Application of Empirical Mode Decomposition in Denoising a Speech Signal

Sonam Maheshwari[1], Dimpy[2]

[1]Assistant Professor. Jagannath University

[2]Assistant Professor, SGI SAMALAKHA, Kurukshetra University

Abstract: ***This paper attempts to review and summarize the use of Empirical mode decomposition (EMD) for denoising a speech signal. EMD, introduced by Huang et al in 1998 gives time-frequency representation of non-linear and non-stationary signals. It decomposes a signal into a sum of finite zero-mean, oscillating components called as intrinsic mode functions based on the local time characteristics of the signal. Main essence of the method is it's adaptive and data driven nature.***

Keywords: Emirical mode decomposition,IMF, Wavelet transform, Hard and soft thresholding,Denoising

1. Introduction

Background noise is a severe problem in communication and other speech related systems. Speech signals often get distorted by the noise during transmission process which considerably reduces the quality and intelligibility of speech signals. Speech signal denoising techniques aim to reduce the noise level to an extent so that speech signal becomes more clear and intelligible. Literature survey reveals a number of techniques to denoise a signal while maintaining its intelligibility and quality. Spectral Subtraction[1] is one of the earliest speech enhancement method based on the idea of estimating the clear signal spectrum by subtracting the noisy spectrum from the corrupted signal spectrum. Noise spectrum is obtained during the pause periods. Main drawback of the method is the presence of residual noise in the estimated signal which is generally more annoying. In signal subspace approach[2], noisy signal is first decomposed into a signal plus noise and an orthogonal noise subspace, then noise subspace is removed to get an estimate of the clean signal. Separating noise from the distorted signal by employing Weiner filter is effective only when the spectrum of noise is known in advance and signal is stationary[3].Wavelet Transform is a popular technique in signal processing and analysis[4].

Properties like MRA and variable size window for different frequency components makes Wavelet more suited for non-linear and non-stationary signals. Requirement of a pre-defined basis functions prior to signal decomposition is a major drawback of wavelet transform. Empirical Mode Decomposition is a relatively new method [5] for signal analysis which decomposes the signal into a finite number of components named as Intrinsic Mode Functions(IMFs) based on the local characteristics of the signal. Biggest advantage of the method is its adaptive nature, ie, it requires no predefined basis functions prior to signal decomposition. Basis functions are derived from the data itself. Since, speech signal is non-linear and transient in nature. Hence EMD based speech enhancement methods can pave the way to new dimension in the field of speech processing [6],[7],[8],[9],[10]. In this paper, potentiality of EMD combined with wavelet thresholding has been studied and investigated. Paper is organized in the following sections: First Section covers the fundamental concepts of EMD. Second section reviews the principle of wavelet thresholding and its extension to EMD. Next section presents the denoising method and contains results and discussions. Last section concludes the paper.

2. Empirical Mode Decomposition

Empirical mode decomposition (EMD) decomposes a signal x(t) into a finite number of Intrinsic Mode Functions (IMFs), $h_{(i)}(t)$, $1\leq i\leq L$,

$$x(t)=\sum_{i=1}^{L} h_i(t)+r(t) \tag{1}$$

where r(t) is a remainder which is a non zero-mean slowly varying function with only few extrema. Decomposition is based on the characteristics of the signal itself.

IMFs are zero-mean oscillating signals satisfying the following conditions:

1) The number of extrema and the number of zero crossings must either be equal or differ at most by one,
2) At any point, the mean value of the envelope defined by local maxima and the envelope defined by the local minima is zero.

Steps for finding the IMFs of a signal are as follows-

1) Identify local maxima and minima of x(t).
2) Form the upper and lower envelope u(t) and l(t) by cubic spline interpolation of the extrema points.
3) Calculate the mean of the upper and lower envelop, $m_1(t)$ using $m_1(t) = u(t) + l(t) / 2$.
4) Subtract mean from the signal x(t) to obtain $d_1(t)$. If $d_1(t)$ is a zero-mean function, then the iteration stops and $d_1(t)$ is accepted as first IMF,ie $h_1(t) = d_1(t)$.
5) If not, use $d_1(t)$ as the new data and repeat steps 1-4 until an IMF is obtained.
6) Once the first IMF $h_1(t)$ is obtained, residual signal is defined as

$$r_1(t) = x(t)- h_1(t) \tag{2}$$

Residual signal contains information about the lower frequency components and is taken as the input signal to obtain next IMFs. At the end, a monotonic function with only few extrema is obtained from which no further decomposition can be done.

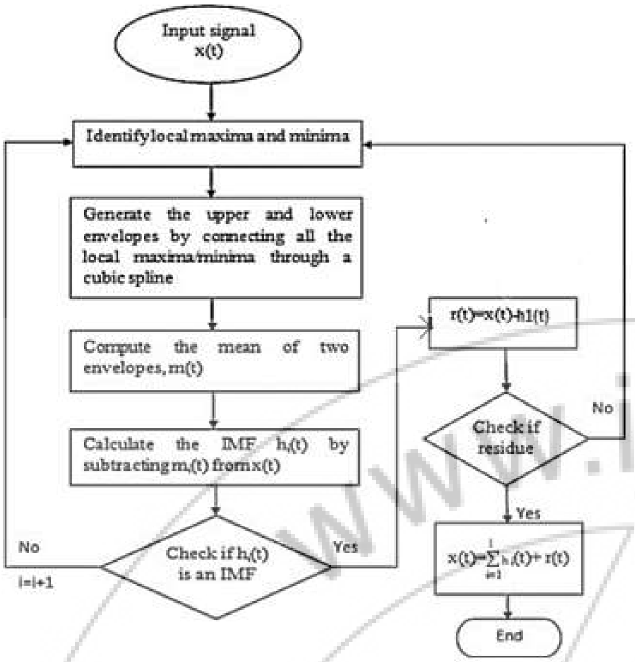

Figure 1: Flowchart of the EMD Process

Figure 2 depicts a signal and its IMFs. It can be observed that lower order IMFs contain higher frequency components and vice versa. Thus overall result of EMD can be seen as to successively remove highest frequencies from a signal[11],[12].

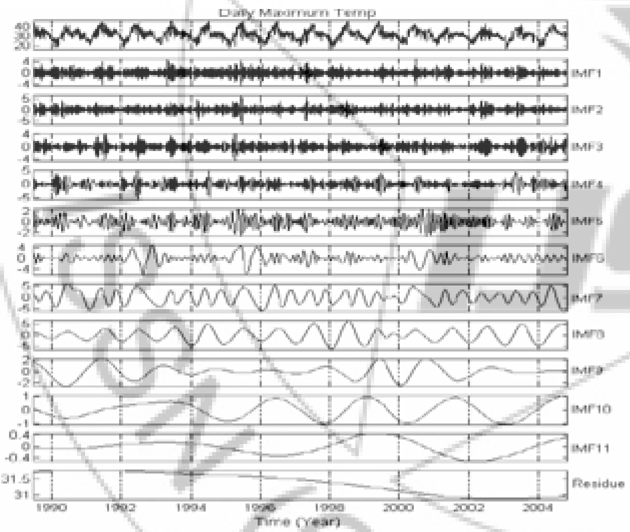

Figure 2: Signal and its IMFs

3. Thresholding

Thresholding method is of two types - hard thresholding and soft thresholding. In the hard thresholding, coefficients below a give value are stetted to zero, while in soft thresholding the coefficients are reduced to the threshold value. Mathematically,

$$\text{For hard thresholding; } w_{ht} = \begin{cases} -|w|, w \geq t \\ |0|, w < t \end{cases} \quad (3)$$

$$\text{Soft thresholding; } w_{st} = \begin{cases} [\mathrm{sgn}(w)](w - t), |w| \geq t \\ 0, |w| < t \end{cases} \quad (4)$$

where w is a coefficient; t is a value of threshold which is applied on the coefficients. To determine the value of threshold, different thresholding rules can be applied. i) Sqtwolog (wtq)- This is a fixed threshold or global thresholding method and it is computed as:

$$\mathrm{wtq} = \sigma\sqrt{2\log(n)/n} \quad (5)$$

where n is the total number of wavelet coefficients. ii) Rigrsure (wtsu)- It is an adaptive thresholding method based on Stein's unbiased likelihood estimation principle. iii) Heursure (wth)- This is the hybrid of SureShrink and Universal threshold and is used when wavelet representation at any level is very sparse. iv) Minimax (wtm) Minimax threshold uses a fixed threshold to obtain a minimum error between wavelet coefficient of noise signal and original signal.

4. Method

The property of EMD to behave as a dydadic filter bank has been exploited in denoising a signal[11],[12] EMD has been found to be a powerful tool for removing noise from signals. It allows to analyze the noise separately at each scale and to adapt the denoising algorithm accordingly. For efficient noise reduction EMD is combined with soft thresholding.

Steps are as follows:
1) The noisy signal is applied to EMD algorithm which decomposes it into a finite sum of IMFs.
2) Obtained IMFs are soft-thresholded to remove the IMFs that contain noise.
3) Thresholded IMFs are added to obtain the denoised signal.

5. Results

To confirm the effectiveness of the method, a male speech is recorded and taken as the original signal. White Gaussian noise is used to model the background noise. Noisy speech signal is applied to the EMD algorithm taken from R-Package. The obtained IMFs are thresholded using the principle of soft thresholding to recover an estimate of the original signal. Dohono's universal thresholding rule is used. All the simulations are done in MATLAB environment (R.2010a).

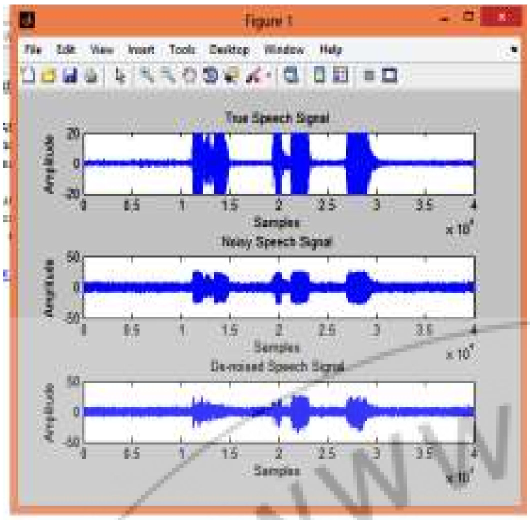

Figure 3: Shows the true, noisy and denoised signals

6. Conclusion and Future Scope

Based on the results, it can be shown that EMD combined with thresholding proves to be a promising technique in the field of speech signal analysis and processing. Decomposition based on the characteristics of the signal itself makes it more suited for the speech signals which are highly non-linear and transient in nature. Main drawback of the EMD-based method is the computational time and lack of mathematical base which needs to be addressed. With the applications and results presented in the paper, it is believed that the presented framework will benefit many applications in speech signal processing, as well as providing a better understanding of the EMD algorithm.

References

[1] S.F.Boll, "Suppression of Acoustic Noise in Speech using Spectral Subtraction", IEEE Transactions on ASSP, Vol. 27, No. 2, pp. 113-120, 1979.

[2] Y. Ephraim, H.L.V.Trees, "A Signal Subspace Approach for Speech Enhancement", IEEE Transactions on Speech and Audio Processing, Vol.3, No. 4, pp. 251-266, 1995.

[3] J.G. Proakis, D.G. Manolakis, "Digital Signal Processing: Principles, Algorithms, and Applications", Prentice-Hall, 3rd Edition, 1996.

[4] S.Mallat, "A wavelet tour of signal processing", Academic press, second edition, 1999.

[5] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, Q. Shin, H.H. Zheng, N.C. Yen, C.C. Tung, and H.H. Liu, "The Empirical Mode Decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis", Proc. Royal Soc. London A, vol. 454, pp. 903-995, 1998.

[6] P. Flandrin, P. Gonçalvès, and G. Rilling, "Detrending and denoising with Empirical Mode Decompositions", Eusipco, 12th European Signal Processing Conference, September 6-10 Vienna, Austria, 2004.

[7] A. O. Boudraa, and J. C. Cexus, "Denoising via empirical mode decomposition", in Proc. ISCCSP 2006.

[8] Boudraa, J.C.Cexus, "EMD based signal filtering", IEEE Trans. On Instrumentation and measurement, Vol.56, No.6, pp.2196-2202.

[9] K. Khaldi, M. Turki-Hadj Alouane, A.-O. Boudraa, "A new EMD denoising approach dedicated to voiced speech signals"", Signals, Circuits and Systems, 2008. SCS 2008. 2nd International Conference on Volume, Issue 7-9 Page(s): 1 – 5, Nov. 2008.

[10] Hamid et al, " Empirical Mode Decomposition for advanced speech signal processing", Journal of Signal Processing, vol.17, no.6, pp.215-229, Nov 2013.

[11] P.Flandrin, G. Rilling, P. Goncalves, "Empirical mode decomposition as a filter bank", IEEE Signal Processing Letters,Vol. 11, pp. 112– 114, 2004.

[12] P.Flandrin, G. Rilling, P. Goncalves, "EMD equivalent filter banks, from interpetation to applications", World Scientific Publishing Company, first edition, 2005.