Hierarchical Prior Regularized Matrix Factorization for Image Completion
Liyu Su

Abstract—The recent low-rank prior based models solve the tensor completion problem efficiently. However, these models fail to exploit the local patterns of tensors, which compromises the performance of tensor completion. In this paper, we propose a novel hierarchical prior regularized matrix factorization model for tensor completion. This model hierarchically incorporates the low-rank prior, total variation prior, and sparse coding prior into a matrix factorization, simultaneously characterizing both the global low-rank property and the local smoothness of tensors. For solving the proposed model, we use the alternating direction method of multipliers to establish our algorithm. Besides, the complexity and convergence are investigated to further validate the algorithm effectiveness. The proposed scheme is then evaluated through various data sets. Experiment results verify the algorithm effectiveness. The proposed scheme is then evaluated through various data sets. Experiment results verify that the proposed method outperforms several state-of-the-art approaches.

Index Terms—Tensor completion, matrix factorization, hierarchical prior, image recovery.

I. INTRODUCTION

Tensors, as multidimensional arrays, attract intensive attention for their natural strength to represent data with high order structure, e.g., a color image is a third-order tensor characterized by two spatial modes, and one color mode respectively. For decades, tensors have emerged in image processing [1]-[4], computer vision [5]-[7], data mining [8]-[10], and machine learning [11]-[14], etc. However, due to the transmission and storage restrictions, incomplete tensors are common concerns in practice. To tackle the concerns, tensor completion methods are performed for estimating the missing entries of incomplete tensors.

Inspired by low-rank matrix completion schemes [15]-[18], low-rank tensor completion methods are popular tools to exploit the global structures of incomplete tensors. Unlike the matrix rank, the definition of tensor rank is still an open problem. Recently, the tensor CANDECOMP/PARAFAC (CP) rank [19]-[22], tensor Tucker rank [23]-[25], tensor tubal rank [26]-[28], and tensor train (TT) rank [29]-[31] are incorporated as the low-rank priors for tensor completion. In detail, the tensor CP rank is defined as the number of rank-one tensors. Unfortunately, minimizing the tensor CP rank is NP-hard yet [34]. The tensor Tucker rank is defined as the ranks of factor matrices, which inevitably destroys the tensor internal structure. To alleviate this structure destruction, the tensor tubal rank uses Fourier transform to effectively characterize the tensor structure, but it is only available for third-order tensors.

Furthermore, the tensor TT rank is proposed via the tensor train decomposition. Despite the improvement, the tensor TT rank lacks representation flexibility, since the border ranks have a fixed pattern.

In addition to the low-rank priors, the smoothness priors are imposed to describe the local patterns of incomplete tensors. Explicitly, the total variation (TV) is a typical prior to enhance piecewise smoothness [35]-[37], minimizing the difference between neighboring tensor entries. Besides, the sparse coding (SC) is another efficient prior to induce the local smoothness [38]-[40]. Compared to the TV prior, the SC prior is often expressed through dictionaries, which are further categorized into learned dictionaries and prespecified dictionaries. The learned dictionaries are mainly obtained via the dictionary learning, showing flexibility and adaptivity to the specific data. Nevertheless, the dictionary learning may be time-consuming when processing high order tensors. The prespecified dictionaries include various transform based dictionaries, e.g., contourlets, wavelets, discrete cosine transforms (DCT), etc. These dictionaries are simple to implement, and require limited computational complexity.

Increasing researchers claim the exclusive employment of low-rank priors or smoothness priors may lead to compromised results for tensor completion. Concerning low-rank priors, each tensor rank describes the tensor structure from different perspectives. Joint utilization of low-rank priors may illustrate improvement [41]. Furthermore, simultaneously using low-rank and smoothness priors to mine both global and local tensor features presents significant progress [42]-[44]. Although the existing approaches show promising development, they independently consider the low-rank and smoothness priors, which yields redundancy and increases complexity.

In this paper, we propose a novel model by simultaneously capturing the global low-rank property and the local smoothness to recover incomplete tensors. Consequently, the corresponding Lagrangian function is formulated, and we solve the function via the program of alternating direction method of multipliers (ADMM). To determine the subproblems in terms of the low-rank prior and the smoothness prior, we optimize one prior term while fix other prior terms at a time. Furthermore, the convergence and complexity are also investigated to further depict the performance of the proposed algorithm. Experiments on different data sets show that, our proposed algorithm is superior to several state-of-the-art algorithms. In summary, the contributions of this paper are listed as follows: 1) We establish a novel hierarchical prior regularized matrix factorization (HPMF) model for tensor recovery, which

Copyright of this work may be transferred without notice, after which, the current version may no longer be accessible.
hierarchically adopts the low-rank prior, TV prior, and SC prior to explore both global structures and local patterns of tensors. This hierarchical prior framework can be generalized to different factorization schemes with other priors.

2) We propose an effective optimization algorithm to solve the HPMF model by the program of ADMM. Additionally, the complexity and convergence analyses are investigated, indicating our method shows acceptable computational efficiency with guaranteed convergence.

3) We utilize various data sets to test the performance of the proposed algorithm. The experimental results demonstrate that, our algorithm outperforms several state-of-the-art algorithms not only quantitatively but also visually.

The remainder of this paper is organized as follows. Section IV. We conduct experiments in section V. Finally, section VI illustrates the conclusion.

II. Preliminaries

This section introduces notations, definitions and reviews related works for tensor completion methods.

A. Notations

In this article, we show scalars by lower case letters, e.g., \( x, y, z \), and depict matrices as upper case letters, e.g., \( X, Y, Z \). The \((i,j)\)-th entry of matrix \( X \) is represented via \( X_{i,j} \). Tensors are denoted as calligraphic letters, e.g., \( \mathcal{X}, \mathcal{Y}, \mathcal{Z} \). Precisely, an \( N \)-th order tensor is denoted as \( \mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N} \), where \( I_n, n = 1, 2, \ldots, N \), is the \( n \)-th mode of \( \mathcal{X} \). We demonstrate \( \mathcal{X}_{i_1,i_2,\ldots,i_N} \) for the \((i_1, i_2, \ldots, i_N)\)-th tensor entry of \( \mathcal{X} \).

B. Definitions

The necessary definitions used in this paper are briefly provided as follows:

Definition 1 (Frobenius norm). The Frobenius norm of an \( N \)-th order tensor \( \mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N} \) is defined as:

\[
\|\mathcal{X}\|_F = \left( \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} X_{i_1,i_2,\ldots,i_N}^2 \right)^{1/2},
\]

where \( \mathcal{X}_{i_1,i_2,\ldots,i_N} \) means a tensor entry with the coordinate of \( i_1, i_2, \ldots, i_N \).

Definition 2 (Tucker decomposition \cite{tucker1966some}). The Tucker decomposition of an \( N \)-th order tensor \( \mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N} \) is defined as:

\[
\mathcal{X} \approx \mathcal{S} \times_1 U_1 \times_2 U_2 \cdots \times_N U_N,
\]

where tensor \( \mathcal{S} \in \mathbb{R}^{r_1 \times r_2 \times \cdots \times r_N} \) is named as the core tensor, and matrix \( U_n \in \mathbb{R}^{I_n \times r_n}, n = 1, 2, \ldots, N \), is called the factor matrix.

Definition 3 (Tensor mode-\( n \)-unfolding). The mode-\( n \)-unfolding of an \( N \)-th order tensor \( \mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N} \) produces a matrix \( X \in \mathbb{R}^{r_n \times \prod_{m=1, m \neq n}^{N} I_m} \). In other words, a tensor entry \( \mathcal{X}_{i_1,i_2,\ldots,i_N} \) maps to a matrix entry \( X_{i_n,j} \), where index follows

\[
j = 1 + \sum_{k=1, k \neq n}^{N} (i_k - 1) \prod_{m=1, m \neq n}^{N} I_m.
\]

For brevity, the mode-\( n \)-unfolding of tensor \( \mathcal{X} \) is represented as \( \mathcal{X}(n) \), and its inverse operation is denoted as \( \text{fold}_n(\cdot) \), namely, \( \mathcal{X} = \text{fold}_n(\mathcal{X}(n)) \).

Definition 4 (Tensor \( n \)-unfolding). The \( n \)-unfolding of an \( N \)-th order tensor \( \mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N} \) leads to a matrix \( X \in \mathbb{R}^{\prod_{m=1}^{N} I_m \times (\prod_{m=1, m \neq n}^{N} I_m)} \). Furthermore, a tensor entry \( \mathcal{X}_{i_1,i_2,\ldots,i_N} \) maps to a matrix entry \( X_{i,j} \), where index follows

\[
i = 1 + \sum_{k=2}^{N} (i_k - 1) \prod_{m=1}^{k-1} I_m,
\]

\[
j = i_n + 1 + \sum_{k=n+2}^{N} (i_k - 1) \prod_{m=n+1}^{k-1} I_m.
\]

The \( n \)-unfolding is a balanced tensor unfolding scheme.

Definition 5 (Kronecker product \cite{vandewalle2000tensor, he2014kronecker}). The Kronecker product is denoted as \( \otimes \). For two arbitrary matrices \( X \in \mathbb{R}^{I_1 \times I_2}, Y \in \mathbb{R}^{J_1 \times J_2} \), we have

\[
Z = X \otimes Y = \mathbb{R}^{(I_1 J_1) \times (I_2 J_2)},
\]

where the matrix entry \( Z_{k_1,k_2} = X_{i_1,i_2} Y_{j_1,j_2}, k_1 = i_1 + (i_2 - 1) J_1, k_2 = j_2 + (j_1 - 1) J_2 \). Particularly, the Kronecker product admits the column-wise form, which is denoted as \( \odot \). For matrices with same column size \( X \in \mathbb{R}^{I \times K}, Y \in \mathbb{R}^{J \times K} \), we have

\[
Z = X \odot Y = \mathbb{R}^{(I J) \times K}.
\]

The column-wise Kronecker product is also named as the Khatri-Rao product.

Definition 6 (Total variation \cite{villani2003optimal}). The Total variation of matrix \( X \in \mathbb{R}^{I_1 \times I_2} \) consists of two formulations, namely, the isotropic TV and anisotropic TV. The isotropic TV of matrix \( X \) is defined as:

\[
\text{TV}_i(X) = \left( \sum_{j=1}^{I_2} \sum_{i=1}^{I_1} |X_{i+1,j} - X_{i,j}|^2 \right)^{1/2}.
\]

The anisotropic TV of \( X \) is defined as:

\[
\text{TV}_a(X) = \left( \sum_{i=1}^{I_1} \sum_{j=1}^{I_2} |X_{i+1,j} - X_{i,j}| + \sum_{j=1}^{I_2} \sum_{i=1}^{I_1} |X_{i,j+1} - X_{i,j}| \right)^{1/2}.
\]

The total variation can be generalized to the tensor form.

Definition 7 (Tensor inner product). The inner product of two tensors \( \mathcal{X}, \mathcal{Y} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N} \) is defined as:

\[
\langle \mathcal{X}, \mathcal{Y} \rangle = \sum_{i_1, i_2, \ldots, i_N} \mathcal{X}_{i_1,i_2,\ldots,i_N} \mathcal{Y}_{i_1,i_2,\ldots,i_N}.
\]

The two tensors are with the same size to decide their inner product.
C. Related Works

This subsection shows the related works on the low-rank prior and the smoothness prior of tensor completion. The low-rank tensor completion for incomplete tensor $X \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is formulated as an optimization problem:

$$\min_{X} \text{rank}(X) \quad \text{s.t.} \quad P_{I}(X) = P_{I}(T), \quad (10)$$

where $\Omega$ suggests the observation set, which tells the coordinates of given tensor entries. $P$ is the operator, mapping the tensor entries from original tensor $T$ to incomplete tensor $X$.

Since the problem in equation (10) is NP-hard, Zhao et al. [22] applied Bayesian inference to decide an appropriate tensor CP rank, but the estimated tensor CP rank may be inaccurate. Furthermore, Liu et al. [25] used tensor mode-$n$ unfolding, transforming the low-rank tensor completion to the low-rank matrix completion:

$$\min_{X(n)} \sum_{n=1}^{N} \alpha_n \|X(n)\|_S \quad \text{s.t.} \quad P_{I}(X) = P_{I}(T), \quad (11)$$

where $\|X(n)\|_S$ decides the matrix nuclear norm of mode-$n$ unfolding of tensor $X$, and weight $\alpha_n$, $n = 1, 2, \ldots, N$, follows $\sum_{n=1}^{N} \alpha_n = 1$. Considering the computational complexity of performing the matrix nuclear norm, Xu et al. [24] employed the parallel matrix factorization to reformulate the problem in equation (11) as:

$$\min_{X(n), U, V_n} \sum_{n=1}^{N} \alpha_n \|X(n) - U_n V_n\|_F^2 \quad \text{s.t.} \quad P_{I}(X) = P_{I}(T), \quad (12)$$

where $U_n \in \mathbb{R}^{I_n \times r_n}$, $V_n \in \mathbb{R}^{r_n \times (\prod_{i=1, i \neq n}^{N} I_i \times I_m)}$, are factor matrices to $X(n)$, and $r_n$ is a prespecified low-rank prior. Both the methods efficiently conduct the tensor completion of incomplete tensors, but the tensor unfolding destroys the global structure. Forremedying the structure damage, Zhang et al. [27] employed the matrix singular value decomposition in Fourier domain to gain the low-rank prior. Although global structure is well preserved, it is only applicable to third-order tensors. Additionally, Bengua et al. [31] utilized $n$-unfolding to replace mode-$n$ unfolding, which better preserves the global structure. However, the $n$-unfolding produces unbalanced border matrices, since the border ranks are fixed.

The tensor low-rank priors focus on the global structure, while the smoothness priors are crucial for the local pattern of tensors. Chen et al. [37] simultaneously considered the low rank prior and the smoothness prior. Based on the Tucker decomposition, they proposed the model as:

$$\min_{X, S, L_n} \sum_{n=1}^{N} \alpha_n \|U_n\|_S + \beta \text{tr}(LL_n)U_n + \gamma \|S\|_F^2 \quad \text{s.t.} \quad X = S \times U_1 \times \cdots \times U_N, \quad P_{I}(X) = P_{I}(T), \quad (13)$$

where $\beta$, $\gamma$ are regularization parameters, $\text{tr}(\cdot)$ is the trace operation, and matrix $L \in \mathbb{R}^{(I_n - 1) \times (I_n - 1)}$ is Laplacian matrix. The model in equation (13) can be interpreted via two parts, i.e., the nuclear norm term for the low-rank prior, the Laplacian term and Frobenius norm term for the smoothness prior. In spite of this comprehensive using of both the low-rank prior and the smoothness prior, the computational complexity may reduce the performance.

For further improvement, Ji et al. [45] fused the TV prior to the problem in equation (12) as follows:

$$\min_{X(n), U_n, V_n} \sum_{n=1}^{N} \alpha_n \|X(n) - U_n V_n\|_F^2 + \mu \text{TV}_I(V_3),$$

s.t. $P_{I}(X) = P_{I}(T)$, where $\text{TV}_I(\cdot)$ indicates the isotropic TV calculation, $\mu$ is the compromise parameter. Concretly, the matrix $U_n$ represents the library, and $V_n$ is treated as the unmixing of $X(n)$. Since they claimed the mode-3 of the tensor might contain the complete information and enjoy the useful structure, the TV prior is only applied to $V_3$. Besides, Jiang et al. [44] designed the framelet transform based dictionary to provide SC prior for $V_3$, which shows obvious progress in smoothness performance. Nevertheless, the local pattern may not be fully exploited due to the partial utilization of the mode-3 unfolding.

III. Model

To sufficiently exploit both the global low-rank prior and the local smoothness prior, we first reformulate the tensor mode-$n$ unfolding, $n = 1, 2, \ldots, N$, as follows:

$$X(n) \approx U_n(U_n \otimes \cdots \otimes U_{n+1} \otimes U_{n-1} \otimes \cdots \otimes U_1)^T, \quad (15)$$

where $U_n \in \mathbb{R}^{I_n \times r_n}$, indicates the matrix with a unified low-rank prior $r_n$ of nth tensor mode, and $\otimes$ implies the Khatri-Rao product. In addition, we denote $V_n = (U_n \otimes \cdots \otimes U_{n+1} \otimes U_{n-1} \otimes \cdots \otimes U_1)^T$, $V_n \in \mathbb{R}^{r_n \times (\prod_{i=1, i \neq n}^{N} I_i \times I_m)}$ for notational brevity.

We consider each tensor mode-$n$ unfolding can be hierarchically characterized via the low-rank prior, TV prior, and SC prior. Accordingly, the following optimization model is established:

$$\min_{X(n), U_n, V_n} \sum_{n=1}^{N} \alpha_n \|X(n) - U_n V_n\|_F^2 + \lambda_1 \|L_n U_n\|_0 + \lambda_2 \|C_n V_n\|_0 + \rho_1 \|B_n U_n\|_0 + \rho_2 \|D_n V_n\|_0, \quad \text{s.t.} \quad P_{I}(X) = P_{I}(T), \quad (16)$$

where $L_n \in \mathbb{R}^{(I_n - 1) \times I_n}$ and $C_n \in \mathbb{R}^{(r_n - 1) \times r_n}$ represent the TV regularization matrices, e.g., entries $L_{k,i} = 1$, $L_{k,i+1} = -1$, $i = 1, 2, \ldots, I_n - 1$, for matrix $U_n$, while $C_n$ is for matrix $V_n$. Matrices $B_n \in \mathbb{R}^{I_n \times I_n}$ and $D_n \in \mathbb{R}^{r_n \times r_n}$ denote the DCT dictionaries for $U_n$, $V_n$ respectively. The weight $\alpha_n$ observes $\sum_{n=1}^{N} \alpha_n = 1$. Compromising parameters are shown as $\lambda_1$, $\lambda_2$, $\rho_1$, $\rho_2$, and $\| \cdot \|_0$ suggests the matrix $l_0$ norm. For example, Fig. 1 explains the proposed hierarchical prior regularized matrix factorization model for a color image.
Unfortunately, the optimization problem in equation (16) is NP-hard owing to the matrix $l_0$ norm. Hence, we use the convex surrogate to further reformulate the problem into:

$$
\begin{align*}
& \min \sum_{n=1}^{N} \frac{\alpha_n}{2} \left\| \mathcal{X}(n) - U_n V_n \right\|_F^2 \\
& \quad + \lambda_n \left\| L_n U_n \right\|_1 + \lambda_n \left\| C_n V_n \right\|_1 \\
& \quad + \rho_n \left\| B_n U_n \right\|_1 + \rho_n \left\| D_n V_n \right\|_1,
\end{align*}
$$

where $\left\| \cdot \right\|_1$ determines the matrix $l_1$ norm. Introducing auxiliary variables $G_n$, $H_n$, $R_n$, $M_n$, $n = 1, 2, \ldots, N$, the problem in equation (17) is equivalent to:

$$
\begin{align*}
& \min \sum_{n=1}^{N} \frac{\alpha_n}{2} \left\| \mathcal{X}(n) - U_n V_n \right\|_F^2 \\
& \quad + \lambda_n \left\| G_n \right\|_1 + \lambda_n \left\| H_n \right\|_1 \\
& \quad + \rho_n \left\| R_n \right\|_1 + \rho_n \left\| M_n \right\|_1,
\end{align*}
$$

where $\left\| \cdot \right\|_1$ determines the matrix $l_1$ norm. Introducing auxiliary variables $G_n$, $H_n$, $R_n$, $M_n$, $n = 1, 2, \ldots, N$, the problem in equation (17) is equivalent to:

The model in equation (18) can be parallelly solved for each mode. Therefore, we establish the corresponding augmented Lagrangian function $f_L$ to the $n$th mode of the tensor as:

$$
\begin{align*}
& f_L(\mathcal{X}(n), U_n, V_n, G_n, H_n, R_n, M_n, \Lambda_n, \Pi_n, \Phi_n, \Gamma_n) \\
& = \frac{\alpha_n}{2} \left\| \mathcal{X}(n) - U_n V_n \right\|_F^2 \\
& \quad + \lambda_n \left\| G_n \right\|_1 + \lambda_n \left\| H_n \right\|_1 > + \frac{\beta_n}{2} \left\| L_n U_n - G_n \right\|_F^2 \\
& \quad + \lambda_n \left\| H_n \right\|_1 + \lambda_n \left\| C_n V_n - H_n \right\|_F^2 \\
& \quad + \rho_n \left\| R_n \right\|_1 + \lambda_n \left\| B_n U_n - R_n \right\|_F^2 \\
& \quad + \rho_n \left\| M_n \right\|_1 + \lambda_n \left\| D_n V_n - M_n \right\|_F^2
\end{align*}

s.t. $P_{\Omega}(\mathcal{X}) = P_{\Omega}(T),$}

where $\Lambda_n$, $\Pi_n$, $\Phi_n$, $\Gamma_n$ are dual variables to the equality constraints, and $\beta_n$, $\lambda_n$, $\omega_n$, $\delta_n$ are augmented Lagrangian parameters for the $n$th mode of the tensor. The formulation in equation (19) are subsequently rewrite as:

$$
\begin{align*}
& f_L(\mathcal{X}(n), U_n, V_n, G_n, H_n, R_n, M_n, \Lambda_n, \Pi_n, \Phi_n, \Gamma_n) \\
& = \frac{\alpha_n}{2} \left\| \mathcal{X}(n) - U_n V_n \right\|_F^2 \\
& \quad + \lambda_n \left\| G_n \right\|_1 + \frac{\beta_n}{2} \left\| L_n U_n - G_n \right\|_F^2 + \Lambda_n \frac{\beta_n}{2} \left\| \right\|_F^2 \\
& \quad + \lambda_n \left\| H_n \right\|_1 + \frac{\beta_n}{2} \left\| C_n V_n - H_n \right\|_F^2 + \Pi_n \frac{\beta_n}{2} \left\| \right\|_F^2 \\
& \quad + \rho_n \left\| R_n \right\|_1 + \frac{\omega_n}{2} \left\| B_n U_n - R_n \right\|_F^2 + \Phi_n \frac{\omega_n}{2} \left\| \right\|_F^2 \\
& \quad + \rho_n \left\| M_n \right\|_1 + \frac{\omega_n}{2} \left\| D_n V_n - M_n \right\|_F^2 + \Gamma_n \frac{\omega_n}{2} \left\| \right\|_F^2 + \Delta,
\end{align*}
$$

where $\Delta$ is a constant.

IV. ALGORITHM

This section demonstrates the algorithm for the proposed model. The convergence and complexity of the algorithm are also investigated.

A. Solution for Algorithm

We address the problem in equation (20) under the framework of ADMM. For the $k$th iteration, we have the following main subproblems:

1) $U_n$ subproblem: We update $U_n$ at $k$th iteration as

$$
\begin{align*}
& U_n^{(k+1)} = \arg \min_{U_n} \alpha_n \left\| \mathcal{X}(n) - U_n V_n \right\|_F^2 \\
& \quad + \frac{\beta_n}{2} \left\| L_n U_n - G_n \right\|_F^2 + \frac{\lambda_n}{2} \left\| H_n \right\|_1 \\
& \quad + \rho_n \left\| R_n \right\|_1 + \frac{\omega_n}{2} \left\| B_n U_n - R_n \right\|_F^2 + \frac{\delta_n}{2} \left\| \Phi_n \right\|_F^2.
\end{align*}
$$

This formulation admits a closed form solution. We then derive that:

$$
\begin{align*}
& \alpha_n U_n V_n^{(k)} V_n^{(k)\dagger} + \beta_n L_n^T L_n + \omega_n B_n^T B_n U_n
\end{align*}

= \alpha_n \mathcal{X}(n) V_n^{(k)\dagger} + \beta_n L_n^T G_n - L_n^T \mathcal{X}(n) \\
& \quad + \omega_n B_n^T R_n^{(k)} - B_n^T \Phi_n^{(k)}.
$$

Fig. 1. Graphical explanation of the proposed model. (a) The hierarchical prior-regularized matrix factorization model for a color image $\mathcal{X}$ of size $256 \times 256 \times 3$. (b) First hierarchy for low-rank prior: Singular values of mode-$n$ unfoldings. (c) Second hierarchy for TV prior: Margin values of mode-$n$ unfoldings. (d) Third hierarchy for SC prior: Absolute coefficient values of mode-$n$ unfoldings by DCT.
where $V_n^{(k),T}$ represents the transpose of $V_n^{(k)}$. By the vectorization operation vec($\cdot$), we have:

\[
(a_n V_n^{(k)}) v_n^{(k),T} \otimes I_n
\]

\[
+ I_\beta \otimes (\beta_n L_n^T L_n + \omega_n B_n^T B_n) \text{ vec}(U_n)
\]

\[
= \text{vec}(a_n X_n^{(k)}) v_n^{(k),T} + \beta_n L_n^T G_n - L_n^T A_n
\]

\[
+ \omega_n B_n^T R_n - B_n^T \Phi_n^{(k)}
\)

(23)

where $I_n \in \mathbb{R}^{n \times I_n}$, $I_\beta \in \mathbb{R}^{n \times \beta}$ represent the identity matrices, $\otimes$ is the Kronecker product, and it provides:

\[
U_n^{(k+1)} = \text{vec}((a_n V_n^{(k)}) v_n^{(k),T} \otimes I_n
\]

\[
+ I_\beta \otimes (\beta_n L_n^T L_n + \omega_n B_n^T B_n) \text{ vec}(U_n)
\]

\[
= \text{vec}(a_n X_n^{(k)}) v_n^{(k),T} + \beta_n L_n^T G_n - L_n^T A_n
\]

\[
+ \omega_n B_n^T R_n - B_n^T \Phi_n^{(k)}
\)

(24)

where vec($\cdot$) is the inverse operation of vec($\cdot$), and $\dagger$ denotes the Moore-Penrose pseudo-inverse.

2) $V_n$ subproblem: At $k$th iteration, we update matrix $V_n$ via

\[
V_n^{(k+1)} = \arg \min_{V_n} \frac{\alpha_n}{2} \|X_n^{(k)} - U_n^{(k+1)} V_n\|_F^2
\]

\[
+ \frac{\beta_n}{2} \|C_n V_n - H_n + \frac{\Pi_n}{\beta_n} \|^2_F
\]

\[
+ \frac{\omega_n}{2} \|D_n V_n - M_n + \frac{\Gamma_n}{\omega_n} \|^2_F
\)

(25)

Owing to the convexity this subproblem, we consequently present that:

\[
V_n^{(k+1)} = (a_n U_n^{(k+1),T} U_n^{(k+1)} + \beta_n C_n^T C_n + \omega_n I_\beta)
\]

\[
+ (a_n U_n^{(k+1),T} X_n^{(k)} + \beta_n C_n^T H_n - C_n^T \Pi_n
\]

\[
+ \omega_n D_n^T M_n - D_n^T H_n)
\)

(26)

where $I_\beta \in \mathbb{R}^{n \times \beta}$ is the identity matrix.

3) $G_n$ subproblem: Matrix $G_n$ is updated at $k$th iteration by solving

\[
G_n^{(k+1)} = \arg \min_{G_n} \lambda_n \|G_n\|_1
\]

\[
+ \frac{\lambda_n}{2} \|L_n U_n^{(k+1)} - G_n + \frac{\Lambda_n}{\lambda_n} \|^2_F
\)

(27)

The matrix $l_1$ norm can be solved via the soft value thresholding operation $\tau_\lambda(\cdot)$ [49]. Subsequently, we have:

\[
G_n^{(k+1)} = \frac{\lambda_n}{\lambda_n + \beta_n \lambda_n} - \frac{\Lambda_n}{\beta_n \lambda_n}
\)

(28)

4) $H_n$ subproblem: Matrix $H_n$ is updated at $k$th iteration by solving

\[
H_n^{(k+1)} = \arg \min_{H_n} \lambda_n \|H_n\|_1
\]

\[
+ \frac{\lambda_n}{2} \|C_n V_n^{(k+1)} - H_n + \frac{\Pi_n}{\lambda_n} \|^2_F
\)

(29)

Similarly, we have:

\[
H_n^{(k+1)} = \frac{\lambda_n}{\lambda_n + \beta_n \lambda_n} - \frac{\Pi_n}{\beta_n \lambda_n}
\)

(30)

5) $R_n$ subproblem: Matrix $R_n$ is updated at $k$th iteration by solving

\[
R_n^{(k+1)} = \arg \min_{R_n} \rho_n \|R_n\|_1
\]

\[
+ \frac{\omega_n}{2} \|B_n U_n^{(k+1)} - R_n + \frac{\Phi_n}{\omega_n} \|^2_F
\)

(31)

Consequently, we have:

\[
R_n^{(k+1)} = \tau_{\alpha_n \rho_n} \left( B_n U_n^{(k+1)} + \frac{\Phi_n}{\omega_n} \right)
\)

(32)

6) $M_n$ subproblem: Matrix $M_n$ is updated at $k$th iteration by solving

\[
M_n^{(k+1)} = \arg \min_{M_n} \rho_n \|M_n\|_1
\]

\[
+ \frac{\omega_n}{2} \|D_n V_n^{(k+1)} - M_n + \frac{\Gamma_n}{\omega_n} \|^2_F
\)

(33)

The solution to this problem is:

\[
M_n^{(k+1)} = \tau_{\alpha_n \rho_n} \left( D_n V_n^{(k+1)} + \frac{\Gamma_n}{\omega_n} \right)
\)

(34)

7) $X_n$ subproblem: We update the tensor mode-$n$ unfolding at $k$th iteration as

\[
X_n^{(k+1)} = \arg \min_{X_n} \frac{\alpha_n}{2} \|X_n - U_n^{(k+1)} V_n^{(k+1)}\|_F^2
\)

(35)

This subproblem has a closed form solution due to the convexity of the formulation:

\[
X_n^{(k+1)} = U_n^{(k+1)} V_n^{(k+1)}
\)

(36)

8) Multiplier subproblem: Dual variables $\Lambda_n$, $\Pi_n$, $\Phi_n$, $\Gamma_n$, are updated following the ADMM procedure, and we have

\[
\Lambda_n^{(k+1)} = \Lambda_n^{(k)} + \beta_n (L_n U_n^{(k+1)} - G_n^{(k+1)})
\]

\[
\Pi_n^{(k+1)} = \Pi_n^{(k)} + \beta_n (C_n V_n^{(k+1)} - H_n^{(k+1)})
\]

\[
\Phi_n^{(k+1)} = \Phi_n^{(k)} + \omega_n (B_n U_n^{(k+1)} - R_n^{(k+1)})
\]

\[
\Gamma_n^{(k+1)} = \Gamma_n^{(k)} + \omega_n (D_n V_n^{(k+1)} - M_n^{(k+1)})
\)

Moreover, we introduce a constant $\mu$ to accelerate the convergence, e.g., $\beta_n^{(k+1)} = \mu \beta_n^{(k)}$

Finally, to employ the constraint provided by the observation set, we update the recovered tensor $\hat{X}$ using all mode-$n$ unfoldings as follows:

\[
\hat{X}^{(k+1)} = P_\Omega(T) + P_\Omega \left( \sum_{n=1}^N \alpha_n \text{fold}_n (X_n^{(k+1)}) \right)
\)

(38)

where $\Omega$ means the observation set of missing components. Furthermore, the convergence criterion of the proposed algorithm is to calculate the absolute relative error of two successive iterations as:

\[
\frac{\|\hat{X}^{(k+1)} - \hat{X}^{(k)}\|_F}{\|\hat{X}^{(k)}\|_F} < \xi
\)

(39)

where $\xi$ is a small positive constant. Our method is summarized in Algorithm [1].
Algorithm 1: Hierarchical Prior Regularized Matrix Factorization Algorithm (HPMF)

\textbf{Input:} Original $N$th-order tensor $T$ with observation set $\Omega$. Parameters $\alpha_n$, $\lambda_n$, $\lambda_{n_2}$, $\rho_n$, $\rho_{n_2}$, $\beta^{(0)}_n$, $\beta^{(0)}_{n_2}$, $\omega_n$, $\omega_{n_2}$, constant $\mu$, maximum iterations $K$, TV regularization matrices $L_n$, $C_n$, DCT matrices $B_n$, $D_n$, $n = 1, 2, \ldots, N$.

1 \textbf{Initialization:} \(P_{\Omega}(X^{(0)}) = P_{\Omega}(T)\), \(P_{\Omega}(X^{(0)}) = 0\), \(U^{(0)}, V^{(0)}, k = 0\);
2 \textbf{while not converged or} \(k < K\) \textbf{do}
3 \hspace{1em} for \(n = 1, 2, \ldots, N\) \textbf{do}
4 \hspace{2em} Update \(U_n^{(k+1)}\) by equation (24);
5 \hspace{2em} Update \(V_n^{(k+1)}\) by equation (26);
6 \hspace{2em} Update \(G_n^{(k+1)}\) by equation (28);
7 \hspace{2em} Update \(H_n^{(k+1)}\) by equation (30);
8 \hspace{2em} Update \(R_n^{(k+1)}\) by equation (32);
9 \hspace{2em} Update \(M_n^{(k+1)}\) by equation (34);
10 \hspace{2em} Update \(\chi^{(k+1)}\) by equation (36);
11 \hspace{2em} Update \(\Lambda_n^{(k+1)}\), \(\Pi_n^{(k+1)}\), \(\Phi_n^{(k+1)}\), \(\Gamma_n^{(k+1)}\) by equation (37);
12 \hspace{1em} end
13 \hspace{1em} Update \(\chi^{(k+1)}\) by equation (38);
14 \hspace{1em} \(\beta_n^{(k+1)} = \mu \beta_n^{(k)}\), \(\beta_{n_2}^{(k+1)} = \mu \beta_{n_2}^{(k)}\);
15 \hspace{1em} \(\omega_n^{(k+1)} = \mu \omega_n^{(k)}\), \(\omega_{n_2}^{(k+1)} = \mu \omega_{n_2}^{(k)}\);
16 \hspace{1em} \(k = k + 1\).
17 \textbf{end}

\textbf{Output:} Recovered tensor \(X^{(k+1)}\).

### B. Complexity and Convergence

We first discuss the overall complexity at each iteration of Algorithm 1. For simplicity, we assume the model lengths $I_n$ and the corresponding mode ranks $r_n$, $n = 1, 2, \ldots, N$, ($N \geq 3$), of an $N$th-order tensor are all equal to $I$ and $r$ respectively, $I \geq r$. Therefore, equation (24) shows the complexity $O(rI^3 + rIN)$. Equation (26) shows the complexity $O(rI^2)$. Equation (28) shows the complexity $O(rI^3 + rIN - 1)$. Equation (30) shows the complexity $O(rI^2 + rIN - 1)$. Equation (32) shows the complexity $O(rI^3 + rIN - 1)$. Finally, equation (38) demonstrates the complexity $O(IN)$. To sum up, the overall complexity of the proposed algorithm is $O(rI^3 + rIN)$ at each iteration. Our algorithm can be parallelly solved along different tensor modes, which further improves the computational efficiency.

For convergence analysis, since the proposed model in equation (20) consists of separable convex subproblems, i.e., the squared $F$-norm and the $l_1$ norm subproblems, the convergence of Algorithm 1 is theoretically guaranteed under the ADMM framework [50].

### V. Experiment

This section conducts experiments on color images. We compare the proposed algorithm with several state-of-the-art methods, including FBCP [22], HalRTC [25], TN [27], SiLRTC-TT [31], STDC [37], and MF-TV [45]. All experiments are implemented by the Intel i5-8500 CPU at 3.0 GHz, 32 GB RAM machine, under the Matlab (R2020a) environment.

We use sampling ratio (SR) to measure the incomplete tensors, and SR is defined as follows:

$$\text{SR} = \frac{\left| \Omega \right|}{\prod_{n=1}^{N} I_n},$$

where $|\Omega|$ denotes the number of uniformly sampled entries at random from an original tensor $T \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$.

For numerical comparison, given the recovered tensor $\hat{T}$, we decide peak signal-to-noise ratio (PSNR) as:

$$\text{PSNR} = 10 \log_{10} \left( \frac{\mathcal{T}_{\text{max}}}{\| \hat{T} - T \|_F} \right) \cdot \prod_{n=1}^{N} I_n,$$

where $\mathcal{T}_{\text{max}}$ denotes the largest value of the original tensor $T$.

Relative standard error (RSE) is computed as:

$$\text{RSE} = \frac{\| \hat{T} - T \|_F}{\| T \|_F}.$$

For visual comparison, structure similarity (SSIM) assesses the similarity between $\hat{T}$ and $T$ through luminance, contrast, and structure [51].

#### A. Installation Settings

For the proposed algorithm, the mode-$n$ unfolding rank $r_n$, $n = 1, 2, \ldots, N$, is estimated as in the work [31], by keeping the singular values which satisfy:

$$\frac{\sigma_i}{\sigma_1} > \delta,$$

where $i = 1, 2, \ldots, r_n$, and the singular value $\sigma_1$ is sorted in descending order. We empirically obtain the threshold $\delta$ from the range of $[0.01, 0.37]$ for the best algorithm performance. Additionally, the convergence criterion $\xi$ is $10^{-5}$, and the maximum iterations $K$ is 500. We decide the acceleration constant $\mu$ from the range of $[1, 1.05]$ to obtain the best algorithm performance. The compared state-of-the-art algorithms utilize their default settings.

#### B. Color Image with Random Sampling

This part uses the typical third-order tensors, i.e., the color images, to test the recovering performance of algorithms. We consider the scenario that the experiments under different SRs. In this scenario, the compromising parameters of the proposed algorithm are set as $\alpha_n = 1/3$, $\lambda_n = 100$, $\lambda_{n_2} = 100$, $\rho_n = 0.1$, $\rho_{n_2} = 100$. In addition, the augmented Lagrangian parameters are determined as $\beta_n^{(0)} = 1$, $\beta_{n_2}^{(0)} = 100$, $\omega_n^{(0)} = 0.001$, $\omega_{n_2}^{(0)} = 1000$, where $n = 1, 2, 3$.

We employ benchmark color images with size of $256 \times 256 \times 3$ to test the recovering performance of algorithms under
In this scenario, the smooth constraint may be redundant, since it takes longer processing time to recover images. In summary, our HPMF method obtains clearer image details, sharper features, and finer textures, than compared schemes.

We provide the average CPU time (in seconds) to process benchmark color images with SRs varying from 5% to 50% in Fig. 3. As illuminated in the figure, HaLRTC takes less CPU time than compared algorithms. Our HPMF costs moderate CPU time, which is less than TNN and MF-TV to recover most color images, except for the Baboon image. Due to the calculation of isotropic TV, MF-TV takes much longer processing time than other algorithms.

C. Color Image with Non-random Sampling

This subsection conducts the experiments under various image masks. We utilize another group of color images with size of $321 \times 481 \times 3$ to exploit the recovering performance of algorithms, and Fig. 3 illustrates the original color images. The settings are the same as in the color image with random sampling.

In Table 2, we describe the recovery performance in terms of PSNR, RSE, SSIM, and average CPU time for color images with different image masks. The best PSNR, RSE, SSIM, and average CPU time are highlighted in boldface. It can be observed that, the proposed HPMF outperforms compared algorithms in terms of PSNR and RSE values for all masks. Furthermore, the SSIM values of HPMF are larger than compared state-of-the-art methods in most cases. All low-rank prior based schemes (FBCP, HaLRTC, TNN, SiLRTC-TT) remove masks efficiently. In detail, TNN and SiLRTC-TT

---

1 https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/bds/.
produce good PSNR, SSIM, and RSE values for all masks. In all experiments, FBCP gives moderate performance in terms of PSNR, SSIM, and RSE. Although HaLRTC shows ordinary PSNR, RSE, and SSIM values, it has the second best performance for removing the Alphabet mask. STDC fails to remove masks in the experiments, and MF-TV only demonstrates good PSNR, RSE, and SSIM values to remove the Circle mask. On average, the PSNR improvement of our HPMF method over the second best algorithm (TNN) is 1.549 dB. The average RSE decline of the HPMF over the second best scheme (TNN) is 0.012, and the average SSIM improvement of the HPMF over the second best algorithm (SiLRTC-TT) is 0.017. Therefore, it can be concluded that, the proposed HPMF for image completion with non-random sampling yields superior performance.

In Fig. 7, we illuminate visual quality on the image recovery with masks. From the figure, except for STDC, all
methods roughly remove different masks, and reconstruct image outlines. Nevertheless, the proposed HPMF recovers clearer image details and textures than compared algorithms. Concretely, low-rank prior based methods (FBCP, HaLRTC) shows limited details, especially when removing masks of Grid, Line, and Circle. This may tell that, masks along the tensor modes severely degrade the recovery accuracy of FBCP and HaLRTC. Other low-rank prior based methods (TNN, SiLRTC-TT) give sharp visual quality for all mask removals. For smoothness prior regularized methods (STDC, MF-TV), we see that, STDC fails to recover images corrupted by masks. MF-TV only presents clear results with masks of Text, Circle, and Alphabet. To sum up, our HPMF demonstrates the best visual quality to remove all masks on average in the experiments.

VI. CONCLUSION

In this article, we propose a novel matrix factorization model to hierarchically employ the low-rank prior, TV prior, and SC prior for tensor completion. The ADMM based algorithm is then utilized to solve the proposed model, and we also investigate the algorithm complexity and convergence. Experiments on various data sets demonstrate the superiority of our algorithm over several state-of-the-art algorithms. For tensor completion, the proposed HPMF framework can be generalized by using different factorization methods and priors. Furthermore, we also expect this framework is applicable to other applications, e.g., the image denoising, the background subtraction, etc.

REFERENCES

[1] Y. Liu, Z. Long, C. Zhu, Image completion using low tensor tree rank and total variation minimization, IEEE Trans. Multimed. (2018), doi:10.1109/TMM.2018.2859026.

![Fig. 5. The average CPU time (in seconds) for recovering color images via algorithms. All compared algorithms are labeled in the legend.](image)

![Fig. 6. The original images with size of 321 × 481 × 3. (a) Farm, (b) Coast, (c) Tiger, (d) People, (e) Port, (f) Lady.](image)
Fig. 7. The observations and the recovered images via algorithms with different masks. Illustrations from left to right: (a) Observations, (b) FBCP, (c) HaLRTC, (d) TNN, (e) SILRTC-TT, (f) STDc, (g) MF-TV, (h) HPMF.

[2] D. H. J. Poot and S. Klein, “Detecting statistically significant differences in quantitative MRI experiments, applied to diffusion tensor imaging,” IEEE Trans. Med. Imag., vol. 34, no. 5, pp. 1164–1176, May 2015.

[3] 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.

[4] B. Du, M. Zhang, L. Zhang, R. Hu, D. Tao, PLTD: patch-based low-rank tensor decomposition for hyperspectral images, IEEE Trans. Multimed. 19 (1) (2017) 67–79.

[5] Y. Liu, L. Chen, and C. Zhu, Improved robust tensor principal component analysis via low-rank core matrix, IEEE J. Sel. Topics Signal Process., vol. 12, no. 6, pp. 1378–1389, Dec. 2018.

[6] V. Vigneron, A. Kodewitz, M.N. da Costa, A.M. Tome, E. Langlang, Non-negative sub-tensor ensemble factorization (NsTEF) algorithm. a new incremental tensor factorization for large data sets, Signal Process. 144 (2018) 77–86.

[7] Y. Pang, X. Li, and Y. Yuan, “Robust tensor analysis with L1-norm,” IEEE Trans. Circuits Syst. Video Technol., vol. 20, no. 2, pp. 172–178, Feb. 2010.

[8] E.E. Papalexakis, C. Faloutsos, N.D. Sidiropoulos, Tensors for data mining and data fusion: Models, applications, and scalable algorithms, ACM Trans. Intell. Syst. Technol. 8 (2) (2017) 16.

[9] L. Saet, I. Jeon, U. Kang, Scalable tensor mining, Big Data Res. 2 (2) (2015) 82-86.

[10] T. G. Kolda and J. Sun, “Scalable tensor decompositions for multi-aspect data mining,” in Proc. 8th IEEE Int. Conf. Data Mining, Dec. 2008, pp. 363–372.

[11] X. Cao, X. Wei, Y. Han, and D. Lin, “Robust face clustering via tensor decomposition,” IEEE Trans. Cybern., vol. 45, no. 11, pp. 2546–2557, Nov. 2015.

[12] J. Chen, B. Jia, and K. Zhang, Trifocal tensor-based adaptive visual trajectory tracking control of mobile robots, IEEE Trans. Cybern., vol. 47, no. 11, pp. 3784–3798, Nov. 2017.

[13] Y. Sun, J. Gao, X. Hong, B. Mishra, and B. Yin, “Heterogeneous tensor decomposition for clustering via manifold optimization,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 38, no. 3, pp. 476–489, Mar. 2016.

[14] S. K. Biswas and P. Milanfar, “Linear support tensor machine with LSK channels: Pedestrian detection in thermal infrared images,” IEEE Trans. Image Process., vol. 26, no. 9, pp. 4229–4242, Sep. 2017.

[15] E.J. Candes, T. Tao, The power of convex relaxation: near-optimal matrix completion, IEEE Trans. Inf. Theory 56 (5) (2010) 2053–2080.

[16] R.H. Keshavan, A. Montanari, S. Oh, Matrix completion from a few entries, IEEE Trans. Inf. Theory 56 (6) (2010) 2980–2998.

[17] Y. Hu, D. Zhang, J. Ye, X. Li, X. He, Fast and accurate matrix completion via truncated nuclear norm regularization, IEEE Trans. Pattern Anal. Mach. Intell. 35 (9) (2013) 2117–2130.

[18] C. Lu, J. Tang, S. Yan, Z. Lin, Generalized nonconvex nonsmooth low-rank minimization, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 4130–4137.

[19] Y. Liu, F. Shang, L. Jiao, J. Cheng, and H. Cheng, Trace norm regularized CANDECOMP/PARAFAC decomposition with missing data, IEEE Trans. Cybern., vol. 45, no. 11, pp. 2437–2448, Nov. 2015.

[20] Y. Yang, Y. Feng, X. Huang, and J. A. K. Suykens, “Rank-1 tensor properties with applications to a class of tensor optimization problems,” SIAM J. Optim., vol. 26, no. 1, pp. 171–196, 2016.

[21] E. Acar, D. Dunlavy, T. Kolda, M. Murup, Scalable tensor factorizations for incomplete data, Chemometr. Intell. Lab. 106 (2011) 41–56, doi:10.1016/j.chemolab.2010.08.004.

[22] Q. Zhao, L. Zhang, A. Cichocki, Bayesian CP factorization of incomplete tensors with automatic rank determination, IEEE Trans. Pattern Anal. Mach. Intell. 37 (9) (2015) 1751–1763.

[23] D. Kressner, M. Steinlechner, and B. Vandereycken, “Low-rank tensor completion with Riemannian optimization,” BIT Numer. Math., vol. 54, no. 2, pp. 447–468, 2014.

[24] 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.

[25] 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.

[26] Z. Zhang, G. Ely, S. Aeran, N. Hao, and M. Kilmer, Novel methods for multilinear data completion and de-noising based on tensorSVD, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2014, pp. 3842–3849.
[27] Z. Zhang, G. Ely, S. Aeron, N. Hao, and M. Kilmer, “Exact Tensor Completion Using t-SVD,” in IEEE Trans. Signal Process. 65 (6) (2017) 1511–1526.

[28] P. Zhou, C. Lu, Z. Lin, C. Zhang, Tensor Factorization for Low-Rank Tensor Completion, IEEE Trans. Image Process. 27 (3) (2018) 1152–1163.

[29] L. Grasedyck, M. Kluge, S. Kramer, Variants of alternating least squares tensor completion in the tensor train format, SIAM J. Sci. Comput. 37 (5) (2015) A2424–A2450.

[30] M. Steinlechner, Riemannian optimization for high-dimensional tensor completion, SIAM J. Sci. Comput. 38 (5) (2016) S461–S484.

[31] J.-A. Bengua, H.N. Phien, H.D. Tuan, M.N. Do, Efficient tensor completion for color image and video recovery: low-rank tensor train, IEEE Trans. Image Process. 26 (5) (2017) 2466–2479.

[32] Z. Long, Y. Liu, L. Chen, C. Zhu, Low rank tensor completion for multiview visual data, Signal Processing, 155 (2019) 301-316.

[33] T. G. Kolda and B. W. Bader, “Tensor decompositions and applications,” SIAM Rev., vol. 51, no. 3, pp. 455–500, 2009.

[34] C. J. Hillar and L.-H. Lim, “Most tensor problems are NP-hard,” J. ACM, vol. 60, no. 6, p. 45, Nov. 2009.

[35] Y. Liu, S. Wu, X. Huang, B. Chen, and C. Zhu, “Hybrid CS-DMRI: Periodic time-variant subsampling and omnidirectional total variation based reconstruction,” IEEE Trans. Med. Imag., vol. 36, no. 10, pp. 2148–2159, Oct. 2017.

[36] 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.

[37] Y.-L. Chen, C.-T. Hsu, H.-Y. M. Liao, Simultaneous tensor decomposition and completion using factor priors, IEEE Transactions on Pattern Analysis and Machine Intelligence 36 3 (2014) 577–591.

[38] Z. Han, C. Leung, L. Huang, and H. C. So, Sparse and truncated nuclear norm based tensor completion. Neural Process. Lett., vol. 45, no. 3, pp. 729–743, 2017.

[39] J. Mairal, M. Elad, and G. Sapiro, “Sparse representation for color image restoration,” IEEE Trans. Image Process., vol. 17, no. 1, pp. 53–69, 2007.

[40] C. Bao, J. F. Cai, and H. Ji, “Fast sparsity-based orthogonal dictionary learning for image restoration,” in Proc. IEEE Int. Conf. Comput. Vision, 2013.

[41] Y. Liu, Z. Long, H. Huang, and C. Zhu, Low CP Rank and Tucker Rank Tensor Completion for Estimating Missing Components in Image Data, in IEEE Trans. Circuits Syst. Video Technol., vol. 30, no. 4, pp. 944–954, Apr. 2020.

[42] Y.-L. Chen, C.-T. Hsu, H.-Y. M. Liao, Simultaneous tensor decomposition and completion using factor priors, IEEE Transactions on Pattern Analysis and Machine Intelligence 36 (3) (2014) 577–591.

[43] B. Madathil and S. N. George, Twist tensor total variation regularized reweighted nuclear norm based tensor completion for video missing area recovery, Inf. Sci., vol. 423, pp. 376–397, Jan. 2018.

[44] 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., vols. 436–437, pp. 403–417, Apr. 2018.

[45] T.-Y. Ji, T.-Z. Huang, X.-L. Zhao, T.-H. Ma, G. Liu, Tensor completion using total variation and low-rank matrix factorization, Inf. Sci. 326 (2016) 243–257.

[46] C. F. Van Loan, The ubiquitous Kronecker product, J. Comput. Appl. Math., 123 (2000), pp. 85–100.

[47] A. Smilde, R. Bro, and P. Geladi, Multi-Way Analysis: Applications in the Chemical Sciences, Wiley, West Sussex, England, 2004.

[48] A. Chambolle, An algorithm for total variation minimization and applications, J. Math. Imag. Vis., vol. 20, no. 1, pp. 89–97, 2004.

[49] J. Cai, E. Candès, Z. Shen, A Singular Value Thresholding Algorithm for Matrix Completion. 2008.

[50] S. Boyd, Distributed optimization and statistical learning via the alternating direction method of multipliers. Found. Trends Mach. Learn., vol. 3, no. 1, pp. 1–122, 2010.

[51] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” IEEE Trans. Image Process., vol. 13, no. 4, pp. 600–612, Apr. 2004.