On ECG reconstruction using weighted-compressive sensing

Dornoosh Zonoobi, Ashraf A. Kassim

Department of Electrical and Computer Engineering, National University of Singapore, Singapore
E-mail: dornoosh@nus.edu.sg

Published in Healthcare Technology Letters; Received on 3rd December 2013; Revised on 6th April 2014; Accepted on 8th April 2014

The potential of the new weighted-compressive sensing approach for efficient reconstruction of electrocardiograph (ECG) signals is investigated. This is motivated by the observation that ECG signals are hugely sparse in the frequency domain and the sparsity changes slowly over time. The underlying idea of this approach is to extract an estimated probability model for the signal of interest, and then use this model to guide the reconstruction process. The authors show that the weighted-compressive sensing approach is able to achieve reconstruction performance comparable with the current state-of-the-art discrete wavelet transform-based method, but with substantially less computational cost to enable it to be considered for use in the next generation of miniaturised wearable ECG monitoring devices.

1. Introduction: Electrocardiograph (ECG) is considered as one of the most important diagnostic tools for assessing the electrical and muscular functions of the heart and thus plays an important role in the battle against cardiovascular diseases which are among the top causes of death in the world [1]. Although resting ECG monitoring is a standard practice in hospitals, major efforts have been spent to realise wireless-enabled low-power ECG monitoring. These ambulatory monitoring devices, however, face many technical challenges including limited wireless connectivity, short battery life, bulkiness and so on [1]. State-of-the-art ECG monitors fall short because either they transmit the uncompressed ECG data over wireless networks, which put much pressure on wireless links; or they compress the data in a compression unit after collecting and storing the full data on the chip, which leads to bulkiness and power inefficiency. In these approaches, the signal of interest is first fully sampled according to the Nyquist rate. It is then compressed, using a discrete wavelet transform (DWT)-based algorithm [2] before being encrypted for transmission.

An alternative approach, however, would be to use the recent paradigm of compressive sensing (CS) [3–5] based on the fact that ECG signals are largely compressible and thus sparse. The CS employs linear sampling operators that map the signal into a relatively small dimensional space which results in a small number of measurements that can be wirelessly transmitted to the remote tele-cardiology centre. The full signal can then be recovered from a much smaller set of measurements than the number of Nyquist-rate samples using complex, yet computationally feasible and numerically stable, nonlinear methods. In CS, not only are the sampling and compression phases unified, but under fairly general conditions, the measurements obtained are considered to be encrypted [6]. Hence, CS essentially moves the burden of computation from the sampling device to the receiver-end where more resources are available.

In recent works [1, 7–10], it has been found that CS-based methods used in ECG sampling devices exhibit the best overall energy efficiency. However, it was reported that they are inferior in terms of compression performance compared with DWT-based compression methods. This is not surprising as these works were carried out using basic ℓ_1 minimisation [3]. It is known that ℓ_1 minimisation is better suited for sampling/reconstruction of static signals and does not exploit the highly structured nature and quasi-periodic property of ECG signals [11].

In this Letter, our focus is on the CS-based methods that are specifically tailored for sampling and reconstruction of time-varying signals such as ECG. We examine the performance of these state-of-the-art CS-based methods when applied to the problem of ECG sampling/reconstruction and compare them with the conventional DWT-based method.

The rest of the Letter is organised as follows. This section ends with a description of the notations used. Section 2 presents the problem of reconstruction of time-varying signals and the current state-of-the-art CS-based approaches that address this problem. Details of our algorithm called ‘weighted-CS’ as applied to ECG reconstruction, together with its efficient implementation, are presented in Section 3. Finally, we present and analyse our experimental results in Section 4 before providing the concluding remarks in Section 6.

Throughout the Letter, vectors are denoted by boldface letters (e.g. $\boldsymbol{f}$, $\boldsymbol{F}$) and f_i is the ith element of the vector $\boldsymbol{f}$. Scalars are shown by small regular letters (e.g. n, k) and matrices are denoted by bold capital Greek letters ($\boldsymbol{\Phi}$, $\boldsymbol{\Psi}$). Superscript (t) added to a vector/matrix refers to that of time t. We use the notation $\boldsymbol{\Psi}|_{\mathcal{S}}$ to denote the sub-matrix containing the columns of $\boldsymbol{\Psi}$ with indices belonging to $\mathcal{S}$. For a vector, the notation $\boldsymbol{y}|_{\mathcal{S}}$ forms a sub-vector that contains elements with indices in $\mathcal{S}$. Other notations are introduced when needed.

2. Reconstruction of time-varying signals: Digitised ECG signals are generally processed in non-overlapping windows of n samples because of the limited on-chip memory and real-time computing constraints. In its noiseless formulation, the problem of reconstruction of such a signal in each window, can be posed as follows: let $\boldsymbol{f}^{(t)} := [f_1^{(t)}, \ldots, f_n^{(t)}]^{\mathrm{T}} \in \mathcal{R}^n$ be the signal of interest at epoch t which is measured using a sampling matrix $\boldsymbol{\Psi} \in \mathcal{R}^{m\times n}$ and $\boldsymbol{y}^{(t)} = \boldsymbol{\Psi}\boldsymbol{f}^{(t)}$ is the observation vector at epoch t. It is assumed that the signal of interest has a sparse approximation ($\boldsymbol{F}^{(t)} = \boldsymbol{\Phi}\boldsymbol{f}^{(t)}$) in some transform domain ($\boldsymbol{\Phi}$), which is incoherent with the sampling domain. Let $\mathcal{S}^{(t)}$ be the support of the vector $\boldsymbol{F}^{(t)}$, that is, $\mathcal{S}^{(t)} := \{k \in \{1, \ldots, n\}: F_k^{(t)} > \varepsilon\}$, it is then assumed that $|\mathcal{S}^{(t)}| \ll n$, where $|\mathcal{S}^{(t)}|$ is the cardinality of $\mathcal{S}^{(t)}$. The problem, at each window t, is then to recover the original signal, $\boldsymbol{f}^{(t)}$, from the corresponding compressive samples ($\boldsymbol{y}^{(t)}$), assuming that $\boldsymbol{F}^{(t)}$ is sparse. A naive approach would be to use the basic CS approach to reconstruct the signal of interest at each time frame separately [1, 12]. However, this approach does not make any use of the fact that the signal of interest in each window is related to the signal of the previous epoch.

Another method to tackle this problem is based on the assumption that the majority of spikes in the transform domain of the signal window at t, would occur at same locations as the spikes

doi: 10.1049/htl.2013.0038

in the transform domain of the previous signal window at $t-1$. Based on this, the modified-CS method [11] uses the support of the previous time instance ($\mathcal{S}^{(t-1)}$) as an estimated support of the signal of interest ($\boldsymbol{F}^{(t)}$) at current time and uses this estimate for reconstruction of $\boldsymbol{f}^{(t)}$, by finding a signal which satisfies the observations and is sparsest outside $\mathcal{S}^{(t-1)}$. It is shown in [11] that under fairly general conditions, the number of samples needed for perfect recovery would be less than the conventional CS.

More recently the authors of [13] proposed a method called regularised modified-CS (RegMod-CS) which is as follows

$$
\begin{aligned}
\boldsymbol{f}^{(t)} = \arg\min \{ & \|(\boldsymbol{\Phi} \boldsymbol{g}^{(t)})|_{\hat{\mathcal{S}}^{(t-1)}}\|_1 \text{ such that} \\
& \boldsymbol{\Psi} \boldsymbol{g}^{(t)} = \boldsymbol{y}^{(t)}, \quad \|(\boldsymbol{\Phi} \boldsymbol{g}^{(t)})|_{\mathcal{S}^{(t-1)}} - (\boldsymbol{\Phi} \boldsymbol{f}^{t-1})|_{\mathcal{S}^{(t-1)}}\|_2 \leq \delta \}
\end{aligned} \tag{1}
$$

where $\hat{\mathcal{S}}^{(t-1)}$ is the complement of $\mathcal{S}^{(t-1)}$. The above optimisation not only ensures that the signal of interest is sparsest outside the estimated support ($\mathcal{S}^{(t-1)}$) but also the spike values in the transform domain are close enough to the ones of the previous time instance. However, in many real-world signals, the spikes may not stay in exactly the same locations as in the previous time instance, but move about the vicinity of these locations. This motivated us to use a weighted-ℓ_1 minimisation method, where weights are set based on the reconstructed signal of the previous epoch. This method is discussed in the next section.

3. Weighted-CS for ECG signal reconstruction

3.1. Sparse domain: In CS methods, reconstruction is carried out in a transform domain ($\boldsymbol{\Phi}$) which is both sparse and incoherent with the sampling domain ($\boldsymbol{\Psi}$). We use the Gini index (GI) [14] and μ to measure sparsity and incoherency of different sparse domains, respectively. Table 1 shows the values of these two measures in different transform domains for an ECG signal. It can be seen that ECG is sparsest in the Wavelet domain (since the GI is closest to 1); however, the sampling domain is least coherent with the discrete cosine transform (DCT) domain. Therefore we select DCT as the sparse basis ($\boldsymbol{\Phi}$) and we set $\boldsymbol{F}^{(t)} = \boldsymbol{\Phi} \boldsymbol{f}^{(t)}$.

3.2. Weighted-CS algorithm: The weighted-CS is a method specifically developed for reconstructing time sequences of sparse signals, where sparsity changes smoothly with time [15]. Based on the assumption that the sparsity of the current signal is closely related to the one of the previous time instance, weighted-CS uses the reconstructed signal of the previous time instance, $\boldsymbol{F}^{(t-1)}$, and estimates the probability of each element of $\boldsymbol{F}^{(t)}$ having a non-zero value, that is, $p_i := \text{Prob}(F_i^{(t)} > \varepsilon)$. It then employs this estimated probability model to guide the reconstruction process.

The above assumption is observed to be valid in the context of ECG signals. The support of ECG signals in the DCT domain changes slowly between successive windows even when there are significant changes in the spatial domain. This is illustrated in Fig. 1, where despite differences between the signal in the spatial domain for two consecutive windows, the frequency spectrum of these windows in the DCT domain are similar. Note that the range shown in Fig. 1*b* is [1–300] as there are no frequency spikes beyond 300. Fig. 1*c* illustrates how spike locations change from one time frame to the next by plotting binary signals $\hat{\boldsymbol{F}}^{(1)}$ and $\hat{\boldsymbol{F}}^{(2)}$ corresponding to DCT signals $\boldsymbol{F}^{(1)}$ and $\boldsymbol{F}^{(2)}$, such

Table 1 Comparison of sparsity and incoherency of different transform domains

Φ	Spatial	DFT	DCT	Wavelets
GI($\boldsymbol{\Phi f}$)	0.292	0.615	0.791	**0.833**
$\mu(\Phi, \boldsymbol{\Psi})$	5.899	4.855	**4.614**	5.049

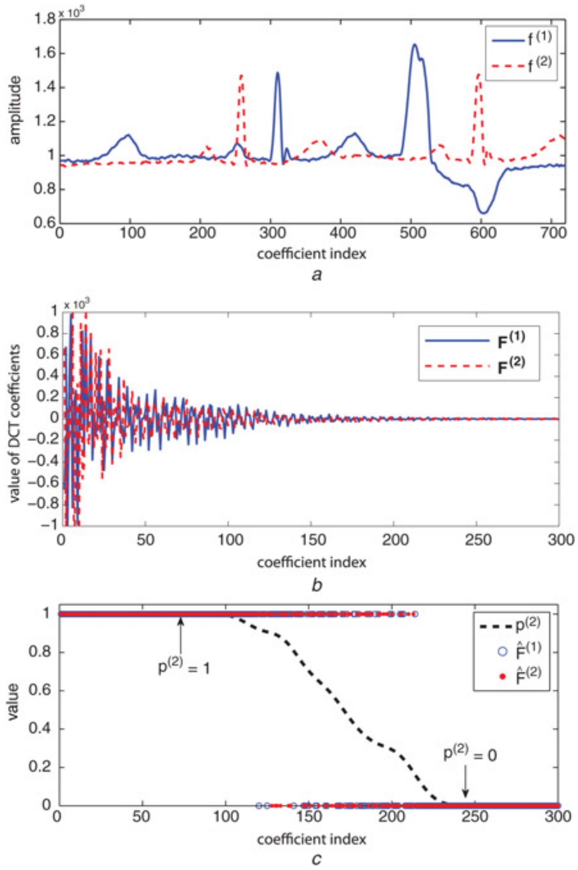

Figure 1 *Two consecutive ECG windows taken from record #119 in the spatial and DCT domains*
a Spatial domain
b DCT domain
c Estimation of $\boldsymbol{P}^{(2)}$ from $\boldsymbol{F}^{(1)}$

that a location is set to one only if a spike exists at that location that is, $\hat{F}_i = 0$ if $F_i \leq \varepsilon$ and $\hat{F}_i = 1$ if $F_i > \varepsilon$. It can be seen that the frequency spikes (•) in the second window, either occur at the same locations, especially in the lower frequency range [1–100], or in the close vicinities of the frequency spikes of the first signal window (○) for higher frequencies in the range [100–250].

3.2.1 Probability model: Based on the observation that the location of spikes in $\boldsymbol{F}^{(t)}$ is closely related to the ones of $\boldsymbol{F}^{(t-1)}$, this section describes the estimation of the sparsity probability model for a signal at time t, from the reconstructed signal at time $t-1$.

Suppose there is a spike in the ith location of $\boldsymbol{F}^{(t-1)}$. In the next time instance, there is a very-high probability that this spike remains in the same location but also some possibility that it moves to some other point in the vicinity. Based on this, the weighted-CS uses a Gaussian distribution to provide an estimate of the probability of the progression of a spike in the next time frame (see Fig. 2*a*); and probability of j^{th} element of $\boldsymbol{F}^{(t)}$ being non-zero $\left(p_j^{(t)}\right)$ is the accumulated probability of the spikes of $\boldsymbol{F}^{(t-1)}$ moving to location j at time t as follows

$$
p_j^{(t)} = \sum_{i \in \mathcal{S}^{(t-1)}} \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-(j-i)^2/2\sigma^2} \tag{2}
$$

where $\mathcal{S}^{(t-1)}$ is the support of $\boldsymbol{F}^{(t-1)}$ and σ is the variance of the Gaussian distribution.

In Fig. 1*c*, it can be seen that the maximum value $\boldsymbol{p}^{(2)}$ (the dashed line) which is extracted from $\boldsymbol{F}^{(1)}$, based on (2), is 1 and coincides with the locations where $F_j^{(1)}$ and all elements in its vicinity are all

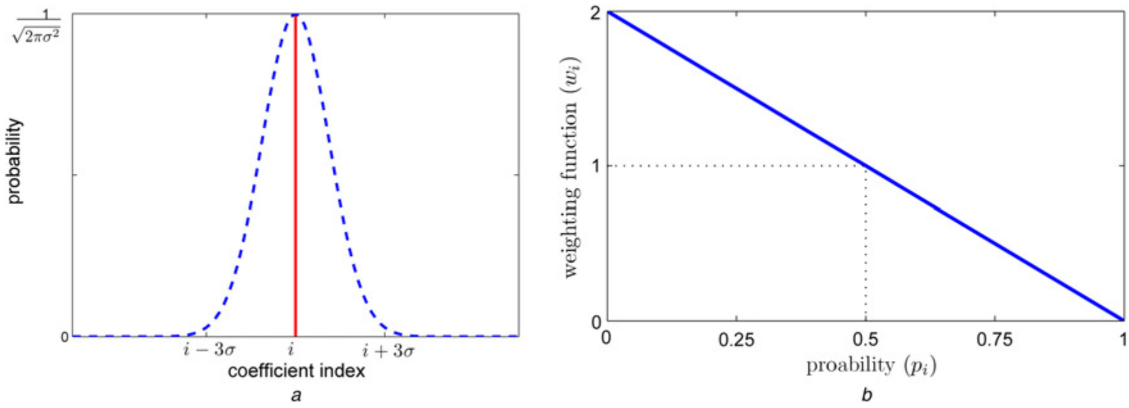

Figure 2 *Illustration of*
a Probability of progression of a spike in the next time instance
b Weights (w_i) against probability (p_i)

non-zero, that is, in the range of [0–100]). The probability decreases from 100 to 250, finally $\boldsymbol{p}^{(2)}$ becomes zero corresponding to locations where $F_j^{(1)}$ and all the elements in its vicinity are zero; from 250 to720. It can be seen that $\boldsymbol{p}^{(2)}$ indeed corresponds well to the frequencies existing in $\boldsymbol{F}^{(2)}$.

3.2.2 Reconstruction: In order to incorporate the probability model of the signal into the process of reconstruction, we minimise a weighted ℓ_1 norm (3), where the weights are adjusted according to the probability of each entry being non-zero

$$\min \| \boldsymbol{W}^{(t)}\boldsymbol{\Phi}\boldsymbol{g}^{(t)} \|_1 \quad \text{s.t.} \quad \boldsymbol{\Psi}\boldsymbol{g}^{(t)} = \boldsymbol{y}^{(\mathbf{t})} \tag{3}$$

where $\boldsymbol{W} = \text{diag}([w_1, w_2, \ldots, w_n])$. Intuitively, a smaller weight should be given to those entries with higher probability of being non-zero whereas those elements with small probability should be penalised with larger weights. Naturally, we want to reward and penalise the elements uniformly using a linear function. Thus, the choice of the weight for each element is

$$w_i = 2(1 - p_i) \tag{4}$$

Fig. 2*b* shows the chosen weights with respect to the value of the probability. It can be seen that as the probability of an element being non-zero increases, its weight decreases accordingly. Note that if $p_i = 0.5$ then $w_i = 1$.

The weighted-CS method for reconstruction of the ECG signals is summarised in Algorithm 1.

Algorithm 1: Reconstruction of ECG signals using weighted-CS:

1. If $t = 1$:$\boldsymbol{g}^{(t)} = \arg\min\{\| \boldsymbol{\Phi}\boldsymbol{g}^{(1)} \|_1 \ \text{s.t.}\boldsymbol{\Psi}^{(1)}\boldsymbol{g}^{(1)} = \boldsymbol{y}^{(1)}\}$;
2. $t = t + 1$;
3. $\mathcal{S}^{(t-1)} := \{k \in \{1, \ldots, n\}: \boldsymbol{\Phi}g_k^{(t-1)} > \varepsilon\}$;
4. Compute $\boldsymbol{p}^{(t)} := [p_1^{(t)}, \ldots, p_n^{(t)}]$ from (2);
5. Compute $\boldsymbol{W}^{(t)}$ from (4);
6. $\boldsymbol{g}^{(t)} = \arg\min\{\| \boldsymbol{W}^{(t)}\boldsymbol{\Phi}\boldsymbol{g}^{(t)} \|_1 \ \text{s.t.}\boldsymbol{\Psi}\boldsymbol{g}^{(t)} = \boldsymbol{y}^{(t)}\}$; and
7. Go to step 2.

It should be noted that at time $t = 1$, since no prior knowledge of the signal is available, all weights are equal to one and therefore (3) becomes the conventional ℓ_1 minimisation.

3.2.3 Efficient implementation of weighted-ℓ_1 minimisation: The weighted-ℓ_1 minimisation in step (6) of Algorithm 1 is a convex optimisation problem, and it is acknowledged that solving it using methods such as interior points has a polynomial running time of $O(n^3)$ [16]. To deal with the complexity, we use a class of greedy-based iterative method which is easy to implement, relatively fast and known [16, 17] to have a reconstruction performance similar to those of convex optimisation-based methods.

In order to find the minimiser of (3), let $\hat{\boldsymbol{g}}^{(t)}$ be $\boldsymbol{W}^{(t)}\boldsymbol{\Phi}\boldsymbol{g}^{(t)}$. Equation (3) can then be rewritten as

$$\min \| \hat{\boldsymbol{g}}^{(t)} \|_1 \quad \text{s.t.} \quad \boldsymbol{\Psi}\boldsymbol{\Phi}^{-1}\boldsymbol{W}^{-1}\hat{\boldsymbol{g}}^{(t)} = \boldsymbol{y}^{(t)}$$

Reformulating the above constrained optimisation problem using Lagrangian multipliers and setting $\Lambda{:} = \boldsymbol{\Psi}\boldsymbol{\Phi}^{-1}\boldsymbol{W}^{-1}$, we obtain

$$\min \lambda \| \hat{\boldsymbol{g}}^{(t)} \|_1 + \frac{1}{2} \| \boldsymbol{\Lambda}\hat{\boldsymbol{g}}^{(t)} - \boldsymbol{y}^{(t)} \|_2 \tag{5}$$

It is shown in [17] that the solution to (5) is given by the limit of the sequence

$$\hat{\boldsymbol{g}}_{it+1}^{(t)} = \mathrm{S}_\lambda(\hat{\boldsymbol{g}}_{it}^{(t)} + \boldsymbol{\Lambda}^{\mathrm{T}}(\boldsymbol{y}^{(t)} - \boldsymbol{\Lambda}\hat{\boldsymbol{g}}_{it}^{(t)})) \tag{6}$$

where S_λ is the soft-thresholding function as

$$\mathrm{S}_\lambda(x) = \begin{cases} x - \lambda & \text{if } x > \lambda \\ 0 & \text{if } |x| \le \lambda \\ x + \lambda & \text{if } x < -\lambda \end{cases}$$

Equation (6) basically minimises the error between the current estimate of the signal and the observations through the gradient decent direction of $\| \boldsymbol{\Lambda}\hat{\boldsymbol{g}}_{it}^{(t)} - \boldsymbol{y}^{(t)} \|_2$. It then imposes sparsity on the estimated signal using the soft-thresholding operator applied component-wise to each element of the input vector. Once (6) converges to $\hat{\boldsymbol{g}}_{\text{conv}}^{(t)}$, the original signal $\boldsymbol{g}^{(t)}$ can be recovered as $\boldsymbol{g}^{(t)} = \boldsymbol{\Phi}^{-1}\boldsymbol{W}^{-1}\hat{\boldsymbol{g}}_{\text{conv}}^{(t)}$ (We set $\boldsymbol{W} = \boldsymbol{W} + \varepsilon\boldsymbol{I}$ so that $\boldsymbol{W}$ is invertible.).

4. Experimental results: ECG signals of 10 min duration were extracted from the following records of the MIT-BIH Arrhythmia database [18]: 100, 101, 102, 103, 107, 109, 111, 115, 117, 118, 119, 201, 207, 208, 209, 212, 213, 214 and 232. These ECG signal datasets were used in experiments to compare the performance of the weighted-CS with the other state-of-the-art CS and DWT-based algorithms. These datasets consist of ECG signals with different rhythms, ectopic beats, QRS complex morphologies and irregularities [19]. The ECG signals were digitised at above-Nyquist sampling rate of 360 Hz and each sample is represented by 11 bits. The processing window of 2 s with length of n = 720 samples, was used for all algorithms.

The ECG signals are then reconstructed from different number of samples using Algorithm 1. We compared the results obtained with

doi: 10.1049/htl.2013.0038

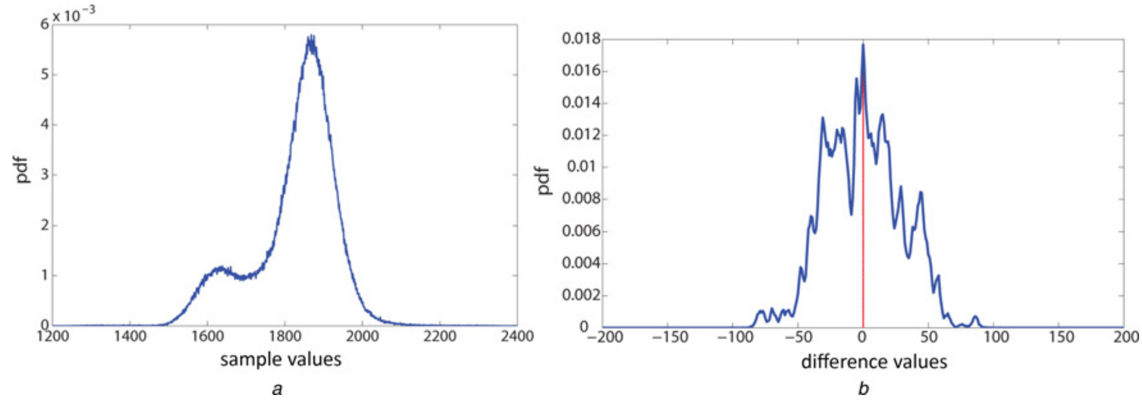

Figure 3 *Distribution of*
a Compressive sample measurements
b Their difference value

our weighted-CS algorithm with those obtained using RegMod-CS, ℓ_1 minimisation and the conventional DWT thresholding-based compression technique [2]. We used the latter as a comparison benchmark since it has been shown to outperform the embedded zerotree wavelet [19] and the set partitioning in hierarchical trees-based methods, for ECG compression at a lower cost [2].

4.1. Performance measures: Since no pre-processing has been done on the ECG datasets, we use the normalised percentage root-mean-square difference (called PRD1) in this Letter. PRD1 is independent of the dc level of the original signal and is defined as

$$\mathrm{PRD1} = \frac{\| \boldsymbol{g}^{(t)} - \boldsymbol{f}^{(t)} \|_2}{\| \boldsymbol{f}^{(t)} - \bar{f}^{(t)} \|_2} * 100$$

where $\bar{f}^{(t)}$ is the mean value of the original signal. The link between the value of PRD1 and the diagnostic quality of the reconstructed ECG signals for the MIT database was established in [20], that is, PRD1 in the ranges 0–2 and 2–9 corresponds to very good and good, respectively. The compression ratio (CR) is defined as

$$\mathrm{CR} = \frac{br_{\mathrm{o}}}{br_{\mathrm{c}}}$$

where br_{o} and br_{c} denote the number of required bits to represent the original and compressed signal, respectively.

4.2. Signal encoding in CS against DWT method: In CS only m linear combinations (referred to as compressive measurements) of the n samples of each window need to be stored and processed. In our experiments, compressive measurements are taken using $\boldsymbol{y}^{(t)} = \boldsymbol{\Psi f}^{(t)}$, where $\boldsymbol{\Psi}$ is of size $m \times n$ and its independent identically distributed (i.i.d.) entries are drawn from a Gaussian distribution $\mathcal{N}(0, 1/m)$. It is known from the CS literature that such matrices satisfy a property called restricted isometery property with an overwhelming probability, which makes them a good choice for the sampling operator [3, 21].

Fig. 3*a* shows the histogram of measurement values of the ECG datasets. Although the original signal requires 11 bits for its representation, removing redundancy between measurements prior to transmission would result in smaller 'difference values' between consecutive ECG measurements as shown in Fig. 3*b*. Just 9 bits would be sufficient to represent these difference values (which range from −200 to 200). Huffman coding with a dictionary of size 512 and average codeword length of 6.3 is then used to encode the difference signal. Huffman coding is used as an entropy coder instead of the run-length coding since the former was found to achieve higher CRs.

In the DWT method, the signal of interest is first decomposed using Daubechies wavelet of order 8 as it was found to be the most appropriate wavelet basis for ECG signals [22]. Different levels of thresholds are selected to achieve different compression rates and coefficients having values smaller than these threshold are set to zero. The run-length algorithm is then used to code the resulting vector. Note that choosing the value of the threshold may generally not be easy in portable ECG sensing devices since the noise variance of the signal is not known a priori [23]. It should be noted that in RegMod-CS and weighted-CS, reconstruction is based on information which is extracted directly from the reconstructed signal of the previous time instance. In the very first time window of each ECG signal, about 75% of the samples are used to fully recover the signal of that window without using any prior knowledge.

4.3. Performance comparison: The mean PRD1 against CR for all test datasets is presented in Fig. 4 for all methods. It can be seen that the performance of the weighted-CS is comparable with DWT and superior to the other CS-based methods in terms of PRD1. This is more prominent in the higher CRs. Figs. 5*a* and *b* show the performance variance of our method and DWT, respectively. The centre of each 'bar' represents the average performance. From Fig. 5, it can be seen that performance variation for both algorithms is more in higher CRs. Moreover, it is observed that DWT exhibits smaller variations in the performance, compared with weighted-CS. Also, for CRs of up to 11 for DWT and 10 for Weighted-CS, the reconstructed signals are of good diagnostic quality (corresponding to PRD1 $\leq$ 9 [20]).

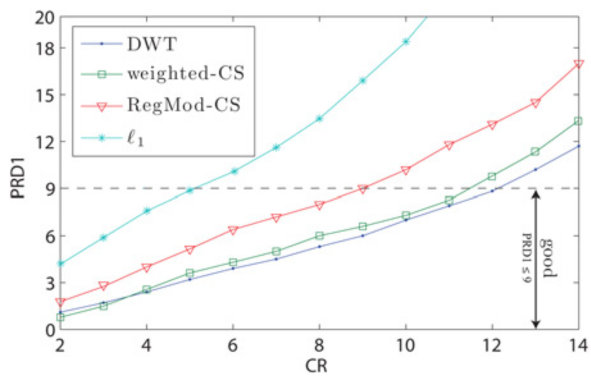

Figure 4 *Mean performance comparison over all records*

doi: 10.1049/htl.2013.0038

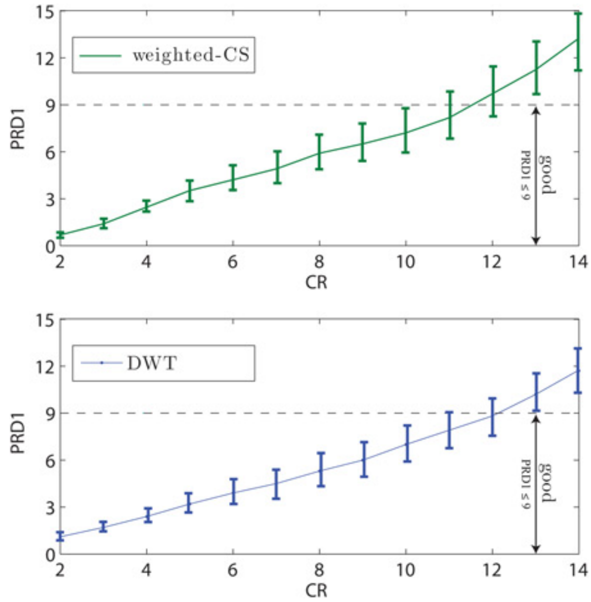

Figure 5 *Performance variance for all records*

Table 2 Average computational time per window (*ms*)

Method	ℓ_1	RegMod-CS	Weighted-CS	DWT
encoder	48	48	48	502
decoder	638	719	269	195

4.4. Running time and memory requirement: Although in DWT-based algorithms, all n samples need to be stored prior to compression, in CS-based methods only m compressive measurements need to be stored and processed ($m \ll n$). Table 2 shows the execution time (all the algorithms are implemented in Matlab 7 on a Intel Core i5 CPU with 1.6 GHz processor.) of encoder and decoder of the methods presented in this Letter for one window. From the table it can be seen that the DWT-based method, unlike CS-based approaches, results in a higher computational load at the sampling/encoder end rather than at the decoding end. In addition, it can be seen that the decoding time of the weighted-CS is comparable with the one of the DWT method.

5. Discussion: It is evident from Fig. 4 that our method significantly out-performs the RegMod-CS in terms of mean PRD1 at different CRs and this is because of two main reasons. Firstly, while RegMod-CS is based on the assumption that the values of spikes in the DCT domain are close enough to the ones of the previous time instance, the weighted-CS only uses the location of the spikes of the previous time instance. The reason is that even though the location of the spikes tend to remain within a close vicinity of the ones of the previous time instance, their values do change and therefore $\|\boldsymbol{F}^{(t)} - \boldsymbol{F}^{(t-1)}\|_2$ is not as small as expected (see Figs. 1*b* and *c*). Secondly, instead of just using the support of the previously reconstructed signal, $\mathcal{S}^{(t-1)}$, as the estimated support of the current frame, weighted-CS obtains a probability vector which remains valid even when the support is changing.

Aside from comparing with the CS-based methods, we have also compared our method with a DWT-based signal compression algorithm which has been shown to be superior to other wavelets-based methods for ECG signal compression and at a lower cost [1, 2]. Based on the mean PRD1 presented in Fig. 4, the performance variances reported in Fig. 5 and running times given in Table 2, it can

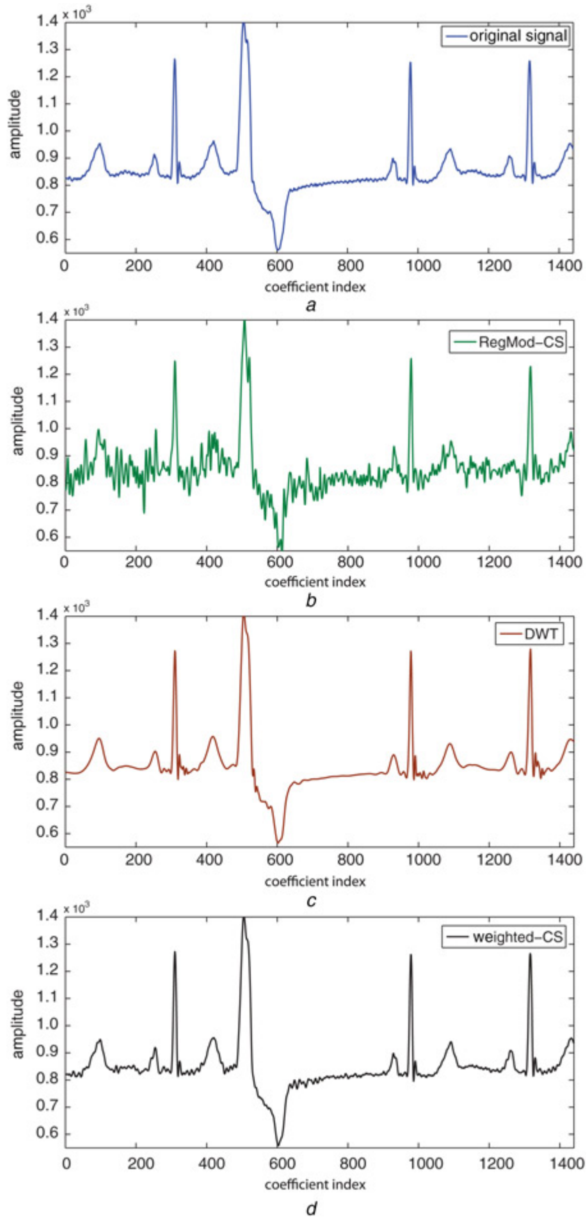

Figure 6 *Reconstructed signals (record 119)*
a Original signal
b RegMod-CS reconstruction, CR = 10.63 and PRD1 = 13.69
c DWT reconstruction, CR = 10.41 and PRD1 = 8.76
d Weighted-CS reconstruction, CR = 10.63 and PRD1 = 9.03

be concluded that the performance of the weighted-CS is quite comparable with that of the DWT-based method in terms of the reconstruction PRD1 but with significantly lower computational complexity and memory requirements imposed on the encoder. From Figs. 6*c* and *d* which shows window segments of reconstructed irregular heartbeat signals, it is evident that our method performs perceptually better with less loss of details although our method is slightly inferior to DWT in terms of the PRD1. The DWT method generally tends to reconstruct over-smoothed signals which could potentially lead to loss of some important diagnostic features.

One of the main differences between CS-based and conventional DWT-based methods is that in DWT compression, all n samples are stored prior to compression and then an optimal representation of the signal, which is equivalent to the m most significant coefficients of the sparse domain, is encoded. The location of these significant coefficients are not known beforehand. Therefore there is a need to

Healthcare Technology Letters, 2014, Vol. 1, Iss. 2, pp. 68–73
doi: 10.1049/htl.2013.0038

either directly encode the whole signal of size n in the sparse domain or encode the m significant bits, along with their corresponding indices using two different codebooks. On the other hand, instead of taking the best m coefficients, m linear combination of all the coefficients are taken in CS-based methods. This sampling matrix could be fixed and known to the receiver device and therefore only the values of the samples need to be encoded and transmitted. It should be noted that in wavelets-based methods the significant bits can be coded in order of their importance (depending on their magnitude) and therefore the encoding process can be stopped at any point once the target quality/compression is achieved. This is however not possible in CS, as compressive measurements are all of the same importance to the reconstruction process.

5.1. Choosing the window size: The window size is of vital importance since it was observed that increasing the window size improves the CR for a given distortion, but on the other hand increases the running time and memory requirements. We observed a good compromise between processing time and CR to be a window size of 720. However, the decision about this is fundamentally open. One may even choose more complicated (adaptive) windows which for example include a certain number of heart beats.

It should be also mentioned that in the last stage of our decoder, we used a relatively simple lossless coding scheme, and therefore there is still much room for improvement. We expect that a more elaborate entropy coder would result in greatly improved performance. In summary, based on the experimental results presented in this letter, weighted-CS appears to be an appealing alternative to the traditional full acquisition/compression methodology. This might potentially have a significant impact on the future of low-power wearable ECG monitoring devices.

6. Conclusions: We have successfully adapted the recently introduced weighted-CS method for reconstructing ECG signals from a small number of samples, by integrating an iterative method to more efficiently solve the minimisation in the weighted-CS. Our weighted-CS algorithm uses information in the DCT domain to efficiently reconstruct ECG signals using a weighted ℓ_1 minimisation. Our experiments on ECG signals, show that our algorithm achieves competitive performance with the state-of-the-art DWT compression method, but at much lower computational cost and memory requirements.

7. Acknowledgments: This work is supported by a grant (TDSI/11-014/1A) from the Temasek Defence Systems Institute (TDSI), Singapore.

8 References

[1] Mamaghanian H., Khaled N., Atienza D., Vandergheynst P.: ‘Compressed sensing for real-time energy efficient ECG compression on wireless body sensor nodes’, *IEEE Trans. Biomed. Eng.*, 2011, **58**, (9), pp. 2456–2465

[2] Benzid R., Marir F., Boussaad A., Benyousef M., Arar D.: ‘Fixed percentage of wavelet coefficients to be zeroed for ECG compression’, *Electron. Lett.*, 2003, **39**, (11), pp. 830–831

[3] Candes E., Wakin M.: ‘An introduction to compressive sampling’, *IEEE Signal Process. Mag.*, 2008, **25**, (2), pp. 21–30

[4] Donoho D.: ‘Compressed sensing’, *IEEE Trans. Inf. Theory*, 2006, **52**, (4), pp. 1289–1306

[5] Venkatesh Y.V., Kassim A.A., Zonoobi D.: ‘Medical image reconstruction from sparse samples using simultaneous perturbation stochastic optimization’. ICIP, Proc. of the ICIP Conf., Hong Kong, September 26–29, 2010

[6] Rachlin Y., Baron D.: ‘The secrecy of compressed sensing measurements’. Forty-sixth Annual Allerton Conf. on Communication, Control and Computing, 2008, pp. 813–817

[7] Fira M., Goras L., Barabasa C., Cleju N.: ‘On ECG compressed sensing using specific overcomplete dictionaries’, *Adv. Electr. Comput. Eng.*, 2010, **10**, (4), pp. 23–28

[8] Cleju N., Fira M., Barabasa C., Goras L.: ‘Robust reconstruction of compressively sensed ECG signals’. Tenth Int. Symp. on Signals, Circuits and Systems (ISSCS), 2011, pp. 1–4

[9] Allstot E., Chen A., Dixon A., Gangopadhyay D., Mitsuda H., Allstot D.: ‘Compressed sensing of ECG bio-signals using one-bit measurement matrices’. IEEE 9th Int. Conf. New Circuits and Systems (NEWCAS), June 2011, pp. 213–216

[10] Dixon A., Allstot E., Gangopadhyay D., Allstot D.: ‘Compressed sensing system considerations for ECG and EMG wireless biosensors’, *IEEE Trans. Biomed. Circuits Syst.*, 2012, **6**, (2), pp. 156–166

[11] Vaswani N., Lu W.: ‘Modified-CS: modifying compressive sensing for problems with partially known support’, *IEEE Trans. Signal Process.*, 2010, **58**, (9), pp. 4595–4607

[12] Senay S., Chaparro L.F., Sun M., Sclabassi R.J.: ‘Compressive sensing and random filtering of EEG signals using Slepian basis’. Proc. EURASIP EUSIPCO’08, 2008

[13] Lu W., Vaswani N.: ‘Regularized modified bpdn for noisy sparse reconstruction with partial erroneous support and signal value knowledge’, *IEEE Trans. Signal Process.*, 2012, **60**, (1), pp. 182–196

[14] Zonoobi D., Kassim A.A., Venkatesh Y.V.: ‘Gini index as sparsity measure for signal reconstruction from compressive samples’, *IEEE J. Sel. Top. Signal Process.*, 2011, **5**, (5), pp. 927–932

[15] Zonoobi D., Kassim A.A.: ‘On the reconstruction of sequences of sparse signals – the weighted-CS’, *J. Vis. Commun. Image Represent.*, 2013, **24**, (2), pp. 196–202

[16] Blumensath M.D.T.: ‘Iterative thresholding for sparse approximations’, *J. Fourier Anal. Appl.*, 2008, **14**, (5), pp. 629–654

[17] Daubechies I., Defrise M., De Mol C.: ‘An iterative thresholding algorithm for linear inverse problems with a sparsity constraint’, *Commun. Pure Appl. Math.*, 2004, **57**, (11), pp. 1413–1457

[18] Harvard-MIT Division of Health Sciences and Technology: ‘MIT-BIH database distribution’. Available: http://www.ecg.mit.edu

[19] Blanco-Velasco M., Cruz-Roldan F., Moreno-Martinez E., Godino-Llorente J.-I., Barner K.E.: ‘Embedded filter bank-based algorithm for ECG compression’, *Signal Process.*, 2008, **88**, (6), pp. 1402–1412

[20] Zigel Y., Cohen A., Katz A.: ‘The weighted diagnostic distortion (wdd) measure for ECG signal compression’, *IEEE Trans. Biomed. Eng.*, 2000, **47**, pp. 1422–1430

[21] Candes E., Tao T.: ‘Near-optimal signal recovery from random projections: universal encoding strategies?’, *IEEE Trans. Inf. Theory*, 2006, **52**, (12), pp. 5406–25

[22] Singh B.N., Tiwari A.K.: ‘Optimal selection of wavelet basis function applied to ECG signal denoising’, *Digit. Signal Process.*, 2006, **16**, (3), pp. 275–287

[23] del Mar Elena M., Quero J.M., Borrego I.: ‘Optimal selection of wavelet coefficients for electrocardiograph compression’, *ETRI J.*, 2007, **29**, (4), pp. 530–532