This is the preprint manuscript of the article that appeared in final form in **Applied Mathematics and Computation** 369 : (2020) // Article ID 124839, which has been published in final form at https://doi.org/10.1016/j.amc.2019.124839.

Online Robust R-peaks Detection in Noisy Electrocardiograms using a Novel Iterative Smart Processing Algorithm

Unai Zalabarria[a], Eloy Irigoyen[a], Raquel Martinez[a], Andrew Lowe[b]

[a]*Department of Systems Engineering and Automation, Faculty of Engineering of the UPV/EHU, Bilbao, ES, 48013. E-mail: unai.zalabarria@ehu.eus*
[b]*School of Engineering, Mathematical and Computer Sciences, Auckland University of Technology, Auckland, NZ, 1010.*

Abstract

Nowadays, many contributions deal with R-peak detection in Electrocardiographic (ECG) signals. Although they present an accurate performance in detection, most of these are presented as offline solutions, both to be processed in high performance platforms (under a big cost), or to be analysed in laboratories without constraints in time, neither in computational load. Owing to this, it is also very important to take one step further, trying to develop new solutions which work in portable/wearable low-cost platforms, with constraints in time and in computational load.

In this work, an accurate and computationally efficient method for online and robust detection of R-Peaks is presented. This method is divided in three main stages: first, in the pre-processing stage, a complete elimination of artifacts is performed based on a noise and signal intensity approach; second, R-peaks detection is carried out through an efficient "area over the curve" method; finally, in the third stage, a novel iterative algorithm consisting in three sequential state machines performs the correct detection of the R-peaks applying heart period distance rules. Moreover, the method is performed over time in short length sliding windows.

The algorithm has been tested using all 48 full-length ECG records of the MIT-BIH Arrhythmia Database, achieving 99.54% sensitivity and 99.60% positive predictivity in R-peak detection.

Keywords: Electrocardiogram, ECG processing, R-peak detection, filtering, smart computing, state machine.

1. Introduction

The electrocardiogram (ECG) is the most used technology for recording complex waveforms generated from heart electrical activity in the myocardial contraction [1, 2]. The electrical propagation is measured from the patient's body surface using a set of electrodes. For a normal human heartbeat, the ECG signal has the characteristic shape shown in Fig. 1.

The R-peak located in the QRS complex is the most characteristic waveform and is usually employed as a reference for the ECG analysis [2]. Once the location of R-peak is defined, then other wave components of ECG signal can be determined. Therefore accurate detection of R-peaks is the most important objective in automatic ECG signal analysis.

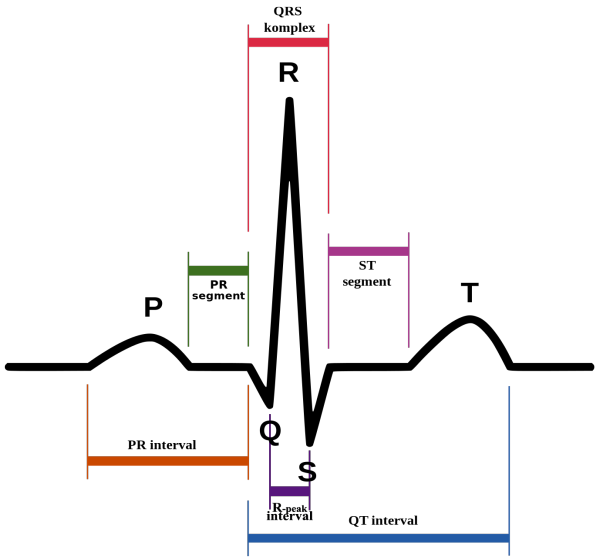

Figure 1: Intervals on the PQRST complex.

The use of increasingly modern technologies is leading to the development of portable devices, which carry out the acquisition of the ECG using a single lead [3], two leads [4] or Einthoven's triangle configuration [5, 6]. These configurations are vulnerable to artifacts, causing the correct beats recognition being impeded by power-line interference, electromyogram noise and baseline wander often present in the ECG signal [7, 8, 9, 10, 11]. These artifacts are

shown in Fig. 2, where graphs (a-d) correspond to different sections of one of the records acquired using Biopac MP150, in which 4 main types of artifacts are distinguished.

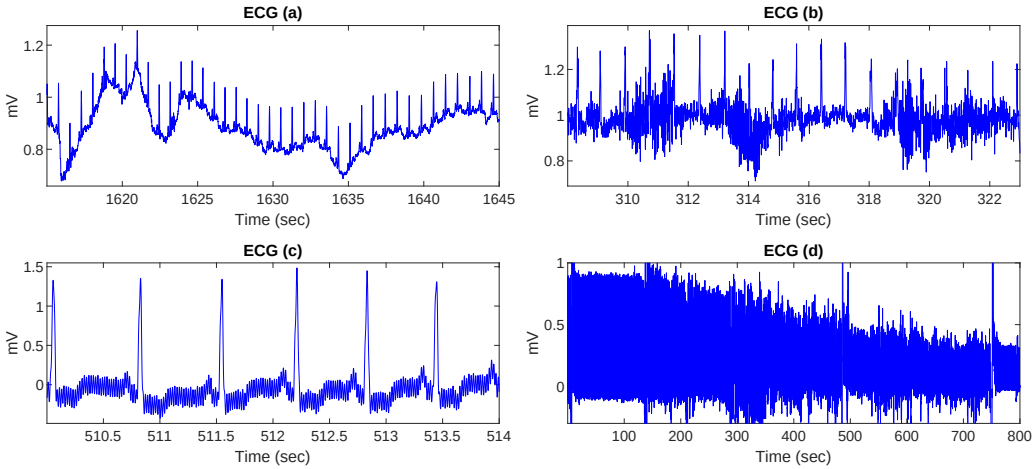

Figure 2: Artifacts present in an ECG signal acquired with Biopac MP150 in a daily situation: a) Baseline wander b) White noise c) Network 50Hz electromagnetic interference d) Loss of sensor conduction results in a loss of signal strength.

Numerous algorithms based on different techniques have been developed during the last decades for ECG analysis [12]. These are derivatives [13, 14], digital filters [15, 16, 17, 18, 19], wavelet-transform [20, 21, 22, 23, 24, 25, 26], neural networks [27], genetic algorithms [28], support vector machine (SVM) [29], k-means [30], combined threshold method [31], moving averaging method [32] and Hilbert Transform method [33] among others.

Additionally, some proposals in the ECG processing use short length sliding windows for the detection of physiological events in real-time [5, 6]. This type of processing requires a reliable detection of R-peaks, since an error in the detection can generate an even greater error in the frequency, nonlinear and temporal parameters extracted from detected R-peaks as the duration of the analyzed window is shorter.

An accurate detection of R-peaks is important in the study of the ECG to enable subsequent signal processing, both to calculate the parameters derived from heart rate and to obtain the locations of the P, Q, S and T waves that shape each PQRST complex. This paper focuses on the robust detection of R-peaks in noisy ECGs by applying a novel iterative computing approach in a low computational load algorithm. As this work studies noisy ECGs, P,

Q, S and T waves detection becomes impracticable as shown in Fig. 2b. A representative scheme of the different phases that compose the full algorithm is shown in Fig. 3. First robust elimination of the artifacts is carried out in the preprocessing stage, followed by an accurate detection of R-peaks using an original area over the QRS complex-based approach. Finally the detected R-peaks are analyzed using a smart iterative method, which is composed of three state machines that are executed sequentially to detect possible false negatives (non-detected R-peaks) and false positives (surplus R-peaks) and correct them. The proposal has been tested on large scale using the standard MIT/BIH Arrhythmia Database [4] and achieved results and computational load have been compared with other existing methods.

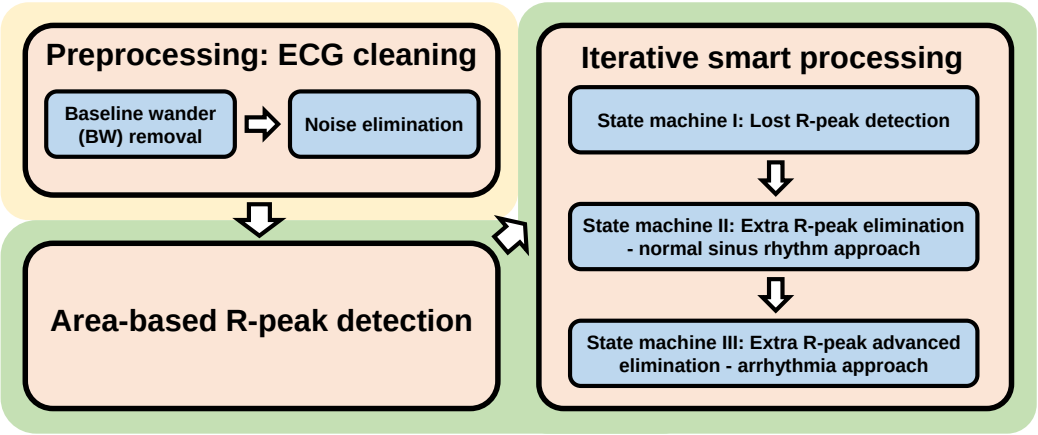

Figure 3: Schematic representation of the proposed R-peak detection algorithm.

This paper is organized as follows. Section 2 explains the datasets used for the development and validation of the algorithm. Section 3 focuses on the proposed methodology. In section 4 the results obtained are presented, discussed and compared with other methods. Finally, section 5 describes conclusions.

2. ECG databases

The proposed algorithm has been created using two different datasets. The first one for development, consisting on noisy ECG records. The second one for validation and comparison.

For development 60 records of 15 minutes each have been collected at a sampling frequency (fs) of 1000 Hz using the Biopac MP150 system (Biopac Systems Inc., USA) with a three-lead Einthoven's triangle configuration. These records were acquired in real life situations and so, they correspond to real ECG fragments as mentioned above. As it is possible to see in Fig. 2, the registers were affected by a wide variety of artifacts.

To validate the algorithm performance against other proposals in the literature, a widely used MIT-BIH Arrhythmia Database has been processed [4]. These 30-minute recordings were sampled at 360 Hz with 11-bit resolution over a 10 mV range. The database is strongly affected by arrhythmia events and includes annotations with beat class information verified by MIT-BIH experts.

3. Material and methods

In general, the R-peak detection is mainly divided in two parts: the first part consist on a preprocessing stage where noise removal is carried out, and second is R-peak detection. This work proposes different methodologies focused on both the preprocessing and the processing of the ECG signal in order to carry out a robust detection of the R-peaks. The methodology proposed in the following steps focuses on the study of the ECG in short duration ECG windows and with low computational load requirements that make it implementable in real-time applications. Specifically, a 20-second sliding window has been used, since it contains enough information to carry out an accurate real-time analysis of the ECG [5, 6].

3.1. Preprocessing: ECG cleaning

In order to carry out the elimination of artifacts that affect the ECG, a two phase signal preprocessing has been developed using time domain methods. Frequency domain methods have been discarded due to the high computational load and the distortions generated in the R-peaks positions when modifying the spectrum [12]. The first phase deals with baseline wander (BW) elimination. In the second phase, the artifacts due to power-line interference and electromyogram noise are removed. Fig. 4 shows the ECG signal before and after the removal of artifacts.

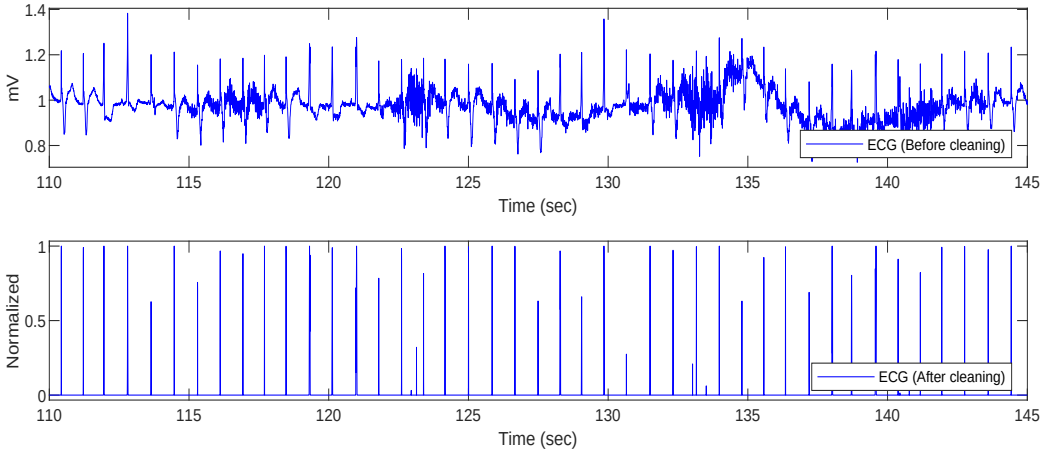

Figure 4: ECG signal from MIT/BIH Arrhythmia Database (record 104) before and after preprocessing. The resulting signal consists of the normalized R-peaks.

3.1.1. Baseline wander removal

The BW is a low-frequency artifact which is caused due to interaction between the electrodes and skin [7, 9, 24]. For the elimination of BW several techniques have been proposed in the literature such as high pass filters, adaptive filters and Wavelet filters among others [7, 8, 10, 34].

In this work a computationally efficient stepped moving median filtering together with a cubic interpolation is proposed for BW removal. The use of a moving median filter is ideal for BW elimination due to R-peaks morphology, which is similar to a high intensity and short-duration wave [7, 10, 24, 33, 34]. Calculating the value of the median for each sample, however, requires a high computational load, which makes it inefficient. In this work an approach based on a stepped median points (p_i) linked together by calculating the cubic interpolation has been proposed according to formulas (1) and (2) respectively, where S_x^i is the section defined by the median filter width in the i^{th} step.

$$p_i(x) = \text{median}_{t \in S_x^i}\{ECG(t)\} \tag{1}$$

$$\begin{aligned} BW_i^{i+1} = (&-\frac{1}{2}p_{i-1} + \frac{3}{2}p_i - \frac{3}{2}p_{i+1} + \frac{1}{2}p_{i+2})x^3 + (p_{i-1} - \\ &\frac{5}{2}p_i + 2p_{i+1} - \frac{1}{2}p_{i+2})x^2 + (-\frac{1}{2}p_{i-1} + \frac{1}{2}p_{i+1})x + p_i \end{aligned} \tag{2}$$

To set up the width of the median filter and the size of the step, the width of the R-peaks and the width of the median filter itself have been considered, respectively. One of the premises is that the width of the median filter must be greater than twice the width of the R-peak so that it is not filtered. At the same time, the width should not be too large to preserve an acceptable computational load. Based on the fact that the normal R-peak width is around 60 milliseconds [35], a 150-millisecond median filter has been configured. For the step width configuration, a value corresponding to half of the width of the median filter has been established, thus maintaining a considerable overlap. Achieved performance is illustrated in Fig. 5 where the BW is perfectly removed according to formula (3), giving ECG′ the resulting filtered ECG.

$$ECG' = ECG - BW. \tag{3}$$

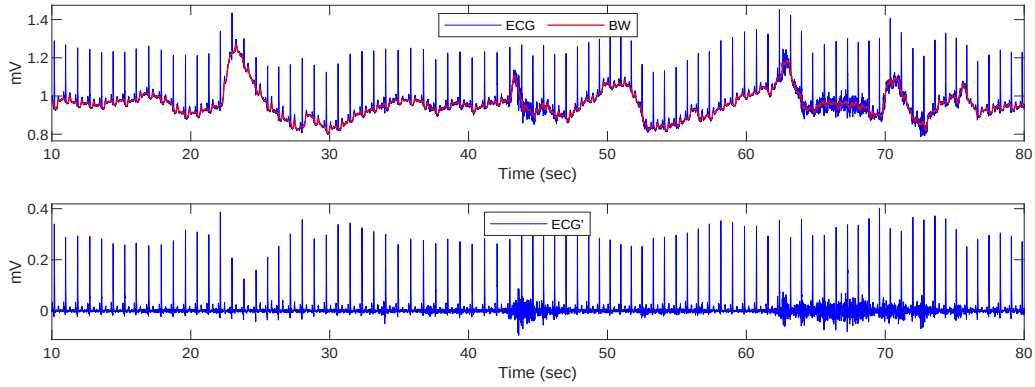

Figure 5: Signal ECG affected by the BW in the upper graph. Result of the elimination of the BW in the lower graph.

3.1.2. Noise elimination

The ECG signal also contains artifacts due to power-line interference and electromyogram noise. This work proposes a noise elimination algorithm based on signal and noise intensity measures in which a dynamic cutting line (CL: black line in Fig. 6) is calculated according to the formula (6). The aim of CL is to cut the ECG, keeping R-peaks (regardless of their morphology) intact and eliminating the noise below based on formula (7) where ECG″ is the resulting cut signal. CL depends on three main parameters:

- Signal Intensity (SI): Consists of a sliding window (width: 1 second, step: 0.5 second) that crosses the signal calculating the maximum value of the ECG according to (4) (green line in Fig. 6). This corresponds to the largest R-peak amplitude in the window.

$$SI_i(x) = \max_{t \in S_x^i} \{ECG'(t)\}. \tag{4}$$

- Noise Intensity (NI): Consists of another sliding window (width: 1 second, step: 0.5 second) where the standard deviation is calculated according to formula (5). Since the standard deviation is proportional to noise, the final value is multiplied by 2 so that 95% of the noise is below the resulting value (pink line in Fig. 6).

$$NI_i = 2 * \sqrt{\frac{1}{fs-1} \sum_{j=i-fs/2}^{i+fs/2} (ECG'_j - \overline{ECG'})^2}. \tag{5}$$

- Aggressiveness level (α): This parameter is used to calibrate the aggressiveness level in the elimination of noise when applying the cut, which is calculated according to formula (6). Best empirical results have been obtained with an α value equal to 5 for the whole dataset. This value should be set according to the quality of the signal acquired by the device in which it is implemented, with 1 the most aggressive value and 10 the least aggressive.

$$CL_i = NI_i + (SI_i - NI_i) * \frac{10 - \alpha}{9} \quad \{\alpha \in R \mid 1 \leq \alpha \leq 10\} \tag{6}$$

$$ECG''_i = \begin{cases} ECG'_i - CL_i & : ECG'_i > CL_i \\ 0 & : ECG'_i \leq CL_i \end{cases}. \tag{7}$$

After calculating ECG″, a normalization is carried out (ECG‴) in a range of 0 to 1 according to formula (8) where ECG‴ is the resulting normalized signal. The result is visible in the lower graph of Fig. 6.

$$ECG'''_i = \frac{ECG''_i}{SI_i - NI_i} \tag{8}$$

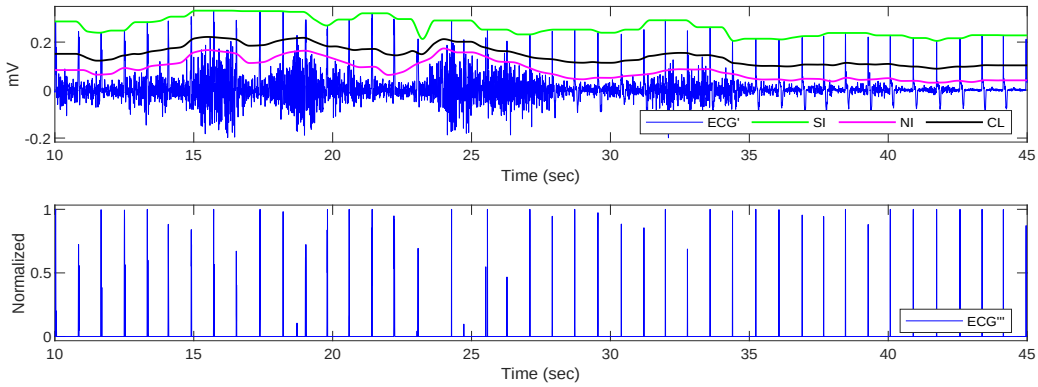

Figure 6: Signal ECG affected by electromagnetic interference in the form of white noise in the upper graph. Result of the elimination of white noise under CL and intensity normalization in the lower graph.

3.2. R-peak detection method

The detection of R-peaks is divided into two complementary phases. The first one consists of an initial detection of R-peaks applying an "area over the curve"-based approach [3]. In the second phase a novel state machine (SM) structure-based methodology has been developed, in which detected R-peaks are evaluated through a set of conditions that decide whether a detected peaks correspond to erroneous detections or not or if the signal presents some inconsistency due to undetected R-peaks.

3.2.1. Area-based R-peak detection

Based on the morphology of the R-peaks, the hypothesis that R-peaks have a high and narrow shape have been considered, defining the neighbors (N) for each local maximum in ECG‴ and letting M be the amplitude of the local maximum [3]. These neighbors correspond to half the value of the duration of the QRS complex (W). Depending on these parameters the area over the curve is calculated according to formula (9), which coincides with the blue area represented in Fig. 7. When a local maximum is high and narrow the corresponding area is high, this is the case of an R-peak. The other peaks correspond to the remaining waves of the PQRST complex that do not have such a large area. In this study, a QRS duration of 100 ms has been

considered (W=100 ms, N=50 ms) based on the literature [2, 3, 35, 36, 37].

$$area = \frac{1}{2 * N} \sum_{i=-N*fs}^{N*fs} (M - ECG_i''') \tag{9}$$

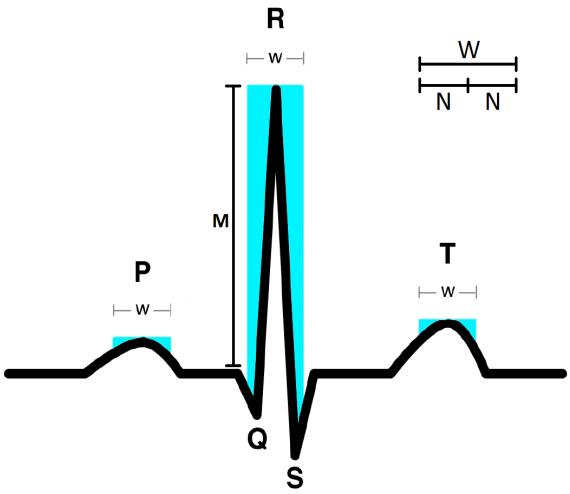

Figure 7: Areas over the curve (blue shading) for each local maximum (P, R, T) in a PQRST complex.

An important parameter to be defined in this approach is the cut-off value of the area from which a local maximum is considered to be an R-peak. To define the optimum value of the cut-off point, all area distributions obtained from the development dataset have been studied and represented in Fig. 8 using a logarithmic scale to improve visualization.
Optimal balance between false positives (FP) and false negatives (FN) has been obtained with a cut-off value of 0.55, reaching the highest accuracy in the detection of R-peaks in the development dataset.

3.2.2. Iterative smart processing method

The result obtained using the area-based R-peak detection method is effective when the morphology of the R-peak corresponds to the theoretical shape shown in Fig. 7. When an important pathology such as premature ventricular contraction (PVC) occurs, this morphology differs from the theoretical shape as illustrated in Fig. 9. This may result in the value of the area

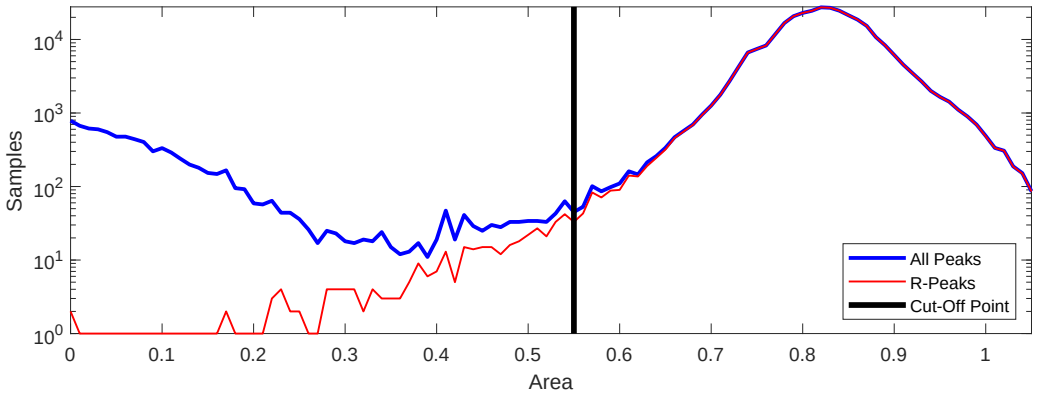

Figure 8: Probability histogram for all detected peaks areas (blue line) and areas corresponding only to R-peaks (red line) in the development dataset.

over the curve being below the defined cut-off point, leading to a FN. Moreover, when very noisy sections are removed, the amplitude of some R-peaks appears very reduced, not reaching the cut-off area value as in the processing represented in Fig. 10.

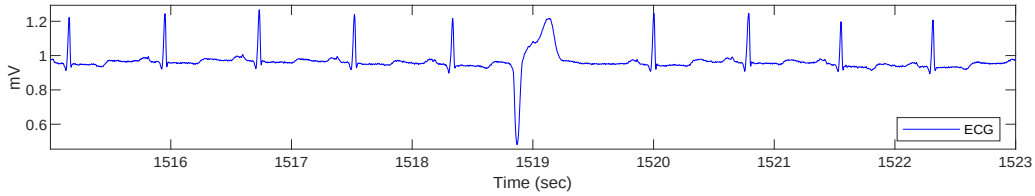

Figure 9: Premature ventricular contraction phenomenon at time 1519 seconds.

In both cases a FN is a serious problem when analyzing short duration windows. An error in the detection leads to significant variations in the parameters calculated from detected R-peaks. The same happens when a FP occurs. In the following subsections, three sequentially executed state-machines are presented for the selective detection and elimination of FN and FP. These algorithms are based on a set of conditions that relate previously detected R-peaks to carry out an intelligent optimization. The first SM carries out the detection of the undetected R-peaks. The second SM eliminates the surplus R-peaks according to a normal sinus rhythm (NSR) HP conditions. Finally, the third SM eliminates the remaining FP considering the possibility of arrhythmias.

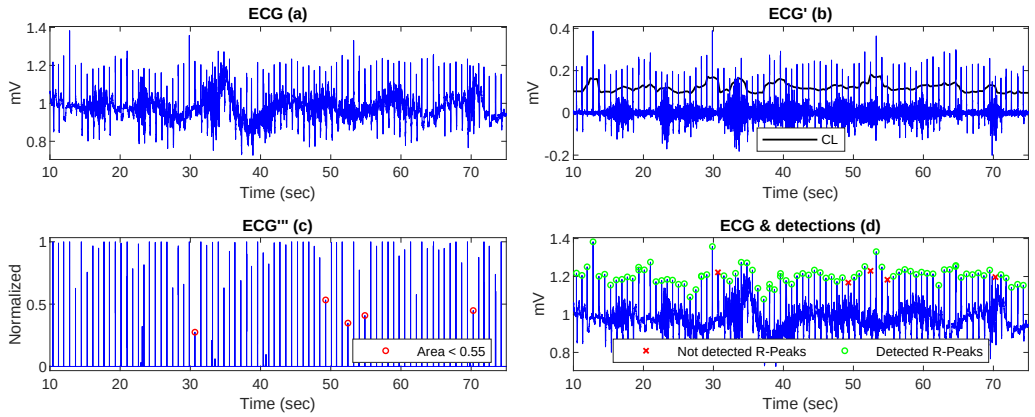

Figure 10: Example of non-detected R-peaks that do not reach the minimum cut-off area value: a) Original ECG signal frame, b) Filtered ECG and cutting line, c) Cut and normalized ECG, d) R-Peak detection with some non-detected peaks.

The conditions rest on heart period (HP) measurements to carry out an analysis based on contiguous R-peaks. HP is defined as an N-1 length array calculated according to equation (10), being R_Peaks an N length array composed of the R-peaks timestamps.

$$HP_i = R\_Peaks_i - R\_Peaks_{i-1} \tag{10}$$

In addition to raw HP, parameters such as median HP ($median(HP)$), maximum HP (HP_{max}) and minimum HP (HP_{min}) are used to normalize the conditions so that they are applicable to every register equally. All conditions are composed of logical operations which are scaled with constant values calculated interactively in the experiments performed on the development dataset.

SM I - Lost R-peak detection

Having all the possible R-peaks detected is a fundamental step that provides enough information to carry out the elimination of those that are FP. In order to ensure the detection of all possible R-peaks, the first SM has been designed as shown in Fig. 11. The processing is based on an iterative analysis of HP that analyzes each of the spaces between R-peaks in search of anomalous situations that indicate the lack of one or several R-peaks.
This first SM depends on condition 1, which is represented in equation (11) and considers the cases in which a FN occurs. When condition 1 is met, it is

considered that the algorithm has detected a FN and a search for the missing R-peak begins between the two consecutive R-peaks where the missing one is supposed to be. The cut-off point is reduced to 0.4 in this interval, allowing to detect possible R-peaks that could have been overlooked with the previous cut-off value. If there exist peaks with an area over the curve greater than 0.4, the largest one is taken. If none exists, the largest local maximum is selected as the missing R-peak.

$$\begin{gathered} Condition_1 = \\ (HP_i > HP_{max}) \cap (HP_i > 1.7 \cdot median(HP)) \cap \\ (1.3 \cdot HP_{i-1} < HP_i) \cup (HP_i > 1.3 \cdot HP_{i+1}) \cap \\ (HP_i > 1.5 \cdot HP_{i+1}) \cap (HP_i > 1.5 \cdot HP_{i-1}) \cap \\ (1.3 \cdot HP_{i-1} < HP_i) \cup (1.3 \cdot HP_{i-2} < HP_i) \cap \\ (1.3 \cdot HP_{i+1} < HP_i) \cup (1.3 \cdot HP_{i+2} < HP_i) \end{gathered} \tag{11}$$

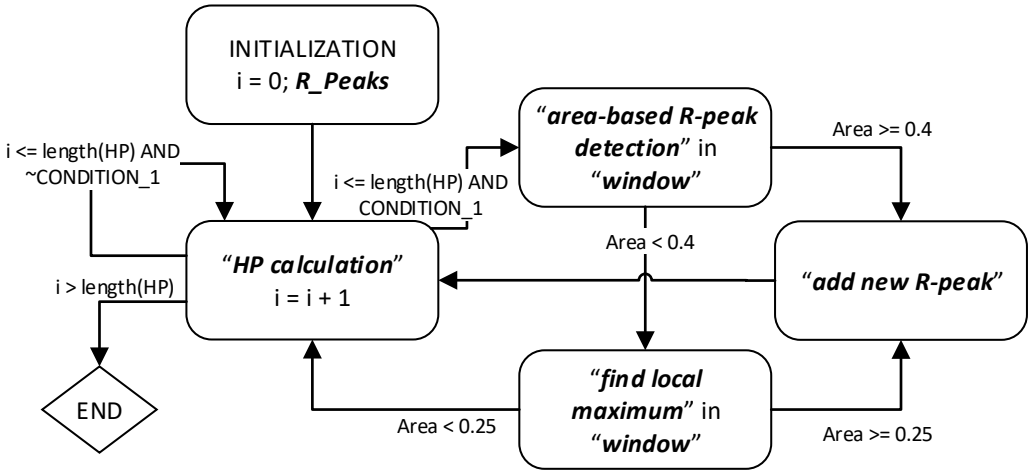

Figure 11: State machine I for lost R-peak detection. "***window***" is the signal fragment that goes from R_Peaks_i to R_Peaks_{i+1}. "***area-based R-peak detection***" is the processing carried out in section 3.2.1. "***find local maximum***" looks for a maximum peak in "***window***". "***add new R-peak***" is a function that adds the new R-Peak in the correct position.

SM II - Extra R-peak elimination: normal sinus rhythm approach

Second SM analyzes HPs obtained from the first SM in search of FP considering that the HP morphology corresponds to a NSR. It is based on

the fact that each HP must be related to the contiguous HPs, so that certain distances between contiguous R-peaks are not exceeded. To control that the distances are within a fixed ranges the Condition 2.1 and Condition 2.2 have been defined in equations (12) and (13) respectively. Both conditions are based on the duration of a parameter called "***window***" (W_t), defined in the SM of Fig. 12. This "***window***" is generated dynamically and contains series of R-peaks corresponding to consecutive HPs in which their values meet the conditions $HP_i < HP_{min}$ and $HP_i < 0.6 * median(HP)$. If Condition 2.1 is true, it is considered that the R-peak corresponding to the current index "i" is a FP. If Condition 2.2 is true, more than one consecutive R-peak are considered to be FP and a selective elimination is carried out.

$$\begin{aligned} Condition_{2.1} = (HP_i < HP_{min}) \cup \\ (HP_i < 0.6 \cdot median(HP)) \cup (W_t < 2.5 \cdot median(HP)) \end{aligned} \tag{12}$$

$$\begin{aligned} Condition_{2.2} = (HP_i < HP_{min}) \cup \\ (HP_i < 0.6 \cdot median(HP)) \cup (W_t \geq 2.5 \cdot median(HP)) \end{aligned} \tag{13}$$

This state machine does not consider possible cases of arrhythmias since it is focused on carrying out a general cleaning of FP in NSR conditions. This requires a third SM to perform a more thorough cleaning, considering the morphology of the HP in arrhythmia situations.

SM III - Extra R-peak advanced elimination: arrhythmia approach

In the third SM an elimination of FP is carried out using a more complex process focused on arrhythmia events due to the difficulty of detecting FP when the ratio of distances between R-peaks becomes irregular. The algorithm analyzes the current HP in a broader context focusing on a wider range of contiguous HPs. This makes the SM very sensitive to an excess of FP in the analyzed range. Hence, a first raw elimination of the surplus R-peaks was necessary in the second SM.

To detect the remaining FP, the morphology of the HP when arrhythmia occurs has been modeled empirically in 5 conditions defined in (14), (15), (16), (17) and (18). Possible irregularities have been considered due to this pathology in which the R-peaks use to be displaced (relative to their expected position in NSR conditions) more or less close to contiguous R-peaks. If any of the 5 conditions is true during the analysis of a particular HP, current R-peak (index "i") is considered to be a FP and is eliminated. A schematic

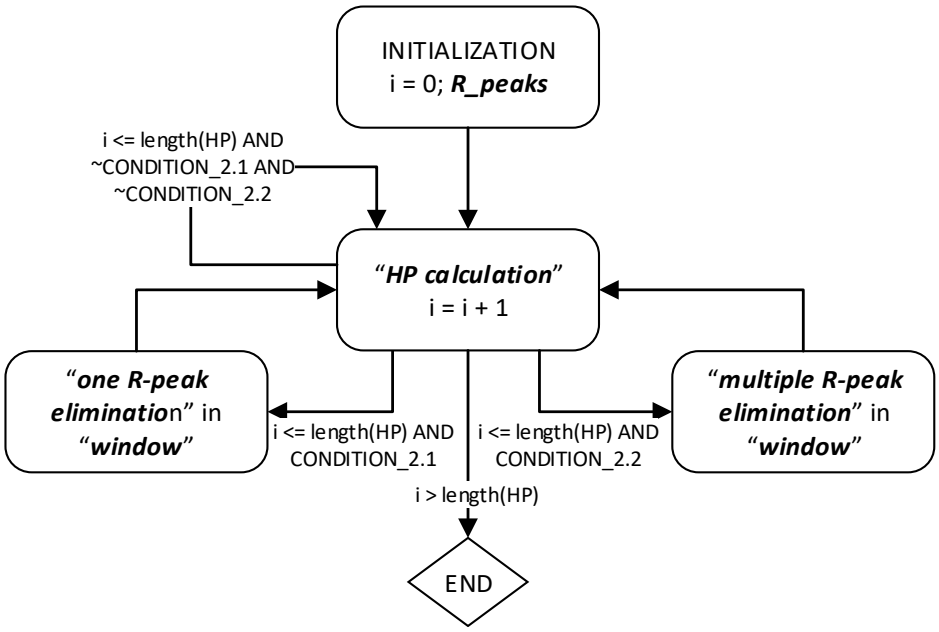

Figure 12: State machine II for extra R-peak elimination: "***window***" is the signal fragment in which consecutive series of anomalous R-peaks have been detected. "***one R-peak elimination***" is executed considering that in "***window***" there is only one erroneous R-peak to find and eliminate. "***multiple R-peak elimination***" is executed considering that in "***window***" there are two or more consecutive erroneous R-peaks.

representation of the SM is shown in Fig. 13.

$$\begin{gathered} Condition_{3.1} = \\ (1.2 \cdot (HP_1 + HP_2) > HP_3 > (HP_1 + HP_2)/1.2) \cup \\ (1.2 \cdot (HP_1 + HP_2) > HP_4 > (HP_1 + HP_2)/1.2) \cup \\ (HP_3 > HP_{min}) \cup (HP_4 > HP_{min}) \end{gathered} \tag{14}$$

$$\begin{gathered} Condition_{3.2} = \\ (1.2 \cdot (HP_{i+1} + HP_{i+2}) > HP_i > (HP_{i+1} + HP_{i+2})/1.2) \cup \\ (1.2 \cdot (HP_{i+1} + HP_{i+2}) > HP_{i+3} > (HP_{i+1} + HP_{i+2})/1.2) \cup \\ (HP_i > HP_{min}) \cup (HP_{i+3} > HP_{min}) \end{gathered} \tag{15}$$

$$
\begin{gathered}
Condition_{3.3} = \\
(1.2 \cdot (HP_{N-1} + HP_N) > HP_{N-2} > (HP_{N-1} + HP_N)/1.2) \cup \\
(1.2 \cdot (HP_{N-1} + HP_N) > HP_{N-3} > (HP_{N-1} + HP_N)/1.2) \cup \\
(HP_{N-2} > HP_{min}) \cup (HP_{N-3} > HP_{min})
\end{gathered} \tag{16}
$$

$$
\begin{gathered}
Condition_{3.4} = \\
(1.2 \cdot (HP_{i+1} + HP_{i+2}) > HP_{i+2} > (HP_{i+1} + HP_{i+2})/1.2) \cup \\
(1.2 \cdot (HP_{i+1} + HP_{i+2}) > HP_{i+3} > (HP_{i+1} + HP_{i+2})/1.2) \cup \\
(1.1 \cdot HP_{i+3} > HP_{i+2} > HP_{i+3}/1.1) \cup \\
(1.1 \cdot HP_{i+2} > HP_{i+3} > HP_{i+2}/1.1) \cup \\
(HP_{i+2} > HP_{min}) \cup (HP_{i+3} > HP_{min}) \cup \\
(HP_i + HP_{i+1} > HP_{min}) \cup (HP_i + HP_{i+1} < HP_{max})
\end{gathered} \tag{17}
$$

$$
\begin{gathered}
Condition_{3.5} = \\
(1.2 \cdot (HP_{i+2} + HP_{i+3}) > HP_i > (HP_{i+2} + HP_{i+3})/1.2) \cup \\
(1.2 \cdot (HP_{i+2} + HP_{i+3}) > HP_{i+1} > (HP_{i+2} + HP_{i+3})/1.2) \cup \\
(HP_i > HP_{min}) \cup (HP_{i+1} > HP_{min})
\end{gathered} \tag{18}
$$

The result of the three state machines implementation is a complete detection of the R-peaks illustrated in Fig. 14. This figure corresponds to the same example previously shown in Fig. 10 prior to the iterative smart processing method.

4. Results and discussion

The designed algorithm has been validated and compared with the methods available in the literature using MIT-BIH Arrhythmia Database [4]. The performance of the algorithm is given in Table 1 based on the number of True Positives (TP), False Positives (FP) and False Negatives (FN) obtained. From these values, sensitivity (Se) (19) and positive predictivity (P^+) (20) have been calculated to analyze the performance of the proposed algorithm according to the formulas:

$$
Se = \frac{TP}{TP + FN} \quad (\%) \tag{19}
$$

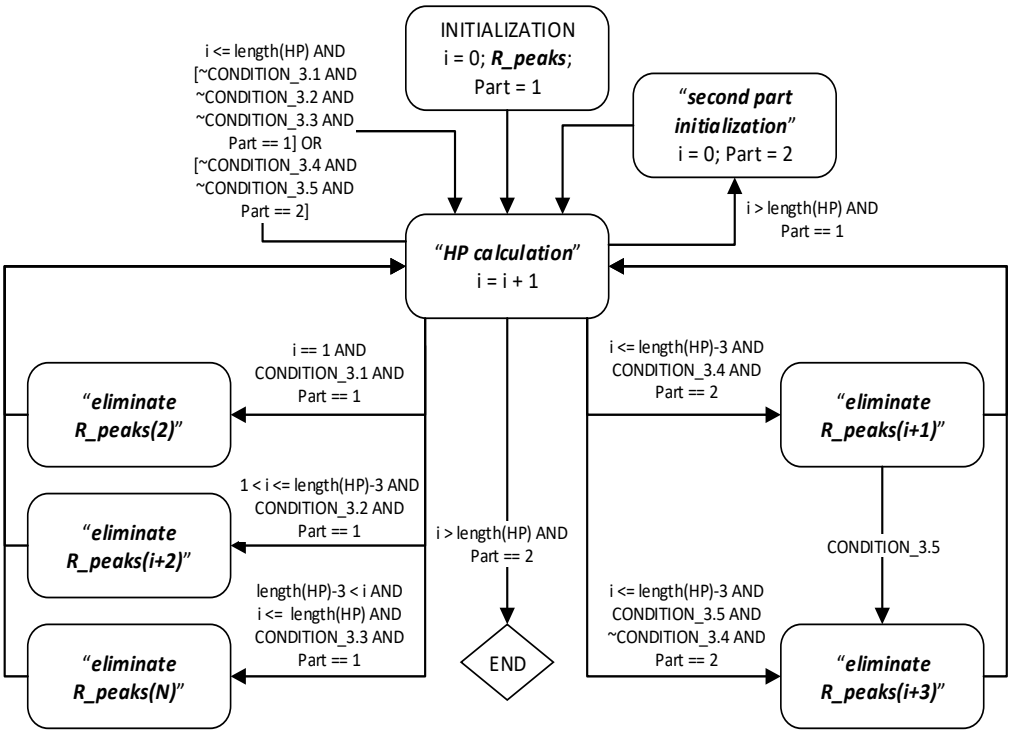

Figure 13: State machine III for extra R-peak advanced elimination. "***second part initialization***" initializes the parameters to execute the second part conditions. "***eliminate R_peaks(x)***" removes R-peaks corresponding to position $\boldsymbol{x}$.

$$P^{+} = \frac{TP}{TP + FP} \quad (\%) \tag{20}$$

Table 1: Results of evaluating the proposed algorithm performance on MIT-BIH Arrhythmia Database.

Tape No.	**#annotations**	**TP**	**FP**	**FN**	**Se(%)**	**P^{+}(%)**
100	2273	2273	0	0	100.00	100.00
101	1865	1864	3	1	99.95	99.84
102	2186	2185	2	1	99.95	99.91
103	2084	2084	0	0	100.00	100.00
104	2228	2227	3	1	99.96	99.87
105	2559	2557	9	2	99.92	99.65

Continued on next page

Table 1 – *Continued from previous page*

Tape No.	**#annotations**	**TP**	**FP**	**FN**	**Se(%)**	**P$^+$(%)**
106	1978	1970	0	8	99.60	100.00
107	2099	2093	3	6	99.71	99.86
108	1755	1753	73	2	99.89	96.00
109	2521	2519	0	2	99.92	100.00
111	2124	2124	0	0	100.00	100.00
112	2539	2539	0	0	100.00	100.00
113	1795	1795	0	0	100.00	100.00
114	1875	1874	183	1	99.95	91.10
115	1953	1953	0	0	100.00	100.00
116	2393	2390	1	3	99.87	99.96
117	1533	1532	0	1	99.93	100.00
118	2275	2275	0	0	100.00	100.00
119	1958	1953	0	5	99.74	100.00
121	1863	1862	0	1	99.95	100.00
122	2476	2476	0	0	100.00	100.00
123	1516	1515	0	1	99.93	100.00
124	1614	1613	0	1	99.94	100.00
200	2587	2585	8	2	99.92	99.69
201	1806	1783	1	23	98.73	99.94
202	2124	2122	1	2	99.91	99.95
203	2557	2496	24	61	97.61	99.05
205	2627	2622	0	5	99.81	100.00
207	1803	1794	89	9	99.50	95.27
208	2424	2348	4	76	96.86	99.83
209	3004	3003	0	1	99.97	100.00
210	2531	2514	1	17	99.33	99.96
212	2748	2748	0	0	100.00	100.00
213	3212	3206	0	6	99.81	100.00
214	2216	2209	1	7	99.68	99.95
215	3304	3295	1	9	99.73	99.97
217	2207	2206	1	1	99.95	99.95
219	2115	2109	0	6	99.72	100.00
220	2036	2034	0	2	99.90	100.00
221	2405	2401	0	4	99.83	100.00
222	2414	2403	0	11	99.54	100.00

Continued on next page

Table 1 – *Continued from previous page*

Tape No.	#annotations	TP	FP	FN	Se(%)	P^+(%)
223	2458	2436	0	22	99.10	100.00
228	2042	2040	9	2	99.90	99.56
230	2254	2253	0	1	99.96	100.00
231	1886	1885	2	1	99.95	99.89
232	1136	1043	12	93	91.81	98.86
233	2472	2385	0	87	96.48	100.00
234	2751	2750	0	1	99.96	100.00
Total	**106581**	**106096**	**431**	**485**	**99.54**	**99.60**

Results with a high percentage of success in the detection of R-peaks have been achieved with a sensitivity value of 99.54%. The number of false detections or FP is low, which corresponds to a 99.60% positive predictivity value.

Some R-peaks are undetected due to their large width, as in PVC. The morphology of PVC differs from the standard R-peak, which hinders the analysis when "area over the curve" method-based detection is performed. This phenomenon can be observed in some of the records processed such as 114, 207 and 208.

Even though the PVCs of register 114 are uniform, their morphology has led to erroneous detections using "area over the curve" method. Record 207 was extremely difficult to process. The predominant rhythm is normal sinus with first-degree atrioventricular block and left bundle branch block. In addition, since PVCs are multiform, they greatly hindered detection. Record 208 contains several uniform PVCs, in addition to ventricular and normal beat fusions, leading to undetected R-peaks. The amount of FP obtained in register 108 is also remarkable, mainly due to first-degree atrioventricular block and multiform PVCs. Finally, during the state machines processing of register 232, several R-peaks were undetected. This happened because distance-based rules were used to perform the detections in a register affected by numerous long intervals between R-peaks (up to 6 seconds in duration).

Table 2 compares the performance of the proposed algorithm with other well-known works. All of them have used signals from the MIT-BIH Arrhythmia Database for validation, which makes a fair comparison possible. Computational load has been considered according to Kohler et al. [12], who suggested measuring the computational load as low, medium and high according to the generation of signal features and the complexity of the used techniques.

Table 2: R-peak detection performance comparison on MIT-BIH Arrhythmia Database (first channel).

Method	**#peaks**	**TP**	**FP**	**FN**	**Se(%)**	**P$^+$(%)**	**Load [12]**
Pan and Tompkins [15]	109809	109208	507	601	99.45	99.54	High
Hamilton and Tompkins [16]	109267	108927	248	340	99.69	99.77	Medium
Saxena et al. [20]	103763	103664	102	99	99.90	99.90	Medium
Martinez et al. [21]	109428	109208	153	220	99.80	99.86	High
Ghaffari et al. [22]	109428	109327	129	101	99.91	99.88	High
Chouakri et al. [23]	109488	108043	3068	1445	98.68	97.24	High
Zhang et al. [38]	109510	109297	204	213	99.81	99.81	Medium
Christov [31]	110050	109548	215	502	99.54	99.80	Medium
Chen et al. [32]	110050	109615	239	435	99.60	99.78	Medium
Afonso et al. [19]	90909	90535	406	374	99.59	99.55	Low
Bahoura et al. [25]	109809	109635	135	174	99.84	99.88	Medium
Li et al. [26]	104182	104070	65	112	99.89	99.94	High
Proposed method	**106581**	**106096**	**431**	**485**	**99.54**	**99.60**	**Low**

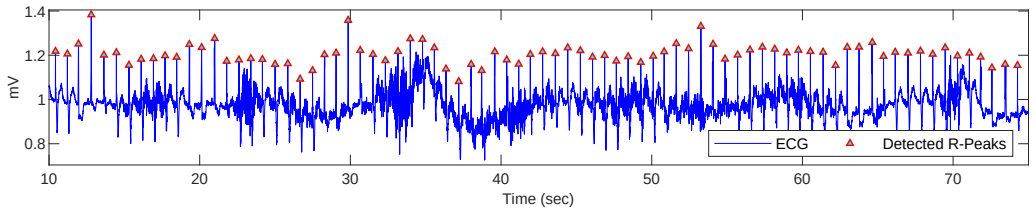

Figure 14: Result of the detection of R-peaks using the proposed algorithm in an ECG signal affected by several artifacts.

Looking at the performance obtained in sensitivity and positive predictivity values in table 2, all are similar for the compared works. Values greater than 99.50% have been achieved in both Se and P^+ measures with the exception of Chooukri et al. [23], who gets somewhat lower results.

Computational load is an important parameter to be considered since many ECG processing applications requires to be implemented in online processes with real-time constraint. This also affects energy expenditure, which can be critical in portable devices focused on long term ECG acquisition. In the proposed work this factor has been considered looking at the complexity of the used methods and verifying its real-time execution in low-cost portable devices such as Raspberry Pi Zero and Arduino Micro models. This leads to an implementation and development of a computationally light algorithm and achieves an effective and robust real-time R-peak detection.

5. Conclusions

This work presents a real-time R-peak detection algorithm based on a robust sliding window strategy. This algorithm reliably detects R-peaks using a computationally light preprocessing for artifacts elimination, as well as an efficient "area over the curve" approach for a raw R-peak detection and a novel iterative analysis on detected R-peaks oriented to FP and FN elimination. The iterative analysis is carried out sequentially through three state machines. For each state machine a set of conditions, which evaluate the HP at every instant searching erroneous detections, has been designed.

Several works propose computationally more expensive techniques based on time and frequency analysis to obtain the position of R-peaks. The proposed algorithm achieves results comparable to those well-known works, and is implementable in real-time with a low computational load and is capable of dealing with a wide range of noise contaminations.

Acknowledgments

This work has been performed thanks to the support of the University of the Basque Country (UPV/EHU), the Intelligent Control Research Group of the UPV/EHU, the Pacific Atlantic Network for Technical Higher Education and Research (PANTHER) program and the Institute of Biomedical Technologies (IBTec) of the Auckland University of Technology, to which the authors are very grateful. The authors also like to thank Javier Mendez for his collaboration in the paper edition.

References

[1] F. G. Cosío, J. Palacios, A. Pastor, A. Núñez, The electrocardiogram, in: The ESC Textbook of Cardiovascular Medicine, Oxford University Press, 2009, pp. 29–82. doi:10.1093/med/9780199566990.003.002. URL https://doi.org/10.1093/med/9780199566990.003.002

[2] J. Hampton, The ECG made easy, Churchill Livingstone/Elsevier, Edinburgh New York, 2013.

[3] Y. Liao, R.-X. Na, D. Rayside, Accurate ecg r-peak detection for telemedicine, in: 2014 IEEE Canada International Humanitarian Technology Conference - (IHTC), 2014, pp. 1–5. doi:10.1109/IHTC.2014.7147524.

[4] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, H. E. Stanley, PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals, Circulation 101 (23) (2000) e215–e220, circulation Electronic Pages: http://circ.ahajournals.org/content/101/23/e215.full PMID:1085218; doi: 10.1161/01.CIR.101.23.e215.

[5] U. Zalabarria, E. Irigoyen, R. Martínez, A. Salazar-Ramirez, Detection of stress level and phases by advanced physiological signal processing based on fuzzy logic, in: International Conference on EUropean Transnational Education, Springer, 2016, pp. 301–312.

[6] A. Salazar-Ramirez, E. Irigoyen, R. Martinez, U. Zalabarria, An enhanced fuzzy algorithm based on advanced signal processing

for identification of stress, Neurocomputing 271 (2018) 48 – 57. doi:https://doi.org/10.1016/j.neucom.2016.08.153.
URL http://www.sciencedirect.com/science/article/pii/S092523121731216X

[7] V. R. Lele, K. S. Holkar, Removal of baseline wander from ecg signal, in: Special Issue of International Journal of Electronics, Communication & Soft Computing Science & Engineering, 2013, pp. 60–65.

[8] R. P. Narwaria, S. Verma, P. K. Singhal, Removal of baseline wander and power line interference from ecg signal - a survey approach, International Journal of Electronics Engineering 3 (2011) 107–111.

[9] M. Shahbakhti, H. Bagheri, B. Shekarchi, S. Mohammadi, M. Naji, A new strategy for ecg baseline wander elimination using empirical mode decomposition, Fluctuation and Noise Letters 15 (2). doi:10.1142/S0219477516500176.

[10] A. Jayant, T. Singh, M. Kaur, Different techniques to remove baseline wander from ecg signal, International Journal of Emerging Research in Management &Technology 2 (6) (2013) 16–19.

[11] G. Han, B. Lin, Z. Xu, Electrocardiogram signal denoising based on empirical mode decomposition technique: An overview, Journal of Instrumentation 12 (3). doi:10.1088/1748-0221/12/03/P03010.

[12] B. U. Kohler, C. Hennig, R. Orglmeister, The principles of software qrs detection, IEEE Engineering in Medicine and Biology Magazine 21 (1) (2002) 42–57. doi:10.1109/51.993193.

[13] Y.-C. Yeh, W.-J. Wang, QRS complexes detection for ECG signal: The difference operation method, Computer Methods and Programs in Biomedicine 91 (3) (2008) 245–254. doi:10.1016/j.cmpb.2008.04.006.
URL https://doi.org/10.1016/j.cmpb.2008.04.006

[14] L. She, G. Wang, S. Zhang, J. Zhao, An adaptive threshold algorithm combining shifting window difference and forward-backward difference in real-time r-wave detection, in: 2009 2nd International Congress on Image and Signal Processing, 2009, pp. 1–4. doi:10.1109/CISP.2009.5304666.

[15] J. Pan, W. J. Tompkins, A real-time qrs detection algorithm, IEEE Transactions on Biomedical Engineering BME-32 (3) (1985) 230–236. doi:10.1109/TBME.1985.325532.

[16] P. S. Hamilton, W. J. Tompkins, Quantitative investigation of qrs detection rules using the mit/bih arrhythmia database, IEEE Transactions on Biomedical Engineering BME-33 (12) (1986) 1157–1165. doi:10.1109/TBME.1986.325695.

[17] M. Adnane, Z. Jiang, S. Choi, Development of QRS detection algorithm designed for wearable cardiorespiratory system, Computer Methods and Programs in Biomedicine 93 (1) (2009) 20–31. doi:10.1016/j.cmpb.2008.07.010.
URL https://doi.org/10.1016/j.cmpb.2008.07.010

[18] R. K. Bal, A. Kumar, Improved qrs detector using hybrid mamemi filter, in: 2016 IEEE International Conference on Recent Trends in Electronics, Information Communication Technology (RTEICT), 2016, pp. 1351–1355. doi:10.1109/RTEICT.2016.7808051.

[19] V. X. Afonso, W. J. Tompkins, T. Q. Nguyen, Shen Luo, Ecg beat detection using filter banks, IEEE Transactions on Biomedical Engineering 46 (2) (1999) 192–202. doi:10.1109/10.740882.

[20] S. C. Saxena, V. Kumar, S. T. Hamde, Feature extraction from ECG signals using wavelet transforms for disease diagnostics, International Journal of Systems Science 33 (13) (2002) 1073–1085. doi:10.1080/00207720210167159.
URL https://doi.org/10.1080/00207720210167159

[21] J. P. Martinez, R. Almeida, S. Olmos, A. P. Rocha, P. Laguna, A wavelet-based ecg delineator: evaluation on standard databases, IEEE Transactions on Biomedical Engineering 51 (4) (2004) 570–581. doi:10.1109/TBME.2003.821031.

[22] A. Ghaffari, M. Homaeinezhad, M. Akraminia, M. Atarod, M. Daevaeiha, A robust wavelet-based multi-lead electrocardiogram delineation algorithm, Medical Engineering & Physics 31 (10) (2009) 1219–1227. doi:10.1016/j.medengphy.2009.07.017.
URL https://doi.org/10.1016/j.medengphy.2009.07.017

[23] S. Chouakri, F. Bereksi-Reguig, A. Taleb-Ahmed, QRS complex detection based on multi wavelet packet decomposition, Applied Mathematics and Computation 217 (23) (2011) 9508–9525. doi:10.1016/j.amc.2011.03.001.
URL https://doi.org/10.1016/j.amc.2011.03.001

[24] P. Sasikala, R. Wahidabanu, Robust r peak and qrs detection in electrocardiogram using wavelet transform, International Journal of Advanced Computer Science and Applications-IJACSA 1 (6) (2010) 48–53.

[25] M. Bahoura, M. Hassani, M. Hubin, DSP implementation of wavelet transform for real time ECG wave forms detection and heart rate analysis, Computer Methods and Programs in Biomedicine 52 (1) (1997) 35–44. doi:10.1016/s0169-2607(97)01780-x.
URL https://doi.org/10.1016/s0169-2607(97)01780-x

[26] Cuiwei Li, Chongxun Zheng, Changfeng Tai, Detection of ecg characteristic points using wavelet transforms, IEEE Transactions on Biomedical Engineering 42 (1) (1995) 21–28. doi:10.1109/10.362922.

[27] R. Silipo, C. Marchesi, Artificial neural networks for automatic ecg analysis, IEEE Transactions on Signal Processing 46 (5) (1998) 1417–1425. doi:10.1109/78.668803.

[28] R. Poli, S. Cagnoni, G. Valli, Genetic design of optimum linear and nonlinear qrs detectors, IEEE Transactions on Biomedical Engineering 42 (11) (1995) 1137–1141. doi:10.1109/10.469381.

[29] S. Mehta, N. Lingayat, SVM-based algorithm for recognition of QRS complexes in electrocardiogram, IRBM 29 (5) (2008) 310–317. doi:10.1016/j.rbmret.2008.03.006.
URL https://doi.org/10.1016/j.rbmret.2008.03.006

[30] S. Mehta, D. Shete, N. Lingayat, V. Chouhan, K-means algorithm for the detection and delineation of QRS-complexes in electrocardiogram, IRBM 31 (1) (2010) 48–54. doi:10.1016/j.irbm.2009.10.001.
URL https://doi.org/10.1016/j.irbm.2009.10.001

[31] I. I. Christov, Real time electrocardiogram qrs detection using combined adaptive threshold, BioMedical Engineering OnLine 3 (1) (2004) 28.

doi:10.1186/1475-925x-3-28.
URL https://doi.org/10.1186/1475-925x-3-28

[32] S.-W. Chen, H.-C. Chen, H.-L. Chan, A real-time QRS detection method based on moving-averaging incorporating with wavelet denoising, Computer Methods and Programs in Biomedicine 82 (3) (2006) 187–195. doi:10.1016/j.cmpb.2005.11.012.
URL https://doi.org/10.1016/j.cmpb.2005.11.012

[33] P. Valluraiah, B. Biswal, Ecg signal analysis using hilbert transform, in: 2015 IEEE Power, Communication and Information Technology Conference (PCITC), 2015, pp. 465–469. doi:10.1109/PCITC.2015.7438211.

[34] F. P. Romero, L. V. Romaguera, C. R. Vázquez-Seisdedos, M. G. F. Costa, J. E. Neto, et al., Baseline wander removal methods for ecg signals: A comparative study, Electrical Engineering and Systems Science at Cornell University.

[35] H. Blackburn, Classification of the electrocardiogram for population studies: Minnesota code, Journal of Electrocardiology 2 (3) (1969) 305 – 310. doi:https://doi.org/10.1016/S0022-0736(69)80120-2.

[36] M. Shenasa, The ECG handbook of contemporary challenges, CardioText, Minneapolis, 2015.

[37] A. Aurobinda, B. P. Mohanty, M. N. Mohanty, R-peak detection of ecg using adaptive thresholding, in: 2016 International Conference on Communication and Signal Processing (ICCSP), 2016, pp. 0284–0287. doi:10.1109/ICCSP.2016.7754140.

[38] F. Zhang, Y. Lian, Qrs detection based on multiscale mathematical morphology for wearable ecg devices in body area networks, IEEE Transactions on Biomedical Circuits and Systems 3 (4) (2009) 220–228. doi:10.1109/TBCAS.2009.2020093.