Rhythm-based Accuracy Improvement of Heart Beat Detection Algorithms

Zoltán Gilián[1], Péter Kovács[1], Kaveh Samiee[2]

[1]Department of Numerical Analysis, Eötvös Loránd University, Budapest, Hungary
[2]Department of Signal Processing, Tampere University of Technology, Tampere, Finland

Abstract

Aims: Our aim is to improve the accuracy of existing heart beat detection algorithms in order to provide reliable heart beat locations in a multi-modal beat detection scheme.

Methods: A rhythm-based algorithm is presented which on top of a base beat detection method processes the detected beats by rejecting annotations and filling in gaps while minimizing a deviation score. A novel beat detection method based on rational modelling of ECG signals is also presented as a base algorithm.

Results: The rhythm-correction algorithm applied to Sachin Vernekar's phase II entry was submitted to the third phase of the PhysioNet/CinC Challenge 2014 contest. The algorithm has 99.98% gross and average sensitivity and 99.96% gross and average positive predictivity compared to 99.92% and 99.94%, respectively, of the base algorithm. Due to run-time performance problems, the rational algorithm was not able to qualify in the contest.

Conclusions: The rhythm-based method improves the results of the base algorithm on the training data set. The hidden records are not yet available at the time of writing of this paper; therefore we are not able to report the final performance of the algorithm. Run-time improvement of the rational algorithm remains future work.

1. Introduction

Heart beat detection algorithms using local, per-beat information as a basis of detection are widely used, as there is a clear performance advantage in computing features based on local signal data. These methods are prone to miss or misplace beats due to noisy data even in cases where a human inspector could correct the mistakes with ease based on neighbouring beats and rhythm. Figure 1 shows such a missed beat. In Figure 2 instead of two in-rhythm beats a misplaced one is detected. The figures 2 and 3 are generated using Sachin Vernekar's second phase solution on the record 133 of the challenge training data set.

Here we propose a method to correct these types of errors as a post-processing step built on top of an arbitrary base algorithm.

The paper is organised as follows. Section 2 introduces the rhythm-based post-processing method. In addition, three particular base algorithms are considered. Section 3 contains detailed information on a particular base algorithm which uses rational modelling of ECG signals. In Section 4 the performance of the rhythm correction step on top of each of the base algorithms are compared. Final conclusions are drawn in Section 5. Section 6 contains a link to the source code of the algorithm.

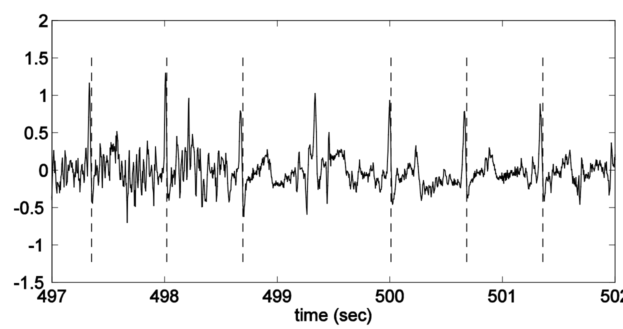

Figure 1. Missed beat.

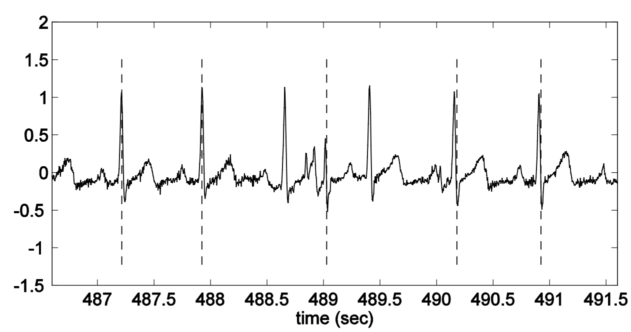

Figure 2. Misplaced and missed beats.

2. Method

The proposed method tries to correct missed and misplaced beats in a set of beat locations found by a base algorithm. Furthermore, a set of candidate locations are considered. The candidate locations influence the performance of the rhythm-based algorithm greatly, since any applied correction originates from the candidates.

The algorithm iterates through the base locations. In each iteration the next location is selected, and a few base locations after the selected one are considered for replacement. These locations are called skipped. The algorithm tries to skip zero to N locations, and perform the best replacement in a sense described later.

The interval defined by the active location and the first non-skipped one is searched for candidates. According to the contest rules the tolerable time difference between a detected and a reference beat location is at most 150 ms. Therefore, no candidates closer than 150 ms to a non-skipped location are selected. The maximum number of selected candidates is restricted by the constant M. Weight values assigned to the candidates determine which ones to choose, supposing the allowed number of candidates is exceeded. Base algorithms using machine learning techniques are able to provide such weights naturally.

All subsets of the selected candidates are evaluated to possibly replace the skipped locations. The basis of evaluation is a score function which measures how well, in terms of rhythm, the candidates fit into the sequence of locations in a sliding window preceding and following the skipped ones.

There is a trade-off between run-time performance and error detecting capabilities when choosing the constants N and M. Based on experimental data, the constants are set to $N = M = 4$.

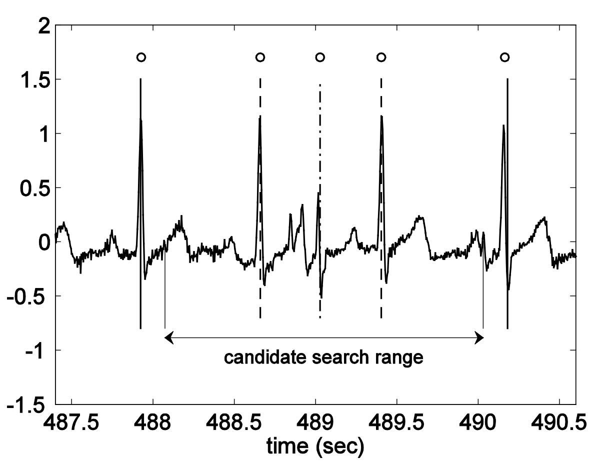

Figure 3. Overview of the rhythm-based algorithm. The dash-dotted location in the middle is being skipped and replaced by the dashed candidates found in the candidate search range. All subsets of the candidates in the search range are examined, and the ones with lowest score are selected. Since the dashed locations are more in-rhythm, the algorithm chooses them instead of the dash-dotted location. The small circles mark the location of the candidates.

Figure 3 illustrates the algorithm on the misplaced location of Figure 2.

In the following subsections the choice behind the particular elements of the algorithm is discussed, namely the base algorithm, candidate generation, candidate weighting and score function.

2.1. Base algorithm and candidates

During the course of the PhysioNet/CinC challenge 2014 contest we considered three particular base algorithms. However any beat detection algorithm can be augmented by the rhythm-based one.

Initially the sample entry was used as a base algorithm. The candidates were chosen by merging the output annotations of the gqrs and wabp functions of the WFDB library [1]. Based on the sample entry, blood pressure annotations were time-shifted by 200 ms. The weights were set to a constant value.

Our second approach was to try a novel beat detection algorithm using rational modelling of ECG signals [2] [3] [4] instead of gqrs. The algorithm uses support vector machine classification which provides weights to the algorithm. The candidates were chosen by lowering the acceptance threshold of the SVM classifier and merging the results with the time-shifted output of wabp if a blood pressure signal is present. This approach had run-time performance problems and couldn't finish in time to be qualified in the contest.

The third version which was submitted to the contest uses the phase II solution of Sachin Vernekar. The algorithm uses deep learning on top features extracted from the ECG and blood pressure signals to find beats. The candidates were chosen by finding sufficiently high peaks in the estimated probabilities. The probabilities were also used as weights.

2.2. Score function

The score function measures the regularity of the given beat locations. The smaller the score, the more in-rhythm the locations are. The rhythm-based algorithm tries to minimize this function.

Here, two score functions are considered. Initially the standard deviation of the beat-to-beat intervals was used. Since standard deviation measures deviation from the mean, missed or inserted beat locations can introduce outliers to the beat-to-beat intervals, causing the mean to shift, erroneously penalizing the regular beat-to-beat intervals. To this end an alternative score function is applied. The score function uses the unbiased standard deviation formula with the mean replaced by the median. Given a sequence of beat-to-beat intervals $d_1, \ldots, d_n$, the modified score function is

$$\text{score}(d) = \sqrt{\frac{1}{n-1}\sum_{k=1}^{n}\left(d_k - \text{median}(d)\right)^2}.$$

Although no significant performance difference can be measured on the training data set (see Table 1), the results on the hidden data set should be examined before drawing a final conclusion, once the data becomes available.

3. QRS detection based on rational modelling of ECG signals

In Section 2.1 three different base algorithms were introduced. This section describes the second one based on rational modelling of ECG signals briefly.

Let's consider the sequence $a_0, a_1, \dots, a_n, \dots$ of complex numbers in the open unit disc, i.e. $a_n \in \mathbb{C}, |a_n| < 1$ $(n \in \mathbb{N})$. The rational functions

$$\Phi_n(z) = \frac{\sqrt{1-|a_n|^2}}{1-\bar{a}_n z}\prod_{j=0}^{n-1}\frac{z-a_n}{1-\bar{a}_n z}$$

are called Malmquist-Takenaka functions (see e.g. [5] [6]). These functions form an orthogonal system with respect to the scalar product

$$\langle f, g\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{it})\overline{g(e^{it})}\,dt$$

where the functions f and g are square integrable on the complex unit circle.

QRS complexes of ECG signals can be approximated using Fourier partial sums with respect to the system (Φ_n, $n \in \mathbb{N}$):

$$\tilde{f}(z) = \sum_{n=0}^{N}\langle f, \Phi_n\rangle\,\Phi_n(z)$$

where the complex unit circle is parameterized by $z = e^{it}$; $t \in [-\pi, \pi]$ is the normalized time variable. Figure 5 shows a concrete example of QRS complexes approximated by rational functions.

The parameters $(a_0, a_1, \dots, a_N)$ make the model very general and adaptable. The parameters most suitable to the signal are found using numerical optimization. To keep the implementation simple and the run-time performance manageable, the parameters are constrained as follows: let $N = 2$, $a_0 = 0$ and $a_1 = a_2 = a$. This way the numerical optimizer only has to find a.

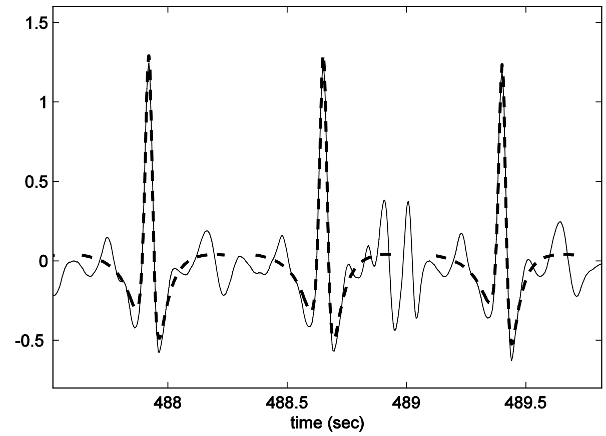

Figure 5. Rational approximation of QRS complexes. The dashed curves are graphs of Malmquist-Takenaka partial sums.

We used the Rational Interpolation and Approximation Toolbox (RAIT) [7], which is an open source toolbox for Matlab and Octave, for computing the partial sums and optimizing the model parameter.

The rational modelling method introduced above is applied to perform ECG-based beat detection. The peaks of the ECG signal are selected, and the ECG signal in a small neighbourhood of each peak is approximated using Malmquist-Takenaka partial sums. The model parameter a and the coefficients $\langle f, \Phi_n\rangle$ form a descriptor of the peak, and are used as features by an SVM classifier to decide if the location is indeed a true beat. To ensure this description captures only local information, the ECG signal is multiplied by a window function before rational modelling.

The training of the classifier was performed as follows. The records of the challenge training data set was divided into 60% training set, 20% cross-validation set and 20% test set. A set of feature locations was selected for each record containing the reference annotations and an equal number of random locations. All training and validation were based on the features of these locations. Model selection was performed by minimization of the error on the cross validation set.

We used the libsvm open source library [8] to train the classifier and to perform predictions.

4. Results

The algorithm submitted to the third phase of the PhysioNet/CinC Challenge 2014 contest applies the rhythm correction step to the solution using deep learning submitted to the second phase of the contest by Sachin Vernekar. The entry achieved 78.17% and 89.19% gross and average sensitivity, furthermore 74.43% and 71.60% gross and average positive predictivity respectively.

The rhythm correction algorithm was improved since the submission of the aforementioned entry. The results presented here reflect the performance of the improved one. Since the hidden data set is not publicly available at the time of writing of this paper, these results consist of measurements taken only on the training data set.

The performance of the base algorithm without rhythm correction and the results of the rhythm-corrected ones equipped with each score function discussed in Section 2.2 are compared in Table 1.

Table 1. The table summarizes the experimental results. Columns correspond to the base algorithm. The rows contain the number of false positive (FP) and false negative (FN) outcomes of the base algorithm without and with rhythm-based correction using standard deviation and the modified score function. The data set has 72413 beat locations in total.

		Sample entry	Rational	Deep learning
Base algorithm	FP	71	46	41
	FN	56	210	60
Std score	FP	18	9	22
	FN	60	107	14
Modified score	FP	22	9	22
	FN	83	111	12

In case of the sample entry rhythm correction can eliminate about 75% of false positive beats; however the number of false negatives is increased. The standard deviation score performs considerably better here.

The algorithm based on rational functions clearly benefits from post processing step; the number of false positives and false negatives are reduced by 80% and nearly 50% respectively.

In case Sachin Vernekar's algorithm, the post-processing step corrects close to 50% of false positives and about 80% of false negatives.

5. Conclusion

The rhythm correction post processing step was able to increase detection accuracy of other beat detection algorithms by considering rhythm. The performance of the algorithm depends heavily on the base algorithm and the candidate generation method, therefore these components should be chosen with care.

Due to the nature of the algorithm performance can severely decrease in the presence of false in-rhythm candidates. E.g. in case of arrhythmia, robust candidates are required to maintain good results.

A novel beat detection method based on rational modelling of QRS complexes is presented. The issue with run-time performance should be addressed in the future.

6. Source code

The source code of the algorithm can be downloaded from gilian.web.elte.hu/cinc2014/rhythm.tar.gz.

References

[1] Moody GB. WFDB Application's Guide. 10th ed. Harvard-MIT Division of Health Sciences and Technology. 2014;30-31,105.

[2] Fridli S, Lócsi L, Schipp F. Rational function systems in ECG processing. In Moreno-Díaz R, Pichler F, Quesada-Arencibia A, editors. Computer Aided Systems Theory - EUROCAST 2011, Lecture Notes in Computer Science. Springer Berlin Heidelberg. 2011;6927:88-95.

[3] Fridli S, Kovács P, Lócsi L, Schipp F. Rational modeling of multi-lead QRS complexes in ECG signals. Annales Univ. Sci. Budapest, Sect. Comp. 2012;36:145-155.

[4] Gilián Z. ECG-based heart beat detection using rational functions. In Conference on Dyadic Analysis and Related Fields with Applications. 2014;13.

[5] Heuberger PS, Van den Hof PM. Modelling and identification with rational orthogonal basis functions. Springer London. 2005.

[6] Pap M, Schipp F. Malmquist-Takenaka systems and equilibrium conditions. Mathematica Pannonica. 2011;12:185-194.

[7] Kovacs P, Locsi L. RAIT: The rational approximation and interpolation toolbox for Matlab. 35th International Conference on Telecommunications and Signal Processing. 2012;671-677.

[8] Chang CC, Lin CJ. LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology (TIST). 2011;2(3):27.

Addresses for correspondence.

Zoltán Gilián
Department of Numerical Analysis,
Eötvös Loránd University,
1/C Pázmány Péter Sétány, 1117 Budapest, Hungary
gilian@caesar.elte.hu

Péter Kovács
Department of Numerical Analysis,
Eötvös Loránd University,
1/C Pázmány Péter Sétány, 1117 Budapest, Hungary
kovika@inf.elte.hu

Kaveh Samiee
Department of Signal Processing
Tampere University of Technology
P.O.Box 553, FI-33101 Tampere, Finland
kaveh.samiee@tut.fi