Baseline correction of ECG using regression estimation method

Jagdeep Rahul[1*], Marpe Sora[2], LakhanDev Sharma[3]
[1]Department of Electronics and Communication Engineering, Rajiv Gandhi University,India.
[2]Department of Computer Science and Engineering, Rajiv Gandhi University ,India.
[3]Department of Electronics and Communication Engineering, MLV Textile & Engineering College, Bhilwara, India.
*Corresponding author
E-mail address: jagdeeprahul11@gmail.com

***Abstract*—The presence of baseline wander in ECG signal severely affects the quality of ECG signal. Baseline wander is generally eliminated at the initial stage of preprocessing of ECG signal. In this paper a method using regression estimation is applied to ECG signal to remove the baseline wander. This method uses two stage median filter and regression estimation for smoothing and correction of baseline. The cross correlation is performed between the baseline wander containing ECG signal with corrected ECG signal. Then cross-correlation coefficient is calculated. The method evolves with the lowess and loess of local regression estimation, rlowess and rloess of robust local regression estimation. We found highest coefficient of 0.9914 for the baseline wander amplitude of 0.5mV using local regression method (lowess) and lowest coefficient of 0.9688 for the amplitude of 0.8mV using robust local regression method (rlowess). Further this method is validated on MIT-BIH, Noise Stress Test Database (NSTD).**

Keywords—Baseline wander, ECG signal, local regression, noise.

I. Introduction

The heart diseases are the main cause of the sudden death in the world due to cardiac arrest. The electrocardiogram (ECG) is the electrical manifestation of the heart activity over the time shown in fig.1 [2]. It can be recorded for the purpose of detecting the various types of abnormalities in the heart [1]. The ECG is a non invasive and most widely used technique for the diagnosis of the heart diseases [10]. The ECG may be recorded in various conditions such as at rest, moving condition, or during stress situation. The recording of ECG at resting mode is performed, when subject is in supine condition for better ECG quality [1]. The moving condition ECG is recorded while patient carries out his daily activities for a period of twenty four hours or more. In this situation, the functioning capacity of heart is evaluated through the ECG recording and it is also capable to detect the arrhythmias such as arterial fibrillation, tachycardia, bradycardia and myocardial infarction [2]. The ECG contaminated with noises, when it is recorded in stress test during the time of acquisition[11,12]. The ECG is mainly affected by the main source of the noise in acquisition process is baseline wander. These are very low frequency noises ranging from 0.05 to 3Hz, inducted due to patient movement, breathing movement, poor body-electrode contact. The baseline wander (BW) noise almost affect the all ECG recording and degrade the accuracy of diagnosis. The presence BW noise in ECG also alters the ST segment characteristics, which play very important rolein determining the acute coronary syndrome. TheBW alsocreates the difficulty in real time monitoring of ECG and arrhythmia detection [3].

Several methods have been proposed by the many researchers to correct the baseline of ECG in order to remove the noise in ECG. In many methods, high pass filters using IIR or FIR have been used to remove the low frequency components by setting the cut off frequency but most of the time frequency spectrum of BW overlaps the ECG frequency spectrum [4]. Hence filtering of low frequency components may cause distortion in ECG. The moving average method is used to extract the BW from the original ECG. Then BW is subtracted from the contaminated ECG to obtain the clean ECG. However it may create distortion in ST segment. The cubic spline curve fitting method is also used to eliminate the BW by using fiducial points. Other methods such as wavelet filtering and empirical mode decomposition can be used to reduce the low frequency noises [13]. In BW removal, these methods give the excellent results but hardly implemented due to theirs cost in real time applications [5].

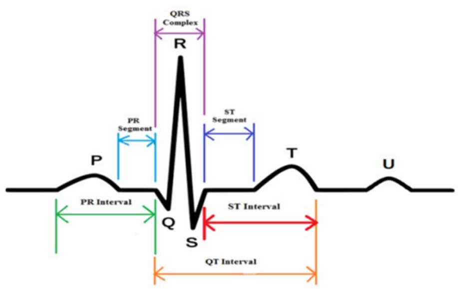

Fig.1 Representation of electrocardiogram (ECG) waveform [2]

In this work, we propose a method using regression estimation for baseline correction for accurate identification of arrhythmia. It accurately estimate the baseline and remove it, same time minimize the distortion and preserve the information of ECG signal. First of all ECG signal pass through the two stage median filter with different window size. Then regression method with and without weights and robustness (rlowess, rloess, lowess, and loess) is applied to estimate the BW. The lowess and loess are derived from the term ‘locally weighted scatter plot smoothening’, proposed by the Cleveland in 1979 [6]. Due to uncertainty in the variation of the BW, it is very hard to accurately estimate the BW by using a particular model. Lowess and loess both are locally weighted linear regression methods and used for smoothening the date. The rlowess and rloess are the robust local regression methods that can be utilized for smoothing the data, which are resistant to outliers. We have considered the standard ECG and baseline wander to verify the performance of lowess, loess, rlowess and rloess. Further it will be validated on MIH-BIH database.

II. METHODOLOGY

The basic preprocessing techniques and filters are mostly used in the ECG signal processing to suppress or remove the noises in ECG. This section is organized in four sub sections artificial ECG signal with baseline wander noise, local weighted linear regression estimation, robust local regression estimation and cross correlation coefficient.

A. ECG signal with Baseline wander noise.

An artificial ECG signal is generated using the ECGSYM software of Physionet, which allows user to configure the ECG parameters like heart rate, sampling rate, frequency and other morphology of ECG waves etc. A ten second duration segment of artificial ECG signal of 70 beats per minutes is chosen for this study, which is shown in fig.2. The baseline wander noise signal is generated and added with artificial signal. The considered BW for this study has assumed that these are generated during the time of breathing and movement of electrodes, which is shown in fig. 3.

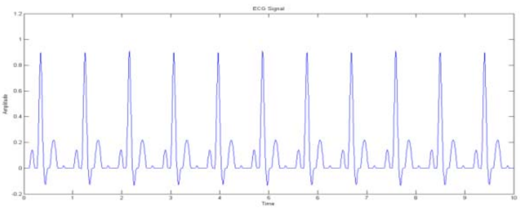

Fig.2. Generated artificial ECG signal

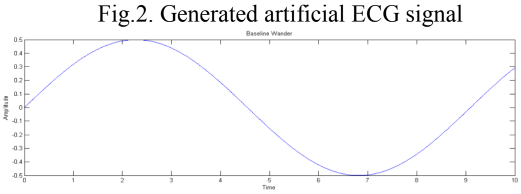

Fig.3. Baseline wander noise.

B. Local Regression Method

Local regression was initially developed for local scatter plot smoothening and also known as local polynomial regression. The lowess, and loess are non-parametric regression and built on classical methods like linear and non-linear least square regression. When each smooth value is generated by the weighted linear least square regression over a span is known as lowess plot. Other hand each smooth value given by weighted quadratic least square regression over a span is known as loess plot. The local polynomials can fit into each subset of the data that are always first or second degree. The higher degree polynomials generally work in theoryonly, numerically unstable and create difficulty in accurate computation [7]. The local regression methods go through the following steps

1. Computation of weights for each data points in a span. Weights are calculated by given below function

$$W_i = \left(1 - \left|\frac{x - x_i}{d(x)}\right|^3\right)^3$$

Where x is the predicted smooth value, xi is the nearest neighbor of x in a defined span and $d(x)$ is the distance from the x, most distant predicted value within the span. Weights should have the following characteristics

(i) The largest weighted data point value to be smoothed within the span.

(ii) The data point outside the span not to be considered and zero weight.

2. First degree polynomial is used in regression of lowess and second degree is used for loess.
3. The final smoothed value is estimated by the weighted linear regression at the predicted value of interest.

C. Local Robust Regression

The local robust regression method is used when data point contains outliers and smoothed value becomes distorted. Due to outlier, smoothed value does not reflect the behavior of bulk of neighboring data points. The small fraction of outlier does not affect the smooth value in robust regression method. The robust lowess and loess are used for smoothening the data. These robust methods use different methods for calculating the weights, which is resistant to outliers [8]. The weight calculating procedure follows these steps

1. Calculation of residuals from smoothening process
2. Computation of robust weights for each data point within the span. Weights are calculated by the equation given below

$$W_i = \begin{cases} (1 - (r_i/6MED)^2)^2 & |r_i| < 6MED, \\ 0 & |r_i| < 6MED \end{cases}$$

Where r_i is the residual point, MED is the absolute deviation of the residuals $|r_i|$.

3. The data is smoothed again using robust weights. Final smooth data values are calculated using both robust weights and regression weights.
4. Last two steps will be repeated for five times.

D. Calculation of cross-correlation coefficient

The cross-correlation coefficient is calculated after removing the baseline wander from the raw ECG signal. The cross-correlation coefficient determines the similarity between the two series as a function of displacement one relative to other [9]. The coefficient (r) is calculated by given below equation.

$$r = \frac{n(\sum XY) - (\sum X)(\sum Y)}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}}$$

Where X is the artificial ECG signal, Y is the ECG signal after removal of baseline wander noise.The cross correlation coefficient value lies between the 0 and 1. The maximum correlation is represented by 1 and minimum correlation is represented by 0.

III. RESULTS AND DISCUSSION

The contaminated ECG signal pass through the two stage median filter to remove the varying baseline in the ECG signal. In the first stage median filter the size of window is equal to the sampling frequency and in second median filter the size window is equal to the half of the sampling frequency. Then regression estimation is performed to smooth the ECG wave for accurate identification and determination of the arrhythmia in the ECG recording. For validating this technique, the cross correlation coefficients are calculated with lowess and loess, using both weight and without weight local regression estimation. The cross correlation coefficient performed between the clean artificial ECG and corrected ECG signal after the subtraction of baseline wander (BW) noise. The fig.4 (a) shows the ECG signal with baseline wander noise or raw ECG signal. The fig.4 (b) shows the output of first median filter having the window size equal to the sampling frequency. The fig. 4(c) shows the output of second stage median filter having the window size equal to the half of the sampling frequency. Then the extracted baseline wander is subtracted from the contaminated ECG signal. After subtraction the regression estimation method is applied to the ECG signal for smoothing and correction of the ECG signal, which is shown in fig. 4(d).

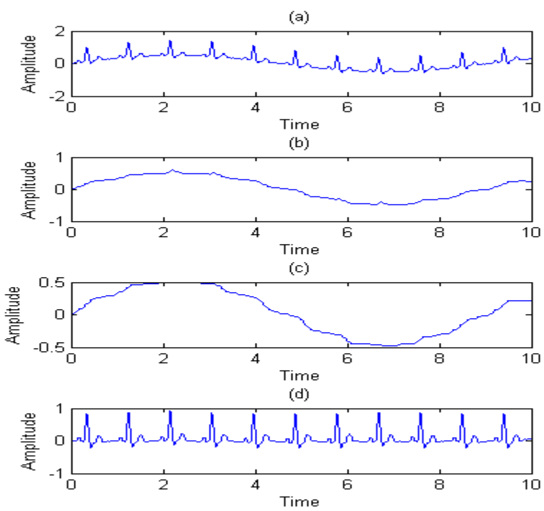

Fig.4. (a) ECG signal with Baseline wander noise. (b) Baseline wander noise after first median filter. (c) Baseline wander noise after second median filter (d) Clean ECG signal after subtraction and regression estimation.

In this study we have also considered six different amplitude baseline wander noises for validation of this technique through the calculation of cross correlation coefficients. The various types of BW are considered in this study, whose amplitude is ranging from 0.5 to 1 mV, are shown in fig.5. The cross correlation coefficients for different baseline wanders are shown in Table.1. All coefficients are found in the acceptable range from 0.9688 to 0.9914. We found highest coefficient for the BW amplitude of 0.5mV using local regression method (lowess) and lowest for the amplitude of 0.8mV using local robust regression method (rlowess).

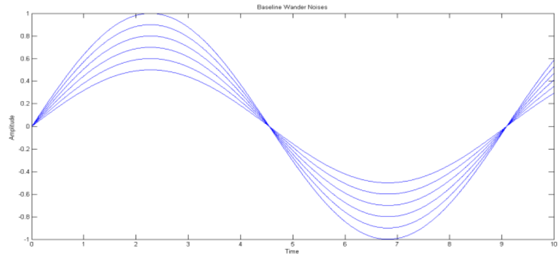

Fig.5. Different amplitude baseline wanders

Table1. Cross-correlation coefficient of baseline wanders with different amplitude.

Sl. No.	Amplitude(mV)	lowess	loess	rlowess	rloess
1	0.5	0.9914	0.9913	0.9841	0.9912
2	0.6	0.9895	0.9894	0.9817	0.9896
3	0.7	0.9874	0.9874	0.9779	0.9873
4	0.8	0.9851	0.9851	0.9688	0.9843
5	0.9	0.9828	0.9828	0.9750	0.9827
6	1.0	0.9802	0.9803	0.9818	0.9804

This method is applied to the MIT-BIH Noise Stress Test Database (NSTD) for further validation of this method using local regression (lowess).

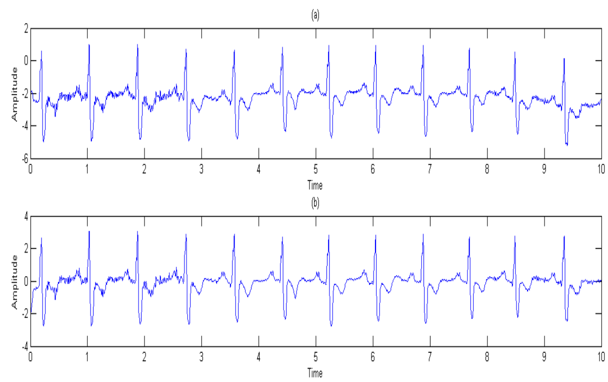

Fig.6. (a) MIT-BIH, ECG signal record 118e00m of NSTD (b) Baseline corrected ECG signal.

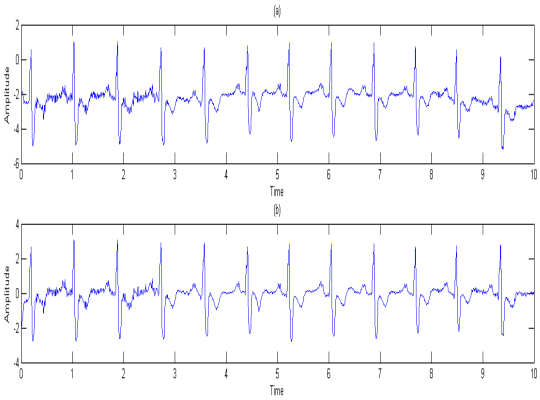

Fig.7. (a) MIT-BIH, ECG signal record 118e06m of NSTD (b) Baseline corrected ECG signal.

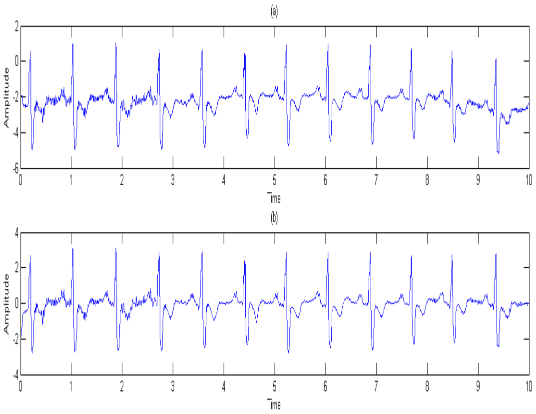

Fig.8. (a) MIT-BIH, ECG signal record 118e12m of NSTD (b) Baseline corrected ECG signal

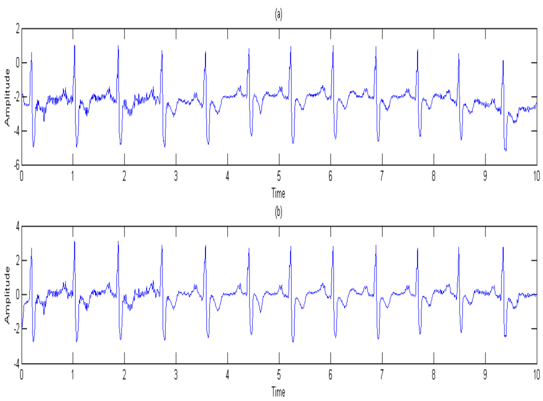

Fig.9. (a) MIT-BIH, ECG signal record 118e18m of NSTD (b) Baseline corrected ECG signal

The regression estimation method on MIT- BIH Noise Stress Test Database (NSTD) applied and gives the good results in acceptable range. The fig.6 (a) shows the 118e00m signal record from the database and fig. 6(b) shows the processed ECG signal after applying the regression method. This method is applied on ECG records 118e06m, 118e12m and 118e18m on MIT-BIH database respectively and results are found in acceptable range.

IV. CONCLUSION

The removal and correction of baseline wander (BW) in ECG is become very important step in ECG signal processing. In this work we proposed a method to remove the baseline wander using regression estimation. The clean ECG signal is obtained after the subtraction of BW from the raw ECG signal. This method is applied on variety of BW having different amplitude ranging from 0.5mV to 1mV. The simulated results shows that proposed method can effectively correct the baseline of ECG signal with minimum distortion. The performance of proposed method is depends on window size of median filters and also on type of local regression. We found that use of 'lowess' in local regression estimation methods gives the better results and suitable for ECG signal processing for devices.

References

1. UditSatija, BarathramRamkumar, M. SabarimalaiManikandan, "A robust sparse signal decomposition framework for baseline wander removal from ECG signal", Region 10 Conference (TENCON) 2016 IEEE, pp. 2470-2473, 2016.
2. Nang AnijaManlong, Jagdeep Rahul, MarpeSora, "ST Segment Analysis for Early Detection of Myocardial Infarction", International Journal of Computer Sciences and Engineering, Vol.6, Issue.6, pp.1500-1504, 2018.
3. X. Tan, X. Chen, R. Ren, X. Hu, B. Zhou, Z. Fang, S. Xia, "Real-time baseline wander removal in ECG signal based on weighted local linear regression smoothing", *IEEE Int. Conf. Inform. Automat. (ICIA)*, pp. 453-456, Aug. 2013.
4. Meyer, C.R. & Keiser, H.N.. (1977). Electrocardiogram baseline estimation and removal using cubic splines and space-state computation techniques. Computers and Biomedical Research. 10. 459-470. 10.1016/0010-4809(77)90021-0.
5. O. Sayadi, M. B. Shamsollahi, "Multi adaptive bionic wavelet transform: application to ECG denoising and baseline wandering reduction", *J. Adv. Signal Process*, vol. 11, 2007.
6. Cleveland, William S. (1979). "Robust Locally Weighted Regression and Smoothing Scatterplots". Journal of the American Statistical Association. 74 (368): 829 836. doi:10.2307/2286407.
7. Cleveland, William S.; Devlin, Susan J. (1988). "Locally-Weighted Regression: An Approach to Regression Analysis by Local Fitting". Journal of the American Statistical Association. 83 (403): 596–610. doi:10.2307/2289282.
8. Cleveland, William S. (1981). "LOWESS: A program for smoothing scatterplots by robust locally weighted regression". The American Statistician. 35 (1): 54. doi:10.2307/2683591.
9. Chaw M. Khamhoo, Jagdeep Rahul, MarpeSora."Algorithm for QRS Complex Detection using Discrete Wavelet Transform", International Journal of Electronics Engineering. Vol.10, Issue.2 , pp. 352-357, 2018.
10. Jagdeep Rahul, MarpSora, LakhanDevSharma."An overview on Biomedical signal analysis", International Journal of Recent Technology and Engineering. Vol. 7, Issue 5, pp.206-209. 2019.
11. Srivastva R., Singh Y.N. "Identifying Individuals Using Fourier and Discriminant Analysis of Electrocardiogram". Communications in Computer and Information Science, vol 834. Springer, Singapore.ICMC 2018
12. Srivastva, R., Singh, Y.N.: Human recognition using discrete cosine transform and discriminant analysis of ECG. In: Proceedings of IEEE 2017 Fourth International Conference on

Image Information Processing, JUIT, Solan, pp. 368–372 (2017)

13. LakhanDev Sharma, Ramesh Kumar Sunkaria. “A robust QRS detection using novel pre-processing techniques and kurtosis based enhanced efficiency”. Measurement 87 (2016) 194–204. http://dx.doi.org/10.1016/j.measurement.2016.03.015.

978-1-7281-1253-4/19/$31.00 © 2019 IEEE