DENOISING VIA EMPIRICAL MODE DECOMPOSITION

A.O. BOUDRAA^{*+} *and J.C. CEXUS*$^{*+}$

*IRENav, Ecole Navale/$^+$E^{3}I^2 (EA 3876) ENSIETA
Groupe ASM, Lanvéoc Poulmic, BP600, 29240 Brest−Armées, France.

ABSTRACT

In this paper a signal denoising scheme based a multiresolution approach referred to as Empirical mode decomposition (EMD) [1] is presented. The denoising method is a fully data driven approach. Noisy signal is decomposed adaptively into intrinsic oscillatory components called Intrinsic mode functions (IMFs) using a decomposition algorithm algorithm called *sifting* process. The basic principle of the method is to reconstruct the signal with IMFs previously filtered or thresholded. The denoising method is applied to one real signal et to four simulated signals with different noise levels and the results compared to Wavelets, Averaging and Median methods. The effect of level noise value on the performances of the proposed denoising is analyzed. The study is limited to signals corrupted by additive white Gaussian random noise.

1. INTRODUCTION

The recovery of a signal from observed noisy data is a classical problem in signal processing. Especially for the case of additive white Gaussian noise a number of filtering methods have been proposed [2]-[4]. Linear methods such as the Wiener filtering [2] are largely used because linear filters are easy to implement and design. However, these methods are not so effective when signals contain sharp edges and impulses of short duration. Furthermore, real signals are often nonstationary. To overcome these difficulties nonlinear methods have been proposed and especially those based on Wavelets thresholding [3]-[4]. The idea of thresholding relies on the assumption that signal magnitudes dominate the magnitudes of the noise in a Wavelets representation, so that Wavelets coefficients can be set to zero if their magnitudes are less than a pre-determined threshold [4]. A limit of the Wavelets approach is that the basis functions are fixed, and thus do not necessarily match all real signals. To avoid this problem time-frequency atomic signal decomposition can be used [5]-[6]. As for Wavelets packets, if the dictionary is very large and rich with a collection of atomic waveforms which are located on a much finer grid in time-frequency space than Wavelets and cosine packet tables, then it should be possible to represent a large class of real signals (for denoising, compression,...). But, in spite of this, the basis functions must be specified (Gabor functions,. . .).
Recently, a new signal decomposition method called Empirical mode decomposition (EMD) has been introduced by Huang et *al*. [1] 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 traditional methods where the basis functions are fixed. In this paper, a denoising scheme using EMD 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 the basic idea of the proposed scheme is to preprocessed each IMF using thresholding, as in Wavelets analysis, of filtering before complete signal reconstruction.

2. EMD ALGORITHM

The EMD involves the adaptive decomposition of given signal, $x(t)$, into a series of oscillating components, IMFs, by means of a decomposition process called *sifting* algorithm. The name IMF is adapted because it represents the oscillation mode embedded in the data. With this definition, the IMF in each cycle, defined by the zero crossings of, involves only one mode of oscillation, no complex riding waves are allowed. The essence of the EMD is to identify the IMF by characteristic time scales, which can be defined locally by the time lapse between two extrema of an oscillatory mode or by the time lapse between two zero crossings of such mode. The EMD picks out the highest frequency oscillation that remains in the signal. Thus, locally, each IMF contains lower frequency oscillations than the one extracted just before. Furthermore, the EMD does not use any pre-determined filter or Wavelet function. It is fully data driven method. Since the decomposition of the EMD is based on the local characteristics time scale of the data, it is applicable to nonlinear and non-stationary processes. The EMD decomposes into a sum of IMFs that: (1) have the same numbers of zero crossings and extrema; and (2) are symmetric with respect to the local mean. The first condition is similar to the narrow-band requirement for a stationary Gaussian process. The second condition modifies a global requirement to a local one, and is necessary to ensure that the IF will not have unwanted fluctuations as induced by a symmetric waveforms. The *sifting* process is defined by the following steps:

Step 1) Fix ϵ, $j \leftarrow 1$ (j^{th} IMF)
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 envelope and lower envelope functions $U_{j,i-1}(t)$ and $L_{j,i-1}(t)$ by interpolating respectively local maxima and minima of $h_{j,i-1}(t)$
(d) Compute the envelopes mean: $\mu_{j,i-1}(t) \leftarrow (U_{j,i-1}(t) + L_{j,i-1}(t))/2$
(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:
$$SD(i) = \sum_{t=0}^{T} \frac{| h_{j,i-1}(t) - h_{j,i}(t) |^2}{(h_{j,i-1}(t))^2}$$
(g) Decision: Repeat Step (b)-(f) until $SD(i) < \epsilon$ and then put $IMF_j(t) \leftarrow h_{j,i}(t)$ (j^{th} IMF)
Step 4) Update residual: $r_j(t) \leftarrow r_{j-1}(t) - IMF_j(t)$
Step 5) Repeat Step 3 with $j \leftarrow j + 1$ until the number of extrema in $r_j(t) \leq 2$

where T is the time duration. The sifting is repeated several times (i) in order to get h to be a true IMF that fulfills the requirements R1 and R2. The result of the *sifting* procedure is that $x(t)$ will be decomposed into $IMF_j(t), j = 1, \ldots N$ and residual $r_N(t)$:

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

To guarantee that the IMF components retain enough physical sens of both amplitude and frequency modulations, we have to determine a criterion for the sifting process to stop. This is accomplished by limiting the size of the standard deviation SD computed from the two consecutive sifting results. Usually, SD is set between 0.2 to 0.3. Note that the EMD does not use any pre-determined filter or Wavelet function. It is a fully data driven method.

3. DENOISING PRINCIPLE

Let $f_j(t)$ be a noiseless IMF and IMF_j its noisy version. Consider a deterministic signal $y(t)$ corrupted by an additive Gaussian white random noise, $b_j(t)$, with a noise level $\sigma_j(t)$ as follows :

$$IMF_j(t) = f_j(t) + b_j(t) \tag{2}$$

where $j = \{1, \ldots, N\}$. An estimation $\tilde{f}_j(t)$ of $f_j(t)$ based on the noisy observation $IMF_j(t)$ is given by

$$\tilde{f}_j(t) = \Gamma[IMF_j(t), \tau_j] \tag{3}$$

where $\Gamma[h_j, \tau_j]$ is a preprocessing function, defined by a set of parameters τ_j, applied to signal h_j. The denoising signal $\tilde{x}(t)$ is given by:

$$\tilde{x}(t) = \sum_{j=1}^{N} \tilde{f}_j(t) + r_N(t) \tag{4}$$

In this work different kinds of preprocessing are used: temporal filtering using Savitzky-Golay[1] [7], Averaging, Median, and nonlinear transformation (hard and soft thresholding) [4].

3.1. EMD-Thresholding

A smooth version of the input signal can be obtained by thresholding the IMFs before signal reconstruction. If $\Gamma[., \tau_j]$ is a thresholding function, then τ_j is the threshold parameter. This threshold can be determined in different ways. Donoho and Johnstone [3] proposed an universal threshold for removing added Gaussian noise τ_j given by

$$\tau_j = \tilde{\sigma}_j \sqrt{2.\log_e(T)} \tag{5}$$
$$\tilde{\sigma}_j = MAD_j/0.6745 \tag{6}$$
$$MAD_j = \text{Median}\,\{|IMF_j(t) - \text{Median}\,\{IMF_j(t')\}|\} \tag{7}$$

where $\tilde{\sigma}_j$ is the estimation of the noise level of the j^{th} IMF (scale level) and MAD_j represents the absolute median deviation of the j^{th} IMF. The soft thresholding shrinks the IMF samples by τ_j towards zero as follows [4]:

$$\hat{f}_j(t) = \begin{cases} IMF_j(t) - \tau_j & \text{If } IMF_j(t) \geq \tau_j \\ 0 & \text{If } |IMF_j(t)| < \tau_j \\ IMF_j(t) + \tau_j & \text{If } IMF_j(t) \leq -\tau_j \end{cases} \tag{8}$$

Hard thresholding is defined as follows [4]:

$$\hat{f}_j(t) = \begin{cases} IMF_j(t) & \text{If } |IMF_j(t)| > \tau_j \\ 0 & \text{If } |IMF_j(t)| \leq \tau_j \end{cases} \tag{9}$$

3.2. EMD-SG

Rather than having its properties defined in the Fourier domain, and then translated to the time domain, Savitzky-Golay (SG) filter derives directly from a particular formulation of the data smoothing problem in the time domain [7]. Here M_L is the number of points used to the left of a data point i, i.e., earlier than it, while M_R is the number used to the right, i.e., later. A so-called causal filter would have $M_R = 0$.

$$\tilde{f}_j(i) = \sum_{m=-M_L}^{m=M_R} \alpha_m . IMF_j(i + m) \tag{10}$$

The idea of the filtering is to find filter coefficients α_m that preserve higher moments. Equivalently, the idea is to approximate the underlying function with the moving window by

[1] Also called DISPO filter (Digital Smoothing Polynomial).

a polynomial of higher order. For each point $IMF_j(i)$, we least-squares fit to a polynomial to all $(M_L + M_L + 1)$ points in the moving window, and then set $IMF_j(i)$ to be the value of that polynomial of that polynomial at position i. We move on the next sample $IMF_j(i + 1)$, to do a whole new least-squares fit using a shifted window. There are particular sets of coefficients α_m for which equation (10) automatically accomplishes the process of polynomial least-squares fitting inside a moving window [8]. To derive such coefficients, consider how $IMF_j(0)$, for example, might be obtained. We fit a polynomial of degree K in i, namely $\alpha_0+\alpha_1 i^1+\alpha_1 i^2+\ldots+\alpha_K i^L$ to the values $IMF_j(M_L), \ldots IMF_j(M_R)$. Then $IMF_j(0)$ will be the value of that polynomial at $i = 0$, namely α_0.

4. RESULTS

To test the denoising scheme, we have performed numerical simulations for four test signals: "Doppler", "Blocks", "Bumps" and "Heavysine" obtained using WAVELAB Software[2]. The method is also tested on one real signal: "ECG". The signals size is $T = 2048$. For synthesized signals the variance of the white Gaussian noise is set so that the original SNR (before denoising) is maintained at constant value (2 dB). The SNR of the "ECG" is -9 dB. The original signals and their noisy versions are shown in figures 1 and 2, respectively. Table I shows comparisons of SNR values for Averaging, Median, Wavelets, EMD-Soft and EMD-SG methods. Each noisy signal is decomposed using the EMD and the derived IMFs are filtered (thresholded) using preprocessing a P (SG, Median, Average,...) method. Hence the corresponding denoising scheme is termed "EMD-P" method. For SG filter the order L is set to 3. Each reconstructed signal plot (black line) is superposed on the corresponding free noise signal (cyan 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. 3(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 10 dB to 28 dB. As indicated in Table I, both the EMD-SG and the EMD-Soft outperform the Averaging and the Median methods. For Bumps and ECG signals both the EMD-SG and the EMD-Soft perform better than the Wavelets method. However, the Wavelets 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 but globally the EMD-SG performs better than Wavelets, EMD-Soft, Median and Averaging methods. For the ECG signal the Averaging method achieves better SNR than the Wavelets method. The oscillations seen in flat regions (Figs. 3(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. 3(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 Wavelets reconstruction.

We have investigated the effect of noise level value on the EMD denoising performances using "Bumps" signal:

	Signals				
	Doppler	***Blocks***	***Heavysine***	***Bumps***	***ECG***
	SNR	SNR	SNR	SNR	SNR
Noise	$2,03$	$2,03$	$2,03$	$2,03$	$-9,02$
Averaging	$9,86$	$9,06$	$9,46$	$12,66$	$7,23$
Median	$10,57$	$10,17$	$10,55$	$10,67$	$4,62$
Wavelets	**14,97**	**11,94**	**14,47**	$18,76$	$5,82$
EMD-Soft	$11,13$	$11,98$	$11,18$	$19,86$	$14,39$
EMD-SG	$13,57$	**12,00**	**14,50**	**20,60**	**17,77**

Table 1. Denoising results in SNR for test and ECG signals corrupted by Gaussian noise.

4.1. Noise effect

Figure (4) summarizes the noise effect analysis of Bumps signal with different values of square root of SNR (before denoising) ranging from 0.2 to 5 (step=0.2). Six methods (EMD-SG, EMD-Soft, EMD-Hard, EMD-Median, EMD-Average, Wavelets) have been compared. For Wavelets method the soft thresholding is used. For EMD-SG, EMD-Median and EMD-Average the window size w is set to 7. Figure (4) shows that as the square root of SNR increases (noise level decreases) the Wavelets and EMD-SG give the same results. At the same time EMD-Soft, EMD-Median, EMD-Average and EMD-Hard are equivalents but performs less better than EMD-SG and Wavelets. For very high noise levels EMD-Soft and EMD-Hard are equivalents to Wavelets.

5. CONCLUSION

This paper presents a signals denoising scheme. This approach, based on EMD method, is simple and fully data-driven. Results obtained for synthetic signals and for one real signal show that our method is effective for noise removal. To run the EMD with SG filter the window size is required. For better signal reconstruction we plan to study the effect of sampling frequency on interpolation to compute the upper and lower envelopes. We also plan how to adapt the SG filter order (L) to each IMF and how to find optimal size, w, window. To confirm the effectiveness of the EMD scheme, the method must be evaluated with a large class of real signals and in different

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

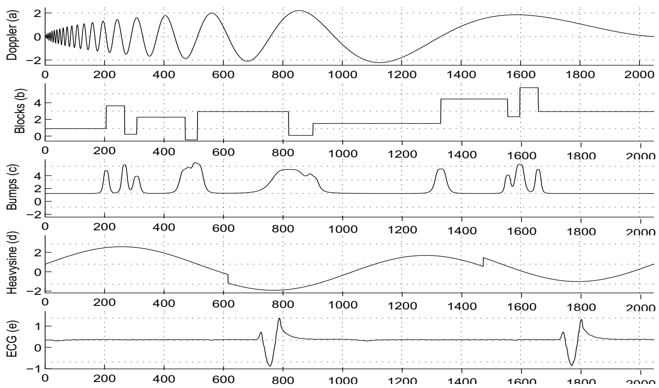

Fig. 1. Synthetic and real signals with T=2048.

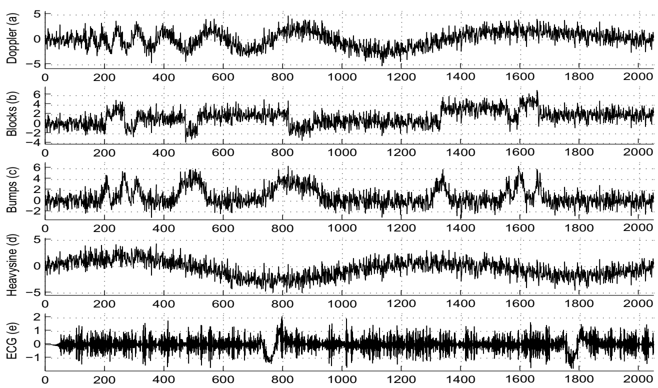

Fig. 2. Noisy signals (SNR=2dB; SNR=-9dB for "ECG").

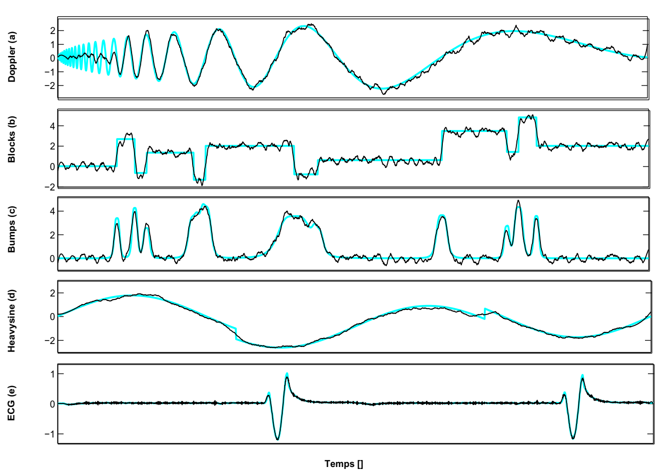

Fig. 3. Denoising results in SNR (dB) of test signals corrupted by Gaussian noise.

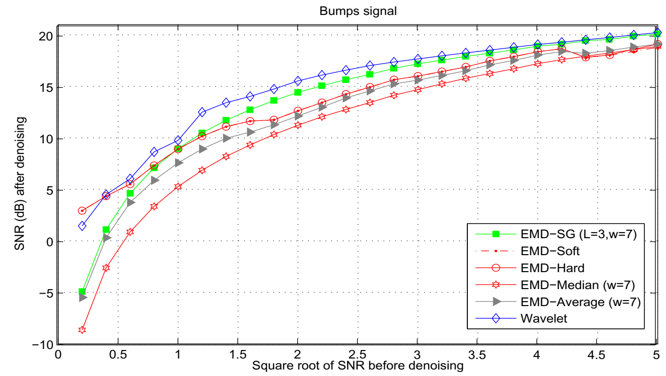

Fig. 4. Noise effect analysis using Bumps signal.

experimental. The presented noise effect analysis of "Bumps" signals must be extended to other signals to confirm the obtained results.

6. REFERENCES

[1] 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.

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

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

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

[5] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Sig. Process.,* vol. 41, pp. 3397-3415, 1993.

[6] M.M. Goodwin and M. Vetterli, "Matching pursuit and atomic signal models based on recursive filter banks," *IEEE Trans. Sig. Process.,* vol. 47, no. 7, pp. 1890-1901, 1999.

[7] 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.

[8] W.H. Press, S.A. Teukolsky, W.T. Vetterling and B. P. Flannery, *Numerical Recipes in C*, Cambridge University Press, 1992.