35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Sleep apnea classification using least-squares support vector machines on single lead ECG*

Carolina Varon, *Member, IEEE*, Dries Testelmans, Bertien Buyse,
Johan A.K. Suykens *Senior Member, IEEE*, and Sabine Van Huffel, *Fellow, IEEE*

***Abstract*— In this paper a methodology to identify sleep apnea events is presented. It uses four easily computable features, three generally known ones and a newly proposed feature. Of the three well known parameters, two are computed from the RR interval time series and the other one from the approximate respiratory signal derived from the ECG using principal component analysis (PCA). The fourth feature is proposed in this paper and it is computed from the principal components of the QRS complexes. Together with a least squares support vector machines (LS-SVM) classifier using an RBF kernel, these four features achieve an accuracy on test data larger than 85% for a subject independent classification, and of more than 90% for a patient specific approach. These values are comparable with other results in the literature, but have the advantage that their computation is straightforward and much simpler. This can be important when implemented in a home monitoring system, which typically has limited computational resources.**

I. INTRODUCTION

Sleep apnea is considered an important factor for morbidity and mortality due to its direct effect on the cardiovascular system [1]. These effects are associated with physiological functions such as systemic hypertension and increased sympathetic activity that compromise the heart.

Different types of respiratory events were identified in [2], and they are characterized by alterations in the airflow. These alterations can consist of a reduced airflow as in hypopneas or complete absence of airflow (during at least 10 seconds) as in apneas. The classification of respiratory events as obstructive, mixed or central is based on the respiratory effort.

*Research supported by Research Council KUL: GOA MaNet, PFV/10/002 (OPTEC), IDO 08/013 Autism, several PhD/postdoc & fellow grants;Flemish Government: FWO: PhD/postdoc grants, projects: G.0427.10N (Integrated EEG-fMRI), G.0108.11 (Compressed Sensing) G.0869.12N (Tumor imaging) G.0A5513N (Deep brain stimulation); IWT: TBM070713-Accelero, TBM080658-MRI (EEG-fMRI), TBM110697-NeoGuard, PhD Grants; iMinds 2012; Flanders Care: Demonstratieproject Tele-Rehab III (2012-2014); Belgian Federal Science Policy Office: IUAP P719/ (DYSCO, 'Dynamical systems, control and optimization', 2012-2017); ESA AO-PGPF-01, PRODEX (CardioControl) C4000103224; EU: RECAP 209G within INTERREG IVB NWE programme, EU HIP Trial FP7-HEALTH/ 2007-2013 (n 260777), EU ITN Transact 2012; ERC ADG A-DATADRIVE-B

C. Varon is with the KU Leuven, Department of Electrical Engineering ESAT, SCD-SISTA, and iMinds Future Health Department, Leuven, Belgium e-mail: carolina.varon@esat.kuleuven.be

D. Testelmans and B. Buyse are with the UZ Leuven, Department of Pneumology , Leuven, Belgium

S. Van Huffel and J.A.K. Suykens are with the KU Leuven, Department of Electrical Engineering ESAT, SCD-SISTA, and iMinds Future Health Department, Belgium.

Currently, sleep apnea is diagnosed using poly(somno)-graphy, which is a sleep test evaluated by a clinician, and performed either in a hospital environment or at home. This test monitors different physiological measurements such as heart rate and respiration, which are significantly affected during episodes of apnea. Although poly(somno)graphy is nowadays the most important tool to diagnose sleep disorders, its elevated costs and reduced comfort cause sleep apnea to be under-diagnosed. To overcome this, several studies like for instance [3], [4], [5] and [6], have considered the possibility of detecting episodes of sleep apnea by using only the ECG signal. In this way, home monitoring systems, which are cheaper and more comfortable, can be developed to detect and diagnose this disorder.

In general home monitoring systems have limited calculation power, and this eliminates the possibility to implement computationally expensive algorithms. That is why methods are needed that call for less processing, but which nonetheless achieve the same grade of accuracy as more demanding algorithms. In this paper such a method is proposed which uses a set of four very simple features. Amongst the features there are three generally known ones: namely the standard deviation of the RR intervals, the serial correlation coefficient of the RR intervals and the standard deviation of the ECG derived respiration using Principal Component Analysis (PCA) [10]. In addition, a new feature is proposed here, based on the percentage of variance explained by the second principal component of a matrix formed by the QRS complexes. This feature will be explained in more detail in the next section. These features were used with a least squares support vector machine (LS-SVM) classifier [7] using an RBF kernel, to separate respiratory events from normal activity on two different datasets.

The remainder of this paper is organized as follows. Section II describes the two datasets used in this study and the methodology implemented to analyze them. The results are presented and discussed in section III and the conclusions in the final section.

II. METHODOLOGY

A. Data

Two datasets were used in this study. The first one was the Apnea-ECG database [3] publicly available on Physionet[1] [8]. This dataset consists of 70 single-lead ECG recordings sampled at 100Hz, and 8 respiratory signals available for 8 of

[1] www.physionet.org

978-1-4577-0216-7/13/$26.00 ©2013 IEEE

the 70 recordings. Each ECG signal was manually annotated by an expert, who indicated on a minute-by-minute basis whether an apnea episode occurred. These annotations do not differentiate between types of respiratory events. In total 34 313 minutes were extracted and annotated.

The second dataset consists of 10 single-lead ECG signals extracted from polysomnographic recordings of 10 different patients of the sleep laboratory at the University Hospital Leuven (UZ Leuven), Belgium. Additionally, respiratory signals are available for 5 patients. All the signals were sampled at 200Hz and their duration ranges from 320 to 469 minutes (384$\pm$47). In total 3 847 minutes were extracted and annotated by a medical doctor experienced on interpreting polysomnographic signals. These annotations correspond to the beginning and duration of the "respiratory events", and they make a distinction between 5 different types: obstructive apnea (OSA), obstructive hypopnea (OSH), hypopnea (HPA), mixed (Mix), and central apnea (Cen).

B. ECG processing and RR interval time series

Initially, each signal was segmented into epochs of one minute, which were then analyzed for artefacts using the algorithm presented in [9]. If an ECG minute contained an artefact, it was removed from the dataset. Then, each remaining ECG minute was processed using the Pan-Tompkins algorithm, which identifies the position of the R-peaks. Additionally, a search back procedure [5] was applied to detect and correct ectopic and missing beats. Once the R-peaks were identified, the RR interval time series were obtained by taking the time intervals between successive beats.

C. ECG derived respiration (EDR)

After identifying the R-peaks, three different methodologies, namely the amplitude of the R-peak, principal component analysis (PCA) [10] and kernel principal component analysis (kPCA) [11], were implemented to compute the ECG derived respiratory (EDR) signals. These three methods are based on the volume changes produced in the lungs during a respiratory cycle, which on its turn modifies the position of the electrodes with respect to the heart. In other words, they exploit the mechanical interaction between respiration and the morphology of the heart beats. Previous studies validated these methodologies [10], [11], [12], and they showed that the computed EDRs closely resemble the real respiratory signals.

D. Feature extraction and Feature selection

To characterize each ECG minute, a set of time and frequency domain parameters were derived from the RR interval time series and three EDR signals. The time domain parameters derived from the RR time series were the mean, standard deviation, root mean square of inter-beat differentials, standard deviation of inter-beat differentials, 5 serial correlation coefficients and 4 fractal Alan factors. From the EDR signals, only the mean and standard deviation were calculated. Concerning the frequency domain, the same set of features was extracted from the EDR and from the RR time series. This set of features consists of variances of 5 levels of wavelet decomposition, 32 points of the power spectrum, power in the low frequency band (0.04-0.15Hz) and power in the high frequency band (0.15-0.4Hz). The computation of this set of features is described in [4].

This feature set was then expanded with an extra parameter to reduce the complexity and the number of features needed to classify respiratory events. This extra parameter is based on the computation of the EDR signal using PCA. Usually, only the end result of this computation is taken into account, however, in this paper it is shown that a variation of the procedure provides information that can be related to apnea events. This becomes clear when this procedure is analyzed in detail.

In [10] a collection of beat features extracted from the single-lead ECG is aligned in a single matrix to which PCA is applied. Once the principal components (PCs) of the new matrix are computed, the first PC is closely related to the respiratory signal. In this work, the selected beat feature is the QRS complex, and the algorithm defines a window of 60ms around each R-peak, removes the mean and organizes these windows in a QRS matrix $X \in \mathbb{R}^{N \times M}$, where N is the number of beats, $M = 0.06 \times fs$, and fs the sampling frequency (see Fig. 1). The matrix X is transformed into a set of principal components by means of the eigendecomposition of its covariance matrix, and the first component is then identified as the ECG derived respiration. The diagram of Fig. 1 summarizes the procedure described above.

The variation proposed in this paper, applies the algorithm to X^T rather than X. The reasoning behind this adaptation can be explained very intuitively. Applying the procedure to X and keeping the first component gives an estimation of the respiratory signal. When considering X^T rather than X, however, a "local" interpretation around each R-peak is obtained. This local interpretation corresponds to the morphology of the QRS complexes, which on its turn is modulated by respiration. Because of this relation with respiration, the local interpretation can be used to distinguish between normal and abnormal minutes. This is clarified in Fig. 2, where a comparison between the eigenvalues of two minutes, one normal and one containing an apneic event, is displayed together with the percentage of variance explained by each component. These percentages of variance are defined as $(\lambda_i / \sum \lambda_i) \times 100$, where λ_i is the ith eigenvalue of the covariance matrix, $i = 1, \ldots, m$, and m the window size. In the figure it is observed that when apnea occurs, the percentage of variance explained by the first eigenvector is reduced, while the variance explained by the second increases. In other words, more components are needed to describe the variations in the morphologies of the QRS complexes. Taking into account the analysis above, this study proposes to use both the percentage of variance explained by the first and second principal components as features.

After deriving the set of features described above, their discrimination level was determined by means of the F1

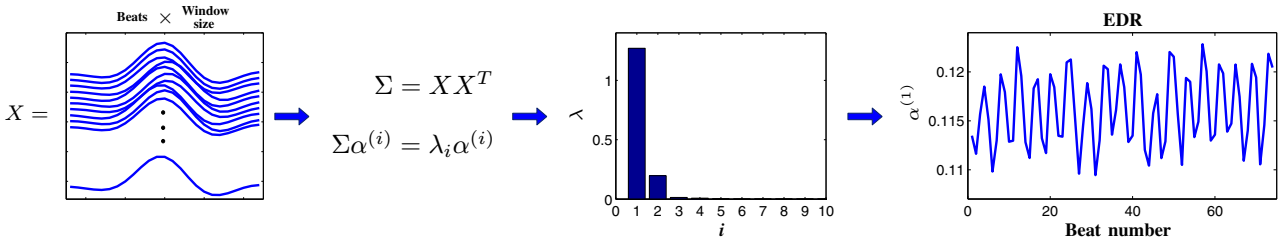

Fig. 1. Computation of the EDR signal using PCA [10]. From left to right: matrix containing the QRS complexes, general procedure to find the principal components, eigenvalues of the covariance matrix, and EDR signal corresponding to the eigenvector with the highest eigenvalue. Note that the first eigenvalue is large compared to the others, which indicates that most of the variance of the data is explained by its corresponding eigenvector.

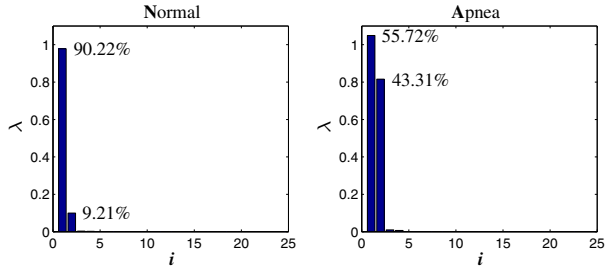

Fig. 2. Eigenvalues (λ_i) of the covariance matrix of X^T for two different ECG segments, one normal and one apneic. Note that the percentage of variance explained by the second principal component becomes larger during an apneic minute.

score, which ranges between 0 and 1. This score is equal to 1 when the corresponding feature allows for perfect classification of the data, and it is equal to 0 when no differentiation between classes can be made.

E. Classification

The dataset was divided into training and test sets. For the training set 2000 minutes were selected by means of the fixed size method proposed in [7], which guarantees that the underlying distribution of the data is approximated. The number of samples in the training set is limited by the memory and computational requirements of the classifier. In this study the least squares version of the standard support vector machines (SVM) [13] was implemented for the first time to classify respiratory events. Least-squares SVM (LS-SVM) simplifies the problem by defining the cost function as a least squares problem, and equality constraints instead of inequality constraints. This implies that a set of linear equations is solved instead of a quadratic programming problem as in traditional SVM. Different studies have shown that for some applications the performance of LS-SVM is comparable to or better than the one of SVM [7], [14], [15]. In [4] SVMs were used to classify sleep apnea, and the work presented in this paper aims to obtain comparable results by means of a simplified algorithm.

In this work, the performance of an LS-SVM classifier using an RBF kernel and a linear kernel, was evaluated on the test set using accuracy, sensitivity and specificity defined as: Sens $= TN/(TN + FP) \times 100$, Spec $= TP/(TP + FN) \times 100$, and Acc $= (TP + TN)/(TP + FP + TN + FN) \times 100$, where TN, TP, FP and FN, correspond to the true negative, true positive, false positive and false negative events respectively.

III. RESULTS AND DISCUSSION

All computations were carried out in MATLAB on an Intel Dual Core, 4GB. To perform classification, the LS-SVMLab[2] toolbox was used.

After the removal of artefacts from each dataset, 30 648 minutes remained in Physionet and 3 495 minutes in the Leuven dataset. Next, 177 features were extracted from each minute of ECG, and evaluated using the F1 score. Only the features with scores larger than 0.3 were used in the classification step, as this threshold provided the most accurate results. This highly discriminative set of features consists of the standard deviation of the RR time series, the serial correlation coefficient of the RR at 3 time lags, the standard deviation of the EDR computed using PCA, and the percentage of variance explained by the second principal component as proposed in this paper. The performance of each of the features individually is shown in Table I. The values in the table already give an indication that the extra parameter has a high discriminative power. With the addition of this parameter, only four very straightforward features are needed to achieve an accuracy comparable to previous results in the literature for subject independent classification such as in [5] with Acc $> 85\%$ with manual verification of the QRS, and [4] with Acc $\approx 89\%$, amongst others. Note that the algorithm presented in this paper is fully automated. No manual verification of the Rpeak detection or of the set of features was performed. From the table it is clear that the LS-SVM classifier with an RBF kernel outperformed the linear case.

The classification of the Physionet dataset for a subject independent approach, gives an accuracy of 85.07% and an area under the receiver operating characteristic curve (AUC) of 0.9186 on test set, while for patient specific classification an accuracy of 91.03% is reached. These results were obtained with an LS-SVM classifier, using an RBF kernel with kernel parameter $\sigma = 0.77$, a regularization parameter $\gamma = 10.45$, and 10-fold cross-validation. As mentioned before, this is in line with values in the literature, but without computational expensive frequency domain features, such as the ones derived using Fourier analysis and wavelet decomposition.

[2]www.esat.kuleuven.be/sista/lssvmlab/

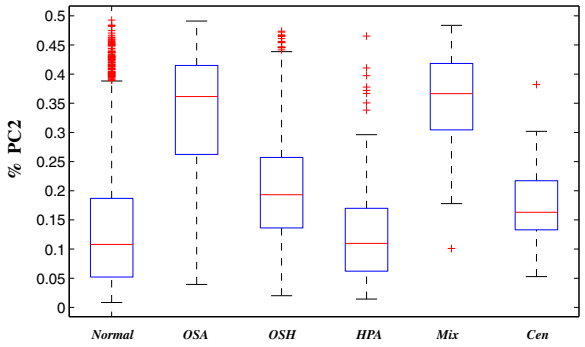

Fig. 3. Boxplot of the percentage of the variance of the data that is explained by the second component (%PC2). Note that there are clear differences when comparing normal and obstructive apneas.

The discriminative power of the feature set becomes more clear when it is applied to a dataset that labels different types of respiratory events such as the Leuven dataset. First, classification with each feature individually was performed, the results of which are shown in Table I. From this table it is clear that the percentage of variance again performs best. Boxplots of this feature are shown in Fig. 3, where it differentiates between obstructive apneas and normal segments. Conclusions on the central and mixed apneas will not be drawn due to the low amount of examples contained in the dataset.

The classification of the dataset using all four features gives an accuracy of 85.13% and AUC of 0.9029, and distinguishes two different classes. The class labeled as "apnea" contained 87.90% of 372 OSA, 65.52% of 496 OSH, only 26.19% of 84 HPA, 95.65% of 23 Mix, 75% of 16 Cen, and 11% of 2504 normal minutes. This group of normal minutes classified as apnea, contains the outliers indicated in the "*Normal*" box of Fig. 3. This should not come as a surprise, since hypopneas consist of a reduction in the airflow, rather than an interruption of breathing. Additionally, these (non-obstructive) hypopneas can manifest themselves while the patient is awake [2]. As a result they are very similar to normal minutes and hence more difficult to separate, even for human experts.

To remedy for this, classification using only normal heart rate and EDR based features should be expanded with different signals such as SpO_2 and CO_2 [2] and [6].

TABLE I

TEST PERFORMANCES OF SLEEP APNEA CLASSIFICATION

	Physionet			Leuven		
Feature	**Sens**	**Spec**	**Acc**	**Sens**	**Spec**	**Acc**
std(RR)	55.77	76.68	70.60	51.66	84.15	74.19
SCorr. Coef	67.88	78.33	74.00	58.06	65.23	61.98
std(EDR)	55.58	70.43	65.63	72.35	95.02	75.14
% PC2	75.54	74.23	74.11	76.79	95.39	77.01
All 4 (linear)	83.92	79.97	81.09	51.97	88.78	78.27
All 4 (RBF)	**88.84**	**83.29**	**85.07**	**70.23**	**91.05**	**85.13**

IV. CONCLUSION

In this paper, a classification using three well-known time domain features and a newly proposed parameter was performed. The new parameter was calculated using the eigenvalues of the covariance matrix computed from the QRS matrix. One big advantage of using this feature set together with an LS-SVM classifier is that even though the features are more "basic", the performance is comparable to the best results in the literature for fully automated methods. Another advantage is that there is no need to perform demanding algorithms, such as Fourier analysis and wavelet decomposition. This becomes particularly important when the algorithm needs to be implemented in home monitoring systems, which typically have limited hardware specifications.

REFERENCES

[1] S. M. Caples, A. Garcia-Touchard, and V. K. Somers, "Sleep-disordered breathing and cardiovascular risk". *Sleep*, vol. 30, pp. 291-303, 2007.

[2] "Sleep-Related Breathing Disorders In Adults: Recommendations for syndrome definition and measurement techniques in clinical research". *The report of an American academy of sleep medicine task force*, vol. 22, pp. 667-689, 1999.

[3] T. Penzel, G. B. Moody, R. G. Mark, A. L. Goldberger, and J. H. Peter, "The apnea-ECG database". *Computers in Cardiology*, pp. 255-258, 2000.

[4] M. Bsoul, H. Minn, and L. Tamil, "Apnea MedAssist: Real-time sleep apnea monitor using single-lead ECG". *IEEE Trans. Inf. Technol. Biomed.*, vol. 15, pp. 416-427, 2011.

[5] P. de Chazal, C. Heneghan, E. Sheridan, R. Reilly, P. Nolan, and M. OMalley, "Automated processing of the single-lead electrocardiogram for the detection of obstructive sleep apnoea". *IEEE Trans. Biomed. Eng.*, vol. 50, pp. 686-696, 2003.

[6] T. Al-Ani, Y. Hamam, D. Novak, P. T. Pozzo Mendoza, L. Lhotska, F. Lofaso, D. Isabey, and R. Ffodil, "Noninvasive automatic sleep apnea classification system". *BioMedSim*, pp. 26-27, 2005.

[7] J. A. K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle, *Least squares support vector machines*. World Scientific Publishing Co. Pte. Ltd., 2002.

[8] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals". *Circulation*, vol. 101, pp. 215-220, 2000.

[9] C. Varon, D. Testelmans, B. Buyse, J. A. K. Suykens, and S. Van Huffel, "Robust artefact detection in long-term ECG recordings based on autocorrelation function similarity and percentile analysis". *Proc. of the 34th annual international conference of the IEEE EMBS*, pp. 3151-3154, 2012.

[10] P. Langley, E. J. Bowers, and A. Murray, "Principal component analysis as a tool for analyzing beat-to-beat changes in ECG features: Application to ECG-derived respiration". *IEEE Trans. Biomed. Eng.*. 2010;57: 821-829.

[11] D. Widjaja, C. Varon, A. Caicedo Dorado, J. A. K. Suykens, and S. Van Huffel, "Application of kernel principal component analysis for single lead ECG-derived respiration". *IEEE Trans. Biomed. Eng.*. 2012;59: 1169-1176.

[12] C. Varon, K. Jansen, L. Lagae, and S. Van Huffel. "Respiratory pattern in infants with West syndrome". *Proc. of the 5th international conference of advances in medical signal and information processing (MEDSIP)* 2012.

[13] V. Vapnik, *Statistical learning theory*. John Wiley & Sons, Inc., 1998.

[14] V. Mitra, C-J. Wang, and S. Banerjee, "Text classification: A least square support vector machine approach". *Applied Soft Computing*, vol. 7, pp. 908-914, 2007.

[15] R. K. Tripathy, A. Acharya, and S. K. Choudhary, "Gender classification from ECG signal analysis using least square support vector machines". *American Journal of Signal Processing*, vol. 2, pp. 145-149, 2012.