Denoising in Biomedical signals using Ensemble Empirical Mode Decomposition

Megha Agarwal1, Richa Priyadarshani2

[1]*(Electronics and Communication, Amity University, Greater Noida India)*
[2]*(Electrical and Electronics, Amity University, Greater Noida India)*

Abstract: *In this paper a novel Ensemble Empirical Mode decomposition (EEMD) and adaptive filtering is proposed to filter out Gaussian noise and contact noise contained in raw biomedical signals. Real Biomedical signals from the MIT-BIH database are used to validate the performance of the proposed method. It has been observed that original signals can be significantly enhanced by using the proposed method where the contact noise is eliminated while useful features of original signals are kept. The results also show that the proposed method is quite effective to reduce noise from ECG signals and many other biomedical signals with a very small mean square error.*
Keywords: *Ensemble Empirical Mode Decomposition; Gaussian noise; Contact noise; NADA; Adaptive filtering.*

I. Introduction

Most of the signals in raw format are represented in time domain signals. But in many cases, the most important structure is hidden in frequency domain. Historically, there are various methods to obtain frequency content from the raw signals; among them Fourier and Wavelet analysis are the most commonly used method. But they have some limitations as Fourier transformation is valid under extremely general conditions, i.e. the system must be linear; and the data must be strictly periodic or stationary. On the other hand Wavelet analysis uses fixed basis function and method is not adaptive. So, to overcome the limitation of existing methods in 1998 Huang introduced a new tool called Empirical Mode Decomposition (EMD) associated with Hilbert transform to perform comprehensive analysis of non-linear and non-stationary data.

Empirical Mode Decomposition (EMD) [1] is an adaptive and fully data driven method, which is developed to analyze non-linear and non-stationary signals. It decomposes the signal in fast and slow oscillations called Intrinsic Mode functions (IMFs). However, EMD suffers from a problem known as 'mode mixing'. Mode mixing means presence of oscillations of very disparate amplitude in a mode, or the presence of very similar oscillations in different modes. Mode mixing is a consequence of signal intermittency. As discussed by Huang et al. (1998 and 1999), the intermittence could not only cause serious aliasing in the time-frequency distribution, but also make the individual IMF devoid of physical meaning. To overcome the scale separation problem without introducing a subjective intermittence test, a new noise-assisted data analysis (NADA) method is proposed, the Ensemble EMD (EEMD) [3], which defines the true IMF components as the mean of an ensemble of trials, each consisting of the signal plus a white noise of finite amplitude. With this ensemble approach, we can clearly separate the scale naturally without any a priori subjective criterion selection.

The addition of white Gaussian noise solves the mode mixing problem by populating the whole time-frequency space to take advantage of the dyadic filter bank behavior of the EMD [2].

II. Ensemble Emd Algorithm

EMD [1] was introduced by Huang et al which decomposes a given signal x(t) into a finite number of sub components called IMFs through an iterative process called sifting. The decomposed signal will be called IMFs if it satisfies two conditions:

a) the number of extrema and the number of zero crossing must be equal or differ at most by one; and
b) the mean value of the upper and lower envelope is zero everywhere.

EEMD defines the "true" IMF [6] components as the mean of the corresponding IMFs obtained via EMD over an ensemble of trials, generated by adding different realizations of white noise of finite variance to the original signal . EEMD algorithm can be described as:

The steps for EEMD are as follows [3].

a) Initialize the number of ensemble I.
b) Generate $x^i[t] = x[t] + w^i[t]$ $(i = 1, \ldots, I)$ are different realization of white Gaussian noise.

c) Each $x^i[t]$ $(i = 1, \ldots, I)$, is fully decomposed by EMD getting their modes $IMF_k^i[t]$, where k=1, 2, K indicates the modes.

d) Assign $\overline{IMF_k}$ as the k-th mode of x[t], obtained as the average of the corresponding

$$IMF_k^i : \overline{IMF_k}[t] = \frac{1}{I}\sum_{i=1}^{I} IMF_k^i[t]. \qquad (1)$$

In this paper just as the EMD method, the given signal, x(t) can be reconstructed according to the equation(2) which has a very small mean square error with the original signal. Table 1 shows these values.

$$x(n) = \sum_{k=1}^{K} \overline{IMF_k}(t) + \bar{r}(t) \qquad (2)$$

where $\overline{IMF_k}[t] = \frac{1}{I}\sum_{i=1}^{I} IMF_k^i[t]$ and $\bar{r}(t) = \frac{1}{I}\sum_{i=1}^{I} r_i(t)$

The EEMD described here employs all the important characteristics of noise. Its principle is simple: when a collection of white noise is added to the target signal it cancels each other out in a time space ensemble mean. The reason is obvious that the added white noise would populate the whole time-frequency space uniformly with the constituting components of different scales separated by the filter bank.

III.   Experiments And Result

Here Real Biomedical signals (ECG, EMG, BP & Neurodegenerative disease signal) from the MIT-BIH database are used to validate the performance of the proposed method. All the simulation is done in MATLAB only.

IV.   Ecg Signal

The electrocardiogram (ECG) has been widely used for diagnosis purposes of heart diseases[5]. Good quality ECG is utilized by the physicians for interpretation and identification of physiological and pathological phenomena. However, in real situations, ECG recordings are often corrupted by artifacts.
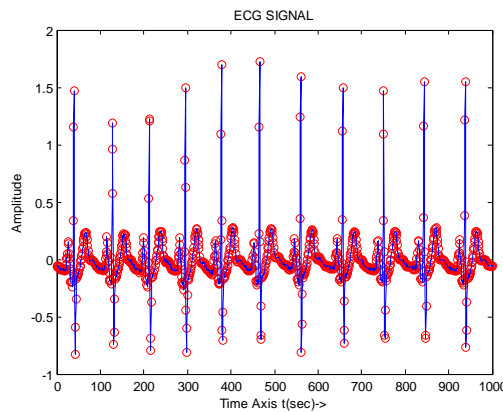

Figure: 1 Typical waveform of ECG

Generally three different patterns of noise are added to ECG signal which are following [4]:
a) Muscle Contractions (EMG): Muscle contractions cause artifactual millivolt level potentials to be generated. Electromagnetic sources from the environment may overlay or cancel the signal being recorded from a muscle. It can be modelled by a random number with normal distribution, originally manipulated with the Matlab code randn.m. The maximum noise level is (1/8) Vpp (peak to peak voltage).

b) Power line interference: Power line interference is modelled by 50 Hz sinusoidal function with multiplication on amplitude derived with Matlab code rand.m. The maximum noise level is (1/4) Vpp.

c) Baseline wander: Baseline wander is modelled by a Baseline wander a 0.333 Hz sinusoidal function. The maximum noise level is (1/8) Vpp.

Following are the typical waveforms of ECG signal (fig. 2 & 3) along with reconstructed signal via EEMD method.
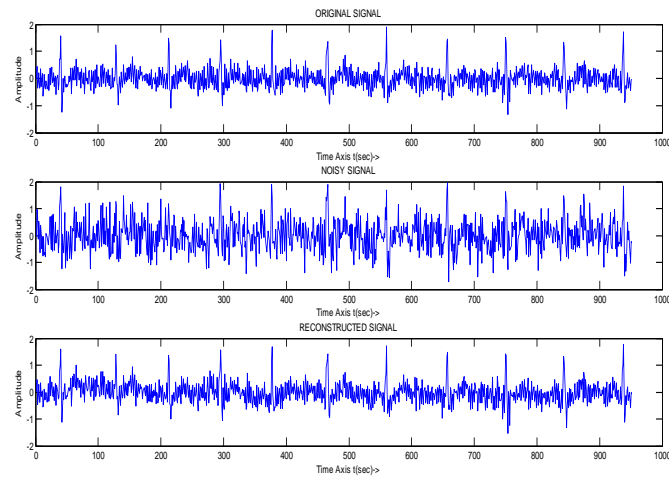
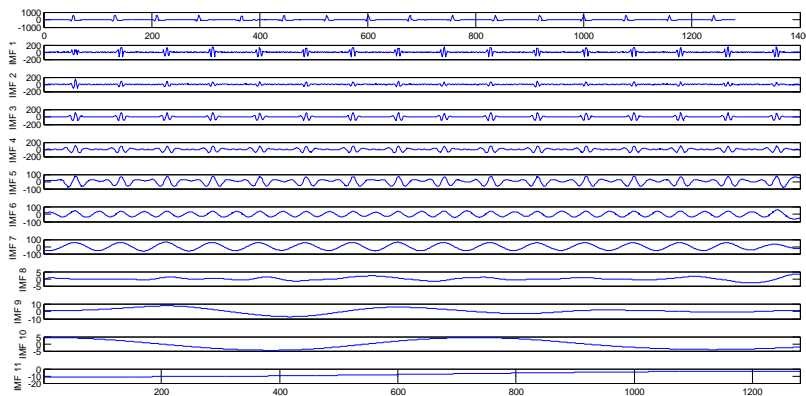Figure: 2 Noisy ECG signal reconstructed via EEMD method having SNR=-1 and I=20

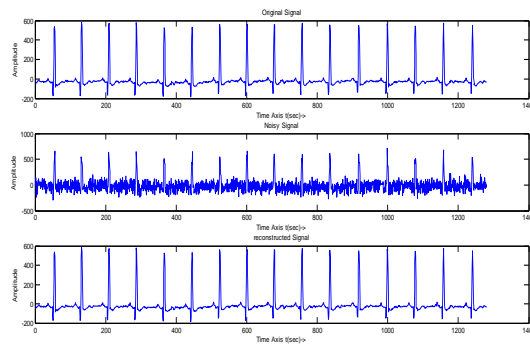Figure: 3 Set of decomposed IMFs of ECG signal via EEMD method

Figure: 4 Noisy ECG signal reconstructed via EEMD method with I=500

V. Bp Signal

Blood pressure (BP), sometimes referred to as arterial blood pressure, is the pressure exerted by circulating blood upon the walls of blood vessels, and is one of the principal vital signs. When used without further specification, "blood pressure" usually refers to the arterial pressure of the systemic circulation. During each heartbeat, blood pressure varies between a maximum (systolic) and a minimum (diastolic) pressure.
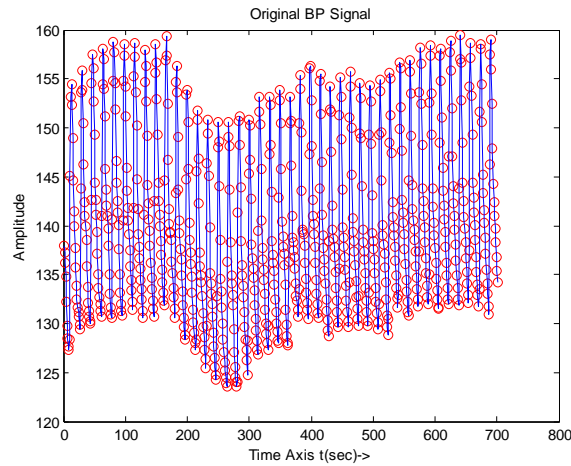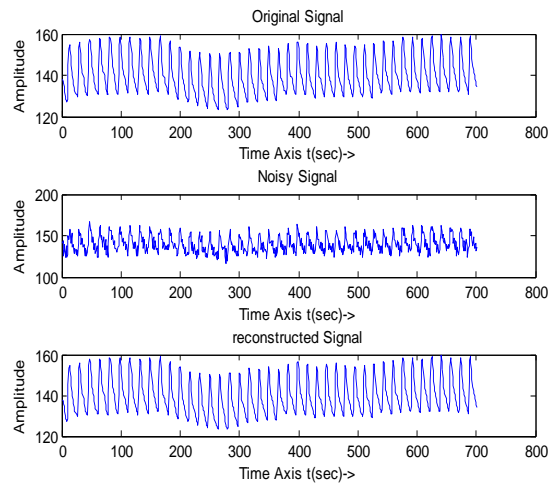
Figure: 5 Typical waveform of BP

Figure: 6 Noisy BP signal reconstructed via EEMD method with I=500

VI. Emg Signal

Electromyography (**EMG**) is a technique for evaluating and recording the electrical activity produced by skeletal muscles. EMG is performed using an instrument called an **electromyograph**, to produce a record called an **electromyogram**. An electromyograph detects the electrical potential generated by muscle cells when these cells are electrically or neurologically activated. The signals can be analysed to detect medical abnormalities, activation level and recruitment order or to analyse the biomechanics of human or animal movement.
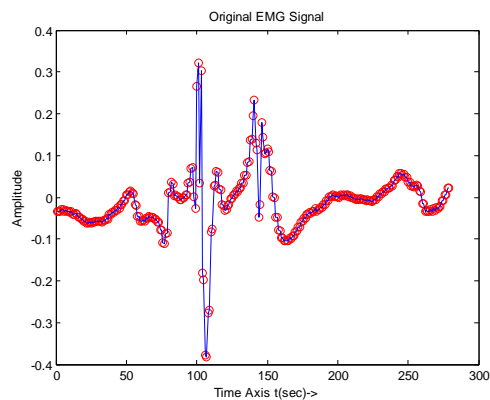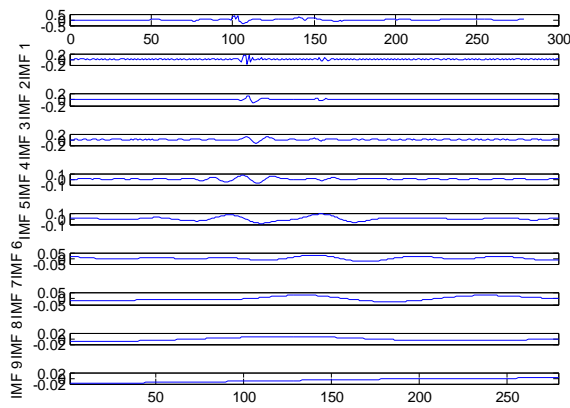

Figure: 7 Typical waveform of EMG

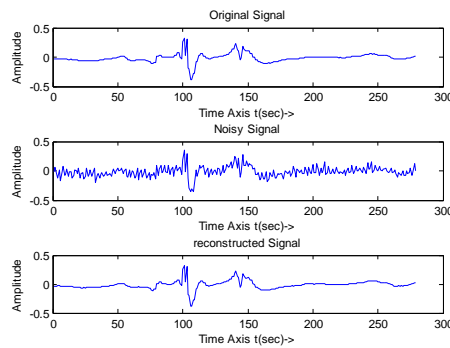Figure: 8 Set of decomposed IMFs of EMG signal via EEMD method

Figure: 9 Noisy EMG signal reconstructed via EEMD method with I=500

VII. Neurodegenerative Disease Signal

Neurodegeneration is the umbrella term for the progressive loss of structure or function of neurons, including death of neurons. Many neurodegenerative diseases including Parkinson's, Alzheimer's, and Huntington's occur as a result of neurodegenerative processes.
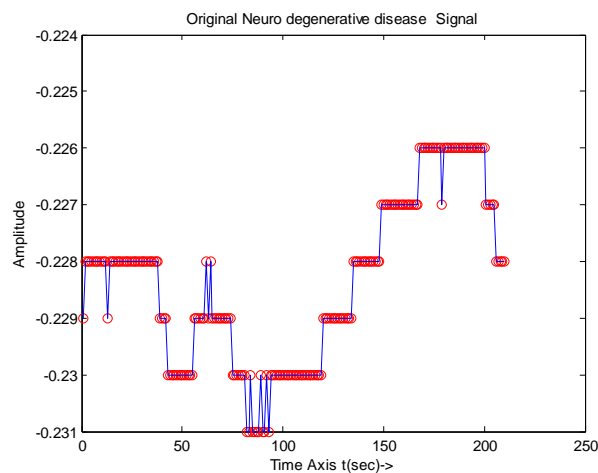

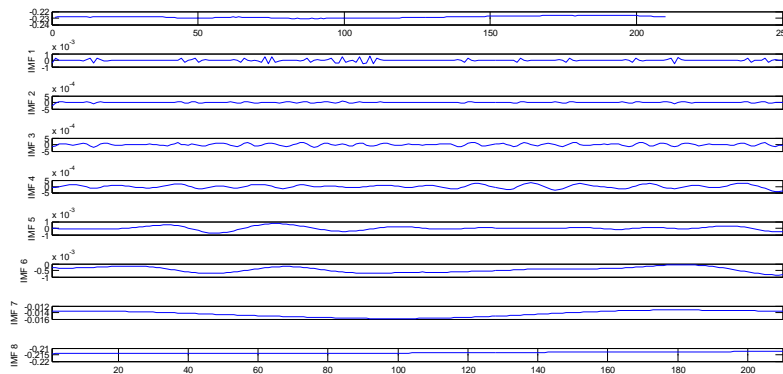Figure: 10 Typical waveform of Neuron-degenerated signal

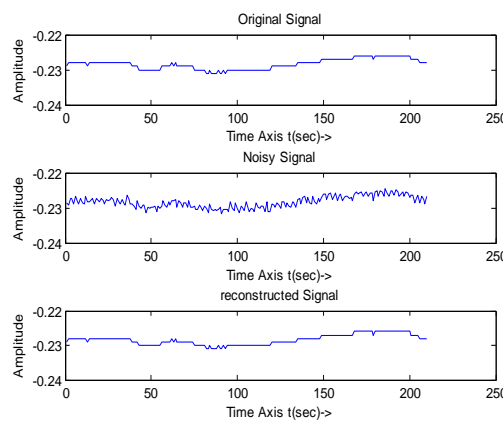Figure: 11 Set of decomposed IMFs of Neuron-degenerated signal via EEMD method

Figure: 12 Noisy Neuron-degenerated signal reconstructed via EEMD method with I=500

VIII. Mean Square Error

In statistics, the **mean squared error (MSE)** of an estimator is one of many ways to quantify the difference between values implied by an estimator and the true values of the quantity being estimated. MSE is a risk function, corresponding to the expected value of the squared error loss or quadratic loss. MSE measures the average of the squares of the "errors." The error is the amount by which the value implied by the estimator differs from the quantity to be estimated.

In each case the mean square error has been calculated which is defined as:

$$MSE(x, \check{x}) := \frac{1}{N} \sum_{1=1}^{N} [x(t_i) - \check{x}(t_i)]^2$$

where x=Original signal without noise
$\check{x}$ =Reconstructed signal
N=Length of the signal

Table I

SIGNAL	MSE
ECG	1.6530 mv
BP	.50896mv
EMG	.0012 mv

IX. Conclusion

Thus above experiments show that EEMD which is a NADA method successfully remove the mode mixing problem which is a major drawback of EMD. From the Matlab simulation it has been shown that EEMD is quite effective to reduce noise from ECG and many other biomedical signals with a very small mean square error.

The major drawback of the algorithm comes in term of computational cost. Since the EMD does not have a mathematical expression researchers are working to derive a theoretical definition of EMD, which if achieved will let the EMD as popular Fourier transform and Wavelet Transform.

References

Journal Papers:

[1] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H.Shih, Q. Zheng, N.C. Yen, C. C. Tung, and H. H. Liu. 'The Empirical Mode Decomposition and Hilbert Spectrum for Non-linear and Non stationary Time Series Analysis. Proceedings of the Royal Society London A., 1998 454,903–995. [URL:http://rspa.royalsocietypublishing.org/content/454/1971/903.full.pdf+html]

[2] J. Beran, Statistics for Long-Memory Processes, Chapman & Hall, 1994. P. Flandrin G. Rilling and P. Goncalves, "Empirical Mode Decomposition as a filter bank," IEEE Sig. Proc. Lett.Vol.11, no. 2, pp. 112–114, 2004.

[3] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," Advances in Adaptive Data Analysis, vol. 1, no. 1, pp. 1–41, 2009.

[4] Friesen, G.M, Jannett, T.C. Jannett, M.A. Jadallah, S.L. Yates, S.R. Quint, H.T. Nagle 'A comparison of the noise sensitivity of nine QRS detection algorithms. IEEE Trans. Biomed. Eng. 1990, pp. 85–98, vol 37.No.1

[5] K. M. Chang, "Arrhythmia ECG noise reduction by ensemble empirical mode decomposition," Sensors, vol. 10, no. 6, pp. 6063–6080, 2010.

[6] Marıa E. Torres, Marcelo A. Colominas, Gaston Schlotthauer, Patrick Flandrin "A complete ensemble empirical mode decomposition with adaptive noise" 978-14577-0539-7/11/2011 IEEE.