Wavelet Transform Use for Feature Extraction and EEG Signal Segments Classification

Aleš Procházka and Jaromír Kukal
Institute of Chemical Technology in Prague
Department of Computing and Control Engineering
Technicka Street 5, 166 28 Prague 6, Czech Republic
Phone: +420 220 444 198 * Fax: +420 220 445 053
E-mail: A.Prochazka@ieee.org * Web: http://dsp.vscht.cz

Oldřich Vyšata
Neurocenter Caregroup
Jiráskova 1389, 516 01 Rychnov nad Kněžnou
Czech Republic
Phone: +420 494 531 544 * Fax: +420 494 531 544
E-mail: vysata@neurol.cz * Web: http://dsp.vscht.cz

Abstract—**Segmentation, feature extraction and classification of signal components belong to very common problems in various engineering, economical and biomedical applications. The paper is devoted to the use of discrete wavelet transform (DWT) both for signal preprocessing and signal segments feature extraction as an alternative to the commonly used discrete Fourier transform (DFT). Feature vectors belonging to separate signal segments are then classified by a competitive neural network as one of methods of cluster analysis and processing. The paper provides a comparison of classification results using different methods of feature extraction most appropriate for EEG signal components detection. Problems of multichannel segmentation are mentioned in this connection as well.**

Index Terms—**Segmentation, change-point detection, feature extraction, classification, multichannel signal processing, discrete wavelet transform, neural networks**

I. INTRODUCTION

The preliminary stage of multichannel signal classification includes its *segmentation* using principal component analysis and signal de-noising [1] in many cases followed by different methods of change-points detection [2], [3]. This stage is fundamental for further data processing both for one-dimensional and multi-dimensional signals.

Signal segments *feature extraction* forms the next step of signal segmentation allowing combination of time-domain and frequency-domain signal features. Commonly used spectral representation of a signal based upon its all-pole model or its discrete Fourier transform provides the same frequency resolution over the whole window function. To allow different resolution the wavelet transform [4], [5], [6] is often used providing its very efficient alternative allowing different levels of decomposition. The basic principle and application of wavelet transform is described in the first part of the contribution resulting in the given signal wavelet feature extraction and feature vector definition.

The task of signal segments *classification* forms another problem that can be solved by neural networks [7], [8] in many cases. The paper presents wavelet signal features classification by self-organizing neural networks and it mentions a possible compression of signal features as well. The method presented in the paper is applied for an EEG signal analysis and its segments classification into the proposed number of classes.

II. EEG SIGNAL PREPROCESSING AND SEGMENTATION

Information content of EEG signals is essential for detection of many problems of the brain and in connection with analysis of magnetic resonance images it forms one of the most complex diagnostic tools. To extract the most important properties of EEG observations it is necessary to use efficient mathematical tools [9], [10] to enable reliable and fast enough processing of very extensive data sets in most cases.

Digital filters can be used in the initial stage of EEG data processing to remove power frequency from the observed signal and to reduce its undesirable frequency components. Fig. 1 presents a sample of a selected EEG channel comparing results of its segmentation by an expert and by a selected Bayesian method [2] detecting changes of its mean value and variation. This approach has been used in this case for a selected channel only even though further channels must be taken into account in the real case as well.

III. WAVELET ANALYSIS AND SIGNAL FEATURE EXTRACTION

Wavelet transform forms a general mathematical tool for signal processing with many applications in EEG data analysis [11], [12], [13], [14], [15] as well. Its basic use includes time-scale signal analysis, signal decomposition and signal compression.
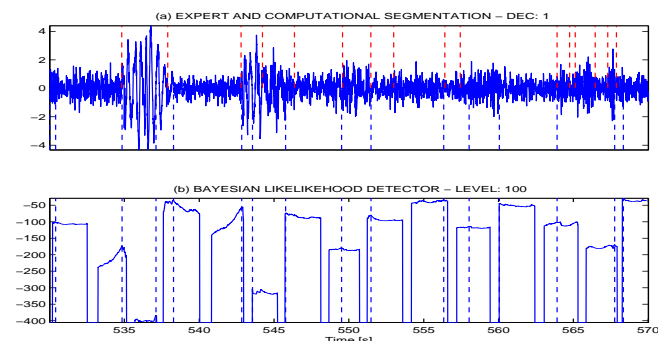

Fig. 1. Results of EEG signal segmentation presenting **(a)** signal segmentation by an expert (red) and by a Bayesian method of change-point detection (blue) and **(b)** a probability function to detect change-points

The set of wavelet functions is usually derived from the initial (mother) wavelet $h(t)$ which is dilated by value $a = 2^m$, translated by constant $b = k\,2^m$ and normalized so that

$$h_{m,k}(t) = \frac{1}{\sqrt{a}}\,h(\frac{t-b}{a}) = \frac{1}{\sqrt{2^m}}\,h(2^{-m}\,t-k) \qquad (1)$$

for integer values of m, k and the initial wavelet defined either by the solution of a dilation equation or by an analytical expression [4], [5]. Both continuous or discrete signals can be then approximated in the way similar to Fourier series and discrete Fourier transform. In case of a sequence $\{x(n)\}_{n=0}^{N-1}$ having $N = 2^s$ values it is possible to evaluate its expansion

$$x(n) = a_0 + \sum_{m=0}^{s-1}\sum_{k=0}^{2^{s-m-1}-1} a_{2^{s-m-1}+k}\,h(2^{-m}\,n-k)$$

Wavelet transform coefficients can be organized in a matrix $\mathbf{T}$ with its nonzero elements forming a triangle structure

$$\begin{bmatrix} a_{2^{s-1}} & a_{2^{s-1}+1} & & \cdots & & a_{2^s-1} \\ a_{2^{s-2}} & & & \cdots & & a_{2^{s-1}-1} \\ & a_4 & a_5 & a_6 & a_7 & \\ & a_2 & & & a_3 & \\ & & a_1 & & & \\ & & a_0 & & & \end{bmatrix}$$

with each its row corresponding to a separate dilation coefficient m. The set of $N = 2^s$ decomposition coefficients $\{a(j)\}_{j=0}^{N-1}$ of the wavelet transform is defined in the way formally close to the Fourier transform but owing to the general definition of wavelet functions they can carry different information. Using the orthogonal set of wavelet functions they are moreover closely related to the signal energy [5].

The initial wavelet can be considered as a pass-band filter and in most cases half-band filter covering the normalized frequency band $\langle 0.25, 0.5\rangle$. A wavelet dilation by the factor $a = 2^m$ corresponds to a pass-band compression. This general property can be demonstrated for the harmonic wavelet function [5] and the corresponding scaling function by expressions

$$h(t) = \frac{1}{j\pi/2\,t}(e^{j\pi\,t} - e^{j\pi/2\,t}) \qquad (2)$$

$$l(t) = \frac{1}{\pi/2\,t}(e^{j\pi/2\,t} - 1) \qquad (3)$$

As both these functions are modified by the scaling index $m = 0, 1, \cdots$ according to Eq. (1), the wavelet is dilated and its spectrum compressed resulting in time and frequency domain representation presented in Fig. 2. Similar approach can be also applied for other wavelet functions defined in either analytical or recurrent form.

The set of wavelets define a special filter bank which can be used for signal component analysis and resulting wavelet transform coefficients can be further applied as signal features for its classification. Signal decomposition performed by a pyramidal algorithm is interpreting wavelets as pass-band filters. Another approach [5] is based upon a very efficient parallel algorithm using the fast Fourier transform.

The basic decomposition of a given column vector $\{x(n)\}_{n=0}^{N-1}$ presented in Fig. 3 assumes a half-band low-pass scaling sequence

$$\{l(n)\}_{n=0}^{L-1} = [l(0), l(1), l(2), \cdots, l(L-1)]$$
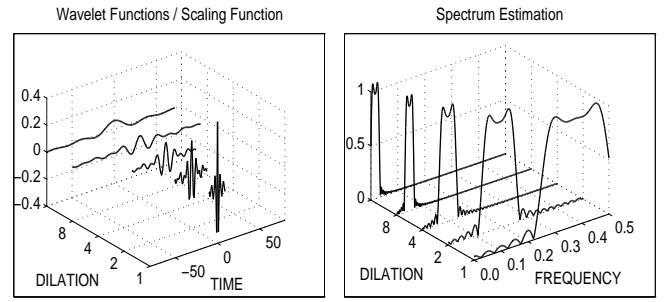
and the complementary orthogonal wavelet sequence

Fig. 2. The real part of the complex wavelet $h(t)=(e^{j\pi\,t}-e^{j\pi/2\,t})/(j\pi/2\,t)$ and the effect of its dilation to the corresponding spectrum compression

$$\{h(n)\}_{n=0}^{L-1} = [l(L-1), -l(L-2), l(L-3), \cdots, -l(0)]$$

Corresponding sequences convolution and subsampling by two provides values

$$p(n) = \sum_{k=0}^{L-1} l(k)\,x(n-k) = \sum_{j=n,n-1,\ldots}^{n-L+1} x(j)\,l(n-j)$$

$$q(n) = \sum_{k=0}^{L-1} h(k)\,x(n-k) = \sum_{j=n,n-1,\ldots}^{n-L+1} x(j)\,h(n-j)$$

for $n = L-1, L+1, ..., N-1$. Subsequent application of the same method provides signal analysis for different scales. Introducing decomposition matrices

$$\mathbf{L}_{N/2,N} = \begin{bmatrix} l(1) & l(0) & 0 & 0 & \cdots \\ l(3) & l(2) & l(1) & l(0) & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & l(L-1) & \cdots & l(0) \end{bmatrix}$$

$$\mathbf{H}_{N/2,N} = \begin{bmatrix} h(1) & h(0) & 0 & 0 & \cdots \\ h(3) & h(2) & h(1) & h(0) & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & h(L-1) & \cdots & h(0) \end{bmatrix}$$

it is possible to decompose the initial signal x into two sequences $\mathbf{p} = \mathbf{L}\,\mathbf{x}$ and $\mathbf{q} = \mathbf{H}\,\mathbf{x}$ standing for subsampled low-pass and high-pass signal components. The elements of vector $\mathbf{q}$ represent wavelet transform coefficients of the initial level. Further wavelet coefficients can be obtained after the application of this process to signal $\mathbf{p}$ according to Fig. 3. Resulting wavelet coefficients $\{a(n)\}_{n=0}^{N-1}$ related to chosen scales can then be used as signal features for its classification.
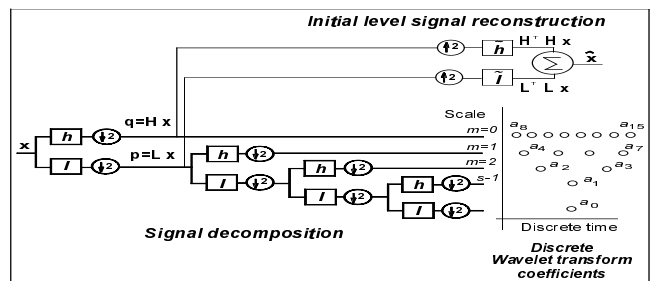

Fig. 3. A pyramidal filter bank structure used to evaluate wavelet transform coefficients for a sequence $\{x(n)\}_{n=0}^{N-1}$ and for values $s=4$ and $N=2^s=16$
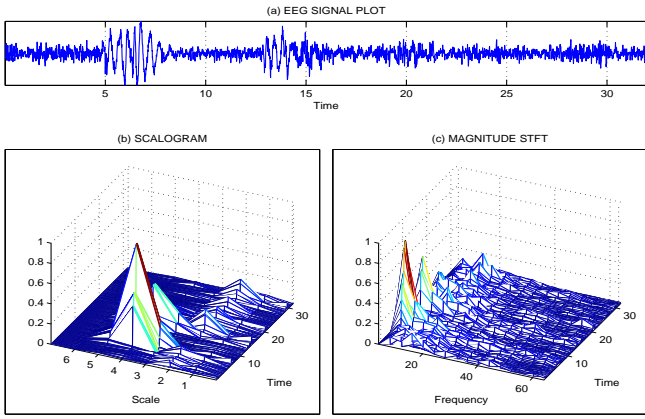
Fig. 4. Presentation of **(a)** a part of a one channel EEG signal plot and the result of its analysis **(b)** by the wavelet transform based on harmonic wavelet functions and **(c)** by the short-time Fourier transform

Fig. 4 presents the scalogram and spectrogram of the selected part of the EEG signal comparing results achieved by the DWT and DFT. It is obvious that owing to the principle of the wavelet transform short time signal components can be better detected and more precisely localized by the DWT comparing to results obtained by the DFT.

IV. SIGNAL SEGMENTS CLASSIFICATION

Classification of signal segments into a given number of classes using segments features can be achieved by various statistical methods. Another approach presented further is based upon the application of self-organizing neural networks.

A. Principle

Two typical principles of signal classification are presented in Fig. 5. In the first case general self-organizing maps employ directly signal features as patterns for an input layer. The number of output layer elements is equal to signal classes and must be either defined in advance or it can be automatically increased to create new classes [8]. During the learning process neural network weights are changed to minimize distances between each input vector and corresponding weights of a winning neuron characterized by its coefficients closest to the current pattern. In case that the learning process is successfully completed network weights belonging to separate output elements represent typical class individuals.
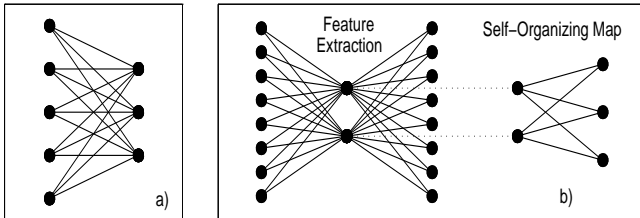

Fig. 5. Principles of signal classification using **(a)** signal features obtained separately and **(b)** signal features defined by a compression method

Another approach to signal classification is based on signal features compression at first to decrease number of patterns for self-organizing maps. The disadvantage of this algorithm in larger time consumption resulting from the application of two different subsequent methods can be minimized by a very fast Levenberg-Marqardt network optimization used to reduce the number of signal features to the number given by the size of the hidden layer. Further improvement can be achieved by a proper choice of signal segments for signal features evaluation.

B. Results

The algorithm presented above has been applied for classification of a real EEG signal. Fig. 6 presents EEG signal segments and their analysis by a harmonic wavelet transform resulting in features standing for scales 1, 2 and 3 respectively covering three frequency bands with different time-scale resolution. Results of signal classification into four classes by a self-organizing neural network are given in Fig. 7 and 8 for two selected signal features allowing a simple visualization of segmentation results. Class boundaries were evaluated and presented in this case as well. Both features clustering and time domain signal segments values show that signals of similar structure belong to the same class.

Results of signal classification into four classes for different features are compared in Tab. I. Each class $i = 1, 2, \cdots, C$ can be characterized by the mean distance of the column feature vector $\mathbf{p}_{j,i}$ belonging to its separate segments j for $j = 1, 2, \cdots, N_i$ from the centre $\mathbf{c}_i$ of individual classes using relation

$$crit = \frac{1}{C} \sum_{i=1}^{C} \frac{1}{N_i} \sum_{j=1}^{N_i} dist(\mathbf{p}_{j,i}, \mathbf{c}_i) \qquad (4)$$

where C stands for the number of classes, N_i represents the number of segments belonging to class i and function $dist$ is used for evaluation of the Euclidean distance between two
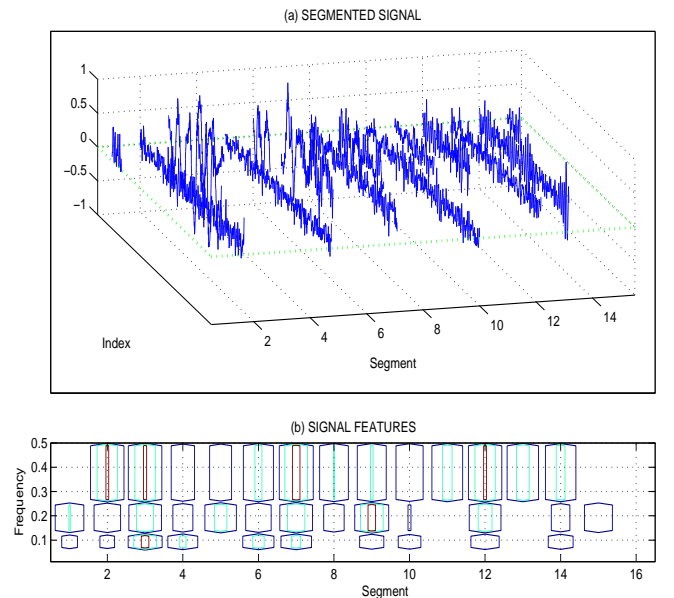

Fig. 6. Results of feature extraction presenting **(a)** EEG signal segments and **(b)** their wavelet features resulting from a harmonic DWT on scales 1, 2, 3

(a) SIGNAL SEGMENTS CLASSIFICATION

(b) SIGNAL CLASSIFICATION

(c) CLASS A   CLASS B   CLASS C   CLASS D

Fig. 7. Results of signal classification into four classes presenting **(a)** given signal, **(b)** the topology of signal segment features with class boundaries, and **(c)** signal segments divided into four classes and typical signal segments closest to the final neuron values at the bottom of each class

CLASS A: 0.043   CLASS B: 0.017   CLASS C: 0.008   CLASS D: 0.039

Fig. 8. The topography of the two EEG signal segments features for their classification into four classes with typical segment features at the bottom

vectors. It is obvious that classification parameters achieved both by the DFT and DWT provide similar results but features of the dominant cluster are closer together in the case of the DWT use in most cases.

V. CONCLUSION

The contribution presents the use of wavelet transform for a given signal classification. Mathematical basis of the DWT and the following numerical experiments proved that signal analysis based on DWT coefficients can be used very efficiently for the estimation of signal segments features. The paper presents results of a classification for signal segments detected by the the Bayesian approach [2] for a selected EEG channel. The following signal classification assumed the knowledge of the range of the number of classes to apply a self-creating classification method [8] to find their optimal value and to exclude the possibility of dead neurons. Results of signal classification can be further improved by various methods but one the most important problems is in the right definition of signal segment features using both its frequency-domain and time-domain properties and it seems that the DWT can be used in this area very efficiently. Further studies will be also devoted to the use of principal component analysis to find fundamental signal components from all observed channels for the following global segmentation.

REFERENCES

[1] S. V. Vaseghi, *Advanced Digital Signal Processing and Noise Reduction*, John Wiley & Sons Ltd, West Sussex, 2006.

[2] W. J. Fitzgerald, J. J. K. O. Ruanaidh, and J. A. Yates, "Generalised Changepoint Detection," Tech. Rep., University of Cambridge Engineering Department, Cambridge, U.K., 1994.

[3] B. E. Brodsky and B. S. Darkhovski, *Nonparametric Methods in Change-Point Problems*, Kluwer Academic Publishers, Boston, U.S.A., 1993.

[4] I. Daubechies, "The Wavelet Transform, Time-Frequency Localization and Signal Analysis," *IEEE Trans. Inform. Theory*, vol. 36, pp. 961–1005, September 1990.

[5] D. E. Newland, *An Introduction to Random Vibrations, Spectral and Wavelet Analysis*, Longman Scientific & Technical, Essex, U.K., third edition, 1994.

[6] A. Procházka, J. Jech, and J. Smith, "Wavelet transform use in signal processing," in *31st International Conference in Acoustics*. 1994, pp. 209–213, Czech Technical University.

[7] S. Haykin, *Neural Networks, A Comprehensive Foundation*, Macmillan College Publishing Company, New York, 1994.

[8] D. I. Choi and S. H. Park, "Self-Creating and Organizing Neural Networks," *IEEE Trans. Neural Networks*, vol. 5, no. 4, pp. 561–575, July 1994.

[9] Saeid Sanei and J. A. Chambers, *EEG Signal Processing*, Wiley - Interscience, 2007.

[10] M. Nixon and A. Aguado, *Feature Extraction & Image Processing*, Elsevier, Amsterdam, 2004.

[11] A. Glavinovitch, M. N. S. Swamy, and E. I. Plotkin, "Wavelet-Based Segmentation Techniques in the Detection of Microarousals in the Sleep EEG," in *48th Midwest Symposium on Circuits and Systems*. 2005, pp. 1302–1305, IEEE.

[12] P. Johankhani, V. Kodogiannis, and K. Revett, "EEG Signal Classification Using Wavelet Feature Extraction and Neural Networks ," in *IEEE John Vincent Atanasoff 2006 International Symposium on Modern Computing (JVA06)*. 2006, pp. 120–124, IEEE.

[13] C. Dimoulas, G. Kalliris, G. Papanikolaou, and A. Kalampakas, "Long-Term Signal Detection, Segmentation and Summarization Using Wavelets and Fractal Dimension: A Bioacoustics Application in Gastrointestinal-Motility Monitoring," *Comput. Biol. Med.*, vol. 37, no. 4, pp. 438–462, 2007.

[14] I. W. Selesnick, R. G. Baraniuk, and N. G. Kingsbury, "The Dual-Tree Complex Wavelet Transform," *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 123–151, 2005.

[15] Nazareth P. Castellanos and Valeri A. Makarov, "Recovering EEG Brain Signals: Artifact Suppression with Wavelet Enhanced Independent Component Analysis," *Journal of Neuroscience Met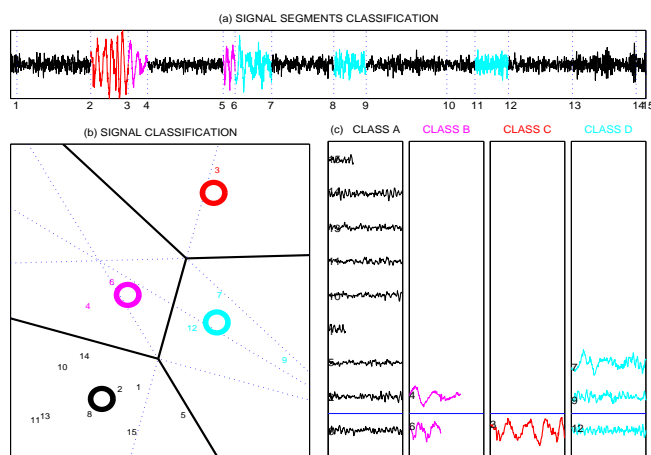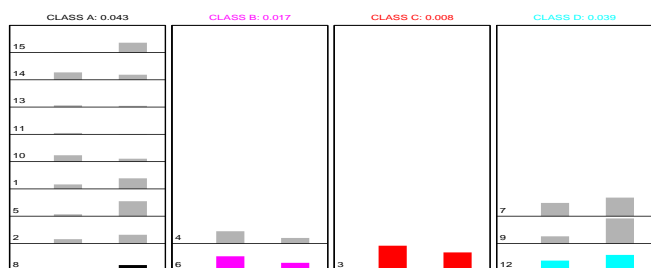hods*, vol. 158, no. 2, pp. 300–312, 2006.