Selection of Measures for Sleep Stages Classification

K. Šušmáková, A. Krakovská

Institute of Measurement Science, Slovak Academy of Sciences
Dúbravská cesta 9, 842 19 Bratislava, Slovak Republic
E-mail: kristina.susmakova@savba.sk

Abstract: *In this work a large amount of features from polysomnographic recordings was tested to find the best set of variables for sleep stages classification. Discriminant analysis was done with Fisher quadratic classifier and forward selection procedure. Resulting set contains 14 measures from EEG, EOG, EMG and ECG signals, some of them are used in this context for the first time (e. g. fractal exponent and entropy of EMG).*

Keywords: Sleep stages classification, Forward selection procedure, rules of Rechtschaffen and Kales, nonlinear, spectral measures

1. Introduction

The evaluation of sleep stages is done after broadly appreciated Rechtschaffen-Kales manual [1], which involves parameters, techniques and wave patterns of three physiological signals: electroencephalogram (EEG), electrooculogram (EOG) and electromyogram (EMG) needed for definitive assignation of sleep stages. In clinical practice also other physiological signals, e. g. electrocardiogram (ECG), blood pressure etc., may be measured.

The main states of vigilance are wakefulness (W), REM sleep and non-REM sleep (NREM). NREM sleep is further divided into four stages from the lightest Stage 1 (S1) to the deepest Stage 4 (S4). Stages 3 and 4 are referred to as slow wave sleep (SWS).
During W there is a low voltage (10−30 μV) and mixed frequency EEG, substantial alpha activity in EEG and relatively high tonic EMG.
S1 is characterized by low voltage, mixed frequency EEG with the highest amplitude in 2-7 Hz range. Alpha activity may be present but it must not take more than 50% of an epoch. Vertex sharp waves may occur; their amplitude can reach the value of about 200 μV. In S1 after wakefulness slow eye movements can be present. The EMG level is lower than in the wakefulness.
S2 is characterized by sleep spindles and K-complexes on a relatively low voltage, mixed frequency background activity and the absence of slow waves. Sleep spindles are bursts of brain waves of 12-16 Hz. A K-complex is a sharp negative wave (the amplitude demand is at least 75 μV) followed by a slower positive one.
SWS is detected if more then 20% of the epoch of EEG record contains delta waves which are characterized with 2 Hz or slower and with the amplitudes above 75 μV. Sleep spindles and K-complexes may also be present.
REM sleep shows low voltage and mixed frequency (similarly to S1) of EEG, sawtooth wave pattern is often present. EMG reaches the lowest level and episodic rapid eye movements occur.
The convenience of developing a computerized system for automated analysis and classification of sleep stages has been recognized by several authors [2-4]. A few commercial systems are also available; however they showed substantial differences from the visually scored polysomnographs in the distribution of the sleep stages. Most of these works were concentrated on the choice of the best kind of classifier, however there are only a few works dealing with selection of proper set of discriminative features [5].

2. Subject and Methods

Data

Data with all-night polysomnographic records were kindly provided by Prof. G. Dorffner, received by The Siesta Group Schlafanalyse GmbH. The records were obtained from 20 healthy subjects, 10 men and 10 women. Ages ranged from 23 to 82 years old with an average 50 ± 21.5 years. Subjects slept at their usual sleeping time, typically from 11 pm, the average recording time was 7.5 hours with sleep efficiency 87.1%. Sleep stages were scored by two experts on 30 s non-overlapping segments, if there was an ambiguity, the third independent expert made the decision. All measures were computed on these 30 s long windows, for 1 channel of EMG, 2 channels of EOG, 6 EEG channels (derivations: Fp1-M2, C3-M2, O1-M2, Fp2-M1, C4-M1, O2- M1, where M1, M2 are the left and right mastoids) and 1 channel of ECG. Following numbers of sleep stages were analyzed: 2069 stages of waking, 1452 stages of S1, 7860 stages of S2, 1586 stages of S3, 1865 stages of S4, and 3226 stages of REM sleep.

Computed measures

Following measures were computed for all 10 channels: average frequency, average amplitude, variance, skewness, kurtosis, normality test [6], spectral moments [7], spectral edge [8], spectral exponent [9], spectral entropy [8], fractal dimension [10], exponent of detrended fluctuation analysis [11, 12], entropy, absolute spectral powers [14], relative spectral powers [14], relative power ratios [14]. Coherency [14], phase angle [14] and mutual information were computed for all 15 combinations of EEG channels and between EOG channels.

Powers, coherences, and phase angles were computed for EEG channels in following frequency bands: delta 1: 0.5 - 2 Hz, delta 2: 2 - 4 Hz, theta 1: 4 - 6 Hz, theta 2: 6 - 8 Hz, alpha 1: 8 - 10 Hz, alpha 2: 10 - 12 Hz, sigma 1: 12 - 14 Hz, sigma 2: 14 - 16 Hz, beta: 16 - 30 Hz, and gamma: 30 - 40 Hz and total power: 0.5 Hz - 40 Hz. Power ratios were computed between the relative spectral powers in the main frequency bands: delta/theta, delta/alpha, delta/sigma, delta/beta, delta/gamma, theta/alpha, theta/sigma, theta/beta, theta/gamma, alpha/sigma, alpha/beta, alpha/gamma, sigma/beta, sigma/gamma, and gamma/beta. For EMG and ECG spectral measures were computed only for frequency bands from 10 Hz due to the high-pass filter used on these signals.

Discriminant analysis

Discriminant analysis was done by Fisher quadratic classifier with forward selection procedure (FSP) [12]. In FSP the choice of best features is realized in iterative way - in the first step the best variable in discrimination of sleep stages is selected. At each step a new variable is added to previous set of variables so that new set maximizes (minimizes) some specific criterion. At each step the significance of change of the value of the criterion is tested with t-test. Null hypothesis is that the value of the criterion is not changed with the new set of variables in comparison with the set in previous step. After the p-value of the t-test exceeds the limit 0.05 the null hypothesis is confirmed and FSP is finished. All measures entered in FSP and the criterion to be minimized was the mean error of classification into all sleep stages. Classification procedure was as follows:

1. With the aim to eliminate effect of various numerosity of sleep stages on the classification error from the whole database the same amount of epochs of each sleep stage was randomly chosen; this amount was equal to the less numerous stage (S1).
2. From this representative data set of all sleep stages 90% of randomly determined values created a training set on which the discrimination function was derived.
3. Testing was done on the rest of the data. The total error rate - the percentage of incorrect classified epochs - and the errors of the respective states were computed.

This procedure was repeated 100 times and the mean values and standard deviations of errors were calculated.

3. Results and conclusions

The result of SFP can be seen in Table 1, where selected measures, p – values of significance of change between two steps, mean and standard deviation of classification errors and also mean errors of respective sleep stage are listed. The mean error of classification was diminished from 42.5% by best single-performing variable (delta/beta C3 derivation) to 19.3% by 14 selected measures. On the first three positions, there are EEG, EOG and EMG measures placed and in the last 14th position, the zero-crossing rate of ECG are placed, which indicates the benefit of measuring all polysomnographic signals.

Table 1: Results of forward selection procedure. Selected measures, p-values of hypotheses testing, means and standard deviations of classification errors in [%] and also mean errors of particular sleep stages are listed. Abbreviations: f. fractal; a. absolute, r, relative; c. coherency, zero-cross. zero-crossing rate.

Measure	Channel	p	mean error[%]	Std error[%]	error W[%]	error S1[%]	error S2[%]	error SWS[%]	error REM[%]
delta/beta	C3	0	42.5	1.1	32	80.4	40.3	11.4	48.4
f. exponent	EMG	0	35.3	0.9	29.4	60.5	46.1	11.6	28.5
variance	EOG2	0	29.9	0.9	25.7	47.1	35.1	9.2	32.4
a.sigma1	C3	0	27	0.9	25.1	43.8	30.9	9.6	25.8
r.delta2	O2	0	24.2	0.9	18.6	39.2	29.8	9.4	24.2
theta/gamma	C3	0	23.2	0.9	17.9	39.1	28.6	8.6	21.4
theta/alpha	O1	0	22.4	0.9	16.4	39.9	26.7	8.1	20.8
sigma/gamma	C4	0	21.7	0.8	15.9	39.2	26.5	7.6	19.1
c.delta1	O1-O2	0.002	21.3	0.8	15.2	38.7	26.7	7.9	18.3
entropy	EMG	0.001	20.9	0.8	15.2	39.7	27.2	7.7	15.1
a.theta2	Fp1	0	20.4	0.9	14.7	38.4	26.2	7.7	14.9
theta/alpha	Fp1	0.009	20.1	0.8	13.7	38.3	26	7.7	14.5
c.sigma2	O1-C3	0	19.6	0.8	13.4	38.1	25.3	7.4	13.8
zero-cross.	ECG	0.016	19.3	0.9	13	37.6	24.8	7.5	13.7
alpha/sigma	O1	0.555	19.3	0.8	12.6	36.8	25.5	7.5	13.9

Table 2: Confusion matrix using results from FSP (first 14 measures from Table 1): labels in the first column denote sleep stages determined by experts and in the row there are percentages of sleep stages how this particular stage was determined by classifier.

	W	**S1**	**S2**	**SWS**	**REM**
W	86.36	12.12	0.48	0.02	1.01
S1	10.68	61.20	11.89	0.29	15.95
S2	1.12	10.42	74.66	9.40	4.40
SWS	0.50	0.39	6.14	92.89	0.08
REM	1.81	8.58	3.35	0.11	86.15

Confusion matrix of these 14 selected measures can be found in Table 2. The best determined stage was SWS with 92.89% classification occurrence. Two less predicable stages were S1 with 61.20% of successful classifications and S2 with 74.66%. S1 is a transient state between wakefulness and first "real sleep" stage S2. From the confusion matrix it can be seen that about 11% is misclassified as wakefulness and about 12% as S2. The high error of S1 detection is not so serious problem with regard to automatic sleep stages classification and its application in sleep medicine because S1 amounts only about 5% of all sleep stages of normal healthy subject. The most numerous stage, S2 (about 50% of all stages), was misclassified with about 10% as neigbouring NREM stages S1 or SWS. The error of S2 could be improved

with combination of resulted measures and algorithms of detection of K-complexes and sleep spindles which occur typically during S2 but could be present in smaller amount also during SWS.

Acknowledgements

This work was supported by Slovak Grant Agency for Science (grant No 2/7087/27).

References

[1] Rechtschaffen A. and Kales A. (Eds.): A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subject, US Government Printing Office, National Institute of Health Publication, Washington DC, 1968.

[2] Park, H.-J., Oh, J.-S., Jeong, D.-U. and Park, K. - S.: Automated sleep stage scoring using hybrid rule- and case-based reasoning. *Computers and Biomedical Research*, 33(5):330349, 2000.

[3] Principe, J. C., Gala, S. K. and Chang, T. G.: Sleep staging automaton based on the theory of evidence. *IEEE Transactions on Biomedical Engineering*, 36(5):503–509, 1989.

[4] Anderer, P., Gruber, G., Parapatics, S., Woertz, M., Miazhynskaia, T., Klösch, G., Saletu, B., Zeitlhofer, J., Barbanoj, J., Danker-Hopfe, M., H. Himanen, S.-L., Kemp, B., Penzel, T., Grözinger, M., Kunz, D., Rappelsberger, P., Schlögl, A. and Dorffner, G.: An E-Health Solution for Automatic Sleep Classification according to Rechtschaffen and Kales: Validation Study of the Somnolyzer 24 x 7 Utilizing the Siesta Database, *Neuropsychobiology*, 51(3):115–133, 2005.

[5] Zoubek, L., Charbonnier, S., Lesecq, S., Buguet, A. and Chapotot, F.: Feature selection for sleep/wake stages classification using data driven methods, *Biomedical Signal Processing and Control*, 2(3): 171–179, 2007.

[6] Galbraith, C. G. and Wong E.: Moment analysis of EEG amplitude histograms and spectral analysis: relative classification of several behavioral tasks, *Perceptual and Motor Skills*, 76, 1993, p. 859-866.

[7] Vogel F., Holm S., and Lingjærde O. C.: Spectral moments and time domain representation of photoacoustic signals used for detection of crude oil in produced water, In Proc. Nordic Symp. on Physical Acoustics, Ustaoset, Norway, 2001.

[8] Fell J., Röschke J., Mann K., and Schäffner, C: Discrimination of sleep stages: a comparison between spectral and nonlinear EEG measures, *Electroencephalography and clinical Neurophysiology*, 98, 1996, p. 401-410.

[9] Pereda E., Gamundi A., Rial R., and González, J.: Nonlinear behaviour of human EEG: fractal exponent versus correlation dimension in awake and sleep stages, *Neurosci. Lett.*, 250, 1998, p. 91–94.

[10] Higuchi T.: Approach to an irregular time series on the basis of the fractal theory, *Physica D*, 31, 1988, p. 277 – 283.

[11] Little M., McSharry P., et al: Nonlinear, biophysically-informed speech pathology detection, In Proceedings of ICASSP, Toulouse, France, 2006, EEE Publishers.

[12] Peng C.-K., Havlin S., Stanley H. E., and Goldberger A. L.: Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series, *Chaos*, 5, 1995, p. 82 – 87.

[13] Thatcher W., North R.D., and Biver C.: EEG and intelligence: Relations between EEG coherence, EEG phase delay and power, *Clinical Neurophysiology*, 116, 2005, p. 2129-2141

[14] Meloun M., Militký, J.: Statistická analýza experimentálních dat (Statistical Analysis of Experimental Data, in Czech). ACADEMIA, Praha, 2004.

[15] Larose, D.: Data Mining Methods and Models. Wiley InterScience, 2006.