, i.e.,

$$\begin{aligned} \mathcal{M} &= \arg \max_{\mathcal{M}} P(\mathcal{M}|\mathcal{T}) = \arg \max_{\mathcal{M}} \frac{P(\mathcal{T}|\mathcal{M})P(\mathcal{M})}{P(\mathcal{T})} \\ &= \arg \max_{\mathcal{M}} \{\log P(\mathcal{T}|\mathcal{M}) + \log P(\mathcal{M})\}. \end{aligned} \quad (3)$$

In (3), the likelihood term $P(\mathcal{T}|\mathcal{M})$ is determined by the observation process. Assuming that the observation of each entry in \mathcal{M} is independent to each other, the likelihood term can be written as

$$P(\mathcal{T}|\mathcal{M}) = \prod_{i_1 i_2 \dots i_l \in \Omega} \delta(\mathcal{M}_{i_1 i_2 \dots i_l} - \mathcal{T}_{i_1 i_2 \dots i_l}), \quad (4)$$

where $\delta(\cdot)$ is the Dirac delta function, i.e.,

$$\delta(x) = \begin{cases} \infty, & \text{if } x = 0, \\ 0, & \text{otherwise,} \end{cases}$$

and it satisfies $\int_{-\infty}^{\infty} \delta(x) dx = 1$. Maximizing (4) enforces that the estimation should strictly agree with the observation in Ω . To avoid the situation where zeros time infinity terms, we directly generalize the Dirac delta function with respect to the support Ω as

$$\delta_\Omega(\mathcal{M} - \mathcal{T}) = \begin{cases} \infty, & \text{if } \mathcal{M}_\Omega = \mathcal{T}_\Omega, \\ 0, & \text{otherwise,} \end{cases}$$

where $\mathcal{M}_\Omega = \mathcal{T}_\Omega$ indicates that \mathcal{M} equals to \mathcal{T} with respect to the support Ω and $\delta_\Omega(\cdot)$ is logarithm invariant. Consequently, the logarithm of the likelihood term turns to be

$$\log P(\mathcal{T}|\mathcal{M}) = \delta_\Omega(\mathcal{M} - \mathcal{T}).$$

Then, employing the global TT low-rankness, local, and nonlocal property of \mathcal{M} , i.e., $\mathcal{M} \sim e^{-(\phi_G(\mathcal{M})+\lambda_1\phi_L(\mathcal{M})+\lambda_2\phi_N(\mathcal{M}))}$, the prior term in (3) can be formulated as

$$\log P(\mathcal{M}) = -(\phi_G(\mathcal{M}) + \lambda_1\phi_L(\mathcal{M}) + \lambda_2\phi_N(\mathcal{M})). \quad (5)$$

Although explicitly formulating the prior distribution of \mathcal{M} is very difficult when ϕ_G , ϕ_L , and ϕ_N are complicated, we can still effectively obtain the MAP estimation as long as we have ϕ_L and ϕ_N .

By using the TT low-rank to depict the global low-rankness prior, the proposed unified low-rank TC framework is formulated as

$$\min_{X, Y, \mathcal{M}} \sum_{i=1}^{l-1} \frac{\alpha_i}{2} \|\mathcal{K}(\mathcal{M})_{[i]} - X_i Y_i\|_F^2 + \Psi_\Omega(\mathcal{M} - \mathcal{T}) + \lambda_1 \Phi_L(\mathcal{M}) + \lambda_2 \Phi_N(\mathcal{M}), \quad (6)$$

where \mathcal{K} is the VDT operator (see Sec. III.B), $\mathcal{K}(\mathcal{M}) \in R^{n_1 \times n_2 \times \dots \times n_l}$ is the augmented tensor by VDT, α_i are positive weight parameters satisfying $\sum_{i=1}^{l-1} \alpha_i = 1$, $X = (X_1, X_2, \dots, X_{l-1})$ and $Y = (Y_1, Y_2, \dots, Y_{l-1})$ are factor matrices, λ_1 and λ_2 are regularization parameters. $\Phi_L(\mathcal{M})$ and $\Phi_N(\mathcal{M})$ denote regularization terms expressing the local prior and the nonlocal prior, respectively. The indicator function $\Psi_\Omega(\cdot)$ is defined as

$$\Psi_\Omega(\mathcal{M} - \mathcal{T}) = \begin{cases} 0, & \text{if } \mathcal{M}_\Omega = \mathcal{T}_\Omega, \\ \infty, & \text{otherwise.} \end{cases}$$

Owing to the specific structure of the Dirac delta function, $\Psi_\Omega(\mathcal{M} - \mathcal{T})$ reaches its minimal value at the same point where $\delta_\Omega(\mathcal{M} - \mathcal{T})$ achieves its maximal value. Meanwhile, their roles are equivalent to constraint term $U_\Omega(\mathcal{M}) = U_\Omega(\mathcal{T})$ in (1).

B. The PAM Solver

To facilitate optimization, we use the half quadratic splitting (HQS) technique [79], [80] and introduce the variable \mathcal{P} , the (6) turns the following problem:

$$\min_{\mathcal{M}, X, Y, \mathcal{P}} \sum_{i=1}^{l-1} \frac{\alpha_i}{2} \|\mathcal{K}(\mathcal{M})_{[i]} - X_i Y_i\|_F^2 + \lambda_1 \Phi_L(\mathcal{M}) + \lambda_2 \Phi_N(\mathcal{P}) + \frac{\beta}{2} \|\mathcal{P} - \mathcal{M}\|_F^2 + \Psi_\Omega(\mathcal{M} - \mathcal{T}), \quad (7)$$

where β is a penalty parameter. Since the optimization problem (7) is nonconvex and has four variables, we design an algorithm based on PAM [81] to resolve the problem (7) into several easier subproblems. In the alternating minimization scheme, we iteratively update X , Y , \mathcal{P} , and \mathcal{M} as follows:

$$\begin{cases} X^{t+1} = \arg \min_X \{Q_1(X|X^t) \\ = F(X, Y^t, \mathcal{P}^t, \mathcal{M}^t) + \frac{\rho}{2} \|X - X^t\|_F^2\}, \\ Y^{t+1} = \arg \min_Y \{Q_2(Y|Y^t) \\ = F(X^{t+1}, Y, \mathcal{P}^t, \mathcal{M}^t) + \frac{\rho}{2} \|Y - Y^t\|_F^2\}, \\ \mathcal{P}^{t+1} = \arg \min_{\mathcal{P}} \{Q_3(\mathcal{P}|\mathcal{P}^t) \\ = F(X^{t+1}, Y^{t+1}, \mathcal{P}, \mathcal{M}^t) + \frac{\rho}{2} \|\mathcal{P} - \mathcal{P}^t\|_F^2\}, \\ \mathcal{M}^{t+1} = \arg \min_{\mathcal{M}} \{Q_4(\mathcal{M}|\mathcal{M}^t) \\ = F(X^{t+1}, Y^{t+1}, \mathcal{P}^{t+1}, \mathcal{M}) + \frac{\rho}{2} \|\mathcal{M} - \mathcal{M}^t\|_F^2\}, \end{cases} \quad (8)$$

where $F(X, Y, \mathcal{P}, \mathcal{M})$ is the objective function in (7), ρ is a positive constant, and t denotes the iteration number. Next, we discuss the details for solving the X -, Y -, \mathcal{P} -, and \mathcal{M} -subproblems.

X -subproblem: as the minimizations with respect to each X_i are decoupled, we decomposed it into $l-1$ independent problems as follows:

$$X_i^{t+1} = \arg \min_{X_i} \frac{\alpha_i}{2} \|X_i Y_i^t - \mathcal{K}(\mathcal{M}^t)_{[i]}\|_F^2 + \frac{\rho}{2} \|X_i - X_i^t\|_F^2, \quad (9)$$

which has the closed-form solution

$$X_i^{t+1} = (\alpha_i \tilde{M}_{[i]}^t (Y_i^t)^T + \rho X_i^t) (\alpha_i Y_i^t (Y_i^t)^T + \rho I)^{-1}, \quad (10)$$

where $I \in R^{r_i \times r_i}$ is an identity matrix.

Y -subproblem: similarly, the Y -subproblem is decomposed into $l-1$ independent problems as follows:

$$Y_i^{t+1} = \arg \min_{Y_i} \frac{\alpha_i}{2} \|X_i^{t+1} Y_i - \mathcal{K}(\mathcal{M}^t)_{[i]}\|_F^2 + \frac{\rho}{2} \|Y_i - Y_i^t\|_F^2, \quad (11)$$

which has the closed-form solution

$$Y_i^{t+1} = (\alpha_i (X_i^{t+1})^T X_i^{t+1} + \rho I)^{-1} ((\alpha_i X_i^{t+1})^T \tilde{M}_{[i]}^t + \rho Y_i^t). \quad (12)$$

\mathcal{P} -subproblem: the \mathcal{P} -subproblem is

$$\mathcal{P}^{t+1} = \arg \min_{\mathcal{P}} \lambda_2 \Phi_N(\mathcal{P}) + \frac{\beta}{2} \|\mathcal{P} - \mathcal{M}^t\|_F^2 + \frac{\rho}{2} \|\mathcal{P} - \mathcal{P}^t\|_F^2. \quad (13)$$

In particular, \mathcal{P}^{k+1} is computed by

$$\mathcal{P}^{k+1} = \arg \min_{\mathcal{P}} \lambda_2 \Phi_N(\mathcal{P}) + \frac{\beta + \rho}{2} \|\mathcal{P} - \frac{\beta \mathcal{M}^t + \rho \mathcal{P}^t}{\beta + \rho}\|_F^2. \quad (14)$$

Letting $(\beta \mathcal{M}^t + \rho \mathcal{P}^t)/(\beta + \rho)$ as the input of a denoiser, then we can get

$$\mathcal{P}^{k+1} = \text{D}_N((\beta \mathcal{M}^t + \rho \mathcal{P}^t)/(\beta + \rho), \sigma_2), \quad (15)$$

where $\sigma_2 = \sqrt{\lambda_2/(\beta + \rho)}$, D_N denotes the denoiser to express the nonlocal prior. Based on the PnP framework, $(\beta \mathcal{M}^t + \rho \mathcal{P}^t)/(\beta + \rho)$ is treated as the observed tensor, and \mathcal{P}^{k+1} is the denoising result by the plugged prior.

\mathcal{M} -subproblem: the \mathcal{M} -subproblem is

$$\mathcal{M}^{t+1} = \arg \min_{\mathcal{M}} \sum_{i=1}^{l-1} \frac{\alpha_i}{2} \|X_i^{t+1} Y_i^{t+1} - \mathcal{K}(\mathcal{M})_{[i]}\|_F^2 + \Psi_\Omega(\mathcal{M} - \mathcal{T}) + \lambda_1 \Phi_L(\mathcal{M}) + \frac{\beta}{2} \|\mathcal{P} - \mathcal{M}^t\|_F^2 + \frac{\rho}{2} \|\mathcal{M} - \mathcal{M}^t\|_F^2. \quad (16)$$

We iteratively solve (16) using ADMM [37]. The \mathcal{M} -subproblem is transformed into the following problem by introducing the auxiliary variable \mathcal{E} :

$$\arg \min_{\mathcal{M}} \sum_{i=1}^{l-1} \frac{\alpha_i}{2} \|X_i^{t+1} Y_i^{t+1} - \mathcal{K}(\mathcal{M})_{[i]}\|_F^2 + \lambda_1 \Phi_L(\mathcal{E}) + \Psi_\Omega(\mathcal{M} - \mathcal{T}) + \frac{\beta}{2} \|\mathcal{P} - \mathcal{M}^t\|_F^2 + \frac{\rho}{2} \|\mathcal{M} - \mathcal{M}^t\|_F^2, \quad (17)$$

s.t. $\mathcal{E} = \mathcal{M}$.

By separating the variables in (17) into two groups \mathcal{M} and \mathcal{E} , (17) fits the framework of ADMM [67], [82]. To tackle the linear constraint, the corresponding augmented Lagrangian function of (17) is

$$\begin{aligned} \mathcal{L}(\mathcal{M}, \mathcal{E}, \mathcal{F}) = & \sum_{i=1}^{l-1} \frac{\alpha_i}{2} \|X_i^{t+1} Y_i^{t+1} - \mathcal{K}(\mathcal{M})_{[i]}\|_F^2 \\ & + \lambda_1 \Phi_L(\mathcal{E}) + \frac{\beta_1}{2} \|\mathcal{E} - \mathcal{M} + \frac{\mathcal{F}}{\beta_1}\|_F^2 \\ & + \Psi_\Omega(\mathcal{M} - \mathcal{T}) + \frac{\beta}{2} \|\mathcal{P} - \mathcal{M}^t\|_F^2 \\ & + \frac{\rho}{2} \|\mathcal{M} - \mathcal{M}^t\|_F^2, \end{aligned} \quad (18)$$

where \mathcal{F} is Lagrangian multiplier and β_1 and β are penalty parameters.

Then, we use ADMM to solve the problem (17) with the following scheme:

$$\begin{cases} \mathcal{M}^{t+1,k+1} = \arg \min_{\mathcal{M}} \mathcal{L}(\mathcal{M}, \mathcal{E}^k, \mathcal{F}^k), \\ \mathcal{E}^{k+1} = \arg \min_{\mathcal{E}} \mathcal{L}(\mathcal{M}^{t+1,k+1}, \mathcal{E}, \mathcal{F}^k), \\ \mathcal{F}^{k+1} = \mathcal{F}^k + \beta_1(\mathcal{E}^{k+1} - \mathcal{M}^{t+1,k+1}). \end{cases} \quad (19)$$

Following, we give the details on each step of ADMM.

ADMM-step 1: update \mathcal{M} . $\mathcal{M}^{t+1,k+1}$ is the solution of the following least square problem:

$$\begin{aligned} \mathcal{M}^{t+1,k+1} = & \arg \min_{\mathcal{M}} \sum_{i=1}^{l-1} \frac{\alpha_i}{2} \|X_i^{t+1} Y_i^{t+1} - \mathcal{K}(\mathcal{M})_{[i]}\|_F^2 \\ & + \Psi_\Omega(\mathcal{M} - \mathcal{T}) + \frac{\beta_1}{2} \|\mathcal{E} - \mathcal{M} + \frac{\mathcal{F}}{\beta_1}\|_F^2 \\ & + \frac{\beta}{2} \|\mathcal{P} - \mathcal{M}\|_F^2 + \frac{\rho}{2} \|\mathcal{M} - \mathcal{M}^t\|_F^2. \end{aligned} \quad (20)$$

To minimize (20), we have $\Psi_\Omega(\mathcal{M} - \mathcal{T}) = 0$, i.e., $\mathcal{M}_\Omega = \mathcal{T}_\Omega$. Then, the elements in the complementary set of Ω can be obtained by solving a least square problem. Thus, $\mathcal{M}^{t+1,k+1}$ can be updated by

$$\mathcal{M}^{t+1,k+1} = U_{\Omega^c} \left(\frac{W}{1 + \beta_1 + \beta + \rho} \right) + \mathcal{T}, \quad (21)$$

where $W = \sum_{i=1}^{l-1} \alpha_i \text{reshape}_{[i]}^{-1}(X_i^{t+1} Y_i^{t+1}) + \beta_1 \mathcal{E} + \beta \mathcal{P} + \mathcal{F} + \rho \mathcal{M}^t$.

ADMM-step 2: update \mathcal{E} . \mathcal{E}^{k+1} is computed by

$$\mathcal{E}^{k+1} = \arg \min_{\mathcal{E}} \lambda_1 \Phi_L(\mathcal{E}) + \frac{\beta_1}{2} \|\mathcal{E} - \mathcal{M}^{k+1} + \mathcal{F}^k / \beta_1\|_F^2. \quad (22)$$

Letting $\mathcal{M}^{k+1} - \mathcal{F}^k / \beta_1$ as the input of a denoiser, then we can get

$$\mathcal{E}^{k+1} = D_L(\mathcal{M}^{k+1} - \mathcal{F}^k / \beta_1, \sigma_1), \quad (23)$$

where $\sigma_1 = \sqrt{\lambda_1 / \beta_1}$, D_L denotes the denoiser for expressing the local image prior.

Algorithm 1 The PAM Algorithm for Solving (7)

Input: The observed tensor $\mathcal{T} \in R^{n_1 \times \dots \times n_l}$, the index set Ω , the initial TT-rank \mathbf{r}^0 , $\Delta \mathbf{r}_n$ and \mathbf{r}^{\max} for the rank increasing scheme, and parameters $\alpha_i, i = 1, \dots, l-1, \beta, \beta_1, \lambda_1, \lambda_2, \rho, \epsilon = 2 \times 10^{-3}$.

Initialization: $X^0, Y^0, \mathcal{P}^0, \mathcal{M}^0, \mathbf{r}^0$.

```

1: while  $\|\mathcal{M}^{t+1} - \mathcal{M}^t\|_F / \|\mathcal{M}^t\|_F > \epsilon$  and  $t \leq 100$  do
2:   for  $i = 1$  to  $l-1$  do
3:     Update  $X_i$  via (10);
4:     Update  $Y_i$  via (12);
5:   end for
6:   Update  $\mathcal{P}$  via (15);
7:   while  $\|\mathcal{M}^{k+1} - \mathcal{M}^k\|_F / \|\mathcal{M}^k\|_F > \epsilon$  and  $k \leq 15$  do
8:     Update  $\mathcal{M}$  by (21);
9:     Update  $\mathcal{E}$  by (23);
10:    Update  $\mathcal{F}$  by (19);
11:   end while
        update  $\mathbf{r}$  if (30) satisfies;
12: end while
Output: The estimated tensor  $\mathcal{M}$ .

```

C. Convergence Analysis

In this part, we establish the convergence analysis of Algorithm 1. For convenience, we define the following formulas:

$$\begin{aligned} F(X, Y, \mathcal{P}, \mathcal{M}) &:= \sum_{i=1}^{l-1} \frac{\alpha_i}{2} \|\mathcal{K}(\mathcal{M})_{[i]} - X_i Y_i\|_F^2 + \lambda_1 \Phi_L(\mathcal{M}) \\ &\quad + \lambda_2 \Phi_N(\mathcal{P}) + \frac{\beta}{2} \|\mathcal{P} - \mathcal{M}\|_F^2 + \Psi_\Omega(\mathcal{M} - \mathcal{T}), \\ W(X, Y, \mathcal{P}, \mathcal{M}) &:= \sum_{i=1}^{l-1} \frac{\alpha_i}{2} \|\mathcal{K}(\mathcal{M})_{[i]} - X_i Y_i\|_F^2 + \frac{\beta}{2} \|\mathcal{P} - \mathcal{M}\|_F^2, \\ f_1(\mathcal{P}) &:= \lambda_2 \Phi_N(\mathcal{P}), \\ f_2(\mathcal{M}) &:= \lambda_1 \Phi_L(\mathcal{M}). \end{aligned}$$

Now, we establish the convergence analysis of the proposed algorithm as follows.

*Theorem 1: Assuming that the sequence $\{X^t, Y^t, \mathcal{P}^t, \mathcal{M}^t\}$ is bounded and $\Phi_L(\mathcal{M})$ and $\Phi_N(\mathcal{M})$ are KL functions, the sequence $\{X^t, Y^t, \mathcal{P}^t, \mathcal{M}^t\}$ generated by **Algorithm 1** (i.e., iterations in (8)) globally converges to a critical point of (7).*

As the process of updating in (8) is factually a special instance of the Algorithm 4 described in [81], the proof of Theorem 1 follows Theorem 6.2 of [81] if the following conditions are satisfied.

- i) F satisfies the KL property at each point,
- ii) the sufficient decrease condition ((64) in [81]),
- iii) the relative error condition ((65)-(66) in [81]).

The road map of our proof also follows this line. Before verifying these conditions, we present some related definitions and lemmas.

Definition 1 (KL property, [81]): A proper lower semi-continuous function $f: R^n \rightarrow R \cup +\infty$ is said to have the KL property at $\bar{x} \in \text{dom}(\partial f)$ if there exist $\eta \in (0, +\infty]$, a neighborhood U of \bar{x} , and a continuous concave function $\phi: [0, \eta) \rightarrow [0, +\infty)$ such that:

- i) $\phi(0) = 0$,
- ii) ϕ is C^1 on $(0, \eta)$,
- iii) ϕ' is positive on $(0, \eta)$,
- iv) for each $x \in U \cap [f(\bar{x}) < f(x) < f(\bar{x}) + \eta]$, the KŁ inequality hold:

$$\phi'(f(x) - f(\bar{x}))\text{dist}(0, \partial f(x)) \geq 1,$$

where the norm involved in $\text{dist}(\cdot, \cdot)$ is $\|\cdot\|_2$ and the convention $\text{dist}(0, \emptyset) := +\infty$ is used.

Lemma 1 (Theorem 3 in [83]): A semi-algebraic real valued function f is a KŁ function, i.e., f satisfies KŁ property at each $x \in \text{dom}(f)$.

We first verify the KŁ property of function $F(X, Y, \mathcal{P}, \mathcal{M})$.

Lemma 2 (KŁ Lemma): Function $F(X, Y, \mathcal{P}, \mathcal{M})$ satisfies the KŁ property at each point.

Proof: We prove that each term of $F(X, Y, \mathcal{P}, \mathcal{M})$ satisfies the KŁ property. According to the assumption of **Theorem 1**, $\Phi_L(\mathcal{M})$ and $\Phi_N(\mathcal{M})$ satisfy the KŁ property. $\|\mathcal{K}(\mathcal{M})_{[i]} - X_i Y_i\|_F^2$ is a polynomial of (X, Y, \mathcal{M}) and polynomials are semi-algebraic functions [83]. Similarly, $\|\mathcal{P} - \mathcal{M}\|_F^2$ is also a semi-algebraic function. Moreover, the indicator function $\Psi_\Omega(\cdot)$ is semi-algebraic [83]. Since the semi-algebraic function satisfies the KŁ property (see Lemma 1), therefore, the function $F(X, Y, \mathcal{P}, \mathcal{M})$ satisfies the KŁ property. \square

Then, we show that the bounded sequence $\{X^t, Y^t, \mathcal{P}^t, \mathcal{M}^t\}$ satisfies the sufficient decrease condition.

Lemma 3 (Sufficient Decrease Lemma): For $\rho > 0$, let $\{X^t, Y^t, \mathcal{P}^t, \mathcal{M}^t\}$ be the sequence generated by **Algorithm 1**, then

$$\begin{aligned} & F(X^{t+1}, Y^t, \mathcal{P}^t, \mathcal{M}^t) + \frac{\rho}{2} \|X^{t+1} - X^t\|_F^2 \\ & \leq F(X^t, Y^t, \mathcal{P}^t, \mathcal{M}^t), \\ & F(X^{t+1}, Y^{t+1}, \mathcal{P}^t, \mathcal{M}^t) + \frac{\rho}{2} \|Y^{t+1} - Y^t\|_F^2 \\ & \leq F(X^{t+1}, Y^t, \mathcal{P}^t, \mathcal{M}^t), \\ & F(X^{t+1}, Y^{t+1}, \mathcal{P}^{t+1}, \mathcal{M}^t) + \frac{\rho}{2} \|\mathcal{P}^{t+1} - \mathcal{P}^t\|_F^2 \\ & \leq F(X^{t+1}, Y^{t+1}, \mathcal{P}^t, \mathcal{M}^t), \\ & F(X^{t+1}, Y^{t+1}, \mathcal{P}^{t+1}, \mathcal{M}^{t+1}) + \frac{\rho}{2} \|\mathcal{M}^{t+1} - \mathcal{M}^t\|_F^2 \\ & \leq F(X^{t+1}, Y^{t+1}, \mathcal{P}^{t+1}, \mathcal{M}^t). \end{aligned} \quad (24)$$

Proof: According to Q_1, Q_2, Q_3 , and Q_4 defined in (8), when $X^{t+1}, Y^{t+1}, \mathcal{P}^{t+1}$, and \mathcal{M}^{t+1} are minimizers of Q_1, Q_2, Q_3 , and Q_4 , we have

$$\begin{aligned} & F(X^{t+1}, Y^t, \mathcal{P}^t, \mathcal{M}^t) + \frac{\rho}{2} \|X^{t+1} - X^t\|_F^2 \\ & = Q_1(X^{t+1}|X^t) \leq Q_1(X^t|X^t) = F(X^t, Y^t, \mathcal{P}^t, \mathcal{M}^t), \\ & F(X^{t+1}, Y^{t+1}, \mathcal{P}^t, \mathcal{M}^t) + \frac{\rho}{2} \|Y^{t+1} - Y^t\|_F^2 \\ & = Q_2(Y^{t+1}|Y^t) \leq Q_2(Y^t|Y^t) = F(X^{t+1}, Y^t, \mathcal{P}^t, \mathcal{M}^t), \\ & F(X^{t+1}, Y^{t+1}, \mathcal{P}^{t+1}, \mathcal{M}^t) + \frac{\rho}{2} \|\mathcal{P}^{t+1} - \mathcal{P}^t\|_F^2 \\ & = Q_3(\mathcal{P}^{t+1}|\mathcal{P}^t) \leq Q_3(\mathcal{P}^t|\mathcal{P}^t) = F(X^{t+1}, Y^{t+1}, \mathcal{P}^t, \mathcal{M}^t), \\ & F(X^{t+1}, Y^{t+1}, \mathcal{P}^{t+1}, \mathcal{M}^{t+1}) + \frac{\rho}{2} \|\mathcal{M}^{t+1} - \mathcal{M}^t\|_F^2 \\ & = Q_4(\mathcal{M}^{t+1}|\mathcal{M}^t) \leq Q_4(\mathcal{M}^t|\mathcal{M}^t) \\ & = F(X^{t+1}, Y^{t+1}, \mathcal{P}^{t+1}, \mathcal{M}^t). \end{aligned}$$

The proof of the sufficient decrease condition is completed. \square

Next, we show the bounded sequence $\{X^t, Y^t, \mathcal{P}^t, \mathcal{M}^t\}$ satisfies the relative error condition.

Lemma 4 (Relative Error Lemma): The sequence $\{X^t, Y^t, \mathcal{P}^t, \mathcal{M}^t\}$ is generated by **Algorithm 1** and $\rho > 0$. Then, there exist $V_1^{t+1}, V_2^{t+1}, V_3^{t+1}$, and V_4^{t+1} satisfying

$$\begin{aligned} & \|V_1^{t+1} + \nabla_X W(X^{t+1}, Y^t, \mathcal{P}^t, \mathcal{M}^t)\|_F \leq \rho \|X^{t+1} - X^t\|_F, \\ & \|V_2^{t+1} + \nabla_Y W(X^{t+1}, Y^{t+1}, \mathcal{P}^t, \mathcal{M}^t)\|_F \leq \rho \|Y^{t+1} - Y^t\|_F, \\ & \|V_3^{t+1} + \nabla_{\mathcal{P}} W(X^{t+1}, Y^{t+1}, \mathcal{P}^{t+1}, \mathcal{M}^t)\|_F \\ & \leq \rho \|\mathcal{P}^{t+1} - \mathcal{P}^t\|_F, \\ & \|V_4^{t+1} + \nabla_{\mathcal{M}} W(X^{t+1}, Y^{t+1}, \mathcal{P}^{t+1}, \mathcal{M}^{t+1})\|_F \\ & \leq \rho \|\mathcal{M}^{t+1} - \mathcal{M}^t\|_F. \end{aligned} \quad (25)$$

Proof: Note that $X^{t+1}, Y^{t+1}, \mathcal{P}^{t+1}$, and \mathcal{M}^{t+1} are, respectively, optimal solutions of Q_1, Q_2, Q_3 , and Q_4 . For each subproblem, we have

$$\begin{cases} 0 = a_i(X_i^{t+1}Y_i^t - \mathcal{K}(\mathcal{M})_{[i]}^t)Y_i^{tT} + \rho(X_i^{t+1} - X_i^t), \\ 0 = a_i X_i^{tT}(X_i^{t+1}Y_i^t - \mathcal{K}(\mathcal{M})_{[i]}^t) + \rho(X_i^{t+1} - X_i^t), \\ 0 \in \partial_{\mathcal{P}} f_1(\mathcal{P}^{t+1}) + \nabla_{\mathcal{P}} W(X^{t+1}, Y^{t+1}, \mathcal{P}^{t+1}, \mathcal{M}^t) \\ \quad + \rho(\mathcal{P}^{t+1} - \mathcal{P}^t), \\ 0 \in \partial \Psi_\Omega(\mathcal{M}^{t+1}) + \partial_{\mathcal{M}} f_2(\mathcal{M}^{t+1}) + \nabla_{\mathcal{M}} \\ \quad \mathcal{P}^{t+1}, W(X^{t+1}, Y^{t+1}, \mathcal{M}^{t+1}) + \rho(\mathcal{M}^{t+1} - \mathcal{M}^t). \end{cases} \quad (26)$$

Then we define V_1, V_2, V_3 , and V_4 as

$$\begin{cases} V_1^{t+1} := 0, \\ V_2^{t+1} := 0, \\ V_3^{t+1} := -\nabla_{\mathcal{P}} W(X^{t+1}, Y^{t+1}, \mathcal{P}^{t+1}, \mathcal{M}^t) \\ \quad - \rho(\mathcal{P}^{t+1} - \mathcal{P}^t), \\ V_4^{t+1} := -\nabla_{\mathcal{M}} W(X^{t+1}, Y^{t+1}, \mathcal{P}^{t+1}, \mathcal{M}^{t+1}) \\ \quad - \rho(\mathcal{M}^{t+1} - \mathcal{M}^t). \end{cases} \quad (27)$$

It is clear that $V_3^{t+1} \in \partial_{\mathcal{P}} f_1(\mathcal{P}^{t+1})$ and $V_4^{t+1} \in \partial \Psi_\Omega(\mathcal{M}) + \partial_{\mathcal{M}} f_2(\mathcal{M}^{t+1})$. Thus, we have

$$\begin{aligned} & \|V_1^{t+1} + \nabla_X W(X^{t+1}, Y^t, \mathcal{P}^t, \mathcal{M}^t)\|_F \\ & \leq \rho \|X^{t+1} - X^t\|_F, \\ & \|V_2^{t+1} + \nabla_Y W(X^{t+1}, Y^{t+1}, \mathcal{P}^t, \mathcal{M}^t)\|_F \\ & \leq \rho \|Y^{t+1} - Y^t\|_F, \\ & \|V_3^{t+1} + \nabla_{\mathcal{P}} W(X^{t+1}, Y^{t+1}, \mathcal{P}^{t+1}, \mathcal{M}^t)\|_F \\ & = \rho \|\mathcal{P}^{t+1} - \mathcal{P}^t\|_F \leq \rho \|\mathcal{P}^{t+1} - \mathcal{P}^t\|_F, \\ & \|V_4^{t+1} + \nabla_{\mathcal{M}} W(X^{t+1}, Y^{t+1}, \mathcal{P}^{t+1}, \mathcal{M}^{t+1})\|_F \\ & = \rho \|\mathcal{M}^{t+1} - \mathcal{M}^t\|_F \leq \rho \|\mathcal{M}^{t+1} - \mathcal{M}^t\|_F. \end{aligned} \quad (28)$$

The proof of the relative error condition is completed. \square

Now, we establish the proof of **Theorem 1**.

Proof: [Proof of **Theorem 1**] According to Lemma 2, $F(X, Y, \mathcal{P}, \mathcal{M})$ satisfies the KŁ property.

Combining Lemmas 3 and 4, we show that the bounded sequence $\{X^t, Y^t, \mathcal{P}^t, \mathcal{M}^t\}$ satisfies the sufficient decrease condition and relative error condition. In fact, Lemmas 3 and 4 correspond to the (64)–(66) in [81]. Under these conditions, this proof conforms to Theorem 6.2 in [81].

Therefore, the bounded sequence $\{X^t, Y^t, \mathcal{P}^t, \mathcal{M}^t\}$ globally converges to a critical point of (7). The proof of **Theorem 1** is completed. \square

In the following, we show that there are a large number of denoisers satisfying the conditions of Theorem 1, such as TV [84], WNNM [73], and BM3D [85].

Remark 1: For many state-of-the-art denoisers, e.g., TV, WNNM, and BM3D, the corresponding regularizers satisfy the KL property.

Proof: The proximal operator of the regularizer $\text{Prox}_\Phi : R^{n_1 \times n_2} \rightarrow R^{n_1 \times n_2}$ is defined as

$$\text{Prox}_\Phi(B) = \arg \min_M \Phi(M) + \frac{\beta}{2} \|M - B\|_F^2. \quad (29)$$

Under the PnP framework, since the problem (29) can be regard as a denoising problem, the proximal operator Prox_Φ can be replaced by denoising algorithms (termed as denoisers), which map the noisy image to the clean image.

The corresponding subproblem of (anisotropic) TV denoiser has the explicit mathematical form

$$\arg \min_M \|DM\|_1 + \frac{\beta}{2} \|M - B\|_F^2,$$

where D is a derivative operator. Since l_1 -norm is semi-algebraic [83] and D is a linear operator, the composition $\|DM\|_1$ is semi-algebraic. Therefore, the regularizer $\Phi_{\text{TV}}(M) := \|DM\|_1$ is a semi-algebraic function and satisfies the KL property.

The corresponding subproblem of WNNM denoiser has the explicit mathematical form

$$\arg \min_M \sum_j \|R_j(M)\|_{w,*} + \frac{\beta}{2} \sum_j \|R_j(M) - R_j(B)\|_F^2,$$

where R_j is the operator which extracts the similar patches with respect to the j -th key patch, $\|R_j(M)\|_{w,*} = \sum_k w_k \sigma_k(R_j(M))$, and $\sigma_k(R_j(M))$ denotes the k -th singular value of $R_j(M)$ [73]. According to Lemma 6 in [86], the $\|R_j(M)\|_{w,*}$ is semi-algebraic. Since finite sums of semi-algebraic functions is semi-algebraic [83], the regularizer $\Phi_{\text{WNNM}}(M) := \sum_j \|R_j(M)\|_{w,*}$ satisfies the KL property.

The corresponding subproblem of BM3D denoiser has the explicit mathematical form

$$\arg \min_M \|\psi(M)\|_p + \frac{\beta}{2} \|M - B\|_F^2,$$

where ψ is the analysis operator of BM3D [85]. Since l_p -norm is a semi-algebraic function [83], [86] and ψ is a linear operator [85], the composition $\|\psi(M)\|_p$ is a semi-algebraic function. Thus, the regularizer $\Phi_{\text{BM3D}}(M) := \|\psi(M)\|_p$ satisfies the KL property. The proof is completed. \square

The Remark 1 can be easily extended to the color image case. For example, the KL property of CBM3D which is an extension of BM3D on color images can be similarly derived.

Inspired by the success of CNN for image denoising, by exploiting the large training dataset, we choose the CNN denoiser, i.e. FFDNet [41], to preserve the local image details. Besides, due to the superior ability of CBM3D [45] in preserving NSS, we use the traditional CBM3D denoiser to express the nonlocal image prior. In the Algorithm 1, we plug in FFDNet denoiser and CBM3D denoiser under the PnP framework to solve their related subproblem. We remark that both two denoisers are applied on spatial slices. For color videos, we feed the data into the denoisers frame by frame. FFDNet is expected to preserve the abundant image spatial local details by bringing in the external prior learned from the

training data of. CBM3D can explore and enhance the nonlocal structure, including fine and repetitive patterns.

Remark 2: How to formulate the corresponding regularizer of the deep denoiser is still an open problem. However, this fact should not prevent us to use such state-of-the-art CNN denoisers. A very recent and related work by Ryu *et al.* [87] shows that the ADMM-PnP algorithm would converge to a stationary point if the CNN denoiser is properly trained, i.e., satisfying the Lipschitz condition. Thus, though we can't guarantee the KL property, the convergence of the inner loop for the FFDNet related subproblem would empirically contributes to the convergence of the outer loop, i.e., the PAM iterations.

D. Rank-Increasing Scheme

The estimation of TT-rank $\mathbf{r} = (M_{[1]}, M_{[2]}, \dots, M_{[l-1]})$ is important to capture global structures of the underlying tensor. Therefore, we adopt a rank-increasing scheme to adjust it automatically [61]. Considering that matricizing the tensor along permutations makes the size of the middle unfolding matrices more balanced than those of the border matrices, the TT-rank mainly depends on the middle unfolding matrices. Thus, we set the initial TT-rank as $\mathbf{r} = (n_1, n_1 n_2, r_3^0, r_4^0, \dots, r_{l-3}^0, n_{l-1} n_l, n_l)$ and increase r_n^{t+1} to $\min(r_n + \Delta r_n, r_n^{\max})$ at iteration $t + 1$ if

$$\left| 1 - \frac{\|P_{\Omega^c}(X_i^{t+1} Y_i^{t+1})\|_F}{\|P_{\Omega^c}(X_i^t Y_i^t)\|_F} \right| < 10^{-2}, \quad i = 3, 4, \dots, l-3, \quad (30)$$

where Δr_n is a positive integer and r_n^{\max} denotes the maximal rank estimate. Specifically, when the r_n increased at iteration $k + 1$, the X_i^{t+1} will be updated to $[X_i^t, \text{rand}(\Pi_{k=1}^i n_k)]$ and Y_i^{t+1} will be updated to $[Y_i^t, \text{rand}(\Pi_{k=i+1}^l n_k)]$, i.e., adding randomly generated columns Δr_n to X_i^t and randomly generated rows Δr_n to Y_i^t . By this scheme, we can obtain a low-TT-rank estimation adapted to the target tensor.

Finally, our algorithm is summarized in Algorithm 1.

V. NUMERICAL EXPERIMENTS

In this section, we test the performance of the proposed method (referred to as GLON) on real-world data including color images and videos. For comparison, we select five state-of-the-art LRTC methods: HaLRTC [33], TSVD [32], KBR [51], TMac-TT [54], and TT-TV [58]. We scale the test data to the interval $[0, 255]$. In this work, all numerical experiments are performed on Windows 10 64-bit and MATLAB platform running on a desktop equipped with an Intel(R) Core(TM) i7-8700 CPU with 3.70 GHz, 8 GB RAM, and a GTX1080 GPU.

The quality of recovered results is measured by the peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM). By calculating the average PSNR and SSIM values of all frames, the PSNR and SSIM values of color video can be obtained. Higher PSNR and SSIM values indicate better image quality.

Implementation details. For outer iteration, the convergence criterion of the proposed algorithm is based on the relative error of \mathcal{M} between two successive iterations as follows:

$$\frac{\|\mathcal{M}^{t+1} - \mathcal{M}^t\|_F}{\|\mathcal{M}^t\|_F} \leq 2 \cdot 10^{-3}. \quad (31)$$



Fig. 2. The original color images.

Also, we set the inner iteration number k and the outer iteration number t to 15 and 100, respectively.

The proposed method contains the following parameters: α_i in (6) controlling the global low-rankness of the desired tensor, λ_1 controlling the weight of the local prior, λ_2 controlling the weight of the nonlocal prior, and penalty parameters β , β_1 , and ρ . For α_i , we set

$$\alpha_i = \frac{\delta_i}{\sum_{i=1}^{l-1} \delta_i} \text{ with } \delta_i = \min(\prod_{k=1}^i n_k, \prod_{k=i+1}^l n_k), \quad (32)$$

where $i = 1, \dots, l-1$. We empirically set $\lambda_1 = p \times 10^{-2}$, $\beta_1 = p \times 10^2$, $p = 1, \dots, 6$, $\lambda_2 = 100$, $\rho = 0.01$, and $\beta = 20$ for all experiments. For color image, we set initial TT-rank as $\mathbf{r} = (4, 16, 5, 5, 5, 5, 12, 3)$, $\Delta \mathbf{r}_n = (0, 0, 5, 5, 5, 5, 0, 0)$, and $\mathbf{r}^{\max} = (4, 16, 30, 50, 50, 30, 12, 3)$. For color video, we set initial TT-rank as $\mathbf{r} = (5, 5, 5)$, $\Delta \mathbf{r}_n = (5, 5, 5)$, and $\mathbf{r}^{\max} = (50, 50, 50)$. All parameters involved in the competing methods are optimally assigned or selected as suggested in the relevant papers to obtain the highest PSNR values. Detailed settings are listed as follows.

- (HaLRTC [33]) There are weighted parameters α_i ($i = 1, 2, 3$) and the penalty parameter β in HaLRTC. We use the parameters $\alpha_i = 1/3$ ($i = 1, 2, 3$) and empirically choose the parameter $\beta \in \{5 \times 10^{-5}, 10^{-4}, 5 \times 10^{-4}, 10^{-3}, 5 \times 10^{-3}\}$.

- (TSVD TSVD [32]) There is only a penalty parameter β . We empirically choose the penalty parameter $\beta \in \{10^{-6}, 5 \times 10^{-6}, 10^{-5}, 5 \times 10^{-5}\}$ for higher PSNR values.

- (KBR [51]) According to the author's suggestion, we take the magnitude balance parameter $\lambda \in \{10^{-3}, 10^{-2}, 5 \times 10^{-2}, 10^{-1}, 5 \times 10^{-1}, 1, 5, 10\}$ and the initial penalty parameter $\mu \in \{10^{-13}, 10^{-11}, 10^{-9}, 10^{-7}, 10^{-5}, 10^{-3}\}$ in their algorithm for good performance.

- (TMac-TT [54]) We set regularization parameters α_i with referred to (32).

- (TT-TV [58]) We select the regularization parameter λ and the penalty parameter β_3 from the candidate set: $\{0.01, 0.03, 0.05, 0.1, 0.3\}$, and set penalty parameters $\beta_1 = 5 \times 10^3$ and $\beta_2 = 0.1$ provided by authors.

Meanwhile, we set the maximum number of iterations of all compared methods to 1000. For reproducibility, the source code will be available on <https://zhaoxile.github.io/>.

A. Color Image Completion

We evaluate the performance of the proposed method on color image completion, involving random missing and structured missing. The test data are of size $256 \times 256 \times 3$; see Fig. 2. In the low-TT-rank term of (6), we transform the third-order tensor into a ninth-order tensor

TABLE II
PSNR, SSIM, AND RUNNING TIME (IN SECONDS) OF DIFFERENT METHODS FOR COLOR IMAGE COMPLETION IN THE RANDOM MISSING CASE. THE BEST VALUES ARE HIGHLIGHTED IN HOLD

Image	Method	SR=0.05			SR=0.1			SR=0.2		
		PSNR	SSIM	Time	PSNR	SSIM	Time	PSNR	SSIM	Time
barbara	HaLRTC	16.87	0.3191	2.85	19.19	0.4561	2.81	22.26	0.6276	2.42
	TSVD	15.65	0.2180	87.80	18.40	0.3324	89.23	22.02	0.5720	87.13
	KBR	17.20	0.2492	13.30	20.50	0.4097	13.43	24.26	0.6270	13.29
	TMac-TT	20.77	0.4878	7.90	22.43	0.5872	3.15	25.06	0.7194	2.55
	TT-TV	20.52	0.5950	527.29	23.54	0.7221	514.36	25.79	0.8086	494.82
	GLON	24.21	0.7231	37.10	28.37	0.8664	17.48	31.04	0.9134	15.56
baboon	HaLRTC	17.16	0.2311	3.00	18.77	0.3295	2.84	20.45	0.4758	2.59
	TSVD	15.45	0.1661	87.35	17.85	0.2562	85.40	20.17	0.4241	86.34
	KBR	17.29	0.1789	12.30	18.86	0.2879	12.34	20.57	0.4602	12.18
	TMac-TT	18.45	0.2527	8.78	19.93	0.3585	3.78	21.28	0.4882	1.10
	TT-TV	17.07	0.3106	509.85	19.96	0.4713	506.30	21.99	0.6185	523.01
	GLON	20.83	0.4668	21.11	22.37	0.6432	31.89	24.59	0.7756	15.09
palace	HaLRTC	16.26	0.3117	2.83	18.25	0.4614	2.77	20.78	0.6360	2.73
	TSVD	15.18	0.2319	86.79	17.92	0.3871	88.57	21.11	0.6000	88.32
	KBR	16.22	0.2684	11.74	18.71	0.4207	12.02	22.21	0.6265	13.03
	TMac-TT	16.96	0.3627	8.14	18.90	0.5016	4.57	21.18	0.6534	4.57
	TT-TV	18.43	0.5224	510.59	20.90	0.6666	505.66	23.42	0.7961	475.00
	GLON	20.36	0.7129	57.14	23.84	0.8486	47.55	27.54	0.9206	20.77
house	HaLRTC	17.82	0.3603	2.76	20.29	0.4979	2.73	23.91	0.6687	2.38
	TSVD	17.68	0.2433	84.29	20.26	0.3851	84.11	24.49	0.6071	85.17
	KBR	19.80	0.3549	13.06	22.20	0.4485	12.59	26.42	0.6568	12.84
	TMac-TT	22.43	0.5707	5.96	24.93	0.6654	3.80	28.18	0.7703	2.97
	TT-TV	22.42	0.6302	474.67	24.02	0.7112	440.93	26.54	0.7871	231.13
	GLON	26.72	0.8012	199.34	28.45	0.8444	14.65	29.61	0.8876	12.73
lena	HaLRTC	18.00	0.3957	2.76	20.47	0.5347	2.42	23.61	0.7110	2.26
	TSVD	15.94	0.2042	86.35	18.82	0.3409	97.11	22.47	0.5531	88.33
	KBR	18.53	0.3291	11.97	20.99	0.4891	10.70	24.69	0.6264	12.90
	TMac-TT	21.94	0.5512	3.97	23.75	0.6420	3.45	26.45	0.7698	4.43
	TT-TV	21.93	0.6364	503.81	23.89	0.7286	476.09	26.04	0.8097	414.64
	GLON	25.57	0.8040	19.37	28.91	0.8812	17.78	32.31	0.9288	8.87
monarch	HaLRTC	16.27	0.3855	2.95	17.38	0.4880	3.00	19.52	0.6354	2.76
	TSVD	15.20	0.2158	85.89	16.89	0.3143	86.87	19.71	0.5155	88.30
	KBR	16.05	0.2963	11.89	17.53	0.3980	12.88	20.79	0.6096	13.89
	TMac-TT	16.96	0.4575	3.45	19.25	0.6066	10.60	22.49	0.7609	10.71
	TT-TV	18.32	0.6954	503.15	20.51	0.7925	522.85	24.09	0.8861	507.27
	GLON	20.52	0.8124	167.14	25.22	0.9204	39.71	30.07	0.9670	22.35
peppers	HaLRTC	15.76	0.3063	2.91	18.05	0.4398	3.04	21.25	0.6535	2.55
	TSVD	14.04	0.1206	85.23	17.15	0.2554	88.01	20.42	0.4363	89.49
	KBR	16.59	0.2157	12.52	19.19	0.3421	12.66	23.45	0.5610	13.85
	TMac-TT	20.32	0.4835	9.01	22.61	0.6130	8.16	25.16	0.7334	3.52
	TT-TV	20.60	0.5858	514.90	22.43	0.6735	498.04	24.44	0.7689	504.44
	GLON	24.15	0.8027	42.71	27.11	0.8752	18.28	30.08	0.9243	16.42

$\mathcal{K}(\mathcal{M}) \in R^{4 \times 4 \times 4 \times 4 \times 4 \times 4 \times 4 \times 4 \times 4}$ to explore the global correlation by using VDT. Meanwhile, we use the pre-trained FFDNet and CBM3D on color images as denoisers to recover color images.

1) *Random Missing*: We randomly sample incomplete images with sampling rates (SRs) 0.05, 0.1, and 0.2, which are sampled element-wise.

Table II shows the PSNR, SSIM, and running time of different methods for color image completion with random missing. We observe that our method obtains the highest PSNR and SSIM values in all cases. More precisely, the proposed method achieves about 3 dB PSNR and 0.2 SSIM gain than the second-best methods. Although our method needs more time, the recovered results have shown significant improvements compared to competing methods in terms of PSNR and SSIM values.

Fig. 3 shows the visual comparison of different methods with $SR = 0.1$. As we can see, HaLRTC, KBR, TSVD, and TMac-TT only recover the coarse structure of the color images but blur image details, which suggests that the global low-rankness prior is not enough to recovery image details. TT-TV produces better results than TMac-TT, but generates the artifacts introduced by TV. In comparison, GLON produces the most visually pleasant results with satisfactorily recovered



Fig. 3. The recovered color images by different methods with $SR = 0.1$. From top to bottom: *palace*, *house*, *lena*, *monarch*, and *peppers*.

image details and regular textures. This demonstrates the advantages of the plugged CNN and CBM3D denoisers.

2) *Structural Missing*: We consider the structural missing, which is more challenging than the random missing. Specifically, we test two structural missing cases: image inpainting (black text and black stripes) and image demosaicing. We obtain the test data by sampling the same spatial location along all RGB channels. The test data in image demosaicing is sampled by Bayer pattern, i.e., each two-by-two cell contains two green, one red, and one blue (grbg).

Fig. 4 shows the results of color images for structural sampling by different methods. For structural missing, especially when the missing entries exhibit low-rank structures, the global LRTC method TSVD almost fails; see the third and fourth rows of Fig. 4. From the enlarged regions, we observe that the results obtained by HaLRTC, KBR, and TMac-TT lost image details and textures. TT-TV yields comparative results, but still misses some fine textures, due to the ignorance of image nonlocal information. As a comparison, the recovered results by GLON exhibit better visual quality with better preservation of local image details and repetitive textures. In addition, we observe that the proposed method obtains the highest PSNR and SSIM values in different images and structural missing cases.

B. Video Completion

To show the flexibility of GLON for multi-dimensional image, we test fourth-order color videos, including *foreman*, *hall*, and *carphone*.² The size of the test videos is $144 \times 176 \times 3 \times 300$. Considering that color videos are already high-order tensors, there is no need to use VDT for dimension

²<http://trace.eas.asu.edu/yuv/>

augmentation. In addition, the spatial slice is fed in the pre-trained FFDNet and CBM3D on color images as denoisers to recover color videos.

Table III shows the numerical performance of the recovered videos by different methods with different SRs, and the best results are denoted in bold. We observe that the proposed method obtains higher PSNR and SSIM values in all cases. Fig. 5 plots the PSNR and SSIM values against the frame number. Clearly, GLON achieves the best performance in most frames.

Fig. 6 shows the restored results of two frames by different methods with $SR = 0.1$. The proposed method outperforms all competing methods in preserving image details and textures from the visual comparison, which are consistent with the quantitative performance in Table III.

C. Discussions

1) *Contributions of Different Priors*: Here, we discuss the contributions of different priors to the performance, based on color video *carphone* completion results shown in Fig. 7 with $SR = 0.2$. We observe that TMac-TT (GLON without FFDNet and CBM3D) suffers from serious block-artifacts. When only employing CBM3D for reconstruction, the result is unsatisfied. We attribute this to that the error textures will accumulate in the iterations if the sampling rate too low to support a desire block matching outcome. Benefiting from the powerful representation ability, FFDNet performs well. However, we can observe some vertical stripes. When taking two prior terms into consideration, the results are better. Although the SSIM of TMac-TT+CBM3D are lower than those of TMac-TT, the result by TMac-TT+CBM3D is more visually pleasure. TMac-TT+FFDNet preserves the image details well. The result by

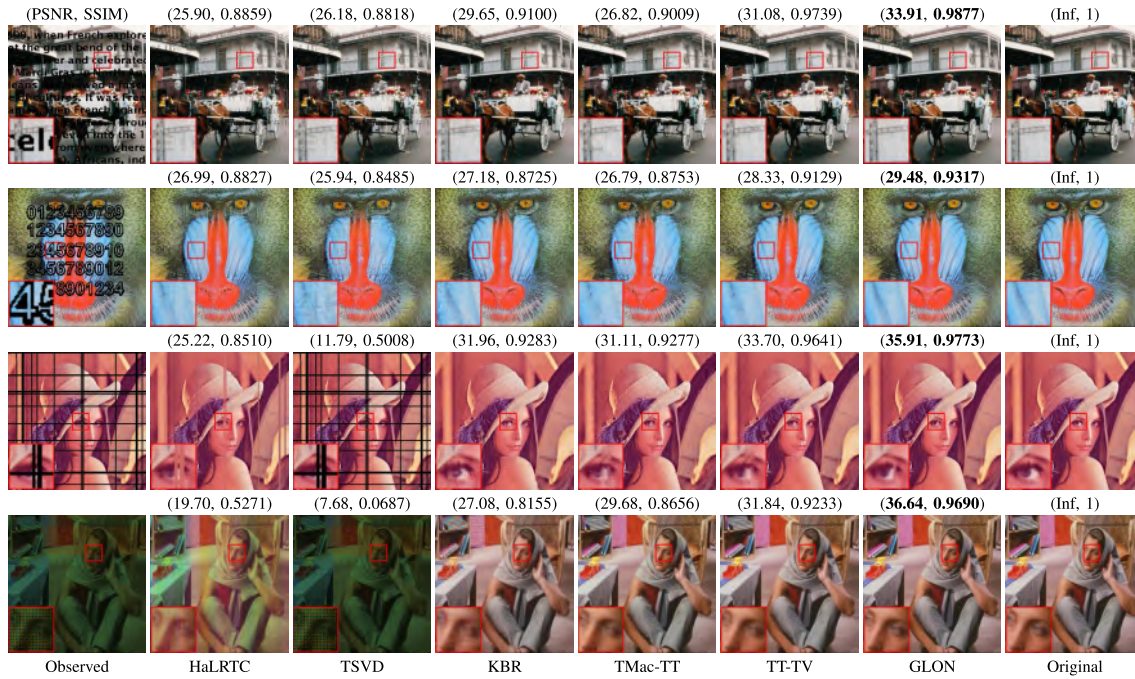


Fig. 4. The recovered color images by different methods for structural missing. From top to bottom: *new*, *baboon*, *lena*, and *barbara*. The numerical indexes at the top of each image indicate the PSNR and SSIM values. The best values are highlighted in hold.

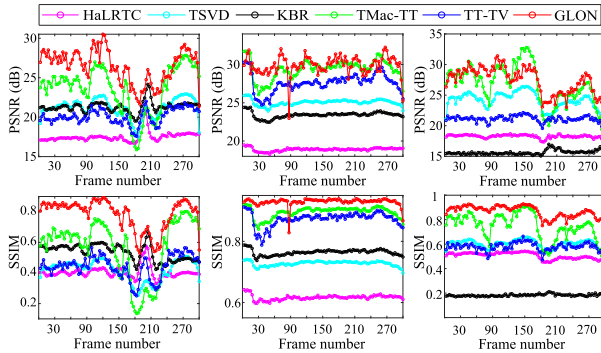


Fig. 5. PSNR and SSIM values with respect to the frame number of the recovered videos *foreman*, *hall*, and *carphone* ($SR = 0.05$) by different methods.

FFDNet+CBM3D is better than that by isolatedly utilizing FFDNet and some image textures are preserved well.

As a whole, the results of the last four methods of Fig. 7 are of higher visual quality. That is, these four methods all perform well for spatial preservation. Fig. 8 shows the pixel values of the video *carphone* restored by different methods in the temporal mode. We can see that our GLON outperforms FFDNet, FFDNet+CBM3D, and TMac-TT+FFDNet, being closer to the original data.

Therefore, from the conjoint analysis on Figs. 7 and 8, we can conclude that these three prior knowledge terms have different effects and are indeed complementary to each other.

2) *Parameters Discussions*: We test the effects of parameters λ_1 and λ_2 on the performance of the proposed method, based on color image completion results shown in Table IV with $SR = 0.1$. In the proposed model (6), λ_1 and λ_2 control the weights of the local prior and the nonlocal prior, respectively. Table IV suggests that the best visual quality is achieved when $\lambda_1 = 0.05$ and $\lambda_2 = 100$. However, too large

TABLE III

PSNR, SSIM, AND RUNNING TIME (IN SECONDS) OF DIFFERENT METHODS ON COLOR VIDEOS WITH THE RANDOM MISSING CASE. THE BEST VALUES ARE HIGHLIGHTED IN HOLD

Image	Method	SR=0.05			SR=0.1			SR=0.2		
		PSNR	SSIM	Time	PSNR	SSIM	Time	PSNR	SSIM	Time
foreman	HaLRTC	17.48	0.4056	105.25	19.86	0.4962	102.36	22.93	0.6501	97.29
	TSVD	21.16	0.4198	2100.60	22.96	0.5191	2602.70	25.34	0.6504	1757.50
	KBR	21.26	0.5354	1570.40	24.04	0.6793	1537.70	27.27	0.8178	1638.70
	TMac-TT	23.64	0.5713	375.12	25.68	0.6921	134.04	27.01	0.7612	70.09
	TT-TV	20.10	0.4463	2076.50	25.49	0.6902	3256.40	26.84	0.7713	2124.50
	GLON	26.46	0.7837	1822.60	30.07	0.8736	2317.10	34.08	0.9341	2340.50
carphone	HaLRTC	17.91	0.5061	126.62	20.98	0.6377	113.20	25.04	0.7906	80.29
	TSVD	24.00	0.6015	1689.80	26.47	0.7397	5444.70	28.72	0.8180	2593.70
	KBR	22.49	0.6690	308.37	25.52	0.7789	1509.00	28.86	0.8770	240.84
	TMac-TT	26.62	0.7439	385.30	29.23	0.8468	142.64	30.59	0.8882	49.42
	TT-TV	20.94	0.5793	1964.20	26.10	0.7892	2291.50	30.18	0.8860	2872.50
	GLON	27.29	0.8651	1843.00	30.53	0.9171	2364.0	34.10	0.9519	2826.00
hall	HaLRTC	18.91	0.6168	105.00	22.05	0.7382	104.55	26.52	0.8676	82.32
	TSVD	25.04	0.7277	2814.20	31.13	0.9210	2349.10	33.63	0.9459	4746.90
	KBR	23.41	0.7640	1538.70	26.98	0.8646	1548.20	30.74	0.9285	1459.70
	TMac-TT	29.42	0.8963	122.04	30.53	0.9214	39.70	31.64	0.9397	27.09
	TT-TV	27.60	0.8728	2022.20	30.57	0.9231	3022.70	31.02	0.9270	4131.20
	GLON	29.97	0.9298	1774.00	34.38	0.9553	2482.60	37.61	0.9656	3739.00

λ_1 and λ_2 lead to over-smooth results; too small λ_1 and λ_2 suffer from severe artifact effects. Specifically, λ_1 controls the local prior constraint, too large λ_1 leads to the blurred restored image, which demonstrate beyond the processing capacity of FFDNet; too small λ_1 leads to some artificial repetitive patterns in the beard part of the *baboon*. λ_2 controls the nonlocal prior constraint, too large λ_2 leads to that textures of the restored image cannot be kept well due to the ability of CBM3D to explore NSS is limited; too small λ_2 leads to some fake textures in the restored image, i.e., some textures in the nose part of the *baboon*. In summary, only when FFDNet and CBM3D cooperate with each other, the abundant details and repetitive textures can be simultaneously preserved well (see the third result in the third row).

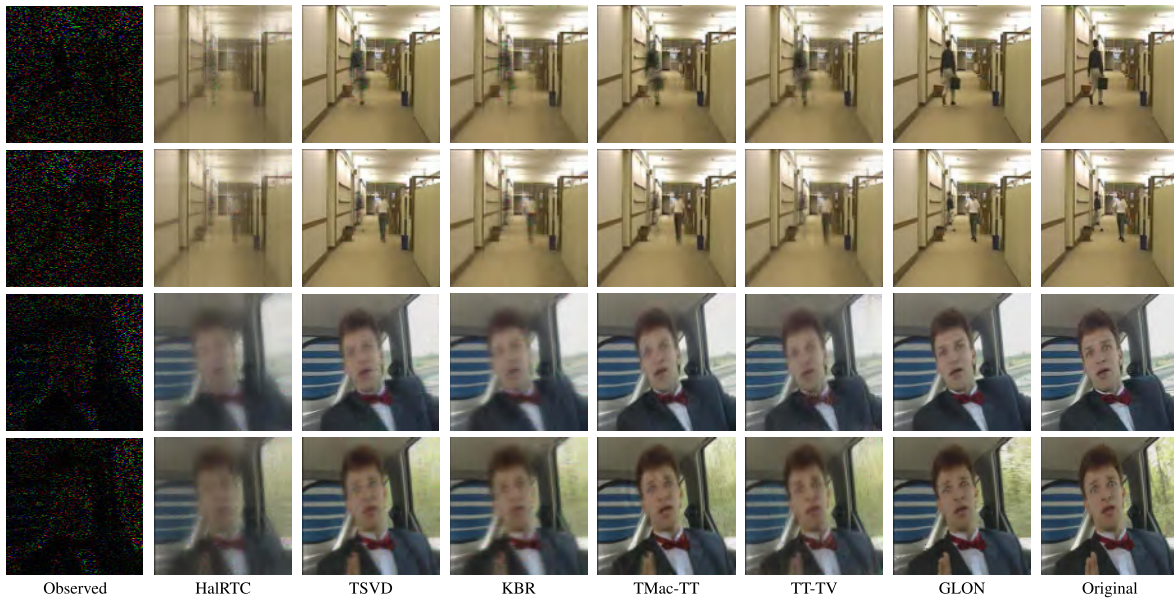


Fig. 6. The 50-th and 230-th frames of the recovered color videos by different methods with SR = 0.1. Top two rows: *hall*, bottom two rows: *carphone*.

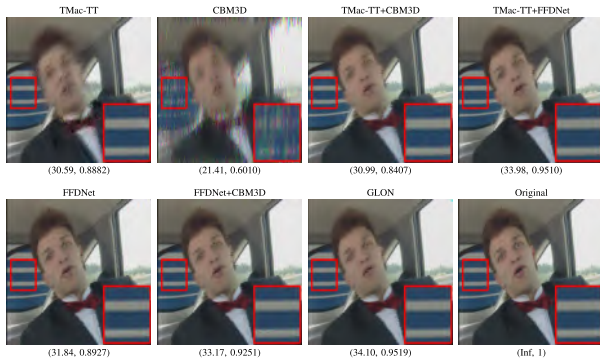


Fig. 7. The 82-th frames of the recovered color video *carphone* by TMac-TT, CBM3D, TMac-TT+CBM3D, TMac-TT+FFDNet, FFDNet, FFDNet+CBM3D, and GLON with SR = 0.2. The numerical indexes at the bottom of each image indicate the PSNR and SSIM values.

Based on the above observations, we recommend adjusting the parameter λ_1 from [0.01, 0.1] with the increment of 0.01 and λ_2 from [100, 500] with the increment of 100 to obtain the highest PSNR restoration result for all datasets. For the penalty parameters β , β_1 , and ρ , we recommend setting $\beta = 20$, $\beta_1 = 0.01$, and $\rho = 0.01$ for all datasets.

Next, we discuss the influence of the inner iteration number k and the outer iteration number t on the performance of the proposed method. Taking the color image *barbara* with SR = 0.1 as an example, Fig. 10 shows the performance of the proposed method with different settings of parameters k and t . In Fig. 10(a), when the outer iteration number t is fixed to 100, we observe that the PSNR curve is stable when the inner iteration number $k \geq 15$. Fig. 10(b) shows that when the inner iteration number k is fixed to 15, the performance is robust for the outer iteration number $t \geq 100$. Considering that the computational time increases with increasing k and t , we fix $k = 15$ and $t = 100$ for all experiments.

3) *Generalization*: The FFDNet is pretrained on the natural images and the CBM3D is originally designed for the natural color images. It is not a surprise that our method could achieve

good performance on the color images and color videos, which are captured from real-world natural scenes. Thus, the generalization ability of the proposed method should be tested and we select the fluorescence images to this end. We implement all the methods on fluorescence images with SR = 0.1. Fig. 9 exhibits the results by our method and the compared methods. It can be found that the recovered results by our method are more visual pleasant. The details and textures are well preserved by our method. It is worth noting that although deep learning-based prior (FFDNet) is trained from a number of natural images, the proposed method can be directly applied to those unnatural images. Therefore, we demonstrate that the proposed method can flexibly complete different kinds of data including natural images and unnatural images.

4) *Convergence Behaviors*: In theory, we have established the convergence guarantee of the proposed algorithm in Theorem 1. Here we show the numerical convergence of the proposed algorithm. Taking the color images *baboon*, *monarch*, and *palace* as examples, Fig. 11 shows the relative error curves of the successive restored image \mathcal{M}^k and \mathcal{M}^{k+1} , i.e., $\|\mathcal{M}^{k+1} - \mathcal{M}^k\|_F / \|\mathcal{M}^k\|_F$. We observe that the relative error keeps decreasing as the iteration number increasing, which demonstrates the numerical stability and the convergence of the proposed algorithm.

5) *Further Discussion*: Here, we compare the proposed GLON method with Tubal-Alt-Min [66] and NGmeet [57]. We test color images *barbara* and *lena* with random sampling rates (SR = 0.8 and SR = 0.2). Fig. 12 (a) shows the recovered results of GLON with Tubal-Alt-Min with SR = 0.8. As a comparison, our method obtains higher PSNR and SSIM values and achieves better visual restoration results. This is because that Tubal-Alt-Min only considers the low-tubal-rankness of data. Fig. 12 (b) shows the recovered results of NGmeet and GLON with SR = 0.2. We can observe that NGmeet could not preserve image details well, due to the ignorance of local information of images and the invalidity of the low-dimensional subspaces assumption for color images. In comparison, the proposed method performs well on the global structures and local image details preservation, which

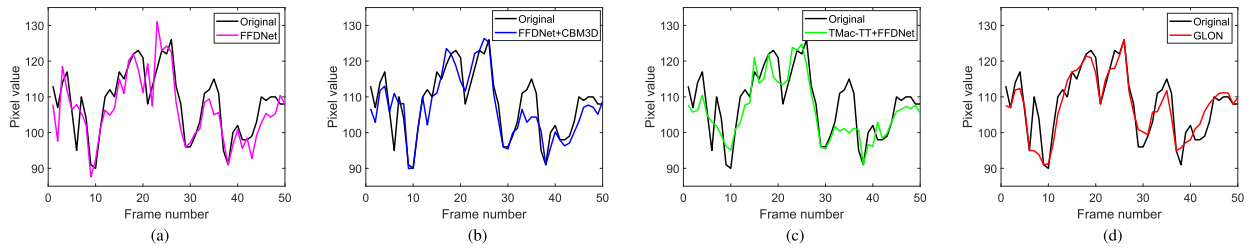


Fig. 8. The pixel values along the temporal dimensional (the same location of each frame) of the recovered video *carphone* by different methods. (a) FFDNet. (b) FFDNet+CBM3D. (c) TMac-TT +FFDNet. (d) GLON.

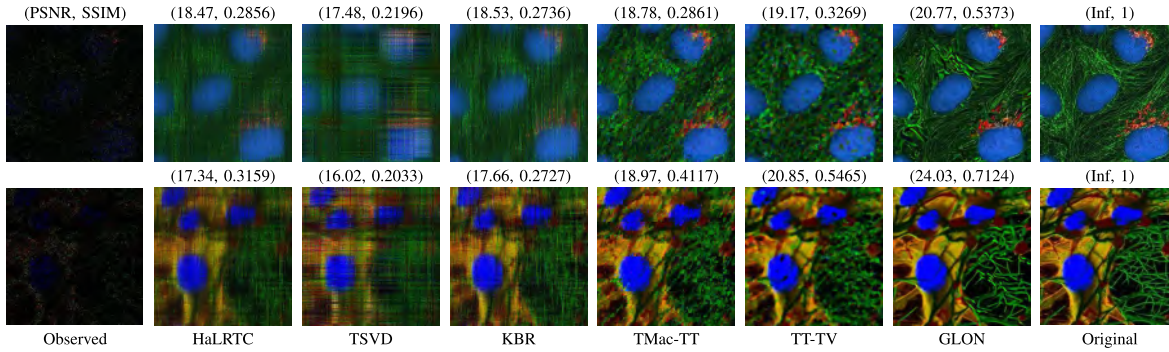


Fig. 9. The recovered fluorescence images by different methods with $SR = 0.1$.

TABLE IV
INFLUENCES OF PARAMETERS λ_1 AND λ_2 ON THE COMPLETION PERFORMANCE. THE NUMERICAL VALUES AT THE TOP OF EACH IMAGE INDICATE THE PSNR AND SSIM VALUES. THE BEST ARE BOXED IN RED RECTANGLE

(PSNR,SSIM)		λ_1				
		0.001	0.01	0.05	0.1	0.5
λ_2	1	(14.87, 0.19)	(16.93, 0.35)	(22.10, 0.62)	(21.34, 0.47)	(17.45, 0.29)
	10	(14.67, 0.19)	(17.58, 0.38)	(22.10, 0.62)	(21.34, 0.47)	(17.45, 0.29)
	100	(14.45, 0.19)	(16.95, 0.37)	(22.37, 0.64)	(21.34, 0.47)	(17.45, 0.29)
	500	(14.38, 0.17)	(17.02, 0.37)	(22.36, 0.63)	(21.34, 0.47)	(17.45, 0.29)
	5000	(12.69, 0.13)	(14.18, 0.19)	(22.05, 0.62)	(21.34, 0.47)	(17.37, 0.29)

suggests that the global, local, and nonlocal priors are originally combined and benefit from each other.

VI. CONCLUSION

We have proposed a flexible low-rank tensor completion framework for multi-dimensional image recovery, in which different types of local and nonlocal priors knowledge can be flexibly and simultaneously utilized. Specifically, we introduce

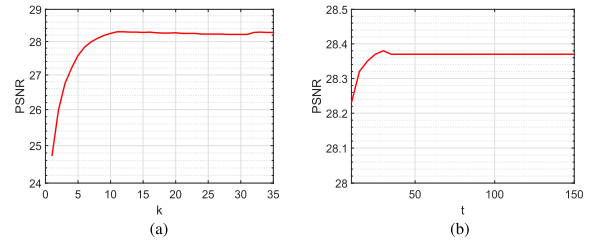


Fig. 10. The PSNR curves with respect to the inner iteration number k and the outer iteration number t , respectively. (a) Inner iteration number. (b) Outer iteration number.

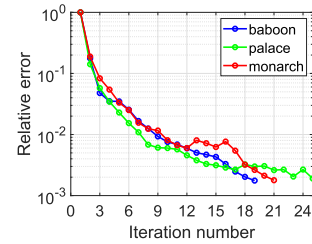


Fig. 11. The curves of relative error versus iterations.

the low-TT-rank regularization and two denoisers CNN and CBM3D to explore the advantages of global-local-nonlocal priors. These three priors are complementary to each other, so that the proposed method can preserve simultaneously the global structures, the local details and features, and the nonlocal textures of the underlying tensor. We develop an efficient PAM algorithm to solve the proposed framework and establish the theoretical convergence guarantee. Extensive numerical experiments including diverse data (color images, color videos, and fluorescence images) and samplings (random and structural samplings) are also reported to validate the uniqueness and complementarity of different types of priors (global TT low-rank prior, nonlocal self-similar prior, local deep prior) for multi-dimensional image recovery. In the



Fig. 12. The recovered color images by different methods with (a) SR = 0.8 and (b) SR = 0.2.

future, we will extend the proposed method to other image processing tasks, such as image denoising [57], [88], image deblurring [89], hyperspectral unmixing or fusion [90], [91], image classification [92], and target detection [93]. For visual data, one remaining challenge for our framework is to recover the large areas missing across all bands without side information (e.g., multi-temporal information). Moreover, since the prior of visual data is considered in our framework, the proposed framework is suitable for visual data and might not be suitable for non-visual data, e.g., traffic data and internet data. We will attempt to address the remaining challenges in our future work.

ACKNOWLEDGMENT

The authors would like to thank the editor and reviewers for giving them many comments and suggestions, which are of great value for improving the quality of this manuscript.

REFERENCES

- [1] L. Sun, "Low rank component induced spatial-spectral kernel method for hyperspectral image classification," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 10, pp. 3829–3842, Oct. 2020.
- [2] Y. Zhang, L. Zhang, B. Du, and S. Wang, "A nonlinear sparse representation-based binary hypothesis model for hyperspectral target detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2513–2522, Jun. 2015.
- [3] Z. Long, Y. Liu, L. Chen, and C. Zhu, "Low rank tensor completion for multiway visual data," *Signal Process.*, vol. 155, pp. 301–316, Feb. 2019.
- [4] X.-L. Zhao, W.-H. Xu, T.-X. Jiang, Y. Wang, and M. K. Ng, "Deep plug-and-play prior for low-rank tensor completion," *Neurocomputing*, vol. 400, pp. 137–149, Dec. 2020.
- [5] S. Gao and Q. Fan, "A mixture of nuclear norm and matrix factorization for tensor completion," *J. Sci. Comput.*, vol. 75, no. 1, pp. 43–64, Apr. 2018.
- [6] X. Zhang, "A nonconvex relaxation approach to low-rank tensor completion," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 6, pp. 1659–1671, Jun. 2019.
- [7] Y.-T. Wang, X.-L. Zhao, T.-X. Jiang, L.-J. Deng, Y. Chang, and T.-Z. Huang, "Rain streaks removal for single image via kernel-guided convolutional neural network," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 8, pp. 3664–3676, Aug. 2021.
- [8] Y. Chang, L. Yan, H. Fang, and C. Luo, "Anisotropic spectral-spatial total variation model for multispectral remote sensing image destriping," *IEEE Trans. Image Process.*, vol. 24, no. 6, pp. 1852–1866, Jun. 2015.
- [9] Z. Xing, M. Zhou, A. Castrodad, G. Sapiro, and L. Carin, "Dictionary learning for noisy and incomplete hyperspectral images," *SIAM J. Imag. Sci.*, vol. 5, no. 1, pp. 33–56, Jan. 2012.
- [10] S. Li, R. Dian, L. Fang, and J. M. Bioucas-Dias, "Fusing hyperspectral and multispectral images via coupled sparse tensor factorization," *IEEE Trans. Image Process.*, vol. 27, no. 8, pp. 4118–4130, Aug. 2018.
- [11] X. Cao, Q. Zhao, D. Meng, Y. Chen, and Z. Xu, "Robust low-rank matrix factorization under general mixture noise distributions," *IEEE Trans. Image Process.*, vol. 25, no. 10, pp. 4677–4690, Oct. 2016.
- [12] Y.-Y. Liu, X.-L. Zhao, Y.-B. Zheng, T.-H. Ma, and H. Zhang, "Hyperspectral image restoration by tensor fibered rank constrained optimization and plug-and-play regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–17, 2022, doi: [10.1109/TGRS.2020.3045169](https://doi.org/10.1109/TGRS.2020.3045169).
- [13] L. Zhuang, M. K. Ng, X. Fu, and J. M. Bioucas-Dias, "Hy-demosaicing: Hyperspectral blind reconstruction from spectral subsampling," *IEEE Int. Geosci. Remote Sens.*, 2021, doi: [10.1109/TGRS.2021.3102136](https://doi.org/10.1109/TGRS.2021.3102136).
- [14] K. Xie, L. Wang, X. Wang, G. Xie, J. Wen, and G. Zhang, "Accurate recovery of internet traffic data: A tensor completion approach," in *Proc. 35th Annu. IEEE Int. Conf. Comput. Commun. (IEEE INFOCOM)*, Apr. 2016, pp. 10–14.
- [15] H. Tan, Y. Wu, G. Feng, W. Wang, and B. Ran, "A new traffic prediction method based on dynamic tensor completion," *Proc. Social Behav. Sci.*, vol. 96, pp. 2431–2442, Nov. 2013.
- [16] J. M. de M. Goulart, A. Y. Kibangou, and G. Favier, "Traffic data imputation via tensor completion based on soft thresholding of Tucker core," *Transp. Res. C, Emerg. Technol.*, vol. 85, pp. 348–362, Dec. 2017.
- [17] G. T. Ely, S. Aeron, N. Hao, and M. Kilmer, "5D seismic data completion and denoising using a novel class of tensor decompositions," *Geophysics*, vol. 80, no. 4, pp. 83–95, 2015.
- [18] J. Feng, X. Li, X. Liu, C. Chen, and H. Chen, "Seismic data denoising based on tensor decomposition with total variation," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 7, pp. 1303–1307, Jul. 2021.
- [19] P. Symeonidis, "Matrix and tensor decomposition in recommender systems," in *Proc. 10th ACM Conf. Recommender Syst.*, Sep. 2016, pp. 429–430.
- [20] W. Wang, V. Aggarwal, and S. Aeron, "Efficient low rank tensor ring completion," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 5697–5705.
- [21] J. Li, Q. Hu, and M. Ai, "Haze and thin cloud removal via sphere model improved dark channel prior," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 3, pp. 472–476, Mar. 2019.
- [22] D. Chu *et al.*, "Long time-series NDVI reconstruction in cloud-prone regions via spatio-temporal tensor completion," *Remote Sens. Environ.*, vol. 264, Oct. 2021, Art. no. 112632, doi: [10.1016/j.rse.2021.112632](https://doi.org/10.1016/j.rse.2021.112632).
- [23] G. T. Herman, A. R. De Pierro, and N. Gai, "On methods for maximum a posteriori image reconstruction with a normal prior," *J. Vis. Commun. Image Represent.*, vol. 3, no. 4, pp. 316–324, Dec. 1992.
- [24] T. G. Kolda, B. W. Bader, and J. P. Kenny, "Higher-order web link analysis using multilinear algebra," in *Proc. ICDM*, 2005, pp. 242–249.
- [25] N. Kreimer and M. D. Sacchi, "A tensor higher-order singular value decomposition for prestack seismic data noise reduction and interpolation," *Geophysics*, vol. 77, no. 3, pp. 113–122, 2012.
- [26] J. M. Bioucas-Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sensing*, vol. 46, no. 8, pp. 2435–2445, Aug. 2008.
- [27] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.
- [28] M. Filipović and A. Jukić, "Tucker factorization with missing data with application to low-rank tensor completion," *Multidimensional Syst. Signal Process.*, vol. 26, no. 3, pp. 677–692, 2015.
- [29] M. E. Kilmer, K. Braman, N. Hao, and R. C. Hoover, "Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 1, pp. 148–172, 2013.

- [30] I. V. Oseledets, "Tensor-train decomposition," *SIAM J. Sci. Comput.*, vol. 33, no. 5, pp. 2295–2317, Sep. 2011.
- [31] Y. Liu, Z. Long, H. Huang, and C. Zhu, "Low CP rank and Tucker rank tensor completion for estimating missing components in image data," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 4, pp. 944–954, Apr. 2020.
- [32] Z. Zhang and S. Aeron, "Exact tensor completion using t-SVD," *IEEE Trans. Signal Process.*, vol. 65, no. 6, pp. 1511–1526, Mar. 2017.
- [33] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 208–220, Jan. 2013.
- [34] Q. Shi, H. Lu, and Y.-M. Cheung, "Tensor rank estimation and completion via CP-based nuclear norm," in *Proc. ACM Conf. Inf. Knowl. Manage.*, Nov. 2017, pp. 949–958.
- [35] G. Zhou, A. Cichocki, Q. Zhao, and S. Xie, "Efficient nonnegative Tucker decompositions: Algorithms and uniqueness," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 4990–5003, Dec. 2015.
- [36] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithm," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, 1992.
- [37] X. T. Li, Y. M. Ye, and X. F. Xu, "Low-rank tensor completion with total variation for visual data inpainting," in *Proc. AAAI*, 2017, pp. 2210–2216.
- [38] T.-Y. Ji, T.-Z. Huang, X.-L. Zhao, T.-H. Ma, and G. Liu, "Tensor completion using total variation and low-rank matrix factorization," *Inf. Sci.*, vol. 326, pp. 243–257, Jan. 2016.
- [39] J.-F. Cai, R. H. Chan, and Z. Shen, "A framelet-based image inpainting algorithm," *Appl. Comput. Harmon. Anal.*, vol. 24, no. 2, pp. 131–149, Mar. 2008.
- [40] Y.-B. Zheng, T.-Z. Huang, T.-Y. Ji, X.-L. Zhao, T.-X. Jiang, and T.-H. Ma, "Low-rank tensor completion via smooth matrix factorization," *Appl. Math. Model.*, vol. 70, pp. 677–695, Jun. 2019.
- [41] K. Zhang, W. Zuo, and L. Zhang, "FFDNet: Toward a fast and flexible solution for CNN-based image denoising," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4608–4622, Sep. 2018.
- [42] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 2, pp. 84–90, Jun. 2017.
- [43] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
- [44] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Proc. CVPR*, vol. 2, Jun. 2005, pp. 60–65.
- [45] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Color image denoising via sparse 3D collaborative filtering with grouping constraint in luminance-chrominance space," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2007, pp. 313–316.
- [46] J. Xu, L. Zhang, D. Zhang, and X. Feng, "Multi-channel weighted nuclear norm minimization for real color image denoising," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 1096–1104.
- [47] J. Xu, L. Zhang, and D. Zhang, "External prior guided internal prior learning for real-world noisy image denoising," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2996–3010, Jun. 2018.
- [48] L. Zhu, C.-W. Fu, M. S. Brown, and P.-A. Heng, "A non-local low-rank framework for ultrasound speckle reduction," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5650–5658.
- [49] M. Maggioni, V. Katkovnik, K. Egiazarian, and A. Foi, "Nonlocal transform-domain filter for volumetric data denoising and reconstruction," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 119–133, Apr. 2013.
- [50] M. Maggioni, G. Boracchi, A. Foi, and K. Egiazarian, "Video denoising, deblocking, and enhancement through separable 4-D nonlocal spatiotemporal transforms," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 3952–3966, Sep. 2012.
- [51] Q. Xie, Q. Zhao, D. Meng, and Z. Xu, "Kronecker-basis-representation based tensor sparsity and its applications to tensor recovery," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 8, pp. 1888–1902, Aug. 2017.
- [52] Y. Chang, L. Yan, X. Zhao, H. Fang, Z. Zhang, and S. Zhong, "Weighted low-rank tensor recovery for hyperspectral image restoration," *IEEE Trans. Cybern.*, vol. 50, no. 11, pp. 4558–4572, Nov. 2020.
- [53] T. Xie, S. Li, L. Fang, and L. Liu, "Tensor completion via non-local low-rank regularization," *IEEE Trans. Cybern.*, vol. 49, no. 6, pp. 2344–2354, Jun. 2019.
- [54] J. A. Bengua, H. N. Phien, H. D. Tuan, and M. N. Do, "Efficient tensor completion for color image and video recovery: Low-rank tensor train," *IEEE Trans. Image Process.*, vol. 26, no. 5, pp. 2466–2479, May 2017.
- [55] T.-X. Jiang, T.-Z. Huang, X.-L. Zhao, T.-Y. Ji, and L.-J. Deng, "Matrix factorization for low-rank tensor completion using framelet prior," *Inf. Sci.*, vol. 436, pp. 403–417, Apr. 2018.
- [56] X.-T. Li, X.-L. Zhao, T.-X. Jiang, Y.-B. Zheng, T.-Y. Ji, and T.-Z. Huang, "Low-rank tensor completion via combined non-local self-similarity and low-rank regularization," *Neurocomputing*, vol. 367, pp. 1–12, Nov. 2019.
- [57] W. He, Q. Yao, C. Li, N. Yokoya, and Q. Zhao, "Non-local meets global: An integrated paradigm for hyperspectral denoising," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 2392–2399.
- [58] M. Ding, T.-Z. Huang, T.-Y. Ji, X.-L. Zhao, and J.-H. Yang, "Low-rank tensor completion using matrix factorization based on tensor rank and total variation," *J. Sci. Comput.*, vol. 81, no. 2, pp. 941–964, Nov. 2019.
- [59] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3551–3582, Jul. 2017.
- [60] Y. Wang, D. Meng, and M. Yuan, "Sparse recovery: From vectors to tensors," *Nat. Sci. Rev.*, vol. 5, no. 5, pp. 756–767, 2018.
- [61] Y. Xu, R. Hao, W. Yin, and Z. Su, "Parallel matrix factorization for low-rank tensor completion," *Inverse Problems Imag.*, vol. 9, no. 2, pp. 601–624, 2015.
- [62] O. Semerci, N. Hao, M. E. Kilmer, and E. L. Miller, "Tensor-based formulation and nuclear norm regularization for multienergy computed tomography," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1678–1693, Apr. 2014.
- [63] T.-X. Jiang, M. K. Ng, X.-L. Zhao, and T.-Z. Huang, "Framelet representation of tensor nuclear norm for third-order tensor completion," *IEEE Trans. Image Process.*, vol. 29, pp. 7233–7244, 2020.
- [64] T.-X. Jiang, X.-L. Zhao, H. Zhang, and M. K. Ng, "Dictionary learning with low-rank coding coefficients for tensor completion," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Aug. 31, 2021, doi: 10.1109/TNNLS.2021.3104837.
- [65] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, and S. Yan, "Tensor robust principal component analysis: Exact recovery of corrupted low-rank tensors via convex optimization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 5249–5257.
- [66] X.-Y. Liu, S. Aeron, V. Aggarwal, and X. Wang, "Low-tubal-rank tensor completion using alternating minimization," *IEEE Trans. Inf. Theory*, vol. 66, no. 3, pp. 1714–1737, Mar. 2020.
- [67] J.-H. Yang, X.-L. Zhao, T.-Y. Ji, T.-H. Ma, and T.-Z. Huang, "Low-rank tensor train for tensor robust principal component analysis," *Appl. Math. Comput.*, vol. 367, no. 15, p. 124783, 2020.
- [68] J.-H. Yang, X.-L. Zhao, T.-H. Ma, Y. Chen, T.-Z. Huang, and M. Ding, "Remote sensing images destriping using unidirectional hybrid total variation and nonconvex low-rank regularization," *J. Comput. Appl. Math.*, vol. 363, pp. 124–144, Jan. 2020.
- [69] K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep CNN denoiser prior for image restoration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 3929–3938.
- [70] W. Dong, L. Zhang, G. Shi, and X. Li, "Nonlocally centralized sparse representation for image restoration," *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1620–1630, Apr. 2013.
- [71] S. Gu, Q. Xie, D. Meng, W. Zuo, X. Feng, and L. Zhang, "Weighted nuclear norm minimization and its applications to low level vision," *Int. J. Comput. Vis.*, vol. 121, no. 2, pp. 183–208, Jan. 2017.
- [72] L. Yuan, Q. Zhao, L. Gui, and J. Cao, "High-order tensor completion via gradient-based optimization under tensor train format," *Signal Process.*, *Image Commun.*, vol. 73, pp. 53–61, Apr. 2019.
- [73] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm minimization with application to image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2862–2869.
- [74] H. C. Burger, C. J. Schuler, and S. Harmeling, "Image denoising: Can plain neural networks compete with BM3D?" in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 2392–2399.
- [75] L. Xu, J. S. Ren, C. Liu, and J. Jia, "Deep convolutional neural network for image deconvolution," in *Proc. NIPS*, 2014, pp. 1790–1798.
- [76] X. Yuan, Y. Liu, J. Suo, and Q. Dai, "Plug-and-play algorithms for large-scale snapshot compressive imaging," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 1447–1457.

- [77] Z. T. Harmany, R. F. Marcia, and R. M. Willett, "This is SPIRAL-TAP: Sparse Poisson intensity reconstruction Algorithms—Theory and practice," *IEEE Trans. Image Process.*, vol. 21, no. 3, pp. 1084–1096, Mar. 2011.
- [78] R. Dian, S. Li, and X. Kang, "Regularizing hyperspectral and multispectral image fusion by CNN denoiser," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 3, pp. 1124–1135, Mar. 2020.
- [79] D. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuities," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 3, pp. 367–383, Mar. 1992.
- [80] M. Nikolova and M. K. Ng, "Analysis of half-quadratic minimization methods for signal and image recovery," *SIAM J. Sci. Comput.*, vol. 27, no. 3, pp. 937–966, Jul. 2006.
- [81] H. Attouch, J. Bolte, and B. F. Svaiter, "Convergence of descent methods for semi-algebraic and tame problems: Proximal algorithms, forward-backward splitting, and regularized gauss–seidel methods," *Math. Program.*, vol. 137, nos. 1–2, pp. 91–129, Feb. 2013.
- [82] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jul. 2011.
- [83] J. Bolte, S. Sabach, and M. Teboulle, "Proximal alternating linearized minimization for nonconvex and nonsmooth problems," *Math. Program.*, vol. 146, nos. 1–2, pp. 459–494, Aug. 2014.
- [84] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, 1992.
- [85] A. Danielyan, V. Katkovich, and K. Egiazarian, "BM3D frames and variational image deblurring," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1715–1728, Apr. 2012.
- [86] Z. Y. Chen, "Efficient rank minimization via solving non-convexPenalties by iterative shrinkage-thresholding algorithm," *Proc. Mach. Learn. Res.*, vol. 65, pp. 1–34, Sep. 2017.
- [87] E. K. Ryu, J. L. Liu, S. C. Wang, X. H. Chen, Z. Y. Wang, and W. T. Yin, "Plug-and-play methods provably converge with properly trained denoisers," in *Proc. PMLR*, 2019, pp. 5546–5557.
- [88] X. Fu, K. Huang, B. Yang, W.-K. Ma, and N. D. Sidiropoulos, "Robust volume minimization-based matrix factorization for remote sensing and document clustering," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6254–6268, Dec. 2016.
- [89] X.-L. Zhao, F. Wang, T.-Z. Huang, M. K. Ng, and R. J. Plemmons, "Deblurring and sparse unmixing for hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 4045–4058, Jul. 2013.
- [90] J. Yao, D. Meng, Q. Zhao, W. Cao, and Z. Xu, "Nonconvex-sparsity and nonlocal-smoothness-based blind hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 28, no. 6, pp. 2991–3006, Jun. 2019.
- [91] R. Dian and S. Li, "Hyperspectral image super-resolution via subspace-based low tensor multi-rank regularization," *IEEE Trans. Image Process.*, vol. 28, no. 10, pp. 5135–5146, Oct. 2019.
- [92] M.-M. Cheng, L. Jing, and M. K. Ng, "Tensor-based low-dimensional representation learning for multi-view clustering," *IEEE Trans. Image Process.*, vol. 28, no. 5, pp. 2399–2414, May 2019.
- [93] C. Gao, D. Meng, Y. Yang, Y. Wang, X. Zhou, and A. G. Hauptmann, "Infrared patch-image model for small target detection in a single image," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4996–5009, Dec. 2013.



Xi-Le Zhao (Member, IEEE) received the M.S. and Ph.D. degrees from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2009 and 2012, respectively. He is currently a Professor with the School of Mathematical Sciences, UESTC. His research interests mainly focus on model-driven and data-driven methods for image processing problems. His homepage is <https://zhaoxile.github.io/>



Jing-Hua Yang (Graduate Student Member, IEEE) received the B.S. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2016. She is currently pursuing the Ph.D. degree with the Faculty of Information Technology, Macau University of Science and Technology. Her current research interests include data mining, image processing, and artificial intelligence.



Tian-Hui Ma received the B.Sc. and Ph.D. degrees in applied mathematics from the University of Electronic Science and Technology of China, Chengdu, China, in 2011 and 2017, respectively. From 2018 to 2020, he was a Post-doctoral Researcher with the School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, China. He is currently a Lecturer with the School of Science, Civil Aviation University of China, Tianjin, China. His current research interests include inverse problems in image processing and sparse optimization.



Tai-Xiang Jiang (Member, IEEE) received the Ph.D. degree in mathematics from the University of Electronic Science and Technology of China (UESTC) in 2019. He was a Co-Training Ph.D. Student with the University of Lisbon supervised by Prof. Jose M. Bioucas-Dias from 2017 to 2018. He was a Research Assistant with the Hong Kong Baptist University supported by Prof. Michael K. Ng in 2019. He is currently an Associate Professor with the School of Economic Information Engineering, Southwestern University of Finance and Economics. His research interests include sparse and low-rank modeling and tensor decomposition for multi-dimensional image processing, especially on the low-level inverse problems for multi-dimensional images. His homepage is <https://taixiangjiang.github.io/>



Michael K. Ng (Senior Member, IEEE) is currently the Director of the Research Division for Mathematical and Statistical Science, and a Chair Professor with the Department of Mathematics, The University of Hong Kong. His research areas are data science, scientific computing, and numerical linear algebra.



Ting-Zhu Huang received the B.S., M.S., and Ph.D. degrees in computational mathematics from the Department of Mathematics, Xi'an Jiaotong University, Xi'an, China. He is currently a Professor with the School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, China. His research interests include scientific computation and applications, numerical algorithms for image processing, numerical linear algebra, preconditioning technologies, and matrix analysis with applications. He is an Editor of *The Scientific World Journal*, *Advances in Numerical Analysis*, the *Journal of Applied Mathematics*, the *Journal of Pure and Applied Mathematics: Advances in Applied Mathematics*, and the *Journal of Electronic Science and Technology*, China.