Detecting Obstructive Sleep Apnea in Children by Self-Affine Visualization of Oximetry

Ainara Garde, Member, IEEE, Parastoo Dekhordi, Student Member, IEEE, Christian L Petersen, J Mark Ansermino, and Guy A Dumont, Fellow, IEEE

Abstract— **Obstructive sleep apnea (OSA), characterized by cessations of breathing during sleep due to upper airway collapse, can affect the healthy growth and development of children. The gold standard for OSA diagnosis, polysomnography (PSG), is expensive and resource intensive, resulting in long waiting lists to perform a PSG.**

Previously, we investigated the time-frequency analysis of blood oxygen saturation (SpO_2) to screen for OSA. We used overnight pulse oximetry from 146 children, collected using a smartphone-based pulse oximeter (Phone Oximeter), simultaneously with standard PSG. Sleep technicians manually scored PSG and provided the average of apnea/hypoapnea events per hour (AHI).

In this study, we proposed an alternative method for analyzing SpO_2, in which a set of contracting transformations form a self-affine set with a 2D attractor, previously developed for qualitative visualization of the photoplethysmogram and electroencephalogram. We applied this technique to the overnight SpO_2 signal from individual patients and extracted features based on the distribution of points (radius and angle) in the visualization. The cloud of points in children without OSA (NonOSA) was more confined than in children with OSA, which was reflected by more empty pixels (radius and angles). The maximum value, skewness and standard deviation of the distribution of points located at different radius and angles were significantly (Bonferroni corrected) higher in NonOSA compared to OSA children. To detect OSA defined at different levels (AHI≥5, AHI≥10 and AHI≥15), three multivariate logistic regression models were implemented using a stepwise feature selection and internally validated through bootstrapping. The models (AHI≥5, AHI≥10, AHI≥15), consisting of 3, 4 and 1 features respectively, provided a bootstrap-corrected AUC of 73%, 81%, 73%. Thus, applying this visualization to nocturnal SpO_2 could yield both visual and quantitative information that might be useful for screening children for OSA.

I. Introduction

The diagnosis of obstructive sleep apnea (OSA) in children presents a challenging problem. OSA is characterized by frequent complete (apnea) or partial (hypopnea) cessations of breathing during sleep due to periodic collapse of the upper airway. OSA affects up to 6% of children [1], [2] and poses a serious threat to their healthy growth and development. The frequent sleep fragmentations and oxygen desaturations experienced during sleep have been associated with daytime sleepiness, growth failure, behavioural problems and developmental delays. The primary cause of pediatric OSA is enlarged tonsils and adenoids [3] and the highest OSA prevalence is in children aged 2-5 years, when adenoids and tonsils are largest relative to the airway [4].

To diagnose and determine OSA severity children must undergo overnight sleep test or polysomnography (PSG), the gold standard for OSA diagnosis. PSG is a resource intensive study involving overnight stay in a well-equipped sleep laboratory and measurement of electrocardiography, chest movement, brain activity, oxygen saturation, nasal airflow, and video/audio data. Thus, a very limited number of children can have access to PSGs each year. In British Columbia, Canada, all children referred for PSG must be tested overnight at BC Children's Hospital in Vancouver, which has the capacity to perform only 250 PSGs annually. As a result, the waitlist is approximately eight months long, delaying diagnosis and treatment for children in need. Examination of PSG data allows identification of apnea/hypopnea events (defined as interruption/reduction of airflow), arousals and associated oxyhemoglobin desaturations. The average number of apnea and hypopnea events per hour (the apnea/ hypopnea index [AHI]), is used to determine OSA severity.

Pulse oximetry, a part of the standard of care for PSG, is a simple, non-invasive method of measuring blood oxygen saturation (SpO_2) using a probe attached to the finger or toe. Overnight oximetry or SpO_2 analysis, has been previously explored as a potential standalone method to screen for subjects with OSA. Nixon et al. developed and validated a severity scoring system, refer to as the McGill Oximetry Score (MOS), using overnight oximetry, as a tool to prioritize the adenotonsillectomy surgical list [5]; Álvarez et al. demonstrated that the characterization of overnight oximetry provided relevant information to identify adults with significant OSA [6].

Smartphones have been adopted widely, offering an ideal platform for vital signs assessment and automatic diagnosis. The Phone Oximeter (a mobile device that integrates a pulse oximeter with a smartphone [7], [8]), provides a very useful platform to record overnight pulse oximetry. SpO_2 fluctuations caused by episodes of apnea modulate the SpO_2 signal. Thus, in previous studies, we have characterized nocturnal SpO_2, recorded by the Phone Oximeter, in time and frequency domain, to identify subjects with OSA [9], [10]. However, SpO_2 exhibits complex non-linear information as

*This work was supported in part by The Natural Sciences and Engineering Research Council of Canada (NSERC), the Canadian Institutes of Health Research (CIHR) and MITACS Elevate Fellowship Program.

A. Garde is with the Biomedical Signals and Systems Department at the Faculty EEMCS, University of Twente, Postbox 217, 7500 AE Enschede, The Netherlands. (email: A.gardemartinez@utwente.nl).

P. Dehkordi, C. L. Petersen, J. M. Ansermino and G. A. Dumont are with the Electrical and Computer Engineering in Medicine Group (ECEM), Departments of Electrical & Computer Engineering and Anesthesiology, Pharmacology & Therapeutics, University of British Columbia, 1L7-4480 Oak Street, Vancouver, BC Canada V6H 3V4.

well. This complexity is often ignored in conventional analysis, where time or frequency domain filtering is used to extract simple numeric trends. Thus, in this study we propose an alternative method of analysis, in which a set of contracting transformations form a self-affine set with a 2D attractor [11]. We investigate the effect of OSA on SpO_2 attractor patterns, and implement a smartphone application for visualization of the SpO_2 attractor.

II. DATASET

We recruited 160 children (59 female and 87 male), 9.1± 4.3 years old with signs of OSA, who were referred to the British Columbia Children's Hospital for a PSG sleep study. All children were recruited per a protocol approved by the University of British Columbia and Children's and Women's Health Centre of British Columbia Research Ethics Board (H11-01769). Parental/guardian written informed consent was obtained for all subjects.

Data acquisition was carried out in the sleep laboratory. Overnight measurement of electrocardiography (ECG), electroencephalography (EEG), pulse oximetry, chest movement, and nasal airflow, as well as video recording, were recorded, using the Embla Sandman S4500.

A second pulse oximeter sensor was applied to the finger adjacent to the one used during the standard PSG. This sensor was attached to the smartphone-based pulse oximeter, and SpO_2 was recorded at 1 Hz alongside PSG.

TABLE I. DATASET BASELINE

Subjects (female, male)	146* (59, 87)
Age (years old)	9.11 ± 4.23
Body Mass Index [BMI] (kg/m^2)	20.95 ± 7.97
AHI (apnea/hypoapnea /hour)	8.42 ± 14.97

*Fourteen children were excluded from analysis based on having a total sleep duration, or signal data duration (from PSG or the smartphone-based pulse oximeter) shorter than 3 hours.

III. METHODS

A. SpO_2 signal pre-processing

Displacement of the sensor, motion artifacts, ambient light, and low blood perfusion can corrupt oximetry recordings. Thus, we implemented a signal quality index (SQI), defined by: $SQI(n) = \begin{cases} 0, & SpO_2 \text{ unreliable} \\ 1, & SpO_2 \text{ reliable} \end{cases}$, to indicate whether the quality of the SpO_2 signal was appropriate for further analysis. The SpO_2 samples below 70% or above 100% were considered as artifacts and consequently the SQI of these samples was set to zero. The SQI of SpO_2 segments shorter than 30s located between unreliable SpO_2 (SQI=0) samples were set to zero. The SQI was be used to identify oximetry segments of sufficient quality to warrant appropriate further signal processing analysis.

B. Self-affine visualization

We considered a unit disk to map the current SpO_2 signal sample $SpO_2(n)$ to a point $P(n)$ on the periphery through, $P(n)=[\cos(SpO_2(n)); \sin(SpO_2(n))]$ and plot a point $p(n)$ midway between the previous point $p(n-1)$ (initially the center of the disk) and the periphery location $P(n)$:

$$p(n)=p(n-1)+\frac{1}{2}(P(n)-p(n-1))$$

The resulting cloud of points encodes the history of the signal and lies on the attractor of its self-affine set. The fractal pattern appearance is determined by the unfiltered self-similar structure (see Figure 1). In this study, six hours of reliable SpO_2 (SQI=1) 1 Hz data at 0.1% and 1% resolution were mapped, using this method, for each patient.

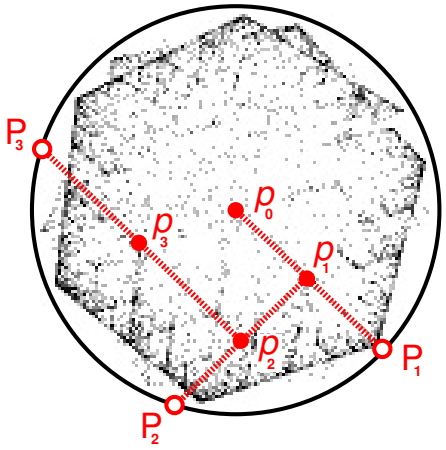

Figure 1. Illustration of the process of calculating the cloud of points using the SpO_2 signal at 1% resolution.

C. Feature extraction

Several features were computed to characterize the dispersion of points in the self-affine visualization. The visualization was divided in a grid of 200x200 pixels, and a matrix with the number of points located at each pixel was computed. In addition, the radius and angle (polar coordinates) of every point was calculated. A vector (Radius_vec) with the percentage of points (number of points/total points) located at different ratios (ranging from 0 to 1 in steps of 0.01) was calculated. A second vector (Angle_vec) with the percentage of points located at different angles (ranging from 0° to 360° in steps of 2°) was computed.

TABLE II. DESCRIPTION OF THE FEATURES EXTRACTED FROM ANALYZING THE SELF-AFFINE VISUALIZATION

Features	**Description**
Empty_pixel	Percentage of empty pixels (number of empty pixels/total pixels)
Empty_radius	Percentage of empty radius (number of empty radius/total radius)
Empty_angle	Percentage of empty angles (number of empty angles/total angles)
KR	Kurtosis of the Radius_vec
SR	Skewness of the Radius_vec
StdR	Standard deviation of the Radius_vec
IqrR	Interquartile range of the Radius_vec
maxR	Maximum value of Radius_vec
KA	Kurtosis of the Angle_vec
SA	Skewness of the Angle_vec
StdA	Standard deviation of the Angle_vec
IqrA	Interquartile range of the Angle_vec
maxA	Maximum value of Angle_vec

D. Data analysis

First, a Mann-Whitney U test was applied to evaluate the differences of each feature in NonOSA and OSA patients with different severities (AHI≥5, AHI≥10 and AHI≥15). Bonferroni correction was used to adjust for multiple (n=13) comparisons. A probability of p-value ≤ 0.05/n was considered significant. Then, a stepwise selection method was applied to identify the most relevant features and to develop three multivariate logistic regression models to detect patients with an AHI≥5, AHI≥10 and AHI≥15.

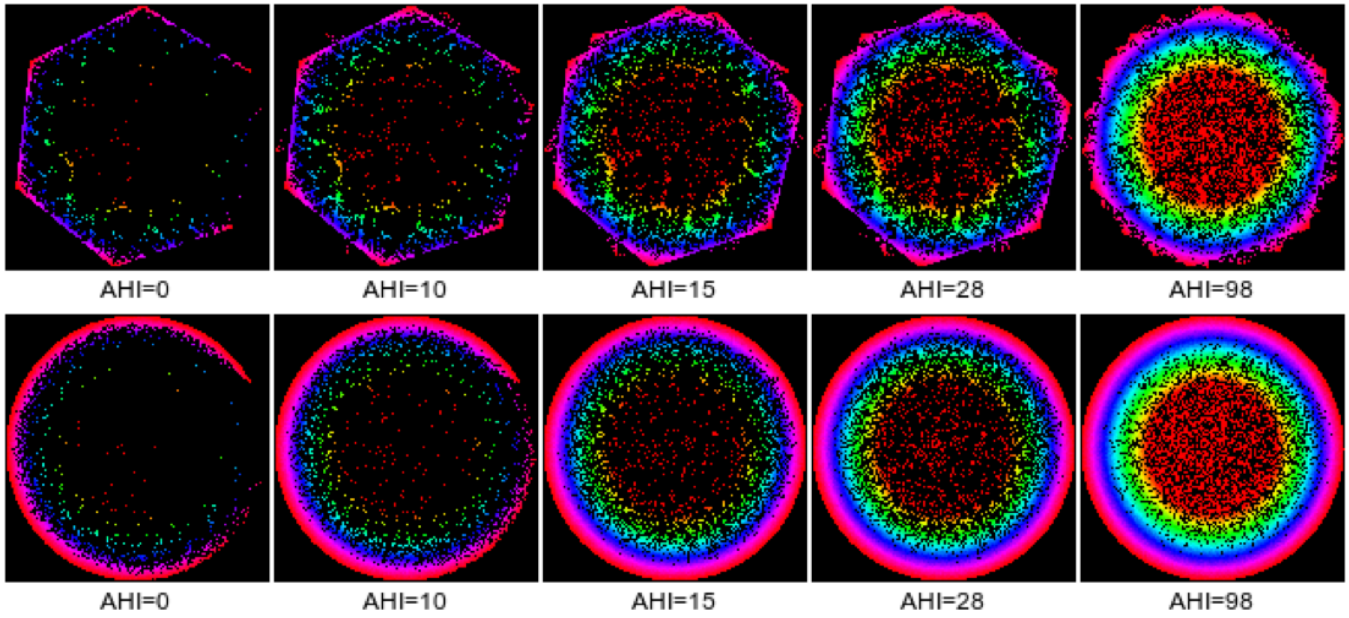

Figure 2. The self-affine visualization applied to SpO_2 readings (at 1% resolution in top figures and 0.1% resolution in bottom figures) for patients with different AHI. The dispersion of points is seen to increase with OSA severity.

The stepwise selection method added/dropped one feature at a time and stopped when further inclusion/exclusion no longer improved the model, as determined by the Akaike information criterion[12]. All features were included as candidate predictors and a minimum value of 40 events per variable was maintained to avoid overfitting [13].

We used the bootstrap method to validate our models, since it is recommended as the optimal technique for estimation of internal validity of a predictive logistic regression models. Using the original data set, 100 bootstrap samples were generated using sampling with replacement. The stepwise selection procedure was applied to each bootstrap sample to develop a multivariate model. This model was applied to predict the outcome in the bootstrap sample and in the original data set. The difference in the prediction performance, quantified by the area under the receiver operating characteristic curve (AUC), was computed to estimate the optimism of the models.

IV. RESULTS

A. Univariate analysis

As it is shown in figure 2, the cloud of points in children without OSA, is confined to some radius whereas in children with OSA (at different AHI thresholds), these points are more dispersed. This was reflected by greater values of empty spaces in the visualization (empty_pixel, empty_angle).

TABLE III. FEATURE DISTRIBUTION (MEAN STANDARD ±DEVIATION) FOR PATIENTS WITH AND WITHOUT OSA.

	AHI threshold=5		AHI threshold=10		AHI threshold=15	
Feat.	***NonOSA***	***OSA***	***NonOSA***	***OSA***	***NonOSA***	***OSA***
Empty_pixel	97.5±2	95.1±3	96.6±1.9	93.9±3.2	96.4±1.9	93.2±3.6
Empty_radius	10.5±2.9	10.3±3.1	10.8±3.1	9.3±1.7	10.6±3.1	9.3±1.9
Empty_angle	7.1±7.8	4.1±6.8	7.24±8.0	1.37±2.5	6.7±7.9	1.4±2.7
KR	93.1±5.73	89.6±6.31	93.1±5.33	87±6.7	92.6±5.51	86.5±7.32
SR	9.48±0.43	9.22±0.47	9.48±0.4	9.03±0.5	9.45±0.41	8.98±0.55
StdR	5.9±1.22	5.04±1.28	5.89±1.19	4.45±1.08	5.77±1.22	4.37±1.18
IqrR	0.53±0.18	0.64±0.19	0.53±0.18	0.72±0.15	0.55±0.18	0.73±0.16
maxR	58.8±12.5	50±13.2	58.7± 12.2	43.9±11.1	57.5±12.5	43.1±12.1
KA	98.4±37	92.8±35.1	100±36.6	83.1±32.1	99.1±35.6	79.6±36.1
SA	9.32±1.97	8.97±1.89	9.41±1.93	8.4±1.79	9.35±1.88	8.19±2.03
StdA	3.15±0.92	2.69±0.92	3.17±0.92	2.3±0.69	3.1±0.91	2.26±0.81
IqrA	0.14±0.09	0.19±0.1	0.14±0.08	0.23±0.09	0.15±0.09	0.24±0.10
maxA	35.8±14.3	30.3±14	36.2±14.4	24.8±10.5	35.3±14.1	24.1±12.3

*Colored background indicates features with p-value <0.05/n

Regarding the distribution of the points (percentage) located at different radius or angles, children without OSA presented higher maximum values, and higher skewness and standard deviation, than children with OSA (table III).

B. Multivariate modeling

Both, the final logistic regression model to identify an AHI≥5 and the final model to identify an AHI≥10 selected 4 features (Table IV and Table V, respectively), while the model to identify an AHI≥15 (see Table VI) selected only one feature (empty_pixel). All the selected features had p-values < 0.05. The three models were well calibrated (Hosmer-Lemeshow goodness of fit test p=0.2, p = 0.4, p = 0.1, respectively) and presented an area under the ROC curve (AUC) of at least 0.8 (see Table VII).

TABLE IV. PARAMETER ESTIMATES FROM THE FINAL MODEL TO IDENTIFY CHILDREN WITH AHI≥5

Features	Parameter Estimate	Standard Error	p-value
Intercept	3.40	2.59	0.19
maxA	0.13	0.04	0.00174
Empty_radius	3.66e-05	1.05e-05	0.00049
SA	1.08	4.12e-01	0.00898

TABLE V. PARAMETER ESTIMATES FROM THE FINAL MODEL TO IDENTIFY CHILDREN WITH AHI≥10

Features	Parameter Estimate	Standard Error	p-value
Intercept	2.53e+01	1.56e+01	0.11
SA	2.42	0.6	6.3e-05
maxA	4.06e-01	1.07e-01	0.00016
Empty_radius	3.06e-05	1.40e-05	0.02857
Empty_pixel	4.09e+01	1.78e+01	0.02187

TABLE VI. PARAMETER ESTIMATES FROM THE FINAL MODEL TO IDENTIFY CHILDREN WITH AHI≥15

Features	Parameter Estimate	Standard Error	p-value
Intercept	38.8	10.0	0.00011
Empty_pixel	-42.6	10.6	5.4e-05

C. Classification

Assuming that the cost of sending a child without OSA for a PSG is not overwhelmingly high, it would be reasonable to allow an unnecessary PSG of three NonOSA children (i.e. three false positives) in order to avoid missing one OSA case (i.e. one false negative). The optimal risk thresholds in this context (3:1) were 0.125, 0.28 and 0.12 for the model identifying an AHI≥5, AHI≥10, and AHI≥15, respectively. The best classification performance was obtained identifying children with AHI≥10 (table VII).

TABLE VII. CLASSIFICATION RESULTS REPRESENTED BY THE MEAN AND 95% CI OF THE ACCURACY (ACC), SENSITIVITY (SN), SPECIFICITY(SP), AREA UNDER THE ROC CURVE (AUC) AND OPTIMISM OF THE AUC

Model to identify	Mean and 95% confidence intervals				
	Acc (%)	Sn (%)	Sp (%)	AUC (%)	Optimism (%)
AHI≥5	74 [67,82]	74 [54,93]	75 [67,83]	80 [72,88]	6.6 [-0.4,13.1]
AHI≥10	77 [70,84]	79 [61,97]	77 [69,84]	87 [79,94]	6.1 [-0.9,12.3]
AHI≥15	75 [67,82]	79 [60, 97]	74 [65,82]	81 [71,91]	7.6 [-1.0,14.4]

An optimal risk threshold was calculated for each model [14] (0.125, 0.28, and 0.12, respectively).

V. CONCLUSION

We have performed quantitative analysis of a novel non-linear method for OSA visualization that is based on a simple self-affine transformation. The results suggest that when applied to nocturnal SpO_2, this innovative visualization can provide useful information to improve OSA screening, using pulse oximetry as a standalone tool. The visualization permits the extraction of features that could help identifying children with OSA, and applying simple pattern-based classifiers to the visualizations yields promising quantitative results. For further research, we plan to extend this analysis to provide information about OSA severity as well.

Furthermore, when integrated with a mobile phone oximeter system, the proposed visualization approach could open for a new regime of visual OSA interpretations that may be more accessible to patients when used in a home setting, allowing people everywhere to better access their own quality of sleep.

ACKNOWLEDGMENT

The authors would also like to especially thank the Pediatric Anesthesia Research Team at the BC Children's Hospital for their efforts in data collection and continuous support. The authors would also like to thank the clinical staff at BC Children's Hospital's sleep laboratory for their assistance with conducting this study and Aryannah Rollinson for her help in revising this manuscript.

REFERENCES

[1] J. H. Wildhaber et al "Sleep and respiration in children: time to wake up!," *Swiss Med. Wkly.*, vol. 137, no. 49–50, pp. 689–694, Dec. 2007.

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

[3] S. J. Chang et al "Obstructive sleep apnea syndrome in children: Epidemiology, pathophysiology, diagnosis and sequelae.," *Korean J. Pediatr.*, vol. 53, no. 10, pp. 863–871, Oct. 2010.

[4] R. Arens et al "Pathophysiology of Upper Airway Obstruction: A Developmental Perspective," *Sleep*, vol. 27, no. 5, pp. 997–1019, 2004.

[5] G. M. Nixon, et al "Planning Adenotonsillectomy in Children With Obstructive Sleep Apnea: The Role of Overnight Oximetry," *Pediatrics*, vol. 113, no. 1, pp. e19–e25, 2004.

[6] D. Alvarez, et al "Improving diagnostic ability of blood oxygen saturation from overnight pulse oximetry in obstructive sleep apnea detection by means of central tendency measure.," *Artif. Intell. Med.*, vol. 41, no. 1, pp. 13–24, Sep. 2007.

[7] C.L. Petersen, et al "Design and evaluation of a low-cost smartphone pulse oximeter.," *Sensors.*, vol. 13, no. 12, pp. 16882–93, Jan. 2013.

[8] J. Hudson, et al "Usability testing of a prototype Phone Oximeter with healthcare providers in high- and low-medical resource environments," *Anaesthesia*, vol. 67, no. 9, pp. 957–967, 2012.

[9] A. Garde, et al "Oxygen Saturation in Children with and without Obstructive Sleep Apnea Using the Phone-Oximeter," *Proc. IEEE Conf. Eng. Med. Biol.*, vol. 4, pp. 3–6, Jul. 2013.

[10] A. Garde, et al "Using oximetry dynamics to screen for sleep disordered breathing at varying thresholds of severity," in *2015 23rd European Signal Processing Conference, EUSIPCO 2015*, 2015.

[11] Petersen CL, et al, "iSwirl: Self-Affine Visualization of PPG and ECG Waveforms.," *Anesth. Analg.*, vol. 119 (6S), p. 72, 2014.

[12] H. Akaike, "Information theory and an extension of the maximum likelihood principle," in *B.N Petrov, F Csaki (Eds.), 2nd int. symposium on information theory, Academiai Kiado, Budapest*, 1973, pp. 267–281.

[13] E. W. Steyerberg, et al "Internal validation of predictive models," *J. Clin. Epidemiol.*, vol. 54, no. 8, pp. 774–781, 2001.

[14] S. Raihana, et al "Development and internal validation of a predictive model including pulse oximetry for hospitalization of under-five children in Bangladesh," *PLoS One*, vol. 10, no. 11, 2015.