QRS Delineation Algorithms Comparison and Model Fine Tuning for Automatic Clinical Classification

Antonio Casañez-Ventura[1], Francisco-Javier Gimeno-Blanes[1], José Luis Rojo-Álvarez[2], José-Antonio Flores-Yepes[1], Juan-Ramón Gimeno-Blanes[3], José-María López-Ayala[3], Arcadi García-Alberola[3]

[1]Universitat Miguel Hernández, Elche, Spain
[2]Universidad Rey Juan Carlos, Fuenlabrada, Spain
[3]Hospital Clínico Universitario Virgen de la Arrixaca, Murcia, Spain

Abstract

QRS complex extraction and wave detection have been paid intense efforts in scientific publications in the last two decades. This work elaborates on different QRS delineation algorithms for classification and for diagnostic indexing. A subset of 150 cases were randomly selected from a full database including over 3500 consecutive ECGs. Three QRS detection methods were implemented and later benchmarked against the information provided by a GE MAC 5000 ECG system, and also against a Gold-Standard manually and carefully developed by clinicians. All implemented methods were applied to the complete ECG signal and to a consolidated single ECG beat template. Better performance was obtained using beat template signals, due to denoising effects. The absolute error of the QRS duration was chosen as a figure of merit. Results showed that all developed methods outperformed information provided by the ECG device, when compared to Gold-Standard: 7,90 ± 6,83 ms of QRS duration for Physionet Method, 8,31 ± 3,07 ms for Chouhan Method, 6,27 ± 4,77 ms for an add-hoc two stages developed method, and 8,63 ± 5,89 ms for the GE device. Individual methods very much rely on one single measurement that does not easily match clinician's criteria. A two stage strategy, with first a initial candidates pre-selection, overcoming ECG local singularities, following by a fist and second momentum analysis, provided a better fit to Gold-Standard.

1. Introduction

QRS complex extraction and delineation have been a constant effort in scientific field in the last two decades [1, 2]. This activity, which can be monotonous and effort intensive, has motivated the development of a number of models to be tested for automatic detection and/or measurement. In particular, the beginning and end of QRS complex, that provides relevant information from a diagnostic perspective [3, 4, 5, 6], has been an issue in this area, as there are no unified criteria, nor formulated mathematical method to identify them univocally. These circumstances have led to the development of an important number of studies. As a singular reference only in the last 15 years it has published over 4,000 articles containing the keywords "QRS detect" in the SCOPUS database, and this has been a growing effort over the years. Other ECG waves and fiducial points such as the P and T waves, present important difficulties for correct detection. It should be mentioned here that in general terms special difficulty in the ECG delineation process is present when trying to find starting (OnSet) and finishing (OffSet) points of any singular wave, such us both limits of P and T waves, or initial and end of the QRS complex. The difficulty in choosing such points or boundaries of these waves lies in the lack of definition of morphological and mathematical characteristic that allow an irrefutable allocation.

It is relevant to mention that any automated model to be applied to real signals will require an important preprocessing for noise reduction in order to be able to analyze existing information from a computerized perspective. In the ECG signal, this could be related to muscles movements, electrodes contacts, breathing, power supply electrical cycle, and glitch, among others. As a consequence, we could state that feature extraction in a real situation is not easy, because the physiological variations from patient and/or disease, together with the different sources of noise, make ECG a non-stationary signal. By the same token, detection, modeling and processing, become relevant topics towards the implementation of systems where the ECG is an important input for clinical diagnosis and especially for automated or semi-automated indexing and analysis. For that reason, clinicians from the HCUVA (Hospital Clínico Universitario Virgen de la Arrixaca of Murcia) motivated this initiative to set automatic or semi-automatic ECG measurements in order to classify and categorize risk.

The objective of this work is to test ECG delineation algorithms to grade them as how closely them match the Gold Standard set by the cardiologists. In additional, we wanted to clearly identify, if any, the model that best fit clinicians' criteria. Many algorithms are described in literature (i.e. [7, 8, 9, 10, 11]). In this work we compared algorithms and results from some acknowledged methods in literature, against the information provided from the ECG system itself and a Gold-Standard developed manually by cardiologists. We also propose a first attempt to improve results by a combined two-stage approach to this very well known problem.

A key contribution of this paper is the benefit of replacing the (always complex) filtering process, that may eliminate relevant information from the signal, with a averaging ECG filtering, as a natural-way of filtering the signal without damaging relevant information while allowing better detection of the fiducial points. Secondly, the use of a hybrid method that best suits cardiologist's requirements, as the best method to fit the manually developed Gold-Standard.

2. Methods

As mentioned in the previous section, an important number of models are been proposed in literature for QRS delineation. They could be classified as follows: (a) methods based on digital filtering; (b) methods based on wavelet transform; (c) methods based on machine learning; and (d) methods based on different time-frequency analysis. The present paper analyzed models and techniques almost covering the full set of those groups, namely, Biosig [11], Physionet [12], Chouhan [9, 10], and other ad-hoc model specially tailored to fit into clinician's approaches. As an exception, time-frequency or time-scale models were not included in this analysis, and they will be addressed in future developments. The Biosig model did not provide any result in a relevant number of our non-synthetic signals used for this evaluation, and hence, it was not possible to statistically benchmarking it. This was also the case for some other in-house developed models, which were quite promising from an academic perspective but did no show good performance when applied to non-synthetic signals.

2.1. Databases

In this work, we used a database of 3,500 ECG taken at Hospital Universitario Virgen de la Arreixaca from Murcia (HUVAM). For algorithm validation and comparisons, a subset of 150 ECG were benchmarked to both, the Gold-Standard developed by cardiologists for every single beat of every lead, and information from GE MAC 5000 ECG System, were signals were registered.

2.2. Preprocessing

Preprocessing has been always an issue for biological signals analysis. This fact conducted us to elaborate on the idea of developing a "natural" denoising and preprocessing of the signal will allow us to remove noise without eliminating real waves that were not part of formally standarized synthetic signals, but they might have interest from a clinical perspective, affecting to automatic measurements (among others). So, we firstly hypothesed that a severe preprocessing and filtering could damage, or even totally remove, relevant ECG information. Secondly, we also considered the possibility that the analysis of every single beat of a fully preprocessed complete signal for a later averaging measurement could incorporate bias from: (i) an eventual over-preprocessing; (ii) averaging of morphological deficiencies in any lead or beat; (iii) ectopic beats considered as normal beats. From a practical perspective, we noticed that clinician's visual inspection of the automatic measurement applied over the whole signal, even after several fine tunings of free parameters and standard filtering, led to a total discard of presented models.

2.3. Single beat template

In order to overcome the previously mentioned points, we applied a preprocessing method based on: (a) applying soft filtering (low pass filter of order 5 and a cut-off frequency of 150 Hz), (b) detrending the complete signal (using spline), (c) removing ectopic beat (and also the subsequent beat), (d) removing artifactual beats (by a full cross-correlation analysis), (e) and finally, an averaging process applied to validated beats for noise reduction. In order to implement (c), (d), and (e), a main peak detection of the QRS complex (namely R or S) was developed by following derivative method mentioned in [**Error! Reference source not found.**]. The correlation method was fine tuned, and correlation window size was set around the main peak, as a percentage of the full beat cycle.

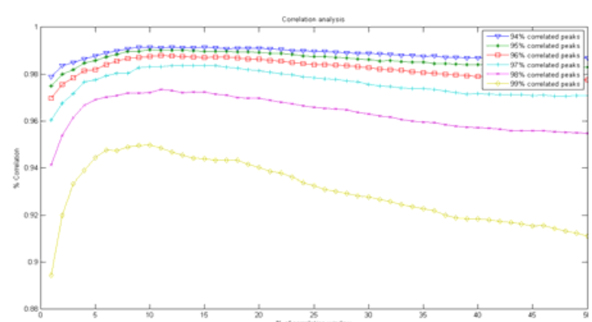

Figure 1. Evaluation of the correlation indexes.

The correlation index was evaluated and maximum of correlations was found over on a 10% window of the total beat (see figure 1). As a result, it was set a window of the 10% of the cycle for beat validation in correlation analysis. Complexes with cross-correlation index over 99.95% were considered valid for the processing.

2.4. Implemented methods

Five implemented methods were compared during this study, namely, Biosig [11], Physionet [12], Chouhan [9, 10], ad-hoc first and second momentum method developed based on Physionet strategy, and a combined method using Chouhan model and our developed ad-hoc method. Here follows a very brief description of these methods, and detailed information could be obtained on respective references.

Physionet is a very well know software-library for physiologic signal processing and analysis, detection of significant events and display and characterization of signals. This software was developed by Boston's Beth Israel Deaconess Medical Centre, Boston University, McGill University, and MIT in 1999 [12]. Algorithms use a morphological approach for QRS delineation, based on first and second derivative of the signal. Full detailed description can be found at http://www.physionet.org.

Chouhan's algorithm develops a complex model substantiated on a number previous publications, where the outcome is a smooth risk index of presence of the QRS, which flows in parallel to the signal. This index generates an easy and efficient approach to set a region for OnSet and OffSet candidates. Full development of this model can be found on [9, 10].

The *combined proposed method* consists of using previous pre-detected candidates for the starting an the end of the QRS complex by Chouhan's method. A second adjustment is performed based on slope detection (ad-hoc developed method) from the first and second momentum.

A number of thresholds and free parameters were set for all mentioned methods, in a joint effort with the medical team, to evaluate accuracy of each method.

2.5. Gold-standard

The Gold Standard was manually and carefully developed by clinicians, an ad-hoc software was created for their convenience. Cases were randomly shuffled and presented to clinician without any reference, to ensure reliable and full independence of the results. Clinicians set fiducial points of the ECG for each lead of all previously selected cases, namely, QRS OnSet and QRS OffSet, among other registered information that will be used in future developments.

3. Results and discussion

3.1. Results

The 150 randomly selected ECGs were analized with the methods described in Section 2, and results were compared against the Gold Standard and the information provided by ECG System GE MAC 5000. Basic statistical analysis (mean and standard deviation) was applied to results (see Tables 1 and 2). A single consolidated QRS duration was provided by the ECG system, and as a consequence, mean value among all lead measurements was be calculated for benchmarking.

From the mean analysis perspective, it could be appreciated that all methods provided larger QRS duration values when compared to clinicians' measurements (See Table 1), special mention required by the GE device being the largest value. In terms of standard deviation, the highest value corresponded again for GE device, and the lowest for an important number of leads in Chouhan method. Only for certain leads the Physionet method overcame the others. The combined method provided a good trade-off in terms of basic statistics across leads.

The total error was selected as a direct merit figure, and defined as the difference between the QRS for each method and the Gold Standard. Table 2 contains this error for each lead and the averaged QRS in terms of mean and standard deviation.

Lead	*Physionet*	*Chouhan*	*Combined*	*GE MAC 5000*	*Gold Standard*
I	96,08±27,52	94,33±19,44	92,85±24,90		81,36±19,14
II	89,09±24,44	95,74±18,14	91,59±21,92		90,05±18,60
III	89,97±19,87	95,03±12,96	89,85±18,58		89,41±16,07
aVR	101,21±37,72	94,40±15,71	90,89±20,28		86,27±20,31
aVL	97,64±21,42	96,63±18,47	95,29±22,68		85,88±18,26
aVF	88,95±22,12	94,71±13,99	90,90±19,63		89,69±17,14
V1	103,38±26,81	101,40±18,44	97,59±22,00		87,20±20,80
V2	104,33±17,85	102,10±16,34	103,03±17,56		91,57±15,92
V3	105,55±21,58	99,33±18,23	105,08±20,61		91,06±16,76
V4	94,36±18,85	93,49±14,93	92,84±19,83		88,45±17,45
V5	89,58±19,85	92,33±14,33	88,52±19,36		86,12±17,72
V6	86,15±21,10	93,01±14,12	89,23±19,02		85,37±18,37
Mean	95,60±15,75	96,01±12,18	93,97±15,75	96,33±17,20	87,70±15,82

Table 1. Mean values and Standard Deviations of QRS onset-offset (ms).

Lead	*Physionet*	*Chouhan*	*Combined*	*GE MAC 5000*
I	14,73±27,11	12,97±21,13	11,49±25,39	
II	-0,06±17,47	5,69±15,36	1,54±14,40	
III	0,56±15,70	5,62±15,62	0,44±16,11	
aVR	14,93±40,92	8,12±17,72	4,62±15,76	
aVL	11,76±19,70	10,75±17,89	9,40±19,56	
aVF	-0,74±16,43	5,02±15,09	1,21±15,30	
V1	16,18±25,55	13,84±15,90	10,39±15,65	
V2	12,76±13,27	10,53±16,66	11,46±15,84	
V3	14,49±18,44	8,27±15,10	14,02±16,18	
V4	5,90±14,80	5,04±15,18	4,38±15,05	
V5	3,45±15,43	6,21±16,30	2,40±14,89	
V6	0,78±14,90	7,64±15,00	3,86±13,77	
Mean	7,90±6,83	8,31±3,07	6,27±4,77	8,63±5,89

Table 2. Errors of the measurements in ms (Method VS

Gold Standard).

A by-lead analysis indicates that accuracy is very much lead dependent, providing in all methods better results on leads II, III, V5, V6, and contrary to the results I, V1, V2, V3, where errors are in mean over 10 ms. Attending to the shown information, the combined method provided consistently better results.

3.2. Discussion of the results

All Implemented methods, applied over ECG beat template (ECG mean), performed better than the GE MAC 5000 ECG device when compared against the Gold Standard. This fact suggested that preprocessing and development over a "clean", and not artifacted, beat template, plays a significant role in detection. Although no material data is shown here, it should be mentioned that clinicians rejected as valid delineation over 80% of the measurements from this very same automatic methods when applied over the full signal. So, we can assume that the removal of not well-correlated complexes helps significantly to method performance.

Among all implemented methods, the combined effort of Chouhan and the empirically developed model, slightly improved individual methods. We considered that Chouhan method provides a valid evaluating a risk slope for QRS pre-detection, while a proper selection of thresholds in derivative signals contributes significantly to set the obtained final result. In that sense, we understand that the medical contribution, especially for thresholds adjustments, plays also an important role.

4. Conclusions and future work

We conclude that: (i) the preprocessing has been revealed as a key element for detection, in particular, selection of valid QRS complexes is relevant in this area; (ii) clinical contributions, for free parameters tuning, were revealed key to set up existing methods; (iii) a blended approach incorporating different methods provides better results.

It is understood that further developments are needed for a fully unsupervised ECG delineation. In that direction, we propose for future developments: (i) to improve the presented methods on lead-by-lead fine tunning; (ii) to evaluate other methods present in literature, such as time-scale; (iii) to extend the results and models for detecting other fiducial points, namely P and T waves.

Acknowledgements

This work has been partially supported by: TEC2010-19263, IPT-2012-1126-300000, DPI2011-27022-C02-01, ACOMP/2013/018, PI11/02459, and RD06/0014/0017.

References

[1] Pahlm O, Sörnmo L. Software QRS detection in ambulatory monitoring—a review. Medical and Biological Engineering and Computing 1984;22:289–297.

[2] Kohler BU, Hennig C, Orglmeister R. The principles of software QRS detection. Engineering in Medicine and Biology Magazine, IEEE 2002;21:42–57.

[3] Engelse W, Zeelenberg C. A single scan algorithm for QRS-detection and feature extraction. Computers in Cardiology 1979;6:37–42.

[4] Ligtenberg A, Kunt M. A robust-digital QRS-detection algorithm for arrhythmia monitoring. Computers and Biomedical Research 1983;16:273–286.

[5] Pan J, Tompkins WJ. A real-time QRS detection algorithm. IEEE Transactions on Biomedical Engineering 1985;3:230–236.

[6] Iuliano S, Fisher S, Karasik P, Fletcher R, Singh S. Department of Veterans Affairs Survival Trial of Antiarrhythmic Therapy in Congestive Heart Failure. QRS duration and mortality in patients with congestive heart failure. Am Heart J 2002;143:1085–91.

[7] Martnez JP, Almeida R, Olmos S, Rocha AP, Laguna P. Á wavelet-based ECG delineator: evaluation on standard databases. Transactions on Biomedical Engineering, IEEE 2004;51:570–581.

[8] Mateo J, Laguna P. Improved heart rate variability signal analysis from the beat occurrence times according to the IPFM model. IEEE Transactions on Biomedical Engineering 2000;47:985–996.

[9] Chouhan V, Mehta S. Detection of QRS complexes in 12-lead ECG using adaptive quantized threshold. IJCSNS 2008;8:155.

[10] Chouhan V, Mehta S. Threshold-based detection of P and T-wave in ECG using new feature signal. IJCSNS 2008;8:144–153.

[11] Schlogl A, Brunner C. BioSig: a free and open source software library for BCI research. Computer 2008;41:44–50.

[12] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation 2000;101:e215–e220.

Address for correspondence.

Francisco-Javier Gimeno-Blanes
Innova Bld., Av. Univesidad, s/n, 03202 Elche Alicante, Spain
javier.gimeno@umh.es