The *Scientific* World JOURNAL

*Research Article*

# A Data-Driven Noise Reduction Method and Its Application for the Enhancement of Stress Wave Signals

**Hai-Lin Feng, Yi-Ming Fang, Xuan-Qi Xiang, Jian Li, and Guan-Hui Li**

*School of Information Engineering, Zhejiang A & F University, Zhejiang, Lin'an 311300, China*

Correspondence should be addressed to Yi-Ming Fang, zjfcfym@gmail.com

Ensemble empirical mode decomposition (EEMD) has been recently used to recover a signal from observed noisy data. Typically this is performed by partial reconstruction or thresholding operation. In this paper we describe an efficient noise reduction method. EEMD is used to decompose a signal into several intrinsic mode functions (IMFs). The time intervals between two adjacent zero-crossings within the IMF, called instantaneous half period (IHP), are used as a criterion to detect and classify the noise oscillations. The undesirable waveforms with a larger IHP are set to zero. Furthermore, the optimum threshold in this approach can be derived from the signal itself using the consecutive mean square error (CMSE). The method is fully data driven, and it requires no prior knowledge of the target signals. This method can be verified with the simulative program by using Matlab. The denoising results are proper. In comparison with other EEMD based methods, it is concluded that the means adopted in this paper is suitable to preprocess the stress wave signals in the wood nondestructive testing.

## 1. Introduction

Recovery of a signal from observed noisy data is usually regarded as an important preprocessing and has been an area of research for decades. Many algorithms for noise reduction have been reported so far in the literature including traditional linear filter, such as Butterworth low pass filter, Wiener filter, and wavelet based thresholding filter [1]. Most of them have been proved to be effective in removing the unwanted components. For example, Hsu et al. succeeded in removing the aliasing on the original step-edge response curve (SRC) caused by the binning of Moire patterns [2]. However, the linear filtering methods are not very effective when the signals contain sharp edges and impulses of short duration [3]. As for wavelet based denoising methods, it's difficult to select the wavelet base, scale, threshold function, and optimal threshold value.

In 1998, empirical mode decomposition (EMD) was designed by Wu and Huang primarily for decomposing the nonlinear and nonstationary signals into a series of intrinsic mode functions (IMFs) [4]. The main advantage of EMD is that it depends entirely on the data itself. Consequently, the

results preserve the full nonstationarity characteristics of the target signals. Seen in this light, the EMD method is superior to the wavelet analysis approach, where the basic functions are fixed and, thus, do not necessarily match all real signals [3]. The property of EMD to behave as a dyadic filter bank resembling those involved in wavelets [5] has been useful in signal denoising and received more and more attention [3, 5–11].

However, the EMD algorithm may encounter the problem of mode mixing when a signal contains intermittency. Therefore the ensemble empirical mode decomposition (EEMD) was introduced [12]. The algorithm defines the IMF set for an ensemble of trials, each one obtained by applying EMD to the signal of interest with added independent identically distributed white noise of the same standard deviation. Taking into account properties of the white noise, the problem of mode mixing can be overcome [13]. Owing to the impressive performance, EEMD has been used to address several problems in the field of science and engineering. It was employed to calculate the residual signal with the purpose of detecting localized gearbox faults of damage at an early stage [14]. In [15], the authors used it to predict

the short term passenger flow with back-propagation neural networks (BPN). Zhang and Xie decomposed the impact echo signals into different spectral composition for defect signal extraction [16].

As a more robust and noise-assisted version of EMD, EEMD can be used as an alternate in EMD based denoising methods. Furthermore, the use of the EEMD process as a filter and its comparisons with the EMD method have just recently been studied in [17, 18]. An improved filtering performance can be achieved by EEMD than EMD with suitable added noise and sufficient trial number. In our earlier paper, we proposed a signal denoising strategy using EEMD combined with the instantaneous half period (IHP) to restore stress wave signal from observed raw data [19].

In this paper, we investigated an improvement based on the above method. The main contribution of this paper is that, utilizing the consecutive mean square error (CMSE), we can determine the optimum threshold adaptively. The method can work well on condition that no prior knowledge is required. The whole procedure is fully data driven.

## 2. Conventional EEMD Based Filtering Approach

*2.1. EMD Algorithm.* The EMD algorithm can be described as follows [4].

(1) Extract all the local maxima and minima of $x(k)$.

(2) Form the upper and lower envelop by cubic spline interpolation of the extrema point developed in step (1).

(3) Calculate the mean function of the upper and lower envelop, $m_1(k)$.

(4) Let $h_1(k) = x(k) - m_1(k)$. If $h_1(k)$ is a zero-mean process, then the iteration stops and $h_1(k)$ is an IMF1, named $c_1(k)$, else go to step (1).

(5) Define $r(k) = x(k) - c_1(k)$.

(6) If $r(k)$ still has least 2 extrema then go to step (1) else decomposition process is finished.

At the end of the procedure, we have a residue $r(k)$ and a collection of $n$ IMF, named from $c_1(k)$ to $c_n(k)$. The original signal can be represented as

$$x(k) = \sum_{i=1}^{n} c_i(k) + r(k). \tag{1}$$

*2.2. EEMD Algorithm.* The steps for EEMD are as follows [12].

(1) Initialize the number of ensemble $J$, the amplitude of the added white noise, and $j = 1$.

(2) Add a white noise series to the targeted signal, $x_j(k) = x(k) + n_j(k)$.

(3) Apply EMD to the noise-added signal $x_j(k)$ to derive a set of IMFs $c_{i,j}(k)$ ($i = 1, 2, \ldots, n$) and residues $r_j(k)$, where $c_{i,j}(k)$ denotes the $i$th IMF of the $j$th trial, and $n$ is the number of IMFs.

(4) Repeat steps (1) and (2) until $j > J$.

(5) Average over the ensemble to obtain the final IMF of decompositions as the desired output:

$$\bar{c}_i(k) = \frac{1}{J} \sum_{j=1}^{J} c_{i,j}(k) \quad (i = 1, 2, \ldots, n),$$

$$\bar{r}(k) = \frac{1}{J} \sum_{j=1}^{J} r_j(k). \tag{2}$$

Just as the EMD method, the given signal, $x(k)$ can be reconstructed according to the following equation:

$$x(k) = \sum_{i=1}^{n} \bar{c}_i(k) + \bar{r}(k). \tag{3}$$

In contrast with EMD, EEMD skillfully eliminates the mode mixing phenomenon and the results obtained by EEMD reflect the nature of signals more accurately [20]. In spite of a heavy computational load, it is still suitable for getting better performance [17, 18]. So the EEMD algorithm is utilized in this study. Furthermore, to illustrate the performance of the present approach, we restricted ourselves to the EEMD method which was used as the mean to decompose the signal in both EMD based method and EEMD based method involved in this study.

*2.3. EEMD Based Filtering Approach.* A conceptual model of EEMD based filtering approach is depicted in Figure 1. It consists of three steps. First, noisy signal is adaptively decomposed into IMFs by means of EEMD algorithm. In the next stage, these IMFs are classified and detected in terms of their different properties under certain criteria and those undesirable IMFs are thereby removed by switching corresponding switches "off", that is, setting the values of the corresponding $c_i(k)$ to zeros, and will not be used in the signal reconstruction. Finally, the recovered signal is reconstructed with only a few IMFs which are signal dominated. Thus it is reasonable to assume that a full restoration is possible by this approach if enough information is available on the characteristics of underlying signals, depending on the selection and rejection of the IMFs.

The second step is typically performed by two operations. Based on this, the existing EEMD based method can be divided into two categories: EEMD based thresholding filter [6, 8, 21] or EEMD based low pass filter [3, 17, 21]. The former method reconstructs the signals with all the IMFs, which use the previous threshold as in wavelet analysis. Because most of the important structures of the signal are often concentrated on lower frequency components (high order IMFs), and they generally decrease towards the high frequency modes (low order IMFs), the noise power can be reduced significantly by adding a suitable threshold on the high frequency modes. However, when applying the threshold on the high order IMFs with little or no noise, the main features of the original signal may be lost or changed accordingly.
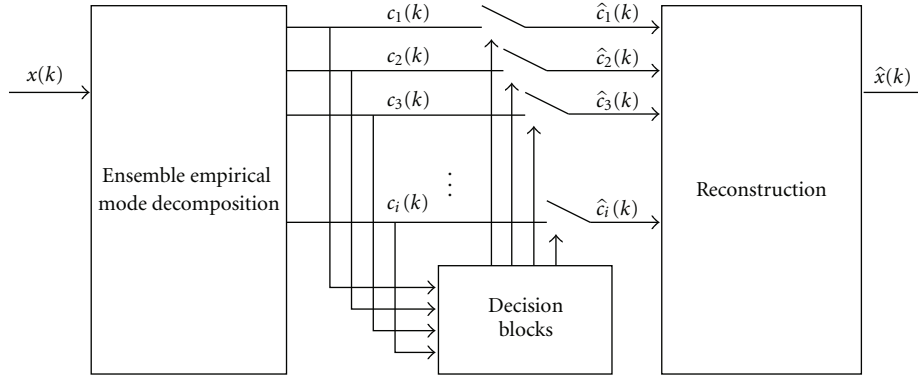
Figure 1: Conceptual model of noise reduction method using EEMD.

The latter method, EEMD-based low pass filter, was developed based on the assumption that the IMFs derived by EEMD can only be divided into two classes: noise-only IMFs and signal-only IMFs. Accordingly, it is feasible to use a criterion to classify and remove the noise-only IMFs, which leads to the result that the signal-only IMFs are partially reconstructed. However, noises are usually distributed over all IMFs. Thus the low pass filter based on EEMD will remove the high frequency components of both the noise and the signal, and the low frequency components of noise also remain.

## 3. The Data-Driven IHP (DIHP) Approach

In our presented method, the $i$th IMF, $i = 1, 2, \ldots, n$, can be denoted as $c_i(k)$, where $n$ is the number of IMFs. Accordingly, we can use mathematical operations to locate the zero-crossings of $c_i(k)$. The symbol $ZP_i^j$ is used to define the $j$th zero-crossing of the $i$th IMF correspondingly. Moreover, the time when the $j$th zero-crossing of the $i$th IMF emerges is defined as $\tau_i^j$. As a result, the time interval between $ZP_i^{j+1}$ and $ZP_i^j$ can be treated as the half period of an oscillation, which is used in our method as a criterion. Considering the half periods may be different from each other, we define it as IHP and the formula can be expressed as

$$T_i^j = \tau_i^{j+1} - \tau_i^j. \tag{4}$$

Usually, the signal structures are corresponding to the slow time variation of data. Besides, the frequency of the signal is often lower than that of the noise structures [3]. Consequently, it can be supposed that the IHP of a signal dominated oscillation is longer than the IHP of a noise dominated oscillation. Based on this assumption, the symbol thr is introduced to be a threshold, which allows us to retrieve the most important structures of the signal from its noisy version. If the IHP is bigger than thr, the waveforms between the two adjacent zero-crossings will be considered as signal dominated oscillations and be retained, whereas the waveforms with smaller LHP will be treated as noise

dominated oscillations and be set to zeros. This process can be described as

$$\hat{c}_i(k) = \begin{cases} c_i(k), & T_i^j \geq \text{thr}, \\ 0, & \text{others}, \end{cases} \quad ZP_i^j < k \leq ZP_i^{j+1}. \tag{5}$$

Finally a reconstruction process of projecting the restored IMF, $\hat{c}_i(k)$, back onto the filtered signals is done as follows:

$$\hat{x}(k) = \sum_{i=1}^{n} \hat{c}_i(k) + \hat{r}(k). \tag{6}$$

Note that the filtering effect is related to the value of thr. A large thr would result in oversmoothing of the target signal, thus removing some low-frequency oscillations while these oscillations are signal dominated. Moreover a small thr might not be able to remove the artifacts, hence resulting in a signal of relatively low quality. We have earlier reported a solution based on the maximum frequency and a constant coefficient which can be determined with experience [19]. In this paper, we concern the method under the condition that no prior knowledge about the target signal is required. It is a very common problem because the prior knowledge is unavailable or can be obtained with high cost in many applications.

In general, the aim of the filtering is to find an approximation reconstructed signal $\hat{x}(k)$ from the observed signal $x(k)$ with minimum errors, that is, with lower distortion measures such as mean square error (MSE), mean absolute error (MAE), or mean square difference (MAD). Unfortunately, these measures can not be calculated because the original signal is unknown in practice. Boudraa proposed a criterion, called CMSE, based on the squared Euclidean distance between two consecutive reconstructions of the signal. It does not require any knowledge of the target signal and is a fully data-driven approach. In this study, we attempted to find the optimum threshold $\text{thr}_{\text{opt}}$ by minimizing the cost function CMSE

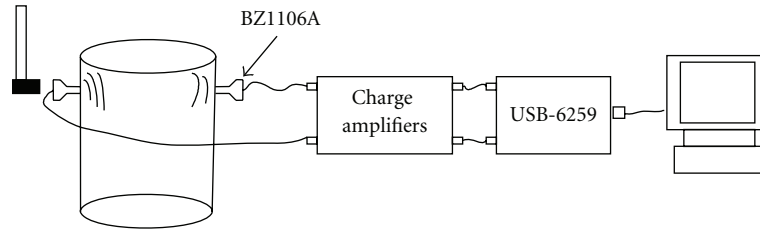$$\text{thr}_{\text{opt}} = \arg\min\{\text{CMSE}(\hat{x}_m(k), \hat{x}_{m+1}(k))\}, \tag{7}$$

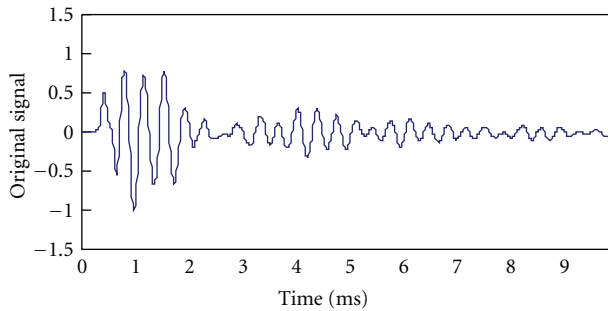FIGURE 2: The self-developed signal collection system.


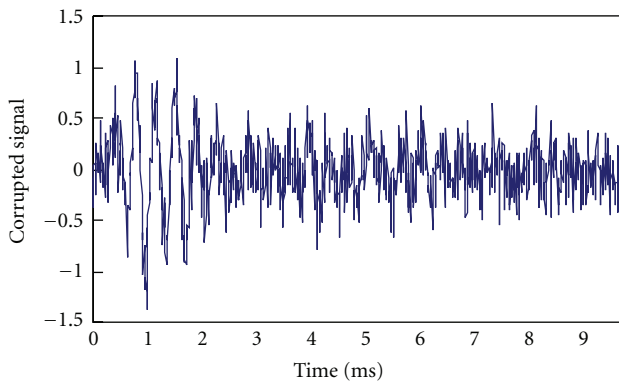
FIGURE 3: A typical stress wave signal used in this study.



FIGURE 4: The corrupted version of the signal shown in Figure 3.

The CMSE is defined as follows:

$$\text{CMSE}(\hat{x}_m(k), \hat{x}_{m+1}(k)) \triangleq \frac{1}{L} \sum_{k=0}^{L-1} [\hat{x}_m(k) - \hat{x}_{m+1}(k)]^2, \quad (8)$$

where $\hat{x}_m(k)$ and $\hat{x}_{m+1}(k)$ are signals that are reconstructed using the thr = $m$ and thr = $m + 1$, respectively.

## 4. Experiments and Results

*4.1. Stress Wave Signals Description.* During the past few decades, many advanced signal processing algorithms have been utilized to analysis the stress wave signal for the wood nondestructive testing, such as spectral analysis [22], wavelet analysis [23]. Unfortunately, no developed method or system is used worldwide until now due to the difficulty to extract useful information directly from the raw signals. The collected signals can be viewed as the result of multiple

interferences and reflections of these two waves fitting the boundary conditions, which interfered with the stress wave information identification [24]. Therefore, noise reduction is a necessary step for any stress wave based wood test technique to pave the way for further discovery in physics and nature [19].

Figure 2 illustrates the signal acquisition system which was used to collected stress wave signals in this study. We selected a typical *Cinnamomum camphora* trunk with a diameter of 27 cm as the sample. The piezoelectric receivers BZ1106A, coupled with the matching charge amplifiers, from Beidaihe Institute of Electrical Automation were employed. The DAQ instrument was USB-6259 card manufactured by National Instruments. A typical signal recorded is shown in Figure 3. The sampling rate is 100 KHz and the duration is 10 ms.

Gaussian White Noise is added as noise with zero mean. A noisy version of the signal shown in Figure 3 is depicted in Figure 4. The SNR is 0 dB.

*4.2. The Performance Evaluation of Finding the Optimum Threshold.* In this section, we should validate the presented method's ability of finding the optimum threshold thr$_{\text{opt}}$. Firstly, the signal shown in Figure 4 is decomposed into several monotonic components (IMFs) of distinct time scales as shown in Figure 5 using EEMD method. The parameters used to run the EEMD are the number of ensemble, $M$, and amplitude of the added white noise, $k$, which is set to 100 and 0.2 time standard deviation [12]. Here, we can see fast oscillations in the lower-order IMFs, and slow oscillations as the order of components increases.

Then CMSE was calculated under different value of $m$. Figure 6 shows the variety of CMSE when the $m$ changed from 1 to 20. It is obviously suggested in the figure that the CMSE reaches the minimum value when $m$ gets the value 10.

After that, the thresholding operation as described in (5) was performed and followed by the reconstruction process defined in (6). Figure 7 shows the results. As a comparable result, the purified signals are also given in the figure. It is obviously obtained that the noise signals have been restrained notably and the recovered signals come close to the original signals.

*4.3. Comparison with Other EEMD-Based Methods.* To verify the efficiency of the presented denoised method under different levels of noises, we designed an experiment to compare the results in different SNR with other two EEMD
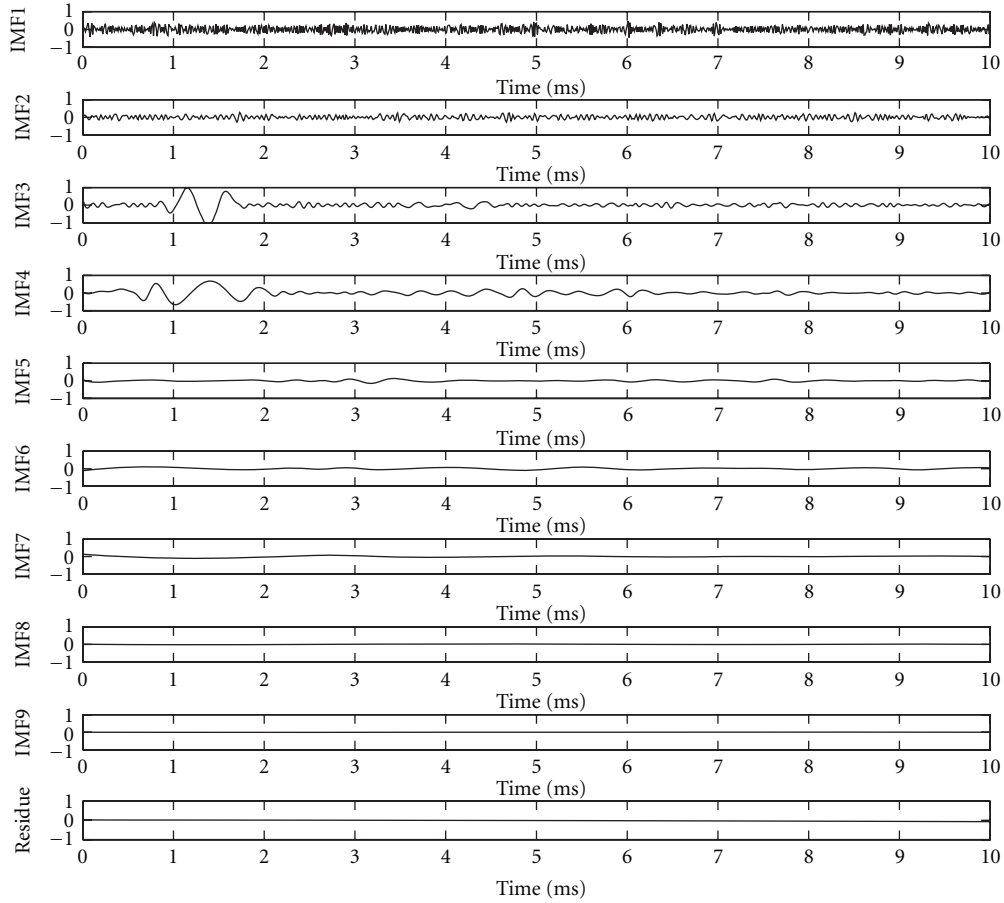
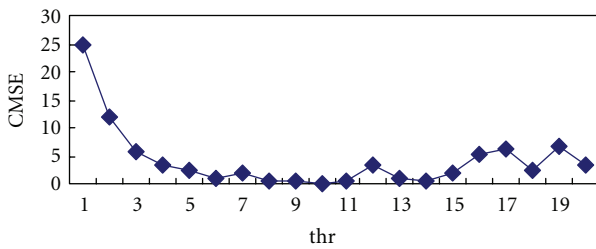FIGURE 5: The decomposition result with EEMD.



FIGURE 6: CMSE versus $m$.



--- Clean signal
— Restored signal

FIGURE 7: The denoised stress wave signal using proposed method.

algorithms. The signals with distinct SNR were first generated and the SNR ranged from $-5$ to 15. The MSE was still selected as the criterion of evaluation. The definition of MSE is described as follows:

$$\text{MSE} = \frac{1}{L} \sum_{k=0}^{L-1} [s(k) - \hat{x}(k)]^2, \tag{9}$$

where $s(k)$ and $\hat{x}(k)$ denote the values of the original signal and restored signal, and $L$ is the length.

The experimental results are shown in Figure 8. The MSE in the figure was gained by computing the average value of 10-time repeat. For comparison, the results of EEMD based low pass filter, EEMD based thresholding filter, and
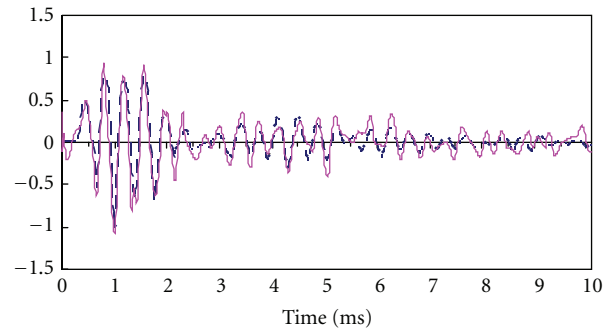
IHP filter [19] were shown in Figure 8. It is evident that the proposed DIHP filter has given a better performance compared to the EEMD based low pass filter and the EEMD based thresholding filter. What should be emphasized here is that the result of DIHP filter is close to that of IHP. It is capable of producing better noise-removal results even in cases where the signal quality is low (SNR value is $-5$ dB). This means that the method is effective for very noisy signals.
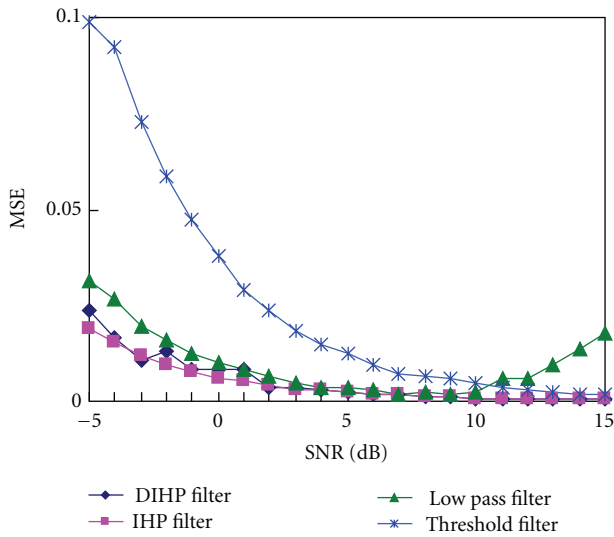
FIGURE 8: MSE obtained with different noise levels by proposed method, IHP filter, EEMD-based low pass filter, and EEMD-based thresholding filter.

It implies we can get the same effect on conditions that any prior knowledge of the target signals is. The proposed method is fully data driven.

## 5. Conclusions

A new technique for the signal enhancement has been proposed and developed. Simulations results of the synthesized signals have expressed the effectiveness of the new algorithm. The technique differs from many EEMD based algorithms as it uses IHP to detect and classify the noise oscillations and utilizes CMSE to compute the optimum threshold adaptively. The filtering method is a fully data-driven approach. As a result, the proposed technique has the ability to be used as a preprocessing step for computerized wood nondestructive test using the stress wave technique.

## Acknowledgments

## References

[1] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, 1995.

[2] S. W. Hsu, C. Y. Chang, Z. Y. Chung, and K. N. Wu, "Improved measurement of dynamic modulation transfer functions on display using pursuit camera method based on wavelet-denoising method," *Optical Review*, vol. 18, no. 1, pp. 153–156, 2011.

[3] A. O. Boudraa and J. C. Cexus, "EMD-based signal filtering," *IEEE Transactions on Instrumentation and Measurement*, vol. 56, no. 6, pp. 2196–2202, 2007.

[4] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hubert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society A*, vol. 454, no. 1971, pp. 903–995, 1998.

[5] P. Flandrin, G. Rilling, and P. Gonçalvés, "Empirical mode decomposition as a filter bank," *IEEE Signal Processing Letters*, vol. 11, no. 2, pp. 112–114, 2004.

[6] A. O. Boudraa, J. C. Cexus, and Z. Saidi, "EMD-based signal noise reduction," *International Journal of Signal Porcessing*, vol. 1, no. 1, pp. 33–37, 2004.

[7] T. Hasan and M. K. Hasan, "Suppression of residual noise from speech signals using empirical mode decomposition," *IEEE Signal Processing Letters*, vol. 16, no. 1, pp. 2–5, 2009.

[8] Y. Kopsinis and S. McLaughlin, "Development of EMD-based denoising methods inspired by wavelet thresholding," *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1351–1362, 2009.

[9] C. S. Qu, T. Z. Lu, and Y. Tan, "A modified empirical mode decomposition method with applications to signal de-noising," *Acta Automatica Sinica*, vol. 36, no. 1, pp. 67–73, 2010.

[10] W. H. Su, C. K. Lee, and C. W. Lee, "Noise-reduction for fringe analysis using the empirical mode decomposition with the generalized analysis model," *Optics and Lasers in Engineering*, vol. 48, no. 2, pp. 212–217, 2010.

[11] P. A. A. Esquef and G. S. Welter, "Audio de-thumping using Huang's empirical mode decomposition," in *Proceedings of the International Conference on Digital Audio Effects*, pp. 19–23, Paris, France, 2011.

[12] 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.

[13] M. T. Lo, V. Novak, C. K. Peng, Y. Liu, and K. Hu, "Nonlinear phase interaction between nonstationary signals: a comparison study of methods based on Hilbert-Huang and Fourier transforms," *Physical Review E*, vol. 79, no. 6, Article ID 061924, 11 pages, 2009.

[14] H. Mahgoun, R. E. Bekka, and A. Felkaoui, "Gearbox fault diagnosis using ensemble empirical mode decomposition (EEMD) and residual signal," *Mechanics and Industry*, vol. 13, no. 1, pp. 33–44, 2012.

[15] Y. Wei and M. C. Chen, "Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks," *Transportation Research C*, vol. 21, no. 1, pp. 148–162, 2012.

[16] Y. Zhang and Z. H. Xie, "Ensemble empirical mode decomposition of impact-echo data for testing concrete structures," *NDT & E International*, vol. 51, pp. 74–84, 2012.

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

[18] K. M. Chang and S. H. Liu, "Gaussian noise filtering from ECG by Wiener filter and ensemble empirical mode decomposition," *Journal of Signal Processing Systems*, vol. 64, no. 2, pp. 249–264, 2011.

[19] Y. M. Fang, H. L. Feng, J. Li et al., "Stress wave signal denoising using ensemble empirical mode decomposition and

an instantaneous half period model," *Sensors*, vol. 11, no. 8, pp. 7554–7567, 2011.

[20] T. Wang, M. C. Zhang, Q. H. Yu et al., "Comparing the application of EMD and EEMD on time-frequency analysis of seimic signal," *Journal of Applied Geophysics*, vol. 83, pp. 29–34, 2012.

[21] Z. K. Peng, P. W. Tse, and F. L. Chu, "An improved Hilbert-Huang transform and its application in vibration signal analysis," *Journal of Sound and Vibration*, vol. 286, no. 1-2, pp. 187–205, 2005.

[22] S. Bozhang and R. F. Pellerin, "Nondestructive evaluation of the degree of deterioration in wood: stress wave frequency spectrum analysis," in *Proceedings of 10th International Symposium on Nondestructive Testing of Wood*, pp. 99–115, Presses Polytechniques et Universitaires Ro-mandes, Lausanne, Switzerland, 1996.

[23] M. Jahed, B. Najafi, A. Khamene, and S. J. Lai-Fook, "Time delay calculation of stress waves using wavelet analysis—application in canine edematous lungs," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'97)*, pp. 2141–2144, April 1997.

[24] M. P. Hayes and J. Chen, "A portable stress wave measurement system for timber inspection," in *Proceedings of the Electronics New Zealand Conference*, pp. 1–6, Hamilton, New Zealand, 2003.