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

Decision tree for smart feature extraction from sleep HR in bipolar patients

Matteo Migliorini, Sara Mariani, and Anna M. Bianchi, *Member, IEEE*

***Abstract*— The aim of this work is the creation of a completely automatic method for the extraction of informative parameters from peripheral signals recorded through a sensorized T-shirt. The acquired data belong to patients affected from bipolar disorder, and consist of RR series, body movements and activity type. The extracted features, i.e. linear and non-linear HRV parameters in the time domain, HRV parameters in the frequency domain, and parameters indicative of the sleep quality, profile and fragmentation, are of interest for the automatic classification of the clinical mood state. The analysis of this dataset, which is to be performed online and automatically, must address the problems related to the clinical protocol, which also includes a segment of recording in which the patient is awake, and to the nature of the device, which can be sensitive to movements and misplacement. Thus, the decision tree implemented in this study performs the detection and isolation of the sleep period, the elimination of corrupted recording segments and the checking of the minimum requirements of the signals for every parameter to be calculated.**

I. Introduction

The PSYCHE system, developed in the frame of the ending PYSCHE project [1][2], is a multidisciplinary platform aiming at providing a personal, cost-effective, multi-parametric monitoring system based on textile platforms and portable sensing devices for long- and short-term data acquisition from patients affected by mood disorders where the patients represent the epicenter of a closed-loop management, for treatment and prevention of depressive and manic episodes. The PSYCHE platform (Fig. 1) is comprised of portable devices for the monitoring, among other parameters, of biomedical signals are processed and fused with medical analysis in order to verify the diagnosis and help in prognosis of the illness. A communication feedback to the patient and the physician is provided through a closed loop by means of a data mining approach for facilitating disease management with a greater autonomy and empowerment of patients. Constant feedback and monitoring are used to manage the illness, to give patients support, to facilitate the interaction between patient and physician as well as to alert professionals in case of patients relapse and depressive or manic episodes incoming. This platform aims at identifying disease trends for detection and prediction of critical events.

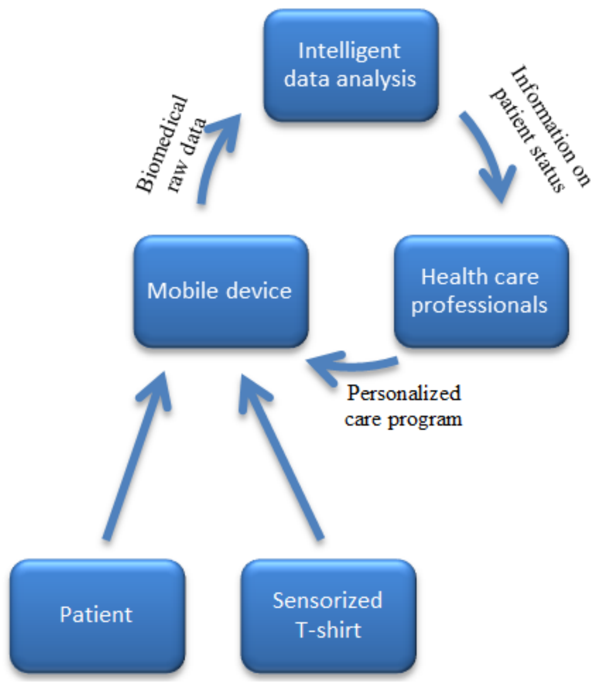

Figure 1: Closed-loop of the PSYCHE platform

Several studies have pointed out the relationships between the bipolar disorder and the Autonomic Nervous System, exploring in particular the mechanisms by which the heart rate can be affected by mood changes [3][4][5]. The low level of body movements and the low external stimuli influence during the sleeping periods, accounts for a privileged condition to record RR series with wearable devices. Despite this, the recordings during sleep are not artifact-proof: in fact, the patient is completely free to do whatever he/she wants, thus meaning that the t-shirt may be worn in the wrong way, misused or taken out, and the recording device can be unplugged. Also, the patient's movements may cause poor contact with the sensor and artifacts in the recorded signals.

The aim of this study is to develop a decision tree designed to allow a completely automatic data processing pipeline, in order to address the big issue represented by the patient freedom and to obtain reliable features for the assessment of the patient's state.

II. Materials and Methods

58 night recordings from 12 bipolar patients were analyzed in this study. The SMARTEX© system [1] records the ECG and the 3-axes acceleration and preprocesses these signals obtaining the body movements, the activity type (classified into resting and non-resting) and the RR series (Fig. 2). The

*This work was supported by Psyche under Project FP7-247777 of the European Community.

M. Migliorini, S. Mariani and A. M. Bianchi are with the Politecnico di Milano, Dept. Of Electronic, Information and Biomedical Engineering, P.zza Leonardo da Vinci 32, Milan, Italy (tel: +390223993322, email: matteo.migliorini@mail.polimi.it, sara1.mariani@mail.polimi.it, annamaria.bianchi@polimi.it).

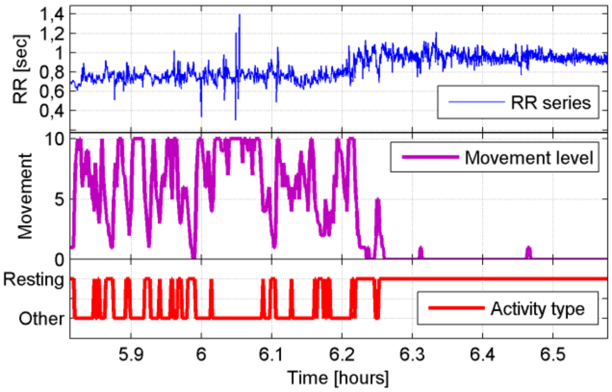

Figure 2: Portion of a recording.

recording protocol envisages the recording start at late afternoon, until the morning wake up.

The developed decision tree (Fig. 3) is composed by the following steps.

The first step consists of the detection of the resting periods, based on the activity and the movement signals (Fig. 2): a threshold is superimposed to the result of the application of a moving average filter with a 1200-second window. The periods over threshold that are longer than 3 hours are classified as resting periods, and will be further analyzed, the other segments are considered as wake periods. If no resting periods are detected, the analysis stops.

Once the resting periods are detected, the distance between them is calculated. Resting segments that are separated by wake segments shorter than 1 hour and 30 minutes are merged. The wake segments in between are considered as Wake After Sleep Onset.

The new segments are modified by means of an epoch-by-epoch (epochs are 30 seconds long) comparison of a filtered version of the movement signal (Fig. 2) with a threshold.

At this point, on the basis of the movement and the activity signals, some sleep parameters of clinical interest can be computed, such as: Starting Time (ST); Time In Bed (TIB); Sleep Onset Latency (SOL); Total Wake After Sleep Onset (WASO); Total Wake Time (TWT=SOL+WASO); Total Sleep Time (TST=TIB-TWT); Sleep Efficiency (SE=TST/TIB*100).

The next step is represented by the quality evaluation of the RR signal within each resting period. Three main problems inside the signals are evaluated: the presence of artifacts, recognized as high percentage of outliers in the segments, the missing of values in the RR series, or the presence of repeated data values. Resting segments, or portions of them, where the problems subsist are eliminated from the analysis. The outliers detection was performed according to the methodology proposed in [6].

At this stage, the clean RR series is modeled using a 9th order Time Variant Autoregressive Model (TVAM) with a Fortescue forgetting factor [7]. Starting from this model, frequency-domain and sleep parameters related to the REM stages can be extracted. The sleep parameters are extracted by

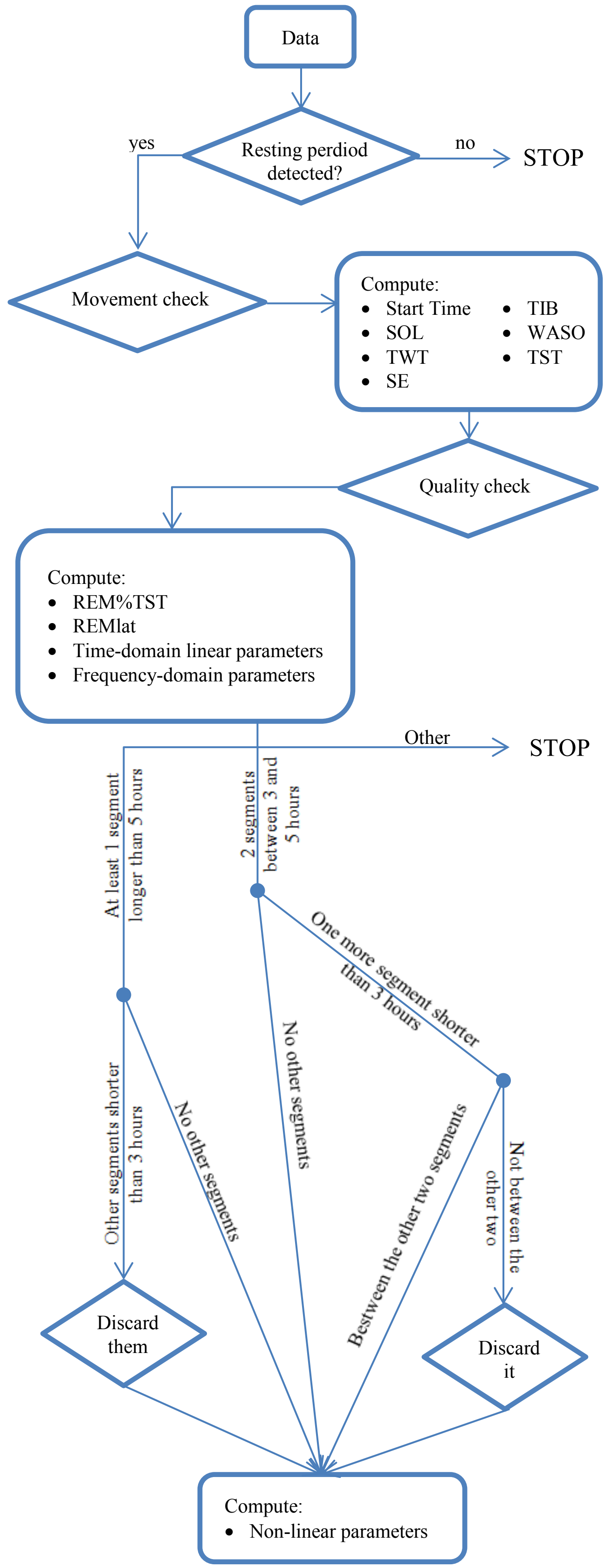

Figure 3: Decision tree overview.

TABLE I. LIST OF PARAMETERS

Group	Name	Description
Sleep parameters	ST	Starting Time
	TIB	Time In Bed
	SOL	Sleep Onset Latency
	WASO	Wake After Sleep Onset
	TWT	Total Wake Time
	TST	Total Sleep Time
	SE	Sleep Efficiency
	REMlat	Latency of the first REM stage
	REM%TST	Percentage of REM on sleep
Linear time-domain parameters	MEANN	Average value of NN intervals
	SDNN	Standard deviation of the NN intervals
	RMSSD	Square root of the mean of the sum of the squares of differences between subsequent NN intervals
Frequency-domain parameters	LF max and min	Maximum and minimum normalized LF power
	deltaLF	LFmax-LFmin
	HF max and min	Maximum and minimum normalized HF power
	deltaHF	HFmax-HFmin
	LF/HF max and min	Maximum and minimum value of the sympathovagal balance
	deltaLF/HF	LF/HFmax-LF/HFmin
Non-linear parameters	SampEn	Sample Entropy
	LZC	Lampel-Ziv Complexity
	DFA	Detrended Fluctuation Analysis scaling exponents
	1/f slope	Slope of the regression line defining the power-law distribution of the RR series

means of an automatic sleep staging, which was performed through a Feed-Forward Neural Network composed of 20 neurons in a single hidden layer and of 3 output neurons [3]. In addition, linear parameters in the time domain can be extracted [4].

In [3] and [4], also non-linear parameters were considered for the evaluation of the clinical state. These parameters describe the scaling and complexity properties of the signal, and are meant to be calculated on long-term recordings [8][9][10]; thus, in order to obtain informative features, in this decision tree non-linear parameters are computed only if there is at least a segment longer than 5 hours or there are two segments with duration between 3 and 5 hours. In this second case, if there is one more segment shorter than 3 hours that is in between of the other two, it is included in the analysis, otherwise it is not considered.

Tab. 1 reports the final list of conditionally computed features, grouped into sleep parameters, parameters in the frequency domain, linear parameters in the time domain, and nonlinear parameters in the time domain.

III. RESULTS

The presented decision tree has been tested on the entire dataset and its performances were visually checked for every recording.

Fig. 4 shows the results of the first step of the tree. The analysis of the activity type during the recording can be used to detect resting periods in a reliable way. The detected resting period in Fig. 4 is proved also by the body movement and by a higher RR mean value. The first detected resting period (after about 2 hours) is too short to be considered as sleep.

After the resting period detection the RR series quality check is performed. Fig. 5 shows the different steps of good quality segments selection. The reliability of the heart beats detection decreases in correspondence of high body movements, therefore those beats detected during high movements are discarded (Fig. 5b). The remaining beats are checked on the means of the presence of outliers, constant data values and missing data. The portions of the signals where these problems subsist are eliminated (Fig. 5c).

IV. DISCUSSION

The results point out the robustness of the decision tree in analyzing automatically data recorded during sleep time through a wearable device. This method permits the extraction

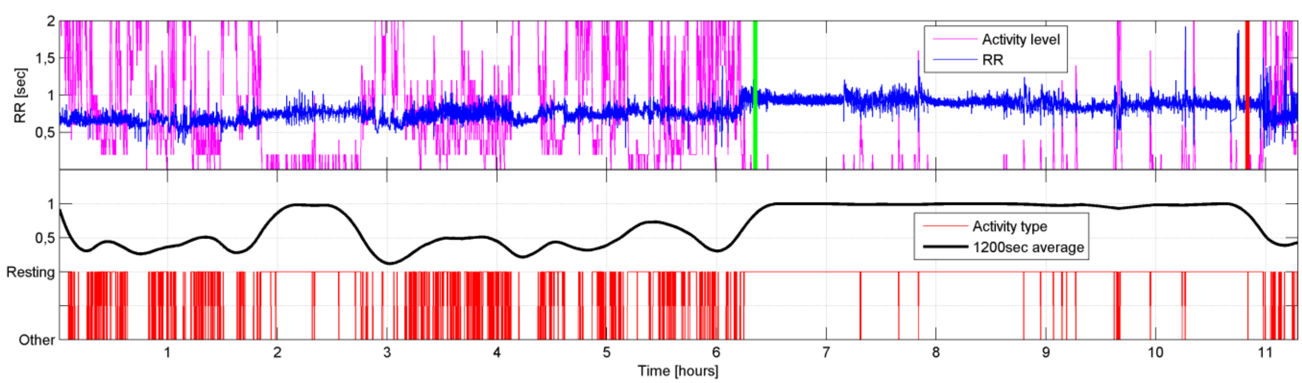

Figure 4: The top graph represents the RR series (blue) and the movement level (magenta) of a night recorded. The bottom one represents the activity type (red) and his filtered version (black). From the top graph the green and the red vertical lines represent the starting and the ending time of the detected resting period respectively.

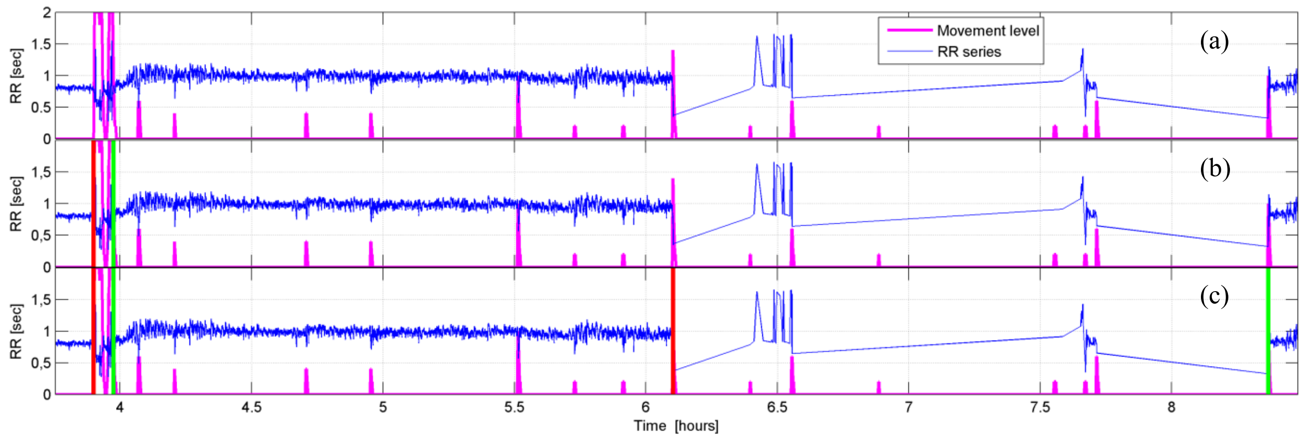

Figure 5: (a) shows the RR series and the movement level of a portion of a sleep period. The green and red vertical lines in (b) and (c) show the starting and the ending point of a selected segment after the movement adjustment (b) and after the RR quality check (c).

of parameters even if the signal quality is not elevated. Some parameters related to sleep can be calculated even if the RR signal is completely corrupted, in addiction, REM-related sleep parameters, time-domain linear parameters and frequency-domain parameters can be calculated even if the sleep period length is quite short.

Further activities will be the comparison of the presented parameters extracted using the decision tree and recorded through the sensorized t-shirt with the ones extracted from data recorded through standard sensors.

ACKNOWLEDGMENT

The present research was partially founded by the PSYCHE 24777 EU project. We would like to thank Dr. Sandra Werner and Dr. Geoffrey Viardot from Forenap and Dr. Claudio Gentili and Dr. Emiliano Ricciardi from University of Pisa.

REFERENCES

[1] R. Paradiso. T. Faetti and S. Werner, “Wearable monitoring systems for psychological and physiological state assessment in a naturalistic environment”, in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, 2011, pp. 2250-2253.

[2] R. Paradiso, A. M. Bianchi, K. Lau and E. P. Scilingo, “PSYCHE: Personalized monitoring systems for care in mental health”, in *Engineering in Medicine and Biology Society, EMBC, 2010 Annual International Conference of the IEEE*, 2010, pp. 3602-3605.

[3] S. Mariani, M. Migliorini, G. Tacchino, C. Gentili, G. Bertschy, S. Werner and A. M. Bianchi, “Clinical state assesment in bipolar patients by means of HRV features obtained with a sensorized T-shirt”, in *Engineering in Medicine and Biology Society, EMBC, 2012 Annual International Conference of the IEEE*, 2012, pp. 2240-2243.

[4] M. Migliorini, M. O. Mendez and A. M. Bianchi, “Study of heart rate variability in bipolar disorder: linear and non-linear parameters during sleep”, *Frontiers in Neuroengineering*, vol. 4, January 2012.

[5] G. Valenza, C. Gentili, A. Lanatà and E. P. Scilingo, “Mood recognition in bipolar patients through the PSYCHE platform: Preliminary evaluations and perspectives”, *Artificial Intelligence in Medicine*, Available online 16 January 2013.

[6] K. J. Kemper, C. Hamilton and M. Atkinson, “Heart rate variability: Impact of differences in outlier identification and management strategies on common measures in three clinical populations”, *Pediatric research* 62(3), pp. 337-342, September 2007.

[7] G. Tacchino, S. Mariani, M. Migliorini and A. M. Bianchi, “Optimization of Time-Variant Autoregressuve Models for tracking REM – non REM transitions during sleep”, in *Engineering in Medicine and Biology Society, EMBC, 2012 Annual International Conference of the IEEE*, 2012, pp. 2236-2239.

[8] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American Journal of Physiology.Heart and Circulatory Physiology*, vol. 278, pp. H2039, 2000.

[9] M. Aboy, R. hornero, D. Abasolo and D. Alvarez, "Interpretation of the Lempel-Ziv complexity measure in the context of biomedical signal analysis," *IEEE Transactions on Biomedical Engineering*, vol. 53, pp. 2282, 2006.

[10] S. Havlin, S.V. Buldyrev, A. L. Goldberger, R. N. Mantegna, S.M. Ossadnik, C. K. Peng, M. Simons and H. E. Stanley, "Fractals in biology and medicine," *Chaos, Solitons Fractals*, vol. 6, pp. 171, 1995.