A Noise Robust QRS Delineation Method Based on Path Simplification

T. Teijeiro, P. Félix and J. Presedo

Centro Singular de Investigación en Tecnoloxías da Información (CITIUS), University of Santiago de Compostela, Santiago de Compostela, Spain

Abstract

In this work we present a new method for the delineation of QRS complexes in ECG signals. The main objective is to provide a robust method that gives acceptable results even under severe noise conditions in the input signal, as often happens in continuous bedside or home monitoring. Our method is based on relevant point selection using a path simplification algorithm, and then a clustering strategy combined with a qualitative description of the waveform is applied in order to select the most promising signal segments to include inside the QRS limits. The validation was performed by adding different levels of noise to the records of a standard database, and then comparing our proposal with a state-of-the art approach. Results show a high sensitivity and stable error evolution even at the highest noise levels.

1. Introduction

ECG is one of the best known and most studied biosignals in the field of intelligent signal processing. There are a large number of databases, algorithms and methods covering from simple heart rate measurement to the characterization of a wide range of diseases, like myocardial ischemia, sleep apnea, or chronic obstructive pulmonary disease, among others [1, 2]. Despite new proposals appear incessantly, computer interpretation of the ECG is still considered an open problem, and it seems we are far from provide sufficiently satisfactory solutions to be transferred to clinical routine, integrated in the bedside instrumentation or in home monitoring settings. A clear proof of this is the topic selected for recent Physionet challenges in 2014 and 2015 [3, 4], encouraging the exploration of methods for biosignal processing in poorly controlled scenarios. The posed problems are relatively simple and well studied, but when existing solutions are transferred from training databases to more demanding scenarios, the number of errors and false alarms is considerably increased.

This work addresses the QRS delineation problem, that is, the identification of the onset and end boundaries of a detected QRS complex. A proper QRS delineation makes easier the identification of heartbeats origin, and therefore empowers arrhythmia characterization and higher level ECG analysis. It also may improve the specificity of QRS detection algorithms by enforcing additional constraints on the QRS waveform.

This problem has been addressed by various studies with quite remarkable results. For example, a number of relevant proposals are reviewed in [5], showing a great precision with errors in the order of a few milliseconds, even exceeding the capacity of human experts. However, if we try to apply these methods in scenarios presenting a low signal quality compared with the reference databases, the performance is greatly degraded. Therefore, our approach poses a bit different objective: improve the stability against high noise levels, as long as keeping the delineation error within acceptable limits.

The rest of the paper is organized as follows: In section 2 we detail the proposed algorithm and the intuition behind it; section 3 describes the evaluation process, including the generation of a test database and comparing the results with a well known algorithm; finally, section 4 provides some conclusions.

2. Algorithm description

The delineation algorithm we propose, that we will call *qrsdel*, is a multi-lead procedure decomposed in various stages, depicted in figure 1.

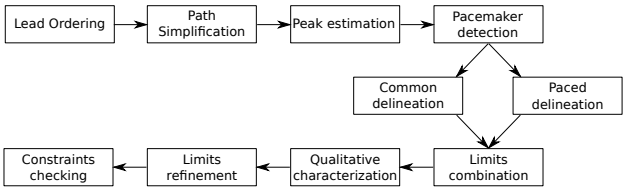

Figure 1. Delineation algorithm steps

As a preliminary step, a temporal window is defined from a punctual beat annotation, that may be obtained by any QRS detection procedure. This window starts 80 ms before the annotation and finishes 200 ms after it, determining the signal fragment used by the subsequent steps. Note that we assume the annotation to be located near the beginning of the QRS complex, as is usual in most of the state of the art QRS detection algorithms, such as those included in the WFDB toolbox [2].

2.1. Lead ordering

The first step of the algorithm is to take all signal fragments from every available ECG lead and sort them using the kurtosis as a signal quality indicator [6]. This ordering will determine the priority in the selection of the delineation points in later stages.

2.2. Path simplification

Path simplification [7] is the underpinning technique of *qrsdel*. The basic idea is to reduce the signal fragment containing the QRS complex to a small number of straight line segments, while preserving the basic waveform of the complex. The assumption this model makes is that, despite how noisy the signal is, the QRS complex is the most prominent element in the region surrounding it.

Figure 2 illustrates this idea. We can see the original signal in light gray color, the s_{k_i} points selected by the path simplification algorithm, and the signal reconstructed using only that points. The solid and dashed segments correspond with the delineation of the QRS complex.

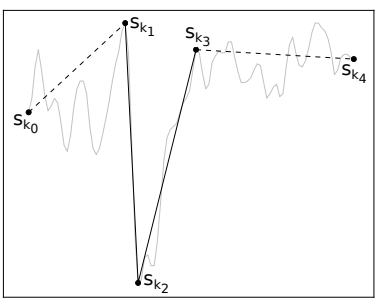

Figure 2. QRS signal and its simplification

The algorithm we selected to perform signal simplification is a minor variation of the Douglas-Peucker method [7], considering only amplitude differences for point selection. The input for this algorithm is a signal fragment s as a sequence of n samples $(s_0, \ldots, s_{n-1})$; an integer p, with $2 \leq p \leq n$, representing the maximum number of points that can be included in the simplification; and a value $d > 0$ representing the minimum amplitude difference that can be considered to include a point in the simplification. As output, the algorithm returns a sequence $rp = (k_0, \ldots, k_{m-1})$, with $2 \leq m \leq p$, representing the indices of the samples selected for the simplification. The operation is detailed in algorithm 1.

In this algorithm, the simplification always contains the endpoints 0 and $n-1$, and it is extended by adding at each step the point with maximum vertical distance to the segment defined by the linear interpolation (INTERP function) between any consecutive pair of points already included in the simplification. The procedure finishes when the maximum number of points is reached, or if none of the points exceeds the minimum amplitude difference.

Algorithm 1 Douglas-Peucker simplification algorithm

function DOUGLAS-PEUCKER(s, p, d)
  let $rp = (0, n-1)$
  let $s' = \text{INTERP}(s_0, s_{n-1}, n)$
  let $M, k = max(|s_i - s'_i|), i \in [0, n-1]$
  while $|rp| < p \wedge M > d$ **do**
    $rp = rp \cup \{k\}$
    $M = 0$
    for all $j \in [0, \ldots, |rp| - 2]$ **do**
      $s' = \text{INTERP}(s_{k_j}, s_{k_{j+1}}, k_{j+1} - k_j)$
      $M', k' = max(|s_i - s'_{i-k_j}|), i \in [k_j, k_{j+1}]$
      if $M' > M$ **then**
        $M, k = M', k'$
  return rp

In the *qrsdel* procedure, the DOUGLAS-PEUCKER algorithm is executed on each lead using as parameters $p = 9$ and $d = 50\mu V$. The value of p is determined by the maximum complexity of the waveform that can be recognized (see for example the *qrSRS* pattern in figure 3), while d has been selected as the minimum peak-to-peak amplitude distance for a wave to have diagnostic implications in routine electrocardiography [8].

2.3. Peak estimation

The next step is to estimate the temporal location of the peak of the QRS complex. In each lead, the peak is selected according to the following equation:

$$max_r \left(\frac{|s_r - b|}{1 + \frac{2 \cdot |r-t|}{150}} \right), r \in [0, \ldots, n-1]$$

where t is the time point of the annotation in ms, and b is the baseline value, estimated as the mode of the signal in a temporal window of 1 second around the time point of the annotation. This equation selects the point with highest deviation from the baseline, but it applies a distance function with respect to the beat annotation. Initially, the peak is selected from the best quality lead, but if the peaks in other leads are within a 40 ms margin, the earliest one is taken as reference.

2.4. Delineation and pacemaker detection

At this point, QRS delineation itself is performed. To this aim, the segments resulting from path simplification are clustered using a *k-means* strategy, using the following two features extracted for each segment $(s_{k_i}, s_{k_{i+1}})$:

$$f_1 = ln\left(1 + \left|\frac{(s_{k_{i+1}} - s_{k_i})}{(k_{i+1} - k_i)}\right|\right), \; f_2 = \begin{cases} r - k_i, & k_i < r \\ k_{i+1} - r, & k_i \geq r \end{cases}$$

The f_1 feature represents the slope of each segment, while f_2 represents the distance to the peak estimation r. The *k-means* procedure is executed taking as centroids $c_1 = (max(f_1), min(f_2))$ and $c_2 = (min(f_1), max(f_2))$, and the limits of the QRS are set as the limits of the longest sequence of segments in the c_1 cluster, requiring the presence of at least one slope change. Figure 2 shows a delineation example, with segments $\{(s_{k_1}, s_{k_2}), (s_{k_2}, s_{k_3})\} \in c_1$, and $\{(s_{k_0}, s_{k_1}), (s_{k_3}, s_{k_4})\} \in c_2$.

The selection of the slope as a key feature causes that the presence of pacemaker spikes, which usually show extreme slopes, strongly affects the clustering results. For this reason, we include a prior stage to determine the presence of a paced QRS complex, by looking for a signal wave satisfying the following constraints: (1) a duration smaller than 30 ms, (2) ascent and descent angles higher than 75°, (3) an amplitude higher than 0.2mV, and (4) a rising edge of higher amplitude than the falling edge. If all these conditions are satisfied, we assume the presence of a paced beat, and the clustering results are only used to determine the end of the QRS complex, setting as its onset the beginning of the spike.

2.5. Limits combination and waveform characterization

At this point, we have the limits estimation for the QRS complex in every available lead. The next step then is to combine them, using the same strategy we followed to set the peak estimation. Initially, the endpoints are considered only from the highest quality lead, and then are expanded using the endpoints in other leads only if the difference with the current estimation is lower than 40 ms. If we have detected a pacemaker spike in any lead, then the beginning of the QRS is always the beginning of the spike.

The next step consists of performing a qualitative characterization of the waveform delineated in each lead. For this, we consider the line segments as a sequence of waves, and each wave is labeled with the common names *q, Q, r, R, S*, etc. based on its position and positive/negative polarity [8]. Then, the longest sequence with a recognizable tag is selected for each lead, and the delineation limits are refined to fit these limits. Figure 3 shows some examples of recognized waveforms, from a total of 27.

Finally, to accept the delineation as valid, we check some common electrocardiographic constraints. Specifically, we require that at least in one lead we have identified a recognizable QRS pattern with an amplitude between 0.5 mV and 6 mV. If this requirement cannot be satisfied, then we consider sufficient to have an identified pattern in most leads. If none of this conditions can be met, then we discard the annotation, considering it as a false positive.

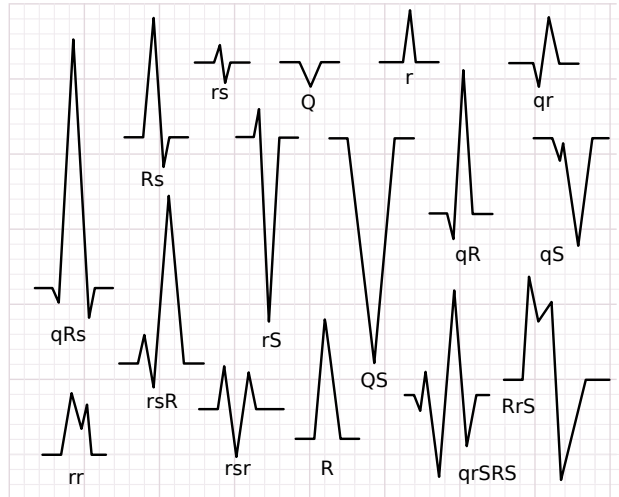

Figure 3. Some recognized QRS qualitative tags

3. Evaluation

The gold standard for the evaluation of ECG delineation algorithms is the QT database [9]. However, the set of records included in this database are more focused on the evaluation of the precision and the capability of the algorithms to deal with a wide variety of waveforms, and the signal quality is usually high. As long as our main objective is to assess the robustness of *qrsdel*, we have generated a new database by adding different noise levels to the original records of the QT database, and then we compared the behaviour of our algorithm with the results of the *ecgpuwave* application [2], since it is the only reference algorithm we are able to test with the new generated records.

To generate the test database, we have relied on the *nst* application [2], using the noise model recommendations from [10]. For each record in the QT database we created 5 new records by adding electrode motion noise at different Signal-to-Noise ratios. Selected SNRs were 24 dB, 12 dB, 6 dB, 3 dB and 0 dB.

Table 1 shows the sensitivity and measurement errors for the *qrsdel* and *ecgpuwave* algorithms at different noise levels. Since *ecgpuwave* is a single-lead algorithm, we calculated the errors using all the measurements that can be obtained in the available leads. Figure 4 illustrates these results, allowing to appreciate the different behaviour of both algorithms as noise level increases. Light-grey shaded regions represent the standard deviation of *ecgpuwave*, while the dark regions represent the standard deviation of *qrsdel*. Even if *ecgpuwave* shows lower errors at reduced noise levels, it can clearly be seen that the error behaviour is far more stable for the *qrsdel* algorithm, both for the mean and for the standard deviation. It is also worth noting the evolution of the sensitivity, showing a significant degradation in the case of *ecgpuwave*, while in *qrsdel* it is barely reduced

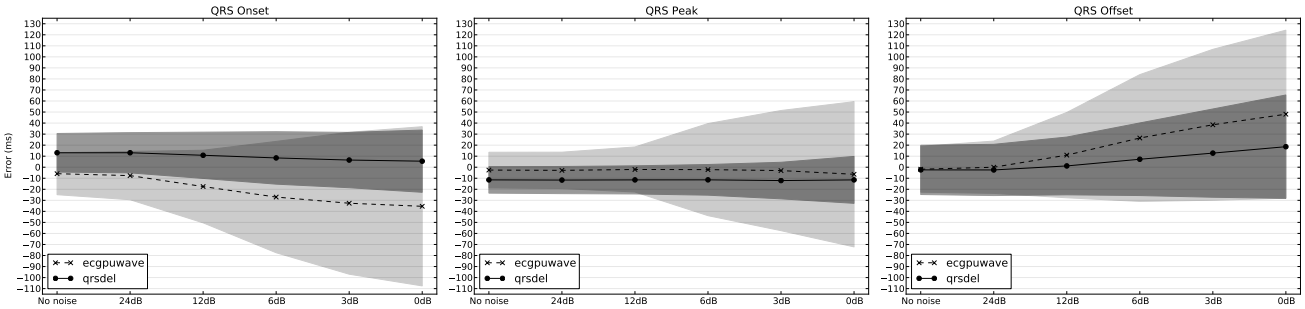

Figure 4. Error comparison for the *qrsdel* and *ecgpuwave* algorithms

SNR		Se	QRS Onset	QRS Peak	QRS Offset
Original	*QD*	100	13.09±17.48	-11.47±12.01	-2.41±22.34
	EC	96	-5.87±19.03	-2.59±16.16	-1.83±20.80
24 dB	*QD*	100	13.12±18.24	-11.64±12.30	-2.47±23.21
	EC	96	-7.60±21.85	-2.84±16.47	-0.01±23.70
12 dB	*QD*	100	10.82±20.96	-11.48±12.75	1.17±26.16
	EC	95	-17.51±32.89	-2.04±20.41	10.96±38.66
6 dB	*QD*	100	8.42±23.76	-11.44±13.89	7.22±32.89
	EC	91	-27.05±50.44	-2.17±41.73	26.46±57.45
3 dB	*QD*	100	6.50±25.07	-12.09±16.56	12.78±39.98
	EC	88	-32.64±64.31	-3.05±54.44	38.45±68.36
0 dB	*QD*	99	5.47±28.19	-11.48±21.24	18.60±46.85
	EC	85	-35.41±72.14	-6.40±65.80	48.04±76.14

Table 1. Sensitivity (%) and delineation errors (ms) for the *qrsdel (QD)* and *ecgpuwave (EC)* algorithms.

from 100% even at the 0 dB SNR noise level.

Regarding the accuracy of the results, although it is not a primary goal of this work, it is necessary to ensure that the errors remain within acceptable limits. In this case, the majority of errors and standard deviations stay within 0.5 mm in common electrocardiographic scale (20 ms), and only for the QRS offset and with a SNR of 0 dB the standard deviation slightly exceeds the range of 1 mm (40 ms). We consider these margins admissible to provide a coarse approximation of the QRS waveform in noisy conditions.

4. Conclusions

In this work we presented *qrsdel*, a new algorithm for robust delineation of QRS complexes in multi-lead ECG signals based on a path simplification technique, and we compared it with a state-of-the art algorithm, proving a more stable behaviour against signals with high noise levels. We believe this proposal can be very valuable in poorly controlled monitoring scenarios, like bedside or home monitoring, and therefore we have released it as free software under the terms of the LGPLv3 license[1]. We also published the source code needed to generate the test database and replicate the experiments shown in this paper.

[1] `https://github.com/citiususc/qrsdel`

Acknowledgements

T. Teijeiro is funded by an FPU grant from the Spanish Ministry of Education (MEC) (ref. AP2010-1012).

References

[1] Gacek A, Pedrycz W. ECG Signal Processing, Classification and Interpretation: A Comprehensive Framework of Computational Intelligence. Springer Science & Business Media, 2012.

[2] Goldberger A, et al. PhysioBank, PhysioToolkit and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation 2000;101:215–220.

[3] Moody G, Moody B, Silva I. Robust Detection of Heart Beats in Multimodal Data: the PhysioNet / Computing in Cardiology Challenge, 2014.

[4] Silva I. Reducing False Arrhythmia Alarms in the ICU: The PhysioNet / Computing in Cardiology Challenge, 2015.

[5] Ghaffari A, Homaeinezhad MR, Akraminia M, Atarod M, Daevaeiha M. A robust wavelet-based multi-lead Electrocardiogram delineation algorithm. Medical Engineering Physics 2009;31(10):1219–1227.

[6] Li Q, Mark RG, Clifford GD. Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter. Physiological Measurement 2008;29(1):15–32.

[7] Douglas D, Peucker T. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. The International Journal for Geographic Information and Geovisualization 1973;10(2):112–122.

[8] CSE WP. Recommendations for measurement standards in quantitative electrocardiography. European Heart Journal 1985;6(10):815–825.

[9] Laguna P, Mark R, Goldberg A, Moody G. A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG. In Computers in Cardiology. 1997; 673–676.

[10] Moody G, et al. A noise stress test for arrhythmia detectors. Computers in cardiology 1984;11:381–384.

Address for correspondence:

Tomás Teijeiro Campo
tomas.teijeiro@usc.es