Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Multiwavelet Design for Cardiac Signal Processing

R.L.M. Peeters†, J.M.H. Karel†, R.L. Westra†, S.A.P. Haddad‡ and W.A. Serdijn‡

Abstract— **An approach for designing multiwavelets is introduced, for use in cardiac signal processing. The parameterization of the class of multiwavelets is in terms of associated FIR polyphase all-pass filters. Orthogonality and a balanced vanishing moment of order 1 are built into the parameterization. An optimization criterion is developed to associate the wavelets with different meaningful segments of a signal. This approach is demonstrated on the simultaneous detection of QRS-complexes and T-peaks in ECG signals.**

I. INTRODUCTION

In cardiac signal processing the detection of the different segments in the ECG is an important activity. However the QRS complex is very dominant in this type of signal. As a result, other segments in the wavelet transform of ECG signals such as the T-wave have relatively little energy. In the presence of noise, this makes it difficult to obtain accurate information on the morphology of the T-wave, which however is known to carry important information on various cardiac pathologies.

In order to overcome this problem an approach is discussed that involves the construction of a number of mutually orthogonal wavelet functions, that are designed by joint optimization with respect to different segments of a prototype signal. An optimization approach for wavelets was previously discussed in [2], but here this approach is extended to the multiwavelet case [4], [5]. The construction of the multiwavelets is based on the theory of filter banks and addressed here in terms of polyphase filtering. Orthogonality of the multiresolution structure and compact support of the wavelets leads to polyphase filters that are FIR all-pass. To parameterize this class in a new convenient way, use is made of the results in [1]. This enables the direct incorporation of a balanced vanishing moment of order 1, which is needed to avoid bias for constant signals.

The potential of this approach is demonstrated with an application concerning the simultaneous detection of both the QRS-complex and the T-wave.

II. MULTIWAVELETS AND FILTER BANKS

We briefly discuss the theory of multiwavelets from the perspective of filter banks, largely based on the expositions in [4], [5]. We consider two real $r \times r$ FIR filters of order $2n-1$, denoted by

$$C(z) = C_0 + C_1 z^{-1} + \ldots + C_{2n-1} z^{-(2n-1)}, \quad (1)$$
$$D(z) = D_0 + D_1 z^{-1} + \ldots + D_{2n-1} z^{-(2n-1)}, \quad (2)$$

and two real $r \times 1$ vector functions with entries that are functions in $L^2(\mathbb{R})$, denoted by

$$\phi(t) = (\phi_1(t), \phi_2(t), \ldots, \phi_r(t))^T, \quad (3)$$
$$\psi(t) = (\psi_1(t), \psi_2(t), \ldots, \psi_r(t))^T, \quad (4)$$

to which the following identities are assumed to apply:

(i) $\phi(t)$ is a *multiscaling function*, satisfying a two-scale vector equation called the *dilation equation*:

$$\phi(t) = \sqrt{2} \sum_{k=0}^{2n-1} C_{2n-1-k}\phi(2t-k), \quad (5)$$

(ii) $\psi(t)$ is a *multiwavelet function*, defined by a two-scale vector equation called the *wavelet equation*:

$$\psi(t) = \sqrt{2} \sum_{k=0}^{2n-1} D_{2n-1-k}\phi(2t-k), \quad (6)$$

(iii) the FIR filters $C(z)$ and $D(z)$ are thought of as low-pass and high-pass filters, respectively, jointly satisfying the Smith-Barnwell *orthogonality conditions*:

$$C(z)C(z^{-1})^T + C(-z)C(-z^{-1})^T = 2I_r, \quad (7)$$
$$D(z)D(z^{-1})^T + D(-z)D(-z^{-1})^T = 2I_r, \quad (8)$$
$$C(z)D(z^{-1})^T + C(-z)D(-z^{-1})^T = 0. \quad (9)$$

Due to the FIR property of the related filters, the functions $\phi(t)$ and $\psi(t)$ both have compact support in the interval $[0, 2n-1]$. They generate a multiresolution structure for $L^2(\mathbb{R})$ consisting of a nested sequence of subspaces $\ldots \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \ldots$ for which $\cap_{k\in\mathbb{Z}} V_k = \{0\}$ and $\overline{\cup_{k\in\mathbb{Z}} V_k} = L^2(\mathbb{R})$. Moreover, it holds that for each $k \in \mathbb{Z}$, the set of functions

$$B_k = \{2^{k/2}\phi_j(2^k t - \ell) | j = 1, 2, \ldots, r; \ell \in \mathbb{Z}\} \quad (10)$$

defines an orthonormal basis for the space V_k, and the set of functions

$$\widetilde{B}_k = \{2^{k/2}\psi_j(2^k t - \ell) | j = 1, 2, \ldots, r; \ell \in \mathbb{Z}\} \quad (11)$$

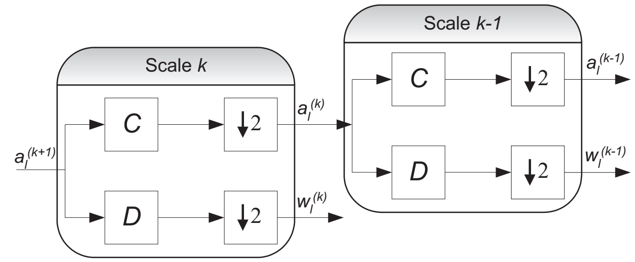
defines an orthonormal basis for the space W_k which is defined as the orthogonal complement of V_k within the space V_{k+1}.

†Department of Mathematics, Universiteit Maastricht. P.O. Box 616, 6200 MD Maastricht, The Netherlands. {ralf.peeters, joel.karel, westra}@math.unimaas.nl

‡Electronics Research Laboratory, Delft University of Technology. Mekelweg 4, 2628 CD Delft, The Netherlands. {s.haddad, w.a.serdijn}@ewi.tudelft.nl

This paper is part of the BioSens project that is funded by STW (Project no. DTC 6418).

Fig. 1. Tree structure for the DWT



Now suppose that $x(t) \in V_{k+1}$ is represented by its sequence of $r \times 1$ coordinate vectors $\{x_\ell\}$ in terms of the basis B_{k+1}, such that

$$x(t) = \sum_{\ell \in \mathbb{Z}} 2^{(k+1)/2} x_\ell^T \phi(2^{k+1} t - \ell). \tag{12}$$

Since $V_{k+1} = V_k \oplus W_k$ the function $x(t)$ can be decomposed into $x(t) = a(t) + w(t)$, with an approximation component $a(t) \in V_k$ having coordinate vectors $\{a_\ell\}$ with respect to the basis B_k, and a detail component $w(t) \in W_k$ having coordinate vectors $\{w_\ell\}$ with respect to the basis $\widetilde{B}_k$. It then holds that

$$a_\ell = \sum_{k=0}^{2n-1} C_{2n-1-k} x_{k+2\ell}, \qquad w_\ell = \sum_{k=0}^{2n-1} D_{2n-1-k} x_{k+2\ell}. \tag{13}$$

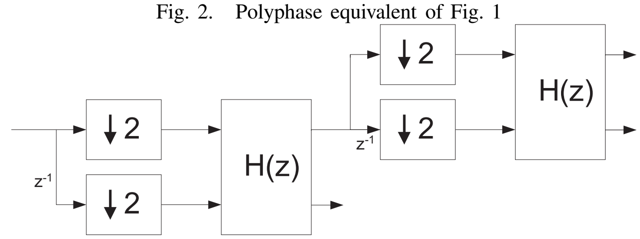
Consequently, the coordinate vector sequences $\{a_\ell\}$ and $\{w_\ell\}$ can be obtained directly from the sequence $\{x_\ell\}$ by means of the following three-step procedure:
(1) advancing $\{x_\ell\}$ by $2n-1$, yielding the sequence $\{x_{\ell+2n-1}\}$;
(2) filtering $\{x_{\ell+2n-1}\}$ by the filters $C(z)$ and $D(z)$, respectively, yielding the sequences $\{u_\ell\}$ and $\{v_\ell\}$;
(3) down-sampling the sequences $\{u_\ell\}$ and $\{v_\ell\}$ by two, yielding the sequences $\{a_\ell\} = \{u_{2\ell}\}$ and $\{w_\ell\} = \{v_{2\ell}\}$.
This procedure can be iterated by reapplying it to the sequence of approximation coordinate vectors $\{a_\ell\}$. It is illustrated in Fig. 1.

Note that an explicit computation of the functions $\phi(t)$ and $\psi(t)$ is not required to carry out this procedure, provided that an initial sequence $\{x_\ell\}$ is available. In practice, such a sequence is naturally obtained from a given measurement sequence by decomposing it into its r phases. The sequences $\{a_\ell\}$ and $\{w_\ell\}$ are about half as long as the sequence $\{x_\ell\}$ due to down-sampling. The (repeated) decomposition of a sequence $\{x_\ell\}$ into the associated sequences of wavelet coefficients for a number of subsequent levels is called the *discrete wavelet transform* (DWT). As for the FFT, there are various methods to deal with finite length issues when computing the DWT, but we do not go into this here.

III. POLYPHASE FILTERING AND PARAMETERIZATION ISSUES

Instead of first passing the vector sequence $\{x_\ell\}$ through the $r \times r$ filters $C(z)$ and $D(z)$ and then applying down-sampling by two, one may alternatively proceed by first splitting $\{x_\ell\}$ into its two phases (i.e., into its even-indexed and its odd-indexed part) and then applying a $2r \times 2r$ polyphase filter $H(z)$. This leads to the *same* sequences $\{a_\ell\}$ and $\{w_\ell\}$ when

Fig. 2. Polyphase equivalent of Fig. 1



$H(z)$ is defined as the following block-partitioned FIR filter of order $n-1$:

$$H(z) = \begin{pmatrix} C_{\text{even}}(z) & C_{\text{odd}}(z) \\ D_{\text{even}}(z) & D_{\text{odd}}(z) \end{pmatrix}, \tag{14}$$

where

$$C_{\text{even}}(z) = C_0 + C_2 z^{-1} + \ldots + C_{2n-2} z^{-(n-1)}, \tag{15}$$
$$C_{\text{odd}}(z) = C_1 + C_3 z^{-1} + \ldots + C_{2n-1} z^{-(n-1)}, \tag{16}$$
$$D_{\text{even}}(z) = D_0 + D_2 z^{-1} + \ldots + D_{2n-2} z^{-(n-1)}, \tag{17}$$
$$D_{\text{odd}}(z) = D_1 + D_3 z^{-1} + \ldots + D_{2n-1} z^{-(n-1)}, \tag{18}$$

so that

$$C(z) = C_{\text{even}}(z^2) + z^{-1} C_{\text{odd}}(z^2), \tag{19}$$
$$D(z) = D_{\text{even}}(z^2) + z^{-1} D_{\text{odd}}(z^2). \tag{20}$$

This alternative set-up is illustrated in Fig. 2. Here it holds that the orthonormality conditions translate into the single condition

$$H(z) H(z^{-1})^T = I_{2r}. \tag{21}$$

Equivalently, $H(z)$ is required to be a real $2r \times 2r$ FIR all-pass filter of order $n-1$.

For the purpose of multiwavelet design through optimization of a criterion function as discussed in Section V, it is useful to have a parameterization available for the set of FIR filters $C(z)$ and $D(z)$ of a fixed order $2n-1$. When the entries of the coefficient matrices C_k and D_k are all chosen as parameters, then the orthogonality conditions provide a set of nonlinear constraints that may be difficult to handle.

A convenient way to deal with the orthogonality constraints in an automated fashion, is by direct parameterization of the class of $2r \times 2r$ all-pass filters, a topic which is studied in much detail in [1] from two different perspectives: that of Schur interpolation theory and that of balanced state space realizations. As a special application, one may restrict that theory to the case of real FIR all-pass filters. This leads to a useful parameterization that is recursive in the order n in the following way.

Theorem 3.1: Let $H(z)$ be a $2r \times 2r$ real FIR all-pass filter of order $n \geq 1$. Then there exists a $2r \times 1$ vector u with norm $\|u\| = 1$, for which $H(z)$ can be factored as

$$H(z) = (I_{2r} + (z^{-1} - 1) u u^T) G(z), \tag{22}$$

where $G(z)$ is a $2r \times 2r$ real FIR all-pass filter of order $n-1$. Note that it is not difficult to parameterize a $2r \times 1$ vector u of norm 1 by means of $2r-1$ free real angular parameters.

Also, for order 0 the corresponding real all-pass function is a $2r \times 2r$ orthogonal matrix, which can easily be parameterized by $r(2r-1)$ free real angular parameters. Thus, the total number of free parameters for order $n-1$ is $(n+r-1)(2r-1)$.

We remark that this parameterization is different from those described in Chapters 4 and 9 of [5]. The present parameterization has the property that $H(1) = G(1)$, which makes it possible to build in a vanishing moment with respect to the multiwavelet $\psi(t)$ in a straightforward manner, as we now describe.

IV. Vanishing Moments

In the literature it has been argued by many authors that *vanishing moments* for $\psi(t)$ are highly desirable for many practical purposes. This means that $\psi(t)$ is designed to be orthogonal to the class of polynomials up to a certain degree. Consequently, polynomial functions pass through the low-pass filter $C(z)$ and are blocked by the high-pass filter $D(z)$. Vanishing moments thus correspond to smoothness properties of the reconstruction signals produced by the synthesis counterpart of the filter bank. To avoid bias in the multiwavelet analysis, one needs to incorporate a balanced vanishing moment of order 1, i.e.: it is required that $\int_{\mathbb{R}} \psi(t)dt = 0$. According to [4, Thm. 2], taking into account the differences in notation and terminology, this is characterized by the conditions:

$$(1, 1, \ldots, 1)C(1) = \sqrt{2}(1, 1, \ldots, 1), \qquad (23)$$

$$(1, 1, \ldots, 1)C(-1) = (0, 0, \ldots, 0). \qquad (24)$$

In terms of the polyphase all-pass filter $H(z)$ we have the following result.

Theorem 4.1: If $H(z)$ is a $2r \times 2r$ real FIR all-pass filter of order $n-1$, associated with the corresponding FIR filters $C(z)$, $D(z)$ and vector functions $\phi(t)$ and $\psi(t)$ as discussed above, then $\psi(t)$ has a balanced vanishing moment of order 1 if and only if

$$(1, 1, \ldots, 1|1, 1, \ldots, 1)H(1)^T = \sqrt{2}(1, 1, \ldots, 1|0, 0, \ldots, 0). \qquad (25)$$

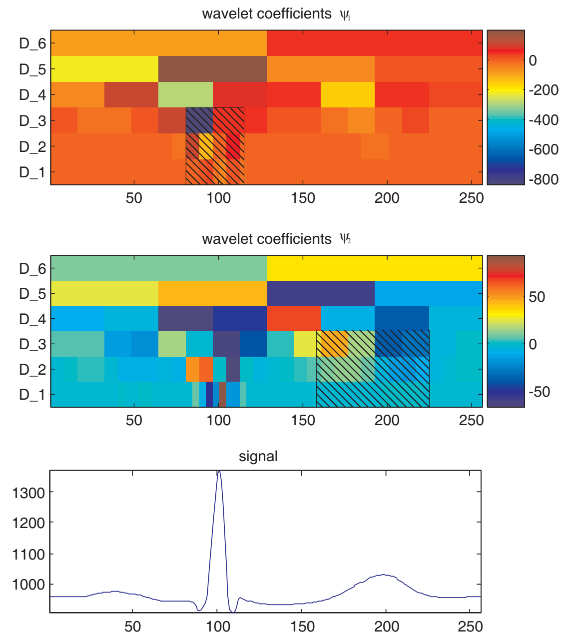
Under this condition it also holds that $\int_{\mathbb{R}} \phi_j(t)dt = 1$ for each $j = 1, 2, \ldots, r$.

From this result one can build a vanishing moment into the parameterized class of polyphase filters $H(z)$, by noting that $H(1)$ is identical to the $2r \times 2r$ orthogonal matrix that serves as the all-pass filter of order 0 in the recursive construction. Therefore $H(1)$ can be factored as

$$H(1) = P \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix} R \qquad (26)$$

where $P = I_{2r} - 2\frac{\alpha\alpha^T}{\alpha^T\alpha}$ with $\alpha^T = (1, 1, \ldots, 1|0, 0, \ldots, 0)/\sqrt{r} - (1, 0, \ldots, 0|0, 0, \ldots, 0)$ and $R = I_{2r} - 2\frac{\beta\beta^T}{\beta^T\beta}$ with $\beta^T = (1, 1, \ldots, 1|1, 1, \ldots, 1)/\sqrt{2r} - (1, 0, \ldots, 0|0, 0, \ldots, 0)$ are two fixed orthogonal Householder matrices, and Q is an arbitrary $(2r-1) \times (2r-1)$ orthogonal matrix. Such a matrix Q can be parameterized by $(r-1)(2r-1)$ free real angular parameters in a straightforward way. Then the total number of free parameters for the class of FIR all-pass filters of order $n-1$ with a balanced vanishing moment of order 1 is $(n+r-2)(2r-1)$.

Fig. 3. Windowing in time and scale



V. Design Criterion

For various applications it is opportunistic for the wavelet transform of a given signal $x(t)$ to have a high variance. This property relates to a sparse representation in the wavelet domain. For compression purposes, it is desirable that a signal can be approximated well with only a few non-zero wavelet coefficients. For detection purposes, a sparse representation is likely to involve a few dominant characteristic points in the wavelet representation of $x(t)$. When using an orthonormal multiresolution structure, the wavelet transform is energy preserving. Maximization of the variance of the absolute values of the wavelet coefficients has been shown in [2] to correspond to minimization of the L_1-norm. Likewise, maximization of the variance of the energies in a wavelet decomposition corresponds to maximization of the L_4-norm. The latter criterion is more suitable for use in practical applications, since it can more easily be combined with weighting of selected scales and locations. For multiwavelet optimization this criterion may require adjustment. Suppose that $r = 2$ and the two wavelets are intended to correspond to characteristic features of different segments of the signal $x(t)$. Two windows on the time-axes can then be employed to optimize the multiwavelet. As an example for a signal of length 256, the first wavelet $\psi_1(t)$ may for instance be optimized on wavelet coefficients that relate to $t \in \{85, \ldots, 115\}$, and the second wavelet $\psi_2(t)$ may be optimized on wavelet coefficients that relate to $t \in \{160, \ldots, 225\}$. Similarly, there may also be a window defined on the scales for each wavelet in order to force a high variance into certain scales. The process of windowing is illustrated in Fig. 3, where the top illustration is the wavelet decomposition of a given prototype signal for the first wavelet $\psi_1(t)$ of the multiwavelet and the bottom figure is the decomposition for the second wavelet $\psi_2(t)$. The marked areas indicate on which wavelet coefficients the criterion is evaluated.

Since not all the wavelet coefficients are taken into account when evaluating the criterion, a simple minimization of the L_1-norm is inappropriate since it will push most of the energy into the areas outside the window of interest. Maximization of the L_4-norm does not suffer from this defect. It is therefore proposed for use in the multiwavelet design case.

The criterion values for the r decompositions corresponding to the r wavelets are all evaluated individually and then combined into a single criterion value by for example taking their sum. It may happen that one wavelet is optimized for a high-energy area in the signal such as the QRS-complex, whereas the other wavelet is optimized for a low-energy area in the signal such as the T-wave. Then the latter criterion value is less dominant than the former and the second wavelet may still exhibit relatively large coefficients in the high-energy area. There are two ways around this problem: 1) using the minimum of the r criteria as global criterion and 2) using a weighted sum of the r criteria. Both approaches can be useful, however the latter one is the most flexible.

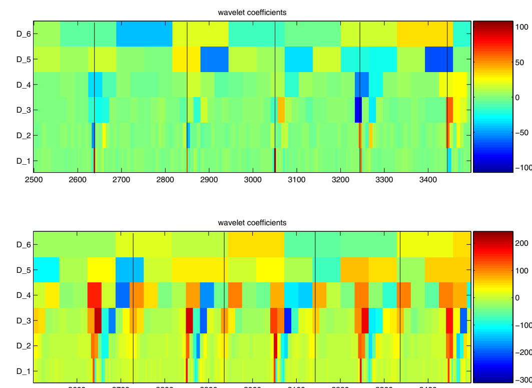
VI. Experiments

In order to assess the potential of the multiwavelet design technique, an initial biwavelet was designed for the prototype signal in Fig. 3 with the goal to differentiate between the QRS-complex and the T-peak simultaneously. As a motivation for this investigation, note that if a T-peak can be detected without reference to neighboring QRS-complexes, this may be very useful for real-time analog cardiac signal processing in devices with limited memory storage capacity.

To achieve this design goal, two windows were defined around the locations of the indicated segments and only the three finest scales were taken into consideration. The first wavelet $\psi_1(t)$ was fitted to the QRS-complex and the second wavelet $\psi_2(t)$ to the T-peak. In both the wavelet decompositions of the prototype signal the QRS-complex was clearly present. However only for $\psi_1(t)$ its energy was concentrated in a small number of wavelet coefficients. In the decomposition for $\psi_1(t)$ the coefficients associated with the T-peak were small compared to those for the QRS-complex, but in the decomposition for $\psi_2(t)$ they were more dominant. Lack of dominance of the T-peak in the first decomposition is helpful to distinguish it from the QRS-complex.

As a second preliminary test for detecting the T-peak a new prototype signal was constructed by averaging a number of beats from an available 250Hz dataset, all centered around the R-peak. For this prototype signal a window around the QRS-complex and the T-peak was defined and a new multiwavelet with $r = 2$ and $n = 2$ was constructed. In Fig. 4 the decomposition of a portion of the test signal is displayed. In a single scale $\sigma = 2$ of the wavelet decomposition for $\psi_1(t)$, the R-peak was detected by thresholding the absolute value. Around this peak a window was then defined where no T-peak is conceived to occur. Next, T-peaks were detected by thresholding with respect to the wavelet decomposition

Fig. 4. Simultaneous QRS and T-peak detection



for $\psi_2(t)$, by requiring that the absolute value of the wavelet coefficients should be above a certain threshold, while in addition the absolute value of the wavelet coefficients at the same locations for $\psi_1(t)$ should not exceed another certain threshold value. This simple detection algorithm was evaluated on a 250Hz dataset of length $2 \cdot 10^{14}$ from the QT-database [3]. The T-peak was detected with a tolerance of 64ms. Preliminary results show that 72 out of 79 T-peaks were correctly detected and that there were 8 false positives. An more advanced detection algorithm is now under development and it is expected to give a substantial increase in performance.

VII. Conclusions

A parameterization of the class of multiwavelets in terms of associated FIR polyphase all-pass filters was introduced. Orthogonality is ensured by the all-pass property and a balanced vanishing moment of order 1 is built into the parameterization in order to avoid a bias. The parameterization allows for criterion optimization over the class of orthonormal multiwavelet structures. For the purpose of cardiac signal processing, a criterion design procedure is developed to associate the wavelets with different meaningful segments of a signal. The potential of this approach has been demonstrated by an application involving the detecting of T-peaks in ECGs. Although the results are preliminary and the detection algorithm is very simple, it already shows a promising performance.

References

[1] B. Hanzon, M. Olivi, and R.L.M. Peeters. Balanced realizations of discrete-time stable all-pass systems and the tangential Schur algorithm. *Linear Algebra and its Applications*, 2006. In press; also available as Technical Report RR-5111, INRIA Sophia-Antipolis, France.

[2] J.M.H. Karel, R.L.M. Peeters, R.L. Westra, K.M.S. Moermans, S.A.P. Haddad, and W.A. Serdijn. Optimal discrete wavelet design for cardiac signal processing. In *EMBC 2005, Shanghai, China*. IEEE, 2005.

[3] P. Laguna, R.G. Mark, A. Goldberger, and G.B. Moody. A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG. *Computers in Cardiology*, 24:673–676, 1997.

[4] J. Lebrun and M. Vetterli. High-order balanced multiwavelets: theory, factorization, and design. *IEEE Trans. Signal Processing*, 49(9):1918–1930, 2001.

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