SPATIALLY REGULARIZED LOW RANK TENSOR OPTIMIZATION FOR VISUAL DATA COMPLETION

Jianchao Gao[⋆] *Hong Shi*[⋆] *Wenwu Wang*[†]

[⋆] School of Computer Science and Technology, Tianjin University, Tianjin, China
[†] Department of Electrical and Electronic Engineering, University of Surrey, Guildford, UK

ABSTRACT

Low-rank tensor completion is a recent method for estimating the values of the missing elements in tensor data by minimizing the tensor rank. However, with only the low rank prior, the local piecewise smooth structure that is important for visual data is not used effectively. To address this problem, we define a new spatial regularization S-norm for tensor completion in order to exploit the local spatial smoothness structure of visual data. More specifically, we introduce the S-norm to the tensor completion model based on a non-convex LogDet function. The S-norm helps to drive the neighborhood elements towards similar values. We utilize the Alternating Direction Method of Multiplier (ADMM) to optimize the proposed model. Experimental results in visual data demonstrate that our method outperforms the state-of-the-art tensor completion models.

Index Terms— tensor completion, S-norm, LogDet function, low rank, visual data processing

1. INTRODUCTION

In practical applications, the data captured often contain missing values due to the failure of storage device, packet loss in transmission, and incomplete measurement in data collection. The goal of completion is to recover the unknown elements based on the known elements. The basic methods for dealing with this problem are local [1], which utilize the adjacent relationship but do not capture the global information. Recently, a global model has been proposed for addressing the completion problem, e.g. using the tensor rank minimization model [2, 3, 4, 5, 6].

For tensor completion, Liu et al. proposed to use the Sum of Nuclear Norm (SNN) to approximate the tensor rank [7]. Since the tensor rank is more sensitive to small singular values, Huang et al. proposed a truncated nuclear norm minimization (TNNM) method for tensor completion [8], in which a higher recovery accuracy is obtained by eliminating smaller singular values. The calculation of the tensor nuclear norm is expensive, therefore a parallel matrix factorization method is proposed in [9]. The tensor train (TT) rank is used to replace Tucker rank in [10], leading to the Tmac-TT algorithm that works better than Tmac. Simultaneous tensor decomposition and completion (STDC) is presented in [11] by utilizing a graph-Laplacian based Tucker decomposition. However, the matrix nuclear norm does not behave well in some cases, e.g. when the elements of matrix are sampled non-uniformly. In order to overcome these limitations, a non-convex rank approximation has been considered by Zhao et al. [12] where the log-determinant (LogDet) function is used to approximate the rank function. The effectiveness of LogDet has been widely verified in several applications, such as subspace clustering [13], recommender system [12], and tensor completion [14].

Due to the presence of the object edges, visual data are often only piecewise smooth. For visual data completion, it is not sufficient to only consider the low rank prior. To address this, total variation (TV) is used for completion problem. For example, TV has been combined with matrix factorization [15], or SNN and Tucker decomposition, leading to LRTC-TV-I, and LRTC-TV-II [16]. Although the tensor completion model based on the TV-norm has achieved great performance, it does not make full use of the neighborhood information of each missing element. Chen et al. proposed a new spectral spatial (SS) regularization method to exploit local smoothness, and applied it successfully to change detection in hyperspectral imagery [17]. Inspired by this work, we generalize SS regularization for the matrix to higher-order tensors and define the S-norm as the spatial regularization for tensor. More specifically, we add the S-norm based spatial regularization to the tensor completion model based on a non-convex LogDet function. To optimize this model, we present an Augmented Lagrangian Multiplier (ALM) method with Alternating Direction Minimizing (ADM) strategy [18]. To summarize, the key contributions of this work are as follows:

- We propose a new spatial regularization S-norm to exploit the property of the local piecewise smoothness of visual data.
- We combine the S-norm with the non-convex LogDet function to build a tensor completion model.
- Experimental results on color image and the medical

The work is partly supported by National Natural Science Foundation of China under Grant 61502335.

datasets show that our model outperforms the state-of-the-art tensor completion methods.

1.1. Notation

In this paper, a scalar is denoted by lowercase letters x, and a matrix is denoted by capital letters, $X \in \mathbb{R}^{m\times n}$ which is composed of column vectors, denoted as $(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n)$, where $\boldsymbol{x}_i \in \mathbb{R}^m, i = 1, \cdots, n$. An Nth-order tensor is denoted by calligraphic letters $\mathcal{X} \in \mathbb{R}^{I_1\times I_2\times\cdots\times I_N}$, and the elements of $\mathcal{X}$ are denoted as $x_{i_1,i_2,\cdots,i_N}$, where $1 \leq i_k \leq I_k$, $1 \leq k \leq N$. The mode-n unfolding of $\mathcal{X}$ is a matrix $\mathcal{X}_{(n)}$ of size $I_n \times I_1 \cdots I_{n-1}I_{n+1} \cdots I_N$, i.e, $\mathcal{X}_{(n)} = Unfold_n(\mathcal{X})$. On the contrary, its reverse operator is denoted as $Fold(\cdot)$, and $\mathcal{X} = Fold_n(\mathcal{X}_n)$. The operation for a tensor is similar to those for a matrix. $\|\mathcal{X}\|_F$ denotes the Frobenius norm of the tensor $\mathcal{X}$.

2. THE PROPOSED METHOD

2.1. Tensor Spatial Regularization

For a matrix data $A \in \mathbb{R}^{M\times N}$, Chen et al. [17] propose the SS regularization that exploits local spatial patterns and its definition is given as follows:

$$\|A\|_{SS} = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{W^2-1} w_{ij}^k \left\|a_{ij}^0 - a_{ij}^k\right\|_2^2, \tag{1}$$

where a_{ij}^k is the kth neighbor of the center element a_{ij}^0, $\|\cdot\|_2$ denotes the ℓ_2-norm of the matrix, and w_{ij}^k is the weight of the kth neighbor of a_{ij}^0. $W^2 - 1$ is the number of neighbors.

For visual data completion, to preserve the local smoothness structure, we generalize the SS regularization from the matrix to higher-order tensors. We propose the following definition for S-norm as a tensor spatial regularization:

$$\|\mathcal{X}\|_S = \sum_{n=1}^{N} \lambda_n \left\|\mathcal{X}_{(n)}\right\|_{SS}, \tag{2}$$

where $\mathcal{X}$ is a tensor of size $I_1 \times I_2 \times \cdots \times I_N$, and λ_n's are tuning parameters. For the third-mode visual data, its mode-1 and mode-2 matricizations help preserve the spatial structure, so we set $\lambda_1, \lambda_2 = 1$, and $\lambda_3 = 0$.

2.2. LogDet and S-norm for Tensor Completion

We consider tensor completion by integrating a non-convex LogDet function with the spatial regularization. Given an Nth-order incomplete tensor $\mathcal{X} \in \mathbb{R}^{I_1\times I_2\times\cdots\times I_N}$, our method starts with both the global information and the local structure of the tensor data. Specifically, the incomplete tensor is reconstructed by the global low rank prior of the data. The S-norm helps to smooth the reconstructed data, by driving the nearby elements to have similar values. The objective function is written as:

$$\min_{\mathcal{X}} \quad \tau\|\mathcal{X}\|_S + \sum_{n=1}^{N}\alpha_n \text{logdet}\left(\left({\mathcal{X}_{(n)}}^{\mathrm{T}}\mathcal{X}_{(n)}\right)^{1/2} + \eta_n I_n\right) \quad s.t. \quad P_\Omega(\mathcal{X}) = P_\Omega(\mathcal{T}), \tag{3}$$

where τ is a regularization parameter. $\mathcal{T}$ is the given incomplete tensor with missing elements and $\mathcal{X}$ has the same size as $\mathcal{T}$. $P_\Omega(\cdot)$ is an operator, which chooses entries that is equal to 1 in Ω. I_n is the n-th identity matrix. Given the constant array $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \cdots, \alpha_N]$, α_n satisfies $\alpha_n \geq 0, \sum_n \alpha_n = 1$. Ω is an index set where the index of observation elements can be either "1" or "0". η_n is a constant, satisfying $\eta_n > 0$.

2.3. Optimization Algorithm

Our objective function (3) consists of a non-convex rank approximation constraint and a spatial prior. Since the unfolded matrices share the same entries, they cannot be solved independently. To address this issue, we introduce additional variables $\{M_n\}_{n=1}^N$ and $\{H_n\}_{n=1}^N$ to separate these interdependent terms. As a result, Eq. (6) can be converted to the following equivalent problem:

$$\min_{\substack{\mathcal{X},M_n,\\H_n}} \tau\sum_{n=1}^{N}\lambda_n\|H_n\|_{SS} + \sum_{n=1}^{N}\alpha_n\text{logdet}\left(\left({M_n}^{\mathrm{T}}M_n\right)^{1/2} + \eta_n I_n\right)$$
$$s.t. \quad P_\Omega(\mathcal{X}) = P_\Omega(\mathcal{T}), M_n = \mathcal{X}_{(n)}, H_n = \mathcal{X}_{(n)}, \tag{4}$$

Each variable can be updated alternatively with other variables fixed. Firstly, we define an effective strategy based on the ALM method with the ADM strategy [18]. Eq. (4) can be solved by the following ALM problem:

$$\begin{aligned}
&\mathcal{L}(\mathcal{X}, M_n, H_n, Y_n, \Lambda_n) = \sum_{n=1}^{N}\lambda_n * \tau\|H_n\|_{SS} \\
&+\sum_{n=1}^{N}\alpha_n*\text{logdet}\left(\left({M_n}^{\mathrm{T}}M_n\right)^{1/2} + \eta_n I_n\right) \\
&+\sum_{n=1}^{N}\alpha_n*\left(\left\langle\mathcal{X}_{(n)} - M_n, Y_n\right\rangle + \frac{\mu_1}{2}\left\|\mathcal{X}_{(n)} - M_n\right\|_{\mathrm{F}}^2\right) \\
&+\sum_{n=1}^{N}\lambda_n*\left(\left\langle\mathcal{X}_{(n)} - H_n, \Lambda_n\right\rangle + \frac{\mu_2}{2}\left\|\mathcal{X}_{(n)} - H_n\right\|_{\mathrm{F}}^2\right),
\end{aligned} \tag{5}$$

where Y_n, Λ_n are Lagrange multipliers, μ_1, μ_2 are positive penalty scalars and $\langle\cdot,\cdot\rangle$ denotes the matrix inner product. Then, we can apply the alternative projection strategy to update $\mathcal{X}, M_n, H_n, Y_n, \Lambda_n$, i.e., updating one of the variables with the others fixed, so we separate a large-scale problem into the following four smaller subproblems.

1. Update M_n^{k+1}: Given $\mathcal{X}^{\mathrm{k}}, H_n^k, Y_n^k, \Lambda_n^k$, we solve the following optimization problem:

$$\begin{aligned}
M_n^{k+1} = \arg\min_{M_n}\sum_{n=1}^{N}\alpha_n*\left(\text{logdet}\left(\left({M_n}^{\mathrm{T}}M_n\right)^{1/2} + \eta_n I_n\right)\right) \\
+\sum_{n=1}^{N}\alpha_n*\left(\frac{\mu_1^k}{2}\left\|\mathcal{X}_{(n)}^k - M_n + \frac{Y_n^k}{\mu_1^k}\right\|_{\mathrm{F}}^2\right).
\end{aligned} \tag{6}$$

In order to solve Eq. (6) effectively, we use the Theorem 1 in [19] to deal with the tensor case. Based on this theorem, define $A = \mathcal{X}^k_{(n)} - \frac{1}{\mu_1^k} Y_n^k$. We can obtain the closed form solution of Eq. (6), written as:

$$M_n^{k+1} = U diag(prox_{f,u_1^k}(\sigma_A) - diag(v^k))V^T, \quad (7)$$

where U, V, σ_A are obtained by the singular value decomposition (SVD) of A. $prox_{f,u_1^k}(\cdot)$ is the Moreau-Yosida operator [19]. $v^k = [1/(\sigma_1^k (M_n^k) + \eta_n), \cdots, 1/(\sigma_1^k (M_n^k) + \eta_n)]$.

2. Update H_n^{k+1}: Given $\mathcal{X}^k, M_n^{k+1}, Y_n^k, \Lambda_n^k$, we solve the following optimization problem:

$$H_n^{k+1} = \arg\min_{H_n} \sum_{n=1}^{N} \lambda_n * \left(\tau \|H_n\|_{SS} + \frac{\mu_2^k}{2} \|H_n - J\|_F^2 \right). \quad (8)$$

Let $J = \mathcal{X}^k_{(n)} - \frac{1}{\mu_2^k} \Lambda_n^k$. Taking the derivative with respect to H_n and setting it to zero, we get

$$\begin{aligned} \boldsymbol{h}_r^{k+1} &= \frac{1}{\sum_i \boldsymbol{c}_i^k} D_r^k \boldsymbol{c}^k \\ H_n^{k+1} &= [\boldsymbol{h}_1^{k+1}, \boldsymbol{h}_2^{k+1}, \cdots, \boldsymbol{h}_{I_n}^{k+1}]^T, \end{aligned} \quad (9)$$

where $i = 1, 2, \cdots, W^2$, $D_r^k \in R^{I_1 \cdots I_{n-1} I_{n+1} \cdots I_N \times W^2}$, and the matrix $D_r^k = [\boldsymbol{h}_r^{k(1)}, \cdots, \boldsymbol{h}_r^{k(W^2-1)}, \boldsymbol{j}_r]$. $\boldsymbol{j}_r$ is column vector of J. The coefficient vector is denoted as $\boldsymbol{c}^k = [w_1, w_2, \cdots, w_{W^2-1}, \frac{1}{2\tau} \mu_2^k]$. So the variable H_n^{k+1} can be easily obtained by solving Eq. (9).

3. Update $\mathcal{X}^{k+1}$: Given $H_n^{k+1}, M_n^{k+1}, Y_n^k, \Lambda_n^k$, the closed form solution to the least squares problem is obtained as:

$$\mathcal{X}^{k+1} = \left\{ \begin{array}{c} \left[\frac{\sum_n \left[Fold_n\left(\mu_1^k M_{(n)} + Y_n^k\right) + Fold_n\left(\mu_2^k H_{(n)} + \Lambda_n^k\right)\right]}{\sum_{n=1}^{N} \left(\alpha_n \mu_1^k + \lambda_n \mu_2^k\right)} \right]_{\overline{\Omega}} . \\ \\ \mathcal{T}_\Omega \end{array} \right. \quad (10)$$

4. Update multipliers: Given $\mathcal{X}^k, M_n^{k+1}, H_n^{k+1}$, we have the following update equations:

$$\begin{aligned} Y_n^{k+1} &= Y_n^k + \mu_1^k \left(\mathcal{X}_{(n)}^{k+1} - M_n^{k+1} \right) \\ \Lambda_n^{k+1} &= \Lambda_n^k + \mu_2^k \left(\mathcal{X}_{(n)}^{k+1} - H_n^{k+1} \right). \end{aligned} \quad (11)$$

The pseudocode is summarized in Algorithm 1.

Algorithm 1: Optimization Procedure of LRTC-S

Input: input an Nth-order corrupted tensor $\mathcal{T}$, and index set Ω.
Output: recovered tensor $\mathcal{X}$

```
Initialize these variables
    H_n^0 = M_n^0 = 0, Y_n^0 = Λ_n^0 = 0, P_Ω(X^0) = P_Ω(T),
    μ_max = 10^10, max_iter = 500, η_n = ε = 10^-6,
    ρ = 1.05, μ_1^0 = μ_2^0 = 10^-4, α, λ;
while not converged do
    Update variables M_n^{k+1} according to Eq. (7);
    Update variables H_n^{k+1} according to Eq. (9);
    Update variables X^{k+1} according to Eq. (10);
    Update Y_n^{k+1}, Λ_n^{k+1} according to Eq. (11);
    Check the convergence conditions;
    Update μ_1^{k+1} = ρμ_1^k; μ_2^{k+1} = ρμ_2^k.
end
```

3. EXPERIMENTS

In this section, we illustrate the experimental results for visual data, namely the color image completion and medical data completion. We compare our method with the existing algorithms, such as FaLRTC and HaLRTC [7], TmacTT [10], STDC [11], LRTC-TV-I and LRTC-TV-II [16] and LogDet [14]. Our model is implemented by the Tensor Toolbox for MATLAB[1]. For convenience, we use three metrics to evaluate the performance of our algorithm, namely the relative squared error (RSE), peak signal-to-noise ratio (PSNR)[16], and structural similarity (SSIM) [20]. In our experiments, we use the relative error of two adjacent iterations as the criterion to show convergence, i.e, $\left\|\mathcal{X}^k - \mathcal{X}^{k-1}\right\|_F / \left\|\mathcal{X}^k\right\|_F \leq \varepsilon$. The other parameters of our algorithm are initialized as follows. We tune τ from set $\{0.001, 0.005, 0.01, 0.1\}$, which gives the best performance. We set $W = 3$, and $w_{1,3,6,8}, w_{2,4,7,9}$ are chosen from values in $\{1, 2, 3, 4\}$.

3.1. Color image completion

First, we choose seven classic color images, namely airplane, lena, baboon, barbara, house, peppers and sailboat. Each image is represented as a third order tensor with size $256 \times 256 \times 3$. In the experiment, we randomly remove many pixels of each image, and the missing rate is ranged from 65% to 95%. In Table 1, we have listed the comparison of seven algorithms on the color image that are recovered at 10%, 20%, and 30% sampling rate (SR), i.e. missing rate at 90%, 80%, and 70% respectively. Each value in the table is the average based on the seven images. It can be seen that LogDet and TmacTT algorithms behave similarly and they are better than FaLRTC and HLRTC at 20% and 30% sampling rate, but their performances are worse than STDC and LRTC-TV-I. However, our proposed model is superior to the other baselines. In particular, when the sampling rate is 10%, the RSE of the proposed model is only 0.1085. In order to intuitively show the recovery performance of our method, Figure 1 presents the recovery results of five algorithms on two color images with 95% missing rate. We can see that LogDet has better performance

[1] http://www.sandia.gov/tgkolda/TensorToolbox/

Table 1. Comparison of different algorithms on color image

SR(%)	Metrics	FaLRTC	HaLRTC	LogDet	TmacTT	STDC	LRTC-TV-I	Proposed
	RSE	0.1894	0.1847	0.1739	0.1467	0.2052	0.1361	**0.1085**
10	PSNR	19.6037	19.8018	20.3600	21.7721	18.6976	22.3243	**24.3231**
	SSIM	0.4573	0.4670	0.4218	0.6072	0.5054	0.7091	**0.7609**
	RSE	0.1320	0.1302	0.1157	0.1112	0.1069	0.1010	**0.0910**
20	PSNR	22.7397	22.8598	24.0388	24.0812	24.7263	24.9529	**25.8891**
	SSIM	0.6281	0.6333	0.6268	0.7135	0.7250	0.7969	**0.8159**
	RSE	0.1010	0.1003	0.0852	0.0970	0.0838	0.0825	**0.0789**
30	PSNR	25.1087	25.0649	26.7999	25.2472	26.8640	26.7656	**27.1592**
	SSIM	0.7412	0.7431	0.7528	0.7684	0.7963	0.8525	**0.8586**

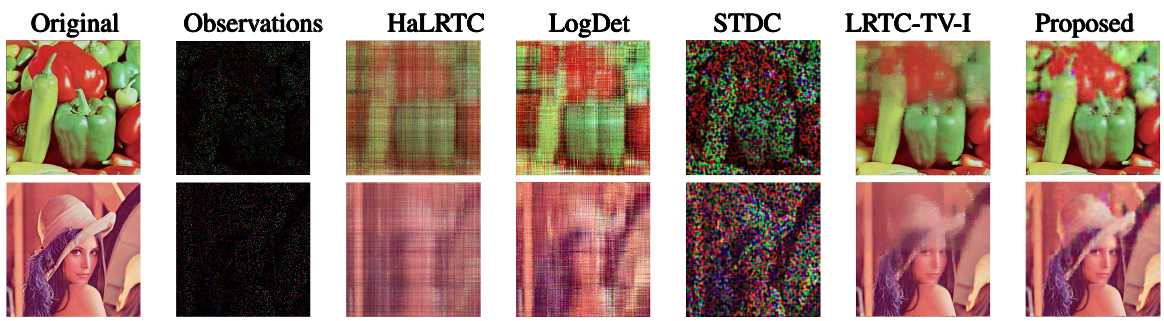
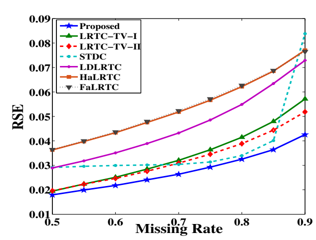


Fig. 1. Recovery results for different algorithms. The first column lists the original image, and the second column lists the image with a missing rate at 95%. The other columns are the recovery results of various methods.

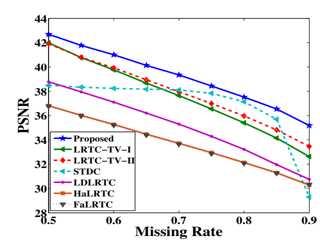
than HaLRTC because the non-convex function approximates tensor rank better than the tensor nuclear norm. However, it does not exploit the local smoothing properties of data, resulting in much worse performance than the LRTC-TV-I. Our method utilizes the S-norm to make full use of the neighborhood information and has better performance than the other baseline methods.

3.2. Medical data completion

In the second experiment, we used medical MRI images, which can be obtained from the OsiriX website[2]. We select an MRI dataset, namely BRAINIX. The dataset contains 22 slices each of 512×512 pixels, so we build a $512 \times 512 \times 22$ tensor. We randomly remove many pixels from the BRAINIX dataset, with the missing rate ranging from 50% to 95%. Figure 2 shows all the simulation results on the BRANIX dataset. Although the BRANIX dataset is bigger than the color image dataset, the performance obtained on this dataset for all the tested algorithms is much better. We observe that STDC has a slight fluctuation, its performance achieved for the missing rate between 70% to 85% is better than LRTC with TV constraint, but it performs worse at the 90% missing rate. As the missing rate increases, our method is more stable than STDC, and it gives better performance.



(a) RSE Metric



(b) PSNR Metric

Fig. 2. Comparison of different methods on MRI dataset.

4. CONCLUSION

We have presented a tensor completion method for visual data by utilizing the local spatial structure of data. We defined a new tensor spatial regularization S-norm, which further smooths low rank visual data, and added it to the tensor completion model based on a non-convex LogDet function. We introduced the ADMM framework to optimize the proposed model. Our empirical study demonstrates that our method gives better recovery performance than several baseline tensor completion methods.

[2]http://www.osirix-viewer.com/datasets/

5. REFERENCES

[1] K. Lundbk, R. Malmros, and E. F. Mogensen, "Image completion using global optimization," in *IEEE Computer Society Conference on Computer Vision Pattern Recognition*, 2006, pp. 442–452.

[2] T.-H. Ma, T.-Z. Huang, and X.-L. Zhao, "Group-based image decomposition using 3-d cartoon and texture priors," *Information Sciences*, vol. 328, pp. 510–527, 2016.

[3] Q. Zhao, D. Meng, X. Kong, and Q. Xie, "A novel sparsity measure for tensor recovery," in *IEEE International Conference on Computer Vision*, 2015, pp. 271–279.

[4] J. F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *Siam Journal on Optimization*, vol. 20, no. 4, pp. 1956–1982, 2008.

[5] E. J. Candès and T. Tao, "The power of convex relaxation: near-optimal matrix completion," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2053–2080, 2010.

[6] Z. Wen, W. Yin, and Y. Zhang, "Solving a low-rank factorization model for matrix completion by a nonlinear successive over-relaxation algorithm," *Mathematical Programming Computation*, vol. 4, no. 4, pp. 333–361, 2012.

[7] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," in *IEEE International Conference on Computer Vision*, 2013, pp. 2114–2121.

[8] L. T. Huang, H. C. So, Y. Chen, and W. Q. Wang, "Truncated nuclear norm minimization for tensor completion," in *IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2014, pp. 417 – 420.

[9] Y. Xu, R. Hao, W. Yin, and Z. Su, "Parallel matrix factorization for low-rank tensor completion," *Inverse Problems Imaging*, vol. 9, no. 2, pp. 601–624, 2017.

[10] J. A. Bengua, H. N. Phiem, H. D. Tuan, and M. N. Do, "Efficient tensor completion for color image and video recovery: Low-rank tensor train.," *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, vol. 26, no. 5, pp. 2466–2479, 2017.

[11] Y. L. Chen, C. T. Hsu, and H. Y. M. Liao, "Simultaneous tensor decomposition and completion using factor priors," *IEEE Transactions on Pattern Analysis Machine Intelligence*, vol. 36, no. 3, pp. 577–591, 2014.

[12] Z. Kang, C. Peng, and Q. Cheng, "Top-n recommender system via matrix completion," in *Thirtieth AAAI Conference on Artificial Intelligence*, 2016, pp. 179–184.

[13] C. Peng, Z. Kang, H. Li, and Q. Cheng, "Subspace clustering using log-determinant rank approximation," in *in Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2015, pp. 925–934.

[14] T.-Y. Ji, T.-Z. Huang, X.-L. Zhao, T.-H. Ma, and L.-J. Deng, "A non-convex tensor rank approximation for tensor completion," *Applied Mathematical Modelling*, vol. 48, pp. 410–422, 2017.

[15] 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," *Information Sciences*, vol. 326, no. C, pp. 243–257, 2016.

[16] X. Li, Y. Ye, and X. Xu, "Low-rank tensor completion with total variation for visual data inpainting," in *Thirty-First AAAI Conference on Artificial Intelligence*, 2017, pp. 2210–2216.

[17] Z. Chen, B. Yang, B. Wang, G.-H. Liu, and W. Xia, "Change detection in hyperspectral imagery based on spectrally-spatially regularized low-rank matrix decomposition," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS17)*, 2017, pp. 157–160.

[18] Z. Lin, R. Liu, and Z. Su, "Linearized alternating direction method with adaptive penalty for low-rank representation," *Advances in Neural Information Processing Systems*, pp. 612–620, 2011.

[19] Z. Kang, C. Peng, and Q. Cheng, "Robust pca via nonconvex rank approximation," in *IEEE International Conference on Data Mining*, 2015, pp. 211–220.

[20] 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, 2004.