IMPROVEMENT OF ENSEMBLE EMPIRICAL MODE DECOMPOSITION BY OVER-SAMPLING

Article *in* Advances in Adaptive Data Analysis · August 2013

DOI: 10.1142/S179353691350012X

2 authors:

Raïs El'hadi Bekka
Ferhat Abbas University of Setif

YAAKOUB BERROUCHE
Ferhat Abbas University of Setif

Advances in Adaptive Data Analysis
Vol. 5, No. 3 (2013) 1350012 (10 pages)

DOI: 10.1142/S179353691350012X

IMPROVEMENT OF ENSEMBLE EMPIRICAL MODE DECOMPOSITION BY OVER-SAMPLING

RAÏS EL'HADI BEKKA* and YAAKOUB BERROUCHE†

Université de Sétif 1, Faculty of Technology
Department of Electronics, LIS
19000 Sétif, Algeria
**bekka_re@yahoo.fr*
†*berrouche_ya@yahoo.com*

Received 21 December 2011
Revised 24 December 2012
Accepted 30 April 2013
Published 26 August 2013

The empirical mode decomposition (EMD) is a useful method for the analysis of nonlinear and nonstationary signals and found immediate applications in diverse areas of signal processing. However, the major inconvenience of EMD is the mode mixing. The ensemble EMD (EEMD) was proposed to solve the problem of mode-mixing with the assistance of added noises producing the residue noise in the signal reconstructed. The residue noise in the IMFs can be reduced with a large number of ensemble trials at the expense of the increase of computational time. Improving the computing time of the EEMD by reducing the number of ensemble trials was thus proposed in this paper by over-sampling the signal to be decomposed. Numerical simulations were conducted to demonstrate proposed approach.

Keywords: Empirical mode decomposition; ensemble empirical mode decomposition, over-sampling.

1. Introduction

The EMD method was introduced by Huang *et al.* [1998] for adaptively decomposing nonlinear and nonstationary signal. Unlike the classical Fourier transform and wavelet decompositions using *a priori* determined basis functions, EMD directly derives the basic functions from the signal itself. The principle of EMD is to decompose a signal into a collection of oscillatory functions named the intrinsic mode functions (IMFs) by a sifting process. The analysis of signal decomposition having a uniform distribution by EMD shows that this method is organized in dyadic filter bank [Flandrin *et al.* (2004); Rilling *et al.* (2005)]. It found immediate applications in diverse areas of signal processing, such as biomedical engineering

*Corresponding author.

[Liang *et al.* (2000); Liang *et al.* (2005); Salisbury and Sun 2007)], mechanical fault diagnosis [Loutridis (2004); Yang *et al.* (2006); Gao *et al.* (2008)] image and geophysical applications [Linderhed (2002); Huang and Wu (2008); Numes and Deléchelle (2009)]. The main drawback of EMD is the mode mixing problem, where oscillations of different time scales can appear in one IMF, or oscillations with the same time scale can appear in different IMFs.

To overcome the mode mixing problem, the ensemble empirical mode decomposition (EEMD) was introduced by Wu and Huang [2009]. EEMD is noise-assisted data analysis method in which the IMFs are defined as the mean of an ensemble of trials, each consisting of the original signal plus a finite white noise. The EEMD has been applied to mechanical fault detection [Lei *et al.* (2009); Lei *et al.* (2011)], analyze experimental space-time data in flows detection [Debert *et al.* (2011)], image data analysis [Kopecky (2010)] and optical fringe pattern analysis [Zhou *et al.* (2009)]. However, the resulting IMFs will be contaminated by added noise. The residue noise can be decreased by using a large number of ensemble trials resulting in a high computation time.

The objective of this study was to reduce the number of ensemble trials in EEMD by over-sampling the original signal while providing more accurate decomposition results.

2. Empirical Mode Decomposition

The EMD method decomposes a signal $x(t)$ into a finite number of IMFs and a nonzero-mean low order polynomial called residue [Huang *et al.* (1998)]:

$$x(t) = \sum_{i=1}^{N} c_i(t) + r_N(t), \tag{1}$$

where N is the number of IMFs extracted for a given accuracy, $c_i(t)$ is the IMF of order i and $r_N(t)$ is the final residue of the decomposition that results after a number N of IMFs.

IMFs are extracted from high frequency to low frequency (IMFs of lower order are high frequencies and IMFs of higher order are bases frequencies).

A component is an IMF if and only if [Huang *et al.* (1998)]:

(1) The number of maximum must equal the number of zeros or differ at most by one.
(2) In each period, it is necessary that the signal average is zero.

The condition (1) ensure that the IMF is a narrow band signal and the condition (2) ensure that the IMF has one component (over a period, a single frequency exists).

The main weakness of the EMD is the mode mixing effect. As a result, the IMF loses its physical meaning.

3. Ensemble Empirical Mode Decomposition

The EEMD defines the true IMF as the mean of an ensemble of trials, each consisting of the investigated signal plus a finite white noise. For a number of ensemble trials N_e given, a finite white noise is added to the signal in each trial and then this noisy signal is decomposed into IMFs by using the EMD method. Eventually, the true IMFs are obtained by averaging the IMFs of the same order [Wu *et al.* (2009)].

$$s_j(t) = x(t) + b_j(t), \qquad 1 \le j \le N_e \tag{2}$$

$$c_i(t) = \frac{1}{N_e}\sum_{j=1}^{Ne} IMF_{ij}(t), \quad 1 \le i \le N \tag{3}$$

$$r_N(t) = \frac{1}{N_e}\sum_{j=1}^{Ne} r_{Nj}(t), \qquad 1 \le i \le N \tag{4}$$

$$x(t) = \sum_{i=1}^{N} c_i(t) + r_N(t), \tag{5}$$

where $s_j(t)$ is the noisy signal by white noise of order j, $IMF_{ij}(t)$ is the IMF of order i of signal $s_j(t)$ and r_N is the rest of signal decomposition $s_j(t)$.

Indeed the EEMD method removes the problem of mode mixing; however the noise amplitude and the number of ensemble trials are the two parameters that must be controlled. An appropriate choice of the noise amplitude is achieved by using the signal to noise ratio *SNR* [Zhang *et al.* (2010)].

A simulated signal $x(t)$ consisting of a permanent component of low-frequency and of a transient component of high-frequency was used to show the resolution of the mode mixing problem and the reduction of residue noise. The signal is described as (Fig. 1):

$$x(t) = \sin(2\pi \times f_1 t) + \alpha \times \sin(2\pi \times f_2 t)\left[e^{\frac{-(t-t_0)^2}{\tau}} + e^{\frac{-(t-t_1)^2}{\tau}}\right], \quad 0 \le t \le 0.025\,\text{s}. \tag{6}$$

With $t_0 = 0.005\,\text{s}$, $t_1 = 0.015\,\text{s}$, $\tau = 6 \times 10^{-6}\,\text{s}$, $\alpha = 0.09$, $f_1 = 200\,\text{Hz}$ and $f_2 = 2000\,\text{Hz}$.

To study the effect of the noise amplitude and the number of ensemble trials, the signal $x(t)$ was sampled at a frequency of 5 kHz. Figure 2 illustrates the relationship between *SNR* and the number of redundant components, and the relationship between the correlation coefficient and number of ensemble trials. Figure 2(a) shows the variation of the number of IMFs redundant according to *SNR*. A value of 0 represents no redundant IMF component and a value of -1 represents mode mixing [Zhang *et al.* (2010)]. It can be seen in Fig. 2(a) that *SNR* within the range of 15–20 dB is a good value for a good EEMD decomposition.

Theoretically, an infinite number of ensemble trials would be required to fully eliminate the impact of the white noise added to the signal. A practicable number of

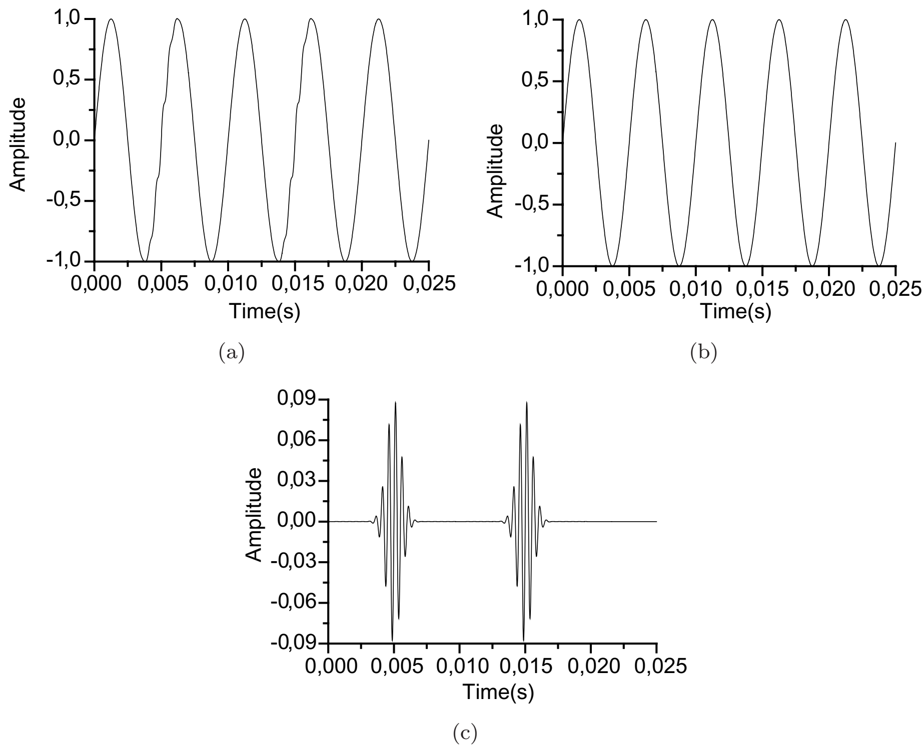


Fig. 1. (a) Signal $x(t)$ and its basic components. (b) The component of low-frequency. (c) The transient component of high-frequency.

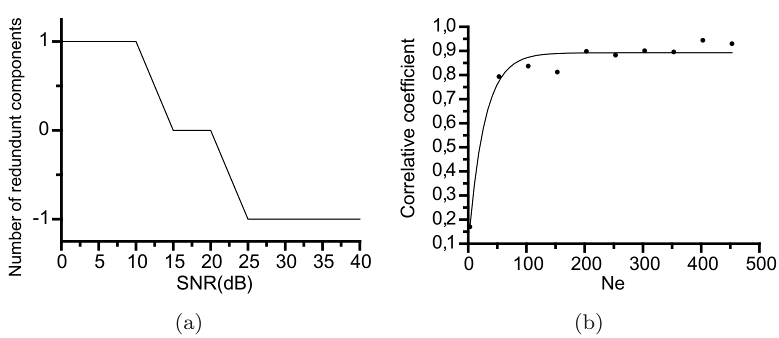


Fig. 2. (a) Relationship between SNR and the decomposition results of the signal $x(t)$ sampled at 5 kHz obtained from EEMD. (b) Relationship between the correlation coefficient and number of ensemble trials obtained from EEMD.

ensemble trials is obtained by the correlation between the intrinsic mode function and the corresponding component in the simulated signal [Zhang *et al.* (2010)]. Figure 2(b) shows in order that the signal components have a high correlation coefficient of 94% (IMF least noisy), it is necessary that the number of ensemble trials is greater than 400.

By decomposing the signal $x(t)$ sampled at the frequency of 5 kHz using EEMD with an SNR of 15 dB and two values of the number of ensemble trials, we obtained the results showed in Fig. 3 for $N_e = 200$ and Fig. 4 for $N_e = 20{,}000$. These figures show that the EEMD succeeded in extracting the two components and therefore the problem of mode mixing has been resolved. In addition, the second IMF corresponding to the permanent component of low frequency of the signal appears to be stable for both values of the number of ensemble trials; and the first IMF representing the transient component of high-frequency is less noisy for the largest value of the number of ensemble trials. Thus, the noise effect on the

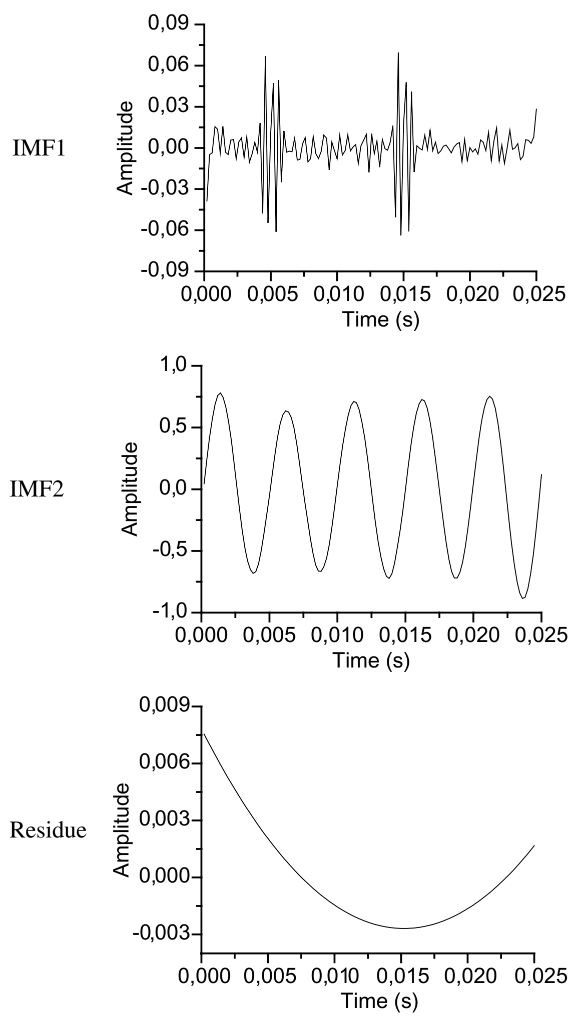


Fig. 3. Decomposition of signal $x(t)$ sampled at a frequency of 5 kHz by EEMD method ($SNR =$ 15 dB, $N_e = 200$).

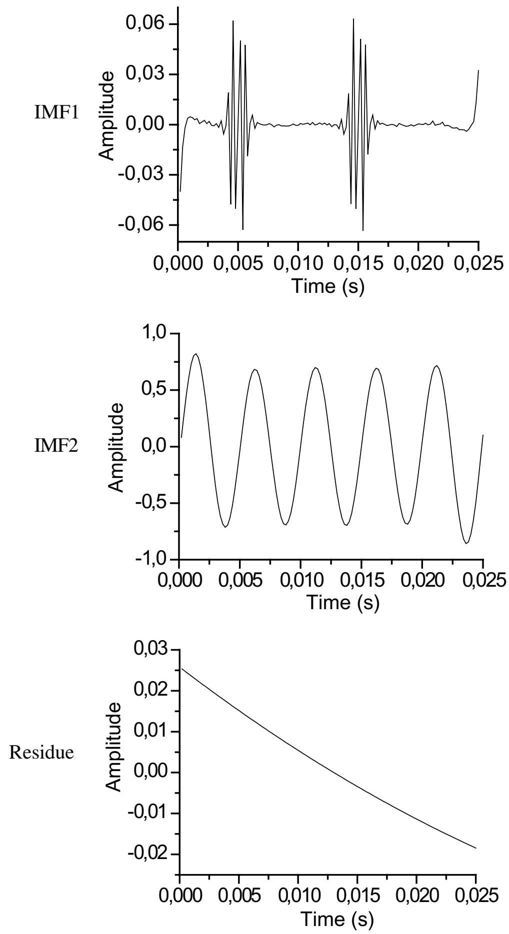


Fig. 4. Decomposition of signal $x(t)$ sampled at a frequency of 5 kHz by EEMD method ($SNR =$ 15 dB, $N_e = 20{,}000$).

IMF decreases significantly with the increase of the number of ensemble trials. Nevertheless, this increase distorts the ends of the transitory component (IMF1).

The accuracy was evaluated by the root mean square error (RMSE) of the differences between the original signal and the reconstructed signal, it is ideal if it is small. Table 1 shows how the RMSE decreases with increasing the number of ensemble trials. Of course, this decrease is achieved at the expense of computing time.

Table 1. RMSE derived from EEMD (sampling frequency of 5 kHz).

N_e	200	450	1,500	8,000	20,000
RMSE	0,0091	0,0052	0,0030	0,0012	0,00080963

4. Over-Sampling in EEMD

4.1. *Study of EEMD according to the sampling frequency*

To examine the behavior of EEMD parameters with respect to the sampling frequency of the signal to decompose, the signal $x(t)$ was sampled with different sampling frequencies (from $f_s = 2f_{\max}$ to $f_s = 200f_{\max}$). The signal to noise ratio and the number of ensemble trials were evaluated according to frequency ratio ($f_s/f_{\max}$) for decomposition constituted of nonredundant IMFs having a correlation coefficient greater than or equal to 98%.

Figure 5(a) shows that the $SNR_{\max}$, corresponding to minimum noise added to have decomposition, according to $f_s/f_{\max}$, follows a logarithmic law which provides a slow transition in the interval (25,200) and an abrupt transition in the interval from (2,25). It also shows that it is necessary to add a large noise (small SNR) for a signal sampled at the limit of the sampling theorem (small $f_s/f_{\max}$) and a low noise for over-sampled signals.

Figures 5(b) and 5(c) show how the number of ensemble trials varies with the sampling frequency. Figure 5(b) shows that the signals sampled at the limit of the

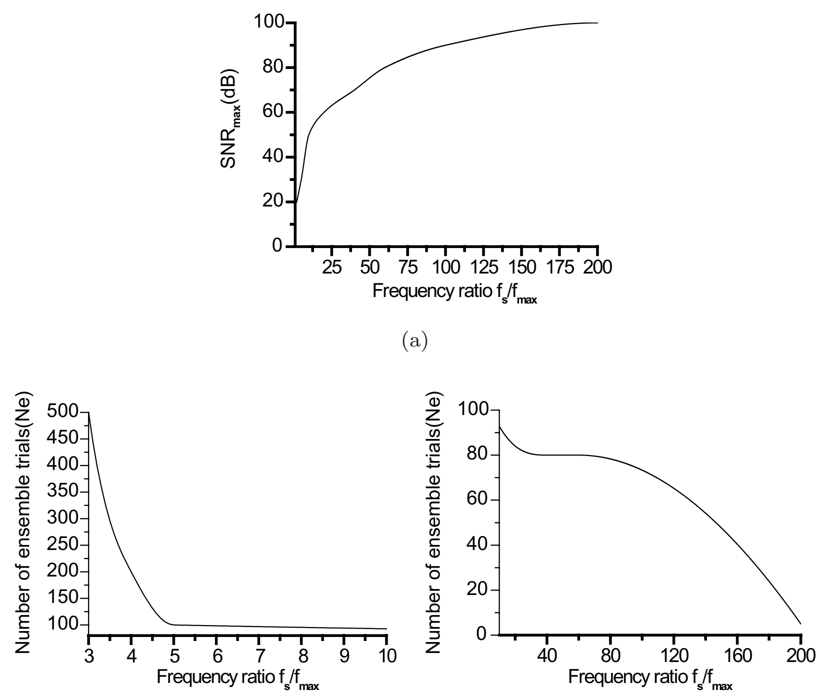
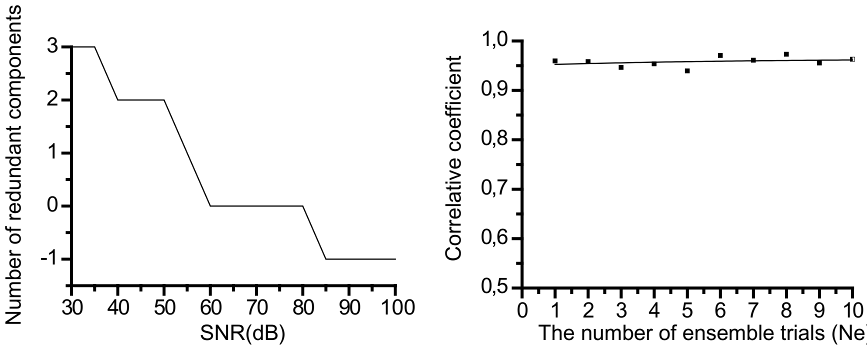


Fig. 5. The effect of sampling on the parameters of the method EEMD: (a) $SNR_{\max}$ according to $f_s/f_{\max}$, (b) and (c) the number of ensemble trials according to $f_s/f_{\max}$.



(a) (b)

Fig. 6. (a) Relationship between SNR and the decomposition results of the signal $x(t)$ sampled at 100 kHz obtained from EEMD. (b) Relationship between the correlation coefficient and number of ensemble trials obtained from EEMD.

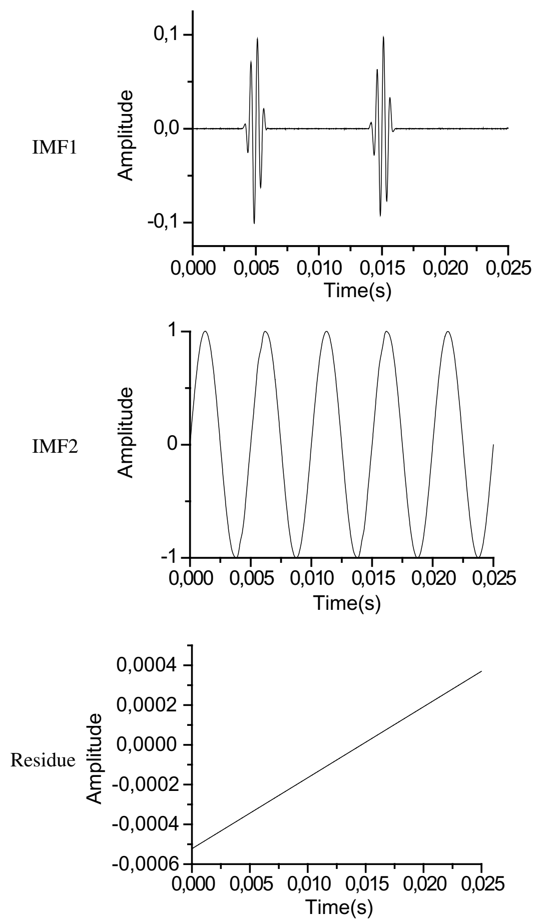


Fig. 7. Decomposition of signal $x(t)$ over-sampled at a frequency of 100 kHz by EEMD method ($SNR = 70\,\mathrm{dB}, N_e = 5$).

sampling frequency require a very large number of ensemble trials to their decomposition, 500 ensemble trials are necessary to decompose the signal $x(t)$ sampled at the frequency $f_s = 3f_{\max}$. It should be noted more than 10,000 ensemble trials are needed to decompose the same signal sampled at the frequency $f_s = 2f_{\max}$. Figure 5(c) shows that for over-sampled signals, the number of trials is lower than 100. Indeed for the weakly sampled signals, the necessity to add an important noise allows the components of the noise of intervening in a majority way on the IMFs, and therefore the elimination of this noise requires an awful lot of ensemble trials. For a standard decomposition with addition of a low amplitude noise it is necessary to over-sample the original signal to reduce the number of trials.

4.2. *Effect of over-sampling*

Although EEMD solved the problem of mode mixing, the large number of ensemble trails to eliminate the residue of added white noises in the signal reconstruction requires a large computation time. The improvement on the computational efficiency of EEMD can be achieved by over-sampling the signal to be decomposed. Figure 6 shows the results of decomposition when the signal $x(t)$ was sampled at the frequency of 100 kHz. To ensure that the signal components have a high correlation coefficient of 95%, Fig. 6(b) shows that it is necessary to have a number of ensemble trials superior to one. By decomposing the signal $x(t)$ over-sampled at a frequency of 100 kHz using EEMD method with a *SNR* of 70 dB and a number of ensemble trials of five, we obtained the result showed in Fig. 7. This figure shows that the EEMD method succeeded in extracting the two components and how the residue has decreased. This represents a reduction of 99,975% of the number of ensemble trails compared to results obtained with $N_e = 20{,}000$ (Fig. 4). In addition the RMSE derived from EEMD to decompose this over-sampled signal (sampling frequency of 100 kHz) was $9{,}9412 \times 10^{-5}$ which is also smaller than that obtained with $N_e = 20{,}000$. This comparison clearly illustrates the improvement provided by the over-sampling of the original signal.

5. Conclusion

Over-sampling of the original signal was proposed to improve processing efficiency of EEMD by reducing the number of ensemble trials. Actually for the weakly sampled signals, the necessity to add an important noise allows the components of the noise of intervening in a majority way on the IMFs, and therefore the elimination of this noise requires an awful lot of ensemble trials. However, this increase of the number of ensemble trials distorts the ends of the component of high-frequency (IMF1). For decomposition with low amplitude of the added noise, a better decomposition of original signals with EEMD was achieved by their over-sampling. In addition the RMSE derived from EEMD to decompose the over-sampled signals is also smaller than those obtained with signals weakly sampled.

References

Debert, S., Pachebat, M., Valeau, V. and Gervais, Y. (2011). Ensemble-empirical-mode-decomposition method for instantaneous spatial-multi-scale decomposition of wall-pressure fluctuations under a turbulent flow. *Exp. Fluids*, **50**: 339–350.

Flandrin, P., Rilling, G. and Goncalces, P. (2004). Empirical mode decomposition as a filter bank. *IEEE Signal Proc. Lett.*, **11**(2): 112–114.

Gao, Q., Duan, C., Fan, H. and Meng, Q. (2008). Rotating machine fault diagnosis using empirical mode decomposition. *Mech. Syst. Signal Process.* **22**(5): 1072–1081.

Huang, N. E. and Wu, Z. (2008). A review on Hilbert-Huang transform: Method and its applications to geophysical studies. *Rev. Geophys.*, **46**: RG2006.

Huang, N. E., Zheng, S. and Steven, R. L. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. A*, **454**: 903–995.

Kopecky, M. (2010). Ensemble empirical mode decomposition: Image aata analysis with white-noise reflection. *Acta Polytechnica*, **50**(6): 49–56.

Lei, Y., He, Z. and Zi, Y. (2009). Application of the EEMD method to rotor fault diagnosis of rotating machinery. *Mech. Syst. Signal Process.*, **23**(4): 1327–1338.

Lei, Y., He, Z. and Zi, Y. (2011). EEMD method and WNN for fault diagnosis of locomotive roller bearings. *Expert Syst. Appl.*, **38**: 7334–7341.

Liang, H., Lin, Q. and Chen, J. D. Z. (2005). Application of the empirical mode decomposition to the analysis of esophageal manometric data in gastroesophageal reflux disease. *IEEE Trans. Biomed. Eng.*, **52**(10): 1692–1701.

Liang, H., Lin, Z. and McCallum, R. W. (2000). Artifact reduction in electrogastrogram based on the empirical model decomposition method. *Med. Biol. Eng. Comput.*, **38**: 35–41.

Linderhed, A. (2002). 2D empirical mode decompositions in the spirit of image compression. *Wavelet and Independent Component Analysis Applications IX, Conf.* No. 9, *SPIE Proc.*, Orlando FL, USA, 4738, pp. 1–8.

Loutridis, S. J. (2004). Damage detection in gear systems using empirical mode decomposition. *Eng. Struct.*, **26**(12): 1833–1841.

Numes, J. C. and Deléchelle, E. (2009). Empirical mode decomposition: Application on signal and image processing. *Adv. Adapt. Data Anal.*, **1**(1): 125–175.

Rilling, G., Flandrin, P. and Goncalces, P. (2005). Empirical mode decomposition, fractional gaussian noise and hurst exponent estimation. *Proc. ICASSP 05 IEEE Int. Conf. Acoustics Speech and Signal Processing*, Philadelphia, USA, **4**(1): 489–492.

Salisbury, J. I. and Sun, Y. (2007). Rapid screening test for sleep apnea using a nonlinear and nonstationary signal processing technique. *Med. Eng. Phys.*, **29**(3): 336–343.

Wu, Z. and Huang, N. E. (2009). Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Adv. Adapt. Data Anal.*, **1**(1): 1–41.

Yang, Y., Yude, J. and Cheng, J. (2006). A roller fault diagnosis method based on EMD energy entropy and ANN. *J. Sound Vib.*, **294**(1): 269–277.

Zhang, J., Yan, R., Gao, R. X. and Feng, Z. (2010). Performance enhancement of ensemble empirical mode decomposition. *Mech. Syst. Signal Process.*, **24**(7): 2104–2123.

Zhou, X., Zhao, H. and Jiang, T. (2009). Adaptive analysis of optical fringe patterns using ensemble empirical mode decomposition algorithm. *Opt. Lett.*, **34**(13): 2033–2035.