World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering
Vol:1, No:2, 2007

EMD-Based Signal Noise Reduction

A.O. Boudraa, J.C. Cexus, and Z. Saidi

***Abstract*—**This paper introduces a new signal denoising based on the Empirical mode decomposition (EMD) framework. The method is a fully data driven approach. Noisy signal is decomposed adaptively into oscillatory components called Intrinsic mode functions (IMFs) by means of a process called *sifting*. The EMD denoising involves filtering or thresholding each IMF and reconstructs the estimated signal using the processed IMFs. The EMD can be combined with a filtering approach or with nonlinear transformation. In this work the Savitzky-Golay filter and shoft-thresholding are investigated. For thresholding, IMF samples are shrinked or scaled below a threshold value. The standard deviation of the noise is estimated for every IMF. The threshold is derived for the Gaussian white noise. The method is tested on simulated and real data and compared with averaging, median and wavelet approaches.

***Keywords*—**Empirical mode decomposition, Signal denoising nonstationary process.

I. INTRODUCTION

Estimating a signal of interest degraded by additive random noise is a classical problem in signal processing. In many applications, signal denoising is used to produce estimates of the original signal from noisy observations. The recovered signal should be as close as possible to the original one while retaining most of its important properties (e.g. smoothness). Traditional denoising schemes are based on linear methods, where the most common choice is the Wiener filtering [1]. Linear methods are frequently used because they are easy to implement and design. However, linear filtering methods are not so effective when signals contain sharp edges and impulses of short duration. Furthermore, linear methods are not so effective when transient nonstationary wide-band components are involved since they have similar spectrum to the noise. In order to overcome these shortcomings nonlinear methods have been proposed and especially those based on wavelets thresholding [2]-[3]. The wavelet schemes rely on the basic idea that the energy of a signal will often be concentrated in a few coefficients in wavelet domain while the energy of noise is spread among all coefficients in wavelet domain. Donoho and Johnstone [2] proposed hard and soft thresholding methods for denoising, where the former leaves the magnitudes of coefficients unchanged if they are larger than a given threshold, while the latter just shrinks them to zero by the threshold value. A main drawback of the wavelet approach is that the basis functions are fixed, and do not necessarily match varying nature of signals.

Recently, Huang *et al*. [4] have introduced the Empirical mode decomposition (EMD) method for analyzing data from nonstationary and nonlinear processes. The major advantage of the EMD is that the basis functions are derived from the signal itself. Hence, the analysis is adaptive, in contrast to the wavelet method where the basis functions are fixed. In this paper, a denoising method based on the EMD approach is proposed. The EMD is based on the sequential extraction of energy associated with various intrinsic time scales of the signal starting from finer temporal scales (high frequency modes) to coarser ones (low frequency modes). The total sum of the IMFs matches the signal very well and therefore ensures completeness. The proposed method relies on the basic idea, as in wavelet analysis, that the first IMFs (finest modes) are dominated by noise than the last ones (corset modes). Thus, the recovered signal can be reconstructed using, filtered or thresholded IMFs.

II. EMD ALGORITHM

The EMD involves the decomposition of a given signal $x(t)$ into a series of IMFs, through the *sifting* process, each with distinct time scale [4]. The major advantage of the EMD is that the basis functions are derived from $x(t)$ itself. The decomposition is based on the local time scale of the signal and yields adaptive basis functions. The EMD can be seen as a type wavelet decomposition whose subbands are built up as needed to separate the different components of $x(t)$. Each IMF replaces then the detail signals of $x(t)$ at a certain scale or frequency band [5]. The EMD picks out the highest frequency oscillation that remains in $x(t)$. An IMF must fulfil two requirements: (R1) the number of extrema and the number of zero crossings are either equal or differ at most by one; (R2) at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero. Thus, locally, each IMF contains lower frequency oscillations than the one just extracted before. The EMD does not use any pre-determined filter or wavelet function and it is fully data driven method [4]. To be successfully decomposed in IMFs, $x(t)$ must has at least two extrema, one minimum and one maximum. The *sifting* process involves the following steps:

A.O. Boudraa, J.C. Cexus, and Z. Saidi are with IRENav, Ecole navale, Lanvéoc-Poulmic, BP 600, 29240, Brest-Armées, France (corresponding author: A.O. Boudraa, e-mail: boudra@ecole-navale.fr).

Step 1) Fix $\varepsilon, j \leftarrow 1$ $\left(j^{th}\text{IMF}\right)$

Step 2) $r_{j-1}(t) \leftarrow x(t)$ (residual)

Step 3) Extract the j^{th} IMF

(a) $h_{j,i-1}(t) \leftarrow r_{j-1}(t)$, $i \leftarrow 1$ (i number of sifts)

(b) Extract local maxima/minima of $h_{j,i-1}(t)$

(c) Compute upper and lower envelopes $U_{j,i-1}(t)$ and $L_{j,i-1}(t)$ by interpolating, using cubic spline, respectively local maxima and minima of $h_{j,i-1}(t)$

(d) Compute the envelopes mean: $\mu_{j,i-1}(t) \leftarrow \left(U_{j,i-1}(t)+L_{j,i-1}(t)\right)/2$

(e) Update: $h_{j,i}(t) \leftarrow h_{j,i-1}(t) - \mu_{j,i-1}(t)$, $i \leftarrow i+1$

(f) Calculate stopping criterion:

$$SD(i)=\sum_{t=0}^{T}\frac{\left|h_{j,i-1}(t)-h_{j,i}(t)\right|^2}{\left(h_{j,i-1}(t)\right)^2}$$

(g) Repeat Step (b)-(f) until $SD(i) < \varepsilon$ and then put $IM_j(t) \leftarrow h_{j,i}(t)$ $\left(j^{th}\ \text{IMF}\right)$

Step 4) Update residual: $r_j(t) \leftarrow r_{j-1}(t) - IM_j(t)$

Step 5) Repeat Step 3 with $j \leftarrow j+1$ until the number of extrema in $r_j(t)$ is < 2

Where "$\leftarrow$" is the affectation operator and T is the time duration. The *sifting* is repeated several times (i) in order to get h to be a true IMF that fulfils the requirements (R1) and (R2). The result of the *sifting* procedure is that $x(t)$ will be decomposed into IMFs, $IM_j(t)$, $j=1,\ldots,N$ and residual $r_N(t)$.

$$x(t)=\sum_{j=1}^{N} IM_j(t)+r_N(t)$$

The *sifting* process has two effects: (a) to eliminate riding waves and (b) to smooth uneven amplitudes. To guarantee that the IMF components retain enough physical sense of both amplitude and frequency modulations, we have to determine a criterion for the *sifting* process to stop. This accomplished by limiting the size of the standard deviation SD, set to ε, computed from the two consecutive *sifting* results. Usually, SD is set between 0.2 to 0.3 [4]. Note that the number of IMFs, N, is determined automatically during the *sifting* process.

III. EMD DENOISING

Let $C_j(t)$ be a clean deterministic IMF with the finite length L and IM_j the corrupted IMF with additive noise $b_j(t)$ with variance $\sigma_j^2(t)$:

$$IM_j(t)=C_j(t)+b_j(t)$$

Let $\hat{C}_j(t)$ be an estimation of $C_j(t)$ based on the noisy observation $IM_j(t)$. The estimation $\hat{C}_j(t)$ is given by

$$\hat{C}_j(t)=\Gamma\left(IM_j,\tau_j\right)$$

where $\Gamma\left(IM_j,\tau_j\right)$ denotes a thresholding function or filtering method, defined by parameters τ_j, applied to signal IM_j. The denoised signal $\hat{x}(t)$ is given by

$$\hat{x}(t)=\sum_{j=1}^{N}\hat{C}_j(t)+r_N(t)$$

For noise reduction, the EMD can be combined with a filtering method such as Savitzky-Golay smoothing [6] or nonlinear transformation such as the soft-thresholding [3].

A. EMD-Soft

A smooth version of the input data can be obtained by thresholding the IMFs before signal reconstruction. If $\Gamma\left(.,\tau_j\right)$ is a thresholding function, then τ_j is the threshold value and this can be determined in many ways [7]. Donoho and Johnstone [8] proposed an universal threshold for removing added Gaussian noise τ_j given by

$$\tau_j=\hat{\sigma}_j\sqrt{2\log(L)}$$

$$\hat{\sigma}_j = MAD_j/0.6745$$

where σ_j is the noise level of the j^{th} IMF. MAD_j represents the absolute median deviation of the j^{th} IMF and is defined by

$$MAD_j = \text{Median}\left\{\left|IM_j(t)\text{-Median}\left\{IM_j(t')\right\}\right|\right\}$$

Instead of using a global thresholding, level-dependent thresholding uses a set of thresholds, one for each IMF (scale level). The soft-thresholding method shrinks the IMF samples by τ_j towards zero as follows [3]

$$\hat{C}_j(t) = \begin{cases} IM_j(t)-\tau_j & \text{if } IM_j(t)\geq\tau_j \\ 0 & \text{if } \left|IM_j(t)\right|<\tau_j \\ IM_j(t)+\tau_j & \text{if } IM_j(t)\leq-\tau_j \end{cases}$$

B. EMD-SG

The Savitzky-Golay (SG) filter method is time-domain smoothing [6]. This method was originally designed to preserve higher moments within time-domain spectral peak data. The SG filter can be considered as piece-by-piece fitting of polynomial function to the signal. The smoothed points are computed by replacing each data point with the value of its fitted polynomial. The fitting is done by least squares approach.

IV. RESULTS

In order to test the EMD denoising method, we have performed numerical simulations for four test signals: "Doppler", "Blocks", "Bumps" and "Heavysine" obtained using WAVELAB Software*. The method is also tested on one real biomedical signal: "ECG". The signals size is L=2048. MAE and SNR are calculated as the measures of efficiency of noise reduction. For synthesized signals the variance of the white Gaussian noise is set so that the original SNR (before denoising) is maintained at 2 dB. The SNR of the "ECG" is -9 dB. The original signals and noise free ones are depicted in Figures 1 and 2, respectively. Figure 3 shows a sequential extraction of local oscillations by the EMD of the signal "Blocks" (Fig. 1(a)). Tables I and II show comparisons of MAE and SNR values for averaging, median, wavelet, EMD-Soft and EMD-SG methods. The EMD decomposed the "Blocks" signal into 5 oscillating modes (IMFs) and a residual. One can remark that the first IMF corresponds to fast oscillation while the fifth IMF corresponds to slow one (Fig. 3). A comparison of the signal (top diagram) and the residue (bottom diagram) of Figure 3 shows that the residue captures the trend of the signal. Each noisy signal is decompose using the EMD and the derived IMFs are filtered (thresholded) using a SG filter of third-order (soft-thresholding). The standard deviation, ε, is set to 0.3 [4]. Thus, the N value is determined automatically based on the ε value. Figure 4 displays the outcome of applying the EMD denoising scheme to the five signals. Each reconstructed signal plot (dot line) is superposed on the corresponding free noise signal (solid line). Globally, the results are qualitatively appealing; the reconstructions jump where the signal jumps and are smooth where the true signal is smooth. The significant results are obtained for “Blocks”, “Heavysine”, “Bumps” and “ECG” (Figs. 4(b)-(e)) which are very close to their corresponding original signals. These findings are confirmed by the SNRs values listed in Table I where significant improvements in SNR range from 9.73 dB to 26.79 dB. As indicated in Tables I and II, both the EMD-SG and the EMD-Soft outperform the averaging and median methods. For “Heavysine” and “ECG” signals both the EMD-SG and the EMD-Soft perform better than the wavelet method. For “Bumps” signal both the EMD-SG and the wavelet method give the same SNR. However, the wavelet method (14.97 dB) performs better than the EMD-SG (13.57 dB) for “Doppler” signal. The efficiency of the compared methods depends on the signal behaviour. In particular, for the “ECG” signal the averaging method achieves better SNR than the wavelet method. The oscillations seen in flat regions (Figs. 4(b)-(c)) may be due to the interpolation scheme used in the *sifting* process and thus it would be interesting to search for other interpolation methods other than cubic splines. A careful examination of the “Doppler” signal (Fig. 4(a)) shows that the beginning of this signal, (oscillations of rapid change), is not well reconstructed. This may be due to the rate sampling used. The same problem is seen in the wavelet reconstruction.

* Available from Stanford Statistics Department, courtesy of D.L. Donoho and I.M. Johnstone.

Table I
Denoising results in SNR of test signals corrupted by Gaussian noise.

	Doppler	Blocks	Bumps	Heavysine	ECG
	SNR	SNR	SNR	SNR	SNR
Noisy	2.06	2.03	2.03	2.03	-9.02
Averaging	9.86	9.06	9.46	12.66	7.23
Median	10.57	10.17	10.55	10.67	4.62
Wavelet	**14.97**	**11.94**	14.47	18.76	5.82
EMD-Soft	11.13	11.18	11.18	19.86	14.39
EMD-SG	13.57	11.76	**14.50**	**20.60**	**17.77**

Table II
Denoising results in MAE of test signals corrupted by Gaussian noise.

	Doppler	Blocks	Bumps	Heavysine	ECG
	MAE	MAE	MAE	MAE	MAE
Noisy	0.81	0.81	0.81	0.81	0.46
Averaging	0.32	0.29	0.32	0.24	0.04
Median	0.30	0.31	0.31	0.30	0.09
Wavelet	**0.16**	**0.20**	0.19	0.12	0.04
EMD-Soft	0.21	0.25	0.24	0.10	0.03
EMD-SG	0.18	0.22	**0.18**	**0.09**	**0.01**

V. CONCLUSION

In this paper a new signal denoising approach is introduced. This denoising scheme, based on EMD method, is simple and fully data-driven. The method does not use any pre- or post-processing. The EMD-Soft is fully automated denoising method. To run the EMD-SG only the analysis window size of the SG filter is required. Results obtained for synthetic signals and for one real signal indicate that both EMD-Soft and EMD-SG are effective for noise removal. The EMD-Soft and the EMD-SG methods outperform averaging and median methods and for three signals EMD-SG performs better than the wavelet method. To confirm the effectiveness of the EMD denoising, the method must be evaluated with a large class of signals and in different experimental conditions such as noise levels, sampling rates and sample sizes. For better signal reconstruction, we plan to study how to adapt the SG filter order to each IMF.

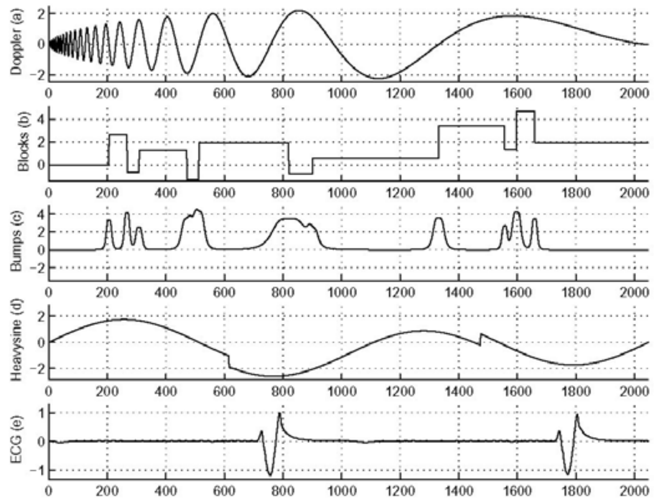

Figure 1. Test signals with L=2048.

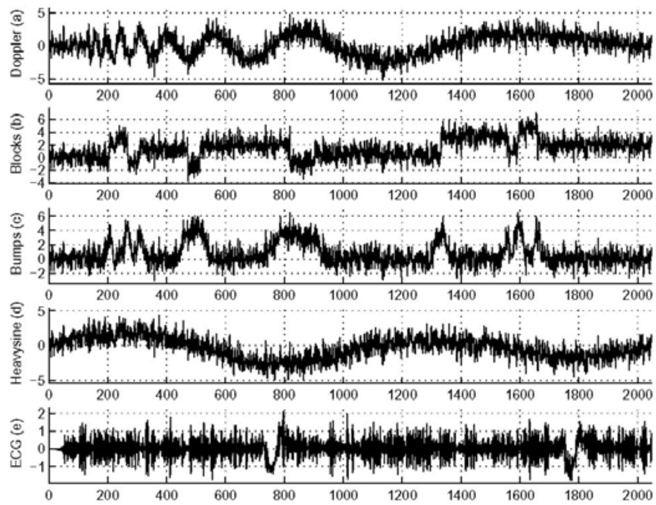

Figure 2. Noisy test signals (SNR=2dB; SNR=-9dB for ECG).

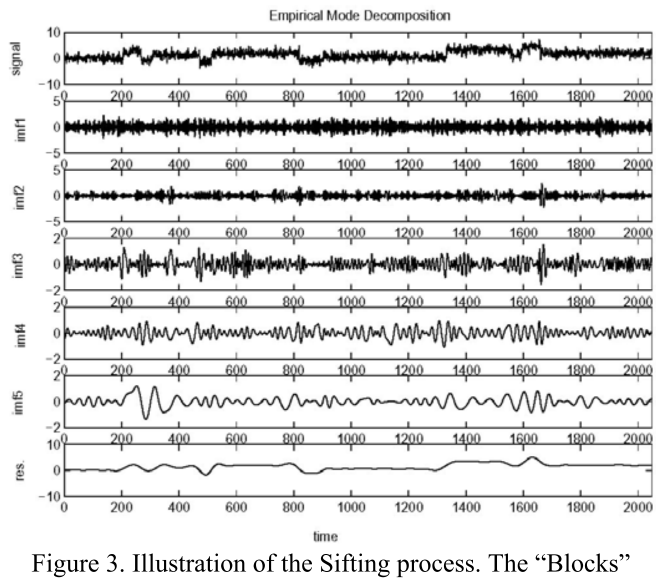

Figure 3. Illustration of the Sifting process. The “Blocks” signal is decomposed into ten IMFs and a trend. (residue).

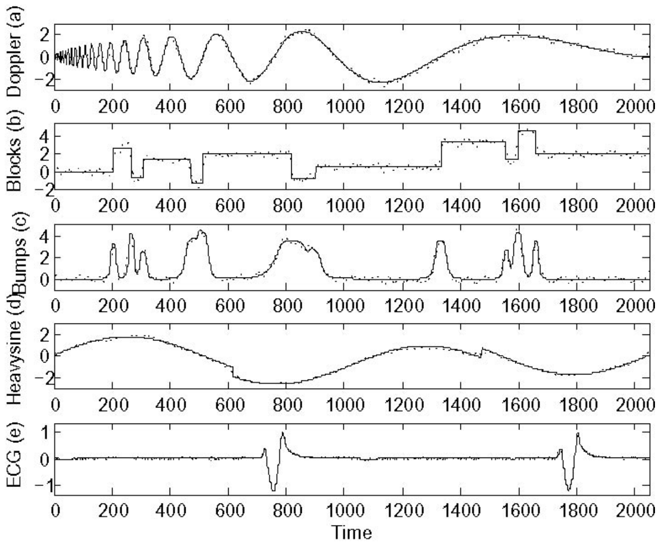

Figure 4.Results of the EMD-SG denoising. The free noise signals (solid line). The reconstructed signals (dot line).

REFERENCES

[1] J.G. Proakis, and D.G. Manolakis, Digital Signal Processing: Principles, Algorithms, and Applications (3rd edition), Prentice-Hall, 1996.

[2] D.L. Donoho, and I.M. Johnstone, "Ideal spatial adaption via wavelet via wavelet shrinkage," *Biometrica*, vol. 81, pp. 425-455, 1994.

[3] D.L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inform. Theory*, vol. 41, no. 3, pp. 613-627, 1995.

[4] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shin, Q. 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.

[5] P. Flandrin, G. Rilling, and P. Gonçalves, "Empirical Mode Decomposition as a Filter Bank," *IEEE Sig. Proc. Lett.*, vol. 11, no. 2, pp. 112-114, 2004.

[6] A. Savitzky, and M.J.E. Golay, "Smoothing and differentiation, of data by simplified least squares procedures," *Analytical chemistry*, vol. 36, pp. 1627-1639, 1964.

[7] A. Antoniadis, J. Bigot, and T. Sapatinas, "Wavelet estimators in nonparametric regression: a comparative simulation," study, *J. Statist. Software*, vol. 6, no. 6, pp. 1-83, 2001.

[8] D.L. Donoho, and I.M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage, " *Biometrika*, vol. 81, no. 3, pp. 425-455, 1994.