Robust Algorithm to Locate Heart Beats from Multiple Physiological Waveforms

Lars Johannesen[1,2], Jose Vicente[1,3], Christopher G Scully[1], Loriano Galeotti[1], David G Strauss[1]

[1]Office of Science and Engineering Laboratories, Center for Devices and Radiological Health, US FDA, Silver Spring, Maryland, USA

[2] Department of Clinical Physiology, Karolinska Institutet and Karolinska University Hospital, Stockholm, Sweden

[2] BSICoS Group, Aragón Institute of Engineering Research (I3A), IIS Aragón, University of Zaragoza, Zaragoza, Aragon, Spain

Abstract

Alarm fatigue is a major issue in patient monitoring that could be reduced by merging physiological information from multiple sensors, minimizing the impact of a single sensor failing. We developed a heart beat detection algorithm that utilizes multi-modal physiological waveforms (e.g. ECG, blood pressure, stroke volume, photoplethysmogram and electro-encephalogram). The 100 record training set from the Physionet challenge 2014 was used for development. The algorithm was evaluated at three testing phases during the 2014 challenge consisting of 100 (phase I), 200 (phase II) and 300 (phase III) hidden records, respectively. A true positive was declared if a beat was detected within 150 ms of a reference annotation. The algorithm had a sensitivity of >99.9%, Positive Predictive Value of 99.7%, and an overall score (average of sensitivity and Positive Predictive Value) of 99.8% when applied to the training set. The best overall performance on the test sets were: 88.9%, 76.3% and 84.4% for phases I, II and III, respectively. We developed a robust heart beat detector that fuses annotations from multiple individual detectors. The algorithm improves the training results compared to ECG detections alone, and performs well on the test sets. Data fusion approaches like this one can improve patient monitoring and reduce false alarms.

1. Introduction

The heart rate of patients in a hospital setting such as an intensive care unit is an important vital sign, and can be derived from different signal types such as the electrocardiogram (ECG) or blood pressure waveform (BP). Today the heart rate is typically obtained from a single signal (e.g. from an ECG lead), and artifacts such as patient movements or sensor disconnect may result in false alarms due to incorrect heart rate measurements. Many of these false alarms could be avoided if using correct beat rate information extracted from other signals (e.g. BP), however this is not usually done by monitoring devices in clinical practice. The goal of the Physionet challenge 2014 [1] was to develop a new method that could be used for measuring the heart rate using information from different signal types (ECG, BP, stroke volume (SV), photoplethysmogram (PPG), electroencephalogram (EEG), electrooculogram (EOG)) recorded at the same time to improve the automatic detection of heart beats. In this paper we propose a solution to this year's Physionet challenge.

2. Methods and data

2.1. Data

For algorithm development, we used mainly the 2014 Physionet challenge training set. It contains 100 records (most 10 minutes in duration) sampled at 250 Hz. Each record contains several signals. The signals in the training set always included ECG and BP, and some records additionally included SV, EEG or EOG. The signal order could vary, and the signal name could be incorrect. However, the first signal included in each record was always an ECG. We also used the Physionet challenge 2009 [2] for algorithm development as the 2014 training set did not include PPG signals.

The algorithm performance was then evaluated using three hidden test sets, each containing 100 records similar to the training set. The evaluation was divided into three testing phases so that Phase I contained only test set 1, Phase II contained test sets 1 and 2 and Phase III contained test sets 1, 2 and 3. Each record could be sampled at any frequency between 100 and 1000 Hz. None of the test records were available to the challenge participants as the validation was done on a dedicated Physionet server.

2.2. Algorithm overview

The algorithm consists of multiple stages: detection of heart beats in the various signals independently, correction for time lag between signals, and combination of the heart beat detections across signals using a voting scheme (Figure 1).

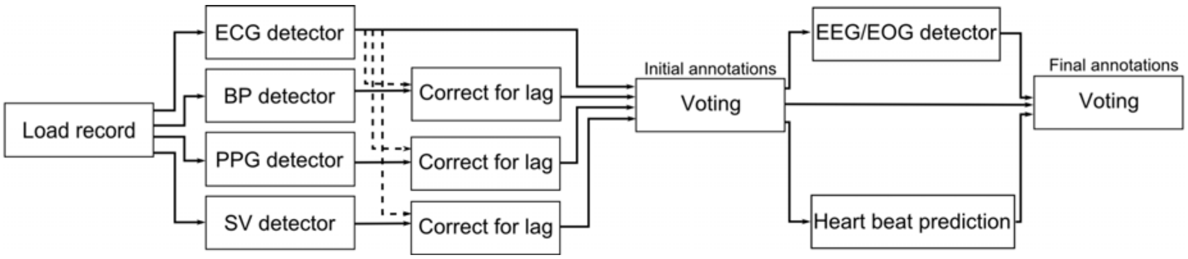

Figure 1. Algorithm overview.

2.2.1. ECG detector

Heart beats were detected in the ECG by using a detector based on the U3 signal [3, 4]. More specifically a modified version of the one proposed in [5]. The modified version has similar performance as evaluated using the MIT-BIH database [6].

2.2.2. Pulsatile heart beat detectors

Heart beats in pulsatile signals such as BP, SV and PPG were detected using a second derivative-based approach. For the blood pressure each beat was associated with a measure of reliability based on the root-mean-square of the beat. Both the BP and SV detectors were developed using the training set, but the PPG detector was developed and tested using records from the Physionet challenge 2009, as the training set provided did not include any PPG signals [2].

2.2.3. EEG/EOG detector

A specialized version of the U3 detector was used to identify heart beats presenting as artifacts in the EEG and EOG signals. Not all EEG/EOG contain noticeable heart beat artifacts. An example of an EEG with detectable heart beat artifacts and the corresponding U3 signal is shown in the left panel of Figure 2 and an example of an EEG without clear heart beat artifacts in the right panel.

To determine if heart beat artifacts were present, the U3 amplitude value of spikes occurring at the same time as the annotation set from the previous stage were compared with the rest of the spikes. This comparison was only performed inside regions of detections. If the spike amplitudes that coincided with the initial detections were greater than the rest of the spike amplitudes, the signal was used for detecting heart beat artifacts. The detection threshold was subsequently defined based on spike amplitude that coincided with other detections.

2.2.4. Correcting for lag time

Beats present in pulsatile signals (BP, SV or PPG) can be substantially delayed compared to those present in the ECG. We corrected this delay (that we will refer as lag

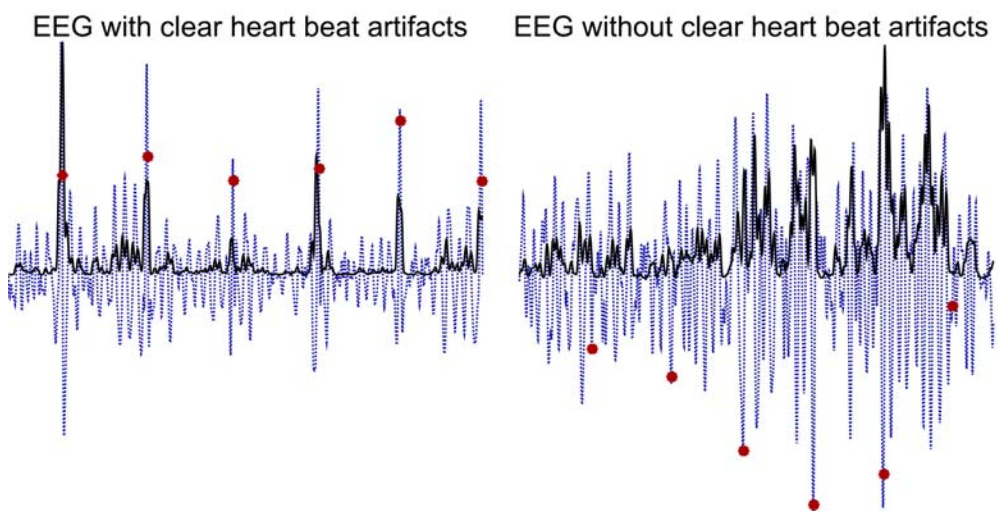

Figure 2. Example of EEG signals (blue dotted line) and the corresponding U3 signal (black solid line) and heart beat annotations obtained in the previous stage (red dots).

time) by using the autocorrelation between the beat annotations from the ECG time series with the greater number of detections and the time series of the signal that needs to be corrected. If the estimated lag results outside a physiologically reasonable range or it was not possible to determine the lag reliably, it was replaced by a default value (average lag observed in the training set).

2.2.5. Combining beat detections

After correcting for lag time, the voting scheme was run twice, once to produce an initial set of locations (used to train the EEG/EOG detectors) and a second time to produce the final annotation set (including EEG/EOG detections).

Both voting schemes were implemented by first converting the individual signal annotations into a time series with a Tukey window (tapered cosine) of 250 ms width centered at each detection. The Tukey window time series for all signal types are then summed together, to create a voting signal.

The Tukey windows were weighted based on signal type. Specifically, ECG, BP and PPG detections were weighted higher than SV, EEG and EOG due to performance on the training set. Thus, detections in either ECG, BP or PPG could trigger final detections on their own.

In contrast, two of the SV, EEG and EOG detections or one of these that overlapped with the location of a predicted heart beat (using linear interpolation) were required to trigger a final detection. Finally, if there were two consecutive detections within 250 ms in the voting signal, the location of the largest was selected. In the case of two detections within 250 ms of equal weight, the average of the two detections was chosen.

2.3. Statistical analysis

The algorithm was applied to the annotated training set. A true positive was declared if a beat was detected within 150 ms of a reference annotation, using bxb from the Physionet library [7]. Gross and average sensitivity (Se) and positive predictive value (PPV) were determined from the true positives, false positives and false negatives. In addition an overall performance is reported, which is an average of the gross and average Se and PPV results.

2.4. Implementation

The algorithm was implemented in MATLAB/Octave with some components written in C++. As per requirement of the challenge, the code is freely available on the Physionet challenge 2014 website [1].

Table 1. Results for different detectors on the training set

Training	Se*	PPV*	Overall
Combined	>99.9 / >99.9	99.7 / 99.7	99.8
ECG	99.9 / 99.9	99.8 / 99.8	99.8
BP	99.6 / 99.6	>99.9 / >99.9	99.8
SV	94.9 / 95.0	97.2 / 96.8	96.0
EEG	61.7 / 62.5	90.1/90.0	76.3
EOG	60.8 / 60.0	88.4/86.6	73.9

* Gross / Average

Table 2. Results for the test set

Test	Se*	PPV*	Overall
Phase I	86.6 / 85.5	95.7 / 88.0	88.9
Phase II	73.4 / 75.3	80.5 / 75.6	76.3
Phase III	84.6 / 82.9	86.8 / 83.5	84.2

* Gross / Average

3. Results

3.1. Performance on training dataset

The training set was used for algorithm development, and each of the main detectors (ECG, BP and SV) all achieved overall performance >95% (Table 1).

The performances for the EEG and EOG detectors were lower, due to a much lower sensitivity (~60%). The lower sensitivity is a result of only 24 (EEG) and 30 (EOG) out of the 40 training records with EEG/EOG signals having heart beat artifacts that were detectable.

3.2. Performance on validation dataset

Unlike the training set, only results for the combined detector were available for the test set. The best results from each of the three phases are listed in Table 2. The results in Table 2 show that the multi-signal detector was able to fuse detections in phases I and III, but less so in phase II. The reason for the lower performance in phase II was due to experimentation with the ECG detector and a signal quality index for the ECG.

Moreover, it should be noted that the results from the different phases are not directly comparable as the test set changed between the phases (see Data section).

4. Discussion

Important clinical measurements such as heart rate can be subject to measurement error and subsequently false alarms when the signal used is noisy. A method to fuse heart beat detections from multi-modal physiological waveforms has the potential to decrease the false alarm rate. We developed a robust heart beat detector that fuses detections from BP, PPG, SV, EEG, and EOG signals with ECG detections, using the training sets from the Physionet challenge 2014 [1] and the Physionet challenge 2009 [2]. The detections are then merged using a voting scheme.

The proposed voting scheme allows for a simple way to combine detections from different signals. Specifically, it allows for lowering the weight of detections when we have less confidence that the detections reflect a true heart beat, e.g. due to noise. In the training set the performance of the multi-detector is slightly better than any detector individually, but in the test set the improvement is not known due to the limited information reported by the testing environment.

However, in the test set the advantage over the sample entry (essentially gqrs, a QRS detector from the Physionet library [7]) was not evident in either of the phases. This could be due to differences in the ECG detectors. The ECG detector used in this work has been tested against MIT-BIH, where it showed a false error rate comparable to other detectors [5]. It is also worth noting, that during the development process we could not obtain any debug information from the testing dataset and thus we were unable to exclude implementation errors that we did not observe on the training set.

While the training set contains relatively “clean” signals sampled at a unique frequency of 250 Hz with the channels labeled consistently, the records in the hidden test sets are sampled at different frequencies, the channels can be mislabeled and the signals are likely more “noisy”. Attempts were made to develop an ECG signal quality metric (in phase II), which performed well at excluding noisy beats (using NSTDB [8]). However, after removal of the signal quality metric in phase III the overall performance increased, suggesting that the metric did not capture the ECG quality problems in the test sets.

We have proposed a novel way of fusing heart beat detections from different signal types, and accounting for patient specific lag between different pulsatile signals and the ECG. Moreover, we have proposed new detectors for BP, SV, PPG and EEG/EOG. Multi-signal detectors and data fusion approaches such as proposed in this article could improve patient monitoring and lead to a reduction of false alarms, but further improvements in the signal quality assessment are necessary.

Acknowledgements

This project was supported in part by FDAs Critical Path Initiative, Office of Women’s Health, Medical Countermeasures Initiative and appointments to the Research Participation Programs at the Oak Ridge Institute for Science and Education through an interagency agreement between the Department of Energy and FDA.

Disclaimer

The opinions presented here are those of the authors, and not necessarily those of the US Food and Drug Administration. No official support or endorsement by FDA is intended nor should be inferred. The mention of commercial products, their sources, or their use in connection with material reported herein is not to be construed as either an actual or implied endorsement of such products by FDA.

References

[1] Physionet. Physionet Challenge 2014. http://physionet.org/challenge/2014. Accessed 14 August 2014.

[2] Physionet. Physionet Challenge 2009. http://physionet.org/challenge/2009/. Accessed 14 August 2014.

[3] Paoletti M, Marchesi C. Discovering dangerous patterns in long-term recordings using a fast QRS detection algorithm and explorative data analysis. Comput Methods Programs Biomed 2006;82:20-30.

[4] Marchesi C, Paoletti M. ECG Processing Algorithms for Portable Monitoring Units. Internet J Med Tech 2004;1(2).

[5] Johannesen L, Vicente J, Galeotti L, Strauss D. ECGlib: Library for Processing Electrocardiograms. Comput Cardiol 2013;40:951-4.

[6] Moody G, Mark R. The impact of the MIT-BIH Arrhythmia Database. IEEE Eng in Med and Biol 2001;20:45-50.

[7] Goldberger A. Physiobank, PhysioToolkit, and PhysioNet Components of a New Research Resource for Complex Physiologic Signals. Circ 2000;101:e212-200.

[8] Moody G, Muldrow W, Mark R. A noise stress test for arrhythmia detectors. Comput Cardiol 1984;11: 381-4.

Address for correspondence:

Lars Johannesen
Building 62, RM 1125C
10993 New Hampshire Ave, Silver Spring, MD, 20910, USA
Lars.johannesen@fda.hhs.gov