R-Peak Estimation using Multimodal Lead Switching

Alistair EW Johnson[1], Joachim Behar[1], Fernando Andreotti[1,2], Gari D Clifford[1,3], Julien Oster[1]

[1] Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, UK
[2] Institute of Biomedical Engineering, Faculty of Electrical and Computer Engineering, TU Dresden, Helmholtzstr. 10, 01069 Dresden, Germany
[3] Departments of Biomedical Informatics & Biomedical Engineering, Emory University & Georgia Institute of Technology, Atlanta, GA, USA

Abstract

Introduction: Intensive care unit patients are heavily monitored, and a number of clinically relevant parameters are routinely extracted from high resolution signals. In particular, heart rate is derived from intervals between pulses in pseudo-periodic signals such as the electrocardiogram (ECG) or arterial blood pressure (ABP) waveforms. However, poor signal quality and high noise levels can unfortunately lead to false localisation of these pulses (or peaks), resulting in incorrect estimates of heart rate. The goal of the 2014 Physionet/CinC Challenge was to encourage the creation of an intelligent system that fused information from different biosignals to create a robust set of peak detections.

Methods: First, a set of peak detectors were evaluated on different cardiovascular signals. The detections were then fused using two different approaches: the first one was based on a calculated measures of signal quality for the ECG and ABP signals and the the second fusion technique was based on the regularity of the derived intervals between subsequent detections made on ECG, ABP, Stroke Volume and Photoplethysmogram signals. These techniques were developed using the MGH/MF database and submitted for scoring on the Challenge test-set.

Conclusion: The best entries for the two approaches obtained an overall score of 87.88% *and* 87.66%*, respectively, in phase III of the challenge, which provided the highest official score.*

1. Introduction

Millions of patients are admitted to Intensive Care Units (ICUs) in the United States every year. These patients require a high level of acute care, with numerous bedside monitors which are continuously monitoring both invasive and non-invasive variables. These monitors provide synchronous waveforms with both independent and complementary information. Huge ICU databases are therefore becoming available, and include parameters such as the electrocardiogram (ECG), the photoplethysmogram (PPG), the arterial blood pressure (ABP) waveform and various surrogates of respiratory effort. In clinical practice these signals are processed individually and derived parameters are frequently set to trigger an alarm when the parameter of interest exceeds a pre-defined range. These alarms are frequently false alarms (FAs) and account for a large majority of all alarms generated in the ICU [1].

It has been demonstrated that switching between signals independently derived from the patient can reduce the rate of FAs, e.g. by checking the heart rate from the ABP waveform when an ECG derived heart rate alarm was triggered [2, 3].

Nevertheless, little has been done fusing pulse estimates from physiological signals other than the ECG, even though this has the potential of improving detections when the ECG is compromised. The Physionet/CinC 2014 challenge (the 'Challenge') provided a multimodal training set of 100 recordings with references for this task, with a hidden set of 300 recordings for testing purpose [4].

2. Material and Methods

2.1. New Training Set

The challenge training set contained 100 10-minute segments of up to 8 physiological signals, all containing one ECG and one ABP channel. Additional signals which were inconsistently present included the electrooculogram (EOG), electroencephalogram (EEG), stroke volume (SV), other blood pressure based signals and respiration signals. Evaluation of the sample entry ("gqrs"), available on Physionet [5], achieved a score of 99.60% on the training set, leaving no real space for improvement. Moreover, the sample entry scored 89.83% for phase I, 85.70% for phase II, and 84.49% for the final phase. The training set was not representative of the complexity of the test set and consequently could not be used for deriving an intelligent multimodal peak fusion algorithm.

In order to alleviate the issues with the training set, a custom database was built. As the challenge training set had both ECG and ABP signals available in all records, it was desirable to create a new training set with one ECG lead and one ABP channel. The MGH/MF Waveform database [6] is a collection of 250 recordings from hemodynamic (ABP, CVP, PAP) and electrocardiographic (three leads) signals acquired from patients in critical care units, operating rooms, and catheterization labs. These recordings were usually one hour in length, but could vary between 12 to 86 minutes. All recordings had manually corrected R-peak annotations. The MGH/MF database was reformatted into 5274 10-minute recordings, containing one ECG and one ABP channel.

2.2. Individual Peak Detector

R-peak detection on ECG signals has been extensively studied. In this work three different peak detectors were evaluated: (1) “gqrs” (available on Physionet [5]), which consists of a QRS matched filter with a custom built set of heuristics (such as search back). (2) “coqrs” [7–9] based on the peak energy (no search back). (3) “jqrs” [10, 11] consists of a window-based peak energy detector but with replacement of the original band-pass filter with a QRS matched filter (Mexican hat) and an additional heuristic ensuring no detections were made during flat lines.

The onset of the pulses in the ABP signal was detected using an open-source algorithm, “wabp”, based on the length transform [12]. These detections occur with a delay relative to the R-peak as the peak of the pressure wave occurs after the heart contraction. The delay was estimated on a patient-specific basis by isolating a one minute segment with more than 80% agreement between the R-peak detections and wabp detections. Agreement was high when a R-peak detection was consistently followed by a wabp one. The mean delay between these ABP peaks and their matched R-peaks was set as the ABP delay for the entire record. If there was no one-minute segment which fulfilled this criteria, a delay of 200 ms was set.

Other physiological signals were potentially available and could provide useful information. SV was processed and peaks were extracted using a zero-crossing procedure after a band-pass filter. Peaks were also detetected from PPG signals using a peak energy technique [11]. As for ABP peaks, SV and PPG peaks were mapped back in order to account for their intrinsic delays.

2.3. Fusion of multiple signals

The most important aspect of the Challenge was the development of an intelligent fusion technique for multiple peak detectors from different physiological signals. Two approaches were implemented: the first based on signal quality indices (SQI) for the different physiological signals and the second based on the regularity of the RR interval time-series. The two approaches are referred as fusion-SQI (FSQI) and fusion-regularity (FREG) in what follows.

For FSQI, the agreement level of two R-peak detectors in a 10-second window, evaluated every second, was used as the SQI (known as “bSQI”). Intuitively, the presence of noise and artifacts will lower the agreement level between two semi-independent detectors. bSQI was recently successfully used on a database with pathological rhythms [3,13]. The ABP signal quality was also evaluated using an open-source algorithm [14] which flags a signal as bad quality if derived parameters from a blood pressure wave are not in reasonable physiological ranges. Only the pressure ranges and the average derivative for a cycle (a subset of the original SQI parameters) were checked for validity, and this SQI is henceforth known as aSQI. The switching proceeded on a second-by-second basis: if bSQI value was > 0.9, the ECG was used. If bSQI value was ≤ 0.9, and if aSQI $= 1$ (all derived parameters were physiologically plausible), the ABP was used. Finally, if both signals were considered bad quality, the ECG was used.

The FREG approach was based on selecting the most regular of a set of RR interval time-series for consecutive windows. The implicit assumption made is that the presence of noise and artifacts will induce missed or extra detections, whose subsequent rhythm is less likely to be regular than the true RR interval time-series (even in presence of arrhythmias). This approach was used previously in foetal ECG extraction for the selection of the best abdominal lead [11]. The surrogate for regularity was chosen as the standard deviation of the RR interval time-series. If a time-series contained less than three detections within a given window, it was considered to be bad quality. The optimal window size was searched for and set at 15 seconds.

2.4. Algorithm evaluation

The challenge score was defined as the average of four quantities: the average sensitivity (Se), average positive predictivity (PPV), gross Se, and gross PPV. The F_1 measure was also evaluated: it is defined as $F_1 = 2(Se \times PPV)/(Se + PPV)$, thus equally penalising false positives and false negative detections which is likely to be a better accuracy measure than the average between Se and PPV [11].

All individual ECG and ABP detectors were evaluated on the bespoke dataset. Five of the best entries in phase III, excluding the method here-in, were also evaluated [15–19]. Finally, two entries corresponding to FSQI and one to FREG were also assessed. The FSQI approach requires two ECG peak detectors, both are used to compute the SQI.The first entry used “gqrs” as the baseline set of annotations, denoted FSQI (gqrs), and the second one used “jqrs” as the baseline, denoted FSQI (jqrs). The third entry used the FREG approach, with three peak detectors applied to the ECG signal and “wabp” applied to the ABP signal, and finally peaks detected on SV and PPG when available. These three entries were also evaluated on the Challenge final test set in phase III.

3. Results

The results for the different techniques using the custom training set are assembled in Table 1. The highest PPV was achieved using jqrs, which only used the ECG signal and did not contain any multimodal switching. The "Challenge" scored entries on the final test set using the depicted techniques are assembled in Table 2. The best result was achieved using FSQI (gqrs). This technique performed slightly better ($+0.22\%$) than FREG, and provided the highest official score in the final phase of the competition.

4. Discussion & Conclusion

In this paper, two methods for fusing peak detected from multiple physiological signals were described. Both approaches obtained similar results on the Challenge test set, with the SQI-based technique being slightly better ($+0.22\%$) whereas it was slightly worse on the custom training set (-0.64%).

Overall, the use of hemodynamic signals for the estimation of peak positions was shown to be beneficial on noisy signals. Nevertheless, such an approach must be taken with care. The presence of an arrhythmia can sometimes only manifest on ECG signals and not on hemodynamic signals. One such example is an early premature ventricular contraction preceding the filling of the ventricles. In this case, the ECG and ABP will inevitably disagree, and a regularity based approach would incorrectly switch to the more regular ABP.

There are some limitations to the use of MGHDB as a custom training set. First, as it only contains ECG and ABP signals, performance on this database are biased to methods which switch between these two signals. As all Challenge training set records had ECG and ABP signals, this bias was desirable, but this should still be acknowledged. Secondly, the MGHDB may not accurately reflect the Challenge test set, as arrhythmias or dramatic rhythm changes may be underrepresented in the MGHDB.

Acknowledgments

JB is supported by the UK Engineering and Physical Sciences Research Council, the Balliol French Anderson Scholarship Fund and MindChild Medical Inc. North Andover, MA. JO is supported by Wellcome Trust Centre Grant No. 098461/Z/12/Z (Sleep, Circadian Rhythms & Neuroscience Institute). FA is supported by CNPq - Brazil and TU Dresden's Graduate Academy. AJ acknowledges the support of the RCUK Digital Economy Programme grant number EP/G036861/1 (Oxford Centre for Doctoral Training in Healthcare Innovation).

References

[1] Chambrin MC, Ravaux P, Calvelo-Aros D, Jaborska A, Chopin C, Boniface B. Multicentric Study of Monitoring Alarms in the Adult Intensive Care Unit (ICU): a Descriptive Analysis. Intensive Care Medicine 1999;25(12):1360–1366.

[2] Aboukhalil A, Nielsen L, Saeed M, Mark RG, Clifford GD. Reducing False Alarm Rates for Critical Arrhythmias using the Arterial Blood Pressure Waveform. Jour of Biomed Info 2008;41(3):442–451.

[3] Li Q, Mark RG, Clifford GD. Robust Heart Rate Estimation from Multiple Asynchronous Noisy Sources using Signal Quality Indices and a Kalman Filter. Phys Meas 2008; 29(1):15.

[4] Moody GB, Moody BE, Silva I. Robust Detection of Heart Beats in Multimodal Data: The PhysioNet/Computing in Cardiology Challenge 2014. In Comput in Cardiol, volume 41. 2014; .

[5] Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. Physiobank, Physiotoolkit, and Physionet Components of a New Research Resource for Complex Physiologic Signals. Circ 2000;101(23):e215–e220.

[6] Welch J, Ford P, Teplick R, Rubsamen R. The Massachusetts General Hospital-Marquette Foundation Hemodynamic and Electrocardiographic Database–Comprehensive Collection of Critical Care Waveforms. Clinic Monitor 1991;7(1):96–97.

[7] Clifford GD. Signal Processing Methods for Heart Rate Variability. Ph.D. thesis, University of Oxford, Michaelmas 2002.

[8] Nygårds ME, Sörnmo L. Delineation of the QRS Complex using the Envelope of the ECG. Med and Biol Engin and Comput 1983;21(5):538–547.

[9] Oster J, Behar J, Colloca R, Li Q, Li Q, Clifford GD. Open source Java-based ECG Analysis Software and Android App for Atrial Fibrillation Screening. In Comput in Cardiol. 2013; 731–734.

[10] Behar J, Johnson A, Clifford GD, Oster J. A Comparison of Single Channel Fetal ECG Extraction Methods. Ann of Biomed Engin 2014;42(6):1340–1353.

[11] Behar J, Oster J, Clifford GD. Combining and Benchmarking Methods of Foetal ECG Extraction without Maternal or Scalp Electrode Data. Phys Measur 2014;35(8):1569.

[12] Zong W, Heldt T, Moody G, Mark R. An open-source algorithm to detect onset of arterial blood pressure pulses. In Comput in Cardiol. 2003; 259–262.

[13] Behar J, Oster J, Li Q, Clifford GD. ECG Signal Quality during Arrhythmia and its Application to False Alarm Reduction. IEEE Trans on Biomed Engin 2013;60:1660–1666.

[14] Sun J, Reisner A, Mark R. A Signal Abnormality Index for Arterial Blood Pressure Waveforms. In Comput in Cardiol. 2006; 13–16.

[15] Vernekar S. PhysioNet/Computing in Cardiology Challenge. In Comput in Cardiol, volume 41. 2014; .

[16] Teo SK, Yang B, Huan D, Christopher M, Su Y. Robust Detection of Heart Beats using ECG, BP and EEG signals. In Comput in Cardiol, volume 41. 2014; .

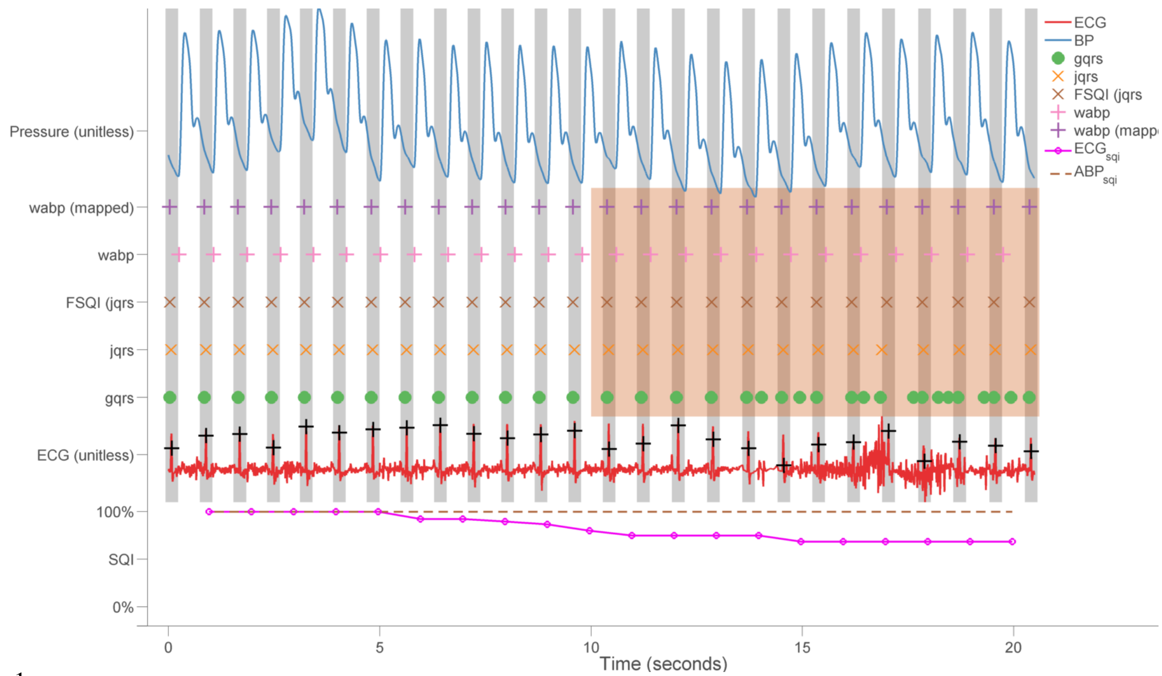

Figure 1. Example of SQI switching when ECG is low quality. The top row contains the ABP signal, while the ECG signal is depicted in red in the third row from the bottom, with the reference annotations marked with black crosses. Vertical grey lines represent the windows for acceptable detected peaks. The two bottom graphs represent both bSQI (plain line with circle markers) and aSQI (dotted lines). The annotations of the different individual detectors are depicted in the middle rows, and the region using ABP annotations instead of ECG annotations is shaded red.

		Average		Gross			
	Algorithm	Se(%)	PPV(%)	Se(%)	PPV(%)	F_1(%)	Score(%)
Single detectors	gqrs	91.75	96.19	91.87	96.63	94.19	94.11
	jqrs	92.66	**96.51**	93.00	96.72	94.82	94.72
	coqrs	91.19	93.46	91.77	93.41	92.58	92.46
	wabpmapped	77.03	80.38	79.79	81.59	80.68	79.70
This work	FSQI (gqrs)	93.94	95.90	94.29	96.39	95.38	95.16
	FSQI (jqrs)	93.73	96.18	94.03	96.28	95.15	95.06
	FREG†	**95.83**	95.56	96.23	95.56	**95.89**	**95.79**
Competitors	sachi [15]	90.43	90.35	95.39	94.30	94.84	92.61
	teosk [16]	89.82	95.66	89.91	96.41	93.05	92.95
	podz [17]	92.94	95.75	92.69	95.64	94.14	94.25
	marcus [18]	95.34	95.33	94.46	96.22	93.84	94.96
	thomas [19]	92.06	96.31	92.16	**96.98**	94.51	94.37

Table 1. Results of the different techniques on the MGH training database.
† It should be highlighted that on this training set no SV or PPG signals are available.

	Average		Gross			
Algorithm	Se(%)	PPV(%)	Se(%)	PPV(%)	F_1(%)	Score
FSQI (gqrs)	89.78	**86.74**	89.73	**85.29**	**87.45**	**87.88**
FSQI (jqrs)	88.66	86.18	88.68	85.09	86.82	87.15
FREG	**91.19**	85.08	**91.18**	83.22	87.02	87.66

Table 2. Results of the different techniques on the Challenge test set.

[17] Gieraltowski JJ, Ciuchcinski K, Grzegorczyk I, Kozna K, Solinski M, Podziemski P. Heart Rate Variability Discovery: ALgorithm for Detection of Heart Rate from Noisy Multimodal Recordings. In Comput in Cardiol, volume 41. 2014; .

[18] Vollmer M. Robust Detection of Heart Beats using Dynamic Thresholds and Moving Windows. In Comput in Cardiol, volume 41. 2014; .

[19] DeCooman T, Goovaerts G, Varon C, Widjaja D, VanHuffel S. Heart Beat Detection in Multimodal Data Using Signal Recognition and Beat Location Estimation. In Comput in Cardiol, volume 41. 2014; .