manuscript received: 10 August 2021
revised: 15 October 2021

accepted: 21 November 2021

Compressed Sensing: A Review

Razieh Keshavarzian
.Department of Electrical Engineering, Heris Branch, Islamic Azad University, Heris, Iran
Email: r_keshavarzian@herisiau.ac.ir

Abstract

Compressed sensing (CS) is a new and promising framework for simultaneous sampling and compression of signals at sub-Nyquist rates. Under certain conditions, the signal can be reconstructed exactly from a small set of measurements via solving an optimization problem. In order to make this possible, compressed sensing is based on two principles of sparsity and incoherence. Compressed sensing takes advantage of the fact that most signals in nature are sparse or compressible, which means that when expressed in a suitable basis called as sparsifying basis, they will have a sparse representation. In the CS, the sparse signal is sampled by a non-adaptive linear sampling matrix. Then, based on the limited measurements obtained from the sampling matrix and using a non-linear algorithm, the original signal is reconstructed. The sparse signal reconstruction problem in the CS is an optimization problem that various algorithms have been proposed to solve it. The compressed sensing has a great application potential and can be used in a wide range of applications. Recently, deep learning has been used to solve the CS problem and its medical applications. In this paper, the generalities of compressed sensing are presented and CS reconstruction algorithms are reviewed. Also, the application of CS in magnetic resonance imaging (MRI) are investigated.

Keywords Compressed sensing, Reconstruction algorithm, Sampling matrix, Sparsity

1. Introduction

Previous methods of signal sampling are usually based on the Shannon-Nyquist theory. According to this theory, the sampling rate of a band-limited signal should be at least twice its bandwidth to guarantee error-free reconstruction. In most applications, the Nyquist rate is so high that many samples are obtained. Due to the existence of limitations such as the hardware memory needed to store samples and the bandwidth of the channels to be sent, compression prior to storage or transmission is a necessity [1]. In the compression stage, less important samples are removed. In this way, the cost and energy used to collect these samples is wasted [2].

In order to reduce the number of samples required to reconstruct signals without reducing their quality, the theory of compressed sensing (CS) has been proposed. In principle, compressed sensing integrates sampling and compression by collecting the least number of samples that contain the most information about the signal [2]. Compressed sensing theory states that certain signals can be reconstructed from much fewer samples or measurements compared to Nyquist theory. In order to make this possible, compressed sensing is based on the following two principles:

- Principle of sparsity: It states that the information rate of a time continuous signal may be much lower than the value specified by the bandwidth of the signal. More precisely, compressed sensing takes advantage of the fact that most signals in nature are sparse or compressible, which means that when expressed in a suitable basis, they will have a sparse representation.

- The principle of incoherence: it states that the signals that have a sparse representation in a base must be expanded in the area where the signal acquisition is performed. In other words, incoherence means that unlike the signal, the sampling waveforms must have a very dense representation at the sparse base.

In compressed sensing, the sparse signal is sampled by a non-adaptive linear sampling matrix. Then, based on the limited measurements obtained from the sampling matrix and using a non-linear algorithm, the original signal is reconstructed. Therefore, in compressed sensing, there are two approaches: the first is the design of an efficient sampling matrix with certain properties that leads to error-free reconstruction of the signal. The second is the improvement of signal sparsity and the development of non-linear reconstruction algorithms that can provide accurate reconstruction of the signal when the sampling matrix is known. In recent years, a variety of sampling matrices have been developed. A category of matrices that are uncorrelated with each sparsifying basis are random matrices that are built based on a specific probability distribution. But these matrices are very expensive in practical applications because they require high computational complexity and a lot of memory. Another category of matrices are a uniform random subset of the rows of an orthonormal matrix, which has a fast and efficient implementation. The third category of matrices are structured non-random matrices that can greatly reduce the memory and computational complexity in the reconstruction process, but they do not lead to the optimal solution. The sparse signal reconstruction problem in compressed sensing is an optimization problem that various algorithms have been proposed to solve it. These algorithms are divided into three general categories: greedy algorithms, convex approximation based algorithms, non-convex approximation based algorithms. A greedy algorithm obtains the approximate solution of the CS problem iteratively and step by step. This work is done by detecting the location of the non-zero components of the signal. The convex approximation based algorithms obtain the approximate solution of the CS reconstruction problem by replacing a convex function instead of the l_0 norm and then solving it iteratively. The non-convex approximation based algorithms replace the l_0 norm with a non-convex function. Applying compressed sensing to images faces several challenges; including the need for a huge memory to store the random sampling matrix and also the high cost of some reconstruction algorithms. So far, various methods have been proposed to reconstruct CS images that have tried to deal with these challenges. Among them, we can mention the block compressed sensing in [3, 4], which was proposed in order to solve the first challenge. In general, it can be said that a CS image reconstruction method, by using previous knowledge about natural images, tries to provide a high quality reconstructed image from the least number of measurements. The compressed sensing has a great application potential and can be used in a wide range of applications, like signal, image and video processing [5, 6, 7], medical image processing [8], wireless sensor network [9], data mining [10], and communication [11].

2. Compressed Sensing Problem

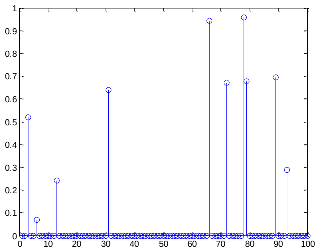
Consider a finite-length signal $\mathbf{u} \in \mathbb{R}^N$. A number of M ($M \ll N$) linear, non-adaptive measurement of $\mathbf{u}$ are acquired through the following linear transformation:

$$\mathbf{y} = \boldsymbol{\Phi}\mathbf{u} + \mathbf{e} \qquad (1)$$

where $\mathbf{y} \in \mathbb{R}^M$ is the measurement vector, $\boldsymbol{\Phi} \in \mathbb{R}^{M\times N}$ is a measurement matrix and $\mathbf{e}$ denotes possible measurement noise vector with $\|\mathbf{e}\|_2 \leq \epsilon$. The usual choice for the measurement matrix $\boldsymbol{\Phi}$ is a random matrix [4]. We wish to reconstruct the signal $\mathbf{u}$ from $\mathbf{y}$ by solving (1). Since $M \ll N$, the reconstruction of $\mathbf{u}$ from $\mathbf{y}$ is ill-posed in general. However, if $\mathbf{u}$ is sparse (or compressible), then exact reconstruction is possible.

A. Sparse signals

A signal $\mathbf{u} \in \mathbb{R}^N$ is called sparse if most of its components are equal to zero, that is, $\|\mathbf{u}\|_0 \ll N$, where $\|\mathbf{u}\|_0$ is the l_0 norm of the signal and represents the number of its non-zero components. If there is a maximum of s non-zero components, i.e. $\|\mathbf{u}\|_0 \leq s$, the signal $\mathbf{u}$ is called sparse of order s or s-sparse. Figure (1) shows an example of a sparse signal.

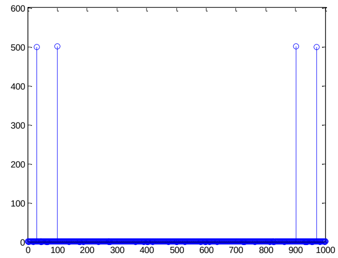


Fg.1. An example of sparse signal

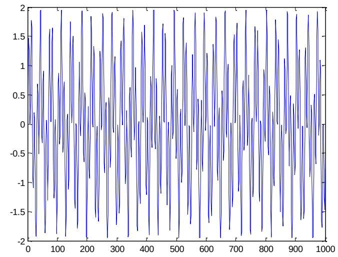
Many signals themselves are not sparse, but when they are expressed in a convenient basis will have sparse representations. In this case, the signal $\mathbf{u}$ can be represented as

$$\mathbf{u} = \boldsymbol{\Psi}\mathbf{x}\,, \qquad (2)$$

where $\boldsymbol{\Psi}$ is a sparsifying basis or dictionary and $\mathbf{x} \in \mathbb{R}^N$ is the coefficient vector that most entries of which are zero or close to zero. In figure (2), an example of a signal is shown, which has a sparse representation in the Fourier basis.



a



b

Fig.2. Representation of sparse signal, a) original signal, b) the signal in Fourier base

B. Compressible signals

An important point in practice is that few signals in nature are exactly sparse; rather, most of them are compressible, which means that most of the components of these signals are small and close to zero either in the time domain or in the $\boldsymbol{\Psi}$ domain. Therefore, these small components can be considered as zero without significant change in the signal. As a result, compressible signals are well approximated by sparse signals.

C. Reconstruction

To reconstruct the signal $\mathbf{u}$ from $\mathbf{y}$, one could search for the sparsest coefficient vector (i.e. the vector $\mathbf{x}$ with the smallest l_0 norm) consistent with the measurement $\mathbf{y}$ by solving the optimization problem

$$\min_{\mathbf{x}} \ \|\mathbf{x}\|_0 \quad \textit{subject to} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon \qquad (3)$$

which $\mathbf{A} = \mathbf{\Phi\Psi}$, and then reconstruct $\mathbf{u}$ using (2). Unfortunately, the optimization problem (3) is NP-hard that can only be solved using a combinatorial approach [12]. Thus, alternative procedures to find out a suboptimal solution have been proposed in recent years. One of these, is to relax the l_0 norm, replacing it by a continuous or even smooth approximation [13]. Examples of such approximations include l_p norms for some $0 < p \leq 1$ [14] or even smooth functions such as Logarithm [15], Exponential [16]. A popular choice for the approximation function is l_1 norm which leads to the convex optimization problem

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad subject\ to \quad \|\mathbf{y} - \mathbf{Ax}\|_2 \leq \epsilon \qquad (4)$$

This optimization problem known as basis pursuit denoising (BPDN) can be recast into the unconstrained optimization problem

$$\min_{\mathbf{x}} \ \frac{1}{2}\|\mathbf{y} - \mathbf{Ax}\|_2^2 + \lambda \|\mathbf{x}\|_1 , \qquad (5)$$

which the second term is a regularization term that reflects prior information of the signal and λ is the regularization parameter. With an appropriate choice of the λ, the problem in (5) will yield the same solution as that in (4) [17]. This problem can be solved by many efficient algorithms which are investigated in the next section.

3. Reconstruction Algorithms

A CS reconstruction algorithm should be able to reconstruct the original signal from the measurement vector when the sampling matrix is known. In this section, the different algorithms presented for the approximate solution of the problem (3) are studied. These algorithms are divided into three general categories: greedy algorithms, algorithms based on convex approximation, algorithms based on non-convex approximation.

A. Greedy algorithms

A greedy algorithm obtains the approximate solution of the problem (3) iteratively and step by step. This is done by detecting the location of the non-zero components of the signal. After detecting the non-zero locations, the signal is reconstructed using the relationship $\mathbf{x} = (\mathbf{A}_s)^{\dagger}\mathbf{y}$, where $\mathbf{A}_s$ is the measurement matrix with columns corresponding to s non-zero locations and $(\mathbf{A}_s)^{\dagger}$ is the pseudo-inverse of the matrix $\mathbf{A}_s$. Among greedy algorithms, matching pursuit (MP) [18] and its improved version orthogonal matching pursuit (OMP) [19] can be mentioned. In the OMP algorithm, in each iteration, the column of the matrix $\mathbf{A}$ that has the highest correlation with the measurement vector is selected. Then, the effect of this column in the measurement vector is removed and a residual vector is obtained. The steps are repeated on this vector until finally s columns of $\mathbf{A}$, which corresponds to s non-zero locations, are determined [20]. This algorithm is very fast but has less reconstruction accuracy. Since this algorithm selects only one column of $\mathbf{A}$ per iteration, at least s iterations are required for an s-sparse vector. But when the vector is not very sparse (s is large), the reconstruction becomes expensive. To speed up the algorithm, several columns can be selected in each iteration [17]. Based on this, its improved versions stagewise orthogonal matching pursuit (StOMP) [21], regularized orthogonal matching pursuit (ROMP) [22] and compressive sampling matching pursuit CoSaMP [23] were presented. There is another class of greedy algorithms that use a threshold level to select non-zero components. Among these algorithms, iterative hard thresholding (IHT) [24 and 25] can be mentioned. The IHT algorithm is resistant to noise and can be solved with the minimum number of measurements, but it is sensitive to the change of the scale of **A**, so that the performance of the algorithm changes in different scales of **A**.

B. *Convex approximation based algorithms*

These algorithms obtain the approximate solution of the problem (3) by substituting a convex function instead of the l_0 norm and then solving it iteratively. These algorithms start from an initial estimate. Then, in each iteration, an estimate of the signal is obtained, and during the iterations, this estimate becomes more accurate to reach the optimal solution. One of these algorithms is iterative shrinkage thresholding algorithm (ISTA) [26] that solves the problem (5) as:

$$\mathbf{x}^{(k+1)} = S_{\lambda\mu,1}(\mathbf{x}^{(k)} + \mu\mathbf{A}^T(\mathbf{y} - \mathbf{A}\mathbf{x}^{(k)})) \quad (6)$$

where μ is a suitable step size and $S_{\lambda\mu,1}(.)$ is a soft thresholding function:

$$S_{\lambda\mu,1}(x_i) = \begin{cases} x_i - \lambda\mu/2 & if \quad x_i \geq \lambda\mu/2 \\ 0 & if \quad |x_i| < \lambda\mu/2 \\ x_i + \lambda\mu/2 & if \quad x_i \leq -\lambda\mu/2 \end{cases} \quad (7)$$

This algorithm is simple to be implemented; however, it converges quite slowly. Some accelerated versions of ISTA have been proposed, including two-step IST (TwIST) [27] and fast IST (FISTA) [28]. In these algorithms, each iteration depends on two previous iterations instead of one previous iteration. The FISTA for solving problem (5) is as follows:

$$\mathbf{x}^{(k+1)} = S_{\lambda\mu,1}(\boldsymbol{z}^{(k)} + \mu\mathbf{A}^T(\mathbf{y} - \mathbf{A}\boldsymbol{z}^{(k)})) \quad (8)$$

$$t^{(1)} = 1, \quad t^{(k+1)} = \frac{1}{2}\left(1 + \sqrt{1 + 4(t^{(k)})^2}\right) \quad (9)$$

$$\boldsymbol{z}^{(k)} = \mathbf{x}^{(k)} + \left(\frac{t^{(k)} - 1}{t^{(k+1)}}\right)(\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}) \quad (10)$$

The subband adaptive iterative shrinkage thresholding algorithm (SISTA) [29] is a generalization of ISTA, in which different step sizes are considered for different subbands of wavelet transform as the sparsifying transform. This increases the convergence speed of the algorithm. Given the vector μ=[μ_1,μ_2,…,μ_J], the algorithm uses the following iteration:

$$\min_{\mathbf{x},\mathbf{z}} f(\mathbf{x}) + g(\boldsymbol{z}) \quad subject\ to\ \boldsymbol{B}\mathbf{x} + \boldsymbol{Cz} = \boldsymbol{D} \quad (11)$$

where $\boldsymbol{\Lambda_\mu}$ is a diagonal matrix whose elements are equal to μ_j. The iteratively reweighted least squares (IRLS) algorithm [30] provides a simple method to solve the problem (5). In this method, $\|\mathbf{x}\|_1$ is replaced by the weighted l_2 norm of **x**.

The alternating direction method of multipliers (ADMM) algorithm [31] solves convex optimization problems by breaking them into smaller pieces. For this, it uses an auxiliary variable and separates the cost function into two parts. In general, this algorithm solves the following problems:

$$\min_{\mathbf{x},\mathbf{z}} f(\mathbf{x}) + g(\boldsymbol{z}) \quad subject\ to\ \boldsymbol{B}\mathbf{x} + \boldsymbol{Cz} = \boldsymbol{D} \quad (12)$$

where $\mathbf{x} \in \mathbb{R}^N$ is the main variable, $\boldsymbol{z} \in \mathbb{R}^m$ is auxiliary variable, $\boldsymbol{B} \in \mathbb{R}^{p\times N}$, $\boldsymbol{C} \in \mathbb{R}^{p\times m}$ and $\boldsymbol{D} \in \mathbb{R}^p$. The augmented Lagrange function is calculated as follows:

$$L_\rho(\mathrm{x}, z, \mathrm{v}) = f(\mathrm{x}) + g(z) + \mathrm{v}^T(B\mathrm{x} + Cz - D) + \frac{\rho}{2}\|B\mathrm{x} + Cz - D\|_2^2 \quad (13)$$

which $\rho > 0$. The augmented Lagrange function is minimized with respect to the variables in an orderly or sequential manner, in such a way that it is minimized first with respect to the variable x (fixed z) and then with respect to the variable z (fixed x). In both steps, the Lagrange coefficient is assumed to be constant. Finally, the obtained answers are used to update the Lagrange coefficient.

C. Non-convex approximation based algorithms

These algorithms obtain the approximate solution of the problem (3) by substituting a non-convex function instead of the l_0 norm and then solving it iteratively. Due to the non-convexity of the resultant problem, identifying its global minimum is challenging. For this reason, the conditions under which the algorithm converges to a local minimum are emphasized. However, algorithms based on the non-convex approximation reconstruct sparse signals from a smaller number of measurements. Also, they may increase noise resistance and lead to stability [17]. Among the suggested approximation functions, we can mention l_p $(0 < p \leq 1)$ [14] norms. The l_p norms, which are the most common approximation for l_0 norm, are defined as follows for $p > 0$:

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{N} |x_i|^p\right)^{\frac{1}{p}} , \quad (14)$$

$\|\mathbf{x}\|_p^p$, when p tends to zero, it will be equal to $\|\mathbf{x}\|_0$ [32]. Therefore, the problem (3) can be approximated by the following problem:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_p^p \quad subject\ to \quad \mathbf{y} = \mathbf{A}\mathbf{x} \quad (15)$$

The IRLS algorithm is also used to solve the problem (15) for $0 < p < 1$. In this method, $\|\mathbf{x}\|_p^p$ is replaced by the weighted l_2 norm of $\mathbf{x}$.

The problem (15) is a non-convex optimization problem for $0 < p < 1$. This problem can be rewritten in the following unrestricted form, which is known as l_p regularization problem:

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda\|\mathbf{x}\|_p^p \quad (16)$$

In [33], this problem is solved for $p = \frac{1}{2}$ with the Half thresholding algorithm. This algorithm uses the following iteration:

$$\mathbf{x}^{(k+1)} = H_{\lambda\mu,\frac{1}{2}}(\mathbf{x}^{(k)} + \mu\mathbf{A}^T(\mathbf{y} - \mathbf{A}\mathbf{x}^{(k)})) \quad (17)$$

where, $H_{\lambda\mu,\frac{1}{2}}(.)$ is the Half thresholding function.

4. Image Compressed Sensing

Applying compressed sensing to images faces several challenges; including the need for a huge memory to store the random sampling matrix and also the high cost of some reconstruction algorithms. So far, various methods have been proposed to reconstruct compressed sensing images that have tried to solve these challenges. Among them, we can mention the block compressed sensing in [3], which was proposed in order to solve the first challenge. In general, it can be said that a CS image reconstruction method, by using previous knowledge about natural images, tries to provide a high-quality reconstructed image from the least number of measurements.

In block compressed sensing, the original image is divided into non-overlapping small blocks, and each block is sampled independently by a similar sampling matrix. This method has advantages, including the following:

• The sampling matrix is easily stored.

• Since each block is processed independently, the initial solution is easily obtained and the reconstruction process is greatly accelerated.

• It is more economical for real-time applications; because the coder does not need to wait for the whole image to be sampled and then send the samples.

5. Compressed Sensing MRI

One of the important applications of compressed sensing is in magnetic resonance imaging (MRI). MRI is a non-invasive imaging method that uses the magnetic resonance properties of hydrogen atoms inside the body and can display a wide range of tissues with high resolution. The main challenge of MRI that limits its use is the relatively slow speed of data collection. This leads to prolongation of imaging time and as a result patient's discomfort, increase of complications caused by radiation and motion distortion. Therefore, improving the speed of MRI imaging is of particular importance. Considering the physical and physiological limitations, the only efficient way to reduce imaging time is to reduce the number of required samples. But reducing the sampling rate violates the Nyquist condition and leads to distortion in the reconstructed image. By introducing the compact measurement theory as an alternative to the Shannon-Nyquist theory, it is possible to reconstruct MR images without distortion from much less data compared to the Nyquist theory. The result of using intensive measurement in MRI is reducing the time required for imaging, reducing costs and patient comfort.

Sampling in MRI is a special case of compact sensing where the sampled linear combinations are Fourier coefficients. In this situation, the compact sensing method claims to be able to accurately reconstruct the original signal from a small subset of k-space.

6. Conclusions

The problem of reconstructing the compressed sensing image is an optimization problem that various algorithms have been proposed to solve. These algorithms are divided into three general categories: greedy algorithms, algorithms based on convex approximation, algorithms based on non-convex approximation. Studies have shown that algorithms based on non-convex approximation require fewer measurements for reconstruction. Also, they may increase noise immunity and lead to stability.

References

[1] Baraniuk, R. G.; “Compressive Sensing,” IEEE Signal Processing Mag., vo. 24, no. 4, p.p. 118-120, July 2007.

[2] Qaisar, S.; Bilal, R. M.; Iqbal, W.; Naureen, M.; and Lee, S.; “Compressive Sensing: From Theory to Applications, A Survey,” Journal of Communication and Networks, nol. 15, no. 5, p.p. 443-456, July 2013.

[3] L. Gan, “Block Compressed Sensing of Natural Images ,” in Proc. of the 15th Intl. Conf. on Digital Signal Processing (DSP 2007), 2007.

[4] S. Mun, J. E. Fowler, Block Compressed Sensing of Images Using Directional Transforms, 16th IEEE Int. Conf. on Image Process. (ICIP 2009), Cairo, Egypt, 2009.

[5] Z. Luo, J. Liang, and J. Ren. "Deep Learning Based Compressive Sensing for UWB Signal Reconstruction." IEEE Transactions on Geoscience and Remote Sensing (2022).

[6] Z. Yin, Z. Wu, and J. Zhang. "A Deep Network Based on Wavelet Transform for Image Compressed Sensing." Circuits, Systems, and Signal Processing, p.p 1-20, 2022.

[7] Y. Sun, X. Chen, M. S. Kankanhalli, Q. Liu, and J. Li. "Video Snapshot Compressive Imaging using Residual Ensemble Network." IEEE Transactions on Circuits and Systems for Video Technology, (2022).

[8] Liu, S.; Cao, J.; Wu, G.; Liu, H.; Tan, X.; and Zhou, X.; “CS-MRI Reconstruction via Group based Eigenvalue Decomposition and Estimation,” Neurocomputing, vol. 283, p.p. 166-180, 2018.

[9] A. Aziz, W. Osamy, O. Alfawaz, and A. M. Khedr. "EDCCS: effective deterministic clustering scheme based compressive sensing to enhance IoT based WSNs." Wireless Networks (2022): 1-17.

[10] C. Wang, X. Shen, H. Wang, and H. Mei. "Energy-efficient collection scheme based on compressive sensing in underwater wireless sensor networks for environment monitoring over fading channels." Digital Signal Processing 127 (2022): 103530.

[11] W. Xie, J. Xiao, P. Zhu, C. Yu, and L. Yang. "Deep Compressed Sensing-Based Cascaded Channel Estimation for RIS-Aided Communication Systems." IEEE Wireless Communications Letters 11, no. 4 (2022): 846-850.

[12] J. J. Fuchs, “On Sparse Representations in Arbitrary Redandants Bases,” IEEE Trans. on Information Theory, vol. 50, no. 6, June 2004.

[13] M. Elad, “Sparse and Redundant Representations,” springer, 2010.

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

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

[16] C. Gao, N. Wang, Q. Yu, and Z. Zhang, “A feasible nonconvex relaxation approach to feature selection,” in Proc. 25th AAAI Conf. Artif. Intell., pp. 356–361, 2011.

[17] K. Hayashi, M. Nagahara and T. Tanaka, “A User's Guide to Compressed Sensing for Communication Systems,” IEEE Trans. on commun., no. 3, March 2013.

[18] Mallat, S.; and Zhang, Z.; “Matching pursuits with time-frequency dictionaries, ” IEEE Trans. on Signal Processing, vol. 41, no. 12, p.p. 3397-3415, 1993.

[19] Tropp, J. A.; and Gilbert, A. C.; “Signal Recovery from Partial Information via Orthogonal Matching Pursuit,” IEEE Trans. on Information Theory, vol. 53, no. 12, p.p. 4655-4666, 2007.

[20] Qaisar, S.; Bilal, R. M.; Iqbal, W.; Naureen, M.; and Lee, S.; “Compressive Sensing: From Theory to Applications, A Survey,” Journal of Communication and Networks, nol. 15, no. 5, p.p. 443-456, July 2013.

[21] Donoho, D.; Tsaig, Y.; Drori, I.; and Starck, J.; “Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit,” Stanford University, Tech. Rep. Mar. 2006.

[22] Needell, D.; and Vershynin, R.; “Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit,” in Foundations of computational mathematics, vol. 9, no. 3, p.p. 317-334, 2009.

[23] Needell, D.; and Tropp, J. A.; “COSAMP: Iterative Signal Recovery from Incomplete and Inaccurate samples,” Appl. Comput. Harmon. Anal., vol. 26, p.p. 301-321, 2009.

[24] Blumensath, T.; and Davies, M. E.; “Iterative Hard Thresholding for Compressed Sensing,” Appl. Comput. Harmon. Anal., vol. 27, no. 3, p.p. 265-274, Jan. 2009.

[25] Blumensath, T.; and Davies, M. E.; “Iterative Thresholding for Sparse Approximations,” Journal of Fourier Anal Appl., p.p. 626-654, 2008.

[26] Daubechies, I.; Defrise, M.; and De-Mol, C.; “An Iterative Thresholding Algorithm for Linear Inverse Problems with a Sparsity Constraint,” Pure Appl. Math., vol. 57, no. 11, p.p. 1413-1457, 2004.

[27] Bioucas-Dias, J.; and Figueiredo, M.; “A new TwIST: Two step iterative shrinkage/ thresholding algorithms for image restoration,” IEEE Trans. on Image Processing, vol.16, no. 12, p.p. 2992-3004, Dec. 2007.

[28] Beck, A.; and Tebulle, M.; “A fast iterative shrinkage-thresholding algorithm for linear inverse problems,” SIAM J. Imaging Sci., vol. 2, no. 1, p.p. 183-202, Jan. 2009.

[29] Bayram, I.; and Selesnick, I.; “A Subband adaptive Iterative shrinkage/thresholding algorithm,” IEEE Trans. on Signal Processing, vol.58, no. 3, March 2010.

[30] Chartrand, R.; and Yin, W.; “Iteratively reweighted algorithms for compressive sensing, ” in Proc. ICASSP, Las Vegas, USA, p.p. 3869-3872, Mar.-Apr. 2008.

[31] Boyd, S.; and et.al.; “Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers,” Foundations and Trends in Machine Learning, vol. 3, no. 1, 2010.

[32] Bruckstein, A. M.; Donoho, D. L. and Elad, M.; “From Sparse Solutions of Systems of Equations to Sparse Modeling of Signals and Images,” SIAM Review, vol. 51, no. 1, p.p. 34-81, 2009.

[33] Zeng, J.; Lin, S.; Wang, Y.; and Xu, Z.; “L0.5 Regularization: Convergence of Iterative Half Thresholding Algorithm,” IEEE Trans. on signal processing, vol. 62, No. 9, 2014.