Performance Study of Different Denoising Methods for ECG Signals

Article *in* Procedia Computer Science · December 2014

DOI: 10.1016/j.procs.2014.08.048

2 authors:

Mohammed Almahamdy
University of Anbar

H. Bryan Riley
Clemson University

Available online at www.sciencedirect.com
ScienceDirect
Procedia Computer Science 37 (2014) 325 – 332
Procedia Computer Science

The 4th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH-2014)

Performance Study of Different Denoising Methods for ECG Signals

Mohammed AlMahamdy[a], H. Bryan Riley[*b] PhD

[a]*Ohio University, Stocker Center EE P.Box #71, Athens 45701, USA*
[b]*Ohio University, Stocker Center 322A, Athens 45701, USA*

Abstract

ECG is an important tool to measure health and disease detection. Due to many noise sources, this signal has to be denoised and presented in a clear waveform. Noise sources may consist of power line interference, external electromagnetic fields, random body movements or respiration. In this project, five common and important denoising methods are presented and applied on real ECG signals contaminated with different levels of noise. These algorithms are: discrete wavelet transform (*universal* and *local* thresholding), adaptive filters (LMS and RLS), and Savitzky-Golay filtering. Their denoising performances are implemented, compared and analyzed in a Matlab environment.

Peer-review under responsibility of the Program Chairs of EUSPN-2014 and ICTH 2014.

Keywords: ECG Signals; Denoising; LMS; RLS; Savitzky-Golay; Wavelet Filtering; Discrete Wavelet Transform;

* Corresponding author. Tel.: +1-740-597-3171; fax: +1-740-593-0007.
E-mail address: rileyh1@ohio.edu

Peer-review under responsibility of the Program Chairs of EUSPN-2014 and ICTH 2014.
doi:10.1016/j.procs.2014.08.048

1. Introduction

Electrocardiography (ECG) is the recording of the electrical activities of the heart. It is very important for screening and diagnosis of many diseases. For its importance, an ECG signal should be presented as clean and clear as possible to support accurate decisions by physicians and doctors. As an electrical signal, ECG is susceptible to different kinds of noise. The main sources of this noise are electrical activities of other body muscles, baseline shift because of respiration, poor contact of electrodes, and equipment or electronic devices [1-2]. Since the ECG is a non-stationary signal [2], normal filters cannot be effective to remove the noise; so, several techniques are used to do so for such types of signals.

There are different techniques available for use in denoising signals, such as adaptive filtering, wavelet, and Savitzky-Golay filtering. These methods are referenced in the literature of ECG denoising.

Many publications present wavelets in denoising ECG signals, such as [3-6]. In Gokhale [7], the wavelet is used to remove the 50/60 Hz Power line interference from ECG signals. Savitzky-Golay is used in smoothing ECG Signals in [8] and [9]. Adaptive filtering is used with ECG signals for noise cancelation in [10] and [11].

In this paper, these methods are presented and applied for ECG denoising with a comparison between their performances. The algorithms are: discrete wavelet transform (universal and local thresholds), adaptive LMS filtering, adaptive RLS filtering, and Savitzky-Golay filtering. The performance measure is based on the difference between the original ECG (without noise) and the denoised version after contamination with noise. Actual recorded ECG signals are available online, such as [12-14]. In this project, these signals are used after adding different levels of noise and then compared to the original and the denoised versions for the prescribed methods.

1.1. Wavelet

Recently, signal processing engineers have used the Wavelet Transform (WT) in various applications including the processing of non-stationary signals. Unlike Fourier transform, WT provides more information in its domain, which includes both frequency and time representation. Discrete Wavelet Transform (DWT) is more popular in application because of the usage of computers. In [15], Mallat's developed a very efficient and reliable algorithm to compute the DWT decompositions using consecutive filters and decimators, as shown in Fig. 1. The HPFs and LPFs determine the corresponding *Details* $d_k[n]$ and *Approximations* $a_k[n]$ coefficients respectively, of the applied signal. The use of the down-samplers is important to remove redundant samples from the output of the filters and to keep the total number of samples the same. According to the characteristics of the applied signal and the required degree of processing, we can reapply the approximation to the most recent decomposition segment several times.

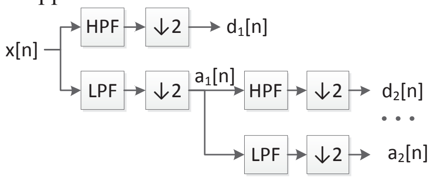


Fig 1. Two levels of Mallat's algorithm for decomposition.

The approximation coefficients a_k are related to the low frequency part of the signal, which contains the main features and information. On the other hand, detail coefficients d_k are important to preserve the perfect shape when reconstruction is invoked. It is noticeable that d_k coefficients at higher levels of noiseless signals are sparse. So, in these high levels, larger coefficients can be assumed to be actual ones plus noise; the remaining coefficients can be considered as pure noise. Fig. 2 illustrates the idea of denoising signals using wavelet thresholding of detail coefficients. *'SureShrink'* and *'NeighBlock'* are the two famous thresholding methods in the use of wavelets for estimation of unknown signals in the presence of noise. These methods are developed by Donoho et.al. [16] and Cai et.al. [17], respectively.

Fig. 2. Denoising steps using Wavelet [3].

The Donoho et.al. algorithm removes the noise by thresholding most of the detail coefficients according to the value of the threshold (λ), which depends on the level of decomposition. There are two approaches to decide whether d_i is a pure noise or an actual coefficient plus noise: soft and hard decisions as given in Equation (1) [18]:

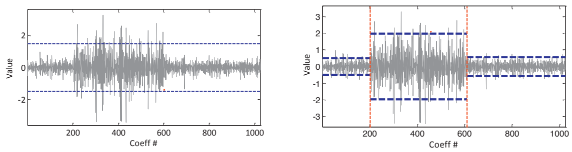
$$\text{Hard: } d_i = \begin{cases} d_i \text{ if } |d_i| > \lambda \\ 0 \text{ else} \end{cases} \quad , \quad \text{Soft: } d_i = \begin{cases} d_i - \lambda \text{ if } d_i > \lambda \\ d_i + \lambda \text{ if } d_i < -\lambda \\ 0 \text{ else} \end{cases} \tag{1}$$

In the hard decision, coefficients are set to zero if they are less than λ or preserved otherwise. This type of decision is modified by the soft decision as all coefficients are considered with noise, so deducting λ from the coefficients which are considered non-pure noise values. Denoising using the soft decision method performs better than the hard decision approach. Thus, in this project the soft decision is considered among the wavelet denoising methods.

Donoho et.al. proposed an approximation to caculate the optimum value of λ, which is based on Stein's Unbiased Risk Estimate (SURE), as: $\lambda = \sqrt{2\log M}$, where M is the number of coefficients. This fixed thresholding per decomposition level is called *SureShrink* [16]. Fig. 3(a) illustrates how a unique threshold is applied to the entire coefficients set. In this Figure, the noise levels are increased, changed several times to show that a single threshold is not always efficient. Therefore utilizing several thresholds for the complete set of coefiecients, improves the denoising performance. The thresholds are chosen according to local noise levels and the method is called the *NeighBlock* [17]. In this algorithm, detail coeffiecients are grouped according to their local properties to find an optimum λ for each group. The coefficients of the j^{th} block are caculated by Equation 2 [19].

$$d_i^j = \max(0, \frac{S_j^2 - \lambda_C z}{S_j^2}) d_i^j \tag{2}$$

Where S_j^2 is the $\|L\|_2^2$ of the j^{th} block, $\lambda_C = 4.505$, and $z = \frac{\log M}{2} + 2\max\left(1, \frac{\log M}{4}\right)$. With this algorithm, the accuracy of signal estimation is improved compared to the *SureShrink*. Fig. 3 (b) shows a different threshod is calculated for each set of set of coefficients according to their noise level .



(a) *SureShrink* (b) *NeighBlock*
Fig. 3. Wavelet thresholding methods.

1.2. Adaptive Filtering

An adaptive filter uses iterative computations to minimize the error "in modelling the relationship between two signals in real time" [20]. Fig. 4 shows a basic diagram of an adaptive filter. Here, the input s_1 represents the ECG which is observed with the additive noise n. The reference signal s is either a pure noise generator or a signal related to n. Since the n and s_1 are uncorrelated, then $E[e^2] = E[(n-y)^2] + E[s_1^2]$. [1]

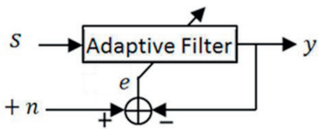


Fig. 4. A general diagram of an adaptive filter.

Let the N coefficients of the filter at the k^{th} iteration be denoted as $W_k = [w_1(k), w_2(k), \dots, w_N(k)]^T$. For an input vector $X_k = [x(k), x(k-1), \dots, x(k-N)]^T$, the output will be given in Equation (3) as:

$$y(k) = \sum_{i=0}^{N} w_i(k) s(k-i) = W_k^T X_k \tag{3}$$

The filter's task is to adjust its weights W iteratively to minimize the mean square error (MSE) between the primary and the reference inputs. This adjustment is mainly achieved by:

Least Mean Squares (LMS)

Because of the important features of LMS: the simplicity [21] and relatively fewer computational operations [20], it is favourable in many applications such as approximation of unknown signals. The weights of the LMS adapting algorithm can be calculated at the k^{th} iteration as in Equation (4):

$$W_{k+1} = W_k + \mu_k e(k) X_k \tag{4}$$

where μ is the step size parameter which controls the rate of convergence. The value of this step size should be optimized empirically to trade off the speed of convergence and instability [1].

Recursive Least-Squares (RLS)

This algorithm is considered more attractive than the LMS in terms of the number of iterations required for convergence, but the price is paid as the number of computations per iteration. Proakis et. al. [22] summarize the iterative steps to update the weights W_k of an RLS filter as in Equation (5):

$$W_k = W_{k-1} + K_k (I - X_k W_{k-1}) \tag{5}$$

where: I is an identity matrix, and the Kalman gain vector (K_k) and correlation matrix (P_k) are given by Equations (6) and (7) respectively:

$$K_k = \frac{P_{k-1} X_k}{\beta + X_k P_{k-1} X_k^*}, \tag{6}$$

$$P_k = \frac{1}{\beta}\left[P_{k-1} - K_k X_k P_{k-1}\right]. \tag{7}$$

$where\ \beta$ is a weighting factor $(0 < \beta < 1)$. Haykin [21] points out that replacing the step-size parameter in LMS by the correlation matrix of the input vector in RLS has a great effect on expediting the convergence rate and achieving the required minimized mean square error at fewer steps.

1.3. Savitzky-Golay filtering

The method of sample smoothing proposed by Savitzky and Golay [23] is based on local least-squares polynomial approximation. By approximating a group of $2M+1$ noisy samples $x[n]$ by the polynomial $p(n) = \sum_{k=0}^{N} a_k n^k$, it is required to minimize the error: $\varepsilon_N = \sum_{n=-M}^{M} (p(n) - x[n])^2$ by selecting the appropriate vector $\mathbf{a} = [a_0, a_1, \dots a_N]^T$. Each considered group of the input has M samples at each side around the central point, or $\mathbf{x} = [x_{-M}, x_{-M+1}, \dots, x_{-1}, x_0, x_1, \dots, x_M]^T$. The output samples $y[n]$ for each of these sets of the input x are computed by the discrete convolution $[n] = \sum_{m=-M}^{M} h[m] x[n-m]$, where the finite impulse response $h[\cdot]$ is equivalent to the least-squares polynomial approximation. The vector **a** can be computed in Equation (8):

$$\mathrm{a} = \left(A^T A\right)^{-1} A^T \mathrm{x} = H\mathrm{x} \tag{8}$$

where the matrix $A = \{n^i\}$, $i = 0,1, \dots N$ and $-M \le n \le M$. For the computation of the impulse response, we set $\mathrm{x} = \mathrm{d} = [0,0, \dots ,0,1,0, \dots ,0,0]^T$. Thus, we obtain the vector: $\tilde{\mathrm{a}} = \left(\mathrm{A}^T \mathrm{A}\right)^{-1} \mathrm{A}^T \mathrm{d}$.

Now, the 0^{th}row of the matrix: $H = \left(\mathrm{A}^T \mathrm{A}\right)^{-1} \mathrm{A}^T$ is $\left[h_{0,-M}, h_{0,-M+1}, \dots, h_{0,0}, \dots, h_{0,M-1}, h_{0,M}\right]$ which equals $[\tilde{p}(-M), \tilde{p}(-M+1), \dots, \tilde{p}(0), \dots, \tilde{p}(M)]$, where $\tilde{p}(n)$ is the polynomial that approximate **d** with least-square error. Thus, $h[-n] = \tilde{p}(n)$.[24-25]. The selection of M and N is important to have the best approximation and smoothing of the noisy data.

2. Current Work and Approach

In this project, ECG signals from [12] are normalized and used after being contaminated with different levels of AWGN. These noisy signals are applied to the denoising techniques: Wavelet (‘*SureShrink’* and *‘NeighBlock’*), Adaptive filtering (LMS and RLS), and Savitzky-Golay filtering. These different algorithms are used to approximate and recover the original ECG signal out of its contaminated one. Next a comparision is made to measure their accuracy of performance. The performace of each method is governed by its operation parameters. These parameters are tuned in this project to achieve the best feasible performance. The following is the summary of the used algorithms and their parameters:

- **Wavelet:** The choice of the mother wavelet function ψ which forms a set of functions (i.e., family of wavelets) and the number of the decomposition levels (lev) is very important in the measure of denoising accuracy. In this project it is found that the combination of the function ψ = ‘sym12’ and lev=3 gives the best performance of denoising for both *universal* and *interval-dependent* thresholds. The soft-decision thresholding is used instead of the hard one because the performance of the former is better.
- **Adaptive Filtering:** for both LMS and RLS, the Adaptive filter length is chosen to be 32 tabs, the number of the coefficients of the LPF (Hamming window-based, linear-phase, and FIR filter) is set to 32 and its normalized cutoff frequency 0.2Hz. Moreover:
  - *LMS:* step size = 0.0276.
  - *RLS:* the size of the Initial correlation matrix inverse is 32x32, and the forgetting factor of 0.91398.
- **Savitzky-Golay filtering:** The polynomial order is 8, and the frame size is31.

3. Discussion of Results

Quantitative performance results of these methods are obtained by looking at the denoised signals in the time domain. Some techniques’ results appear more appealing to the eye than others. Figures 5 and 6 show the denoised signals for a noisy ECG signal at SNR=10 and 20dB respectively. The “err” numbers shown for each method are calculated in Equation (9) as:

$$\text{err} = \sum_i \left(x_i - r_i\right)^2 \tag{9}$$

The comparison can be accomplished using mathematical equations to rank each method’s performance at a given SNR. The introduction of these equations requires the definition of all variables. Let L samples of the measured signal x_i represent the summation of the desired unknown ECG signal samples, s_i and noise samples n_i, or $x_i = s_i + \sigma n_i$, where $i = 1,2, \dots L$, σ is the noise level, and the n_i are i.i.d. N(0,1) (Independent and identically distributed random variables)[17]. The vector $R = \{r_i\}$ is the denoised version of the measured vector $X = \{x_i\}$. The reconstructed signal can be expressed as: $R = \mathbb{C}\{X\}$, where $\mathbb{C}$ represents the denoising algorithm used.

Most of the published papers [3-5,7-8] use the Equations (10) and (11) for comparison between different methods of denoising, (for clarity, denote the summation $\Sigma_{i=1}^{L}$ by Σ_i). Equations (10) and (11) respectively are the Percentage Root-mean-square Difference (PRD) and the estimated Signal to Noise Ratio (SNR_E):

$$\text{PRD} = 100 \times \sqrt{\sum_i \left(x_i - r_i\right)^2 \Big/ \sum_i x_i^2} \tag{10}$$

$$SNR_E = 10 \times \log\left\{\sum_i x_i^2 \Big/ \sum_i \delta_i^2\right\} \tag{11}$$

where δ_i is the deformation in the ECG reconstruction. It is determined for this project that the maximum error is also a good indicator for evaluating the performance of a denoising algorithm. This normalized measure is denoted in Equation (12) by $\Delta_{\max}$:

$$\Delta_{\max} = \max\left|r_i - s_i\right| \tag{12}$$

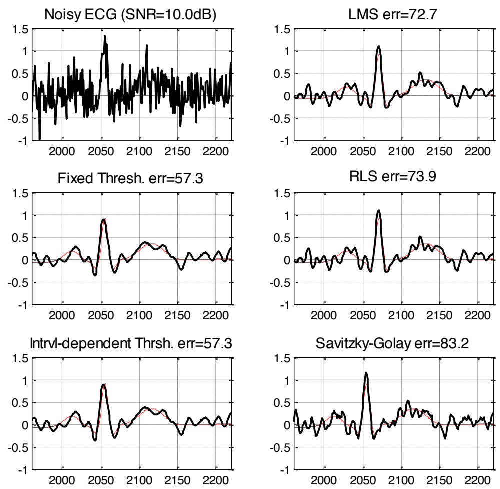


Fig. 5. A time segment showing one noisy ECG pattern at SNR=10dB and the corresponding denoised signals.

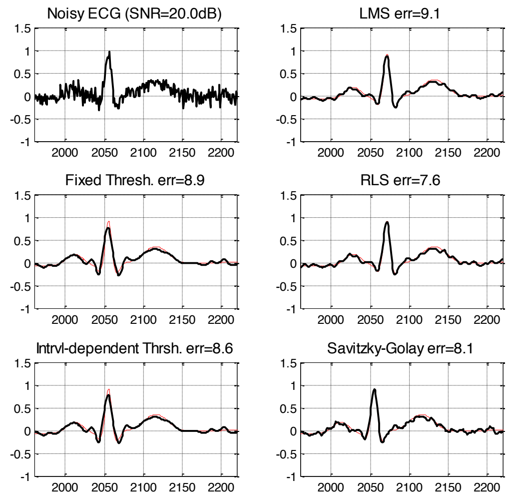


Fig. 6. Time segment shown in Fig. 5 where the SNR=20dB.

It is important to discuss Fig. 7 prior to the discussion of the graphs for Equations (10)-(12), since it supports a greater understanding for these measures. Unlike RLS, the LMS algorithm requires more iteration steps to acquire the desired signal and to achieve the least squared error output, which is clear in Fig. 7 subplots (a, c).

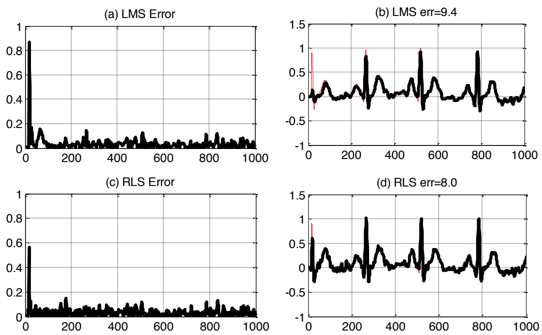


Fig. 7. LMS and RLS startup errors and the corresponding denoised samples of an ECG with SNR=20dB.

Equations (10)-(12) are evaluated and plotted in Fig. 8. These plots provide insight regarding the "*goodness of performance*" for the denoising algorithm at selected SNRs compared to other methods.

In LMS and RLS, the denoised signals accordingly have larger error magnitudes at the beginning of the sequence rather than later in the remaining segments of the signal stream. Since the performance in equations (9)-(12) calculate the total errors for the entire set of samples, these measures for the LMS and RLS are degraded at the start of the adaptation. Although the performance measures in these Figures show that LMS and RLS are the worst along the SNR range, their denoised signals in the time domain are still acceptable and sometimes are very close to those of the other methods.

The results in Fig. 8, indicate that at higher levels of noise (smaller SNRs), the performance of all the methods are similar. This is not surprising because the noise is so high that it extensively distorts the original signal. This is not the case in the actual measurements of ECGs, but it should be included nevertheless for a comprehensive comparison. We notice in Fig. 8 that the interval-dependent threshold wavelet denoising is slightly better than that with universal threshold, which is already expected. That difference is not very noticeable because we apply the same noise level to the entire signal samples. But when this is not the case, we observe a much better performance for the *NeighBlock* algorithm compared to the *SureShrink* one. It is also noticed that at the mid-range SNRs ~15 to 25 dB, both RLS and the Savitzky-Golay filtering perform better than the wavelet.

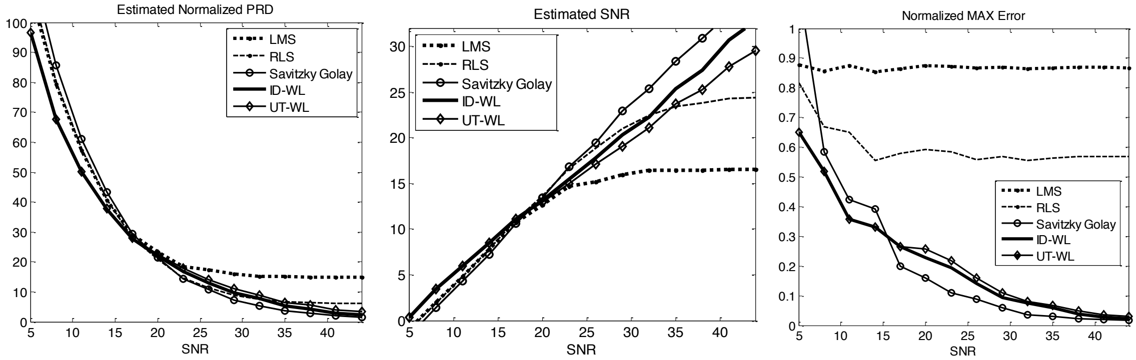


Fig. 8. (a) The Percentage Root-mean-square Difference. (b) Estimated SNR. (c) Maximum error between the original and the denoised signals Where: ID-WL= Interval-dependent threshold wavelet denoising , and UT-WL= universal threshold wavelet denoising.

When the noise of the input signal becomes negligible, the output of the wavelet technique used follows the input without any distortion. Because RLS and Savitzky-Golay methods are based on a filtering approach, the outputs of noiseless inputs have a slight distortion compared to the wavelet outputs.

4. Wavelet Improvement

As stated earlier, we have better performances for the filter based methods at some mid-range SNRs. Therefore, if the SNR of a noisy signal was within that range, a signal can be applied to a LPF prior to the wavelet denoising to obtain a relative improvement in denoising. As a result we see in Fig. 9, the improvement in wavelet denoising with the LPF is only within the mentioned SNR range. Thus, to utilize this option, an SNR estimator can be used to decide when to switch the input to the LPF before the wavelet or to bypass it.

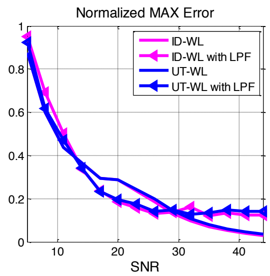
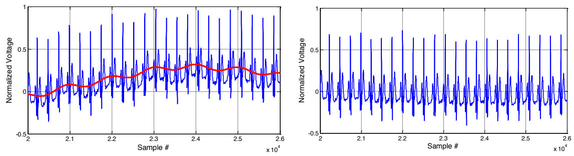


Fig. 9. The effect of the LPF before the wavelet denoising.



(a) before. (b) after.

Fig. 10. Baseline removal using a notch filter around the DC.

5. Baseline removal

The authors in [26] point out that sometimes ECG is recorded with baseline wanders, which mainly results from patient movement or respiration; the frequency of this baseline ranges from 0.5Hz to several Hertz for some periods. As it is suggested in [1], a notch filter around zero is designed in this project and is utilized to remove the DC component. Additionally, very few frequencies occur and result in baseline drifting, as shown in Fig. 10. In this project a LPF with a cutoff frequency very close to the zero is used to remove low frequency content producing this drift.

6. Concluding Remarks

In this study, several methods of denoising are presented and applied to actual ECG signals at noise levels from 5dB SNR to 45dB SNR. The comparison shows that the *NeighBlock* wavelet algorithm performs better denoising than the others. However the RLS and Savitzky-Golay filters perform better in some mid-range SNRs.

References

[1] N. V Thakor and Y. S. Zhu, "Applications of adaptive filtering to ECG analysis: noise cancellation and arrhythmia detection.," *IEEE Trans. Biomed. Eng.*, vol. 38, no. 8, pp. 785–94, Aug. 1991.

[2] B. Arvinti, M. Costache, D. Toader, M. Oltean, and A. Isar, "ECG statistical denoising in the wavelet domain," in *2010 9th International Symposium on Electronics and Telecommunications, IEEE*, 2010, pp. 307–310.

[3] K. Daqrouq and A.-R. Al-Qawasmi, "ECG Enhancement using Wavelet Transform," *WSEAS Trans. Biol. Biomed.*, vol. 7, no. 2, pp. 62–72, 2010.

[4] V. Prasad, T. S. Latha, and M. Suresh, "Denoising of Biological Signals using Wavelets," *Int. J. Curr. Eng. Technol.*, vol. 3, no. 3, pp. 863–866, 2013.

[5] G. Georgieva-Tsaneva and K. Tcheshmedjiev, "Denoising of Electrocardiogram Data with Methods of Wavelet Transform," in *International Conference on Computer Systems and Technologies*, 2013, pp. 9–16.

[6] M. F. M Kania, "Wavelet denoising for multi-lead high resolution ECG signals," *Meas. Sci. Rev.*, vol. 7, no. 2, 2007.

[7] P. S. Gokhale, "ECG Signal De-noising using Discrete Wavelet Transform for removal of 50Hz PLI noise," *Int. J. Emerg. Technol. Adv. Eng.*, vol. 2, no. 5, pp. 81–85, 2012.

[8] A. Awal, S. S. Mostafa, and M. Ahmad, "Performance Analysis of Savitzky-Golay Smoothing Filter Using ECG Signal," *IJCIT*, vol. 1, no. 2, pp. 24–29, 2011.

[9] S. Hargittai, "Savitzky-Golay least-squares polynomial filters in ECG signal processing," in *Computers in Cardiology*, 2005, pp. 763–766.

[10] G. M. S. Sajjad, H. Rahman, A. K. Dey, A. M. Biswas, Z. Islam, and A. K. M. J. Hoque, "Performance Comparison of Modified LMS and RLS Algorithms in Denoising of ECG Signals," *Int. J. Eng. Technol.*, vol. 2, no. 3, pp. 466–468, 2012.

[11] C. Chandrakar and M. K. Kowar, "Denoising ECG Signals Using Adaptive Filter Algorithm," *Int. J. Soft Comput. Eng.*, vol. 2, no. 1, pp. 120–123, 2012.

[12] "PhysioNet." [Online]. Available: http://www.physionet.org/cgi-bin/atm/ATM. [Accessed: 27-Mar-2014].

[13] "DIT Bioelectromagnetism." [Online]. Available: http://bioelectromagnetism.wordpress.com/2012/11/28/data-from-ecg-recording-in-todays-class/. [Accessed: 16-Apr-2014].

[14] "Biomedical Engineering - ECG Assignment." [Online]. Available: http://eleceng.dit.ie/tburke/biomed/assignment1.html. [Accessed: 16-Apr-2014].

[15] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.

[16] D. L. Donoho and I. M. Johnstone, "Adapting to Unknown Smoothness via Wavelet Shrinkage," *J. Am. Stat. Assoc.*, vol. 90, no. 432, p. 1200, Dec. 1995.

[17] T. Cai and B. Silverman, "Incorporating information on neighboring coefficients into wavelet estimation," *Sankhya*, vol. 63, no. Series B, 2001.

[18] T. Ogden, *Essential Wavelets for Statistical Applications and Data Analysis*. Birkhäuser, 1996.

[19] R. Cohen, "Signal Denoising Using Wavelets," 2012.

[20] S. C. Douglas, "Introduction to Adaptive Filters," in *Digital Signal Processing Handbook*, V. K. Madisetti and D. B. Williams, Eds. CRC Press LLC, 1999.

[21] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Prentice Hall, 2001.

[22] J. Proakis and M. Salehi, *Digital Communications*, 5th Ed. McGraw-Hill, 2007.

[23] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Anal. Chem.*, vol. 36, no. 8, pp. 1627–1639, 1964.

[24] R. Schafer, "What Is a Savitzky-Golay Filter? [Lecture Notes]," *IEEE Signal Process. Mag.*, vol. 28, no. 4, pp. 111–117, Jul. 2011.

[25] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd editio. Johns Hopkins University Press, 1996.

[26] N. Rastogi and R. Mehra, "Analysis of Savitzky-Golay Filter for Baseline Wander Cancellation in ECG Using Wavelets," *Int. J. Eng. Sci. Emerg. Technol.*, vol. 6, no. 1, pp. 15–23, 2013.