Comparison of ECG Baseline Wander Removal Techniques and Improvement Based on Moving Average of Wavelet Approximation Coefficients

Mounaim Aqil[1*], Atman Jbari[2], Abdennasser Bourouhou[2]

[1]*Engineering and Applied Technologies Laboratory*
Superior School of Technology
Sultan Moulay Slimane University in Beni Mellal, Morocco
E-mail: mounaimaqil@hotmail.fr

[2]*Electronic Systems Sensors and Nanobiotechnologies Research Team*
High School of Technical Education
Mohammed V University in Rabat, Morocco
E-mails: atman.jbari@um5.ac.ma, a.bourouhou@um5s.net.ma

Received: January 26, 2020 **Accepted: February 02, 2021**

Published: June 30, 2021

Abstract: *The baseline wander is among the artifacts that corrupt the ECG signal. This noise can affect some signal features, in particular the ST segment, which is an important marker for the diagnosis of ischemia. This paper presents a study on the effectiveness of several methods and techniques for suppressing the baseline wonder (BW) from the ECG signals. As a result, a new technique called moving average of wavelet approximation coefficients (DWT-MAV) is proposed. The techniques concerned are the moving average, the approximation of the baseline by polynomial fitting, the Savitzky-Golay filtering, and the discrete wavelet transform (DWT). The comparison of this techniques is performed using the main criteria for assessing the BW denoising quality criteria such mean square error (MSE), percent root mean square difference (PRD) and correlation coefficient (COR). In this paper, three other criteria of comparison are proposed namely the number of samples of the ECG signal, the baseline frequency variation and the time processing. Two of these new indices are related to possible real time ECG denoising. To improve the quality of BW suppression including the new indices, a new method is proposed. This technique is a combination of the DWT and the moving average methods. This new technique performs the best compromise in terms of MSE, PRD, coefficient correlation and the time processing. The simulations were performed on ECG recording from MIT-BIH database with synthetic and real baselines.*

Keywords: *ECG signal, Baseline wander, Discrete wavelet analysis, Mean square error, Percent root mean square difference, Baseline removal techniques.*

Introduction

The electrocardiogram (ECG) is the recording of the electrical cardiac activity. It's an essential tool for diagnosis of heart diseases. The acquisition of ECG signal often introduces different kind of noise. The main sources of such artefacts are:

- The baseline wander (BW), caused by respiration and body movements, can cause problems to analysis, especially when examining the low-frequency components of ECG signal [2].
- The electromyography (EMG) mainly caused by the muscle activity, is considered as a high-frequency noise [2].
- The Power line interferences (PLI) due to differences in the electrode impedances and the parasitic currents that pass through the body and the electrode. Thus, the false differential potential created cannot be suppressed even by instrumentation amplifier with infinite common mode rejection ratio (CMRR) [13].

The BW removal is considered as a classical problem since this noise can produce atrifactual data when measuring the ECG parameters. Especially, this noise affects the ST segment. In the other hand, the guidelines for diagnosis of myocardial ischemia are based on specific, small changes in the ST segment [9, 10].

Many methods and techniques have been reported in literature to eliminate the BW noise. Among these approaches, we can report the approaches based on linear and nonlinear filter banks [3, 12, 19], moving average [4, 11, 21], adaptive filtering [1], approximation of baseline morphology [6, 7, 14] and wavelet transform analysis [2, 5, 18].

Considering the power spectra of ECG and its artifact given in Fig. 1 [6], the majority of baseline wander removal techniques cancel the low frequency components of the ECG signal.

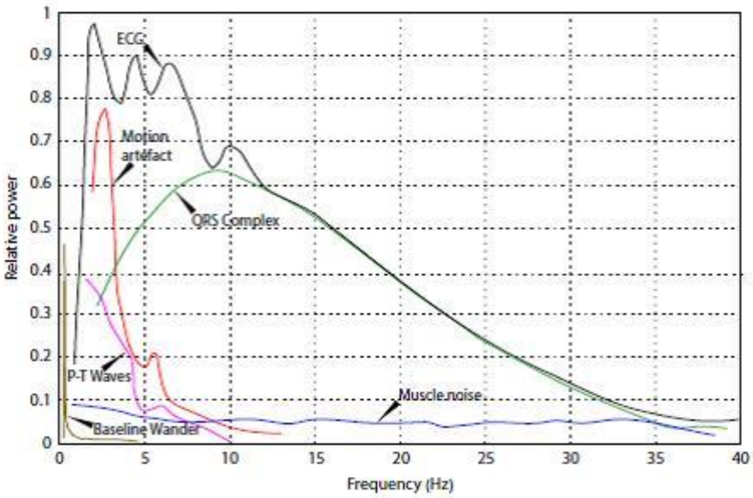


Fig. 1 Power spectra of ECG component and artifacts

The comparison of different techniques suggested is usually based on some criteria of baseline denosing quality such as the mean square error (MSE), root mean square error (RMSE) and percent root mean difference (PRD). These parameters are often evaluated for a fixed duration of the ECG signal, i.e. for a fixed number of samples.

In the other hand, these parameters are computed only for one frequency of signal synthesized according to a sinusoidal model.

The first aim of this work is to establish a comparison between different methods of baseline denoising based on the number of samples. The methods concerned by these studies are the moving average (MAV), approximation of the baseline by polynomial fitting (POL), Savitzky-Golay (SGO) filtering, and discrete wavelet transform (DWT). This study will allow determining the appropriate approach giving the best quality criteria with a minimum number of samples of the signal. This idea is interesting when we look for a real time ECG processing. The second objective is to compare the performance of denoising taking into account different frequencies of the baseline. The third purpose is to improve the performance of baseline cancellation. This last consists to propose a new method based on the combination of DWT analysis and moving average techniques.

This paper is organized as follows: Section 2 presents the workflow of baseline wander; Section 3 describes the background theory of different methods; Section 4 presents the methodology and simulation results; the proposed technique is presented in Section 5, while a conclusion is given in Section 6.

Baseline wander removal model

In this study the compared techniques are applied to an ECG signal $x(n)$ downloaded from MIT-BIH database [15]. The baseline fluctuation signal $b(n)$ is then added to the ECG signal as follow:

$$y(n) = x(n) + b(n) \tag{1}$$

where $x(n)$ is the original ECG signal, and $y(n)$ is the corrupted one.

An example of corrupted ECG signal is illustrated in Fig. 2.

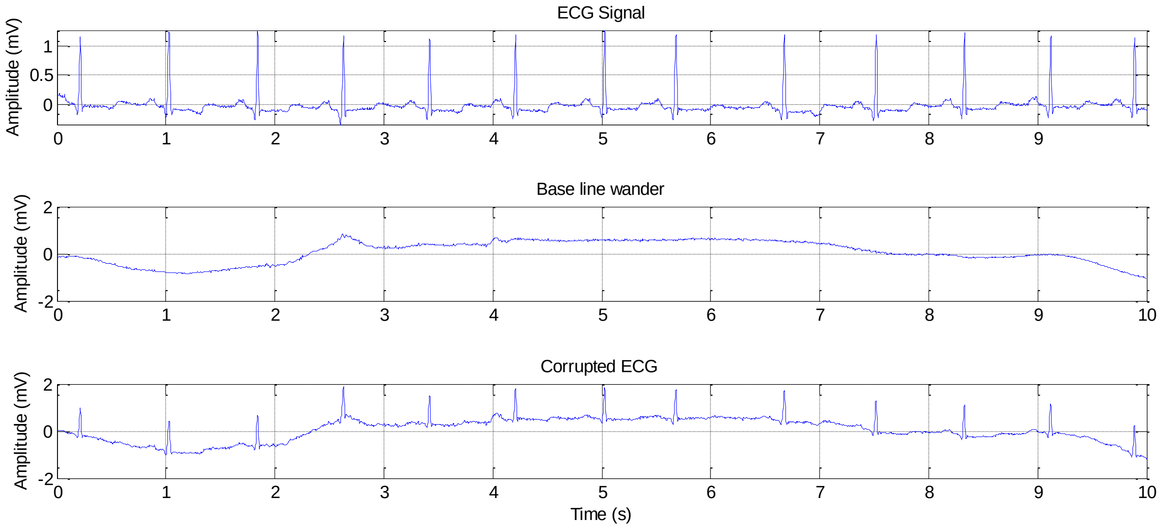


Fig. 2 An example of corrupted signal

The signal $y(n)$ will be filtered using four different baseline wander removal methods. After filtering, the reconstructed signal $\tilde{x}(n)$ will be compared to the original signal $x(n)$. Fig. 3 shows the flow diagram of the process.

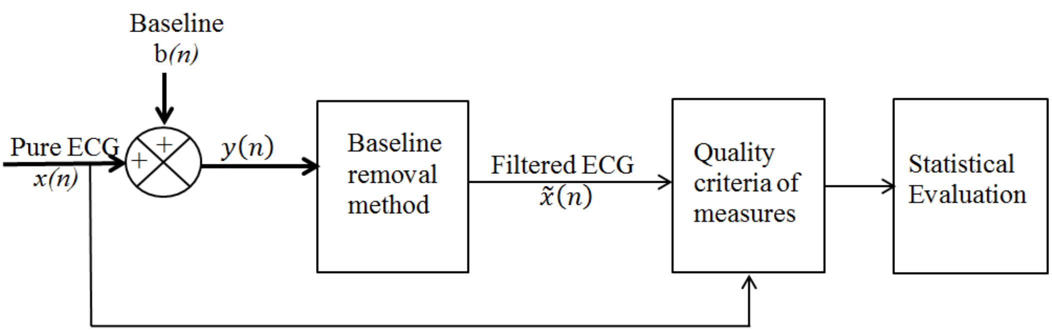


Fig. 3 Signal filtering workflow

In order to evaluate the quality of filtering process, we use some parameters of performance, the most important are:

- The mean square error:

$$MSE = \frac{1}{N}\sum_{n}\left(\tilde{x}(n) - x(n)\right)^2 \tag{2}$$

- The percent root mean square difference:

$$PRD = 100 \times \sqrt{\frac{\sum_n\left(\tilde{x}(n) - x(n)\right)^2}{\sum_n\left(x(n)\right)^2}} \tag{3}$$

- The correlation coefficient:

$$COR = \frac{E\{(x - \mu_x)(\tilde{x} - \mu_{\tilde{x}})\}}{\sigma_{\tilde{x}}\sigma_x} \tag{4}$$

where x is the original ECG signal, $\tilde{x}$ is the filtered signal, $E\{.\}$ is the expected value operator, σ_x is the standard deviation of x, μ_x is the expected value of x.

Theory of baseline wander cancellation techniques

Moving average method

The principle of a moving average filter (MAV) is to estimate the baseline noise by replacing each data sample with the average of the neighbor data samples defined within a window. The estimated baseline can be expressed by the following equation [8, 9]:

$$\tilde{b}(n) = \frac{1}{2N + 1}\sum_{i=-N}^{N} y(n + i) \tag{5}$$

where $\tilde{b}(n)$ is the estimated baseline, $\mathrm{y}(n)$ is the corrupted ECG, and $2N + 1$ is the length of the observation window. The filtered signal can be reconstructed using the Eq. (6):

$$\tilde{x}(n) = y(n) - \tilde{b}(n) \tag{6}$$

Polynomial fitting technique

An alternative to estimate the baseline with a moving average method is to fit a polynomial to representative samples (knots) of ECG, with one knot being defined for each beat. The knots are chosen from the isoelectric line. The last is represented by the PR segment. The polynomial estimating the baseline is fitted by requiring it to pass through each of the knots smoothly [14, 22]. The most popular of this method is the cubic spline baseline estimation, which has its starting point in a Taylor series expansion. The knots of successive beats located at times t_i are denoted $y(t_i)$, $i = 0, 1, 2, ...$. The estimated baseline is computed for the interval $[t_i; t_{i+1}]$ by incorporating the three knots $y(t_i)$, $y(t_{i+1})$, $y(t_{i+2})$ into the k^{th} order Taylor series expanded around t_i:

 doi: 10.7546/ijba.2021.25.2.000770

$$\tilde{b}(t) = \sum_{l=0}^{k} \frac{(t-t_i)^l}{l!}\, \tilde{b}^{(l)}(t_i) \tag{7}$$

The third-order polynomial description is truncated to the Eq. (8):

$$\tilde{b}(t) = \tilde{b}(t_i) + (t-t_i)\tilde{b}^{(1)}(t_i) + \frac{(t-t_i)^2}{2}\tilde{b}^{(2)}(t_i) + \frac{(t-t_i)^3}{6}\tilde{b}^{(3)}(t_i) \tag{8}$$

where $\tilde{b}^{(l)}$ denotes the l^{th} derivative of $\tilde{b}(t)$.

By applying the series expansion for the first derivative $\tilde{b}^{(1)}(t)$, we obtain the following equation [14, 22]:

$$\tilde{b}^{(1)}(t) = \tilde{b}^{(1)}(t_i) + (t-t_i)\tilde{b}^{(2)}(t_i) + \frac{(t-t_i)^2}{2}\tilde{b}^{(3)}(t_i) \tag{9}$$

Under the two following conditions:

- the estimated baseline $\tilde{b}(t)$ must pass through the knots $y(t_i)$, i.e. $\tilde{b}(t_i) = y(t_i)$;
- the first derivative $\tilde{b}^{(1)}(t_i)$ can be approximate by the

$$\tilde{b}^{(1)}(t_i) \approx \frac{x(t_{i+1}) - x(t_i)}{t_{i+1} - t_i} \tag{10}$$

In order to find the remaining two variables $\tilde{b}^{(2)}(t_i)$ and $\tilde{b}^{(3)}(t_i)$ we repeat the same steps bellow for the time point t_{i+1}.

The estimation of baseline can be defined by the following system:

$$\begin{aligned}
\tilde{b}(t) &= \tilde{b}(t_i) + (t-t_i)\tilde{b}^{(1)}(t_i) + \frac{(t-t_i)^2}{2}\tilde{b}^{(2)}(t_i) + \frac{(t-t_i)^3}{6}\tilde{b}^{(3)}(t_i) \\
\tilde{b}^{(1)}(t) &= \tilde{b}^{(1)}(t_i) + (t-t_i)\tilde{b}^{(2)}(t_i) + \frac{(t-t_i)^2}{2}\tilde{b}^{(3)}(t_i) \\
\tilde{b}^{(2)}(t) &= \tilde{b}^{(2)}(t_i) + (t-t_i)\tilde{b}^{(3)}(t_i) \\
\tilde{b}^{(3)}(t) &= \tilde{b}^{(3)}(t_i)
\end{aligned} \tag{11}$$

where [22]:

$$\begin{aligned}
\tilde{b}^{(2)}(t_i) &= \frac{6(\tilde{b}(t_{i+1}) - \tilde{b}(t_i))}{(t_{i+1}-t_i)^2} - \frac{2\left(2\,\tilde{b}^{(1)}(t_i) + \left(\frac{\tilde{b}(t_{i+2}) - \tilde{b}(t_i)}{t_{i+2}-t_i}\right)\right)}{t_{i+1}-t_i} \\
\tilde{b}^{(3)}(t) &= -12\frac{\left(\tilde{b}(t_{i+1}) - \tilde{b}(t_i)\right)}{(t_{i+1}-t_i)^3} + \frac{6\left(\tilde{b}^{(1)}(t_i) + \left(\frac{\tilde{b}(t_{i+2}) - \tilde{b}(t_i)}{t_{i+2}-t_i}\right)\right)}{(t_{i+1}-t_i)^2}
\end{aligned} \tag{12}$$

By choosing t equal to one sample interval, we obtain the discrete form of Eq. (11) as expressed in the following system of equations:

 doi: 10.7546/ijba.2021.25.2.000770

$$\begin{aligned}
\tilde{b}(n+1) &= \tilde{b}(n) + \tilde{b}^{(1)}(n) + \frac{1}{2}\tilde{b}^{(2)}(n) + \frac{1}{6}\tilde{b}^{(3)}(n) \\
\tilde{b}^{(1)}(n+1) &= \tilde{b}^{(1)}(n) + \tilde{b}^{(2)}(n) + \frac{1}{2}\tilde{b}^{(3)}(n) \\
\tilde{b}^{(2)}(n+1) &= \tilde{b}^{(2)}(n) + \tilde{b}^{(3)}(n) \\
\tilde{b}^{(3)}(n+1) &= \tilde{b}^{(3)}(n)
\end{aligned} \tag{13}$$

Once the baseline is estimated, the filtered signal is obtained using the Eq. (6).

The performance of this technique is critically dependent on the accuracy of the knot determination.

Savitzky-Golay filtering

Savitzky and Golay showed that fitting a polynomial to a set of input samples and then evaluating the resulting polynomial at a single point within the approximation interval is equivalent to discrete convolution with a fixed impulse response [17, 20].

In order to filter a data vector $X = [x_{-M}, \dots, x_{-1}, x_0, x_1, \dots . x_M]^{\mathrm{T}}$ having M samples around the sample y_0; the method fits the data samples by a polynomial of degree d as follow:

$$\tilde{x}_m = c_0 + c_1 m + \cdots + c_d m^d, -M \le m \le M \tag{14}$$

After, we define a $(d+1)$-dimensional polynomial basis vectors s_i to have components:

$$s_i(m) = m^i, -M \le m \le M \tag{15}$$

The smoothed values of Eq. (14) can be expressed as:

$$\tilde{X} = \sum_{i=0}^{d} c_i s_i = [s_0, s_1, \dots, s_d] \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_d \end{bmatrix} \tag{16}$$

The steps to realize the SGO filters are as giving [17]:

- we define the symmetric matrix $F = S^{\mathrm{T}}S$;
- we calculate the matrix $G = SF^{-1}$;
- we evaluate the matrix $B = GS^{\mathrm{T}} = [b_{-M}, \dots ., b_0, \dots, b_M]$.

The components of the matrix B are the coefficients of SG filter of length $N = 2M + 1$ and order d. The filtered signal can be write as follows:

$$\tilde{X} = BY \Leftrightarrow \tilde{x}_m = b_m^{\mathrm{T}} Y, -M \le m \le M \tag{17}$$

where Y is the corrupted signal.

 doi: 10.7546/ijba.2021.25.2.000770

Discrete wavelet transform based baseline cancellation

The DWT is a powerful tool for the processing of non-stationary signals. This transform is widely used in ECG de-noising. The aim of the DWT is to decompose a signal into different resolutions using high pass and low pass filters. Regarding the equations of decomposition, consider [2]:

$$A(k) = \sum_n y(n)h(2k - n) \tag{18}$$

$$D(k) = \sum_n y(n)g(2k - n) \tag{19}$$

where $h(n)$ is the half band low pass filter; $g(n)$ is the half band high pass filter; $A(k)$ are the approximation coefficients; $D(k)$ are the detail coefficients; $y(n)$ is the discrete form of the corrupted signal.

Considering that f_s is the sampling frequency of the signal $y(t)$, the detail $D(k)$ contains the information on the high frequency components of the signal within the interval $\left[\frac{f_s}{2^{k+1}}, \frac{f_s}{2^k}\right]$. The approximation coefficient $A(k)$ contains the low frequency components in the interval $\left[0, \frac{f_s}{2^{(n+1)}}\right]$. An example of the DWT decomposition at level 2 can be represented by the block given in Fig. 4 [2].

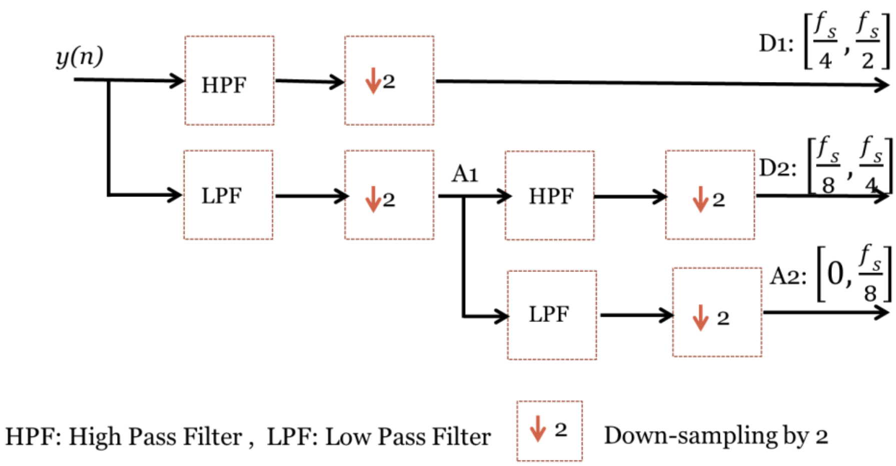


Fig. 4 Second level DWT decomposition filter model

The cancellation of baseline scheme involves two main steps:

- L-level DWT decomposition of input corrupted signal. The decomposition level L for the DWT was chosen such that the approximation coefficients in that level corresponded to the frequency band where the artifact was located. Since in every decomposition level the signal is down-sampled by a factor of two, its Nyquist frequency is also halved. Thus, the necessary decomposition level to achieve a cut-off frequency f_c depending of the sampling frequency f_s can be given by:

$$f_c = \frac{f_s}{2^{L+1}} \Rightarrow L = \log_2\left(\frac{f_s}{f_c}\right) - 1 \tag{20}$$

- Cancellation of the approximation coefficients which correspond to the frequency band of baseline.

Methodology and simulation results

Test data

The ECG signals used in the simulations that will be presented below are downloaded from MIT-BIH database [15]. This database contains 48 half-hour excerpts of two-channel ambulatory ECG recordings, obtained from 47 subjects. The recordings were digitized at 360 samples per second ($f_s = 360$ Hz) per channel with 11-bit resolution over a 10 mV range. Also, we worked with two categories of the baseline artifact:

- Synthetic data, in this case three models of baseline wander are used:
  - ✓ Sinusoidal model (sbw1): $b(t) = A\sin(2\pi f t)$; where the amplitude $A = 0.4\,mV$ and the frequency $f \in [0 - 0.5$ Hz$]$.
  - ✓ Linear model (sbw2): $b(t) = \alpha * t$; where the slope $\alpha = 0.04$.
  - ✓ Square model (sbw3): $b(t) = \beta * t^2$; where the coefficient $\beta = 0.04$.
- Real data: is downloaded from the MIT-BIH Noise Stress Database [16]. These datasets (bwm1.mat and bwm2.mat) are baseline wander recordings, sampled at $f_s = 360$ Hz with 11-bit resolution.

Comparison based on the number of samples

The aim of this study is to compare the different baseline removal techniques considering the size of the ECG signal recording. Indeed, the signal is segmented using a window whose size increases and the performance of each method is then evaluated according to the size of each segment.

Synthetic baseline

The simulations are done with the following conditions for the studied methods:

- Moving average technique: the length of the observation window equals to 100 samples.
- Polynomial fitting method: the polynomial of degree 6 is used to fit the ECG signal.
- Savitzky-Golay filtering: the filter of order 3 and a frame size (length N) equal 49 is exploited.
- DWT approach: the DWT decomposition is done at level 8 with the mother wavelet of symlet 8 which provides best performance in ECG denoising [2].

The ECG recording used in these simulations is the 100m.mat of MIT-BIH database [15].

The simulation result of the different techniques is presented in Fig. 5. The amplitude and frequency of the baseline used in these simulations are: $A = 0.4$ mV, $f = 0.1$ Hz.

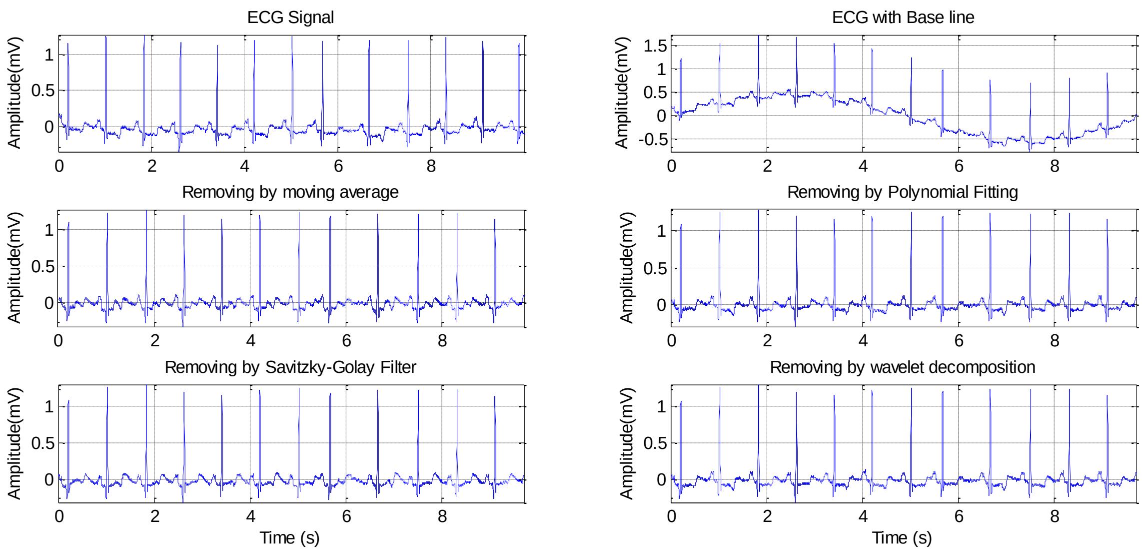


Fig. 5 Simulation result of different methods for synthetic baseline

The comparison between the different methods can be established on the estimated baseline. Indeed, the Fig. 6 shows this comparison.

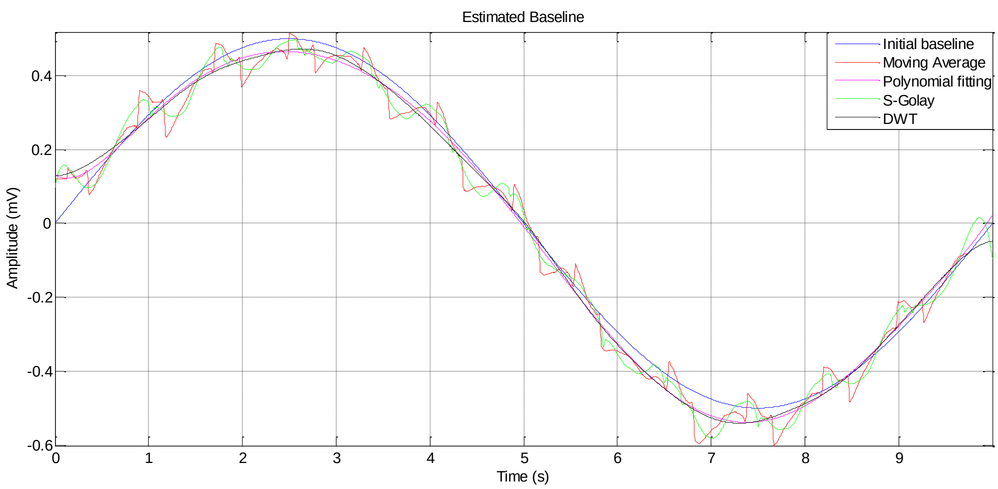


Fig. 6 Estimated baseline for the different techniques

The performance parameters of synthetic baseline suppression are summarized in the Table 1.

Table 1. Quality criteria of synthetic baseline removal

Method	**MSE×10^{-3}**			**PRD, (%)**			**Cor**		
	sbw1	**sbw2**	**sbw3**	**sbw1**	**sbw2**	**sbw3**	**sbw1**	**sbw2**	**sbw3**
MAV	2.1	2.1	2.2	3.97	3.88	4.23	0.96	0.96	0.96
POL	0.8	0.8	0.8	3.21	3.21	3.21	0.98	0.98	0.98
SGO	2	2	2	3.82	3.82	3.82	0.96	0.96	0.96
DWT	1	1	1	3.56	3.59	3.62	0.98	0.98	0.98

In order to compare the different methods, the Fig. 7 gives the quality criteria comparison for sinusoidal baseline (sbw1) removal.

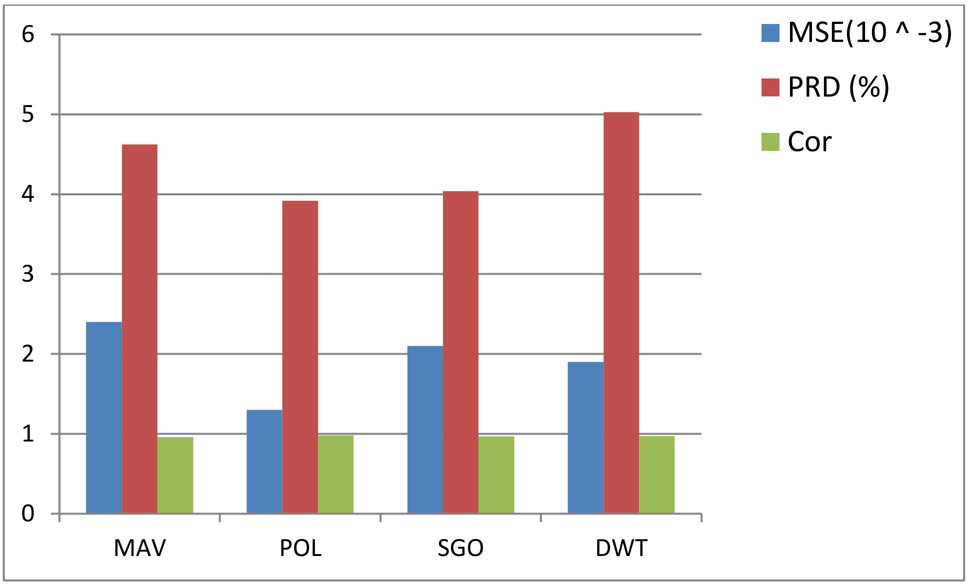


Fig. 7 Comparison of quality criteria for synthetic baseline removal

In this step of our study, we can confirm that all the different techniques achieve good performance parameters for baseline cancellation. Indeed, the MSE is less than 2.5×10^{-3} for all the techniques, the correlation coefficient is greater than 0.96 and the PRD parameter is less than 5%. We can note that the polynomial fitting and DWT based techniques achieve coefficient parameters better than the others.

The quality comparison of according to the segment size is illustrated in Fig. 8. This comparison shows that 1000 samples of ECG signal are sufficient to remove the synthetic baseline with best quality. This result is interesting for a real time ECG processing. This study shows also that the DWT approach gives a quality of baseline suppression better than the others.

Real baseline

In this case we used the baseline bwm2.mat, the simulations are done with the conditions of synthetic data. The simulation result of the process of baseline cancellation is shown in Fig. 9. The estimated baseline is given in Fig. 10.

The performance parameters of synthetic baseline suppression are summarized in the Table 2.

Table 2. Quality criteria of real baseline removal

Method	**MSE×10^{-3}**		**PRD, (%)**		**Cor**	
	bwm1	**bwm2**	**bwm1**	**bwm2**	**bwm1**	**bwm2**
MAV	3	2.4	6.5566	4.6235	0.9517	0.9615
POL	10.6	1.3	17.5844	3.9195	0.8564	0.9803
SGO	3.1	2.1	7.1670	4.0383	0.9505	0.9660
DWT	7	1.9	13.7376	5.0287	0.8993	0.9707

In order to better compare the different methods, the Fig. 11 presents the quality comparison for different approaches. These results concern the bwm2 baseline.

We can notice that the polynomial fitting method presents the best performances.

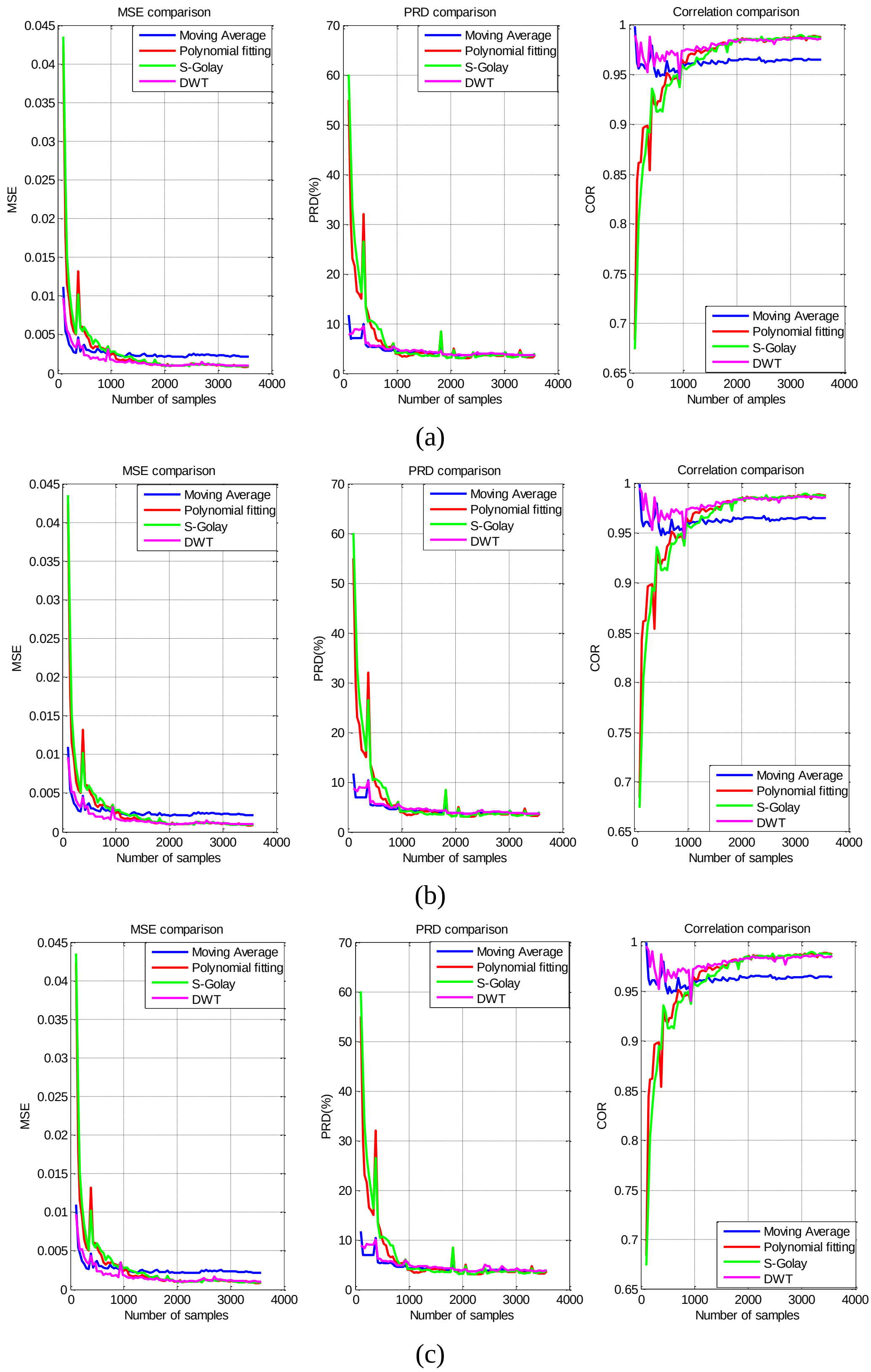


Fig. 8 Quality criteria comparison according to segment size for synthetic baseline: (a) sbw1, (b) sbw2, (c) sbw3.

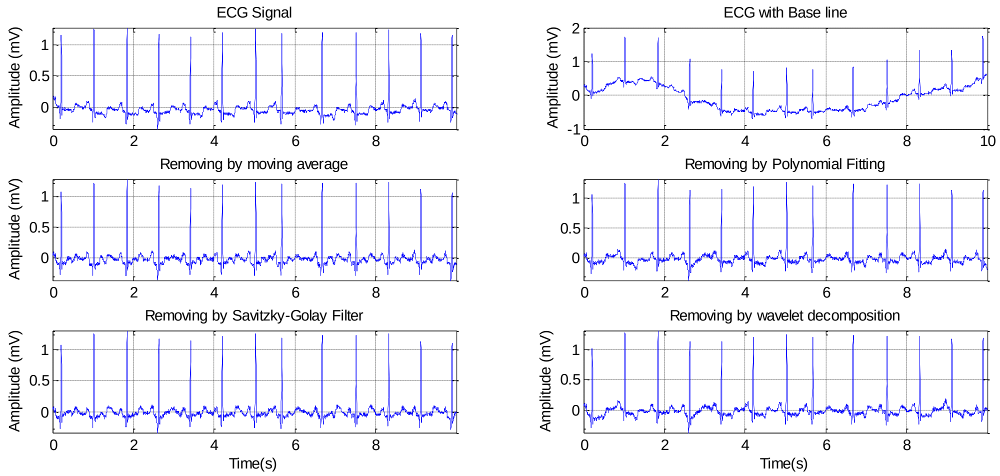


Fig. 9 Simulation result of different methods for real baseline

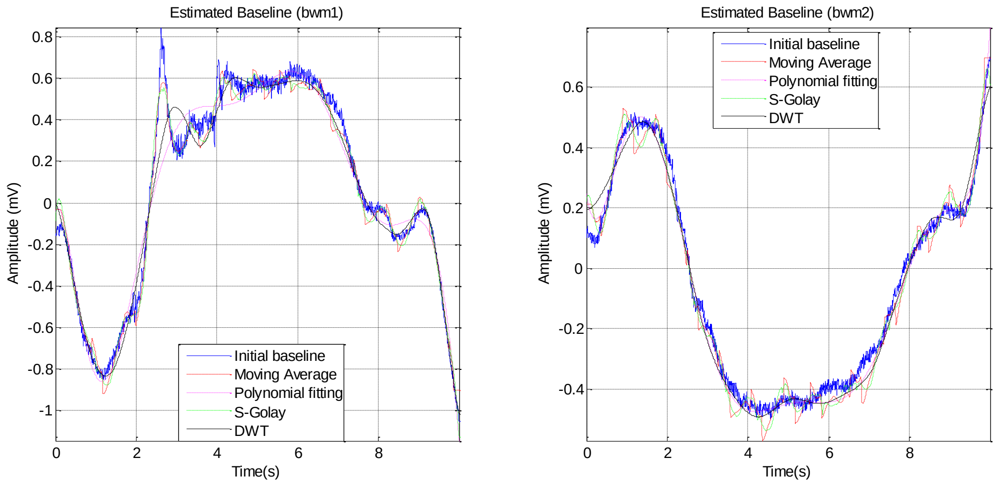


Fig. 10 Estimated Baseline for the different techniques

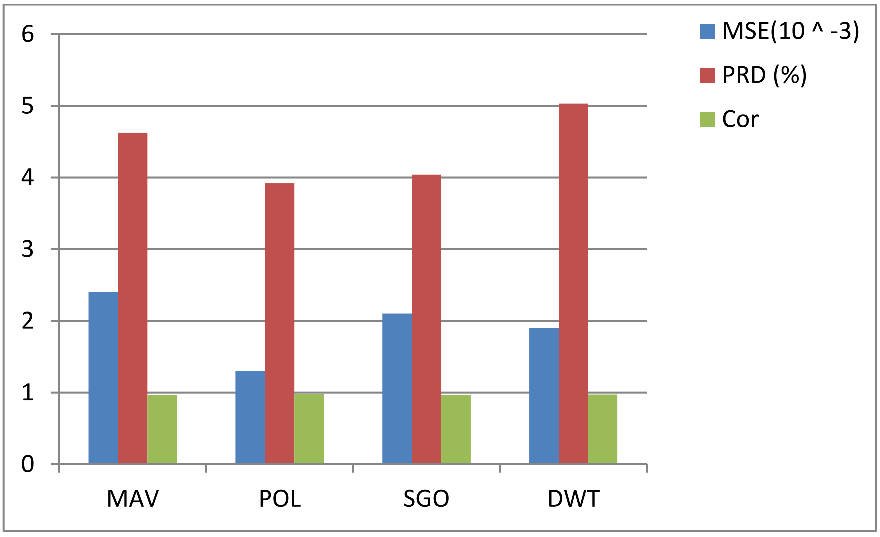


Fig. 11 Comparison of correlation coefficient for real baseline

In this case, we can confirm that the quality criteria are lower than the previous ones. This result is normal, since in the first case the signal presents only one frequency component; whereas the real signal can have several frequency components.

The study of quality criteria according to the segment size is illustrated in Fig. 12. Following these results, we can conclude that 1000 samples of corrupted ECG signal are sufficient to remove the real baseline.

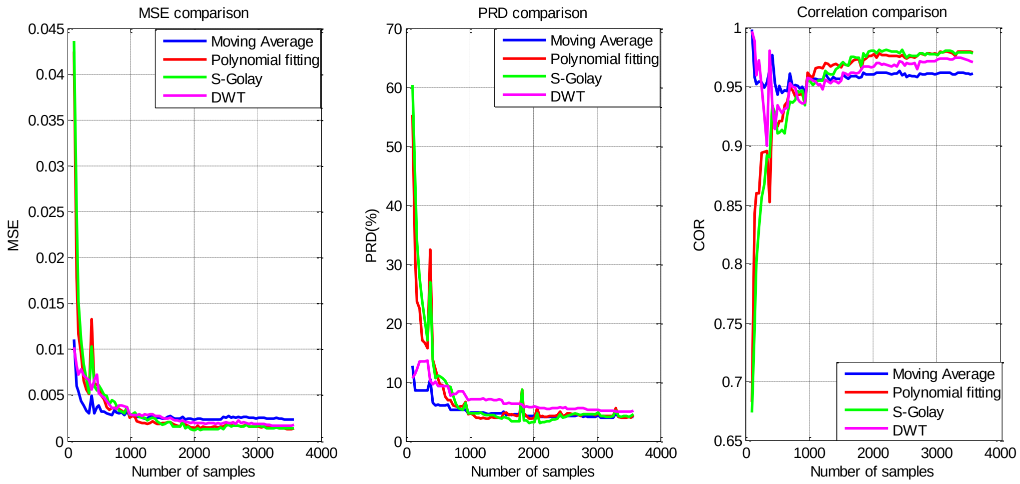


Fig. 12 Quality criteria comparison according to segment size for real baseline bwm2

Comparison based on baseline frequency

The goal of this study is to compare the performance of the different baseline removal techniques according to variation of baseline frequency. The idea is to determine which method allows keeping the best performances even if the baseline frequency changes within the range $[0 - 0.5 \text{ Hz}]$.

In this study the comparison is established using sinusoidal baseline (sbw1) with a frequency variation of 0.001 Hz. The comparison is shown in Fig. 13.

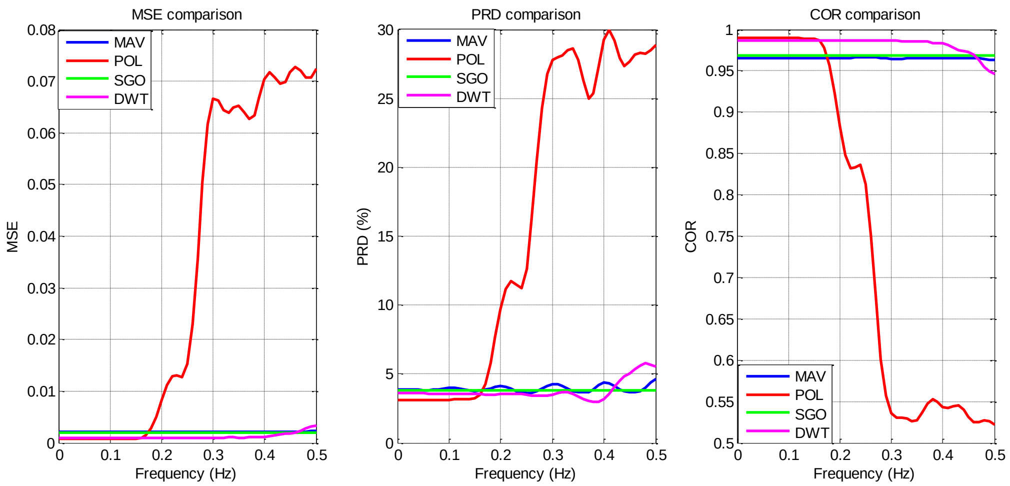


Fig. 13 Quality criteria comparison according to baseline frequency

According to these results, except the polynomial fitting method, all the other techniques have constant performances in the frequency range of the baseline.

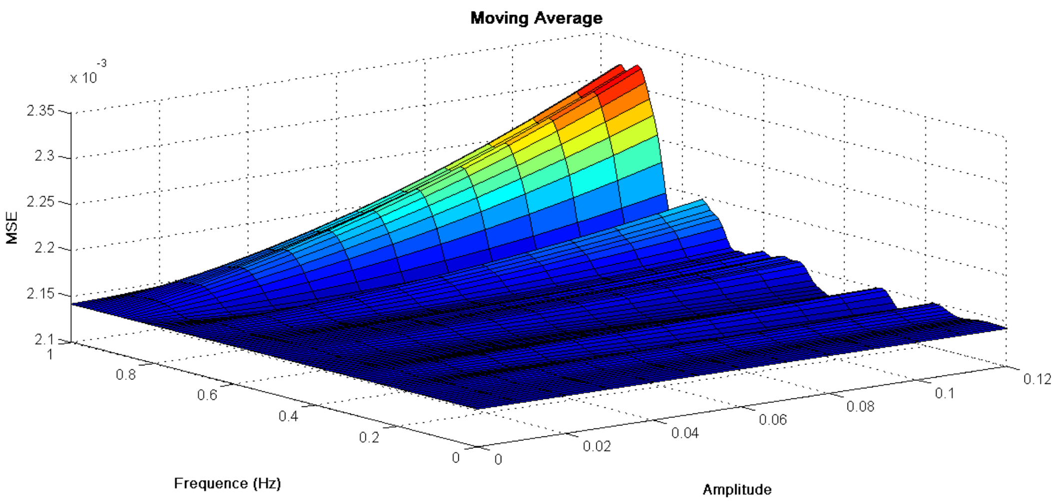
Another comparison of the studied methods can be made. This last is based on the variation of both the amplitude and the frequency of the sinusoidal baseline. Indeed, the simulation result of MSE comparison is giving in Fig. 14.

These simulations are performed with the following values:

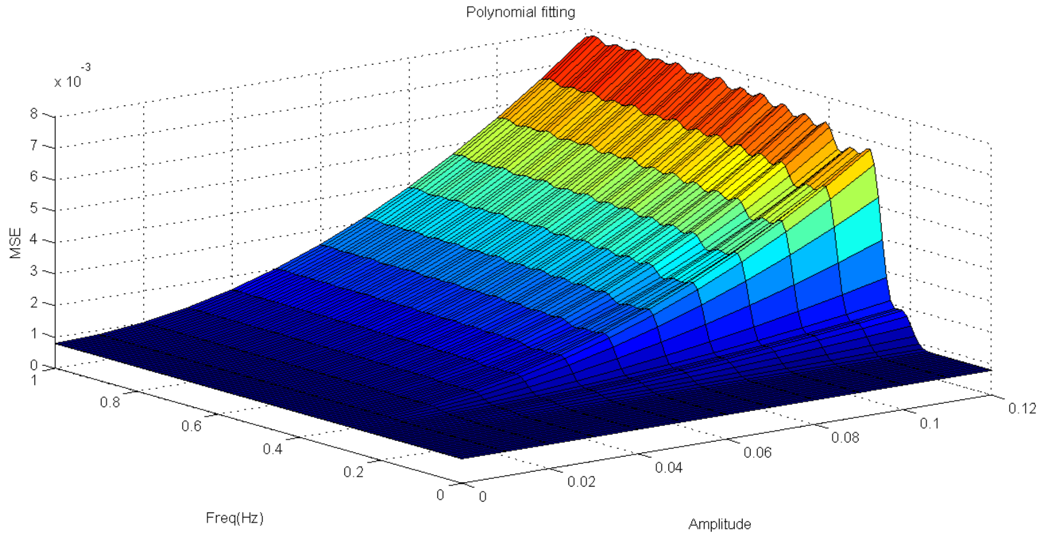
- frequency range of $[0-1\text{ Hz}]$ and a variation of 0.01 Hz;
- amplitude interval of $\left[0-\frac{\max(x)}{10}\right]$ and a variation of 0.01 mV, where x is the original ECG signal.

The simulations give the following deductions:

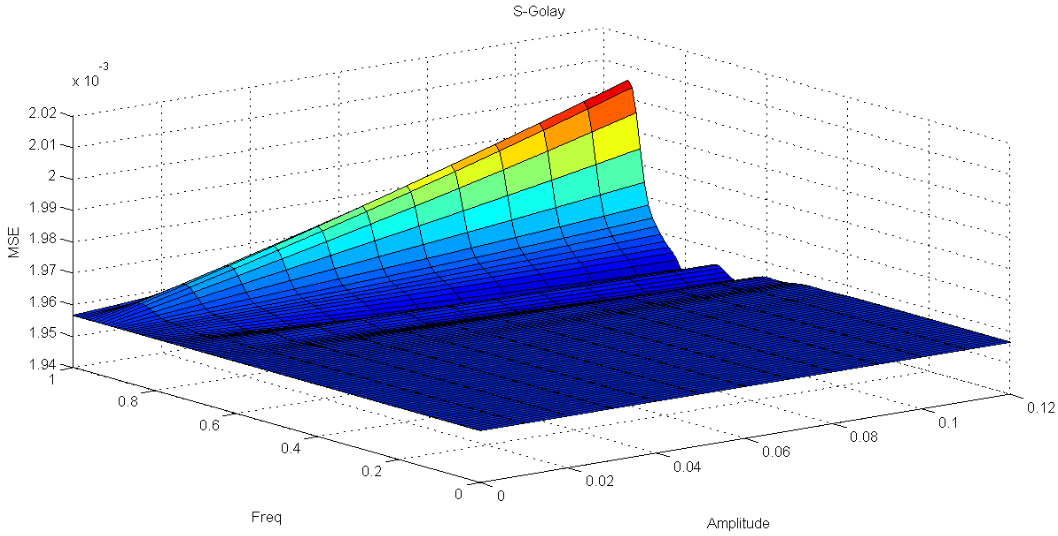
- The MAV technique performs a low MSE for all the amplitude interval when baseline frequency is less than 0.4 Hz;
- The POL method gives a best MSE when amplitude is less than 0.012 mV for all the frequencies in the range $[0-1\text{ Hz}]$;
- The SGO filtering returns a good MSE for all the amplitude interval when baseline frequency is less than 0.8 Hz;
- The DWT method retains a low MSE for a wide range of amplitude and frequency variation.



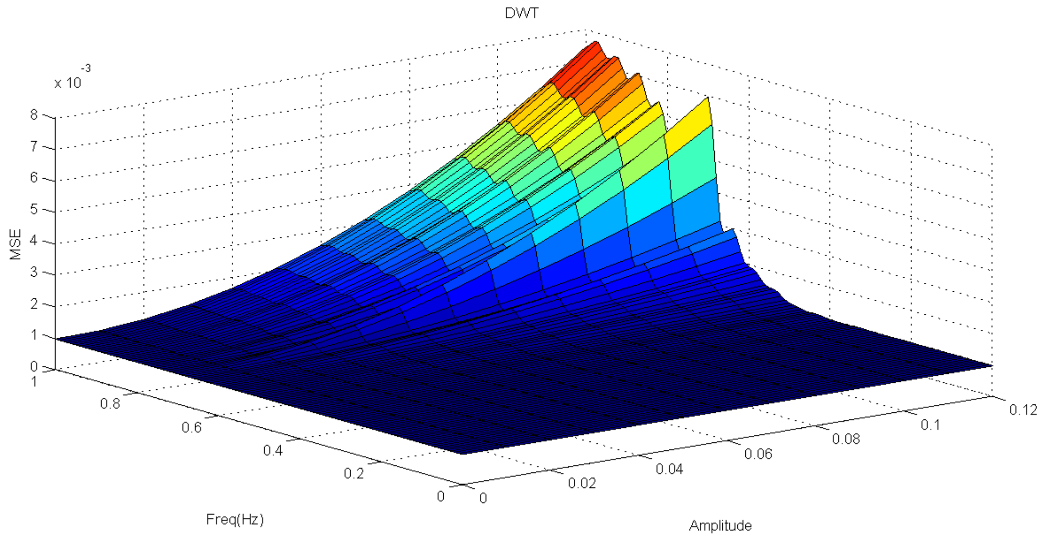
(a)



(b)



(c)



(d)

Fig. 14 MSE comparison according to amplitude and frequency variation:
(a) MAV, (b) POL, (c) SGO, (d) DWT.

Comparison based on time processing

Taking into account the results of the study carried out so far, we have shown that the DWT technique presents the best performances with respect to the variations of the amplitude and the frequency of the baseline. However, this method has a major drawback which is the processing time. Indeed, the comparison between the different methods based on processing time is giving in Table 3. In this table, the sbw1 baseline is the synthetic one modeled by sinusoidal signal; bwm1 and bwm2 are the real baselines. These values are computed on i5-2520M CPU with 8 GB of RAM memory.

Table 3. Time processing comparison

Method	**Processing time, (s)**		
	sbw1	**bwm1**	**bwm2**
MAV	0.0625	0.0781	0.0625
POL	0.2813	0.1406	0.1094
SGO	0.25	0.1406	0.1406
DWT	0.4219	0.3281	0.3125

These results show that the DWT method is the slowest while the moving average method is the fastest. From this last deduction came the idea of combining these two methods.

Moving average of discrete wavelet coefficients method

Principle

The proposed technique, which we call the DWT-MAV, is based on following steps:

- First level DWT decomposition of the corrupted signal $y(n)$: allows computing the coefficients of detail $D(1,k)$ and approximation $A(1,k)$. These last contain the low frequency components of the signal; while the coefficients detail – the high frequency ones.
- A moving average technique is applied to the approximation coefficients $A(1,k)$. This step allows canceling the baseline components from the coefficients $A(1,k)$. It can be expressed by the:

$$\tilde{A}(1,k) = A(1,k) - \tilde{b}(k) \tag{21}$$

and

$$\tilde{b}(k) = \frac{1}{2N+1}\sum_{i=-N}^{N} A(k+i) \tag{22}$$

where $\tilde{A}(1,k)$ are the filtered approximation coefficients, $\tilde{b}(k)$ is the estimated baseline, $2N+1$ is the length of the observation window.

- The filtered approximation coefficients $\tilde{A}(1,k)$ are used to reconstruct the filtered signal $\tilde{x}(n)$ by inverse DWT (IDWT).

These steps are summarized in the flow diagram given in Fig. 15.

Fig. 15 Flow diagram of the proposed method

Simulation results

We performed the simulations of the DWT-MAV method under the following conditions:

- the ECG recoding used is 100m of MIT-BIH database;
- we used the real baseline bwm2.mat;
- the DWT decomposition is done at level 1 with the mother wavelet of symlet 8;
- the moving average is done with a length of the observation window equals to 100 samples.

The simulation result is giving in Fig. 16.

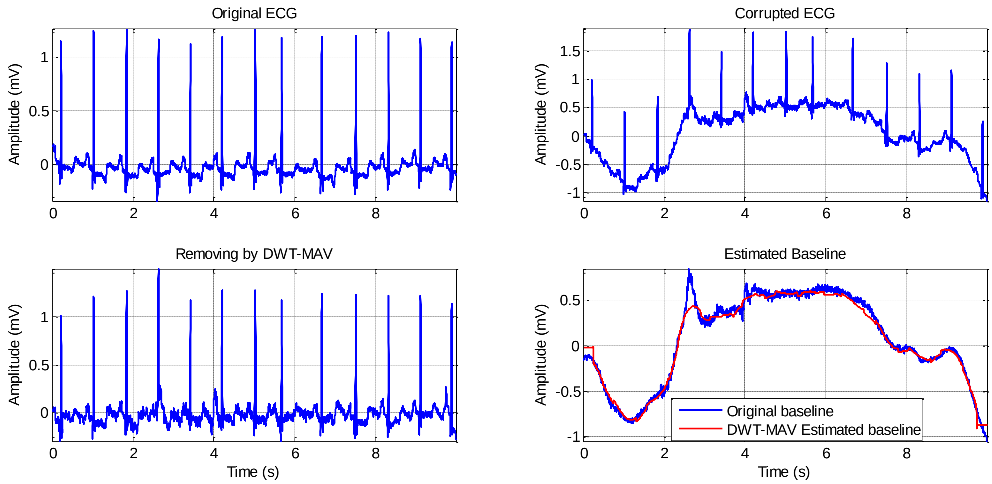


Fig. 16 Simulation of the DWT-MAV method

The performance parameters of the DWT-MAV technique are summarized in Table 4. These parameters are evaluated for the three types of baseline: sinusoidal baseline (sbw1), real baselines (bwm1 and bwm2).

Table 4. Performance parameters of the proposed technique

Parameter	**sbw1**	**bwm1**	**bwm2**
MSE	$0.9*10^{-3}$	$2.9*10^{-3}$	$1.6*10^{-3}$
PRD	3.2199%	8.7%	3.7790%
Cor	0.9870	0.9543	0.9752
Processing time	0.125 s	0.125 s	0.0938 s

The comparison of this new method with the others is presented in Table 5 and Fig. 17. The values presented in this table are available for the bwm1 baseline suppression.

The comparison according to the time processing is shown in Fig. 18. It can be concluded that DWT-MAV method achieves the best compromise between a good quality criteria and a short processing time.

Table 5. Performance parameters comparison

Method	MSE×10^{-3}	PRD, (%)	Cor	Time, (s)
MAV	3	6.5566	0.9517	0.0781
POL	10.6	17.5844	0.8564	0.1406
SGO	3.1	7.1670	0.9505	0.1406
DWT	7	13.7376	0.8993	0.3281
DWT-MAV	2.9	8.7	0.9543	0.125

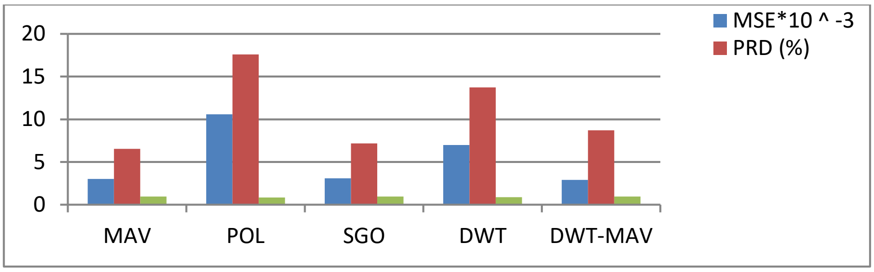


Fig. 17 Performance parameters comparison

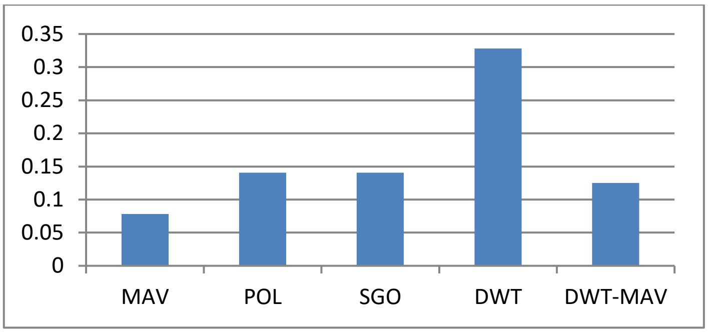


Fig. 18 Time processing comparison

In addition, this method allows a better smoothing of the estimated baseline compared to the MAV technique. In fact, the MAV method has a significant disadvantage. This is the appearance of ripples in the estimated baseline. These ripples are due to the intervention of the values of the peaks R in the evaluation of the average for some windows. This result is presented in Fig. 19.

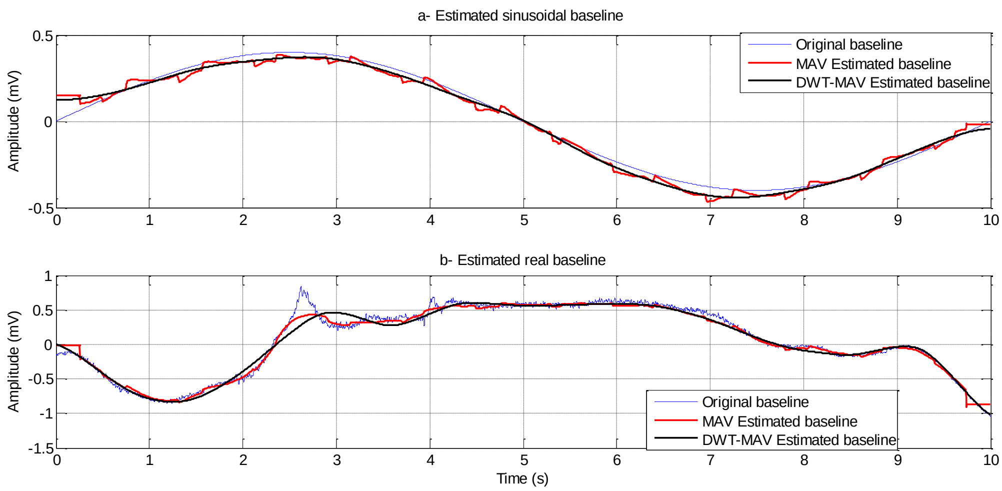


Fig. 19 Estimated baseline comparison

In order to show the efficiency of the proposed method, it is interesting to apply it on different records and compare it with the other techniques. In fact, Table 6 gives a comparison of the performances for 10 different records. These records are chosen with a zero baseline. The results are obtained by adding the same real baseline bwm2.

Table 6. Performance parameters comparison for different records

Record Id		100.m	101.m	102.m	103.m	104.m
MAV	MSE	0.0024	0.0053	0.0051	0.0043	0.0152
	PRD (%)	4.6235	7.0593	15.8778	2.9602	17.9853
	COR	0.9615	0.9414	0.9385	0.9774	0.8965
	Time (s)	0.0781	0.0625	0.0625	0.0625	0.0781
POL	MSE	0.0013	0.0030	0.0031	0.0017	0.0026
	PRD (%)	3.9195	5.3848	9.5568	2.7403	5.5541
	COR	0.9803	0.9718	0.9727	0.9911	0.9837
	Time (s)	0.1094	0.1250	0.0938	0.1406	0.1250
SGO	MSE	0.0021	0.0048	0.0043	0.0040	0.0140
	PRD (%)	4.0383	6.5177	11.6678	3.3386	15.3499
	COR	0.9660	0.9478	0.9546	0.9792	0.9067
	Time (s)	0.0938	0.1406	0.1250	0.1250	0.0938
DWT	MSE	0.0019	0.0043	0.0033	0.0019	0.0033
	PRD (%)	5.0287	6.8230	9.6800	2.7717	7.0709
	COR	0.9707	0.9545	0.9703	0.9904	0.9797
	Time (s)	0.3438	0.3594	0.3438	0.3750	0.3125
DWT-MAV	MSE	0.0018	0.0042	0.0034	0.0026	0.0057
	PRD (%)	4.7839	6.4966	9.2879	2.7522	9.8000
	COR	0.9712	0.9543	0.9683	0.9861	0.9633
	Time (s)	0.0938	0.0938	0.0938	0.1250	0.1250

Record Id		105.m	106.m	111.m	114.m	115.m
MAV	MSE	0.0096	0.0212	0.0046	0.0095	0.0103
	PRD (%)	10.3047	10.2735	11.7002	11.1122	6.6873
	COR	0.9464	0.9269	0.9420	0.9412	0.9401
	Time (s)	0.0781	0.0781	0.0781	0.0625	0.0625
POL	MSE	0.0025	0.0064	0.0026	0.0026	0.0053
	PRD (%)	4.8747	5.7596	8.0702	5.4380	4.4633
	COR	0.9873	0.9784	0.9680	0.9853	0.9697
	Time (s)	0.1250	0.1250	0.1406	0.1094	0.1094
SGO	MSE	0.0075	0.0226	0.0043	0.0056	0.0103
	PRD (%)	10.6511	10.7079	10.7274	6.6065	6.7875
	COR	0.9594	0.9214	0.9465	0.9667	0.9401
	Time (s)	0.1563	0.1250	0.1563	0.1406	0.1094
DWT	MSE	0.0023	0.0085	0.0037	0.0017	0.0053
	PRD (%)	4.1376	8.2490	11.3083	2.4566	4.5483
	COR	0.9884	0.9714	0.9535	0.9916	0.9700
	Time (s)	0.3125	0.3125	0.3906	0.3438	0.3125
DWT-MAV	MSE	0.0041	0.0128	0.0030	0.0039	0.0071
	PRD (%)	4.3438	8.0308	9.7870	3.2496	4.9437
	COR	0.9783	0.9560	0.9636	0.9776	0.9592
	Time (s)	0.1250	0.1406	0.0938	0.0938	0.0938

According these results, it can be noted that the performance of baseline suppression changes with the type of ECG recording and also with the type of the baseline drift.

The proposed method shows good performance in terms of execution time which makes it suitable for real-time processing of the ECG signal.

Conclusion

The purpose of this work is to study the influence of the number of signal samples on the performance of a baseline suppression process as well as its frequency. It has been shown that 1000 samples of corrupted ECG signal can provide a good baseline cancellation performance for the different methods studied. This result is interesting for real time ECG processing.

Moreover, except the polynomial technique, the performances remain unchanged for the whole frequency band [0-0.5 Hz].

Unfortunately, in the case of a real baseline, the performances of baseline suppression are lower than the ones obtained for synthetic baseline. This result is predictable since the synthetic signal is a modeling of the real signal.

It has also been noted that the DWT method performs better than other techniques, which augurs well for upstream processing by extracting ECG signal features.

In order to improve the method based on the DWT, we proposed the new DWT-MAV approach. This last is obtained by combining the DWT and the moving average methods. We have shown that this technique can be a new alternative to traditional methods of base line cancellation. Indeed, this approach performs better than the DWT with a shorter processing time. Also, the DWT-MAV method reduces the ripples which appear in the estimated baseline by moving average technique.

As a future scope, we will consider measuring the impact of this method on the morphology of the ECG signal, especially the protection of the ST segment against unacceptable fluctuations.

References

1. Almenar V., A. Albiol (1999). A New Adaptive Scheme for ECG Enhancement, Signal Processing, 75(9), 253-263.
2. Aqil M., A. Jbari, A. Bourouhou (2017). ECG Signal Denoising by Discrete Wavelet Transform, Int J Online Eng, 13(9), 51-68.
3. Chavan M. S., R. A. Agarwala, M. D. Uplane (2008). Comparative Study of Chebyshev I and Chebyshev II Filter Used For Noise Reduction in ECG Signal, Int J Circuits Syst Signal Process, 2(1), 1-17.
4. Dai M., S. L. Lian (2009). Removal of Baseline Wander from Dynamic Electrocardiogram Signals, Proceedings of the 2nd Int Congr Image Signal Process – CISP’09, 1-4.
5. Ercelebi E. (2004). Electrocardiogram Signals De-noising Using Lifting-based Discrete Wavelet Transform, Comput Biol Med, 34(6), 479-493.
6. Guven O., A. Eftekhar, R. Hoshyar, G. Frattini, W. Kindt, T. G. Constandinou (2014). Realtime ECG Baseline Removal: An Isoelectric Point Estimation Approach, Proceedings of the IEEE 2014 Biomed Circuits Syst Conf – BioCAS’2014, 29-32.

7. Guven O., W. Kindt, A. Eftekhar, T. G. Constandinou (2016). Computationally Efficient Real-time Interpolation Algorithm for Non-uniform Sampled Biosignals, Health Technol Lett, 3(2), 105-110.
8. Hu X., Z. Xiao, N. Zhang (2011). Removal of Baseline Wander from ECG Signal Based on a Statistical Weighted Moving Average Filter, J Zhejiang Univ-Sci C (Comput Electron), 12(5), 397-403.
9. Kaur M., B. Singh, Seema (2011). Comparison of Different Approaches for Removal of Baseline Wander from ECG Signal, Proceedings of the Int Conf Work Emerg Trends Technol – ICWET'11, 1290-1294.
10. Lenis G., N. Pilia, A. Loewe, W. H. W. Schulze, O. Dössel (2017). Comparison of Baseline Wander Removal Techniques Considering the Preservation of ST Changes in the Ischemic ECG: A Simulation Study, Comput Math Methods Med, Vol. 2017, Article ID 9295029.
11. Łęski J. M. (2002). Robust Weighted Averaging, IEEE Trans Biomed Eng, 49(8), 796-804.
12. Łęski J. M., N. Henzel (2005). ECG Baseline Wander and Powerline Interference Reduction Using Nonlinear Filter Bank, Signal Processing, 85(4), 781-793.
13. Levkov C., G. Mihov, R. Ivanov, I. Daskalov, I. Christov, I. Dotsinsky (2005). Removal of Power-line Interference from the ECG: A Review of the Subtraction Procedure, BioMedical Engineering OnLine, 4(50), 1-18.
14. Meyer C. R., H. N. Keiser (1977). Electrocardiogram Baseline Noise Estimation and Removal Using Cubic Splines and State-space Computation Techniques, Comput Biomed Res, 10(5), 459-470.
15. MIT-BIH Arrhythmia Database, Available at www.physionet.org/physiobank/database/mitdb/ (Access data 18 June 2021).
16. MIT-BIH Noise Stress Database, Available at https://physionet.org/physiobank/database/nstdb/ (Access data 18 June 2021).
17. Orfanidis S. J. (2010). Introduction to Signal Processing, Pearson Education, Inc.
18. Poornachandra S., N. Kumaravel (2005). Hyper-trim Shrinkage for Denoising of ECG Signal, Digit Signal Process, 15, 317-327.
19. Rastogi N., R. Mehra (2013). Analysis of Savitzky-Golay Filter for Baseline Wander Cancellation in ECG Using Wavelets, Int J Eng Sci Emerg Technol, 6(1), 2231-6604.
20. Schafer R. W. (2011). What Is a Savitzky-Golay Filter? [Lecture Notes], In: IEEE Signal Processing Magazine, 28(4), 111-117.
21. Schulze H. G., R. B. Foist, K. Okuda, A. Ivanov, R. F. B. Turnera (2012). A Small-window Moving Average-based Fully Automated Baseline Estimation Method for Raman Spectra, Appl Spectrosc, 66(7), 757-764.
22. Sörnmo L., P. Laguna (2005). Bioelectrical Signal Processing in Cardiac and Neurological Applications, 1st Ed., Academic Press.

Prof. Mounaim Aqil, Ph.D.
E-mail: mounaimaqil@hotmail.fr

Mounaim Aqil is currently a Professor at the Mechatronics Department of High School of Technology, Sultan Moulay Slimane University in Beni Mellal, Morocco. In 2018, he received his Ph.D. in Electrical Engineering from Mohammed V University. Prof. Aqil is a Member of Engineering and Applied Technologies Laboratory and an Associate Member of Electronic Systems, Sensors and Nano-biotechnology Research Team. His current research interests include signal processing, embedded electronic systems and IoT for health care.

Prof. Atman Jbari, Ph.D.
E-mail: atman.jbari@um5.ac.ma

Atman Jbari is currently a Professor at the Electrical Engineering Department of Ecole Normale Supérieure de l'Enseignement Technique (ENSET), Mohammed V University in Rabat, Morocco. In 2009, he received his Ph.D. in Computer and Telecommunications from Mohammed V-Agdal University. Prof. Jbari is a Member of Electronic Systems, Sensors and Nanobiotechnology Research Group. His current research interests include signal processing, blind source separation and embedded electronic systems.

Prof. Abdennaser Bourouhou, Ph.D.
E-mail: a.bourouhou@um5s.net.ma

Abdennaser Bourouhou is a Professor at the Higher School of Technical Education (ENSET) of Mohammed V University of Rabat, Rabat, Morocco. He obtained his Ph.D. in Physics from Ibn Tofail University of Kenitra, Morocco, in 2008. Prof. Bourouhou is a Member of the E2SN team of the STIS research center of UM5 Rabat. He has published in the fields of signal processing and sensor networks. His current interests are in wireless sensor networks for environmental protection and medical monitoring.