A Comparative Study of WPD and EMD for Shaft Fault Diagnosis

Zhiqiang Huo[1,2], Yu Zhang[1], Lei Shu[2,1*]

[1]School of Engineering, University of Lincoln, Lincoln, UK

[2]Guangdong Provincial Key Laboratory on Petrochemical Equipment Fault Diagnosis,
Guangdong University of Petrochemical Technology, Maoming, China

Email: {zhuo, yzhang, lshu}@lincoln.ac.uk

Abstract—Fault diagnosis of incipient crack failure in rotating shafts allows the detection and identification of performance degradation as early as possible in industrial plants, such as downtime and potential injury to personnel. The present work studies the performance and effectiveness of crack fault detection by means of applying wavelet packet decomposition (WPD) and empirical mode decomposition (EMD) on fault diagnosis of rotating shafts using multiscale entropy (MSE). After WPD and EMD, the most sensitive reconstruction vectors and intrinsic mode functions (IMFs) are selected using Shannon entropy. Then, these feature vectors are fed into support vector machine (SVM) for fault classification, where the entropy features represent the complexity of vibration signals with different scales. Experimental results have demonstrated that WPD combined with MSE can achieve an accuracy of 97.3% for crack fault detection in rotating shafts, whilst EMD combined with MSE has shown a higher detection rate of 98.5%.

Index Terms—Shaft fault diagnosis, Wavelet packets decomposition, Empirical mode decomposition, Multi-scale entropy, Support vector machine.

I. INTRODUCTION

In industrial plants, crack fault is one of commonly failures which is defined as any unintentional discontinuities in the shaft material. Consequences of shaft failure can lead to enormous costs in downtime, damage to equipment and potential injury to personnel. Consequently, real-time monitoring and diagnosis of rotating shafts is essential to detect and prevent performance deterioration in rotating shafts, which enables to increase the confidence of safety and reliability of the process system. For this purpose, continuing efforts have been dedicated to reducing magnitude and redundant information in original signals by using signal processing approaches. Fast Fourier transform and short time Fourier transform are two classical approaches for signal decomposition; however their performances are limited by the finite window size which is not suitable to analyze non-stationary and non-linear vibration signals. Instead, wavelet transform and empirical mode decomposition (EMD) have shown their effectiveness in providing high resolution in both time and frequency domain, which have been successfully applied in the field of fault diagnosis of shafts, such as continuous wavelet transform coefficients were used in works [1]–[5]. Apart from that, wavelet packets

decomposition (WPD) benefits from effectively decomposing frequency bands into detail and approximate coefficients with multi-levels [6], which has been applied in [7], [8]. The advent of EMD, proposed by Huang and Wu, provides a powerful self-adaptive signal processing method for analyzing non-stationary and non-linear signals by decomposing the signals into a set of intrinsic mode functions (IMFs) [9]–[11], which are determined by the signal itself rather than the pre-defined kernels compared with wavelet analysis. Generally, after signal decomposition, desired coefficient vectors and IMFs can be further selected to characterize fault symptoms hidden in machine's signals measured from different conditions of rotating shafts.

Entropy analysis has been successfully applied in the field of fault diagnosis in the recent decade, the concept of which takes into account the probability of measuring state information with respect to underlying probability distribution. In mechanical systems, the occurrence of failures in rotating components changes their energy distributions and frequency characteristics, which results in increasing harmonic frequency components. Hence, entropy methods can be used in the step of feature extraction to discriminate fault information in various conditions of rotating shafts as appropriate indicators. As an example, Shannon entropy and sample entropy have been successfully applied in the field of fault diagnosis of rotating components [12]–[15]. Nevertheless, monitoring signals measured from rotating machinery usually contain multiple intrinsic oscillatory modes produced by interaction or coupling effects between rotating components, which may lead to more non-linear and non-stationary signals. Consequently, to overcome this limitation, multiscale entropy (MSE), proposed by Costa et al. [16], was developed to analyze the time series from multiple scales through the coarse-grain process based on the concept of sample entropy. It is needed to point out that although MSE has been applied in diagnosing components in rotating machinery [17], [18]. However, there is still a need to further study the performance of multiple-scale analysis by combining WPD and EMD with MSE techniques respectively in crack fault diagnosis of rotating shafts.

In an attempt to detect and identify crack fault in rotating shafts, this paper investigates and compares the effectiveness of WPD and EMD methods using MSE respectively. In this

study, original vibration signals are first decomposed using WPD and EMD. After that, the most sensitive reconstruction vectors and IMFs are selected using Shannon entropy criteria. Afterwards, MSE was then adopted to characterize fault symptoms and establish feature vectors, which are then fed into support vector machine (SVM) for fault classification, where the entropy features represent the complexity of vibration signals with different scales. The main contributions in this paper are concluded as follows:

- The effectiveness of WPD and EMD based MSE approaches were studied respectively for crack fault diagnosis of rotating shafts.
- Useful vectors containing dominant fault information can be effectively selected with the criteria of Shannon entropy. The most high accuracies were obtained using decomposed vectors in accordance with the order of vectors whose entropy values are greater.
- Experimental results have shown that, the number of wavelet vectors that achieved accuracies over 80% is larger than that of IMFs generated from EMD, whilst the later achieved highest detection rate.

The rest of this paper is organized as follows: Section II presents the proposed diagnostic methods using WPD and EMD based on MSE and SVM. Experimental validation and results are given in Section III. Finally, conclusions are drawn in Section IV.

II. PROPOSED METHOD

Aiming at the characteristics of vibration signals obtained from rotating shafts, in this study, a hybrid method was proposed to investigate and compare the effectiveness of WPD and EMD based on MSE applied for crack fault diagnosis of shaft in rotating machinery. Fig. 1 presents the procedure for crack fault classification on the basis of this proposed approach. WPD and EMD are first used to decompose original signals and obtain a set of reconstruction vectors and IMFs. Afterwards, the first m with the largest energy in those decomposed vectors are separately selected using Shannon entropy criteria. After that, MSE is applied to extract features from multi-scale levels using selected vectors or IMFs. Feature vectors are finally fed into SVM for fault classification. The theories behind these techniques are briefly reviewed as following:

A. Wavelet Package Decomposition

The essential function of WPD is the filtering operation that allows the signal $x(t)$ to be respectively separated to approximation and detail through low-pass and high-pass filters. For each partitions, decomposition is repeated until the desired resolution [6]. Let $W(i,j)$ denote the coefficients of the signal in each packet, i the decomposition level, and j the order of the packet within the decomposition level. Consequently, each decomposition vector $W(i,j)$ has the following structure:

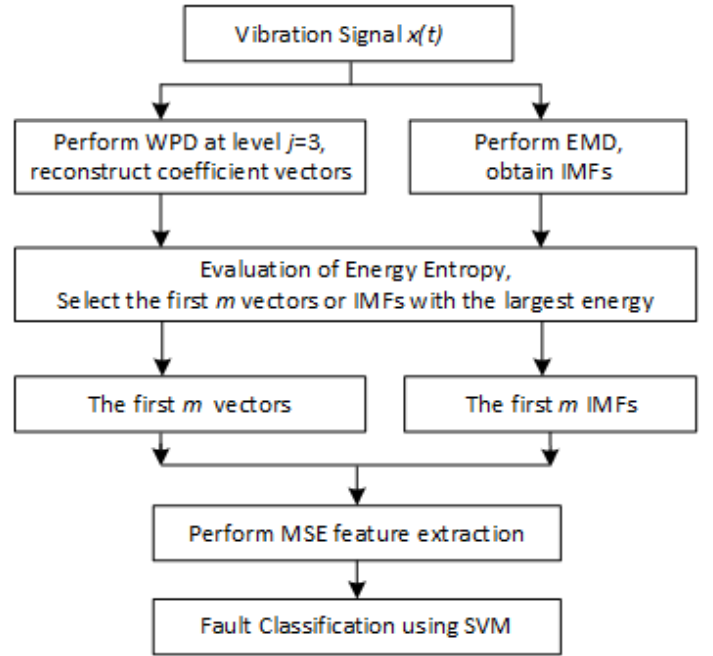$$W(i,j) = \{w(i,0), w(i,1), \cdots, w(i,N-1)\} \quad (1)$$

Fig. 1: Flow chart of the proposed approach.

where the number of packets within k- decomposition level is $N = 2^i$. Therefore, the reconstruction of the signal, denoted as $W(0,0)$, which considers the totality of the packets within a given decomposition level can be obtained. Such as the $W(0,0)$ of the signal from level four can be obtained as following:

$$W(0,0) = \sum_{j=0}^{15} \{w(4,j)\} \quad (2)$$

B. Empirical Mode Decomposition

The EMD method is one of the most powerful methods proposed for analyzing non-stationary and non-linear signal, which can be decomposed into a set of IMFs that are supposed to satisfy the following two conditions [9]:

- in the whole data set, the number of extrema and the number of zero-crossings must either equal or differ at most by one.
- at any point, the mean value of the envelope defined by local maxima and the envelope defined by the local minima is zero.

Steps of decomposing signals using EMD are briefly introduced as following:

Step 1: Identify all the local extrema from signal $x(t)$, and then separately connect all the local maxima and minima to produce upper and lower envelope.

Step 2: The mean value of upper and lower envelope is marked as m_1, a new data series, h_1, can be obtained by subtracting m_1 from the original signal $x(t)$.

$$h_1 = x(t) - m_1, \quad (3)$$

Step 3: Normally, if h_1 does not satisfy the IMF conditions, and then step 1 - 2 can be repeated until the first IMF, c_1, can be obtained.

$$c_1 = h_{1(k-1)} - m_{1(k-1)}, \qquad (4)$$

where k means after k iterations $h_{1(k-1)}$ satisfies IMF conditions, and at this moment the first IMF component, c_1, can be acquired from x_t.

Step 4: Residue r_1 can be obtained by separating c_1 from $x(t)$, that is

$$r_1 = x(t) - c_1, \qquad (5)$$

By repeating the above process until r_n becomes a monotonic function; therefore IMF can no more be extracted. After n times iterations, summing up all IMFs and the final residue r_n, $x(t)$ can be represented by

$$x(t) = \sum_{i=1}^{n} c_i + r_n, \qquad (6)$$

The IMFs $c_i (i = 1, 2, \cdots, n)$ consist different frequency bands from high to low. The frequency components contained in each frequency band are different and they change with the variation of the signal $x(t)$, while r_n represents the central tendency of the signal $x(t)$. A more detailed explanation about EMD can be found in [9].

C. Shannon Entropy

Shannon entropy, denoted as E_i, is applied to select the reconstruction vectors $W(i, j)$ and IMFs with the first finite largest values [20] after WPD and EMD decomposition, which means they may contain more dominant fault information than that with lower values. E_i of each vector and IMF can be obtained by

$$E_i = -\sum_{i=1}^{M} p_i \log p_i \qquad (7)$$

where M is the total number of vectors within j-th level or IMFs obtained from original signals, $\{p_i = E_i/E, i = 1, 2, \cdots, M\}$ is the percent of the power energy in whole vectors or IMFs, and $E = \sum_{i=1}^{M} E_i$ is the sum energy of E_i. Given a time series, S with the length N, p_i can be obtained using $E_i = \sum_{i=1}^{N} |S_k|^2, i = 1, 2, \cdots, N$. Hence, the entropy energy can be respectively obtained from reconstruction vectors and IMFs. After that, a group of wavelet vectors and IMFs with the first m largest values can be finally selected and further applied in the step of feature extraction.

D. Multiscale Entropy

In fault feature extraction, MSE is adopted to represent the complexity of time series in different scales through the coarse-grain process [16]. For the sake of brevity, only major steps of MSE algorithm are discussed as follows:

Step 1: Initialize the embedding dimension m and the

tolerance $r = 0.15 \times$ SD, where m and r are the required parameters to be used to compute MSE, and SD herein means standard deviation of the original data.

Step 2: Segment original time series into several data sets with length τ by using the mean values of the segmented data, and then new series data sets y_j^τ can be obtained.

$$y_j^\tau = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i, 1 \leqslant j \leqslant \frac{N}{\tau}, \qquad (8)$$

where τ is also called the scale factor. Apparently, as τ increases the length of the resulting coarse-grained time series decreases. As a matter of fact, the coarse-grain process is used to remove the high frequency or order components by changing a sliding window of length τ and averaging the data series within the window in a way of non-overlap.

Step 3: Calculate sample entropy for new data sets y_j^τ at each scale through the coarse-grain process. Therefore, the irregularity and complexity of time series in different scales can be obtained.

E. Fault Classification

SVM classifier is adopted for fault classification in this study. A supervisor mode is used for data training and testing using LIBSVM Matlab Toolbox [21]. A grid search method is applied to search best parameters, namely the cost parameter c and the width parameter g in the training phrase. Herein, c and g are respectively set between 2^{-10} to 2^{10}. In addition, after the data training, a 5-fold cross-validation method is used for the validation of the proposed shaft fault diagnosis approach. In a k-fold cross-validation method, the data sets are divided into k subsets, and the holdout method is repeated k times. After that, the average error for all k trials can be obtained to guarantee the reliable efficiency in fault classification using SVM.

III. EXPERIMENTAL STUDY

In this section, the experimental test rig is first introduced, and experimental results of the proposed comparative method are then presented and discussed.

A. Experiment Setup

In this study, PT 500 machinery diagnostic system is applied to collect vibration signals respectively generated from healthy and cracked shafts [19], as shown in Fig. 2. This test rig allows to closely simulate the characteristic behaviour of a shaft with a crack using asymmetrical flange connection. Crack detection in rotating shafts kit is composed of motor assembly, motor control unit, shaft, flange, belt drive kit, computerised vibration analyser. Tightening the flange connection with spacer sleeves gives a connection that is either loose or secure, which can very closely resembles the behaviour of a crack in the shaft. Besides, vibration data was captured at a sampling frequency of 8 kHz under 1500 r.p.m. The representative vibration signals of two conditions of rotating shafts are presented in Fig. 3.
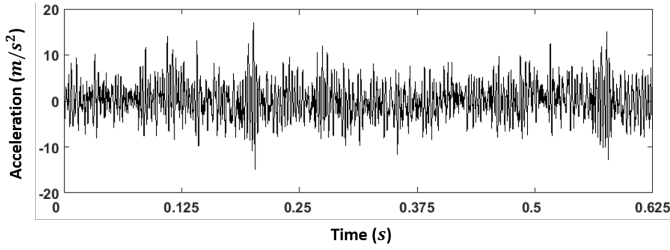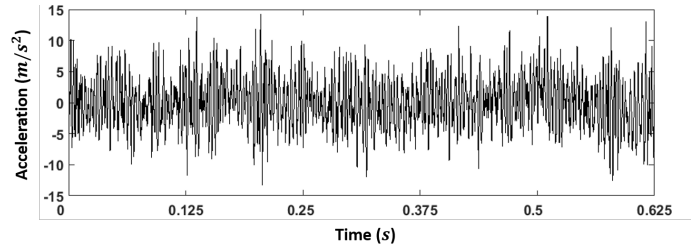
(a) PT 500 test rig
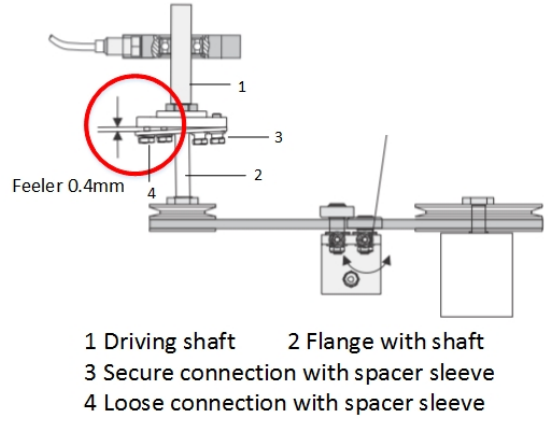
1 Driving shaft   2 Flange with shaft
3 Secure connection with spacer sleeve
4 Loose connection with spacer sleeve

(b) layout plan of test rig [19]

Fig. 2: PT 500 experimental test rig and corresponding layout plan.

(a) Vibration signal obtained from healthy condition.

(b) Vibration signal obtained from faulty condition.

Fig. 3: Representative original vibration signals generated from two conditions of rotating shafts.

The experimental data set contains two conditions of rotating shafts, namely healthy condition and 0.4mm crack condition. For each condition, 190 data sets were used, and therefore the total number of data set includes 380 data sets samples, each of which is a section of vibration signal containing 5000 sampling points. Hence, the entire data set was split into two data sets, namely 190 data sets for training and 190 data sets for testing respectively.

B. Experimental Results

Original signals were first decomposed using WPD and EMD, after which wavelet decomposition vectors and IMFs were obtained respectively. The mother wavelet 'Daubechies 4' and 3-th decomposition level were selected for WPD analysis in this study. After WPD analysis, coefficient vectors were then reconstructed to vectors with the same length of original signals. After that, the most sensitive reconstruction vectors and IMFs were selected using Shannon entropy criteria, which are generally considered that they have prominent fault information [20]. For this purpose, vectors and IMFs with first 8 largest values of Shannon entropy were finally selected and compared in this study, which is illustrated in Fig. 4 (a) and (c). It can be seen that 1-th, 2-th, 4-th reconstruction vectors and 2-th, 3-th, 4-th IMFs contain the largest energy values, which means that they contain the most discriminatory information with respect to fault symptoms

than that with lower values. Interestingly, by comparing Fig. 4 (a) and (b), and Fig. 4 (c) and (d), it is not hard to find that the most high accuracies can be obtained using decomposed vectors in accordance with the order of vectors whose Shannon entropy values are greater. Furthermore, it can be seen from Fig. 4 (b) and (d) that the first four decomposed vectors generally achieved high classification accuracy as well as Shannon entropy values.

In addition, MSE was performed to extract fault features from selected vectors and IMFs in this study. As described in Section II, three parameters, namely the tolerance r, the embedding dimension m, and the scale factor τ were used in the calculation of MSE. A small r leads to a poor conditional probability, whilst a larger r results in losing too much system information. In practice, the value of $r = 0.15 \times$ SD can achieve positive performance for the analysis of irregularity in time series [16]. Furthermore, feature vectors were constructed using MSE values with different scales in purpose of investigating the impact of scale number on the accuracy of fault classification. Specifically speaking, one feature vector has τ features composed of MSE energy values separately extracted from scale 1 to τ. In practice, the number of τ approaching to zero normally depends on the the length of sampling points in one data set. In this study, the largest number of τ is set to 40 since the energy value of MSE scale is almost approaching
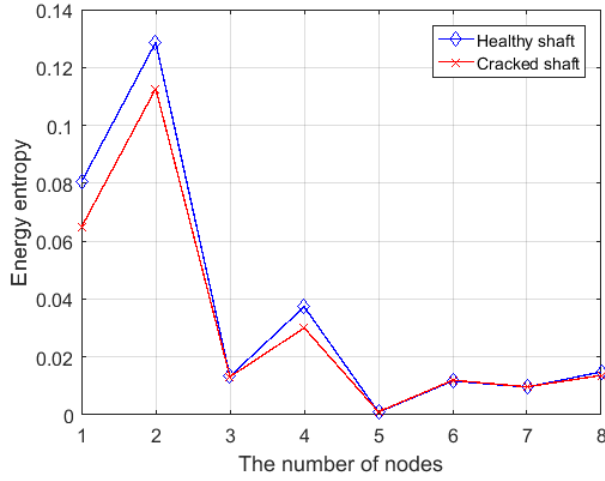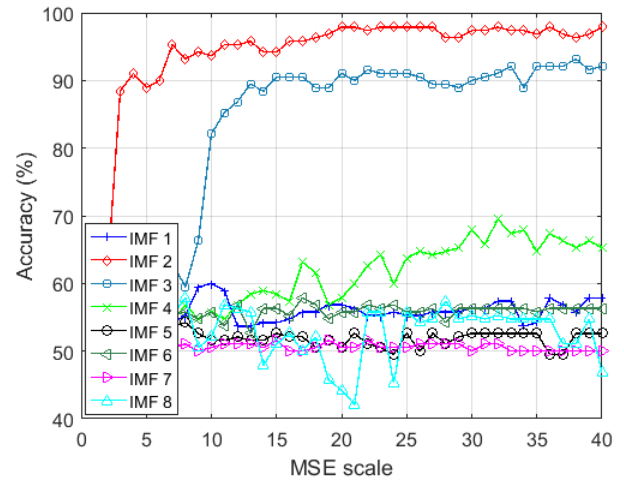
(a) Shannon entropy of each vector after WPD
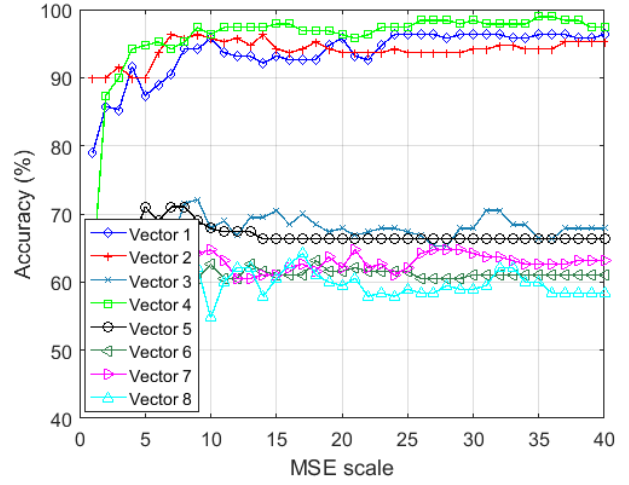
(b) Accuracy results using WPD, MSE and SVM
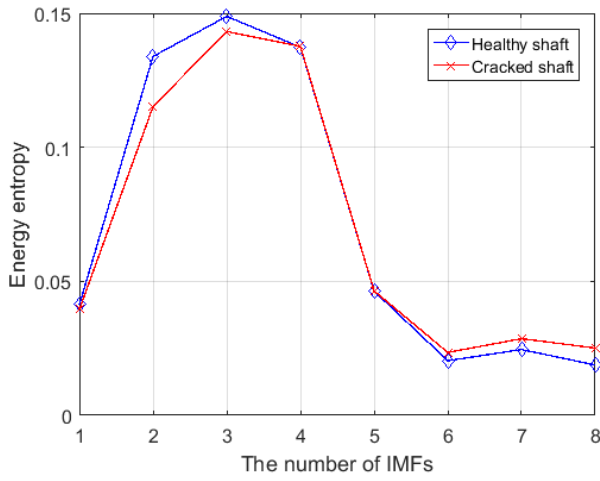
(c) Shannon entropy of each IMF after EMD

(d) Accuracy results using EMD, MSE and SVM

Fig. 4: The first 8 energy of vectors and IMFs and corresponding classification accuracy using MSE with various scales.

to zero when τ is larger than 40. Finally, those feature vectors were then fed into SVM respectively to perform fault classification. Accuracy of the proposed approach resulting from fault classification using selected reconstruction vectors are presented in Fig. 4 (b) and (d). From this figure, it can be found that there are three vectors achieved over 80% accuracies, whilst only two IMFs achieved that comparatively. Moreover, it is not hard to find that the higher classification accuracy can be achieved when the increasing number of MSE scales adopted to construct feature vectors. Table I presents classification accuracy of each reconstruction vector and IMF using two different numbers of MSE values with various scales, namely $\tau = 20$ and $\tau = 40$ respectively. As can be seen from this table, when τ is selected as the largest number, namely 40, the classification accuracy can achieve the best detection rate, which means more discriminatory fault information were characterized in fault classification by estimating complexity in signals; however, it is important to note that this may meanwhile aggravate the calculation burden.

Additionally, 4-th reconstruction vector combined with MSE can achieve an accuracy of 97.3% for crack fault detection in rotating shafts, whilst 2-th IMF combined with MSE has shown a higher detection rate of 98.5%. Apart from that, it can be concluded that lower order of decomposed vectors can hardly to identify crack fault in this study, the reason of which is that, with a belt connected with a shaft, the emerging crack failure in a shaft can increase the number of high coupling frequency components and finally lead to more non-linear and non-stationary characteristics in signals; therefore, vectors with low-frequency may contain less useful discriminatory information for crack fault diagnosis of rotating shafts in this study.

IV. CONCLUSIONS

The paper presents a comparative study on effectiveness and performance of WPD and EMD using MSE as feature indicators for crack fault diagnosis of rotating shafts. Experimental measurements were taken from a fault simulation machine

TABLE I: The classification accuracy using MSE with $m = 2$ and τ ranging from 1 to 20 and 40.

Signal Processing	Scale(τ)	Vector 1	Vector 2	Vector 3	Vector 4	Vector 5	Vector 6	Vector 7	Vector 8	Best Accuracy
WPD	1,2,···,20	95.7%	93.6%	67.8%	96.3%	66.3%	61.5%	62.1%	59.4%	97.3%
	1,2,···,40	96.3%	95.2%	67.8%	97.3%	66.3%	61.0%	63.1.0%	58.4%	
EMD	Scale(τ)	IMF 1	IMF 2	IMF 3	IMF 4	IMF 5	IMF 6	IMF 7	IMF 8	98.5%
	1,2,···,20	56.8%	98.5%	91.5%	57.8%	50.5%	55.8%	50.5%	44.2%	
	1,2,···,40	57.9%	98.5%	92.1%	65.2%	52.6%	56.3%	50.0%	46.8%	

with two conditions of shafts, namely 0.4mm crack failure and healthy condition. The 'Daubechies 4' wavelet kernel and 3-th decomposition level were selected, meanwhile the first 8 reconstruction vectors and IMFs were respectively chosen for the comparative purpose. In this study, the criteria of Shannon entropy was adopted to select desired decomposed vectors with dominant fault information after WPD and EMD decomposition. Experimental results have proved that positive classification accuracies can be achieved by using selected vectors whose entropy values are greater, which are the first 4 decomposed vectors obtained from multi-resolution analysis (i.e., WPD and EMD). Moreover, it was found that WPD has more useful vectors than that of EMD, the classification accuracies of which can achieve more than 80%. In addition, experimental results have demonstrated that WPD combined with MSE can achieve an accuracy of 97.3% for crack fault detection in rotating shafts, whilst EMD combined with MSE has shown a higher detection rate of 98.5%.

ACKNOWLEDGEMENT

REFERENCES

[1] T. R. Babu and A. Sekhar, "Shaft crack identification using artificial neural networks and wavelet transform data of a transient rotor," *Adv. Vib. Eng*, vol. 9, pp. 207–214, 2010.

[2] Y. Zhang, C. Bingham, M. Gallimore, Z. Yang, and J. Chen, "Machine fault detection during transient operation using measurement denoising," in *Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), 2013 IEEE International Conference on*. IEEE, 2013, pp. 110–115.

[3] M. J. Gómez, C. Castejón, and J. C. García-Prada, "Review of recent advances in the application of the wavelet transform to diagnose cracked rotors," *Algorithms*, vol. 9, no. 1, p. 19, 2016.

[4] N. H. Chandra and A. Sekhar, "Fault detection in rotor bearing systems using time frequency techniques," *Mechanical Systems and Signal Processing*, vol. 72, pp. 105–133, 2016.

[5] Z. Huo, Y. Zhang, P. Francq, L. Shu, and J. Huang, "Incipient fault diagnosis of roller bearing using optimized wavelet transform based multi-speed vibration signatures," *IEEE Access*, 2017.

[6] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.

[7] F. Al-Badour, M. Sunar, and L. Cheded, "Vibration analysis of rotating machinery using time–frequency analysis and wavelet techniques," *Mechanical Systems and Signal Processing*, vol. 25, no. 6, pp. 2083–2101, 2011.

[8] C. Castejón, M. J. Gómez, J. C. García-Prada, A. J. Ordõnez, and H. Rubio, "Automatic selection of the WPT decomposition level for condition monitoring of rotor elements based on the sensitivity analysis of the wavelet energy," *Mech. Int. J. Acous. Vib*, vol. 20, pp. 95–100, 2015.

[9] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, 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," in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971. The Royal Society, 1998, pp. 903–995.

[10] Y. Lei, J. Lin, Z. He, and M. J. Zuo, "A review on empirical mode decomposition in fault diagnosis of rotating machinery," *Mechanical Systems and Signal Processing*, vol. 35, no. 1, pp. 108–126, 2013.

[11] Y. Zhang, C. Bingham, Z. Yang, B. W.-K. Ling, and M. Gallimore, "Machine fault detection by signal denoising – with application to industrial gas turbines," *Measurement*, vol. 58, pp. 230–240, 2014.

[12] P. K. Kankar, S. C. Sharma, and S. P. Harsha, "Fault diagnosis of ball bearings using continuous wavelet transform," *Applied Soft Computing*, vol. 11, no. 2, pp. 2300–2312, 2011.

[13] D. Gu, J. Kim, T. Kelimu, S.-C. Huh, and B.-K. Choi, "Evaluation of the use of envelope analysis and DWT on AE signals generated from degrading shafts," *Materials Science and Engineering: B*, vol. 177, no. 19, pp. 1683–1690, 2012.

[14] K. Zhu, X. Song, and D. Xue, "Fault diagnosis of rolling bearings based on IMF envelope sample entropy and support vector machine," *Journal of Information &Computational Science*, vol. 10, no. 16, pp. 5189–5198, 2013.

[15] J. Liang and Z. Yang, "A novel wavelet transform–empirical mode decomposition based sample entropy and SVD approach for acoustic signal fault diagnosis," in *International Conference in Swarm Intelligence*. Springer, 2015, pp. 232–241.

[16] M. Costa, A. L. Goldberger, and C.-K. Peng, "Multiscale entropy analysis of complex physiologic time series," *Physical Review Letters*, vol. 89, no. 6, p. 068102, 2002.

[17] A. K. Verma, S. Sarangi, and M. Kolekar, "Misalignment faults detection in an induction motor based on multi-scale entropy and artificial neural network," *Electric Power Components and Systems*, vol. 44, no. 8, pp. 916–927, 2016.

[18] S. Pan, T. Han, A. C. Tan, and T. R. Lin, "Fault diagnosis system of induction motors based on multiscale entropy and support vector machine with mutual information algorithm," *Shock and Vibration*, vol. 2016, 2016.

[19] "PT 500 machinery diagnostic system," (Date last accessed 31-July-2016). [Online]. Available: www.gunt.de/static/s3680_1.php

[20] Y. Yu, C. Junsheng *et al.*, "A roller bearing fault diagnosis method based on EMD energy entropy and ANN," *Journal of Sound and Vibration*, vol. 294, no. 1, pp. 269–277, 2006.

[21] "LIBSVM Matlab Toolbox," (Date last accessed 31-July-2017). [Online]. Available: https://www.csie.ntu.edu.tw/~cjlin/libsvm/