Time-Frequency Classification of continuous sleep stages

PDSB 2019 – Group 9
António Gonçalves - IST181569
Francisco Cachado- IST181641

1. ABSTRACT

Sleep medicine has developed a method to classify different sleep stages in order to better understand what is going on during sleep. This classification is based on the visual scoring of polysomnography data by a physician and is very time consuming due to the large amount of data. To address this problem, automatic computational classifiers were designed. Most of the classifiers designed for this classification were built based on frequency analysis or based on time analysis but only a few consider both dimension and perform a time-frequency analysis. In our project we use this strategy and build our own unsupervised sleep stage classifier. We use the moving average algorithm for noise removal followed by the calculation of the instantaneous frequency. We then use the values of the instantaneous frequency as features and classify the different sleep stages based on them. We were able to achieve an accuracy of 34,29% which is better than choosing a label at random but still has a poor performance that might be explained by the difficulty of detecting sleep spindles and K complexes in the Non-REM stage 2 in automatic/continuous methods.

2. PROBLEM AND MOTIVATION

Sleep is an activity that is common to all living animals. It is clear that sleeping is extremely important for proper functioning during waking hours and its disregard leads to cognitive damage, immunological imparity, compromised psycho-motor capabilities, etc.[1]

What is unclear, is its biological function, why it is necessary and what really happens when we sleep and its relation to the waking hours. Why do we need to sleep a third of our life? Is it better to sleep only once per day or should we consider naps? What is the psychological and biological effect of long-term sleep deprivation or sleep reduction? [2]

These are all fascinating question and deserve to be looked into. It is also estimated that more than 50 million people in the US alone suffer from some kind of sleep disorder such as insomnia, narcolepsy, nightmare disorders, etc and there a medical and social need and interest to look into them.[3,4]

During the investigation of what happen when we sleep different sleep phases were perceived and therefore a need of characterizing and identifying them based on physiological parameter was created. The first attempt to create such tool was the based on visual scoring of the pattern of the values provided by physiological measurements over time. This first method was developed by Rechtschaffen and Kales in 1968, 15 years after the first REM sleep was found.[5]

This method persisted till in 2007, after 28 years of medical progress and innovation, a taskforce made of physician and specialists made a review of all the literature and took in their clinical experience and came up with a new method of classification based on visual scoring. This is the method that is currently used in clinical practice.[6]

Since 2007, technology has evolved and with new developments in electronics and signal acquisition, more specifically, the ability to record big amounts of data from many different biomedical sensors, fast processing power in newer computers and new developments in signal processing theory and its algorithms; it is now possible to make this once laborious task, a more specific and automatic one instead of manually visual scoring every recording from a sleep study.[7]

This paper presents another way of performing this visual classification that could be applied to current clinical practice sleep evaluations.

3. BACKGROUND AND RELATED WORK

The different stages of sleep classification are relaxed wakefulness or awake, non-REM stage 1 or N1, non-REM stage 2 or N2, non-REM stage 3 or N3 or REM or rapid eye movement. Between every stage there a change in the physiological measurements performed and they can be captured in a polysomnography (PSG) test.

The stage awake is described by having present in its EEG recordings alpha waves, brain waves that are characterized by a frequency between 8 and 13 Hz. In this stage there also the presence of some muscular movement in the EMG and some very slow movement recorded by the EOG.

The stage N1 marks the official beginning of the sleep phenomenon and is described by having slower brain waves than the previous stage, which are the theta waves, waves that have a frequency between 4 and 8 Hz. There is also some muscular movement on the EMG.

The stage N2 is the one that follows and is characterized by the presence of gamma waves and, most importantly, sleep spindle and K complexes. There should be no movement on the EMG.

The stage N3 is where delta waves, waves with frequency bellow 4Hz, start to appear and the sleep spindles and K complexes disappear. There should be no movement on the EMG.

The REM stage is a characterized by, as the name mentions, rapid eye movements detected on the EOG and, also, by sawtooth waves which are a mix of alpha and theta waves in the EEG. There should be no movement on the EMG.[6]

Most methods for the automatic classification of sleep stages according to this classification use either time or frequency characteristics but not both. In this project we will use time-frequency analysis for the automatic classification of sleep stages. Our project consists on data integration, noise removal, feature extraction and final classification.

4. APPROACH AND UNIQUENESS

4.1 Material

The first part of our project was data integration.

We received the files of a recorded PSG which started at 01h:20m:29s and finished at 07h:59m:49s having a duration of 6h:39m:20s which is equal to 21 600 + 2340 + 20 = 23 960 seconds. The sampling frequency was 1024 Hz so the total number of samples was 24 535 040 data points. In the classification system described above an epoch is made up of 30 seconds which mean 30 720 data points. This tells us that we have 799 epochs in our data. To each epoch a sleep stage was attributed after visual scoring classification.

The files received were: an xml file, which contained metadata about the PSG like the classification of each epoch and its durations among with other not so relevant information, another xml, which makes the translation of the previous xml into a readable format for the classification stages, and an edf file which is where the actual values for each datapoint was stored.

This edf file was made of 10 channel: 6 EEG (electroencephalography) channels, F3, F4, C3, C4, O1 and O2, 1 EMG (electromyography) channel, 1 EOG (electrooculography) channel and 2 ECG (electrocardiography) channels, as represented in fig.1.

This makes our data a matrix of 10 rows by 24 535 040 columns classified every 30 720 data points into 1 of 5 sleep stages.

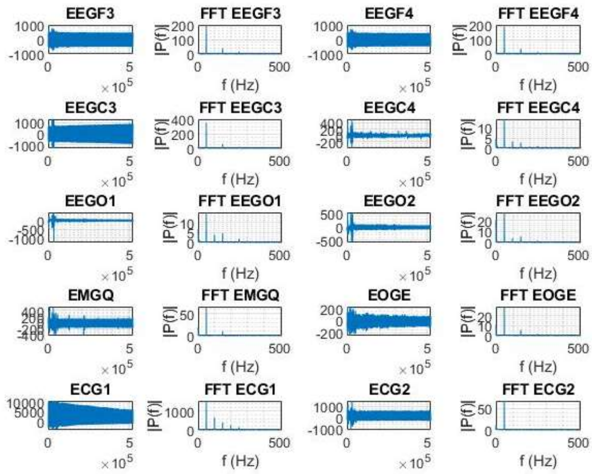

Fig. 1 - Input signal

4.2 Methods

As mentioned before, our project consists on data integration, noise removal, feature extraction and final classification. Data integration was previously described. 2 different methods were followed.

First Method:

For noise removal, the algorithms used were Moving Average followed by filtration using a Butterworth filter.

Moving Average is a correction that receives as inputs a predefined number of samples, sums all of them and divides by the number of samples:

$$x_j = \frac{1}{n}\sum_i^n x_i$$

This correction is done for every single data point in the PSG signals and help to remove random noise from our signal by averaging high values with the rest of the signal. We performed this algorithm twice consecutively with two different time windows (n) being our choice was 20 and then 15 data points because this was the best way to remove the noise from our data. The matlab function used to perform the moving average is MA.[8]

After that we use the Fast Fourier Transform (FFT) to see how the signal behaved in the frequency domain and it was noted that there was still some frequencies above 50 Hz that were most likely due to noise. The FFT is an algorithm that implements the DFT in way that its performance is faster changing from $O(n^2)$ to $O(n\log(n))$. The formula for the discrete Fourier Transform is:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{i2\pi}{N}kn}$$

The matlab function used was fft.[9]

In order to correct for that a Butterworth filter was designed. This specific type of filter was designed to have a frequency response as flat as possible in the passband. The filter was designed using the matlab function designfilt [10] and then the proper filtration was performed using the matlab function filtfilt [11].

For feature extraction, we first had to know which parameters to would be necessary, our features. Since we are performing a time-frequency analysis our features must be related to both time and frequency parameters. Because of that, we chose to utilize the instantaneous frequency of our nonstationary signal, as it is a parameter that calculates the average frequency of the signal as it evolves in time, being this way time varying. The instantaneous frequency is calculated as the first conditional spectral moment of the time-frequency distribution of the input signal:

$$f_{inst}(t) = \frac{\int_0^\infty fP(t,f)df}{\int_0^\infty P(t,f)df}$$

As it can be seen in the formula above, the calculation of the instantaneous frequency requires the calculation of the power spectrum P(t,f). This computation is done by: First, the signal is divided in equal length segments that may or may not overlap. Then the spectrum of each segment is calculated using the short-time Fourier Transform:

$$STFT\{x[n]\}(m,\omega) = \sum x[n]w[n-m]e^{-j\omega}$$

Finally the segment spectrum is used to create the spectrogram.

The matlab function used to calculate the instantaneous frequency is instfreq[12].

Second Method:

For noise removal, the algorithms used were Butterworth filter with lowpass configuration in order to remove all frequency higher than 45 Hz, then was applied a Moving Average with a length of 5 samples, and for the EOG signal was applied consecutively a Moving Average with 20 samples length, resulting in the signal represented in fig. 2.

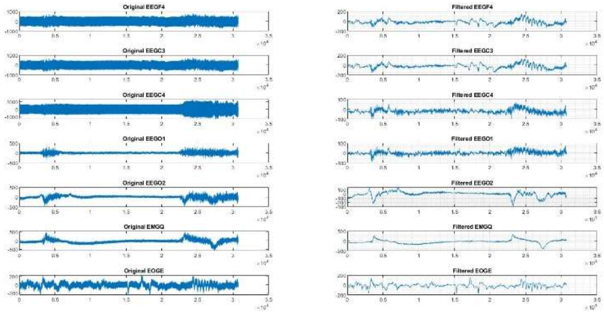

Fig. 2 - Filtered Signal

After the noise removal, was applied DWT, Discrete Wavelet Transform. In the EEG signal was used *db4* as the coefficients used for further analysis was D5, D6, D7, D8, D9, D10. For the EOG signal was used *bior 3.3* [15] and the selected coefficients was D4, D5, D6.

The following statistical features are derived from the coefficients of the DWT referred above using the mathematical equation (1)-(3).

(i) Mean value (MV):
Mean value is a measure of frequency information of the signal. This can be calculated using equation (1):

$$MV = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad (1)$$

(ii) Average power (AP):

This feature provides information about the frequency content of the signal and the mathematical expression:

$$AP = \frac{1}{N}\sum_{i=1}^{N} |x_i|^2 \qquad (2)$$

(iii) Standard deviation (SD):
Standard deviation represents the amount of change in the
frequency of the signal and calculated using the equation (3):

$$SD = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu)^2 \qquad (3)$$

4.3 Proposed solution

Classification of the first method:

After feature extraction we now have a matrix of 10 rows by 799*518 = 413 882 columns. We will now begin the classification part.

For each epoch in the EEG channels, the number of times the IF was either in the range between 13 to 8 Hz (alpha waves), in the range between 8 and 4 Hz (theta waves) and in the range bellow 4Hz (delta waves) was counted. This gave us a view per epoch of the percentage of each wave type. The value was then averaged across the six EEG channels for the same epochs.

For the EMG and EOG channels it was both calculated the average of the whole signal and the average per epoch.

The classification was this one:

If an epoch had a majority of alpha waves and an EMG signal average higher than the EMG whole signal average, then it was classified as Motion.

If an epoch had a majority of alpha waves and a low EMG signal but an EOG signal average higher than the whole EOG signal average then that signal is classified as REM.

If an epoch had a majority of alpha waves but didn't fill the previous two categories, then it was classified as Awake.

If an epoch had a majority of theta waves and if the percentage of alpha waves plus theta waves was higher than the percentage of theta waves plus delta waves, then it was classified as N1.

If an epoch had a majority of theta waves and if the percentage of alpha waves plus theta waves was lower than the percentage of theta waves plus delta waves, then it was classified as N2.

If an epoch had a majority of delta waves and an EOG signal average higher than the whole EOG signal average, then that signal is classified as REM.

If an epoch had a majority of delta waves and an EOG signal average lower than the whole EOG signal average, then that signal is classified as S3.

Classification of the second method:

After the features acquirement all the values were organized in a table, in each row is a different epoch and each column is one different feature. The table with all epochs is a table of 799 rows by 117 columns.

The table is then clustered using the k-means algorithm already available in matlab [13].

5. RESULTS AND CONTRIBUTIONS

The results of the previously described classifier in the first method were compared to the classification made by visual scoring (labels) in order to determine the accuracy of our method. Accuracy for the case when the same label was assigned to all the data and for the case when a label was picked at random (using the matlab function randi) were also computed. The results are shown in table 1.

Table 1. Accuracies achieved by the our method and by assigning the different labels to all our samples.

	Accuracy
Classifier	34,29%
Motion	2,38%
Awake	7.88%
N1	4,51%
N2	27.66%
N3	40,05%
REM	17,52%
Random Classification	16,39%

The results of the second method when compared to visual scoring are described in the follow image, fig.3, a confusion matrix where class 1 to 7 correspond to the following classes {Motion, Awake, S1, S2, S3, S4, REM}.

Truth data

Classifier results

	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Classification overall	Producer Accuracy (Precision)
Class 1	1	2	0	6	0	0	9	18	5.556%
Class 2	0	29	0	8	1	0	25	63	46.032%
Class 3	1	0	2	10	1	0	22	36	5.556%
Class 4	1	3	2	146	19	1	49	221	66.063%
Class 5	1	1	0	65	58	5	12	142	40.845%
Class 6	0	1	0	62	25	88	2	178	49.438%
Class 7	1	2	0	8	0	0	129	140	92.143%
Truth overall	5	38	4	305	104	94	248	798	
User Accuracy (Recall)	20%	76.316%	50%	47.869%	55.769%	93.617%	52.016%		

Fig. 3 - Confusion Matrix

As it is possible to see in the image the classifier had a 93% of success classifying the REM sleep, however had a poor result classifying motion and S1 stage, with only 5% success.

The overall accuracy of k-means is 56,77%.

6. DISCUSSIONS AND CONCLUSIONS

Continuous classification of sleep stage is a much-desired technology due its time-consuming nature. But for that to be possible, the method should have a high accuracy. The accuracy we obtained in the first method was 34,29 % and in the second one was 56,77%, although having a much higher accuracy it is very small when compared to other classifiers developed for the same task.[14]

Despite this, it was better at classification than randomly assigning a label to the data and better than selecting one of the labels for the whole dataset except for the N3 label.

This might be due to the classification technique used in our method which makes it hard to gather all relevant features specially the time-frequency analysis of sleep spindles and K complexes.

These two complex waves are very determinant in the distinction between non-REM stage 2 and stage 1 which account for more less 30 percent of our dataset labels and might be a reason why we got a low performance.

Another reason for such low performance might be the difficulty in correctly assessing the REM stage due to its physiological variability.

Other methods in the literature that are successful in this classification do not focus only on certain properties of the signal like we did for the time-frequency, but borrow information across multiple parameters like temporal-spatial information, z-scores and other statistical measurements, etc...[7]

Other reason might be the choice of the classifier since most of the successful classifiers used more sophisticated techniques such as neural networks, support vector machines and other machine learning techniques. [7,14]

7. REFERENCES

[1] Vyazovskiy, V. V., & Harris, K. D. (2013). Sleep and the single neuron: the role of global slow oscillations in individual cell rest. Nature reviews. Neuroscience, 14(6), 443–451. doi:10.1038/nrn3494

[2] Cirelli C, Tononi G. Is sleep essential? PLoS Biol. 2008; 6:e216. [PubMed: 18752355]

[3] Institute of Medicine. Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem. Washington, DC: The National Academies Press; 2006.

[4] Treatment of sleep dysfunction and psychiatric disorders. Becker PM, Department of Psychiatry, University of Texas Southwestern Medical Center at Dallas, 5477 Glen Lakes Drive, Curr Treat Options Neurol. 2006 Sep;8(5):367-75.

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

[6] Silber MH; Ancoli-Israel S; Bonnet MH et al. The visual scoring of sleep in adults. J Clin Sleep Med 2007:3(2);121-131

[7] Sen, Baha & Peker, Musa & Cavuşoğlu, Abdullah & Celebi, Fatih. (2014). A Comparative Study on Classification of Sleep Stage Based on EEG Signals Using Feature Selection and Classification Algorithms. Journal of medical systems. 38. 18. 10.1007/s10916-014-0018-0.

[8] https://www.mathworks.com/help/finance/movavg.html

[9] https://www.mathworks.com/help/matlab/ref/fft.html

[10] https://www.mathworks.com/help/signal/ref/filtfilt.html

[11] https://www.mathworks.com/help/signal/ref/designfilt.html

[12] https://www.mathworks.com/help/signal/ref/instfreq.html

[13] https://www.mathworks.com/help/stats/kmeans.html

[14] V. Nigam and D. Graupe, A neural-network-based detection of epilepsy, Neur. Res. 26(2004) 55–60.

[15] A, Dr. Naga Rajesh & Chandralingam, S & Anjaneyulu, Tangirala & Satyanarayana, K. (2012). Denoising EOG Signal using Stationary Wavelet Transform. Measurement Science Review. 12. 46-51. 10.2478/v10048-012-0010-0.

Author(s) Bio:

António Maria Forsado Martins Gonçalves

ist181569

Francisco Marques Sousa Cachado

ist181641