Sparse Signal Recovery based on Hybrid Genetic Algorithm

Jawad Ali Shah[1*], I.M. Qureshi[2], Amir A. Khaliq[1] and Hammad Omer[3]

[1]Department of Electronic Engineering, Faculty of Engineering and Technology International Islamic University Islamabad, PAKISTAN
[2]Department of Electronic Engineering, Air University, Institute of Signals, Systems and Soft computing (ISSS), Islamabad, PAKISTAN
[3]Department of Electrical Engineering, COMSATS Institute of Information Technology, Islamabad, PAKISTAN

Abstract

This paper introduces a novel technique of recovering an S-sparse signal from lesser number of random measurements (under sampled signal) using hybrid genetic algorithm (GA). The proposed method uses a chromosome-based cross over to produce the off springs with better fitness. The l_0minimizationconstraint is directly incorporated in population to achieve the desired sparsity level. The slow and premature convergence of GA is prevented with the help of modified parallel coordinate descent (PCD) algorithm. This hybrid mechanism of GA along with modified PCD makes the sparse signal recovery faster and accurate. The performance of the proposed algorithm is verified by comparing its results with those of PCD and separable surrogate functional (SSF) algorithms. The effectiveness of hybrid GA is further demonstrated experimentally by faithfully recovering a synthetic one-dimensional sparse signal. The simulation results show that an accurate sparse signal reconstruction is possible with hybrid GA using only a small number of observations.

Keywords: Sparse signal recovery, compressed sensing, genetic algorithms, iterative shrinkage.

Introduction

A sparse signal contains less information as compared to its size or dimensions with the energy concentrated in only a few components[1]. Mathematically, the sparsity of a signal is represented by a number linked to the l_0 norm of the signal[2]. Assuming $0^0 = 0$, the l_0 norm of a signal or vector $x \in R^n$ is defined as: $\|x\|_0 = \sum_{i=1}^{n} |x_i|^0$. Strictly speaking l_0 is not a norm as it does not satisfy the homogeneity property[3]. Thus the counting function $\|.\|_0 : R^n \to R$ gives the number of non-zero elements in its argument. A signal x is said to be *S-sparse* if $\|x\|_0 \leq S$. For two vectors $x_1, x_2 \in R^n$, we say x_2 is sparser than x_1 if $\|x_2\|_0 < \|x_1\|_0$. If a signal is not sparse in the canonical domain, then depending on the application, one can use a sparsifying transform like wavelet, discrete cosine etc. to get a sparse version of the signal in that specific domain[4]. Sparse signals have the advantages of fast computation and less storage requirements. Practical signals are not exactly sparse, but are compressible. Compressible signals have coefficients that tend to decay exponentially when sorted in a descending order by magnitude[1].

In most of the engineering and signal processing applications such as filtering, linear regression, least square estimation etc. the discrete linear system takes on an input signal $x \in R^n$ to produce another signal $y \in R^m$ through some linear transformation[5]. Mathematically,
$y_i = \sum_{j=1}^{n} \varphi_{ij} x_j \quad 1 \leq i \leq m$ Or in matrix form $y = \Phi x$ (1)

Where $\Phi : R^n \to R^m$ is the desired transformation matrix. We assume it to have a full rank throughout.

The problem of sparse signal recovery comes under the general umbrella of linear inverse problems[6]. Given the observation signal y and the measurement matrix Φ, the algorithms of linear inverse problems look for the best solution under certain constraints. If the matrix Φ is square and invertible, then the unique solution can be found through matrix inversion. However, in most of the practical scenarios the sensing matrix Φ is ill-conditioned with $m \ll n$ which leads to an under-determined system of linear equations. This can also be thought of as a dimensionality reduction problem. In some applications related to signal acquisition (such as magnetic resonance imaging (MRI) etc.), it may be highly desirable to have Φ as fat matrix (less number of rows than columns) resulting in a small number of samples thus reducing the acquisition and scan time[7]. However, for such an underdetermined system, a unique solution does not exist as there are more number of unknowns than equations or equivalently the matrix Φ has a null space. In order to get the original signal back, we need to impose proper constraints (priors) on the solution. One of the many possible options will be to compute the minimum norm solution using pseudo-inverse. This classical estimator uses the least square technique that minimizes the energy or l_2 (Euclidian) norm of the estimated signal and has a closed form:

$$\hat{x} = \arg\min_{x} \|y - \Phi x\|_2^2 = \Phi^T (\Phi \Phi^T)^{-1} y \quad (2)$$

The minimum norm solution is generally non-sparse with energy distributed among large number of components of the estimated signal[8]. As an example, consider $\Phi = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ and $y = \begin{bmatrix} 1 & 0 \end{bmatrix}^T$ then the minimum norm solution is

$\hat{x} = [2/3 \quad -1/3 \quad 1/3]^T$ while the sparse solution is $x_s = [1 \quad 0 \quad 0]^T$ with sparsity $S=1$.

The problem of sparse signal estimation arises in many important areas of signal and image processing, neural networks, statistics and linear regression, to name a few. The challenging job in sparse signal approximation is to find a low complexity algorithm with acceptable performance. This involves finding both the locations of the nonzero elements and their values in the recovered or target signal. The term l_0 minimization refers to the problem of finding the sparsest solution given the observation vector y, the sparsity S and the measurement matrix Φ. Thus we can formulate the sparse signal reconstruction problem (for noise free case) as:

$$\hat{x} = \arg\min_{x} \|y - \Phi x\|_2^2 \text{ subject to } \|x\|_0 \leq S \tag{3}$$

The field of sparse signal approximation was revolutionized by the idea of compressed sensing or sampling (CS)[9-11]. CS has challenged the well-known signal acquisition Nyquist sampling theory with the claim that sparse signals can be recovered exactly from far less measurements than the conventional methods[4]. It addresses the issue of compression algorithms such as JPEG, JPEG-2000 where a large number of samples of a compressible signal are acquired first and after sparsifying it in a certain domain, most of the coefficients are discarded. CS replaces the conventional sample-then-compress framework with the idea to acquire less but sufficient number of samples (non-adaptive linear measurements) at the data acquisition that can faithfully reproduce the original signal using some reconstruction algorithm. This requires certain conditions to be imposed on the sensing matrix.

For the sparse solution to be unique, the sensing matrix Φ should map two different S-sparse signals (x_1 and x_2) to different locations (y_1 and y_2) after transformation by preserving their Euclidian distance. i.e. $\|x_1 - x_2\|_2^2 \approx \|\Phi x_1 - \Phi x_2\|_2^2 = \|y_1 - y_2\|_2^2$. This property is known as restricted isometry property (RIP)[4,8]. Random Gaussian and Bernoulli matrices have been shown to obey RIP if the number of rows of the matrix satisfies the criteria of $m \geq cS \log(n/S)$ for some constant c[12]. So in this context, the random Gaussian measurement matrices can be considered as universal. Thus CS theory has sparse estimation as a sub-problem besides the stable design of the measurement matrix[13].

Solving the l_0 minimization problem defined in (3) directly to find the sparse signal estimation is NP-hard and is practically not tractable as it requires exhaustive search of $\binom{m}{S}$ possible combinations for the locations of nonzero entries in the estimated vector[13]. In this paper we propose a new technique to solve the problem given in (3) using the combination of GA along with a randomized version of iterative thresholding algorithm to find the sparse solution.

Sparse signal approximation algorithms

Various computational techniques have been used to make the sparse reconstruction problems tractable. One simple approach is to recast the non-convex problem of (3) as a convex optimization problem. If proper measurements are taken, the l_0 norm in formulation (3) can be replaced with the l_1 norm making it a convex optimization problem. Least absolute shrinkage and selection operator (LASSO)[14] recast the problem of (3) for parameter α as:

$$\hat{x} = \arg\min_{x} \|y - \Phi x\|_2^2 \text{ subject to } \|x\|_1 \leq \alpha \tag{4}$$

Basis Pursuit (BP)[15] reformulates the sparse signal estimation by the following program:

$$\hat{x} = \arg\min_{x} \|x\|_1 \text{subject to } y = \Phi x \tag{5}$$

Various software tools such as CVX[16], $l_1 \; magic$[17] are available online that can be used for solving the formulation of (4) and (5) as a convex optimization problem.

Different variants of greedy pursuit algorithms such as Orthogonal Matching Pursuit (OMP)[18] and compressive sampling matching pursuit (CoSaMP)[19] have been proposed by the researchers which iteratively estimate a better version of the sparse signal. For approximating an S sparse signal, these algorithms look for the best S columns of the measurement matrix that explains the target vector.

$$y = \sum_{i=1}^{n} \varphi_i x_i = \sum_{<S>} \varphi_i x_i$$

Where φ_i is the ith column (or atom) of the sensing matrix (dictionary) Φ.

The S best columns are found through a sequential search and the nonzero entries of the estimated vector are computed by solving the overdetermined system of linear equations[20].

Another popular class of sparse signal approximation is iterative shrinkage. These methods do not involve matrix inversion and the iterations comprise of multiplication of the measurement matrix Φ (or Φ^T) with a vector[21]. Iterative hard thresholding[22] (IHT), fast iterative shrinkage thresholding[6] (FISTA), separable surrogate functionals (SSF) and parallel coordinate descent[21] (PCD) belong to the family of algorithms which minimizes the following objective function for sparse signal approximation:

$$\hat{x} = \arg\min_{x} (\|y - \Phi x\|_2^2 + \beta \|x\|_1) \tag{6}$$

Where $\beta \geq 0$ is known as the regularization parameter whose value determines the sparsity of the solution.

Other major classes of algorithms related to sparse signal reconstruction include Bayesian framework[3,23], nonconvex optimization[24], brute force[25] and smooth l_0 norm[26].

Genetic Algorithm

GA is a general purpose heuristic optimization technique which is based on the principles of genetics. The algorithm starts with a random group of interbreeding individuals known as population or generation. Each individual in the population acts as a candidate solution and is referred to as a chromosome. The elements of a chromosome are called genes. The effectiveness or cost of each chromosome is determined through a fitness function. With the help of cross over, the genes of different chromosomes (parents) can be combined in a variety of way to produce the offsprings having different fitness values. The new population is formed with the natural selection by combining the best (in terms of fitness) parents and offsprings. In this way the algorithm proceeds to search for the best candidate solution (chromosome)[27]. Depending onthe applications, different variants of GA have been developed.

Heuristic algorithms are preferred for the problems which are NP-hard. However, when the optimization involves constraints, the application of GA becomes more difficult as the cross overs are blind to the constraints[28]. Thus solving (3) with the conventional GA is not possible as offsprings may not follow the sparsity constraint even if it is fulfilled by the parents.

Constraints can be incorporated in the fitness function (indirect constraint handling) as well as in the chromosomes (direct constraint handling). However, indirect constraint handling does not work well for the sparse problems[28]. Therefore in our proposed algorithm, the direct constraint handling has been used to ensure the desired sparsity level before and after cross over through hard thresholding.

In spite of many achievements, one of the main problems of GA is the premature convergence whichis related to the loss of genetic diversity of the population[29]. One way to avoid this problem is through mutation. However, for sparse signal recovery, the ordinary mutation will not work, making the chromosome denser and compromising the sparsity constraint. The proposed method avoids this problem by using the modified PCD algorithm when the population tends to converge prematurely.

Modified Parallel Coordinate Descent (PCD) Algorithm

PCD is an algorithm that minimizes the formulation given in (6) with the idea taken from coordinate descent method in which the cost function minimizes one coordinate at a time. In order to get the new solution, it updates all the coefficients in parallel instead of doing it sequentially[30]. The update equation of the algorithm is given by[21]:

$$x_k = x_{k-1} + \gamma(e_s - x_{k-1}) \tag{7}$$

Where γ is a constant and is computed usually through line search. In the proposed algorithm we replace it with a random number. The starting value of the solution can be either an estimate of the least square solution or a zero vector. In (7) the term e_s is first computed by the equation:

$$e_s = \rho_{T,\beta}\left(x_{k-1} + \mathrm{diag}(\Phi^T\Phi)^{-1}\Phi^T r_{k-1}\right) \tag{8}$$

Here $r_k = y - \Phi x_k$ is the residue and $\rho_{T,\beta}$ is the soft-thresholding or shrinkage operator defined by:

$$\rho_{T,\beta}(u) = \begin{cases} 0 & if \ |u| \leq \beta \\ \frac{(|u|-\beta)}{|u|}u & if \ |u| \geq \beta \end{cases} \tag{9}$$

The hybrid algorithm proposed in this paper modifies the regular PCD to accelerate the convergence of GA. In order for PCD to interact with GA, randomness is introduced in it at various levels. Instead of taking the diagonal values of $(\Phi^T\Phi)^{-1}$ as the weights, the proposed algorithm uses random weights thus avoiding the matrix inversion. PCD is used to update achromosome when the fitness of the best chromosome does not change in a few consecutive iterations thereby preventing the convergence issue. So the interaction of PCD in the proposed algorithm is again a random phenomenon. Every time the modified PCD algorithm is accessed by GA, the residue is computed using the currently best chromosome while in (7) and (8) a chromosome from the current generation is selected randomly to replace r_{k-1}. These modifications allow the hybrid GA to recover the sparse signal with an acceptable level of accuracy.

Proposed Algorithm

Figure-1 lists the detailed description of the proposed hybrid genetic algorithm. The symbol $card(z)$ is used to denote the cardinality of the vector z while $indx$ represents a vector containing the indices of the array when sorted in the descending order. The notation $\otimes$ represents the element-by-element product and the operator $\lfloor g \rfloor_s$ sets all except the *S* largest elements of vector g to zero.

The program can be halted either after achieving the desired fitness value evaluated in step-2 (i.e. $f_i < threshold$) or after a fixed number of iterations.

Results and Discussion

We use Matlab as a simulation tool to verify the validity of the proposed algorithm. Based on the theory of CS, a random matrix with $m = 256$ and $n = 512$ is taken as them easurement matrix Φ. First an $n \times n$ square matrix is constructed with random entries of ± 1. The sensing matrix is then formed bytaking m rows from the matrix generated through Gram-Schmidt orthogonalizationfrom the $n \times n$ square matrix.The

sparsity level is set to *S*=85 during the simulation. A one dimensional *S*-sparse test signal $x_0 \in R^{512}$ is produced having random support and magnitude for non-zero entries. All the results are produced for maximum of 500 iterations to reconstruct the desired signal from an under sampled measurement vector $y = \Phi x_0 \in R^{256}$. For GA, the population size consisting of *N*=100 chromosomes is formed with each chromosome comprising of 512 genes. The initial population is produced by merging the minimum norm solution $(\hat{x}_0 = \Phi^+ y)$ with the randomly scaled basis of the null space of Φ. The desired sparsity level is randomly produced in the chromosomes belonging to the initial population to ensure $\|g_i\|_0 \leq S \ \forall \ i \in \{1,2,...N\}$. For comparison of reconstruction accuracy, we use original PCD and SSF algorithms to recover the same signal from the same measurements with thresholding parameter $\beta = 0.001$. These algorithms are initiated with the best chromosomes found in the initial population. The normalized mean-square-error(MSE) is determined at *jth* iteration as

$$\frac{\|\hat{x}_j - x_0\|_2^2}{\|x_0\|_2^2} \quad (10)$$

Input: Sensing matrix $\Phi \in R^{m \times n}$, measurement vector $y \in R^m$, sparsity level *S*, population size *N*, thresholding parameter β for modified PCD only

Output: An *S*-sparse vector $x \in R^n$

   1)   **Population Generation:** Randomly generate *N* chromosomes

$G = [g_1, g_2, ... g_N], g_i \in R^n$ and $card(g_i) \leq S \quad \forall \ i \leq 1 \leq N$

   2)   **Fitness Evaluation of parents & Sorting:** Evaluate the fitness of each chromosome based on (3) and sort them in the descending order (the lower the fitness, the better the chromosome)

$f_p = fit(g_1, g_2, ... g_N)$

$\quad = [f_{p1}, f_{p2}, ... f_{pN}], f_{pi} = (\Phi g_i - y)^T (\Phi g_i - y)$

$[f_{ps} \ indx \ ] = sort(f_p, descend)$

$f_{ps} = [f_1, f_2, ... f_N]$ With $f_1 < f_2 ... ... < f_N$

$G_s = G(indx)$

$\quad = [g_{s1}, g_{s2}, ... g_{sN}]$ where g_{s1} has fitnes f_i

   3)   **Modified PCD:** If f_1 remains the same during the specified consecutive iterations then execute:

$e_s = \rho_{T,\beta} (g_{srnd} + w_{rand} \otimes (\Phi^T(y - \Phi g_{s1})))$

$g_{srnd} \in G_s$ is a randomly selected chromosome from sorted population

$w_{rand} \in R^n$ is randomly generated vector.

$g_{s2} = g_{s2} + \gamma(e_s - g_{srnd}), \gamma =$ randomly generated number

$g_{s2} = \lfloor g_{s2} \rfloor_s$

   4)   **Cross over:** offsprings of size half of the population are generated in random fashion:

$C = xover(G_s)$

$\quad = [c_1, c_2, ... c_{\frac{N}{2}}]$

$c_j = \lfloor g_{sj} + \gamma(g_{sj} - g_{srndi}) \rfloor_s \qquad 1 \leq j \leq \frac{N}{2} \ and \ 1 \leq i \leq N$

   5)   **Fitness Evaluation of children & Sorting:** Same as step-2 but executed for offspring.

$f_c = fit\left(c_1, c_2, ... c_{\frac{N}{2}}\right)$

$[f_{cs} \ indx \ ] = sort(f_c, descend)$

$C_s = C(indx)$

$\quad = \left[c_{s1}, c_{s2}, ... c_{s\frac{N}{2}}\right]$

   6)   **New population:** Generate new population using half of the best parents and all children.

$G = \left[g_{s1}, g_{s2}, ... g_{s\frac{N}{2}} c_{s1}, c_{s2}, ... c_{s\frac{N}{2}}\right]$

Repeat (2) - (6) until the stopping criteria meet.

   7)   **Output:** The chromosome with best fitness is the candidate solution $x = g_{s1}$

Figure-1
The proposed hybrid genetic algorithm for sparse signal recovery

---

Results based on various parameters are shown in the Figures. Figure-2 depicts the slow convergence problem of GA. The modified PCD algorithm is not accessed during the first 30 iterations so the value of the cost function decreases in the same way for both algorithms. After 30th iteration, modified PCD is randomly used by the GA to decrease the fitness of a chromosome when required thus accelerating the convergence. The correlation, MSEand fitness values achieved for the reconstructed signals are given in table-1.

Table-1
Parameters achieved by PCD, GA and proposed algorithm

Algorithm Used	Correlation	MSE	Fitness
PCD	0.8752	0.2516	2.4620
GA	0.6102	0.6276	11.9948
Hybrid GA	1.0000	8.0656e-06	1.5115e-04

Figure-3 shows the performance comparison of the hybrid GA, SSF and PCD in terms of mean square error computed through (10) clearly indicating that the proposed algorithm is reconstructing the signal more effectively as compared to the other two algorithms.

Figure-4 shows the signal approximation using PCD alone while Figure-5 depicts the signal reconstruction of GA for the same number of iterations. The sparse signal recovery with the proposed hybrid GA shown in Figure-6 indicates that not only the support of the signal is recovered accurately but also its magnitude is approximated to an acceptable limit.

Conclusion

The direct application of genetic algorithm to solve the sparse signal reconstruction problem has the issue of slow and premature convergence. However, using a local optimizer with GA can easily overcome this inefficiency. Incorporating sparsity constraint handling in GA is also challenging. The proposedalgorithm uses a modified version of PCD with direct sparsity constraint handling for sparse signal approximation to address these issues. The simulation results show that the proposed hybrid genetic algorithms can recover the sparse signal accurately from less number of samples. For a fixed number of iterations, the reconstruction accuracy of hybrid algorithm is better than the recovery of GA and PCD.

A lot of work can be done to improve the performance of GA for sparse signal recovery. One way will be to devise a smart method of mutation to address the conversion issue of GA. Performance of GA integrated with other optimizers such as SSF and FISTA can be explored as well. The analysis of the algorithm under noisy cases can also be investigated.
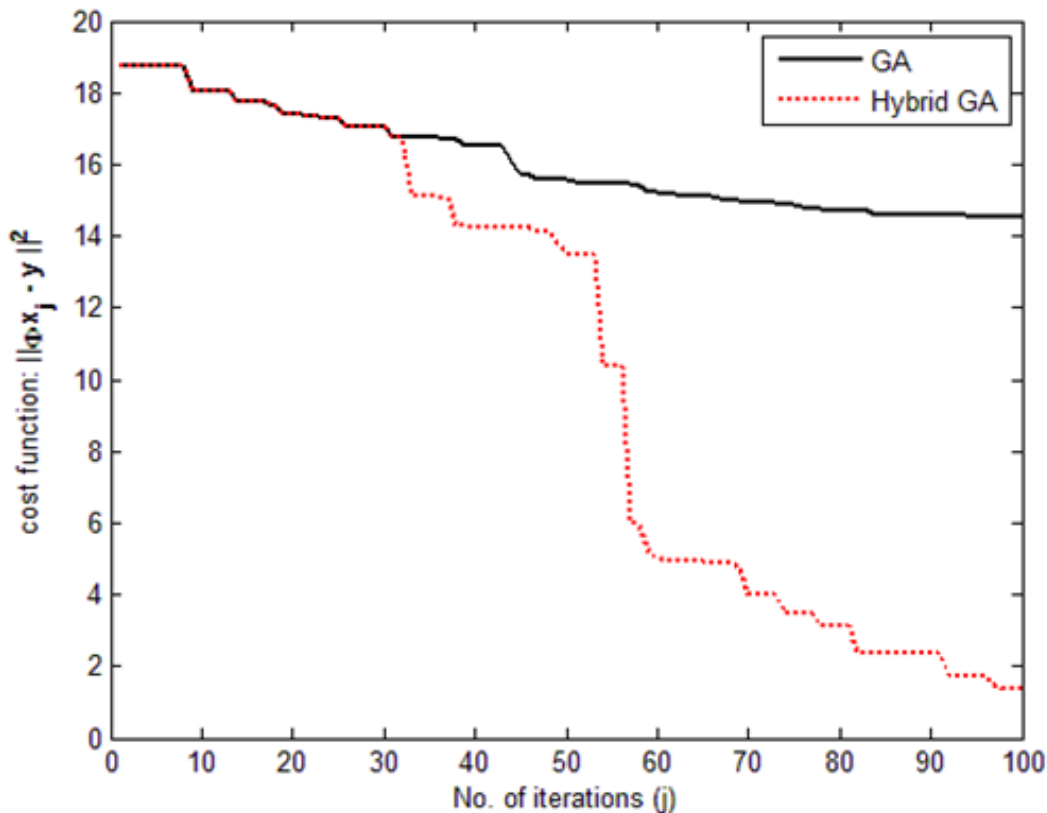

Figure-2
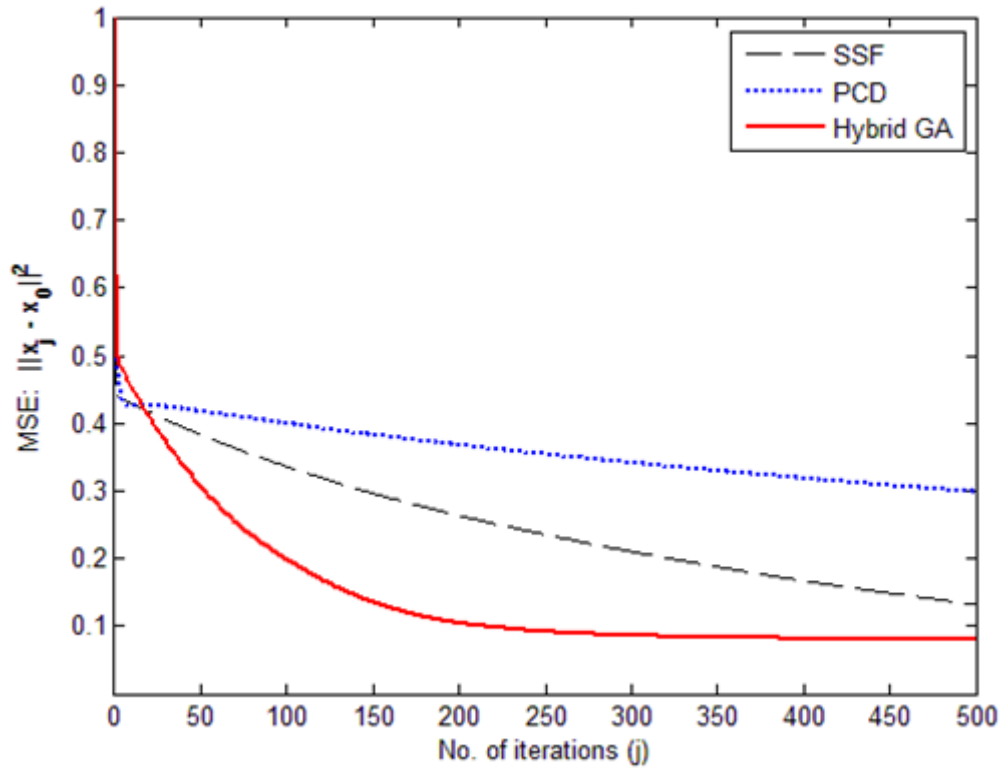**Improvement in convergence of GA using proposed method**

Figure-3
Comparison (based on MSE) of proposed hybrid GA with SSF & PCD
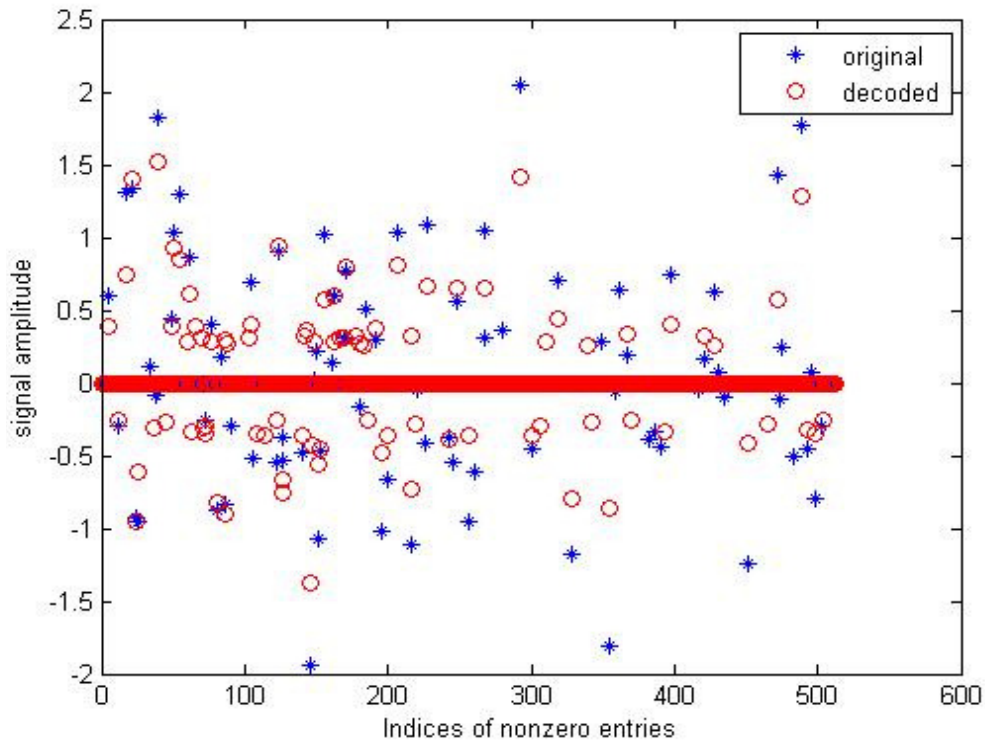

Figure-4
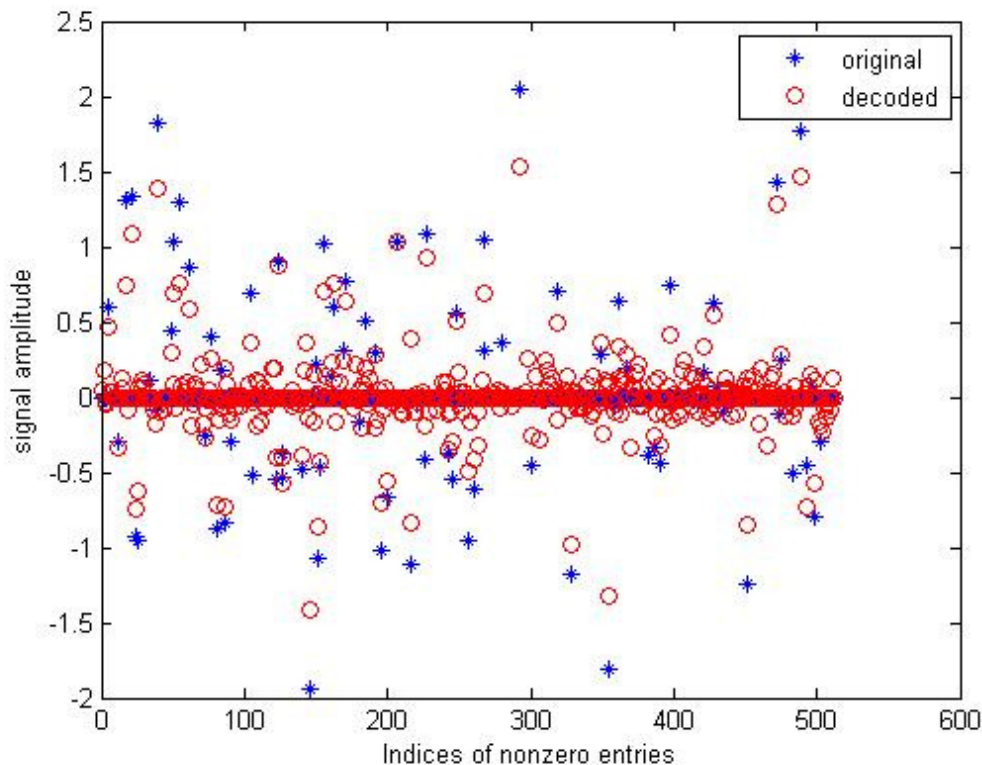**Signal reconstruction through PCD**

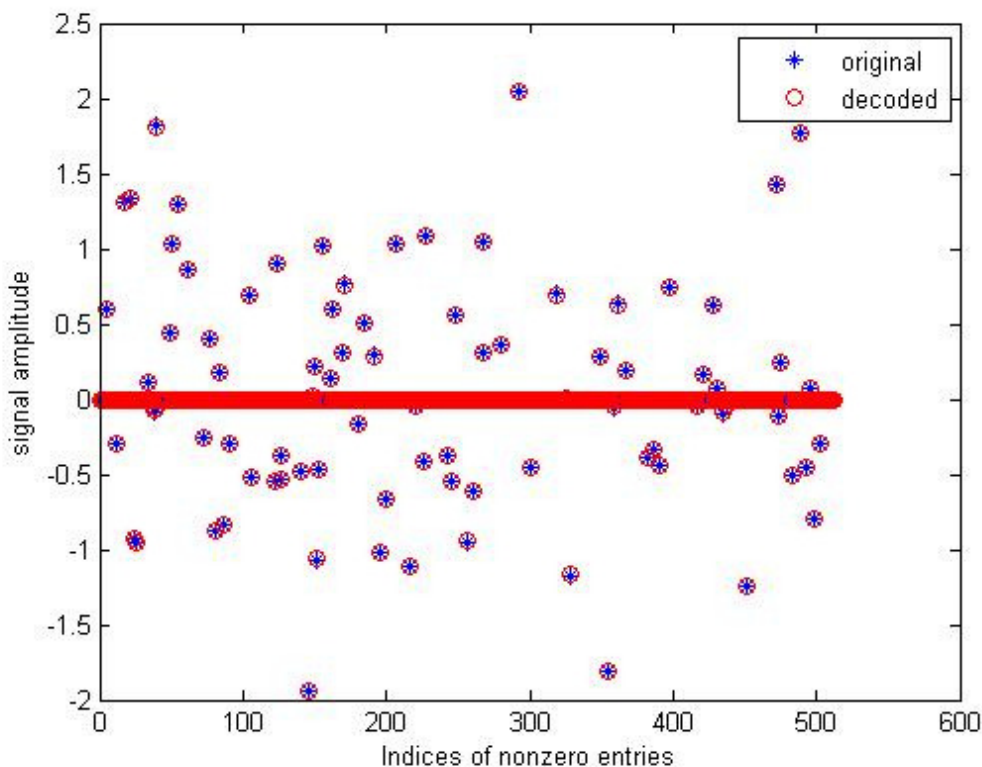**Figure-5**
Signal reconstruction through GA

Figure-6
Signal reconstruction with the proposed algorithm

References

1. Needell D., Tropp J. and Vershynin R. Greedy signal recovery review, *Signals, Systems and Computers, 2nd Asilomar Conference*, 1048-1050, IEEE **(2008)**

2. Tropp J.A. and Wright S.J., Computational methods for sparse solution of linear inverse problems, *Proc. IEEE,* **98(6)**, 948 -958 **(2010)**

3. Olver Shakiban, Applied linear algebra, ISBN-13: 9780131473829 **(2005)**

4. Cand E.J., egrave;s and M.B. Wakin, An introduction to compressive sampling, *IEEE Signal Process Mag.*, **25(2)**, 21-30 **(2008)**

5. Schmidt M., Least squares optimization with 11-norm regularization, *Technical report*, **(2005)**

6. Beck Amir and Marc Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, *SIAM Journal on Imaging Sciences*, **2(1)**, 183-202 **(2009)**

7. Lustig M., Donoho D.L., Santos J.M. and Pauly J.M., Compressed sensing MRI, *Signal Processing Magazine, IEEE*, **25(2)**, 72-82 **(2008)**

8. Gorodnitsky Irina F., Bhaskar D. Rao, Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm, *Signal Processing, IEEE Transactions,* **45(3)**, 600-616 **(1997)**

9. Candès E., Romberg J. and Tao T., Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, *IEEE Trans.Inform. Theory,* **52(2),** 489–509 **(2006)**

10. Candès E., Romberg J. and Tao T., Stable signal recovery from incomplete and inaccurate measurements, *Comm. Pure Appl. Math*. **59(8)**, 1207–1223 **(2006)**

11. D. Donoho, Compressed sensing, *IEEE Trans. Inform. Theory,* **52(4)**,1289–1306 **(2006)**

12. Mendelson S., Pajor A. and Tomczak-Jaegermann N., Uniform uncertainty principle for Bernoulli and subgaussian ensembles. *Constructive Approximation*, **28(3),** 277-289 **(2008)**

13. Baraniuk, Richard G. Compressive sensing [lecture notes], *Signal Processing Magazine, IEEE,* **24(4)**, 118-121 **(2007)**

14. Tibshirani R., Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society,* Series B (Methodological), 267-288 **(1996)**

15. Chen S.S., Donoho D.L. and Saunders M.A., Atomic decomposition by basis pursuit, *SIAM journal on scientific computing*, **20(1)**, 33-61 **(1998)**

16. Grant Michael and Stephen Boyd, CVX: Matlab software for disciplined convex programming (web page and software), URL http://stanford. edu/boyd/cvx **(2008)**

17. Candes E., Romberg J., 11-magic: Recovery of sparse signals via convex programming, URL: www. acm. caltech. edu/l1magic/downloads/l1magic. pdf, 4 **(2005)**

18. Tropp J.A. and Gilbert A.C., Signal recovery from random measurements via orthogonal matching pursuit, *Information Theory, IEEE Transaction*, **53(12)**, 4655-4666 **(2007)**

19. Needell D. and Tropp J.A., Cosamp: iterative signal recovery from incomplete and inaccurate samples, *Communications of the ACM*, **53(12),** 93-100 **(2010)**

20. Tropp J.A., Greed is good: Algorithmic results for sparse approximation, *Information Theory, IEEE Transactions*, **50(10)**, 2231-2242 **(2004)**

21. Elad M., *Sparse and redundant representations: from theory to applications in signal and image processing,* Springer, **(2010)**

22. Blumensath T. and Davies M.E. Iterative hard thresholding for compressed sensing. *Applied and Computational Harmonic Analysis*, **27(3)**, 265-274 **(2009)**

23. Wipf David P. and Bhaskar D. Rao., Sparse Bayesian learning for basis selection, *Signal Processing, IEEE Transactions,* **52(8)**, 2153-2164 **(2004)**

24. Chartrand R., Exact reconstruction of sparse signals via nonconvex minimization, *Signal Processing Letters, IEEE,* **14(10),** 707-710 **(2007)**

25. Hastie T., Tibshirani R. and Friedman J., *Linear Methods for Regression*, Springer New York, **(2009)**

26. Mohimani, G. Hosein, MassoudBabaie-Zadeh, Christian Jutten, Complex-valued sparse representation based on smoothed ℓ_0 norm, *Acoustics, Speech and Signal Processing, IEEE International Conference*, 3881-3884 **(2008)**

27. Haupt Randy L., Sue Ellen Haupt. *Practical genetic algorithms,* John Wiley & Sons, **(2004)**

28. Craenen B.G.W., A.E. Eiben, E. Marchiori, How to handle constraints with evolutionary algorithms, *Practical Handbook of Genetic Algorithms: Application*s, 341-361,**(2001)**

29. Rocha, Miguel, José Neves, Preventing premature convergence to local optima in genetic algorithms via random offspring generation, *Multiple Approaches to Intelligent Systems, Springer Berlin Heidelberg*, 127-136 **(1999)**

30. Elad M., Why simple shrinkage is still relevant for redundant representations?, *Information Theory, IEEE Transactions*, **52(12)**, 5559-5569 **(2006)**