USING WAVELET TRANSFORM FOR FEATURE EXTRACTION FROM EEG SIGNAL

Lenka Lhotska, Vaclav Gerla, Jiri Bukartyk
Gerstner Laboratory, Czech Technical University in Prague, Technicka 2, 166 27 Prague 6, Czech Republic

Vladimir Krajca, Svojmil Petranek
University Hospital Bulovka, Budínova 2, 18081 Prague 8, Czech Republic

Keywords: EEG processing, wavelet transform, feature extraction.

Abstract: Manual evaluation of long-term EEG recordings is very tedious, time consuming, and subjective process. The aims of automated processing are on one side to ease the work of medical doctors and on the other side to make the evaluation more objective. This paper addresses the problem of computer-assisted sleep staging. It describes ongoing research in this area. The proposed solution comprises several consecutive steps, namely EEG signal pre-processing, feature extraction, feature normalization, and application of decision trees for classification. The work is focused on the feature extraction step that is regarded as the most important one in the classification process.

1 INTRODUCTION

The electroencephalogram (EEG), describing the electric activity of the brain, contains a lot of information about the state of patient health. It has the advantage of being non-invasive and applicable over longer time span (up to 24 hours if necessary). This is an important feature in case we want to follow disorders that are not permanently present but appear incidentally (e.g. epileptic seizure) or under certain conditions (various sleep disorders). Although the attempts to support EEG evaluation by automatic or semi-automatic processing have been made for a long time, there are still many problems to be solved. We try to contribute by our research to this effort. The main objective of the described work is the identification of the most informative features from sleep EEG records that could be used for automated (or semi-automated) sleep stage classification. Our approach to the analysis of human sleep uses wavelet transform (WT) and statistics for feature extraction and construction. The extracted and computed features are used as inputs for a decision tree (Quinlan, 1990) that is learned to classify individual sleep stages. We use for our experiment EEG sleep records rated by an expert, freely available and downloadable from the Internet (Kemp, 2007).

The paper is organized as follows. Section 2 describes sleep EEG signal and approaches to its evaluation. Methods used in our research are presented in Section 3. Section 4 is devoted to description of performed experiments. In Section 5 the results of experiments are discussed and the conclusion is presented in Section 6.

2 SLEEP AND ITS COMPUTER SUPPORTED CLASSIFICATION

Sleep is a non-uniform biological state that has been divided into several stages based on polysomnographic (PSG) measurements that include EEG, EMG, EOG, ECG, temperature, SpO_2 (oxygen saturation of the blood, recorded on the finger), respiration signals, as well as movement or body position. Polysomnography is usually performed over the duration of an entire night, or at least 6.5 hours, in order to investigate normal and disturbed sleep or vigilance (Bloch, 1997). Normal healthy sleep is organized into sequences of stages that typically cycle every 60 – 90 min. The most widely used standard for terminology and scoring of sleep

Lhotska L., Gerla V., Bukartyk J., Krajca V. and Petranek S. (2008).
USING WAVELET TRANSFORM FOR FEATURE EXTRACTION FROM EEG SIGNAL.
In *Proceedings of the First International Conference on Bio-inspired Systems and Signal Processing*, pages 236-241
DOI: 10.5220/0001062002360241

stages is the manual by Rechtschaffen and Kales (RK) (Rechtschaffen and Kales, 1968). A standard summary method is the hypnogram that graphically represents sleep stages in 20-30 second epochs. The PSG can be generally divided into epochs of 10, 20, 30, or 60 s, which are then visually classified into one of RK stages by a sleep technologist. The resulting time evolutionary description of sleep in terms of stages, termed hypnogram, is used by physicians for diagnosis. The Rechtschaffen and Kales manual details a complete process of recording and analysing sleep, which is followed by the vast majority of sleep laboratories, worldwide. On the basis of EEG (plus EOG and EMG), epochs can be scored into sleep stages:

- Stage 1 – shallow/drowsy sleep;
- Stage 2 – light sleep;
- Stage 3 – deepening sleep;
- Stage 4 – deepest sleep;
- Stage REM – dreaming sleep.

Stages 1 to 4 are frequently described as non-REM sleep, and stages 3 and 4 are described as slow wave sleep (SWS). Other scores are Wake (W) and Movement Time (MT). Since the depth of sleep changes continuously, the artificial demarcation of sleep stages by the RK classification is a simplification. The exact time of change of state is highly subjective and leaves room for interpretation by the physician who scores transitional epochs (e.g., Stage 1 and Stage 3) differently on different occasions (Schaltenbrand, 1996).

Studies have shown agreement between physicians performing scoring that ranges from 67% to 91% (Gaillard and Tissot, 1973), (Stanus et al., 1987), (Kim et al., 1992), depending on different scoring epoch lengths and number of readers. However it is necessary to remark that most data on interscorer agreement are based on the study of normal subjects. Processing of sleep recordings requires elaborate training and is time consuming and expensive. No generally accepted standard exists for automatic sleep staging, but computerization can improve efficiency and reduce cost (Doman, 1995), and enhance collaboration between laboratories (Kemp, 1993).

Various approaches to computer classification of PSGs have been used. Johnson et al. (Johnson et al., 1969) presented a spectral analysis study of the EEG in different stages, which was subsequently used by Larsen and Walter (Larsen and Walter, 1970) to develop an automated staging technique based on multiple-discriminant analysis. Agarwal and Gotman (Agarwal and Gotman, 2001) use a method based on the segmentation and self-organization technique. The following five steps are necessary to perform computer-assisted staging: segmentation; feature extraction; clustering; assignment of stages to different clusters of patterns; and optional smoothing of the hypnogram. The study (Agarwal and Gotman, 2001) shows that the greatest discrepancy occurs in Stage 1. The sensitivity and the specificity are 38.6% and 43.4%, respectively. This is to be expected in the highly transitional Stage 1. Stage 1 also has significant similarities to REM stage and can be considered as one stage away from Stage 1. Moreover, it is accepted that manual scoring of Stage 1 is the most subjective due to its transitional nature.

3 METHODS

In our study we have used similar procedure as Agarwal and Gotman and the same we used in one of our previous studies (Gerla, Lhotska, and Krajca, 2005). The sleep EEG signal classification comprises several steps: segmentation, feature extraction, feature normalization, feature selection, and generation of decision trees.

We have applied wavelet transform (Daubechies, 1992) to sleep EEG signal preprocessing. Mean of the signal is calculated and subtracted from a signal before WT is applied. Discrete Wavelet Transform (DWT) represented by a filter bank is employed for wavelet decomposition. Before the decomposition starts it is necessary to select a mother wavelet used for defining FIR filters and a level of a decomposition tree. For deciding which mother wavelet should be selected we consider the impulse response and amplitude frequency characteristics of the FIR filter specified by the corresponding mother wavelet. After the DWT is done we get approximation and detail coefficients as input data for further processing. Then the segmentation is performed.

Segmentation. The non-adaptive segmentation is employed. Non-adaptive or constant segmentation divides a signal into segments of a constant length. This kind of segmentation is basically the easiest one. The disadvantage of this method is that the segments are not necessarily stationary. The length of a segment is chosen regarding the character of data.

Feature extraction is the second most important part after wavelet decomposition. It is a process which changes representation of segments by extracting features from them. The aim is to select those features which carry most information about the segment. The statistic parameters are in principle very suitable for this purpose. We use autoregressive features and computed wavelet coefficients as well. We use the following parameters: average absolute amplitude, maximal positive amplitude, maximal negative amplitude, maximal absolute amplitude, frequency weighted energy, sample mean, sample central moment, sample variance, statistical median, energy, and entropy. The autoregressive features are calculated from the transfer function of an *autoregressive model*, in which a present value x_n or future values x_{n+I} , i=1,2,… are estimated by using the previous values $\{x_{n-m}, \ldots, x_{n-1}\}$ (Therrien, 1992). We can extract features from each source (an original signal, its first and second derivation) independently.

Feature normalization. Mean and standard deviation of extracted features are different. That could have a negative influence to the classification process, when a classifier uses distances between points in n-dimensional space. Before we start classification the features must be normalized to have the same mean and standard deviation. The features have normal distribution N(0,1).

Feature reduction. There are several different ways in which the dimension of a problem can be reduced. In this work Principal Component Analysis (PCA) (Smith, 2002) approach is used which defines new features (principal components or PCs) as mutually-orthogonal linear combinations of the original features.

Feature selection is considered successful if the dimensionality of the data is reduced and the accuracy of a learning algorithm improves or remains the same. Decision tree algorithms such as C4.5 can sometimes overfit training data, resulting in large trees. In many cases, removing irrelevant and redundant information can result in C4.5 producing smaller trees. The Chi-squared statistic is used for feature selection.

Classification. We have decided to use decision tree algorithms because they are robust, fast, and what is important especially in medical domain their results are very easy to interpret. In particular, the C4.5 algorithm has been applied, namely its J48 variant available in the Weka software tool (Frank et al., 2007).

Success rate of classification. As a measure of success rate we have used the overall accuracy of the classification. The overall accuracy is calculated as the relative number of correct decisions.

4 EXPERIMENTS

The main purpose of our experiments has been to find the most suitable wavelet decomposition and the most discriminative features to achieve good classification results. The analyzed EEG recordings are presented in the next section and then our experiments with EEG data are described.

4.1 Source of EEG Recordings

Our source of EEG recordings is The Sleep-EDF Database (Kemp, 2007). Four EEG recordings from different subjects were downloaded. The recordings were obtained from Caucasian males and females (21 - 35 years old) without any medication. They contain horizontal EOG, Fpz-Cz and Pz-Oz EEG, each sampled at 100 Hz. The recordings also contain the submental-EMG envelope, oro-nasal airflow, rectal body temperature and an event marker, all sampled at 1 Hz. Hypnograms are also added which are manually scored according to Rechtschaffen & Kales based on Fpz-Cz / Pz-Oz EEG instead of C4-A1 / C3-A2 EEG (Sweden et al., 1990).

Subjects, recordings and hypnogram scoring for the 4 sc* recordings are described in (Mourtazaev, 1995). Subjects and 4 st* recordings are more extensively described in (Kemp et al., 2000). The sleep stages Wake, Stage1, Stage2, Stage3, Stage4, REM and 'unscored' are coded in the file as binaries 0, 1, 2, 3, 4, 5, 6 and 9.

After reviewing the data we have found out that the classes in data are unevenly represented. Class 1 (Wake) is the most frequent one and class 5 (stage4) occurs sporadically. We have generated the training set in which all classes are equally represented. That means that a classification error caused by an unequal distribution of classes should be reduced.

4.2 Experiment 1

A goal of this experiment is to find features which contain the information about classes included in data. In other words the features should be highly correlated with the class. In our case we have six classes (wake, stage1, stage2, stage3, stage4, REM). This is a complex task and it is quite impossible to find only one feature to correlate with all classes.

We modify our goal to examine all features for every combination of two different classes and select the most significant feature for discriminating the classes from each other. There are 15 combinations so we get 15 features in total. We have chosen EEG sample (sc4002e0), which includes all 6 classes; each having 200000 samples (2000 seconds). For WT, the following setting has been used: level of decomposition tree 7; mother wavelet db6; wavelet coefficients used for feature extraction (2,1), (3,1), (4,1), (5,1), (6,1), (7,1), (7,0); segment length 10s.

The results of this experiment and the best features selected for classification of every combination of two different classes are shown in

Table 1: Results of experiment 1 and the best features selected for differentiation between couples of classes.

Stage	Wake	Stage 1	Stage 2	Stage 3	Stage 4	REM
class	1	2	3	4	5	6
1		96% - f1	97.5% - f2	99.5% - f3	99.5% - f5	98.9% - f1
2	96% - f1		*85% - f7*	91.5% - f8	98.5% - f9	*70% - f10*
3	97.5% - f2	*85% - f7*		*73% - f11*	94% - f12	*85% - f4*
4	99.5% - f3	91.5% - f8	*73% - f11*		*85% -f3*	94.5% - f13
5	99.5% - f5	98.5% - f9	94% - f12	*85% -f3*		99.4% - f6
6	98.9% - f1	*70% - f10*	*85% - f4*	94.5% - f13	99.4% - f6	

Table 2: Description of the used features.

feature	original name of a feature	source for extraction	wavelet coefficient	full name of the feature
f1	MeaAbV_1d_d2_Fpz-Cz	first derivation	D2 (2.1)	average absolute amplitude
f2	Energy_sg_d4_Pz-Oz	signal	D4 (4.1)	energy
f3	MeaAbV_Sg_d5_Pz-Oz	signal	D5 (5.1)	average absolute amplitude
f4	Energy_1d_d5_Fpz-Oz	first derivation	D5 (5.1)	energy
f5	Energy_1d_d5_Pz-Oz	first derivation	D5 (5.1)	energy
f6	FrWeiE_Sg_d6_Fpz-Cz	signal	D6 (6.1)	frequency weighted energy
f7	FrWeiE_1d_d5_Pz-Oz	first derivation	D5 (5.1)	frequency weighted energy
f8	FrWeiE_Sg_d5_Pz-Oz	signal	D5 (5.1)	frequency weighted energy
f9	MeaAbV_Sg_d5_Fpz-Cz	signal	D5 (5.1)	average absolute amplitude
f10	MeaAbV_Sg_d3_Pz-Oz	signal	D3 (3.1)	average absolute amplitude
f11	FrWeiE_Sg_d7_Pz-Oz	signal	D7 (7.1)	frequency weighted energy
f12	Energy_Sg_d6_Pz-Oz	signal	D6 (6.1)	energy
f13	Energy_1d_d6_Fpz-Cz	first derivation	D6 (6.1)	energy

Table 1. The names and sources of these features are presented in Table 2. There are five classification results below 90% as it is shown in Table 1. It means that we are not able to extract any single feature which can separate these particular combinations of two classes. There must be used a combination of features. We can see that there are two features (f1, f3) occurring not only once as most discriminative. Each of them is chosen to be the discriminative feature for two combinations. A set of features is therefore reduced and we have 13 features. Unfortunately 5 of the features (marked in italics in Table 1) are not good enough for classification and thus we have decided to perform another experiment where different wavelet coefficients are decomposed and other features are examined.

4.3 Experiment 2

The goal is implicated by the result of the previous experiment. There are five combinations of two classes (4x3, 6x2, 3x2, 5x4, 6x3) which are classified with success rate lower than 90% by using features extracted from the wavelet coefficients. Now we try to achieve more accurate results by extracting features from such wavelet coefficient that have the same frequency resolution. Frequency resolution of a wavelet coefficient depends on sample frequency of data (100Hz) and on the level of the wavelet coefficient. We may be able to find more specific features carrying more information about separability of classes. Two different settings and wavelet decomposition trees are used: 1. level of decomposition tree 4; mother wavelet db15; wavelet coefficients used for feature extraction (4,0), (4,1), (4,2), (4,3), (4,4), (4,5), (4,6), (4,7); segment length 10s; 2. level of decomposition tree 5; mother wavelet db20; wavelet coefficients used for feature extraction (5,0), (5,1), (5,2), (5,3), (5,4), (5,5), (5,6), (5,7), (5,8), (5,9), (5,10), (5,11), (5,12), (5,13), (5,14), (5,15); segment length 10s. Wavelet coefficients from the highest level of decomposition trees are used for feature extraction. They have the highest frequency resolution compared with others in the wavelet decomposition tree. We can assume that the features extracted from these coefficients carry different piece of information about classes.

The results of this experiment have not been so successful as we have expected. The only feature that has brought relatively significant improvement of differentiation between two classes (by 7.5%) has been average absolute amplitude FPz-Cz (wavelet coefficient (5,1)). The experiment has shown that application of other mother wavelets (db20 or higher) and different wavelet decomposition trees could result in finding new more discriminative features.

4.4 Experiment 3

The final experiment has been divided into three parts, namely using different groups of classes. EEG recordings (sc4012e0, st7022j0, sc4102e0) have been used as testing sets for this purpose.

Part 1. We have classified data into all six classes using all features as described in experiment 1. The results have verified our assumption that the features f3, f4, f7, f10 and f11 which do not separate classes well (see Table 1) decrease final classification accuracy.

Part 2. Based on the experiment 1 we have tried to distinguish among four classes, organized in two groups, namely (1, 3, 5, 6) and (1, 4, 5, 6). We have used six features from the original 15 for each group. The classification results for the first group have been negatively influenced by the feature f3 and for the second group by the feature f4.

Part 3. We have verified well discriminating features discovered in experiment 1. For this purpose we have selected three groups of three classes each that can be separated very well by these features. The three groups are composed of the following classes (1, 5, 6), (1, 2, 5) and (1, 4, 6).

All results are summarized in Table 3. The record sc4102e0 has not been used in those experiments where class 5 has been tested because it does not contain any segment belonging to class 5. The classification results are as we have assumed. They are mainly affected by low discriminability between classes 2 (stage1) and 3 (stage2) and classes 2 (stage1) and 6 (REM).

Table 3: Results of experiment 3 (success rate of classification).

classes	sc4002e0	sc4012e0	st7121j0	sc4102e0
1,2,3,4,5,6	72.1%	69.5%	63%	x
1,3,5,6	87.2%	87.6%	78.4%	x
1,4,5,6	87.8%	82.9%	73.2%	x
1,5,6	98.3%	94.5%	92%	x
1,2,5	97.2	92%	87%	x
1,4,6	96.5%	91.3%	90%	87%

5 DISCUSSION

Tables 1 and 2 are used for the interpretation. When we look at Table 2 we can see that all features extracted for the classification task in the experiment 1 are based on energy, mean absolute amplitude and frequency weighted energy. These features reflect the changes of energy in the given wavelet coefficient which is related to a specific frequency spectrum. This is very important as we see later. Now we try to explain why we have got the results of classification shown in Table 1. When we look at this table we can see that successful classifications are for the classes classified with features extracted from wavelet coefficients which have the frequency spectrum same as the frequency spectrum only a single class in the set of two classes has. That means that the feature used for such classification has high energy for this class and small energy for the other one. Then we can simply use a threshold to separate these two classes from each other. When we look at Table 1 again we can see that all successful classification results (success rate higher than 90%) are achieved between classes with mutual distance more than one class, for example, classes 1x3, 1x4, 2x5 etc. It is because the distance between these classes is quite long which is required for successful classification. An exception is the class 1 which is classified correctly in all cases. When we examine frequency spectra of classes 1 (Wake) and 2 (stage1), we find out that they are well separable. However we have to note that there exists overlapping (some frequencies occur in neighbouring stages). Therefore poorer classification result (below 90%) is for classes just next to each other (2x3, 3x4, 4x5 and 6x2). Unfortunately we have not yet found any feature better describing the classes by using different wavelet decomposition (experiment 2). The results of classification in experiment 3 are affected by this fact as well. In the following paragraph we suggest some ideas which could improve classification of sleep EEG data.

The frequency resolution of wavelet coefficients in level 5 of a wavelet decomposition tree is 3.12Hz. This decomposition is used in experiment 2. It was not detailed enough for distinguishing incorrectly classified classes. So we propose to make the frequency resolution higher by getting wavelet coefficients from level 6 (frequency resolution 1.57Hz) or even level 7 (0.78Hz). For these purposes we must ensure that the filter used for decomposition has steep frequency characteristic. We would recommend to use mother wavelets db20 and higher. If this condition is satisfied the results would not be influenced by leakage of other frequency components (antialiasing).

6 CONCLUSIONS AND FUTURE WORK

Sleep problems belong to the most common serious neurological disorders. Reliable and robust detection of these disorders would improve the quality of life of many people. The implemented methods allow automatic classification of EEG signals. The approach has been tested on real sleep EEG recording for which the classification has been known. We have focused on discovering the most significant features which would be highly correlated with classes of data. Our experiments have been based on the selection of a single feature to separate data belonging to two classes. There have been many other features with good selection results. The most frequent ones have been autoregressive features representing the order of used AR model and error of AR model. We have determined some features and wavelet coefficients which are best suited for classification of sleep EEG data. The future work will be focused on exploitation of other types of mother wavelets, using higher level of wavelet coefficients as source of features, and more sophisticated classifiers.

ACKNOWLEDGEMENTS

This work has been supported by the research program "Information Society" under Grant No. 1ET101210512 "Intelligent methods for evaluation of long-term EEG recordings".

REFERENCES

Agarwal, R., Gotman, J., 2001. Computer-assisted sleep staging, *IEEE Trans. on Biomed. Engineering*, 48 (12), pp. 1412-1423.

Bloch, K.E., 1997. Polysomnography: a systematic review, *Technology and Health Care*, 5, pp. 285-305.

Daubechies, I., 1992. *Ten lectures on Wavelets*, CBMS-NSF, SIAM, 61, Philadelphia, Pennsylvania, USA.

Doman, J., Detka, C., Hoffman, T., Kesicki, D., Monahan, J.P., Buysse, D.J., Reynolds III, C.F., Coble, P.A., Matzzie, J., Kupfer, D.J., 1995. Automating the sleep laboratory: implementation and validation of digital recording and analysis, *Int. J. Biomed. Comput*., 38, pp. 277-290.

Frank, E., Hall, M., Trigg, L. Weka – Data Mining Software in Java. Internet site address: http://www.cs.waikato.ac.nz.

Gaillard, J.M., Tissot, R., 1973. Principles of automatic analysis of sleep records with a hybrid system, *Comput. Biomed. Res*., 6, pp. 1-13.

Gerla, V., Lhotská, L., Krajča, V., 2005. Utilization of Time Dependence in EEG Signal Classification. *The 3rd European Medical and Biological Engineering Conference EMBEC'05 Prague.*

Johnson, L., Lubin, A., Naitoh, P., Nute, C., Austin, M., 1969. Spectral analysis of the EEG of dominant and nondominant alpha subjects during waking and sleeping, *Electroencephalogr. Clin. Neurophysiol.*, 26, pp. 361-370.

Kemp, B., 1993. A proposal for computer-based sleep/wake analysis, *J. Sleep Res*., 2, pp. 179-185.

Kemp, B., Zwinderman, A., Tuk, B., Kamphuisen, H., Oberyé, J., 2000. Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG. *IEEE-BME* 47(9), pp. 1185-1194.

Kemp, B., 2007. Sleep Recordings and Hypnograms in European Data Format. The Netherlands. Internet site address: http://www.physionet.org/ physiobank/ database/sleep-edf.

Kim, Y., Kurachi, M., Horita, M., Matsuura, K., Kamikawa, Y., 1992. Agreement in visual scoring of sleep stages among laboratories in Japan, *J. Sleep Res*., pp. 58-60.

Larsen, L.E., Walter, D.O., 1970. On automatic methods of sleep staging by EEG spectra, *Electroencephalogr. Clin. Neurophysiol*., 28, pp. 459-467.

Mourtazaev, M., Kemp, B., Zwinderman, A., Kamphuisen, H., 1995. Age and gender affect different characteristics of slow waves in the sleep EEG. *Sleep* 18(7), pp. 557-564.

Quinlan, J.R., 1990. Decision trees and decision making. *IEEE Trans System, Man and Cybernetics*, 20(2), pp. 339-346.

Rechtschaffen, A., Kales, A. (eds.), 1968. A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects, Brain Inform. Service/Brain Res. Inst., Univ. California, Los Angeles.

Schaltenbrand, N. et al., 1996. Sleep stage scoring using the neural network model: Comparison between visual and automatic analysis in normal subjects and patients, *Sleep*, 19, pp. 26-35.

Smith, L.I., 2002. *A tutorial on Principal Components Analysis*, University of Otago, New Zealand, 2002.

Stanus, E., Lacroix, B., Kerkhofs, M., Mendlewicz, J., 1987. Automated sleep scoring: A comparative reliability study of algorithms, *Electroencephalogr. Clin. Neurophysiol*., 66, pp. 448-456.

Sweden, B., Kemp, B., Kamphuisen, H., Velde, E., 1990. Alternative electrode placement in (automatic) sleep scoring (Fpz-Cz / Pz-Oz versus C4-A1 / C3-A2). *Sleep* 13(3), pp.279-283.

Therrien, C.W., 1992. *Discrete Random Signals and Statistical Signal Processing.* Englewood Cliffs (NJ): Prentice Hall.