34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Training using Short–time Features for OSA Discrimination

L.M. Sepulveda-Cano* A.M. Alvarez-Meza* and G. Castellanos-Dominguez*

Abstract— **Heart rate variability (HRV) is one of the promising directions for a simple and noninvasive way for obstructive sleep apnea syndrome detection (OSA). The interaction between the sympathetic and parasympathetic systems on the HRV recordings, gives rise to several non-stationary components added to the signal. Aiming to improve the classifier accuracy for obstructive sleep apnoea detection, the use of more appropriated techniques for leading with non–stationarity and mixed dynamics, are needed. This work aims at searching a convenient training strategy of combining the feature set to be further fed in to the classifier, which should take into consideration the different dynamics in the HRV signal. Therefore, a set of the short-time features, extracted from a given HRV time-varying decomposition, and selected by spectral splitting is considered. Additionally, three methods of projection are used: none, simple, and multivariate. Finally, the different approaches are tested and compared, using *k*-nn and support vector machines (SVM) classifiers. Attained results show that using continuous wavelet transform with short–time features and multivariate projection, followed by a SVM classifier, allow to obtain a suitable OSA detection.**

Keywords: Obstructive Sleep Apnea, Spectrogram, Scalogram, Cepstral Coefficients, Support Vector Machines

I. INTRODUCTION

The obstructive sleep apnea syndrome (OSA) is a common sleep disorder, characterized by an obstruction in the airflow. To perform an OSA diagnosis, detection of repetitive episodes of apnea and hypopnea during sleep is carried out, mostly, by attended overnight polysomnography in a sleep laboratory. The main disadvantage of a standard polysomnography test is the large amount of information required to be analyzed [1], [2]. One of the promising directions for a simple, less costly, noninvasive screening method for OSA detection is provided by an analysis based on the heart rate variability (HRV) [3]. The interaction between the sympathetic and parasympathetic systems on the HRV recordings, gives rise to several non-stationary components added to the signal [4].

Aiming at finding the informative features in the HRV signal for OSA diagnosis, some authors have used different approaches. The time–frequency (*t*–*f*) representations has been also proposed, planned to determine the energy distribution along the frequency axis at each time instant, to investigate the time–variant properties of the spectral parameters during either transient physiological or pathological episodes [5]. In this line of analysis, the use of relevant and non–redundant dynamic features is presented in [6] based on the Short Time Fourier Transform (STFT) and a linear variant of the Continuous Wavelet Transform (CWT); nevertheless, there are some normal *t*–*f* maps whose waveform are similar to the pathological ones, and vice versa; so, the spatial distribution of the energy in each sub–band is not clear. In [7], a spectral splitting is proposed for stochastic features extraction based on the STFT; different relevance measures are proved for boundaries selection. However, the dynamic of the HRV signal is so complex that it is necessary to develop a robust methodology, from the decomposition of the signal to the classifier, handling properly this dynamics, and thus, improving the performed accuracy by a given classifier. This study proposes a convenient training strategy of combining the feature set to be further fed in to the classifier, which should take into consideration the different dynamics in the HRV signal.

The rest of this paper is organized as follows: first, the signal decomposition is introduced, then, the methodology for stochastic features is explained. Lastly, the effectiveness of each decomposition is illustrated for the OSA detection using a *k*-nearest neighbors (*k*-nn), and support vector machines (SVM) classifiers, followed by a discussion of the results.

II. MATERIALS AND METHODS

A. Enhanced HRV Representation

a) Time–Scale and Time–Frequency Decompositions: The signal decomposition through continuous wavelet can be expressed in two dimensions, as follows [8]:

$$x(t) = \int\int \boldsymbol{S}(a,b)\phi(t,a,b)\frac{dadb}{a^2}, \tag{1}$$

where $\boldsymbol{S}(a,b)$ are the coefficients of the continuous wavelet, a is scale factor, b is the position factor, and $\phi(t,a,b)$ is the mother wavelet. The CWT becomes a convenient time–scale representation when Morlet is chosen as the mother wavelet due to the similitude with the HRV biosignal.

On the other hand, the commonly used time–frequency representation calculated by Short Time Fourier Transform introduces a time localization concept by using a tapering window function, $w(t-\tau)$, (being τ the shifting factor) of short duration going along the studied signal, $x(t)$, with the window length remaining constant, namely:

$$x(t)w(t-\tau) = \frac{1}{2\pi}\int \boldsymbol{S}(\tau,\omega)\phi(t,\omega)d\omega, \tag{2}$$

where $\phi(t,\omega) = e^{j\omega t}$ is the basis and $\boldsymbol{S}(\tau,\omega)$ are the decomposition coefficients.

The coefficient matrix $\boldsymbol{S}$ of enhanced representations describing HRV signal dynamics are the *scalogram*, in case of

* Signal Processing and Recognition Group, Universidad Nacional de Colombia, sede Manizales. lmsepulvedac,amalvarezme,cgcastellanosd@unal.edu.co

CWT, and *spectrogram* for STFT; both of them, respectively, defined as:

$$\boldsymbol{S}(t,\omega)=\left|\int_T x(\tau)w(\tau-t)e^{-j\omega t}d\tau\right|^2 \tag{3}$$

$$\boldsymbol{S}(a,b)=\int_T \frac{1}{\sqrt{a}}x(t)\phi^*(t,a,b)dt, \tag{4}$$

where $\phi^*(t,a,b)$ is the conjugated mother wavelet.

B. Spectral Splitting upon Enhanced Representations

To extract the set of band-pass feature vectors, the sub–band boundaries are to be determined, which in the concrete case are estimated by introducing a filter–type measure of relevance to evaluate the whole enhanced representation (either spectrogram or scalogram), as proposed in [7]. Since each extracted feature vector may hold a different amount of useful information for OSA detection, the proposed relevance–based splitting scheme emphasizes the most relevant sub–bands. That is, the more relevant the set of spectral components $\{\boldsymbol{s}_f(t)\}$ of a given sub–band the more important the derived stochastic feature. The following unsupervised measure of time–variant relevance is assessed [6]:

$$\boldsymbol{\rho}(\boldsymbol{S};\tau)=[\chi(1)\ \cdots\ \chi(\tau)\ \cdots\ \chi(pT)]^\top, \tag{5}$$

where $\chi(\tau)=\mathscr{E}\{|\lambda_j^2 v_j(\tau)|\}$, $\{\lambda_j : j=1,\dots,q\}$ is the set of most relevant eigenvalues of matrix $\boldsymbol{S}$, and scalar $v_j(\tau)$ is the respective element at τ moment, and $\tau=1,\dots,pT$ indexes each one of the relevance values computed for the whole time–variant data set. Notation $\mathscr{E}\{\cdot\}$ stands for expectation operator.

To determine distinctly the relevance related to each one of the stochastic variables, Eq. (5) can be reallocated to the relevance matrix, $[\boldsymbol{\rho}_1(\boldsymbol{S};t)\cdots\boldsymbol{\rho}_f(\boldsymbol{S};t)\cdots\boldsymbol{\rho}_{\Delta F}(\boldsymbol{S};t)]^\top$, where each row $\boldsymbol{\rho}_f(\boldsymbol{S};t)=[\chi((f-1)T+1)\dots\chi(t)\dots\chi(fT)]\in\mathbb{R}^{T\times 1}$ that is a sectioned version of vector $\boldsymbol{\rho}(\boldsymbol{S};\tau)$ plainly holds the contribution of the $\boldsymbol{s}_f(t)$ stochastic feature along the fixed moments of time. To measure the contribution of each spectral component, a simple average is accomplished, i.e., $\rho_f(\boldsymbol{s}_f)=\mathscr{E}\{\boldsymbol{\rho}_f(\boldsymbol{S};\tau):\forall\tau\}$. Because of high level of correlation existing between each pair of adjacent spectral components $\{\boldsymbol{s}_f,\boldsymbol{s}_{f+1}\}$, the main assumption is that the minimum values of their measured contribution should be considered as the boundaries of the spectral sub–bands. Each assessed frequency or scale sub–band from end to end along the time domain holds the boundary of a single stochastic feature. In turn, each vector feature is attained by filter bank modeling. For the sake of simplicity, in this study the set of *Cepstral Coefficients* (CC) is estimated, as proposed in [7].

C. Training Feature Representation

Since each $i-$extracted stochastic feature is a vector $\boldsymbol{x}_i\in\mathbb{R}^{1\times T}$, a proper feature set representation within training framework should be carried out. It is expected that the dynamic behavior of each narrow–band stochastic process should remain clearly slow. In fact, time–variant feature vectors of lower orders, which correspond to lower spectral bandwidths and that are assumed to carried out most of the information, are expected to behave slowly enough along the time axis so that usually imposed stationary restrictions on short–term estimations should work out better than if computing over the raw input signal. On this regard, several methods of representation can be used, among others the following are considered:

– A simple projection into a single real-valued escalar, i.e.,
$$y_i=\mathscr{E}\{\boldsymbol{x}_i:\forall t\in T\},\ y_i\in\mathbb{R} \tag{6}$$

– A multivariate projection into a reduced latent vector. Because of high computational cost of stochastic feature-based training, dimension reduction of the input space is carried out by means of a time–evolving version of the standard Principal Component Analysis (PCA), as follows:
$$\boldsymbol{y}_i=\mathrm{PCA}\{\boldsymbol{x}_i\},\ \boldsymbol{y}_i\in\mathbb{R}^{1\times\tau},\tau\le T \tag{7}$$

– The raw extracted vector is to be fed into the classifier,
$$\boldsymbol{y}_i=\boldsymbol{x}_i,\boldsymbol{y}_i\in\mathbb{R}^{1\times T}, \tag{8}$$

This work uses a couple of clustering-based classifiers: the k-nn, for which classification is grounded on how close observation falls to the cluster centroid, and SVM, where clusters are hyperplanes segregating one region from another.

III. EXPERIMENTAL SETUP

The proposed training methodology using short–time features for OSA discrimination is divided into three steps: *i*) Computation of enhanced representation, *ii*) Dynamic features estimation by spectral splitting, *iii*) Dimension reduction and classification.

A. Database

This collection holds $M=70$ electrocardiographic recordings from PhysioNet, each one including a set of reference annotations added every minute of the recording indicating either the presence or absence of apnoea during each segment of time. The recordings were subdivided in three groups: apneic patients, with more than 100 min in apnea, borderline patients, with total apnea duration more than 5 and less than 99 min and control or normal patients, with less than 5 min in apnea. From the database, 25 recordings were used as a training set for the classification algorithms. A second group with 25 recordings was used as a test set to measure the performance of the algorithms, as considered in [9].

B. Computation of Enhanced Representations

Parameter tuning for a considered TFR is achieved by the procedure developed for biosignals, discussed in [9]. the STFT–based quadratic spectrogram is computed by sliding Hamming windows for the following set of estimation parameters: 32.5 *ms* processing window length, 50% of overlapping, and 512 frequency bins.

Regarding to CWT alternative, the respective time–scale representation is performed by using Morlet wavelet mother function that is likely to be an efficient means for HRV studies [10]. Based on the dominant frequency components of the underlying HRV signal, the number of WT decomposition levels is chosen within the interval $[2-512]$, where the value 2 is regarded to the highest assumed frequency value while value of 512 reflects the lowest frequency.

C. Dynamic Feature Extraction by Spectral Splitting

Figure 1 shows the relevance map of the *t–f* distribution. Table I shows the selected frequency bands. So, computed relevance maps of HRV representations reveal a large difference in terms of dynamic behavior between sub–bands. Generally, the stochastic variability contribution of every one of the stationary spectral components, should remain constant throughout the time axis. In this regard, the sub-bands do not entirely hold the stationary assumption. Former relevance matrix patterns stochastic variability of spectral components with some lack of uniformity. From the above observations, one may confirm the difference between both HRV bands of interest in terms of stochastic properties. Because the amount of information obtained by the relevance maps, the underlying bands with similar stochastic behavior are clustered through spectral slitting. It must be noted that because of easier medical interpretation, the splitting over *t–f* maps is carried out separately for each one of the two bands of interest (LF and HF), as recommended in [1], [7].

TABLE I
FREQUENCY BANDS SELECTION

Band [*Hz*]	**Selected Frequency Sub–Bands** [*Hz*]
LF	$0.04-0.06, 0.06-0.08, 0.8-0.15$
HF	$0.15-0.21, 0.21-0.250.25-0.32, 0.32-0.50$

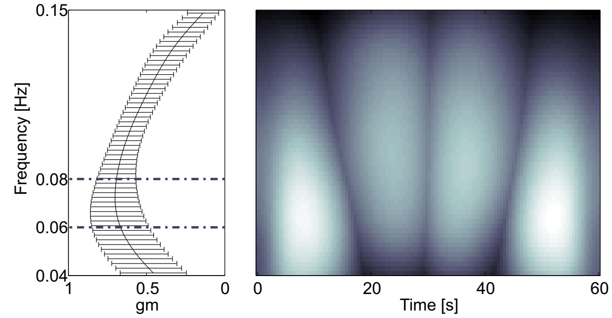
The splitting of the scale axis is achieved in a similar way, using the relevance obtained for each scale set. Figure 2 shows the relevance map of the *t–s* distribution. In this case, only three bans are chosen at following scales: $2-413$; $413-485$; $485-512$.

According to the sub–band selection by spectral splitting, both representations of the HRV are decomposed, using CC. Within this procedure at each time, a triangular response filter is applied, extracting 7 stochastic features from spectrogram (see Fig. 3(a)), while just 3 stochastic features are obtained from scalogram (see Fig. 3(b)).Validation of accomplished results is carried out by well–known cross–validation methodology.

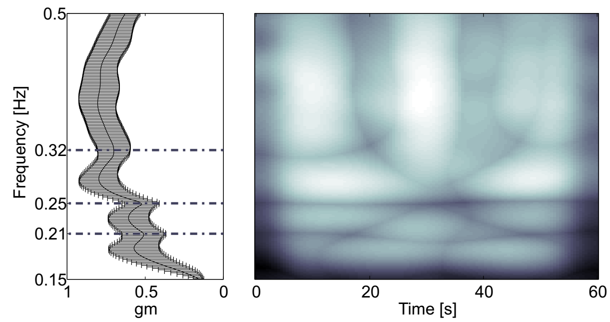
D. Results and Discussion

Tuning of the different stages throughout considered training methodology (signal decomposition, characterization, and dimensionality reduction) is carried out by using the average classification accuracy for the automatic OSA detection. The different approaches are tested and compared, using *k*-nn and SVM classifiers, in accordance with above discussed schemes of feature representation.

Table II summarizes the performed minute–by–minute classification accuracy for each training strategy. As seen



(a) Low frequencies



(b) High frequencies

Fig. 1. Relevance of the time–frequency plane and frequency axis obtained and Spectral splitting.

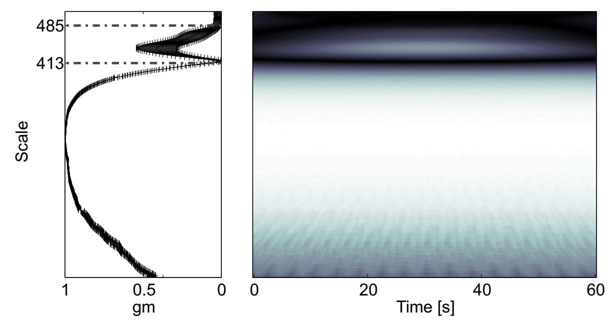
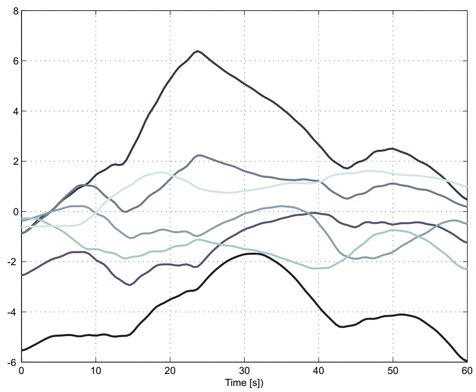
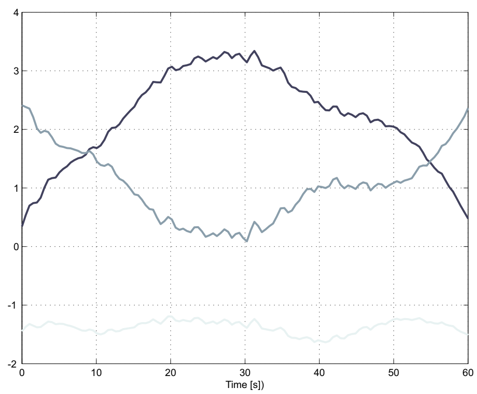


Fig. 2. Relevance of the time–scale plane and scale axis obtained and Spectral splitting.

in Fig. 3, the stochastic features dynamic associated to both decomposition techniques needs a different training approach. In the case of STFT, the dynamic of the short–term features remains stronger and the enhancements can not handling properly its behavior; so, the best classifier accuracy is obtained with a simple scheme, i.e. a projection into a single real valued scalar and *k*–nn as classifier. For CWT, the short-term features are softer, and the improvements are proportional to the complexity of the proposed scheme. In this approach, it is clear the contribution of temporal evolution, due to the fact that the simple projection do not provides enough information about the studied phenomena. Then, multivariate projection is added, improving the performance accuracy for *k*–nn as for SVM classifiers. Finally, the best training strategy, is constituted by CWT and its respective short–time features with multivariate projection, followed



(a) spectrogram



(b) scalogram

Fig. 3. Extracted short–time features (short-time ceptral coefficients) from respective enhanced representation

by a SVM classifier, with an accuracy of 78%, specificity rounded 72% and sensibility of 75.12%.

IV. CONCLUSIONS AND FUTURE WORK

The training methodology is explored, which is based short-time features, extracted from HRV signal by using STFT and CWT decompositions for OSA detection. In addition, the use of dimensionality reduction is analyzed. The methodology lies on the hypothesis that using techniques more appropriate for deal with non–stationarity and mixed dynamics, a better accuracy could be obtained. With relevance–based splitting scheme over enhanced representation of HRV signals, a set of dynamic filter–banked features can be extracted providing an appropriate OSA segment classification accuracy; nevertheless, the classifier choose has a significant role in the accuracy. The best performance is carried out by CWT with PCA and support vector machines as classifier, followed by the same scheme without PCA as dimensionality reduction technique. With aim to improve the segment classification performance, some aspects should be thoroughly studied. Particularly, it would be of benefit to explore the needed enhancement by using more elaborated approaches like Gaussian Mixture Models, Hidden Markov Models and Gaussian Process, in order to obtain a more accurate tracking of the strong dynamics on the HRV signals. Additionally, the scheme should be proved in another databases as recommended in [11].

TABLE II
CLASSIFICATION ACCURACY

k–nn Accuracy [%]		
STFT	Simple projection	75.97
	Multivariate projection	73.17
	None	74.85
CWT	Simple projection	73.46
	Multivariate projection	77.10
	None	76.98
SVM Accuracy [%]		
STFT	Simple projection	74.65
	Multivariate projection	74.95
	None	74.47
CWT	Simple projection	77.27
	Multivariate projection	78.00
	None	77.45

V. ACKNOWLEDGMENTS

This work is supported by "*Centro de Investigación e Innovación de Excelencia – ARTICA*", financed by COLCIENCIAS, and "*Becas para estudiantes sobresalientes de posgrado* ", financed by Universidad Nacional de Colombia.

REFERENCES

[1] P. Chazal, C. Heneghan, E. Sheridan, R. Reilly, P. Nolan, and M. O'Malley, "Automated processing of the single-lead electrocardiogram for the detection of obstructive sleep apnoea," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 6, pp. 686–696, 2003.

[2] H. Abdullah, G. Holland, I. Cosic, and D. Cvetkovic, "Correlation of sleep EEG frequency bands and heart rate variability," in *31st Annual International Conference of the IEEE EMBS*, 2009, pp. 5014–5017.

[3] Z. Deng, C. Poon, N. Arzeno, and E. Katz, "Heart rate variability in pediatric obstructive sleep apnea," in *Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE*, aug. 2006, pp. 3565 –3568.

[4] R. Bailón, L. Mainardi, M. Orini, L. Sörnmo, and P. Laguna, "Analysis of heart rate variability during exercise stress testing using respiratory information," *Biomedical Signal Processing And Control*, no. 4, pp. 299–310, 2010.

[5] M. Sun, M. Scheuer, and R. Sclabassi, "Decomposition of biomedical signals for enhancement of their time–frequency distributions," *Journal of the Franklin Institute*, vol. 337, pp. 453–467, 2000.

[6] L. Sepulveda-Cano, C. Acosta-Medina, and G. Castellanos-Dominguez, "Relevance analysis of stochastic biosignals for identification of pathologies," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, 2011.

[7] J. Martínez-Vargas, L. Sepúlveda-Cano, and G. Castellanos-Domínguez, "On determining available stochastic features by spectral splitting in obstructive sleep apnea detection," in *33rd Annual International Conference of the IEEE EMBS, Boston, Massachusetts USA*, IEEE, Ed., August 30 - September 3, 2011, pp. 6079–6082.

[8] J. Sadowsky, "Investigation of signal characteristics using the continuous wavelet transform," *JOHNS HOPKINS APL TECHNICAL DIGEST,*, vol. 17, no. 3, pp. 258 – 269, 1996.

[9] L. Sepulveda-Cano, E. Gil, P. Laguna, and G. Castellanos-Dominguez, "Selection of Nonstationary Dynamic Features for Obstructive Sleep Apnoea Detection in Children," *EURASIP Journal on Advances in Signal Processing*, vol. 2011, p. 10, 2011.

[10] D. Newandee and S. Reismad, "Wavelet representation comparison for heart rate variability analysis," in *Bioengineering Conference, 2003 IEEE 29th Annual*, 2003, pp. 112–113.

[11] M. Lado, X. Vila, L. Rodríguez-Liñares, A. Méndez, D. Olivieri, and P. Félix, "Detecting sleep apnea by heart rate variability analysis: Assessing the validity of databases and algorithms," *Journal of Medical Systems*, vol. DOI:10.1007/s10916-009-9383-5, 2009.