

Data Fusion for QRS Complex Detection in Multi-Lead Electrocardiogram Recordings

Carlos A. Ledezma^a, Gilberto Perpiñan^a, Erika Severeyn^a and Miguel Altuve^{*b}

^aGrupo de Bioingeniería y Biofísica Aplicada, Universidad Simón Bolívar, Caracas, Venezuela;

^bFaculty of Electronic Engineering, Pontifical Bolivarian University, Bucaramanga, Colombia

ABSTRACT

Heart diseases are the main cause of death worldwide. The first step in the diagnose of these diseases is the analysis of the electrocardiographic (ECG) signal. In turn, the ECG analysis begins with the detection of the QRS complex, which is the one with the most energy in the cardiac cycle. Numerous methods have been proposed in the bibliography for QRS complex detection, but few authors have analyzed the possibility of taking advantage of the information redundancy present in multiple ECG leads (simultaneously acquired) to produce accurate QRS detection. In our previous work we presented such an approach, proposing various data fusion techniques to combine the detections made by an algorithm on multiple ECG leads. In this paper we present further studies that show the advantages of this multi-lead detection approach, analyzing how many leads are necessary in order to observe an improvement in the detection performance. A well known QRS detection algorithm was used to test the fusion techniques on the St. Petersburg Institute of Cardiological Technics database. Results show improvement in the detection performance with as little as three leads, but the reliability of these results becomes interesting only after using seven or more leads. Results were evaluated using the detection error rate (DER). The multi-lead detection approach allows an improvement from $DER = 3.04\%$ to $DER = 1.88\%$. Further works are to be made in order to improve the detection performance by implementing further fusion steps.

Keywords: QRS Complex Detection, ECG Signal Processing, Optimal Data Fusion, Digital Filtering.

1. INTRODUCTION

Heart diseases are the main cause of death worldwide, being responsible for over 7 million fatalities per year.¹ The diagnose of these illnesses begins with the analysis of the electrocardiographic (ECG) signal. This analysis may be made manually, by a trained specialist, or automatically, by using digital signal processing techniques. When following an automatic scheme, heart beats are usually found through the detection of the QRS complexes, which are the biggest waves and those containing the most of the signal's energy. Other waveforms may then be located and advanced features may be determined in order to perform a more advanced interpretation of the physiological phenomena that are reflected on the ECG signal. It is so that the implementation of accurate QRS complex detectors is of high interest.

Numerous methods have been designed in the search of an automatic and precise QRS complex detector. Different methods use specific techniques that try to minimize false alarms and failed detections. QRS complex detectors may be classified according to the signal processing technique they use² in filter based, wavelet transform based, machine learning based or additional approaches.

The first successful QRS complex detectors were those that used filtering techniques that enhanced particular features of the ECG signal. The main detectors proposed in this field are those of Okada³ and Pan & Tompkins;⁴ these were the first highly effective, real-time, QRS complex detectors. These algorithms used digital filters to eliminate as much noise as possible and take advantage of features such as slope and energy to accurately detect the QRS complex. A novel approach to QRS complex detection was the wavelet-based detection,^{5,6} which allows the use of filter banks to detect interesting features of the ECG. The multiresolution analysis given by the use of these filter banks permits not only QRS complex detection but full ECG delineation. Analysis based in

*Further information: (Send correspondence to M. Altuve.)

M. Altuve: E-mail: miguel.altuve@upb.edu.co; Phone +57 7 679 62 20 ext. 203

mathematical morphology,⁷ dynamic time warping⁸ and Hilbert transform⁹ have also proven their worth in QRS complex detection.

Even when numerous methods have been developed in this field, the topic is still of interest to the scientific community. As computers gain processing capacity, more complex routines may be implemented. More recent approaches, such as the use of machine learning techniques,^{10,11} advanced spectral analysis^{12,13} and complex mathematical techniques,¹⁴ have been made possible thanks to the latest technological advances. Also, recent contributions have taken into account the power requirements of the microcontroller where these routines will be implemented. These contributions have developed methods for their use in smart phones¹⁵ and in low power applications.¹⁶

In this work we are interested in making a contribution to the QRS complex detection topic. We wish to develop a detection method capable of working in real time and that takes advantage of the redundancy of information present in multi-lead ECG recordings. Other studies have explored this possibility,^{17,18} but they are not correctly adapted to real-time applications, since their algorithms are too complicated or have not been properly tested with multiple channel ECG recordings. Real-time applications require simple routines, capable of processing data and producing decisions with minimal latency.

Data fusion techniques have been used in biomedical engineering with different goals.¹⁹⁻²¹ The success of these studies, particularly in event detection, motivates the use of data fusion techniques to perform an accurate QRS complex detection.

This paper is the continuation of the algorithm fusion for QRS complex detection scheme presented previously by us.²² In the following, we take further steps in this research and reach interesting results concerning the use of data fusion techniques for QRS complex detection. Specifically, we show here the results of testing the fusion algorithm with 3 to 11 leads and compare the results to those previously obtained when testing with 2 and 12 leads.²²

The remainder of this paper is organized as follows: section 2 presents the methodology followed to design and test a novel method that combines various ECG leads to produce a single, robust, detection signal; further on, section 3 presents the results of applying the proposed methodology and section 4 discusses these results and measures their validity; finally, conclusions and proposed future works are presented on section 5.

2. METHODOLOGY

2.1 Multi-Lead detection scheme

The block diagram shown in figure 1 is proposed to implement the multi-lead QRS complex detection algorithm. We propose the application of a single detector to M leads of a single ECG recording and an algorithm that uses fusion techniques to decide, using the information coming from each lead, if a detection made in one channel is true or false.

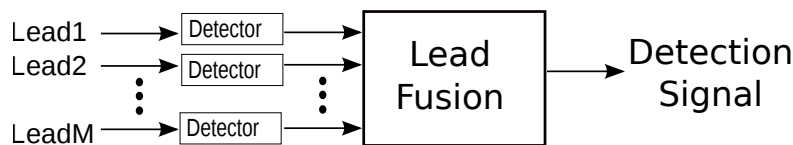


Figure 1. Block diagram for performing multi-lead QRS complex detection.

The fusion algorithm works as follows:

- M detection signals $(d_i[n], i \in 1, \dots, M)$ are monitored at the same time.
- If a signal shows detection a time lapse w_D is opened.
- Detections signaled from different channels during this window are considered simultaneous.
- During w_D a fusion rule is applied, this rule will decide if the detection signal will show detection or not.

- If detection is shown close w_D and open a “blind-eye window” during which new detections are ignored; otherwise close w_D .
- Once w_D is closed, repeat.

Time window w_D is used because having all channels show detection at the exact same time is nearly impossible, so detections are considered simultaneous if they all happen within a time window. Furthermore $w_D = 100$ ms since the mean duration of a QRS complex is between 70–110 ms; the use of this window ensures that detections will be considered simultaneous only if they are within the same QRS complex.

The blind-eye window is of 250 ms since it is not physiologically possible for two heartbeats to occur within this time interval. The use of this window ensures that oversized T waves will not be taken as detections when a QRS complex is correctly detected.

One detector and two fusion techniques were used to test the scheme presented in figure 1, these are explained in the next subsections.

2.2 QRS complex detector

The QRS detection algorithm implemented to test the fusion scheme was the one proposed by Pan and Tompkins.⁴ This detector takes advantage of the QRS complex slope and energy to accurately detect it, it shows a very good performance and has the ability to work in real time applications. This was the detector implemented in our previous work,²² those results will serve to compare the performance when using different amounts of leads. This detector’s block diagram is shown in figure 2. The original algorithm included a searchback for missed beats; this was not implemented in this work since it introduces a detection delay that is undesirable in real-time applications. We expect to compensate the drop in detection performance by implementing the multi-lead approach explained before.

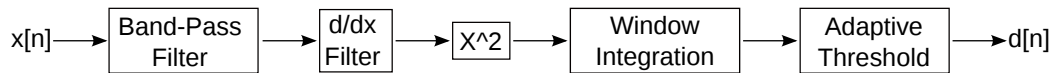


Figure 2. Processing block diagram proposed by Pan and Tompkins⁴ for QRS complex detection.

The detection signal $d[n]$ is a discrete signal containing the sample number where a detection was made. It represents the time when the QRS complexes occur and it may be used to perform further, more advanced, analysis concerning the ECG signal.

The algorithm was tested on each channel of the database, without using the multi-lead detection scheme. These results were used as a base to evaluate the gain of using the multi-lead detection scheme.

2.3 Fusion techniques

As was determined by previous results,²² there are two fusion techniques that prove to be interesting when performing multi-lead QRS complex detection.

The first is the Vote fusion, ruled by equation 1 where u_i is the detection signal coming from lead i , $u_i = 1$ when a detection is made and $u_i = 0$ when no detection is made. This fusion rule establishes that a detection is made when an algorithm shows detection in half or more of the channels.

$$u = \begin{cases} 1, & \sum_{i=1}^M u_i \geq \frac{M}{2} \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

The second fusion technique to be tested in order to perform multi-lead QRS complex detection is the Optimal fusion method proposed by Chair and Varshney.²³ This technique uses equation 2 to determine if a detection is accepted or rejected. This technique measures the probability of occurrence of the event being looked for (QRS complex) by using the a_0 coefficient and the ability of the detector to correctly detect the event on each channel using the a_i coefficients, where each coefficient corresponds to a given lead ($i \in 1, \dots, M$).

$$u = \begin{cases} 1, & a_0 + \sum_{i=1}^M a_i u_i > 0 \\ -1, & \text{otherwise} \end{cases} \quad (2)$$

The a_0 coefficient is calculated as shown in equation 3 where P_1 is the probability of occurrence of the QRS complex and P_0 is the probability of absence of such event.

$$a_0 = \log \frac{P_1}{P_0} \quad (3)$$

Since this fusion technique has never been used in applications such as this, we propose calculating P_1 as the number of samples corresponding to QRS complexes in the database divided by the total number of samples in the database; this corresponds to the probability of a sample being part of a QRS complex, which may be interpreted as the probability of occurrence of a QRS complex. The probability of absence of the QRS complex is complementary to its probability of occurrence, so $P_0 = 1 - P_1$.

The a_i coefficients are calculated as shown in equation 4. One can observe that these coefficients take two values, one when the detector says a QRS complex occurred in a given lead ($u_i = 1$) and another when the detector says the QRS complex didn't occur on a given lead ($u_i = -1$). These coefficients are defined by the probability of missed detection (P_M) and the probability of false alarm (P_F), calculated as shown in equations 5 and 6, respectively.²⁴

$$a_i = \begin{cases} \log \frac{1-P_{M_i}}{P_{F_i}}, & u_i = 1 \\ \log \frac{1-P_{F_i}}{P_{M_i}}, & u_i = -1 \end{cases} \quad (4)$$

$$P_M = 1 - Se \quad (5)$$

$$P_F = 1 - P+ \quad (6)$$

These coefficients must not be learned using all the records of the database since this would cause an overtraining. This would result in a very good performance but a very bad generalization capability. To avoid overtraining, these coefficients were learned using the cross-validation procedure explained in the following section.

2.4 10-fold cross-validation

The *k-fold cross-validation* is a technique used in machine learning applications to train a model before using it on real data. We decided to use $k = 10$ to implement the cross-validation procedure. Furthermore, 75% of the data was randomly selected to perform the cross-validation; from the 10 iterations, the coefficients that resulted in the best evaluation results were saved and tested on the remaining 25% of the data.

Performing the cross-validation in this manner guarantees that the coefficients will have a good generalization capability. It is so that the coefficients found using this cross-validation may be used in further applications without the need of training once again.

2.5 Performance evaluation

The *St. Petersburg Institute of Cardiological Technics 12-lead Arrhythmia Database* (INCART), freely available in Physionet,²⁵ was used to evaluate the method. This database has 75, 12-lead, 30 min, ECG recordings sampled at 257 Hz. Four recordings (I02, I03, I57 and I58) were left out of this study because one lead didn't show any signal and recording I75 was left out to use 70 recordings (which simplifies training and evaluation).

Detection performance was evaluated using sensitivity (Se), positive predictive value ($P+$) and detection error rate (DER) as shown in equations 7, 8 and 9 respectively. In these equations TP corresponds to the sum of all true positives (correctly detected QRS complexes), FP corresponds to the sum of all false positives (detections marked where there was no QRS complex) and FN corresponds to the sum of all false negatives (QRS complex not marked when it should have been).

$$Se = \frac{TP}{TP + FN} \quad (7)$$

$$P+ = \frac{TP}{TP + FP} \quad (8)$$

$$DER = \frac{FP + FN}{TP + FN} \quad (9)$$

To determine TP , FP and FN , the detections made by the proposed algorithm were compared to those supplied by the INCART database as professional annotations. An automatic detection was considered to be a TP if it was within 150 ms of a reference annotation.

2.6 Lead selection

We are interested in testing the multi-lead detection algorithm with $M = 3, \dots, 11$ channels. Since only one database will be used to test the results, it is necessary to define a procedure to select which leads will be used to test the multi-lead detection algorithm.

To avoid bias, leads are to be selected randomly. This random selection may result in a favorable or unfavorable outcome, depending on the noise characteristics of the selected leads. So, we propose to perform the random selection and testing 10 times and report Se , $P+$ and DER as box plots of the 10 tests made for each M .

3. RESULTS

Table 1 shows the result of running the detector on each of the databases without applying the multi-lead fusion algorithm. These results serve as a base to compare if the fusion scheme helps improving the detection results.

Figures 3-5 show the performances (sensitivity in figure 3, positive predictive value in figure 4, and detection error rate in figure 5) of the multi-lead detection approach through Vote and Optimal data fusion, using $M = 3, \dots, 11$ leads.

4. DISCUSSION

Results show that to achieve an improved performance a minimum amount of leads is needed, *i.e.* a minimum amount of information is needed to properly take advantage of the fusion techniques. In table 1 one can observe that the best performance of the Pan and Tompkins algorithm is $DER = 3.04\%$; this is only improved, in average, with 8 or more leads when using fusion through vote and 4 or more leads when using optimal fusion.

On the other hand, figure 5 shows an important advantage in comparison to the single lead approach. The worst test results shown in this figure, $DER = 7.4\%$ in figure 5b, is still an acceptable detection error rate. In contrast, when performing single lead detections, the Pan and Tompkins algorithm shows DER as high as 16%. This result shows that performing fusion effectively counters the effect of noisy leads, improving a detector's performance even in the presence of noise on a particular channel.

Table 1. Results of applying Pan and Tompkins⁴ detector on the 12 channels of the INCART database. In bold the best results.

Channel	Se (%)	$P+$ (%)	DER (%)
1	92.51	97.24	10.12
2	93.63	98.42	7.87
3	92.36	97.94	9.58
4	95.71	98.92	5.34
5	85.50	97.38	16.81
6	92.43	98.17	9.30
7	97.64	99.32	3.04
8	94.75	99.47	5.75
9	94.08	99.76	6.15
10	95.08	99.75	5.15
11	95.13	99.90	4.96
12	96.97	99.81	3.21

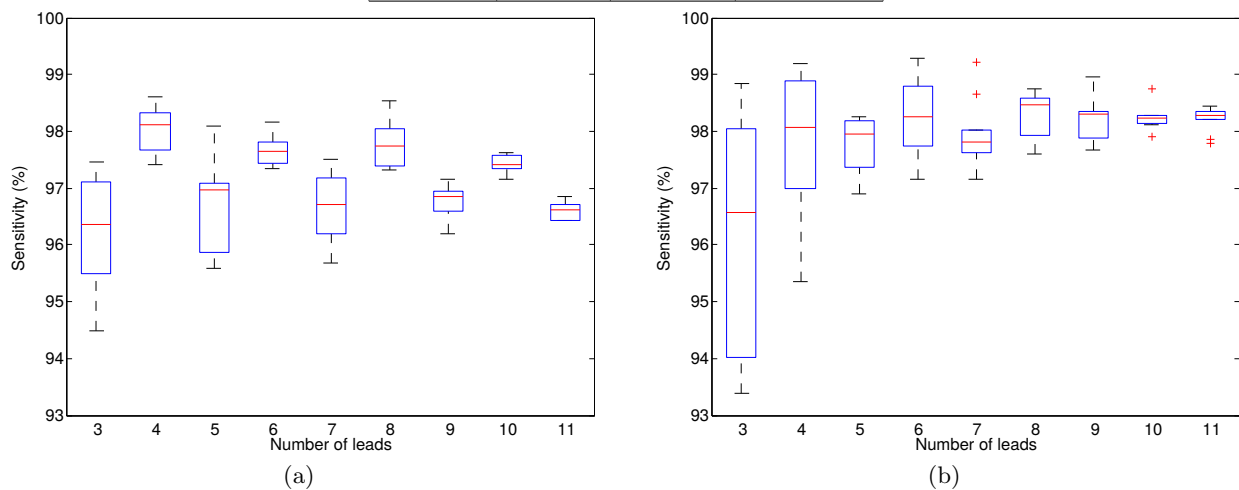


Figure 3. Sensitivity of the detector using Vote (a) and Optimal (b) fusions for $M = 3, \dots, 11$ leads.

Figures 3-5 show another fact that may be noted. The confidence of the results increases as the amount of leads increases. This may be seen in the reduction of the dispersion in the results, which becomes smaller as the amount of leads increases. This result is to be expected, when having less leads to work with the effect of a noisy channel will damage the overall performance in a more direct way; on the other hand, when working with more leads, a noisy channel may be neutralized more easily.

Results also show that the optimal detection method is, as was expected, better than the vote method. It achieves a better performance for any amount of leads. In the best case, when using 11 leads, the optimal fusion shows a DER improvement of 45% over the vote fusion.

Table 2 presents the comparison between the best results obtained in this work, some methods that use data fusion techniques and some methods that have been recently proposed for QRS complex detection.

The detection error rates produced in this work are slightly higher than those of other methods proposed in the literature. It is, nevertheless, important to highlight the advantages of our method over the others shown in table 2. First, our method has a simple implementation, which makes it convenient for real time applications, allowing to increase the reliability of any detector capable of performing such action; the other methods shown in the table have not been tested in real time applications or use complicated algorithms that make them difficult to reproduce. In addition, what was proved with these two algorithms may work as well with different QRS complex detection approaches. Taking advantage of the redundancy present in different ECG leads may improve the performance of any detector. The algorithm proposed in this paper to combine the decisions made on different

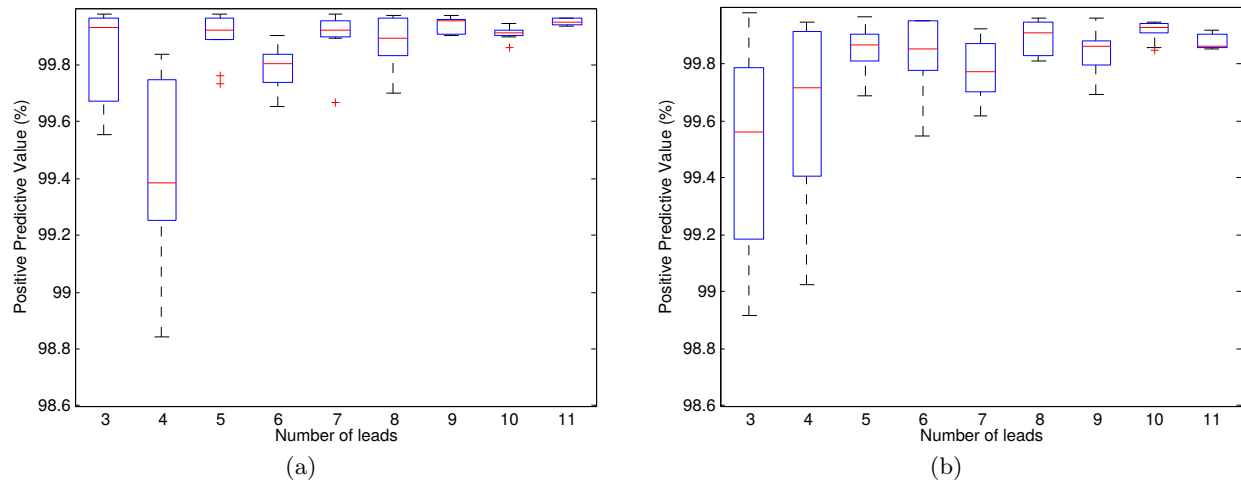


Figure 4. Positive predictive value of the detector using Vote (a) and Optimal (b) fusions for $M = 3, \dots, 11$ leads.

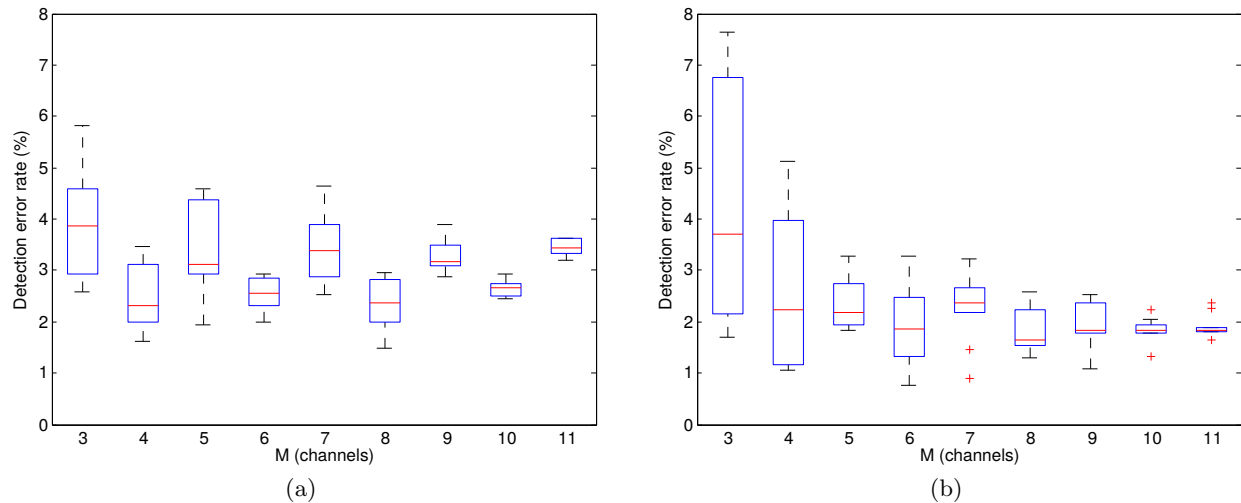


Figure 5. Detection error rate of the detector using Vote (a) and Optimal (b) fusions for $M = 3, \dots, 11$ leads.

leads may be applied to any detection algorithm. Finally, the implementation of the fusion techniques here presented is rather simple, the fusion techniques consist of arithmetic operations that can be implemented with very low computational cost. This makes our work easily reproducible and implementable on microcontrollers that could be included in monitoring devices.

5. CONCLUSIONS AND FUTURE WORKS

In this paper we presented a fusion method that may be used to transform any single-lead QRS complex detection algorithm into a multi-lead QRS complex detection algorithm. The method was tested using a well known QRS complex detection algorithm and a widely used database to validate the results. Additionally, two different fusion techniques (Vote and Optimal) were used to test the fusion algorithm.

The tests made proved that performing data fusion helps improve the detection performance by eliminating false positives and false negatives and thus incrementing the sensitivity and positive predictive value of the detector. This increase in performance is reflected in the decrease of the detection error rate.

Furthermore, results showed that using as little as three leads to perform the fusion is equally advantageous, since using many leads to look for detections helps neutralize the effect of noisy ECG channels. Nevertheless, the more leads available, the more reliable the detections will be.

Table 2. Comparison between the proposed methods and similar and recent approaches. N/R indicates the author did not report the value.

Detection Method	Se (%)	$P+$ (%)	DER (%)
11-Lead Pan and Tompkins (Optimal)	98.21	99.86	1.88
11-Lead Pan and Tompkins (Vote)	96.85	99.96	3.18
Wavelet+PT ²⁶	N/R	N/R	0.63
PCA+WaveletEntr ¹⁷	99.99	99.98	0.02
Paired Windows ¹⁸	99.85	99.86	0.29
Feature template ²⁷	99.81	99.70	N/R
Cuadratic Filter ²⁸	99.82	99.81	0.38

Future works will concentrate in the fusion of the detection results of many multi-lead adapted detectors, using the method described in this work, to produce a multi-lead/multi-detector approach that should produce better detection results and may produce better results than those proposed up to the date in the bibliography.

In the future, we will also concentrate on performing tests on other databases, such as those available in GBBANet²⁹ to test the algorithms here presented using new signals.

REFERENCES

- [1] WHO, "The top 10 causes of death," (May 2014). <http://www.who.int/mediacentre/factsheets/fs310/en/>.
- [2] Kohler, B.-U., Hennig, C., and Orglmeister, R., "The principles of software QRS detection," *Engineering in Medicine and Biology Magazine, IEEE* **21**(1), 42–57 (2002).
- [3] Okada, M., "A digital filter for the QRS complex detection," *Biomedical Engineering, IEEE Transactions on* (12), 700–703 (1979).
- [4] Pan, J. and Tompkins, W. J., "A real-time QRS detection algorithm," *Biomedical Engineering, IEEE Transactions on* (3), 230–236 (1985).
- [5] Li, C., Zheng, C., and Tai, C., "Detection of ECG characteristic points using wavelet transforms," *Biomedical Engineering, IEEE Transactions on* **42**(1), 21–28 (1995).
- [6] Martínez, J. P., Almeida, R., Olmos, S., Rocha, A. P., and Laguna, P., "A wavelet-based ECG delineator: evaluation on standard databases," *Biomedical Engineering, IEEE Transactions on* **51**(4), 570–581 (2004).
- [7] Trahanias, P., "An approach to QRS complex detection using mathematical morphology," *Biomedical Engineering, IEEE Transactions on* **40**(2), 201–205 (1993).
- [8] Vullings, H., Verhaegen, M., and Verbruggen, H., "Automated ECG segmentation with dynamic time warping," in [*Engineering in Medicine and Biology Society, 1998. Proceedings of the 20th Annual International Conference of the IEEE*], 163–166, IEEE (1998).
- [9] Benitez, D., Gaydecki, P., Zaidi, A., and Fitzpatrick, A., "A new QRS detection algorithm based on the Hilbert transform," in [*Computers in Cardiology 2000*], 379–382, IEEE (2000).
- [10] Saini, I., Singh, D., and Khosla, A., "QRS detection using K-Nearest Neighbor algorithm (KNN) and evaluation on standard ECG databases," *Journal of Advanced Research* **4**(4), 331–344 (2013).
- [11] Arbateni, K. and Bennia, A., "Sigmoidal radial basis function ANN for QRS complex detection," *Neurocomputing* (2014).
- [12] Sahoo, J. P., Das, M. K., Ari, S., and Behera, S., "Autocorrelation and Hilbert transform-based QRS complex detection in ECG signal," *International Journal of Signal and Imaging Systems Engineering* **7**(1), 52–58 (2014).
- [13] Zidelmal, Z., Amirou, A., Ould-Abdeslam, D., Moukadem, A., and Dieterlen, A., "QRS detection using S-Transform and Shannon energy," *Computer methods and programs in biomedicine* **116**(1), 1–9 (2014).
- [14] Nallathambi, G. and Principe, J., "Integrate and Fire Pulse Train Automaton for QRS detection," *IEEE Transactions on Biomedical Engineering* (2014).
- [15] Wu, X., Chu, C.-H., Yue, D., Kim, J., and Li, S., "A simple and effective algorithm for R-wave detection using smartphones," in [*Electronics, Computer and Applications, 2014 IEEE Workshop on*], 902–905, IEEE (2014).

- [16] Górriz, J. M., Ramírez, J., Olivares, A., Padilla, P., Puntonet, C. G., Cantón, M., and Laguna, P., “Real Time QRS Detection Based on M-ary Likelihood Ratio Test on the DFT Coefficients,” *PLoS one* **9**(10), e110629 (2014).
- [17] Huang, B. and Wang, Y., “QRS complexes detection by using the principal component analysis and the combined wavelet entropy for 12-lead electrocardiogram signals,” in [*Computer and Information Technology, 2009. CIT'09. Ninth IEEE International Conference on*], **1**, 246–251, IEEE (2009).
- [18] Torbey, S., Akl, S. G., and Redfearn, D. P., “Multi-lead QRS detection using window pairs,” in [*Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*], 3143–3146, IEEE (2012).
- [19] Hernández, A. I., Carrault, G., Mora, F., Thoraval, L., Passariello, G., and Schleich, J.-M., “Multisensor fusion for atrial and ventricular activity detection in coronary care monitoring,” *Biomedical Engineering, IEEE Transactions on* **46**(10), 1186–1190 (1999).
- [20] Kannathal, N., Acharya, U. R., Ng, E., Krishnan, S., Min, L. C., and Laxminarayan, S., “Cardiac health diagnosis using data fusion of cardiovascular and haemodynamic signals,” *Computer methods and programs in biomedicine* **82**(2), 87–96 (2006).
- [21] Shamim, N., Atul, M., Gari D, C., et al., “Data fusion for improved respiration rate estimation,” *EURASIP journal on advances in signal processing* **2010** (2010).
- [22] Ledezma, C. A. and Altuve, M., “Fusión de datos para detectar complejos QRS en registros electrocardiográficos multicanal,” in [*V Congreso Venezolano de Bioingeniería, Mérida, Venezuela*], 118–121 (2015).
- [23] Chair, Z. and Varshney, P., “Optimal data fusion in multiple sensor detection systems,” *Aerospace and Electronic Systems, IEEE Transactions on* (1), 98–101 (1986).
- [24] Sörnmo, L. and Laguna, P., [*Bioelectrical signal processing in cardiac and neurological applications*], Academic Press (2005).
- [25] Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C.-K., and Stanley, H. E., “Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals,” *Circulation* **101**(23), e215–e220 (2000).
- [26] Fernandez, J., Harris, M., and Meyer, C., “Combining algorithms in automatic detection of R-peaks in ECG signals,” in [*Computer-Based Medical Systems, 2005. Proceedings. 18th IEEE Symposium on*], 297–302, IEEE (2005).
- [27] Karimipour, A. and Homaeinezhad, M. R., “Real-time electrocardiogram P-QRS-T detection–delineation algorithm based on quality-supported analysis of characteristic templates,” *Computers in biology and medicine* **52**, 153–165 (2014).
- [28] Phukpattaranont, P., “QRS detection algorithm based on the quadratic filter,” *Expert Systems with Applications* **42**(11), 4867–4877 (2015).
- [29] Ledezma, C. A., Severeyn, E., Perpignan, G., Altuve, M., and Wong, S., “A new on-line electrocardiographic records database and computer routines for data analysis,” in [*Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*], 2738–2741, IEEE (2014).