Progress In Electromagnetics Research C, Vol. 87, 241–253, 2018

A New Non-Convex Regularized Sparse Reconstruction Algorithm for Compressed Sensing Magnetic Resonance Image Recovery

Xiangjun Yin, Linyu Wang, Huihui Yue*, and Jianhong Xiang

Abstract—Compressed sensing (CS) relies on the sparse priorin posed on the signal to solve the ill-posed recovery problem in an under-determined linear system (ULS). Motivated by the theory, this paper proposes a new algorithm called regularized re-weighted inverse trigonometric smoothed function approximating L_0-norm minimization (RRITSL0) algorithm, where the inverse trigonometric (IT) function, iteratively re-weighted scheme and regularization mechanism constitute the core of the proposed RRITSL0 algorithm. Compared with other state-of-the-art functions, our proposed IT function cluster can better approximate the L_0-norm, thus improving the reconstruction accuracy. And the new re-weighted scheme we adopted can promote sparsity and speed up convergence. Moreover, the regularization mechanism makes the RRITSL0 algorithm more robust against noise. The performance of the proposed algorithm is verified via numerical experiments with additive noise. Furthermore, the experiments prove the superiority of the RRITSL0 algorithm in magnetic resonance (MR) image recovery.

1. INTRODUCTION

MR image recovery [1, 2] plays an essential role in clinical diagnosis. However, at present, the quality of MR image recovery needs to be improved. Fortunately, CS [3, 4], as a new sampling technology, was introduced to MR imaging to significantly improve image recovery accuracy. The CS first acquires very few k-space data (also known as Fourier coefficients) to shorten the image recovery time, and then it reconstructs the MR image from the undersampled data. Image sparsity is assumed to make it possible that we can recover the underlying image from only a few Fourier coefficients. Therefore, we need to find an image that is sparse in a transform domain to fit the undersampled k-space data. MR images can be expressed sparsely by function transformation, such as DCT, Fourier, Wavelet, Curvelet, and Gabor [5], which makes it possible for CS to apply to MR image recovery. Fig. 1 shows the framework of CS model in noiseless case, if we consider the case of noise, then the CS model for MR image recovery can be written as,

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b} \tag{1}$$

where $\mathbf{y} \in \mathbb{R}^m$ is the compressed signal (or image), and $\mathbf{x} \in \mathbb{R}^n$ is the original signal (or image), $m \ll n$. $\mathbf{A} = \mathbf{\Phi\Psi} \in \mathbb{R}^{m\times n}$ is a sensing matrix, where $\mathbf{\Phi} \in \mathbb{R}^{m\times n}$ represents an observation matrix, which is generally composed of a Gaussian matrix or a Bernoulli matrix or a Topplitz matrix. Furthermore, $\mathbf{\Psi} \in \mathbb{R}^{n\times n}$ is a basis function, which is made of DWT basis, DCT basis, etc. $\mathbf{b} \in \mathbb{R}^m$ denotes noise that obeys a Gaussian distribution.

From Eq. (1), we try to recover the sparse signal $\mathbf{x}$ from given $\{\mathbf{y}, \mathbf{A}\}$. In this case, the sensing matrix $\mathbf{A}$ contains more columns than rows, which means that there will be more than one solution that

Received 21 July 2018, Accepted 15 October 2018, Scheduled 27 October 2018
* Corresponding author: Huihui Yue (yuehuihui@hrbeu.edu.cn).
The authors are with the College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China.

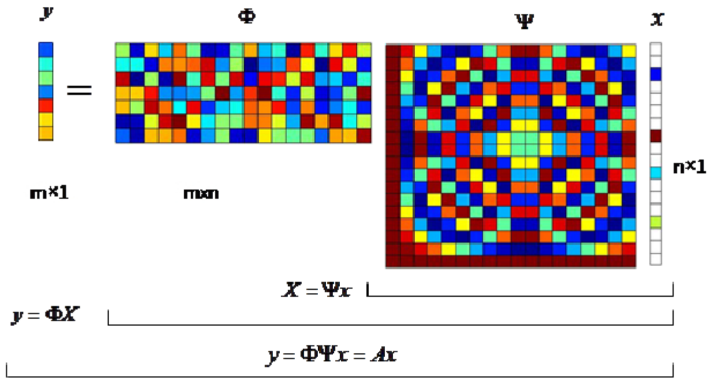

Figure 1. Framework of CS model in the noiseless case.

satisfies the constrain $||\mathbf{Ax}-\mathbf{y}||_2^2 \leq \epsilon$ (ϵ is a very small constant). This makes the recovery of sparse signal $\mathbf{x}$ an ill-posed problem. Luckily, since the target signal itself is sparse, the most straightforward method is to use its sparsity to improve the problem. So this problem is transformed into solving the L_0-norm minimization problem.

$$\arg\min_{\mathbf{x}\in\mathbb{R}^n} ||\mathbf{x}||_0 \quad s.t. \quad ||\mathbf{Ax}-\mathbf{y}||_2^2 \leq \epsilon \tag{2}$$

where $||\cdot||_0$ is L_0-norm, which represents the number of nonzero elements (sparsity). In order to minimize sparsity $||\mathbf{x}||_0$ and constraint term $||\mathbf{Ax}-\mathbf{y}||_2^2$, this optimization problem can be transformed from a constrained problem to an unconstrained problem which can be reformulated as a regularized least squares problem (RLSP) [6],

$$\arg\min_{\mathbf{x}\in\mathbb{R}^n} \frac{1}{2}||\mathbf{Ax}-\mathbf{y}||_2^2 + \lambda||\mathbf{x}||_0 \tag{3}$$

Unfortunately, $||\mathbf{x}||_0$ in Eqs. (2) and (3) are not directly processable, because the gradient of $||\mathbf{x}||_0$ cannot be solved, resulting in uncertainty of optimization direction and cannot be optimized. Therefore, only ergodic method can be used to solve L_0-norm, thus greatly increasing the computational complexity. In fact, this problem is usually relaxed into other forms. For example, $||\mathbf{x}||_0$ is replaced with $\Omega(\mathbf{x})$.

$$\arg\min_{\mathbf{x}\in\mathbb{R}^n} \frac{1}{2}||\mathbf{Ax}-\mathbf{y}||_\mathbf{2}^\mathbf{2} + \lambda\boldsymbol{\Omega}(\mathbf{x}). \tag{4}$$

For Eq. (4), there are many popular techniques for solving this problem. ISTA [7], FISTA [8], ADMM [9, 10], SpaRSA [11], ISpaRSA [12] and BPDN [13, 14] replace $\Omega(\mathbf{x})$ with L_1-norm or re-weighted L_1-norm. In noiseless case, L_1-norm is equivalent to L_0-norm, and L_1-norm is the only norm with sparsity and convexity, hence, it can be optimized by convex optimization methods. However, in noisy case, L_1-norm is not exactly equivalent to L_0-norm, so the effect of promoting sparsity is not obvious. Therefore, in the case of noise, only the algorithm that optimizes the approximation of L_0-norm can improve the accuracy of the algorithm. In [15], the authors proposed a new algorithm called L_p-RLS, which converts $\Omega(\mathbf{x})$ into $(\mathbf{x}^2+e)^{\frac{p}{2}}$. In [15], when $p\to 0$ and $e\to 0$, the objective function will approximate the form in Eq. (3). In [16], the authors proposed a smoothed function $F_\sigma(\mathbf{x})$ to replace $\Omega(\mathbf{x})$, because the smoothed function approximates the L_0-norm. In addition to the algorithms described above, in recent years, new CS algorithms such as EPRESS [17] and EWISTA [18] algorithms have emerged in the MRI field, which greatly improves the reconstruction performance.

Based on the above-mentioned state-of-the-art algorithms, this paper proposes a new algorithm called RRITSL0 algorithm. In this algorithm, we first propose an IT function approximating $||\mathbf{x}||_0$. Then a new iterative re-weighted function is proposed to promote signal sparsity. Finally, conjugate gradient (CG) method is used to implement the optimization process. On this basis, the proposed RRITSL0 algorithm is applied to MR image recovery.

This paper is organized as follows. Section 2 introduces the theories of the proposed RRITSL0 algorithm. Then we verify the performance of the RRITSL0 algorithm through simulation experiments and apply this algorithm to recover MR image in Section 3. Section 4 concludes this paper.

2. RELATE WORK

2.1. New Smoothed L_0-Norm Function Model

Mohimani et al. [19] proposed that the problem of finding the sparsest vector in the set $\{\mathbf{x}|\mathbf{y} = \mathbf{A}\mathbf{x}+\mathbf{b}\}$ can be interpreted as the task of approximating the Kronecker delta function, which is described as

$$\delta(x_i) = \begin{cases} 1, & \text{for } x_i = 0 \\ 0, & \text{otherwise} \end{cases}, \ i = 1, 2, \dots, n \tag{5}$$

Therefore, the L_0-norm of $\mathbf{x}$ is equal to $||\mathbf{x}||_0 = \sum_{i=1}^{n}[1-\delta(x_i)]$ and can be approximated by $\sum_{i=1}^{n} f(x_i)$ in which $f(x_i)$ denotes a smoothed function that acts as a delta approximating (DA) function. Based on this, we propose IT function

$$\begin{aligned} F_\sigma(\mathbf{x}) &= \sum_{i=1}^{n} f_\sigma(x_i), \\ f_\sigma(x_i) &= \frac{2}{\pi}\arctan\left(\frac{|x_i|}{\sigma}\right) \end{aligned} \tag{6}$$

where σ is a smoothed factor, and x_i is an independent variable. Obviously, $\lim_{\sigma\to 0} f_\sigma(x_i) = \begin{cases} 0, & \text{for } x_i = 0 \\ 1, & \text{otherwise} \end{cases}$ can be regarded as a DA function, so $||\mathbf{x}||_0$ can be approximated as $||\mathbf{x}||_0 \approx F_\sigma(\mathbf{x}) = \lim_{\sigma\to 0}\sum_{i=1}^{n} f_\sigma(x_i)$. Similarly, there are other smoothed functions proposed, such as the Gaussian function $f_\sigma(x_i) = 1 - e^{-\frac{x_i^2}{2\sigma^2}}$ in [19] and the hyperbolic tangent function (tanh) $f_\sigma(x_i) = \frac{e^{\frac{x_i^2}{2\sigma^2}} - e^{-\frac{x_i^2}{2\sigma^2}}}{e^{\frac{x_i^2}{2\sigma^2}} + e^{-\frac{x_i^2}{2\sigma^2}}}$ in [20]. We

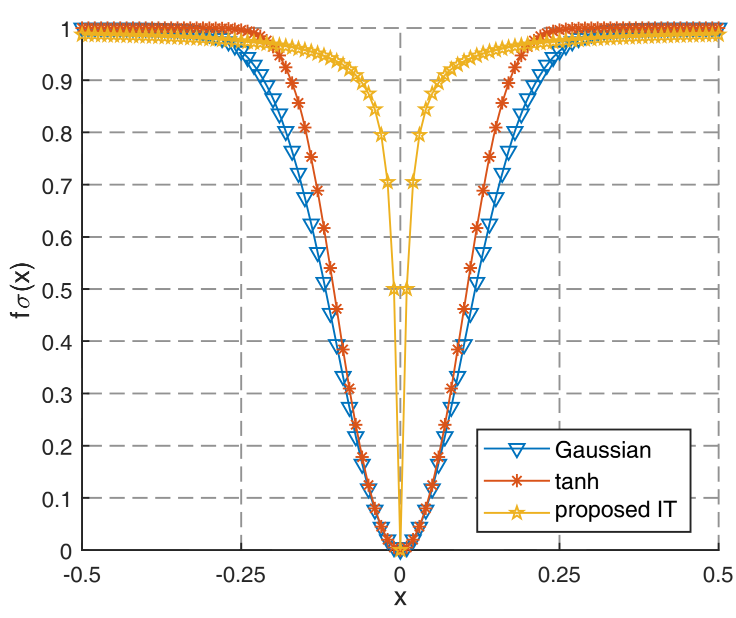

Figure 2. Different DA functions are plotted in this figure for comparison in 2D space when $\sigma = 0.1$.

can know that the three smoothed functions are closer to the DA function with respect to smaller σ, but the IT function can approximate the DA function more than the other two functions at the same σ, as shown in Fig. 2.

Figure 2 shows that the situation of IT function clusters approximates the L_0-norm. Obviously, the IT function makes a better approximation than the Gaussian and tanh functions. In general, this proposed smoothed function has two obvious merits:

1) Its clusters closely approximate L_0-norm;

2) It is simpler than tanh function.

These two merits can reduce the computational complexity on the premise of ensuring the reconstruction accuracy.

2.2. New Re-Weighted Function Design

Candès et al. [21] proposed the re-weighted L_1-norm minimization method, which employs the re-weighted norm to enhance the sparsity of the solution. And they provided an analytical result of the improvement in the sparsity recovery by incorporate re-weighted function to the objective function. Pant et al. [22] applied another re-weighted smoothed L_0-norm minimization method, which used a similar re-weighted function to improve sparsity. The re-weighted functions can be summarized as follows:

- Candès, et al.: $w_i = \begin{cases} \dfrac{1}{|x_i|}, & \text{for } x_i \neq 0 \\ \infty & \text{for } x_i = 0 \end{cases}$.
- Pant et al.: $w_i = \frac{1}{|x_i|+\zeta}$, ζ is a small enough positive constant.

From the two re-weighted functions, we can find a phenomenon: a large signal entry x_i is re-weighted with a small w_i. On the contrary, a small signal entry x_i is re-weighted with a large w_i. By analysis, the large w_i forces the solution $\mathbf{x}$ to concentrate on the indices where w_i is small, and by construction these correspond precisely to the indices where $\mathbf{x}$ is nonzero.

Combined with the above idea, we propose a new re-weighted function, which is given by

$$w_i = \frac{1}{e^{|x_i|}}, \quad i = 1, 2, \ldots, n \tag{7}$$

.

In Eq. (7), when $x_i \to 0$, $w_i \to 1$, and when $x_i \to -\infty$ or $x_i \to +\infty$, $w_i \to 0$. w_i is an even function that monotonically decreases on $x_i \in [0, \infty]$, which shows that the re-weighted function has maximum in location $x_i = 0$ and minimum in location x_i approximating negative infinity or positive infinity. By computation, the range of w_i in Eq. (7) is $[0, 1]$, while the range of Candès et al. is $[0, +\infty]$, and that of Pant et al. is $[0, \frac{1}{\zeta}]$. As for Candès et al., when signal entry is zero or close to zero, the value of w_i will be very large, which is not suitable for computation by computer. Although Pant et al. noticed the problem and improved the re-weighted function to avoid this problem, the constant ζ depends on experience. Luckily, the proposed re-weighted function can avoid this problem. In conclusion, the proposed re-weighted function has two merits:

1) It has a proper range that can give each signal component a proper re-weighted value, and when the signal component is close to zero, the re-weighted value will not be large.

2) It need not adjust parameters like ζ, and the denominator does not equal zero.

2.3. The New Proposed RRITSL0 Algorithm and Its Steps

As explained above, the objective function can be described as

$$\arg\min_{\mathbf{x}\in\mathbb{R}^n} \frac{1}{2}||\mathbf{y} - \mathbf{A}\mathbf{x}||_2^2 + \lambda \mathbf{W} F_\sigma(\mathbf{x}) \tag{8}$$

where λ is a regularized factor that revises the original objective function. Re-weighted function $\mathbf{W} = \text{diag}\{w_1, w_2, \ldots w_n\}$, and w_i is illustrated in Eq. (7) and can show the difference of each component of signal. The differentiable smoothed accumulated function $F_\sigma(\mathbf{x}) = \lim_{\sigma\to 0}\sum_{i=1}^{n} f_\sigma(x_i) = \lim_{\sigma\to 0}\sum_{i=1}^{n}\frac{2}{\pi}\arctan(\frac{|x_i|}{\sigma})$ is used to approximate $||\mathbf{x}||_0$.

Let

$$\mathbf{g} = \bigtriangledown F_\sigma(\mathbf{x}) = \sum_{i=1}^{n}\frac{\partial f_\sigma(x_i)}{\partial x_i} = \sum_{i=1}^{n}\frac{2}{\pi}\frac{1}{\left(\frac{|x_i|}{\sigma}\right)^2+1}\frac{\partial |x_i|}{\partial x_i} \tag{9}$$

In fact, $\frac{\partial |x_i|}{\partial x_i}$ does not exist, and the main reason is that $|x_i|$ cannot find the derivative at zero. In order to solve the problem, we make $\left.\frac{\partial |x_i|}{\partial x_i}\right|_{x_i=0} = 0$. Hence, $\frac{\partial |x_i|}{\partial x_i}$ can be represented as

$$\frac{\partial |x_i|}{\partial x_i} = \left\{\begin{array}{rl} 1, & \text{for}\ \ x_i > 0 \\ 0, & \text{for}\ \ x_i = 0 \\ -1, & \text{for}\ \ x_i < 0 \end{array}\right\} = \text{sign}(x_i) \tag{10}$$

From Eq. (10), $\frac{\partial |x_i|}{\partial x_i} = \text{sign}(x_i)$, hence, $\mathbf{g} = \sum_{i=1}^{n} g_i = \sum_{i=1}^{n}\frac{2}{\pi}\frac{1}{(\frac{|x_i|}{\sigma})^2+1}\text{sign}(x_i)$. Then the gradient for Eq. (8) can be written as

$$\mathbf{G} = \mathbf{A}^T(\mathbf{y}-\mathbf{A}\mathbf{x}) + \lambda\mathbf{W}\mathbf{g} \tag{11}$$

According to the objective function, the hessian of Eq. (8) can be readily expressed in closed form as

$$\mathbf{H} = \mathbf{A}^T\mathbf{A} + \lambda\mathbf{W}\mathbf{U}, \tag{12}$$

where

$$\mathbf{U} = \text{diag}\{u_1, u_2, \ldots, u_n\}, \tag{13}$$

$$u_i = \frac{\partial g_i}{\partial x_i} = \frac{4}{\pi}\frac{x_i\sigma^2}{x_i^2+\sigma^2}\text{sign}(-x_i), \quad i = 1, 2, \ldots, n \tag{14}$$

In fact, the problem of solving the objective function in Eq. (8) is translated into an optimization problem. This paper applies CG method to RRITSL0 algorithm to optimize the objective function. The problem can be firstly solved by using a sequential $\sigma - continuation$ strategy as detailed in the next paragraph.

Given a small target value σ_T and a sufficiently large initial value of parameter σ, i.e., σ_1, monotonically decreasing sequence $\{\sigma_t : t = 1, 2, 3, \ldots, T\}$ is generated as

$$\sigma_t = \sigma_1\theta^{-\alpha(t-1)}, \quad t = 1, 2, \ldots, T \tag{15}$$

where $\alpha = \frac{\log_\theta(\sigma_1/\sigma_T)}{T-1}$, and T is the maximum of iterations.

In the CG algorithm [23], iteratively, $\mathbf{x}_{(\mathit{\Gamma})}$ ($\mathit{\Gamma}$ denotes the number of inner loop iterations) is updated as

$$\mathbf{x}_{(\mathit{\Gamma}+1)} = \mathbf{x}_{(\mathit{\Gamma})} + \varrho_{(\mathit{\Gamma})}\mathbf{d}_{(\mathit{\Gamma})}, \tag{16}$$

where the parameter $\mathbf{d}_{(\mathit{\Gamma})}$ can be given by

$$\mathbf{d}_{(\mathit{\Gamma})} = -\mathbf{G}_{(\mathit{\Gamma})} + \eta_{(\mathit{\Gamma}-1)}\mathbf{d}_{(\mathit{\Gamma}-1)}, \tag{17}$$

the parameter $\eta_{(\mathit{\Gamma}-1)}$ is given as

$$\eta_{\mathit{\Gamma}-1} = \frac{\sum_{i=1}^{n}(G_{\mathit{\Gamma},i})^2}{\sum_{i=1}^{n}(G_{\mathit{\Gamma}-1,i})^2}, \tag{18}$$

and the parameter $\varrho_{(\Gamma)}$ is updated as

$$\varrho_{(\Gamma)} = \frac{\sum_{i=1}^{n}(G_{\Gamma,i})^2}{\mathbf{d}_{(\Gamma)}^T \mathbf{H}_{(\Gamma)} \mathbf{d}_{(\Gamma)}} \tag{19}$$

where $\mathbf{G}_{(\Gamma)}$ and $\mathbf{H}_{(\Gamma)}$ are the gradient and hessian of objective function in Eq. (8) evaluated at $\mathbf{x} = \mathbf{x}_{(\Gamma)}$ using Eqs. (11) and (12), respectively. As shown in Eq. (19), $\varrho_{(\Gamma)}$ is positive if $\mathbf{H}_{(\Gamma)}$ is positive definite (PD). And we can see from Eq. (12) that $\mathbf{A}^T\mathbf{A}$ is PD, and $\mathbf{W}$ is PD, so $\mathbf{H}_{(\Gamma)}$ is PD if $\mathbf{U}_{(\Gamma)}$ is PD. To get the PD of $\mathbf{U}_{(\Gamma)}$, we can make the following processing:

$$u_i = \begin{cases} u_i, & \text{for } u_i > \xi \\ \xi & \text{for } u_i \leqslant \xi \end{cases} \tag{20}$$

where ξ is a small positive constant (about 10^{-5}). The denominator in Eq. (19) can be evaluated efficiently as

$$\mathbf{d}_{(\Gamma)}^T \mathbf{H}_{(\Gamma)} \mathbf{d}_{(\Gamma)} = ||\mathbf{A}\mathbf{d}_{(\Gamma)}||_2^2 + \lambda||\mathbf{E}_{(\Gamma)}||_{\mathbf{2}}^{\mathbf{2}}, \tag{21}$$

$$\mathbf{E}_{(\Gamma)} = \mathbf{Q}_{(\Gamma)} \mathbf{d}_{(\Gamma)} \tag{22}$$

where $\mathbf{Q}_{(\Gamma)} = [q_{\Gamma,1}, q_{\Gamma,2}, \ldots, q_{\Gamma,n}]^T$ with $q_{\Gamma,i} = \sqrt{w_{\Gamma,i}u_i}$, $w_{\Gamma,i}$ is the component of w_i evaluated at $x = x_{(\Gamma)}$ using Eq. (7), thereby ϱ_Γ can be expressed as

$$\varrho_{(\Gamma)} = \frac{\sum_{i=1}^{n}(G_{\Gamma,i})^2}{||\mathbf{A}\mathbf{d}_{(\Gamma)}||_{\mathbf{2}}^{\mathbf{2}} + \lambda||\mathbf{E}_{(\Gamma)}||_{\mathbf{2}}^{\mathbf{2}}} \tag{23}$$

Based on the above explanation, we can conclude the steps of the proposed RRITSL0 algorithm, which is given in Table 1. As shown in Table 1, $\lambda = 0.1\lambda_{\max}$ and $\lambda_{\max} = 2||\mathbf{A}^T\mathbf{y}||_\infty$ are the same as the value in [24]. As for σ, it can be shown that function $F_\sigma(\mathbf{x})$ remains convex in the region where the largest magnitude of the component of $\mathbf{x}$ is less than σ. Based on this, a reasonable initial value of σ can be chosen as $\sigma_1 = \max(|x_i|) + \tau$ (τ is defaulted as 0.01) to ensure the optimization starts in a convex region. This greatly facilitates the convergence of the RRITSL0 algorithm.

Table 1. Regularized re-weighted inverse trigonometric smoothed function approximating L_0-norm minimization (RRITSL0) algorithm.

Initialization: $\mathbf{A}$, $\mathbf{x}$, $\mathbf{y}$, ξ, η, τ, σ_T, T, λ **and** $\mathbf{x}^* = \mathbf{x}$
Step1: Set $\sigma_1 = \max(|x_i|) + \tau$, $t = 0$**;**
Step2: Compute W using Eq. (7);
Step3:
For $t = 1, 2, 3, \ldots, T$
compute σ_t using Eq. (15);
1) Set $\sigma = \sigma_t$, $\Gamma = 0$, $\mathbf{x}_{(\Gamma)} = \mathbf{x}^*$
2) Compute Residual $Res = ||\varrho_{(\Gamma)}\mathbf{d}_{(\Gamma)}||_2^2$, and iterative termination threshold err
3) While $Res > err$
a) Compute $\mathbf{x}_{(\Gamma+1)}$ using Eqs. (16)-(23), $\mathbf{W}$ using Eq. (7)
b) Set $\Gamma = \Gamma + 1$
c) Compute $Res = ||\varrho_{(\Gamma)}\mathbf{d}_{(\Gamma)}||_2^2$
4) Set $\mathbf{x}^* = \mathbf{x}_{(\Gamma)}$
Step4: Output $\mathbf{x} = \mathbf{x}^*$

3. NUMERICAL SIMULATION AND ANALYSIS

In this section, we will verify the performance of the RRITSL0 algorithm in the case of noise and apply the algorithm to MR image recovery. The numerical simulation platform is MATLAB 2017b, which is installed on the WINDOWS 10, 64-bit operating system. The CPU is Inter (R) Core (TM) i5-3230M, and the frequency is 2.6 GHz.

First, we analyze the performance of the proposed RRITSL0 algorithm in sparse signal recovery and compare it with SL0, L2-SL0 [25] and L_p-RLS [15] algorithms. We fix $n = 256$ and $m = 100$ and the sparsity $k = 4s + 1$, $s = 1, 2, \dots, 15$, or let $n = [170, 220, 270, 320, 370, 420, 470, 520]$, $m = n/2$, $k = n/5$. For every experiment, we randomly generate a pair of $\{\mathbf{x}, \mathbf{A}, \mathbf{b}\}$: $\mathbf{A}$ is an $m \times n$ random Gaussian matrix with normalized and centralized rows; the *nonzero* entries of the sparse signal $\mathbf{x} \in \mathbb{R}^n$ are *i.i.d.* generated according to the Gaussian distribution $\mathcal{N}(0, 1)$; $\mathbf{b}$ is randomly formed and follows the *Gaussian* distribution of $\mathcal{N}(0, \delta_N)$.

Given the measurement vector $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}$, sensing matrix $\mathbf{A}$ and noise $\mathbf{b}$, we try to recover the signal $\mathbf{x}$. Choose the parameters that give the best performance for each method: for the SL0 algorithm, $\sigma_{\min} = 10^{-2}$, scale factor is set as $S = 10$, $\rho = 0.8$; for L2-SL0 algorithm, $\sigma_{\min} = 0.01$, $S = 10$, $\rho = 0.8$; for L_p-RLS algorithm, $p_1 = 1$, $p_T = 0.1$, $\epsilon_1 = 1$, $\epsilon_T = 10^{-2}$; and for RRITSL0 algorithm, $\sigma_T = 10^{-2}$, $err = 10^{-8}$. All experiments are based on 100 trials.

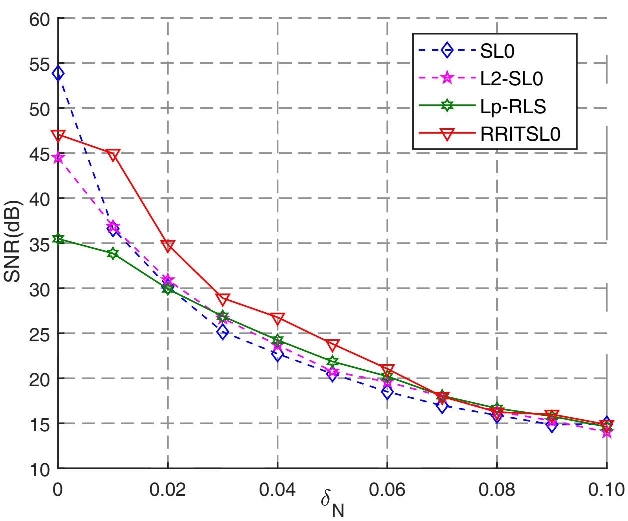

Figure 3. Signal SNR analysis for the SL0, L2-SL0, L_p-RLS algorithms and the proposed RRITSL0 algorithm with noise power factor δ_N equaling [0, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1] while 100 runs.

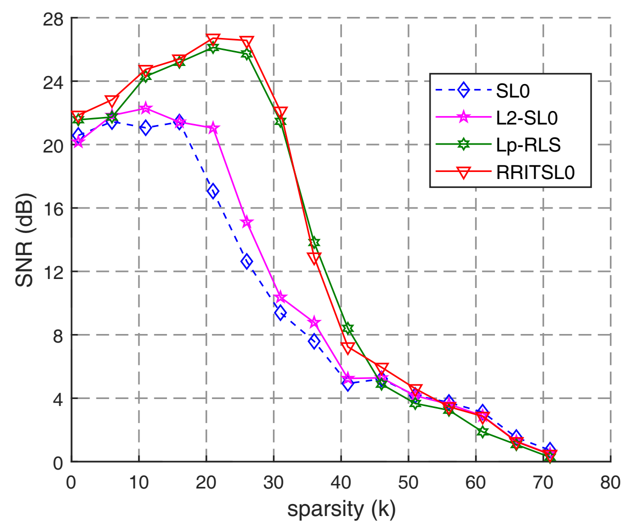

Figure 4. Signal SNR analysis for the SL0, L2-SL0, L_p-RLS algorithms and the proposed RRITSL0 algorithm with noise power factor $\delta_N = 0.05$ and the sparsity $k = 4s + 1$, $s = 1, 2, \dots, 15$ while 100 runs.

Next, we analyze the convergence of proposed algorithm by experiments.

At last, we apply the proposed algorithm to recover MR image. For MR image recovery, we conclude the recovery performance when the Compression Ratio (CR) is certain or changed. CR is defined as m/n.

3.1. The Recovery Performance Comparison of the Algorithms

In this section, we evaluate the recovery performance of the RRITSL0 algorithm to recover the noisy signal by Signal to Noise Ratio (SNR), Normalized Mean Square Error (NMSE) and CPU Running Time (CRT). SNR and NMSE are two aspects of signal reconstruction performance and have the same effect. They are respectively defined as $20 \log(||\mathbf{x} - \widehat{\mathbf{x}}||_2/||\mathbf{x}||_2)$ and $||\mathbf{x} - \hat{\mathbf{x}}||_2/||\mathbf{x}||_2$. CRT is measured with *tic* and *toc*.

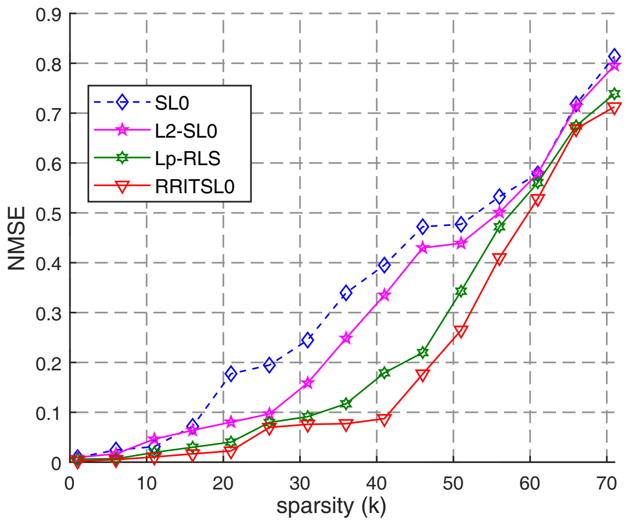

Figure 5. Signal NMSE analysis for the SL0, L2-SL0, L_p-RLS algorithms and the proposed RRITSL0 algorithm with noise power factor $\delta_N = 0.05$ and the sparsity $k = 4s + 1$, $s = 1, 2, \ldots, 15$ while 100 runs.

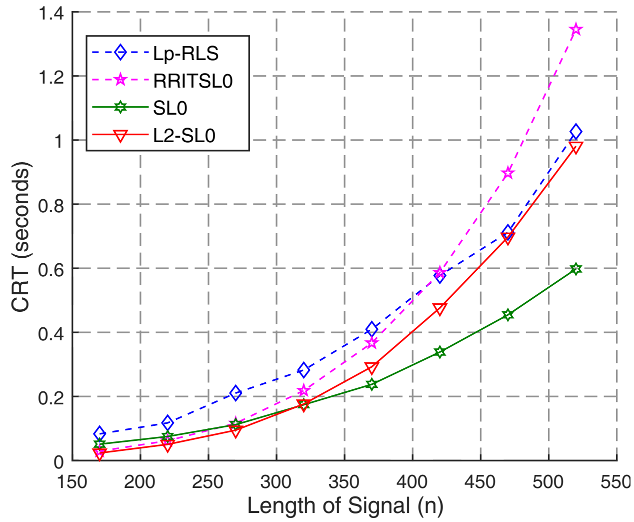

Figure 6. Signal CRT analysis for the SL0, L2-SL0, L_p-RLS algorithms and the proposed RRITSL0 algorithm while 100 runs with $\delta_N = 0.05$.

SNR of recovered signal is shown in Fig. 3 when $n = 256$, $m = 100$, $k = 20$. The SNR of all algorithms decreases sharply with increase of δ_N, which shows that the noise can seriously affect the performance of algorithms. Despite this, the RRITSL0 gets the largest SNR, followed by three other algorithms with similar SNR, which proves that the de-noise performance of the RRITSL0 is better than the other three algorithms.

Figures 4 and 5 show the SNR and NMSE of all chosen algorithms with noise power factor δ_N equaling 0.05 and the sparsity $k = 4s+1$, $s = 1, 2, \ldots, 15$. From the two figures, the proposed RRITSL0 outperforms other chosen algorithms.

The CRT is shown in Fig. 6 when $n = [170, 220, 270, 320, 370, 420, 470, 520]$, $m = n/2$, $k = n/5$. The figure shows that in terms of CRT performance, although the RRITSL0 is superior to the L_p-RLS, it is inferior to the SL0 and L2-SL0. Therefore, improving the reconstruction speed is one of the main directions of the RRITSL0 algorithm in the future.

3.2. The Convergence Performance Comparison of the Algorithms

To illustrate the convergence of the proposed RRITSL0 algorithm, we present the performance of NMSE with the iterations in Figs. 7 and 8. For this section, the signal is set as random signal $\mathbf{x} \in \mathbb{R}^n$ and measurement vector $\mathbf{y} \in \mathbb{R}^m$ with $n = 256$, $m = 100$, $k = 20$.

Figure 7 shows the NMSE of the SL0, L2-SL0, L_p-RLS and RRITSL0 algorithms with iterations. It can be seen that these algorithms eventually converge to a very small value, but obviously, RRITSL0 with re-weighted function has the fastest convergence rate. In addition, the NMSE of the RRITSL0 algorithm is always minimal at any time. This fully proves that the proposed RRITSL0 can promote the sparsity of the signal and thus improve the convergence speed.

Figure 8 shows the NMSE of the proposed RRITSL0 algorithm at different k with iterations. As k increases, the convergence speed of the RRITSL0 gradually decreases. But when k is less than 25, the RRITSL0 has a faster convergence speed. This shows that the RRITSL0 has good convergence when k is not large.

Through the above simulation experiments, we prove that the RRITSL0 algorithm can accurately reconstruct the sparse signal under noise conditions and has good convergence performance, which provides a basis for the application of the RRITSL0 algorithm in MR image processing.

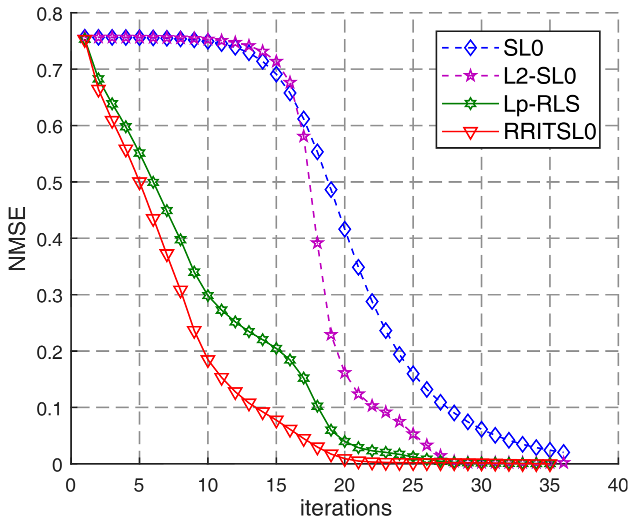

Figure 7. NMSE of recovery signal changes with iterations, the figure shows the comparison between the SL0, L2-SL0, L_p-RLS algorithms and the proposed RRITSL0 algorithm.

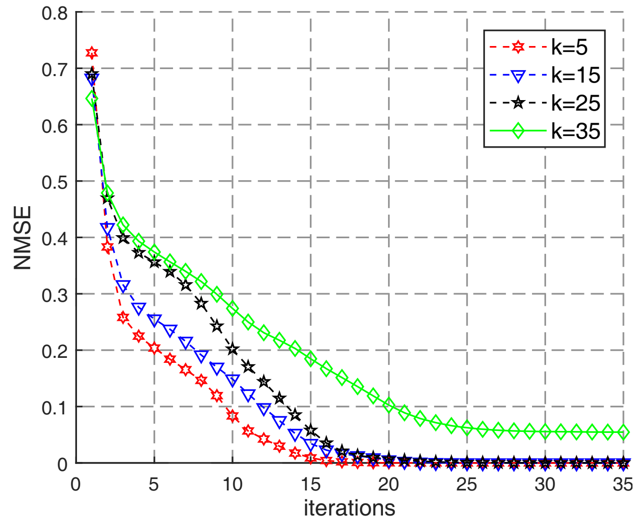

Figure 8. NMSE of recovery signal changes with iterations of the proposed RRITSL0 algorithm, the figure shows the comparison with sparsity $k = [5, 15, 25, 35]$.

3.3. MR Image Recovery Performance Comparison of the Algorithms

Real images are considered approximately sparse under some proper basis, such as the DCT basis and DWT basis. Here we choose DWT basis as sparse basis of *brain* MR image. The size of this MR image is 256×256. First, we verify the recovery performance of the proposed RRITSL0 algorithm by comparing it with the SL0, L2-SL0 and L_p-RLS algorithms. The noise δ_N equals 0.01, and CR is 0.4, 0.5, 0.6. Then, we fix CR to 0.5 and compare the MR image recovery effects of the RRITSL0 at $\delta_N = [0, 0.05, 0.1, 0.2, 0.5]$. For performance of MR image recovery, we valuate it by Peak Signal to Noise Ratio (PSNR) and Structural Similarity index (SSIM). PSNR is defined as

$$\mathrm{PSNR} = 10 \log(255^2/\mathrm{MSE}) \tag{24}$$

where $\mathrm{MSE} = ||\mathbf{x} - \widehat{\mathbf{x}}||_2^2$, and SSIM is defined as

$$\mathrm{SSIM}(p, q) = \frac{(2\mu_p + \mu_q + c_1)(2\sigma_{pq} + c_2)}{(\mu_p^2 + \mu_q^2 + c_1)(\sigma_p^2 + \sigma_q^2 + c_2)} \tag{25}$$

Among this, μ_p is the mean of MR image p, μ_q the mean of image q, σ_ps the variance of MR image p, σ_q the variance of MR image q, and σ_{pq} the covariance between MR image p and MR image q. Parameters $c_1 = z_1 L$ and $c_2 = z_2 L$, in which $z_1 = 0.01$, $z_2 = 0.03$, and L is the dynamic range of pixel values. The range of SSIM is $[-1, 1]$, and when these two images are same, the SSIM equals 1.

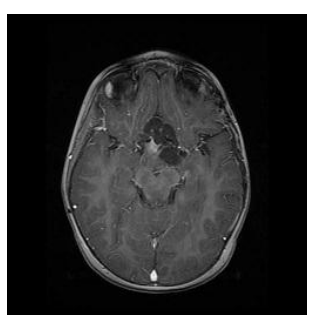

Figure 9. Original MR image.

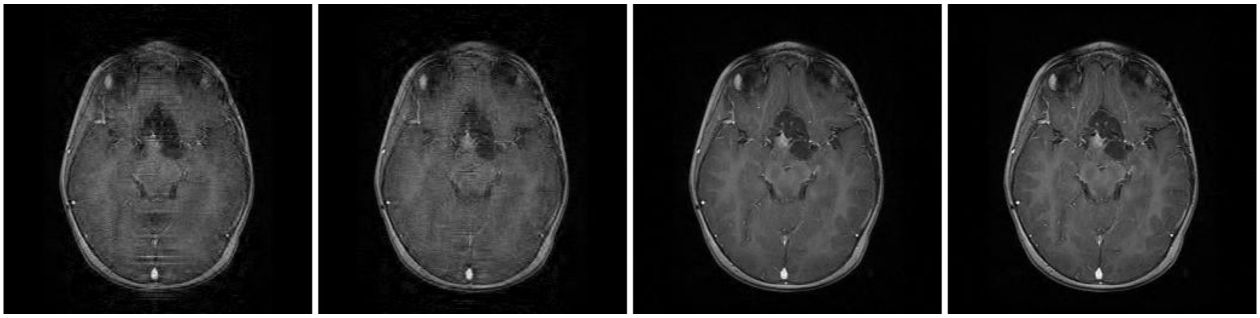

Figure 10. MR Image recovery effect by the SL0, L2-SL0, L_p-RLS and the proposed RRITSL0 algorithms with CR = 0.4. From left to right in the figure are: recovered MR image by the SL0, recovered MR image by the L2-SL0, recovered MR image by the L_p-RLS and recovered MR image by the proposed RRITSL0.

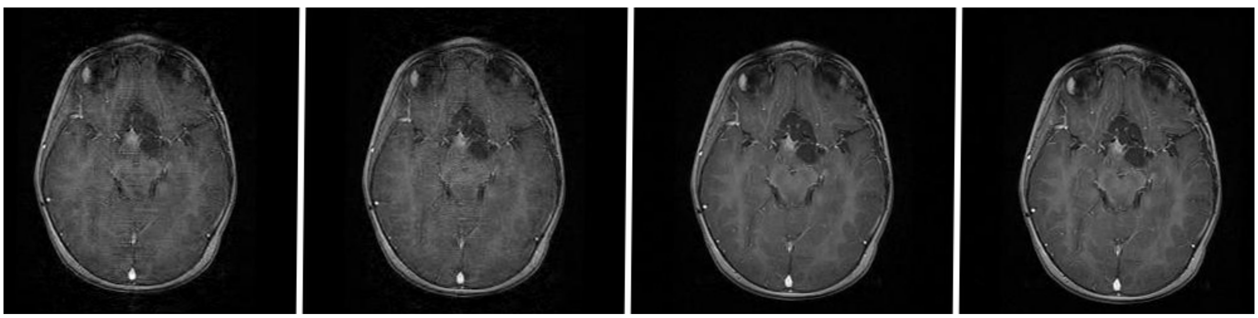

Figure 11. MR Image recovery effect by the SL0, L2-SL0, L_p-RLS and the proposed RRITSL0 algorithms with CR = 0.5. From left to right in the figure are: recovered MR image by the SL0, recovered MR image by the L2-SL0, recovered MR image by the L_p-RLS and recovered MR image by the proposed RRITSL0.

Figure 9 shows an original MR image. Figs. 10–12 show the situation about MR image recovery when CR is respectively 0.4, 0.5, 0.6. Table 2 shows the PSNR and SSIM of Figs. 10–12. As shown in Figs. 10–12, all algorithms can clearly recover MR image when CR is over 0.5. Furthermore, we can see the difference of each algorithm in detail by scientific data in Table 2. From the table, the RRITSL0 has better PSNR and SSIM than other three selected algorithms, which verifies the good MR image recovery performance of the proposed RRITSL0. So, the proposed RRITSL0 can be used for MR image recovery.

Figure 13 shows de-noise performance of the proposed RRITSL0 algorithm when recovering sparse images. When δ_N is less than 0.2, the difference in MR image recovery is not obvious. But when δ_N

Table 2. PSNR and SSIM of recovery MR image by the SL0, L2-SL0 and L_p-RLS algorithms and the proposed RRITSL0 algorithm.

CR	PSNR (dB)				SSIM (%)			
	SL0	L2-SL0	L_p-RLS	RRITSL0	SL0	L2-SL0	L_p-RLS	RRITSL0
0.4	29.065	30.206	37.368	39.237	97.46	98.06	99.63	99.78
0.5	31.967	32.710	38.965	40.183	98.72	98.92	99.74	99.79
0.6	34.875	35.282	40.243	41.031	99.35	99.41	99.81	99.86

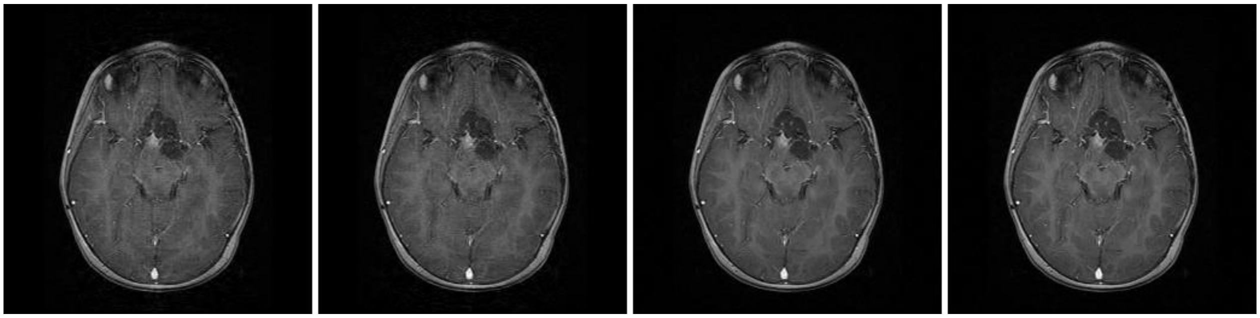

Figure 12. MR Image recovery effect by the SL0, L2-SL0, L_p-RLS and the proposed RRITSL0 algorithms with CR = 0.6. From left to right in the figure are: recovered MR image by the SL0, recovered MR image by the L2-SL0, recovered MR image by the L_p-RLS and recovered MR image by the proposed RRITSL0.

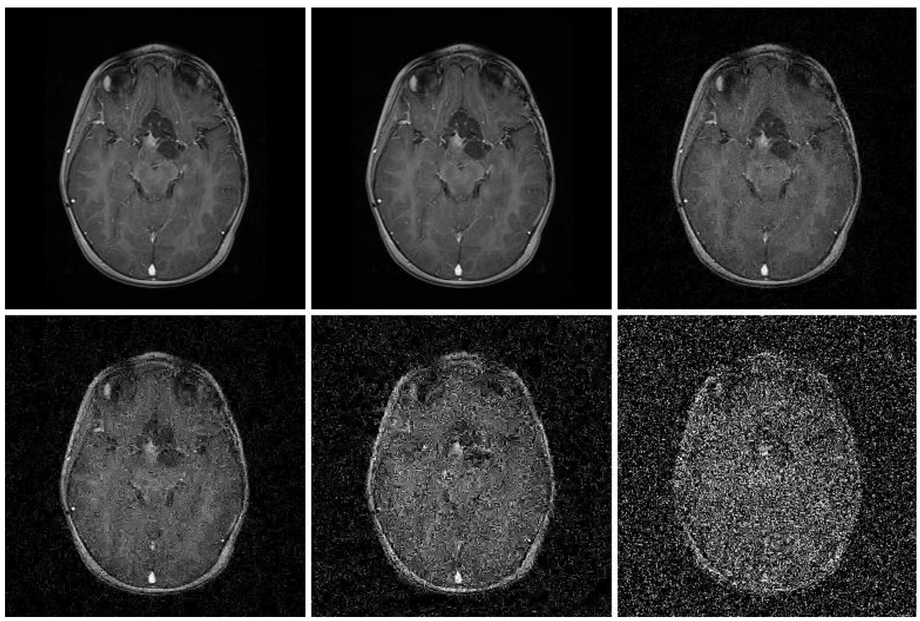

Figure 13. MR image recovery effect by the proposed RRITSL0 algorithm when noise is incrementing according to a sequence $\delta_N = [0, 0.05, 0.1, 0.2, 0.5]$. And the above three sub-figures in this figure are from left to right: original MR image, recovered MR image with $\delta_N = 0$, 0.05. And the following three sub-figures in this figure are from left to right: recovered MR image with $\delta_N = 0.1$, 0.2, 0.5.

Table 3. PSNR and SSIM of recovered MR image by the the proposed RRITSL0 algorithm with $\delta_N = [0, 0.05, 0.1, 0.2, 0.5]$.

δ_N	PSNR (dB)	SSIM (%)
0	47.150	99.92
0.05	30.148	98.31
0.10	24.282	92.79
0.20	18.375	76.88
0.50	12.518	35.52

is over 0.2, the effect of MR image recovery is significantly reduced. These show that the proposed RRITSL0 has a certain ability to de-noise, but under high noise conditions, the effect needs to be improved. Table 3 gives the scientific data. From the table, we can see that as δ_N gradually increases from 0 to 0.5, both PSNR and SSIM decrease.

Through experiments, we can know that the proposed RRITSL0 algorithm can obtain better results in sparse image recovery. This is mainly because the IT function cluster used by it can approximate the L_0-norm well; the re-weighted function can promote the sparsity; the regularization mechanism has the ability to resist noise.

4. CONCLUSIONS

This paper proposes an RRITSL0 algorithm for reconstructing sparse signals and images. The RRITSL0 is based on the IT function as a DA function, and its cluster can better approximate the L_0-norm. Then, we use re-weighted function to promote sparsity and apply the CG algorithm to optimize. Furthermore, experiments show that: (1) the RRITSL0 algorithm can improve accuracy and has certain de-noise performance; (2) RRITSL0 has a faster convergence speed; (3) the RRITSL0 satisfies the needs of sparse signals and MR image recovery, and improves chance of CS applied to other fields. In our future research, the RRITSL0 algorithm will be optimized for operating rates and high de-noise performance.

ACKNOWLEDGMENT

This paper is supported by the National Key Laboratory of Communication Anti-jamming Technology (No. 614210202030217).

REFERENCES

1. Chen, Z., Y. Fu, Y. Xiang, and R. Rong, "A novel iterative shrinkage algorithm for CS-MRI via adaptive regularization," *IEEE Signal Processing Letters*, Vol. 99, 1–1, 2017.
2. Yazdanpanah, A. P. and E. E. Regentova, "Compressed sensing MRI using curvelet sparsity and nonlocal total variation: CS-NLTV," *Electronic Imaging*, Vol. 2017, No. 13, 5–9, 2017.
3. Candès, E. J., "Compressive sampling," *IEEE Proceedings of the International Congress of Mathematicians*, 2006.
4. Candès, E. J. and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, Vol. 25, No. 2, 21–30, 2008.
5. Li, S., H. Yin, and L. Fang, "Remote sensing image fusion via sparse representations over learned dictionaries," *IEEE Transactions on Geoscience & Remote Sensing*, Vol. 51, No. 9, 4779–4789, 2013.
6. Zhang, J., D. Zhao, F. Jiang, and W. Gao, "Structural group sparse representation for image compressive sensing recovery," *Data Compression Conference*, Vol. 6, No. 3, 331–340, 2013.
7. Josá, M., D. Bioucas, and M. A. T. Figueiredo, "A new TwIST: Two-step iterative shrinkage/thresholding algorithms for image restoration," *IEEE Transactions on Image Processing*, Vol. 16, No. 12, 2992–3004, 2007.
8. Beck, A. and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *Siam Journal on Imaging Sciences*, Vol. 2, No. 1, 183–202, 2009.
9. Ghadimi, E., A. Teixeira, and I. Shames, "Optimal parameter selection for the alternating direction method of multipliers (ADMM): quadratic problems," *IEEE Transactions on Automatic Control*, Vol. 60, No. 3, 644–658, 2015.
10. Ramdas, A. and R. J. Tibshirani, "Fast and flexible ADMM algorithms for trend filtering," *Journal of Computational and Graphical Statistics*, Vol. 25, No. 3, 839–858, 2014.
11. Wright, S. J., R. D. Nowak, and M. A. T. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2479–2493, 2008.

12. Ye, X., W. Zhu, A. Zhang, and Q. Meng, "Sparse channel estimation in MIMO-OFDM systems based on an improved sparse reconstruction by separable approximation algorithm," *Journal of Information & Computational Science*, Vol. 10, No. 2, 609–619, 2013.
13. Figueiredo, M. A. T., R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application To Compressed Sensing And Other Inverse problems," *IEEE Journal of Selected Topics in Signal Processing*, Vol. 1, No. 4, 586–597, 2008.
14. Zibulevsky, M. and M. Elad, "L1-L2 optimization in signal and image processing," *IEEE Signal Processing Magazine*, Vol. 27, No. 3, 76–88, 2010.
15. Pant, J. K., W. S. Lu, and A. Antoniou, "New improved algorithms for compressive sensing based on Lp NORM," *IEEE Transactions on Circuits & Systems II Express Briefs*, Vol. 61, No. 3, 198–202, 2014.
16. Ye, X., W. P. Zhu, A. Zhang, and J. Yan, " Sparse channel estimation of MIMO-OFDM systems with unconstrained smoothed L0 -norm-regularized least squares compressed sensing," *Eurasip Journal on Wireless Communications & Networking*, Vol. 2013, No. 1, 282, 2013.
17. Zhang, Y., B. S. Peterson, G. Ji, and Z. Dong, "Energy preserved sampling for compressed sensing MRI," *Computational and Mathematical Methods in Medicine*, Vol. 2014, No. 5, 546814, 2014.
18. Zhang, Y., S. Wang, G. Ji, and Z. Dong, "Exponential wavelet iterative shrinkage thresholding algorithm with random shift for compressed sensing magnetic resonance imaging," *Information Sciences*, Vol. 10, No. 1, 116–117, 2015.
19. Mohimani, H., M. Babaie-Zadeh, and C. Jutten, "A fast approach for overcomplete sparse decomposition based on smoothed L0 norm," *IEEE Transactions on Signal Processing*, Vol. 57, No. 1, 289–301, 2009.
20. Zhao, R., W. Lin, H. Li, and S. Hu, "Reconstruction algorithm for compressive sensing based on smoothed L_0 norm and revised newton method," *Journal of Computer-Aided Design & Computer Graphics*, Vol. 24, No. 4, 478–484, 2012.
21. Candès, E. J., M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by re-weighted L1 minimization," *Journal of Fourier Analysis and Applications*, Vol. 14, No. 5, 877–905, 2008.
22. Pant, J. K., W. S. Lu, and A. Antoniou, "Reconstruction of sparse signals by minimizing a re-weighted approximate L0-norm in the null space of the measurement matrix," *IEEE International Midwest Symposium on Circuits and Systems*, 430–433, 2010.
23. Zibetti, M. V. W., C. Lin, and G. T. Herman, "Total variation superiorized conjugate gradient method for image reconstruction," *Inverse Problems*, Vol. 34, No. 3, 2017.
24. Wen, F., Y. Yang, P. Liu, R. Ying, and Y. Liu, "Efficient ℓ_q minimization algorithms for compressive sensing based on proximity operator," *Mathematics*, 2016.
25. Ye, X. and W. P. Zhu, "Sparse channel estimation of pulse-shaping multiple-input—multiple-output orthogonal frequency division multiplexing systems with an approximate gradient L_2SL_0 reconstruction algorithm," *Iet Communications*, Vol. 8, No. 7, 1124–1131, 2014.