AUTOMATIC SLEEP APNEA DIAGNOSIS SYSTEM USING NONINVASIVE VITAL SIGNALS RECORDS

P. T. Pozzo Mendoza*, T. Al-ani**, D. Novak*, L. Lhotska*, L. Rigaux***

* Czech Technical University in Prague/Gerstner Laboratory, Technicka 2, Prague, Czech Republic **ESIEE - A²SI, Noisy-Le-Grand, 93162 Cedex, France ***TIS (Health Information Technologies), Polytech'Grenoble, France

pozzop1@feld.cvut.cz, t.alani@esiee.fr, xnovakd1@centrum.cz, lhotska@labe.felk.cvut.cz

Abstract: Sleep Apnea Syndrome (SAS) is a very common sleep disorder. SAS is considered as clinically relevant when the breath stops during more than 10 seconds due to different factors and occurs more than five times per sleep hour. In this paper, we present an automatic approach to sleep apnea classification. This system uses only noninvasive records of the respiratory and cardiac activities (Nasal Airway Flow (NAF) and Pulse Transit Time (PTT)) issued by the technique of PolySomnoGraphy (PSG) are considered for the detection of the different sleep apnea syndromes: obstructive, central and hypopnea. Experimental results using clinical data are presented.

Introduction

Sleep Apnea Syndrome (SAS) is a very common sleep disorder. SAS is considered as clinically relevant when the breath stops during more than 10 seconds and occurs more than five times per sleep hour [1]. These non breathing episodes may sometimes occur more than 300 times a night. Health studies affirm that more than 30 of these non breathing episodes per night should be considered abnormal [1]. There exist two kinds of apneic events that may cause insufficient pulmonary ventilation during sleep, Apnea and Hypopnea. Apnea is defined as the total absence of airflow, followed by the reduction of oxygen levels in arterial blood. The term hypopnea is used when the breath doesn't stop but decrease over 50% of its normal value , followed by the reduction of oxygen levels in arterial blood. The SAS is present mainly in adults and in 11% of children especially in the male population [1, 2]. Different types of apnea-hypopnea may be distinguished: obstructive sleep apnea (OSA), central sleep apnea (CSA), and mixed sleep apnea (MSA).

The treatment of SAS depends on the cause of the apnea. That's why we need to make classification. Nowadays the sleep apneas are classified manually by the expert physician thanks to the nocturnal polysomnographic monitoring that simultaneously records several vital signals during the entire sleeping process (Nasal Air Flow (NAF), Electrocardiogram (ECG), Electroencephalogram (EEG), Electromyogram (EMG), Esophageal Pressure (Peso), Gastric Pressure (Pgas), Oxygen Saturation $(SaO₂)$, ...) [1, 3]. A sleep apnea diagnosis is a very time consuming, expensive and tedious task consisting of expert visual evaluation all 10 minutes pieces of approximately 8 hour recording with a setting of many channels. In a previous work [3], it was demonstrated that sleep apnea classification may be done automatically using three simultaneous records of NAF, Peso and Pgas.

Current techniques of investigating patients with suspected sleep disordered breathing are inadequate. The OSA episodes are not usually difficult to detect even when only a basic measure of respiratory effort such as thoracic and abdominal movement is used. On the other hand, correctly identifying obstructive hypopneas and episodes of upper airway resistance needs a sensitive measure of airflow and inspiratory effort. The measurement of swings in pleural pressure by esophageal manometry is the current gold standard techniques for detecting changes in respiratory effort. However, the placement of an esophageal catheter is often uncomfortable and unacceptable, it may modify the upper airway dynamics, and some believe that it contributes to the sleep disturbance during the sleep study. Furthermore, this technique is available in only a proportion of sleep laboratories and, if performed, adds significantly to the cost of the sleep study. For all these reasons, other new techniques for detecting and classifying sleep apneas and other breathing disorders are developed using mainly the ECG [4] or Pulse Transit Time (PPT) [2, 5]. In this study, we are investigating the use of the neural networks based on simultaneous NAF and PPT records as noninvasive signals, to classify and detect automatically the different types of apnea.

The aim of this work is to find a new technique for detecting and classifying automatically sleep apnea syndrome with the help of the PTT, to validate the use of PTT as a method to automatic diagnosis of sleep apnea syndrome.

The Polysomnography (PSG)

The most important diagnostic tool in any medical condition is for the physician to take the time to obtain a good history and physical examination. A chest x-ray along with laboratory tests are usually performed to evaluate other possible contributing factors, such as diabetes or hypothyroidism. The definitive diagnostic

exam is a polysomnograph where the patient stays in a sleep laboratory or even at home overnight while measurements of his brain activity, respiratory activity, oxygen levels, and cardiac activity are performed (see Figure 1).

Figure 1: A polysomnograph example.

From the clinical point of view, episodes of sleep apnea are often detected first using only the respiration signals, such as nasal airflow and abdominal and thoracic movements. The nasal airflow measures the Apnea/Hypopnea Index (AHI), which indicates the number of apnea/hypopnea during one sleeping hour. The abdominal and thoracic signals show the difference between OSA and CSA. When the type of sleep apnea is not able to be determined from the PSG signals, the esophageal pressure signal has to be measured. This signal shows the diaphragm activity, but to record it requires the application of an invasive method, which increases the risk of arousal due to patient discomfort. The blood oxygen saturation $(SaO₂)$ signal measured by the oxymetry allows distinguishing between apnea and hypopnea. The EEG signals aim is to evaluate arousals. Furthermore, the physiologic range varies during sleep. For example, during the low stable sleep phase the thoracic movements increase, the abdominal movements decrease, and the esophageal pressure increases 50% in comparison to the situation of waking state.

The PSG is uncomfortable for the patient and involves a considerable capital investment for the healthcare system in equipment, bed space and specialized technical support. Interpretation of the test data is also complex and time consuming and consequently the overall cost of performing a PSG is estimated to be around 1000-2000 dollars [6]. Each PSG has its own signal analysis software. Most of them use time domain algorithms, which examine the amplitude and time, with a precision of about 80 to 90%. Even if it is quite efficient and allows a good diagnosis in most of the case it is still difficult to detect the patient's type of sleep apnea with good precision. Even experts do not always agree on the PSG results interpretation. Another difficulty is the dependence between the signals interpretation and the sleep stage and/or the environment. These are the most important reasons why automatic detection is needed.

State-of- the-art and Our Previous Work

To try to solve the PSG interpretation problems, many different approaches have been suggested. These approaches are based on the analysis of different PSG signals with the help of the most common kinds of artificial intelligence (AI) and stochastic methods [3, 712, 18]. Each approach uses different types of signals. In this section, breif introduction to our previous approach [3] is presented.

The classification of sleep apnea syndrome today is not perfect. On one hand the way to diagnose SAS can be improved in order to be more comfortable for the patient and on the other, actual systems do not correctly distinguish the different kinds of SAS. These are the main motivations for researchers to actively work in the createation of a new system to detect and classify the different types of sleep apnea syndrome.

Obtaining non-linear analytical models for the different Sleep Apnea Syndromes is difficult problem. A good approach to be applied could be a method with the help of a expert system, due to their large impact in automated diagnostic systems. The main reasons for not choosing this approach are :

- The difficulty of building and maintaining large rule bases;
- The difficulty to act in real time.

The approach presented in this paper is based on Artificial Neural Network. The main reason to choose this approach is because it has showed good results in physiologic applications and in SAS applying different kinds of other signals than the PTT signal (which is the main one applied in our work) [12]. It should be noted that in the ANN it is important to have enough good data; because these data are needed in the learning process.

Our previous work [3]: In this work, an automatic diagnosis system based on Hidden Markov Models (HMMs) [13] is proposed to help clinicians in the diagnosis of sleep apnea syndrome, Figure 2.

Figure 2: Sleep apnea diagnosis system structure [3].

Some of the measurements of the respiratory activity issued by the technique of polysomnography (NAF, Peso and Pgas) are considered for off-line and on-line detection of the different sleep apnea syndromes: obstructive, central and hypopnea, Figure 3.

Figure 3: An example of CSA event. From top to down: NAF, Peso and Pgas [3].

The inference method of this system translates parameter values into interpretations of physiological and pathophysiological states, Figure 4.

Figure 4: An example of data HMMs based segmentation of the air flow into different states and their interpretations. Event model contains three states: state 1: week airflow, state 2: snoring during expiration and state 3: snoring during inspiration [3].

The interpretation is extended to sequences of states in time to obtain a state-space trajectory. Experimental results using respiratory clinical data showed a good differentiation between the different SAS events. The objective of the current and future work is to improve our system by replacing the invasive signals Peso and Pgas by a non-invasive signals such as the Pulse Transit Time (PTT).

Pulse Transit Time

The measurement of swings in pleural pressure, for detecting changes in respiratory effort is usually assessed by measuring esophageal pressure (Peso), trough an esophageal balloon catheter. This technique has several disadvantages; it causes some discomfort, due to the placement of the esophageal catheter, can lead to fragmentary sleep, and may modify the upper airway dynamics.

A new noninvasive method for measuring respiratory effort has been proposed [2, 5]. It is based on the estimation of the Pulse Transit Time (PTT) signal, which has been demonstrated that its oscillations yield a valid measure of inspiratory effort [2]. The PTT signal is a method to measure the variations in blood pressure. It is the time needed for the arterial pulse pressure wave to travel from the aortic valve to the periphery, generally the ear or the finger. This time is estimated as the delay between the R wave in the ECG and the arrival of the pulse wave at the periphery as determined by pulse oxymetry (about 200-250 ms). We measure one value of PTT by heart beat. There is a link between the PTT and the esophageal pressure (Pes). If the esophageal pressure increases, then the amplitude of the PTT oscillations falls. A decrease of the esophageal pressure corresponds to an increase of the blood pressure (BP), which is directly due to arousals and not to hypoxemia. The PTT signal provides a good measure of respiratory effort, quantification of the obstruction, and therefore a classification of the type of apneahypopnea. There is no respiratory effort in central apnea. If there is an increase of PTT oscillations then there is an OSA or upper airway resistance, because the respiratory effort increases. If there is a decrease of PTT oscillations then a CSA occurs. In this work the PTT was estimated as the interval between the ECG R-wave and the point at which the pulse wave at the finger reached 50% amplitude, Figure 5.

where

 t_1 : the point where ECG-R wave occurs,

 t_2 : the point at which the pulse wave (PW) reaches 50% amplitude (this percentage could change in each study. some of the more recent studies chose a percentage of 25%).

 $PTT = t_2-t_1$

Figure 5: PTT estimation [5].

It should be emphasized that the PTT signal is not able to make a differentiation between apnea and hypopnea [2]. To be able to distinguish between these two apneic events we need to use another signal in addition. It has been decided to use the NAF, which allows a good distinction between apnea and hypopnea [3, 18].

System Overview

To carry out the automatic diagnostic system for sleep apnea classification, it was needed to realize four main steps, shown in Figure 6.

Figure 6: System overview [18].

All these steps are described in [18]. The feature extraction and selection approach is one of the most important in the automatic diagnostic system, because most of the neural network accuracy depends on it. The ANN training is the set of examples used for learning. Enough data is necessary in the training to fit the parameters of the ANN to be efficient. The ANN testing is the set of examples used to assess the performance of the ANN.

Data Acquisition

The data files used for training and testing the neural networks are given by a sleep laboratory at the hospital Raymond Poincaré in France. These data files are all in European Data Format (EDF) and consist of three patients PSG record. Each record measured the different vital parameters: EOG, EEG, ECG, abdominal and thoracic movements, EMG, NAF, phonograph, pulse and $SaO₂$. A more detailed description for each patient is presented in [18]. For reading the data, it was necessary to use other toolboxes: EEGLAB toolbox [13] and the EDF toolbox [14].

Pre-processing of the data: The ECG baseline is the interference that appears due to different reasons, such as: patient movement, breathing, physical exercise, etc. The baseline wandering can make the inspection of the ECG difficult; therefore it is very important to reduce as much as possible its effect. The method used to remove baseline in this work is based on wavelet transform, which removes the low frequency artifacts [15]. The QRS detection by Tompkins and Pan, improved by Fokapu and Girard algorithm [16] was chosen because it showed the best detection accuracy.

Feature extraction: For efficient pattern classification, measurements that could lead to disjoint sets of features vectors are desired. This point underlines the importance of the preprocessing and features extraction procedures. This section aim is to choose features, which will be part of the feature vectors. It is necessary to perform a feature selection in order to have only the most relevant features.

There are two stages that must be fulfilled, the feature extraction stage, which decides how the features will be generated, and the feature selection stage, which decides the best amount of features to be used. In this work two signals are selected for features extraction: PTT and NAF signals.

In order to avoid having too many inputs in the neural network, several windows of the signal were taken. According to the definition of sleep apnea syndrome, the length of this window should be at least 10 seconds in order to detect apnea within a single time window. But in order to assure that some events are not missed, it is better to take a window longer than 10 seconds. Since signals may contain short apnea or hypopnea episodes (lasting only several seconds) which are not pathological. Therefore it was decided to use a window of 14 seconds length. The use of a longer window would unnecessarily increase the complexity of the system.

NAF extraction: The first approach applied for NAF signal extraction was the decreasing of sampling rate from 200 Hz to 20 Hz, in order to have less input in the network. This change of sampling rate was done using a simple moving average filter. These signals are then resampled at 2 Hz using the antialiasing filter. At the end of all the re-sampling the selected 14 second window consisted only of 28 sample points for the NAF instead of 2,800. It has been noticed that the decreasing of sampling rate was not a suitable approach to be applied, because the data was re-sampled too much. Because of this re-sampling some of the information was lost. Therefore a new approach was needed for the features extraction of this signal. The most important information in the NAF signal is the amplitude. This information can be used to determine whether normal breathing, apnea, or hypopnea are present. The details of the method used for calculating the amplitude may be found in [18], Figures 7.

PTT extraction: The first approach used was based on re-sampling the data in the same way done as for the NAF signal. For the PTT signal case the re-sampling was performed first from 200 Hz to 22.22 Hz and at last to 2.469 Hz. It was obtained between 22 and 30 sample

points. It should be noted that the PTT signal is a stair function and not a continuous function, i.e. the numbers of sample points are not the same for each part of the signal. This method is not suitable for the PTT signal, because this reduction does not keep all the information in the signal. Because of this a new approach called PTT segmentation was suggested in [18], Figure 8. Note that a third approach based on the PTT signal feature extraction that correlates with the NAF signal was suggested [19]. This approach calculates the PTT amplitude and duration of each breathing cycle.

Figure 7 : From top to down, Maximas and minimas in a NAF signal for : normal, apnea and hypopnea events [18].

Figure 8: From top to down, PTT segmentation for : normal, apnea and hypopnea events [18].

Training and Testing Sets Creation

The whole training is composed of noisy and non noisy data. In this work the noisy data is consider to be the one that is in REM sleep stage, because during this period the PTT signal is difficult to understand. Two training sets were created for this work, each of them containing 45 windows, 15 windows for each event (i.e. normal breathing, apnea, hypopnea) these data were taken from the PSG records of three patients. The testing set is a set of examples used only to assess the performance (generalization) of a fully-specified classifier. For testing the network performance, 36 windows were taken, 12 windows for each event case (i.e. normal breathing, apnea, and hypopnea) were taken from the PSG records of three patients. It should be noted that the selected test record sections contain no training patterns. Three testing set were created each containing one class of desired outputs. For more details, see [18].

Table 1. ANN Input specification (Yes: used, No: not used). A: Average of NAF amplitude times total maximum peaks, B: Average of each NAF amplitude duration, C: Average of PTT for each amplitude segmentation, D: Average of PTT duration of each segmentation, E: Average of PTT amplitude for each breathing cycle, F: Average of each PTT amplitude duration.

Network parameters: The network outputs were chosen to be binary values, coding the respiration patterns corresponding to the actual input training pattern: [1; 0; 0]: Normal breathing event, [0; 1; 0]: Apneic event (OSA, CSA, MA), [0; 0; 1]: Hypopnea. It should be noted that a hidden layer should be needed to add for further analysis such as classification between OSA, CSA or MA.

The feature extraction and selection was used to build the neural network input. Four simple different neural networks were created using three structures, Figure 9.

These structures are composed of an array of one row and four columns, where: LW $\{1, 1\}$ = First layer weight coefficient, $b \{1\}$ = First layer threshold, LW $\{1,$ 2 } = Second layer weight coefficient, b $\{2\}$ = Second layer threshold.

Figure 9: **A**rchitecture of the first neural network [18].

Table 1 describes neural networks. The input vector of the first neural network is composed of four elements, that of the second neural network is composed of two elements. The third and the fourth neural networks have three elements in their input vector.

Results

The feature extraction and selection helped us to build these neural network inputs. The testing set used was the same for all the neural networks in order to compare the results between them. Table 2 $(1st$ Neural net.) shows that, the first neural network had a problem differentiating between normal breathing and hypopnea. The reason is that the average of NAF amplitude in these cases is almost the same as the one in a normal breathing event or a hypopnea event. One of the reasons could be that one of the elements of the neural network input, listed in Table 1, is the average of NAF amplitude multiply by the total amount of maximum peaks in 14 second window. The reason to choose this element to be part of the neural network input is that when the window contains the beginning or the end of an apneic event, it can be taken the first or last point from a normal breathing. In order to avoid the misclassification between apnea and normal breathing the amplitude average was multiplied by the amount of maximum peaks presented on the window, so this assures us to have a good differentiation between the normal breathing and apnea. However, this multiplication could lead to a misclassification between apnea and hypopnea. Another way of solving these misclassifications is modifying the window start time. From the left part of Figure 13, it could be seen that the amplitude average multiplied by the amount of peaks can still give a result similar for the hypopnea case. It may also appear a misclassification between normal breathing and hypopnea, because sometimes the 14 second window contains more maximum peaks for the hypopnea event than the normal breathing event. The replacement of this element to the average of NAF amplitude would not help to solve those misclassifications. Instead, it showed worst results.

From Table 2 $(2nd$ neural net.), it can be seen that some apneic events were misclassified as hypopneic events. One of the reasons of these misclassifications was mentioned above. Table 2 ($3rd$ neural net.) shows also some hypopneic events misclassification. The third neural network was created to show that the window start time affects in the misclassifications. Table 2 $(4th$ neural net.) shows a small misclassification between an apneic event and a hypopneic event and yields better results than the 3rd neural network. This proofs that the window start time affects the neural networks inputs.

Table 2. Neural networks accuracy results for the three events: Normal, Apnea and Hypopnea.

To make a better comparison among all of the neural networks, their diagnostic accuracy [18, 19] is calculated for each of them. The fourth neural network presents the best diagnostic accuracy of 99.97%, followed by the second and third ones with an accuracy of 94.50%. The first neural network has an accuracy of 86.10%. These proves that the window start time affects the neural network accuracy, the extraction segmentation approach for the PTT signal is better than the amplitude extraction of each breathing cycle approach for the PTT.

Table 3. Diagnostic accuracy calculation for each ANN.

Conclusions

In this preliminary work, our objective was to examine the possibility to use the NAF and PTT records for automatic sleep apnea classification.

We manage to make a differentiation between normal breathing, apnea, and hypopnea.

The first neural network showed worse results than the third neural network. This proves that the extraction by segmentation approach for the PTT signal is better than the amplitude extraction [18] of each breathing cycle for the PTT.

The feature extraction and selection of the PTT signal and the NAF signal was the most difficult part to implement in the artificial neural network design. We believed that further improvements can be achieved in order to get better input data for neural network. This part is one of the most important, because most of the neural network accuracy depends on the step of feature extraction and selection.

The authors are presently investigating the use of Hidden Markov Models (HMMs) and make some comparisons between the approaches presented in this present work and in [3].

Acknowledgements

The authors would like to thanks Daniel Coquelle, manager engineer in MEDATEC France, the 'Service De Physiologie et d'Exploration Fonctionnelle' at Hôpital Raymond Poincaré, France in particular Dr. MA Duera Salva for providing the data used in this work and their help.

The work of P. T. Pozzo Mendoza and L. Lhotska has been supported by the research program "Information Society" under Grant No. 1ET101210512 "Intelligent methods for evaluation of long-term EEG recordings". The work of D. Novak has been supported by the research programme No. MSM 6840770012 "Transdisciplinary Research in the Field of Biomedical Engineering II".

References

- [1] AASM,(1999): 'Sleep-Related Breathing Disorders in Adults: Recommendations for Syndrome Definition and Measurement Techniques in Clinical Research', The Report of AASM Task Force, Sleep, **22**, No. 5, pp. 667-689.
- [2] KALTZ E.S., LUTZ J., BLACK C., and MARCUS C.L., (2003): 'Pulse Transit Time as a measure of arousal and respiratory effort in children with sleep-disorder breathing', in Pediatric research, April 1, **53**, No.4, pp. 580-588.
- [3] AL-ANI T., HAMAM Y., FODIL R., LOFASO F., and ISABEY D., (2004): 'Using Hidden Markov Models For Sleep Disorderd Breathing Identification'. Simulation Practice and Theory, Special Issue, **12**, Issue 2, Elsevier.
- [4] SHINAR Z., BAHARAV A., and AKSELROD S., (2000): 'Obstructive Sleep Apnea Detection Based

on Electrocardiogram Analysis', IEEE Trans. on Computers in Cardiology, **27**, pp. 757-760.

- [5] ROBIN P S., et al. (1999) 'Pulse transit time: an appraisal of potential clinical applications'. THORAX, **54**, No. 5, pp. 452-457.
- [6] RONALD J., DELAIVE K., ROOS L., MANFREDA J., KRYGER M.H., (1998): 'Obstructive Sleep Apnea patients use more health care resources ten years prior to diagnosis',Sleep Res. Online, **1**, pp.71-74.
- [7] VÁRADY P., BONGÁR S., BENYÓ Z., (2003) 'Detection of Airway Obstructions and Sleep Apnea by analyzing the Phase Relation of Respiration Movements Signals', IEEE Trans. Instrument. and Measurement, **52**, No.1, pp. 2-6.
- [8] YEN F.-C., BEHBEHANI K., LUCAS E.A., BURK J.R., AXE J.R., (1997): 'A Noninvasive Technique for Detecting Obstructive and Central Sleep Apnea', IEEE Trans. Biomed. Eng., **44**, pp. 1262-1268.
- [9] AGUIRRE L.A., BARROS V.C., SOUZA A.V.P., (1999): 'Nonlinear multivariable modeling and analysis of sleep apnea time series', Computers in Biology and Medicine, **29**, No. 3, pp. 207-228.
- [10] CABRERO-CANOSA M., CASTRO-PEREIRO M., GRANA-RAMOS M., HERNANDEZ-PEREIRA E., MORET-BONILLO V., MARTIN-AGANA M., VEREA-HERNANDO H., (2003): 'An intelligent system for the detection and interpretation of sleep apneas', Expert Systems with Appl., **24**, No. 4, pp. 335-349.
- [11] VÁRADY P., MICSIK T., BENEDEK S., and BENYÓ Z., (2002): 'A Novel Method for the Detection of Apnea and Hypopnea Events in Respiratory Signals', IEEE Trans.Biomed.Eng.,**49**, pp.936-942.
- [12] RABINER L. R., (1989): A Tutorial on Hidden Markov Models and Selected Application in Speech Recognition. Proc. IEEE, **77**, no. 2, pp. 267-296.
- [13] EegLab, Internet site address: http://www.sccn.ucsd.edu/eeglab/
- [14] EDF, Internet site address: http://www.hsr.nl/edf/companies/companies.html
- [15] NOVAK D., CUESTA FRAU D., (2000) 'Baseline Wandering Removal Using Wavelets", 17th international EURASIP conference, BIOSIGNAL, Brno,Czech Republic.
- [16] PAN J., TOMPKINS W.J., (1985): 'A Real-Time QRS Detection Algorithm', IEEE Trans. Biomed. Eng., **32**, No.3, pp. 230-237.
- [17] PAGANI J., VILLA M.P., CALCAGNINI G., LOMBARDOZZI E., CENSI F., POLI S., BARTOLINI P., BARBARO V., and RONCHETTI R., (2002): 'Detection of Central and Obstructive Sleep Apnea in Children using Pulse Transit Time', Comp. in Cardiol, **29**, pp. 529-532.
- [18] POZZO MENDOZA P.T., (2005): 'Automated diagnosis system for sleep apnea classification', Master Thesis, Czech Technical University in Prague, Fac. El. Eng., Dept. Cybernetics, Prague.
- [19] RANGAYYAN R.M., (2002): 'Biomedical Signal Analysis: A Case-Study Approach' (ED. Wiley-IEEE Press), IEEE Press Series on Biomedical Engineering.