БЪЛГАРСКА АКАДЕМИЯ НА НАУКИТЕ . BULGARIAN ACADEMY OF SCIENCES

ПРОБЛЕМИ НА ТЕХНИЧЕСКАТА КИБЕРНЕТИКА И РОБОТИКАТА, 50
PROBLEMS OF ENGINEERING CYBERNETICS AND ROBOTICS, 50

София . 2000 . Sofia

Local Signal Feature Detection by Wavelet Transforms*

Atanas Gotchev, Zdravko Nikolov, Elena Rangelova, Nikolay Nikolaev

Institute of Information Technologies, 1113 Sofia

I. Introduction

In signal analysis the detection of relevant feature is of great importance because they carry out the information about signal and genesis. Most often, the important signal features, such as edges, spikes, transients, etc., are localized either in time or in frequency (scale) or in both. For this reason the wavelet bases are very attractive for the local feature detection tasks [1], [2]. They offer well organized and flexible coordinate systems which can decompose the signal in the time-frequency plane using fast algorithms. The most widely used wavelet bases are the orthonormal bases with compact support and maximum number of vanishing moments [3]. The orthonormality is very important for perfect signal reconstruction and for fast algorithms construction and that is why these bases are used especially for image and speech compression [4], [5]. However, they have two drawbacks complicating their application for transient detection:

– they are not shift-invariant, and

– they have asymmetric shape.

The shift-invariance has been got over using redundant representations without downsampling at each scale, continuous wavelet transforms, etc. [6], [7], also called orthonormal wavelet shell.

The symmetric basic functions are preferred for their linear phase and for easy finding of zero-crossings and extrema. The symmetric basis functions, however, are either without compact support or biorthogonal [8], [9].

The present work studies a wavelet approach for local signal features detection, based on segmental wavelet transform with basis biorthogonal of the Haar basis which uses the information about the (potentially) transition point.

*This work is partially supported by the Institute of Inf. Technologies (project 010016) and by the National Science Fund under contract No 1626/96.

II. Segmented wavelet transform

Segmented Wavelet Transform (SWT) is signal decomposition into a basis of biorthogonal wavelets, which is dependent on preliminary choice of the segmentation point which is in the vicinity of signal abrupt transition. Wavelets are obtained using average-interpolation refinement scheme [12].

II.1. Average-interpolating multiresolution

Segmented multiresolution analysis (SMA) V_j^t for the interval $[0,1]$ is multiresolution analysis (MRA), where intervals $[0, t]$ and $[t, 1]$ are considered separately and it is not necessary that functions in V_j^t are continuous in t.

Let consider an array $(a_{j,k})_{k\in\mathbb{Z}}$ which consists of averages of a function f in dyadic intervals $I_{j,k} = [k/2^j, (k+1)/2^j]$. There is selected polynomial $\pi_{j,k}$ of even degree D, which allows reception of interpolated sequences of finer scale by refinement procedure for average interpolating described in [13].

The vector space V_j of functions obtained by refining sequences $(a_{i,k})_k$ has an alternative description. Refining the Kronecker sequence $a_{0,k} = \delta_{k,0}$, we obtain the fundamental functions $\phi = \phi_D$. These functions and their integer translations and dyadic dilations $\phi_{j,k}(t) = 2^{j/2}\phi(2^j t - k)$ generate the spaces $V_j = \{f: f = \sum_k \beta \phi_{jk}(t)\}$. The space V_j makes up MRA, which is biorthogonal to the classical Haar basis.

If we have averages in scale $j+1$ we can obtain averages in scale j because $a_{j,k} = (a_{j+1,2k} + a_{j+1,2k+1})/2$. The vector space W_j obtained by refinement of sequences $(a_{j+1,k})_k$, where $a_{j+1,2k} = -a_{j+1,k+1}$, consists entirely of functions which averages in the coarsest scale are zero; in fact this is the difference space $W_j = V_{j+1} - V_j$.

There are 2^j basis functions both in the spaces V_j and in W_j which are denoted $\phi_{j,k}$ and $\psi_{j,k}$ respectively. Here j_0 is fixed so that $2^{j_0} > 2(D+2)$. Then every function f in $L^2[0, 1]$ has an expansion:

$$f = \sum_{k=0}^{2^{j_0}-1} \beta_{j_0,k}\,\phi_{j_0,k} + \sum_{j\ge j_0} \sum_{k=0}^{2^{j-1}} \alpha_{j_0,k}\,\psi_{j_0,k}\,. \tag{1}$$

The mapping from f to its coefficients $((\beta_{j_0,k})_k, (\alpha_{j_0,k}), (\alpha_{j_0+1,k})_k, \ldots)$ is *non-segmented wavelet transform*. There is a fast algorithm associated with it. Given $\beta_{j_1,k}$ in scale $2^{-j_1} \ll 2^{-j_0}$, the coefficients $(\beta_{j_0,k})$ and $(\alpha_{j,k})$ of all coarser levels $j_0 \le j < j_1$ can be computed for time n, where $n = 2^{j_1}$.

II.2. Segmented refinement

We consider segmented refinement with segmentation point τ, which is in the "heart" of the observed interval, so that $D/2^j < \tau < (2^j - D/2^j)$. Given a sequence of averages $(a_{j,k})$, $0 \le k < 2^j$, we synthesise "pseudo" averages by the following procedure:

1) At each point k, which is more than $D/2$ points away from the boundaries 0 and 1 and more than $D/2$ points away from the segmentation point τ, we use interpolation procedure to find the polynomial $\pi_{j,k}$ of degree D which generates the same values in the neighbourhood $(a_{j,k'}, k' = k - D/2, \ldots, k + D/2)$ [13].

2) At each point k, which is at most $D/2$ points away from the boundaries 0 and

1 we find the polynomial $\pi_{j,k}$ of degree D, which generates the same averages in the neighbourhood $(a_{j,k'})_{k' \in N(k)}$ where $N(k)$ consists of $D+1$ nearest neighbours of k.

3) At each point k, which is at no more than $D/2$ points from the segmentation point t we consider two cases:

3a) If $\tau \notin [2^{-j}k, 2^{-j}(k+1)]$ then find the polynomial $\pi_{j,k}$ of degree D, which generates the same averages in the neighbourhood $(a_{j,k'})_{k' \in N(k)}$, where $N(k)$ consists of $D+1$ nearest neighbours of k, which are at one and the same side of the segmentation point as k.

3b) If $\tau \in [2^{-j}k, 2^{-j}(k+1)]$ then construct by constrained least squares polynomials $\pi^R_{j,k}$ and $\pi^L_{j,k}$ of degree D for block averages in the neighbourhoods on the right and left of the segmentation point; the constraint is that the polynomial $\pi^\tau_{j,k}$, which is π^L on the left of t and π^R on the right of t should have an average equal to the average a_{jk}.

4) In cases 1), 2) and 3a) we define "pseudo" averages at the next fine scales as averages of the polynomial $\pi_{j,k}$. On the left half of the sub-interval we get:

$$a^\tau_{j+1,2k} = Ave_{j+1,2k}\,\pi_{j,k}, \tag{2}$$

and on the right half:

$$a^\tau_{j+1,2k+1} = Ave_{j+1,2k+1}\,\pi_{j,k}. \tag{3}$$

In case 3b) the steps are the same, only using the piecewise polynomial $\pi^\tau_{j,k}$.

5) After having all the $a^\tau_{j+1,2k}$ set $j:=j+1$ and go to 1).

The basic property of SWT is that if the is the piecewise polynomial of degree D with one knot at τ, then the refinement process recovers π^τ exactly.

II. 3. Minimum entropy segmentation principle

Segmented wavelet transform depends essentially on the exact choice of the segmentation point t. Unfortunately in most cases we do not have information about it.

Donoho proposes adaptive determination of an appropriate segmentation depending on data [13]. Let denote with $\varepsilon=\varepsilon(0)$ the *entropy* considered as functional having low values for sparse vectors (with a few non-zero elements) and high values for vectors with many non-zero components.

The basic *Minimum Entropy Segmentation Principle* (MESP) is described by this procedure: with x we denote a vectorc with dyadic length $n=2^j$ which contains block averages at scale 2^{j_1}, with $W^\tau_n x$ we denote a vector with dyadic length $n=2^j$, which contains block averages at scale 2^{j_1}, with we denote segmented wavelet coefficients $((\beta_{j_\theta,k})_k, (\alpha_{j_{\theta+1},k})_k)$, obtained with segmentation point τ. The idea of MESP is to select from all possible basses one which gives coefficients with smallest entropy in the wavelet domain.

$$\hat{\tau} = \arg\min_{\tau \in [0,1]} \varepsilon(W^\tau_n x). \tag{4}$$

Coifman and Wickerhauser used a Shannon entropy as an information measure

for best-basis selection $\varepsilon^{CW}(\theta) = -\sum_{j,k} p_{j,k} \lg(p_{j,k})$ where $p_{j,k} = (\alpha^{\tau}_{j,k})^2$.

In addition to the original C-W entropy Donoho proposes a family of entropy measures $\varepsilon^{\alpha}, \alpha \in [0,2]$:

(5) $$\varepsilon^{\alpha}(\theta) = \sum_{j,k} p_{j,k}^{\alpha/2} .$$

We are interested in entropies ε^{j}, $\varepsilon^{j/2}$ and ε^{2}.

II. 4. Fast computation of all segmentations

The SWT needs n of order operations and seems that we need procedure of order $O(n^2)$ for calculation all pixel-level segmentations. There is a fast algorithm for calculating all possible n pixel-level segmentations in time proportional to $n \lg(n)$. The method is based on the assumption that in the segmented refinement only the blocks at a distance not more than $D/2$ away from the block containing the segmentation point are influenced by the latter. Therefore for calculation of each pixel-level segmentation $t = i/n$, $i = 0, \ldots, n-1$, we need only the coefficients which differ from the non-segmented transform. Let denote these coefficients with:

(6) $$\delta_{j,l}(t), j = j_0, \ldots, j_1 - 1, \; l = -D/2, \ldots, D/2.$$

where index j denotes the resolution level, l – the offset in the block $[t2^j]$, containing the segmentation point. Suppose we have coefficients at a given segmentation point t and an unsegmented transform. By copying values from the array n into the appropriate location of the array of non-segmented wavelet coefficients we obtain the segmented wavelet transform at the segmentation point t.

Let θ^0 denote the coordinates of non-segmented wavelet transform. We define *differential entropy* as:

(7) $$\delta\varepsilon(t) = \varepsilon(\theta^t) - \varepsilon(\theta^0).$$

The minimum of the entropy $\varepsilon(\theta^t)$ will be at the same point t as the minimum of the differentiated entropy, so it is sufficient to minimise the latter. If we denote coefficients in the non-segmented transform that could be replaced with

(8) $$\mu_{j,l}(t), j = j_0, \ldots, j_1 - 1, \; l = -D/2, \ldots, D/2,$$

it is necessary to compute the entropies of the v-coefficients and of the μ-coefficients and to minimise difference between them as a function of t.

The improved algorithm is as follows:

1) Calculate the ordinary non-segmented transform.

2) For $i = 0, \ldots, n-1$ do:

2a) Calculate the v- coefficients for $t = i/n$.

2b) Evaluate the entropy difference between the v-coefficients and the corresponding μ-coefficients from the non-segmented transform.

3) After obtaining the best i at step 2 compute segmented transform for $t = i/n$.

III. Application of the methods for ECG denoising

III. 1. The problem with ECG signal denoising

One of the most serious problems in the registration of electrocardiographic (ECG) signals is the parasite interference of muscle active potentials – electromiographic (EMG) signals. This leads to difficulties in determining of the signal parameters and making diagnoses.

Electromiographic (EMG) signal is a result of superposition of the biopotentials of the muscles, arising between the electrocardiograph's electrodes. The EMG signal has widespread spectrum – 99% of its spectral energy is located within the frequency band from 10 Hz to 1 kHz [14]. Since the spectral energy of ECG signal is located within the frequency band from 0.05 Hz to 125 kHz, the ECG and EMG spectra are overlapped.

The problem is partially avoided by lowpass filtering of the signal. This approach improves the SNR but decreases the amplitudes of the high frequency Q, R and S waves, which can be fatal in diagnostics of some diseases.

Our goal has been to suppress the parasite EMG and in the same time to preserve the parameters of the ECG.

Solving the task of denoising, ECG signal could be considered as a superposition of the signal waves – P waves, QRS complexes and T waves. Each wave carryes information for diferent signal parameters. Since the characteristics of the waves - the frequency band and duration – are diferent for each kind of waves it is important to know its location in order to process them in diferent ways.

III. 2. Denoising via wavelet transform

Let us consider the denoising task. Suppose we have N noisy samples of function f

$$y_i = f(t_i) + z_i, \quad i = 1, \ldots, N, \tag{9}$$

where $t_i = (i-1)/N$, and the z_i is a white noise with distribution $N(0, \sigma^2)$. Our goal is to estimate the vector $\mathbf{f} = f(t_i)|_{i=1}^{N}$ with small mean-squared error i.e. to find an estimation $\hat{\mathbf{f}}$ depending on $\mathbf{y}$ with small *risk* $R(\hat{\mathbf{f}}, \mathbf{f}) = N^{-1} E \| \hat{\mathbf{f}} - \mathbf{f} \|_2^2 = EAve_i (\hat{\mathbf{f}}(t_i) - f(t))^2$.

The basic property of the wavelet transforms is its orthogonality which leads to important statistical consequence: the white noise remains white noise in the transform domain. Therefore if $(y_{j,k})$ denotes wavelet coefficients of $(y_i)|_{i=0}^{N-1}$ according to the model (9) and $w_{j,k}$ denotes wavelet coefficients of $f(t_i)$, then

$$y_{j,k} = w_{j,k} + z_{j,k}, \tag{10}$$

where $z_{j,k}$ is noisy sequence with distribution $N(0, \sigma^2)$. Therefore the wavelet coefficients of a noisy sequence are simply noisy versions of the original wavelet coefficients.

So the problem of recovering f becomes the problem of recovering by using only these coefficients (often a few in number) which are essentially non-zero in the background of white Gaussian noise. This leads to using a threshold dependant scheme which "cuts" the little $y_{j,k}$ and "preserve" the big ones $y_{j,k}$.

The classical Wavelet Shrinkage (WS) denoising approach has three steps [15]:

i. Pyramid wavelet decomposition of the noisy signal;

ii. Shrinkage of the noisy wavelet coefficients with threshold depending on the statistical parameters of the signal;

iii. Inverse wavelet transform, producing the estimated signal.

III. 3. Denoising using wavelet shrinkage and time-frequency dependent threshold

In the present study the algorithm for ECG signal denoising via WS is further modified by time-frequency dependent threshold [16]. The algorithm is based on finding a threshold appropriate for ECGs and depending on the scale and time position of the transform coefficients. Fig.2 shows the threshold as a function of the wavelet coefficients' position in scale j.

The parameters of the threshold function are:

– b_i, e_i – beginning and end of the i-th QRS complex

– τ_{qrs} – threshold of the wavelet coefficients, describing QRS complexes

– τ_{st} – threshold of the coefficients, describing the areas outside of QRS complexes.

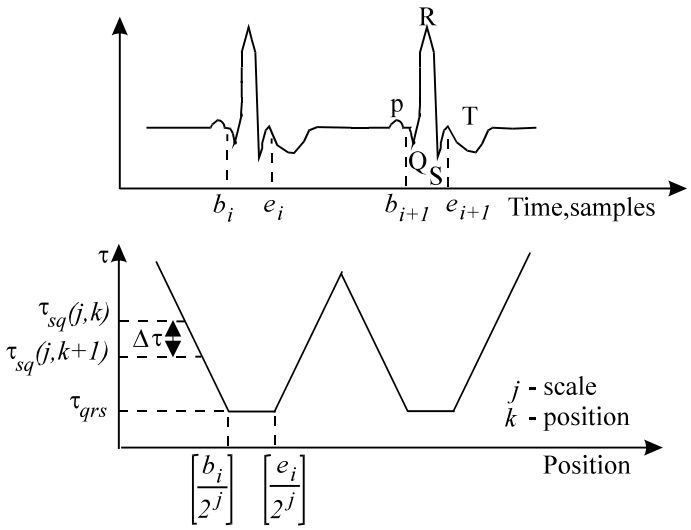

Fig.1. Threshold versus the wavelet coefficients' position in scale j

The exact determination of the beginnings and the ends of QRS complexes is essential for preservation of the QRS complexes' important signal parameters. In the present study the detection of b_i and e_i is based on the algorithm proposed by Li, Zeng and Tai [17] modified by us with the use of autocorelation shell expansion instead of orthonormal sell.

In the decomposition of the signal via Continuous Wavelet Transform the local maxima/minima in the time domain correspond to the local maxima/minima in the scales of the orthonormal shell. The positions of the extrema throughout the scales are shifted in reference to the position of the extremum in the time domain. The shift depends on the QMF length, the number of the scale and the steepness of the fronts of the extremum. The presence of shifts compicates the determination of the position and the type of signal feature.

Table 1 shows the variance of the residual signal $\hat{z} = f - \hat{f}'$ for different levels of noise variance σ_z. The variance of the noised signal y is $\sigma_y = 1$. The results are compared with the results of lowpass filtering with 20 order FIR filter with cut-off frequency 40Hz.

Table 1

No	Variance of the residual signal after denoising with the proposed technique			Variance of the residual signal after lowpass filtering		
	σ_z=0.1	σ_z=0.2	σ_z=0.3	σ_z=0.1	σ_z=0.2	σ_z=0.3
174	0.05	0.08	0.11	0.09	0.11	0.14
157	0.05	0.09	0.13	0.09	0.11	0.14
137	0.05	0.09	0.14	0.06	0.09	0.12
106	0.06	0.10	0.14	0.06	0.09	0.12
147	0.07	0.10	0.13	0.09	0.12	0.14
154	0.05	0.09	0.13	0.06	0.09	0.12

IV. Conclusions

A wavelet method for local signal feature detection has been presented. Autocorrelation functions of compactly supported wavelets seem to be the best choice for signal edges and transients detection. They have compact support and lead to redundant decomposition preserving exactly the transition points without time shifts. They are superior in comparison to the quadratic and cubic spline wavelets for detection of QRS complexes of ECG signals.

Further studies will involve the investigation of ACF derived from other orthonormal wavelets and other interpolation schemes for construction of wavelet bases appropriate for segmented transforms.

VI. References

1. Daubechies, I. The wavelet transform, time-frequency localization and signal analysis. – In: IEEE Trans. on Inf. Theory, 1990, **36**, 961-1005.
2. Mallat, S., W. L. Hwang. Singularity detection and processing with wavelets. – In: IEEE Trans. Inf. Theory, 1992, **38**, 617-643.
3. Daubechies, I. Orthonormal bases of compactly supported wavelets. – In: Comm. in Pure and Applied Math., 1988, **41**, 909-996.
4. Barnard, H. Image and video coding using a wavelet decomposition. Ph. D. Thesis, Tech. University Delft, 1994.
5. Sinha, D., A. H. Tevfik. Low bit rate transparent audio compression using adapted wavelets. – In: IEEE Trans. Signal Processing, 1993, **41**, 3463-3480.
6. Rioul, O., P. Duhamel. Fast algorithms for discrete and continuous wavelet transforms. – In: IEEE Trans. Inf. Theory, 1992, **38**, 569-586.
7. Shensa, M. The discrete wavelet transform: wedding the a trous and Mallat algorithms. – In: IEEE Trans. SP, 1992, **40**, 2464-2482.
8. Meyer, Y. Ondelettes et operateurs. Vol.1, Hermann, Paris, 1990.
9. Cohen, A., I. Daubechies, J. C. Feauveau. Biorthogonal bases of compactly supported wavelets. – In: Comm. Pure Appl. Math., 1992, **45**, 485-560.
10. Mallat, S. A theory for multiresolution signal decomposition: the wavelet representation. – In: IEEE PAMI, 1989, **11**, 674-693.

11. Saito, N., G. Beylkin. Multiresolution representations using the auto-correlation functions of compactly supported wavelets. – In: IEEE Trans. SP, 1993, **41**, 3584-3590.
12. Deslauriers, G., S. Dubic. Symmetric interpolation processes. – Constructive Approximation, **5**, 1989, 49-68
13. Donoho, D. L.. On minimum entropy segmentation. – In: Wavelets: Theory, Algorithms and Applications. Academic Press, San Diego, 1994.
14. N. V. Thakor, J. G. Webster, W. J. Tompkins. Estimation of QRS complex power spectra for design of a QRS filter. – In: IEEE Trans. on Biomedical Engineering, **31**, 1984, 702-705.
15. Donoho, D. L. Wavelet shrinkage and W.V.D. – A Ten-Minute Tour. Progress in Wavelet Analysis and Applications, 1993, 109-128.
16. N. Nikolaev, A. Gotchev. Denoising of ECG signals using wavelet shrinkage with time- frequency dependant threshold. – In: EUSIPCO - 98.
17. C. Li, C. Zheng, C. Tai. Detection of ECG characteristic point using wavelet transform. – In: IEEE Trans. on Biomedical Engineering, **42**, 1995, 21-28

Детектирование локальных особеностей сигналов при помощи волновых трансформаций

Атанас Гочев, Здравко Николов, Елена Рангелова,
Николай Николаев

Институт информационных технооги, 1113 София

(Резюме)

Предложен общий подход к детектировании локальных особеностей в разных класссах сигналов при помощи волновых трансформаций (wavelet transforms). Метод использует автокорреляционные функции волн с компактным носителем. Базисные функции детектируют точно моменты переходов, что доказывается экспериментами с ЕКГ сигналами.