International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 7 (2015) pp. 17201-17212

http://www.ripublication.com

Review of Different Compressive Sensing Algorithms and Recovery Guarantee of Iterative Orthogonal Matching Pursuit

P. Vimala and G. Yamuna*
*Assistant Professor, * Professor, Department of Electrical Engineering, Faculty of Engineering and Technology, Annamalai University, Annamalai Nagar, Chidambaram, TN.*

Abstract

Recently compressive sensing (CS) technique places a major role in many research areas like Image Processing, Medical and Seismic Imaging, Analog to Information Conversion, Wireless Communication and Networks. It predicts the sparse or approximately sparse high dimensional data or signal from the highly incomplete linear measurements. This paper reviews the various compressive sensing iterative algorithms. Performance of those algorithms are evaluated in terms of Mean Square Error (MSE) and inferred that Orthogonal Matching Persuit has lower MSE for Gaussian sparse signal. Recovery guarantee of OMP algorithm is also stated.

Index Terms: Compressive Sensing, Reconstruction algorithms, Orthogonal Matching Pursuit, Sensing Matrix, Mutual Incoherence.

Introduction

The Shannon Nyquist Sampling theorem states that a band limited signal can be reconstructed from the samples without distortion if the samples are taken at least twice the highest frequency spectral component. But if the signal of interest is sparse or approximately sparse in some suitable basis, then a new technique called compressive sensing provides an alternate means for lower sampling rate and its recovery. A signal is said to be *k* sparse if it has only *k* non zero entries alternatively support of a signal is *k*. The CS measurement system for recovering k sparse signal $x \in R^n$ can be modelled mathematically, as

$$y = \phi x$$

where $\phi \in R^{dxn}$ called sensing matrix and $y \in R^d$ represent measurements. If the measurement vector y is equal to the length of the signal *x* then it can be perfectly recovered provided matrix ϕ is invertible. In case of measurement vector y with size

$d << n$ and also greater than the number of nonzero entries of the original signal x then from these few incomplete measurements x can be recovered using CS theory.

For the recovery, projection basis ϕ should be incoherent with the basis in which the signal has a sparse representation [1]. The natural attempt to recover x by solving an optimization problem of the form

$$\hat{x} = \arg\min \|z\|_0 \; subject\ to\ z \in B(y)$$

where $B(y) = \{ z : \phi z = y \}$. It is a NP hard problem. One avenue for translating this problem into track able is replace l_0 minimization with l_1 minimization based on linear programming techniques. l_1 minimization techniques provides accurate method for sparse signal recovery if it satisfies Restricted Isometry Property(RIP). But the computational complexity of l_1 minimization is highly impractical. It leads to the need of faster recovery algorithm that works in linear time. Recently several low complexity iterative algorithms are proposed.

The rest of the paper is organized as follows. The basics of Iterative Compressive Sensing are described in section II. Various iterative algorithms are discussed in section III. Implementation of those algorithms and results are given in section IV and also describes the recovery guarantee of OMP. Conclusion is given in section V.

Basics of Iterative Compressive Sensing

A set $\{\phi_i\}_{i=1}^n$ is called a basis for R^n if the vectors in the set span R^n and are linearly independent. Each vector in the space has unique representation as a linear combination of these vectors. For any $x \in R^n$, there exist coefficient $\{c_i\}_{i=1}^n$ such that

$$x = \sum_{i=1}^{n} c_i \phi_i$$

It can be represented by matrix as $x = \phi c$ where ϕ^{nxn} matrix with columns given by φ_i and c length n vector. Note that if the columns of ϕ are orthonormal then the coefficient c can be easily calculated as

$$c_i = \langle x, \varphi_i \rangle \text{ or } c = \phi^T x$$

Basis can be generalized to set of possibly linearly dependent vectors known as frame. Mathematically, a frame is a set of vectors $\{\phi_i\}_{i=1}^n$ in R^d, $d<n$ corresponding to a matrix $\phi \in R^{dxn}$. Frame is the richer representation of data due to their redundancy. That is for a given x, there exist infinitely many coefficients vectors c such that $x = \phi c$, since number of unknowns are greater than number of equations. If ϕ is well defined and have linearly independent rows that ensures $\phi\phi^T$ is invertible then one of the ways to obtain the coefficient is

$$c = (\phi\phi^T)^{-1}\phi x$$

CS literatures refer $\phi \in R^{dxn}$ as dictionary or sensing matrix. CS recovery algorithms are guaranteed for the perfect recovery of the sparse signal *x* via RIP [2-4] given by

A sensing matrix ϕ satisfies the RIP of order k if there exists a constant δ such that

$$(1-\delta)\|x\|_2^2 \le \|\phi x\|_2^2 \le (1+\delta)\|x\|_2^2$$

for any *k* sparse vector.

The minimum of all constants δ satisfying the above condition is called isometry constant δ_k. But verifying a sensing matrix ϕ for RIP involves combinatorial computation complexity, since in each case one must essentially consider $\binom{n}{k}$ sub matrices. Other than RIP the widely used condition is mutual incoherence given by Donoho et al.[5]

The mutual incoherence μ of a sensing matrix ϕ , is the largest absolute inner product between any two columns φ_i, φ_j of ϕ

$$\mu(\phi) = \max_{1\le i\le j\le n} \frac{|\langle \varphi_i, \varphi_j \rangle|}{\|\varphi_i\|_2 \|\varphi_j\|_2}$$

Iterative Reconstruction Algorithms

The underdetermined system of equations $y = \phi x$ can be perfectly reconstructed from the properly designed recovery algorithms with known signal sparsity. From the various CS approaches available, iterative greedy algorithms received a significant attention due to their low complexity. Greedy iterative algorithm searches the support of the sparse signal. In each greedy iteration search, correlation between each columns of ϕ and residual are compared to identify the element of the support. This approach includes Orthogonal Matching Pursuit (OMP), Single Step Orthogonal matching pursuit(SSOMP), Regularized Orthogonal Matching Pursuit(ROMP), S tagewise Orthogonal Matching Pursuit (StOMP) and Compressive Sampling Matching Pursuit (CoSaMP). The reconstruction probability of OMP from given y is stated by following theorem [6].

Theorem

$\delta \in (0,0.36)$, and choose $d \ge Ck \ln(n/\delta)$. Suppose *x* is an arbitrary *k* sparse signal in R^n. Draw *d* from the standard Gaussian distribution on R^d. Given the data, OMP can reconstruct the signal with probability exceeding $1\text{-}2\delta$. For this theoretical results, it suffices that $C=20$. When k is large, $C=4$ is enough.

The basic iterative algorithm is Orthogonal Matching Pursuit (OMP), which was popularised and analyzed in [6][7]under a hypothetical assumption that columns ϕ are orthonormal.

A. OMP

To identify the original signal *x*, it is needed to determine which columns ϕ participate in the measurement vector y. The algorithm picks the columns in greedy fashion. At each iteration a column of ϕ is selected which is strongly correlated with y. Then subtract off its contribution to y and iterate on residue. The OMP algorithm is followed [6]

Input:

- Measurements *y*
- Sensing matrix ϕ
- Sparsity *k*

Initialization

- Residual vector $r_t=y$,
- Estimated support set ϕ_0=null matrix,
- Iteration $t=1$
- Index set I_o= null matrix

1. Find the index λ_t that solves the optimization problem
   $$\lambda_t = \underset{j=1,2,....n}{\arg\max} \left| < r_{t-1}, \phi_j > \right|$$
2. Augment the index set $I_t = I_{t-1} \cup \{\lambda_t\}$ and $\phi_t = [\phi_{t-1}\phi_{I_t}]$
3. Solve a least squares problem to obtain a new signal estimate
   $$x_t = \underset{x}{\arg\min} \|\phi_t x - y\|$$
   $$= \left(\phi_t^T \phi_t\right)^{-1} \phi_t^T y$$
4. Calculate the new approximation of the data and the new residue
   $a_t = \phi_t x_t$, $r_t = y - a_t$
5. Increment t, and return to step 2 if $t<k$
6. The estimate $\hat{x}$ for the signal has non zero indices at the components listed in I_k. The value of the estimate $\hat{x}$ is component I_j equals j^{th} component of x_t

---

The computational complexity of OMP depends on the number of iterations needed for exact reconstruction. The OMP always run through *k* iterations and its complexity is in *O(kdn)*[8].

B. Single step OMP

Algorithm acquires the participated *d* columns of ϕ in the measurement vector *y* in a single step by picking *d* columns which have the highest correlation magnitudes developed by Manjumdar et al.[9]. It leads to form linear matrix equation where concerned matrix is almost full rank.

Input:

- Measurements y
- Sensing matrix ϕ
- Sparsity k

Initialization:

- Index set I

1. Compute
   $I_i = |< y\phi_i >| \; 1 \le i \le n$
2. Form a matrix A whose columns are $\phi_i, j \in S$ where
   $S = \{i : I_i \ge I_j \forall j \notin S\} \, and \, |S| = d$
3. Compute $\hat{x} = A^{-1}y$ as the signal estimate

SSOMP reduces the number of iterations required for exact recovery to single iteration, in some sparsity levels. The range over which SSOMP effective is different from the effective range of OMP. And also it is proved in Avishek Manjumdar et al.[9] the asymptotic performance of SSOMP is poor where the recovery error decays logarithmically with signal dimension *n* and *d=klogn*. However, for more realistic ranges of parameters, SSOMP can have good performance.

C. ROMP

The main difference between OMP and ROMP algorithm is the identification and regularization steps given by Needell & Vershynin [10]. Instead of choosing only one strongly correlated column at each iteration in OMP, choose set of *k* biggest absolute coordinates of the observation vector $u = \phi^* y$. By this way, ROMP can recover signals perfectly without going through all k iterations. And also it is able to make mistakes in the support set by selects more than one coordinate, while still correctly reconstructing the original signal. This is accomplished because the number of incorrect choices the algorithm can make. Once the algorithm chooses an incorrect coordinate, however, there is no way for it to be removed from the support set [10][11].

Input:

- Measurements y
- Sensing matrix ϕ
- Sparsity k

Initialization:

- Residual vector $r=y$
- Iteration $t=1$
- Index set I_o= null matrix

1. Choose a set J for the k biggest coordinates in magnitude of the observation vector $u = \phi^* r$
2. Divide the set J into subsets J_0 which satisfies
   $|u(i)| \le 2\ |u(j)|$ for all $i, j \in J_0$
   and choose the subset J_0 with maximum energy of $\|u_{j_0}\|_2$
3. Set $I_t \in I_{t-1} \cup J_0$ & calculate the new approximation by solving the least square equations
   $\hat{x} = \arg\min \|\phi_{I_t} x - y\|_2$
   $r = y - \phi\hat{x}$
4. Go to step 1 if r ≠ 0 then keep increasing $t = t + 1$

The algorithm ROMP reconstructs a sparse signal in a number of iterations linear to the sparsity that is it runs with at most *O(k)* iterations and the reconstruction is exact provided that the RIP holds with parameter $\delta_{2k} \le 0.06/\sqrt{\log k}$. It also demonstrates that the number of iterations needed for sparse compressible is higher than the number needed for sparse flat signals.

The total running time of ROMP is *O(ndk)* which is the same bound as for OMP[10].

D. StOMP

The modified version of OMP is called StOMP. The OMP find solution by adding one vector at a time but StOMP uses several vectors. StOMP compares the values of the dot product of y with the columns of ϕ. Then selects all vectors above a preset threshold value and uses a least squares method to find an approximation. This is then repeated with a residue vector. The algorithm iterates through a fixed number of stages and then terminates given by Donoho et al. [12].

Input:

- Measurements y
- Sensing matrix ϕ
- Sparsity k

Initialization:

- Residual vector $r_t=y$,

- Estimated support set ϕ_0=null matrix,
- Iteration *t=1*
- Index set I_o= null matrix

1. The t^{th} stage applies matched filtering to the current residual, getting a vector or residual correlation
$$C_t = \phi^T r_{t-1}$$
2. Next perform hard thresholding to find the significant non zeros.
$$J_t = \{ j : |C_t(j)| > th_t \sigma_t$$
where J_t-Set contain large coordinates, σ_t -Formal noise level and *th*-Threshold parameter
3. Update the newly selected coordinate with the previous support estimate
$$I_t = I_{t-1} \cup J_t$$
4. New approximation $\hat{x}_t$ supported in I_t with coefficient supported in I_t and estimate residual are given by
$$(\hat{x}_t)_{I_t} = \left(\phi_{It}^T \phi_{It}\right)^{-1} \phi_{It}^T y \qquad r_t = y - \phi \hat{x}_t$$
5. Check stopping condition. If not *t=t+1*and go to the next stage of procedure. If time to stop $\hat{x}_t$ set as the final output.

The algorithm selects coordinates whose values are above a specified threshold in each iteration rather than selecting the largest component in OMP. Then it solves the least squares problem to update the residual. An advantage of using a method like this is that it can produce a good approximation with a small number of iterations. A disadvantage is in determining an appropriate value for the threshold as different threshold values could lead to different results.

E. CoSaMP

An alternative approach for ROMP would be to allow the algorithm to choose incorrectly as well as fix its mistakes in later iterations. Needell &Tropp [13] developed a new variant of OMP, Compressive Sampling Matching Pursuit (CoSaMP). In this case, each iteration selects a slightly larger support set, reconstruct the signal using that support, and use that estimation to calculate the residual. At each iteration, the algorithm runs through five steps. At the first step using the current sample y, the algorithm computes a vector that is highly correlated with the signal which is called as proxy. Then the current approximation is coupled with the newly indentified components. Next the algorithm solves the least squares problem to approximate the target signal. Through pruning, the algorithm provides a new approximation by retaining only the largest entries in the least squares approximation. Finally the samples are updated. The algorithm's major steps are described as follows

---

Input:

- Measurements y
- Sensing matrix ϕ
- Sparsity k

Initialization:

- Residual vector $r=y$
- set $a^0=0$
- Iteration $t=1$

Repeat the following steps and increment t until the halting criterion is true.

Signal Proxy: *set* $v=\phi^*r$, $\Omega=supp\ v_{2k}$, and merge the supports $T=\Omega \cup supp a^{t-1}$

Signal Estimation: Using least squares, set $b_T=\phi^{\dagger}_T y$ *and* $b_{T^c}=0$.

Prune: To obtain the next approximation, set $a^t=b_k$.

Sample Update: Update the current samples, $r=y-\phi a^t$.

The CoSaMP recovery algorithm delivers the same guarantees as the best optimization-based approaches[13]. Moreover, this algorithm offers rigorous bounds on computational cost and storage. It is likely to be extremely efficient for practical problem because it requires only matrix-vector multiplies with the sampling matrix. For compressible signals, the running time is just $O(n \log^2 n)$, for signal of length n. Satpathi et al. [14] demonstrate that the CoSaMP algorithm can successfully reconstruct a k sparse signal from a compressed measurement y by a maximum of $5k$ iterations if the sensing matrix ϕ satisfies the Restricted Isometry Constant of $\delta_{4k} < 1/\sqrt{5}$

Implementation and Results

The Sensing matrix ϕ^{50x250} was randomly generated with normally distributed elements with mean equal to 0 and variance to 1, then its columns were normalised to 1 in the l_2 norm. The input signal $x \in R^{256}$ was also generated randomly with gaussian distribution. Tropp's[6] OMP, Needell's[10] ROMP, Donoho's[12] StOMP and Needell &Topp's[13] CoSaMP algorithms are subjected to MATLAB and the performances of those algorithms are evaluated interms of mean square error computed by $\|x-\hat{x}\|_{l2}$ over 1000 trial signals. For halting, Zero threshold uniformly fixed for all algorithms as $10\wedge(-4)$ and maximum iteration are fixed as per the recommendations of the authors. For stomp, threshold is fixed according to normal inverse cumulative distribution function. Figure1 shows Mean Square Error versus Sparsity for all the algorithms. The result of OMP with respect to Gaussian amplitude is quite encouraging.

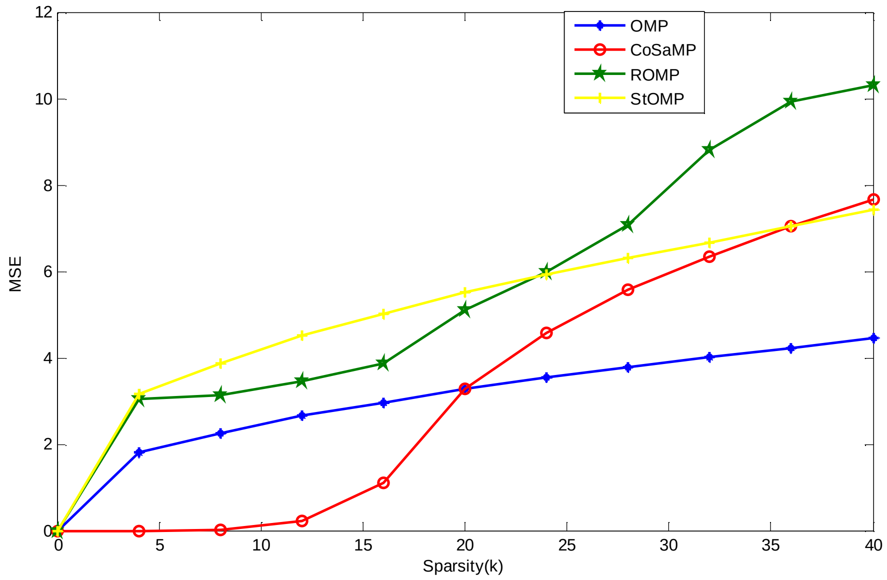

Figure1: Mean Square Error of The Reconstructed Signal Versus Sparsity

Number of iterations required by algorithm for various sparsity levels is presented in table1. From the table it is observed that OMP requires iterations on the order of k. But when the signal is not very sparse, OMP may poor choice because the cost of orthogonalization increases quadratically with the number of iterations. CoSaMP requires in the order of $5k$ iterations. But ROMP and StOMP terminates with 3 and 2 iterations respectively. It was studied OMP is efficient when the signal is highly sparse. In cases where unknown vector x is not very sparse, OMP is not a practical algorithm to find the solution vector x because the computational complexity of OMP increases linearly with the number of nonzeros k. Therefore, OMP has been revised in a way to find less accurate solutions for such large scale underdetermined linear systems in a reasonable time using ROMP or StOMP.

Table 1: Comparison of iterations for OMP, CoSaMP, ROMP and StOMP algorithms for various sparsity level

Sparsity	**No of Iterations**			
	OMP	**CoSaMP**	**ROMP**	**StOMP**
4	4	3	3	2
8	6	4	3	2
12	8	15	3	2
16	10	39	3	2
20	12	64	3	2
24	14	85	3	2
28	16	105	3	2

32	18	124	3	2
36	20	143	3	2
40	22	161	3	2

Among all, the OMP algorithm is significant interms of simplicity and competitive reconstruction performance but with complexity as linear function of sparsity. The guarantee to recover sparse signal by OMP algorithm is already ascertained by Ganesh et al.[15] for block sparse signal now the same is proved for sparse signal.

Theorem

Suppose that a signal is k sparse and $k < 0.5(1+\frac{1}{\mu})$ where μ incoherence between the columns of ϕ, then OMP guarantees to find the k columns of ϕ.

Proof:

Normalize each column of matrix ϕ with a coefficient diagonal matrix C so that $\phi = D.C$ where D^{dxn} has unit norm columns. C diagonal matrix constructed by length of each column in ϕ

The guarantee to find the k columns

$\langle a_1 y\rangle > \langle a_t y\rangle$ where $t>k$

The lower bound of $\langle a_1 y\rangle$ is given by

$\langle a_1 y\rangle$

$=a_1^T y$

$= a_1^T \sum_{i=1}^{k} w_i z_i$ where w_i - length of the vector &

z_i - unit vector

$= w_1 a_1^T z_1 + \sum_{i-2}^{k} w_i a_i^T z_i$

$\geq w_1 a_1^T z_1 - \sum_{i=2}^{k} w_i a_1^T z_i$

$\geq w_i a_1^T z_i - w_i \mu(1-k)$

$\geq w_i(1-\mu(k-1))$

The upper of $\langle a_t y\rangle$ is given by

$\langle a_t y\rangle = \sum_{i=1}^{k} w_i a_t^T z_i = k\mu w_i$

using the upper and lower bounds

$$w_i - w_i \mu(k-1) \geq k\mu w_i$$

$$1+\mu \geq 2k\mu$$

$$\frac{1+\mu}{\mu} \geq 2k$$

$$\frac{1}{2}(1+\frac{1}{\mu}) \geq k$$

which leads to the condition $\langle a_1 y \rangle > \langle a_t y \rangle$. Hence the theorem is proved. Next uniqueness of the solution guaranteed by OMP by the following theorem

Theorem
If y can be represented as a linear sum of k distinct columns from ϕ such that $k < 0.5 spark(\phi)$, then sparse decomposition is necessarily unique that is y cannot be represented as linear sum of elements from a different set of k columns.

Proof:
Les us consider there exist two solutions x_1 and x_2 with $\|x_1\|_0 = \|x_2\|_0 = k$ as $y = \phi x_1$ and $y = \phi x_2$ respectively. Which yields $\phi x_1 = \phi x_2$ therefore the vector $\phi(x_1 - x_2) = 0$ is in the null space of ϕ. Since both x_1, x_2 contain only k nonzero entries, x_1 - x_2 contains at most *2k* nonzero entries. It implies that, linear combination of *2k* columns of ϕ is zero means *2k* columns of ϕ are linearly dependent which is less than spark (ϕ) which is contradiction to the definition of spark. Therefore $\phi(x_1 - x_2) = 0$ if and only if $x_1 = x_2$. Hence the solution is unique.

Conclusion
This paper reviews the various greedy iterative algorithms for obtaining the sparse solutions. The performance of various algorithms is compared on the basis of mean square error displaying the efficacy of OMP over other algorithms. If the signal is highly sparse then OMP is the best choice. It represents the uniqueness and correctness guarantee of OMP algorithm in terms of mutual incoherence.

References

[1] D. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[2] R. Baraniuk, M. Davenport, R. Devore, and M. Wakin, "A Simple Proof of the Restricted Isometry Property for Random Matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, 2008.

[3] E. J. Candes, “The Restricted Isometry Property and its Implications for Compressed Sensing”, Comptes Rendus Mathematique, vol.346, no.9-10,pp. 589-592, 2008.

[4] W. Dai and O. Milenkovic, “Subspace pursuit for compressive sensing signal reconstruction,” *IEEE Trans on Inform. Theory*, vol. 55, no. 5, pp. 2230–2249, May 2009.

[5] D. L. Donoho and Huo, “Uncertainty Principles and Ideal Atomic Decomposition”, IEEE trans.Inf. Theory, 47, pp2845 -2862,2001.

[6] J. A. Tropp and A. C. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit,” *IEEE Trans. Inform. Theory*, vol. 53, pp. 4655–4666, Dec. 2007.

[7] J. A.Tropp, “Greed is good: Algorithmic results for sparse approximation”, *IEEE Trans. Inf. Theory* 50, pp.2231-2242, 2004.

[8] J.Wang and b.Shim, “On the recovery Limit of Sparse Signals using Orthogonal Matching Pursuit”, Signal Processing, *IEEE Trans* Vol16, Issue 9,pp. 4973-4976,2012.

[9] A. Majumdar, N. Krishan, S. R. B. Pillai, R.Velmurugan, “ Extension to Orthogonal Matching Pursuit for Compressed Sensing”, National Conference on communications (NCC), pp28-30 Jan2011.

[10] D. Needell and R. Vershynin, “Signal Recovery from Incomplete and Inaccurate Measurements via Regularized Orthogonal Matching Pursuit,” *IEEE J. Sel. Topic Signal Process.*, vol. 4, no. 2, pp. 310–316, April 2010.

[11] Deanna Needell, “Topics in Compressed Sensing”, Dissertation for Doctor of philosophy in Mathematics, University of California, Davis, 2009.

[12] D. L. Donoho, I. Drori, Y. Tsaig, and J. L. Starck, “Sparse Solution of Underdetermined Linear Equations by Stagewise Orthogonal Matching Pursuit”, Citeseer, 2006.

[13] D. Needell and J. A. Tropp, “CoSaMP: Iterative Signal Recovery from Incomplete and Inaccurate Samples,” Applied and Computational Harmonic Analysis, vol. 26, no. 3, pp. 301–321, 2009.

[14] S. Satpathi and M. Chakraborty, “On the number of iterations for convergence of CoSaMP and SP algorithm” . CSmath.IT ar Xiv:1404.4927 v3 Apr. 2014.

[15] A. Ganesh, Z. Zhou and Yi Ma, “Separation of a subspace signal: algorithms and conditions”, In IEEE International Conference on Acoustics, Speech and Signal Processing , pp.3141—3144, 2009.