On the application of optimal wavelet filter banks for ECG signal classification

On the application of optimal wavelet filter banks for ECG signal classification

2014 J. Phys.: Conf. Ser. 490 012142

(http://iopscience.iop.org/1742-6596/490/1/012142)

Journal of Physics: Conference Series **490** (2014) 012142 doi:10.1088/1742-6596/490/1/012142

On the application of optimal wavelet filter banks for ECG signal classification

S Hadjiloucas[1*], N Jannah[1] F. Hwang[1] and R K H Galvão[2]
[1]School of Systems Engineering the University of Reading
[2]Divisão de Engenharia Eletrônica, Instituto Tecnológico de Aeronáutica, São José dos Campos, SP, 12228-900 Brazil

E-mail: s.hadjiloucas@reading.ac.uk

Abstract. This paper discusses ECG signal classification after parametrizing the ECG waveforms in the wavelet domain. Signal decomposition using perfect reconstruction quadrature mirror filter banks can provide a very parsimonious representation of ECG signals. In the current work, the filter parameters are adjusted by a numerical optimization algorithm in order to minimize a cost function associated to the filter cut-off sharpness. The goal consists of achieving a better compromise between frequency selectivity and time resolution at each decomposition level than standard orthogonal filter banks such as those of the Daubechies and Coiflet families. Our aim is to optimally decompose the signals in the wavelet domain so that they can be subsequently used as inputs for training to a neural network classifier.

Introduction

Signal decomposition using perfect reconstruction quadrature mirror filter banks can provide a very parsimonious representation of ECG signals. In previous works [1] we have shown that optimal wavelets can be used for the post-processing of ECG signals so that classifiers can operate directly in the wavelet domain as opposed to the time or frequency domains. Our approach extends the wavelet parametrization approach proposed by Sherlock and Monro [2] to ensure that the derived wavelets have at least two vanishing moments. In the current work, the filter parameters are adjusted by a numerical optimization algorithm in order to minimize a cost function associated to the filter cut-off sharpness. The goal consists of achieving a better compromise between frequency selectivity and time resolution at each decomposition level than standard orthogonal filter banks such as those of the Daubechies and Coiflet families.

Wavelet filter bank parametrization

In the signal decomposition using the DWT, both a low pass (LPF) and a high pass (HPF) filter bank are used to generate time domain responses, these are convolved with the time domain ECG signal. Convolving the response function of the chosen filter (corresponding to a particular mother wavelet) with the signal provides an output which has different energy at different scales. Approximation coefficients relate to the low frequency components of the signal whereas detail coefficients relate to the higher frequency components in the signal. Wavelet decomposition using the DWT provides essentially a multi-resolution representation of the input signal. The user normally retains coefficients

[1*] To whom any correspondence should be addressed.

up to a particular scale whereas more detailed decompositions become redundant as their incorporation have a negligible effect on the signal. The convolution operation may be conveniently performed in the frequency domain where it is implemented through a simple multiplication process.

In this filter bank, the low-pass filtering result undergoes successive filtering iterations with the number of iterations N_{it} chosen by the analyst. The final result of the decomposition of data vector **x** is a vector resulting from the concatenation of row vectors $\mathbf{c}(N_{it})$ (termed approximation coefficient at the largest scale level) and $\mathbf{d}(s)$ (termed detail coefficients at the s^{th} scale level, $s = 1,..., N_{it}$) in the following manner:

$$\mathbf{t} = [\mathbf{c}(N_{it}) \mid \mathbf{d}(N_{it}) \mid \mathbf{d}(N_{it}-1) \mid \ldots \mid \mathbf{d}(1)] \tag{1}$$

with coefficients in larger scales (e.g. $\mathbf{d}(N_{it}),\mathbf{d}(N_{it}-1),\mathbf{d}(N_{it}-2),\ldots$) associated with broad features in the data vector, and coefficients in smaller scales (e.g. $\mathbf{d}(1),\mathbf{d}(2),\mathbf{d}(3),\ldots$) associated with narrower features such as sharp peaks. The filter bank transform can be regarded as a change in variables from $\Re^J$ to $\Re^J$ performed according to the following operation,

$$t_j = \sum_{n=0}^{J-1} x_n v_j(n), \quad j = 0, 1, ..., J-1 \tag{2}$$

where t_j is a transformed variable and $v_j(n) \in \Re$ is a transform weight. It proves convenient to write the transform in matrix form as:

$$\mathbf{t}_{1\times J} = \mathbf{x}_{1\times J}\mathbf{V}_{J\times J} \tag{3}$$

where $\mathbf{x} = [x_0\ x_1\ \ldots\ x_{J-1}]$ is the row vector of original variables, **t** is the row vector of new (transformed) variables and **V** is the matrix of weights. Choosing **V** to be unitary (that is, $\mathbf{V}^{\mathrm{T}}\mathbf{V} = \mathbf{I}$), the transform is said to be orthogonal and it, therefore, consists of a simple rotation in the coordinate axes (with the new axes directions determined by the columns of **V**).

Let $\{h_0, h_1, \ldots, h_{2N-1}\}$ and $\{g_0, g_1, \ldots, g_{2N-1}\}$ be the impulse responses of the low-pass and high-pass filters respectively. Assuming that filtering is carried out by circular convolution, the procedure for generating the approximation coefficients from the data vector **x** is illustrated in Table 1. The convolution consists of flipping the filtering sequence and moving it alongside the data vector. For each position of the filtering sequence with respect to the data vector, the scalar product of the two is calculated (with missing points in the filtering sequence replaced with zeros). For instance, if $N = 2$, the third row in Table 1 shows that $c_1' = x_1h_3 + x_2h_2 + x_3h_1 + x_4h_0$. Dyadic down-sampling is then performed to c_{2i}' to generate coefficients c_i. The detail coefficients d_i are obtained in a similar manner by using the high-pass filtering sequence.

Table 1. Convolution procedure for low-pass filtering showing results before and after dyadic down-sampling.

x_0	x_1	$\cdots$	x_{2N-1}	x_{2N}	$\cdots$	x_{J-1}	x_0	x_1	...	x_{2N-2}	Before	After
h_{2N-1}	h_{2N-2}	$\cdots$	h_0				$\cdots$				c_0'	
	h_{2N-1}	$\cdots$	h_1	h_0			$\cdots$				c_1'	c_0
		$\cdots$		$\vdots$				$\cdots$			$\vdots$	$\vdots$
						h_{2N-1}	h_{2N-2}	h_{2N-3}	$\cdots$		c_{J-2}'	
							h_{2N-1}	h_{2N-2}	$\cdots$	h_0	c_{J-1}'	$c_{J/2-1}$

If the approximation **c** and detail **d** coefficients are stacked in vector $\mathbf{t} = [\mathbf{c} \mid \mathbf{d}]$, the wavelet transform can be expressed in the matrix form with the transformation matrix given by:

$$\mathbf{V}=\left[\begin{array}{ccccc|ccccc} 0 & 0 & \cdots & h_{2N-4} & h_{2N-2} & 0 & 0 & \cdots & g_{2N-4} & g_{2N-2} \\ h_{2N-1} & 0 & \cdots & h_{2N-5} & h_{2N-3} & g_{2N-1} & 0 & \cdots & g_{2N-5} & g_{2N-3} \\ h_{2N-2} & 0 & \cdots & h_{2N-6} & h_{2N-4} & g_{2N-2} & 0 & \cdots & g_{2N-6} & g_{2N-4} \\ h_{2N-3} & h_{2N-1} & \cdots & h_{2N-7} & h_{2N-5} & g_{2N-3} & g_{2N-1} & \cdots & g_{2N-7} & g_{2N-5} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ h_0 & h_2 & \cdots & 0 & 0 & g_0 & g_2 & \cdots & 0 & 0 \\ 0 & h_1 & \cdots & 0 & 0 & 0 & g_1 & \cdots & 0 & 0 \\ 0 & h_0 & \cdots & 0 & 0 & 0 & g_0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & h_{2N-2} & 0 & 0 & 0 & \cdots & g_{2N-2} & 0 \\ 0 & 0 & \cdots & h_{2N-3} & h_{2N-1} & 0 & 0 & \cdots & g_{2N-3} & g_{2N-1} \end{array}\right] \quad (4)$$

A requirement for the transform to be orthogonal (i.e., $\mathbf{V}^{\mathrm{T}}\mathbf{V}=\mathbf{I}$) is that the sum of the squares of each column must be equal to one and the scalar product of different columns must be equal to zero [3]. Therefore, for a filter bank that utilizes low-pass and high-pass filters, the following conditions ensure orthogonality of the transform so that no information is lost in the decomposition process [4]:

$$\sum_{n=0}^{2N-1-2l} h_n h_{n+2l} = \begin{cases} 1, & l=0 \\ 0, & 0<l<N \end{cases} \quad (5a)$$

$$g_n = (-1)^{n+1} h_{2N-1-n},\ n=0,1,\ldots,2N-1 \quad (5b)$$

Under these conditions, the filter bank is said to enjoy a perfect reconstruction (PR) property, because **x** can be reconstructed from **t** which means that there is no loss of information in the decomposition process. Although other non-orthogonal filter bank transforms can also enjoy a PR property, provided that they are associated to a non-singular matrix **V**, the analysis in the present work is restricted to orthogonal transforms. In fact, the orthogonality of the transform (with the consequent PR property) ensures that no information that may be potentially useful for classification purposes is lost in the decomposition process. Moreover, convenient parameterisation schemes may then be employed to cast the transform filters into forms amenable to optimization.

Parametrization and optimization approach.

Let:

$$h_0^{(1)} = \cos(a_1), \qquad h_1^{(1)} = \sin(a_1)$$

$$h_0^{(N+1)} = \cos(a_{N+1})\, h_0^{(N)} \qquad h_1^{(N+1)} = \sin(a_{N+1})\, h_0^{(N)}$$

$$h_{2i}^{(N+1)} = \cos(a_{N+1})\, h_{2i}^{(N)} - \cos(a_{N+1})\, h_{2i-1}^{(N)}, \quad i=1,2\ldots N-1 \qquad h_{2i+1}^{(N+1)} = \sin(a_{N+1})\, h_{2i}^{(N)} + \cos(a_{N+1})\, h_{2i-1}^{(N)}, \quad i=1,2\ldots N-1$$

$$h_{2N}^{(N+1)} = -\sin(a_{N+1})\, h_{2N-1}^{(N)} \qquad h_{2N+1}^{(N+1)} = \cos(a_{N+1})\, h_{2N-1}^{(N)}$$

As stated in [5, 6], in order to ensure two vanishing moments for the resulting transform,

$$a_N = \frac{\pi}{4} - \sum_{i=1}^{N-1} a_i \quad (6)$$

$$a_{N-1} = \frac{1}{2}\arcsin\left\{-\frac{1}{2} - \sum_{k=1}^{N-2}\left[\sin\sum_{i=1}^{k} 2a_i\right]\right\} - \sum_{i=1}^{N-2} a_i \quad (7)$$

In what follows, an optimization process is proposed that maximizes the selectivity of the pair of high-pass/low-pass orthonormal wavelet filters with a given length.

As discussed in [5, 6] the expression in (7) has a real value solution if the set of angles a_i where $1 \le i \le N-2$ satisfy a set of constrains that define a non-convex region in R^{N-2}. Additional constrains are imposed to ensure this, by invoking a new parameter χ_i so that:

$$\chi_i = \sin \sum_{k=1}^{i} 2a_k, \quad 1 \le i \le N-2 \tag{8}$$

$$-\frac{3}{2} \le \sum_{i=1}^{N-2} \chi_i \le \frac{1}{2}, \quad -1 \le \chi_1, \chi_2 \cdots \chi_{N-2} \le 1 \tag{9}$$

Rearranging (8) and (9) we have:

$$a_1 = \frac{1}{2}\arcsin(\chi_1) \tag{10}$$

$$a_i = \frac{1}{2}\arcsin(\chi_i) - \sum_{k=1}^{i-1} 2a_k, \quad 2 \le i \le N-2 \tag{11}$$

The cost J is defined from the frequency response of the low pass filter as:

$$J = \frac{\int_{0.25\omega_s}^{0.5\omega_s} \left| H^{(N)} e^{j\omega T} \right| d\omega}{\int_{0}^{0.5\omega_s} \left| H^{(N)} e^{j\omega T} \right| d\omega} \tag{12}$$

with $\omega_s = 2\pi / T$ being the sampling frequency. As discussed in [6], optimization of the cost function with respect to the χ_i parameters is accomplished by using sequential quadratic programming (SQP) which uses local quadratic approximations of the cost function and local linear approximations of the restrictions. An interesting variance to the above algorithm where additional constrains are imposed to ensure a third vanishing moment can be found in [7].

Example of a signal decomposition process

The signal in Fig. 1 represents the first 1000 points out of a 3600 data points record from patient number 100 (lead 1) from the MIT database. The patient does not have a pathogenic condition and his record (among others from that database) is normally used as a training set to different classifiers to discriminate from other pathogenic patient records. A typical decomposition of the signal to approximation and detail coefficients at the first decomposition level is shown as an inset to that figure. Normally, a much smaller filter tap is generated by the user as shown in the Figure 1b. Figure 2 depicts the angular parameter alpha associated with a standard db6 filter bank as well as for the filter bank generated on the basis of the proposed procedure. Figure 2b shows the difference in the function of these filters in the frequency domain. The introduction of a vanishing moment in the random parameter filter shown in Fig. 2b ensures that its gain drops to zero at high frequencies.

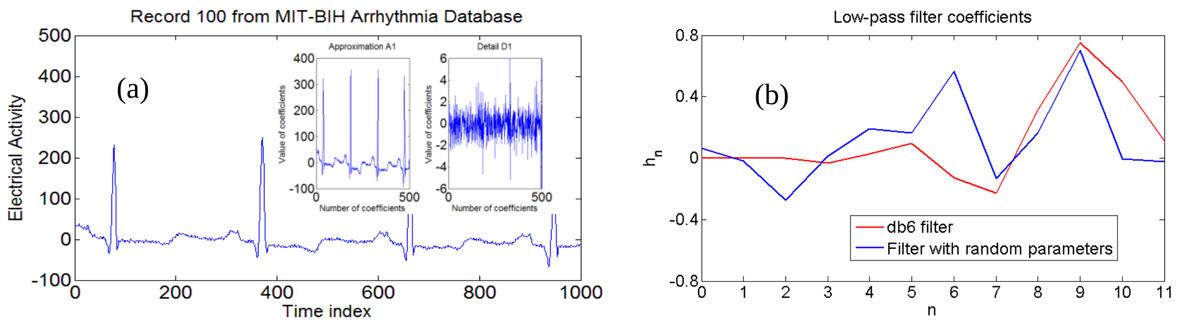

Figure 1. a) Typical signal from the MIT-BIH database with corresponding reconstruction on the basis of approximation and wavelet coefficients at the first decomposition level and b) comparison of filter coefficients impulse response function assuming 12 taps.

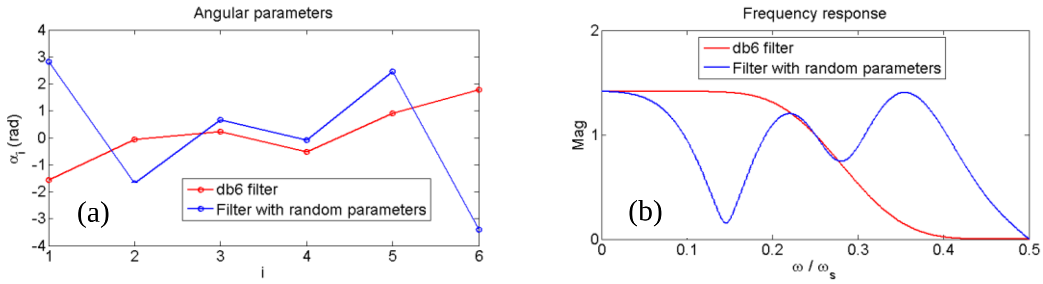

Figure 2. a) Comparison of angular parameters for a standard db6 filter bank with those of a filter bank generated using the proposed procedure and b) normalized frequency response for the two filter banks depicting the difference in their function.

Conclusion

The current work advances previous ECG wavelet decompositions by placing recent work on wavelet parametrizations [5, 7] within a biomedical applications context. A test waveform was decomposed to qualitatively describe the new parametrization process being implemented. This decomposition was contrasted to that of a standard db6 parametrization. We intend to further explore this approach as a feature reduction method before presenting ECG signals from the MIT-BIH database to different classifiers (e.g. a successive projections algorithm [8]). Further work intends to optimize the number of wavelet coefficients as well as number of decomposition levels presented to the classifier. Contrary to previous works in the ECG literature, this needs to be performed in a more systematic manner (e.g., using Kohonen maps), evaluating the performance of the classifier output when it is presented with normal beats, premature ventricular contraction beats, paced beats, left and right bundle branch block beats, atrial premature contraction beats, ventricular flutter wave and ventricular escape beats.

Acknowledgment

The authors acknowledge CNPq for the provision of post-doctoral and research fellowships and FAPESP for the provision of grants 2011/13777-8 and 2011/17610-0 that enabled this work to be performed. N Jannah acknowledges scholarship support from the Ministry of Higher Education, Kingdom of Saudi Arabia.

References

[1] Froese T, Hadjiloucas S, Galvão R K H, Becerra V M and Coelho C J 2006 Comparison of extrasystolic ECG signal classifiers using Discrete Wavelet Transforms. *Pattern Recognition Letters* **27**(5) 393-407.

[2] Sherlock B G, and Monro D M 1998 On the space of orthonormal wavelet. *IEEE Transactions on Signal Processing* **46** 1716-1720.

[3] Vaidyanathan P P 1993 *Multirate Systems and Filter Banks*. Prentice Hall, New Jersey.

[4] Strang G, and Nguyen T, 1996 *Wavelets and Filter Banks*. Wellesley-Cambridge Press Massachusetts.

[5] Paiva H M, Marins M N, Galvão R K H and Paiva J P L M 2009 On the space of orthonormal wavelets: additional constrains to ensure two vanishing moments. *IEEE Signal Processing Letters*, **16** 101-104.

[6] Paiva H M and Galvão R K H 2012 Optimized orthonormal wavelet filters with improved frequency separation. *Digital Signal Processing* **22** 622-627.

[7] Uzinski J C, Paiva H M, Alvarado F V, Duarte M A Q and Galvão R K H 2013 Additional Constraints to Ensure Three Vanishing Moments for Orthonormal Wavelet Filter Banks *Anais do Congresso de Matemática Aplicada e Computacional* CMAC Centro-Oeste 2013, 16-19.

[8] R. K. H. Galvão, M. C. U. Araújo, W. D. Fragoso, E. C. Silva, G. E. José, S. F. C. Soares, H. M. Paiva, 2008 A variable elimination method to improve the parsimony of MLR models using the successive projections algorithm *Chemom. Intell. Lab. Syst.*, **92**, 83-91.