Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 85 (2016) 483 – 489

Procedia Computer Science

International Conference on Computational Modeling and Security (CMS 2016)

Smoothening and Segmentation of ECG Signals Using Total Variation Denoising –Minimization-Majorization and Bottom-Up Approach

Om Prakash Yadav*, Shashwati Ray

Deptt. Of Electronics and Telecommunication,BIT Durg,491001, India
Deptt. Of Electrical Engineering, BIT Durg, 491001,India

Abstract: An ECG Signal records electrical activity of heart. It includes information on heart's rhythm and is useful for diagnosis of heart related diseases. It encounters with various artifacts during acquisition and transmission. The unwanted signals/noises present in ECG signals disturb the clinical information present in it. This paper tries to reduce unwanted signals through Majorization-Minorization approach to optimize total variation in the signals. The denoised signal is then segmented using bottom up approach. The results show significant improvement in signal to noise ratio and successful segmentation of sections of ECG signals.

Keywords: ECG Signal, Total Variation, Majorization –Minimization, Bottom –Up approach.

1. Introduction

The Electrocardiogram (ECG) is a graphical representation of electrical signals generated by heart. These signals are time varying and are result of expansion and contraction of heart muscles. The surface ECG is obtained by recording the potential difference between two electrodes placed on the surface of the skin. A single normal cycle of the ECG represents the successive atria depolarization/repolarization and ventricular depolarization /repolarization which occur with every heartbeat[1]. These can be approximately associated with the peaks and troughs of the ECG waveform labelled P, Q, R, S, and T as shown in Fig. 1.

* Corresponding author. Tel.: 08966000053;
E-mail address:opyadav@csitdurg.in

Peer-review under responsibility of the Organizing Committee of CMS 2016
doi:10.1016/j.procs.2016.05.195

Fig. 1 An ECG Signal

Any disorder in heart rate or rhythm or change in the morphological pattern of ECG is an indication of arrhythmia. Real signals encounters with noise. Noisy signal results in inaccurate diagnosis of disease. So there is a need to denoise the signals. Each wave of ECG is clinically important. Signal segmentation may also not be accurately analysed. Therefore, reducing noise is an important issue. This has become increasingly important in Cardiac related issues. Filters not only can reduce unwanted noise and short-term components of a signal, but also increases the processing time[2].

Rest of the paper is organized as follows: In section 2, we give the required background related to ECG smoothening and Total Variation Denoising (TVD) using Majorization-Minorization (MM) approach. In section 3, we describe signal segmentation and how Bottom –Up approach is applied for signal segmentation. Section 4 deals with the implementation and results obtained. The final section is concluded with discussions regarding the presented approach.

2. ECG Smoothening

The signal encounters various types of artifacts during acquisition, transmission and storage. The noises introduced are due to power line interference (PLI), body movement's, electrode contacts, electromagnetic field interference, respiration movements[1] etc. Presence of noises in ECG signals degrades the signal quality and thus affects the visual diagnosis and feature extraction. Several denoising techniques are available in literature. Many researchers have used digital Infinite Impulse Response filters[3] (IIR) filters to reduce the effect of PLI noise and baseline wander from ECG signals. In some related works, filtering approaches based on time-frequency, wavelet transform have been proposed. Median filters, Empirical Mode Decomposition (EMD), Non-linear Bayesian filtering techniques were also used to reduce noises present in ECG signals. Statistical techniques like Principal Component Analysis[4] (PCA), Independent Component Analysis[5, 6] (ICA), Neural Networks[7] have also been used to extract clinically important signals.

TVD using MM is an optimized approach for noise reduction and preservation of sharp edges of signals[7]. TVD estimates the signal *x(n)* by solving the optimization problem :

$$\arg\min\left\{F(x) = \frac{1}{2}\sum_{n=0}^{N-1}|y(n) - x(n)|^2 + \lambda\sum_{n=1}^{N-1}|x(n) - x(n-1)|\right\} \tag{1}$$

where *x(n)*- piecewise constant signal, *y(n)*-noisy signal, *n* represent number of samples.

The first term represents the mean square error between the noisy signal and the reconstructed signal. The regularization parameter λ controls how much smoothening is performed. Increasing λ gives more weight to the second term which, measures fluctuation of the signal *x(n).*

The MM[8,9] approach to minimize the function F(x) can be summarized as :

Algorithm # 1

1. Store the ECG signal as data and set the number of data points
2. Set λ =2.
3. Set number of iterations as 40
4. Set k=0. Initialize x(0)
5. Choose $G_k(x)$ such that
   $G_k(x) > F(x)$ for all x
   $G_k(x) = F(x(k))$
6. Set X(k+1) as the minimizer of $G_k(x)$
7. Set k=k+1 and go to step 5.

The ECG signals are taken from MIT-BIH arrhythmia database[10,11]. The duration of the signals were taken as 10 sec and sampled as 720Hz sampling frequency with 11 bits per sample of resolution. The signals were corrupted with three different types of noises (generated in Matlab) namely power line Interference (PLI), White Gaussian (WG) and Random Noise (RN). The signal to noise ratio of corrupted signal was made to about 12 dB.

This technique has advantages over simple techniques, that it is remarkably effective at simultaneous preserving edges whist smoothing away noise in flat regions, even at low signal to noise ratios.

Performance of the technique was measured in terms of Signal to Noise Ratio (SNR), Root Mean Square Error (RMSE) and Correlation Coefficients (CC). Application of this technique showed significant improvement in SNR while maintaining the shape of signal. RMSE also was found to be in satisfactory limits. The optimization function for denoising using TVD-MM depends upon smoothening parameter too. It is observed that error increases with increase in smoothening parameter if it is applied to complete wave. So there is a need to segment ECG signals for better denoising results.

3. Segmentation

The purpose of the segmentation is dividing a signal to several with the same statistical characteristics such as amplitude and frequency. Since statistical characteristic of ECG changes with time, so ECG signals are considered as non-stationary signals. Analysis of stationary signal is easier as compared to non- stationary signal, signal segmentation is applied as pre-processing step for non-stationary signal analysis[12].

A segmentation of the integers {1, 2, ..., T} is a sequence $t = (t_0, t_1, ..., t_K)$ which satisfy $0=t_0< t_1 \ldots\ldots < t_{K-1}< t_K=T$. The intervals of integers $[t_0 +1, t_1]$, $[t_1+1 ..., t_2]$, ..., $[t_{K-1} +1, t_K]$ are called segments, the times $t_0, t_1, ..., t_K$ are called segment change points and *K*, the number of segments, is called the order of the segmentation. In many applications a time series *x1, x2, ..., xT* is given and we seek a segmentation of {1, 2, ..., T} which corresponds to changes of the behaviour of *x1, x2, ..., xT* . So segmentation can be formulated as an optimization problem. The segmentation cost function J (t) can be formulated as follows[13]:

$$J(t) = \sum_{k=1} d_{t_{k-1}+1,t_k} \quad (2)$$

where $d_{s,t}$ (for $0 < s < t < T$) is the segment error corresponding to segment $[s,t]$. The optimal segmentation, denoted as $\hat{t} = (\hat{t}_0, \hat{t}_1 ... \hat{t}_k)$ **is** defined as:

$$\hat{\boldsymbol{t}} = \mathbf{ar\,g\,min}_{t \epsilon T} \boldsymbol{J}(\boldsymbol{t})$$

(3)

The segment error $d_{s,t}$ depends on the data x1, x2, ..., xT .

Time series segmentation is referred to as a pre-processing step for variety of data mining tasks such as a trend analysis technique, as a discretisation problem in function of dimensionality reduction, as a component in data mining applications in various fields, etc.

The various sections of an ECG signal have different physiological meaning and the presence, timing, and duration of each segment have diagnostic and biophysical importance. The problem is made more difficult as the shape of an ECG is variable both within and across patients. This variability depends upon various factors like position of electrodes placed on the body, muscle artifacts noise etc. Thus the challenge is to develop automatic segmentation tools that are robust to inherent variability in the signal while relying on heuristics as minimally as possible[14].

Segmentation of ECG is a procedure to find out various waves like P wave, ORS complex, S waves etc present in it. Differentiation forms the basis of many QRS detection algorithms. Since it is a high-pass filter, the derivative function amplifies the higher frequencies characteristic of the QRS complex and attenuates the lower frequencies of the P and T waves. An algorithm based on first and second derivatives[15] originally developed in 1977 was modified for use in high-speed analysis of recorded ECGs by Ahlstrom and Tompkins[16] in 1983. Algorithms to detect noise sensitivity among QRS detectors were also developed and compared[17].

3.1 Bottom-Up Approach

The Bottom–Up algorithm, often called as iterative merge, begins by dividing the original time series data, of length *n*, into a large number of very small segments with equal lengths. In the next step, based on the comparison of each pair of consecutive segments, the pairs that cause the smallest increase in the error are being identified, and consequently merged in one new, bigger segment. The algorithm repeats until some of the defined stopping criteria (a) *k* number of segments, and/or, (b) approximation error > specified threshold, is satisfied[18].

Algorithm # 2.

1) Load the signal
2) Set the numbers of segments(Required number of segments) ⟶ K
3) Set the number of points (window length) for initialization of segments ⟶ M M<<K
4) Approximate Each M with polynomials of order (here we have taken 1) ⟶ N
5) Calculate Error of each M with N
6) If N > K
   a. Find Minimum Error and corresponding segment
7) Repeat step 4 until N < K-1

The bottom-up segmentation algorithm is based on the assumption that two adjacent time series segments can only be merged in case of a low reconstruction error, implying that the correlation structure among the observed variables does not considerably change within the two contiguous segments.

4. Implementation and results

The ECG signals are taken from MIT-BIH database. The duration of the signal was 10 seconds and is sampled at 720 Hz sampling frequency with 11 bits/sample resolution. The signals are corrupted by adding noises (generated in Matlab).

$$y(n) = x(n) + w(n) \quad, n = 0,1,\ldots,N-1 \tag{4}$$

where *x(n)* is original signal taken from MIT-BIH database and *w(n)* is White Gaussian Noise.

We have added three different types of noises namely PLI of 50Hz, White Gaussian noise and random noise. The signals are distorted so as to make their SNR as about 12 dB. The noisy ECG signals are then denoised using TVD-MM approach (Refer Algorithm # 1). The fig 2 indicates original ECG, White Gaussian noise, Corrupted signal and reconstructed signal.

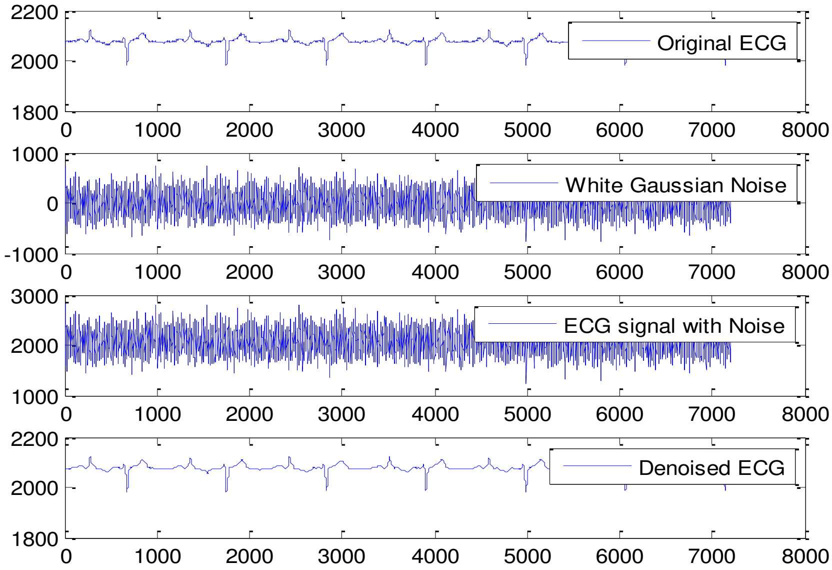

Fig. 2 Original ECG, White Gaussian Noise, ECG with the Noise, Denoised ECG

The performance the technique is evaluated in terms of improvement in SNR, RMSE and CC. Table 1 shows the results obtained for three signals. For better denoising technique there should be improvement in SNR, RMSE should be as minimum as possible and the value of CC should be close to 1. From the table 1, we can see a significant improvement in SNR (about 18 dB) for all the ECG signals. RMSE obtained is more than 0.5 which can be considered good for denoising. CC values are also nearly equal to 1 indicating shape preservation.

Table 1.Noise, ECG record and Performance Matrices

Noise	ECG record	SNR(dB)	SNR(dB)	RMSE	CC
PLI	aami3am.mat	12.03	30.59	0.62	0.99
	aami3cm.mat	11.99	30.71	0.63	0.99
	aami3dm.mat	12.01	30.80	0.60	0.99
WGN	aami3am.mat	12.20	29.96	0.62	0.99
	aami3cm.mat	12.01	30.09	0.63	0.99
	aami3dm.mat	12.25	30.80	0.59	0.99
RN	aami3am.mat	12.40	29.84	0.62	0.99
	aami3cm.mat	10.91	30.76	0.62	0.99
	aami3dm.mat	10.42	31.12	0.60	0.99

The denoised signal was then applied to Bottom-Up approach (Refer Algorithm #2) for segmentation. Initially number of segments was fixed to 10. Results were not satisfactory. Then we have increased the number of segments to 25. Results were quite satisfactory.

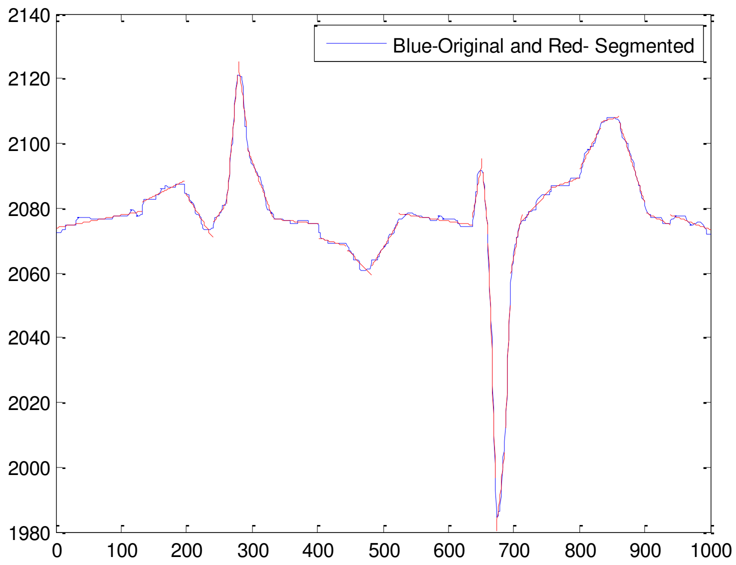

Fig.3 Denoised signal And Segmented ECG signal

Fig 3 shows the denoised signal and the segmented section. It is observed that first order polynomial is not sufficient to approximate all sections of ECG specially QRS complex. Even P, S and T wave requires second order polynomial.

5. Conclusion

In this paper TVD-MM approach is used for denoising of ECG signals. The Signals were taken from MIT-BIH database. The signals were then made noisy. These noisy signals were then denoised using TVD-MM approach. It is observed that the approach successfully denoise ECG signals. The extent of smoothness depends upon smoothening parameter λ too. Since the approach was applied over the complete signal so only 18 dB of SNR improvement was observed. The value of RMSE was also found to be in limits. The Value of CC was near to 1. The denoised signal is then segmented into various sections by Bottom-Up approach. It is observed that accuracy increase with increase in number of segments. Here we have approximated all sections by first order polynomials. More accurate results may be obtained by proper selection of polynomials for approximation of sections i.e. for linear shapes- linear polynomial and for curved shapes-higher order polynomials. The individual sections may provide better results for TVD-MM techniques.

References

1. Acharya U R, Suri J S, Spaan J A E, Krishnan S M. *Advances In Cardiac Signal Processing*. New York: Springer Berlin Heidelberg;2007.
2. Gao J, Sultan H, Hu J, Tung W W. *Denoising Nonlinear Time Series by Adaptive Filtering and Wavelet Shrinkage: A Comparison*. IEEE Signal Processing Letters 2010;17(3):237-240.
3. Sorensen J, Johannesen L, Grove U, Lundhus K, Couderc J P, Graff C. *A Comparison of IIR and Wavelet Filtering for Noise Reduction of the ECG*: Computing in cardiology. 2010(37):489-492.
4. Moody G B, Mark R G. *QRS Morphology Representation and Noise Estimation using the Karhunen-Loeve Transform*:Computers in Cardiology.1989:269-272.
5. Borros A K, Ali M, Ohnishi N. *Removing Artifacts from ECG Signals using Independent Component* Neurocomputing.1998(22):173-186.
6. He T, Clifford G., Tarassenko L. *Application of Independent Component Aanalysis in Removing Artifacts from the Electrocardiogram*. Neural Computing. & Applications. 2006; 15(2): 105-116.
7. Clifford G, Tarassenko L, Townsend N. *One-pass training of optimal architecture auto-associative neural network for detecting ectopic beats*. Electronic Letters.2001; .37(18):1126-1127.
8. Rudin L I, Osher S, Fatemi E. *Nonlinear Total Variation Based Noise Removal Algorithms*. Physica D.1992; 60:259-268.
9. Selesnick, I W, Ilker B. *Sparse Signal Estimation by Maximally Sparse Convex Optimization*. Signal Processing, IEEE Transactions. 2014;62(5): 1078-1092.
10. Moody G B ,Roger G M. *The MIT-BIH arrhythmia database on CD-ROM and software for use with it*. Computers in Cardiology. 1990:185-188.
11. Goldberger A L, Amaral LAN, Glass L, Hausdorff J M, Ivanov P Ch, Mark RG, Mietus J E, Moody G B, Peng C K, Stanley H E. *PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals*. Circulation.2000; 101(23):e215-e220.
12. Kehagias A, Nidelkou E, Petridis V. *A dynamic programming segmentation procedure for hydrological and environmental time series*. Stochastic Environmental Research and Risk Assessment.2006; 20(1-2): 77-94.
13. Keogh E, Chu S, Hart D, Pazzani M. *An Online Algorithm For Segmenting Time Series*. Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on. IEEE, 2001.
14. Pan J, Tompkins W J. *A Real-Time QRS Detection Algorithm*. Biomedical Engineering, IEEE Transactions.1985;3: 230-236.
15. Balda R A, Diller G, Deardorff E , Doue J,Hsieh P . The *HP ECG Analysis Program*. Trends in computer-processed electrocardiograms.1977: 197-205.
16. Ahlstrom M L, Tompkins W J. *Digital Filters For Real-Time ECG Signal Processing Using microprocessors*. Biomedical Engineering, IEEE Transactions.1985;9:708-713.
17. Friesen G M, Jannett T C, Jadallah M A, Yates S L, Quint S R, Nagle H T . *A Comparison Of The Noise Sensitivity Of Nine QRS Detection Algorithms*. Biomedical Engineering, IEEE Transactions.1990; 37(1): 85-98.
18. Spiegel S, Gaebler J, Lommatzsch A, De Luca E, Albayrak S. *Pattern Recognition And Classification For Multivariate Time Series*. In Proceedings of the fifth international workshop on knowledge discovery from sensor data.ACM. 2011: 34-42.