FinJeHeW Finnish Journal of eHealth and eWelfare

SCIENTIFIC PAPERS

VERTAISARVIOITU
KOLLEGIALT GRANSKAD
PEER-REVIEWED
www.tsv.fi/tunnus

Experiments of the sonification of the sleep electroencephalogram

Pedro Franco B.Sc.[1], Alpo Värri Dr.Tech.[2]

[1] Instituto Superior Técnico, Lisbon, Portugal, [2] Tampere University of Technology, Finland

Alpo Värri, Department of Signal Processing, Tampere University of Technology, P.O. Box 553, FI-33101 Tampere, FINLAND. Email: Alpo.Varri@tut.fi

Abstract

It is becoming possible to perform sleep recordings at home with equipment targeted for the regular consumers. This alleviates the pressures to increase capacity in sleep clinics. The interpretation of the sleep recordings is not very easy for the laymen and alternative assisting methods should be sought for this. Sonification is a method by which a phenomenon is converted to a sound for human listeners. This paper describes experiments made for the sonification of the electric activity of the brain, the electroencephalography (EEG) for the purpose of recognizing the presence and absence of the necessary refreshing components of sleep, deep sleep and rapid eye movement (REM) sleep. The methods are based on the calculation of features of the EEG signal which are characteristic to the deep and REM sleep as well as wakefulness. The features are converted to amplitude modulation functions of artificial and musical instrument sounds by using mathematical transforms such as Principal Component Analysis and Linear Discriminant Analysis. The results indicate that modulated sinusoidal signals are not appropriate for the sonification of sleep EEG but that modulating the sound of musical instruments could be a viable option for making the recognition of good and bad sleep possible.

Keywords: sonification, sleep, electroencephalogram, sleep stage scoring

Introduction

The recent developments in the recoding technologies of physiological signals make it possible to record the electroencephalogram (EEG) with inexpensive equipment targeted for consumers. It is becoming possible to consumers to perform polysomnographic sleep studies at home which was previously possible only for professionals in sleep clinics. This is good because there is not enough capacity in the clinics to record and analyze everybody who is interested in their sleep quality for various reasons. When a number of people can be satisfied with the results of the home sleep analyzers, the sleep clinics can focus on those patients who need sleep laboratory level analysis most. The interpretation of these home recordings is, however, by no means trivial. Methods are therefore needed to simplify the interpretation of these recordings for the laymen.

The purpose of the recording of EEG during sleep is to determine objectively how well the subject slept. The present sleep stage classification [1] divides the sleep-wakefulness continuum into the following classes: Stage W (Wakefulness), Stage N1 (NREM 1 sleep, light sleep), Stage N2 (NREM 2 sleep), Stage N3 (NREM 3 sleep, deep sleep, slow-wave-sleep, SWS), Stage R (REM sleep, rapid-eye-movement sleep). The previous sleep stage classification [2], also used partly in this study, too, divided the non-REM sleep into stages S1 to S4 and particularly Stage N3 into two stages S3 and S4 of which S4 represented the deepest sleep stage. Both classifications are based on the recording of the EEG, electrooculography (EOG) and surface electromyography (EMG). The time course of the changes of the stages during the sleep recording is called the hypnogram. According to the present knowledge, the normal, restorative sleep consists of a sufficient duration of SWS and REM sleep and it is not disturbed by frequent awakenings, caused by, e.g. sleep apnea [3, 4, 5].

Stage W is characterized by low amplitude mixed frequency activity (also so-called beta activity in the 16 to 30 Hz range when concentrating on a task) when the eyes are open and typically trains of 8-13 Hz activity (alpha activity) recorded in the occipital regions of the brain when the eyes are closed. Stage N1 contains low amplitude mixed frequency EEG activity predominantly in the 4 to 7 Hz range (theta activity) and slow eye movements. Stage N2 EEG contains typically trains of more than 500 ms long sleep spindles of 11 to 16 Hz activity and so-called K-complexes with a relatively high amplitude single negative and positive wave deflection from the base line lasting for more than 500 ms. Stage N3 consists of waves of frequency 0.5 to 2 Hz and peak-to-peak amplitude above 75 microvolts, the more dominantly the deeper the sleep is. Stage R EEG can consist of trains of 2 to 6 Hz triangular waves but they are not always present making the stage difficult to recognize based on EEG alone. The stage R is therefore recognized by the more or less frequent presence of rapid eye movements and a lowered muscle tone in the chin EMG [1].

Sonification is a method to use non-speech audio to convey information or perceptualize data [6]. It makes use of the fine characteristics of human hearing in making sense of audio material. With some training, the listener can usually make a difference between two sensory signal recordings better than by observing the visual waveforms of the signals [5]. In practice, sonification means that the input signal is processed in such a way that is transformed from its original form, mapping its frequency range to the audible frequency range (20-20 000 Hz) for human listening. For most applications of sonification, not all audible band is used as some people have difficulty hearing at both ends of the scale.

Sonification [7] is not a new technique. Some experiments were performed already in the 1950'ies and a patent was granted to a method of the sonification of multiple physiological signals in 1998 [8]. The Ambient Intelligence group in the University of Bielefeld has published several papers on the sonification of the EEG [9, 10, 11] and they believe that sonification can change clinical procedures even in the near future [12]. The methods of EEG sonification have not, however, been standardized yet and there is still room for innovation in this area. Sleep sonification has been studied a little, too. The study of Tulilaulu et al. used a movement sensor signal as the input to the sonification process [13]. This study did not, however aim at so direct use of the sonification output for determining the sleep quality because some random elements were also used to composing the musical output. Olivan et al. used straightforward audification of the signals recorded during sleep [14]. In order to convert the signals to the audible frequency range, the authors accelerated the speed by a factor of 200.

The nature of this field is such that purely mathematical algorithm is not sufficient to yield good results because listening to the output of the methods is a crucial part of the method development. Therefore we set out for some experiments of EEG sonification. The general objective was to develop a robust method by which people with no previous experience with EEG and only little training can listen to the produced sound and distinguish if the sleep was good or bad. Moreover, the audio files generated by the proposed methods should be pleasant to human hearing. These methods must be independent of the EEG sampling frequency and skip the technical artifacts characteristic of EEG like muscular and ocular artefacts.

Materials and methods

Material

Eight recordings from SIESTA polygraphic data base [15] were used for our experiments. Four of the recordings were from normal subjects and four were from subjects who had a diagnosed anxiety disorder. These recordings were not selected completely blindly because the visual inspection of the hypnograms was used to verify that the recordings were not severely abnormal. Table 1 shows how many epochs of 30 seconds each subject spent in each sleep stage classified according to the R&K system [2].

The C4-M1 and O2-M1 derivations were used for the EEG analysis. EOG activity was analyzed from the Fp1-M2 EEG channel.

Table 1. Number of epochs of 30 s of each sleep stage for each subject.

Subject	Wakefulness	S 1	S2	S3	S4	REM	Movem	Artifact
1	39	73	519	69	40	221	3	2
2	30	72	535	71	66	173	18	0
3	305	102	322	60	27	141	1	3
4	120	71	314	90	105	182	3	3
5	72	102	426	35	139	220	0	2
6	244	87	396	36	105	102	0	3
7	151	87	454	63	5	200	1	3
8	33	83	578	54	33	171	6	3

Methods

General

There are three main types of EEG sonification: Audification, Parameter Mapping and Model based. Audification is probably the simplest and oldest form of sonification technique. This technique corresponds to map (with a certain function) the range of EEG frequency and amplitude to respective audio frequency range and amplitude. However, this method has several limitations as it is very difficult to create a good audification technique that helps to understand essential phenomenon like for example: lack of sleep quality or epilepsy. In fact, most of the novel literature found about EEG sonification doesn´t follow this approach as it is considered obsolete. On the other hand, parameter mapping is nowadays one the most researched and popular forms of EEG sonification. Here, extracted parameters, also called as features extracted from the EEG signal serve to modulate sound characteristics such as amplitude and frequency. For last, model based sonification relies on mathematical or physical models to generate sound depending on the EEG input data. Although it is raising in popularity in data sonification, the method is not optimal for time series data representation like EEG data. Therefore, the choice for the project development was parameter mapping which is the most used method nowadays [7,16,17].

The sonification methods we used can be divided to two groups. The first group consisted of methods which required that an automatic sleep stage classification would be performed first. In the next step each stage would be converted to a predetermined sound relating to that class. The amplitude of the sound would be modulated by the certainty of belonging to that class.

The second group consisted of methods which did not require this classification. Instead, separate modulation amplitude values were calculated for the modulation of the musical instrument sounds for the classes slow wave sleep, REM sleep and wakefulness using formulas which had highest output values when the corresponding class was present in the current epoch. The output of these methods would be the combination sound of the sounds associated with each of these three classes and their amplitude would be weighted by their probability. In the following the methods in both groups are described in more detail. Both methods relied on the extraction of descriptive features and the selection of these features is explained first. Different variations of the methods explained here were also tested but they are not described here to keep this article shorter.

Feature extraction

Feature extraction is a critical part of all pattern recognition systems. In order to obtain the most descriptive features a number of candidate features were extracted from the EEG and their power to discriminate the relevant sleep stages were compared. The candidate features were Relative Power Density from 6 different EEG bands: Delta (0.5-4 Hz), Theta (4-8 Hz), Alpha (8-13 Hz), Sigma (13-16 Hz), Beta (16-30 Hz) and Gamma (30-45 Hz), Vigilance Index (VI), Spectral Asymmetry (SISA), Spectral Edge Frequency (SEF), Median Frequency (MF), Slow Wave Index (Slow Wave Index), TWI (Theta Wave Index), AWI (Alpha Wave Index), total STFT power, band entropy and the Hjorth Parameters: Activity, Mobility and Complexity [18]. The band powers were calculated with the short-time Fourier transform (STFT) in 15 s epochs such as the other features. A Blackman window was applied to the EEG signal before the STFT calculation.

Spectral Edge Frequency (SEF) represents the frequency, *freq*, which the range of 0 to *freq* contains 95% of the total power of the signal. Median Frequency (MF) represents the frequency, freq, which the range of 0 to freq contains 50% of the total power of the signal. The Vigilance Index indicates is represented by the fraction of higher frequencies over lower frequencies (Pgamma+Pbeta)/(Ptheta+Pdelta). Spectral Asymmetry represents asymmetry of the spectral power from the left side of the Alpha band and the right side of the Alpha band: (Pgamma+Pbeta+Psigma-Pdelta-Ptheta) / (Pgamma+Pbeta+Psigma+Pdelta+Ptheta). Slow Wave Index (SWI = Pdelta/(Ptheta+Palpha), Theta Wave Index (TWI = Ptheta/Ptheta+Pdelta) and Alpha Wave Index (AWI = Palpha/(Ptheta+Pdelta) are indexes that traduce different relationships between the lower frequency bands.

Entropy represents how concentrated the STFT power is. In other words, if the power is equally distributed between bands, entropy is maximum while if all power is concentrated in one band, entropy will be 0 (total power organization).

$$Entropy = \sum_{i=1}^{Nbands} -power(i) * \log(power(i)) \quad (1)$$

where *power (i)* represents each band power.

Eight features from Fp1 channel were extracted posteriorly in order to perform REM sonification: Hjorth parameters (Activity, Mobility and Complexity), lower, medium and high frequency band of EOG, signal's vigilance index and band entropy. Despite that eight features were extracted, only Activity was actually used.

Feature selection

As the objective is to separate the SWS and REM sleep from the other stages, a method was required to compare the performance of the extracted features in this respect. The best features chosen for non-classification sonification (SWS and REM) were selected through the Fisher Ratio (FR). The FR tells how separable two classes by their respective means and variance. In this feature selection method, it is assumed that the features have Gaussian distributions and the method evaluates how separable are the two Gaussians. The highest FRs of the studied features allow maximum distinguishability and are chosen for posterior non-classification models [19].

A number of plots were dawn in which the features' FRs were depicted in pairwise comparisons of class separability, e.g. REM vs NREM, Awake vs. Sleep etc. In addition to these plots, the feature outputs were plotted on the time axis and these plots were compared to the hypnograms of the same subject. The experiments showed that generally a FR above one was required to obtain sufficient separation. These experiments led to the choice of the following features shown in Table 2.

For the sonification method which did not use a sleep stage classifier, the feature selection was somewhat different. It was found out that only one parameter was not optimal but a formula of combining a few features worked better. The order of best formulas for SWS detection was (SWI*Complexity)/(Mobility*Entropy), (SWI*Complexity)/Entropy, and SWI/Entropy. The best formula for REM detection was TWI/Activity(Fp1). The Vigilance Index (VI) appeared to work well for wakefulness detection.

Temporal smoothing of the feature outputs was applied to avoid too short transitions from one stage to another. The outputs were first low-pass filtered with a moving average filter of order 30 and then with a median filter of order seven.

Table 2. The most descriptive features in pairwise sleep stage classification (FR >1).

Experiment	Selected features
Awake vs Sleep	Mobility, Delta relative power, Beta Relative Power and SISA
Deep sleep vs light sleep	Complexity, Delta Relative Power, Entropy
Deep sleep vs Rest	Mobility, Complexity, Delta Relative Power, SWI, TWI and Entropy
REM vs NREM	No feature reached FR > 1
Rem vs each other class	Mobility, Delta, Theta and Alpha relative power, SISA, TWI and Entropy

Sonification method based on sleep stage classifier

Various pairwise sleep classifiers were designed based on the results of the feature comparisons described above. The classifiers that were finally taken into use were the classifiers deep sleep stage S4 vs the rest of the classes and stage REM vs all other classes. After the optimal features had been calculated, a Principal Component Analysis transform was applied to improve the classification result. Although dimension reduction was attempted with PCA, better results were obtained when all the eigenvectors of the transformed space were used for classification.

Linear Discriminant Analysis (LDA) was used to project the outputs of the PCA to a one-dimensional line. The classification threshold is set to the point where the classification result is the optimal. This threshold is used as a parameter for the modulation functions (MFs) used to amplitude modulate the output sound for SWS and REM sleep. The MF is called to the range 0..1 with a sigmoid-like function

$$y = \frac{1}{1+e^{-(x-th)}} \quad (2)$$

where x is the output of the LDA classifier, th is the threshold and y is the resulting modulation function value. When x is large, the y value approaches one, whereas y approaches zero when x is small. At the point of the threshold the modulation function output is 0.5.

The output y was felt to need some smoothing, too. Therefore these modulation function outputs were low-pass filtered with a moving average filter of order 40.

Two types of signals were used as “carrier waves” modulated by the modulation function: sine waves with a frequency of 400 Hz and jazz music recording. The sounds were produced separately for REM sleep output and SWS output.

The classifier using all the 18 extracted features with PCA and LDA was tested by dividing the 7614 epochs of recordings from the eight subjects randomly to a training set of 70 % of the data and test set of 30 % of the data. The randomization was performed four times.

Sonification methods based on descriptive features for SWS, REM sleep and wakefulness

Contrary to the previous method PCA transform and LDA classifier were not needed. The outputs of the modulation function were obtained directly from the following formulas with the addition that they were scaled to the range 0..1 in a linear fashion. Three musical instruments, piano to indicate REM, drums to indicate wakefulness and xylophone to indicate SWS were modulated with different modulation functions and the outputs were summed together to a stereophonic output as follows:

$$Left\ channel = drums * VI(O2) + piano * \frac{TWI(O2)}{Activity(Fp1)}$$

(3)

$$Right\ channel = drums * VI(O2) + xylophone * \frac{SWI(O2) \cdot Complexity(O2)}{Mobility(O2) \cdot Entropy(O2)}$$

(4)

The O2 channel was preferred over C4 because it produced slightly better results. Note also the use of the EOG activity in Fp1 in the formula to modulate the piano sound.

Results

The separation of S4 from the other sleep stages succeeded with an accuracy of approximately 96 % in the better O2 channel. This gives a good starting point for the generation of a modulation function for the SWS. The separation of the sleep stages S1 and REM was not very successful and therefore they were combined into one class in classification. When this was done, the combined class S1+REM could be separated from the rest with an accuracy of approximately 93 % in the better O2 channel.

The success in the generation of a good modulation function for SWS can be seen by comparing figures 1 and 2 below. Figure 1 is the hypnogram of test subject number 5. The sleep is normal containing sufficient amounts of both deep SWS sleep and REM sleep. Figure 3 shows the REM classification modulation function.

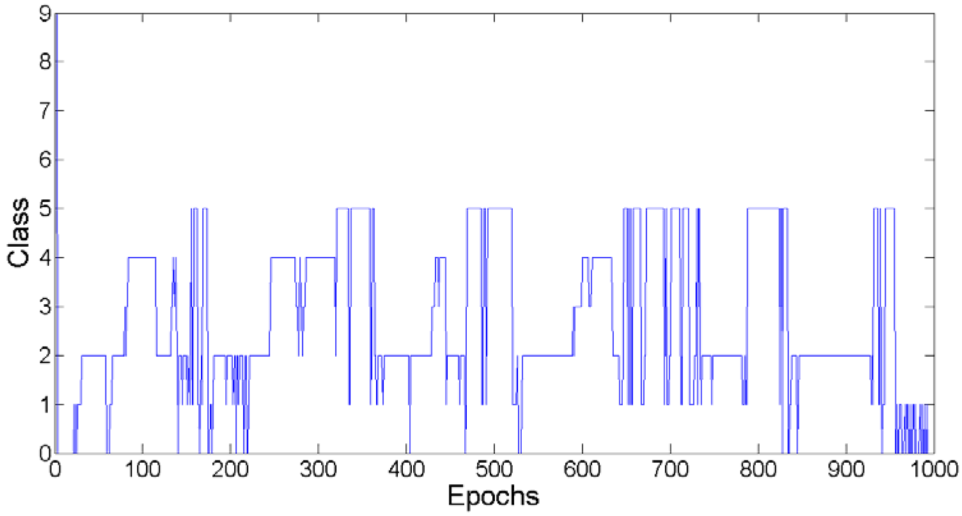

Figure 1. Hypnogram of subject number 5. The classes are 0 = Awake, 1 = S1, 2 = S2, 3 = S3, 4 = S4, 5 = REM. This hypnogram is the result of a visual scoring by experts and the gold standard to which other results are compared.

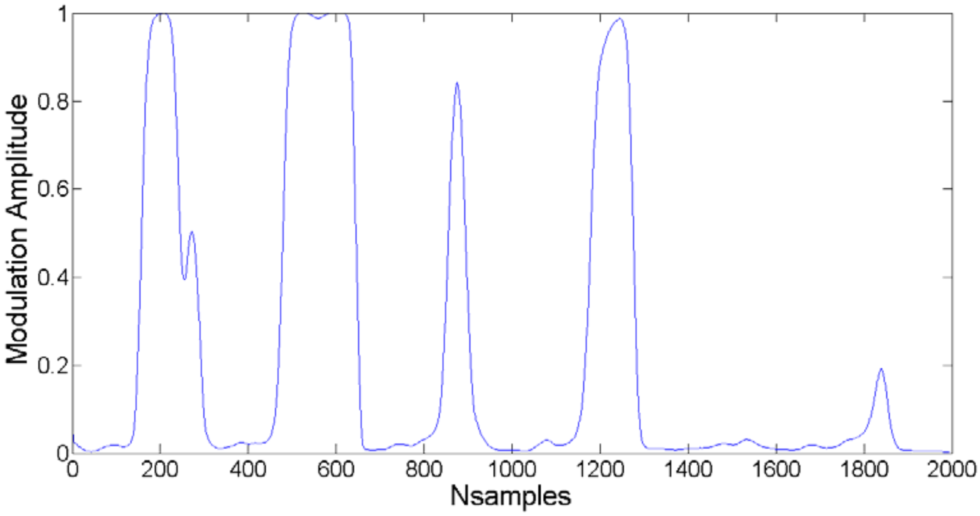

Figure 2. The SWS modulation function generated for subject 5. Note the different horizontal scale to Figure 1 which has 30 s epochs while the epochs are 15 s long in this figure. Note that the modulation function reaches values close to one during S4 sleep which indicates a desired result.

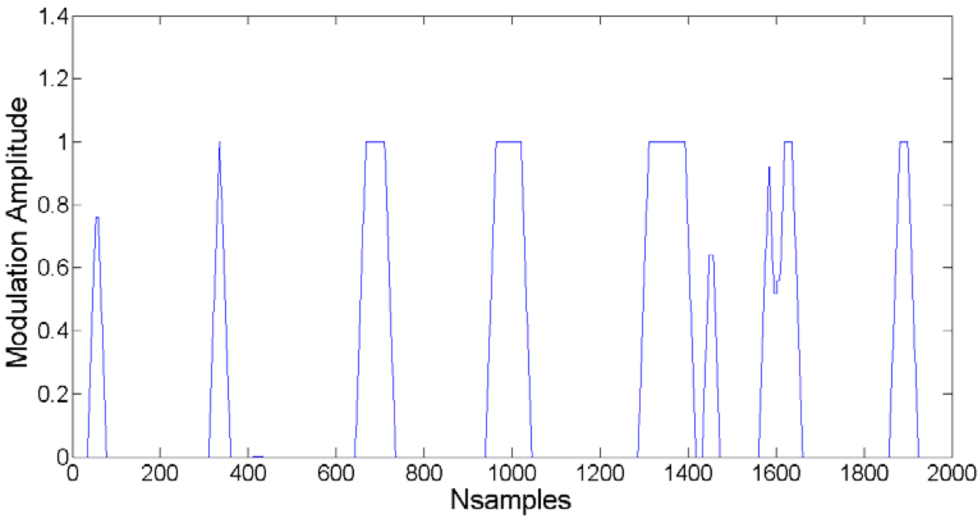

Figure 3. The REM modulation function output for subject 5. The time scale is the same as in Figure 2. Note that the high output values often coincide with the occurrence of stage REM in Figure 1 which indicates a desired result.

The produced sounds were evaluated subjectively. The amplitude modulated sinusoidal signal with different frequencies for REM and SWS outputs indicated the locations of these sleep stages but they were not pleasant to the ear. Moreover, when the amplitudes of both output were low, the listener got the impression that nothing is going on. The experience with the sinusoids did not encourage to continue their use further.

The outputs consisting of piano, drums and xylophone were more pleasant to the listener. It was easy to recognize the phase of the sleep from these recordings. It may, however, require some training to distinguish poor sleep from good sleep with this methodology.

Discussion

The experiments provided invaluable insight into the EEG sonification issue to the authors. For example, hearing the results of the first more primitive methods indicated quickly that the modulated sinusoidal signal output would not be the method of choice approved by human listeners. The use of musical instruments made the listening experience more pleasant. There are, however, moments in the sound output when no sound appears to be active which is a somewhat disturbing phenomenon. Perhaps there should be some background sound indicating S2, for example to fill these empty periods.

We are not yet sure that the sleep disorders can be recognized with the methods described here. Alarms are usually used to indicate the presence of an alarming situation but here the listener should observe the absence of sufficient amount of SWS and REM stage. Maybe this can somehow be improved in the following studies.

The number and variety of subjects from whom we used the recordings was quite small. The number would not be sufficient to test if disturbed sleep can be distinguished from normal sleep, but it was sufficient to recognize that some of the tested sonification methods do not work well and that some deserve to be developed further.

The sonification method based on classification could be developed further. The separation of stages S1 and REM should be improved by extracting more REM related features from EOG and EMG. If the classifier can be made reliable enough, it could, in principle speak out the states in sequence like: "Awake, Stage 1, Stage 1, Stage 2, Stage 2,… Stage REM, Awake.". Taking this idea further could lead to an interpretative statement of the recording like: "In general the sleep recording contained sufficient amounts of SWS and REM sleep but slightly too frequent awakenings towards the end of the recording may have disturbed the sleep somewhat." These methods would then not make use of the human pattern recognition capabilities at all.

Our sonification method differs from [13] by not trying to reach equally high musical quality at the expense of more accurate diagnostic quality and repeatability. Compared to [14] our method aims at making the sleep stage detection easier to the listener by letting the computer leave out some irrelevant details from the produced sound.

Although a lot of experiments were made, we feel that we are still nearer the beginning than the end. The methods contain a lot of tunable parameters which can be adjusted for better performance. We have only started testing different musical instruments but other sounds could be tested as well. The instruments which produce a constant sound during their sound production, like an organ, are probably preferred to those which have a decaying amplitude like an acoustic guitar as some stages can last for a longish period and the guitar sound may get attenuated although the sleep stage continues. After we have made some more progress with this work, we could think of arranging listening tests to laymen to see how easily they can be trained to distinguish poor sleep from good sleep with sonification.

Conflict of interest statement

The authors have no conflicts of interests with the producers of the equipment or producers of the data used in this study. The pseudonymized physiological signal data used in this study was collected according to all ethical and legal requirements of the time of the recordings in the SIESTA project during 1996-2000.

References

[1] Iber C, Ancoli-Israel S, Chesson AL, Quan SF. The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications, 1st ed. Westchester, Illinois, USA: American Academy of Sleep Medicine; 2007.

[2] Rechtschaffen A, Kales A, eds. A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. Los Angeles, California, USA: BI/BR; 1968.

[3] Healthy Sleep. Harvard University. Available from: http://healthysleep.med.harvard.edu/healthy/ [cited 31.12.2014].

[4] Sleep: A Dynamic Activity. American Sleep Association. Available from: http://www.sleepassociation.org/patients-general-public/what-is-sleep/ [cited 31.12.2014]

[5] Brillante R, Cossa G, Liu PY, Laks L. Rapid eye movement and slow-wave sleep rebound after one night of continuous positive airway pressure for obstructive sleep apnoea. Respirology 2012 Apr;17(3):547-53. doi: 10.1111/j.1440-1843.2012.02147.x.

[6] Kramer G, ed. Auditory Display: Sonification, Audification, and Auditory Interfaces. Santa Fe Institute Studies in the Sciences of Complexity. Proceedings Volume XVIII. Reading, MA: Addison-Wesley; 1994. ISBN 0-201-62603-9.

[7] Hermann T, Hunt A, Neuhoff JG, eds. The Sonification Handbook. Berlin, Germany: Logos Verlag; 2011.

[8] Fitch WTS. Sonification system using synthesized realistic body sounds modified by other medically-important variables for physiological monitoring. United States Patent 5730140, March 24, 1998.

[9] Baier G, Hermann T, Stephani U. Event-based sonification of EEG rhythms in real time Clinical Neurophysiology 2007; 118(6):1377-1386.

[10] Baier G, Hermann T, Stephani U. Multi-channel Sonification of Human EEG. In: Martens, Bill (Ed.), Proceedings of the 13th International Conference on Auditory Display p. 491-496, International Community for Auditory Display (ICAD), ICAD, Montreal, Canada, 2007.

[11] Baier G, Hermann T, Sahle S, Stephani U. Sonified Epileptic Rhythms. In: Stockman, Tony (Ed.). Proceedings of the International Conference on Auditory Display (ICAD 2006) p. 148-151. International Community for Auditory Display (ICAD), Department of Computer Science, Queen Mary, University of London, London, UK, 2006.

[12] Hermann T, Baier G. Sonification of the Human EEG. In: Sonic Interaction Design. Franinovic, Karmen and Serafin, Stefania (Eds.). MIT Press, Cambridge, MA, chapter 13, p. 285-297, 2013. Hear samples in http://techfak.uni-bielefeld.de/ags/ami/publications/HB2012-TOT/

[13] Tulilaulu A, Paalasmaa J, Waris M, Toivonen H. Sleep Musicalization: Automatic Music Composition from Sleep Measurements. In: The Eleventh International Symposium on Intelligent Data Analysis (IDA 2012); 2012.

[14] Olivan J, Kemp B, Roessen M. Easy listening to sleep recordings: tools and examples. Sleep Medicine 2004;5:601–603.

[15] Klösch G, Kemp B, Penzel T, Schlögl A, Rappelsberger P, Trenker E et al. The SIESTA Project Polygraphic and Clinical Database. IEEE Engineering in Medicine and biology 2001;20(3):51-57.

[16] Väljamäe A, Steffert T, Holland S, Marimon X, Benitez R, Mealla S, Oliveira A, Jordà S. A review of real-time EEG sonification research. International Conference on Auditory Display; 2013.

[17] Kaniwa T, Terasawa H, Matsubara M, Makino S, Rutkowski TM. Electroencephalogram steady state response sonification focused on the spatial and temporal properties. The 20th International Conference on Auditory Display; 2014.

[18] Hjorth B. EEG analysis based on time-domain properties. Electroencephalography and Clinical Neurophysiology 1970;29:306-310.

[19] Oh SH, Lee YR, Kim HN. A Novel EEG Feature Extraction Method Using Hjorth Parameter, International Journal of Electronics and Electrical Engineering 2014 Jun;2(2):106-110. doi: 10.12720/ijeee.2.2.106-110.