Usefulness of discrete wavelet transform in the analysis of oximetry signals to assist in childhood sleep apnea-hypopnea syndrome diagnosis

Fernando Vaquerizo-Villar, *Student Member, IEEE*, Daniel Álvarez, Gonzalo C. Gutiérrez-Tobal, *Member, IEEE*, Verónica Barroso-García, Leila Kheirandish-Gozal, Andrea Crespo, Félix del Campo, David Gozal, Roberto Hornero, *Senior Member, IEEE*

Abstract—Sleep apnea hypopnea syndrome (SAHS) is a highly prevalent respiratory disorder that may cause many negative consequences for the health and development of children. The gold standard for diagnosis is the overnight polysomnography (PSG), which is a high cost, complex, intrusive, and time-demanding technique. To improve the early detection of pediatric SAHS, we propose an automated analysis of the SpO_2 signal from nocturnal oximetry. A database composed of 298 SpO_2 recordings from children ranging from 0 to 13 years old was used for this purpose. Due to the abrupt changes caused by respiratory events in the SpO_2 signal, our goal was to evaluate the diagnostic ability of this by means of the discrete wavelet transform (DWT). To achieve this objective, we conducted a signal processing approach divided into two main stages: (i) feature extraction, where features from the DWT detail coefficients were computed, and (ii) feature classification, where a logistic regression (LR) model was used to classify children into SAHS negative or SAHS positive. Our results showed that respiratory events introduced more variability in two detail levels of the DWT from SpO_2: 0.024-0.049 Hz and 0.012-0.024 Hz. Moreover, the LR classifier achieved an 81.9% accuracy (79.1% sensitivity and 84.1% specificity) in an independent test set for a clinical cutoff point of 5 events/h, as derived from PSG. These results suggest that DWT analysis may be a useful tool to analyze SpO_2 recordings in the context of childhood SAHS.

I. Introduction

The American Academy of Pediatrics (AAP) defines the sleep apnea-hypopnea syndrome (SAHS) as a breathing disorder characterized by a prolonged partial obstruction of the upper airway (hypopnea) and/or intermittent complete cessation (apnea) of airflow during sleep [1]. Pediatric SAHS is highly prevalent (in the range of 1% to 5%) and may lead to multiple negative consequences in children's health and development, such as cardiometabolic dysfunction, neurocognitive deficit, and stunting [1].

The gold standard test for childhood SAHS diagnosis is overnight polysomnography (PSG), where multiple biomedical signals are recorded during sleep [1]. Nevertheless, PSG is costly, it requires the stay of the patient and qualified staff during the whole night in a specialized sleep laboratory [1]. Furthermore, PSG is highly intrusive due to the use of multiple sensors, and shows limited availability, which results in long waiting lists since initial referral by primary care physicians, and until the patients are ultimately diagnosed and treated [2], [3].

These limitations, together with the high prevalence of the disease, have prompted exploration on the use of simplified diagnosis techniques [2]. The AAP guidelines recommend performing alternative tests when overnight PSG is not available [1]. Thus, an interesting approach is to evaluate the diagnostic ability of a reduced set of signals included in the PSG.

In this regard, the usefulness of overnight oximetry as a simplified alternative method is being widely evaluated. Oximetry records the pulse rate and the blood oxygen saturation (SpO_2) using a pulse-oximeter probe usually placed on the child's finger, toe, or earlobe [4]. Due to its simplicity, reliability and suitability for children [2], [5], we propose an automated analysis of SpO_2 recordings in order to simplify pediatric SAHS diagnosis.

Previous research has supported the potential usefulness of automated analysis of SpO_2 recordings in the context of childhood SAHS diagnosis [5]–[9]. Several of the previous studies assessed frequency domain features from power spectral density (PSD) [5], [6], [8], which contains information related with the recurrence and duration of apnea events. Nevertheless, PSD is based in Short-Time Fourier Transform (STFT), which offers a fixed time-frequency resolution. It is not appropriate for detecting abrupt changes in biomedical signals, such as SpO_2 desaturations elicited by apneic events [10], [11]. By contrast, wavelet transform (WT) offers good frequency resolution at low frequencies and good time resolution at high frequencies [10]. This good resolution at low frequencies makes WT a well suited tool for

F. Vaquerizo-Villar, G. C. Gutiérrez-Tobal, V. Barroso-García and R. Hornero, are with the Biomedical Engineering Group, E.T.S.I. Telecomunicación, Universidad de Valladolid, Paseo de Belén 15, 47011, Valladolid, Spain (phone: +34 983 423000, ext. 3708) (e-mail: fernando.vaquerizo@gib.tel.uva.es).

D. Álvarez, F. del Campo and A. Crespo are with the Hospital Universitario Río Hortega of Valladolid, Spain (e-mail: fsas@telefonica.net).

L. Kheirandish-Gozal, and D. Gozal, are with the Section of Sleep Medicine, Dept. of Pediatrics, Pritzker School of Medicine, Biological Sciences Division, The University of Chicago, Chicago, IL, 60637, USA (e-mail: dgozal@uchicago.edu).

analyzing the non-stationary properties of the SpO_2 signal caused by respiratory events. Nevertheless, there are no large previous studies assessing its screening capability in the context of pediatric SAHS by means of automated analysis of oximetric recordings.

Based on aforementioned considerations, our hypothesis is that wavelet analysis can be useful in the diagnosis of childhood SAHS. Therefore, our objective was to evaluate the diagnostic performance of oximetric features derived from DWT analysis in the context of pediatric SAHS. This analysis was conducted in two phases: feature extraction and feature classification. In the first stage, features from DWT detail coefficients were computed and, in the second phase, these features were combined by means of a logistic regression (LR) classifier with the purpose of improving the diagnostic ability of individual features.

II. SUBJECTS AND SIGNALS UNDER STUDY

In this study, the dataset was composed of 298 children (166 boys and 132 girls) ranging from 0 to 13 years of age. All patients were referred to the Pediatric Sleep Unit at the University of Chicago Medicine Comer Children's Hospital (Chicago, IL, USA) due to clinical suspicion of SAHS. In all cases, an informed consent to participate in the research was obtained and the Ethical Committee approved the protocol.

Children underwent overnight in-lab PSG between 22.00 and 08.00. Sleep was monitored with a digital polysomnography system (Nihon Kohden America Inc., CA, USA). SpO_2 recordings were obtained from PSG at a sampling rate of 25 Hz. In a preprocessing stage, artifacts were removed from SpO_2 signals by eliminating drops to zero, sudden changes between consecutive SpO_2 sampling intervals $\geq 4\%$/second [12] and SpO_2 values below 60%.

According to the American Academy of Sleep Medicine rules, sleep and cardiorespiratory events were quantified and the apnea-hypopnea index (AHI) was derived [13]. An AHI cutoff of 5 events/h was considered as a positive SAHS because surgical treatment is routinely recommended in these cases [14].

The population was divided into a training set (149 subjects, 50%), used for training the LR classifier with the DWT features, and a test set (149 subjects, 50%), employed for assessing the diagnostic performance of DWT individual features and the classifier. Table I shows demographic and clinical data of the population under study (median [interquartile range]).

III. METHODS

The methodology proposed in this study is illustrated in Fig 1. Firstly, DWT is applied in order to extract features from the SpO_2 recordings. Then, a LR model was created with these features to classify the children into SAHS positive or SAHS negative. The training set was used to optimize the LR model and the test was used in order to evaluate the classifier in an independent set.

A. Discrete Wavelet Transform and feature extraction

WT is seen as the decomposition of a signal into a set of basis functions, called wavelets, which are obtained from a mother wavelet by scaling and time translations. Therefore,

TABLE I.   DEMOGRAPHIC AND CLINICAL DATA

|	All subjects		
	All	**SAHS negative**	**SAHS positive**
Subjects (n)	298	164	134
Age (years)	6 [4-9]	7 [5-10]	5 [3-9]
Males (n)	166 (55.70%)	91 (55.49%)	75 (55.97%)
BMI (kg/m²)	18.37 [16.33-23.04]	18.16 [16.32-22.27]	18.66 [16.36-24.08]
AHI (e/h)		1.95 [0.96-3.45]	11.11 [7.53-18.86]
	Training		
	All	**SAHS negative**	**SAHS positive**
Subjects (n)	149	82	67
Age (years)	6 [4-9]	7 [5-10]	6 [3-9]
Males (n)	88 (59.06%)	51 (62.20%)	37 (55.22%)
BMI (kg/m²)	18.34 [16.40-23.18]	17.90 [16.35-21.67]	19.03 [16.50-24.22]
AHI (e/h)		2.01 [0.95-3.39]	11.09 [7.49-19.59]
	Test		
	All	**SAHS negative**	**SAHS positive**
Subjects (n)	149	82	67
Age (years)	7 [4-9]	7 [5-10]	5 [3-9]
Males (n)	78 (52.35%)	40 (48.78%)	38 (56.72%)
BMI (kg/m²)	18.42 [16.26-22.76]	18.33 [16.30-22.66]	18.42 [16.05-22.80]
AHI (e/h)		1.90 [0.97-3.49]	11.13 [7.56-18.41]

BMI: Body Mass Index; AHI: Apnea Hypopnea Index
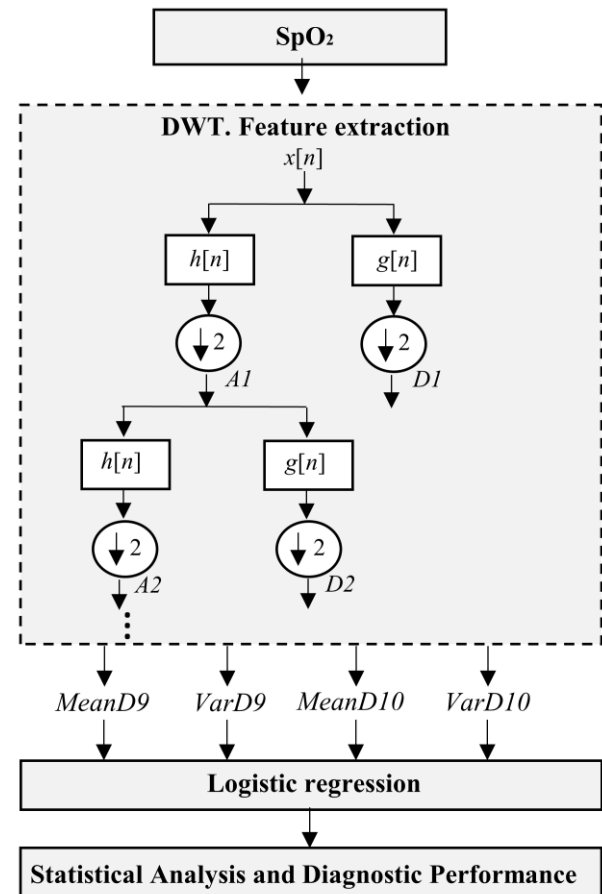
Figure 1. Block diagram of the proposed methodology for SAHS diagnosis

WT is a time-scale representation [10]. Depending on the scale and translation values, WT can be continuous or discrete [15]. Continuous Wavelet Transform (CWT) calculates WT for every possible scale, while DWT only computes wavelet coefficients for dyadic scales, so it is computationally less expensive. As a consequence, DWT was chosen for this work [15].

The decomposition process of a signal $x[n]$ with the DWT is a filter bank tree, as it can be seen in Fig. 1. At each stage, a high pass filter $g[n]$, the mother wavelet, a low pass filter $h[n]$, the mirror version of the wavelet, and two downsamplers by a factor 2 are implemented [15]. In the first level, the signal $x[n]$ is decomposed in an approximation signal, *A1*, and a detail signal, *D1*. Then, *A1* is decomposed in an approximation signal, *A2*, and a detail signal, *D2*, and this process continues until reaching the maximum desired detail level of the signal to analyze, as shown in Fig.1 [15].

In this study, DWT was applied to $N=2^{14}$ sample segments and daubechies44 was the mother wavelet selected, as recommended for biomedical signals [16]. Ten-level DWT analysis was performed and the following frequency bands for the detail coefficients were obtained for each of them:

- Level 1. From 6.25 Hz to 12.5 Hz.
- Level 2. From 3.125 Hz to 6.25 Hz.
- Level 3. From 1.563 Hz to 3.125 Hz.
- Level 4. From 0.781 Hz to 1.563 Hz.
- Level 5. From 0.391 Hz to 0.781 Hz.
- Level 6. From 0.195 Hz to 0.391 Hz.
- Level 7. From 0.098 Hz to 0.195 Hz.
- Level 8. From 0.049 Hz to 0.098 Hz.
- Level 9. From 0.024 Hz to 0.049 Hz.
- Level 10. From 0.012 Hz to 0.024 Hz.

Related work in the context of pediatric SAHS diagnosis determined the following frequency bands of interest for childhood SAHS: 0.0137-0.0473 Hz and 0.021-0.040 Hz [6], [8]. Therefore, we have analyzed the detail coefficients of the levels 9 (*D9*) and 10 (*D10*), which contain the bands of previous works [6], [8]. The features extracted from *D9* and *D10* levels were the mean (*MeanD9* and *MeanD10*) and the variance (*VarD9* and *VarD10*) of the detail coefficients. Features were computed for each segment of $N=2^{14}$ samples and subsequently averaged over all epochs of each recording to obtain a single value per subject.

B. Logistic Regression

LR estimates the posterior probability of a given instance (subject) belonging to one of two mutually exclusive classes (SAHS negative vs. SAHS positive) giving a set of input features (in our case, DWT features) [17]. Therefore, it has been widely used as a binary classifier in the context of computer aided diagnosis systems [17]. LR models the probability density function as a Bernouilli distribution and it uses the maximum likelihood ratio to optimize coefficients of the input features of the classifier [17].

C. Statistical analysis and diagnostic performance

The non-parametric Mann-Whitney *U* test was used to search for statistical differences in the DWT features between the groups under study. A *p*-value under 0.01 was considered significant. Diagnostic performance of the individual features and the LR model in the test set was assessed by means of sensitivity (*Se*, percentage of SAHS positive patients correctly classified), specificity (*Sp*, percentage of SAHS negative children correctly classified), positive predictive value (*PPV*, proportion of subjects classified as positive that are true positives), negative predictive value (*NPV*, proportion of subjects classified as negative that are true negatives), positive likelihood ratio (*LR+*, likelihood ratio for subjects classified as positive), negative likelihood ratio (LR-, likelihood ratio for subjects classified as negative) and accuracy (*Acc*, percentage of subjects correctly classified).

IV. RESULTS

A. Training set

DWT features were computed in the training set and the optimum classification threshold was determined by means of a ROC analysis for each single feature. Table II displays the average values of the parameters for both groups (median [interquartile range]). The variance in both detail levels, *VarD9* and *VarD10*, were significantly higher in SAHS positive, which suggest that the variability in this detail bands is related with apneic events. On the other hand, no significant differences were found linked with the mean values in these detail levels, *MeanD9* and *MeanD10*. Then, coefficients of the LR model were subsequently estimated using all the single features.

B. Test set

Table III summarizes the diagnostic performance of the individual features and the LR model in the test set. Regarding every single feature, the parameter *VarD9* reached the highest accuracy (79.9%), while the LR model including the 4 DWT features achieved 79.1% *Se*, 84.1% *Sp* and an accuracy of 81.9%, outperforming the highest individual feature (*VarD9*).

V. DISCUSSION

In this study, a novel approach for pediatric SAHS diagnosis based on a frequency domain analysis of SpO$_2$ recordings through DWT was developed. This approach was divided in two stages: (i) feature extraction, where mean and variance of the DWT coefficients were computed in the detail levels *D9* and *D10*, and (ii) feature classification, where a LR model was built with these features.

It was observed that variance of the coefficients in both detail levels were higher in SAHS positive cases, which is consistent with the effect of respiratory events on the frequency bands of these detail levels[6], [8]. *VarD9* showed a remarkable diagnostic performance, reaching an accuracy of 79.9% in the test set, with a balanced *Se-Sp* pair (80.6%-79.3%). Moreover, the joint analysis of all the features by means of a LR classifier improved the diagnostic ability of each of the single features, achieving an accuracy of 81.9% in the test set, also with a balanced *Se-Sp* pair (79.1%-84.1%).

TABLE II. MEDIAN, INTERQUARTILE RANGE AND P-VALUE OF THE DWT FEATURE VALUES FOR SAHS NEGATIVE AND SAHS POSITIVE GROUPS IN THE TRAINING SET

	SAHS negative	SAHS positive	*p*-value
MeanD9(10^{-3})	0.97 [-1.53 - 3.43]	1.21 [-3.11 - 4.96]	0.97
MeanD10(10^{-4})	-1.71 [-9.47 - 5.21]	3.32 [-1.29 - 1.07]	0.20
VarD9	3.61 [2.97 - 4.41]	5.85 [4.51 - 7.88]	$p<0.01$
VarD10	4.26 [3.43 - 5.18]	6.63 [5.19 - 9.20]	$p<0.01$

TABLE III. DIAGNOSTIC ASSESMENT OF EACH SINGLE FEATURE AND THE LR MODEL IN THE TEST SET

	Se	Sp	PPV	NPV	LR+	LR-	Acc
MeanD9	37.3	65.9	47.2	56.3	1.09	0.95	53.0
MeanD10	47.8	53.7	45.7	55.7	1.03	0.97	51.0
VarD9	80.6	79.3	76.1	83.3	3.89	0.24	79.9
VarD10	77.6	74.4	71.2	80.3	3.03	0.30	75.8
LR	79.1	84.1	80.3	83.1	4.99	0.25	81.9

These results agree with related work focusing on pediatric SAHS diagnosis from SpO$_2$ recordings. Garde et al. [5] achieved 88.4% *Se* and 83.6% *Sp*, by combining PSD and time domain features from 146 SpO$_2$ and pulse rate recordings in a LDA model. The study by Gutiérrez-Tobal et al. combined PSD features from 50 airflow recordings with oxygen desaturation index (ODI) from SpO$_2$ signal using a LR model, achieving 85.9% *Se* and 87.4% *Sp*, and 86.3% *Acc* [7]. Similarly, Cohen and de Chazal analyzed a database composed of 396 subjects, achieving an accuracy of 66.7%, with 58.1% *Se* and 67.0% *Sp* combining time domain features from the SpO$_2$ signal and time domain and PSD features from ECG recordings [9]. In our work, we use one single biomedical signal, SpO$_2$ from nocturnal oximetry, using a large database (298 recordings).

This study presents some limitations. First, only binary classification has been performed, while clinical studies recommend estimating the degree of severity. Second, it would be interesting to extract more features from the DWT detail coefficients and to use conventional oximetry indices. Moreover, a feature selection stage should be applied, in order to obtain a subset of features that are complementary. Finally, more classifiers should be assessed with the objective of improving the diagnostic performance of the SpO$_2$ signal.

VI. CONCLUSIONS

In summary, we have developed a method in the context of pediatric SAHS based on the automated analysis of SpO$_2$ recordings by means of DWT. Our results showed increased variance in the DWT coefficients of childhood SpO$_2$ recordings in the frequency bands described in the literature (0.024-0.049 Hz and 0.012-0.024 Hz), which appears to be related to slow variations in the SpO$_2$ signals due to SAHS.

Furthermore, a LR model trained with the mean and variance of DWT coefficients in these bands reached high diagnostic ability in an independent test set. Therefore, we conclude that DWT could be a useful tool for analyzing SpO$_2$ recordings and assist in the process of achieving a pediatric SAHS diagnosis in high pre-test probability cases.

REFERENCES

[1] C. L. Marcus *et al.*, "Diagnosis and management of childhood obstructive sleep apnea syndrome.," *Pediatrics*, vol. 130, no. 3, pp. 576–84, 2012.

[2] G. M. Nixon, A. S. Kermack, G. M. Davis, J. J. Manoukian, A. Brown, and R. T. Brouillette, "Planning adenotonsillectomy in children with obstructive sleep apnea: the role of overnight oximetry," *Pediatrics*, vol. 113, no. 1, pp. e19–e25, 2004.

[3] E. S. Katz, R. B. Mitchell, and C. M. D. Ambrosio, "Obstructive Sleep Apnea in Infants," *Am. J. Respir. Crit. Care Med.*, vol. 185, no. 8, pp. 805–816, 2012.

[4] N. Netzar, A. H. Eliasson, C. Netzar, and D. A. Kristo, "Overnight Pulse Oximetry for Sleep- Disordered Breathing in Adults," *CHEST J.*, vol. 120, no. 2, pp. 625–633, 2001.

[5] A. Garde, P. Dehkordi, W. Karlen, D. Wensley, J. M. Ansermino, and G. A. Dumont, "Development of a screening tool for sleep disordered breathing in children using the phone oximeterTM," *PLoS One*, vol. 9, no. 11, 2014.

[6] D. Álvarez *et al.*, "Automated Analysis of Nocturnal Oximetry as Screening Tool for Childhood Obstructive Sleep Apnea-Hypopnea Syndrome," in *37th Annual International Conference of the IEEE In Engineering in Medicine and Biology Society (EMBC)*, 2015, pp. 2800–2803.

[7] G. C. Gutiérrez-Tobal, M. L. Alonso-Álvarez, D. Álvarez, F. del Campo, J. Terán-Santos, and R. Hornero, "Diagnosis of pediatric obstructive sleep apnea: Preliminary findings using automatic analysis of airflow and oximetry recordings obtained at patients' home," *Biomed. Signal Process. Control*, vol. 18, pp. 401–407, 2015.

[8] G. C. Gutiérrez-Tobal *et al.*, "Analysis and Classification of Oximetry Recordings to Predict Obstructive Sleep Apnea Severity in Children," in *37th Annual International Conference of the IEEE In Engineering in Medicine and Biology Society (EMBC)*, 2015, pp. 4540–4543.

[9] G. Cohen and P. De Chazal, "Automated detection of sleep apnea in infants : A multi-modal approach," *Comput. Biol. Med.*, vol. 63, pp. 118–123, 2015.

[10] O. Rioul and M. Vetterli, "Wavelets and signal processing.pdf," *IEEE signal processing magazine*, vol. 8, no. 4, pp. 14–38, 1991.

[11] I. Daubechies, "The Wavelet Transform , Time-Frequency Localization and Signal Analysis," *IEEE Trans. Inf. theory*, vol. 36, no. 5, pp. 961–1005, 1990.

[12] U. J. Magalang and J. Dmochowski, "Prediction of the Apnea-Hypopnea Index From Overnight Pulse Oximetry," *CHEST J.*, vol. 124, no. 5, pp. 1694–1701, 2003.

[13] R. B. Berry, R. Budhiraja, G. D.J, and E. Al., "Rules for scoring respiratory events in sleep: update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events. Deliberations of the Sleep Apnea Definitions Task Force of the American Academy of Sleep Medicine," *J Clin Sleep Med*, vol. 8, no. 5, pp. 597–619, 2012.

[14] H.-L. Tan, D. Gozal, H. M. Ramirez, H. P. R. Bandla, and L. Kheirandish-Gozal, "Overnight polysomnography versus respiratory polygraphy in the diagnosis of pediatric obstructive sleep apnea," *Sleep*, vol. 37, no. 2, pp. 255–260, 2014.

[15] D. Cvetkovic, E. D. Übelyi, and I. Cosic, "Wavelet transform feature extraction from human PPG , ECG , and EEG signal responses to ELF PEMF exposures : A pilot study," *Digit. Signal Process.*, vol. 18, no. 5, pp. 861–874, 2008.

[16] J. Rafiee, M. A. Rafiee, N. Prause, and M. P. Schoen, "Wavelet basis functions in biomedical signal processing," *Expert Syst. Appl.*, vol. 38, no. 5, pp. 6190–6201, 2011.

[17] D. Hosmer and S. Lemeshow, *Applied logistic regression*. John Wiley & Sons, 2004.