A New Non-Convex Approach for Compressive Sensing MRI

Huihui Yue and Xiangjun Yin*

Abstract—Compressive sensing (CS) is an effective method for reconstructing magnetic resonance imaging (MRI) image from under-determined linear system (ULS). However, how to improve the accuracy of MRI image reconstructed by CS is still a serious problem, especially in noisy conditions. To solve this problem, in this paper, we propose a novel approach, dubbed as regularized maximum entropy function (RMEF) minimization algorithm. Specifically, motivated by the entropy function in information theory, we propose a maximum entropy function (MEF) to approximate L_q-norm $(0 < q < 1)$ as sparsity promoting objectives, and then the regularization mechanism for improving the de-noising performance is adopted. Combining the above two ideas, a new objective function of RMEF method is proposed, and the global minimum is iteratively solved. We further analyze the convergence to verify the robustness of the RMEF algorithm. Experiments demonstrate the state-of-the-art performances of the proposed RMEF algorithm and show that the RMEF achieves higher PSNR and SSIM than other widely-adopted methods in MRI image recovery.

1. INTRODUCTION

MRI is a noninvasive imaging technology which plays an important role in clinical diagnosis. MRI uses different characteristics of energy attenuation released by different material structures in the environment, detects the information emitted by gradient magnetic fields, and realizes the mapping of the internal structure of the object, so as to achieve imaging. However, the limitation of the Shannon-Nyquist sampling theorem and the physical mechanism of magnetic resonance data encoding leads to the slow speed of conventional MRI. This is mainly because the MRI scanner captures not the original image pixel, but the global Fourier transform image in the frequency domain. Each frequency domain pixel is a linear combination of time domain pixels, which makes MRI data redundancy more obvious and thus increases the sampling time. In view of this, the accurate reconstruction of MRI with low sampling rate is of great significance.

As an undersampling MRI technique, compressive sensing magnetic resonance imaging (CS-MRI) breaks through the limitation of Shannon-Nyquist sampling theorem and offers us an efficient framework to deal with the above challenge. CS-MRI technology first undersamples data in k-space (i.e., Fourier space), then obtains image using CS theory [1–3]. Therefore, CS tasks eventually boil down to the MRI image recovery problem in the following ULS [4]:

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

where $y \in \mathbb{R}^m$ is an under-sampled k-space data, $x \in \mathbb{R}^n$ the MRI image to be reconstructed, and w the additive Gaussian noise. $\mathbf{A} = \mathbf{\Phi D} \in \mathbb{R}^{m \times n}$ is a measurement matrix, $\mathbf{\Phi} \in \mathbb{R}^{m \times n}$ an undersampling matrix, and $\mathbf{D} \in \mathbb{R}^{n \times n}$ a Fourier transform, $m \ll n$.

In CS, the MRI image x is recovered from given y and $\mathbf{A}$. For this, $\mathbf{A}$ has more columns than rows, which results in multiple solutions to recover x, thus making recovery of x an ill-posed problem.

In addition, since the image of interest itself is sparse in k-space, the most straight-forward way is to search for the solution that also shares this sparse property [4]. Combining the above two points, we apply the L_0-norm to choose the sparsest one:

$$\min_x \|x\|_0, \quad s.t. \ y = \mathbf{A}x. \tag{2}$$

This fairly simple form of Equation (2) is actually supported by theories [4]. According to these theories, when x is sparse and $\mathbf{A}$ satisfies the restricted isometry property (RIP) [4, 5], x can be well restored. In fact, solving the Equation (2) is a nonconvex NP-hard problem whose solutions require an intractable combinatorial search [6]. To solve this problem, scholars have developed two alternative solutions:

- Greedy search method when $\|x\|_0 \leq k$;
- Relaxation method for $\|x\|_0$.

The sparsity of x is k, which represents the upper bound on the number of nonzero entries in x. Greedy search method is mainly composed of a series of algorithms focusing on matching pursuit algorithms [7–10]. These algorithms have good results in noiseless environments, but they do not perform well in noisy environments. The main purpose of relaxation method for $\|x\|_0$ is to convert L_0-norm into other forms [11–13] equivalently so as to avoid the NP-hard problem. This kind of method is relatively scalable. Here this paper focuses on relaxation method that tries to solve the following unconstrained recovery problem:

$$\min_x \|y - \mathbf{A}x\|_2^2 + \lambda g(x). \tag{3}$$

where λ is the positive parameter used to balance the trade-off between data error $\|y - \mathbf{A}x\|_2^2$ and the sparsity regularizer $g(x)$. The role of $g(x)$ is to promote sparsity, and the function of $\|y - \mathbf{A}x\|_2^2$ is to ensure that the recovered image has certain fidelity whether it is noisy or noiseless. The combination of the two can improve the reconstruction accuracy, especially in the noisy case.

For relaxation methods, some convex relaxation methods are proposed, such as basis pursuit de-noise (BPDN) [14], SpaRSA [15], FISTA [16], and D-AMP [17, 18], which relax the L_0-norm regularization into the L_1-norm regularization. In other words, let $g(x)$ be the L_1-norm. This L_1-regularized problem can be efficiently tackled by convex optimization techniques if they meet some conditions, such as the RIP, null space property (NSP) [19], and incoherence condition (IC) [20]. However, due to the relaxation, the recovery accuracy is often degraded in noiseless case, e.g., it often introduces extra bias [21] and cannot reconstruct an image with the least observations [22]. Furthermore, for some applications, the result of the L_1-norm minimization is not sparse enough, and the original images cannot be recovered [23, 24].

To solve this problem, a number of state-of-the-art algorithms is proposed, which replace $g(x)$ with the L_q-norm $(0 < q < 1)$ [25]. Algorithms in this category include StSALq [26], UALp [27], and L_q-RLS [28]. All these algorithms use a de-noising model in Equation (3) and solve an unconstrained L_q-norm regularized least squares (L_q-LS) problem. Compared with the L_1-norm, the L_q-norm is a closer approximation of the L_0-norm. It has been shown in [29] that under certain RIP condition of $\mathbf{A}$, L_q-norm minimization algorithms require fewer sampling data but gain a better recovery performance than L_1-norm minimization algorithms. Moreover, the sufficient conditions in terms of RIP for L_q-norm minimization are weaker than those for L_1-norm minimization [30, 31]. However, in general, relative to L_1-norm minimization, L_q-norm minimization is more difficult to directly tackle due to its nonsmoothing. To solve this problem, L_q-RLS algorithm, which uses an efficient conjugate gradient (CG) method to solve a sequence of smooth subproblems, makes it succeed to handle large scale problem.

In fact, most of algorithms solve the L_q-norm minimization problem via smoothing (approximating) it, and for example, the works [32, 33] use an approximation of $\|x\|_q^q$ as

$$\|x\|_{q,\epsilon}^q = \sum_{i=1}^n \left(x_i^2 + \epsilon^2\right)^{q/2}, \tag{4}$$

where ϵ is a small enough positive constant. Furthermore, the iteratively re-weighted (IR) algorithm [34]

uses the following two penalties

$$\|x\|_{q,\epsilon}^q = \sum_{i=1}^n (|x_i| + \epsilon)^{q-1} |x_i|, \tag{5}$$

$$\|x\|_{q,\epsilon}^q = \sum_{i=1}^n (|x_i|^2 + \epsilon^2)^{q/2-1} |x_i|^2, \tag{6}$$

which explicitly relate to the L_q-norm approximation. Based on this, we obtain a smoothing function, MEF, to approximate L_q-norm, and the sparsity regularizer $g(x)$ in Equation (3) is then replaced with MEF to form a new objective function. Moreover, iterative optimization method is used to tackle this nonsmooth problem. Hence, this paper develops a highly efficient numerical method for the nonconvex L_q-minimization problem.

The rest of the paper is organized as follows: Section 2 introduces the maximum entropy function for nonconvex L_q-norm functions and then proposes a new RMEF minimization algorithm to solve the ULS. Then the convergence of the RMEF minimization algorithm is proved in Section 3. In Section 4, the effect of the RMEF minimization algorithm in MRI image reconstruction with noise is verified by comparing it with L2-SL0 [35], UALp, and L_q-RLS approaches. Section 5 concludes this paper.

2. RELATED WORK

2.1. Maximum Entropy Function to Approximate L_q-Norm

Constructing a smooth L_q-norm is of great significance for CS, as shown in Equations (4–6), and L_q-norm can be represented as $\|t\|_q^q = \sum_{i=1}^n |t|^q$, where $t \in \mathbb{R}^n$, $t \in \mathbb{R}$, and

$$|t| = \begin{cases} t, & t > 0 \\ -t, & t \leq 0 \end{cases}, \tag{7}$$

so it can be converted as

$$\tau(t) = |t| = \max\{t, -t\}. \tag{8}$$

Here, we introduce a maximum entropy function which can be given by

$$\tau(t, \alpha) = \frac{1}{\alpha} \left\{ \log \left(10^{\alpha t} + 10^{-\alpha t} \right) \right\}, \alpha > 0. \tag{9}$$

Then, for any $\alpha > 0$,

$$\tau(t, \alpha) = \tau(t) + \frac{1}{\alpha} \left\{ \log \left(10^{\alpha(t - \tau(t))} + 10^{\alpha(-t - \tau(t))} \right) \right\}, \tag{10}$$

where $0 \leq \{\log(10^{\alpha(t - \tau(t))} + 10^{\alpha(-t - \tau(t))}) \leq \log 2$, which implies

$$\tau(t) \leq \tau(t, \alpha) \leq \tau(t) + \frac{\log 2}{\alpha}. \tag{11}$$

From Equation (11), when $\alpha \to \infty$, $\tau(t, \alpha) \to \tau(t)$. Thus, the function $\tau(t, \alpha)$ defined in Equation (9) is a smoothing approximation of the function $\tau(t)$, which is an MEF.

Here let $x = [x_1, x_2, \ldots, x_n]^T$, and moreover we make

$$\Gamma(x, \alpha) = \lim_{\alpha \to \infty} \sum_{i=1}^n \tau(x_i, \alpha)^q = \lim_{\alpha \to \infty} \sum_{i=1}^n \left(\frac{1}{\alpha} \left\{ \log \left(10^{\alpha x_i} + 10^{-\alpha x_i} \right) \right\} \right)^q, \tag{12}$$

and rewrite the L_q-norm as:

$$\Gamma(x) = \sum_{i=1}^n \max\{x_i, -x_i\}^q. \tag{13}$$

Property 1 $\forall x \in \mathbb{R}^n$, $\alpha > 0$, $\Gamma(x, \alpha)$ is an axisymmetric function, that is to say, $\Gamma(x, \alpha) = \Gamma(-x, \alpha)$.

Proof sketch. *Firstly, the definition field of* $\Gamma(x, \alpha)$ *is a real number field, which is a symmetric domain. Then,* $\Gamma(-x, \alpha) = \lim\limits_{\alpha \to \infty} \sum\limits_{i=1}^{n} (\frac{1}{\alpha} \{\log(10^{\alpha(-x_i)} + 10^{-\alpha(-x_i)})\})^q = \lim\limits_{\alpha \to \infty} \sum\limits_{i=1}^{n} (\frac{1}{\alpha} \{\log(10^{\alpha x_i} + 10^{-\alpha x_i})\})^q.$

$$
\begin{aligned}
\Gamma(x, \alpha) - \Gamma(-x, \alpha) &= \lim_{\alpha \to \infty} \sum_{i=1}^{n} \left(\frac{1}{\alpha} \left\{ \log\left(10^{\alpha(x_i)} + 10^{-\alpha(x_i)}\right) \right\} \right)^q - \lim_{\alpha \to \infty} \sum_{i=1}^{n} \left(\frac{1}{\alpha} \left\{ \log\left(10^{\alpha(-x_i)} + 10^{-\alpha(-x_i)}\right) \right\} \right)^q \\
&= \lim_{\alpha \to \infty} \sum_{i=1}^{n} \left(\frac{1}{\alpha} \left\{ \log\left(10^{\alpha x_i} + 10^{-\alpha x_i}\right) \right\} \right)^q - \lim_{\alpha \to \infty} \sum_{i=1}^{n} \left(\frac{1}{\alpha} \left\{ \log\left(10^{\alpha x_i} + 10^{-\alpha x_i}\right) \right\} \right)^q \\
&= \Gamma(x, \alpha) - \Gamma(x, \alpha) \\
&= 0.
\end{aligned}
$$

Therefore, $\Gamma(x, \alpha) = \Gamma(-x, \alpha)$. *To summarize, Property 1 is proved.*

Theorem 1 $\forall x \in \mathbb{R}^n, \alpha > 0$, *function* $\Gamma(x, \alpha)$ *is a bounded function whose upper bound converges to* $\Gamma(x) + n(\frac{\log 2}{\alpha})^q$ *and the lower bound converges to* $\Gamma(x)$.

Proof *Please see Appendix A.*

From Theorem 1, it can be seen that

$$
0 \leq \Gamma(x, \alpha) - \Gamma(x) \leq n \left(\frac{\log 2}{\alpha} \right)^q. \tag{14}
$$

Then we let $\epsilon = n(\frac{\log 2}{\alpha})^q$, and ϵ is a small enough constant, so

$$
0 \leq \Gamma(x, \alpha) - \Gamma(x) \leq \epsilon. \tag{15}
$$

Therefore, $\Gamma(x, \alpha)$ is equivalent to $\Gamma(x)$, that is to say, the MEF can replace L_q-norm. Therefore, the proposed maximum entropy function model can be used for MRI image recovery.

2.2. Regularized Maximum Entropy Function Minimization Algorithm for MRI

Performing MRI image recovery under noise conditions is an important observational measure of the application of CS technology in practical fields. Traditional CS algorithms are extremely sensitive to noise due to the effects of noise folding [36, 37], thus result in poor recovery performance. Regularization mechanism can offset the impact of noise, so as to achieve the purpose of suppressing noise. This paper adopts this idea, and then the corresponding compression-aware objective function can be improved to

$$
L_q(x, \lambda, \alpha) = \frac{1}{2} \|\mathbf{A}x - y\|_2^2 + \lambda \Gamma(x, \alpha), \tag{16}
$$

where $\lambda > 0$. The above equation is the model of the proposed RMEF minimization algorithm. As shown in this equation, this minimization problem must have a solution because $L_q(x, \lambda, \alpha)$ is continuous with respect to x, and it can achieve the minimum over a bounded set $\{x \mid \|x\|_2 \leq \triangle\}$, where $\triangle$ is a positive constant. In addition, $L_q(x, \lambda, \alpha)$ blows up as $\|x\|_2 \to \infty$. Let $\beta = \frac{1}{\alpha} > 0$, and $x_{\beta, \lambda, q}$ denotes a critical point and it satisfies the first-order optimality condition

$$
\lambda \frac{q}{\alpha^{q-1}} \frac{10^{\frac{x}{\beta}} - 10^{\frac{-x}{\beta}}}{10^{\frac{x}{\beta}} + 10^{\frac{-x}{\beta}}} \log^{q-1} \left(10^{\frac{x}{\beta}} + 10^{\frac{-x}{\beta}} \right) + \mathbf{A}^T(\mathbf{A}x - y) = 0. \tag{17}
$$

Due to the underdetermination, there is no straightforward method to solve the above system of equations unless for specific instances, such as $\mathbf{A}^T\mathbf{A}$ is a diagonal matrix. Combining with [32], we approximately solve the system with a sequence of β, and the method is summarized as follows.

In the steps of proposed algorithm above, β is a descending order execution sequence formed by annealing mechanism. If $|x^{(t+1)} - x^{(t)}| < \xi$, we choose $x^{(t+1)}$ to be an approximation of the MRI image and stop the iteration. Otherwise, we stop the computation within a reasonable time and return the last $x^{(t+1)}$. It is easy to know that the ULS in Equation (1) is invertible for any $x^{(t+1)}$ if $\beta > 0$. Once β is small enough, the iteration can be stopped. Thus, the algorithm is well reasonable.

Algorithm 1 : Regularized Maximum Entropy Function (RMEF) Minimization Algorithm for MRI Image Reconstruction

---

Input: under-sampled k-space data y and matrix $\mathbf{A}$.

Output: MRI image x.

Choose appropriate parameters $\lambda > 0, q \in (0,1)$

Initialize $x^{(0)}$ such that $\mathbf{A}x^{(0)} = y$ and $\beta_0 = (\frac{\epsilon}{n})^{\frac{1}{q}}/\log 2$, $\beta = \beta_0$, $\beta_\Upsilon = 10^{-6}$.

For $t = 0, 1, 2, \ldots, \Upsilon$

    Solve the following linear system for $x^{(t+1)}$

$$\frac{q}{\alpha^{q-1}} \frac{10^{\frac{x^{(t)}}{\beta}} - 10^{\frac{-x^{(t)}}{\beta}}}{10^{\frac{x^{(t)}}{\beta}} + 10^{\frac{-x^{(t)}}{\beta}}} \log^{q-1}(10^{\frac{x^{(t)}}{\beta}} + 10^{\frac{-x^{(t)}}{\beta}}) + \frac{1}{\lambda}\mathbf{A}^T(\mathbf{A}x^{(t+1)} - y) = 0$$

    or equivalently

$$x^{(t+1)} = \frac{1}{\mathbf{A}^T\mathbf{A}}(\mathbf{A}^T y - \frac{q\lambda}{\alpha^{q-1}} \frac{10^{\frac{x^{(t)}}{\beta}} - 10^{\frac{-x^{(t)}}{\beta}}}{10^{\frac{x^{(t)}}{\beta}} + 10^{\frac{-x^{(t)}}{\beta}}} \log^{q-1}(10^{\frac{x^{(t)}}{\beta}} + 10^{\frac{-x^{(t)}}{\beta}}))$$

    Update $\beta = \beta_0 10^{-\mu(t+1)}$, where $\mu = \frac{\log(\beta_0/\beta_\Upsilon)}{\Upsilon+1}$.

    if $|x^{(t+1)} - x^{(t)}| < \xi$ or $\beta = \beta_\Upsilon$, then ends the cycling. ξ is a small eough positive constant.

---

3. CONVERGENCE OF THE PROPOSED RMEF MINIMIZATION ALGORITHM

In this section, we discuss the convergence of the proposed RMEF minimization algorithm. Before analysis, we derive two simple properties of the first derivative of the function $\tau(x, \alpha)$, which will be used in the following analysis.

Proposition 1 *Given $\alpha \in \mathbb{R}^+$, we have that*

(i) $|\tau'(x, \alpha)| > 0.5(1 - 10^{-2})$ holds for any $x \in \mathbb{R}$ satisfying $|x| > \beta$;

(ii) $\lim_{\beta \to 0} \tau'(x, \alpha) = \text{sign}(x)$ holds for any $x \in \mathbb{R}$.

where $\beta = \frac{1}{\alpha}$, and x is a scalar.

Proof sketch. *(i) For any $|x| > \beta$, there will be either $x > \beta$ or $x < -\beta$.*

    • *If $x > \beta$, there will be $-x < -\beta$, hence, $\frac{x}{\beta} > 1$ and $\frac{-x}{\beta} < -1$. Furthermore,*

$$10^{\frac{x}{\beta}} > 10^1 \quad and \quad 0 < 10^{\left(\frac{-x}{\beta}\right)} < 10^{-1},$$

which imply that $10^{\frac{x}{\beta}} - 10^{\left(\frac{-x}{\beta}\right)} > 10^1 - 10^{(-1)} > 0$. Thus,

$$|\tau'(x, \alpha)| = \frac{10^{\frac{x}{\beta}} - 10^{\left(\frac{-x}{\beta}\right)}}{10^{\frac{x}{\beta}} + 10^{\left(\frac{-x}{\beta}\right)}} > \frac{10^{\frac{x}{\beta}} - 10^{(-1)}}{2 * 10^{\frac{x}{\beta}}}$$
$$= 0.5\left(1 - \frac{10^{(-1)}}{10^{\frac{x}{\beta}}}\right) > 0.5\left(1 - 10^{-2}\right).$$

    • *If $x < -\beta$, there will be $-x > \beta$. By the fact that $\tau'(x, \alpha)$ is odd function and the above inequality, we have that*

$$|\tau'(x, \alpha)| = |-\tau'(-x, \alpha)| > 0.5\left(1 - 10^{-2}\right).$$

  (ii) • *If $x > 0$, there will be $10^{\frac{x}{\beta}} \to \infty$ and $10^{\left(\frac{-x}{\beta}\right)} \to 0$ when $\beta \to 0$. Hence,*

$$\lim_{\beta \to 0} \tau'(x, \alpha) = \lim_{\beta \to 0} \frac{10^{\frac{x}{\beta}} - 10^{\left(\frac{-x}{\beta}\right)}}{10^{\frac{x}{\beta}} + 10^{\left(\frac{-x}{\beta}\right)}} = \lim_{\beta \to 0} \left\{ 1 - \frac{2 * 10^{\left(\frac{-x}{\beta}\right)}}{10^{\frac{x}{\beta}} + 10^{\left(\frac{-x}{\beta}\right)}} \right\} = 1.$$

.

• *If $x < 0$, there will be $-x > 0$. By the odd function property of $\tau'(x, \alpha)$ and the above equality, we can get*

$$\lim_{\beta \to 0} \tau'(x, \alpha) = -\lim_{\beta \to 0} \tau'(-x, \alpha) = -1.$$

.

• *If $x = 0$, there will be $10^{\frac{x}{\beta}} = 10^{(\frac{-x}{\beta})} = 1$, hence $\tau'(x, \alpha) = 0$. Therefore, the Proposition 1 is proved.*

In addition, we need the following notation and results. Denote

$$\Gamma'(x, \alpha) := \nabla \Gamma(x, \alpha) = \left(\frac{\partial \Gamma(x, \alpha)}{\partial x_1}, \frac{\partial \Gamma(x, \alpha)}{\partial x_2}, \ldots, \frac{\partial \Gamma(x, \alpha)}{\partial x_n} \right)^T,$$

$$\frac{\partial \Gamma(x, \alpha)}{\partial x_i} = q\tau^{q-1}(x_i, \alpha)\tau'(x_i, \alpha), i = 1, \ldots, n.$$

The first-order necessary condition of $L_q(x, \lambda, \alpha)$ is given in Equation (17), which means χ that satisfied with the above equation is a stationary point of $L_q(x, \lambda, \alpha)$ if it satisfies Equation (17).

Moreover, for given $q \in (0, 1)$ and $\lambda > 0$, the definition of L_q regularized problem is shown as follows:

$$L_q(x, \lambda) = \lambda \|x\|_q^q + \|\mathbf{A}x - y\|_2^2,\tag{18}$$

then $x \in \mathbb{R}^n$ is called a stationary point of L_q regularized problem if it satisfies [38]

$$\lambda q |x|^q + \begin{bmatrix} x_1 & 0 & \ldots & \cdot & 0 \\ 0 & x_2 & \ldots & \cdot & \\ & \cdot & \ldots & \cdot & \\ & \cdot & \ldots & x_{n-1} & 0 \\ 0 & \cdot & \ldots & 0 & x_n \end{bmatrix} \mathbf{A}^T(\mathbf{A}x - y) = 0.\tag{19}$$

Theorem 2 *Suppose that the sequence $\{x_{\beta t}\}$ generated by proposed RMEF minimization algorithm is contained in the level set $\{x | L_q(x, \lambda, \alpha) \leq L_q(x^0, \lambda, \alpha)\}$ for an arbitrary given initial point x^0.*

(i) For any accumulation point of $\{x_{\beta t}\}$ is a stationary point of $L_q(x, \lambda)$.

(ii) Let $\{x_{\beta t}\}$ be a global minimizer of $\{x | L_q(x, \alpha, \lambda)\}$ for any given t, then any accumulation point of $\{x_{\beta t}\}$ is a global minimizer of $L_q(x, \lambda)$.

Proof *Please see Appendix B.*

4. RESULT AND ANALYSIS

Here we verify the MRI image recovery performance of the proposed RMEF minimization algorithm by comparing it with the state-of-the-art L2-SL0, UALp, and L_q-RLS algorithms on *Brain* (256×256) datasets. The numerical simulation platform is MATLAB 2017b, which is installed on the WINDOWS 10, 64-bit operating system. The type of CPU is Inter (R) Core (TM) i5-3230M, and the frequency is 2.6 GHz.

Given a pair of $\{y, \Phi, D, w\}$, we try to obtain the original MRI image. The recovery process of the MRI image by the RMEF minimization algorithm is shown in Figure 1, while the image recovery performance is valuated by *Peak Signal to Noise Ratio* (PSNR) and *Structural Similarity Index* (SSIM). PSNR is defined as

$$\text{PSNR} = 10 \log \left(255^2 / \text{MSE} \right),\tag{20}$$

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

$$\text{SSIM}(p, q) = \frac{(2\mu_p + \mu_q + c_1)(2\sigma_{pq} + c_2)}{(\mu_p^2 + \mu_q^2 + c_1)(\sigma_p^2 + \sigma_q^2 + c_2)}.\tag{21}$$

where μ_p is the mean of image p, μ_q the mean of image q, σ_p the variance of image p, σ_q the variance of image q, and σ_{pq} the covariance between image p and image q. Parameters $c_1 = z_1 L$ and $c_2 = z_2 L$,
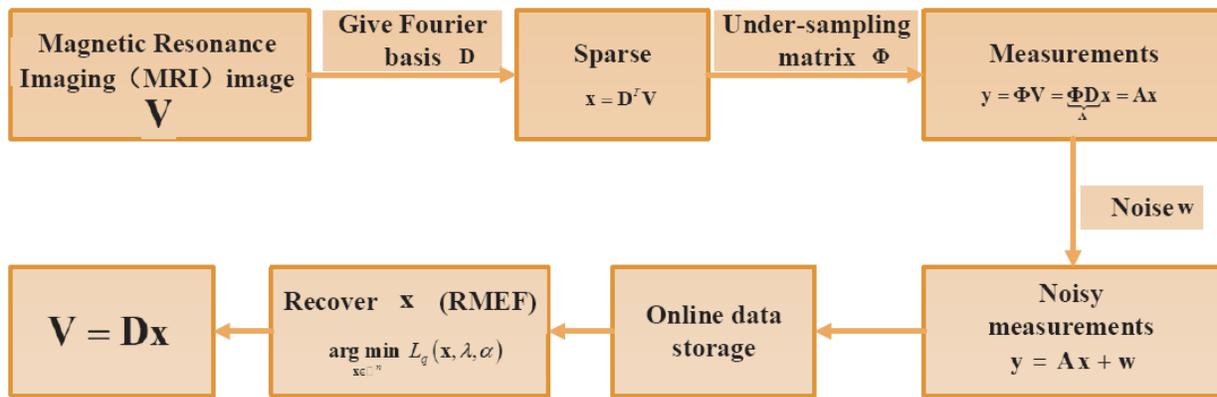
Figure 1. Recovery process of the MRI image by the RMEF minimization algorithm.

in which $z_1 = 0.01$, $z_2 = 0.03$ and L is dynamic range of pixel values. Range of SSIM is $[-1, 1]$, and if the two compared images are same, the SSIM equals 1.

Figures 2, 3, and 4 show the MRI image recovery effect of the selected algorithms when compression ratio (CR, which is defined as m/n) is respectively 0.4, 0.5, 0.6, and noise intensity δ_N equals 0.01. As shown in these figures, under any same CR, the proposed RMEF minimization algorithm can recover a clearer MRI image than the other three algorithms. However, the differences of each algorithm in image recovery performance are not obvious, so the recovery performance cannot be determined. Further, Table 1 shows the difference in detail through scientific data. From this table, it can be seen that the proposed RMEF minimization algorithm performs better than the other three methods in *Brain* recovery, which is consistent with the results in Figures 2, 3, and 4. Compared with L_q-RLS algorithm, which is the best one among the other three algorithms, the proposed RMEF minimization algorithm can improve approximately 0.7038 dB, 0.6027 dB, and 0.0249 dB on PSNR and 0.5%, 0.31%, and 0.01%
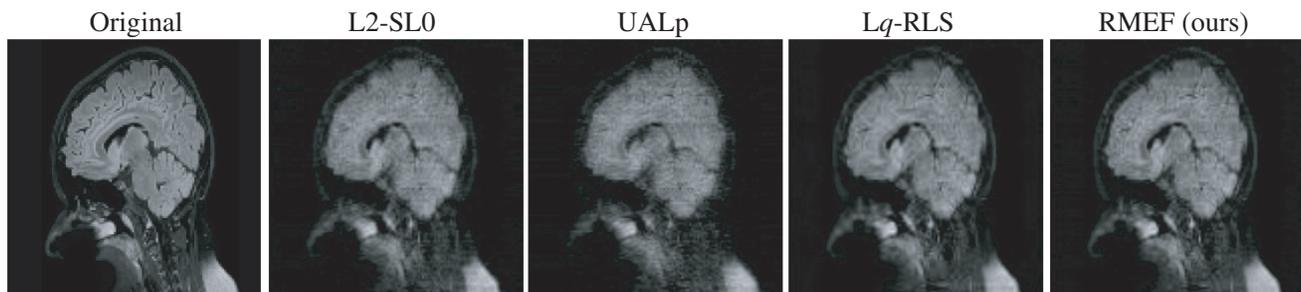

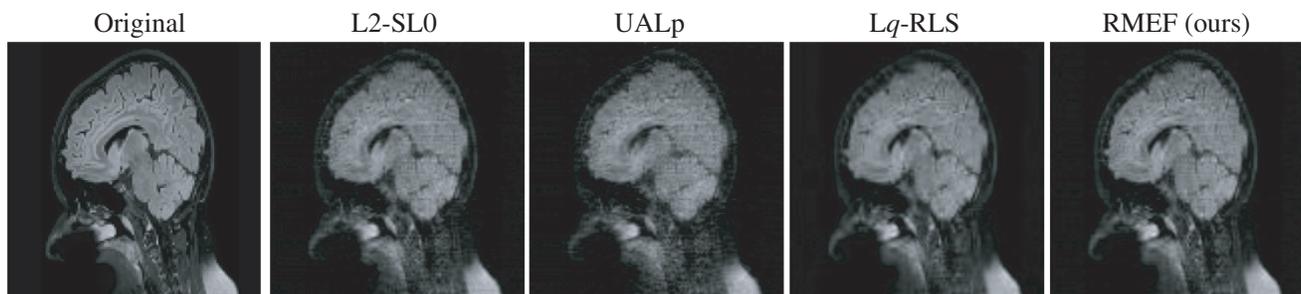**Figure 2.** *Brain* image recovery effect by different algorithms with $\delta_N = 0.01$, $CR = 0.4$.
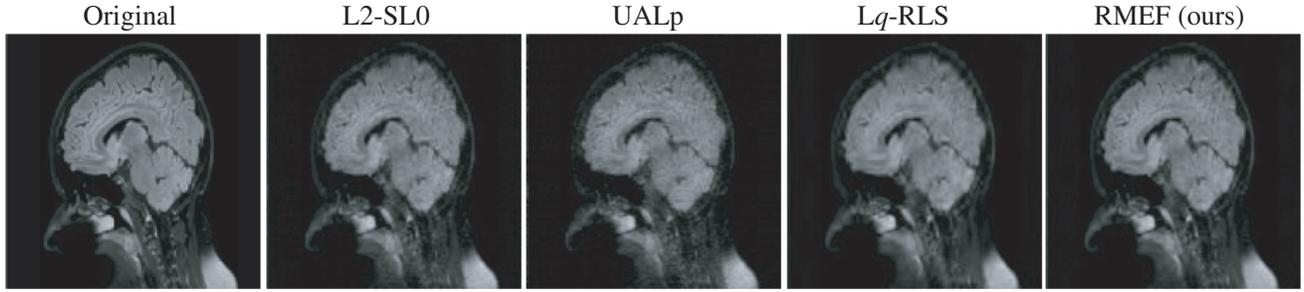

Figure 3. *Brain* image recovery effect by different algorithms with $\delta_N = 0.01$, $CR = 0.5$.

| Original | L2-SL0 | UALp | Lq-RLS | RMEF (ours) |

Figure 4. *Brain* image recovery effect by different algorithms with $\delta_N = 0.01$, $CR = 0.6$.

Table 1. PSNR and SSIM analysis of *Brain* recovered by the L2-SL0, UALp, L_q-RLS, and proposed RMEF minimization algorithms with $CR = 0.4, 0.5, 0.6$ and $\delta_N = 0.01$.

| CR	PSNR (dB)				SSIM (%)			
	L2-SL0	UALp	L_q-RLS	RMEF	L2-SL0	UALp	L_q-RLS	RMEF
0.4	25.4844	25.9210	26.1200	**26.8258**	95.98	96.30	96.53	**97.03**
0.5	26.6514	26.9360	27.6133	**28.2160**	96.94	97.14	97.55	**97.86**
0.6	26.9159	27.0416	28.7090	**28.7339**	97.15	97.23	98.09	**98.10**

on SSIM with respect to CR = $[0.4, 0.5, 0.6]$, respectively. Therefore, these experimental results prove that the proposed RMEF minimization algorithm can reconstruct MRI image more accurately than other state-of-the-art approaches.

To further evaluate the effect of the proposed RMEF minimization algorithm in MRI image recovery under different noise intensities, Figure 5 shows the vivid reconstruction results at $\delta_N = [0, 0.05, 0.1, 0.2, 0.3, 0.5]$, and CR is fixed to 0.5. Table 2 shows the scientific data. From Figure 5 we can see that when δ_N is less than 0.2, the difference in image recovery of the RMEF minimization algorithm is not obvious. But when δ_N is over 0.2, the effect of image reovery is significantly reduced. The data in Table 2 also fully illustrate this result. As can be seen from Table 2, when δ_N is lower than 0.2, the SSIM of MRI images recovered by the proposed RMEF minimization algorithm is always higher than 90%. These show that the proposed RMEF minimization algorithm has a certain ability to denoise, but the effect needs to be improved under large noise conditions.

Table 2. PSNR and SSIM analysis of *Brain* recovered by the proposed RMEF minimization algorithm under different δ_N.

δ_N	MRI image	PSNR (dB)	SSIM (%)
0	*Brain*	30.9261	98.85
0.05	*Brain*	27.8016	97.64
0.1	*Brain*	24.5298	95.12
0.2	*Brain*	18.4506	82.59
0.3	*Brain*	13.6842	60.92
0.5	*Brain*	8.1555	29.49

In conclusion, experiments verify the most advanced MRI image recovery performance of the RMEF minimization algorithm. Therefore, it is feasible for the RMEF minimization algorithm to be applied to MRI image recovery.
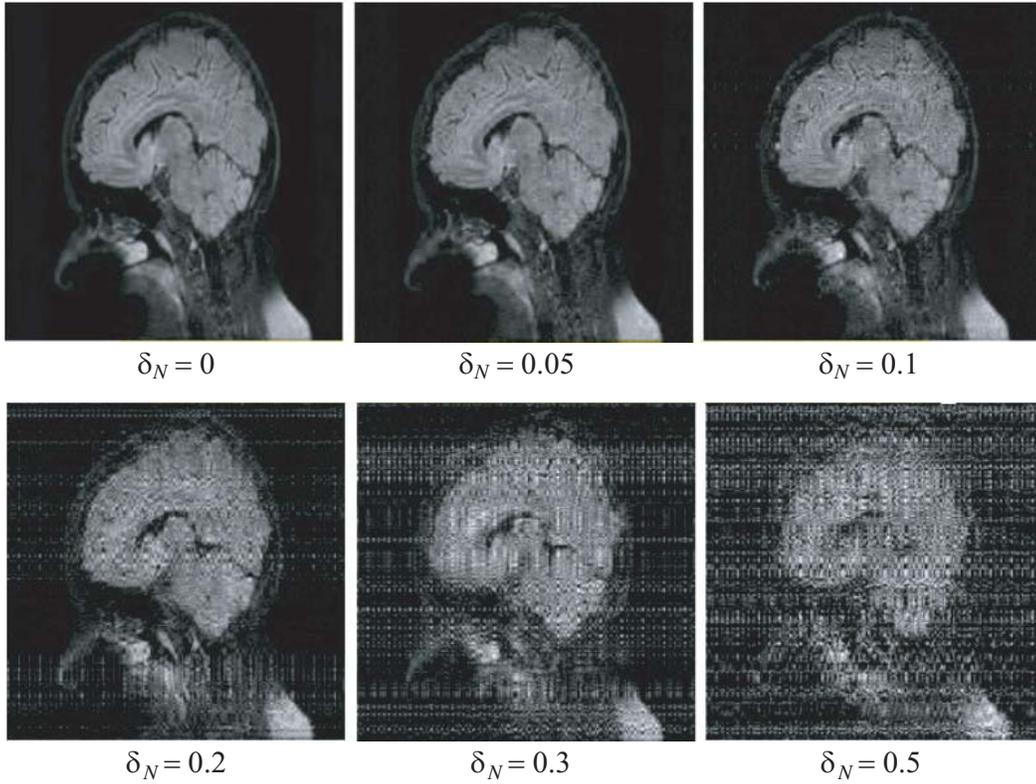
Figure 5. *Brain* image recovery effect by the proposed RMEF minimization algorithm with $\delta_N = [0, 0.05, 0.1, 0.2, 0.3, 0.5]$.

5. CONCLUSIONS

In this paper, a non-convex RMEF minimization algorithm is proposed to reconstruct MRI images with noise. We firstly propose a MEF and use the property that MEF approximates to L_q-norm to promote the sparsity of the recovered images. On this basis, we replace sparsity regularizer $g(x)$ with proposed MEF and form a new approach called RMEF minimization algorithm. We also prove that the proposed RMEF minimization algorithm can converge to optimal solution of MRI images. Finally, experiments show that the proposed RMEF performs better than the popular $\|x\|_q^q$ regularization algorithms in MRI image recovery. Additionally, in the future, we would also like to apply the RMEF minimization algorithm to other CS applications such as the Blind Source Separation (BSS) [39], Robust Principal Component Analysis (RPCA) [40], and Dictionary Learning [41, 42].

APPENDIX A. PROOF OF THEOREM 1

Proof *Known by Equation (11),* $\forall x > 0, \alpha > 0$, *there will be*

$$\max\{x, -x\} \leq \frac{1}{\alpha} \left\{ \log \left(10^{\alpha x} + 10^{-\alpha x} \right) \right\} \leq \max\{x, -x\} + \frac{\log 2}{\alpha}. \tag{A1}$$

According to the definitions of $\Gamma(x, \alpha)$ and $\Gamma(x)$, we can get

$$\Gamma(x, \alpha) - \Gamma(x) = \lim_{\alpha \to \infty} \sum_{i=1}^{n} \left(\frac{1}{\alpha} \left\{ \log \left(10^{\alpha x_i} + 10^{-\alpha x_i} \right) \right\} \right)^q - \sum_{i=1}^{n} \max\{x_i, -x_i\}^q$$

$$\leq \sum_{i=1}^{n} \left(\frac{1}{\alpha} \left\{ \log \left(10^{\alpha x_i} + 10^{-\alpha x_i} \right) \right\} - \max\{x_i, -x_i\} \right)^q$$

$$\leq \sum_{i=1}^{n} \left(\frac{\log 2}{\alpha} \right)^q$$

$$= n \left(\frac{\log 2}{\alpha} \right)^q,$$

and according to Equation (11), we can obtain

$$\Gamma(x) \leq \Gamma(x, \alpha) \leq \Gamma(x) + n \left(\frac{\log 2}{\alpha} \right)^q. \tag{A2}$$

Moreover, according to Property 1, $\Gamma(x, \alpha)$ is an axisymmetric function. For any $x_i \leq 0$, the result of Equation (A2) is still obtained. Therefore, the Theorem 1 is proved.

APPENDIX B. PROOF OF THEOREM 2

Proof *Since $\{x_{\beta t}\} \subseteq \{x | L_q(x, \lambda, \alpha) \leq L_q(x^0, \lambda, \alpha)\}$, we always have $L_q(x_{\beta t}, \lambda, \alpha) \leq L_q(x^0, \lambda, \alpha)$ for any t. It can be obtained that the sequence $\{x_{\beta t}\}$ is bounded. In fact, suppose that $\{x_{\beta t}\}$ is unbounded, then for any given $q \in (0, 1)$, $\{\|x_{\beta t}\|_q^q\}$ is unbounded. Thus, by combining $L_q(x, \lambda, \alpha)$, we have that for any given $q \in (0, 1)$ and $\lambda > 0$, $L_q(x_{\beta t}, \lambda, \alpha) \to +\infty$ when $t \to \infty$, which contradicts the result that $L_q(x_{\beta t}, \lambda, \alpha) \leq L_q(x^0, \lambda, \alpha)$ for all t.*

Thus, there exists at least a convergent subsequence of $\{x_{\beta t}\}$. Let $\tilde{x}$ be an accumulation point of $\{x_{\beta t}\}$, and denote $\lim_{t \to \infty} \{x_{\beta t}\} = \tilde{x}$ without loss of generality. Here we denote $\hat{x}_{\beta t} = \mathrm{diag}(x_{\beta t})$ and $\hat{\tilde{x}} = \mathrm{diag}(\tilde{x})$.

(i) We can see that $x_{\beta t}$ satisfies the first-order necessary condition of $L_q(x, \alpha, \lambda)$. Thus, by Equation (17), we have

$$\nabla L_q(x_{\beta t}, \alpha, \lambda) = \lambda \Gamma'(x_{\beta t}, \alpha) + \mathbf{A}^T (\mathbf{A} x_{\beta t} - \mathbf{y}) = 0.$$

where

$$\Gamma'(x_{\beta t}, \alpha) = \left(\frac{\partial \Gamma(x_{\beta t}, \alpha)}{\partial (x_{\beta t})_1}, \ldots, \frac{\partial \Gamma(x_{\beta t}, \alpha)}{\partial (x_{\beta t})_n} \right)^T,$$

and for any $i = 1, \ldots, n$,

$$\frac{\partial \Gamma(x_{\beta t}, \alpha)}{\partial (x_{\beta t})_i} = q \tau^{q-1}((x_{\beta t})_i, \alpha) \tau'((x_{\beta t})_i, \alpha).$$

Moreover,

$$\hat{x}_{\beta t} \nabla L_q(x_{\beta t}, \alpha, \lambda) = \lambda \hat{x}_{\beta t} \Gamma'(x_{\beta t}, \alpha) + \hat{x}_{\beta t} \mathbf{A}^T (\mathbf{A} x_{\beta t} - \mathbf{y}) = 0,$$

and

$$\left[\hat{x}_{\beta t} \Gamma'(x_{\beta t}, \alpha) \right]_i = q(x_{\beta t})_i \tau^{q-1}((x_{\beta t})_i, \alpha) \tau'((x_{\beta t})_i, \alpha)$$
$$\to q \tilde{x}_i \tau^{q-1}(\tilde{x}_i) \mathrm{sign}(\tilde{x}_i)$$
$$= q|\tilde{x}_i|^q,$$

where the second step follows from Equation (8) and Proposition 1. Therefore,

$$0 = \lim_{t \to \infty} \lambda \hat{x}_{\beta t} \Gamma'(x_{\beta t}, \alpha) + \lim_{t \to \infty} \hat{x}_{\beta t} \mathbf{A}^T (\mathbf{A} x_{\beta t} - \mathbf{y}) = \lambda q |\tilde{x}|^q + \hat{\tilde{x}} \mathbf{A}^T (\mathbf{A} \tilde{x} - \mathbf{y}).$$

i.e., $\tilde{x}$ satisfies Equation (18), which implies that $\tilde{x}$ is a stationary point of $L_q(x, \lambda)$.

(ii) Let x^* be a global minimizer of $L_q(x^0, \lambda)$, then, we have

$$L_q(x^*, \lambda) \le L_q(x_{\beta t}, \lambda) \le L_q(x_{\beta t}, \alpha, \lambda) \le L_q(x^*, \alpha, \lambda) \le L_q(x^*, \lambda) + \lambda n (\beta \log 2)^q.$$

Take $t \to \infty$, we obtain that $L_q(\tilde{x}, \lambda) = L_q(x^*, \alpha)$. Hence, any accumulation point of $\{x_{\beta t}\}$ is a global minimizer of $L_q(x, \lambda)$. So the proof is complete. According to the Theorem, we can see that the convergence of $L_q(x, \alpha, \lambda)$ and $L_q(x, \lambda)$ is equivalent. Hence, our proposed algorithm can converge to global minimizer of $L_q(x_{\beta t}, \lambda)$. So, the convergence is simply analyzed.

REFERENCES

1. Donoho, D. L., "Compressed sensing," *IEEE Transactions on Information Theory*, Vol. 4, 1289–1306, 2006.

2. Baraniuk, R. G., E. Candes, R. Nowak, and M. Vetterli, "Compressive sampling," *IEEE Signal Processing Magazine*, Vol. 25, 12–13, 2008.

3. Candès, E. J. and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, Vol. 2, 21–30, 2008.

4. Huang, S. and T. D. Tran, "Sparse signal recovery via generalized entropy functions minimization," *Eprint Arxiv*, 2017.

5. Candès, E. J., "The restricted isometry property and its implications for compressed sensing," *Comptes rendus-Mathématique*, Vol. 346, 589–592, 2008.

6. Natarajan, B. K., "Sparse approximate solutions to linear systems," *SIAM Journal on Computing*, Vol. 2, 227–234, 1995.

7. Wang, J., S. Kwon, P. Li, and B. Shim, "Recovery of sparse signals via generalized orthogonal matching pursuit: A new analysis," *IEEE Transactions on Signal Processing*, Vol. 64, 1076–1089, 2016.

8. Lee, J., G. T. Gil, and H. L. Yong, "Channel estimation via orthogonal matching pursuit for hybrid mimo systems in millimeter wave communications," *IEEE Transactions on Communications*, Vol. 64, 2370–2386, 2016.

9. Wen, J., Z. Zhou, J. Wang, X. Tang, and Q. Mo, "A sharp condition for exact support recovery with orthogonal matching pursuit," *IEEE Transactions on Signal Processing*, Vol. 6, 1370–1382, 2017.

10. Liu, J., C. Zhang, and C. Pan, "Priori-information hold subspace pursuit: A compressive sensing-based channel estimation for layer modulated TDS-OFDM," *IEEE Trans. Broadcast.*, Vol. 99, 1–9, 2018.

11. Foucart, S. and M. J. Lai, "Sparsest solutions of underdetermined linear systems via L_p-minimization for $0 < p \le 1$, *Applied and Computational Harmonic Analysis*, Vol. 3, 395–407, 2009.

12. Hurley, N. and S. Rickard, "Comparing measures of sparsity," *IEEE Transactions on Information Theory*, Vol. 10, 4723–4741, 2009.

13. Kose, K., O. Gunay, and A. E. Ceti, "Compressive sensing using the modified entropy functional," *Digital Signal Processing*, Vol. 24, 63–70, 2014.

14. 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, 586–597, 2008.

15. Qiao, B., X. Zhang, C. Wang, H. Zhang, and X. Chen, "Sparse regularization for force identification using dictionaries," *J. Sound Vib.*, Vol. 368, 71–86, 2016.

16. Wang, Y., S. Xie, and Z. Xie, "Fista-based papr reduction method for tone reservation's OFDM system," *IEEE Wireless Communications Letters*, Vol. 7, No. 3, 300–303, 2017.

17. Metzler, C. A., A. Maleki, and R. G. Baraniuk, "From denoising to compressed sensing," *IEEE Transactions on Information Theory*, Vol. 62, 5117–5144, 2016.

18. Metzler, C. A., A. Mousavi, and R. G. Baraniuk, "Learned D-AMP: Principled neural network based compressive image recovery," *Advances in Neural Information Processing Systems*, 2017.

19. Chen, L. and Y. Gu, "On the null space constant for L_p minimization," *IEEE Signal Processing Letters*, Vol. 10, 1600–1603, 2015.

20. Wang, J. and B. Shim, "A simple proof of the mutual incoherence condition for orthogonal matching pursuit," *Mathematics*, 2011.

21. Barnett, A. G., J. Beyersmann, and A. Allignol, "The time-dependent bias and its effect on extra length of stay due to nosocomial infection," *Value in Health*, Vol. 2, 381–386, 2011.

22. Bolyog, B. and G. Pap, "On conditional least squares estimation for affine diffusions based on continuous time observations," *Statistical Inference for Stochastic Processes*, Vol. 4, 1–35, 2017.

23. Abgrall, R., D. Amsallem, and R. Crisonovan, "Robust model reduction of hyperbolic problems by L_1-norm minimization and dictionary approximation," *Advanced Modeling and Simulation in Engineering Sciences*, Vol. 1, 1, 2016.

24. Zhao, Y., Z. Liu, and Y. Wang, "Sparse coding algorithm with negentropy and weighted L_1-norm for signal reconstruction," *Entropy*, Vol. 1, 599, 2017.

25. Zheng, L., A. Maleki, H. Weng, X. Wang, and T. Long, "Does ℓ_p-minimization outperform ℓ_1-minimization," *IEEE Transactions on Information Theory*, Vol. 63, 6896–6935, 2017.

26. Saab, R., R. Chartrand, and O. Yilmaz, "Stable sparse approximations via nonconvex optimization," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 3885–3888, 2008.

27. Pant, J. K., W. Lu, and A. Antoniou, "Unconstrained regularized L_p norm based algorithm for the reconstruction of sparse signals," *IEEE International Symposium on Circuits & Systems*, 1740–1743, 2011.

28. Pant, J. K., W. Lu, and A. Antoniou," New improved algorithms for compressive sensing based on L_p-norm," *IEEE Trans. Circuits and Systems II: Express Briefs*, Vol. 3, 198–202, 2014.

29. Boyd, S., N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, Vol. 1, 1–122, 2011.

30. Xie, Q., C. Ma, and C. Guo, "Image fusion based on the $\triangle^1$-TV energy function," *Entropy*, Vol. 16, 6099–6115, 2014.

31. Du, S. and M. Chen, "A new smoothing modified three-term conjugate gradient method for L_1-norm minimization problem," *Journal of Inequalities and Applications*, Vol. 1, 105, 2018.

32. Lai, M. J., Y. Xu, and W. Yin, "Improved iteratively reweighted least squares for unconstrained smoothed L_q minimization," *SIAM Journal on Numerical Analysis*, Vol. 51, 927–957, 2013.

33. Li, Q., S. Y. Liang, and Q. Li, "Incipient fault diagnosis of rolling bearings based on impulse-step impact dictionary and re-weighted minimizing nonconvex penalty L_q regular technique," *Entropy*, Vol. 19, 2017.

34. Wipf, D. and S. Nagarajan, "Iterative reweighted L_1 and L_2 methods for finding sparse solutions," *IEEE Journal of Selected Topics in Signal Processing*, Vol. 4, 317–329, 2013.

35. Ye, X., W. Zhu, and A. Zhang, "Sparse channel estimation of MIMO-OFDM systems with unconstrained smoothed L0-norm-regularized least squares compressed sensing," *EURASIP Journal on Wireless Communications and Networking*, Vol. 1, 282, 2013.

36. Arias-Castro, E. and Y. C. Eldar, "Noise folding in compressed sensing," *IEEE Signal Processing Letters*, Vol. 18, 478–481, 2011.

37. Yang, X., Q. Cui, E. Dutkiewicz, and X. Huang, "Anti-noise-folding regularized subspace pursuit recovery algorithm for noisy sparse signals," *Proceeding of the IEEE Wireless Communications and Networking Conference*, 275–280, Istanbul, Turkey, April 6–9, 2014.

38. Lu, Z., "Iterative reweighted minimization methods for l_p regularized unconstrained nonlinear programming," *Mathematical Programming*, Vol. 147, 277–307, 2014.

39. Mourad, N., J. P. Reilly, and T. Kirubarajan, "Majorization-minimization for blind source separation of sparse sources," *Signal Processing*, Vol. 131, 120–133, 2017.

40. Oh, J. and N. Kwak, "Generalized mean for robust principal component analysis," *Pattern Recognition*, Vol. 54, 116–127, 2016.

41. Zhou, W., Y. B. Sun, Q. S. Liu, and W. U. Min, "L_0 group sparse RPCA model and algorithm for moving object detection," *Acta Electronica Sinica*, Vol. 44, No. 3, 627–632, 2016.

42. Xu, D., X. Gao, X. Fan, D. Zhao, and W. Gao, "ODD: An algorithm of online directional dictionary learning for sparse representation," *Proceedings of the Pacific Rim Conference on Multimedia*, Vol. 10736, 939–947, Harbin, China, September 28–29, 2017.