ECG-Waves: Analysis and Detection by Continuous Wavelet Transform

Mounaim AQIL, Atman JBARI , Abdennasser BOUROUHOU
Electrical Engineering Research Laboratory
High School of Technical Education (ENSET), Mohamed V University
Rabat, Morocco
mounaim_aqil@um5.ac.ma

Abstract— **In this work, we have developed a new algorithm for electrocardiogram (ECG) features extraction. This algorithm was based on continuous wavelet transform (CWT). The core of the process involved analyzing the signal using the CWT coefficients with a selection of scale parameter corresponding to each ECG wave. The entry point of our method was the R peak detection. The next step was the Q and S point localization, after we identified the P and T waves. We evaluated our algorithm on apnea and MIT-BIH databases recording. The algorithm achieved a good performance with the sensitivity of 99.84 % and the positive predictive value of 99.53 %.**

Index Terms— **Continuous Wavelet Transform; ECG; Time-Scale Analysis; Wave Detection.**

I. INTRODUCTION

The electrocardiogram (ECG) signal reflects the activity of the heart. This signal provides important and relevant information about the heart state. This is why the analysis of this signal has significant importance in cardiology. A typical ECG tracing consists of different waves: P wave, QRS complex, T wave and U wave. A normal ECG graph is shown in figure 1.
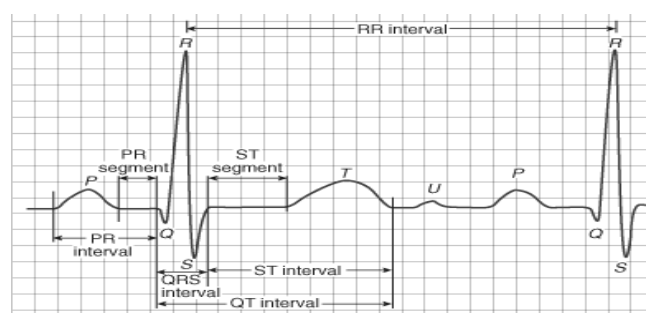

Figure 1 ECG representation in Time-domain

The QRS complex is the most prominent waveform within the ECG. It serves several functions, which are as the basis for the automated determination of the heart rate and an entry point for classification schemes of the cardiac cycle. Often, it is also used in ECG data compression algorithms.

The development of software that allows the detection of the QRS, P and T waves has been a research topic for more than 40 years. In the literature, many methods and tools have been developed. The main approaches for ECG features detection are such as the QRS complex, P and T waves, which are given as follows [1-9]:

- Approaches based on signal derivatives and digital filters: Almost all QRS detection algorithms use a high-pass filter to attenuate P and T waves as well as baseline drift, the suppression of noise is accomplished by a low-pass filter.
- Neural networks approaches: Neural networks have been used as adaptive non-linear predictors. The objective is to predict the current signal value from its past values, and therefore, apply suitable filters to attenuate the noise.
- Hidden Markov Models: HMMs model observes the data sequence by a probability function that varies according to the state of an underlying (hidden) Markov chain. The objective of the algorithm is to infer the underlying state sequence from the observed signal.
- Genetic algorithms: The aim is to get optimal polynomial filters, for preprocessing stage, and parameters for decision stage.
- Hilbert transform based QRS detection: In these approaches, the ideal Hilbert transform is the approximation by a band limited FIR filter. The Hilbert transform of the ECG signal is used for the computation of the signal envelope. In order to remove ripples from the envelope and to avoid ambiguities in the peak level detection, the envelope is low-pass filtered.
- Syntactic methods: In ECG processing, the signal is split into short segments of a variable or a fixed length. Each segment is then represented by a primitive and coded component using a predefined alphabet.
- Wavelet based QRS detection: Some algorithms are based on the wavelet transform because ECG signals are intrinsically non-stationary. Wavelet transforms have been applied to ECG signals to enhance QRS to delineate the ECG features and to reduce computation time.

The method described in this work is based on continuous wavelet transform (CWT) for the following reasons:
- The discrete wavelet transform (DWT) could lose frequency resolution due to resampling at each decomposition level [10].
- The CWT keeps a good frequency resolution.
- With CWT, the appropriate and dominant scale for each component of ECG signal can be extracted, which makes it possible to detect each component separately [11]

In this paper we proposed a new algorithm based on the selection of a scale parameter of CWT coefficients

corresponding to each wave of ECG signal. The proposed algorithm was tested on several ECG records from apnea database with 100 samples per second and 16 bits per sample [12].

On the other hand, the CWT is certainly a powerful tool to accurately locate signal features under the condition of choosing a good wavelet mother. This is why, in Section II, a method for wavelet mother choice has been proposed. The proposed algorithm is described in Section III, whereas the results and discussion are given in Section IV. Finally, Section V summarizes our work and gives some future scope.

The analysis of ECG signal is often divided into a preprocessing stage, including linear and nonlinear filtering and a decision stage, including the detection of ECG wave. In this paper, we are interested only in an algorithm of the detection of ECG wave.

II. BASES OF CWT AND CRITERION CHOICE OF MOTHER WAVELET

A. Continuous wavelet transform

An efficient method to provide localization of a signal in both time and frequency domain is given by Wavelet Transform (WT). This transform is basically a time-scale representation of continuous-time signals. The CWT reflects the correlation between the analyzed continuous-time signal $x(t)$ and a function referred to as wavelets and is defined by the following formula [13 -15] :

$$C(\tau,\sigma) = \int_{-\infty}^{+\infty} x(t)\,\psi_{\tau,\sigma}^*(t)dt \qquad (1)$$

Where:

$$\psi_{\tau,\sigma}(t) = \frac{1}{\sqrt{\sigma}}\psi\left(\frac{t-\tau}{\sigma}\right) \qquad (2)$$

The σ is the dilatation of wavelet (scale) and τ defines a translation of the wavelet and indicates the time localization, $\psi^*(t)$ is the complex conjugate of the analyzing mother wavelet $\psi(t)$. The coefficient $\frac{1}{\sqrt{\sigma}}$ is an energy normalized factor (the energy of the wavelet must be the same for different σ value of the scale).

The scale σ can be converted to frequency ($1/\sigma$ reflects the frequency). The τ parameter represents the location of the wavelet along the time axis.

In order to be a wavelet, the function $\psi(t)$ must be limited in time and satisfy the following conditions [16, 17]:

$$\begin{aligned}\Psi(\omega)|_{\omega=0} &= \int_{-\infty}^{+\infty}\psi(t)e^{-j\omega t}dt\,|_{\omega=0}\\ &= \int_{-\infty}^{+\infty}\psi(t)dt\\ &= 0\end{aligned} \qquad (3)$$

B. Criterion choice of mother wavelet for ECG analysis

In literature, there is no predefined rule to select a wavelet for a particular application, except the selection is application oriented. It is a common practice to select a wavelet function that has similar shape as of the subject signal, which is given in Figure 2.
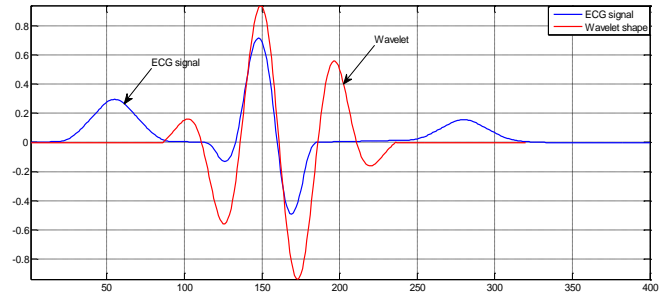

Figure 2: Representation of ECG's waveform and wavelet's shape

In this section, we proposed a criterion choice of mother wavelet based on the inequality of Cauchy-Schwartz. Indeed, this inequality is given as follow: [18, 19]

$$\begin{aligned}&\left|\int_{-\infty}^{+\infty} x(t)\,\psi_{\tau,\sigma}^*(t)dt\right|\\ &\leq \int_{-\infty}^{+\infty}|x(t)|^2 dt\,.\int_{-\infty}^{+\infty}\left|\psi_{\tau,\sigma}^*(t)\right|^2 dt\end{aligned}^2 \qquad (4)$$

Where:
- $\int_{-\infty}^{+\infty}|x(t)|^2 dt$: The energy of the signal
- $\int_{-\infty}^{+\infty}|\psi_{\sigma,\tau}^*(t)|^2 dt$: The energy of the wavelet.

The equality of the two members of the inequality implies the colinearity of $x(t)$ and $\psi(t)$.
Let us define the collinearity ratio L given below:

$$L = \frac{\left|\int_{-\infty}^{+\infty} x(t)\,\psi_{\tau,\sigma}^*(t)dt\right|^2}{\int_{-\infty}^{+\infty}|x(t)|^2 dt\,\int_{-\infty}^{+\infty}\left|\psi_{\tau,\sigma}^*(t)\right|^2 dt} \qquad (5)$$

The discrete form of the equation (5) is given in the following formula:

$$L = \frac{\sum_{n=1}^{N}(\sum_{k=1}^{K}|C_x(n,\sigma_k)|^2)}{\sum_{n=1}^{N}|x(n)|^2\,.\sum_{n=1}^{N}|\psi(n)|^2} \qquad (6)$$

Where:
- $C_x(n,\sigma_k)$ is the CWT coefficients.
- N is the number of samples.
- K is the number of scales.
- $x(n)$ is the sample of signal $x(t)$ at time t_n
- $\psi(n)$ is the sample of psi function of the wavelet.

The criterion choice of mother wavelet involves computing the collinearity ratio L for each type of mother wavelet and choosing which one gives the best ratio. Table 1 summarizes the result of the method with the following conditions:
- The ECG signal used is the *a01m* record of apnea database [12];
- The scale parameters are σ_k ; $k= 1...32$ and $\sigma_1 = 1$, $\sigma_k = \sigma_{k-1} + 1$

Table 1
Collinearity ratio for different mother wavelet

Mother wavelet name	Collinearity Ratio L
Morl	0.9621
Meyr	0.9508
Mexh	0.8721
Gaus 4	0.5897
Coif 4	0.0606
Sym 4	0.0598
Db 4	0.0598
Haar	0.0549

As shown in the Table 1, Morlet's mother wavelet gives a best collinearity ratio L to conserve the maximum of localized energy splint along the acquisition period. For this important result, we used the Morlet's mother wavelet to compute the CWT coefficients of the ECG signal in our paper.

III. ALGORITHM OVERVIEW

The algorithm is based on CWT coefficients with a selection of scale parameter for each wave of the ECG signal. The process starts with a detection of the R peaks; next a localization of the Q and S points followed by the identification of the P and T waves. The steps of the algorithm are given in the Figure 3.

A. R peak detection

The signal was analyzed in a set of scales using a Morlet's wavelet.

The wavelet coefficients were exploited to compute the following distributions: [20]

- Time-frequency content: evaluate the energy $E_x(n)$ of the signal $x(n)$ presented at the time t_n based on the result coefficients $C_x(n, \sigma_k)$, $k= 1...K$ of the applied transform. It can be expressed as,

$$E_x(n) = \sum_{k=1}^{K} |C_x(n, \sigma_k)|^2 \qquad (7)$$

- Relative Time-frequency content: In order to study the most important contribution at each t_n, we proposed to evaluate the rate of absolute maximum coefficient and the total energy as follow:

$$RE_x(n) = \frac{max_n(|C_x(n, \sigma_k)|^2)_{k=1,\cdots,K}}{E_x(n)} \qquad (8)$$

With the aim for automatically selecting a scale parameter for R peak detection, we proceeded as follows:

- We computed the CWT coefficients $C(n, \sigma_k)$
- We extracted the maximum of the $C(n, \sigma_k)$ at each time point and store them in a vector named y_m, the simulation of this step as shown in Figure 4.
- We computed the peaks of y_m and looked for all the scales corresponding to these peaks. The average of these scales gives us the selected scale called σ_R.
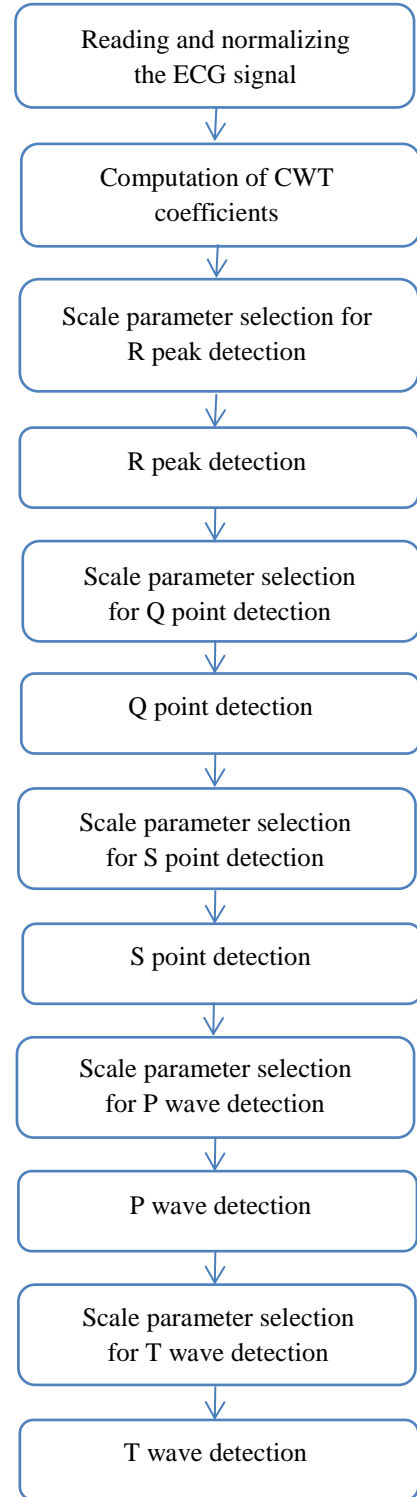

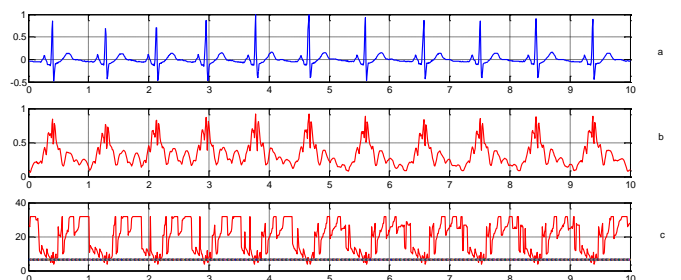Figure 3: Description of the algorithm.

Figure 4: (a) ECG signal. (b) Maximum of wavelet coefficients.
(c) Selection of scale parameter (selected scale is plotted in black line)

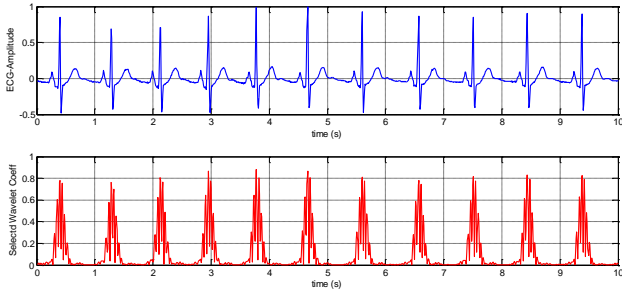The analysis of the signal at σ_R is given in Figure 5.

Figure 5: (a) ECG signal. (b) Wavelet coefficients at the selected scale $\boldsymbol{\sigma_R}$

With the objective to detect the R peak starting from the analysis of the signal at the selected scale σ_R, we did as follows:

- We computed the vector M_C containing the mean value $m(k)$ of wavelet coefficients at σ_R along a sliding window *w*.
- We computed the histogram of vector y_m to determine the distribution of the wavelet coefficients for the different samples of the ECG signal. The histogram is given in figure 6.
- We defined threshold using the centroid of the histogram method as given bellow:

$$th = \frac{\sum_{i=1}^{n}(C_i * N_i)}{\sum_{i=1}^{n} N_i} \tag{9}$$

Where *th* is the threshold, C_i is the wavelet coefficient value, N_i is the number of samples and *n* is the histogram range.
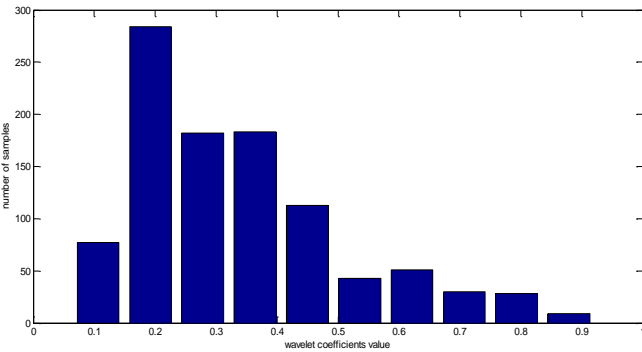

Figure 6: Distribution of wavelet coefficient at a selected scale

We detected the peaks in M_C using the threshold value defined in formula 9. These peaks deal with the R peaks of ECG signal. The results of this step are given in Figure 7.
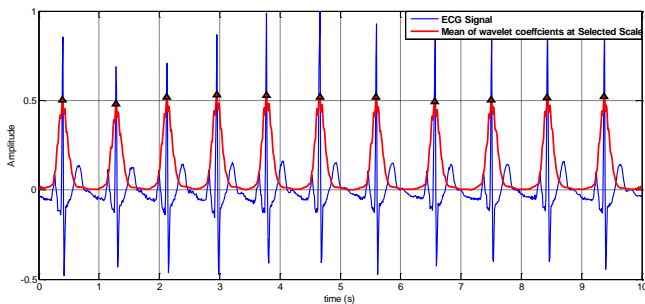

Figure 7: R peak detection

B. Q and S points detection

Once the R peaks are detected, the Q and S points are the next features to be located in order to complete the QRS complex. Generally, the Q and S waves have high frequency and low amplitude.

In order to select the scale parameter corresponding to Q wave called σ_Q , we followed the given steps:

- We localized the local minima of relative time-frequency content $RE_x(n)$ before each R peak position as given in figure 8.
- We seek the positions of those local minima in the vector y_m.
- We looked for the scales corresponding to those positions.
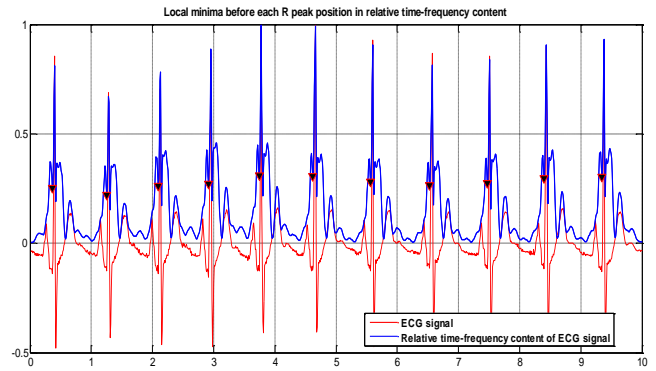- The average of the obtained scales gives us σ_Q.

Figure 8: Local minima before each R peak position in relative time-frequency content

To select the scale parameter corresponding to S wave called σ_S we conducted the same process that we performed for Q wave. The only difference is to localize the local minima after each R peak, as given in Figure 9.
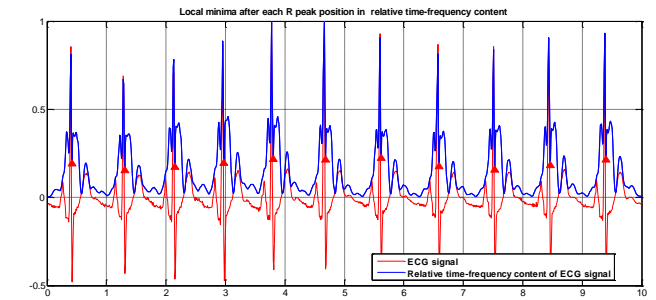

Figure 9: Local minima after each R peak position in relative time-frequency content

Once the scale parameters σ_Q and σ_S are defined we computed the CWT coefficients C_{σ_Q} and C_{σ_S}.
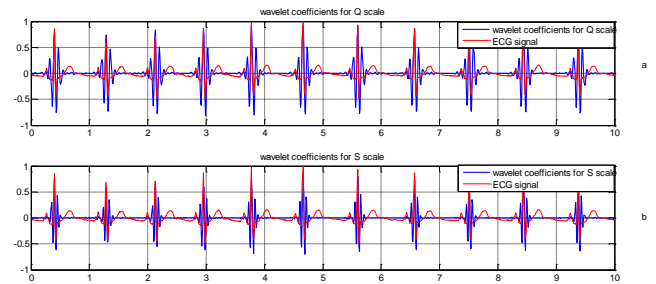The plots of those coefficients are given in Figure 10.

Figure 10: (a) Wavelet coefficients $\mathbf{C_{\sigma_Q}}$. (b) Wavelet coefficients $\mathbf{C_{\sigma_S}}$

Q and S points are the inflexion points in either side of R peak. It is, thus enough to locate the first zero slope in either side of R peak.

For this purpose we did a differentiation of C_{σ_Q} and C_{σ_S}, then we detected the zero slopes around the position of R peak. The result of this process is given in Figure 11.
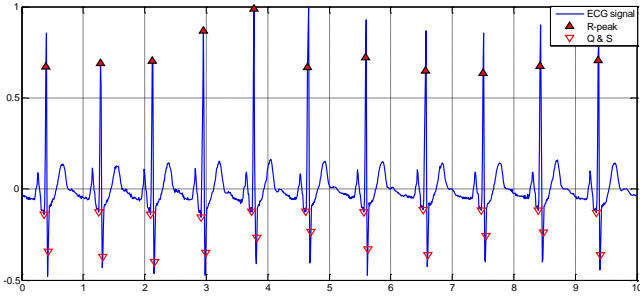

Figure 11: Detection of Q and S waves

C. Detection of P and T waves

Once the QRS complex is completely localized, we proceeded to the next step, which consists of locating the P and T waves of ECG signal.

In order to select the scale parameter corresponding to P wave called σ_P , the following steps were carried out:

- We computed the energy of the signal locally in a sliding window w according to the following formula:

$$E_w = \sum_{n \epsilon w} \sum_{k=1}^{K} |C(n, \sigma_k)|^2 \qquad (10)$$

- We seek the first maxima before each Q point position in E_w. The simulation result of this step is shown in Figure 12.
- We localized the positions of those maxima in the vector y_m.
- We looked for the scales corresponding to those positions.
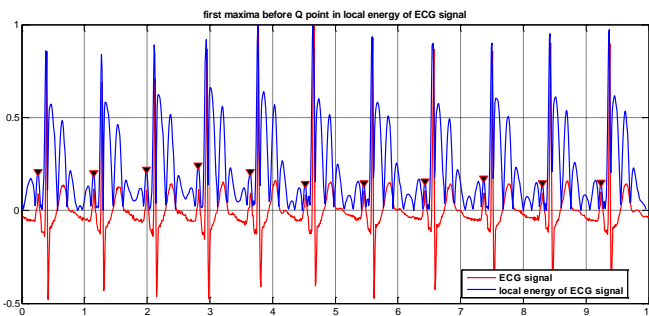- The average of obtained scales gives us σ_P .

Figure 12: Local maxima before each Q point position

To select the scale parameters corresponding to T wave called σ_T , we followed the same steps given below and we looked for the first maxima after the S point positions. The simulation result is shown in Figure 13.
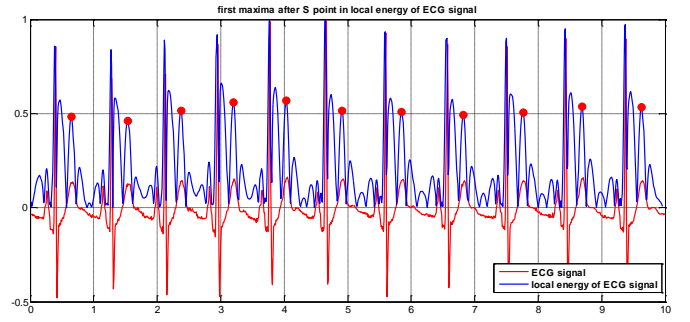

Figure 13: Local maxima after each S point position

Once the scale parameters corresponding to P and T waves were selected, we computed the CWT coefficients at these scales. The plot of CWT coefficients C_{σ_P} and C_{σ_T} are given in Figure 14.
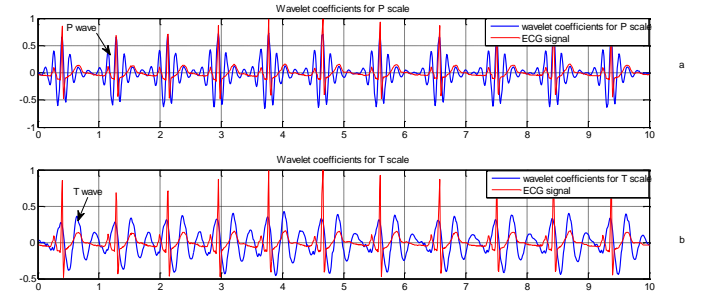
Figure 14: (a) Wavelet coefficients $\mathbf{C_{\sigma_P}}$. (b) Wavelet coefficients $\mathbf{C_{\sigma_T}}$

The P wave is identified as the first maxima before each Q point position in the plot of C_{σ_P}. For this goal, we did the derivative of C_{σ_P} named C'_{σ_P}. Then we looked for the point P corresponding to the following conditions:

- The point P is before the Q point.
- $C'_{\sigma_P}(P) = 0$.
- C_{σ_P} is increasing before P and decreasing after.

The result of this step is given in Figure 15.

Figure 15: P wave detection

In order to detect the T wave, we applied the same method above since the T wave is identified as the first maxima after each S point position in the plot of C_{σ_Q}. The result of this detection is shown in Figure 16.

Figure 16: T wave detection

IV. RESULT AND DISCUSSION

In the previous section, we presented the different steps of the proposed algorithm. It appears that the R peak detection is an entry point for de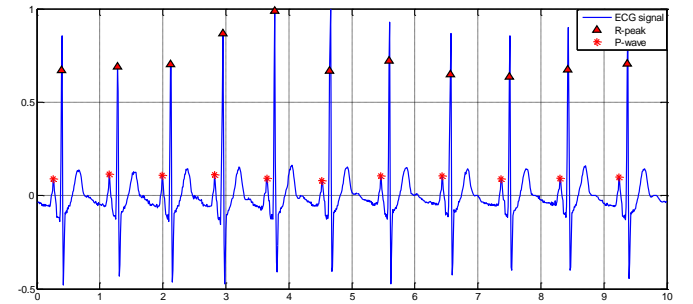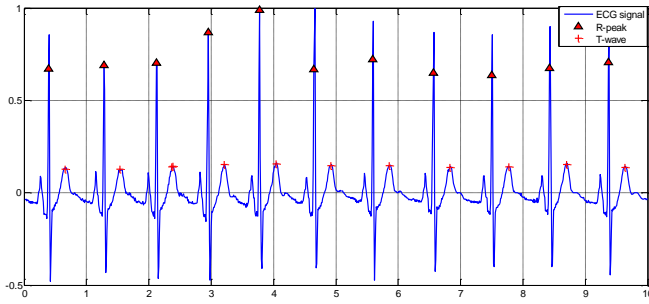tection of the other waves. Thus, in order to validate our method, we applied it to several records from apnea database. We took only 10 seconds of duration from each recording. The result is summarized in Table 2.

Table 2:
Recognition accuracy of R peak detection

Record No	Total beats	Beats correctly Recognized	Recognition accuracy in %
a01	11	11	100
a04	14	14	100
a05	11	11	100
a06	9	8	88.89
a07	12	12	100
a08	15	15	100
a09	12	12	100
a10	10	10	100
a11	10	10	100
a12	14	14	100
a13	13	13	100
a14	12	12	100
a15	10	10	100
a16	13	13	100
a17	12	12	100
a18	11	11	100
a19	12	12	100
Mean of recognition accuracy			99.35 %

On the other hand, in the apnea database, an annotation file named *.qrs* is associated to each recording. This file was created using the toolkit WFDB [21] of physiobank. The *.qrs* can give the position time of QRS complex in each recording. Using on this fact, we can compared the position time of QRS provided by our algorithm and those given in the annotation file. Table 3 shows the time position of QRS complex in seconds for a01 recording.

Table 3
Time position of QRS complex

qrs file	Algorithm	Error
0.37 s	0.37s	0
1.25 s	1.25 s	0
2.09 s	2.09 s	0
2.92 s	2.92 s	0
3.75 s	3.74 s	0.01
4.63 s	4.62 s	0.01
5.57 s	5.56 s	0.01
6.54 s	6.53 s	0.01
7.48 s	7.47 s	0.01
8.41 s	8.4 s	0.01
9.35 s	9.34 s	0.01
Mean of error		0.0063

As given in Table 3, we can say that our algorithm localizes approximately the QRS complex at the same position time of the annotation file.

The detection of different waves for a01 recording is given in Figure 17
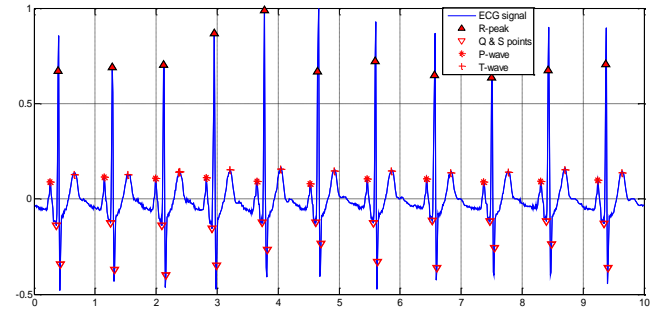

Figure 17: P,QRS and T waves detection

In order to evaluate our algorithm and make a comparison with other methods, we did the test on records from MIT-BIH database [22]. Each record has its manually annotation file. With the objective to quantify the evaluation, the following statistical parameters of performance of QRS detection have been used:

- Sensitivity:

$$Se(\%) = \frac{TP}{TP + FN} \times 100 \qquad (11)$$

- Positive predictive value:

$$P^+(\%) = \frac{TP}{TP + FP} \times 100 \qquad (12)$$

- Detection Error Rate (DER):

$$DER(\%) = \frac{FN + FP}{total\ beats} \times 100 \qquad (13)$$

Note that TP is the number of true positive QRS detections, FN stands for the number of false negative detections and FP represents the numbers of false positive misdetections.

Table 4 illustrates the QRS detection rates for records of the MIT-BIH database. The algorithm achieved a good performance with the sensitivity of 99.84 % and the positive predictive value of 99.53 %.

Table 4
Performance of QRS detection

record	Beats	TP	FN	FP	Se (%)	P+ (%)	DER (%)
100	2274	2273	1	1	99,95	99,96	0,088
101	1874	1874	0	14	100	99,26	0,747
102	2192	2191	1	15	99,95	99,32	0,729
103	2091	2085	4	2	99,80	99,90	0,286
104	2311	2304	4	16	99,82	99,31	0,865
105	2691	2682	4	16	99,85	99,41	0,743
106	2098	2094	4	3	99,80	99,86	0,333
107	2140	2137	3	12	99,85	99,44	0,700
108	1824	1819	5	25	99,72	98,64	1,644
109	2535	2534	1	11	99,96	99,57	0,473
111	2133	2126	7	13	99,67	99,39	0,93
112	2550	2547	3	10	99,88	99,61	0,5
113	1796	1795	1	10	99,94	99,45	0,61
114	1890	1879	7	10	99,62	99,47	0,89
115	1962	1951	7	2	99,64	99,90	0,45
116	2421	2410	7	4	99,71	99,83	0,45
117	1539	1539	0	10	100	99,35	0,64
118	2301	2300	1	16	99,95	99,31	0,73
119	2094	2084	5	27	99,76	98,72	1,52
121	1876	1876	0	11	100	99,42	0,58
122	2479	2477	2	0	99,91	100	0,08
123	1519	1519	0	15	100	99,02	0,98
124	1634	1632	2	4	99,87	99,76	0,36
200	2792	2779	5	10	99,82	99,64	0,53
201	2039	1978	8	0	99,59	100	0,39
202	2146	2141	5	4	99,76	99,81	0,41
203	3108	3104	4	10	99,87	99,68	0,45
205	2672	2660	5	0	99,81	100	0,18
All	60981	60790	96	271	99,84	99,53	0,62

.

Table 5 compares the performance of our algorithm with other well-known works. [10]

Table 5
Comparison of QRS detector performance

Algorithm	Se (%)	P+(%)	DER(%)
Karimipour-Homaeinezhad	99.81	99.7	0.49
Pan-Tompkins	99.75	99.53	0.675
Li et al.	99.89	99.94	0.14
Poli et al.	99.6	99.5	0.9
Madeiro et al.	99.15	99.18	1.69
Zidelman et al.	99.64	99.82	0.54
Yochum et al.	99.85	99.48	0.67
Martinez et al.	99.8	99.86	0.34
Hamilton-Tompkins	99.69	99.77	0.54
Our work	99.84	99.53	0.62

V. CONCLUSION

In this paper, we presented a criterion to choose the mother wavelet for ECG analysis. The idea of this criterion is to find the type of wavelet given the best collinearity with the ECG signal. Indeed, we proved that Mortlet's mother wavelet is best suited to the analysis of the ECG signal since it has the best collinearity ratio.

We also presented in this work, a new algorithm based on continuous wavelet transform for detection of QRS, P and T waves of ECG signal. This algorithm exploits a technique of selection of scale parameter for the wavelet coefficients. The selection has been carried out according to the power spectra of different waves of the signal and taking into account the fact that smaller scales correspond to high frequency components and higher scales correspond to low frequency components of the signal.

The advantages of this selection are multiple, which are the elimination of the noise and interference with other part of signal while extracting a specific wave, and the reduced computing time of wavelet coefficients that makes a reduced duration of the analysis.

In future, the perspective of this work can be considered to implement the proposed algorithm on DSP and to test the features extraction in real time.

REFERENCES

[1] B.-U. Kohler, C. Hennig, and R. Orglmeister," *The principles of software QRS detection*", Engineering in Medicine and Biology Magazine, IEEE, vol. 21, no. 1, pp. 42–57, 2002.

[2] Y. Min & al, "*Design of wavelet-based ECG detector for implantable cardiac peace maker*", IEEE transactions on biomedical ciruits and systems, vol. 7, pp. 426-436, 2013.

[3] G. Vega-Martínez, C. Alvarado-Serrano, L. Leija-Salas, " *Wavelet packet based algorithm for QRS region detection and R/S wave identification*", IEEE conference publications,2015.

[4] J. Pan, W. J. Tompkins," *A Real-time QRS detection algorithm*", IEEE Trans. Biomed. Eng., vol. BME-32, pp.230-236, 1985.

[5] P. Laguna & al.," *New algorithm for QT interval analysis in 24-hour Holter ECG: performance and applications*", Medical & Biological Engineering & computing, Vol-28, pp.67-73, 1990.

[6] Rémi Dubois," *Application of new learning methods for earlier detection of abnormalities in ECG*", Ph.D. dissertation, Dept. Elect. Eng., Paris 6 University, 2004.

[7] S. Mallat," *Zero-crossings of a wavelet transform*", IEEE Trans. Inform. Theory, vol. 37, pp. 1019-1033, 1991.

[8] S. Barro, M. Fernandez-Delgado, J.A. Vila-Sobrino, C.V. Regueiro, and E. Sanchez," *Classifying multichannel ECG patterns with an adaptive neural network*", IEEE Eng. Med. Biol. Mag., vol. 17, pp. 45-55, 1998.

[9] D.A.Coast, R.M.Stern, G.G.Cano and S.A. Briller,"*An approach to cardiac arrhythmia analysis using hidden Markov model*", IEEE Trans. Biomed. Eng., vol. 37, pp. 826-836, 1990.

[10] M.Yochum, Ch. Renaud, S. Jacquir," *Automatic detection of P, QRS, and T patterns in 12 leads ECG signal based on CWT*", Biomedical signal processing and control, Elsevier, vol. 25, pp 46-52, 2016.

[11] A.Ghaffari, H.Golbayani, M.Ghasemi, *"A new mathematical based QRS detector using continuous wavelet transform"*, Computers and electrical engineering, Elsevier, vol.34,pp.81-91, 2008.

[12] Physiobank databases [Online]. *http://www.physionet.org/physiobank/database/apnea-ecg.*

[13] C. S. Burrus, R. A. Gopinath, H. Guo," *Introduction to wavelets and wavelet transforms: A Primer*", Prentice Hall Inc, 1998.

[14] S. Qian, D. Chen,"*Joint time-frequency analysis*", Englewood Cliffs, N.J., Prentice-Hall, 1996

[15] L.Cohen," *Time–frequency analysis*", Prentice-Hall, New York, 1995.

[16] L. Cohen, P. Loughlin," *Recent developments in time-frequency Analysis*", Springer, 1998.

[17] C. Heil, David F. Walnut," *Fundamental papers in wavelet theory*", Princeton University Press, 2009.

[18] S. MALLAT , "*une exploration des signaux en ondelettes*", pp160-170 Éditions de l'École Polytechnique, 2000.

[19] Ch. Bernard, " *Wavelets and ill posed problems: optic flow and scattered data interpolation*", PhD thesis of polytechnic school, November 1999.

[20] M. Aqil, A. Jbari, A. Bourouhou," *Adaptive ECG Wavelet analysis for R-peaks detection",* IEEE Conference Publications, pp.164-167, May 2016

[21] Physiotools [Online], *https://physionet.org/physiotools/wfdb/*

[22] Physiobank databases [Online]. *http://www.physionet.org/physiobank/database/mitdb.*