2019 CMBEC42
Ottawa, ON
May 21-24, 2019
CMBES/SCGB

A Comparison of Two ECG Inter-beat Interval Measurement Methods for HRV-Based Mental Workload Prediction of Ambulant Users

Abhishek Tiwari[1], Isabela Albuquerque[1], Mark Parent[1], Jean-Francois Gagnon[2], Daniel Lafond[2], Sebastien Tremblay[3] and Tiago H. Falk[1]

[1] INRS-EMT, Université du Québec, Montréal, Québec, Canada
[2] Thales Research and Technology Canada, Québec, Québec, Canada
[3] Université Laval, Québec, Québec, Canada

Abstract— **Heart rate variability (HRV) has been studied in the context of human behavior analysis and many features have been extracted from the inter-beat interval (RR) time series and tested as correlates of constructs such as mental workload, stress and anxiety. Extraction of inter-beat interval series requires processing of the electrocardiogram (ECG) signal. This processing is critical for high quality RR series extraction and overall HRV measurement. Typically, the Pan-Tomkins peak detection algorithm is used. Recently, however, innovative modulation spectral based heart rate detection methods have been proposed. In this paper, we compare the performance of both algorithms and their effects on HRV measurement for mental workload assessment under physical activity. Experiments were conducted with 45 participants while they performed the NASA Revised Multi-Attribute Task Battery II (MATB-II) under different types and levels of physical activity. We show that modulation spectrum based methods perform better than conventional peak detection methods for mental workload prediction in lower levels of physical activity, particularly in the bike riding condition.**

Keywords— **mental workload, HRV, ECG, modulation spectrum, Pan-Tompkins**

I. Introduction

Wearable biomedical devices have burgeoned in recent years with a special focus on the heart rate modality, as heart rate variability (HRV) has shown to be useful in prognostics for cardiovascular diseases [1]. Moreover, HRV is also found to be an important correlate of several quality-of-life indices, such as psycho-social workload [2] (i.e., job stressors), mental workload and anxiety [3], as well as mental fatigue [4].

Heart rate variability is an indicator of the changes in the autonomic nervous system and has traditionally been quantified using time- and/or frequency-domain features computed from the inter-beat interval (RR) time series [5]. This inter-beat interval (RR) series is extracted from the peaks of the QRS complex of an electrocardiogram (ECG) signal. ECG measurement, however, can be corrupted by different artifacts, such as motion, power-line interference, and muscle artifacts, to name a few. Such artifacts cause errors in RR series detection, thus could cause problems for HRV calculation and, consequently, mental state evaluation [6].

Typically, to counter such detrimental effects, ECG enhancement is performed. Typical examples can include adaptive filtering [7], wavelet filtering [8] and empirical mode decomposition [9]. After artifact removal, peak detection algorithms, such as the ubiquitous Pan-Tompkins (PT) algorithm [10] are used. Once an RR series is extracted, an additional cleaning step is performed to remove outliers caused by misdetections.

Recently, the so called modulation spectrum (MS) [11] has been shown to be a relevant tool in studying bio-signals. The modulation spectrum can exploit second order periodicity in a signal which occurs due to amplitude modulation processes. This type of amplitude modulation in bio-signals is suspected to be a direct consequence of the processes of control, synchronization, regulation and inter-system interaction found in biological systems [12]. Modulation spectrum based filtering has been successfully applied for the enhancement of ECG [13] and to better detect heart rate values from noisy ECGs [14]. The effectiveness of these two RR extraction methods, however, has yet to be explored for practical applications. We aim to fill this gap.

In this paper, we analyze ECG data collected from 45 participants while they performed the NASA Revised Multi-Attribute Task Battery II (MATB-II) [15] to elicit mental workload. Tests were conducted while participants engaged in two different physical activities (bike or treadmill), each at three different physical workload levels (PWL: none, medium, high). Three ECG-processing pipelines were explored, namely: (i) classical Pan-Tompkins based peak processing, (ii) Pan-Tompkins followed by RR filtering to remove outliers, and (iii) modulation spectrum based processing. Standard HRV measures are then extracted from the RR series and explored as correlates of mental workload.

II. Materials and Methods

A. Participants

Data was collected from 45 participants (23 female, average age of 27.7 ± 6.6 years old). Twenty-one participants performed the experiment on a treadmill and twenty-four on a stationary bike. Workload was modulated by changing the difficulty of the task (easy or difficult, corresponding to low and high workload, respectively). Each difficulty levels was performed at rest (no movement), at medium activity (3km/h: treadmill, 50 rpm: bike) or at high activity (5 km/h, 70 rpm). This resulted in a total of 6 combinations of mental and physical workload. These 6 sessions lasted 10 minutes each. Ordering of the sessions was counterbalanced. The experiment was preceded by a 10 minute tutorial. Moreover, prior to each session, two baselines (no physical/no mental workload; just physical workload) of 1 and 2 minutes, respectively, were collected. At the end of each session the subject filled the NASA-Task Load Index (TLX) questionnaire [16] to evaluate different dimensions of workload during a 5 minute minute break period between sessions. ECG data were recorded using a wearable device (Bioharness BH3, Zephyr) at a sampling rate of 250 Hz.

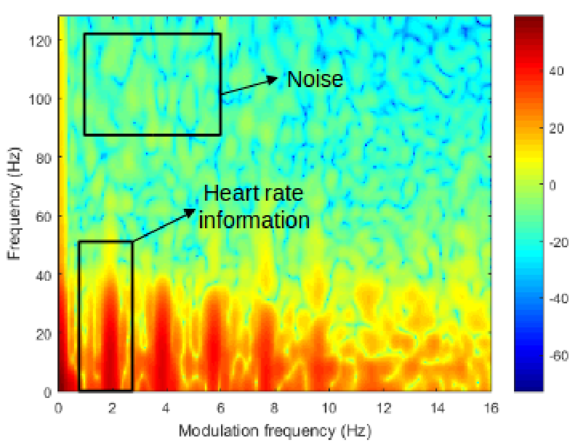

Figure 1: Modulation spectral representation of noisy ECG signal showing the lobes with heart rate information

B. ECG modulation spectral representation

The modulation spectrum (MS) corresponds to the Fourier transform of the different spectral elements in the spectrogram of a signal, thus quantifying the rate of change of the signal spectral components over time. The ECG signal is modulated by heart rate, hence this information is concentrated as lobes in the MS, centred at the heart rate, with lobes appearing in harmonics of the heart rate. This can be clearly seen in Fig 1. As heart rate is encoded in the modulation spectral lobes, artifacts appear in other regions, as depicted in the figure, thus suggesting improved robustness to artifacts. One limitation of the MS processing method is that it requires longer analysis windows (e.g., in the order of 5 seconds), thus the extracted RR series are more coarse grained relative to the instantaneous heart rate measures obtained with PT. This may have some detrimental effect on applications based on HRV measurement, which typically require finer-grained RR time series, especially those based on frequency-domain measures. As such, this work also explores if the gains obtained with improved RR series detection outweigh the disadvantages of HRV measurement from coarse-grained RR series.

C. Processing pipelines and feature extraction

The RR series was extracted using three different processing pipelines, as shown in Fig. 2. All of the pipelines make use of QRS enhancement bandpass filter followed by the different methods. RR detection was done for 5 minute segments of ECG signals with 4 minute overlap, resulting in six RR series for each of the 10-minute experimental sessions. The pipelines are:

- Pipeline 1 (PT): The classical Pan-Tompkins RR detection [10] [17] was applied on the ECG signal. This consists of QRS complex enhancement with different filters followed by thresholding for peak detection.
- Pipeline 2 (PT + RR filter): For this pipeline, the RR peaks detected using Pan-Tompkins algorithm were further filtered to remove outliers using range based detection ($\geq 280ms$ and $\leq 1500ms$), moving average outlier detector and a filter based on percent change in consecutive RR values ($\leq 20\%$) as implemented in [18].
- Pipeline 3 (MS): The modulation domain processing pipeline computes heart rate directly from the MS using the method detailed in [14] after modulation based enhancement [13]. The heart rate was then converted into RR interval values to keep the output features from the two pipeline comparable and allowing for HRV feature extraction.

Standard time- and frequency-domain HRV metrics were extracted and used as benchmark measures. A complete list of these conventional measures can be found in Table 1. The majority of these benchmark features have been shown in the literature to correlate with mental workload [19] and anxiety [3]. Complete details about these measures can be found in [5].

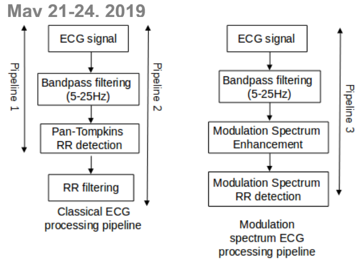

Figure 2: RR series extraction pipelines explored herein

Table 1: Different groups of HRV features extracted

Time domain HRV features
mean, standard deviation, coefficient of variation, rmsdd, pNN50, mean of 1^{st} diff., standard deviation of absolute of 1^{st} diff., normalized mean of absolute 1^{st} diff
Frequency domain HRV features
High frequency power (HF), normalized HF, Low frequency power (LF), normalized LF, very low frequency power, HF/LF

D. *Mental workload classification*

To measure the effects of the three pipelines on mental workload assessment, classification was performed 50 times, where five-fold cross validation was performed each time. This provides 250 unique test and training sets, thus helping guarantee the robustness of the prediction results. Binary classification (low/high workload) was performed on the ground truth labels of the MATB-II test. A support vector machine (SVM) classifier with an radial basis function (RBF) kernel is used. F1-score and accuracy (ACC) are used as classifier performance figures-of-merit. The open-source scikit-learn toolbox was used [20].

III. Experimental Results and Discussion

The mental workload prediction at no, medium and high physical workload levels is shown in Table 2 for the combined treadmill and bike riding experiments. These results are further broken down for just the bike (Table 3) and for the treadmill (Table 4).

As can be seen, for the combined bike and treadmill case, we observe that the MS pipeline outperforms the other two methods for no and medium physical activity cases ($p < 0.01$), particularly for the no-movement case. For the high physical workload case, none of the three pipelines resulted in accuracy above chance. The RR filtering method also reduced performance relative to using PT alone. This could be due to the change in HRV values caused by removing RR intervals, as mentioned in [6], as well as reducing any spurious connections due to noise in the data.

When looking at only the bike scenario, we observe that the overall performance for all the three physical activity cases was better than chance for the MS pipeline, though the accuracy dropped for higher levels of physical workload. For the PT + RR filtering pipeline, in turn, it performed better than chance for only the no and medium physical activity cases and was improved by RR filtering. We also observe that the MS pipeline did significantly better ($p < 0.01$) than the other two in case of no and high physical workload while achieving similar accuracy for medium PWL compared to the PT pipeline, though with a significantly higher F1-score ($p < 0.01$).

Lastly, when looking only at the treadmill scenario, we observe that with no physical workload, the PT+ RR filter pipeline performs better than chance with the MS pipeline outperforming both ($p < 0.01$). This behaviour is reversed for the medium PWL, and the PT pipeline outperforms the other two ($p < 0.01$). Reducing the resolution of the RR series via RR smoothing or MS processing may cause detrimental effects leading to such behavior, thus removing high-frequency content that has been shown useful for mental workload assessment. On the other hand, for high physical workload, we see that RR smoothing helps to remove artifacts that are more pronounced. The MS pipeline, in turn, performed near chance levels. This could be caused by the increased heart rate during high physical activity, thus high-frequency dynamics are not properly captured by the coarse grained heart rate extracted using the modulation spectrum. As such, the noise robustness properties of the RR extraction method may be overshadowed by this limited temporal resolution of the RR series, which has been shown essential for mental workload measurement.

Table 2: Predictions for different physical workload for combined bike and treadmill (* shows cases with $p < 0.01$ compared to chance)

PWL	Pipeline	Acc	F1
No	PT	0.5324*	0.4899
	PT+RR filter	0.5605*	0.4367
	MS	0.6371*	0.5986*
Medium	PT	0.5591*	0.5738*
	PT+RR filter	0.5031	0.5393*
	MS Processing	0.5788*	0.635*
High	PT	0.5090	0.6205*
	PT+RR filter	0.5090	0.5507*
	MS	0.4954	0.4845

Table 3: Predictions for different physical workload for bike (* shows cases with $p < 0.01$ compared to chance)

PWL	Pipeline	Acc	F1
No	PT	0.5948*	0.5018
	PT+RR filter	0.6445*	0.5399*
	MS	0.6737*	0.6273*
Medium	PT	0.5499*	0.5363*
	PT+RR filter	0.5715*	0.5947*
	MS	0.5666*	0.6477*
High	PT	0.4812	0.5808*
	PT+RR filter	0.5049	0.5828*
	MS	0.5525*	0.5982*

Table 4: Predictions for different physical workload for treadmill (* shows cases with $p < 0.01$ compared to chance)

PWL	Pipeline	Acc	F1
No	PT	0.5022	0.4767
	PT+RR filter	0.5947*	0.5421*
	MS	0.6146*	0.5961*
Medium	PT	0.5959*	0.6362*
	PT+RR filter	0.4891	0.5207
	MS	0.5281*	0.5199
High	PT	0.5667*	0.6559*
	PT+RR filter	0.5867*	0.5779
	MS	0.5180	0.4232

IV. CONCLUSION

In this paper, we compare three ECG inter-beat interval extraction pipelines to explore their advantages for HRV measurement for mental workload assessment of ambulant users. We show that a new modulation spectrum based pipeline performs better than classical pipelines for mental workload prediction in low to medium physical activity levels, thus showing its robustness to mild levels of movement artifacts and activity levels. In high physical activity levels, on the other hand, in particular for treadmill running in which movement artifacts are numerous, traditional methods which extract instantaneous heart beat information showed improved performance, thus suggesting that the noise robustness capability of the modulation domain based method may be overshadowed by the low-grained temporal resolution it provides.

ACKNOWLEDGEMENT

The authors acknowledge funding from NSERC via its CRD program and Thales Canada.

REFERENCES

1. Melillo Paolo, others . Automatic Prediction of Cardiovascular and Cerebrovascular Events Using Heart Rate Variability Analysis *PloS one.* 2015;10:e0118504.
2. Togo Fumiharu, Takahashi Masaya. Heart Rate Variability in Occupational Health-A Systematic Review *Industrial health.* 2009;47:589–602.
3. Wen Wan-Hui, others . Toward Constructing a Real-time Social Anxiety Evaluation System: Exploring Effective Heart Rate Features *IEEE Transactions on Affective Computing.* 2018.
4. Patel Miteshkumar, Lal Sara KL, Kavanagh Diarmuid, Rossiter Peter. Applying Neural Network Analysis on Heart Rate Variability Data to Assess Driver Fatigue *Expert systems with Applications.* 2011;38:7235–7242.
5. Camm AJMM, others . Heart Rate Variability: Standards of Measurement, Physiological Interpretation and Clinical Use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology *Circulation.* 1996;93:1043–1065.
6. Peltola Mirja. Role of Editing of RR Intervals in the Analysis of Heart Rate Variability *Frontiers in physiology.* 2012;3:148.
7. Rahman Muhammad, Shaik Rafi, Reddy D. Efficient Sign Based Normalized Adaptive Filtering Techniques for Cancelation of Artifacts in ECG Signals: Application to Wireless Biotelemetry *Signal Processing.* 2011;91:225–239.
8. Singh Brij N, Tiwari Arvind K. Optimal Selection of Wavelet Basis Function Applied to ECG signal Denoising *Digital signal processing.* 2006;16:275–287.
9. Kabir Md Ashfanoor, Shahnaz Celia. Denoising of ECG Signals Based on Noise Reduction Algorithms in EMD and Wavelet domains *Biomedical Signal Processing and Control.* 2012;7:481–489.
10. Pan Jiapu, Tompkins Willis J. A Real-time QRS Detection Algorithm *IEEE Trans. Biomed. Eng.* 1985;32:230–236.
11. Cassani Raymundo, Falk Tiago H. Spectrotemporal Modeling of Biomedical Signals: Theoretical Foundation and Applications 2014.
12. Fedotchev AI, others . Concerning the Amplitude Modulation of the Human EEG *Human Physiology.* 2000;26:393–399.
13. Tobón Diana P, Falk Tiago H. Adaptive Spectro-temporal Filtering for Electrocardiogram Signal Enhancement *IEEE J of Biomedical and Health Informatics.* 2018;22:421–428.
14. Tobón Diana P, Jayaraman Srinivasan, Falk Tiago H. Spectro-Temporal Electrocardiogram Analysis for Noise-Robust Heart Rate and Heart Rate Variability Measurement *IEEE J Translational Eng. in Health and Medicine.* 2017;5:1–11.
15. Santiago-Espada Yamira, Myer Robert, Latorella Kara, Comstock Jr James. The Multi-attribute Task Battery ii (MATB-II) Software for Human Performance and Workload Research: A User's Guide 2011.
16. Hart Sandra G, Staveland Lowell E. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research in *Advances in psychology*;52:139–183Elsevier 1988.
17. Sedghamiz Hooman. Matlab Implementation of Pan Tompkins ECG QRS Detector *Code available at the File Exchange site of MathWorks. URL https://fr. mathworks. com/matlabcentral/fileexchange/45840-complete-pan-tompkins-implementationecg-qrs-detector.* 2014.
18. Behar Joachim A, Rosenberg Aviv A, Weiser-Bitoun Ido, et al. PhysioZoo: A Novel Open Access Platform for Heart Rate Variability Analysis of Mammalian Electrocardiographic Data *Frontiers in physiology.* 2018;9:1390.
19. Castaldo Rossana, Melillo Paolo, Bracale Umberto, Caserta M, Triassi Maria, Pecchia Leandro. Acute Mental Stress Assessment via Short Term HRV Analysis in Healthy Adults: A Systematic Review with Meta-Analysis *Biomedical Signal Processing and Control.* 2015;18:370–377.
20. Pedregosa Fabian, others . Scikit-learn: Machine Learning in Python *Journal of machine learning research.* 2011;12:2825–2830.