Biol. Cybern. 80, 259–267 (1999)

Biological Cybernetics

Nonlinear wavelet and wavelet packet denoising of electrocardiogram signal

P.E. Tikkanen

Department of Physical Sciences, Division of Biophysics, University of Oulu, FIN-90570 Oulu, Finland

Received: 27 April 1998 / Accepted in revised form: 24 November 1998

Abstract. The performance of different wavelet- and wavelet packet-based methods for removing simulated noise was studied using an electrocardiogram (ECG) signal. A non-linear denoising approach was investigated by applying soft and hard thresholding methods, in which thresholds were chosen using four different methods. Coiflet wavelet and wavelet packet functions were used to build up the dyadic wavelet and optimized wavelet packet decompositions. This study involves the quantitative comparison of different denoising approaches by means of optimized error measures and visual inspection of the denoised ECG and the error signal. The localization of the denoising error within the cardiac cycle was studied by visual inspection of the denoised signal and extracting the error measures during the QRS-complex. The results showed that wavelet denoising approaches were generally more efficient than wavelet packet approaches in all cases, but with HEURISTIC SURE threshold selection rule as hard thresholding for white noises was used. Denoising errors tend to concentrate within the QRS-area when the wavelet approach was employed. Moreover, soft and hard non-linearities showed different balances in denoising the high-frequency parts of an ECG.

1 Introduction

The wavelet transform (WT) is a recently introduced time-scale representation which has found applications in a variety of fields in biomedical signal processing. The value of WT as a signal analysis tool has been demonstrated by its application e.g. to image compression (Mallat 1989) and the study of evoked potentials (Thakor et al. 1993b). At the moment, the interest in using WT for the processing of electrocardiogram (ECG) signal is increasing. WT has been used in the detection of ventricular late potentials (VLP) (Dickhaus et al. 1994; Meste et al. 1994), ECG analysis during angioplasty (Gramatikov et al. 1995) and arrhythmia analysis (Senhadji et al. 1996). The approach can also be applied to automatic waveform detection (Li et al. 1995) and signal compression (Bradie 1996; Ramakrishnan and Saha 1997; Thakor et al. 1993a).

One application of WT is in removing noise from signals. The noise filtering scheme, generally called denoising, has been studied in both simulated (Bruce and Gao 1996; Donoho and Johnstone 1994; Hilton and Ogden 1997) and real signals (Hilton and Ogden 1997) involving the selection of the thresholding rule and choosing the thresholding non-linearity. In the wavelet packet-based approach, the signal composition structure is optimized, for example, by an entropy-based rule, which means more adaptivity of the decomposition scheme to the signal characteristics as compared with the pure wavelet-based approach. The selection of wavelet or wavelet packet function should also be considered.

In biomedical signal processing wavelet denoising has been applied in tomographic (Kolaczyk 1996) and functional magnetic resonance imaging (MRI) (Hilton et al. 1996) data. An example of developing a multirate adaptive filtering scheme based on wavelet packets can be found in Karrakchou and Kunt (1996), where it has been used in cancelling respiratory interference from the pulmonary capillary pressure signal. In many data acquisition settings, ECG may contain noise of a technical and/or a physiological origin. The ambulatory ECG, for example, is corrupted by powerline interference, electrosurgical and instrumentation noise, as well as electromyogram and motion artifact (Pahlm and Sörnmo 1984; Thakor and Zhu 1991). ECG noise reduction has been investigated by means of adaptive filtering (Thakor and Zhu 1991; Xue et al. 1992; Hamilton 1996). Various solutions for cancelling the power line interference have also been proposed (Pei and Tseng 1995; Hamilton 1996). Time-varying filters have been used to correct the baseline wander in ECG (Sörnmo 1993). A Wiener filtering approach has been proposed and applied to high-resolution ECG (Lander 1997).

Waveform detection algorithms naturally need as noise-free an ECG as possible to reduce the effect of

Correspondence to: P.E. Tikkanen
(e-mail: Pauli.Tikkanen@oulu.fi)

artifacts in such analyses as RR interval time series (Tikkanen et al. 1999). However, the original ECG should be distorted as little as possible by the applied filtering scheme. For these reasons, ECG noise removing is an interesting area, and moreover, the application of wavelet- and wavelet packet-based approaches has not yet been largely studied. As the selection of the thresholding rule and the thresholding non-linearity depend on the signal properties, they also offer the possibility to adapt the denoising scheme to the prevailing conditions. Furthermore, it is interesting to investigate the performances of the different denoising procedures, as the ECG signal may be considered a prototype for selecting the proper noise removal approach based on wavelet and wavelet packet transforms.

If noise contributes the same frequency bands as the original signal, conventional filtering approaches run into serious difficulty. Consequently, it is beneficial to use wavelet-based methods, where the signal is decomposed into a number of frequency bands (scales), the transform coefficients are interpreted and processed scale by scale, and finally the inverse transform is performed. In addition, the component of interest may fall into the gap regions of the filters and may then be cancelled or enhanched. Therefore, the wavelet packet approach with optimal base selection establishing a decomposition adapted to the signal might be beneficial when dealing with non-white noise. Further, subband adaptive filtering may suffer from non-ideal features of the filters, e.g. an imperfect rejection of stopband, which cause an aliasing phenomena of components on the stopband at the downsampling stage (Karrakchou and Kunt 1996). These points make wavelet denoising a promising choice to bring out better noise cancelling schemes.

The end-point is to evaluate the novel wavelet- and wavelet packet-based denoising methods to remove simulated noise from an ECG signal. The noise realizations include normally and uniformly distributed white noises and non-white noise created using an autoregressive signal model. The study also involves the quantitative comparison of different denoising approaches by means of optimized error measures, and visual inspection of the denoised ECG and the error signal between the original and filtered signal. The localization of the denoising error within the cardiac cycle is studied by visual inspection of the denoised signal and extracting the error measures in the QRS-complex. The QRS-complex refers to successive waveforms seen in an ECG including Q-, R-, and S-waves. The duration of the QRS-complex is defined as from the onset of the Q-wave to the offset of the S-wave. At this stage, the goal is to focus on measuring the denoising performance within the high-frequency part of the ECG.

2 Methods

2.1 Continuous wavelet transform

As a starting point, the continuous wavelet transform (CWT) is introduced and then extended to the theory of discrete (multiresolution) wavelet and wavelet packet transform, which are used to decompose the signal. CWT is defined for a signal $x(t)$ by Daubechies (1992)

$$\mathscr{W}x(a,b) = \int_{-\infty}^{+\infty} x(t)\psi_{a,b}(t)\,dt \tag{1}$$

where a and b are the scaling and translation factors. Different versions of wavelet functions $\psi_{a,b}(t)$ are obtained from the basic wavelet by

$$\psi_{a,b}(t) = |a|^{-\frac{1}{2}}\psi\left(\frac{t-b}{a}\right) \tag{2}$$

where a and b are real ($a \neq 0$). A large value of the factor a stretches the basic wavelet function and allows the analysis of low-frequency components of the signal. A small value of a gives a contracted version of the basic wavelet and then allows the analysis of high-frequency components.

2.2 Discrete wavelet transform

By choosing fixed values $a = a_0^m$ and $b = nb_0a_0^m$, $m, n = 0, \pm1, \pm2, \ldots$, we obtain for the discrete wavelet transform (DWT):

$$\begin{aligned}\mathscr{W}x(m,n) &= \int_{-\infty}^{+\infty} x(t)\psi_{m,n}(t)\,dt \\ &= a_0^{-m/2}\int_{-\infty}^{+\infty} x(t)\psi(a_0^{-m}t - nb_0)\,dt\end{aligned} \tag{3}$$

Both in continuous and discrete cases the wavelet function should satisfy $\int \psi(t)\,dt = 0$. Values $a_0 = 2$ and $b_0 = 1$ construct discrete wavelets $\psi_{m,n}(t) = 2^{-m/2}\psi(2^{-m}t - n)$ used in multiresolution analysis constituting an orthonormal basis for $L^2(\mathbb{R})$ (Daubechies 1992).

In discrete wavelet analysis, $x(t)$ is decomposed on different scales as follows:

$$x(t) = \sum_{j=1}^{K}\sum_{k=-\infty}^{\infty} d_j(k)\psi_{j,k}(t) + \sum_{k=-\infty}^{\infty} a_K(k)\phi_{K,k}(t)$$

where $\psi_{j,k}(t)$ are discrete analysis wavelets and $\phi_{K,k}(t)$ are discrete scaling functions, $d_j(k)$ are the detailed signals (wavelet coefficients) at scale 2^j, and $a_K(k)$ is the approximated signal (scaling coefficients) at scale 2^K. In Fig. 1a, the idea of discrete wavelet analysis is presented by means of a wavelet decomposition tree.

The discrete wavelet transform can be implemented by the scaling and wavelet filters

$$h(n) = \frac{1}{\sqrt{2}}\langle\phi(t), \phi(2t-n)\rangle \tag{4}$$

$$g(n) = \frac{1}{\sqrt{2}}\langle\psi(t), \phi(2t-n)\rangle = (-1)^n h(1-n) \tag{5}$$

being quadrature mirror filters (QMF) (Daubechies 1992). The estimation of the detail signal at level j will be done by convolving the approximate signal at level

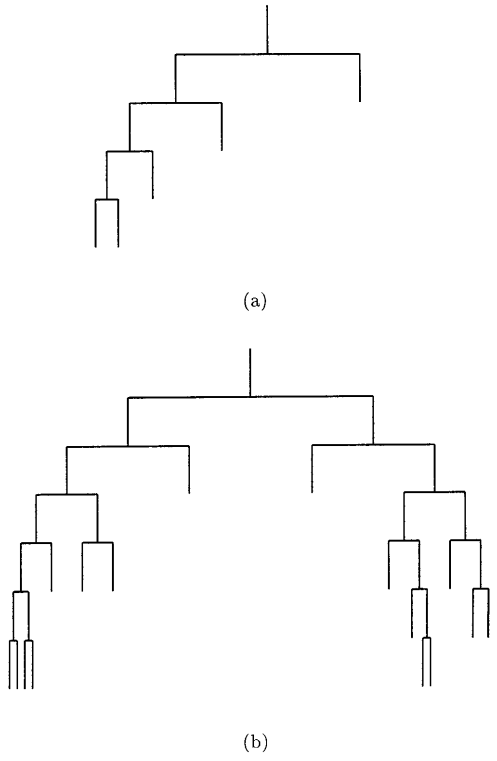

Fig. 1. **a** A wavelet decomposition tree of depth 4. **b** An optimized wavelet packet decomposition tree of depth 6

$j-1$ with the coefficients $g(n)$. Convolving the approximate signal at level $j-1$ with the coefficients $h(n)$ gives an estimate for the approximate signal at level j. The decomposition scheme involves retaining every other sample of the filter output.

2.3 Wavelet packet analysis

Let us define the scaling function $W_0(t) = \phi(t)$ and the wavelet function $W_1(t) = \psi(t)$. Then we can write functions $W_m(t)$, $m = 0, 1, 2, \ldots$, as

$$W_{2m}(t) = 2\sum_{n=0}^{2N-1} h(n) W_m(2t-n) \tag{6}$$

$$W_{2m+1}(t) = 2\sum_{n=0}^{2N-1} g(n) W_m(2t-n) \tag{7}$$

The analysing functions called wavelet packet atoms are given in an orthogonal case as

$$W_{j,m,n}(t) = 2^{-j/2} W_m(2^{-j}t - n) \tag{8}$$

where j is a scale parameter, n is a time-localization parameter, and parameter m gives roughly the number of 'cycles' included in the oscillating waveform. A wavelet packet can be considered as a waveform whose oscillations persists for many cycles but is still finite. With fixed value of j, the function $W_{j,m,n}(t)$ analyses the signal around the position $2^j \cdot n$ at the scale 2^j. The analysed frequencies are roughly given by $n/2N$ with $n = 0, 1, \ldots, (2j-1)$.

Wavelet packet analysis is a generalization of wavelet analysis offering a richer decomposition procedure. Both detail and approximation signals are split at each level into finer components. A set of details and approximations is called the wavelet packet decomposition tree.

2.4 Optimization of the wavelet packet decomposition

Discrete wavelet decomposition allows searching an optimal decomposition among L trees if a signal of length $N = 2^L$ has been decomposed at L levels. Wavelet packet analysis involves the selection of an optimal decomposition tree among at most 2^L different subtrees of depth L. The optimization can be based on e.g. the minimization of the entropy of the analysed signal, where the optimized decomposition is called the best tree. The idea is to look at each node of the decomposition tree and quantify the information to be gained by performing each split. The entropy can be obtained by many approaches; we calculated the Shannon entropy (Coifman and Wickerhauser 1992) defined as $\mathscr{E}(x) = -\sum_t x^2(t) \cdot \log(x^2(t))$. In Fig. 1b, an optimized wavelet packet decomposition tree is shown, which schematically presents the idea of this procedure.

2.5 'Denoising' the signal

A possible application of the discrete wavelet analysis is to remove undesired components (noise) from the signal through a denoising approach. Basically, the procedure includes decomposing the signal into the detail components described above, identifying the noise components, and reconstructing the signal without those components. This is called the linear denoising approach. The linear denoising approach assumes that the noise can be found within certain scales, for example, at the finest scales when the coarsest scales are assumed to be noise-free. More sophisticated denoising can be done by applying the non-linear thresholding approach, which involves discarding the details exceeding a certain limit. This approach assumes that every wavelet coefficient contains noise and is distributed over all scales.

The non-linear denoising by both soft- and hard-thresholding methods was performed (Donoho 1995). The soft-thresholded wavelet coefficients will be

$$\eta(d_j(k)) = \begin{cases} \text{sign}\ (d_j(k)) \cdot (|d_j(k)| - p), & \text{if } |d_j(k)| > p \\ 0, & \text{if } |d_j(k)| \le p \end{cases} \tag{9}$$

where p is the applied threshold. The wavelet coefficients whose absolute values are lower than the threshold are first set to zero, and then the remaining non-zero

coefficients are shrunk towards zero. With hard thresholding, the thresholded coefficients will be

$$\eta(d_j(k)) = \begin{cases} d_j(k), & \text{if } |d_j(k)| > p \\ 0, & \text{if } |d_j(k)| \le p \end{cases} \quad (10)$$

which simply means setting to zero the absolute coefficients lower than the threshold p.

The assumed model for a noisy signal was $x(t) = f(t) + e(t)$, where $f(t)$ is the noise-free signal and $e(t)$ is the white or non-white noise of variance δ^2. The performance of the methods was evaluated from the simulations with L_2-norm given by the equation

$$||f_o - \hat{x}_i||_2 = \left(\sum_t |f_o(t) - \hat{x}_i(t)|^2 \right)^{1/2} \quad (11)$$

where f_o denotes the original ECG signal being the same for all simulations, and $\hat{x}_i$ denotes the ECG signal with added noise after noise removal.

2.6 Selection of the threshold

The threshold p was selected for each signal using four threshold estimation procedures: SURE, HEURISTIC SURE, FIXTHRES and MINIMAX principles. The aim was to compare the performance obtained by different methods in the noise removal of an ECG signal. Stein's unbiased risk estimate (SURE) (Donoho 1993; Donoho and Johnstone 1995) is an adaptive threshold selection rule defined as $p = \sqrt{2 \cdot \log_e(n \cdot \log_2(n))}$, where n is the number of samples in the signal vector. With this approach, obtaining risks and minimizing them with respect to p values give a threshold selection. The method is adaptive through searching a threshold level for each wavelet decomposition level. A fixed threshold approach FIXTHRES calculates the threshold with respect to the length of the signal, and the estimated threshold is given by $p = \sqrt{2 \cdot \log_e(n)}$ (Donoho and Johnstone 1994). The HEURISTIC SURE approach, being a variant of the first, replaces in very noisy conditions the SURE with FIXTHRES estimate (Misiti et al. 1996). Further, the MINIMAXI procedure applies a fixed threshold $p = 0.3936 + 0.1829 \cdot \log(n)$ (Misiti et al. 1996) to produce a so-called minimax performance for mean square error against an ideal case (Bruce and Gao 1996, Donoho and Johnstone 1994).

The underlying signal model assumes the noise is normally distributed with zero mean and variance of 1, which means that we have to rescale the threshold values when dealing with unscaled and non-white noise. When normally and uniformly distributed noises were studied, calculated thresholds were rescaled by the standard deviation of noise estimated from the finest level of the decomposition of each signal so that $\hat{p} = p \cdot \hat{\delta}$. Further, with AR(4)-noise, the noise level was estimated scale by scale to take account of the obviously strong high-frequency content. In the wavelet approach, this was done by calculating $\hat{\delta}$ for all scales. In the wavelet packet case, $\hat{\delta}$ was estimated from the first node at each subdecomposition band which gave the best statistics for the noise level estimation. As a robust estimate of the standard deviation $\hat{\delta} = \text{Median}(|d_j(k)|)/.6745$ was used (Donoho and Johnstone 1994).

2.7 Experimental setup

In this study, 50 independent simulations were used to evaluate the performances of the applied denoising methods. Simulations were created by adding three types of noise to the noise-free ECG: Gaussian and uniformly distributed white noise, and non-white noise generated by an autoregressive (AR) model of order 4. The noise amplitudes were scaled so that the signal-to-noise ratio was 5 dB for all signals. The performances of the methods were studied by obtaining errors within an ECG including 2600 samples and, more specifically, obtaining errors within 6 QRS-complexes extracted manually from the whole ECG. The latter approach allowed the investigation of the performance of the denoising methods to handle the high-frequency parts of the ECG. Matlab software was utilized with Wavelet Toolbox to perform wavelet analysis for the digitized ECG signal obtained from an anaesthesized monkey. A 512 Hz sampling frequency was used with a resolution of 12 bits.

When considering the compactly supported orthogonal wavelet families (Daubechies, Symlets, Coiflets) with discrete transform, the Coiflet wavelet basis was found to be most suitable. The denoising performances were very close between these families, however, with Coiflets showing a slightly better performance. Coiflet and Symlet wavelets were created to improve the features of Daubechies wavelets (Daubechies 1992). The Coiflet wavelet (Coif5) of order $N = 5$ was used which had lowest denoising error among Coiflet functions. The Coif5 function is a near symmetrical wavelet, which is compactly supported with the maximum number of vanishing moments for a given support width $(6N - 1)$ (Fig. 2a). The analysis done by Coif5 wavelet is orthogonal. Wavelet packet analysis was made by Coiflet wavelet packet function $W_{j,m,n}(t)$ with $m = 5$, see Fig. 2b.

3 Results

Denoising performances of the four threshold selection methods are reported in Tables 1 and 2 with optimal decomposition depths. Generally, the denoising error first decreased as the decomposition depth increased. The results were optimized in respect of the minimum value of the error averages giving an optimal downsampling depth, which varied between denoising approaches. Methods were compared in wavelet and wavelet packet analyses applying both soft and hard thresholding. First the performance within the whole ECG strip including six cardiac cycles was measured. The results show that wavelet denoising approaches had better overall denoising performances than wavelet packet approaches in all cases except with the HEURISTIC SURE

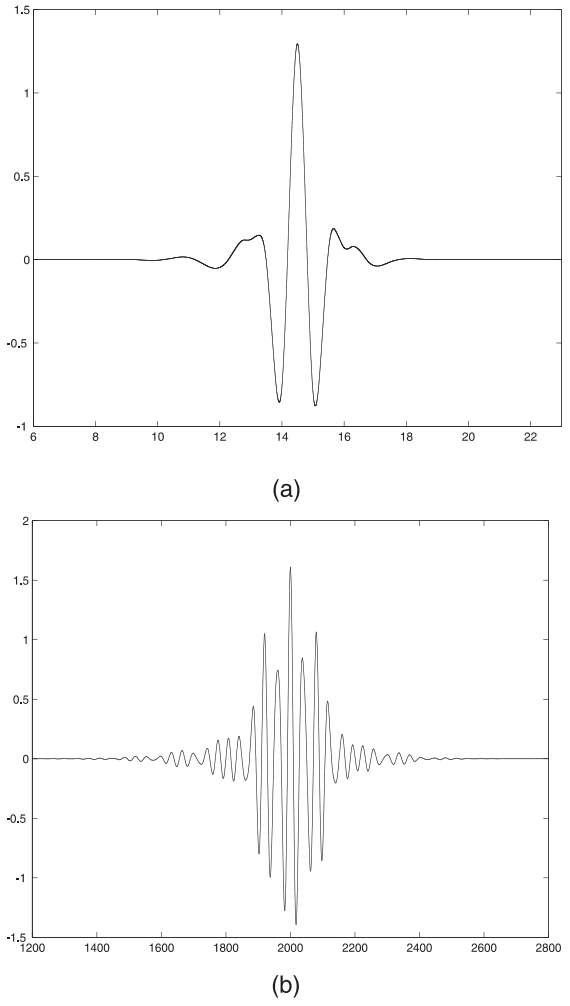

Fig. 2. **a** A Coiflet wavelet function used for a discrete wavelet analysis. **b** A Coiflet wavelet packet function used for a discrete wavelet packet analysis

rule when using hard thresholding for white noises. The wavelet methods were preferable in removing especially the AR(4)-noise, when the errors were generally 2–5 times greater with wavelet packets (Tables 1 and 2). With other noise types the difference varied from a few to a few tens of percent for wavelet methods.

In the wavelet-based approach, the most efficient noise-removing method with soft thresholding was HEURISTIC SURE, which gave the lowest error averages (Table 1). The FIXTHRES rule showed the best performance with hard thresholding. When comparing soft and hard thresholding, the result depended upon the threshold selection rule and the added noise. The highest errors for all noise types with soft thresholding were produced by the FIXTHRES and MINIMAX methods and with hard thresholding by the SURE and HEURISTIC SURE methods.

In the wavelet packet-based approach, the HEURISTIC SURE and FIXTHRES rules produced the lowest denoising errors (Table 2). FIXTHRES and MINIMAX approaches gave the largest errors when soft thresholding was used. Furthermore, with hard thresholding, the SURE and MINIMAX methods had the highest error averages except for AR(4)-noise, when the FIXTHRES and MINIMAX indicated the poorest performance. When comparing the soft and hard thresholding methods, in all cases except FIXTHRES with all noise types and MINIMAX with AR(4)-noise, the soft thresholding was better in denoising the ECG.

The ability of denoising methods to remove noise from the high-frequency parts of an ECG was studied by determining the error values within QRS-complexes, see Tables 3 and 4. The wavelet methods were more efficient to remove the AR(4)-noise, when the errors were up to 4 times greater with wavelet packets. However, wavelet packet approaches showed better performances than wavelet approaches in removing normally or uniformly distributed noise within the QRS-area, especially as hard thresholding was used. The SURE method produced the lowest proportional error for the wavelet denoising using the hard thresholding method for all noise types. With soft thresholding, except AR(4)-noise, FIXTHRES had the lowest proportional error within the QRS-area. In the wavelet packet approach, the lowest proportional error was indicated most often by SURE or HEURISTIC SURE rules. Generally, denoising errors seem to concen-

Table 1. Denoising performance of wavelet denoising approach. Values are means and standard deviations of $||f_o - \hat{x}_i||_2$. $N(0, \delta^2)$ is normally distributed and zero mean noise with variance δ^2, $U[a, b]$ stands for a uniformly distributed noise, and AR(4) is the non-white noise generated by an autoregressive model of order 4. d_{opt} is the optimal decomposition depth minimizing the denoising error

Noise type	Thresholding non-linearity	Thresholding selection rule			
		SURE (d_{opt})	HEURISTIC SURE (d_{opt})	FIXTHRES (d_{opt})	MINIMAX (d_{opt})
$N(0, \delta^2)$	Soft	449.3 ± 16.2 5	445.7 ± 12.7 5	573.0 ± 11.6 2	538.1 ± 16.2 4
	Hard	660.8 ± 74.0 4	539.9 ± 19.8 4	444.5 ± 12.0 4	531.9 ± 24.8 4
$U[a, b]$	Soft	445.3 ± 13.6 5	444.6 ± 13.4 5	576.2 ± 10.7 2	532.5 ± 17.7 4
	Hard	646.4 ± 64.6 4	543.5 ± 19.4 4	446.3 ± 17.0 4	526.0 ± 25.6 4
AR(4)	Soft	363.5 ± 20.7 4	363.3 ± 14.1 4	394.2 ± 13.3 2	390.2 ± 13.5 2
	Hard	526.4 ± 93.0 4	382.6 ± 15.7 4	365.4 ± 15.8 4	481.9 ± 25.2 4

Table 2. Denoising performance of wavelet packet denoising approach. Values are means and standard deviations of $||f_o - \hat{x}_i||_2$. See Table 1 for abbreviations

Noise type	Thresholding non-linearity	Thresholding selection rule			
		SURE (d_{opt})	HEURISTIC SURE (d_{opt})	FIXTHRES (d_{opt})	MINIMAX (d_{opt})
$N(0, \delta^2)$	Soft	466.1 ± 22.7 5	447.8 ± 20.9 5	866.3 ± 27.3 7	594.1 ± 20.4 6
	Hard	705.4 ± 66.1 3	526.8 ± 21.0 5	505.7 ± 19.5 5	699.1 ± 27.6 3
$U[a, b]$	Soft	462.2 ± 19.7 5	448.7 ± 20.2 5	859.1 ± 25.7 7	591.1 ± 17.7 6
	Hard	685.3 ± 61.5 3	526.8 ± 20.1 4	507.7 ± 18.9 5	695.5 ± 31.4 3
AR(4)	Soft	745.6 ± 38.7 6	600.9 ± 222.1 6	2074.5 ± 114.6 7	1371.2 ± 101.4 7
	Hard	903.3 ± 23.0 6	650.5 ± 144.6 6	1416.1 ± 117.5 7	1020.5 ± 45.3 6

Table 3. Denoising performance of wavelet denoising approach measured within QRS-complexes. Values presented are means and standard deviations of $||f_o - \hat{x}_i||_2$. The decomposition depths are the same as in Table 1. See Table 1 for abbreviations

Noise type	Thresholding non-linearity	Thresholding selection rule			
		SURE	HEURISTIC SURE	FIXTHRES	MINIMAX
$N(0, \delta^2)$	Soft	286.2 ± 21.4	300.5 ± 18.5	258.5 ± 10.6	445.8 ± 18.4
	Hard	274.6 ± 25.4	259.9 ± 12.2	319.7 ± 16.7	287.9 ± 17.9
$U[a, b]$	Soft	283.6 ± 19.6	297.3 ± 19.0	259.1 ± 11.9	439.5 ± 19.2
	Hard	269.5 ± 20.0	262.1 ± 12.3	321.7 ± 18.1	289.1 ± 17.5
AR(4)	Soft	208.6 ± 13.9	230.0 ± 9.9	214.3 ± 7.7	200.7 ± 8.4
	Hard	221.9 ± 28.2	209.7 ± 10.0	224.4 ± 10.1	226.2 ± 22.6

Table 4. Denoising performance of wavelet packet denoising approach measured within QRS-complexes. Values presented are means and standard deviations of $||f_o - \hat{x}_i||_2$. The decomposition depths are the same as in Table 2. See Table 1 for abbreviations

Noise type	Thresholding non-linearity	Thresholding selection rule			
		SURE	HEURISTIC SURE	FIXTHRES	MINIMAX
$N(0, \delta^2)$	Soft	271.2 ± 19.1	301.8 ± 28.7	607.1 ± 22.4	381.3 ± 19.4
	Hard	270.9 ± 27.6	253.9 ± 17.0	272.4 ± 16.6	287.6 ± 23.4
$U[a, b]$	Soft	267.1 ± 18.0	300.5 ± 27.7	602.7 ± 23.2	377.5 ± 18.6
	Hard	264.9 ± 24.8	264.4 ± 16.4	276.5 ± 16.1	286.6 ± 18.2
AR(4)	Soft	324.0 ± 34.1	308.8 ± 39.9	866.2 ± 61.6	562.5 ± 32.2
	Hard	320.8 ± 23.2	258.3 ± 42.7	493.8 ± 36.9	391.4 ± 33.7

trate on the QRS-area when the pure wavelet approach is employed. The QRS-complex area included proportionally less error when hard thresholding was used, except when the FIXTHRES rule as the wavelet decomposition was used.

When comparing the performance within QRS-complexes using absolute error measurement, HEURISTIC SURE or FIXTHRES gave most often the lowest error values for the wavelet approach. In the wavelet packet approach, the best performances were found by SURE and HEURISTIC SURE rules. Further, the hard thresholding method gave lower absolute errors with all variations of denoising methods when the wavelet packet approach was used. With pure wavelet denoising, results were the same except with FIXTHRES, when soft thresholding gave a lower error. SURE and MINIMAX also showed better denoising performance for AR(4)-noise with soft thresholding.

Because numerical error measures do not necessarily tell everything about noise removal, it is useful to check the denoised signals visually. Strange signal patterns probably exist which cannot be predicted from error values. It is important to see the error signal between the noisy and the denoised signal, because then one can observe how the error is localized within the cardiac cycle. In Fig. 3, the successful result of denoising an ECG with normally distributed noise by the wavelet approach using the MINIMAX method is shown. The error between the original and denoised ECG is mainly concentrated within the QRS-complexes.

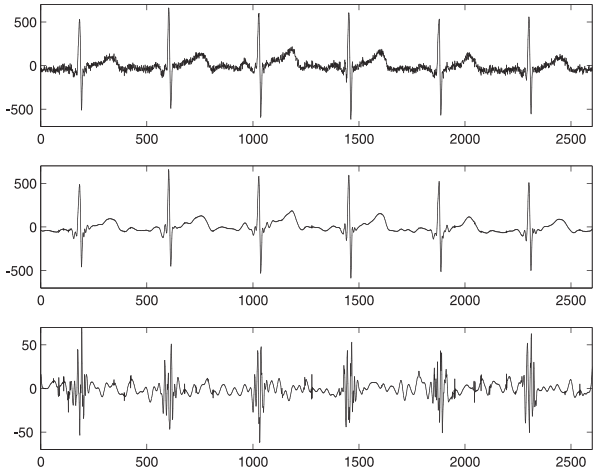

Fig. 3. *Top panel*: Electrocardiogram (ECG) with added normally distributed white noise. *Centre panel*: Result of the noise removal performed by wavelet-based approach using MINIMAX thresholding selection rule with soft thresholding non-linearity. *Bottom panel*: Error signal between the original and denoised ECG

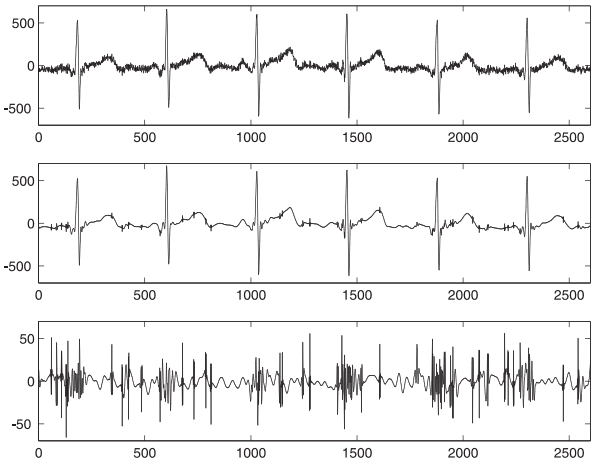

Fig. 5. *Top panel*: ECG with added normally distributed white noise. *Centre panel*: Result of the noise removal performed by wavelet-based approach using MINIMAX thresholding selection rule with hard thresholding non-linearity. *Bottom panel*: Error signal between the original and denoised ECG

Sometimes performing signal denoising does not mean that only the added noise has been removed. In Fig. 4, the ECG is very clean after denoising, and the error signal has a large amplitude, which means that the denoising method has not only been robust in removing added noise but also has seriously altered the ECG signal. This result was observed particularly with the wavelet packet approach when the FIXTHRES approach was applied. With wavelet-based denoising, hard thresholding showed spiky patterns in the error signal which are also seen in the denoised ECG, which was typical with all noise types when the SURE and MINIMAX rules were used (Fig. 5). The wavelet packet-based approaches had difficulties in removing the noise, as can be seen in Fig. 6, where the error signal includes a large random component due to non-white noise. This was common particularly when SURE and MINIMAX were applied, and also often with other rules with hard thresholding.

4 Discussion

In this work new wavelet- and wavelet packet-based noise removal schemes were studied using ECG with simulated noises. The performances of several variations of denoising including thresholding rules and the type of non-linearity were compared. A level-dependent scaling

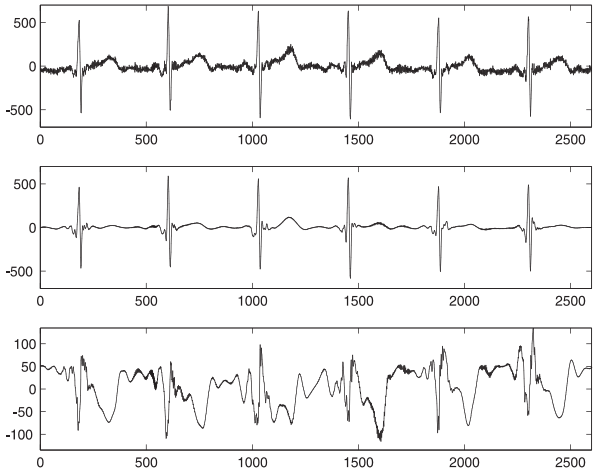

Fig. 4. *Top panel*: ECG with added noise generated by an autoregressive model of order 4. *Centre panel*: Result of the noise removal performed by wavelet packet-based approach using FIXTHRES thresholding selection rule with soft thresholding non-linearity. *Bottom panel*: Error signal between the original and denoised ECG

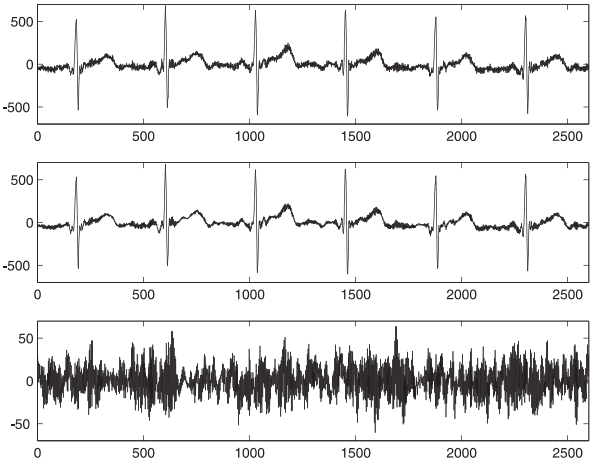

Fig. 6. *Top panel*: ECG with added noise generated by an autoregressive model of order 4. *Centre panel*: Result of the noise removal performed by wavelet packet-based approach using SURE thresholding selection rule with hard thresholding non-linearity. *Bottom panel*: Error signal between the original and denoised ECG

of the thresholds was used for adjusting to the non-white noise structure.

Wavelet and wavelet packets showed different results, which is mainly due to the different division strategies of the signal decomposition structures. Furthermore, these analysing functions also differ in shape. The wavelet-based approach produces a dyadic decomposition structure which is constant for all signals. Correspondingly, the wavelet packet approach is an adaptive method using an optimization of the best tree decomposition structure independently for every signal, which can be quite irregular and gain very fine features. Generally, this kind of adaptivity did not offer an improved overall denoising performance compared with a simpler wavelet approach. Only applying the HEURISTIC SURE rule with soft thresholding produced a superior result. Inside the high-frequency parts of the ECG, the situation varied even more. Visual examination of the error signal proved remarkable, showing the localization of the error within the cardiac cycle as well as its nature. The obtained error values cannot directly indicate the improvement for the ECG waveform detection. However, a large error within a certain area suggests an impaired accuracy of the waveform measurement, which can only be quantified by the appropriate tests.

The performance of the wavelet packet-based noise removal may be improved by adapting the signal decomposition structure to the changing signal characteristics as presented by Xiong et al. (1997), where the signal was divided into segments of variable lengths using dynamic programming setting. The approach involves calculating optimized wavelet packet decompositions independently for each segment. This procedure would probably be useful for ECGs which are corrupted by different types of noises with time-varying magnitudes.

Bruce and Gao (1996) studied soft and hard non-linearities and derived a theoretical result that soft thresholding has a higher bias, but lower variance than hard thresholding. This was also supported by their experiments. This is probably due to the basic properties of these two approaches: the hard thresholding function has a discontinuity, and the soft thresholding function shrinks all big coefficients towards zero. Their results were obtained for FIXTHRES and MINIMAX rules as the wavelet denoising approach was applied.

The observations found in this work support those findings partly, as with these two threshold selection rules soft thresholding tends to give higher overall error values. Nevertheless, hard thresholding gave constantly bigger errors within the QRS-area when the FIXTHRES rule was applied. The wavelet packet approach showed different results, indicating larger error rates in all cases for soft than hard thresholding within QRS-complexes. With SURE and HEURISTIC SURE rules, the soft thresholding non-linearity tends to give a more acceptable overall denoising result compared with hard thresholding. However, it should be noted that when using soft non-linearity, the error between the original and denoised ECG was concentrated within the QRS-complexes. In that case, the absolute error values were generally bigger than using hard non-linearity. Only the FIXTHRES rule with hard non-linearity showed with a wavelet approach a proportionally higher error within the QRS-area than soft non-linearity. It is apparent that soft and hard thresholding cause different high-frequency balances. This is due to the fact that the soft thresholding in general produces proportionally a larger error within the QRS-area by rounding off towards zero the coefficients bigger than the threshold, which obviously affects the coefficients including a remarkable amount of information about the original ECG.

Acknowledgements. I thank Astra Hässle for providing ECG data. The work was financially supported by Astra Finland.

References

1. Bradie B (1996) Wavelet packet-based compression of single lead ECG. IEEE Trans Biomed Eng 43:493–501
2. Bruce AG, Gao HY (1996) Understanding waveshrink: variance and bias estimation. Biometrika 83:727–745
3. Coifman R, Wickerhauser M (1992) Entropy-based algorithms for best basis selection. IEEE Trans Inform Theory 38:713–718
4. Daubechies I (1992) Ten lectures on wavelets. SIAM
5. Dickhaus H, Khadra L, Brachmann J (1994) Quantification of ECG late potentials by wavelet transformation. Methods Programs Biomed 43:185–192
6. Donoho D (1993) Nonlinear wavelet methods for recovery of signals, densities, and spectra from indirect and noisy data. Proc Symp Appl Math 47:173–205
7. Donoho D (1995) De-noising by soft-thresholding. IEEE Trans Inform Theory 41:612–627
8. Donoho D, Johnstone I (1994) Ideal spatial adaptation by wavelet shrinkage. Biometrika 81:425–455
9. Donoho D, Johnstone I (1995) Adapting to unknown smoothness via wavelet shrinkage. J Am Stat Assoc 90:1200–1224
10. Gramatikov B, Yi-chun S, Rix H, Caminal P, Thakor N (1995) Multiresolution wavelet analysis of the body surface ECG before and after angioplasty. Ann Biomed Eng 23:553–561
11. Hamilton PS (1996) A comparison of adaptive and non-adaptive filters for reduction of power line interference in the ECG. IEEE Trans Biomed Eng 43:105–109
12. Hilton M, Ogden R (1997) Data analytic wavelet threshold selection in 2-D signal denoising. IEEE Trans Signal Proc 45:496–500
13. Hilton M, Ogden T, Hattery D, Eden G, Jawerth B (1996) Wavelet denoising of functional MRI data. In: Aldroubi A, Unser M (eds) Wavelets in medicine and biology. CRC Press, Boca Raton
14. Karrakchou M, Kunt M (1996) Interference canceling in biomedical systems: the mutual wavelet packets approach. In: Aldroubi A, Unser M (eds) Wavelets in medicine and biology. CRC Press, Boca Raton
15. Kolaczyk E (1996) An application of wavelet shrinkage to tomography. In: Aldroubi A, Unser M (eds) Wavelets in medicine and biology CRC Press, Boca Raton
16. Lander P (1997) Time-frequency plane Wiener filtering of the high-resolution ECG: development and application. IEEE Trans Biomed Eng 44:256–265
17. Li C, Zheng C, Tai C (1995) Detection of ECG characteristic points using wavelet transforms. IEEE Trans Biomed Eng 42:21–28
18. Mallat S (1989) Multifrequency channel decompositions of images and wavelet models. IEEE Trans Acoust Speech Signal Proc 37:2091–2110
19. Meste O, Rix H, Caminal P, Thakor N (1994) Ventricular late potentials characterization in time-frequency domain by means of a wavelet transform. IEEE Trans Biomed Eng 41:625–634

20. Misiti M, Misiti Y, Oppenheim G, Poggi JM (1996) Wavelet toolbox. For use with MATLAB. Wavelet Toolbox User's Guide. The MathWorks, Natick
21. Pahlm O, Sörnmo L (1984) Software QRS detection in ambulatory monitoring – a review. Med Biol Eng Comput 22:289–297
22. Pei SO, Tseng CC (1995) Elimination of AC interference in electrocardiogram using IIR notch filter with transient suppression. IEEE Trans Biomed Eng 42:1128–1132
23. Ramakrishnan AG, Saha S (1997) ECG coding by wavelet-based linear prediction. IEEE Trans Biomed Eng 44:1253–1261
24. Senhadji L, Thoraval L, Carrault G (1996) Continuous wavelet transform: ECG recognition based on phase and modulus representations and hidden Markov models In: Aldroubi A, Unser M (eds) Wavelets in medicine and biology. CRC Press, Boca Raton
25. Sörnmo L (1993) Time-varying digital filtering of ECG baseline wander. Med Biol Eng Comput 31:503–508
26. Thakor N, Zhu YS (1991) Applications of adaptive filtering to ECG analysis: noise cancellation and arrhythmia detection. IEEE Trans Biomed Eng 38:785–794
27. Thakor N, Sun YC, Rix H, Caminal P (1993a) Multiwave: a wavelet-based ECG data algorithm. IEICE Trans Inf Syst E76-D:1462–1469
28. Thakor N, Xin-Rong G, Yi-Chun S, Hanley D (1993b) Multiresolution wavelet analysis of evoked potentials. IEEE Trans Biomed Eng 40:1085–1093
29. Tikkanen P, Kinnunen H, Nissilä S, Heikkilä I, Mela M (1999) Ambulatory heart rate variability analysis methods. Automedica 17
30. Xiong Z, Ramchandran K, Herley C, Orchard M (1997) Flexible tree-structured signal expansions using time-varying wavelet packets. IEEE Trans Signal Proc 45:333–345
31. Xue Q, Hu YH, Tompkins WJ (1992) Neural-network-based adaptive matched filtering for QRS detection. IEEE Trans Biomed Eng 39:317–329