Optimal discrete wavelet design for cardiac signal processing

Citation for published version (APA):

Document status and date:
Published: 01/01/2005

DOI:
10.1109/IEMBS.2005.1617046

Document Version:
Publisher's PDF, also known as Version of record

Document license:
Taverne

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.umlib.nl/taverne-license

Take down policy
If you believe that this document breaches copyright please contact us at:
repository@maastrichtuniversity.nl
providing details and we will investigate your claim.

Download date: 21 Oct. 2023
Optimal discrete wavelet design for cardiac signal processing

J.M.H. Karel, R.L.M. Peeters and R.L. Westra

Abstract—The question of designing the best wavelet for a given signal is discussed from the perspective of orthogonal filter banks. Two performance criteria are proposed to measure the quality of a wavelet, based on the principle of maximization of variance. The method is illustrated and evaluated by means of a worked example from biomedicine in the area of cardiac signal processing. The experimental results show the potential of the approach.

I. INTRODUCTION

In recent years wavelets have been successfully used in a large number of biomedical applications. The multi-resolution framework makes wavelets into a very powerful compression [5] and filter tool [7], and the time and frequency localization of wavelets makes it into a powerful tool for feature detection [2]. For more advanced applications, however, the success of these techniques depend to a considerable extent on the choice of a good wavelet.

This paper addresses the important question of how to design a good wavelet for an application at hand. Various authors have made suggestions about how to approach this question, see [6], [3], [4]. However, hardly any worked applications are available in the literature. It is the purpose of this paper to describe an approach to this problem based on the theory of orthogonal filter banks and to investigate the performance of this approach for the purpose of cardiac signal processing.

II. ORTHOGONAL WAVELET DESIGN FROM FILTER BANKS

To deal with the question of designing a good wavelet in a systematic way, we employ the theory of wavelets derived from filter banks, see [6]. This provides an elegant and well understood framework for discrete-time (digital) signal processing to which we shall be confined in this paper. There are four important aspects of filter banks to take into account (in decreasing order of importance): (i) perfect reconstruction, (ii) orthogonality of the filter bank and the underlying wavelet based multi-resolution structure, (iii) flatness of the filters and vanishing moments in the wavelets, (iv) smoothness of the wavelets. Once a filter bank has been selected, which corresponds to an underlying scaling function \( \phi(t) \) and an associated mother wavelet \( \psi(t) \), the question needs to be addressed how well it is suited to perform its intended signal processing task. This requires a criterion function, which allows one to measure the performance and to compare the different filter banks.

A. Orthogonal wavelets and filter banks

The idea of filter bank theory is to perform signal processing by means of a bank of (digital) filters in combination with down-sampling. This process is illustrated in Fig. 1. At each stage of the filtering process, the given signal \( x \) is passed through a low pass filter with transfer function \( C(z) \) and a high pass filter with transfer function \( D(z) \) after which down-sampling takes place. These two filters must meet certain requirements to enable perfect reconstruction from the two output signals after down-sampling and to yield an orthogonal underlying wavelet basis. To end up with a corresponding mother wavelet \( \psi(t) \) having compact support, \( C(z) \) and \( D(z) \) must be finite impulse response (FIR) filters:

\[
C(z) = c_0 + c_1 z^{-1} + \ldots + c_N z^{-N},
\]

\[
D(z) = d_0 + d_1 z^{-1} + \ldots + d_N z^{-N}.
\]

Here \( N \) is an integer determining the order of the filters.

Orthogonality of the underlying wavelet basis involves the following conditions on the filter coefficients: (a) normalization: \( \sum_{k=0}^{N} c_k^2 = 1 \) and \( \sum_{k=0}^{N} d_k^2 = 1 \); (b) double shift orthogonality: \( \sum_{k=0}^{N} c_k c_{k-2\ell} = 0 \) and \( \sum_{k=0}^{N} d_k d_{k-2\ell} = 0 \) for all integers \( \ell \neq 0 \); and (c) double shift orthogonality between the two filters: \( \sum_{k=0}^{N} c_k c_{k-2\ell} = 0 \) for all integers \( \ell \). (Here, negative indexed coefficients are all zero by convention.)

Double shift orthogonality implies that \( N \) is odd, say \( N = 2n-1 \). The alternating flip construction allows one to specify the high pass filter coefficients \( d_0, d_1, \ldots, d_{2n-1} \) in terms of the low pass filter coefficients \( c_0, c_1, \ldots, c_{2n-1} \) in such a way that condition (c) is automatically fulfilled:

\[
d_k = (-1)^k c_{2n-1-k} \quad (k = 0, 1, \ldots, 2n - 1).
\]

It is described in [6, Ch. 4-5] how the remaining orthogonality constraints can be handled by reparameterization of the \( 2n \) low pass filter coefficients \( c_0, c_1, \ldots, c_{2n-1} \) in terms of \( n \) new parameters \( \theta_1, \theta_2, \ldots, \theta_n \). The theory of polyphase matrices offers a convenient way to achieve this.
For $k = 1, \ldots, n$, let $R(\theta_k) = \begin{bmatrix} \cos(\theta_k) & -\sin(\theta_k) \\ \sin(\theta_k) & \cos(\theta_k) \end{bmatrix}$ and let $\Lambda(z) = \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix}$. Then consider the $2 \times 2$ matrix product

$$H(z) = \Lambda(-1)R(\theta_n)\Lambda(z)R(\theta_{n-1})\Lambda(z) \cdots R(\theta_2)\Lambda(z)R(\theta_1).$$

Let $H(z)$ be partitioned as $H(z) = \begin{bmatrix} C_0(z) & C_1(z) \\ D_0(z) & D_1(z) \end{bmatrix}$ where

$$C_0(z) = c_0 + c_2z^{-1} + c_4z^{-2} + \ldots + c_{2n-2}z^{-(n-1)}$$
$$C_1(z) = c_1 + c_3z^{-1} + c_5z^{-2} + \ldots + c_{2n-1}z^{-(n-1)}$$
$$D_0(z) = d_0 + d_2z^{-1} + d_4z^{-2} + \ldots + d_{2n-2}z^{-(n-1)}$$
$$D_1(z) = d_1 + d_3z^{-1} + d_5z^{-2} + \ldots + d_{2n-1}z^{-(n-1)}$$

from which $C(z)$ and $D(z)$ are formed as $C(z) = C_0(z^2) + z^{-1}C_1(z^2)$ and $D(z) = D_0(z^2) + z^{-1}D_1(z^2)$.

It then holds that $C(z)$ and $D(z)$ satisfy all the orthogonality constraints mentioned above, as well as the relationship described by the alternating flip construction. Conversely, for all such $C(z)$ and $D(z)$ there exist parameters $\theta_1, \theta_2, \ldots, \theta_n$ which achieve this decomposition of the polyphase matrix $H(z)$.

**B. Vanishing moments**

Apart from orthogonality, an important desirable property for wavelets is to have vanishing moments. This requires $C(z)$ to have zeros at $z = -1$, thus exhibiting a corresponding degree of flatness at the low and high frequencies (and likewise for the high-pass filter $D(z)$).

To have one vanishing moment amounts to the condition $c_0 - c_1 + c_2 - c_3 + \ldots + c_{2n-2} - c_{2n-1} = 0$ which is equivalent to the commonly imposed condition that the integral of the mother wavelet $\psi(t)$ is equal to zero: $\int_{-\infty}^{\infty} \psi(t)dt = 0$. In terms of $\theta_1, \theta_2, \ldots, \theta_n$, this is well known to translate into the condition $\theta_1 + \theta_2 + \ldots + \theta_n = \pi/4 \mod 2\pi$. This allows one to build a vanishing moment into the filter bank by simply eliminating one of the free variables.

The conditions corresponding to more vanishing moments are not difficult to obtain in terms of the filter coefficients $c_0, c_1, \ldots, c_{2n-1}$ by requiring also some derivatives of $C(z)$ to have a zero at $z = -1$. They amount to a set of linear conditions in terms of these filter coefficients, but in terms of the parameters $\theta_1, \theta_2, \ldots, \theta_n$ the expressions attain a more difficult form. We state here that the additional condition in case of a second vanishing moment is given by: $\cos(2\theta_n) + \cos(2\theta_{n-1} + 2\theta_n) + \ldots + \cos(2\theta_2 + \ldots + 2\theta_{n-1} + 2\theta_n) + \frac{1}{2} = 0$.

**C. Computing $\phi(t)$ and $\psi(t)$**

The scaling function $\phi(t)$ is obtained as the solution of the *dilation equation* and $\psi(t)$ then follows from the wavelet equation:

$$\phi(t) = \sqrt{2} \sum_{k=0}^{N} c_k \phi(2t - k), \quad (5)$$
$$\psi(t) = \sqrt{2} \sum_{k=0}^{N} (-1)^k c_{N-k} \phi(2t - k), \quad (6)$$

An iteration scheme which allows for the exact computation of $\phi(t)$ at all the dyadic points up to an arbitrary resolution can be found, for instance, in [6]. We have adopted this method in this paper. It is important to note that it may well happen that these functions exhibit a discontinuous and fractal structure.

**D. Two criteria for wavelet design**

For denoising and compression and also for various detection purposes, it is attractive to measure the quality of a filter bank and its underlying wavelet in terms of the way the given signal is represented by the wavelet at various positions and on various scales. For the orthogonal wavelet bases discussed in this paper, this is conveniently described by the wavelet coefficients (and the approximation coefficients) obtained at the various scales by applying the filter bank.

At the finest level, the signal is represented by the time series $x = (x_0, x_1, x_2, \ldots)$ having the total energy (or variance) $E = \sum_k \sigma_k^2$. At the first level, after application of the filters $C(z)$ and $D(z)$ and down-sampling, two sequences of coefficients are obtained: the approximation coefficients $a^{(1)} = (a_0^{(1)}, a_1^{(1)}, a_2^{(1)}, \ldots)$ and the detail coefficients $d^{(1)} = (d_0^{(1)}, d_1^{(1)}, a_2^{(1)}, \ldots)$. Because of orthogonality of the filter bank the total energy is preserved: $E = \sqrt{\sum_k (d_k^{(1)})^2} + \sum_k (d_k^{(1)})^2$. Then at the second level, the filters $C(z)$ and $D(z)$ are applied to the approximation sequence $a^{(1)}$ to produce the next approximation sequence $a^{(2)}$ and the next detail sequence $d^{(2)}$. Again, the total energy remains preserved, and this process continues for as many levels as desired. See Fig. 2.

Since the wavelet coefficients $a_k^{(s)}$ (together with the approximation coefficients at the coarsest level) represent the contribution to the representation of the signal $x$ in terms of the multi-resolution wavelet basis, it is natural to strive for a filter bank which represents the signal $x$ by means of just a few large wavelet coefficients containing most of the energy and many small wavelet coefficients containing only little energy. The $L_2$-norm is not able to achieve this because of preservation of energy. A guiding principle proposed here is to aim for maximization of the variance, either maximization of the variance of the absolute values of the wavelet coefficients, or maximization of the variance of the squared wavelet coefficients i.e. of the energy distribution over the wavelet contributions at the various scales.
The variance of the sequence of absolute values $|w_k|$ is given by $\sum_k |w_k|^2 - (\sum_k |w_k|)^2 = E - (\sum_k |w_k|)^2$ in which $E$ is constant.

**Proof.** The variance of the sequence of absolute values $|w_k|$ is given by $\sum_k |w_k|^2 - (\sum_k |w_k|)^2 = E - (\sum_k |w_k|)^2$ in which $E$ is constant. Hence maximization of this quantity is equivalent to minimization of $\sum_k |w_k|^2$.

(b) The variance of the sequence of energies $|w_k|^2$ is given by $\sum_k |w_k|^4 - (\sum_k |w_k|^2)^2 = \sum_k |w_k|^4 - E^2$ in which $E$ is constant. Hence maximization of this quantity is equivalent to maximization of the $L_4$-norm $\left(\sum_k |w_k|^4\right)^{1/4}$. $\square$

The criteria $V_1$ and $V_4$ have both been investigated in the worked examples to design wavelets for various given signals. When all the wavelet coefficients are taken into account and no weighting is applied, both of these criteria tend to produce similar results. However, when only a few scales are taken into account (e.g. by weighting) the results may become quite different: in case of minimization of the $L_1$-norm energy tends to be placed in scales not taken into account, whereas in case of maximization of the $L_4$-norm this does not happen.

### III. EXPERIMENTATION

#### A. Theoretical test

For the first test, as a proof of principle, we have designed a test signal that is sparse in the wavelet domain by setting only a few wavelet coefficients to a non-zero value. A wavelet was chosen by selecting some values $\{\theta_1, \ldots, \theta_3\}$ and the corresponding filter coefficients were used to construct the test signal in the time domain. This signal was then used for wavelet design. Since the signal has a sparse wavelet representation it is likely to have a small $L_1$-norm and a large $L_4$-norm. It was then experimentally confirmed that values close to the original $\{\theta_1, \ldots, \theta_3\}$ are reobtained when using these criteria. Next, additive white noise was added to the signal with a signal-to-noise ratio (SNR) equal to 1. Again the original $\{\theta_1, \ldots, \theta_3\}$ was well recovered.

#### B. Reference signal

As a practical example from biomedical science, a reference signal was created by averaging heartbeats from an episode with ECG signals from Physionet’s MIT normal sinus rhythm database [1]. This produced a smooth signal that was upsampled to yield the signal displayed in Fig. 3. This signal is used as a typical ECG signal such that the wavelet will be able to capture the common essence of this type of signal. It was used in all the examples below.

#### C. Local optima

When optimizing a wavelet for the reference signal with respect to a certain norm, the optimization may terminate in a local optimum. In order to gain insight in the existence of these local optima, the situation with three $\theta$’s ($\theta = 3$), and thus with two degrees of freedom since $\theta_1 + \theta_2 + \theta_3 = \pi/4$, was investigated for $L_1$-minimization. $\theta_2$ and $\theta_3$ are set on a megapixel grid in the range $(-\pi, \pi]$, as shown in Fig. 4. Some of the local optima have been marked in this figure. When considering some of the filter coefficients corresponding to the local optima, the resemblance to the Daubechies 2 filter coefficients is evident. When the sum of $\theta_2$ and $\theta_3$ already is close to $\pi/2$, only one degree of freedom is effectively used. This observation gives a rationale for the use of the Daubechies 2 wavelet for cardiac signals. Considering this observation it will not come as a surprise that some of the other filter coefficients resemble the Daubechies 3 wavelet, however with less success.

#### D. Choosing the number of $\theta$’s

The filter size of the wavelet filter is determined by the number $n$ of $\theta$’s used. A larger number of $\theta$’s gives more freedom to fit the wavelet to the signal but also increases the complexity. To determine the effect of the choice of the number of $\theta$’s on the criterium value that can be achieved
for $n = 1, \ldots, 25$, each time 1000 random starting points were generated and the best criterium value using the $L_1$-norm was stored. See Fig. 5. The results indicate that $n = 8$ is a reasonable choice to work with, which was used in the further experiments below.

E. Practical evaluation

Since the practical application in mind is the use of wavelets in cardiac signal processing, a simple test case was designed to investigate the potential of this approach. As a test set episode 103 of the MIT-BIH arrhythmia database [1] was used. This is a 360 Hz annotated, freely available ECG signal. Two wavelets were designed using 8 $\theta$’s: one by minimizing the $L_1$-norm of the wavelet transform of the reference signal, and another one by maximizing its $L_4$-norm. The two resulting wavelet functions are illustrated in Fig. 6. The wavelet function that was obtained with $L_4$-maximization has a fractal structure. This is may be a consequence of the fact that the available degrees of freedom are not used to place many poles at $z = -1$ as is the case with the Daubechies wavelets, which caused the iteration process of finding a wavelet function from the filter coefficients to fail. Despite its fractal nature, the effectiveness of the $L_4$ optimized wavelet is high, as is illustrated by the wavelet decomposition of the reference signal in Fig. 7.

As an experimental setup the wavelet decomposition of the testset with the three wavelets ($L_1$, $L_4$ and Daubechies) was calculated, but only a single level (scale) was considered. This level was selected for each wavelet individually to maximize performance. Next a binary vector was constructed of all the wavelet coefficients of which the absolute value exceeded a certain threshold. These locations were then related to locations of the original signal.

The beat annotations in the testset were used as a reference to locate the QRS-complex. There are 1688 normal QRS-complexes in the testset. If the binary vector corresponding to the wavelet transform has detected a peak within 20 samples (56ms) of the marker, it is assumed that the QRS-complex is detected. If a peak is detected but no marker is within 20 samples, a false detection is assumed. The results of this experiment are illustrated in Table I. This table shows that the performance of the Daubechies 2 wavelet is quite good. The $L_4$ optimized wavelet however shows superior performance. Furthermore the $L_4$-wavelet is more robust with respect to the choice of threshold value, which may be a large advantage in practical applications.

REFERENCES