WAVELET TRANSFORM USE FOR SIGNAL CLASSIFICATION BY SELF-ORGANIZING NEURAL NETWORKS

A. Procházka, M. Štorek

Prague Institute of Chemical Technology, Department of Computing and Control Engineering
Czech Republic

Abstract. Detection and classification of signal components belongs to very common problems in various engineering, economical and biomedical applications. To recognize groups of similar input vectors it is possible to use various methods of cluster analysis and mathematical models of neural networks studied in the paper as well. Characteristic signal features forming network patterns during the learning stage can be based either upon signal segments identification and modelling or its frequency components analysis. The paper is devoted to the use of the discrete wavelet transform (DWT) for their evaluation providing an alternative to the commonly used discrete Fourier transform (DFT). Fundamentals of wavelet analysis and signal decomposition with different resolution both in the time and frequency domains are presented at first. Resulting signal features are then used for signal classification. The method is verified for simulated signals and then applied for a given encephalogram (EEG) signal classification.

INTRODUCTION

Signal classification assumes *signal segments* definition at first. These segments can be of the same length or they can be defined by indices pointing to signal feature changes. There are various methods used to solve this problem described in literature i.e. by Godsill and Rayner (1), Fitzgerald (2), Schlang and Tresp (3) or Brodsky and Darkhovski (4). The paper assumes segments of a chosen length but the same algorithm of signal classification can be used for more precisely defined segments as well.

The following stage results in *feature extraction*. The paper is restricted to signals characterized by their frequency components used i.e. by Juang and Rabiner (5). 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. As the detection of short time signal components is complicated in some cases other methods of signal analysis can be applied. The wavelet transform studied recently among many others by Daubechies (6), Rioul and Vetterli (7), Fliege (8), Newland (9), Procházka (10), (11), (12) provides a very efficient alternative allowing different frequency resolution for different levels of decomposition. The basic principle and application of wavelet transform is described in the first part of the contribution resulting in a given signal wavelet feature extraction.

Classification of signal segments forms another problem of signal analysis studied by many authors including Wu and Fallside (13), Kindermann and Windheuser (14), Zhan and Fu (15), Wann and Thomopulos (16) and Choi and Park (17). This problem solution is based in many cases upon the application of neural networks analyzed recently i.e. by Haykin (18) or Karayinnis and Venetsanopoulos (19). 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 a given number of classes as an alternative to various other methods described i.e. by Elo and Saarinen (20), Roberts and Tarassenko (21) or Brückner, Franz and Richter (22).

WAVELET ANALYSIS

Wavelet functions and their applications in signal processing are described in various papers published recently. Their basic use include time-scale signal analysis, signal decomposition, signal compression etc. The following paper presents their possible use in signal classification.

Principle

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

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

where m and k are integers. Initial wavelet can be defined either the solution of a dilation equation described by Daubechies (6) or in an analytical form used both by Daubechies (6) and Newland (9). 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 in the form

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

Wavelet transform coefficients can be organized in a matrix $\mathbf{T}$ with its nonzero elements in the form

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

and each its row standing for a separate dilation coefficient m corresponding to the pyramidal evaluation process presented further. 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 as described in Newland (9).

Using the digital signal processing point of view it is possible to consider the initial wavelet as a pass-band filter and in most cases half-band filter covering 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 introduced by Newland (9) and defined together with a 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 according to Fig. 1. Similar approach can be also applied for other wavelet functions defined in either analytical or recurrent form.

Signal Features Extraction

Wavelets presented above 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 described by Rioul and Vetterli (7) and performed by a pyramidal algorithm is interpreting wavelets as pass-band filters. Another approach introduced by Newland (9) is based upon a very efficient parallel algorithm using the fast Fourier transform.

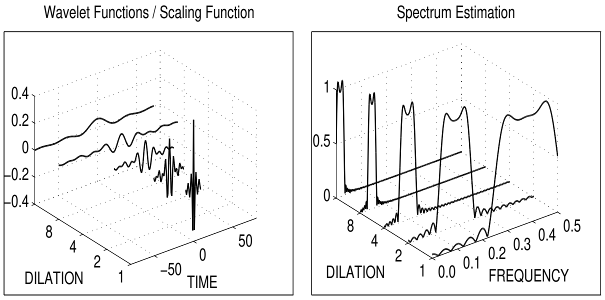

Figure 1: The real part of the complex harmonic initial wavelet $h(t) = (e^{j\pi\ t} - e^{j\pi/2\ t})/(j\pi/2\ t)$ and the effect of its dilation by value $a = 2^m$ for $m = 0, 1, 2, 3$ to the corresponding spectrum compression

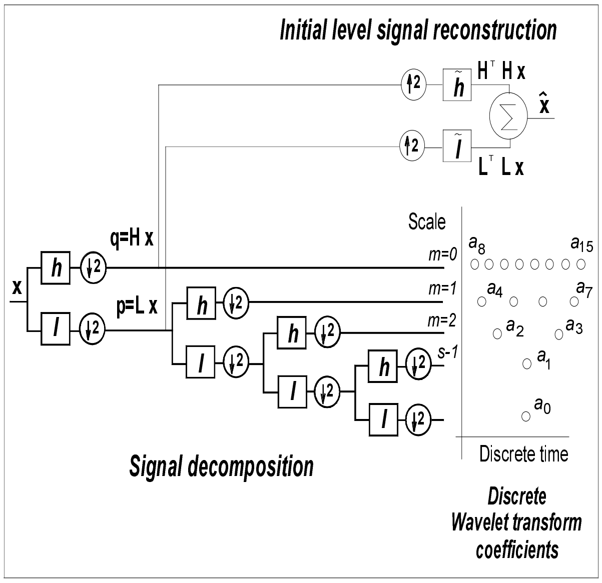

Figure 2: A pyramidal filter bank structure used to evaluate wavelet transform coefficients for a given sequence $\{x(n)\}_{n=0}^{N-1}$ and final coefficients presentation for $s = 4$ and $N = 2^s = 16$

The basic decomposition of a given column vector $\{x(n)\}_{n=0}^{N-1}$ presented in Fig. 2 assumes a half-band low-pass scaling sequence

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

of an even length L and the complementary orthogonal wavelet sequence

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

Corresponding sequences convolution and subsampling by two provides values

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

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

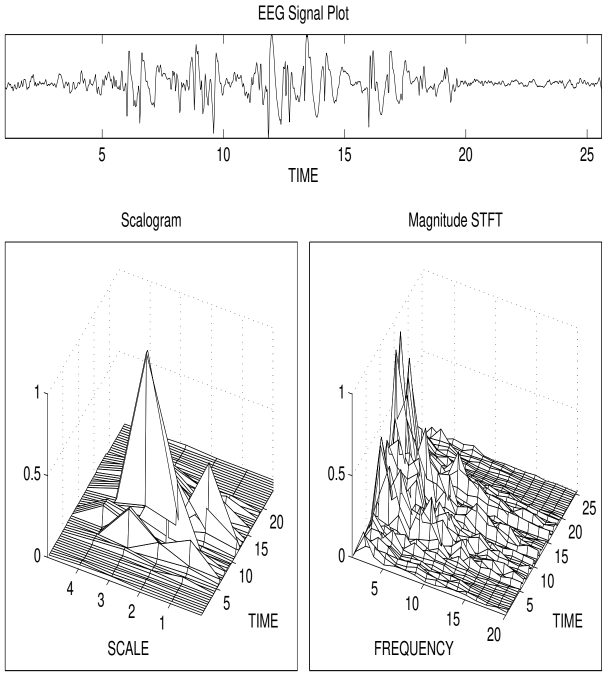

Figure 3: Part of a one channel EEG signal plot and the result of its analysis by the wavelet transform based on harmonic wavelet functions and by the short-time Fourier transform

for $n = L-1, L+1, ..., N-1$. Subsequent application of the same method provides signal analysis on 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 define one stage of the separation process resulting in signals $\mathbf{p} = \mathbf{L}\ \mathbf{x}$ and $\mathbf{q} = \mathbf{H}\ \mathbf{x}$ standing for subsampled low-pass and high-pass signal components. The elements of vector $\mathbf{q}$ moreover represent wavelet transform coefficients of the initial level. Further wavelet coefficients can be obtained after the application of the same process to signal $\mathbf{p}$ according to Fig. 2. In case of orthogonal wavelet functions values of these wavelet transform decomposition coefficients $\{a(n)\}_{n=0}^{N-1}$ are closely related to signal energy mentioned by Newland (9). Wavelet coefficients belonging to chosen scales can be further used as signal features for its classification.

A given EEG signal scalogram and spectrogram is presented in Fig. 3. 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 FFT.

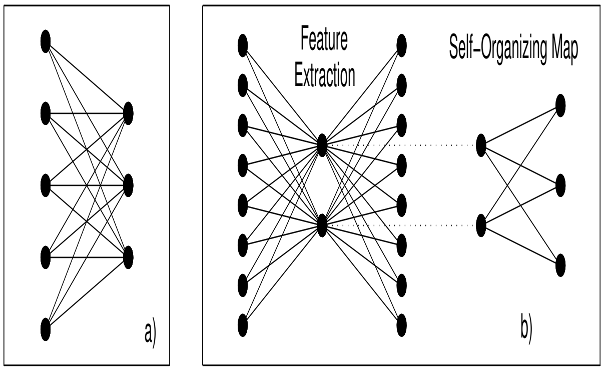

Figure 4: Basic principles of signal classification (**a**) Using signal features as a self-organizing map inputs (**b**) Using signal features compression by a three layer neural network

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.

Principle

Two typical principles of signal classification are presented in Fig. 4. 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 by a special method recommended by Choi and Park (17). 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.

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. The main features of this algorithm were described by Hagan and Menhaj (23).

Further improvement can be achieved by a proper choice of signal segments for signal features evaluation.

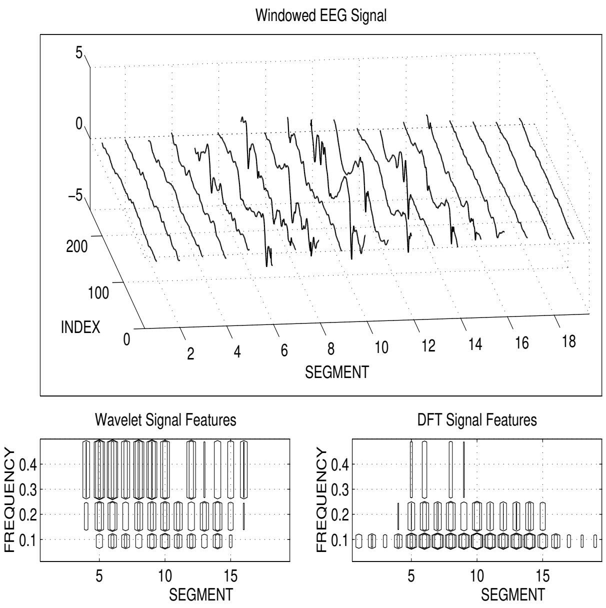

Figure 5: Windowed EEG signal and its wavelet features resulting from a harmonic DWT on scales $1, 2$ and 3 in comparison with DFT features resulting from the same frequency bands

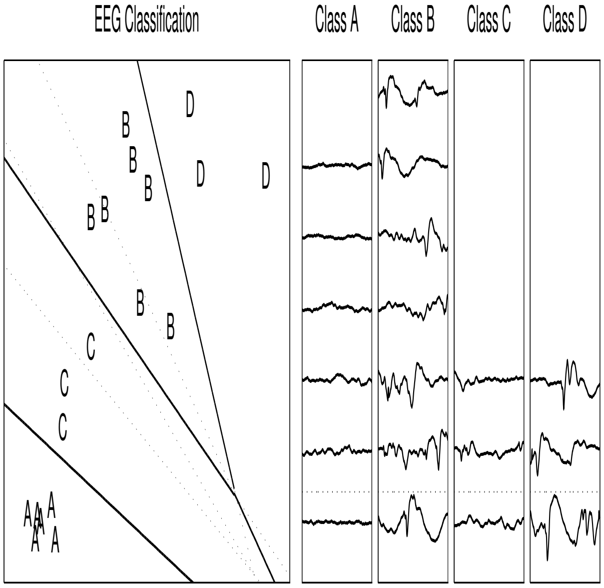

Figure 6: The topography of the two selected EEG signal segments features and results of signal classification into four classes providing class boundaries and presenting typical signal segments closest to the final neuron values under dotted lines

TABLE 1: Comparison of signal segments classification into four classes using two features resulting from a chosen signal segments analysis.

Feature Source	*Class Mean Square Errors*			
	A	*B*	*C*	*D*
DFT	0.066	0.126	0.044	0.186
Harmonic DWT	0.057	0.064	0.074	0.135
D4 DWT	0.104	0.085	0.067	0.174
D16 DWT	0.103	0.123	0.106	0.165

Results

The method presented above has been verified for artificial signals at first and then applied for a real signal classification. Fig. 5 presents EEG signal segments analysis by a harmonic wavelet transform and resulting features standing for scales 1, 2 and 3 respectively covering three frequency bands. These features are compared with that obtained by the DFT for corresponding frequency components. It is obvious that the DWT provides more information owing to its different time-scale resolution. Results of signal classification into four classes by a self-organizing neural network are given in Fig. 6 for two selected signal features allowing in this case class boundaries presentation as well. Both features clustering and time domain signal segments values show that signals of similar structure belong to the same class. Signal classification is affected by the chosen window length in this case.

Results of signal classification into four classes for different features are compared in Tab. 1. Each class is characterized by the variation of signal features distances from a typical class element. It is obvious that corresponding features obtained by the DFT and DWT both for harmonic and Daubechies wavelets of order 4 (D4) and 16 (D16) provide similar results but for the lower number of operations in the wavelet case.

CONCLUSION

The novelty of the contribution is in the use of wavelet transform for a given signal classification. Mathematical basis of the DWT and the following numerical experiments proved that signal features based on wavelet transform coefficients can be used very efficiently for short time signal peaks detection and classification. The paper presents results of such a classification for signal segments of the same length. Better results can be achieved in case of more precise signal segmentation using various methods of change-point detection and namely the Bayesian principle presented i.e. by Fitzgerald (2). All these methods are very efficient but they assume the preliminary knowledge of a signal component used for segmentation. Signal classification used in the paper assumes a given number of classes. Not knowing the number of output classes for signal classification in advance it is moreover possible to apply a self-creating classification mentioned by Choi and Park

(17) excluding the possibility of the dead neuron introduction as well. Results of signal classification can be improved by various methods but further very important step is in the right interpretation of typical segment features.

ACKNOWLEDGEMENT

The paper has been supported by a grant No. 201/94/013 of the Grant Agency of Czech Republic. All EEG signals were kindly provided by the Neurological Department of the Bulovka Hospital in Prague.

REFERENCES

1. Godsill S. J. and Rayner P. J. W., 1992, "A Bayesian Approach to the Detection and Correction of Error Burst in Audio Signals", IEEE, in ICASSP Proceedings 1992, II/261–II/264
2. Fitzgerald W. J., Ruanaidh J. J. K. O. and Yates J. A., 1994, "Generalised Changepoint Detection", Tech. Rep., University of Cambridge Engineering Department, Cambridge, U.K.
3. Schlang M. and Tresp V., 1992, "Neural Networks for Segmentation and Clustering of Biomagnetical signals", in Neural Networks on Signal Processing II, S. Y. Kung, F. Fallside, and J. A. Sorenson, Eds., IEEE, Proceedings of the 1992 IEEE Workshop, 343–349
4. Brodsky B. E. and Darkhovski B. S., 1993, "Nonparametric Methods in Change-Point Problems", Kluwer Academic Publishers, Boston, U.S.A.
5. Juang B. H. and Rabiner L. R., 1992, "Spectral Representation for Speach Recognition by Neural Networks - A Tutorial", in Neural Networks on Signal Processing II, S. Y. Kung, F. Fallside, and J. A. Sorenson, Eds., IEEE, Proceedings of the 1992 IEEE Workshop, 214–222
6. Daubechies I., 1990, "The Wavelet Transform, Time-Frequency Localization and Signal Analysis", IEEE Trans. Inform. Theory, 36, no. 9, 961–1005
7. Rioul O. and Vetterli M., 1991, "Wavelets and Signal Processing", IEEE SP Magazine, October 1991, 14–38
8. Fliege N. J., 1994, "Multirate Digital Signal Processing: Multirate Systems, Filter Banks, Wavelets", John Wiley & Sons Ltd, Sussex, U.K.
9. Newland D. E., 1994, "An Introduction to Random Vibrations, Spectral and Wavelet Analysis", Longman Scientific & Technical, Essex, U.K., Third edition
10. Procházka A. and Sýs V., 1994, "Time Series Prediction Using Genetically Trained Wavelet Networks", in Neural Networks for Signal Processing 3 - Proceedings of the 1994 Workshop, IEEE, 195–203
11. Procházka A. and Sýs V., 1994, "Application of Genetic Algorithms for Wavelet Networks Signal Modelling", in VIIth European Signal Processing Conference EUSIPCO-94, European Association for Signal Processing, II/1078–II/1081
12. Procházka A., Jech J., and Smith J., 1994, "Wavelet Transform Use in Signal Processing", in 31st International Conference in Acoustics, Czech Technical University, 209–213
13. Wu L. and Fallside F., 1992, "Pattern Classification with a Codebook-excited Neural Network", in Neural Networks in Signal Processing II, S. Y. Kung, F. Fallside, and J. A. Sorenson, Eds., IEEE, Proceedings of the 1992 IEEE Workshop, 223–232
14. Kindermann J. and Windheuser Ch., 1992, "Unsupervised Sequence Classification", in Neural Networks on Signal Processing II, S. Y. Kung, F. Fallside, and J. A. Sorenson, Eds., IEEE, Proceedings of the 1992 IEEE Workshop, 184–193
15. Zhan J. and Fu L., 1991, "A Self-Organization Neural Network and Detection of Signals", in International Conference on Neural Networks, IEEE, INNC, IV/748–IV/751
16. Wann Ch-D. and Thomopulos S., 1991, "Comparative Study of Self-Organization Neural Networks Models", in International Conference on Neural Networks, IEEE, INNC, II/549–II/552
17. Choi D. I. and Park S. H., 1994, "Self-Creating and Organizing Neural Networks", IEEE Trans. Neural Networks, 5, no. 4, 561–575
18. Haykin S., 1994, "Neural Networks, A Comprehensive Foundation", Macmillan College Publishing Company, New York
19. Karayiannis N. B. and Venetsanopoulos A. N., 1993, "Artificial Neural Networks, Learning Algorithms, Performance Evaluation and Applications", Kluver Academic Publishers, Massachusetts
20. Elo P., Saarinen J., Varri A., Nieminen H., and Kaski K., 1992, "Classification of Epileptic EEG Using Self-Organization Maps", in Artificial Neural Networks, I. Aleksander and J. Taylor, Eds., Elsevier Science Publishers, 1147–1150
21. Roberts S. and Tarassenko L., 1991, "EEG Analysis Using Self-organisation", in Second International Conference on Artificial Neural Networks, 210–213
22. Brückner B., Franz M., and Richter A., 1992, "A Modified Hypermap Architecture for Classification of Biological Signals", in Artificial Neural Networks 2, I. Aleksander and J. Taylor, Eds., Elsevier Science Publishers B.V., 1167–1170
23. Hagan M. T. and Menhaj M. B., 1994, "Training Feedforward Networks with the Marquardt Algorithm", IEEE Trans. Neural Networks, 5, no. 6, 989–993