Robust PCA Using Matrix Factorization for Background/Foreground Separation

Article *in* IEEE Access · March 2018

DOI: 10.1109/ACCESS.2018.2818322

6 authors, including:

Shuqin Wang
Beijing Jiaotong University
6 PUBLICATIONS **43** CITATIONS

SEE PROFILE

Yongyong Chen
Harbin Institute of Technology Shenzhen Graduate School
30 PUBLICATIONS **273** CITATIONS

SEE PROFILE

Robust PCA using Matrix Factorization for Background/Foreground Separation

Shuqin Wang, Yongli Wang, Yongyong Chen, Peng Pan, Zhipeng Sun, and Guoping He

Abstract—Background/foreground separation has become an inevitable step in numerous image/video processing applications, such as image/video inpainting, anomaly detection, motion segmentation, and augmented reality, etc. Recent low-rank based approaches, such as robust principal component analysis separating a data matrix into a low-rank matrix with a sparse matrix, have achieved encouraging performance. However, these approaches usually need relatively high computation cost, mainly due to calculation of full or partial singular value decomposition of large matrices. On the other hand, the nuclear norm is widely exploited as a convex surrogate of the original rank function, while it is not a tighter envelope of the original rank function. To address these above mentioned issues, this paper proposes a fast background/foreground separation algorithm in which the low-rank constraint is solved by a matrix factorization scheme, thus heavily reducing the computation cost. We further adopt two non-convex low-rank approximations to improve the robustness and flexibility of the traditional nuclear norm. Comparing with the state-of-the-art low-rank reconstruction methods, experimental results on challenging datasets which contain different real datasets show our superior performance in both image clarity and computation efficiency.

Index Terms—Background/foreground separation, matrix factorization, robust PCA.

1 Introduction

As an important step, background/foreground separation has been widely used in various image/video applications, including image/video inpainting, anomaly detection, motion segmentation, behavior recognition, and augmented reality [1], [2], [3], [4], and attracted increasing attention in recent years. For example, the moving objects detection, which is a basic operation in video analysis, aims at subtracting the "foreground" and "background" from the video.

Various approaches for background/foreground separation can be roughly categorized into the following classes [5]: the basic approach, the statistical approach, the fuzzy approach, the neural and neuro-fuzzy approach, and the subspace learning approach [1], [6], [7]. However, the traditional background/foreground separation methods may suffer from some issues. For example, [8] first proposed to use the principal component analysis (PCA) to estimate the background, whereas, PCA is brittle to non-Gaussian noise (such as outliers), which may be prevalent in our real applications. On the other hand, some useful video structure knowledge are neglected by these classical methods, leading to the poor performance. To address this shortcoming, Candès *et al.* [9] developed robust PCA (RPCA), in which the static background is modeled as a low-rank matrix due to the high correlation between frames, and moving objects such as pedestrians, shaking leaves, fountains, are represented as a sparse matrix based on the striking fact that the moving objects just occupy a small proportion in video foreground [10]. RPCA can be formulated as

$$\min_{L,S} rank(L) + \lambda\|S\|_0, \quad s.t. \ D = L + S. \tag{1}$$

However, solving (1) is NP-hard and intractable, because rank function and l_0-norm are non-convex. The most popular choice is to utilize the nuclear norm (sum of all singular values) and l_1-norm instead of the rank function and l_0-norm, respectively, leading to the following formulation:

$$\min_{L,S} \|L\|_* + \lambda\|S\|_1, \quad s.t. \ D = L + S, \tag{2}$$

where $\|L\|_* = \sum_i \sigma_i$ is the nuclear norm, which is the tightest convex relaxation of the rank function; σ_i is the i-th singular value of the matrix L and $\|S\|_1 = \sum_{i,j} |S_{i,j}|$.

Many efforts, such as singular value thresholding [11], exact and inexact augmented lagrange method [12], alternating direction method of multipliers [13], fast alternating linearization method [14], have been developed for (2). However, nuclear norm solution needs computation of full or partial singular value decomposition (SVD). Therefore, with the increase of matrix dimension, time consuming is further extended. On the other hand, the nuclear norm treats each singular value equally, which may lead to overshrinking of the rank component [15]. The main reason is that the nuclear norm is essentially an l_1 norm of the singular values which is a biased estimation [16]. Hence, solving the convex optimization model (2) may get only a suboptimal solution of the original model (1). Specially, when some singular values are very large, the rank estimation will be too large, so that only a part of noise can be removed. In order to make up for the inadequacy of the nuclear norm, several

- *This work was supported by National Natural Science Foundation of China (Grant No. 11626143) and Natural Science Foundation of Shandong Province (Grant No. ZR2017MF054).*
- *Shuqin Wang, Yongli Wang (corresponding author), Peng Pan, Zhipeng Sun are with the College of Mathematics and Systems Science, Shandong University of Science and Technology, Qingdao 266590, China (E-mail: wangyongli@sdkd.net.cn).*
- *Yongyong Chen is with the Department of Computer and Information Science, University of Macau, Macau 999078, China (E-mail: YongyongChen.cn@gmail.com).*
- *Guoping He is with the Shandong Academy of Sciences, Jinan 250014, China (E-mail: hegp@263.net).*

recent studies have investigated some non-convex sparsity-inducing regularizers, such as the weighted nuclear norm [15], capped norm [17], weighted Schatten p-norm [18], log-determinant penalty [16], γ-norm [19], etc. Extensive experiments on background subtraction, face image shadow removal, image inpainting, multispectral/hyperspectral image denoising validate the promising performance of non-convex sparsity-inducing regularizers. Furthermore, Lu *et al.* [20] proposed a general framework of matrix rank minimization and used the iteratively reweighted nuclear norm algorithm to solve the corresponding problem. More recently, some extra structural information priors are considered in many research studies [21], [22], [23]. For example, Sobral *et al.* [21] proposed a shape and confidence map-based RCPA model to enhance the performance of the object foreground detection. Javed *et al.* [23] developed a spatiotemporal sparse subspace clustering based approach for background-foreground modeling in order to address the failure in degradation performance in the presence of dynamic backgrounds. This approach draws inspiration from the observation that these backgrounds in surveillance video may span multiple manifolds.

Another way to low-rank matrix approximation is low-rank matrix factorization [24], [25], [26], *i.e.*, $L = UV^T$, where $U \in R^{m\times r}$, $V \in R^{n\times r}$ and $r \ll \min\{m, n\}$. The low-rank property is based on the fact that $rank(L) = rank(UV^T) \leq \min\{rank(U), rank(V)\}$. The traditional RPCA model (2) is optimized by a batch mode manner, which severely limits its practical applicability. One may be eager to process one sample per time to handle the real-time background-foreground separation issue. Therefore, [27] proposed an online RPCA model, termed OR-PCA, where the original nuclear norm was decomposed into an explicit product of two thin matrices. Following this line, Javed *et. al* [28] modified OR-PCA [27] via stochastic optimization method for background subtraction. And He *et al.* [29] proposed a Grassmannian robust adaptive subspace tracking algorithm, which used a l_1 norm loss instead of the squared loss [30] to handle outliers for robust estimation. Other advanced approaches along this line include [31], [32], [33]. One bottleneck of low-rank matrix factorization based methods is that they usually need to precisely predefine r, which is difficult in practice.

To avoid the defects of nuclear norm and adopt the advantages of matrix factorization which can greatly reduce the cost complexity, in this paper, we propose a fast background/foreground separation model under two special non-convex sparsity-inducing regularizers for RPCA. We highlight the main contributions of this paper as follows:

- We develop a fast background/foreground separation method. By factorizing the low-rank term, the number of iterations and the computation time are effectively reduced. Two non-convex sparsity-inducing regularizers are incorporated in our model, allowing our method to just give an upper bound of the true rank.
- We utilize the augmented Lagrangian multiplier (ALM) method to solve the proposed model and apply our proposed method to the background/foreground separation on surveillance video, hyperspectral image, and medical image. Compared with some state-of-the-art methods, extensive results validate the superior performance.

The remainder of this paper is organized as follows. Section 2 provides a brief review of RPCA including some non-convex models. Section 3 presents our novel matrix factorization-based RPCA model and proposes our algorithm to solve the associated model. We evaluate the performance of our proposed method in section 4 and summarize this paper finally.

2 Related Works

Several enhanced low-rank matrix approximations [15], [16], [18], [25], [26], [34], [35] have been proposed in recent years. For example, Zhou *et al.* [34] developed a non-convex low-rank and sparse matrix decomposition named "Go Decomposition (GoDec)", which is formulated as follows:

$$\min_{L,S} \|D - L - S\|_F^2, \;\; s.t. \;\; rank(L) \leq r, \|S\|_0 \leq k. \tag{3}$$

In order to avoid the time-consuming SVD, GoDec exploited the bilateral random projections, thus obtained a fast background/foreground separation algorithm. One challenge of (3) is that the true rank r and cardinality must be known in advance. In [35], Xu *et al.* proposed to replace l_1-norm with $l_{2,1}$-norm, *i.e.*,

$$\min_{L,S} \;\; \|L\|_* + \lambda\|S\|_{2,1} \;\; s.t. \;\; D = L + S, \tag{4}$$

where $\|S\|_{2,1} = \sum_{j=1}^{n}\sqrt{\sum_{i=1}^{m} S_{ij}^2}$ can detect outliers with column-wise sparsity. Here, the nuclear norm may be not a very good approximation of the original rank function [15], especially when the low-rank matrix has large singular values. Therefore, Gu *et al.* [15] proposed a weighted nuclear norm *i.e.*, $\|L\|_{w,*} = \sum_i w_i * \sigma_i$, where $w = [w_1, w_2, \cdots, w_n]^T$ and each entry is a non-negative number. Then, the weighted nuclear norm was introduced into (2), leading to the following non-convex model:

$$\min_{L,S} \|L\|_{w,*} + \lambda\|S\|_1, \;\; s.t. \;\; D = L + S. \tag{5}$$

3 Models and Algorithms

In this section, we first develop a novel matrix factorization-based RPCA model. Then, we exploit the augmented Lagrangian multipliers method to solve the associated model and derive the closed-form solution of each sub-problem. The keys to the proposed model are to use the matrix factorization scheme, which could mitigate the computation cost of performing SVDs, and two special non-convex functions, which was proven to be a tighter approximation of the rank function.

3.1 Model formulation

As mentioned before, the nuclear norm-based algorithms always suffer from high computation cost of SVDs at each iteration. Hereby, we adopt the matrix factorization scheme to model the low-rank property of the underlying clean data. Therefore, we write the observed data matrix D as

$D = UV^T + S$, where S represents the sparse error matrix, yielding the following optimization problem:

$$\begin{aligned} &\min_{U,V,S} \|S\|_1 \\ &s.t. \;\; UV^T + S = D. \end{aligned} \tag{6}$$

Following the idea in [36], we enforce an equality $U^TU = I$, where I is an identity matrix, to facilitate the uniqueness of solution. Then, the model (6) could be rewritten as the following problem:

$$\begin{aligned} &\min_{U,V,S} \|S\|_1 + \lambda\|V\|_* \\ &s.t. \;\; UV^T + S = D, \;\; U^TU = I. \end{aligned} \tag{7}$$

To improve the flexibility and robustness of the nuclear norm, we introduce two non-convex low-rank approximations and obtain a much smaller-scale minimization problem:

$$\begin{aligned} &\min_{U,V,S} \|S\|_1 + \lambda\|V\|_\gamma \\ &s.t. \;\; UV^T + S = D, \;\; U^TU = I. \end{aligned} \tag{8}$$

Where $\|V\|_\gamma = \sum_{i=1}^{r} \phi\big(\sigma_i(V)\big)$ and $\phi(x)$ is a non-convex sparsity-inducing regularizer as shown in Table 1. When $\phi(x) = |x|$, the model (7) is a special case of model (8). Compared with the convex model (2), the main advantage of our model (8) lies in the fact that the convex model (2) is converted into a small-scale matrix minimization problem (8), resulting in lower computation cost. Another merit is that, as shown in our experimental results, our proposed model (8) just needs to know an upper bound of the true rank by exploiting two non-convex sparsity-inducing regularizers.

TABLE 1
Two particular non-convex surrogate funtions $\phi(x)$.

Function name	Formula
Laplace [37]	$1 - e^{\frac{-x}{\gamma}}, \quad \gamma > 0$
Geman [16]	$\frac{(1+\gamma)x}{\gamma+x}, \quad \gamma > 0$

3.2 Algorithm

The partial augmented Lagrangian function of model (8) is given by:

$$\begin{aligned} \mathcal{L}(U,V,S,\Lambda;\rho) = \|S\|_1 + \lambda\|V\|_\gamma + \\ \langle \Lambda, D - UV^T - S\rangle + \frac{\rho}{2}\|D - UV^T - S\|_F^2, \end{aligned} \tag{9}$$

where $\Lambda \in R^{m\times n}$ is Lagrangian multiplier and ρ is a penalty parameter. Here, we adopt an efficient alternating direction method to iteratively optimize each variable by fixing all the others as follows:

$$U_{k+1} = \arg\min_{U^TU=I} \mathcal{L}(U, V_k, S_k, \Lambda_k; \rho_k); \tag{10}$$

$$V_{k+1} = \arg\min_{V} \mathcal{L}(U_{k+1}, V, S_k, \Lambda_k; \rho_k); \tag{11}$$

$$S_{k+1} = \arg\min_{S} \mathcal{L}(U_{k+1}, V_{k+1}, S, \Lambda_k; \rho_k); \tag{12}$$

$$\Lambda_{k+1} = \Lambda_k + \rho_k(D - U_{k+1}V_{k+1}^T - S_{k+1}); \tag{13}$$

$$\rho_{k+1} = \min\{\beta * \rho_k, \rho_{max}\}. \tag{14}$$

where U_k denotes U in the k-th iteration, and β is set to 1.5 to further facilitate the convergence speed.

In particular, for solving the variable U, the sub-problem is formulated as:

$$U_{k+1} = \arg\min_{U^TU=I} \frac{\rho_k}{2}\|UV_k^T - (T - S_k)\|_F^2 \tag{15}$$

where $T = D + \Lambda_k/\rho_k$. Fortunately, the optimal solution of this sub-problem can be efficiently obtained by resorting to the classical Orthogonal Procrustes problem [38]. Suppose that $(T - S_k)V = A\Sigma B^T$ is the thin SVD of matrix $(T - S_k)V$, then the optimal solution of sub-problem (15) is

$$U_{k+1} = AB^T. \tag{16}$$

To solve the variable V, the optimization problem is formulated as:

$$\begin{aligned} V_{k+1} &= \arg\min_{V} \|V\|_\gamma + \frac{\rho_k}{2\lambda}\|U_{k+1}V^T - (T - S_k)\|_F^2 \\ &= \arg\min_{V} \|V\|_\gamma + \frac{\rho_k}{2\lambda}\|V - (T - S_k)^TU_{k+1}\|_F^2. \end{aligned} \tag{17}$$

Note that sub-problem (17) is non-convex due to the concave sparsity-inducing regularizers. Note that the sub-problem (17) is a combination of concave and convex functions, which can be solved by the difference of convex programming [39]. And the sub-gradient of $\|V\|_\gamma$ is given in Lemma 1.

Lemma 1. *The sub-gradient of $\|V\|_\gamma$ is*

$$\partial\|V\|_\gamma = \{A_V diag(l) B_V^T\}, \tag{18}$$

- *when* $\phi\big(\sigma_i(V)\big) = 1 - e^{\frac{-\sigma_i(V)}{\gamma}}, \quad l_i = \frac{e^{\frac{-\sigma_i(V)}{\gamma}}}{\gamma}$
- *when* $\phi\big(\sigma_i(V)\big) = \frac{(1+\gamma)\sigma_i(V)}{\gamma+\sigma_i(V)}, \quad l_i = \frac{(1+\gamma)\gamma}{\big(\gamma+\sigma_i(V)\big)^2}$

where the columns of A_V and B_V are the left and right singular matrices of V, respectively.

Based on Lemma 1, at the $(k+1)$-th iteration the sub-problem (17) can be relaxed as

$$V_{k+1} = \arg\min_{V} \;\; \langle \partial\|V_k\|_\gamma, V\rangle + \frac{\rho_k}{2\lambda}\|V - (T - S_k)^TU_{k+1}\|_F^2 \tag{19}$$

Compute the derivative of (19) with respect to V and set the derivative to 0, we get

$$\partial\|V_k\|_\gamma + \frac{\rho_k}{\lambda}\Big(V - (T - S_k)^TU_{k+1}\Big) = 0. \tag{20}$$

It is easy to prove that the solution of (20) can be obtained by

$$V_{k+1} = (T - S_k)^TU_{k+1} - \lambda\partial\|V_k\|_\gamma/\rho_k. \tag{21}$$

To solve the variable S, the optimization problem is formulated as:

$$S_{k+1} = \arg\min_{S} \|S\|_1 + \frac{\rho_k}{2}\|S - (T - U_{k+1}V_{k+1}^T)\|_F^2,$$

which has a closed-form solution S_{k+1} by resorting to the element-wise shrinkage-thresholding operator, *i.e.*,

$$[S_{k+1}]_{i,j} = \max\{|[W_k]_{i,j}| - 1/\rho_k, 0\} sign([W_k]_{i,j}) \tag{22}$$

where $W_k = T - U_{k+1}V_{k+1}^T$.

The whole process of our proposed algorithm is presented in Algorithm 1.

Algorithm 1 Our algorithm

Input: The observed data matrix D, parameters λ, ρ_0, β.
Initialize: $\lambda,\ \beta,\ \epsilon,\ \rho_0,\ S_0,\ V_0,\ \Lambda_0,\ k=0$.

```
while not converged do
    Compute T = D + Λ_k/ρ_k;
    Update U_{k+1} by (16);
    Update V_{k+1} by (21);
    Compute W_k = T − U_{k+1} V_{k+1}^T;
    Update S_{k+1} by (22);
    Update Λ_{k+1} by (13);
    Update ρ_{k+1} by (14);
    Check the convergence condition
        ‖D − U_{k+1} V_{k+1}^T − S_{k+1}‖_F ≤ ϵ * ‖D‖_F.
end while
```

Output: The low-rank matrix $L = U_{k+1}V_{k+1}^T$.

3.3 Complexity analysis

The computational complexity of our proposed algorithm is dominated by updating variables: U and V. Updating U and V have a running time of $\mathcal{O}(mr^2 + mnr)$ and $\mathcal{O}(nr^2 + mnr)$, respectively. To update S and Λ all need $\mathcal{O}(mn)$ cost per iteration. Therefore the total complexity of our proposed algorithm is $\mathcal{O}(mr^2 + mnr)$. Without loss of generality, we assume that $r \ll n \leq m$. And the computational complexity of our proposed algorithm would be reduced to $\mathcal{O}(mnr)$ per iteration, which is much less than that of the existing methods for RPCA (2), such as IALM [12].

4 Numerical Experiments

All the experiments were performed on a lenovo laptop with an Intel Core i3-3240T 2.30GHz CPU that has 4 cores and 4GB of memory, running with Windows 8 and MATLAB (R2013a).

4.1 Experimental settings and implementation details

To throughly test the performance of our proposed method, we consider the following three applications:

(1) background extraction in surveillance video: In subsection 4.2, we compare our proposed method with the open source state-of-the-art RPCA-based approaches: IALM [12], NcRPCA [40], GoDec [34], and RBF [36];
(2) contrast-filled vessels extraction: In subsection 4.3, we also compare the proposed method with other open source approach: motion coherency regularized RPCA (MCR-RPCA) [41] to extract contrast-filled vessels from the complex dynamic background;
(3) hyperspectral image (HSI) denoising: In subsection 4.4, we conduct experiment to verify the effectiveness of the proposed method with the hyperspectral image denoising approaches: the video block matching 3-D filtering (VBM3D) [42], GoDec-based low rank matrix recovery (LRMR) [43], noise-adjusted low rank matrix approximation (NAILRMA) [44].

Reasons: The reasons why we choose IALM, NcRPCA, GoDec, and RBF algorithms are as follows. First, to the best of our knowledge, IALM and LRSD (Low Rank and Sparse matrix Decomposition) [13] are the most commonly used algorithms for solving the convex formulation (2). The main difference between the two methods is that IALM performs a partial SVD by only calculating the dominated singular values and vectors while LRSD computes the full SVD. Although the partial SVD scheme requests to estimate the rank of involved matrix, it is faster than the full SVD scheme as the matrix dimensions grow. That is, IALM performs better than LRSD except that the dimension of the involved matrix is low. The state-of-the-art algorithms designed to solve the non-convex RPCA include GoDec, NcRPCA, SpaRCS, etc, where SpaRCS solved a similar problem as NcRPCA while NcRPCA has low computational complexity and gives provable global convergence guarantees. Besides, GoDec, which applied a bilateral random projection scheme to compute the low-rank matrix approximation in order to avert SVD, has outperformed in both noisy and noiseless cases. Hence, we choose GoDec and NcRPCA for comparison in the non-convex formulation case. We select the representative matrix-factorization-based low-rank method: RBF [36] for comparison. In the application of hyperspectral image denoising, we select the recently developed methods under the low-rank modeling framework: LRMR and NAILRMA. Note that LRMR applied the GoDec algorithm to further remove the mixed noise. We download all codes from the authors' websites using the set of default parameters. In order to distinguish two non-convex sparsity-inducing regularizers, we called our algorithm using Laplace and Geman functions as OurLaplace and OurGeman, respectively.

Quantity index: As discussed above, the main advantage of our proposed algorithm is the lower computational cost compared with some existing low-rank approaches, such as IALM, NcRPCA, GoDec, RBF. Following [12], [45], [46], we use the computational time and the number of iterations to measure the computational performance of different methods under the same platform: Matlab. The computational time is the most direct and simplest way to measure the computational performance of different methods. Meanwhile, the computational cost of IALM, RBF, OurLaplace and OurGeman is dominant by performing SVD. Therefore the number of iterations can also reflect the computational performance of different methods. Relative error (RelErr) is used to measure the difference between the original dataset and the recovered foreground and background components, which is defined as $RelErr = \frac{\|D-L-S\|_F}{\|D\|_F}$. The rank prior of matrix L is scored by rank.

4.2 Background extraction in surveillance video

Background extraction in surveillance video [47] is an important application of RPCA. The surveillance video is decomposed into two components: the (approximately) static background which could be represented as a low-rank matrix and the sparse foreground. The parameters in our method are listed as follow: $\gamma = 0.05,\ \ \beta = 1.5,\ \ \rho_0 = 0.01,\ \ \lambda = 20$. In order to estimate the robustness of our method, we set $r = 5$ which is an upper bound of the true rank 1. In this subsection, we evaluate the competing methods on several challenging datasets[1] as shown in Table 2.

1. http://perception.i2r.a-star.edu.sg/bk_model/bk_index.html
http://www.tjucs.win/faculty/likun/projects/bf_separation/index.htm

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2818322, IEEE Access

Fig. 1. Background extraction in five surveillance videos: Car, Escalator, Curtain, Water surface and Highway. The first column: the original frames of Car, Escalator, Curtain, Water surface and Highway, respectively. Second column to the last column: estimated backgrounds and foregrounds by IALM, NcRPCA, GoDec, RBF, OurLaplace, and OurGeman, respectively. The figure is viewed better in zoomed PDF.

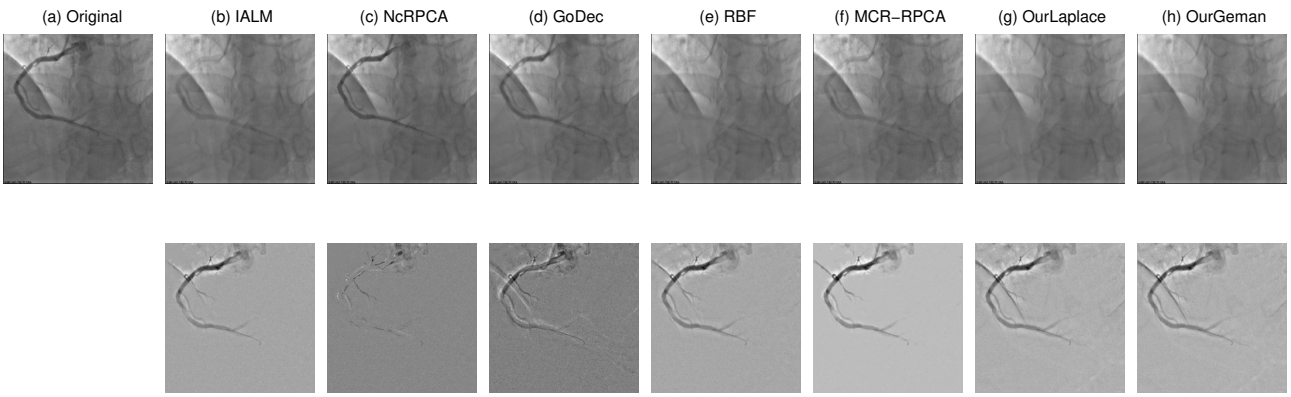

Fig. 2. Contrast-filled vessels extraction on X-ray. The first column: the original frame of X-ray coronary angiograms. Second column to the last column: estimated backgrounds and foregrounds by IALM, NcRPCA, GoDec, RBF, MCR-RPCA, OurLaplace, and OurGeman, respectively. The figure is viewed better in zoomed PDF.

TABLE 2
Quantitative evaluation of the different background extraction algorithms on five real datasets

Data (size)	Algorithm	Rank(L)	$\frac{\|S\|_0}{mn}$	RelErr	Iter	Time(s)
Car ($240\times320\times24$)	IALM	8	0.644	9.47e-4	16	4.66
	NcRPCA	5	**0.624**	7.13e-4	58	27.02
	GoDec	5	0.977	**1e-6**	101	50.57
	RBF	5	0.717	7.95e-4	24	3.54
	OurLaplace	**1**	0.772	7.98e-4	**12**	1.92
	OurGeman	**1**	0.772	7.98e-4	**12**	**1.89**
Escalator ($130\times160\times400$)	IALM	158	0.646	8.96e-4	19	51.04
	NcRPCA	5	0.917	**1.66e-4**	42	49.91
	GoDec	5	**0.216**	2.21e-2	101	158.77
	RBF	**1**	0.883	5.99e-4	23	11.89
	OurLaplace	**1**	0.861	9.47e-4	**11**	**5.81**
	OurGeman	**1**	0.861	9.47e-4	**11**	5.87
Curtain ($128\times160\times633$)	IALM	100	0.745	**6.33e-4**	20	54.60
	NcRPCA	5	0.823	7.62e-4	41	42.84
	GoDec	5	**0.219**	1.04e-2	101	155.65
	RBF	2	0.901	8.72e-4	25	13.35
	OurLaplace	**1**	0.845	9.08e-4	**12**	**6.53**
	OurGeman	**1**	0.845	9.08e-4	**12**	6.81
Water surface ($128\times160\times633$)	IALM	211	0.733	8.67e-4	20	139.83
	NcRPCA	5	0.909	**2.38e-4**	46	76.45
	GoDec	5	**0.139**	2.29e-2	101	224.82
	RBF	**1**	0.837	9.78e-4	23	26.91
	OurLaplace	**1**	0.855	8.05e-4	**12**	10.93
	OurGeman	**1**	0.855	8.05e-4	**12**	**10.60**
Highway ($120\times160\times400$)	IALM	138	0.814	**5.24e-4**	20	49.53
	NcRPCA	5	0.847	6.37e-4	45	73.51
	GoDec	5	**0.234**	1.21e-2	101	193.96
	RBF	**1**	0.874	5.67e-4	24	11.72
	OurLaplace	**1**	0.838	8.95e-4	**12**	5.91
	OurGeman	**1**	0.838	8.95e-4	**12**	**5.89**

The visual results of representative frames in five different surveillance videos are shown in Fig. 1. It is easy to see that for all the testing data, OurLaplace and OurGeman are able to produce clear background and reconstruct a complete foreground as well. However, the competing methods more or less generate residuals in the background or cannot completely detect the moving objects. As shown in the last two rows in Fig. 1, backgrounds obtained by IALM, NcRPCA and GoDec still remain the shadow of moving trucks. Similar observation could be drawn from other three datasets.

We also give the quantitative evaluation by different background/foreground approaches in Table 2 (The best results are highlighted in bold). As it can be seen, the number of iterations (Iter) of OurLaplace and OurGeman are much less than that of other methods. Meanwhile, Ourlaplace and OurGeman are the fastest algorithms, thanks to the matrix factorization scheme and non-convex low-rank regularizers. Since the background is static, the rank of the separated background should be approximately 1. Although we set $r = 5$, the low-rank matrix L recovered by our proposed method is still with the true rank 1 for all the cases, indicating that our proposed method is more robust than other competing methods. Note that for Escalator, Water surface and Highway datasets, RBF has the similar performance with our proposed method. However, the running time of RBF is two times as long as that of our methods.

4.3 Contrast-filled vessels extraction

Due to the rapid advance in modern computer technology, computer-aided medical image processing has been widely used in clinics to facilitate objective disease diagnosis. However, caused by the movement of diaphragm, lung, bones, *ets*, it is still a quite challenging task to extract contrast-filled vessels from the complex dynamic background [41]. Therefore, we also compare our method with the recently proposed method: MCR-RPCA for this task.

The original X-ray coronary angiograms can be downloaded online[2]. For the background (see the first row of Fig. 2), in addition to our algorithms, the backgrounds obtained by other algorithms still contain foreground information, in which NcRPCA and GoDec perform worst. For the foreground (see the second row of Fig. 2), MCR-RPCA performs best with the clearest foreground vessel detection and almost no background information, followed by our proposed method. We give the quantitative evaluation of the different algorithms on X-ray real dataset as shown in Table 3. We can see that the results are similar with that of Table 2. It is worth noting that the computational time of our proposed algorithm is only 1/190 of MCR-RPCA algorithm, which further verifies the lower computation cost of our method.

2. http://www.escience.cn/people/bjqin/research.html

TABLE 3
Quantitative evaluation of the different algorithms on X-ray real dataset

Data (size)	Algorithm	Rank(L)	$\frac{\|S\|_0}{mn}$	RelErr	Iter	Time(s)
X-ray ($256 \times 256 \times 50$)	IALM	18	0.841	**6.20e-4**	17	4.66
	NcRPCA	5	0.847	7.69e-4	48	23.18
	GoDec	5	**0.549**	3.93e-3	101	53.80
	RBF	5	0.882	9.45e-4	38	5.65
	MCR-RPCA	50	0.998	9.49e-4	40	781.85
	OurLaplace	**1**	0.896	6.98e-4	**26**	4.03
	OurGeman	**1**	0.895	7.03e-4	**26**	**3.95**

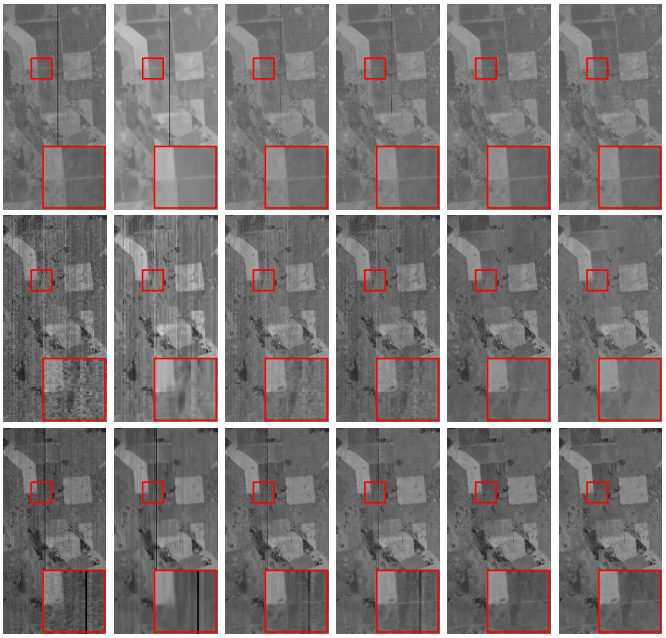

Fig. 3. Restoration results on real data: EO-1 Hyperion Australia. The three rows from top to bottom are the images located at bands 83, 88 and 123; the six columns from left to right are the original images and the restored images obtained by: VBM3D, LRMR, NAILRMA, OurLaplace, and OurGeman, respectively. This figure can be viewed better in zoomed PDF.

4.4 Hyperspectral image denoising

Hyperspectral image denoising has aroused increasing attention on various fields, including environmental studies, military surveillance, and biomedical imaging in recent years. However, HSIs are inevitably contaminated by Gaussian, impulse noise, stripes, dead lines, and many others [19], [43], [44]. Therefore, it is very necessary and important to remove HSI noise. And the denoising methods based on low-rank matrix approximation [19], [43], [44] have achieved promising performance. In this experiment, we select two real-world HSI datasets, *i.e.*, the EO-1 Hyperion Australia dataset[3] and HYDICE urban images[4] to investigate the performance of our proposed method. The original size of EO-1 Hyperion Australia dataset is $3858 \times 256 \times 242$. After removing the overlapping bands between visual near-infrared and shortwave infrared ranges, only a subregion of size $400 \times 200 \times 150$ is used in our experiment as the first testing dataset. While, in the second real-world HSI experiment, we select all 210 bands of size 307×307 as the testing dataset to evaluate the performance for severe noise situation. In [43], 21 bands contaminated by atmosphere and water absorption were discarded. Therefore, our second experimental environment is much difficult and challenging for HSI denoising task. And we use the default parameters for all competing methods, while the parameters of our method are presented in Table 4.

3. http://remote-sensing.nci.org.au/
4. http://www.tec.army.mil/hypercube

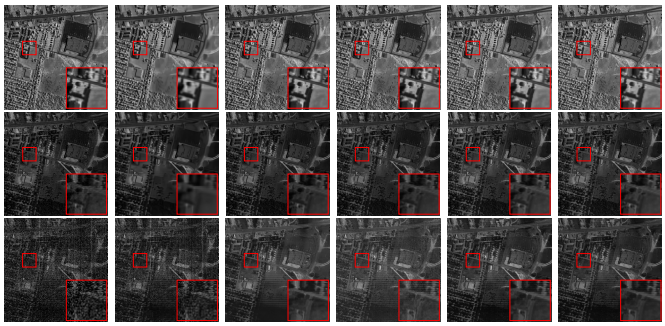

Fig. 4. Restoration results on real data: HYDICE urban images. The three rows from top to bottom are the images located at bands 87, 110 and 207; the six columns from left to right are the original images and the restored images obtained by: VBM3D, LRMR, NAILRMA, OurLaplace, and OurGeman, respectively. This figure can be viewed better in zoomed PDF.

TABLE 4
Parameters setting in our algorithm

Algorithm	λ	ρ_0	γ	patch size	step size
OurLaplace	1	$5e^{-3}$	$5 * 10^{-2}$	50	16
OurGeman	1	$5e^{-3}$	$5 * 10^{-2}$	50	16

The restoration results of six typical bands of two HSIs are shown in Figs. 3 and 4. The original images, see the first column of Fig. 3, are contaminated by the mixture of Gaussian noise, impulse noise, stripes, and dead lines. It is easy to see that our proposed method can effectively remove the mixed noise, and meanwhile, preserve the essential structures of HSIs. Whereas, all the competing methods could remove the mixed noise to a certain extent. More precisely, VBMBD algorithm suffers from oversmoothing the results and fails to restore images under heavy noise. LRMR, NAILRMA can only remove part of the noise. When handling the dead lines, all the restored results are presented ghost shadow more or less at band 123, as shown in the third row of Fig. 3. Only our proposed method achieves a desirable result. Similar observations can be seen from Fig. 4. The experimental results show that our algorithm has great advantage on the removal of stripes and dead lines.

TABLE 5
Comparison of running time (in seconds) for hyperspectral image denoising task on two real-world HSI datasets

Algorithm	VBM3D	LRMR	NAILRMA	OurLaplace	OurGeman
EO-1	240.38	2405.33	260.42	168.81	169.17
HYDICE	425.89	3866.24	362.76	175.73	175.85

We also report the running time of all denoising approaches as shown in Table 5. It is easy to see that

OurLaplace and OurGeman cost the least time, which is much faster than the existing state-of-the-art HSI denoising approaches: NAILRMA and LRMR.

5 Conclusions and Future Work

In this paper, we propose a fast background/foreground separation model, in which we factorize the low-rank matrix into two small-scaled factor matrices and adopt two non-convex functions instead of the nuclear norm to surrogate the rank function. Our proposed method not only overcomes the defects of the nuclear norm but also greatly reduces the computational time and complexity. Experimental results on three important applications demonstrate the efficiency, effectiveness, and robustness of our method. In future, we will extend our method for multi-dimensional data.

Acknowledgment

The authors would like to thank the editor and anonymous reviewers for providing the insightful comments and suggestions which greatly help improve the paper quality.

References

[1] T. Bouwmans, A. Sobral, S. Javed, S. K. Jung, and E.-H. Zahzah, "Decomposition into low-rank plus additive matrices for background/foreground separation: A review for a comparative evaluation with a large-scale dataset," *Computer Science Review*, vol. 23, pp. 1–71, 2017.

[2] X. Zhou, C. Yang, H. Zhao, and W. Yu, "Low-rank modeling and its applications in image analysis," *ACM Computing Surveys (CSUR)*, vol. 47, no. 2, p. 36, 2015.

[3] D. Van Krevelen and R. Poelman, "A survey of augmented reality technologies, applications and limitations," *International Journal of Virtual Reality*, vol. 9, no. 2, p. 1, 2010.

[4] R. Poppe, "A survey on vision-based human action recognition," *Image and vision computing*, vol. 28, no. 6, pp. 976–990, 2010.

[5] W. Cao, Y. Wang, J. Sun, D. Meng, C. Yang, A. Cichocki, and Z. Xu, "Total variation regularized tensor *RPCA* for background subtraction from compressive measurements," *IEEE Transactions on Image Processing*, vol. 25, no. 9, pp. 4075–4090, 2016.

[6] M. Piccardi, "Background subtraction techniques: a review," in *Systems, man and cybernetics, 2004 IEEE international conference on*, vol. 4. IEEE, 2004, pp. 3099–3104.

[7] T. Bouwmans, "Recent advanced statistical background modeling for foreground detection-a systematic survey," *Recent Patents on Computer Science*, vol. 4, no. 3, pp. 147–176, 2011.

[8] N. M. Oliver, B. Rosario, and A. P. Pentland, "A bayesian computer vision system for modeling human interactions," *IEEE transactions on pattern analysis and machine intelligence*, vol. 22, no. 8, pp. 831–843, 2000.

[9] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Journal of the ACM*, vol. 58, no. 3, p. 11, 2011.

[10] A. E. Waters, A. C. Sankaranarayanan, and R. Baraniuk, "Sparcs: Recovering low-rank and sparse matrices from compressive measurements," in *Advances in neural information processing systems*, 2011, pp. 1089–1097.

[11] 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, 2010.

[12] Z. Lin, M. Chen, and Y. Ma, "The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices," *arXiv preprint arXiv:1009.5055*, 2010.

[13] X. Yuan and J. Yang, "Sparse and low rank matrix decomposition via alternating direction method," *Pacific Journal of Optimization*, vol. 9, no. 1, 2009.

[14] D. Goldfarb, S. Ma, and K. Scheinberg, "Fast alternating linearization methods for minimizing the sum of two convex functions," *Mathematical Programming*, pp. 1–34, 2013.

[15] S. Gu, Q. Xie, D. Meng, W. Zuo, X. Feng, and L. Zhang, "Weighted nuclear norm minimization and its applications to low level vision," *International journal of computer vision*, vol. 121, no. 2, pp. 183–208, 2017.

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

[17] Q. Sun, S. Xiang, and J. Ye, "Robust principal component analysis via capped norms," in *Proc. of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD'13)*, Chicago, Illinois, USA, Aug. 2013, pp. 311–319.

[18] Y. Xie, S. Gu, Y. Liu, W. Zuo, W. Zhang, and L. Zhang, "Weighted schatten p-norm minimization for image denoising and background subtraction," *IEEE transactions on image processing*, vol. 25, no. 10, pp. 4842–4857, 2016.

[19] Y. Chen, Y. Guo, Y. Wang, D. Wang, C. Peng, and G. He, "Denoising of hyperspectral images using nonconvex low rank matrix approximation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 9, pp. 5366–5380, 2017.

[20] C. Lu, J. Tang, S. Yan, and Z. Lin, "Nonconvex nonsmooth low rank minimization via iteratively reweighted nuclear norm," *IEEE Transactions on Image Processing*, vol. 25, no. 2, pp. 829–839, 2016.

[21] A. Sobral, T. Bouwmans, and E.-h. ZahZah, "Double-constrained rpca based on saliency maps for foreground detection in automated maritime surveillance," in *Advanced Video and Signal Based Surveillance (AVSS), 2015 12th IEEE International Conference on*. IEEE, 2015, pp. 1–6.

[22] S. Javed, A. Mahmood, T. Bouwmans, and S. K. Jung, "Spatiotemporal low-rank modeling for complex scene background initialization," *IEEE Transactions on Circuits and Systems for Video Technology*, 2016.

[23] ——, "Background–foreground modeling based on spatiotemporal sparse subspace clustering," *IEEE Transactions on Image Processing*, vol. 26, no. 12, pp. 5840–5854, 2017.

[24] N. Wang, T. Yao, J. Wang, and D.-Y. Yeung, "A probabilistic approach to robust matrix factorization," in *European Conference on Computer Vision*. Springer, 2012, pp. 126–139.

[25] D. Meng and F. De La Torre, "Robust matrix factorization with unknown noise," in *Proceedings of the IEEE International Conference on Computer Vision*, 2013, pp. 1337–1344.

[26] Q. Zhao, D. Meng, Z. Xu, W. Zuo, and L. Zhang, "Robust principal component analysis with complex noise," in *International Conference on Machine Learning*, 2014, pp. 55–63.

[27] J. Feng, H. Xu, and S. Yan, "Online robust PCA via stochastic optimization," in *Advances in Neural Information Processing Systems*, 2013, pp. 404–412.

[28] S. Javed, S. H. Oh, J. Heo, and S. K. Jung, "Robust background subtraction via online robust pca using image decomposition," in *Proceedings of the 2014 Conference on Research in Adaptive and Convergent Systems*. ACM, 2014, pp. 105–110.

[29] J. He, L. Balzano, and A. Szlam, "Incremental gradient on the grassmannian for online foreground and background separation in subsampled video," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE, 2012, pp. 1568–1575.

[30] L. Balzano, R. Nowak, and B. Recht, "Online identification and tracking of subspaces from highly incomplete information," in *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*. IEEE, 2010, pp. 704–711.

[31] H. Guo, C. Qiu, and N. Vaswani, "Practical ReProCS for separating sparse and low-dimensional signal sequences from their sumpart 1," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 4161–4165.

[32] P. Rodriguez and B. Wohlberg, "Incremental principal component pursuit for video background modeling," *Journal of Mathematical Imaging and Vision*, vol. 55, no. 1, pp. 1–18, 2016.

[33] F. Seidel, C. Hage, and M. Kleinsteuber, "pROST: a smoothed ℓ_p -norm robust online subspace tracking method for background subtraction in video," *Machine Vision and Applications, Special Issue on Background Modeling for Foreground Detection in Real-World Dynamic Scenes*, vol. 25, no. 5, pp. 1227–1240, 2014.

[34] T. Zhou and D. Tao, "Godec: Randomized low-rank and sparse matrix decomposition in noisy case." in *International Conference on Machine Learning*, 2011, pp. 33–40.

[35] H. Xu, C. Caramanis, and S. Sanghavi, "Robust pca via outlier pursuit." pp. 2496–2504, 2010.

[36] F. Shang, Y. Liu, J. Cheng, and H. Cheng, "Recovering low-rank and sparse matrices via robust bilateral factorization," in *2014*

IEEE International Conference onData Mining (ICDM). IEEE, 2014, pp. 965–970.
[37] Y. Chen, Y. Wang, M. Li, and G. He, "Augmented lagrangian alternating direction method for low rank minimization via non-convex approximation," *Signal, Image and Video Processing*, vol. 11, no. 7, pp. 1271–1278, 2017.
[38] P. H. Schönemann, "A generalized solution of the orthogonal procrustes problem," *Psychometrika*, vol. 31, no. 1, pp. 1–10, 1966.
[39] P. D. Tao and L. T. H. An, "Convex analysis approach to DC programming: theory, algorithms and applications," *Acta Math. Vietnamica*, vol. 22, no. 1, pp. 289–355, 1997.
[40] P. Netrapalli, U. Niranjan, S. Sanghavi, A. Anandkumar, and P. Jain, "Non-convex robust pca," in *Advances in Neural Information Processing Systems*, 2014, pp. 1107–1115.
[41] M. Jin, R. Li, J. Jiang, and B. Qin, "Extracting contrast-filled vessels in x-ray angiography by graduated rpca with motion coherency constraint," *Pattern Recognition*, vol. 63, pp. 653–666, 2017.
[42] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," *IEEE Transactions on image processing*, vol. 16, no. 8, pp. 2080–2095, 2007.
[43] H. Zhang, W. He, L. Zhang, H. Shen, and Q. Yuan, "Hyperspectral image restoration using low-rank matrix recovery," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 4729–4743, Aug. 2014.
[44] W. He, H. Zhang, L. Zhang, and H. Shen, "Hyperspectral image denoising via noise-adjusted iterative low-rank matrix approximation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 6, pp. 3050–3061, 2015.
[45] R. Liu, Z. Lin, and Z. Su, "Linearized alternating direction method with parallel splitting and adaptive penalty for separable convex programs in machine learning," in *Asian Conference on Machine Learning*, 2013, pp. 116–132.
[46] J. Yang and X. Yuan, "Linearized augmented lagrangian and alternating direction methods for nuclear norm minimization," *Mathematics of computation*, vol. 82, no. 281, pp. 301–329, 2013.
[47] L. Li, W. Huang, I. Y. Gu, and Q. Tian, "Statistical modeling of complex backgrounds for foreground object detection." *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, vol. 13, no. 11, pp. 1459–72, 2004.

SUPPLEMENTARY MATERIAL

Model analysis

In this section, for clearer representation, we discuss how these parameters affect to the performance of our proposed algorithm. There are four important parameters, including $\gamma, \lambda, \rho_0, \beta$.

- parameter γ: As stated in [19], the non-convex low-rank matrix approximation $\|V\|_\gamma$ has several good properties, one of which is $\lim_{\gamma \to 0} \|V\|_\gamma = rank(V)$. Which means that the smaller γ is, the tighter $\|V\|_\gamma$ approximates the original rank function. Therefore, we consider as $\gamma = [1e^{-3}, 5e^{-2}, 1e^{-1}, 5e^{-1}, 1]$, and compute the rank and relative error, as shown in Fig. 5. From these figures, we can see that when $\gamma = 1e^{-3}$, the rank of matrix background matrix is zero, which is wrong. In other words, all entries of this matrix are zero. On the other hand, the smaller γ is, the smaller the relative error is. Considering the above two conclusions drawn from Fig. 5, we choose the best value of $\gamma = 5e^{-2}$.
- regularization parameter λ: The regularization parameter λ is used to trade off the contributions between sparse term $\|S\|_1$ and low-rank component $\|V\|_\gamma$. We consider $\lambda = [1e^{-3}, 1e^{-2}, 1e^{-1}, 1e^{1}, 2e^{1}, 5e^{1}, 1e^{2}, 5e^{2}, 1e^{3}, 5e^{3}]$, and compute the rank and relative error, as shown in Fig. 6. From Fig. 6, we can see that when $\lambda = 1e^{1}, 2e^{1}, 5e^{1}, 1e^{2}$, the rank of background matrix is one, which satisfies the fact that the rank of the separated background should be approximately 1 since the background is static. Although $\lambda = 5e^{2}, 1e^{3}, 5e^{3}$, the relative error is small. But the rank of background matrix is zero. So, we choose the best value of $\lambda = 2e^{1}$.
- penalty parameter ρ_0: The initialization penalty parameter ρ_0 should be set to a relative small value. For example, in [19], ρ_0 is selected from the interval $[1e^{-2}, 5e^{-1}]$. Therefore, we select $\rho_0 = 1e^{-2}$.
- parameter β: β is to further facilitate the convergence speed of the iterative algorithm. By consulting other literature, we found that β is usually selected from the interval $[0, \frac{1+\sqrt{5}}{2}]$. In this paper, we select the simplest and straightforward way to set $\beta = 1.5$.

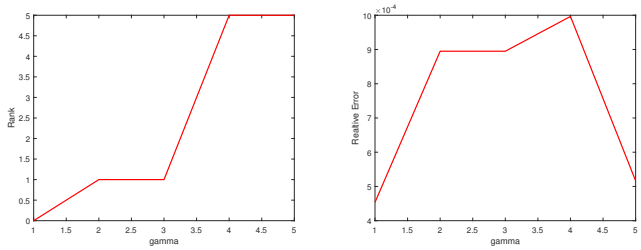

Fig. 5. Left: rank of background VS γ; Right: relative error VS γ.

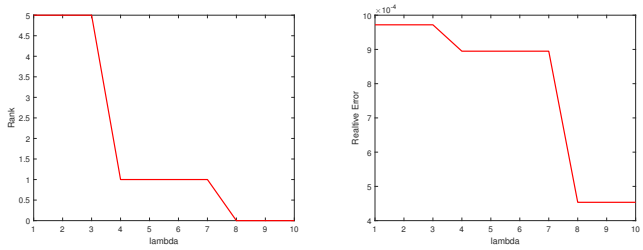

Fig. 6. Left: rank of background VS λ; Right: relative error VS λ.