

Signal Reconstruction of Compressed Sensing Based on Alternating Direction Method of Multipliers

Yanliang Zhang¹ · Xingwang Li¹✉ · Guoying Zhao² · Bing Lu¹ · Charles C. Cavalcante³

Abstract The sparse signal reconstruction of compressive sensing can be accomplished by l_1 -norm minimization, but in many existing algorithms, there are the problems of low success probability and high computational complexity. To overcome these problems, an algorithm based on the alternating direction method of multipliers is proposed. First, using variable splitting techniques, an additional variable is introduced, which is tied to the original variable via an affine constraint. Then, the problem is transformed into a non-constrained optimization problem by means of the Augmented Lagrangian Multiplier method, where the multipliers can be obtained using the gradient ascent method according to dual optimization theory. The l_1 -norm

This work was supported by Ph.D Fund of Henan Polytechnic University with Grant Number B2012-100 and Open Fund of Network and Exchange Technology State Key Laboratory with Grant Number SKLNST-2016-1-02

✉ Xingwang Li
lixingwangbupt@gmail.com

Yanliang Zhang
ylzhang1199@foxmail.com

Guoying Zhao
guoying.zhao@oulu.fi

Bing Lu
13393860892@163.com

Charles C. Cavalcante
charles@gtel.ufc.br

¹ School of Physics and Electrical Information Engineering, Henan Polytechnic University, Jiaozuo 454001, China;

² Center for Machine Vision and Signal Analysis, University of Oulu, Oulu 90014, Finland;

³ Wireless Telecommunications Research Group, Federal University of Ceará Campus do Pici, Bl. 722, ZIP 60455-760, Fortaleza-CE, Brazil .

minimization can finally be solved by cyclic iteration with concise form, where the solution of the original variable could be obtained by a projection operator, and the auxiliary variable could be solved by a soft threshold operator. Simulation results show that a higher signal reconstruction success probability is obtained when compared to existing methods, while a low computational cost is required.

Keywords compressive sensing (CS) · signal reconstruction · l_1 -norm minimization · alternating direction method of multipliers · dual optimization

1 Introduction

In early 2000, several authors reported important advances on the signal compression theory, otherwise referred to as compressive sensing (CS) [6, 7, 10], which breaks through the Shannon sampling theorem on the sampling frequency. According to the theory, if a signal is sparse in an orthogonal space, it can be projected onto a low-dimensional space with an observation matrix which is not related to the orthogonal basis. By solving an optimization problem, the original signal can be reconstructed with high probability, from this small number of projection samples.

1.1 l_0 -norm Minimization

One of the core questions of CS is how to reconstruct sparse signals from finite measurements [1, 8, 31]. This issue has been reformulated as the l_0 -norm minimization problem and it is a classic problem in the signal and information community. Mallat et al. proposed the Matching Pursuit (MP) approach to approximately solve the l_0 -norm minimization problem [18]. This approach involves finding the best matching projections of multidimensional data onto the span of an over-complete dictionary. However, the main problem of MP is the computational complexity of the encoder. In the basic version of the algorithm, the large dictionary has to be searched at each iteration. Improvements include the use of approximate dictionary representations and suboptimal ways of choosing the best match at each iteration [17]. A popular extension of MP is its orthogonal version, which is called orthogonal matching pursuit (OMP) [23]. The main difference between MP and OMP is that after every step, all of the coefficients extracted thus far are updated, by computing the orthogonal projection of the signal onto the set of atoms selected so far. These algorithms are greedy methods with low computational complexity. However, their success probability of reconstruction is lower and often require more measurements when compared with standard MP.

Other extensions, such as multichannel MP [15] and multichannel OMP [24], enable the processing of multicomponent signals. An obvious extension of MP is over multiple positions and scales, by augmenting the dictionary to be that of a wavelet basis. This can be performed efficiently using the convolution operator without changing the core algorithm [23].

The l_0 -norm minimization is a discrete optimization NP-hard problem which is very difficult to solve. In order to overcome this issue, a common way is to convert it

into a convex optimization problem. It is known that l_1 -norm is the convex function which is the closest to the l_0 -norm. Hence, the sparse signal could be approximated using the l_1 -norm minimization method.

1.2 l_1 -norm Minimization

The l_1 -norm minimization can be cast as a linear programming problem and many commercially off-the-shelf optimization tools are available to solve them. Such a problem is of great interest to the numerical optimization community and there are many practical modern convex optimization strategies [4] that are used to solve them. Interior-point methods [5], homotopy methods [19], and other techniques have been used to address the problem of l_1 -norm minimization.

Primal-dual interior-point algorithm (PDIPA) [29] is an interior-point method for the l_1 -norm minimization problem. The framework of PDIPA is to iteratively formulate the inequality constrained problem as an equality constrained problem, which can be solved by using the barrier method based on Newton's algorithm. Its computational complexity is dominated by the iteration of the Newton update, which is of the order of about $O(n^3)$, where n represents the dimensions of the original signal. Therefore, it is very sensitive to the size of the problem n and is not suitable for high dimension and large scale applications. The Homotopy algorithm [19] also contains a Newton update step. PDIPA, it only involves some of the coefficients and the complexity is bounded by order $O(km^2 + kmn)$, where m denotes the dimensions of the observation signal and k is the sparsity level of the original signal. However, the computational complexity is still bounded by $O(n^3)$, when the sparsity level k and the observation dimension m grows proportionally with the signal dimension n .

Owing to its low computational complexity, the first-order method, which makes use of the sub-differential structure of the l_1 -norm, is of concern to researchers. The computational complexity for each iteration of the first-order methods is greatly reduced, and is achieved at the expense of an increased number of iterations when compared with the interior-point methods. Each of its iteration involves only simple operations such as vector addition, matrix-vector multiplication or soft threshold operator, which have low computational cost. The convergence speed of the representative algorithm fast iterative soft-thresholding (FISTA) is $O(1/k^2)$ [29], while the success probability of reconstruction needs to be improved.

1.3 ADMM Based Methods

Owing to the rapid increase in the size and complexity of modern datasets, it is increasingly important to be able to solve the problem of CS with low computational complexity and high success probability. The augmented Lagrangian multiplier (ALM) [4] is another special class of first-order methods, with superior convergence properties. The cyclic minimizer (CM) [22] was a numerical optimization paradigm which partitions the variable vector of the objective function into several sub-vectors. Then, the function is alternately minimized with respect to one sub-vector with other

sub-vectors fixed. By integrating ALM and CM, the alternating direction method of multiplier (ADMM) was summarized by Boyd in [3]. ADMM is a simple but powerful framework in numerical optimization, which is suited for sparse signal recovery in CS [14, 21, 25–28]. Based on ADMM, Junfeng Yang *et al.* [30] introduced a called alternating direction method (ADM) approach to the problem of l_1 -norm minimization. Independently, Allen Y. Yang *et al.* [29] proposed two algorithms called PALM and DALM to address the l_1 -norm minimization problem. Essentially, PALM has a similar form to ADM.

1.4 Motivation and Outline

To solve the problem of sparse signal recovery in CS, the main technical challenge is that the sparse signal vector is coupled using a sensing matrix. This makes the recovery of sparse signal somewhat difficult when extra constraint with non-smooth l_1 -norm is introduced. In principle, if the variables can be decoupled, the problem could be solved element-wise or group-wise and the sparse signal may become easier to recover. In ADM or PALM, variable-splitting is achieved by introducing an auxiliary variable, which is defined as the residual error of the data fitting model. The result of this scheme of introducing variables is that each loop of original variables being updated has an imprecise solution. Moreover, it would require more runtime to converge, and would have a lower success probability of sparse signal reconstruction.

In this paper, we propose an efficient algorithm to reconstruct sparse signals in CS by using the ADMM. Compared with ADM or PALM, we introduce an additional variable, which is tied to the original variable via an affine constraint. In this framework, the solver of three cyclic iterations has concise form. The solution of the original variable could be obtained by using the projection operator, and the auxiliary variable could be solved by using the soft threshold operator. The simulation results show that the higher success probability of signal reconstruction is obtained than existing methods, while low computational cost is required.

The remaining parts of this paper are organized as follows: The problem of sparse signal recovery in CS and the ADM algorithm are discussed in Section II. In section III, we describe our new method to the signal reconstruction of CS based on ADMM. Then, Section IV presents our experimental results and discussion. Finally, conclusions are presented in Section V.

2 Background and Related Work

2.1 Sparse Signal Recovery in CS

CS is a signal processing technique which is employed to efficiently acquire and reconstruct a signal by finding solutions to underdetermined linear systems [10]. This is based on the principle that, through optimization, the sparsity of a signal can be exploited to recover it from far fewer samples than what is required by the Shannon-Nyquist sampling theorem [10]. There are two conditions under which recovery is

possible [13]. The first one is the *sparsity*, which requires the signal to be sparse in some domain. Then, the second one is *incoherence* which is applied using the isometric property, which is sufficient for sparse signals.

The signal $\mathbf{e} \in R^n$ is k -sparsity on the orthogonal basis $\Phi \in R^{n \times n}$, that is $\mathbf{e} = \Phi \mathbf{x}$, $\|\mathbf{x}\|_0 = k \ll n$. Observations are then written as $\mathbf{b} = \Psi \mathbf{e} = \Psi \Phi \mathbf{x}$, where measurement matrix $\Psi \in R^{m \times n}$, $n > m$. According to the theory of CS, the Restricted Isometry Property (RIP) condition of sensing matrix \mathbf{A} , where $\mathbf{A} = \Psi \Phi$, is sufficient to guarantee the reconstruction of the sparse signal of \mathbf{x} . The problem of estimating \mathbf{x} can be formulated as

$$\min \|\mathbf{x}\|_0 \text{ s.t. } \mathbf{b} = \mathbf{A}\mathbf{x} \quad (1)$$

The above formula is also called the l_0 -norm minimization problem.

For the number of rows in the matrix \mathbf{A} less than columns, (1) is underdetermined. The global optimal solution can be found by considering the C_n^k different arrangements of non-zero elements of \mathbf{x} using the exhaustive method. Therefore, it is an NP-hard problem [1, 10, 11]. To render (1) as more tractable, Donoho et al. regularized the highly discontinuous l_0 -norm and replaced it by l_1 -norm [12]

$$\min \|\mathbf{x}\|_1 \text{ s.t. } \mathbf{b} = \mathbf{A}\mathbf{x} \quad (2)$$

Because \mathbf{b} often contains nuisance noise, the equality constraint can be relaxed with $\|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2 \leq \delta$, where $\delta > 0$ represents a pre-determined noise level.

$$\min \|\mathbf{x}\|_1 \text{ s.t. } \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2 \leq \delta \quad (3)$$

For an appropriate scalar weight λ , we can obtain the following variant of the (3):

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2 + \lambda \|\mathbf{x}\|_1 \quad (4)$$

In this paper, we refer to (2), (3) and (4) as the l_1 -norm minimization problem. If \mathbf{x} is sufficiently sparse and the measurement matrix Ψ is incoherent with orthogonal basis Φ , then \mathbf{x} can be reconstructed by solving the l_1 -norm minimization problem.

2.2 Alternating Direction Method for Sparse Signal Reconstruction

This section discusses how the algorithms of ADM [29] or PALM [30] solve the l_1 -norm minimization problem.

In these algorithms, the auxiliary variable $\mathbf{r} \in R^m$ is defined as the residuals error of the data fitting model, that is $\mathbf{r} = \mathbf{b} - \mathbf{A}\mathbf{x}$. Then the problem of (2) could be converted into the form of (5)

$$\min_{\mathbf{r}, \mathbf{x}} \frac{1}{2\tau} \|\mathbf{r}\|_2 + \|\mathbf{x}\|_1 \text{ s.t. } \mathbf{r} = \mathbf{b} - \mathbf{A}\mathbf{x} \quad (5)$$

Using the ALM method, (5) could be transformed into below unconstrained problem

$$\min_{\mathbf{r}, \mathbf{x}, \mathbf{y}} L(\mathbf{r}, \mathbf{x}, \mathbf{y}) = \frac{1}{2\tau} \|\mathbf{r}\|_2 + \|\mathbf{x}\|_1 + \mathbf{y}^T (\mathbf{r} + \mathbf{A}\mathbf{x} - \mathbf{b}) + \frac{\mu}{2} \|\mathbf{r} + \mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 \quad (6)$$

where $\mathbf{y} \in R^m$ is a Lagrangian multiplier vector, and $\mu > 0$ is the penalty parameter. From the paradigm of the alternative direction method, (6) could be converted into three sub-problems as follows.

$$\begin{cases} \mathbf{r}_{k+1} = \arg \min L(\mathbf{r}, \mathbf{x}_k, \mathbf{y}_k) & (a) \\ \mathbf{x}_{k+1} = \arg \min L(\mathbf{r}_{k+1}, \mathbf{x}, \mathbf{y}_k) & (b) \\ \mathbf{y}_{k+1} = \mathbf{y}_k - \mu (\mathbf{r}_{k+1} + \mathbf{A}\mathbf{x}_{k+1} - \mathbf{b}) & (c) \end{cases} \quad (7)$$

Sub-problem(7.a) is investigated as follows.

$$\arg \min L(\mathbf{r}, \mathbf{x}_k, \mathbf{y}_k) = \frac{1}{2\tau} \|\mathbf{r}\|_2 - \mathbf{y}_k^T \mathbf{r} + \frac{\mu}{2} \|\mathbf{r} + \mathbf{A}\mathbf{x}_k - \mathbf{b}\|_2^2 + \|\mathbf{x}_k\|_1 - \mathbf{y}_k^T (\mathbf{A}\mathbf{x}_k - \mathbf{b}) \quad (8)$$

The solution of (7.a) is given by

$$\mathbf{r}_{k+1} = \frac{\tau}{\mathbf{1} + \mu\tau} [\mathbf{y}_k - \mu (\mathbf{b} - \mathbf{A}\mathbf{x}_k)] \quad (9)$$

Then, the optimization problem (7.b) is considered

$$\begin{aligned} \arg \min L(\mathbf{r}_{k+1}, \mathbf{x}, \mathbf{y}_k) &= \arg \min \left\{ \|\mathbf{x}\|_1 - \mathbf{y}_k^T (\mathbf{r}_{k+1} + \mathbf{A}\mathbf{x} - \mathbf{b}) + \right. \\ &\quad \left. \frac{\mu}{2} \|\mathbf{r}_{k+1} + \mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 \right\} \\ &= \arg \min [\|\mathbf{x}\|_1 + j(\mathbf{x})] \end{aligned} \quad (10)$$

where $j(\mathbf{x}) = \frac{\mu}{2} \|\mathbf{r}_{k+1} + \mathbf{A}\mathbf{x} - \mathbf{b} - \mathbf{y}_k/\mu\|_2^2$. Use the second order Taylor expansion to approximate $j(\mathbf{x})$ and get

$$\begin{aligned} \arg \min L(\mathbf{r}_{k+1}, \mathbf{x}, \mathbf{y}_k) &\approx \arg \min \left\{ \|\mathbf{x}\|_1 \right. \\ &\quad \left. + (\mathbf{x} - \mathbf{x}_k)^T \mathbf{A}^T (\mathbf{r}_{k+1} + \mathbf{A}\mathbf{x}_k - \mathbf{b} - \mathbf{y}_k/\mu) + \frac{1}{2\tau} \|\mathbf{x} - \mathbf{x}_k\|_2^2 \right\} \end{aligned} \quad (11)$$

The solution of (11) can be obtained by the soft thresholding operator

$$\mathbf{x}_{k+1} = S_{\tau/\mu} \left\{ \mathbf{x}_k - \tau \mathbf{A}^T (\mathbf{r}_{k+1} + \mathbf{A}\mathbf{x}_k - \mathbf{b} - \mathbf{y}_k/\mu) \right\} \quad (12)$$

$S_{\tau/\mu}$ is the soft threshold operator [9], which acts on each component of the vector $\left\{ \mathbf{x}_k - \tau \mathbf{A}^T (\mathbf{r}_{k+1} + \mathbf{A}\mathbf{x}_k - \mathbf{b} - \mathbf{y}_k/\mu) \right\}$.

The definition of the soft threshold operator is

$$S_c(a) = \text{sgn}(a) \max(|a| - c, 0) \quad (13)$$

where sgn is the sign function which is defined as

$$\text{sgn}(a) = \begin{cases} -1 & \text{if } a < 0 \\ 0 & \text{if } a = 0 \\ 1 & \text{if } a > 0 \end{cases} \quad (14)$$

From the above discussion, we can see that the ADM addresses the l_1 -norm minimization problem(5) by cyclic iteration (9), 12 and 7.c.

Because the algorithm presented above utilizes the second order Taylor expansion to approximately solve the sub-problem (7.b), it is denoted as an inexact ADM. Moreover, an ADM can also efficiently address the dual problem of (6) and obtains the DALM algorithm.

3 Signal Reconstruction Based on Alternating Direction Method of Multipliers

3.1 Problem Reformulation

ADMM is a simple but powerful algorithm which is well suited to distributed convex optimization [3]. It blends the decomposability of dual ascent and fast convergence of the augmented Lagrangian multipliers methods and takes the form of a decomposition-coordination procedure to address a given optimization problem.

It can solve the problem in the following form:

$$\min l(\mathbf{u}) + g(\mathbf{v}) \text{ s.t. } \mathbf{B}\mathbf{u} + \mathbf{C}\mathbf{v} = \mathbf{d} \quad (15)$$

where $l(\cdot) : R^r \rightarrow R, g(\cdot) : R^l \rightarrow R$ are convex functions and $\mathbf{u} \in R^r, \mathbf{v} \in R^l, \mathbf{B} \in R^{p \times r}, \mathbf{C} \in R^{p \times l}$.

Using the variable splitting technique, we introduce a new variable \mathbf{z} , which is tied to the original variable \mathbf{x} . Then (2) can be transformed into

$$\min f(\mathbf{x}) + \|\mathbf{z}\|_1 \text{ s.t. } \mathbf{x} - \mathbf{z} = 0 \quad (16)$$

where $f(\cdot)$ is the indicator function of the non-empty convex hull $\Theta = \{\mathbf{x} \in R^n | \mathbf{A}\mathbf{x} = \mathbf{b}\}$. The definition of $f(\cdot)$ is given as

$$f(\mathbf{x}) = \begin{cases} 0 & \mathbf{x} \in \Theta \\ +\infty & \mathbf{x} \notin \Theta \end{cases} \quad (17)$$

Obviously, $f(\cdot)$ is a convex function.

In this way, Eq.(16) has the same form as (15), which is the general formula of the ADMM.

3.2 Augmented Lagrangian Multiplier Method and Dual Ascend

Our description of the method here is based on the description provided in [4].The augmented Lagrangian for Eq. (16) is given as

$$L_\rho(\mathbf{x}, \mathbf{z}, \mathbf{y}) = f(\mathbf{x}) + \|\mathbf{z}\|_1 + \mathbf{y}^T(\mathbf{x} - \mathbf{z}) + \frac{1}{2}\rho \|\mathbf{x} - \mathbf{z}\|_2^2 \quad (18)$$

where \mathbf{y} is the Lagrangian multiplier or dual variable, and $\rho > 0$ is called the penalty parameter. The dual function of (18) is

$$g_\rho(\mathbf{y}) = \inf_{\mathbf{x}, \mathbf{z}} L_\rho(\mathbf{x}, \mathbf{z}, \mathbf{y}) \quad (19)$$

and the dual problem for (16) is

$$\max_{\mathbf{y}} g_\rho(\mathbf{y}) \quad (20)$$

According to the theory of dual optimization [3], the optimal solution of the original problem (16) is the same as that of the dual problem in (10). That is, $\mathbf{y}^* = \arg \max_{\mathbf{y}} g_\rho(\mathbf{y})$. Therefore, the optimal solution of (18) i.e. $(\mathbf{x}^*, \mathbf{z}^*)$ can be obtained by

$$(\mathbf{x}^*, \mathbf{z}^*) = \arg \min_{\mathbf{x}, \mathbf{z}} L_\rho(\mathbf{x}, \mathbf{z}, \mathbf{y}^*) \quad (21)$$

The optimization of (20) can be addressed by the gradient ascent method, and the gradient is given by $\nabla g_\rho(\mathbf{y}) = \mathbf{x} - \mathbf{z}$. Thus, the minimization of (18) can be solved by the following iteration process:

$$(\mathbf{x}_{k+1}, \mathbf{z}_{k+1}) = \arg \min_{\mathbf{x}, \mathbf{z}} L_\rho(\mathbf{x}, \mathbf{z}, \mathbf{y}_k) \quad (22)$$

$$\mathbf{y}_{k+1} = \mathbf{y}_k + \rho(\mathbf{x}_{k+1} - \mathbf{z}_{k+1}) \quad (23)$$

The dual variable update uses a step size which is equal to the augmented Lagrangian parameter ρ [3].

3.3 Alternating Minimization

To solve problem (22), we utilise the method of alternating minimization and Gauss-Seidel iteration [9], which alternates the update of \mathbf{x} or \mathbf{z} , while considering the other variable as a constant. The following two formulas are obtained:

$$\mathbf{x}_{k+1} = \arg \min_{\mathbf{x}} L_\rho(\mathbf{x}, \mathbf{z}_k, \mathbf{y}_k) \quad (24)$$

$$\mathbf{z}_{k+1} = \arg \min_{\mathbf{z}} L_\rho(\mathbf{x}_{k+1}, \mathbf{z}, \mathbf{y}_k) \quad (25)$$

In order to find the analytical expression of (24) and (25), we reformulate formula (18) as following.

$$\begin{aligned} \left\| (\mathbf{x} - \mathbf{z}) + \frac{1}{\rho} \mathbf{y} \right\|_2^2 &= \left\langle (\mathbf{x} - \mathbf{z}) + \frac{1}{\rho} \mathbf{y}, (\mathbf{x} - \mathbf{z}) + \frac{1}{\rho} \mathbf{y} \right\rangle \\ &= \langle \mathbf{x} - \mathbf{z}, \mathbf{x} - \mathbf{z} \rangle + 2 \left\langle \mathbf{x} - \mathbf{z}, \frac{1}{\rho} \mathbf{y} \right\rangle + \left\langle \frac{1}{\rho} \mathbf{y}, \frac{1}{\rho} \mathbf{y} \right\rangle \\ &= \|\mathbf{x} - \mathbf{z}\|_2^2 + \left\| \frac{1}{\rho} \mathbf{y} \right\|_2^2 + \frac{2}{\rho} \mathbf{y}^T (\mathbf{x} - \mathbf{z}) \end{aligned}$$

From the above, we can derive that $\mathbf{y}^T (\mathbf{x} - \mathbf{z}) + \frac{1}{2} \rho \|\mathbf{x} - \mathbf{z}\|_2^2 = \frac{\rho}{2} \left\| \mathbf{x} - \mathbf{z} + \frac{1}{\rho} \mathbf{y} \right\|_2^2 - \frac{\rho}{2} \left\| \frac{1}{\rho} \mathbf{y} \right\|_2^2$. Defining $\mathbf{q} = \frac{1}{\rho} \mathbf{y}$, formula (18) can then be reformulated as:

$$L_\rho(\mathbf{x}, \mathbf{z}, \mathbf{q}) = f(\mathbf{x}) + \|\mathbf{z}\|_1 + \frac{\rho}{2} \|\mathbf{x} - \mathbf{z} + \mathbf{q}\|_2^2 - \frac{\rho}{2} \|\mathbf{q}\|_2^2 \quad (26)$$

Hence, formulas (24), (25) and (23) can be respectively rewritten as

$$\mathbf{x}_{k+1} = \arg \min_{\mathbf{x}} \left[f(\mathbf{x}) + \frac{\rho}{2} \|\mathbf{x} - (\mathbf{z}_k - \mathbf{q}_k)\|_2^2 \right] \quad (27)$$

$$\mathbf{z}_{k+1} = \arg \min_{\mathbf{z}} \left[\|\mathbf{z}\|_1 + \frac{\rho}{2} \|\mathbf{z} - (\mathbf{x}_{k+1} + \mathbf{q}_k)\|_2^2 \right] \quad (28)$$

$$\mathbf{q}_{k+1} = \mathbf{q}_k + \mathbf{x}_{k+1} - \mathbf{z}_{k+1} \quad (29)$$

3.4 Projection Operator and Soft Threshold Operator

Considering that $f(\mathbf{x})$ in (27) is an indicator function of set Θ , the optimum selection of \mathbf{x}_{k+1} is the point which is projected onto Θ of $\mathbf{z}_k - \mathbf{q}_k$. The optimization of (28) can be addressed using the soft threshold operator. Therefore

$$\begin{cases} \mathbf{x}_{k+1} = \Gamma_{\Theta}(\mathbf{z}_k - \mathbf{q}_k) \\ \mathbf{z}_{k+1} = S_{1/\rho}(\mathbf{x}_{k+1} + \mathbf{q}_k) \\ \mathbf{q}_{k+1} = \mathbf{q}_k + \mathbf{x}_{k+1} - \mathbf{z}_{k+1} \end{cases} \quad (30)$$

where Γ_{Θ} is projection operator [3], that is

$$\Gamma_{\Theta}(\mathbf{z}_k - \mathbf{q}_k) = \left[\mathbf{I} - \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{A} \right] (\mathbf{z}_k - \mathbf{q}_k) + \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{b} \quad (31)$$

$S_{1/\rho}$ is a soft threshold operator [9, 16], which acts on each component of the vector $\mathbf{x}_{k+1} + \mathbf{q}_k$. The definition of the soft threshold operator is given in formula (13) in Section 2.2.

Thus, using the ADMM, we have transformed the problem of sparse signal reconstruction into a loop iteration containing three sub-formulae. The iterative process of Eq. (30) is repeated until the termination condition is satisfied. The result is the solution of the constraint optimization problem (16) as well as the l_1 -norm minimization problem (2).

3.5 Algorithm Flowchart and Computational Complexity Analysis

The proposed algorithm is summarized as follows:

- 1) Initialize parameters ρ and $\{\mathbf{x}_0, \mathbf{z}_0, \mathbf{q}_0\}$; Set the relative error ε and let $k = 0$
- 2) Update \mathbf{x}_{k+1} :

$$\mathbf{x}_{k+1} = \left[\mathbf{I} - \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{A} \right] (\mathbf{z}_k - \mathbf{q}_k) + \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{b}$$

- 3) Update \mathbf{z}_{k+1} :

$$\mathbf{z}_{k+1} = S_{1/\rho}(\mathbf{x}_{k+1} + \mathbf{q}_k) = \text{sgn}(\mathbf{x}_{k+1} + \mathbf{q}_k) \max(|\mathbf{x}_{k+1} + \mathbf{q}_k| - 1/\rho, 0)$$

- 4) Update \mathbf{q}_{k+1} :

$$\mathbf{q}_{k+1} = \mathbf{q}_k + \mathbf{x}_{k+1} - \mathbf{z}_{k+1}$$

- 5) Exit loop if $\frac{\|\mathbf{z}_k - \Phi\mathbf{x}_{k+1}\|}{\|\mathbf{z}_k\|_2} \leq \varepsilon$; Otherwise, let $k = k + 1$ and go to step 2).

Obviously, for the update \mathbf{x}_{k+1} in the step 2), the values of $\mathbf{I} - \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{A}$ and $\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{b}$ are the same for each iteration. Therefore, the value of the upper two formula obtained from the first iteration can be saved for the subsequent iteration. This can greatly reduce the computational burden of the algorithm.

The computational complexity of the algorithm is low. Although the computation of $\mathbf{I} - \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{A}$ and $\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{b}$ is of order $O(n^3)$, they only need to be computed once during the whole algorithm execution. The complexity of steps 2), 3) and 4), that is, the iteration of the algorithm, is $O(n)$.

4 Numerical Experiments

Both in the noise and noiseless cases, we compared the proposed algorithm, called ADMM, with the four other algorithms, namely PALM [29], DALM [29], FISTA [2], and OMP [23]. We implemented the above reconstruction algorithms to recover signal \mathbf{e} according to the observation vector \mathbf{b} .

4.1 Experiment Settings

In our experiments, Φ was set as a normalized cosine basis, and observation matrix Ψ was constructed by drawing its components from a standard normal distribution, i.e., $N(0, 1)$. The k -sparse signal \mathbf{x} of length n was constructed as follows [20, 25]:

- 1) A zero vector \mathbf{x} of length n was generated;
- 2) A random vector of length k was constructed whose components were drawn from a standardized normal distribution;
- 3) The components of the resulting vector were set to k randomly chosen locations of vector \mathbf{x} .

The sensing matrix was constructed by $\mathbf{A} = \Psi\Phi$. In all of the experiments, we generated observation vector \mathbf{b} by MATLAB scripts $\mathbf{b} = \mathbf{A} * \mathbf{x} + \text{sigma} * \text{randn}(m, 1)$, where in the noiseless case, $\text{sigma} = 0$, and in the noise case, $\text{sigma} = 10^{-4}$ [20].

The relative error of the reconstruction is defined as

$$r(\hat{\mathbf{e}}) = \frac{\|\hat{\mathbf{e}} - \mathbf{e}\|_2}{\|\mathbf{e}\|_2} \quad (32)$$

where the recovered signal $\hat{\mathbf{e}}$ is obtained by $\hat{\mathbf{e}} = \Phi\hat{\mathbf{x}}$, and $\hat{\mathbf{x}}$ is the estimation of \mathbf{x} solved by the considered algorithms. The success probability of reconstruction is defined as

$$PER = \frac{\text{Number of successful reconstructing signal}}{\text{Total number of simulations}} \quad (33)$$

In the noiseless case, the signal was deemed to be successfully reconstructed when $r(\hat{\mathbf{e}}) \leq 10^{-6}$; in the noise case, we have found that smaller values of $r(\hat{\mathbf{e}})$ do not necessarily or consistently improve the relative error results, while also increasing the required number of iterations [30]. Thus, we set $r(\hat{\mathbf{e}}) \leq 10^{-3}$.

4.2 Relation between Success Probability of Reconstruction and Sparsity of \mathbf{x} in ADMM Algorithm

Some numerical results have been added into Experiment 1 to investigate the relation between the success probability of reconstruction and the sparsity of \mathbf{x} in the ADMM algorithm, where both the noiseless and noise case are considered. The length of the sparse vector \mathbf{x} was set to $n = 256$ and the number of observations was set to $m = 26, 52, 78, 94$, and 128. Under different number of observations m , the sparsity of \mathbf{x} was set to an interval between 1 and $\text{ceil}(0.75 \cdot m)$, with an interval of 5 samples, where $\text{ceil}(0.75 \cdot m)$ implies rounding each element of $(0.75 \cdot m)$ to the nearest integer

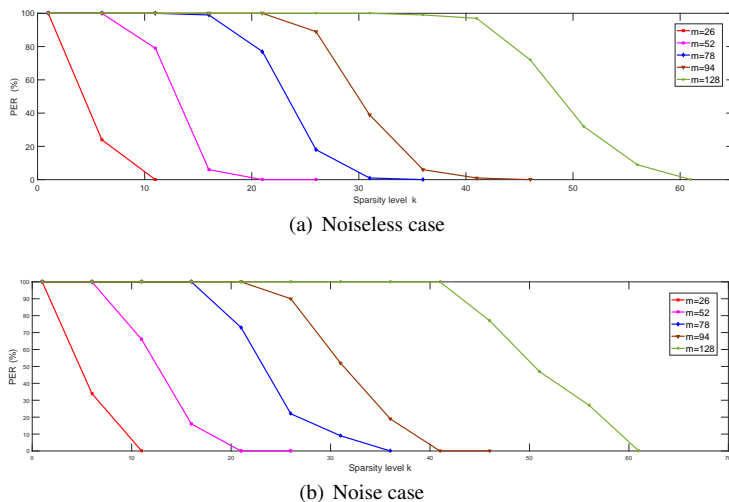


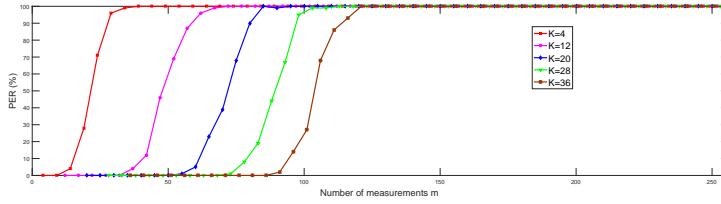
Fig. 1: Average PER versus sparsity level for ADMM algorithm (Experiment 1)

greater than or equal to that element. The quantities $\mathbf{x} \in R^n$, Φ , and Ψ were generated respectively according to the previously described methods. \mathbf{b} was generated in the noise and noiseless case. For each combination of (m, n, k) , the proposed ADMM algorithm was implemented for 1000 runs to reconstruct signals \mathbf{x} and \mathbf{e} . The success probability of reconstruction of ADMM is plotted versus sparsity in Fig.1 over 1000 trails, while m was set to different numbers.

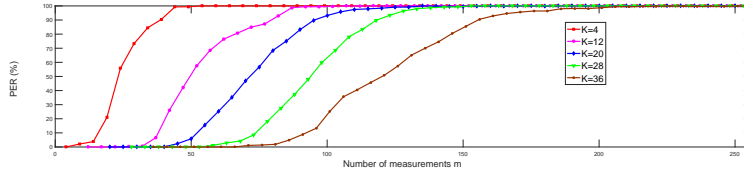
As can be seen from Fig.1, for any number m , the success probability of reconstruction PER increases with a decrease of the sparsity level k . When the sparsity level k drops to a certain value, the success probability of the algorithm increases at a faster rate. When k falls below a certain number, the success probability of reconstruction PER reaches 100%. Consider $m = 128$ in the noiseless case as an example. The success probability is 0 while $k \geq 61$. However, when k is set to $k = 56, 51, 46, 41$, the corresponding success probability values were 9%, 32%, 72% and 97%, respectively. When $k \leq 36$, the success probability reaches 100%. Experiment results show that, at any number of m , the proposed ADMM algorithm possesses good capability of sparse signal reconstruction in compressive sensing.

4.3 Relation between Success Probability of Reconstruction and Length of Observation Vector \mathbf{b} in ADMM Algorithm

In Experiment 2, the simulation was designed for both the noiseless and noise case to demonstrate the relation between the success probability of reconstruction and the length of the observation vector \mathbf{b} in the ADMM algorithm. The length of \mathbf{x} was set as in Experiment 1, and the sparsity k was set to values $k = 4, 12, 20, 28, 36$. Under various levels of sparsity k , the number of observations m was set from k to n ,



(a) Noiseless case



(b) Noise case

Fig. 2: Average PER versus number of measurements for ADMM algorithm (Experiment 2)

with intervals of 5. Under each combination of (m, n, k) , ADMM was implemented for 1000 trails to reconstruct signals \mathbf{x} and \mathbf{e} . The average success probability of reconstruction of ADMM is plotted versus the number of observations m in Fig.2. over 1000 trails, while k was set to different numbers.

As can be seen from Fig.2, for any k , the success probability of reconstruction PER rises with the increase in the number of measurements m . When m reaches a certain value, the success probability of the algorithm increases sharply. When the number of measurements m increases above a certain number, the success probability of reconstruction PER reaches 100%. As an example, consider the noiseless case, $k = 20$. The success probability is lower than 20% while $m \leq 60$. However, when m is set to $m = 65, 75, 85, 95$, the corresponding success probability values were 35%, 63%, 78% and 90%, respectively. When $m > 95$, the success probability reaches 100%. The experiment results show that for any k , the proposed ADMM algorithm possesses good capability of sparse signal reconstruction in compressed sensing.

4.4 Success Probability of Reconstruction Comparisons of ADMM and Other Algorithms

In Experiment 3, both in the noise and noiseless case, we compare the proposed ADMM algorithm with the other four algorithms in the success probability of reconstruction.

In this simulation, the length of the sparse vector n , number of observation m and the sparsity of \mathbf{x} were, respectively, set to $n = 256, m = 100$, and $k = 5a - 4, a = 1, 2, \dots, 15$, $\mathbf{x} \in R^{256}$, $\boldsymbol{\psi} \in R^{256 \times 100}$, and $\boldsymbol{\Phi} \in R^{256 \times 256}$ were respectively generated

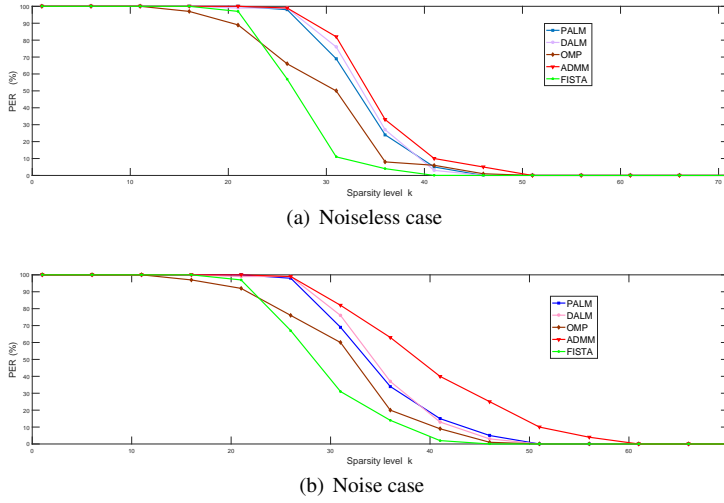


Fig. 3: Average PER versus sparsity for ADMM, PALM, DALM, FISTA and OMP algorithms (Experiment 3)

Table 1: Reconstruction success probability of three algorithms under different sparsity levels

label	$k = 56$	$k = 41$	$k = 36$	$k = 31$	$k = 26$
ADMM	5%	10%	33%	82%	99%
PALM	0	3%	24%	69%	98%
DALM	0	3%	27%	76%	99%

according to the methods already described, respectively. For each set of (m, n, k) , we performed 1000 trials to reconstruct \mathbf{x} and \mathbf{e} for each of the five algorithms, i.e., ADMM, PALM, DALM, FISTA and OMP. The success probability of reconstruction of the five algorithms is plotted versus sparsity in Fig.3 over 1000 trials.

As seen in Fig.3, for different levels of sparsity, the success probability of algorithms based on the Lagrangian multiplier method, i.e., PALM, DALM and ADMM, is clearly higher than that of FISTA and OMP. Moreover, the success probability of ADMM is the highest in PALM, DALM and ADMM. Specifically, the corresponding success probability of the noiseless case in the rising stage of the three algorithms with better performance is listed in Table 1. From the above, it can be seen that the success probability of the proposed ADMM algorithm has more obvious advantages than the other four algorithms.

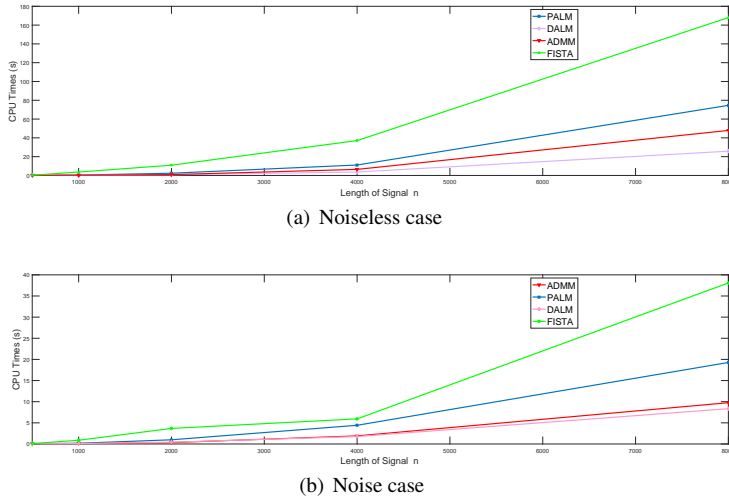


Fig. 4: Average CPU time versus signal length for ADMM, PALM, DALM and FISTA algorithms(Experiment 4)

4.5 Computational Time Comparison of ADMM against Other Algorithms

In Experiment 4, a comparison of the computational time of ADMM and those of four other algorithms were obtained both in the noiseless and noise case. The length of the sparse vector n , the number of observations M , and the sparsity of \mathbf{x} were set to $n = 500, 1000, 2000, 4000, 8000$, $m = n/2$, and $k = \text{round}(m/3)$, respectively, where $\text{round}(m/3)$ means rounding each element of $m/3$ to the nearest integer. For each set of (m, n, k) , we perform 1000 trials to reconstruct \mathbf{x} and \mathbf{e} for each of the five algorithms. Hence, we calculated the average time required for each algorithm to successfully reconstruct the signal \mathbf{e} at different values of n . In the noiseless case, the corresponding time of OMP is 0.0234s, 0.14s and 2.03s, when the length of the sparse vector n was set to $n = 500, 1000, 2000$. However, OMP fails to reconstruct the signal when $n = 4000, 8000$. The results were also observed in the noise case. This is because the OMP method often requires additional measurements to reconstruct the sparse signal. The other four algorithms, i.e., PALM, DALM, ADMM and FISTA, successfully reconstruct the signal \mathbf{x} and \mathbf{e} . The time curves of the four other algorithms are shown in Fig.4.

As can be seen from Fig.4, for a different set of n , the average time for the successfully reconstruction of ADMM is less than that of PALM and more than that of DALM. The average time of algorithms based on the Lagrangian multiplier method, i.e., PALM, DALM and ADMM, is significantly less than that of FISTA. Although FISTA and the Lagrangian multiplier methods are considered first order methods, FISTA requires a larger number of iterations to converge. Specifically, in the noiseless case, while the length of sparse vector n was set to $n = 500, 1000, 2000, 4000, 8000$,

the average time of FISTA is 0.21s, 3.62s, 10.9s, 37.1s and 168s, and that of ADMM is 0.0585s, 0.276s, 0.891s, 6.44s and 47.9s.

Experiment results show that the average computational time for the successfully reconstructing of ADMM is considerably lower than that of FISTA and PALM, while OMP fails to reconstruct the signal. Although the computational time of ADMM is higher than that of DALM, the success probability of reconstruction of DALM is lower than that of ADMM, as is shown in Experiment 3.

5 Conclusion

Sparse signal reconstruction is a key problem in CS. In this paper, in order to achieve signal reconstruction, we solve a l_1 -norm minimization problem using the alternating direction multiplier method. In the proposed algorithm, we introduce new variable by employing the variable splitting technique and transform the constrained optimization problem into an unconstrained one using the Lagrangian multiplier method. Multiplier variables can be obtained by maximizing the dual function of multiplier formulas, which can be solved using the gradient ascent method. Finally, the problem of sparse signal reconstruction can be transformed into a loop iteration containing three sub-formula by employing the methods of alternating minimization and Gauss-Seidel iteration. We carried out extensive experiments, performing comparisons with four other existing algorithms, which showed that our proposed approach outperformed the other methods. The simulation results show that the proposed algorithm can improve the reconstruction success probability while reducing the computational complexity.

References

1. R.G.Baraniuk. Compressive Sensing [Lecture Notes]. IEEE Signal Processing Magazine,24(4):118-121(2007).
2. A.Beck, M.Teboulle. A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems. Siam Journal on Imaging Sciences,2(1):183-202(2009).
3. S.Boyd, N.Parikh, E.Chu, B.Peleato, J.Eckstein. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. Foundations and Trends in Machine Learning,3(1):1-122(2011).
4. S.Boyd, L.Vandenberghe. Convex optimization. Cambridge university press, Cambridge(2004).
5. R.H.Byrd, M.E.Hribar, J.Nocedal. An Interior Point Algorithm for Large-Scale Nonlinear Programming. SIAM Journal on Optimization,9(4):877-900(1999).
6. E.J.Candès. Compressive sampling. Proceedings of the International Congress of Mathematicians, Madrid, Spain(2006).
7. E.J.Candès, J.K.Romberg, T.Tao. Stable signal recovery from incomplete and inaccurate measurements. Comm on Pure and Applied Math.59(8):1207-1223(2006).
8. E.J.Candès, T.Tao. Decoding by linear programming. IEEE Transactions on Information Theory,51(12):4203-4215(2005).
9. P.L.Combettes, V.R.Wajs. Signal recovery by proximal forward-backward splitting. Multiscale Modeling and Simulation,4(4): 1168-1200(2005).
10. D.L.Donoho. Compressed sensing. IEEE Transactions on information theory.52(4):1289-1306(2006).
11. D.L.Donoho, M.Elad, V.N.Temlyakov. Stable recovery of sparse overcomplete representations in the presence of noise. IEEE Transactions on Information Theory, 52(1):6-18(2005).

12. D.L.Donoho. For Most Large Underdetermined Systems of Linear Equations the Minimal l_1 -norm Solution Is also the Sparsest Solution. *Communications on Pure and Applied Mathematics*,59(6):797-829(2006).
13. M.Elad. *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. Springer Publishing Company, West Berlin(2010).
14. L.Feng, H.Sun, J.Zhu. Robust image compressive sensing based on half-quadratic function and weighted Schatten-p norm. *Information Sciences*, 2019, 477: 265-280(2019).
15. R.Gribonval. Piecewise linear source separation. *Proceedings of the SPIE*, 5207, 297-310(2003).
16. A.G.Nikitin, Y.Karadzhev. *Matrix Superpotentials: Lie Theory and Its Applications in Physics*. Springer, Japan,477-483(2013).
17. J.L.Lin, W. L.Hwang, S.C.Pei. Fast Matching Pursuit Video Coding by Combining Dictionary Approximation and Atom Extraction. *IEEE Transactions on Circuits and Systems for Video Technology*, 17(12):1679-1689(2007).
18. S.G.Mallat, Z.Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Trans on Signal Processing*, 41(12):3397-3415(1993).
19. D. M.Malioutov, M.Cetin, A.S.Willsky. Homotopy continuation for sparse signal representation. *IEEE International Conference on Acoustics, Speech, and Signal Processing*,5:733-736(2005).
20. J.K.Pant, W.S.Lu, A.Antoniou. New Improved Algorithms for Compressive Sensing Based on l_p Norm. *IEEE Transactions on Circuits and Systems II Express Briefs*,61(3):198-202(2014).
21. D.S.Pham, S.Venkatesh. Efficient algorithms for robust recovery of images from compressed data. *IEEE transactions on image processing*,22(12): 4724-4737(2013).
22. P.Stoica, Y.Selen. Cyclic minimizers majorization techniques, and the expectation-maximization algorithm: a refresher. *IEEE Signal Processing Magazine*, 21(1):112-114(2005).
23. J.A.Tropp, A.C.Gilbert. Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit. *IEEE Transactions on Information Theory*, 53(12):4655-4666(2007).
24. J.A.Tropp, A.C.Gilbert, M.J.Strauss. Simultaneous sparse approximation via greedy pursuit. *IEEE International Conference on Acoustics, Speech, and Signal Processing*,5:721-724(2005).
25. F.Wen, P.Liu, Y.Liu, Robert C.Qiu, W.Yu. Robust Sparse Recovery in Impulsive Noise via l_p - l_1 Optimization. *IEEE Transactions on Signal Processing*, 65(1):105-118(2017).
26. F.Wen, L.Pei, Y.Yang, W.Yu, P.Liu. Efficient and Robust Recovery of Sparse Signal and Image Using Generalized Nonconvex Regularization. *IEEE Transactions on Computational Imaging*,99:1-1(2017).
27. Y.Xiao, H.Zhu, S.Y.Wu. Primal and dual alternating direction algorithms for l_1 - l_1 -norm minimization problems in compressive sensing. *Computational Optimization and Applications*,54(2):441-459(2013).
28. J.Xie, A.Liao, Y.Lei. A new accelerated alternating minimization method for analysis sparse recovery. *Signal Processing*,145:167-174(2018).
29. Allen Y.Yang, Z.Zhou, A.G.Balasubramanian, S.Shankar Sastry, Y.Ma. Fast l_1 -Minimization Algorithms for Robust Face Recognition. *IEEE Transactions on Image Processing*, 22(8):3234-3246(2013).
30. J.Yang, Y.Zhang. Alternating Direction Algorithms for l_1 -Problems in Compressive Sensing. *Siam Journal on Scientific Computing*,33(1):250-278(2011).
31. Z.Zhang, Y.Xu, J.Yang, X. Li, D.Zhang. A Survey of Sparse Representation: Algorithms and Applications. *IEEE Access*, 3:490-530(2016).