A Robust Estimation of the Cardiorespiratory Coupling in the Presence of Abnormal Beats

John Morales[1], Pascal Borzée[2], Dries Testelmans[2], Bertien Buyse[2], Sabine Van Huffel[1], Raquel Bailón[3], Carolina Varon[1,4]

[1] Department of Electrical Engineering-ESAT, STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, KU Leuven, Leuven, Belgium
[2] Department of Pneumology, UZ Leuven, Leuven, Belgium
[3] BSICoS Group, Aragón Institute of Engineering Research (I3A), IISAragon, University of Zaragoza and CIBER-BBN, Zaragoza, Spain
[4] Circuits and Systems (CAS) group, Delft University of Technology, the Netherlands

Abstract

Respiratory sinus arrhythmia (RSA) is one of the forms of cardiorespiratory coupling. It has been suggested as a potential biomarker for diverse illnesses and conditions. In general, methods for non-invasive quantification of the RSA combine information from heart rate variability (HRV) and respiratory signals. Abnormal beats, which commonly occur in different populations, alter the reliability of the HRV and thus hinder the quantification of the RSA. To overcome this problem, several methods for detection and correction of irregular beats have been reported in literature. However, the effect of each of these methods on the quantification of the RSA is not well understood yet. For this reason, an approach that avoids this step might be useful. This paper presents an alternative based on robust regression models. For comparison purposes, an algorithm to detect and correct for irregular beats, in combination with a state-of-the-art RSA estimate, are tested. A similar performance is achieved with both approaches. These results show that the proposed robust methodology is able to capture the strength of the RSA, even when irregular beats are present, avoiding the irregularities correction step.

1. Introduction

The development of methods for respiratory sinus arrhythmia (RSA) quantification is an active research topic. Reasons for this include the fact that this form of cardiorespiratory interaction has been shown to be a good biomarker for the diagnosis and follow-up of certain conditions and diseases.

The quantification of the RSA is based on the amount of information shared between the heart rate variability (HRV) signal, derived from the ECG, and the respiratory signal. In general, the methods for RSA quantification work well when ECG signals without abnormalities are used. However, their application and interpretation in signals recorded during cardiac abnormalities is still challenging. One of the reasons is that, in these cases, ectopic beats are often present in the HRV signal, which lead to an unreliable estimation of the RSA. This paper proposes and testes a new approach to characterize the RSA in the presence of irregular beats. The method is based on robust regression models built with least squares support vector machines (LS-SVM). Furthermore, the robustness of a state-of-the-art RSA estimator is tested with and without a phase of irregular heartbeat correction.

2. Methods

2.1. Dataset and preprocessing

The data used in this paper was taken from 110 Polysomnography (PSG) recordings of patients with different obstructive sleep apnea (OSA) severities and OSA associated comorbidities. The ECG and thoracic respiratory inductive plethysmograph signals were acquired with a sampling frequency of 500 Hz. The sleep apneas were annotated according to the AASM 2012 scoring rules [1]. The R-peaks in the ECG were detected using the approach described in [2]. Afterwards, these were used to calculate the RR intervals, which were then interpolated to a sampling frequency of 2 Hz, and used as the HRV signal. The respiratory signals were downsampled to 2 Hz after applying an antialiasing filter. Afterwards, both HRV and respiration were filtered to preserve only frequency components

between 0.03 and 1 Hz with a 4th order butterworth filter. This filter was applied in forward and backward directions to have a zero-phase distortion. Next, The respiratory signals and HRV were segmented into 5-minutes epochs and those containing apneas were discarded.

2.2. RSA estimates

The RSA was quantified using the robust and the non-robust approaches described below.

2.2.1. Non-robust RSA estimation

The non robust RSA estimate approach is a state-of-the-art method based on orthogonal subspace projections [3]. Let's denote $\boldsymbol{x}$ and $\boldsymbol{y}$ two vectors containing the samples of the HRV and respiration, respectively. This approach decomposes $\boldsymbol{y}$ into two components: one linearly related to $\boldsymbol{x}$ and another with residual information. To this end, a subspace $\boldsymbol{Q}$ is built as a time-delay embedding of $\boldsymbol{x}$. $\boldsymbol{Q}$ is then used to calculate a projection matrix $\boldsymbol{P}$, given by,

$$\boldsymbol{P} = \boldsymbol{Q}(\boldsymbol{Q}^T\boldsymbol{Q})^{-1}\boldsymbol{Q}^T, \tag{1}$$

which in turn is used to derive the information in the HRV, linearly correlated with the respiration as,

$$\boldsymbol{y}_r = \boldsymbol{P}\boldsymbol{y}. \tag{2}$$

With this, the relative power of the linear respiratory influences on the HRV, is calculated as,

$$\mathcal{P}_x = ({\boldsymbol{y}_r}^T\boldsymbol{y}_r)/(\boldsymbol{y}^T\boldsymbol{y}). \tag{3}$$

2.2.2. Robust RSA estimation

An alternative robust approach is proposed. It is based on measuring the quality of a robust autoregressive model built with LS-SVM [4]. To this end, a training data set of $M = N - L + 1$ points, $\{\hat{\boldsymbol{x}}_k; y_k\}_{k=1}^{M}$ is defined. Here, the input data $\hat{\boldsymbol{x}}_k \in \mathbb{R}^{\mathrm{L}}$ are vectors of L consecutive samples of the respiratory signal built as an embedding, where L is the embedding dimension. The output data point $y_k \in \mathbb{R}$ corresponds to the kth sample of the HRV signal. These are used to formulate the following optimization problem,

$$\begin{aligned} &\min_{w^*,b^*,e^*} J_P(\boldsymbol{w}^*, \boldsymbol{e}^*) = \tfrac{1}{2}\boldsymbol{w}^{*T}\boldsymbol{w}^* + \gamma\frac{1}{2}\sum_{k=1}^{M} v_k e_k^{*2}, \\ &\text{s.t. } y_k = \boldsymbol{w}^{*T}\varphi(\hat{\boldsymbol{x}}_k) + b^* + e_k^*,\ k = 1, ..., M. \end{aligned} \tag{4}$$

Here, $\varphi(\hat{\boldsymbol{x}}_k)$ is a function that maps the vector $\hat{\boldsymbol{x}}_k$ onto a higher dimensional space, $\boldsymbol{w}^*$ is a vector of weights, e_k^* is an error variable, b^* is a bias term, and γ is a term determining the importance of the error. The weights v_k are a variable to correct for outliers, given by,

$$v_k = \begin{cases} 1 & \text{if } \mid e_k/\hat{s} \mid \leq c_1 \\ \dfrac{c_2 - \mid e_k/\hat{s} \mid}{c_2 - c_1} & \text{if } c_1 < \mid e_k/\hat{s} \mid \leq c_2 \\ 10^{-4} & \text{otherwise} \end{cases}, \tag{5}$$

where the error terms, e_k, are estimated by first solving the non-robust LS-SVM problem, i.e not considering the weighting factor, v_k. $IQR(\boldsymbol{e})$ stands for the interquartile range of the errors, and $\hat{s} = IQR(\boldsymbol{e})/(2 \times 0{,}6745)$ is a robust estimation of the standard deviation of the errors. The solution to this problem becomes,

$$\left[\begin{array}{c|c} 0 & \mathbf{1}_v^T \\ \hline \mathbf{1}_v & \boldsymbol{\Omega} + \boldsymbol{V}_\gamma \end{array}\right] \left[\begin{array}{c} b^* \\ \hline \boldsymbol{\alpha}^* \end{array}\right] = \left[\begin{array}{c} 0 \\ \hline \boldsymbol{y} \end{array}\right], \tag{6}$$

where $\boldsymbol{\alpha}^*$ are the Lagrange multipliers, $\boldsymbol{\Omega}$ is the Kernel matrix, and the diagonal matrix $\boldsymbol{V}_\gamma$ is defined as,

$$\boldsymbol{V}_\gamma = diag\left(\left[\frac{1}{\gamma v_1}; ...; \frac{1}{\gamma v_M}\right]\right). \tag{7}$$

The system in 6 is solved for $\boldsymbol{\alpha}^*$ and b^*. The HRV, y, can then be predicted as,

$$y(\boldsymbol{x}) = \sum_{k=1}^{M} \alpha_k^* K(\boldsymbol{x}, \hat{\boldsymbol{x}}_k) + b^*. \tag{8}$$

More details of the methodology to build the regression model can be found in [4].
The residuals of the prediction, $\boldsymbol{e}^*$, are calculated. Finally, a parameter called η is derived, using the samples in $\boldsymbol{y}$ and $\boldsymbol{e}^*$ as,

$$\eta = 1 - IQR(\boldsymbol{e}^*)/IQR(\boldsymbol{y}). \tag{9}$$

2.3. Segment extraction

The cardiorespiratory coupling was estimated using $\mathcal{P}_x$ in the segments described in section 2.1. The epochs were then grouped by their $\mathcal{P}_x$ level in 9 bins of 0.1, ranging from 0 to 0.9. Finally, 50 randomly selected epochs per bin were visually chosen ensuring that they did not contain artifacts, irregular beats, or respiratory signals with an irregular pattern.

2.4. Simulation of irregular beats

The R-peak locations in the 5-minutes epochs were artificially contaminated with irregularities. As explained in [5], the occurrence of ectopic beats can be simulated changing the length of the original RR intervals according to:

$$RR'_n = \beta RR_{n-1}, \tag{10}$$

$$RR'_{n+1} = RR_{n+1} - RR'_n. \quad (11)$$

Here, β is a random value in the interval [0.3 0.8]. RR' are the simulated irregularities in the RR intervals. The available segments were contaminated 25 times with an increasing number of simulated ectopic beats each time. An example of an HRV before and after contamination is shown in Figure 1.

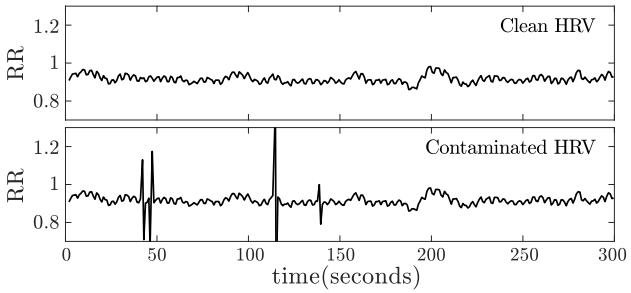

Figure 1. HRV contaminated with synthetic ectopic beats

2.5. HRV representation and irregular beats correction

After introducing the synthetic irregularities, HRV representations were built in three scenarios. First, a spline interpolation was applied to the RR intervals before contamination, to generate a clean HRV with an even sampling frequency (RR_c). The second scenario consisted of the application of a spline interpolation to the contaminated RR interval time series (RR_e). In the third one, the simulated ectopic beats were detected and corrected using the integral pulse frequency modulation model (IPFM) and then used to estimate the heart timing signal (m), which was used as an HRV representation [6]. These three HRV representations were sampled at 2 Hz.

2.6. Testing the RSA estimates

In order to perform the comparison of the RSA estimates, the values of $\mathcal{P}_x$ and η, calculated using RR_c, were first derived and used as the reference for the coupling level of the HRV and respiration. Afterwards, the two parameters were also calculated with RR_c and m. Differences between the RSA estimated with each HRV representation, for each $\mathcal{P}_x$ level and with the non-robust and robust methodologies, were evaluated using the Kruskal-Wallis tests, with a significance level of 5 % ($p \leq 0{,}05$) and with Bonferroni correction.

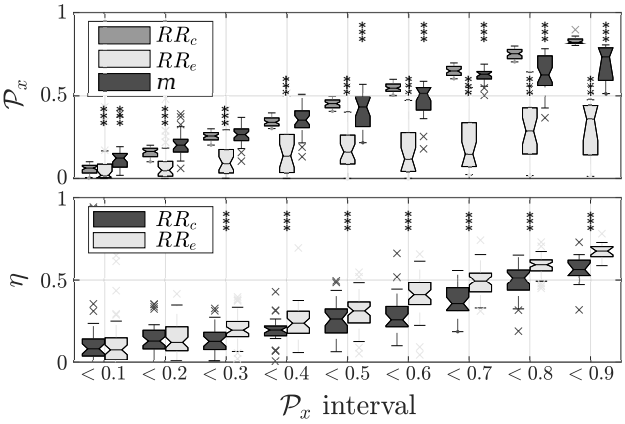

Figure 2. RSA estimates allowing up to 8 irregularities

3. Results and discussion

Figure 2 shows the results of the calculation of the cardiorespiratory coupling with the robust and the non-robust approaches in the 5-minutes epochs when up to 8 irregularities are included. Significant differences are observed between the contaminated/corrected case and the clean case. The top plot shows the results with the non-robust estimator (i.e., $\mathcal{P}_x$) using the three HRV representations. It is observed that $\mathcal{P}_x$ is significantly affected when RR_e is used, up to the point that the increasing trends with a stronger coupling are lost. This occurs because the calculation of this parameter is very sensitive to outliers, which in this case correspond to irregular beats. In addition, the figure shows that a method that takes into account the irregularities to build the HRV, such as the IPFM model, has a positive effect in the quantification of the RSA. Despite that the differences with the clean case are significant, the trends towards a higher $\mathcal{P}_x$ value with a stronger coupling are recovered after correction. The bottom plot shows the proposed robust parameter, η, using either RR_c or RR_e. Despite that the quantifications are significantly different using either HRV representation, it is observed that the trends towards higher η values with a stronger coupling are preserved when irregularities are allowed.

Figure 3 shows the results when a different number of irregularities are included. The Figure depicts the results for two coupling levels, i.e. $0 \leq \mathcal{P}_x \leq 0{,}1$ and $0{,}8 \leq \mathcal{P}_x \leq 0{,}9$. It is observed that all the RSA quantifications are more sensitive to outliers when the RSA is stronger. In this case, the estimates are trustworthy up to 10 irregular beats. With more than 10 irregularities, the computations of η and $\mathcal{P}_x$ with either m or RR_e are very different to the clean case. Similarly to the previous figure, the use of the HRV representation generated with the IPFM model had the same beneficial effect on the RSA estimation than using the

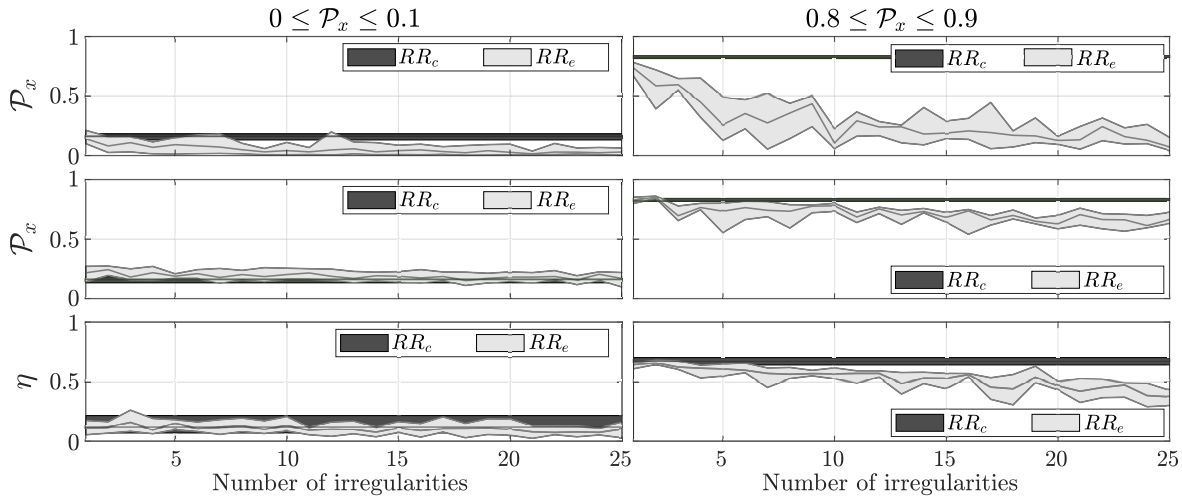

Figure 3. Estimates in function of the number of irregularities. Weak (left) and strong (right) RSA, quantified using $\mathcal{P}_x$.

robust approach. Here, it is important to mention that the computational cost of calculating η is significantly higher because a regression model needs to be built. To apply the correction of irregularities and calculate $\mathcal{P}_x$, only matrix multiplications and operations with vectors are required.

4. Conclusions

A robust approach for RSA estimation was tested when irregularities are simulated in the HRV representations. It was shown that this method is able to capture the cardio-respiratory coupling when irregular beats occur. It was also shown that the use of a non-robust RSA estimate, in combination with a technique for irregular beats corrections, achieved a similar performance. These results suggest that the proposed robust method can be used avoiding a step of irregular beats correction, which is not always straight forward. As future work, a surrogate analysis will be included. In addition, the proposed approach will be tested in real data with real irregular beats and will be compared with the computations using different strategies to correct for irregularities in the HRV.

Acknowledgments

Funding: BOF, C24/18/097. FWO. VLAI, 150466: OSA+, O&O HBC 2016 0184 eWatch. IMEC funds 2020. Flemish Government (AI Research Program). SVH, CV, JM are affiliated to Leuven.AI - KU Leuven institute for AI, B-3000, Leuven, Belgium. KU Leuven Stadius acknowledges the financial support of imec. CIBER, Gobierno de Aragón, projects RTI2018-097723-B-100, LMP44-18, T39-20R.

References

[1] Berry RB, Budhiraja R, Gottlieb DJ, Gozal D, Iber C, Kapur VK, Marcus CL, Mehra R, Parthasarathy S, Quan SF, et al. Rules for scoring respiratory events in sleep: update of the 2007 AASM manual for the scoring of sleep and associated events. Journal of Clinical Sleep Medicine ;8(05).

[2] Varon C, Caicedo A, Testelmans D, Buyse B, Van Huffel S. A novel algorithm for the automatic detection of sleep apnea from single-lead ECG. IEEE Transactions on Biomedical Engineering 2015;62(9):2269–2278.

[3] Varon C, Lázaro J, Bolea J, Hernando A, Aguiló J, Gil E, Van Huffel S, Bailón R. Unconstrained estimation of HRV indices after removing respiratory influences from heart rate. IEEE Journal of Biomedical and Health Informatics 2019;.

[4] Suykens JA, De Brabanter J, Lukas L, Vandewalle J. Weighted least squares support vector machines: robustness and sparse approximation. Neurocomputing 2002;48(1-4):85–105.

[5] Clifford GD, Tarassenko L. Quantifying errors in spectral estimates of HRV due to beat replacement and resampling. IEEE Transactions on Biomedical Engineering 2005; 52(4):630–638.

[6] Mateo J, Laguna P. Analysis of heart rate variability in the presence of ectopic beats using the heart timing signal. IEEE Transactions on Biomedical Engineering 2003;50(3):334–343.

Address for correspondence:

John Morales
Kasteelpark Arenberg 10 - box 2446, 3001 Leuven
jmorales@esat.kuleuven.be