# Robust Identification of Heartbeats with Blood Pressure Signals and Noise Detection

Bo Yang, Soo Kng Teo, Bart Hoeben, Christopher Monterola, Yi Su

Institute of High Performance Computing, A*STAR, Singapore

## Abstract

*Accuracy in detection of electrocardiographic (ECG) heart beats can be vastly improved with the aid of blood pressure (BP) monitoring. Cross validation between ECG and BP signals is used to identify the part of signals not contaminated by large, high frequency noises, where we can extract accurately the delay between the QRS peaks and BP peaks (defined in the "Methods" section). The delay is used to identify the QRS peaks even when the signal is very noisy. We also present a simple noise detection algorithm for the ECG signals. This complementary algorithm leads to high success rate in identifying aberrant ECG beats including the supraventricular premature beats (SVPB), premature ventricular contraction (PVC) and other unclassifiable beats.*

## 1. Introduction

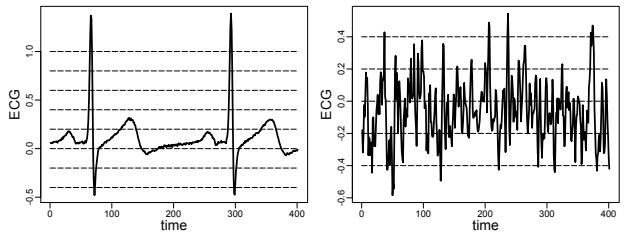
Efficient detection of heart beats from ECG signals for preliminary classification and diagonosis of heart problems remains a challenge[1, 2] due to the large amount of data involved; visual detection can be tedious and very time consuming, as the time series of ECG signals for each patient needs to be long enough to capture cardiac events. Despite significant improvement in performance within the last few years, the effectiveness of such algorithms is still suboptimal[3–9]. The main challenge of processing the ECG signals is not the intrinsic systematic noises including, for example, the power line interference and baseline wandering, as the QRS complex of the ECG signals has large signal-to-noise ratio. It is the sporadic, accidental noises coming from motion artifacts, sweating and muscle contractions, as well as temporary machine malfunctions such as detachment of electrodes and sensors, that interfere most with the heart beat detection.

While most heart beats occur at fairly regular intervals and are easy to predict even for noisy signals, it is difficult to locate aberrant heart beat signals such as SVPB and PVC that may occur at irregular beat intervals. It is nontrivial for the algorithm to differentiate these beats apart from noises, as both of them only occur occasionally, and some of the PVC's do not have complete QRS complex.

In this work we use an independent and concurrent measurement of the patient's blood pressure to complement the QRS peak (defined in the "Methods" section) detection of the ECG signals and improve the detection of the heart beats especially when the signal is noisy. Like the ECG signals, the BP signals also enjoy large signal-to-noise ratio. In addition, the BP signal measurement does not depend on the electrical conductivity between the leads and the patient's skin, it is thus less susceptible to sweating. Thus in general the BP signal is much less noisy. It is also relatively rare for both the ECG signal and the BP signal to be noisy at the same time period, since the former is the measurement of the electrical signals, while the latter is of the mechanical signals.

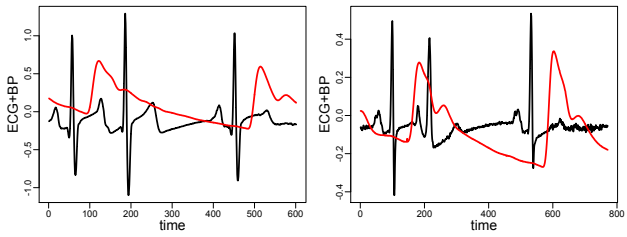
Based on these characteristics of the ECG and BP signals, it is possible to treat the two channels separately and obtain accurately the position of the majority of QRS and BP peaks. This is followed by the extraction of the time delay between the heart beat and the increase in the blood pressure. This time delay is mainly determined by the physical charateristics of the patient (e.g. the distance between the heart chamber and the point of blood pressure measurement), and is in general independent of the patient's heart condition. We thus use this time delay to remove the false positive and false negative "QRS peaks" that are obtained from the ECG signals alone. Special algorithms are also implemented to treat the cases where both the ECG signals and BP signals are noisy, and cases where aberrant heart beats occurs.

The over-arching principle of our algorithm is simplicity, with as few tuning parameters as possible to avoid over-fitting our code during optimization with the training set. We also emphasized accurate detection of the aberrant heart beats. In rare cases our algorithm may sacrifice the identification of noisy normal beats so as not to miss the crucial aberrant ones. This is because we believe the aberrant ones are also the clinically most important ones for doctors to diagnose the symptoms of the patients.




(a) (b)

Figure 1: (a). Two consecutive QRS peaks at regular intervals. (b). QRS peaks contaminated by noises.The horizontal dotted lines are for noise detection, as explained in the main text.




(a) (b)

Figure 2: (a). Supraventricular premature beat (SVPB) sandwiched between two normal beats. (b). Premature ventricular contraction (PVC) sandwiched between two normal beats. The black plots are ECG signals, and the red plots are BP signals.

## 2. Signal Characteristics

It is instructive to first study different types of signal profiles commonly encountered in various parts of ECG and BP signals, as these are the only two channels we focus on. A clean and noisy ECG signal is shown in Fig.(1). While the QRS peaks of clean ECG signals are easy to capture, the noisy signals contain multiple spurious peaks with comparable amplitudes; thus accurately locating the QRS peaks with the ECG signal alone is difficult. For most cases, however, only around $1 \sim 2\%$ of the ECG signal is contaminated with high frequency noises with large magnitude. Using a butterworth filter can be tricky, as it tends to distort the shape of the QRS peaks, risking possible misidentification of the bulk of the ECG signals.

Another important feature of the ECG signals is the presence of aberrant heart beats as shown in Fig.(2). For both the cases, the heart beat intervals between a normal beat and an aberrant beat are significantly smaller than those between two normal peaks. The SVPB in general has a well-defined regular QRS complex, while a PVC tends to be less regular, in some cases only with a pronouced "R" peak.

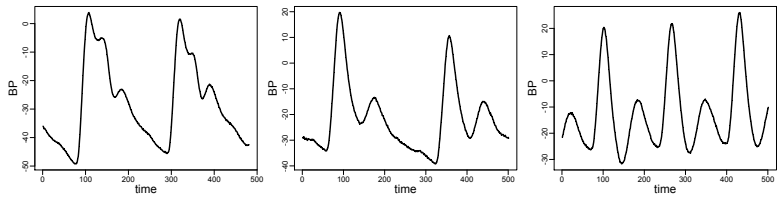
A regular BP peak is characterized by a percussion peak closely followed by a dicrotic peak with a smaller amplitude, as shown in Fig.(3a) and Fig.(3b). It is very rare for BP signals to be contaminated by high frequency noises like the one in Fig.(1). Mis-identification of the BP peaks may occur when the amplitude of the dicrotic peak is significant, and proximate to the primary peak, as shown in Fig.(3c).

## 3. Methods

Our methods depend heavily on the performance of WFDB toolbox[10] routines *gqrs* and *wabp*. The former gives the QRS peak (defined as the peak of the "R" structure of the QRS complex) from the ECG channel, while the latter gives the BP peak (defined as the percussion peak) from the BP channel. Both routines are reasonably robust for clean signals (defined as signals not contaminated by sporadic high frequency, large magnitude noises). We define signals other than clean signals as *contaminated signals*. False postives and false negatives can occur because the QRS complex is distorted or incomplete, or because of the presence of noise like Fig.(1b). BP signals like Fig.(3c) are also likely to lead to false positives in BP peaks. The routines also cannot distinguish cases when the signal is wildly fluctuating or missing. For the rest of the section, we use "QRS peaks" and "BP peaks" with quotation marks to indicate peaks given by *gqrs* and *wabp* respectively; they may or may not be the real QRS and BP peaks (which we indicate without quotation marks).

### 3.1. Pre-treatment of ECG and BP signals

We keep the pre-treatment of the ECG and BP signals to a minimum, due to our algorithm's reliance on the WFDB toolbox routines *gqrs* and *wabp*. As we have mentioned in the introduction, for most part of the signals, the signal-to-noise is sufficiently large for the routines to work extremely well. In the absence of contaminations, the performance of *gqrs* and *wabp* is not significantly affected by the small amplitude, high frequency noises present throughout the signals. Both ECG and BP signals also contain large, long wavelength baseline wandering, but this also does not significantly affect the performace of *gqrs* and *wabp*, as they both detect the local signal profiles for the QRS and BP complex. Any pre-treatment that focus on elimitating the rare occurance of sporadic noises risks affecting the performance of *gqrs* and *wabp* on the majority part of relatively clean signals.




(a) (b) (c)

Figure 3: Different profiles of the BP signals. The signal in (c) tends to be mis-identified because the dicrotic peak has large amplitude and is not proximate to the percussion peak.

### 3.2. QRS and BP peak detections with WFDB toolbox

Using the raw ECG and BP signals from the patients' cardiac recording, the QRS peaks and BP peaks are extracted with *gqrs* and *wabp* respectively for further processing. For the routine *gqrs*, false positive detections may result from noisy part of the signals similar to that shown in Fig.(1b). This type of high frequency, large amplitude noises may also lead to mis-identification of the location of the QRS peak (which differs from the true QRS peak by more than 150ms). Another main source of false negative detections is from the some of the PVC's, where the "Q" and "S" part of the QRS peak is not well-developed.

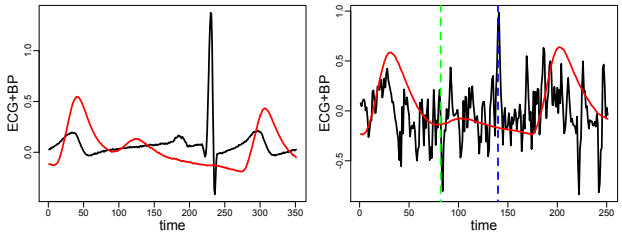
For the routine *wabp*, the main source of the error comes from the mis-identification of the dicrotic peak as the percussion peak, leading to false postive BP peaks with significantly small intervals between consecutive prescribed BP peaks by *wabp*.

Naturally, both routine also fail to work when part of the ECG or BP signals are missing or wildly fluctuating, most probably due to accidental detachment of the measuring device from the patient. One should also note the routine *gqrs* alone can achieve a success rate of $\sim 98\%$ for the 100 records from the training set of the CinC Challenge 2014.

### 3.3. Identification of clean ECG and BP signals

Using the results from *gqrs* and *wabp*, we employ a simple and robust way to identify part of the clean ECG and BP signals where we are confident *gqrs* and *wabp* produce correct results. From the "QRS and BP peaks" reported by *gqrs* and *wabp*, we enforce two simple criteria: a) only one "QRS peak" is sandwiched by two consecutive "BP peaks"; b) only one "BP peak" is sandwiched between two consecutive "QRS peaks". The physical motivation of such two criteria is obvious, and an example is shown in Fig.(4a).

This simple algorithm is not completely error proof. The only source of error we observe from the training set is




(a) (b)

Figure 4: a). An example of clean ECG (black dots) and BP (red curve) signals. b). An example where the ECG signal is noisy, but *gqrs* only gives one (erroneous) QRS peak, given by the vertical green dotted line. The correct location of the QRS peak is given by the vertical blue dotted line.

the rare occasion when *gqrs* mislocated the position of the QRS peak due to the noise, without producing additional false postives or false negatives. For noisy ECG signals like the one shown in Fig.(4b), *gqrs* can produce any number of "QRS peaks" at any random positions. In this particular rare case *gqrs* happened to produce *one* QRS peak at the position of the dotted green line. Even though the position is not correct, our algorithm will mistakenly identify the ECG signal at that part to be clean. This rare mistake can be easily corrected in the next step.

### 3.4. QRS peak validation and prediction

Once the clean part of ECG and BP signals are identified, we can calculate the moving average of the delay between the "QRS peak" and the subsequent "BP peak" from the *clean* part of the signals only. The moving average helps to eliminate the effect of long wavelength baseline wandering, and a window size of *ten* consecutive delays is enough to minimize the distortion introduced by rare cases like those in Fig.(4b).

The validation of the "QRS peaks" at the clean part of the signal and the prediction of the QRS peaks at the noisy part of the signal with the QRS-BP delay is separated into

the following three cases:

*Case 1:* For two consecutive "BP peaks" sandwiching only one "QRS peaks" (previously identified as the clean part of the ECG signal), the closest moving average QRS-BP delay is used to correct those cases in Fig.(4b) if they do occur.

*Case 2:* For two consecutive "BP peaks" sandwiching more than one "QRS peaks", the closest moving average QRS-BP delay is used to predict the correct position of the QRS peak between these two "BP peaks".

*Case 3:* For two consecutive "BP peaks" with no "QRS peaks" in between, no action is taken.

For *Case 2*, the multiple "QRS peaks" are most likely the false positive peaks identified by *gqrs* from noisy ECG signals. It could also happen where part of the BP signal is missing. This rare situation will be dealt with later. For *Case 3*, the most likely cause could be the *wabp* routine mis-identifying the dicrotic peak as the percussion peak (see Fig.(3c), so no action is taken.

### 3.5. Aberrant heart beats and missing BP signals

In the rare cases of aberrant heart beats including SVPB and PVC, there are two "QRS signals" sandwiched between two consecutive BP peaks. Similarly, when part of the BP signals are missing, many "QRS peaks" can be sandwiched between two consecutive BP peaks. It is thus hard to distinguish these legitimate QRS peaks from noisy ECG signals.

We employ a simple "horizontal line check" for cases when more than one "QRS peaks" are sandwiched between two consecutive BP peaks. As one can see from Fig.(1), a sequence of regularly spaced horizontal lines will intersect a well-defined QRS complex at most six times (see Fig.(1a), while this number can be substantially larger for noisy signals (see Fig.(1b)). We thus impose the stringent criterion that for *every* "QRS peak" sandwiched between two consecutive "BP peaks", the maximum number of intersections after applying the "horizontal line check" cannot exceed *six* times. If this criterion is satisfied, all these "QRS peaks" are identified as real QRS peaks.

## 4. Results and conclusions

Our algorithm has a sensitivity score of 99.9% and predictivity score of 99.96% for the 100 training sets. For the Phase III of the CinC Challenge 2014, which includes all the data sets from Phase I and Phase II, our algorithm has a gross sensitivity score of 87.8% and gross predictivity score of 85.15%, with an overall score of 86.73.

Our algorithm depends highly on the performance of *wabp* and especially that of *gqrs*. We generally assume that while *gqrs* is susceptible to false positives due to artifacts, it rarely produce false negatives; thus our algorithm tends to miss a heartbeat when *gqrs* does produce false negatives.

The performance of our algorithm can be improved when a better version of *gqrs* and *wabp* is available, or when a sensible way to pre-treat ECG and BP signals can be applied. Other physiological signals including electromyograph (EMG) and electroencephalograph (EEG) can also be helpful for QRS peak prediction. Both signals are electrical and correlates positively with the ECG signals. While EMG and EEG signals tend to have small signal-to-noise ratio, with proper processing they can potentially be useful especially when both ECG and BP signals are noisy.

## References


[1] O. Sayadi and M. B. Shamsollahi. A model-based bayesian framework for ecg beat segmentation. Physiol Meas, 30(2009), 335-352.

[2] J. Willis Hurst. Interpreting Electrocardiograms. Marcel Dekker, 2001.

[3] N.M. Arzeno, Zhi-De Deng, and Chi-Sang Poon. Analysis of rst-derivative based QRS detection algorithms. 55(2), 2008:478-484.

[4] Ivaylo I. Christov. Real time electrocardiogram qrs detection using combined adaptive threshold. Biomed Eng Online, 3(1), 2004:28.

[5] Ivan A. Dotsinsky and Todor V. Stoyanov. Ventricular beat detection in single channel electrocardiograms. Biomed Eng Online, 2008, 3:3.

[6] Peng Qiu and K.J.R. Liu. A robust method for QRS detection based on modi
ed p-spectrum. In Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on, pages 501-504, 2008.

[7] F. Portet, A.I. Hernandez and G. Carrault, Evaluation of real-time qrs detection algorithms in variable contexts. Med Biol Eng Comput, 43(3), 2005:379-385.

[8] R. Poli, S. Cagnoni, and G. Valli. Genetic design of optimum linear and nonlinear QRS detectors. 42(11), 1995:1137-1141.

[9] Jiapu Pan and Willis J. Tompkins. A real-time QRS detection algorithm. BME-32(3), 1985:230-236.

[10] http://physionet.org/physiotools/wfdb.shtml



Address for correspondence:

Bo Yang
Complex Systems Group
Institute of High Performance Computing
1 Fusionopolis Way #16-16
A*STAR Singapore, 138632
tel./fax: ++(65)64191111/(65)64630200
yangbo@ihpc.a-star.edu.sg