

Sparse Signal Recovery Based on Concentration Measures and Genetic Algorithm

Miloš Brajović, *Student Member, IEEE*, Budimir Lutovac, *Member, IEEE*, Irena Orović, *Member, IEEE*, Miloš Daković, *Member, IEEE* and Srdjan Stanković, *Senior Member, IEEE*

Abstract — In this paper genetic algorithm is applied in the reconstruction of signal with missing samples, sparse in a transformation domain. DFT is considered as a domain of sparsity, without loss of generality. The reconstruction is performed as a minimization of the ℓ_1 -norm based concentration measure, with missing samples acting as minimization variables. Parameters of the genetic algorithm are set based on a numerical study, taking into account the nature of the considered minimization problem. The proposed genetic algorithm parameters setup provides an efficient reconstruction of missing samples under the assumption that the standard reconstruction conditions are met.

Keywords — Compressed sensing, Concentration measures, Digital signal processing, Genetic algorithms, Sparse signal processing.

I. INTRODUCTION

COMPRESSED sensing (CS) and sparse signal processing are closely related areas in the modern signal processing [1]-[17]. The reconstruction of missing samples is possible under the assumption that the signal is sparse in a transformation domain, meaning that it can be represented with a small number of non-zero transformation coefficients [1]-[9]. This number of non-zero transform coefficients is referred to as sparsity.

In the compressed sensing, missing samples arise as a consequence of a specifically chosen sampling strategy, in order to reduce the number of samples needed for storage and transmission of the signal, aiming to preserve the full information contained within the signal at the same time [1], [2]. In some applications signal samples can be lost during the transmission over the communication channels. In many real scenarios, signal samples can be corrupted with strong disturbances (noise) and consequently, it is more adequate to intentionally omit this corrupted data rather than engage it in the further processing. Illustrative example is the application of the robust signal processing and L-statistics [10], [16], [17]. Signal reconstruction algorithms and methods developed within the framework of the compressed sensing and sparse signal processing can deal with all these scenarios in the same manner [10].

This work is supported by the Montenegrin Ministry of Science, project grant: “New ICT Compressive sensing based trends applied to multimedia, biomedicine and communications (CS-ICT)” (Montenegro Ministry of Science, Grant No. 01-1002).

Authors are with the Faculty of Electrical Engineering, University of Montenegro, Cetinjski put bb, 81000 Podgorica, Montenegro (phone: +382-69-486-639; e-mail: milosb@ac.me, budo@ac.me, irenao@ac.me, milos@ac.me, srdjan@ac.me)

During the last decade of extensive research in the areas of compressive sensing and sparse signal processing, many different reconstruction approaches have been developed. The basic principle, however remains the same: solve the undetermined system of linear equations describing the available signal measurements, by adding the constraint that the solution of this system has to be the sparsest among the all possible solutions. Hence, the reconstruction of missing samples can be observed as the problem of the sparsity minimization of transformation coefficients, under the constraints dictated by the system of linear equations representing the available data [1]-[5], [10], [17].

As it is equal to the number of non-zero transform coefficients, the ℓ_0 -norm is a natural choice for measuring the sparsity [1]-[9]. However, a direct minimization of this measure is not computationally feasible. Hence, the reconstruction constraint is relaxed by engaging the ℓ_1 -norm as a measure of signal sparsity, as the developed theory proves the equality of the obtained minimization results under certain conditions [1]-[8], [16], [17]. As the engaged ℓ_1 -norm is convex, different reconstruction approaches have been exploited and new developed, including linear programming approaches, primal-dual interior point methods, CoSaMP, Orthogonal Matching Pursuit (OMP), Gradient Pursuit, Gradient-based reconstruction algorithm etc. [1]-[3], [5], [9], [13], [14]. It is important to emphasize that in the gradient-based reconstruction algorithm originally proposed in [13], ℓ_1 -norm is minimized by engaging the values of missing samples as minimization variables. Recently, a proof of the convexity of the ℓ_1 -norm with respect to the missing samples as variables was presented [15].

Genetic algorithms (GA) are widely used stochastic optimization methods [19]-[24]. However, there are just few interesting applications in signal processing. GAs have been applied in signal parameter estimation [22]. In the processing of radar signals, after the removal of the micro-Doppler, GA was applied in the rigid body reconstruction [24]. Starting from an initial population consisted of a certain number of possible problem solutions represented by chromosomes, the algorithm emulates the natural process of the evolution, providing the chance of surviving and recombination to the best adopted individuals, having the highest fitness. Through the generations the average fitness of the new populations increases, as they are consisted of the best individuals

from previous generations, and their recombinations [24].

The paper is organized as follows. In Section II the basic theory dealing with the considered CS problem is provided. Section III presents the GA setup for the reconstruction of missing samples. Numerical results are presented in Section IV, while the concluding remarks are given in Section V.

II. BASIC THEORY

The sparse signal reconstruction can be observed as a sparsity measure minimization under the constraints defined by the available samples. Let us observe the signal $x(n)$ of length M , with M_A available samples, that is, $M - M_A$ missing samples at random positions. Let us denote with \mathbf{x} the vector consisted of all signal samples. It is assumed that the signal is sparse in the transformation domain, whereas the signal transform is defined by the transform matrix Ψ , and $\mathbf{X} = [X(0), X(1), \dots, X(M-1)]^T$ is consisted of transform coefficients calculated by $\mathbf{X} = \Psi\mathbf{x}$. Let us denote with \mathbf{x}_{cs} the vector consisted of available samples at random positions

$$n \in \mathbf{M}_A = \{n_1, n_2, \dots, n_{M_A}\} \subset \mathbf{M} = \{0, 1, \dots, M-1\}.$$

The compressive sensing based random selection of signal values can be modeled as:

$$\mathbf{y}_{cs} = \Phi\mathbf{x} = \Phi\Psi^{-1}\mathbf{X} = \mathbf{A}_{cs}\mathbf{X}, \quad (1)$$

where Φ represents a random measurement matrix and \mathbf{A}_{cs} is formed from the inverse transform matrix Ψ^{-1} by omitting rows corresponding to missing samples positions. The CS-based reconstruction algorithms try to find the set of values of missing samples that minimize the sparsity of the corresponding transformation coefficients, and this can be formulated as the following optimization problem:

$$\min \sum_{k=0}^{M-1} |X(k)| \quad \text{subject to } \mathbf{y}_{cs} = \mathbf{A}_{cs}\mathbf{X}. \quad (2)$$

The solution of (2) can be found by varying the possible values of missing samples, starting from some initial values, e.g. zeros, while preserving the values of available samples. For a large number of missing samples this procedure, however, is not computationally feasible, and consequently, solvers based on convex optimization are utilized. Another possible reconstruction approach is based on the sparsity measure minimization by varying the missing samples values in a steepest descent manner [13], [14].

III. GENETIC ALGORITHM-BASED RECONSTRUCTION OF MISSING SAMPLES

The optimization problem (2) can be solved by applying the genetic algorithm. A detailed GA overview as well as realization details are given in [19] and [20]. The algorithm simulates the natural process of evolution, by selecting the individuals i.e. the possible problem solutions represented by chromosomes that have the highest match to the selected fitness function. In the considered problem framework, the population is consisted of individuals which represent the possible values of missing samples.

The idea is to vary these values, in order to minimize the sparsity measure, i.e. the fitness function (2), while the values of available samples remain unchanged [13]. The algorithm can be briefly described as follows.

TABLE I: PROPOSED GENETIC ALGORITHM SETUP

Max. No. of generations	500
Fitness function	ℓ_1 -norm of DFT coefficients calculated for the signal whose missing samples are equal to the values contained within the individual's chromosomes.
Population size	3×30
Population type	double vector
Initial population distribution	uniform
Initial population range	$[\min(\mathbf{y}_{cs}), \max(\mathbf{y}_{cs})]$
Fitness scaling type	shift linear
Fitness scaling constant	2
Selection function	remainder selection
Elite count	1
Crossover type	scattered
Mutation type	Gaussian
Mutation initial standard dev.	0.3
Migration direction	forward
Hybrid function	fminunc

First, the initial population of 90 individuals is formed. The problem is represented by a multi-population, with 3 sub-groups of 30 individuals. Our numerical analysis has shown some results improvement when this concept is used. Each individual represents the set of possible values at the positions of missing samples $n \in \mathbf{M} \setminus \mathbf{M}_A$, coded by a vector of real numbers with double precision, $h(n)$. Initially, each value is chosen randomly with uniform distribution within the bounds defined by minimal and maximal values in the set of available samples. Then, for 500 generations, the following steps are taken. The fitness of each individual is calculated as the ℓ_1 -norm of the transform vector of the analyzed signal, whose missing values are equal to the individual's chromosomes. The available samples remain unchanged. In other words, for each individual in the observed generation, the transform coefficients are calculated for signals of the form:

$$y(n) = \begin{cases} y_{cs}(n), & n \in \mathbf{M}_A \\ h(n), & n \in \mathbf{M} \setminus \mathbf{M}_A \end{cases} \quad (3)$$

in order to obtain

$$\|\mathbf{Y}\|_1 = \sum_{k=0}^{M-1} |Y(k)|, \quad (4)$$

where \mathbf{Y} is the vector of transform coefficients of signal

(3) with $h(n)$ being chromosome values, that is, the possible values of the missing samples, representing individuals of the GA population. The raw fitness scores are then scaled by applying the scaling function known as ‘shift linear’, which enables the fitness scaling in the way that the expected result for the fittest individual is equal to twice of the average score [21]. After the fitness is scaled, the selection function is applied in order to choose the parents for the next generation. Our numerical simulations have shown that the best results are obtained by using the so-called ‘remainder selection function’. It takes the integer part of each individual (which is previously scaled by the scaling function) to determine the number of its appearance in the list of possible parents. Next, the roulette selection is used on the remaining fractional part [21]. The ‘scattered’ crossover of the winners is performed to form the children, where a random binary vector is formed. At the positions of ‘1’ genes from the first parent are used to form the new individual, and at the positions of ‘0’ genes of the second parent are used. The mutation is Gaussian, and it adds a random number with zero mean value to each entry of the parent vector [21], with initial standard deviation 0.3 which linearly decreases over the generations, until it reaches the zero value in the last generation. The migration is done in the forward direction, towards the last subpopulation.

The presented GA setup is based on a large number of numerical experiments, and it is summarized in the Table I. Each parameter is chosen by varying its value over the possible range, and selecting the one that minimizes the reconstruction error.

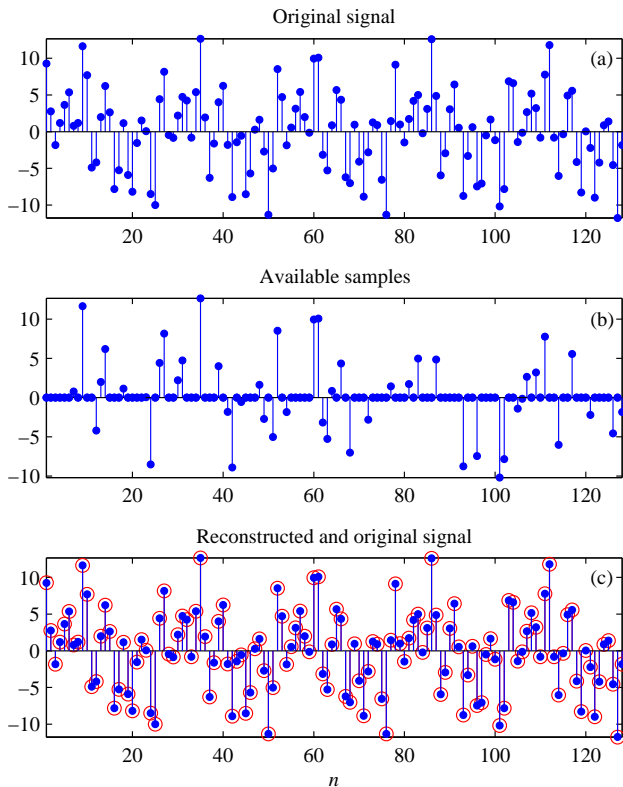


Fig. 1. The reconstruction of signal with missing samples: (a) the original signal with all samples available, (b) the signal with zeros at missing samples positions, (c) the reconstructed signal (dots) and the original signal (circles).

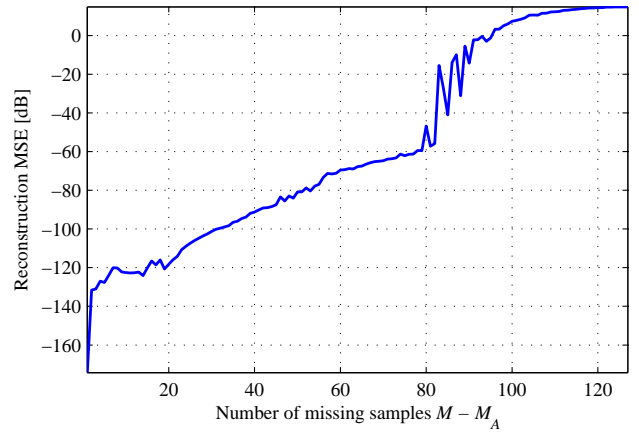


Fig. 2. MSE in the reconstruction of the four-component signal with missing samples from Example 1. The MSE is shown versus the number of missing samples $M - M_A$. The results are shown in logarithmic scale.

IV. NUMERICAL RESULTS

In order to illustrate and verify the presented theory, as well as to confirm that the GA-based reconstruction results are in accordance with the CS theory, the following numerical experiments were taken.

Example 1: To illustrate the reconstruction capability of the proposed approach, we observe a four-component signal sparse in DFT domain, defined as:

$$x(n) = 4 \sin\left(\frac{10\pi n}{M}\right) + 3 \sin\left(\frac{40\pi n}{M} + \frac{\pi}{5}\right) + 5.7 \cos\left(\frac{60\pi n}{M}\right) + 2 \cos\left(\frac{90\pi n}{M} + \frac{\pi}{7}\right), \quad (5)$$

and with length $M = 128$. The $M - M_A = 80$ samples are missing at random positions, i.e. 37.5% signal samples are available. The signal sparsity is $K = 8$, as each sinusoid has two non-zero coefficients in the DFT domain (it is assumed that the components are placed on the frequency grid).

The GA with the proposed parameter setup is applied in the reconstruction of missing samples. The reconstruction results are shown in Fig. 1, where the reconstruction MSE is -60.22 dB.

Example 2: The reconstruction results are highly dependent on the number of missing samples. Hence, in order to additionally verify the results presented in the Example 1, the following experiment was carried out. The signal (5) was considered again. The number of missing samples $M - M_A$ was varied from 1 to $M - 1$, and MSE calculated for each $M - M_A$, based on 40 independent realizations of the signal with random missing samples positions. The reconstruction MSE calculated as

$$MSE = \frac{1}{M} \sum_n |x(n) - x_{rec}(n)|^2 \quad (6)$$

is shown in Fig. 2, confirming that with the increase of the number of missing samples, the reconstruction accuracy decreases, as expected by the CS theory, and that the GA-based reconstruction results match the theoretical expectations.

Example 3: Consider the signal of the form:

$$x(n) = \sum_{i=1}^s A_i \sin\left(2\pi k_i \frac{n}{M} + \theta_i\right) \quad (7)$$

of length $M = 64$, with sparsity $K = 2s$, where s is the number of real sinusoids. The number of real sinusoids is varied from 1 to $M / 2$, with step 2. Amplitudes, frequencies and phases are chosen randomly, with uniform distribution, within the intervals defined by $0 \leq A_i \leq 4$, $0 \leq k_i \leq M / 2 - 1$ and $0 \leq \theta_i \leq 2\pi$ respectively.

The number of missing samples is also varied from 0 to $M - 1$ with step 2. For each possible value of s and for each observed number of missing samples the reconstruction MSE was calculated based on 40 independent realizations of (7) with given number of randomly positioned missing samples. The results are shown in Fig. 3. Blue values correspond to a small reconstruction MSE. The reconstruction is successful when $M_A \geq 2K$ is satisfied, as it is theoretically expected [2].

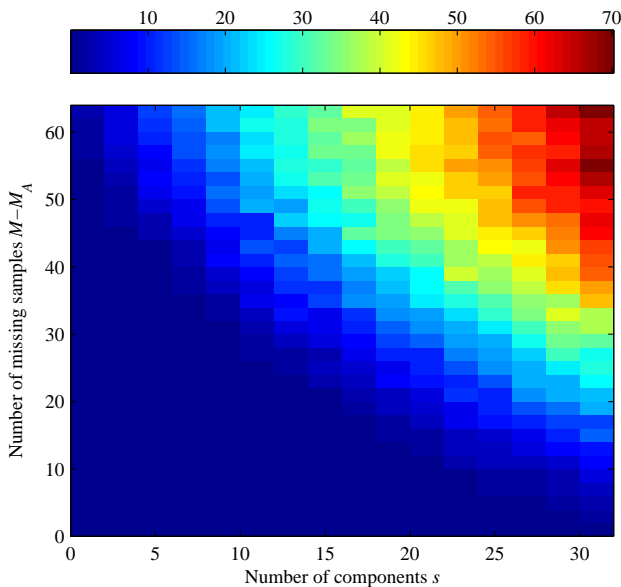


Fig. 3. MSE in the reconstruction of the s -component signal with $M - M_A$ randomly positioned missing samples, shown versus s and $M - M_A$. The MSE takes values defined by the color bar presented above.

V. CONCLUSION

The reconstruction of signals with missing samples is considered. The concentration i.e. the sparsity measure is minimized by varying the missing samples values. The genetic algorithm, well known for its efficient searching capabilities of large spaces of possible solutions is applied in the considered problem framework. A suitable GA parameter setup is proposed for solving the particular minimization problem arising in the CS and sparse signal processing. The presented theory is numerically verified on signals sparse in DFT domain. The application of the proposed approach on real-life signals sparse in a transform domain is straight-forward. Our further research is oriented towards the comparison of the numerical complexity between the presented GA-based approach and existing CS reconstruction algorithms, applied in calculation demanding CS problems, such as in the reconstruction of large digital images with missing pixels.

REFERENCES

- [1] E. J. Candès, M. B. Wakin, "An Introduction To Compressive Sampling", *IEEE Signal Processing Magazine*, 2008.
- [2] D. Donoho, "Compressed sensing," *IEEE Trans. on Information Theory*, 52, (4), pp. 1289 – 1306, 2006.
- [3] R. Baraniuk, "Compressive sensing," *IEEE Signal Processing Magazine*, 24, (4), pp. 118-121, 2007.
- [4] E. Candès, J. Romberg, and T.Tao, "Robust Uncertainty Principles: Exact Signal Reconstruction From Highly Incomplete Frequency Information," *IEEE Transactions on Information Theory*, 52(2), pp. 489-509, 2006.
- [5] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 20, no. 3, pp. 301-321, 2009
- [6] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*, Springer, 2010
- [7] H. Rauhut, "Stability Results for Random Sampling of Sparse Trigonometric Polynomials," *IEEE Tran. on Information theory*, 54(12), 2008, pp. 5661 – 5670
- [8] C. Studer, P. Kuppinger, G. Pope, H. Bolcskei, "Recovery of sparsely corrupted signals," *IEEE Transactions on Information Theory* vol.58, no.5, 2012, pp.3115–3130.
- [9] Davenport, M., Duarte, M., Eldar, Y., Kutyniok, G., 'Introduction to compressed sensing', Chapter in *Compressed Sensing: Theory and Applications*, Cambridge University Press, 2012
- [10] S. Stankovic, L.J. Stankovic, I. Orovic, "A Relationship between the Robust Statistics Theory and Sparse Compressive Sensed Signals Reconstruction," *IET Signal Processing, Special issue on Compressive Sensing and Robust Transforms*, May 2014.
- [11] E. Sejdić, M. A. Rothfuss, M. L. Gimbel, M. H. Mickle, "Comparative Analysis of Compressive Sensing Approaches for Recovery of Missing Samples in an Implantable Wireless Doppler Device," *IET Signal Processing*, vol. 8, no. 3, pp. 230-238, May 2014.
- [12] X. Li, G. Bi, "Image reconstruction based on the improved compressive sensing algorithm," *IEEE Int. Conference on Digital Signal Processing (DSP)*, pp. 357 – 360, Singapore 2015
- [13] L.J. Stanković, M. Daković, and S. Vujović, "Adaptive Variable Step Algorithm for Missing Samples Recovery in Sparse Signals," *IET Signal Processing*, vol. 8, no. 3, 2014.
- [14] M. Brajović, I. Orović, M. Daković, and S. Stanković, "Gradient-based signal reconstruction algorithm in the Hermite transform domain," *Electronics Letters, in print*, 2015.
- [15] M. Brajović, M. Daković, and L.J. Stanković, "Convexity of the l_1 -norm based Sparsity Measure with Respect to the Missing Samples as Variables," *5th Mediterranean Conference on Embedded Computing, MECO 2016*, Bar, 2016
- [16] S. Stanković, I. Orović, and E. Sejdić, *Multimedia signals and systems*. Springer - Verlag, 2012.
- [17] L.J. Stanković, *Digital Signal Processing with Selected Topics*. CreateSpace Independent Publishing Platform, An Amazon.com Company, November 4, 2015
- [18] L.J. Stanković, "A measure of some time-frequency distributions concentration," *Signal Processing*, Vol.81, No.3, Mar.2001, pp.621-631.
- [19] K. Tang, K. Man, S. Kwong, Q. He, "Genetic algorithms and their applications," *IEEE Signal Process. Mag.* 13 (6) (1996) 22–37.
- [20] J.P. Rennard, Introduction to genetic algorithms, September 2016; <http://www.rennard.org/alife/english/gavintrgb.html>.
- [21] Global optimization toolbox, September 2016; <http://www.mathworks.com/products/global-optimization/index.html>
- [22] I. Djurović, M. Simeunović, and B. Lutovac, "Are genetic algorithms useful for the parameter estimation of FM signals?," *Digital Signal Processing*, vol. 22, no. 6, pp. 1137-1144, 2012
- [23] M. Mitchell, *An introduction to genetic algorithms*, MIT press, 1998.
- [24] L.J. Stanković, V. Popović-Bugarin, and F. Radenović, "Genetic algorithm for rigid body reconstruction after micro-doppler removal in the radar imaging analysis," *Signal Processing*, Volume 93, Issue 2013, Jan 2013.