32nd Annual International Conference of the IEEE EMBS
Buenos Aires, Argentina, August 31 - September 4, 2010

Real-Time Sleep Quality Assessment Using Single-Lead ECG and Multi-Stage SVM Classifier

Majdi Bsoul, *Member*, *IEEE*, Hlaing Minn, *Senior Member*, *IEEE*, Mehrdad Nourani, *Senior Member*, *IEEE*, Gopal Gupta, *Senior Member*, *IEEE* and Lakshman Tamil, *Senior Member*, *IEEE*

***Abstract*—Sleep efficiency measures provide an objective assessment to gauge the quality of individual's sleep. In this study we present a home–based, automated and non–intrusive system that is based on Electrocardiogram (ECG) measurements and uses a multi–stage Support Vector Machines (SVM) classifier to measure three indices for sleep quality assessment per 30 s epoch segment: Sleep Efficiency Index, Delta–Sleep Efficiency Index and Sleep Onset Latency. This method provides an alternative to the intrusive and expensive Polysomnography (PSG) and scoring by Rechtschaffen and Kales visual method.**

I. Introduction

SLEEP, a physical and mental resting state, is primarily a restorative process that influences the homeostatic regulation of the autonomic, neuroendocrine, and immune systems. The efficiency of a person is directly proportional to the amount and quality of sleep that a person had in the prior night. The behavioral habits, sleep related breathing disorders such as apnea, drugs such as sleeping pills and alcoholic beverages can suppress certain stages of sleep leading to poor sleep quality or even *sleep deprivation* that have serious effects on individual's health and wellness and lead to various medical problems like cognitive impairment and heart diseases [1].

In this paper we present a home–based automated and a non–intrusive system, based on an *Electrocardiogram* (ECG) recording and a multi–stage *Support Vector Machines* (SVM) classifier [3]. For each 30 s interval, the system detects sleep state from the 4 sleep states: Wake or non–sleep (S_W), REM (S_{REM}), delta deep sleep (stages 3 and 4) known as slow wave sleep (S_{SWS}), and shallow sleep (S_{12}) (stages 1 and 2). Three quantitative measures are used to assess an individual's sleep quality [2]. Sleep Efficiency Index (S_{EI}) is the proportion of sleep in the period potentially filled by sleep or the ratio of the total sleep time to the time in bed. For a normal sleep, S_{EI} should at least be 85 % (of total bed–time). Delta–Sleep Efficiency Index (DS_{EI}) is the proportion of delta–sleep relative to total time in bed. The delta–sleep is the most restorative stage of sleep, which normally is about 20 % of the night. In this stage, the electroencephalographic delta waves are predominant (sleep stages 3/4). Sleep Onset Latency (S_{OL}) is the duration of time from 'lights out' or bedtime, to the onset of sleep, which is normally ~15 minutes.

Manuscript received June 24, 2010.

Majdi Bsoul is with the Alcatel–Lucent, Plano, TX 75075 USA and also with Quality of Life Technology Laboratory (QoLT), University of Texas at Dallas, TX 750 USA (e-mail: majdi@ieee.org).

Hlaing Minn is with QoLT and Dept. of Electrical Engineering, University of Texas at Dallas, TX 75080 USA (e-mail: hlaing.minn@utdallas.edu).

Mehrdad Nourani is with QoLT and Dept. of Electrical Engineering, University of Texas at Dallas, TX 75080 USA (e-mail: nourani@utdallas.edu)

Gopal Gupta is with QoLT and Dept. of Computer Science, University of Texas at Dallas, TX 75080 USA (e-mail: gupta@utdallas.edu)

Lakshman Tamil is with QoLT and Dept. of Electrical Engineering, University of Texas at Dallas, TX 75080 USA (e-mail: laxman@utdallas.edu).

II. Methodology

A. System Architecture

The increased processing power available in today's smartphones and its capability to connect locally through Bluetooth (IEEE 802.15.1) and to the internet through Wi-Fi (IEEE 802.11) or 3G make it an attractive platform to implement a simplified personal sleep apnea monitor. The architecture and the various system modules of *"Apnea MedAssist"* monitor are described in Fig. 1.

The system is implemented in a client/server configuration. The real–time monitoring, filtering, characteristic points detection and feature extraction are implemented on a smartphone (Android–based smartphone). The client communicates the feature vector ($\mathbf{x}_k$) comprising of both Heart Rate Variability (HRV) and surrogate ECG–Derived Respiratory (EDR) features (total of 112) to the server to perform sleep stage identification using a multi–stage SVM classifier (SVC) and to calculate sleep quality indices. The server can best handle the high computational cost of SVM–based classification in both memory requirement and cimputational counts.

B. Automated ECG Processing

Fig. 2 illustrates the processing components of the proposed automated system, which are described in the fol-

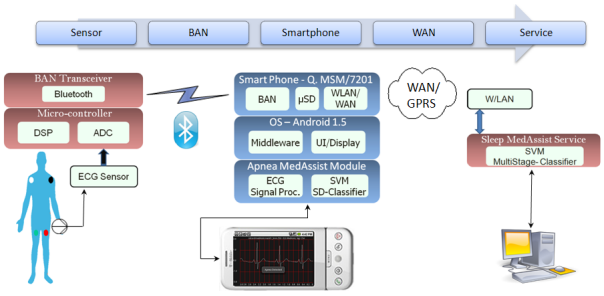

Fig. 1. System architecture of *Sleep MedAssist* showing components and functionalities.

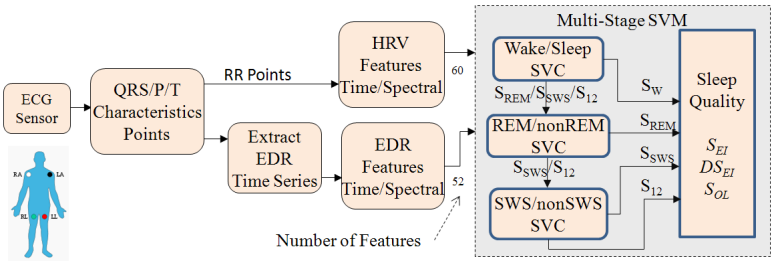

Fig. 2. Functional flowdiagram of automated system for sleep quality assessment using single–lead ECG measurements.

lowing sections. The ECG measurements with a sampling period of 4 ms are segmented into 30 s epochs and then analyzed using signal processing module which is a wavelet–based analysis stage for denoising, detrending and detection of ECG characteristic points: the QRS complex, P and T waves. The wavelet transform algorithm used here is based on the undecimated lifting scheme (ULWT) [6]. The ULWT has reduced computational cost compared to basic FIR implementation. Single–decomposition phase with 7–stages yields details $\{D^s\}_{s=1}^{7}$. To separate QRS–complex from PT–waves, we extract two signals by reconstructing two groups of ULWT sub–bands: $\{D^s\}_{s=2}^{5}$ are used for QRS–complex signal reconstruction and$\{D^s\}_{s=5}^{7}$ are used for P and T waves reconstruction. We performed adaptive thresholding to all details before each reconstruction stage to minimize spectral overlap between the QRS and the PT signals. A rules-based detection algorithm is used to detect the 5 characteristic points of QRS–complex (PQ, Q, R, S and J), and the 6 characteristic points of P and T waves (P_{start}, P_{peak}, P_{end}, T_{start}, T_{peak} and T_{end}). The derived RR–interval time–series is denoted as $\{\text{RR}(m) : rr_i,\ i = 1 \cdots m\}$, where rr_i is the *i*th RR interval.

We then used the T–wave characteristic points to extract the surrogate EDR signal [7]. The derived respiratory time series is denoted as $\{\text{EDR}(q) : edr_i,\ i = 1 \cdots q\}$, where q is the number of elements for the segment's EDR time series after preprocessing.

C. Feature Measures

Based on the extracted *RR*(*m*) and *EDR*(*q*) time series, we consider the following time–based features [8]:

- mean and standard deviation of RR–interval,
- NN50 and pNN50 measures: variant 1 defined as the number of pairs of adjacent RR–intervals where the first RR–interval exceeds the second RR–interval by more than 50 ms, and variant 2 where the second RR–interval exceeds the first RR–interval by more than 50 ms,
- the SDSD measure and its root-mean-square (RMS), SDSD defined as the standard deviation of the differences between adjacent RR–intervals,
- the first five serial correlation coefficients and the Allan factor A(T) evaluated at a time scale T of 5 s, 10 s and 15 s of the RR–intervals, and
- mean and standard deviation of the EDR amplitude.

In addition, we considered spectral–based features for the *autonomous system* (ANS) frequency ranges [4]: VLF, LF and HLF, based on both the variances of decimated wavelet transform (DWT) coefficients and the power spectral density estimates using the Fast Fourier Transform (FFT) for both *RR*(*m*) and *EDR*(*q*) time series.

Each 30 s ECG segment is now mapped to a full set of 112 extracted feature measures; 60 for RR time series and 52 for EDR time series.

D. Multi–Stage Support Vector Classifier

The sleep stage detection is a multi–class classification problem with four mutually exclusive sleep classes/states: S_W, S_{REM}, S_{SWS} and S_{12}. Applying SVMs to multiclass classification problems usually decomposes the multi–class problems into several two–class problems that can be addressed using several SVMs.

In this paper we use a binary decision tree (BDT) technique [9] for our $N = 4$ classes with a modified root node selection due to the order of sleep states classification. In the original BDT method, *N*-1 SVMs are needed to be trained for an *N*–classes problem, but at the most only $\lceil \log_2 N \rceil$ SVMs are required to classify a sample. The root node is selected to build a balanced tree which will reduce the number of levels required to arrive at a final decision on the tested sample class. Fig. 3 shows our three–stage multi–class SVMs for detecting the four classes S_W, S_{REM}, S_{12} and S_{SWS}. BDT–SVM takes advantage of both the efficient computation of the tree architecture and the high classification accuracy of SVMs, leading to an improvement in the speed of the test phase.

The Gaussian radial basis function (RBF) kernel was used because it was shown to perform similar to or better than

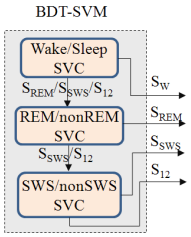

Fig. 3. Three–stage binary decision tree (BDT) SVM.

linear or polynomial kernels for appropriately selected values of kernel parameters (C,γ) [3]. We use here K–fold cross–validation (CV) method to evaluate the performance of the classifier after normalizing all features to be within [-1 +1].

Training was performed using the MIT–BIH polysomnographic database [5]. It is a collection of recordings for 16 subjects (Slp*xx*: *xx* is subject id) monitored in a sleep laboratory using ECG, EEG and respiration signals and annotated with respect to sleep stages and apnea for every 30 s epoch. The sleep state distribution over the total training 9450 epochs of 30 s were: $S_W \rightarrow 2773$, $S_{REM} \rightarrow 686$, $S_{12} \rightarrow 5280$ and $S_{SWS} \rightarrow 711$.

The SVM$_W$ was trained with un–balanced soft–margin penalties (C^+,C^-) biased towards (–1: "sleep–states") class to increase the specificity performance by an order of 1:5. This will decrease the false wake "non–sleep" state and increase true sleep state in SVM$_W$ classifier which enhances the performance of the next two SVM stages by reducing the outliers rippled into them. We used LIBSVM [10] for SVM training and classification.

E. Performance and Sleep Indicators

The three classification efficiency parameters: predictivity, specificity and accuracy, are calculated for all subjects in our sleep database for the classifier stages SVM$_W$, SVM$_{REM}$ and SVM$_{SWS}$. We also use a more robust measure of the reliability for our classifiers. Cohen's Kappa coefficient (κ) is used to assess the inter-rater agreement with the referenced sleep–studies experts and gives an indication if it was partially due to chance [11]. It is defined as

$$\kappa = (\Pr(a) - \Pr(e)) \,/\, (1 - \Pr(e)) \tag{1}$$

where Pr(*a*) is the relative observed agreement among classifiers and Pr(*e*) is the hypothetical probability of chance agreement using the observed data to calculate the probabilities of each observer randomly choosing a category. If the raters are in perfect agreement, then κ = 1. When there is no agreement among the raters (other than what would be expected by chance), then $\kappa \leq 0$.

Using just SVM$_W$ stage alone, we can calculate the Sleep efficiency index (S_{EI}) for each subject as

$$S_{EI} = 1 - \text{Tot}\,(S_W)\,/N_c \tag{2}$$

where Tot(S_W) is the total number of S_W classified epochs and N_c is the number of total number of classified samples. The Sleep Onset Latency (S_{OL}) is the duration of time of the total initial S_W states (in minutes) till first S_{nonW}. S_{OL} can be adversely affected by mis-classified epochs or outliers. We require 2–continuous epochs of S_{nonW} stages to declare a subject as being in that sleep state.

All three SVM stages are needed for (DS_{EI}) calculation

$$DS_{EI} = \text{Tot}\,(S_{SWS})\,/N_c \tag{3}$$

where Tot(S_{SWS}) is the total number of S_{SWS} classified epochs. The efficiency index errors are defined as $E_{S_{EI}} = S_{EI} - S_{EI}^{'}$ for the sleep efficiency and $E_{DS_{EI}} = DS_{EI} - DS_{EI}^{'}$ for the deep sleep efficiency, where $(\bullet)'$ is the actual efficiency index. We also assess the reliability of our SVM–based sleep indices using the classification efficiency rate E_{ff} defined as

$$E_{ff} = 1 - (\text{FP} + \text{FN})\,/\,(\text{TP} + \text{TN})\,. \tag{4}$$

III. Results

Table I shows the cross–validation performance results computed using our automated sleep quality assessment system for both S_{EI} and DS_{EI} on the sleep data from the MIT–BIH polysomnographic database. It lists the efficiency index predicted by the SVM classifier, the actual index calculated from sleep data annotations, the difference between predicted and actual indices, and the Cohen's kappa coefficient (κ). The table details the results for the various subjects and shows the performance variations exhibited by the calculated efficiency indices.

First examining the sleep efficiency (S_{EI}) determined by SVM$_W$ stage, we observe less variation and more consistency of the classifier performance for various subjects. This is also indicated by the kappa coefficient (κ) results. Very high κ–values ($0.6 < \kappa \leq 0.8$) are indicative of substantial agreement between the S_{EI} predictor and the actual measurement [11]. Note that for slp01 subject, we observe a high efficiency index error, but with a high κ result. This is mostly related to noisy ECG data measurements and not to the classifier performance.

Regarding the deep sleep efficiency (DS_{EI}) determined by the SVM$_{SWS}$ stage, we observe more variation and less consistency of the classifier performance for various subjects. Note that κ values of ($0.4 < \kappa \leq 0.6$) are indicative of moderate agreement between the DS_{EI} –predictor and actual measurement.

Table II shows the cross–validation performance (sensitivity, specificity and accuracy) results for the various subjects and the performance variations exhibited by each of the three–stage sleep state classifiers. The SVM$_W$'s degraded sensitivity (~78 %) increases false positives though it does not affect our overall system because it does not contribute to the predictivity of our efficiency indices.

IV. Conclusions

Our study presented an automated real–time system providing three sleep quality and efficiency measures: sleep efficiency, deep sleep efficiency and sleep onset latency. Sleep Efficiency Index (S_{EI}) determined by SVM$_W$ stage demonstrated high classification efficiency (E_{ff} ~87 %) and substantial agreement with R&K method (κ ~0.68) mostly due to the presence of no prior SVMs to inject outliers (misclassifications) into the input stream of this SVM. On the other hand, Delta–Sleep Efficiency Index (DS_{EI}) determined by SVM$_{SWS}$ stage demonstrated an E_{ff} ~78 % and

Table I
PERFORMANCE OF SLEEP EFFICIENCY (S_{EI}) AND DEEP SLEEP EFFICIENCY (DS_{EI}) COMPUTED USING OUR ALGORITHM ON MIT–BIH DATABASE

Record	S_{EI}					DS_{EI}				
	SVM_W (%)	Actual (%)	Error (%)	E_{ff} (%)	Kappa (κ)	SVM_{SWS} (%)	Actual (%)	Error (%)	E_{ff} (%)	Kappa (κ)
Slp01	68.78	50.25	-18.53	68.64	0.52	27.71	18.86	8.85	53.40	0.41
Slp02	79.46	76.06	-3.40	91.53	0.77	1.19	1.19	0.00	93.75	0.24
Slp03	81.23	81.66	0.43	88.68	0.66	32.95	11.17	21.78	52.32	0.34
Slp04	77.50	77.78	0.28	85.51	0.64	11.11	4.64	6.47	84.48	0.45
Slp14	55.13	54.85	-0.28	79.70	0.66	7.88	5.91	1.97	82.78	0.61
Slp16	54.70	63.10	8.39	81.48	0.68	2.60	3.47	-0.87	95.56	0.60
Slp32	40.80	37.46	-3.34	92.06	0.85	20.90	10.03	10.87	50.00	0.39
Slp37	87.96	90.30	2.34	95.45	0.78	0.33	0.00	0.33	99.60	0.00
Slp41	71.99	81.65	9.66	83.90	0.62	3.64	1.82	1.82	91.40	0.40
Slp45	99.03	96.94	-2.09	96.99	0.27	34.77	21.42	13.35	53.56	0.42
Slp48	70.33	68.52	-1.81	90.38	0.79	5.85	0.28	5.57	85.77	0.08
Slp59	69.43	77.07	7.64	86.35	0.70	24.02	17.47	6.55	80.58	0.66
Slp60	51.86	49.83	-2.03	86.59	0.76	4.56	0.00	4.56	84.21	0.00
Slp61	82.16	81.87	-0.29	89.86	0.69	23.88	14.82	9.06	66.67	0.47
Slp66	59.05	60.45	1.39	91.21	0.83	5.29	0.56	4.74	90.86	0.18
Mean	69.96	69.85	4.13	87.22	0.68	13.78	7.44	6.45	77.66	0.35

Table II
SLEEP ASSESSMENT PERFORMANCE {SVM_W, SVM_{REM} AND SVM_{SWS}}

Record	SVM_W			SVM_{REM}			SVM_{SWS}		
	sens (%)	spec (%)	accu (%)	sens (%)	spec (%)	accu (%)	sens (%)	spec (%)	accu (%)
Slp01	91.44	69.17	76.13	34.29	96.62	89.37	98.73	56.29	68.21
Slp02	89.26	92.95	92.19	97.14	28.86	44.87	100	94.06	94.12
Slp03	71.76	94.00	89.83	58.90	89.94	85.96	100	61.60	67.71
Slp04	71.25	92.01	87.34	86.96	72.45	73.06	100	85.56	86.56
Slp14	81.50	84.44	83.12	97.22	49.72	54.10	100	82.31	85.31
Slp16	73.48	93.39	84.37	98.44	62.90	68.12	100	95.59	95.74
Slp32	96.61	86.89	92.64	100	92.41	92.41	100	54.30	66.67
Slp37	72.22	98.86	95.65	100	91.85	91.85	100	99.60	99.60
Slp41	58.00	97.08	86.13	93.98	47.00	53.69	100	91.85	92.08
Slp45	57.14	97.47	97.08	78.90	62.59	65.14	100	53.21	68.29
Slp48	88.26	92.48	91.23	87.10	68.76	69.92	100	87.46	87.54
Slp59	67.86	96.86	87.99	82.86	75.47	76.20	100	77.27	83.74
Slp60	89.82	86.64	88.17	85.71	72.66	73.90	100	86.36	86.36
Slp61	75.00	94.22	90.79	88.14	71.57	73.29	100	68.90	75.00
Slp66	88.44	94.34	91.92	100	93.55	93.55	100	91.54	91.63
Mean	78.14	91.39	88.97	85.98	71.76	73.70	99.92	79.06	83.24

moderate agreement (κ ~0.35) because the prior two SVM stages can inject outliers to degrade its performance.

REFERENCES

[1] M. Vitiello, "Sleep in Normal Aging," *Sleep Medicine Clinics*, pp. 171–176, 2006.

[2] American Academy of Sleep Medicine, "The International Classification of Sleep Disorders: Diagnostic and Coding Manual," Revised 2001.

[3] C. Cortes and V. Vapnik, "Support-Vector Machines," *Machine Learning*, vol. 20, pp. 273–297, 1995.

[4] P. Busek, J. Vankova, J. Opavsky, J. Salinger and S. Nevsimalova, "Spectral Analysis of Heart Rate Variability in Sleep," *Institute of Physiology*, vol. 54, pp. 369–376, 2005.

[5] Y. Ichimaru, G. Moody, "Development of the Polysomnographic Database on CD-ROM," *Psychiatry and Clinical Neurosciences*, vol. 53, pp. 175–177, Apr. 1999.

[6] C. S. Lee, C. K. Lee and K. Y. Yoo, "New Lifting Based Structure for Undecimated Wavelet Transform," *IEEE Electronic Letters*, vol. 36, 2000.

[7] G. D. Furman, Z. Shinar, A. Baharav and S. Akselrod, "Electrocardiogram Derived Respiration during Sleep," *Computers in Cardiology*, vol. 32, pp. 351–354, 2005.

[8] M. C. Teich, S. B. Lowen, B. M. Jost, K. Vibe-Rheymer and C. Heneghan, "Heart rate variability: Measures and models, Nonlinear Biomedical Signal Processing," *IEEE Press*, vol. II, 2001.

[9] B. Fei and J. Liu, "Binary Tree of SVM: A New Fast Multiclass Training and Classification Algorithm," *IEEE Trans. on neural networks*, vol. 17, no. 3, May 2006.

[10] Chih-Chung Chang and Chih-Jen Lin, "LIBSVM: a library for support vector machines." [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm.

[11] J. Landis and G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, no. 1, pp. 159–174, 1977.