Overcomplete Multi-scale Dictionaries for Efficient Representation of ECG Signals

David Meltzer[1] , David Luengo[1(✉)] , and Tom Trigano[2]

[1] Universidad Politécnica de Madrid, 28031 Madrid, Spain
{david.meltzer,david.luengo}@upm.es
[2] Shamoon College of Engineering, 77245 Ashdod, Israel
thomast@sce.ac.il

Abstract. The electrocardiogram (ECG) was the first biomedical signal subject of extensive digital signal processing techniques. Essentially, the ECG consists of a cyclic sequence of relevant activations embedded into inactivity time sequences combined with interferences and noise. By its nature, it can be subject of representation as a sparse signal. This work describes an efficient method to create overcomplete multi-scale dictionaries that can be used for sparse ECG representation. Whereas most of the proposed methods to date use fixed waveforms that somehow resemble actual ECG shapes, the main innovation in our approach is selecting ECG waveforms recorded from actual patients. A relevant result of our method is the ability to process long lasting recordings from multiple patients. Simulations on patient actual records from Physionet's PTB Diagnostic ECG Database confirm the good performance of the proposed approach.

Keywords: Electrocardiogram (ECG) · LASSO · Overcomplete multi-scale signal representation · Dictionary construction · Sparse inference

1 Introduction

It is a well known fact that, since the viable introduction of the electrocardiogram (ECG) recording, its use has become so widespread that it is now commonly used in the health-care sector for obtaining indicators of patient health status. The use and interpretation of the ECG has been dramatically extended by means of the application of digital signal processing (DSP) techniques. Noise and interference removal, extraction of specific signals from the composite ECG, detection and characterization of waveforms are examples of solutions provided by DSP techniques [3].

The time representation of an ECG shows cyclic waveforms corresponding to QRS complexes, as well as P and T waveforms, embedded into non active time sequences corresponding to isoelectric intervals, all this together with noise

and interferences [12]. Therefore, the ECG can be considered a sparse signal. A number of sparse inference and representation techniques have been proposed for many types of signals (images, audio, recordings, biomedical waveforms, etc.) [4] since the introduction of the Least Absolute Shrinkage and Selection Operator (LASSO) regularizer [13].

Most of the works done regarding electrocardiographic signal processing and aiming for sparse representation of single-channel or multi-channel ECGs [1,9,11, 14] use fixed arbitrary waveforms that somehow resemble the characteristic shape of a QRS complex (e.g., a mexican hat wavelet). This type of approach leads to practical positive results that usually include a number of spurious activations which need a removal procedure before their potential use as physiologically interpretable signals. This is typically performed using a post-processing stage [9,11] or by minimizing a complex non-convex cost function [10].

In this work, the authors propose a method to conform a multi-scale dictionary derived from a set of atoms that have their origin in ECG waveforms recorded from actual patients. The aforementioned dictionary is then used to obtain spare representations of ECGs using the LASSO regularizer. Numerical simulations show that the proposed approach leads to very sparse representations that both represent al QRS complexes without misses while lacking spurious activations, which eliminates the need for any post-processing.

The rest of this document is organized as follows. The next section states the problem of sparse representation of ECG signals, with special emphasis in the relevance of counting on an appropriate dictionary. Then, Sect. 3 describes the procedure followed to derive a multi-scale dictionary from one or more atoms that have their origin in ECG waveforms recorded from actual patients. Finally, the next section provides the results of experiments that validate the proposed approach using ECG records from the Physikalisch-Technische Bundesanstalt (PTB) [5] database from PhysioNet and the conclusions close the paper.

2   Sparse Representation Problem Formulation

Let us consider a single-lead discrete-time ECG, $y[n]$, obtained from a properly filtered and amplified continuous-time single-lead ECG, $y_c(t)$, using uniform sampling with a period $T_s = 1/f_s$, i.e., $y[n] = y_c(nT_s)$. Assuming that this signal is obtained as the superposition of a number of waveforms of interest (essentially QRS complexes, P and T waveforms) plus noise and interferences, then $y[n]$ can be expressed as:

$$y[n] = \sum_{k=-\infty}^{\infty} E_k \Phi_k(t_n - T_k) + \varepsilon[n], \quad n = 0, \ldots, N-1, \tag{1}$$

where T_k denotes the arrival time of the k-th electrical pulse; E_k its amplitude; Φ_k the associated, unknown pulse shape; and $\varepsilon[n]$ an additive, white Gaussian noise (AWGN) with variance σ^2. Equation (1) is known in the literature as the *shot-noise model*, and it is commonly used in the digital signal processing of

biomedical signals. Note that, in real-world applications neither Φ_k, T_k nor E_k are known. Even so, if the most common shapes and durations of Φ_k are known, each one can be approximated using a time-shifted, multi-scale dictionary of waveforms from actual patients recordings with finite support $M \ll N$ and this approximation can then be used to infer the values of E_k and T_k. More precisely, let us define a set of P candidate waveforms, Γ_p for $p = 1, \ldots, P$, with a finite support of M_p samples such that $M_1 < M_2 < \cdots < M_P$ and $M = \max_{p=1,\ldots,P} M_p = M_P$. If properly chosen, these waveforms will provide a good approximation of the local behavior of the signal around each sampling point, thus allowing us to approximate (1) through the following model:

$$y[n] = \sum_{k=0}^{N-M-1} \sum_{p=1}^{P} \beta_{k,p} \Gamma_p[n - k] + \varepsilon[n], \quad n = 0, \ldots, N - 1, \qquad (2)$$

where the $\beta_{k,p}$ indicate the amplitude of the p-th waveform shifted to the k-th time instant, $t_k = kT_s$. It is possible to group all the candidate waveforms into a single matrix $\mathbf{A} = [\mathbf{A}_0 \ \mathbf{A}_1 \ \cdots \ \mathbf{A}_{N-M-1}]$, where the $N \times P$ matrices $\mathbf{A}_k$ (for $k = 0, \ldots, N - M - 1$) have column entries equal to $\Gamma_p[m - k]$ for $m = k, \ldots, k + M - 1$ and 0 otherwise. With this grouping, the model of (2) can be expressed in matrix form as follows:

$$\boldsymbol{y} = \mathbf{A}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \qquad (3)$$

where $\boldsymbol{y} = [y[0], \ldots, y[N-1]]^\top$ is an $N \times 1$ vector with all the ECG samples, $\boldsymbol{\beta} = [\beta_{0,1}, \ldots, \beta_{0,P}, \ldots, \beta_{N-M-1,1}, \ldots, \beta_{N-M-1,P}]^\top$ is an $(N - M)P \times 1$ coefficients vector, and $\boldsymbol{\varepsilon} = [\varepsilon[0], \ldots, \varepsilon[N-1]]^\top$ is the $N \times 1$ noise vector. Note that matrix $\boldsymbol{A}$ can be seen as a global dictionary that contains $N - M$ replicas, $\mathbf{A}_k$ for $k = 0, 1, \ldots, N - M - 1$, of the candidate waveforms time shifted to $t_0 = 0$, $t_1 = T_s, \ldots, t_{N-M-1} = (N - M - 1)T_s$. Two measures can now be taken to cope with the uncertainty about the shape and duration of the pulses that can be found in $y[n]$: using several different waveforms with distinct timescales. The result is a time-shifted, multi-scale overcomplete (as $(N-M)P > N$) dictionary. It is important to remark that, once the dictionary signals are defined, the only remaining unknown term in (3) is $\boldsymbol{\beta}$. In this case, since the presence of relevant waveforms in $y[n]$ is typically sparse, the usual approach consists in enforcing sparsity in $\boldsymbol{\beta}$ by applying the so called LASSO, which minimizes a convex cost function composed of the L_2 norm of the reconstruction error and the L_1 norm of the coefficient vector,

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \ \|\mathbf{y} - \mathbf{A}\boldsymbol{\beta}\|_2^2 + \lambda\|\boldsymbol{\beta}\|_1, \qquad (4)$$

where λ is a parameter defining the trade-off between the sparsity of $\boldsymbol{\beta}$ and the precision of the estimation.

3   Multi-scale Dictionary Derivation

As noted in the introduction, most previous works related to construction of sub-dictionaries $\boldsymbol{A}_k$ typically use a single waveform unrelated to actual ECG waveforms, like Gaussians [11] with different variances or the mexican hat wavelet

[9]. This work is an extension of [8], where a single waveform was used to construct the dictionary. This extension aims to populate a dictionary with multiple waveforms derived from actual ECG patients recordings, selected on the basis of a distance metric. Let $Q = 52$ be the recording sessions from different healthy control subjects in the database where, for simplicity reasons, a single channel is considered corresponding to the precordial V4 electrode. The following steps have been applied to each one of the sessions:

1. Identification of the fiducial points of every QRS complex in the session following the approach described in [6]. As a result, the characterization set of all the P-QRS-T waveforms is available in the session.
2. Discard of those records for which fiducial points cannot be reliably found. For the remaining records, all identified QRS complexes are extracted, centered with respect to R and resampled in such a way that they all end up being described with the same number of samples as the longest one, L_i samples.
3. Derivation of the average QRS waveform corresponding to a subject. The standard deviation is also computed as a safety measure: if it is too large, the resulting waveform is discarded.

At this point, a number of $Q' \leq Q$ average waveforms are made finally available from different patients. All of them can potentially be used to build sub-dictionaries. All Q' waveforms are highly correlated and thus using them all would result in a dictionary that would provide a poor performance and lead to high computation consumption. Therefore, in order to obtain a reduced dictionary composed of a few waveforms as differentiated as possible, we perform the following procedure:

1. Resample all the average QRS waveforms in such a way that they all have exactly the same number of samples as the longest one, $L = \max_{i=1,...,Q'} L_i$ samples.
2. Normalize all the waveforms by removing their means and dividing by the square root of their ℓ_2 norms. Remove border effects following the technique described in [7].
3. Compute the correlation coefficient among each pair of waveforms,[1]
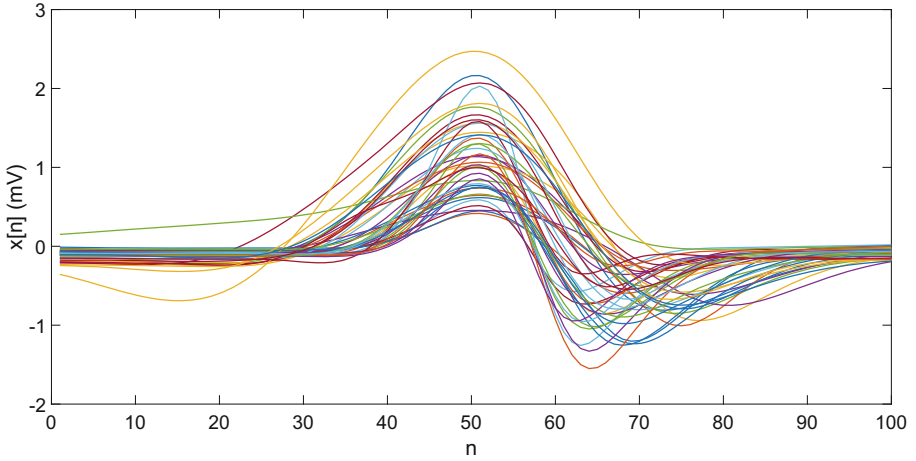
$$\rho_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}}, \tag{5}$$

where C_{ij} denotes the cross-covariance between the i-th and j-th waveforms at lag 0 (i.e., without any time shift).
4. Select the waveform with the highest average correlation (in absolute value) with respect to the rest of candidate waveforms, i.e.,

$$\ell = \arg\max_{i=1,...,Q'} \sum_{j=1}^{Q'} |\rho_{ij}|, \tag{6}$$

---

[1] In practice, this only has to be done $Q'(Q'-1)/2$ times, since $\rho_{ii} = 1 \ \forall i$ thanks to the normalization and $\rho_{ij} = \rho_{ji}$.

Fig. 1. $Q' = 44$ reliable average QRS complexes extracted from the $Q = 52$ healthy patients in the PTB database after resampling, normalization and removal of border effects.

which corresponds to the most representative waveform of all candidate waveforms.

5. If a number $K > 1$ of atoms is needed, additional waveforms can be sequentially selected from the remaining waveforms set:
   (a) Select the waveform with the highest average correlation (in absolute value) with respect to the remaining candidate waveforms (to obtain representative dictionary atoms).
   (b) Compute the correlation of the selected waveform w.r.t. the already accepted waveforms. If this correlation (in absolute value) is below a pre-established threshold γ (to avoid similar atoms), then accept it. Otherwise, discard it. In any case, remove the selected waveform from the pool of candidates.

   Once K waveforms have been selected, all of them are resampled in order to obtain a multi-scale sub-dictionary composed of T different time scales. Note that the total number of available waveforms in the resulting dictionary is thus $P = KT$. Note also that the global dictionary is simply obtained by performing $N - M$ different time shifts on the resulting sub-dictionary [9].

4   Experimental Results

In this section we present the experimental results obtained both for building the dictionary and to perform sparse reconstruction of a number of actual ECG signals using it.
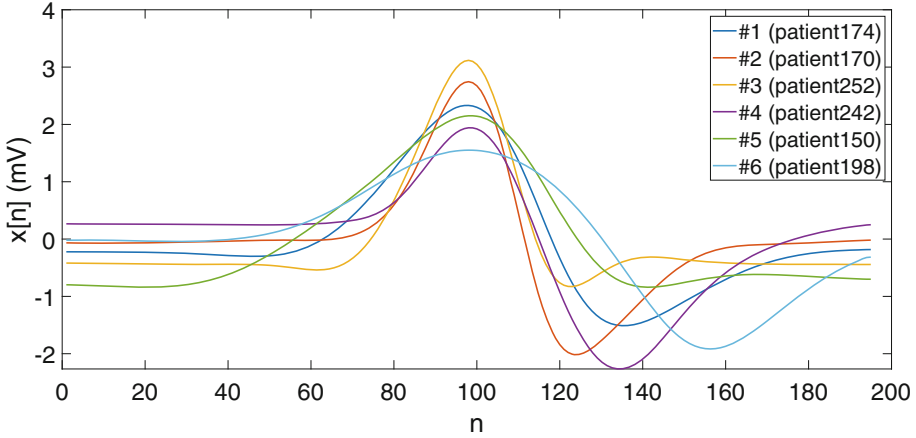
Fig. 2. Sub-dictionary composed of six R-centered average QRS waveforms.

4.1 Dictionary Construction

In order to construct the dictionary, we use the Physikalisch-Technische Bundesanstalt (PTB) database, compiled by the National Metrology Institute of Germany for research, algorithmic benchmarking and teaching purposes [2]. The ECGs were collected from healthy volunteers and patients with different heart diseases by Prof. Michael Oeff, at the Dep. of Cardiology of Univ. Clinic Benjamin Franklin in Berlin (Germany), and can be downloaded from Physionet [5]. The database contains 549 records from 290 subjects (aged 17 to 87 years), where each record includes 15 simultaneously measured signals: the 12 standard leads plus the 3 Frank lead ECGs [3,12]. Each signal is digitized using a sampling frequency $f_s = 1000$ Hz with a 16 bit resolution.

Out of the 268 subjects for which the clinical summary is available, we selected the $Q = 52$ healthy control patients in order to build the dictionary. From those patients, we were able to obtain reliable average QRS complexes in $Q' = 44$ patients. The average waveforms for these 44 patients, after resampling to $L = 100$ samples, normalization and removal of border effects, can be seen in Fig. 1. Note the large similarity among all waveforms when they are centered around the R peak. Given this set of average waveforms, the previously described selection process for atoms is performed. Figure 2 shows a sub-dictionary composed of the 6 atoms obtained by setting $\gamma = 0.9$. To finalize the construction of the dictionary, from each sub-dictionary atom a number of related waveforms are derived by resampling, so that each one has a duration that ranges from 60 ms to 160 ms in 10 ms steps. Hence, the resulting dictionary contains $P = 6 \times 11 = 66$ waveforms that can be easily time shifted in order to construct the sparse ECG representations performed in the following section.

4.2 Sparse ECG Representation

Now we test the constructed dictionary on 5 healthy patients from the PTB database. In order to solve (4), we use the CoSA (Convolutional Sparse Approximation) algorithm recently proposed in [15], which allows us to process the whole signals (approximately 115200 samples, nearly 2 min of recorded time each) at once (i.e., without having to partition them into several segments that have to be processed separately) in a reasonable amount of time. Since several signals showed a significant degree of baseline wander, before applying CoSA all signals were previously filtered using a third-order high-pass IIR (infinite impulse response) Butterworth filter designed using Matlab's `filterDesigner`: stop-band frequency $f_{stop} = 0.1\,\mathrm{Hz}$, pass-band frequency $f_{pass} = 1\,\mathrm{Hz}$, minimum stop-band attenuation $A_{stop} = 40\,\mathrm{dB}$, and maximun pass-band attenuation $A_{pass} = 1\,\mathrm{dB}$.

The results are displayed in Table 1, where each row is identified by the record patient number, the degree of sparsity (measured as the percentage of non-zero coefficients in $\boldsymbol{\beta}$), and the relative error (measured as the percentage of the ℓ_2 norm of the residual error w.r.t. the ℓ_2 norm of the signal). Note the large degree of sparsity attained, which is slightly lower in all cases than the one achieved in [8]. The large relative error shown is due to the fact that the dictionary is optimized for the detection of the QRS complexes (which is usually the first step in ECG signal processing) and does not recover other existing waveforms in the signal (like P and T waves). This can be easily corrected by incorporating additional waveforms to the dictionary in the future. However, note also that this error is substantially lower than the one achieved in [8] using $K = 1$, thus confirming the interest of building dictionaries with multiple waveforms.

Table 1. Sparsity and relative error for 5 signals from the PTB database using the derived dictionary for $K = 1$ [8] and $K = 6$.

Signal	$K = 1$		$K = 6$	
	Sparsity	Rel. error	Sparsity	Rel. error
104	2.84%	30.14%	2.18%	9.89%
105	1.82%	57.74%	1.10%	30.93%
116	1.24%	49.69%	1.05%	14.10%
121	1.53%	24.39%	1.28%	4.94%
122	3.44%	36.23%	2.51%	13.58%

5 Conclusions

In this paper, we have presented a simple mechanism to build multi-scale dictionaries using several distinct QRS complexes directly extracted from real-world

signals. The proposed approach has been used to model ECG signals from healthy patients, showing promising results in terms of the achieved sparsity and the reconstruction error. Future works will focus on the extension of this procedure to patients with cardiac diseases (where robust methods for fiducial point extraction will be essential) and the construction of dictionaries for the other relevant waveforms in the ECG (P and T waves).

References

1. Billah, M.S., Mahmud, T.B., Snigdha, F.S., Arafat, M.A.: A novel method to model ECG beats using Gaussian functions. In: 4th IEEE International Conference on Biomedical Engineering and Informatics (BMEI), vol. 2, pp. 612–616 (2011)
2. Bousseljot, R., Kreiseler, D., Schnabel, A.: Nutzung der EKG-signaldatenbank CARDIODAT der PTB über das internet. Biomedizinische Technik/Biomed. Eng. **40**(s1), 317–318 (1995)
3. Clifford, G.D., Azuaje, F., McSharry, P., et al.: Advanced Methods and Tools for ECG Data Analysis. Artech House, Boston (2006)
4. Elad, M.: Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing. Springer, New York (2010). https://doi.org/10.1007/978-1-4419-7011-4
5. Goldberger, A.L., et al.: PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation **101**(23), e215–e220 (2000)
6. Israel, S.A., Irvine, J.M., Cheng, A., Wiederhold, M.D., Wiederhold, B.K.: ECG to identify individuals. Pattern Recogn. **38**(1), 133–142 (2005)
7. Luengo, D., Meltzer, D., Trigano, T.: An efficient method to learn overcomplete multi-scale dictionaries of ECG signals. Appl. Sci. **8**(12), 2569 (2018)
8. Luengo, D., Meltzer, D., Trigano, T.: Sparse ECG representation with a multi-scale dictionary derived from real-world signals. In: 41st IEEE International Conference on Telecommunications and Signal Processing (TSP), pp. 1–5 (2018)
9. Luengo, D., Monzón, S., Trigano, T., Vía, J., Artés-Rodríguez, A.: Blind analysis of atrial fibrillation electrograms: a sparsity-aware formulation. Integr. Comput.-Aided Eng. **22**(1), 71–85 (2015)
10. Luengo, D., Vía, J., Monzón, S., Trigano, T., Artés-Rodríguez, A.: Cross-products LASSO. In: IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 6118–6122 (2013)
11. Monzón, S., Trigano, T., Luengo, D., Artes-Rodriguez, A.: Sparse spectral analysis of atrial fibrillation electrograms. In: IEEE International Workshop on Machine Learning for Signal Processing, pp. 1–6 (2012)
12. Sörnmo, L., Laguna, P.: Bioelectrical Signal Processing in Cardiac and Neurological Applications, vol. 8. Academic Press, Boston (2005)
13. Tibshirani, R.: Regression shrinkage and selection via the LASSO. J. Roy. Stat. Soc.: Ser. B (Methodol.) **58**(1), 267–288 (1996)
14. Trigano, T., Kolesnikov, V., Luengo, D., Artés-Rodríguez, A.: Grouped sparsity algorithm for multichannel intracardiac ECG synchronization. In: 22nd European Signal Processing Conference (EUSIPCO), pp. 1537–1541 (2014)
15. Trigano, T., Shevtsov, I., Luengo, D.: CoSA: an accelerated ISTA algorithm for dictionaries based on translated waveforms. Sig. Process. **139**, 131–135 (2017)