Multi-Channel EEG Signal Segmentation and Feature Extraction

Aleš Procházka*, Martina Mudrová*, Oldřich Vyšata**, Robert Háva*, and Carmen Paz Suárez Araujo***
* Institute of Chemical Technology in Prague, Department of Computing and Control Engineering, Czech Republic
** Neurocenter Caregroup, Czech Republic, *** Universidad de Las Palmas de G.C., Spain
A.Prochazka@ieee.org, Martina.Mudrova@vscht.cz, vysata@neurol.cz, Robert.Hava@vscht.cz, cpsuarez@dis.ulpgc.es

***Abstract*—Signal analysis of multi-channel data form a specific area of general digital signal processing methods. The paper is devoted to application of these methods for electroencephalogram (EEG) signal processing including signal de-noising, evaluation of its principal components and segmentation based upon feature detection both by the discrete wavelet transform (DWT) and discrete Fourier transform (DFT). The self-organizing neural networks are then used for pattern vectors classification using a specific statistical criterion proposed to evaluate distances of individual feature vector values from corresponding cluster centers. Results achieved are compared for different data sets and selected mathematical methods to detect and to classify signal segments features. Proposed methods are accompanied by the appropriate graphical user interface (GUI) designed in the MATLAB environment.**

I. Introduction

Processing of parallel time series resulting from multi-sensor observation of engineering or biomedical systems form an important area of general digital signal processing methods. Mathematical tools used for their analysis include discrete Fourier transform, wavelet transform and de-noising methods [1], [2], [3] followed by methods for signal segmentation, feature extraction and classification in some cases.

The paper presents the use of the double moving window for signal segmentation and its application for multi-channel signal segmentation analysing its first principal component. Feature vectors [4], [5] of signal segments evaluated by the wavelet transform are then compared with those obtained by the discrete Fourier transform and the most compact clusters are then classified by self-organizing neural networks.

Fig. 1 presents a selected multi-channel EEG signal observed with additive noise component of 50 Hz that can be digitally rejected by different kinds of digital filters either in the time or frequency domains. Results obtained are presented on Fig. 2. Fig. 3 maps the typical localization of EEG electrodes with interpolation of brain activity for three successive time instants.

The EEG signal processing [6], [7], [8], [9], [10], [11] is fundamental for analysis of the brain activity and in connection with magnetic resonance methods and magnetoencephalography [12] it forms one of the most complex diagnostical tools. Both the space localization of brain activities and its time evolution is important for diagnostical purposes which can be moreover correlated with MR observations in the next stage.

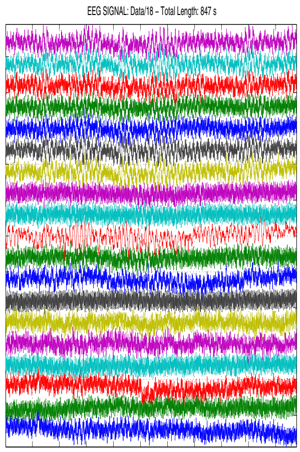

Fig. 1. EEG signal record presenting 19 channels with the additive noise component of 50 Hz

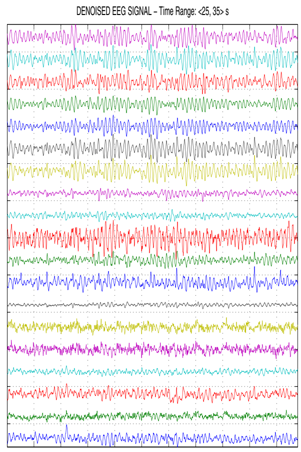

Fig. 2. Processed EEG signal after the digital rejection of signal noise components in all its channels

Change-point detection methods applied to given time series can be based upon Bayesian methods [13] detecting changes of its statistical properties or it is possible to detect changes in signal frequency components. Owing to the multi-channel basis of EEG signals this method has been applied after the principal component analysis (PCA) of an observed and preprocessed multi-channel signal to its first principal component. A special graphical user interface has been proposed and developed in this connection to enable more convenient way of real signal processing.

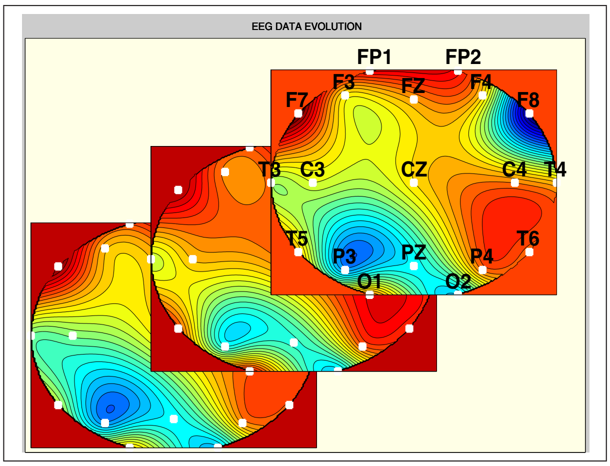

Fig. 3. The evolution of EEG energy distribution

II. Signal De-noising

Signal de-nosing is very important in EEG signal analysis and it forms its fundamental step eliminating the necessity to observe brain activities in special rooms free of disturbing signals. Both wavelet thresholding methods or time domain signal processing can be used. Fig. 4 presents a typical compact algorithm for such a processing of an EEG data segment observed with the sampling frequency $F_s = 200\ Hz$ and bi-directional filtering.

```
%% EEG Data Filtering %%%
   b1=fir1(L,[48 52]/Fs,'stop')
   b2=fir1(L,[0.5 60]/Fs,'bandpass')
   E1=filtfilt(b1,1,EEG);
   E=filtfilt(b2,1,E1));
```

Fig. 4. Algorithm of multi-channel time-domain filtering of an EEG signal (with the sampling frequency F_s) using FIR filter of order L

Results of the rejection of the net frequency of 50 Hz and frequency components outside the frequency band of $\langle 0.5, 60 \rangle\ Hz$ for a selected signal segment are presented in Fig. 5.

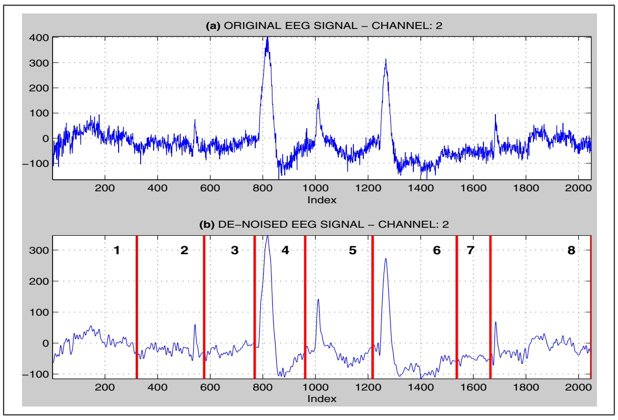

Fig. 5. The EEG signal segment presenting **(a)** original signal and **(b)** de-noised and segmented signal

III. Principal Component Analysis

As the segmentation of multi-channel signals should be performed over all channels in most cases it is important to extract information present in all parallel time series at first. The principal component analysis [5] can be used to perform this task using results of linear algebra to reduce the dimension of a matrix $\mathbf{E}_{N,M}$ containing multi-channel data. In the case of EEG records each its column $j = 1, 2, \cdots, M$ represents channel index and its row $i = 1, 2, \cdots, N$ stand for observations. It is possible to transform this matrix into a new one using matrix $\mathbf{P}_{M,M}$ to find values

$$\mathbf{Y}_{N,M} = \mathbf{E}_{N,M}\ \mathbf{P}_{M,M} \tag{1}$$

In the case of principal component analysis matrix $\mathbf{P}_{M,M}$ is orthonormal evaluating columns of matrix $\mathbf{Y}_{N,M}$ with the decreasing variance using MATLAB commands presented in Fig. 6 in the simplest case.

```
%% Principal Component Analysis %%%
   P = princomp(E);
   Y=E*P;
```

Fig. 6. Algorithm of the PCA applied to matrix $\mathbf{E}_{N,M}$

Results of PCA for a selected segment of EEG observations after their de-noising is presented in Fig. 7 together with resulting signal variances.

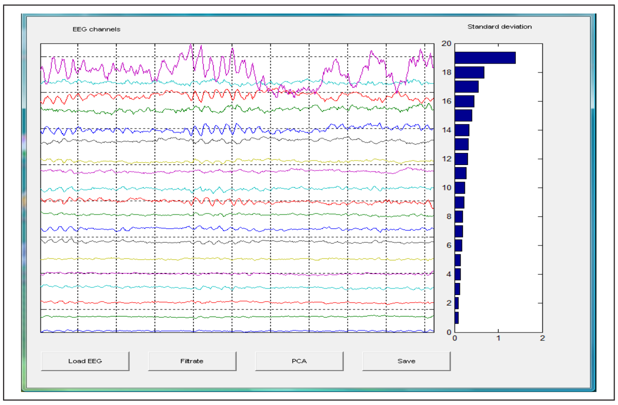

Fig. 7. Graphical user interface for principal component analysis of observed EEG channels and their variances

IV. Signal Segmentation

The proposed method of signal segmentation is based upon the two sliding overlapping windows and the detection of signal properties changes. Owing to the properties of the data source changes of average frequency components in selected frequency bands were chosen for EEG signal segmentation. Suggested algorithm combines raw and fine changes in window positions to detect changes precisely enough and to reduce computational time for long time series processing. Results for a selected signal are presented in Fig. 5.

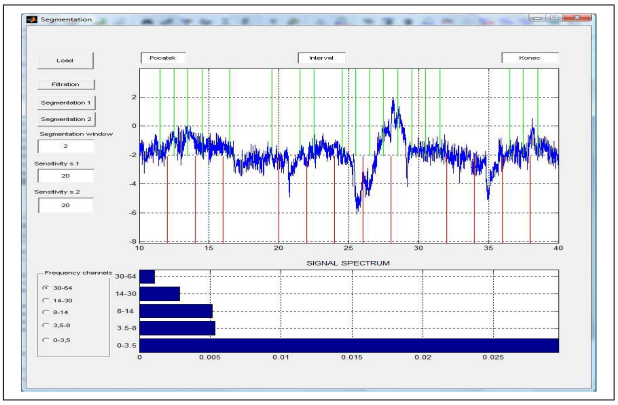

Fig. 8. The graphical user interface for segmentation of the EEG signal using the average signal energy in given frequency bands

Owing to the necessity of multi-channel signal processing the first principal component has been further used for segmentation of the whole set of observed time-series. Fig. 8 presents the proposed graphical user interface designed to find signal segments of similar properties in the frequency domain.

V. Feature Extraction

The selection of the most efficient and reliable method of feature extraction forms a very important problem of signal segments classification. Methods applied are usually based upon the time-domain or frequency-domain signal analysis. The following study is devoted to the wavelet domain signal feature extraction and comparison of results achieved.

The discrete wavelet transform (DWT) forms a general mathematical tool for signal processing with many applications in data processing [14], [1] using time-scale signal analysis, signal decomposition and signal reconstruction. The set of wavelet functions is usually derived from the initial time-limited (mother) wavelet $h(t)$ which is dilated by value $a = 2^m$, translated by constant $b = k\ 2^m$ and normalized so that

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

for integer values of m, k and the initial wavelet defined either by the solution of a dilation equation or by an analytical expression [15]. The initial wavelet can be considered as a pass-band filter and in most cases half-band filter covering the normalized frequency band $\langle 0.25, 0.5)$. A wavelet dilation by the factor $a = 2^m$ corresponds to a pass-band compression. The set of wavelets define a special filter bank in this way.

Suggested algorithm is based upon the wavelet decomposition of signal segments and evaluation of its coefficients for estimation of segment features. Fig. 9 presents application of this method to EEG signal segments and their analysis by a harmonic wavelet transform [15] resulting in features standing for scales 1, 2 and 3 respectively covering three frequency bands.

The discrete wavelet transform enables estimation of signal segment features with changing resolution in time and frequency for different scales and in this way it is possible to obtain the complex description of data segments.

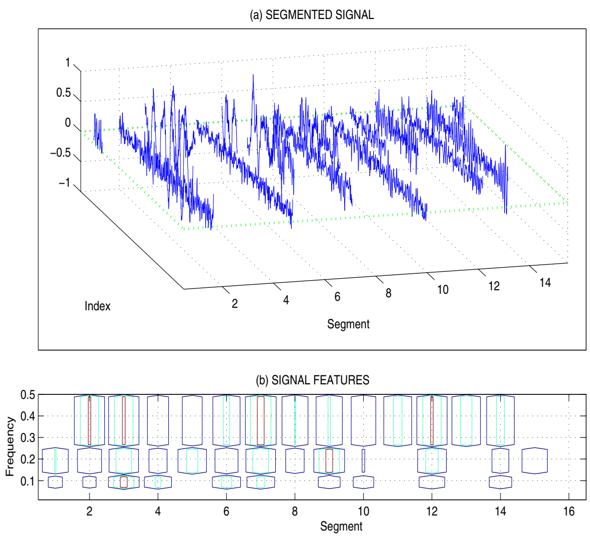

Fig. 9. Feature extraction presenting **(a)** EEG signal segments and **(b)** their DWT features on scales $1, 2, 3$

VI. Classification

Each signal segment can be described by R features specified in the pattern matrix $\mathbf{P}_{R,Q}$ and forming clusters in the R-dimensional space. The proposed algorithm for their classification has been based upon the application of self-organizing neural networks [16], [17] using Q feature vectors as patterns for the input layer of neural network. The number S 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 [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. Fig. 10 presents basic steps of this process with results of classification of $R = 2$ features to $S = 3$ classes presented in Fig. 11.

```
%% Pattern values classification %%%
   net = newc(minmax(P),S);
   net = train(net,P);
   W1=net.IW{1,1};
   A =sim(net,P); Ac=vec2ind(A);
```

Fig. 10. Algorithm for the classification of pattern matrix P into S classes and optimization of its weights $\mathbf{W1}$

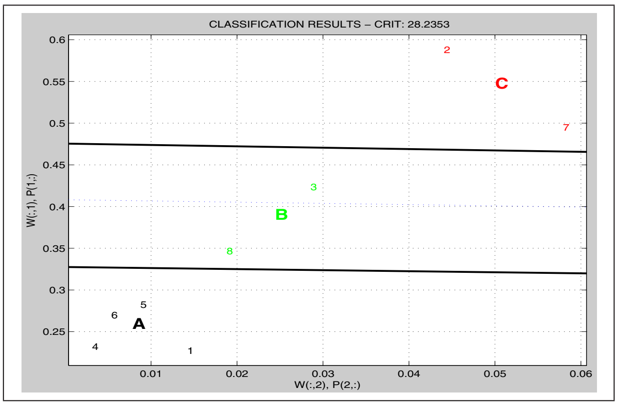

Fig. 11. Visualization of 2 element feature vectors and results of their classification into $S = 3$ classes

The graphical user interface designed for this process is presented in Fig. 12 for classification into four classes by a self-organizing neural network for two selected signal features allowing a simple visualization of segmentation results and visualization of typical class representatives with their features closest to the corresponding cluster centers as well.

To compare results of classification of Q signal segments with feature matrix $\mathbf{P}_{R,Q} = [\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_Q]$ for the selection of different sets of $R = 2$ features and C classes a specific criterium has been designed. Each class $i = 1, \cdots, C$ has been characterized by the mean distance of column feature vectors $\mathbf{p}_{j_k}$ belonging to class segments j_k for indices $k = 1, 2, \cdots, N_i$ from the class centre in the $i-$th row of matrix $\mathbf{W}_{C,R} = [\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_C]'$ by

TABLE I
CLUSTER COMPACTNESS EVALUATION FOR SIGNAL SEGMENTS CLASSIFICATION INTO 3 CLASSES FOR DIFFERENT FEATURE EXTRACTION METHODS AND SELECTED THRESHOLD VALUES

Method	*Thresh1 (0.12)*	*Thresh2 (0.08)*	*Thresh3 (0.01)*
DFT	0.289	0.274	0.307
Haar DWT (level 1,2)	0.150	0.179	0.211
Haar DWT (level 2,3)	0.160	0.238	0.317

relation

$$ClassDist(i) = \frac{1}{N_i}\sum_{k=1}^{N_i} dist(\mathbf{p}_{j_k}, \mathbf{w}_i) \qquad (3)$$

The value N_i represents here the number of segments belonging to class i and function $dist$ is used for evaluation of the Euclidean distance between two vectors. Results of classification can then be numerically characterized by the mean value of average class distances related to the mean value of class centers distances obtained after the learning process according to relation

$$crit = mean(ClassDist)/mean(dist(\mathbf{W}, \mathbf{W}')) \qquad (4)$$

This proposed Cluster Segmentation Criterion (CSC) provides low values for compact and well separated clusters while close clusters with extensive dispersion of cluster vectors provide high values of this criterion.

Selected results of numerical experiments for different feature extraction methods are summarized in Tab. I presenting the criterial values for different signal segments. It is obvious that classification parameters achieved both by the DFT and DWT provide similar results but slightly better in the case of wavelet features selection. This result corresponds to visual evaluation of EEG segments as well.

VII. CONCLUSION

The paper presents selected aspects of multi-channel signal processing with application to EEG signal denoising, segmentation and feature extraction. The principal component analysis is mentioned in connection with segmentation of parallel signals to find joint change-points valid for all channels.

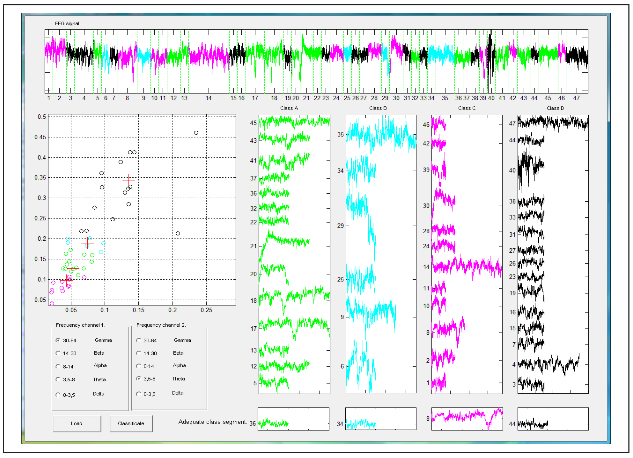

Fig. 12. Graphical user interface for signal segments classification into the given number of classes and detection of typical signal segments closest to cluster centers

A special attention is paid to comparison of the efficiency of feature extraction using signal segments properties estimated both by the FFT a DWT transforms. Cluster compactness is evaluated by the proposed criterial function. Signal components are then classified by the self-creating neural network structures [17] enabling to find the optimal value of classes and to exclude the possibility of dead neurons. The graphical user interface is designed for this study as well.

Further research will be devoted to further methods of signal preprocessing and to the estimation of signal segments features to form compact clusters enabling more reliable signal segments classification.

ACKNOWLEDGEMENT

The paper has been supported by the Research grant No.MSM 6046137306 of the Czech Ministry of Education and by the Science and Innovation Ministry of the Spanish Government under Research Projects TIN2009-1389.

REFERENCES

[1] 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.

[2] C. Chaux, J. C. Pesquet, and L. Duval, "Noise Covariance Properties in Dual-Tree Wavelet Decomposition," *IEEE Trans. on Information Theory*, vol. 53, no. 12, pp. 4690–4700, Dec. 2007.

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

[4] C. M. Bishop, *Neural Networks for Pattern Recognition*, Clarendon Press, Oxford, U.K., 1995.

[5] J. Shlens, "A Tutorial on Principal Component Analysis," http://www.snl.salk.edu/ ˜shlens/pub/notes/ pca.pdf, 2005.

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

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

[8] Ch. V. Latchoumane, D. Chung, S. Kim, and J. Jeong, "Segmentation and Characterization of EEG During Mental Tasks Using Dynamical Nonstationarity," in *CIMED'07 Workshop on Advances in Early Diagnostics andCAre for Alzheimer's Disease*. 2007, IET.

[9] Z. J. Wang, P. W. Lee, and M. J. McKeown, "A Novel Segmentation, Mutual Information Network Framework for EEG Analysis of Motor Tasks," *BioMedical Engineering Online, http://www.biomedical-engineering-online.com*, vol. 8, no. 9, 2009.

[10] U. Graichen, H. Witte, and J. Haueisen, "Analysis of Induced Components in Electroencephalograms Using a Multiple Correlation Method," *BioMedical Engineering Online, http://www.biomedical-engineering-online.com*, vol. 8, no. 21, 2009.

[11] D. J. Krusienski, "A Method for Visualizing Independent Spatio-Temporal Patterns of Brain Activity," *EURASIP Journal on Advances in Signal Processing, http://www.hindawi.com/journals/asp/2009/948961.html*, vol. 2009, no. 948961, 2009.

[12] C. Gómez, R. Hornero, D. Abásolo, A. Fernández, and J. Escudero, "Analysis of MEG Background Activity in Alzheimers Disease Using Nonlinear Methods and ANFIS," *Annals of Biomedical Engineering*, vol. 37, no. 3, pp. 586–594, 2009.

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

[14] P. Johankhani, V. Kodogiannis, and K. Revett, "EEG Signal Classification Using Wavelet Feature Extraction and Neural Networks," in *Int. Symp. on Modern Computing, JVA06*, 2006, pp. 120–124.

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

[16] S. Haykin, *Neural Networks*, IEEE Press, NY, 1994.

[17] 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.