Detection of Obstructive Sleep Apnea by Empirical Mode Decomposition on Tachogram

B. Mijović[1], J. Corthout[1], S. Vandeput[1], M. Mendez[2], S. Cerutti[2], S. Van Huffel[1]

[1] Dept. Of Electrical Engineering – ESAT/SCD, Katholieke Universiteit Leuven, Leuven, Belgium
[2] Dept. of Biomedical Engineering, Politecnico di Milano, Milano, Italy

Abstract — **In this study we proposed a method for screening the obstructive apnea in an easy and non-invasive way by following the tachogram. Ensemble Empirical Mode Decomposition (EEMD), as proposed in [6] is applied, Intrinsic Mode Functions (IMF) are derived and amplitudes and frequencies from different modes are extracted. Frequencies are computed by applying a new Generalized Zero Crossing method on IMF as proposed by Huang [7]. Once frequencies and amplitudes of different modes have been extracted, the optimal performing subset of features has been derived. Linear discriminant analysis (LDA) was used to classify apneic events on a minute-by-minute basis. Our algorithm showed fairly good performance with sensitivity of nearly 89% and accuracy higher than 83% which is at least comparable with the best performing methods. However, the set of features used in this work was significantly smaller than in other studies, e.g. Corthout et al. [5]. In this study only EMD-based methods are used in order to show the power of that family of algorithms, and only the tachogram is processed, so the computational power of the algorithm is enhanced.**

Keywords — **Empirical Mode Decomposition, Obstructive Sleep Apnea, Heart Rate Variability, Automatic Detection**

I. Introduction

Sleep Apnea is a sleep disorder characterized by pauses in breathing during sleep. There are three different types of sleep apnea: central, obstructive and complex. Obstructive sleep apnea (OSA) is the most common of the different types of all sleep-related breathing disorders with around 84% of the cases [1]. Breathing is interrupted by a collapse of the upper airway at the level of tongue or soft palate, which causes cessations of respiratory flow despite respiratory effort.

Several epidemiological studies have identified OSA as an important risk factor for systematic hypertension, myocardial infraction, stroke, and sudden death [2-4]. Given the high prevalence of sleep-disordered breathing, the potential importance of OSA in contributing to cardiovascular morbidity and mortality is considerable.

The frequent arousals that terminate apneas lead to sleep fragmentation, which is characterized by disruption or loss of deep and REM sleep. Consequently sleep looses its restorative function on physical and mental performance. Patients suffering from sleep apnea report excessive daytime drowsiness and non-restorative sleep even when the sleep period is prolonged, while not being aware that they are suffering from it. Thus, epidemiological and clinical studies are limited, since the patients with OSA usually come to seek clinical attention after the disorder has been present for several years and has left long-term sequels. In addition to significant social and emotional problems caused by poor mental performance, patients can also suffer from a series of psychological problems.

The gold standard procedure for sleep apnea diagnosis is overnight polysomnography (PSG). However, it requires sophisticated systems and attending personnel and therefore is not applicable in home monitoring systems. Due to the costs and scarcity of sleep laboratories, it is estimated that sleep apnea is widely under-diagnosed. In the last decade different techniques which provide apnea diagnosis on simpler measurements than PSG have been explored. Since OSA is a condition with increased activation of the Autonomic Nervous System (ANS) [2], ECG is the one of the most interesting signals to investigate as this is easily measured in a non-invasive way with a high signal-to-noise ratio and it has a close relationship with the activity of ANS.

In an earlier study of Corhout et. al [5], a recent technique Empirical Mode Decomposition (EMD) was applied on Heart Rate Variability (HRV) data and on the signal that consists of the area under the QRS complex, both in order to automatically detect apneic events. That study yielded fairly good results with an accuracy of 88.8%. In this study, an improved EMD technique, called Ensemble EMD (EEMD) [6], is used which is less sensitive to noise. In addition, it uses a new technique for extracting instantaneous frequencies from intrinsic mode functions (IMF's) by means of Generalized Zero Crossing (GZC) [7] in order to provide the diagnosis of OSA solely on HRV without the need to compute the area under the QRS complex data, which would double the computational efficiency. Besides the computational efficiency, many different advanced techniques for precise automatic HRV signal extraction are available, contrary to the amplitude of the area under the QRS complex. One of these algorithms, proposed by Mateo and Laguna [8], is used in our work.

J. Vander Sloten, P. Verdonck, M. Nyssen, J. Haueisen (Eds.): ECIFMBE 2008, IFMBE Proceedings 22, pp. 247–251, 2008
www.springerlink.com © Springer-Verlag Berlin Heidelberg 2009

II. Materials and methods

A. Data Base Description

The data set used in this study comes from the PhysioNet data base freely available via internet. Initially the data base has been made to support Computers in Cardiology Challenge 2000. Data was collected by the sleep laboratory at the Phillips Universität in Marburg, Germany. The data were divided into a learning set and a test set, each consisted of 35 overnight recordings, of approximately 8 hours duration. Each recording contained a single ECG signal digitized at 100 Hz with 12-bit resolution. The learning set recordings were accompanied by the set of reference annotations that indicate presence or absence of apnea on the minute-by-minute basis. These annotations were made by sleep experts on standard criteria. For the test subjects, apnea annotations were withheld for the duration of the challenge [9].

Recordings were subdivided in three groups: Apneic patients, Borderline patients and Control or normal patients. Both learning and test group contain 20 apneic, 5 borderline and 10 control patients.

We randomly selected 25 patients from the learning set to train our classification algorithm. From the test group, another 25 patients were chosen to measure performance of our algorithm. Only 50 recordings were selected due to the bad quality of the data in other recordings which would require additional manual correction and preprocessing of the data. The whole data set may be preferably included in some future studies.

B. Used Methodologies

In contrary to Corthout et al. [5], in this study we tried to detect apnea using only the inverse of HRV signal, which is the time interval between two consecutive R peaks. This signal is referred to as RR signal further in this paper. The RR signal shows characteristic oscillations (brady-tachycardia) during apneic events, producing a low frequency peak in the signal spectrum while rising in the signal amplitude in time domain. The EEMD technique was applied to each RR signal and IMF's have been derived. From IMF's, instantaneous amplitudes and frequencies have been calculated from which certain features have been extracted, used afterwards for classification. From the feature set, the best performing subset was selected and used as an input to LDA.

C. Automatic RR correction algorithm

Each recording in PhysioNet data base was accompanied by *qrs annotations* which contained positions of the R peak in all ECG recordings. This signal is further referred to as a Heart Timing (HT) signal. These annotations were derived by an automatic algorithm, freely available at the PhysioNet website, and were not manually checked. In this study we used an algorithm [8] which automatically corrects HT signal. This algorithm models beat occurrence times from a generalized, continuous time integral pulse frequency modulation (IPFM) model and minimizes the effects of the presence of ectopic beats. First the algorithm detects irregular sinus beats (due to ectopic beats or QRS complex misdetection) that can affect the PSD estimate of the HRV. Then, it minimizes the effect of these types of anomalies in the series of beats and corresponding PSD-derived clinical-indices. The solution is valid both if the ectopic beat has reset the activity of the SA node and if it wasn't affected.

D. Ensemble Empirical Mode Decomposition

Empirical Mode Decomposition (EMD) is a recently developed technique, proposed by Huang et. al. [10] for decomposing any complicated time series into a finite set of Intrinsic Mode Functions (IMF). Intrinsic mode functions are meant to be mono-component, orthogonal to each other and a set of IMF's should be complete. Here the term *mono-component* means that all the IMF's contain only one frequency at the time, which is called *Instantaneous Frequency* (IF). Orthogonal property states that different IMF's do not have similar frequency content. The amplitude of the IMF is the power of the frequency component at the particular time instant and is called *instantaneous amplitude*.

One of the major drawbacks of the original EMD is the frequent appearance of mode mixing, which is defined as a single IMF either consisting of widely disparate scales, or a signal of a similar scale residing in different IMF components. To overcome this problem, the advanced noise-assisted data analysis method is proposed, called Ensemble Empirical Mode Decomposition (EEMD) [6], which defines the IMF as an ensemble of trials, each consisting of the signal plus a white noise of finite amplitude. The algorithm is pretty simple:

1. An amount of white noise of a finite amplitude (in our case 20% of the standard deviation of the signal) is added to the original signal itself
2. EMD is applied and IMF's are derived
3. Steps 1 and 2 are repeated a few times (in our case 30), which leads to ensemble of 30 different sets of IMF's
4. Average over the ensemble in order to obtain a set of averaged IMF's

Taking into account properties of white noise, noisy components are expected to cancel out and a set of noise-free IMF's is derived. Increasing the ensemble would give more accurate estimates of IMF's, but the computational efficiency is decreased.

E. Extracting Frequencies and Amplitudes

After the EEMD has been applied and IMF's have been extracted, frequencies and amplitudes have been computed. The regular way to compute amplitudes and frequencies from the IMF's is to apply the Hilbert Transform (HT). However, in time instants where large and fast changes in amplitude occur, cubic splines during sifting are not able to follow these changes and hence IMF's are not always ideal. Due to this phenomenon, the HT will give non-accurate instantaneous frequencies in those places. Therefore, frequencies in this study were computed using Generalized Zero Crossing (GZC) approach as described in [7]. The GZC approach is the most direct, local, and also the most accurate method in the mean. This mean is localized down to quarter of a wave period, so it is already a well-accepted result, since we anyway detected apnea scores on a minute-by-minute basis which is much greater than the wave period of the used IMF. The algorithm consists in detecting local maxima **1**, local minima **2** and zero-crossings **3**. The time intervals between all the combinations of the critical points (**1,2,3**) are considered as the whole, half or a quarter of a wave period. The mean frequency is then computed as

$$\omega = \frac{1}{7}\left\{\frac{1}{4T_4}+\left(\frac{1}{2T_2^1}+\frac{1}{2T_2^2}\right)+\left(\frac{1}{T_1^1}+\frac{1}{T_1^2}+\frac{1}{T_1^3}+\frac{1}{T_1^4}\right)\right\} \quad (1)$$

Wherein ω is the mean frequency, T_1^x are full periods (x=1, 2, 3, and 4), T_2^y are half periods (y=1 and 2), and T_4 is a quarter period, all enclosing the point under consideration.

F. Extracting Features

After EEMD is performed, 15 IMF's are obtained. From this set we retain only the first 5, since they contain the most useful information about the signal, both in amplitude and frequency. Ratios of a particular mode amplitude (or sum of particular modes amplitudes) mean to the sum of means of all modes amplitudes and standard deviations of particular mode frequencies (or of mean frequency of 2 or 3 modes, as described in [7]) are computed on a minute-by-minute basis.

G. Subset and Classifier Selection

First a broad set of features was formed, and then a limited subset of features was selected in order to achieve the optimal performance, namely an as good as possible accuracy with an as small as possible number of features. Since the initial set of features had 20 different features, all combinations of features were tried to work together. First we tried all combinations of 2 features, then all combinations of 3 features and so on.

We tried several classifiers in order to choose an appropriate one. Linear, Quadratic and Mahalanobis discriminant analysis were explored. LDA classifier showed to be optimal, since it outperformed the others in terms of accuracy, and more significantly in terms of sensitivity.

III. Results

After decomposing a signal, extracting the features, and selecting the best performing subset, LDA classification provided us with accuracy higher than 83%. Classification performance obtained on a test set is shown in Table 2, and compared to results of Corthout et al. [5], since they used the same data base as used in this study. Subsets of chosen features which showed the best performance are shown in Table 1. Concerning the feature set, the biggest difference compared to Corthout's work is in the second feature where the sum of amplitudes of the 4th and 5th mode is used instead of using them separately, since apneic events provide the same effect on these two amplitudes, so in this way the effect is enhanced.

Although the classification performance in this study is still a little bit lower, only 4 instead of 20 or 40 features are used, only one signal instead of 2 is processed using EEMD, and standard HRV measures like pNN50, HRV triangular index and TINN are not used here in contrary of Corthout's study. Apneic and normal/control patients are easily separated, as shown in Figure 1. Apneic and borderline subjects are not so well separated compared to the work of Corthout et al., where it was not possible to separate normal/control from borderline subjects. This is obvious when sensitivity and specificity from these two studies are compared (see Table 2).

Table 1 Selected feature set

Used Four Features	Feature Name
1	Mean of amplitude of the 3rd IMF
2	Mean of sum of amplitudes of the 4th and 5th IMF
3	Standard deviation of frequency of 2nd IMF
4	Standard Deviation of frequency of 3rd IMF

Table 2 Comparison of Classification Performances in this study and in the study of Corthout et al. (RAS stands for Re-Assigned Spectrogram)

	Our study	EMD and HT	EMD and RAS
Used Features	4	20	20
Sensitivity	0.8881	0.8056	0.8182
Specificity	0.8031	0.9444	0.9439
Accuracy	0.8331	0.8888	0.8945

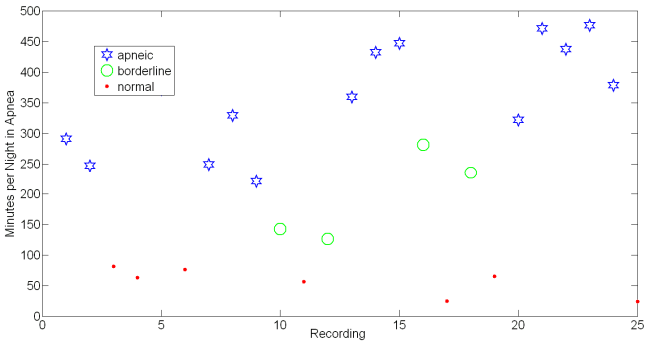


Fig. 1 Subject classification using our method. Applying threshold at around 100 apnea minutes per night shows good separation among patients.

IV. Conclusion

In this study the goal was to show the power of the recently proposed EMD algorithm and its improvements (like EEMD and advanced techniques for extracting frequencies and amplitudes from IMFs) in detecting the Obstructive Sleep Apnea without using any other algorithm or HRV measure. EEMD was applied only to the tachogram (directly derived from ECG), which means that it is an easy way to screen the obstructive apnea in a cheap and non-invasive way. The classification performance was at the best performances published before, although by using a noticeably smaller set of features. Compared to the classification algorithm of Corthout et al. (which we use for estimating this classification algorithm since in both studies the same Physionet data base is used), sensitivity is significantly enhanced, although at the cost of specificity, resulting into a little bit lower accuracy (Table 2). However, a noticeably smaller feature set is used in this study (only 4 features in stead of 20), and only one signal is processed.

Generally, used features are amplitudes and frequencies of the IMFs on the minute-by-minute basis. The effect of apneic events to IMFs was an increase in the amplitude and a decrease of standard deviation in frequency. This means that the amplitude of tachogram is rising, while changes in frequency are small, i.e. the signal is more predictable.

The difference in feature set compared to the work of Corthout et al., was introducing the second feature, using the sum of amplitudes of 4th and 5th mode (Table 1), instead of taking them together, since an apneic event has the same influence on both. This new feature showed a better performance than taking the two original features separately.

This study indicates that apnea detection on tachogram can be done reliably with a sensitivity of around 89%. Besides that, it is also possible to easily separate between normal and apneic subjects. However, separation between apneic and borderline patients was not completely possible, although borderline patients could be wrongly classified as apneic rather than normal, which is more precise.

In future work, we will try to merge this study with the study of Corthout et al., to include more standard HRV measures, since their algorithm significantly outperformed this one in specificity, although sensitivity is much better in this work (Table 2). This way we are trying to enhance the accuracy by finding the “perfect match”.

V. Acknowledgments

Research supported by Research council KUL: GOA-AMBioRICS, CoE EF/05/006, by DWTC: IUAP P6/04 (DYSCO); by EU: Neuromath (COST-BM0601), by ESA: Cardiovascular Control (Prodex-8 C90242)

References

1. Morgenthaler T I, Kagramanov V, Hanak V, and Decker P A (2006) Complex sleep apnea syndrome: Is it a unique clinical syndrome? Sleep 29(9):1203-1209
2. Young T, Palta M, Dempsey J, Skatrud J, Weber S, and Badr S (1993) The occurrence of sleep-disordered breathing among middle-aged adults. N Engl J Med 328:1230-1235.
3. Kodkrnvuo M, Kaprio J, Telakivi T, Partinen M, Heikkila K, Sarna S (1987) Snoring as a risk factor for ischemic heart disease and stroke in men. Br Med J 294:16-19
4. He J, Kryger M H, Zorick F J, Conway W, Roth T (1988) Mortality and apnea index in obstructive sleep apnea. Expirience in 385 male patients. Chest 1:9-14
5. Corthout J, Mendez M O, Bianchi A M, Penzel T, Van Huffel S, Cerutti S (2008) Automatic screening of Obstructive Sleep Apnea from the ECG based on Empirical Mode Decomposition and Wavelet Analysis. EMBC Proc. vol. 30, International Conference of the IEEE Eng. in Med. and Biol. Soc., Vancouver, Canada, 2008, 4 pg
6. Zhaohua W, Huang N E (2005) Ensemble empirical mode decomposition: A noise assisted data analysis method. Centre for Ocean-Land-Atmosphere Studies, Tech Rep 193.
7. Huang N E, (2006) Computing frequency by using generalized zero-crossing applied to intrinsic mode functions. US Patent 6990436
8. Mateo J, Laguna P (2003) Analysis of heart rate variability in the presence of ectopic beats using the heart timing signal. IEEE Trans on Biomed Eng 50(3):334-343

9. Penzel T, Moody G B, Mark R G, Goldberger A L, Peter J H (2000) The apnea-ECG database. Computers in Cardiology 2000 255-258 DOI 10.1109/CIC.2000.898505.
10. Huang N E, Shen Z, Long H R, Wu M C, Shih H H, Zheng Q, Yen N C, Thung C C, Liu H H (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proc R Soc London 454:903-995

Author: Bogdan Mijović
Institute: Katholieke Universiteit Leuven
Street: Kasteelpark Arenberg 10
City: Leuven
Country: Belgium
Email: bogdan.mijovic@esat.kuleuven.be