ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1.
Available from http://www.i6doc.com/en/.

Outlining a simple and robust method for the automatic detection of EEG arousals

Isaac Fernández-Varela[1], Diego Álvarez-Estévez[2],
Elena Hernández-Pereira[1] and Vicente Moret-Bonillo[1] *

1- Universidade da Coruña - Departamento de Computación
Facultade de Informática, Campus de Elviña, A Coruña - Spain

2- Sleep Center & Clinical Neurophysiology - Haaglanden Medisch Centrum
Lijnbaan 32 2512 VA, The Hague - Netherlands

Abstract. This work proposes a new technique for the automatic detection of electroencephalographic (EEG) arousals in sleep polysomnographic recordings. We have developed a non-computationally complex algorithm with the idea of providing an easy integration into different software platforms. The approach combines different well-known signal analyses to identify relevant arousal patterns. Special emphasis is carried out to produce a robust, artifact tolerant algorithm. The resulting approach was tested using a database of 6 polysomnographic recordings from real patients, achieving an average kappa index of 0.77 with respect to the visual scorings made by clinical experts.

1 Introduction

Analysis of the sleep microstructure has a fundamental importance in the diagnosis of sleep disorders. Polysomnography (PSG) is the standard clinical test for the monitoring of the patient's biosignals during sleep. As with this test, reporting on the occurrence of electroencephalographic (EEG) arousals is of relevance for the good assessment of the patient's sleep continuity [1].

According to the American Academy of Sleep Medicine (AASM), an EEG arousal is an abrupt shift in the EEG frequency including alpha, theta and/or frequencies greater than 16 Hz (but not spindles) that lasts at least 3 seconds, with at least 10 seconds of previous stable sleep. Also, during Rapid Eye Movement (REM) sleep, a concurrent increase in the submental electromyogram (EMG) is required [1]. Visual analysis of the whole night neurophysiological signals is complex and very time-consuming for the clinician. To solve this problem, during the last years different works have explored the possibility of developing automatic detection methods [2, 3, 4, 5].

This work proposes a new method based on a multi-channel analysis context. The goal is to build a robust, efficient, but at the same time simple method, to automatically score EEG arousals and help the clinician in the PSG examination task. The method is designed to allow its easy integration into different applications. To meet this goal classical signal processing routines, which can be implemented in any programming language, are used. The resulting algorithm is configurable and can be executed using one EEG and one chin EMG signal.

*This work was partially funded by the Xunta de Galicia (Grant Code GRC2014/035) and by the Spanish Ministerio de Economía y Competitividad, MINECO, under Research Project TIN2013-40686P, both partially supported by the European Regional Development Fund, ERDF.

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1.
Available from http://www.i6doc.com/en/.

2 Proposed Method

The proposed method is structured according to the schema presented in Figure 1. At the beginning, the algorithm carries out a signal conditioning, followed by a preliminary detection of arousal events using a frequency based identification criterion. Then EEG arousal patterns are detected studying the features of each event individually. Finally, false positives are discarded by examining the resulting patterns within the context of clinical definitions.

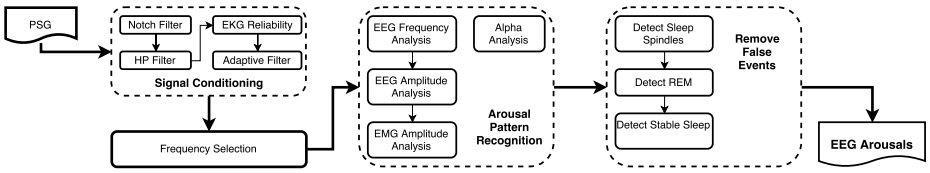

Fig. 1: Proposed method's structure.

2.1 Signal Conditioning

At this step both the EEG and the EMG signals are Notch filtered for mains interference (in our case 50 Hz), and further a high-pass filter (cut-off 15 Hz) is applied to get rid of low frequencies not related to the chin muscle activity. Detailed information over the implementation of these two filters can be found in [6].

Removal of electrocardiogram (EKG) artifacts follows in the resulting EEG and EMG signals, and for this purpose an adaptive filtering algorithm is implemented. First, the EKG beat series is obtained by using a standard QRS detection algorithm. An adaptive filtering template T is then used updating the corresponding signal (EEG or EMG) values, 0.5 s around each QRS peak, using a memory factor α of 0.1, i.e., $T[t] = (1-\alpha) * T[t-1] + \alpha * Signal[t]$. At each step, the resulting filtering template which contains the current EKG artifact is subtracted from the original signal.

2.2 Frequency Based Event Selection

To select the events with a frequency change, power analysis of the EEG is carried out using the Short-Time Fourier Transform. For this purpose a $3\,\mathrm{s}$ sliding window (hamming transformation, $0.2\,\mathrm{s}$ shifting step) is used and, for each window, an estimation of the power contained in four different frequency bands is obtained, namely theta (4-7 Hz), alpha (8-12 Hz), spindle (12-16 Hz), and beta ($>$ 16 Hz); power content estimation is carried out using the periodogram.

For each frequency band, the corresponding baseline is created averaging the values from the previous 10 s. Finally, preliminary EEG arousal events are marked depending on the power increases over the corresponding baseline. Specifically, relative increments in the alpha (2.5 times over the alpha baseline) or the beta (2 times over the beta baseline) bands are used to mark the events.

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1.
Available from http://www.i6doc.com/en/.

2.3 EEG Arousal Pattern Recognition

Considering that the beginning of each EEG arousal is triggered by the corresponding frequency change, found in the previous step, we need to determine its real end. For this purpose we try to add consecutive windows to the marked event end, checking after each addition that the resulting event is still a recognizable arousal pattern. We follow different approaches depending on the main band where the event was found (beta or alpha) as pattern characteristics might differ on each case.

2.3.1 Arousals in the Beta Band

To analyze the events in the beta band we use 3 methods. The first two study the EEG signal, analyzing its frequency and amplitude, while the third one studies the EMG amplitude.

EEG Frequency Analysis From the list of preliminary arousal events detected in the beta band we characterize their main event frequency $x[n]$ as:

$$\mathrm{f}(x[n]) = \frac{\mathrm{c}(x[n])}{j-i}, i \leq n \leq j \tag{1}$$

where

$$\mathrm{c}(x[n]) = |\{\mathrm{d}_1(\frac{|\mathrm{d}_1(x[n])|}{\mathrm{d}_1(x[n])}) > 0\}|, \mathrm{d}_1(x[n]) = x[n] - x[n-1]$$

Every event that last less than 3 s is extended, appending to its end consecutive $0.5\,\mathrm{s}$ windows, while $|f_0 - f_e| < th_\beta$, where f_0 is the original event main frequency, f_e is the extended event main frequency, and $th_\beta = 0.6$ is an empirically selected threshold.

EEG Amplitude Analysis Even though the AASM definition for an EEG arousal does not establish any requirement regarding the amplitude of the EEG, these events are often accompanied by a change in the EEG amplitude. Using this fact, an amplitude increase is detected when there is a positive difference between the peak-to-peak amplitude of the event and the reference mean peak-to-peak amplitude (μ_A) from the 5 previous $1\,\mathrm{s}$ windows. We extend the original event, appending $1\,\mathrm{s}$ consecutive windows until the peak-to-peak amplitude of the extended event does not exceed $th_{EEG} \times \mu_A$, where $th_{EEG} = 4$ is an empirically selected threshold.

EMG Amplitude Analysis EEG arousals frequently (but not always) cause the concurrent activation (an amplitude change) of the chin EMG. To detect it, we compare the EMG amplitude during the last second of the EEG event ($\mu_{EMGevent}$) to the signal amplitude of the segment surrounding the event ($\mu_{EMGref\pm 15s}$). In both cases, amplitude reference values are obtained by taking the median of 0.1 s window's peak-to-peak amplitudes. We assume there is EMG activation and extend the event if μ_{EMGref} is higher than $th_{EEG} \times \mu_{EMGevent}$. The extension is made appending the following $0.1\,\mathrm{s}$ windows to the end of the original event while the condition holds.

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1.
Available from http://www.i6doc.com/en/.

2.3.2 Arousals in the Alpha Band

EEG arousals in the alpha band are characterized by wax and wane oscillations in the alpha range with seldom interference from other frequency bands. This pattern is recognized following a similar approach to the aforementioned frequency analysis. In this case, to improve the computational performance, alpha events with less than $1\,\mathrm{s}$ duration are directly discarded, limiting the event's end search period to 6 s after its start. Given an initial event with frequency f_0, we find the longest possible extended event with frequency f_e that verifies $|f_0 - f_e| < 0.05 \times f_0$.

2.4 False Positive Events Removal

To discard possible false positives, several methods are applied. These post-processing methods are inspired by the EEG arousal standard definition found in clinical guideline [1].

Discarding by presence of Sleep Spindles Each arousal pattern previously detected is examined using a 0.5 s sliding window (the minimum spindle duration) with a 0.1 s shifting step. Using (1) the periods within the 12-15 Hz band are marked. The detected arousal pattern is discarded if there is no segment lasting at least $3\,\mathrm{s}$ once the spindle periods are removed.

Discarding by absence of EMG activity during REM periods During REM sleep periods a EEG arousal requires the presence of at least $1\,\mathrm{s}$ of concurrent EMG activation. We perform the detection of REM periods using the low muscle tone characteristic, under the assumption that REM sleep is the one with the lowest EMG amplitude [1]. For that purpose we use a rectified version of the filtered EMG signal and calculate the overall minimum peak-to-peak amplitude on an epoch ($30\,\mathrm{s}$ window) basis. Epochs with a peak-to-peak amplitude lower than 2.5 times the aforementioned minimum are estimated to belong to REM sleep. Using this reference, EEG arousal events marked during REM periods that do not show sufficient EMG activation are discarded. To detect this activation we compare the corresponding event's EMG peak-to-peak amplitude against the overall amplitude during the whole epoch. If the event amplitude is equal or lower than the epoch's amplitude, we assume there is no activation and we discard the event.

Discarding by absence of 10 seconds of previous stable sleep We discard the events that appear after another event happening in the previous 10s. Also, we discard events detected within an epoch marked as awake having the previous 10s in an awake stage. Determination of sleep and wake epochs is done using as input the the expert's hypnogram information.

3 Validation and Results

A dataset of 6 PSG recordings was used for the validation of our method. The recordings were randomly selected from the patient database at the Sleep Center, Haaglanden

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1. Available from http://www.i6doc.com/en/.

Medisch Centrum (The Netherlands). The study obtained the approval of the Medical Ethics Committee of the Southwest Holland region under reference METC 16-027. Signals were analyzed offline by clinical experts, scoring the sleep stages and the EEG arousal events. All the procedures were performed according to the last version of the AASM guidelines [1]. For the validation of our method one EEG derivation (C_4/A_1), the submental EMG, and the single-channel modified lead II EKG derivation (for artifact rejection) were used. All signals were sampled at 256 Hz. In total, the dataset contains 2728 minutes of data, with 458 arousals scored by experts.

The validation was made on a $30\,\mathrm{s}$ epoch basis. Every EEG arousal was assigned to an unique epoch, using the event's middle point. Table 1 shows the validation results.

	# Arousals										
RN	Expert	System	TP	FP	TN	FN	Error	Sens	Spec	F_1 score	Kappa
1	51	57	43	14	877	8	0.023	0.843	0.984	0.796	0.784
2	148	137	103	34	634	45	0.097	0.696	0.949	0.723	0.664
3	45	45	34	11	976	17	0.027	0.667	0.989	0.708	0.694
4	24	27	22	5	925	2	0.007	0.917	0.995	0.863	0.859
5	61	52	41	11	818	20	0.035	0.672	0.987	0.726	0.707
6	129	124	115	9	678	14	0.028	0.891	0.987	0.909	0.892
Total	458	442	358	84	4908	106	0.034	0.781	0.982	0.787	0.767

Table 1: Epoch-based validation for the detection of EEG arousals. RN = Recording Number; TP = True Positive; FP = False Positive; TN = True Negative; FN = False Negative

Threshold selection in our method was optimized to increase the sensitivity while maintaining the number of false positives reasonably low. This is because usually the number of EEG arousals contained in a sleep recording is highly reduced in relation with the total recording time. In this context, validation measures that take into account the imbalance of the classes are preferred, such as the F1 score or the Kappa index. Moreover, achieving a high specificity or a low error is relatively easy, but not representative. As an example, for the same dataset, a trivial method that classifies all the epochs as "free of arousal" obtains an error of 0.084 and a specificity of 1.00.

4 Discussion and Conclusions

EEG arousals are one of main the causes of sleep fragmentation and thus their quantification is important in the field of sleep medicine. Visual inspection of the PSG to score these events is a complex task. This work presents an automatic method to help the clinician during the scoring task. The proposed solution tries to avoid computationally intensive techniques that limit its application. Thus, we have relied in the use of well-known signal processing techniques which can be easily implemented in any programming language.

In our method, preliminary events are detected finding abrupt frequency changes. Then, each initial event is studied to find a recognizable EEG arousal pattern that

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1.
Available from http://www.i6doc.com/en/.

matches the standard AASM definition [1]. Within the described steps, some empirically established thresholds are used. For the preliminary event marking, in general, higher threshold values would imply the necessity of more abrupt frequency changes to detect an event. Besides, by using higher threshold values we loose sensitivity towards the less clear cases. The selected thresholds, therefore, were optimized to maximize the sensitivity-precision trade-off. Moreover, as usually the number of EEG arousals is highly reduced in relation with the total recording time, the validation procedure should take into account the imbalance of the classes. Therefore we have introduced the use of the F_1 score and the Kappa Coefficient as reference validation measures.

The comparison of our results to previous works in the literature is not easy due to the general lack of a standard benchmark or methodology. To minimize this problem we followed the epoch based validation used in [3] and in [4], so our results are somehow comparable to theirs. Specifically in [3], a sensitivity and specificity values of 0.86 and 0.76 are reported, with a best classification error of 0.20. Even though the sensitivity achieved in this work is lower than theirs, we outperformed their classification error and specificity. In [4], the authors reported a sensitivity of 0.72, a specificity of 0.89 and an error value of 0.13. All these measures are outperformed with our current method. Unfortunately, we cannot obtain their F_1 score or Kappa coefficient to compare them with ours.

The results achieved in this work are encouraging, yet there is room for improvement. More emphasis on the signal conditioning stage for artifact removal is needed, as well as extending the testing dataset for a broader evaluation of the method.

References

[1] American Academy of Sleep Medicine, R.B Berry, et al. *The AASM manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications, Version 2.3*. American Academy of Sleep Medicine, 2016.

[2] T. Sugi, F. Kawana, and M. Nakamura. Automatic eeg arousal detection for sleep apnea syndrome. *Biomedical Signal Processing and Control*, 4(4):329–337, 2009.

[3] Diego Alvarez-Estevez and Vicente Moret-Bonillo. Identification of electroencephalographic arousals in multichannel sleep recordings. *IEEE Transactions on Biomedical Engineering*, 58(1):54–63, 2011.

[4] Isaac Fernandez-Varela, Elena Hernandez-Pereira, Diego Alvarez-Estevez, and Vicente Moret-Bonillo. Automatic detection of EEG arousals. In *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2016)*, pages 235–240, apr 2016.

[5] Dorothée Coppieters t Wallant et al. Automatic artifacts and arousals detection in whole-night sleep EEG recordings. *Journal of Neuroscience Methods*, 258:124–133, 2016.

[6] Diego Alvarez-Estevez. A new automatic method for the detection of limb movements and the analysis of their periodicity. *Biomedical Signal Processing and Control*, 26:117–125, 2016.