Noname manuscript No.
(will be inserted by the editor)

Robust PCA Using Nonconvex Rank Approximation and Sparse Regularizer

Jing Dong · Zhichao Xue · Wenwu Wang

Received: date / Accepted: date

Abstract We consider the robust principal component analysis (RPCA) problem where the observed data is decomposed to a low-rank component and a sparse component. Conventionally, the matrix rank in RPCA is often approximated using a nuclear norm. Recently, RPCA has been formulated using the nonconvex ℓ_γ-norm, which provides a closer approximation to the matrix rank than the traditional nuclear norm. However, the low-rank component generally has sparse property, especially in the transform domain. In this paper, a sparsity-based regularization term modeled with ℓ_1-norm is introduced to the formulation. An iterative optimization algorithm is developed to solve the obtained optimization problem. Experiments using synthetic and real data are utilized to validate the performance of the proposed method.

Keywords Robust principal component analysis · ℓ_γ-norm · Sparse prior · Low-rank

1 Introduction

Many applications in signal processing and machine learning involve data of high dimensions, and various dimensionality reduction methods have been developed by projecting the original high-dimensional spaces to low-dimensional spaces [16]. Among these methods, robust principal component analysis (RPCA) is one of the most efficient algorithms, and it reduces the dimensionality of the data based on the low-rank structure of the data and the sparsity of the outliers. RPCA is extended from principal component analysis (PCA) [15] by enhancing the robustness to outliers, and is also known as low-rank and sparse decomposition (LRSD) [7], [31]. RPCA has been

J. Dong (✉) · Z. Xue
College of Electrical Engineering and Control Science, Nanjing Tech University, Nanjing, Jiangsu, China.
E-mail: jingdong@njtech.edu.cn, xuezhichao@njtech.edu.cn

W. Wang
Centre for Vision, Speech and Signal Processing, University of Surrey, Guildford, GU2 7XH, UK.
E-mail: w.wang@surrey.ac.uk

applied to various problems , such as pattern recognition [28], image processing [23], video surveillance [5], [7], background subtraction [8], and image alignment [24].

Assuming the observed data $X \in \mathbb{R}^{m\times n}$ has an underlying low-rank structure, RPCA aims to decompose the data matrix X to a low-rank component $Z \in \mathbb{R}^{m\times n}$ and a sparse component $E \in \mathbb{R}^{m\times n}$. Generally, this problem can be formulated as

$$\begin{aligned} &\min_{Z,E} \ \mathrm{rank}(Z) + \lambda \left\|E\right\|_l \\ &s.t.\ X = Z + E, \end{aligned} \tag{1.1}$$

where $\mathrm{rank}(Z)$ denotes the function that returns the rank of the matrix Z, and $\|E\|_l$ denotes a regularization term like ℓ_0-norm [4], ℓ_1-norm [4], or $\ell_{2,0}$-norm [21] for promoting the sparsity of E. The parameter λ is employed to balance the low-rank and sparse components in X. The optimization problem (1.1) is generally NP-hard as the rank function is discrete and nonconvex. Thus, the rank function is usually relaxed as a convex surrogate. In particular, the nuclear norm, which is defined as the sum of all singular values of a matrix, can be employed as a convex relaxation to address the rank minimization problem [4], [6]. For example, using nuclear norm as the convex surrogate of the rank of Z and ℓ_1-norm to promote the sparsity of E, the RPCA problem can be reformulated as a convex optimization task [10], [29], as both the nuclear norm and ℓ_1-norm in the objective function are convex and the constraint $X = Z + E$ is also convex [3]. In this case, the RPCA problem can be addressed effectively using convex optimization techniques [3], e.g. alternating direction augmented Lagrangian method [10] and proximal gradient method [29].

The low-rank prior involved in RPCA is also widely used in the matrix completion problem [9], however, they are actually two different problems. Firstly, the aim of matrix completion is to recover the original matrix from an incomplete observation, while RPCA aims to recover both the low-rank component and the sparse component from the observed data. Secondly, in low-rank matrix completion, the indices corresponding to the observed entries of the low-rank matrix are given, while related information about the low-rank component in RPCA is unknown.

It should be noted that when nuclear norm is used to approximate the matrix rank, the summation of all singular values is minimized and thus the nonzero singular values make different degrees of contributions to the rank of the matrix. In fact, all nonzero singular values have the same degree of impact on matrix rank. This indicates that the matrix rank cannot be well approximated by the nuclear norm [11], and existing RPCA methods using the nuclear-norm-based relaxation may lead to biased results. Variations of the nuclear norm have been proposed recently to approximate the rank operator more accurately and improve the results of RPCA. For example, the truncated nuclear norm, which is proposed originally for matrix completion [11], [9], has been employed to formulate the rank of a matrix in the RPCA problem [7], and achieved better results as compared with the nuclear-norm-based methods [4],[6]. Kang et al. in [16] present a nonconvex ℓ_γ-norm that can be used as a tighter approximation to the rank of a matrix than the nuclear norm. Although this approximation is nonconvex, an iterative optimization method has been developed, which is shown theoretically to converge to a stationary point.

The existing algorithms for RPCA consider the low-rank property of data in high-dimensional space to model its underlying low-dimensional structure. Note that data in real applications is generally sparse [30], [13], [14], which also reflects the low-dimensional characteristic of the data. The employment of the sparse prior has been demonstrated to be effective in rank minimization related problems, including low-rank matrix completion [9] and RPCA [31]. In this paper, we propose a novel formulation for RPCA by introducing an additional sparsity-based regularizer. In particular, the sparsity-based regularizer promotes the underlying sparse structure of the low-rank component, and ℓ_γ-norm is utilized to model the rank of the matrix to provide a more accurate approximation to matrix rank than the traditional nuclear norm. In addition, we develop an iterative optimization algorithm to solve the nonconvex optimization problem resulting from the proposed formulation.

The rest of this paper is organized as follows. Section 2 introduces related work. Section 3 provides the details of the proposed formulation and the corresponding optimization algorithm. Experimental results are presented in Section 4, and conclusions are drawn in Section 5.

2 Related Work

In general, a typical formulation of the RPCA problem uses the nuclear norm as the convex relaxation of matrix rank, i.e.

$$\begin{aligned} &\min_{Z,E} \ \|Z\|_* + \lambda \|E\|_1 \\ &s.t.\ X = Z + E. \end{aligned} \tag{2.1}$$

Here $\|Z\|_* = \sum_i \sigma_i(Z)$ denotes the nuclear norm of Z where $\sigma_i(Z)$ is the ith largest singular value of Z, and $\|E\|_1 = \sum_{ij} |E_{ij}|$ represents the ℓ_1-norm of E. Many existing RPCA algorithms are based on this formulation and various optimization approaches have been developed to solve this problem. Based on a fast iterative shrinkage-thresholding (FIST) algorithm [1], an accelerated proximal gradient (APG) algorithm is proposed in [26]. The inexact augmented Lagrange multipliers (IALM) method proposed in [20] achieves a trade-off on time and precision. In [32] and [25], the alternating direction method (ADM) is also utilized to solve the RPCA problem via updating the variables alternately.

Since the nuclear-norm-based formulation may lead to biased solutions as explained in Section 1, variations of the nuclear norm have been proposed or employed to formulate the RPCA problem. In [7], Cao et al. applies the truncated nuclear norm (TNN) to the RPCA problem and proposes a novel method named as low-rank and sparse decomposition using truncated nuclear norm (LRSD-TNN), whose formulation is as follows

$$\begin{aligned} &\min_{Z,E} \ \|Z\|_r + \lambda \|E\|_1 \\ &s.t.\ X = Z + E, \end{aligned} \tag{2.2}$$

where $\|Z\|_r$ denotes the truncated nuclear norm of the matrix Z, defined as the summation of the smallest $\min(m,n)-r$ singular values of Z. This truncated nuclear norm based method can obtain better results than the nuclear-norm-based methods.

Based on the LRSD-TNN algorithm [7] and our previous work on low-rank matrix completion [9], we have also introduced the sparse assumption to the formulation of LRSD-TNN, i.e. equation (2.2), and proposed an RPCA algorithm named as low-rank and sparse decomposition using truncated nuclear norm and sparse regularizer (LRSD-TNNSR) [31]. In particular, the low-rank component Z is assumed to be sparse in a transform domain, and the formulation of LRSD-TNNSR is

$$\begin{aligned} &\min_{Z,E} \ \|Z\|_r + \lambda \|E\|_1 + \gamma \|\mathcal{G}(Z)\|_1 \\ &s.t. \ X = Z + E, \end{aligned} \tag{2.3}$$

where the truncated nuclear norm is used to model the rank of the matrix, $\mathcal{G}(\cdot)$ denotes the transform operator, and $\|\mathcal{G}(Z)\|_1$ promotes the sparsity of Z in the transform domain. This algorithm provides better performance than LRSD-TNN in many cases.

As mentioned in Section 1, the truncated nuclear norm approximates the rank of a matrix more accurately than the traditional nuclear norm by only considering the summation of a few smallest singular values and suppressing the influence of the remaining larger singular values on the matrix rank. However, as the truncated nuclear norm is also based on the summation of singular values, larger singular values considered in the summation will still make higher degrees of contributions to the rank of the matrix. Thus, the truncated nuclear norm based methods [7], [9] cannot completely overcome the shortcomings of nuclear-norm-based methods [4],[6].

More recently, Kang et al. in [16] propose a nonconvex function, i.e., γ-norm, as a surrogate of the rank function and present a new nonconvex RPCA (noncvxRPCA) method. In this method, the RPCA problem is formulated as

$$\begin{aligned} &\min_{Z,E} \ \|Z\|_\gamma + \lambda \|E\|_l \\ &s.t. \ X = Z + E, \end{aligned} \tag{2.4}$$

where $\|Z\|_\gamma$ denotes the γ-norm of Z and it is defined as

$$\|Z\|_\gamma = \sum_i \frac{(1+\gamma)\sigma_i(Z)}{\gamma + \sigma_i(Z)}, \ \gamma > 0. \tag{2.5}$$

It is clear that $\lim_{\gamma \to 0} \|Z\|_\gamma = rank(Z)$ and $\lim_{\gamma \to \infty} \|Z\|_\gamma = \|Z\|_*$. With a small value of γ, the γ-norm approximates the rank function more closely than the nuclear norm. In fact, γ-norm can be seen as a scaled version of the traditional nuclear norm. The employment of the factor γ in its definition helps balance the contributions of different singular values. The noncvxRPCA approach is demonstrated to outperform state-of-the-art RPCA algorithms in recovery accuracy [16].

3 Proposed Method

3.1 Problem Formulation

As inherent sparse structures have been revealed in real data under many circumstances, we introduce a sparse prior to the low-rank component of RPCA. In partic-

ular, the low-rank component Z is assumed to be sparse in a transform domain, the proposed formulation for the RPCA problem is as follows

$$\begin{aligned} &\min_{Z,E} \ \|Z\|_\gamma + \lambda \|E\|_1 + \beta\|W\|_1 \\ &s.t.\ X = Z + E \\ &\quad W = \mathcal{G}(Z), \end{aligned} \tag{3.1}$$

where $\mathcal{G}$ denotes the forward transform and $W = \mathcal{G}(Z)$ is the transformed data. The sparsities of W and E are both promoted using the ℓ_1-norm, and the γ-norm is utilized as an approximation to the rank of Z.

It should be noted that this proposed formulation is different from the formulation of the LRSD-TNNSR algorithm [9]. In particular, the proposed formulation (3.1) employs the nonconvex γ-norm as the approximation of matrix rank, while in LRSD-TNNSR, the truncated nuclear norm is utilized instead as shown in equation (2.3). As the γ-norm has the potential to balance the contributions of different singular values to matrix rank better than the truncated nuclear norm, it is used as the surrogate of the rank function in the proposed formulation.

3.2 Optimization Method

The proposed formulation (3.1) is nonconvex, and it is not trivial to obtain the optimal solution. To address this problem, an efficient optimization method based on the framework of the alternating direction method of multipliers (ADMM) is developed. By introducing two multipliers Y and P, and the quadratic penalty terms corresponding to the constraints in (3.1), the augmented Lagrangian function of (3.1) can be obtained, that is

$$\begin{aligned} \mathcal{L}(Z,E,W,Y,P,\mu) = &\|Z\|_\gamma + \lambda\|E\|_l + \beta\|W\|_1 \\ &+ \langle Y, Z+E-X\rangle + \frac{\mu}{2}\|Z+E-X\|_F^2 \\ &+ \langle P, W-\mathcal{G}(Z)\rangle + \frac{\mu}{2}\|\mathcal{G}(Z)-W\|_F^2, \end{aligned} \tag{3.2}$$

where μ is the positive penalty parameter, $\langle\cdot,\cdot\rangle$ returns the inner-product of two matrices, and $\|\cdot\|_F$ denotes the Frobenius norm of a matrix.

Based on ADMM, the solution to the problem (3.2) can be obtained in an iterative way, by only updating one variable at a time and keeping the others fixed. Specifically, in the kth iteration, the variables and the penalty parameter are updated based on the

following steps

$$\begin{cases} Z_{k+1} = \arg\min\limits_{Z} \mathcal{L}\left(Z, E_k, W_k, Y_k, P_k, \mu_k\right), \\ E_{k+1} = \arg\min\limits_{E} \ \mathcal{L}\left(Z_{k+1}, E, W_k, Y_k, P_k, \mu_k\right), \\ W_{k+1} = \arg\min\limits_{W} \mathcal{L}\left(Z_{k+1}, E_{k+1}, W, Y_k, P_k, \mu_k\right), \\ Y_{k+1} = Y_k + \mu_k\left(Z_{k+1} - X + E_{k+1}\right), \\ P_{k+1} = P_k + \mu_k\left[W_{k+1} - \mathcal{G}(Z_{k+1})\right], \\ \mu_{k+1} = \rho\mu_k, \end{cases} \tag{3.3}$$

where $\rho > 1$ is a constant. The details for updating the variables Z, E and W will be presented in the following subsections.

3.2.1 Update Z

The update of Z involves solving the sub-problem as follows

$$\begin{aligned} Z_{k+1} &= \arg\min_{Z} \mathcal{L}\left(Z, E_k, W_k, Y_k, P_k, \mu_k\right) \\ &= \arg\min_{Z} \|Z\|_\gamma + \langle Y_k, Z + E_k - X\rangle \\ &\quad + \frac{\mu}{2}\|Z + E_k - X\|_F^2 + \langle P_k, W_k - \mathcal{G}\left(Z\right)\rangle \\ &\quad + \frac{\mu}{2}\|\mathcal{G}\left(Z\right) - W_k\|_F^2 \\ &= \arg\min_{Z} \|Z\|_\gamma + \frac{\mu_k}{2}\left\|Z - \left(X - E_k - \frac{1}{\mu_k}Y_k\right)\right\|_F^2 \\ &\quad + \frac{\mu_k}{2}\left\|W_k - \mathcal{G}\left(Z\right) + \frac{1}{\mu_k}P_k\right\|_F^2, \end{aligned} \tag{3.4}$$

where $\mathcal{G}(\cdot)$ is assumed to be a unitary transform and its corresponding inverse transform is denoted as $\mathcal{S}(\cdot)$ According to Paseval's theorem [22], a unitary transform $\mathcal{H}(\cdot)$ (e.g., Discrete Fourier Transform, Discrete Cosine Transform (DCT) and Hadamard Transform) can conserve the energy of the original matrix $\boldsymbol{u}$, that is $\|\mathcal{H}(\boldsymbol{u})\|_F^2 = \|\boldsymbol{u}\|_F^2$. Therefore, applying the inverse transform $\mathcal{S}$ to $\left\|W_k - \mathcal{G}\left(Z\right) + \frac{1}{\mu_k}P_k\right\|_F^2$, (3.4) can be recast as

$$\begin{aligned} Z_{k+1} &= \arg\min_{Z} \|Z\|_\gamma + \frac{\mu_k}{2}\left\|Z - \left(X - E_k - \frac{1}{\mu_k}Y_k\right)\right\|_F^2 \\ &\quad + \frac{\mu_k}{2}\left\|Z - \mathcal{S}\left(W_k + \frac{1}{\mu_k}P_k\right)\right\|_F^2 \\ &= \arg\min_{Z} \|Z\|_\gamma + \mu_k\left\|Z - \frac{1}{2}\left[\left(X - E_k - \frac{1}{\mu_k}Y_k\right)\right.\right. \\ &\qquad \left.\left. + \mathcal{S}\left(W_k + \frac{1}{\mu_k}P_k\right)\right]\right\|_F^2. \end{aligned} \tag{3.5}$$

To address the γ-norm minimization problem (3.5), the following theorem can be used [16].

Theorem 1 *Let $A = Udiag(\sigma_A)V^T$ denote the SVD of $A \in \mathbb{R}^{m\times n}$, and $F(Z) = f \circ \sigma_Z$ denote a unitarily invariant function, where σ_A and σ_Z denote the singular values of A and Z, respectively. The optimal solution to the problem*

$$\arg\min_X F(X) + \frac{\mu}{2}\|X - A\|_F^2 \tag{3.6}$$

is $X^* = Udiag(\sigma^*)V$, where $\sigma^* = prox_{f,\mu}(\sigma_A)$ is the proximity operator of f with penalty μ, defined as

$$prox_{f,\mu}(\sigma_A) := \arg\min_{\sigma\geqslant 0}\ f(\sigma) + \frac{\mu}{2}\|\sigma - \sigma_A\|_2^2. \tag{3.7}$$

Based on the theorem above, the optimal solution to (3.5) is

$$Z_{k+1} = Udiag(\sigma^*)V, \tag{3.8}$$

where σ^*, the solution to (3.7), can be approximated by linearizing the concave term $f(\sigma)$ iteratively. Specifically, in the $(l+1)$th inner iteration, σ can be updated as follows

$$\begin{aligned}\sigma_{l+1} &= \arg\min_{\sigma\geqslant 0}\ \langle \nabla_\sigma f(\sigma^l), \sigma\rangle + \frac{\mu}{2}\|\sigma - \sigma_A\|_2^2 \\ &= \max\{\sigma_A - \frac{\nabla_\sigma f(\sigma^l)}{\mu}, 0\},\end{aligned} \tag{3.9}$$

where

$$A = \frac{1}{2}\left[\left(X - E_k - \frac{1}{\mu_k}Y_k\right) + \mathcal{S}\left(W_k + \frac{1}{\mu_k}P_k\right)\right], \tag{3.10}$$

$\nabla_\sigma f(\sigma^l)$ is the gradient of f at σ^l, and $\mu = 2\mu_k$.

3.2.2 Update E

The variable E is updated by solving the problem as follows

$$\begin{aligned}E_{k+1} &= \arg\min_E \mathcal{L}\left(Z_{k+1}, E, W_k, Y_k, P_k, \mu_k\right) \\ &= \arg\min_Z \lambda\|E\|_1 + \langle Y_k, Z_{k+1} + E - X\rangle \frac{\mu}{2}\|Z_{k+1} + E - X\|_F^2 \\ &= \arg\min_E \lambda\|E\|_1 + \frac{\mu_k}{2}\left\|E - \left(X - Z_{k+1} - \frac{1}{\mu_k}Y_k\right)\right\|_F^2.\end{aligned} \tag{3.11}$$

The solution to this problem is [2]

$$E_{k+1} = ST_{\frac{\lambda}{\mu_k}}\left[X - Z_{k+1} - \frac{1}{\mu_k}Y_k\right], \tag{3.12}$$

and here $ST_{\frac{\lambda}{\mu_k}}$ denotes the element-wise soft-thresholding operator which is defined as

$$ST_\tau(x) = \mathrm{sgn}(x)\cdot\max\{|x| - \tau, 0\}, \tag{3.13}$$

with the function $\mathrm{sgn}(\cdot)$ returning the sign of the given operand.

3.2.3 Update W

Based on the steps given in (3.3), W is updated by addressing the following problem

$$\begin{aligned} W_{k+1} &= \arg\min_{W} \mathcal{L}\left(Z_{k+1}, E_{k+1}, W, Y_k, P_k, \mu_k\right) \\ &= \arg\min_{Z} \beta \left\|W\right\|_1 + \left\langle P_k, W_k - \mathcal{G}\left(Z_{k+1}\right)\right\rangle + \frac{\mu}{2}\left\|\mathcal{G}\left(Z_{k+1}\right) - W_k\right\|_F^2 \\ &= \arg\min_{W} \beta\|W\|_1 + \frac{\mu_k}{2}\left\|W - \mathcal{G}\left(Z_{k+1}\right) + \frac{1}{\mu_k}P_k\right\|_F^2. \end{aligned} \tag{3.14}$$

Similar to (3.11), the above problem has the closed-form solution as follows [2]

$$W_{k+1} = ST_{\frac{\beta}{\mu_k}}\left[\mathcal{G}\left(Z_{k+1}\right) - \frac{1}{\mu_k}P_k\right]. \tag{3.15}$$

3.2.4 Summary of the optimization method

The complete procedure to solve the proposed model (3.1) is summarized in Algorithm 1.

Algorithm 1 Optimization method to address the proposed model (3.1)

Input: $X, \lambda, \beta, \gamma, \mu_1, \rho, \epsilon$.

Initialization: Initialize the iteration number $k = 1$, $Z_1 = X$ and E_1, Y_1, W_1, P_1 as zero matrices.

Repeat

1. Update Z:
   Obtain σ^* iteratively based on equation (3.9), and update Z using $Z_{k+1} = Udiag(\sigma^*)V$.
2. Update E:
   $E_{k+1} = ST_{\frac{\lambda}{\mu_k}}\left(X - Z_{k+1} - \frac{Y_k}{\mu_k}\right)$.
3. Update W:
   $W_{k+1} = ST_{\frac{\beta}{\mu_k}}\left[\mathcal{G}\left(Z_{k+1}\right) - \frac{P_k}{\mu_k}\right]$.
4. Update Y: $Y_{k+1} = Y_k + \mu_k(Z_{k+1} - X + E_{k+1})$.
5. Update P: $P_k + \mu_k[W_{k+1} - \mathcal{G}\left(Z_{k+1}\right)]$.
6. Update μ: $\mu_{k+1} = \rho\mu_k$.

Until $\|Z_{k+1} - Z_k\|_F \leq \epsilon$, or $\|E_{k+1} - E_k\|_F \leq \epsilon$

Return Z and E

4 Simulation Results

Experiments with synthetic data and real data are performed to demonstrate the effectiveness of the proposed approach. The applications to real data contain face image shadow removal, singing voice separation, and video background subtraction. The proposed algorithm is compared with several state-of-the-art algorithms including LRSD-TNNSR [31], noncvxRPCA [16], LRSD-TNN [7] and IALM [20][1].

4.1 Experiments with synthetic data

In this experiment, randomly generated matrices are used to evaluate the performance of the proposed algorithm. Each synthetic matrix X_0 of size $m \times n$ is composed of a low-rank matrix Z_0 and a sparse matrix E_0, i.e. $X_0 = Z_0 + E_0$, where the rank of Z_0 is r and the sparse ratio of E_0 is spr. In particular, the low-rank matrix Z_0 is generated based on equation $Z_0 = LR^T$ where the matrices $L \in \mathbb{R}^{m\times r}$ and $R \in \mathbb{R}^{n\times r}$ are randomly generated using Gaussian distribution with zero mean and unit variance. The non-zero entries of the sparse matrix E_0 are independently and uniformly distributed in the range $[-t, t]$, where t denotes the maximum of the absolute values of all elements in Z_0.

The performances of the algorithms are measured with total reconstruction error (*Totalerr*), low-rank reconstruction error (*LRerr*), and sparse reconstruction error (*Sperr*). These measurements are computed as follows

$$Totalerr = \frac{\|X - X_0\|_F}{\|X_0\|_F}, \tag{4.1}$$

$$LRerr = \frac{\|Z - Z_0\|_F}{\|Z_0\|_F}, \tag{4.2}$$

$$Sperr = \frac{\|E - E_0\|_F}{\|E_0\|_F}, \tag{4.3}$$

where X_0, Z_0, and E_0 denote the ground-truth matrices in the generated synthetic data, and X, Z and E denote the matrices recovered using the algorithms.

In the proposed algorithm, the regularization parameters are set as $\lambda = 0.4$ and $\beta = 0.1$. The initial penalty parameter of the quadratic penalty terms is set as $\mu_1 = 0.63$ empirically, and the coefficient for updating μ is set as $\rho = 1.1$. The parameters of noncvxRPCA are empirically set as $\lambda = 0.1$, $\mu_1 = 0.9$, and $\rho = 1.1$. The parameter γ in the γ-norm term in both the proposed algorithm and noncvxRPCA is set as 0.01. In the LRSD-TNNSR method, the parameters are set as $\lambda = 0.9/\sqrt{max(m,n)}$ and $\gamma = 0.9/\sqrt{max(m,n)}$. The parameters of the LRSD-TNN and IALM algorithms are set as the values suggested in the original papers respectively.

[1] The codes of noncvxRPCA were downloaded from the website `https://github.com/sckangz/noncvx-PRCA`. As the codes of the LRSD-TNN algorithm are not available, we implemented this algorithm by ourselves. The codes of IALM were downloaded from `http://perception.csl.illinois.edu/matrix-rank/sample_code.html`.

Various parameters related to the synthetic matrices are used for illustrating the performances of the algorithms in different situations. Specifically, the size of the matrices is set as $m = n \in \{200, 500\}$, the rank of the low-rank component Z_0 is $r = 5$ or 10, and the sparse ratio of the sparse component E_0 is $0.01mn$ or $0.05mn$. The results obtained by different algorithms are summarized in Tables 1-4. In general, all algorithms can obtain good results in decomposing the low-rank component and the sparse component. The LRSD-TNNSR algorithm achieves the best performance in most cases, except when $spr = 0.05mn$. The proposed algorithm outperforms noncvxRPCA and IALM in all cases, which demonstrates the effectiveness of the proposed method.

Table 1 $m = n = 200, rank(Z_0) = 5, sparse\ ratio = 0.01mn.$

Algorithm	Totalerr	LRerr	Sperr	Iteration	Time
Proposed	1.63×10^{-7}	1.96×10^{-7}	9.36×10^{-8}	46	0.6516
LRSD-TNNSR	1.96×10^{-10}	2.27×10^{-10}	3.40×10^{-10}	68	0.7329
noncvxRPCA	4.48×10^{-7}	5.51×10^{-7}	3.24×10^{-7}	20	0.2622
LRSD-TNN	1.21×10^{-10}	1.50×10^{-10}	9.30×10^{-11}	38	0.3544
IALM	8.81×10^{-6}	2.43×10^{-6}	1.83×10^{-5}	8	0.7678

Table 2 $m = n = 500, rank(Z_0) = 5, sparse\ ratio = 0.01mn.$

Algorithm	Totalerr	LRerr	Sperr	Iteration	Time
Proposed	1.93×10^{-7}	2.34×10^{-7}	8.20×10^{-8}	44	4.1578
LRSD-TNNSR	9.34×10^{-11}	1.11×10^{-10}	1.97×10^{-11}	69	4.9170
noncvxRPCA	3.49×10^{-7}	4.31×10^{-7}	1.97×10^{-7}	11	1.2281
LRSD-TNN	4.35×10^{-11}	4.96×10^{-11}	1.56×10^{-11}	36	2.2404
IALM	7.35×10^{-6}	2.44×10^{-6}	1.39×10^{-5}	8	1.4911

4.2 Face image shadow removal

Face images of the same subject under different illumination conditions generally lie in a low-dimensional subspace, while the outliers resulting from lighting variations can be assumed to be sparse [29], [4]. Therefore, RPCA algorithms can be used to deal with the task of face shadow image removal [29], [7]. In this subsection, we use this application of RPCA to evaluate the performances of the algorithms.

Face images from the Extended Yale B dataset [18] are used in our experiments. This dataset contains face images of 39 subjects and for each subject there are 64

Table 3 $m = n = 200, rank(Z_0) = 5, sparse\ ratio = 0.05mn.$

Algorithm	Totalerr	LRerr	Sperr	Iteration	Time
Proposed	1.14×10^{-7}	2.23×10^{-7}	6.67×10^{-8}	52	0.7194
LRSD-TNNSR	1.21×10^{-11}	0.89	0.61	215	2.0218
noncvxRPCA	9.27×10^{-7}	3.85×10^{-2}	2.64×10^{-2}	61	1.0188
LRSD-TNN	1.17×10^{-10}	1.54	1.06	210	1.6060
IALM	1.93×10^{-6}	1.76×10^{-6}	2.60×10^{-6}	11	0.8309

Table 4 $m = n = 200, rank(Z_0) = 10, sparse\ ratio = 0.01mn.$

Algorithm	Totalerr	LRerr	Sperr	Iteration	Time
Proposed	8.49×10^{-8}	1.09×10^{-7}	7.50×10^{-8}	53	0.7187
LRSD-TNNSR	3.51×10^{-10}	4.11×10^{-10}	6.50×10^{-11}	65	0.7331
noncvxRPCA	6.80×10^{-7}	1.10×10^{-2}	1.81×10^{-2}	60	1.0462
LRSD-TNN	1.07×10^{-10}	1.26×10^{-10}	1.80×10^{-10}	197	1.6446
IALM	7.03×10^{-6}	3.86×10^{-6}	1.49×10^{-5}	10	0.6963

images with resolution 192×168 captured with various environmental illuminations. In the experiments, each sample of a subject is reshaped as a column vector of size 32256×1, and a matrix of size 32256×64 corresponding to this subject is constructed by using each of the samples as one column. This matrix is assumed to be composed of a low-rank matrix corresponding to the face images without shadows and a sparse matrix reflecting shadows in the images, and RPCA algorithms are employed to remove shadows by recovering the low-rank component from the observed data.

The parameters in the proposed algorithm are set as $\lambda = 10^{-3}$, $\beta = 10^{-4}$ and $\mu_1 = 0.3$. The parameters in noncvxRPCA method are empirically set as $\lambda = 10^{-3}$ and $\mu_1 = 0.5$. The parameters ρ and γ in these two algorithms are the same as in experiments with synthetic data. The parameters of LRSD-TNNSR, LRSD-TNN and IALM algorithm are set as in the original papers.

Experimental results for subjects *yaleB01* and *yaleB05* in the Extended Yale B dataset are shown in Figs. 1 and 2, respectively. We can find that both LRSD-TNN and IALM can only remove light shadows, as shown in the first samples of yaleB01 and yaleB05. For strong shadows in face images, as the second samples of the subjects, the proposed algorithm, LRSD-TNNSR and noncvxRPCA outperform LRSD-TNN and IALM significantly. For the subject yaleB05 shown in Fig. 2, the results of the proposed algorithm are similar to those of the noncvxRPCA and LRSD-TNNSR algorithms. For the subject yaleB01 in Fig. 1, the proposed algorithm can achieve much better results than noncvxRPCA. This demonstrates the superiority of the proposed algorithm as compared with the baselines.

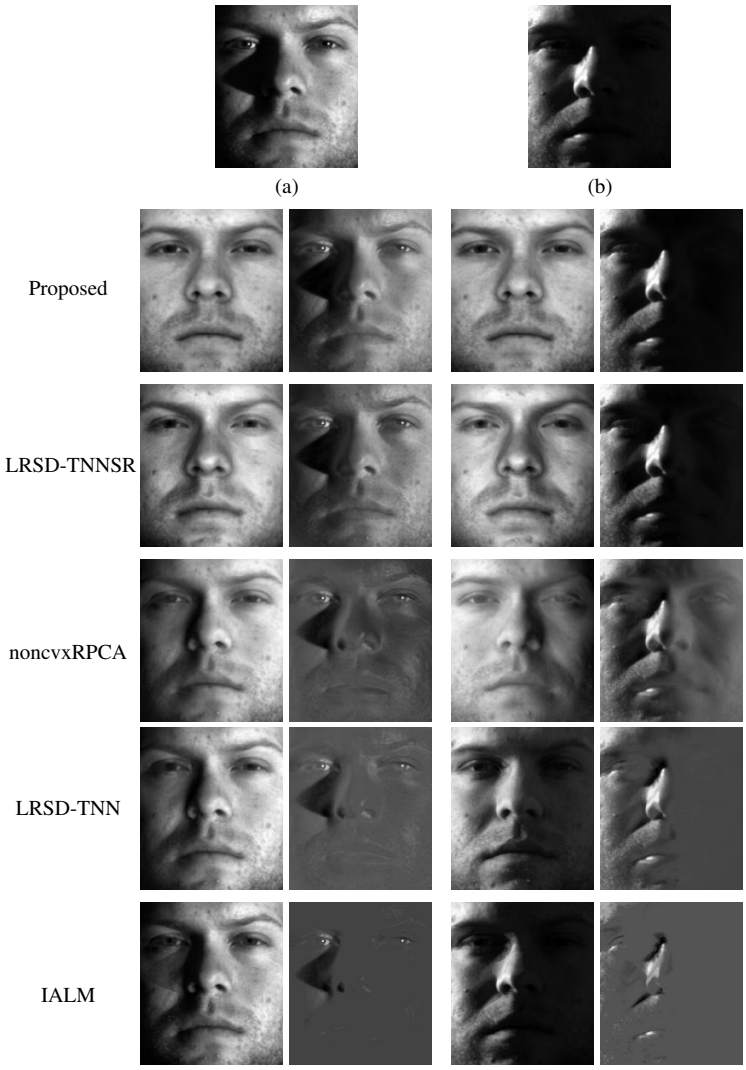


Fig. 1 Experimental results of shadow removal for face images of subject *yaleB01*. The subfigures (a) and (b) are two sample images of *yaleB01*. The subfigures below the samples are the low-rank and the sparse components recovered from the corresponding samples, using the proposed method, LRSD-TNNSR, noncvxRPCA, LRSD-TNN, and IALM, respectively.

4.3 Singing voice separation

Music accompaniment in a song can be assumed to lie in a low-rank subspace due to the repetition structure, and singing voices with more variations can be considered to be sparse. Based on this assumption, RPCA can be used to solve the singing voice separation problem [12].

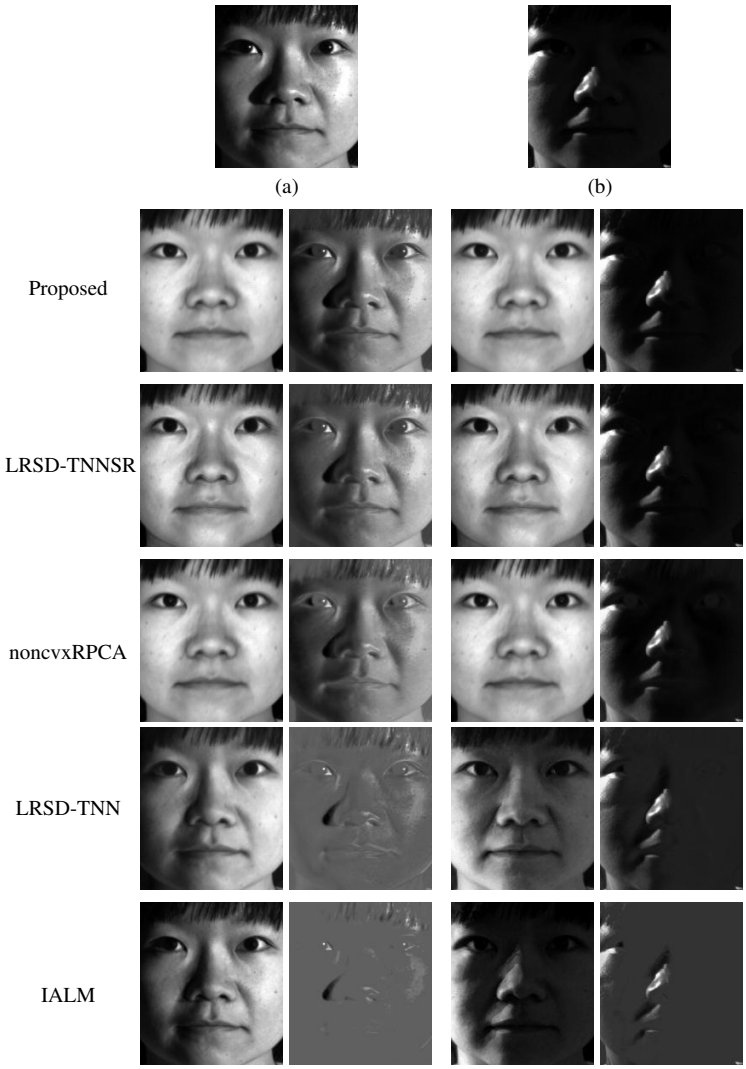


Fig. 2 Experimental results of shadow removal for face images of subject *yaleB05*. The subfigures (a) and (b) are two sample images of *yaleB05*. The subfigures below the samples are the low-rank and the sparse components recovered from the corresponding samples, using the proposedmethod, LRSD-TNNSR, noncvxRPCA, LRSD-TNN, and IALM, respectively.

In the experiment, MIR-1K[2] database is employed as test data. The singing voice and the music accompaniment are mixed at 5 dB signal to noise ratio (SNR). Following the experiments in [12], the spectrogram of the mixture is computed via the Short-Time-Fourier Transform (STFT) with window size being 1024 and hop size

[2] http://perception.i2r.astar.edu.sg/bk_model/bk_index.html.

being 256, and the RPCA methods are applied to the obtained spectrogram matrix to estimate the singing voice component.

In order to evaluate the separation results of the algorithms, we compute energy ratios utilizing BSS-EVAL [27], [12] in terms of source to distortion ratio (SDR), source to interference ratio (SIR), source to artifacts ratio (SAR) [17], and the normalized SDR (NSDR), which are defined as

$$SDR = 10\log 10\frac{\|S_{target}\|^2}{\|e_{interf} + e_{noise} + e_{artif}\|^2}, \tag{4.4}$$

$$SIR = 10\log 10\frac{\|S_{target}\|^2}{\|e_{interf}\|^2}, \tag{4.5}$$

$$SAR = 10\log 10\frac{\|S_{target} + e_{interf} + e_{noise}\|^2}{\|e_{artif}\|^2}, \tag{4.6}$$

$$NSDR(\hat{v}, v, x) = SDR(\hat{v}, v) - SDR(x, v). \tag{4.7}$$

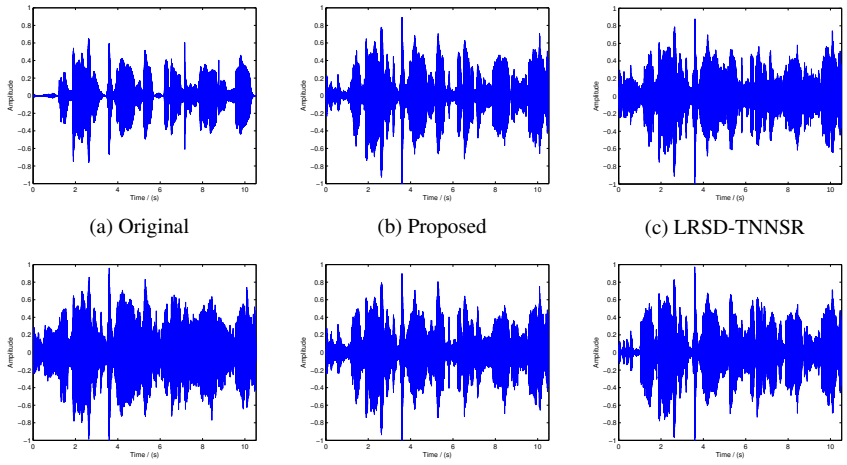
Here S_{target} denotes the energy of the true component of target signal from the separation results, e_{interf}, e_{noise} and e_{artif} are the interference, noise, and artifact error terms, respectively. $\hat{v}$ and v denote the reconstructed singing voice and the original clean singing voice respectively, and x denotes the mixture. In addition, the *Totalerr*, which has been used in the experiments with synthetic data, is employed to evaluate the overall performance of the algorithms.

The parameters in the proposed algorithm are set as $\lambda = 28/\sqrt{max(m,n)}$, $\beta = 0.3$, and $\mu_1 = 0.003$. The parameters in the LRSD-TNNSR method are set as $\lambda = 0.0095/\sqrt{max(m,n)}$ and $\gamma = 0.003/\sqrt{min(m,n)}$. The parameters in noncvxRPCA method are set as $\lambda = 1/\sqrt{max(m,n)}$ and $\mu_1 = 0.1$. The parameters ρ and γ in these two algorithms are the same as in the previous experiments. The parameters of LRSD-TNN and IALM algorithm are set as suggested in the original papers.

Table 5 Experimental results for singing voice separation.

Algorithm	SDR	SIR	SAR	NSDR	Totalerr
Proposed	8.18	18.32	8.69	10.74	9.10×10^{-9}
LRSD-TNNSR	4.14	12.08	5.16	6.70	3.30×10^{-6}
noncvxRPCA	3.75	7.73	6.66	6.32	8.46×10^{-8}
LRSD-TNN	6.06	14.36	6.91	8.62	2.61×10^{-4}
IALM	6.33	12.74	7.67	8.89	2.72×10^{-6}

Table 5 shows the results of the proposed method, the LRSD-TNNSR, noncvxRPCA, LRSD-TNN and IALM algorithms on singing voice separation. Fig. 3 shows the waveform of the original signing voice and waveforms of singing voices recovered by different algorithms. In terms of SDR, SIR, SAR, and NSDR, the proposed



(a) Original (b) Proposed (c) LRSD-TNNSR

(d) noncvxRPCA (e) LRSD-TNN (f) IALM

Fig. 3 Results for singing voice separation. (a) shows the waveform of the original singing voice, subfigures (b), (c), (d), (e) and (f) present waveforms of the singing voice separated by the proposed method, LRSD-TNNSR, noncvxRPCA, LRSD-TN, and IALM, respectively.

method outperforms the baseline algorithms, and it also achieves a higher reconstruction accuracy according to *Totalerr*. From Fig. 3, it can be seen that the voice waveform separated by the proposed method is much closer to the original waveform.

4.4 Video background subtraction

Video background subtraction is another important application of RPCA algorithms, as video frames captured by a fixed camera can be regarded as the sum of low-rank background and sparse foreground [16]. Two scenes *escalator* and *hall* from Perception Test Images Sequences [19] are used as the test data in this experiment. The video data of scene *escalator* which consists of 3417 frames of resolution 160×130 is converted to an observed matrix of size 20800×3417, and the video data of scene *hall* containing 3584 frames of resolution 192×144 is converted to an observed matrix of size 27648×3584.

The parameters of the proposed algorithm are set as $\lambda = 0.2$, $\beta = 0.1$ and $\mu_1 = 0.39$. The parameters of noncvxRPCA are set as $\lambda = 10^{-3}$ and $\mu_1 = 0.5$. Other parameters of these two algorithms are the same as the settings of experiments in the previous subsections. The parameters of LRSD-TNNSR, LRSD-TNN and IALM algorithm are set as in the original papers.

The results of video background subtraction using different algorithms are given in Figs. 4 and 5. It can be observed that all algorithms can decompose the video frames into two distinct parts. In the results for the scene *escalator* as shown in Fig. 4, the background reconstructed by the proposed algorithm has better quality as compared with those from the baselines which still contain some contents from

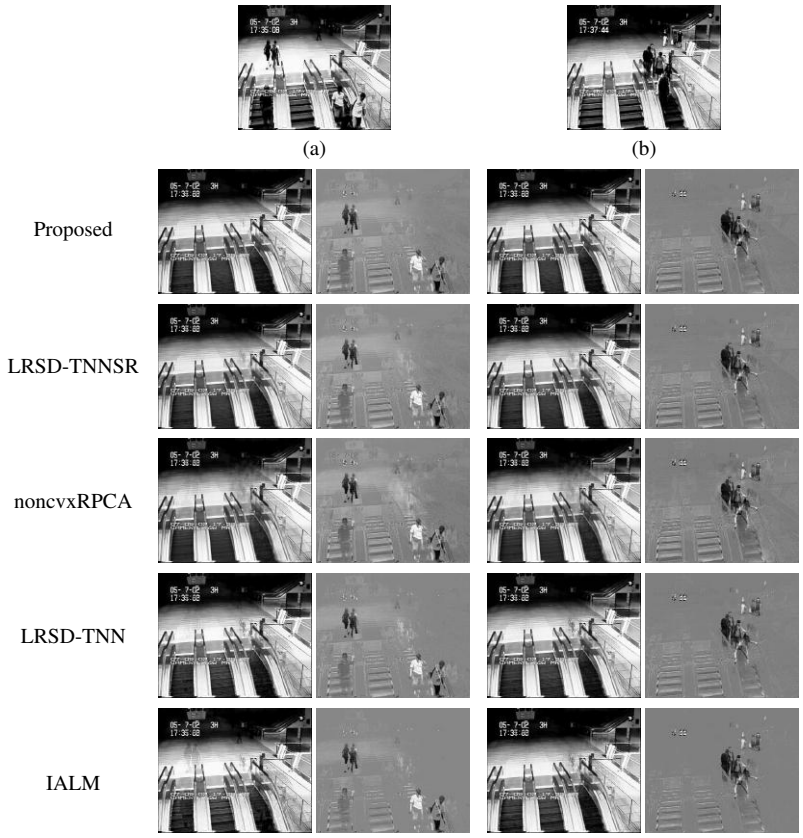


Fig. 4 The results of background subtraction for scene *escalator*. The subfigures (a) and (b) are two sample frames of the video. The subfigures below the samples are the low-rank background components and the sparse foreground components of the corresponding samples, which are decomposed by the proposed method, LRSD-TNNSR, noncvxRPCA, LRSD-TNN, and IALM, respectively.

the foreground, e.g. people on the escalator. For the results of scene *hall*, in the background components obtained by LRSD-TNNSR, noncvxRPCA and LRSD-TNN for the frame (a), as shown in Fig. 5, there is a person with a suitcase near the reception desk, which does not exist in the original sample frame to be decomposed. This probably results from the influence of other frames, e.g. sample frame (b), in the video. The results obtained by IALM for the frame (a) of *hall* also have been affected by other frames, and there contain some foreground in the background component extracted from the frame (b). The proposed algorithm does not introduce any extra contents that do not exist in the original frame and achieves the best performance in general.

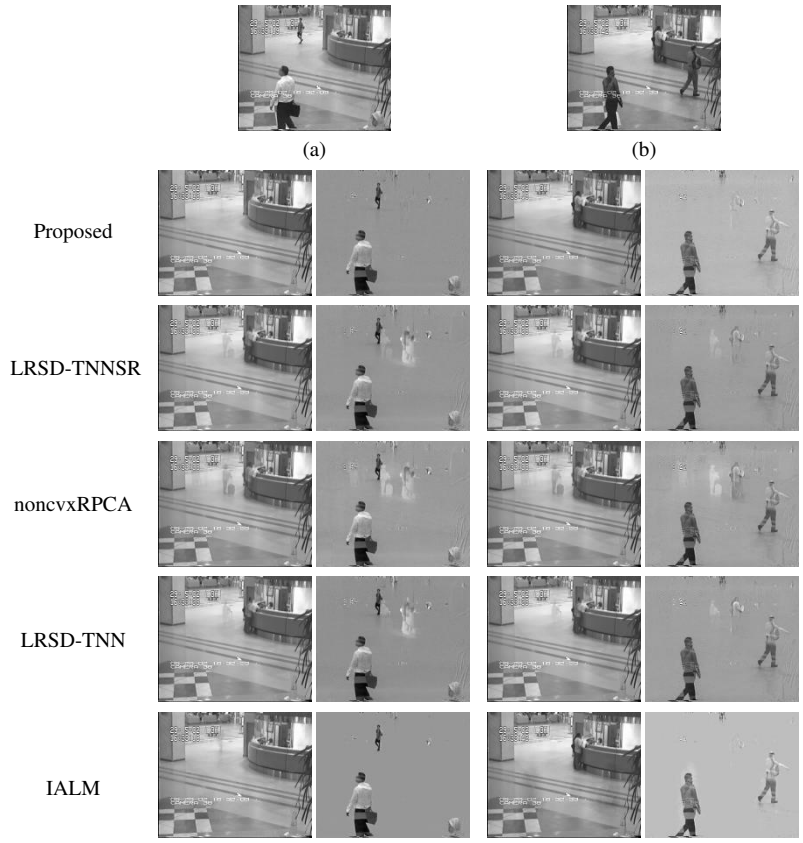


Fig. 5 The results of background subtraction for scene *hall*. The subfigures (a) and (b) are two sample frames of the video. The subfigures below the samples are the low-rank background components and the sparse foreground components of the corresponding samples, which are decomposed by the proposed method, LRSD-TNNSR, noncvxRPCA, LRSD-TNN, and IALM, respectively.

5 Conclusion

We have proposed a novel formulation for the RPCA problem and the corresponding optimization method. By exploiting the sparse property of the low-rank component, a sparse regularizer represented as the form of ℓ_1-norm is introduced to the formulation. Simultaneously, ℓ_γ-norm is applied to approximate to the rank function. To address the proposed optimization problem, we have developed an optimization algorithm by introducing dummy variables and updating variables alternatively. Experimental results on synthetic and real applications including face image shadow removal, singing voice separation and video background subtraction have demonstrated the superiority of the proposed method as compared with several baseline RPCA methods.

Acknowledgements This work was supported by the National Natural Science Foundation of China (61906087), the Natural Science Foundation of Jiangsu Province of China (BK20180692) and the Natural Science Foundation of the Higher Education Institutions of Jiangsu Province of China (17KJB510025). The authors thank the Associate Editor and the anonymous reviewers for their contributions to improving the quality of the paper.

References

1. Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM Journal on Imaging Sciences **2**(1), 183–202 (2009)
2. Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J.: Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends in Machine Learning **3**(1), 1–122 (2011)
3. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge Univeristy Press (2004)
4. Candès, E.J., Li, X., Ma, Y., Wright, J.: Robust principal component analysis? Journal of the ACM (JACM) **58**(3), 11 (2011)
5. Candès, E.J., Recht, B.: Exact matrix completion via convex optimization. Foundations of Computational Mathematics **9**(6), 717 (2009)
6. Candes, E.J., Tao, T.: The Power of Convex Relaxation: Near-Optimal Matrix Completion. IEEE Press (2010)
7. Cao, F., Chen, J., Ye, H., Zhao, J., Zhou, Z.: Recovering low-rank and sparse matrix based on the truncated nuclear norm. Neural Networks **85**, 10–20 (2017)
8. Cao, W., Wang, Y., Sun, J., Meng, D., Yang, C., Cichocki, A., Xu, Z.: Total variation regularized tensor RPCA for background subtraction from compressive measurements. IEEE Transactions on Image Processing **25**(9), 4075–4090 (2016)
9. Dong, J., Xue, Z., Guan, J., Han, Z.F., Wang, W.: Low rank matrix completion using truncated nuclear norm and sparse regularizer. Signal Processing: Image Communication **68**, 76–87 (2018)
10. Goldfarb, D., Ma, S., Scheinberg, K.: Fast alternating linearization methods for minimizing the sum of two convex functions. Mathematical Programming **141**(1-2), 349–382 (2013)
11. Hu, Y., Zhang, D., Ye, J., Li, X., He, X.: Fast and accurate matrix completion via truncated nuclear norm regularization. IEEE Transactions on Pattern Analysis and Machine Intelligence **35**(9), 2117–2130 (2013)
12. Huang, P.S., Chen, S.D., Smaragdis, P., Hasegawa-Johnson, M.: Singing-voice separation from monaural recordings using robust principal component analysis. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 57–60. IEEE (2012)
13. Jianchao, Y., John, W., Thomas, H., Yi, M.: Image super-resolution via sparse representation. IEEE Transactions on Image Processing **19**(11), 2861–2873 (2010)
14. John, W., Yang, A.Y., Arvind, G., S Shankar, S., Yi, M.: Robust face recognition via sparse representation. IEEE Transactions on Pattern Analysis and Machine Intelligence **31**(2), 210–227 (2009)
15. Jolliffe, I.: Principal component analysis. Springer (2011)
16. Kang, Z., Peng, C., Cheng, Q.: Robust PCA via nonconvex rank approximation. In: IEEE International Conference on Data Mining (ICDM), pp. 211–220. IEEE (2015)
17. Kumar, V.A., Rao, C.V.R., Dutta, A.: Performance analysis of blind source separation using canonical correlation. Circuits, Systems, and Signal Processing **37**(2), 658–673 (2018)
18. Lee, K.C., Ho, J., Kriegman, D.J.: Acquiring linear subspaces for face recognition under variable lighting. IEEE Transactions on Pattern Analysis and Machine Intelligence **27**(5), 684–698 (2005)
19. Li, L., Huang, W., Gu, I.Y.H., Tian, Q.: Statistical modeling of complex backgrounds for foreground object detection. IEEE Transactions on Image Processing **13**(11), 1459–1472 (2004)
20. Lin, Z., Chen, M., Ma, Y.: The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices. arXiv preprint arXiv:1009.5055 (2010)
21. Liu, G., Lin, Z., Yu, Y.: Robust subspace segmentation by low-rank representation. In: Proceedings of the 27th International Conference on Machine Learning (ICML-10), pp. 663–670 (2010)
22. Merhav, N., Kresch, R.: Approximate convolution using DCT coefficient multipliers. IEEE Transactions on Circuits and Systems for Video Technology **8**(4), 378–385 (1998)
23. Otazo, R., Candès, E., Sodickson, D.K.: Low-rank plus sparse matrix decomposition for accelerated dynamic mri with separation of background and dynamic components. Magnetic Resonance in Medicine **73**(3), 1125–1136 (2015)

24. Peng, Y., Ganesh, A., Wright, J., Xu, W., Ma, Y.: RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated images. IEEE Transactions on Pattern Analysis and Machine Intelligence **34**(11), 2233–2246 (2012)
25. Shen, Y., Wen, Z., Zhang, Y.: Augmented lagrangian alternating direction method for matrix separation based on low-rank factorization. Optimization Methods and Software **29**(2), 239–263 (2014)
26. Toh, Kim-Chuan, Yun, Sangwoon: An accelerated proximal gradient algorithm for nuclear norm regularized least squares problems. Pacific Journal of Optimization **6**(3), 615–640 (2010)
27. Vincent, E., Gribonval, R., Févotte, C.: Performance measurement in blind audio source separation. IEEE Transactions on Audio, Speech, and Language Processing **14**(4), 1462–1469 (2006)
28. Werner, R., Wilmsy, M., Cheng, B., Forkert, N.D.: Beyond cost function masking: RPCA-based nonlinear registration in the context of vlsm. In: International Workshop on Pattern Recognition in Neuroimaging (PRNI), pp. 1–4. IEEE (2016)
29. Wright, J., Ganesh, A., Rao, S., Peng, Y., Ma, Y.: Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization. In: Advances in Neural Information Processing Systems, pp. 2080–2088 (2009)
30. Wright, J., Ma, Y., Mairal, J., Sapiro, G., Huang, T.S., Yan, S.: Sparse representation for computer vision and pattern recognition. Proceedings of the IEEE **98**(6), 1031–1044 (2010)
31. Xue, Z., Dong, J., Zhao, Y., Liu, C., Chellali, R.: Low-rank and sparse matrix decomposition via the truncated nuclear norm and a sparse regularizer. The Visual Computer pp. 1–18 (2018)
32. Yuan, X., Yang, J.: Sparse and low rank matrix decomposition via alternating direction method. Pacific Journal of Optimization **9**(1), 1–11 (2009)