IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 5, 2016, accepted December 22, 2016, date of publication January 24, 2017, date of current version March 13, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2657645

Dynamic Magnetic Resonance Imaging via Nonconvex Low-Rank Matrix Approximation

FEI XU[1,2], JINGQI HAN[3], YONGLI WANG[4], MING CHEN[4], YONGYONG CHEN[4], GUOPING HE[5], AND YUNHONG HU[6]

[1]College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China
[2]College of Mathematics and Physics, Qingdao University of Science and Technology, Qingdao 266061, China
[3]Department of Radiology, Traditional Chinese Medicine Hospital of Huangdao District of Qingdao City, Qingdao 266500, China
[4]College of Mathematics and Systems Science, Shandong University of Science and Technology, Qingdao 266590, China
[5]Shandong Academy of Sciences, Jinan 250014, China
[6]Department of Applied Mathematics, Yuncheng University, Yuncheng 044000, China

Corresponding author: Y. Wang (wangyongli@sdkd.net.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 11241005, in part by the Shanxi Scholarship Council of China under Grant 2015-093, in part by the Natural Science Foundation of Shandong Province under Grant ZR2015FM013, in part by the Scientific and Technological Program of Huangdao District under Grant 2014-1-28, and in part by the Science and Technology Program of Qingdao, China, under Grant No. 14-2-4- 56-jch.

ABSTRACT Reconstruction of highly accelerated dynamic magnetic resonance imaging (MRI) is of crucial importance for the medical diagnosis. The application of general robust principal component analysis (RPCA) to MRI can increase imaging speed and efficiency. However, conventional RPCA makes use of nuclear norm as convex surrogate of the rank function, whose drawbacks have been mentioned in plenty of literature. Recently, nonconvex surrogates of the rank function in RPCA have been widely investigated and proved to be tighter rank approximation than nuclear norm by the massive experimental results. Motivated by this, we propose a nonconvex alternating direction method based on nonconvex rank approximation to reconstruct dynamic MRI data from undersampled $k - t$ space data. We solve the associated nonconvex model by the alternating direction method and difference of convex programming. The convergence analysis provided guarantees the effectiveness of our algorithm. Experimental results on cardiac perfusion and cardiac cine MRI data demonstrate that our method outperforms the state-of-the-art MRI reconstruction methods in both image clarity and computation efficiency.

INDEX TERMS Magnetic resonance imaging, nonconvex, low-rank matrix approximation.

I. INTRODUCTION

Dynamic magnetic resonance imaging (MRI) is extremely important for medical research and diagnosis. However, MRI usually requires longer scanning time than patients can endure. For example, it often costs 40-60s in cardiac perfusion imaging [1]. Hence, MR perfusion imaging evaluation is inevitably disturbed by the inconsistent myocardial motion mainly caused by the breath of the examined patient. Specially, the long scanning time has become a bottleneck to restrict the clinical application of MRI. Therefore, how to achieve clear images through fast reconstruction is an urgent problem to be solved.

With the help of the theory of compressed sensing (CS) [2] and breaking through the traditional sampling limitation, researchers utilize the undersampled data instead of fully-sampled data to reconstruct the MR images fast and efficiently. In the context of dynamic MRI, RPCA decomposes each image into a low-rank matrix plus a sparse matrix since a series of images have spatiotemporal correlations which necessarily lead to producing a low-rank matrix corresponding to the background of images. On the other hand, a sparse matrix can depict the pathological lesion on the top of the background [3]. The existing MRI reconstruction methods exactly use the above-mentioned characteristic, consequently, a variety of optimization algorithms based on RPCA have been proposed for improving the reconstruction speed, see [1], [3]–[5]. Otazo *et al.* [3] utilized the undersampled dynamic MRI to formulate the multicoil low-rank plus sparse reconstruction by a convex optimization algorithm, where the nuclear-norm and l_1-norm were used as the convex surrogate functions for the rank function and l_0-norm, respectively. Lustig *et al.* [4] exploited the sparsity property

implying in MR images to reduce the scanning time and improve the reconstruction of MRI. Lingala *et al.* [5] proposed a variable splitting algorithm to reconstruct dynamic MRI data from under-sampled k-t space data. Gao *et al.* [6] presented a new spatiotemporal model named the robust PCA-based 4D computed tomography model, applicable in some imaging problems including cardiac MRI. Zonoobi *et al.* [7] proposed a re-formulation of the L and S decomposition to take some priori information into account and used a soft-thresholding based algorithm to reconstruct a time sequence of 3D cardiac MRI.

As we mentioned above, the state-of-the-art RPCA applied in MRI utilizes the dynamic MRI data to formulate low-rank plus sparse reconstruction and most literature considers convex surrogate of the rank function, i.e., nuclear norm. However, this kind of convex formulation has inevitably drawbacks despite their success in the previous applications. We summarize them as follows:

a) The practical data matrix may have not incoherence guarantee, therefore, under this circumstance, the optimal solution of the traditional RPCA convex model may deviate seriously from the truth [8].

b) The nuclear norm (which is the sum of singular values of the matrix in question) requires a computationally expensive singular value decomposition (SVD) of matrix at each iteration. Moreover, it is unreasonable to replace the rank function with the nuclear norm since this kind of simple summation means that all the singular values are dealt with equally. As a matter of fact, the larger singular value should be less penalized otherwise the nuclear norm may be decided by a few large singular values.

c) The matter that nuclear norm cannot approximate the rank function well may derive from the following problem in CS. Conventional CS recovers original vector-form data by replacing l_0-norm with l_1-norm. However, unfortunately, it has been shown that the l_1-norm is just a loose approximation of the l_0-norm even under certain incoherent conditions [9], what is more, the approximation error can not be neglected in most cases [10]. For a vector, l_0-norm is the number of nonzero elements, and l_1-norm is the sum of absolute values of all the elements. In fact, if we rearrange all the singular values of a matrix into a vector, then the nuclear norm of this matrix can be viewed as the l_1-norm of the above vector [7], [8], while the rank of this matrix can be viewed as its l_0-norm. For the reason that l_1-norm is not a satisfactory approximation of l_0-norm, we are aware that the nuclear norm may be not an appropriate approximation of the rank in the RPCA problems.

Therefore, it is necessary to look for more accurate surrogates of the rank function and to design related algorithms for solving the RPCA problems in the reconstruction of MRI. To be exciting, some researchers have considered the drawbacks of convex approximation and have been devoted to looking for new surrogates of the rank function. Among them, Xiu *et al.* [1] utilized the rank-one and transformed sparse decomposition for dynamic cardiac MRI, where the rank-one matrix corresponded to the background of an MR image. However, it needed the priori knowledge that the rank of low-rank matrix was one. Hu *et al.* [11] obtained a better rank approximation by a truncated nuclear norm where the largest few singular values were removed. One other idea is motivated by the recent success in using nonconvex surrogates of l_0-norm [12]–[20]. For example, Sun *et al.* [10] proposed a nonconvex model of RPCA by using the capped trace norm and the capped l_1-norm. Kang *et al.* [21] exploited a log-determinant function as a tighter but nonconvex approximation of the rank function. Similar to the Geman function [19], Kang *et al.* defined a new norm called γ-norm in [8] to approximate the rank function, and experimental results showed that it was a tighter approximation than the nuclear norm. Xie *et al.* [22] introduced a nonconvex regularizer named weighted Schatten p-norm to replace the rank function and achieved better approximation. Lu *et al.* [12] used a family of nonconvex surrogates of l_0-norm on the singular values of the matrix to approximate the rank function for solving nonconvex low rank minimization problem. Yu *et al.* [23] derived a general thresholding representation in the form of a recursive function for general l_p regularization problem and obtained the related filtering algorithm applicable to medical image reconstruction.

In this paper, we propose a novel method to speed up the reconstruction of MRI by using two kinds of nonconvex surrogate functions of l_0-norm as rank approximation. The main contributions of this paper are as follows:

1) The algorithms proposed in this paper are based on the nonconvex approximation of the rank function. Then it is used in RPCA to reconstruct the dynamic MRI data, which is naturally of low-rank and sparse structure.
2) We utilize the alternating direction method (ADM) and difference of convex(DC) programming to solve the resultant optimization problem, in which the objective function is a combination of concave and convex functions.
3) We propose two nonconvex algorithms based on the alternating direction method by using two kinds of nonconvex surrogate functions of l_0-norm for rank approximation, and prove the convergence of our algorithms. In addition, experimental results on cardiac perfusion and cardiac cine MRI data demonstrate the validity and efficiency of our methods.

The remainder of this paper is organized as follows. Section II provides a brief review of RPCA and the popular nonconvex surrogates of l_0-norm. Section III presents the nonconvex approximation of rank function and proposes our algorithms to solve the associated models. Section IV gives convergence analysis. Section V describes the experimental results on cardiac perfusion and cardiac cine MRI data. Section VI summarizes this paper.

II. REVIEW OF RPCA AND NONCONVEX SURROGATES

In this section, we present a brief review of RPCA and the nonconvex surrogate functions.

A. REVIEW OF RPCA

The general form of RPCA problem is as follows [24]:

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

where $L, S, D \in \mathbf{R}^{m\times n}$, L and S are unknown, but L is known to be a low-rank matrix, S is known to be a sparse matrix, D is the observed data matrix, and λ is the regularization parameter. The ideal RPCA problem is to recover L by (1) for appropriate λ.

Unfortunately, (1) is a highly nonconvex optimization problem, and there is no efficient solution to be known [25]. A general method to solve this problem is to replace the objective function with a convex function. By replacing the rank function with the nuclear norm and the l_0-norm with the l_1-norm, we can obtain the convex surrogate reformulation of (1):

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

where the nuclear norm of L is defined as $\|L\|_* = \Sigma_i\sigma_i(L)$, i.e., the sum of all its singular values, and $\|S\|_1 = \Sigma_{ij}|S_{ij}|$. Since $\|L\|_* + \lambda\|S\|_1$ is the convex envelope of $rank(L) + \lambda\|S\|_0$, the above reformulation (2) is a convex approximation of (1).

Although the relaxed convex model (2) based on nuclear norm has been widely used and achieved success, there still exist two intrinsic drawbacks which we have mentioned in section I, b) and c). This motivates us to find the tighter approximation of the rank function.

There are many popular nonconvex surrogate functions of l_0-norm, which can be used for the nonconvex surrogates of the rank function since the rank may be viewed as the number of nonzero singular values. We give a brief review of nonconvex surrogates in the following subsection.

B. REVIEW OF NONCONVEX SURROGATE FUNCTIONS

The popular nonconvex surrogate functions of l_0-norm include capped l_1 [17], Geman [19], Laplace [20] and so on. To obtain a comprehensive summary, see [12]. Here, we only list the forementioned ones in Table 1.

TABLE 1. Nonconvex surrogate functions of $\|\sigma\|_0$.

Name	Function ($\sigma \geq 0, \ \lambda > 0$)
Capped L_1	$f(\sigma) = \begin{cases} \lambda\sigma & \sigma < \gamma \\ \lambda\gamma & \sigma \geq \gamma \end{cases}$
Geman	$f(\sigma) = \frac{\lambda\sigma}{\sigma+\gamma}$
Laplace	$f(\sigma) = \lambda(1 - exp(-\frac{\sigma}{\gamma}))$

Numerical experiments in [26] demonstrate that the nonconvex optimization often outperforms its convex counterpart in the application areas including image processing.

III. MODELS AND ALGORITHMS

In this section, we construct an algorithm structure in which different nonconvex surrogates can be used for approximating the rank function. We call our algorithm nonconvex alternating direction method, abbreviated as NADM, which uses the nonconvex surrogate of the rank function and exploits the alternating direction method framework.

A. MODELS

The low-rank and sparse decomposition model given in [3] to reconstruct undersampled dynamic MRI is as follows:

$$\min_{L,S} \ \|L\|_* + \lambda\|TS\|_1 \quad \text{s.t. } E(L + S) = d, \tag{3}$$

where $L, S \in \mathbf{C}^{m\times n}$, T is a sparsifying transform for S, E is the encoding or acquisition operator and d is the undersampled k-t data.

To solve problem (3), firstly, we introduce two definitions. One is defined in [8] called γ-norm of L therein. The other is also some kind of γ-norm. These two definitions are based on the last two functions in Table 1. In this paper, we name them Geman-norm and Laplace-norm, respectively. We will replace the nuclear norm with these two norms in (3).

Definition 1 (Geman-Norm [8], [19]): Suppose that L is an $m \times n$ matrix, γ is a positive parameter, define

$$\|L\|_\gamma = \sum_i \frac{(1+\gamma)\sigma_i(L)}{\gamma + \sigma_i(L)}, \quad \gamma > 0, \tag{4}$$

where $\sigma_i(L)$ denotes the i-th singular value of matrix L.

Definition 2 (Laplace-Norm): Suppose that L is an $m \times n$ matrix, γ is a positive parameter, define

$$\|L\|_\gamma = \sum_i (1 - e^{-\sigma_i(L)/\gamma}), \quad \gamma > 0, \tag{5}$$

where $\sigma_i(L)$ denotes the i-th singular value of matrix L.

Although neither of them is real norm, they still have satisfying properties. We summarize them in the following proposition:

Proposition 1:

1) $\lim_{\gamma\to 0} \|L\|_\gamma = rank(L)$.
2) $\|L\|_\gamma$ *is unitarily invariant, i.e.,* $\|L\|_\gamma = \|ULV\|_\gamma$ *for any orthonormal* $U \in \mathbf{C}^{m\times m}$ *and* $V \in \mathbf{C}^{n\times n}$.
3) *positive definiteness:* $\|L\|_\gamma \geq 0$ *for any* $L \in \mathbf{C}^{m\times n}$ *and* $\|L\|_\gamma = 0$ *if and only if* $L = 0$.

Proof:

1) For $f(\sigma_i(L)) = \frac{(1+\gamma)\sigma_i(L)}{\gamma+\sigma_i(L)}$ and $f(\sigma_i(L)) = 1 - e^{\frac{-\sigma_i(L)}{\gamma}}$, we have $\lim_{\gamma\to 0} f(\sigma_i) = \begin{cases} 0 & \sigma_i(L) = 0 \\ 1 & \sigma_i(L) \neq 0 \end{cases}$ which consequently deduces the conclusion.
2) Suppose that $(\cdot)^H$, $(\cdot)^{-1}$and λ denote the transpose, inverse and eigenvalue of matrix, respectively. E is a unit matrix. Then the eigenpolynomial $|(ULV)^H ULV - \lambda E| = |V^H L^H U^H ULV - \lambda E| = |V^H L^H LV - \lambda E| = |V^{-1} L^H LV - \lambda E| = |L^H L - \lambda E|$. This means that matrices $(ULV)^H ULV$ and $L^H L$ have the same eigenvalues.

Consequently, ULV and L have the same singular values. Hence, the conclusion is true.

3) The conclusion exactly holds for the reason that $f(\sigma_i(L))$ takes the values of 0 if and only if $L = 0$, alternatively, 1 when $L \neq 0$. □

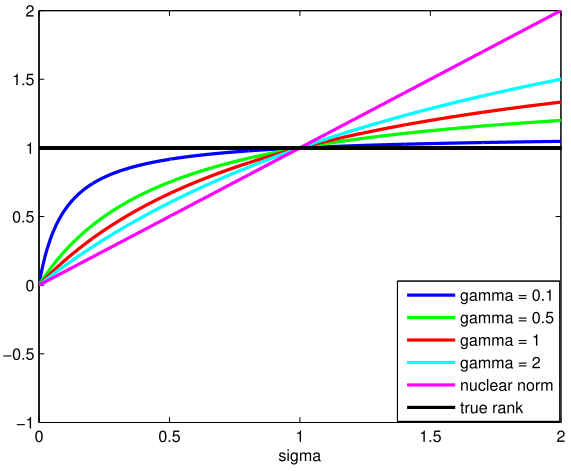

FIGURE 1. The changing curves of Geman-norm for different values of γ. Comparison with nuclear norm and the true rank respecting to an increasing singular value $\sigma_i(L)$.

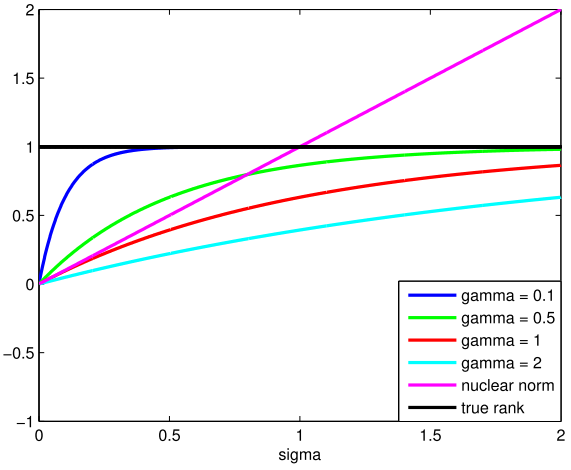

FIGURE 2. The changing curves of Laplace-norm for different values of γ. Comparison with nuclear norm and the true rank respecting to an increasing singular value $\sigma_i(L)$.

Fig. 1 and Fig. 2 show that the smaller the γ is, the closer the Geman and Laplace norm approach the true rank. This phenomenon coincides with the first property of γ-norm in Proposition 1. We can also observe that the nuclear norm deviates seriously from the true rank when the singular value is far beyond one.

Replacing the nuclear norm $\|L\|_*$ in problem (3) with the well-defined γ-norms $\|L\|_\gamma$, we can obtain the following model:

$$\min_{L,S} \|L\|_\gamma + \lambda\|TS\|_1 \quad \text{s.t. } E(L+S) = d, \tag{6}$$

Actually, the real MRI data used in our paper are three-dimensional. Here, we transform them into matrices considering the spatio-temporal correlations, in other words, we unfold the 3-way tensor by mode-3, i.e., along the time dimension. Therefore, similar to (3), L and S in (6) are also matrices. Next, our work is to design algorithms for solving problem (6). Then, in the following subsection, we will discuss the properties of the algorithms involved.

B. ALGORITHMS

In this subsection, we present the details of solving the problem (6). Additionally, different γ-norms can be used in our algorithm according to definition 1 and 2. We distinguish them by G-NADM and L-NADM, respectively.

1) NONCONVEX ALTERNATING DIRECTION METHOD (NADM)

a: ALTERNATING DIRECTION METHOD FRAMEWORK

By using regularization rather than strict constraints, the unconstrained formulation of (6) is as follows:

$$\min_{L,S} \frac{1}{2}\|E(L+S) - d\|_2^2 + \lambda_1\|L\|_\gamma + \lambda_2\|TS\|_1 \tag{7}$$

where the parameters λ_1 and λ_2 are used to balance the data consistency versus the complexity of the solution.

This unconstrained optimization problem can be solved by alternating direction method which is very efficient and extensible for many large-scale programming problems arising in machine learning.

The alternating direction method for (7) takes the form:

$$L_{k+1} = \arg\min_L \left\{ \frac{1}{2}\|E(L+S_k) - d\|_2^2 + \lambda_1\|L\|_\gamma \right\} \tag{8}$$

$$S_{k+1} = \arg\min_S \left\{ \frac{1}{2}\|E(L_{k+1}+S) - d\|_2^2 + \lambda_2\|TS\|_1 \right\} \tag{9}$$

In the following, we will exploit some strategies to solve the subproblems (8) and (9).

b: DIFFERENCE OF CONVEX PROGRAMMING TO UPDATE L

Note that (8) is a combination of convex and nonconvex functions with the first term being convex and the second term being nonconvex. Hence, we resort to the difference of convex(DC) programming. Similar to [27], we have the following lemma.

Lemma 1: The subdifferential of $\|L\|_\gamma$ is given by

$$\partial\|L\|_\gamma = \{Udiag(l)V^T : l_i = \frac{df(\sigma_i)}{d(\sigma_i)}, i = 1, 2, \ldots, rank(L)\} \tag{10}$$

where the columns of U and V are the left and right singular vectors of L, respectively, and $l_i = \frac{\gamma(1+\gamma)}{(\sigma_i(L)+\gamma)^2}$ as in definition 1, or $l_i = \frac{e^{-\sigma_i(L)/\gamma}}{\gamma}$ as in definition 2.

By Lemma 1, the $(k+1)$-th iteration of L in (8) can be reformed as

$$L_{k+1} = \arg\min_L \left\{ \frac{1}{2}\|E(L+S_k) - d\|_2^2 + \lambda_1\langle\partial\|L_k\|_\gamma, L\rangle \right\} \tag{11}$$

Let the derivative of its objective function with respect to L equal zero, we can derive:

$$L_{k+1} = (E^H E)^{-1}(E^H d - E^H E S_k - \lambda_1 \partial \|L_k\|_\gamma) \quad (12)$$

c: DATA CONSISTENCY STRATEGY [3]

We can solve the optimization problem (9) by using the iterative soft-thresholding of TS. Firstly, we introduce the definition of singular value thresholding (SVT) operator.

Definition 3 (SVT Operator):

$$SVT_\lambda(A) = U \Lambda_\lambda(A) V^H \quad (13)$$

where $A = U\Sigma V^H$ is any singular value decomposition of A, and $\Lambda_\lambda(A)$ denotes applying the following shrinkage operator $\Lambda_\lambda(\cdot)$ for each element of A, i.e.,

$$\Lambda_\lambda(x) = \frac{x}{|x|} \max(|x| - \lambda, 0) \quad (14)$$

where $x \in \mathbf{C}$, $\lambda \in \mathbf{R}$.

With this, the $(k+1)$-th iteration of S in (9) can be reformed as

$$S_{k+1} = T^{-1}(\Lambda_{\lambda_2}(T(M_k - L_k))) \quad (15)$$

where $M_0 = E^H d$, and the matrix M_k is used to maintain data consistency. By applying the forementioned shrinkage operator to $M_k - L_k$, we obtain M_{k+1} as follows:

$$M_{k+1} = L_{k+1} + S_{k+1} - E^H(E(L_{k+1} + S_{k+1}) - d) \quad (16)$$

In (16), the aliasing artifacts corresponding to the residual $E^H(E(L_{k+1} + S_{k+1}) - d)$ are subtracted from $L_{k+1} + S_{k+1}$.

At last, we summarize all the details in Algorithm 1.

Algorithm 1 NADM: Alternating Direction Method for Non-Convex RPCA

Input: d : undersampled k-t data
E : encoding operator
T : sparsifying transform
λ_1: singular-value threshold
λ_2: sparsity threshold

Initialize: $L_0 = M_0 = E^H d, \quad S_0 = 0, \quad k = 0$

while not converge **do**
Compute the gradient $\partial \|L_k\|_\gamma$ with (10);
Update L_{k+1} with (12);
Update S_{k+1} with (15);
Update M_{k+1} with (16).
end while

Output: L_{k+1}, S_{k+1}.

IV. CONVERGENCE ANALYSIS

In this section, we give the convergence analysis for Algorithm 1. First, we prove the following lemma:

Lemma 2: Let $\{L_k\}$and$\{S_k\}$ be the sequences generated by Algorithm 1, then $\{L_k\}$and$\{S_k\}$ are bounded.

Proof: First, we rewrite the objective function of (7) as follows:

$$W(L, S) = \frac{1}{2}\|E(L + S) - d\|_2^2 + \lambda_1 \|L\|_\gamma + \lambda_2 \|TS\|_1 \quad (17)$$

We have the following conclusions from (8) and (9):

$$W(L_{k+1}, S_k) \le W(L_k, S_k)$$
$$W(L_{k+1}, S_{k+1}) \le W(L_{k+1}, S_k)$$

then

$$W(L_{k+1}, S_{k+1}) \le W(L_k, S_k) \quad (18)$$

therefore, the sequence $\{W(L_k, S_k)\}$ is nonincreasing.

Next, we will show that the sequences $\{L_k\}$ and $\{S_k\}$ are bounded. We first prove the boundedness of $\{L_k\}$.

Assume to the contrary, suppose $\{L_k\}$ to be unbounded, then for an arbitrary positive real number M, there always exists a positive integer N, such that for $k \ge N$ and any norm $\|\cdot\|$, we have $\|L_k\| \ge M$, i.e., $\|L_k\| \to \infty$. Hence, the first term of (17), $\frac{1}{2}\|E(L_k + S_k) - d\|_2^2 \to \infty$.

Since $\|L\|_\gamma$ is a pseudo norm, we discuss it in two cases:

i) If $\sigma_i(L_k)$ is unbounded when $\{L_k\}$ is unbounded, then from definition 1 and definition 2 respectively, we have

$$\lim_{\sigma_i(L_k)\to\infty} \frac{(1+\gamma)\sigma_i(L_k)}{\gamma + \sigma_i(L_k)} = 1 + \gamma,$$
$$\lim_{\sigma_i(L_k)\to\infty} 1 - e^{-\frac{\sigma_i(L_k)}{\gamma}} = 1.$$

Note that γ takes the values approaching to 0, it follows that $\|L\|_\gamma$ is bounded.

ii) If $\sigma_i(L_k)$ is bounded when $\{L_k\}$ is unbounded, then obviously, $\|L\|_\gamma$ is also bounded from definition 1 and definition 2.

Hence, $\|L\|_\gamma$ is always bounded.

Similar to the proof of $\{L_k\}$, suppose $\{S_k\}$ to be unbounded, then the first and the third term of (17), i.e., $\frac{1}{2}\|E(L_k + S_k) - d\|_2^2$ and $\|TS\|_1$ are unbounded.

In summary, either $\{L_k\}$ or $\{S_k\}$ is unbounded, then $\{W(L_k, S_k)\}$ is unbounded, together with its nonnegativity, we have $W(L_k, S_k) \to \infty$, which contradicts with the conclusion in (18) that $\{W(L_k, S_k)\}$ is monotonically nonincreasing. Hence, $\{L_k\}$ and $\{S_k\}$ are both bounded. □

Theorem 1: Suppose that the sequence $\{(L_k, S_k)\}$ is generated by Algorithm 1, then any accumulation point of $\{(L_k, S_k)\}$ is a local minimizer of (6).

Proof: By Lemma 2, $\{(L_k, S_k)\}$ are bounded. So $\{(L_k, S_k)\}$ has at least one accumulation point denoted by $\{(L^*, S^*)\}$. Together with the conclusion in (18) that $\{W(L_k, S_k)\}$ is monotonically nonincreasing, then $\lim_{k\to\infty} W(L_k, S_k)$ exists, and

$$\lim_{k\to\infty} W(L_k, S_k) = W(L^*, S^*).$$

Therefore, $\{(L^*, S^*)\}$ is a local minimizer of (6). □

Our experimental results on cardiac perfusion and cardiac cine MRI data confirm the convergence of Algorithm 1.

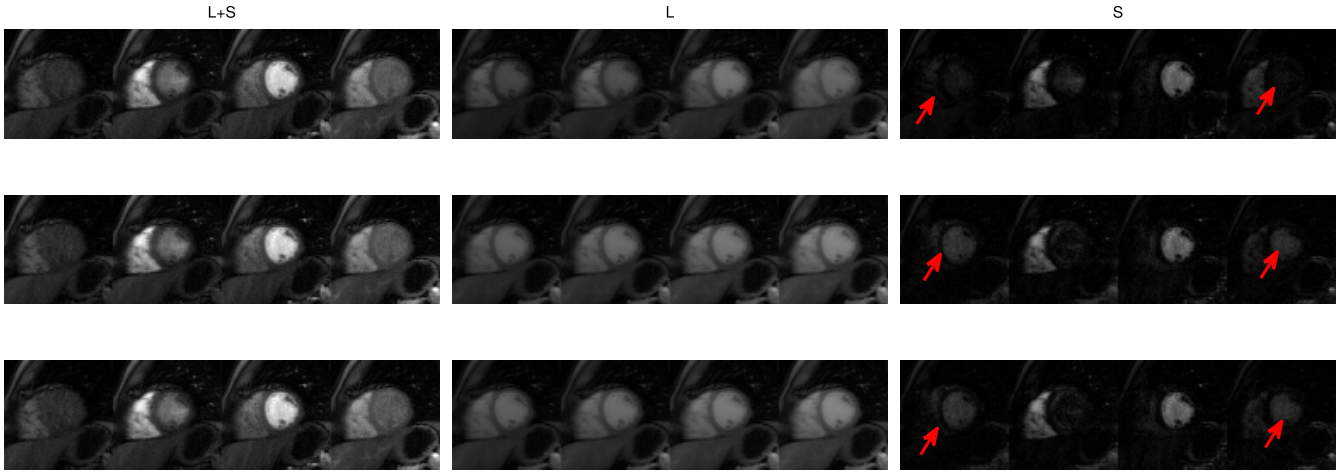

FIGURE 3. Performance of different algorithms for dynamic cardiac perfusion. The three rows from top to bottom correspond to the algorithms: IST, G-NADM, L-NADM; the three columns from left to right correspond to $L + S$, low-rank component L and sparse component S of the observed data D. This figure can be better viewed in zoomed PDF.

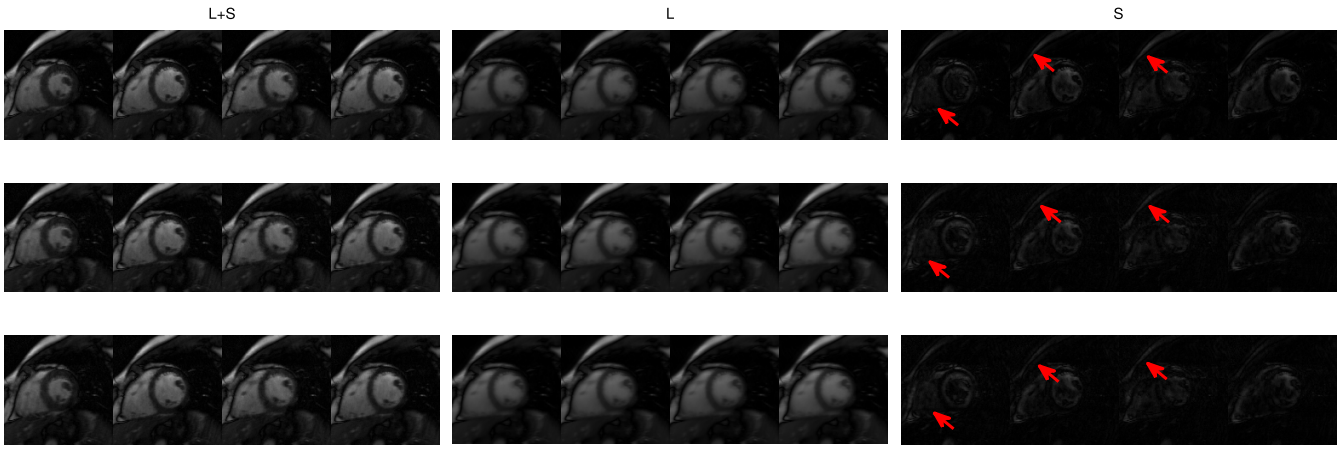

FIGURE 4. Performance of different algorithms for dynamic cardiac cine. The three rows from top to bottom correspond to the algorithms: IST, G-NADM, L-NADM; the three columns from left to right correspond to $L + S$, low-rank component L and sparse component S of the observed data D. This figure can be better viewed in zoomed PDF.

V. NUMERICAL EXPERIMENTS

In this section, we compare our proposed algorithms with IST [3] for dynamic cardiac MRI data. We apply these two algorithms on MR images: cardiac perfusion and cardiac cine, which can be downloaded online at http://cai2r.net/resources/software/ls-reconstruction-matlab-code. All the experiments are run in MATLAB R2013a on a PC with a COREi3 2.40GHz CPU and 4GB memory.

The stopping criterion of our algorithms is the relative error being less than 2.5×10^{-3}, namely,

$$\text{RelErr} = \frac{\|L_{k+1} + S_{k+1} - (L_k + S_k)\|_F}{\|L_k + S_k\|_F} \leq 2.5 \times 10^{-3} \tag{19}$$

We choose $\lambda_1 = 1, \lambda_2 = 0.05, \gamma = 0.005$ on dynamic cardiac perfusion data, and $\lambda_1 = 1, \lambda_2 = 0.0001, \gamma = 0.005$ on dynamic cardiac cine data. Experimental results and the comparison among these three algorithms (IST, G-NADM, L-NADM) are presented in Figs. 3-6, Tables 2 and 3.

Furthermore, we analyze how our algorithms perform with the different values of γ for both of the two data sets. Concretely, we compute the root mean square error (RMSE) defined as:

$$\text{RMSE} = \frac{\|D - L - S\|_F}{\|D\|_F}, \tag{20}$$

where D, L, S represent the observed data, low-rank and sparse matrices, respectively. Take the cardiac perfusion data set as an example, by setting $\gamma = 0.001$, step size $= 0.001$ and stopping when $\gamma = 0.03$, we find that the RMSE takes the value between 0.733 and 0.735 when $\gamma \in [0.005, 0.03]$, and the RMSE takes the value of 1 when $\gamma \in [0.001, 0.004]$, which means that our algorithms fail to recover the low-rank matrix L from the data D. Therefore, smaller γ can better approximate the rank but performs worse in convergence.

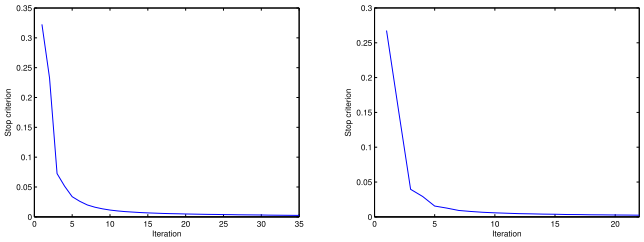

FIGURE 5. Convergence curves with respect to iterations. The two subfigures from left to right are for cardiac perfusion and cardiac cine data sets, respectively. This figure can be better viewed in zoomed PDF.

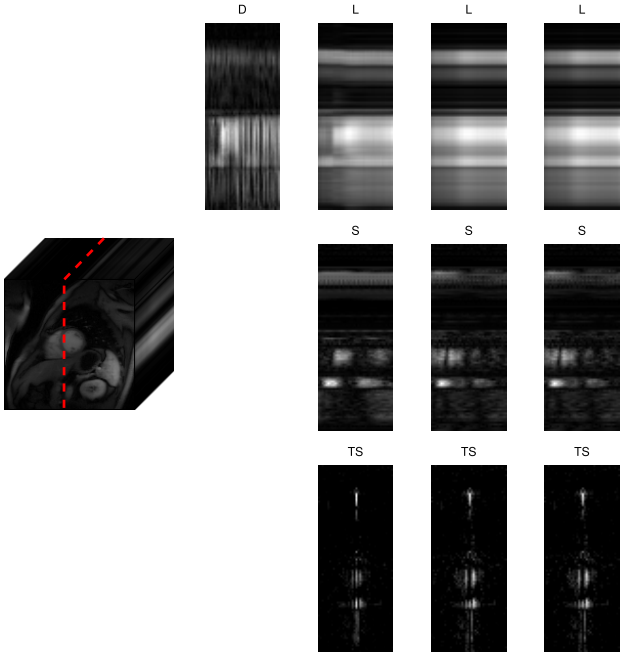

FIGURE 6. $L + S$ decomposition of cardiac perfusion data set corresponding to the 58th. column (signed with the red dashed). The middle column represents the observed data D; the rightmost three columns are the results of decomposition by: IST, G-NADM and L-NADM, respectively; the three rows from top to bottom are L, S and TS, respectively.

TABLE 2. Computational results for dynamic cardiac perfusion.

Algorithm	Time(s)	Iteration	Rank(L)	RMSE	Relative Error
IST	202.4	36	5	0.8044	$2.50 * 10^{-3}$
G-NADM	144.4	35	1	0.7350	$2.49 * 10^{-3}$
L-NADM	144.5	35	1	0.7370	$2.47 * 10^{-3}$

TABLE 3. Computational results for dynamic cardiac cine.

Algorithm	Time(s)	Iteration	Rank(L)	RMSE	Relative Error
IST	243.8	26	1	0.5063	$2.47 * 10^{-3}$
G-NADM	195.6	22	1	0.4829	$2.41 * 10^{-3}$
L-NADM	190.4	22	1	0.4839	$2.42 * 10^{-3}$

The parameter γ with the lowest RMSE and best rank approximation is employed in the reconstruction, here, $\gamma = 0.005$.

A. DYNAMIC CARDIAC PERFUSION

In Fig. 3, the display window is $[0, 1]$, the dynamic range of the images are $128 \times 128 \times 40$; images from top to bottom represent the results of IST, G-NADM and L-NADM, respectively; the first column $L + S$ per row are the superposition of the other two columns L and S. The experimental results demonstrate that the reconstructed images using G-NADM and L-NADM are more clear in both background and foreground (see the red arrows which point to the obvious differences), which means that we can observe cardiac disease easier.

Furthermore, in Table 2, the quantitative indexes of images reconstructed through different method are analyzed. For this end, we compute the RMSE defined as (20). It is clearly seen that our RMSE values are smaller than that of IST, meaning that we have better reconstruction results.

Besides, it is evident from Table 2 that the iteration times and relative errors of G-NADM and L-NADM are less than those of IST, and computation time is much shorter. Additionally, consistent with the expected results, the rank of the low-rank matrix is one. The improvement in the results of experiment is because we use the nonconvex approximation of the rank function, which is closer to the true rank than the nuclear norm. Fig. 5 displays the convergence curves with respect to iterations, our algorithms G-NADM and L-NADM have the similar shape of curve.

Fig. 6 shows $L+S$ decomposition of cardiac perfusion data set corresponding to the central location (see the red dashed sign). L captures the background along the time dimension, S includes the remaining dynamic information and TS shows the increased sparsity by the sparsifying transform T for S. It is clearly seen that the low-rank components L by G-NADM and L-NADM are cleaner.

B. DYNAMIC CARDIAC CINE

In Fig. 4, the display window is $[0, 1]$, the dynamic range of the images are $256 \times 256 \times 24$. Similarly, experimental effects on the dynamic cardiac cine are satisfactory as well. Fig. 4 and Table 3 show that our nonconvex rank surrogate based algorithms (G-NADM, L-NADM) can reconstruct better images than IST on the dynamic cardiac cine data since our nonconvex rank approximation can approach the rank function more closely.

VI. CONCLUSIONS AND FUTURE WORK

In this paper, we propose a nonconvex RPCA formulation (6) to reconstruct the dynamic MRI. Unlike the conventional nuclear norm based methods, our method deduces a combination of convex and nonconvex objective function in the involved problem (7). We make use of the alternating direction method and difference of convex programming to solve the transformed optimization problem. Experimental results on dynamic cardiac MRI data demonstrate the accuracy and efficiency of our algorithms. This verifies the applied ability of nonconvex RPCA for dynamic MRI. Furthermore, as

forementioned, the MRI data used in our paper are three-dimensional. Here, we transform them into matrices through unfolding the 3-way tensor by mode-3. Actually, the matrix unfoldings of a tensor are nearly low-rank [28], no matter by mode-1, mode-2 or mode-3. Therefore, we may also consider the MRI data more from the view of tensor, namely, unfolding them by different modes. This is the future work we will do.

ACKNOWLEDGMENTS

At last, the authors would like to thank anonymous editor and reviewers who give really professional suggestions that have helped to improve the quality of this paper.

REFERENCES

[1] X. Xiu and L. Kong, "Rank-one and transformed sparse decomposition for dynamic cardiac MRI," vol. 2015, Jun. 2015, Art. no. 169317. [Online]. Available: http://dx.doi.org/10.1155/2015/169317

[2] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[3] R. Otazo, E. J. Candès, and D. K. Sodickson, "Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components," *Magn. Reson. Med.*, vol. 73, no. 3, pp. 1125–1136, Mar. 2015.

[4] M. Lustig, D. L. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn. Reson. Med.*, vol. 58, no. 6, pp. 1182–1195, Dec. 2007.

[5] S. G. Lingala, Y. Hu, E. V. R. DiBella, and M. Jacob, "Accelerated dynamic MRI exploiting sparsity and low-rank structure: *k-t* SLR," *IEEE Trans. Med. Imag.*, vol. 30, no. 5, pp. 1042–1054, May 2011.

[6] H. Gao, J.-F. Cai, Z. Shen, and H. Zhao, "Robust principal component analysis-based four-dimensional computed tomography," *Phys. Med. Biol.*, vol. 56, no. 11, pp. 3181–3198, May 2011.

[7] D. Zonoobi, S. F. Roohi, and A. A. Kassim. (2014). "Low-rank and sparse matrix decomposition with a-priori knowledge for dynamic 3D MRI reconstruction." [Online]. Available: https://arxiv.org/abs/1411.6206

[8] Z. Kang, C. Peng, and Q. Cheng, "Robust PCA via nonconvex rank approximation," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Atlantic City, NJ, USA, Nov. 2015, pp. 211–220.

[9] D. L. Donoho, "For most large underdetermined systems of linear equations the minimal ℓ^1-norm solution is also the sparsest solution," *Commun. Pure Appl. Math.*, vol. 59, no. 6, pp. 797–829, 2006.

[10] Q. Sun, S. Xiang, and J. Ye, "Robust principal component analysis via capped norms," in *Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, Chicago, IL, USA, Aug. 2013, pp. 311–319.

[11] Y. Hu, D. Zhang, J. Ye, X. Li, and X. He, "Fast and accurate matrix completion via truncated nuclear norm regularization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 9, pp. 2117–2130, Sep. 2013.

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

[13] L. E. Frank and J. H. Friedman, "A statistical view of some chemometrics regression tools," *Technometrics*, vol. 35, no. 2, pp. 109–135, 1993.

[14] J. Fan and R. Li, "Variable selection via nonconcave penalized likelihood and its oracle properties," *J. Amer. Statist. Assoc.*, vol. 96, no. 456, pp. 1348–1360, 2001.

[15] J. H. Friedman, "Fast sparse regression and classification," *Int. J. Forecasting*, vol. 28, no. 3, pp. 722–738, Jul. 2012.

[16] C.-H. Zhang, "Nearly unbiased variable selection under minimax concave penalty," *Ann. Statist.*, vol. 38, no. 2, pp. 894–942, 2010.

[17] T. Zhang, "Analysis of multi-stage convex relaxation for sparse regularization," *J. Mach. Learn. Res.*, vol. 11, pp. 1081–1107, Mar. 2010.

[18] C. Gao, N. Wang, Q. Yu, and Z. Zhang, "A feasible nonconvex relaxation approach to feature selection," in *Proc. 25th AAAI Conf. Artif. Intell. (AAAI)*, San Francisco, CA, USA, Aug. 2011, pp. 356–361.

[19] D. Geman and C. Yang, "Nonlinear image recovery with half-quadratic regularization," *IEEE Trans. Signal Process.*, vol. 4, no. 7, pp. 932–946, Jul. 1995.

[20] J. Trzasko and A. Manduca, "Highly undersampled magnetic resonance image reconstruction via homotopic ℓ_0-minimization," *IEEE Trans. Med. Imag.*, vol. 28, no. 1, pp. 106–121, Jan. 2009.

[21] Z. Kang, C. Peng, J. Cheng, and Q. Cheng, "Logdet rank minimization with application to subspace clustering," *Comput. Intell. Neurosci.*, vol. 2015, Jan. 2015, Art. no. 68.

[22] Y. Xie, Y. Qu, D. Tao, W. Wu, Q. Yuan, and W. Zhang, "Hyperspectral image restoration via iteratively regularized weighted schatten *p*-norm minimization," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4642–4659, Aug. 2016.

[23] H. Yu and C. Miao, "General thresholding representation for ℓ_p $(0 < p < 1)$ regularization," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5455–5468, Dec. 2015.

[24] J. Wright, A. Ganesh, S. Rao, Y. Peng, and Y. Ma, "Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization," in *Proc. Adv. Neural Inf. Process. Syst.*, Vancouver, BC, Canada, Dec. 2009, pp. 2080–2088.

[25] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004, p. 11.

[26] E. J. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *J. Fourier Anal. Appl.*, vol. 14, nos. 5–6, pp. 877–905, 2008.

[27] 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.

[28] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 208–220, Jan. 2013.

FEI XU received the B.S. degree in mathematics education from Qingdao University, Qingdao, China, in 2001, and the M.S. degree in operations research and control theory from the Dalian University of Technology, Dalian, China, in 2004. She is currently pursuing the Ph.D. degree with the College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao.

She is currently a Lecturer with the College of Mathematics and Physics, Qingdao University of Science and Technology. Her research interests include medical image processing and image reconstruction.

JINGQI HAN received the B.S. degree in medical image from Mudanjiang Medical University, Mudanjiang, China, in 1999, and the M.S. degree in imaging and nuclear medicine from Qingdao University, Qingdao, China, in 2015.

In 1999, he joined the Department of Radiology of Traditional Chinese Medicine of hospital of Huangdao District, where he is currently an Associate Chief Physician. His research interests focus on medical imaging diagnosis and interventional therapy.

YONGLI WANG received the B.S. degree in operations research and control theory from the Shandong University of Science and Technology, Qingdao, China, in 2001 and the Ph.D. degree in applied mathematics from Shanghai Jiao Tong University, Shanghai, China, in 2006.

Her current research interests include optimization theory and method, low-rank representation, image processing, and distributed computation.

MING CHEN received the B.S. degree in mathematics education from Liaocheng University, Liaocheng, China, in 2002, and the M.S. degree and the Ph.D. degree in applied mathematics from Capital Normal University, Beijing, China, in 2005 and 2008, respectively.

In 2008, she joined the College of Mathematics and System Science, Shandong University of Science and Technology, where she is currently an Associate Professor. Her research interests focus on image reconstruction and image processing.

YONGYONG CHEN received the B.S. degree in information and computing sciences from the Shandong University of Science and Technology, Qingdao, China, in 2014, where he is currently pursuing the M.S. degree with the College of Mathematics and Systems Science.

He is currently visiting the National Key Laboratory for Novel Software Technology, Nanjing University, as an Exchange Student. His research interests include (non-convex) low-rank and sparse matrix/tensor decomposition models, with applications to image processing, MR imaging, and computer vision.

GUOPING HE received the B.S. degree in computational mathematics from the Shandong University of Science and Technology, China, in 1982, and the M.S. degree and the Ph.D. degree in operations research and control theory from the Chinese Academy of Sciences, Beijing, China, in 1988 and 1995, respectively.

He is currently the Vice-President with the Shandong Academy of Sciences, Jinan, China. He has over 100 research papers and is the holder of eight foundations. His research interests include nonlinear optimization theory, numerical computation, and data mining.

YUNHONG HU received the B.S. degree in applied mathematics from Shanxi University, Taiyuan, China, in 1996, and the Ph.D. degree in computer software and theory from the Shandong University of Science and Technology, China, in 2011.

She is currently an Associate Professor with the Department of Applied Mathematics, Yuncheng University. Her research interests include data mining, machine learning, and optimization algorithms.