RECONSTRUCTION OF ECG SIGNALS FOR COMPRESSIVE SENSING BY PROMOTING SPARSITY ON THE GRADIENT

Jeevan K. Pant and Sridhar Krishnan

Department of Electrical and Computer Engineering, Ryerson University
350 Victoria Street, Toronto, ON M5B 2K3, CANADA
Email:jpant@rnet.ryerson.ca, krishnan@ee.ryerson.ca

ABSTRACT

A new algorithm for the reconstruction of signals in compressive sensing framework is proposed. The algorithm is based on a least-squares method which incorporates a regularization to promote sparsity on the gradient of the signal. It uses a sequential basic conjugate-gradient method, and it is especially suited for the reconstruction of signals which exhibit temporal correlation, e.g., electrocardiogram (ECG) signals. Simulation results are presented which demonstrate that the proposed algorithm yields upto 80.28% reduction in mean square error and from 49.95% to 65.64% reduction in the required amount of computation, relative to the state-of-the-art block sparse Bayesian learning bound-optimization algorithm.

Index Terms— Compressive sensing, electrocardiogram, conjugate gradient, sparse gradient

1. INTRODUCTION

Compressive sensing (CS) is a novel technique for the acquisition of signals in terms of a small number of measurements [1] [2] [3]. Sparse signal recovery is a problem associated with CS, and the state-of-the-art signal recovery algorithms are based on ℓ_1 and ℓ_p minimization [4] [5] [6], ℓ_0 minimization [7] [8], greedy approximation [9] [10], iterative shrinkage [11] [12], and Bayesian learning [13]. These algorithms are not very effective for the reconstruction of signals (i) which exhibit temporal correlation and (ii) which are not sparse in either time domain, transform domain, or with respect to a dictionary. Recently, so called block sparse Bayesian learning bound-optimization (BSBL-BO) algorithm [14] has been effectively applied for the reconstruction of temporally correlated ECG signals [15].

In this paper, a new algorithm, namely, ℓ_p^d-regularized least-squares (ℓ_p^d-RLS) algorithm, is proposed. The algorithm is based on the minimization of an ℓ_p^d-pseudonorm regularized squared error. The ℓ_p^d pseudonorm is used to promote sparsity on the gradient of the signal and a sequential basic conjugate-gradient method is applied for the optimization. The ℓ_p^d-RLS algorithm yields improved reconstruction performance for temporally correlated ECG signals relative to the BSBL-BO algorithm.

The authors would like to thank the Canada Research Chairs program for supporting this research.

2. BACKGROUND AND PREVIOUS WORK

A discrete-time signal $\boldsymbol{x}$ of length N is said to be K *sparse* with respect to an orthonormal basis $\mathbf{D}$, if it can be expressed as a linear combination of total K columns of $\mathbf{D}$, i.e., $\boldsymbol{x} = \sum_{i=1}^{K} a_i \boldsymbol{d}_i$ typically with $K \ll N$ where $\boldsymbol{d}_i$ is a column of $\mathbf{D}$.

The signal acquisition procedure in CS can be characterized by

$$\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x} = \boldsymbol{\Phi}\mathbf{D}\boldsymbol{a},$$

where $\boldsymbol{y}$ is a measurement vector of size M, $\boldsymbol{\Phi}$ is a measurement matrix of size $M \times N$, typically with $M \ll N$, and $\boldsymbol{a}$ is the coefficient vector of length L.

In principle, signal $\boldsymbol{x}$ can be recovered from the measurement $\boldsymbol{y}$ by solving the optimization problem

$$\begin{array}{ll} \underset{\boldsymbol{a}}{\text{minimize}} & ||\boldsymbol{a}||_0 \\ \text{subject to:} & \boldsymbol{y} = \boldsymbol{\Phi}\mathbf{D}\boldsymbol{a} \end{array} \tag{1}$$

and using the resulting solution in $\boldsymbol{x} = \mathbf{D}\boldsymbol{a}$. Unfortunately, the computational complexity involved in solving the above problem grows exponentially with the signal length N.

The ℓ_1-minimization based *basis pursuit* algorithm is a computationally tractable algorithm which solves the optimization problem

$$\begin{array}{ll} \underset{\boldsymbol{a}}{\text{minimize}} & ||\boldsymbol{a}||_1 \\ \text{subject to:} & \boldsymbol{y} = \boldsymbol{\Phi}\mathbf{D}\boldsymbol{a} \end{array}, \tag{2}$$

where $||\boldsymbol{a}||_1 = \sum_{i=1}^{N} |a_i|$ is the ℓ_1 norm of $\boldsymbol{a}$. The solution of the problem in (1) can be found by solving the problem in (2) provided that a condition which requires the sparsity value of $\boldsymbol{a}$ to be sufficiently small is satisfied [1] [2] [3].

The ℓ_p-pseudonorm minimization based algorithms, which solve the optimization problem

$$\begin{array}{ll} \underset{\boldsymbol{a}}{\text{minimize}} & ||\boldsymbol{a}||_p^p \\ \text{subject to:} & \boldsymbol{y} = \boldsymbol{\Phi}\mathbf{D}\boldsymbol{a} \end{array}, \tag{3}$$

where $||\boldsymbol{a}||_p^p = \sum_{i=1}^{N} |a_i|^p$ with $p < 1$, have been shown to offer improved signal reconstruction performance relative to the techniques based on solving the problem in (2) [5] [6].

3. ℓ_p^d-REGULARIZED LEAST-SQUARES ALGORITHM

3.1. ℓ_p^d pseudonorm and problem formulation

We define the ℓ_p^d pseudonorm of signal $\boldsymbol{x}$ as

$$||d\boldsymbol{x}||_p = \left[\sum_{i=1}^{N-1} |x_i - x_{i+1}|^p\right]^{1/p} \tag{4}$$

which is essentially the ℓ_p pseudonorm of the first-order difference $d\boldsymbol{x}$ whose ith component is given by $x_i - x_{i+1}$. With $p \leq 1$, function $||d\boldsymbol{x}||_p$ gives a measure of the sparsity in $d\boldsymbol{x}$. In other words, it gives an approximate measure of the number of elements of the set $\{i : 1 \leq i < N\}$ for which $x_i \not\approx x_{i+1}$. Therefore, a signal $\boldsymbol{x}$ obtained by minimizing the function $||d\boldsymbol{x}||_p$ tends to exhibit increased correlation between its succeeding components. Consequently, the algorithm to be presented, which minimizes the function $||d\boldsymbol{x}||_p$, is expected to be more effective for the reconstruction of temporally correlated signals.

It can be shown that for $p \leq 1$, function $||d\boldsymbol{x}||_p$ is not differentiable. To make it differentiable, and thereby, to facilitate the optimization to be presented below, we consider the approximate ℓ_p^d pseudonorm given by

$$||d\boldsymbol{x}||_{p,\epsilon}^p = \sum_{i=1}^{N-1} \left[(x_i - x_{i+1})^2 + \epsilon^2\right]^{p/2} \tag{5}$$

where $\epsilon > 0$ is the approximation parameter. We should mention that an approximation similar to this has recently been used to approximate the ℓ_p pseudonorm [6] [16] [17]. To the best of our knowledge, the pseudonorm given in (5) has not been used in the context of signal reconstruction in CS.

We propose to recover signal $\boldsymbol{x}$ from measurements $\boldsymbol{y}$ by solving the ℓ_p^d-regularized least-squares problem

$$\underset{\boldsymbol{x}}{\text{minimize}} \quad f(\boldsymbol{x}) = \tfrac{1}{2}\,||\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{y}||_2^2 + \lambda\,||d\boldsymbol{x}||_{p,\epsilon}^p \tag{6}$$

with $p \leq 1$ and small ϵ, where λ is the regularization parameter.

We end this subsection with the remark that the ℓ_p^d pseudonorm is related to the widely used total-variation (TV) norm [18] [19] [20] in that the function $||d\boldsymbol{x}||_{p,\epsilon}^p$ reduces to the one-dimensional version of the TV norm when $p = 1$ and $\epsilon = 0$.

3.2. Optimization

As long as $\epsilon > 0$, the function $f(\boldsymbol{x})$ in (6) remains differentiable whose gradient $\boldsymbol{g}$ and Hessian $\mathbf{H}$ are given by

$$\boldsymbol{g} = \boldsymbol{\Phi}^T(\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{y}) + \lambda \boldsymbol{g}_p^d \tag{7}$$

and

$$\mathbf{H} = \boldsymbol{\Phi}^T\boldsymbol{\Phi} + \lambda\mathbf{U}. \tag{8}$$

In (7), $\boldsymbol{g}_p^d$ is the gradient of function $||d\boldsymbol{x}||_{p,\epsilon}^p$ given by $\boldsymbol{g}_p^d = \left[g_{p1}^d \; g_{p2}^d \; \cdots \; g_{pN}^d\right]^T$ where

$$g_{pi}^d = \begin{cases} p \cdot c_i & \text{for} \quad i = 1 \\ p \cdot (-c_{i-1} + c_i) & \text{for} \quad 1 < i < N \\ p \cdot (-c_{i-1}) & \text{for} \quad i = N \end{cases} \tag{9}$$

and

$$c_i = \left[(x_i - x_{i+1})^2 + \epsilon^2\right]^{p/2-1}(x_i - x_{i+1}) \tag{10}$$

for $i = 1, 2, \ldots, N-1$.

In (8), $\mathbf{U}$ is the Hessian of function $||d\boldsymbol{x}||_{p,\epsilon}^p$ whose $\{i,j\}$th component is given by

$$u_{i,j} = \begin{cases} h_i & \text{for } i = j = 1 \\ h_{i-1} + h_i & \text{for } i = j, 1 < i < N \\ h_{i-1} & \text{for } i = j = N \\ -h_i & \text{for } i = j-1, 1 < j \leq N \\ -h_j & \text{for } j = i-1, 1 < i \leq N \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

where

$$h_i = p\left[(x_i - x_{i+1})^2 + \epsilon^2\right]^{p/2-2}\left[(p-1)(x_i - x_{i+1})^2 + \epsilon^2\right] \tag{12}$$

for $i = 1, 2, \ldots, N-1$.

It can be shown that the Hessian $\mathbf{H}$ in (8) is positive definite as long as $h_i > 0, \forall i$. Consequently, the objective function $f(\boldsymbol{x})$ in (6) is convex over a region characterized by $\{\boldsymbol{x} : |x_i - x_{i+1}| < \frac{\epsilon}{\sqrt{1-p}} \;\text{ for }\; 1 \leq i < N\}$, where x_i is the ith component of $\boldsymbol{x}$. This implies that, for a sufficiently large value of ϵ the objective function $f(\boldsymbol{x})$ has a large convex region, and it is easy to locate the minimizer of such an objective function. However, the global minimizer of $f(\boldsymbol{x})$ is the desired solution only when ϵ is sufficiently small. Therefore, we apply a sequential optimization strategy whereby the problem in (6) is solved sequentially for a set of decreasing values of ϵ. Such an optimization procedure can be described as follows:

- Select two sufficiently large values, one of ϵ and the other of λ. Solve the problem in (6) using the zero vector as an initializer.

- Reduce the values of ϵ and λ and solve the problem in (6) again using the solution obtained from the previous optimization as an initializer.
- Repeat this procedure until the problem in (6) is solved for a sufficiently small values of ϵ and λ.
- Output the final solution and stop.

Although the optimization problem of the form (6) is usually solved with a fixed value of λ, we have proposed to solve the problem in (6) for a set of decreasing values of λ. This helps to improve the convergence of the algorithm.

To solve the problem in (6) for fixed values of ϵ and λ, we use the basic conjugate-gradient (BCG) algorithm described in [21], which is based on gradient $\boldsymbol{g}$ and Hessian $\mathbf{H}$ given in (7) and (8), respectively. The BCG algorithm is a simple, efficient, and well tested algorithm which has been found to be effective to solve the problem in (6). In the kth iteration of the BCG algorithm, the iterate $\boldsymbol{x}_k$ is updated as [21]

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k, \tag{13}$$

where

$$\boldsymbol{d}_k = -\boldsymbol{g}_k + \beta_{k-1}\boldsymbol{d}_{k-1}, \tag{14}$$

$$\beta_{k-1} = \frac{||\boldsymbol{g}_k||_2^2}{||\boldsymbol{g}_{k-1}||_2^2}, \tag{15}$$

$$\alpha_k = \frac{||\boldsymbol{g}_k||_2^2}{\boldsymbol{d}_k^T \mathbf{H}_k \boldsymbol{d}_k}. \tag{16}$$

The Hessian $\boldsymbol{H}_k$ in (16) need not be evaluated explicitly. Instead, its denominator term can be computed efficiently as

$$\boldsymbol{d}_k^T \mathbf{H}_k \boldsymbol{d}_k = ||\boldsymbol{\Phi}\boldsymbol{d}_k||_2^2 + \lambda\, ||\boldsymbol{p}_k||_2^2\,, \tag{17}$$

where $\boldsymbol{p}_k = [p_1\ p_2\ \cdots\ p_{N-1}]^T$ and

$$p_i = \sqrt{h_i}\, d_i^d \quad \text{for} \quad i = 1, 2, \ldots, N-1. \tag{18}$$

In (18), h_i can be computed using (12) and $d_i^d = d_{ki} - d_{k(i+1)}$ where d_{ki} is the ith component of $\boldsymbol{d}_k$.

Note that, in order for the Hessian matrix $\mathbf{H}_k$ to be positive definite, we must have $h_i > 0, \forall i$. Therefore, before using in (18), h_i's are thresholded as

$$h_i = \max\{h_i, \delta\}, \quad \text{for} \quad i = 1, 2, \ldots, N-1,$$

where δ is a small positive scalar.

3.3. Algorithm

The proposed ℓ_p^d-regularized least-squares (ℓ_p^d-RLS) algorithm for the reconstruction of temporally correlated sparse signals is summarized in Table 1. The algorithm takes parameters T, p, ϵ_1, ϵ_T, λ_1, λ_T, E_t, L_b, r, and δ as inputs. Total $T-2$ values of ϵ lying in between ϵ_1 and ϵ_T are computed as

$$\epsilon_t = \epsilon_1 \exp(-\alpha(t-1)) \quad \text{for} \quad t = 2, 3, \ldots, T-1 \tag{19}$$

where $\alpha = \log(\epsilon_1/\epsilon_T)/(T-1)$. Similarly, total $T-2$ values of λ lying between λ_1 and λ_T are computed as

$$\lambda_t = \lambda_1 \exp(-\gamma(t-1)) \quad \text{for} \quad t = 2, 3, \ldots, T-1 \tag{20}$$

where $\gamma = \log(\lambda_1/\lambda_T)/(T-1)$.

Table 1. ℓ_p^d-RLS Algorithm

Step 1
Input: T, p, ϵ_1, ϵ_T, λ_1, λ_T, $\boldsymbol{\Phi}$, $\boldsymbol{y}$, E_t, L_b, r, and δ.
Set $\boldsymbol{x}_s = \mathbf{0}$.
Step 2
Compute ϵ_t for $t = 2, 3, \ldots, T-1$ using (19) and
λ_t for $t = 2, 3, \ldots, T-1$ using (20).
Step 3
Repeat the following for $t = 1, \ldots, T$
i) Set $\epsilon = \epsilon_t$, $\lambda = \lambda_t$, $k = 0$, $\boldsymbol{x}_0 = \boldsymbol{x}_s$, $E_r = 10^{10}$.
ii) Repeat the following while $E_r > E_t$,
a) Compute $\boldsymbol{x}_{k+1}$ using (13), (14), (15), (16), (17).
b) Compute $L = L_b + \text{round}(t/r)$.
c) Set $k = k + 1$.
d) Exit loop if $k > L$.
e) Compute $E_r = \|\|\alpha_k \boldsymbol{d}_k\|\|_2$.
iii) Set $\boldsymbol{x}_s = \boldsymbol{x}_k$.
Step 4
Output $\boldsymbol{x}^* = \boldsymbol{x}_s$ and stop.

4. SIMULATION RESULTS

In the first simulation, an ECG signal $\boldsymbol{x}$ of length $N = 250$ was constructed by retaining the first 250 samples of the second channel of the cutaneous potential recordings of a pregnant woman [22] which contains eight channel recordings. The number of measurements was set to $M = 125$. A sparse measurement matrix was constructed as follows: a) a matrix $\boldsymbol{\Phi}$ of size $M \times N$ with all zero elements was constructed and b) randomly chosen 15 components of each column of $\boldsymbol{\Phi}$ were set to unity. As suggested in [15] and [23], a measurement matrix constructed using this approach can have an energy efficient implementation for CS. Measurement was taken as $\boldsymbol{y} = \boldsymbol{\Phi x}$. The ℓ_p^d-RLS algorithm was run with $p = 1$, $\epsilon_1 = 100$, $\epsilon_T = 1e-2$, $\lambda_1 = 100$, $\lambda_T = 1e-2$, $T = 20$, $E_t = 1e{-}25$, $L_b = 15$, $r = 4$, and $\delta = 1e{-}5$. Matlab implementation of the BSBL-BO algorithm was downloaded from [24]. In [15], the BSBL-BO algorithm is shown to offer superior reconstruction performance for the ECG signals relative to several state-of-the-art signal reconstruction algorithms including compressive sampling matching pursuit (CoSaMP) [25], basis pursuit (BP) [4], smoothed ℓ_0 (SL0) [7], and block orthogonal matching pursuit (BOMP) [26] algorithms. Therefore, we compare the performance of the proposed ℓ_p^d-RLS algorithm with that of the BSBL-BO algorithm. The original signal and the signals reconstructed using the two algorithms are shown in Fig. 1. As can be seen, the signal reconstructed

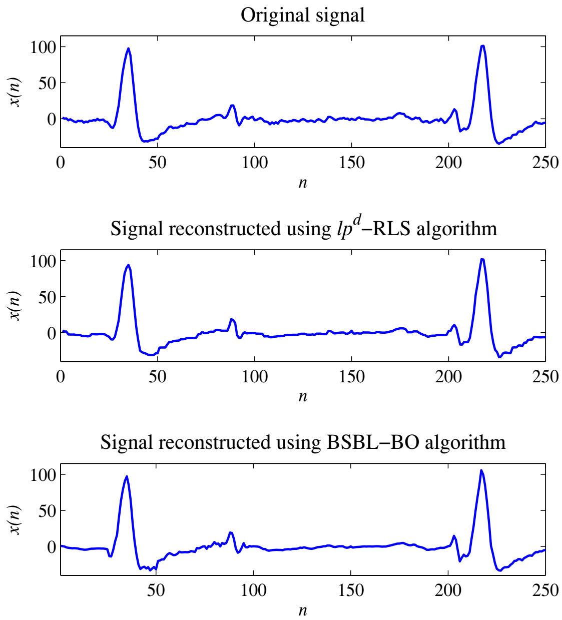

Fig. 1. Original signal and signals reconstructed using the ℓ_p^d-RLS and BSBL-BO algorithms.

using ℓ_p^d-RLS algorithm appears better than that reconstructed using BSBL-BO algorithm. Mean square error (MSE) was measured as $(1/N)\sum_{i=1}^{N}(x_i - \hat{x}_i)^2$, where x_i and $\hat{x}_i$ are the ith components of $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$, respectively, and $\hat{\boldsymbol{x}}$ is the reconstructed signal. CPU time required by the both algorithms were measured by using MATLAB commands *tic* and *toc*. The CPU time and MSE for both the algorithms are shown in Table 2. One can notice that the ℓ_p^d-RLS algorithm requires

Table 2. Comparison of the ℓ_p^d-RLS and BSBL-BO algorithms

Algorithm	CPU time, seconds	MSE
ℓ_p^d-RLS	0.17377	2.2523
BSBL-BO	0.85073	3.0863

much less CPU time and yields less MSE compared to the BSBL-BO algorithm.

In the second simulation, the length of a signal was set to $N = 256$ and total eight values of the number of measurements were chosen as $M = round(t \times N)$ with $t = 0.1, 0.2, \ldots, 0.8$. Test signals were obtained from the *normal sinus rhythm* database record number 16265 from Physionet [27]. This record contains a two channel ECG signal of 24 hours duration sampled at the rate of 128 samples per second. A signal $\boldsymbol{x}$ of length N was constructed by selecting N consecutive samples of the first channel ECG signal from a random location. A measurement matrix $\boldsymbol{\Phi}$ was constructed and measurement was taken using the procedure used in the first simulation except that each column of $\boldsymbol{\Phi}$ had $round(0.06 \times N)$ ones and $M - round(0.06 \times N)$ zeros. The ℓ_p^d-RLS algorithm was run with the same parameters as in the first simulation except $\sigma_1 = 600$, $\lambda_1 = 600$, and $T = 30$. Both the algorithms were applied 1000 times with a different $\boldsymbol{x}$ and a different $\boldsymbol{\Phi}$ each time. Average MSE and average CPU time are plotted in Fig. 2; the left panel shows average MSE and the right panel shows average CPU time over 1000 runs. As can be seen, both the MSE and CPU time for the ℓ_p^d-RLS algorithm is less than that for the BSBL-BO algorithm. The MSE for the ℓ_p^d-RLS algorithm was less than that for the BSBL-BO algorithm by a maximum of 80.28% for $M = 128$. The CPU time for the ℓ_p^d-RLS algorithm was less than that for the BSBL-BO algorithm by a minimum of 49.95% and a maximum of 65.64%.

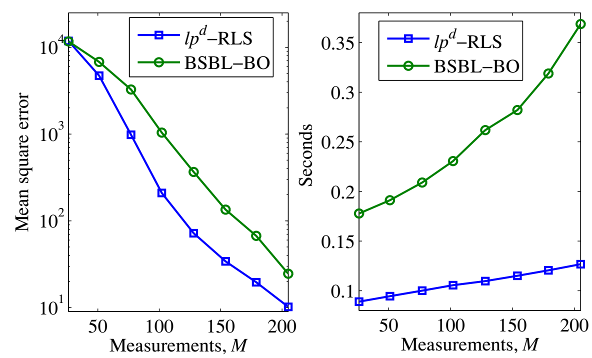

Fig. 2. Average mean square error (left panel) and average CPU time (right panel) for the ℓ_p^d-RLS and BSBL-BO algorithms over 1000 runs.

5. CONCLUSION

A new algorithm for compressive sensing, namely, the ℓ_p^d-RLS algorithm, for the reconstruction of signals has been proposed. The algorithm is based on minimizing an ℓ_p^d-regularized squared-error, and it is especially suited for the recovery of temporally correlated signals. As demonstrated using simulation results, the ℓ_p^d-RLS algorithm yields reduced MSE and reduced CPU time for the reconstruction of ECG signals relative to the state-of-the-art BSBL-BO algorithm.

6. RELATION TO PRIOR WORK

The presented work takes advantage of smoothing of ℓ_p pseudonorm and sequential Fletcher-Reeves' conjugate-gradient method which were recently applied in the algorithms for the reconstruction of sparse signals for compressive sensing [6] [17]. These algorithms promote sparsity on signal coefficients. Thus, they are not very effective for the reconstruction of temporally correlated ECG signals. On the other hand, the proposed algorithm promotes sparsity on the gradient. Consequently, this algorithm encourages temporal correlation on the signal and it is effective for the reconstruction of ECG signals.

7. REFERENCES

[1] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489-509, Feb. 2006.

[2] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, pp. 1289-1306, Apr. 2006.

[3] E. Candès and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406-5425, Dec. 2006.

[4] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Scientif. Comput.*, vol. 20, no. 1, pp. 33-61, 1999.

[5] R. Chartrand and W. Yin, "Iteratively reweighted algorithms for compressive sensing," in *Proc. IEEE Inter. Conf. Acoustics, Speech, Signal Process.*, pp. 3869-3872, 2008.

[6] J. K. Pant, W.-S. Lu, and A. Antoniou, "Unconstrained regularized ℓ_p-norm based algorithm for the reconstruction of sparse signals," in *Proc. IEEE Iter. Symp. Circuits and Syst.*, pp. 1740-1743, May 2011.

[7] H. Mohimani, M. Babie-Zadeh, and C. Jutten, "A fast approach for overcomplete sparse decomposition based on smoothed ℓ_0-norm," *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 289-301, Jan. 2009.

[8] J. K. Pant, W.-S. Lu, and A. Antoniou, "Reconstruction of sparse signals by minimizing a re-weighted approximate ℓ_0-norm in the null space of the measurement matrix," in *Proc. IEEE Inter. Midwest Symp. on Circuits-Syst.*, pp. 430-433, Aug. 2010.

[9] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proc. 27th Annu. Asilomar Conf. Signals Syst. Comput.*, pp. 40-44, vol. 1, Nov. 1993.

[10] D. Needell and J. A. Torpp, "CoSaMP: iterative signal recovery from incomplete and inaccurate samples," in *Commun. ACM*, vol. 53, no. 12, pp. 93-100, Dec. 2010.

[11] S. Sardy, A. Bruce, and P. Tseng, "Block Coordinate Relaxation Methods for Nonparametric Wavelet Denoising," *Jour. Comput. and Graph. Stat.*, vol. 9, no. 2, pp. 361-379, Jun., 2000.

[12] M. Elad, "Why simple shrinkage is still relevant for redundant representations?," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5559-5569, Dec., 2006.

[13] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2346-2356, Jun., 2008.

[14] Z. Zhang and B. D. Rao, "Sparse signal recovery with temporally correlated source vectors using sparse Bayesian learning," *IEEE Jour. Select. Topics Sig. Process.*, vol. 5, no. 5, pp. 912-926, Sept., 2011.

[15] Z. Zhang, T.-P. Jung, S. Makeig, and B. D. Rao, "Compressed sensing for energy-efficient wireless telemonitoring of noninvasive fetal ECG via block sparse Bayesian learning," *IEEE Trans. Biomed. Engg*, vol. 60, no. 2, pp. 300-309, Feb. 2013.

[16] J. K. Pant, W.-S. Lu, and A. Antoniou, "Recovery of sparse signals from noisy measurements using an ℓ_p regularized least-squares algorithm," in *Proc. IEEE Pac. Rim Conf. in Comm., Computers, and Signal Process.*, pp. 48-53, Aug. 2011.

[17] J. K. Pant, W.-S. Lu, and A. Antoniou, "Reconstruction of block-sparse signals by using an $\ell_{2/p}$ regularized least squares algorithm," in *Proc. IEEE Iter. Symp. Circuits and Syst.*, May 2012.

[18] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2419-2434, Nov. 2009.

[19] A. Chambolle, "An algorithm for total variation minimization and applications," *J. Math. Imag. Vis.*, vol. 20, no. 1-2, pp. 89-97, 2004.

[20] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, pp. 259-268, 1992.

[21] A. Antoniou and W.-S. Lu, *Practical Optimization: Algorithms and Engineering Applications,* Springer, 2006.

[22] "DaISy: Database for the identification of systems," `http://homes.esat.kuleuven.be/~smc/daisy/daisydata.html` [Accessed online: 10-July-2012]

[23] H. Mamaghanian, N. Khaled, D. Atienza, and P. Vandergheynst, "Compressed sensing for real-time energy-efficient ECG compression on wireless body sensor nodes," *IEEE Trans. Biomed. Eng.* vol. 58, no. 9, pp. 2456-2466, Sept. 2011.

[24] "Matlab Codes for Compressed Sensing and ICA," `http://dsp.ucsd.edu/~zhilin/Software.html` [Accessed online: 10-July-2012]

[25] 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.

[26] Y. C. Eldar, P. Kuppinger, and H. Bolcskei, "Block-sparse signals: uncertainty relations and efficient recovery," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3042-3054, Jun. 2010.

[27] "Physionet: MIT-BIH normal sinus rhythm database," `http://www.physionet.org` [Accessed online: 10-July-2012]