

Crack Detection in Rotating Shafts using Wavelet Analysis, Shannon Entropy and Multi-class SVM

Zhiqiang Huo^{1,2}, Yu Zhang¹, Zhangbing Zhou³, and Jianfeng Huang²

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

² Guangdong University of Petrochemical Technology, Maoming, China,

³ China University of Geosciences, Beijing, China,

{zhuo, yzhang}@lincoln.ac.uk, zbzhou@cugb.edu.cn, jianfeng.huang@outlook.com,

Abstract. Incipient fault diagnosis is essential to detect potential abnormalities and failures in industrial processes which contributes to the implementation of fault-tolerant operations for minimizing performance degradation. In this paper, an innovative method named Self-adaptive Entropy Wavelet (SEW) is proposed to detect incipient transverse crack faults on rotating shafts. Continuous Wavelet Transform (CWT) is applied to obtain optimized wavelet function using impulse modelling and decompose a signal into multi-scale wavelet coefficients. Dominant features are then extracted from those vectors using Shannon entropy, which can be used to discriminate fault information in different conditions of shafts. Support Vector Machine (SVM) is carried out to classify fault categories which identifies the severity of crack faults. After that, the effectiveness of this proposed approach is investigated in testing phrase by checking the consistency between testing samples with obtained model, the result of which has proved that this proposed approach can be effectively adopted for fault diagnosis of the occurrence of incipient crack failures on shafts in rotating machinery.

Key words: Fault diagnosis, Shaft, Continuous wavelet transform, Shannon entropy, Multi-class SVM

1 Introduction

Fault diagnosis of rotating machinery has been drawn a great deal of attentions for meeting the increasing demands on reliability and safety of industrial systems subjected to latent process and component failures due to complex working conditions (e.g., increasing speed, power and load). As a result, significant development in operating stress can lead to premature shaft failures in rotating machinery which may cause performance degradation and even dangerous situations [1]. Hence, it is paramount to continuously monitor and diagnose conditions of rotating shafts and finally enhance the confidence of industrial system reliability and safety. In recent decades, vibration monitoring has been widely

applied in the field of fault diagnosis in rotating machinery [2]. Continuing efforts has been dedicated to reducing magnitude and redundant information in original signals by means of using advanced techniques of signal processing [3]. 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. To solve this problem, the advent of Wavelet Transform (WT) provides a powerful technique for characterizing a variety of conditions from both time and frequency domains. To be more specific, WT enjoys good time and poor frequency resolutions at high frequencies, and good frequency and poor time resolutions at low frequencies. Therefore, WT has been successfully applied for fault detection and diagnosis of rotating machinery [4, 5]. Among wavelet analysis, Continuous Wavelet Transform (CWT), Discrete Wavelet Transform (DWT), Wavelet Packets Transform (WPT) are widely used for decomposing signals into specific frequency spectrum ranges of interest. For analyzing crack faults on shafts, CWT coefficients were applied in [6], [7] by analyzing sub-critical peaks and phase angles. Crack depth and growth were analyzed and identified using envelop analysis and DWT based on acoustic emission in [8]. Apart from CWT and DWT, WPT was also adopted with time-frequency features for crack detection in [9–11], which benefits from appropriately decomposing frequency spectrum into the same bands with detail and approximate coefficients.

Generally, in signal-based fault diagnostic approaches, features of monitored signals are extracted to reduce redundant information and analyze corresponding patterns, which can be in time and/or frequency domains, and nonlinear features [12]. Among those, entropy analysis is an appropriate and widely used indicator that enjoys advantages in discriminating changes in different conditions of shafts by quantifying the uncertainty in times series. One of the mostly encountered feature is Shannon Entropy (*ShanEn*) [13], which is a quantitative measure of unpredictability and irregularity in a system. It only takes into account that the probability of observing a specific event, in fault diagnosis, which is the underlying signal probability distribution used for representing fault characteristics generated from the change of different conditions of shafts.

In this paper, a Self-adaptive Entropy Wavelet (SEW) approach is proposed for fault diagnosis of rotating shafts with breathing cracks by using CWT, *ShanEn* and multi-class support vector machine (SVM). An optimized wavelet kernel, based on impulse modelling, is applied for decomposing signals into multi-scale coefficients vectors, which can be obtained using Particle Swarm Optimization (PSO) and quasi-Newton minimization algorithms in training step for achieving highest fault classification accuracy. After that, *ShanEn* features are extracted from coefficient vectors, and then fed into SVM for training diagnostic pattern, which is carried out for fault identification and classification. The main contributions in this paper are concluded as follows:

- An optimized wavelet model named SEW was presented for crack fault detection and fault diagnosis of rotating shafts by means of applying techniques of

- CWT, *ShanEn*, and SVM. The experimental results have shown that 99.3% fault classification accuracy can be achieved by using this proposed approach.
- The effectiveness of feature vectors obtained from coefficient vectors having different scales on classification accuracy is evaluated and analyzed. In addition, the classification performance of SVM with four kinds of kernel functions based the proposed method are also investigated.

The rest of this paper is organized as follows: Section 2 briefly summarizes related work on shaft fault diagnosis based on wavelet analysis. Section 3 presents the proposed approach for crack fault detection and identification on rotating shafts. Experimental validation and results are given in Section 4. Finally, conclusions are drawn in Section 5.

2 Related Work

Advances in wavelet have paved the way for the emergence of plentiful techniques to be applied in the field of fault diagnosis. In recent decades, great efforts have been devoted to fault diagnosis of rotating shafts in industrial machinery. A detection method based on CWT was proposed by Babu et al. [6], in which the amplitudes of sub-critical and critical peaks of the CWT coefficients are considered as the crack indicators which were then fed as input to Artificial Neural Network (ANN) for crack identification. Nagaraju et al. [7] presented a shaft diagnosis method based on 3D wavelet, which considers both time and frequency features of the crack faults. To be more specific, this method considers both phase angles of different frequency components and sub-critical peaks in a CWT plot, the experiment of which showed that this method performs better than when only sub-critical is considered in the case. In addition, apart from been directly applied in signal decomposition, DWT and WPT are also adopted as preprocess techniques in signal analysis, such as band-pass filters and wavelet de-noising tools. DWT was applied in purpose of detecting crack growth on a shaft in [8], in collaboration with envelope analysis and acoustic emission monitoring. In addition, a method combining WPT and Empirical Mode Decomposition (EMD) was proposed by Bin et al. [9] to precisely characteristic features of crack defects in specific frequency bands. Similarly, features related to WPT have been applied in [10], [11]. By using WPT, it is great convenient to decompose a signal into approximation and detail information having the same frequency resolution. On the other hand, WPT can be applied for de-noising signals by means of either reconstructing coefficient vectors or thresholding methods. Table I summarizes the related work on fault diagnosis of shafts using wavelet analysis in rotating machinery.

To sum up, wavelet analysis is one of the most powerful methods of signal processing to be used for effectively detecting and identifying crack faults in rotating shafts. In practice, the selection of wavelet kernels, decomposition levels, feature vectors, and classification methods has great influences on the performance of fault identification and classification. Hence, there is still strong need

Table 1: Comparison of related work for crack fault diagnosis rotating shafts using wavelet techniques.

	Nagaraju et al. [7]	Gu et al. [8]	Bin et al. [9]	Gómez et al. [10]	Gómez et al. [11]
Year	2009	2012	2012	2016	2016
Objects considered	crack depth and position	crack depth and growth	10 types of rotor failures	crack depth and growth	crack depth
Wavelet methods used	CWT	DWT and HHT	WPD and EMD	WPD	WPD
Parameters applied	Morlet, scale 30	Daubechies 8, level 1 to 5	Daubechies 8, node(4,3)	Daubechies 8, level 9	Daubechies8, level 2 to 9
Features selected	phase angles difference between transverse vibrations	peak, mean, RMS, entropy estimation	bandwidth energy of vibration signal spectrum	mean, RMS	power energy
Classification methods	ANN	N/A	ANN	N/A	ANN
Signals monitored	vibration	acoustic emission	vibration	vibration	vibration

to provide innovative approaches to optimize diagnostic procedures and improve the reliability of fault diagnostic approaches on the basis of wavelet analysis. For this purpose, in this study, a new diagnostic method is proposed and investigated by combining optimized wavelet function, Shannon entropy and SVM for identifying incipient breathing crack failures on rotating shafts.

3 Proposed Method

Aiming at fault diagnosis of rotating shafts with crack defects, in this study, a new data-driven model was proposed named SEW approach by combining optimized wavelet function using impulse modelling based CWT, *ShanEn* and SVM techniques. A supervisor model is used for training and testing SEW model, which are applied respectively for optimizing SEW model and checking the consistency between testing samples and obtained pattern. In the former step, PSO and quasi-Newton minimization algorithms are applied to obtain optimized wavelet kernel function. Fig. 1 presents the procedure for crack fault detection and identification using the proposed approach.

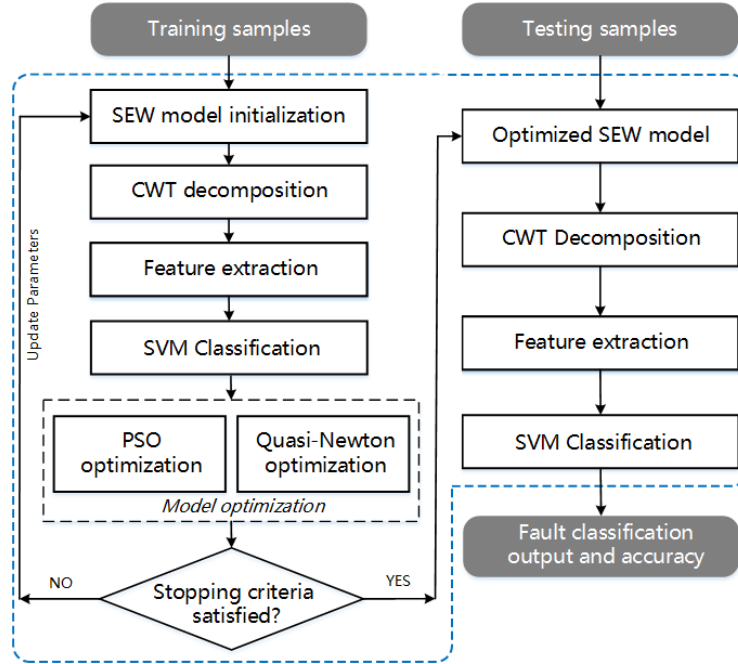


Fig. 1: Flow chart of proposed diagnostic approach.

3.1 Self-adaptive Entropy Wavelet (SEW)

Wavelet analysis is one of the most powerful signal processing techniques which enjoys high resolutions in both time and frequency domains [4], which enjoys the ease of interpretation at the cost of saving space using groups of non-orthogonal wavelet frames to generate general symptoms. The CWT wavelet transform is defined as follows:

$$CWT_x^\psi(\gamma, s) = \Psi_x^\psi(\gamma, s) = \frac{1}{\sqrt{s}} \int x(t) \Psi\left(\frac{t-\gamma}{s}\right) dt. \quad (1)$$

where $x(t)$ is the signal, s is the scale factor, γ is the translation parameter, $\Psi(t)$ is the wavelet kernel function, and it is also called the mother wavelet.

In practice, when the rotating shaft with cracked defects is operating, the resonance or harmonic frequency components corresponding to base frequency are most likely to appear, which show the correlation with the severity of crack defect to some extent. Therefore, these kinds of changes in harmonic frequency can be captured with the help of analyzing impulse models in which differences corresponding to the change of frequency in waveform can be detected. To illustrate the effectiveness of this change case that can be effectively used for breathing crack defects in shaft, SEW is proposed in this study. The response of the system to the instant δ -impulse in vibrodiagnostics can be represented using

a pattern depicted as a response of the single-degree-of-freedom-system, which can be formulated as follows [14]:

$$f(x) = \alpha e^{-\beta x} \cos(wx + \varphi) \quad (2)$$

where $f(x)$ is the displacement, α is the premier amplitude, w is the resonance frequency, which is the frequency of the system fluctuation without resistance. Taking assumption that at the impulsive start the system was at rest into consideration, Eqs. (2) can be applied as a mother wavelet in CWT, which can be expressed as follows:

$$\Psi(x) = \sin(\alpha x + \beta) e^{-\gamma|x|} \quad (3)$$

For keeping minimum parameters in the mother wavelet, α , β , and γ are kept to represent a system's working state based on the theory of impulse modelling; therefore, in this study, Eqs. (2) is selected as a wavelet kernel for helping to detect and diagnose the changes bringing in vibration signals by the occurrence of harmonic frequency. By using CWT, a signal can be respectively decomposed into numbers of coefficient vectors with different scales. Then the power energy of each vector, denoted as E_i , can be obtained from following form:

$$E_i = \int_{-\infty}^{+\infty} |c_i(t)|^2 dt \quad (4)$$

where $c_i(t)$ is the coefficient values of a signal with the i -th scale obtained from a vibration signal. Hence the total energy, $E = \{E_1, E_2, \dots, E_m\}$ (m represents the total number of vectors, namely the first m vectors), which generates energy distribution in different frequency domains of rotating shafts. Then the corresponding Shannon entropy can be defined as:

$$H_i = - \sum_{i=1}^m p_i \log p_i \quad (5)$$

where H_i is the *ShanEn* value of i -th coefficient vector in total m vectors with different scales obtained from CWT method, $\{p_i = E_i/E, i = 1, 2, \dots, m\}$ is the percent of the power energy in the total energy E . Therefore, feature vector, denoted as F , can be constructed with the first m *ShanEn* features, $\{H_i(i = 1, 2, \dots, m)\}$. After that, F can be then fed into SVM for fault classification. During this step, fitness function can be formed to optimize the proposed wavelet model by locating best solutions that can achieve the best accuracy. In this study, desired classification accuracy can be achieved when the values of fitness approach to minimized solutions by means of defining the fitness function as a negative number of resulting accuracy obtained from fault classification using SVM classifier.

3.2 SEW Model Optimization

In the step of parameter optimization in proposed SEW model, PSO and quasi-Newton minimization algorithms are used to locate desired parameters

that can achieve the best classification accuracy. The former is first applied to locate the global best solutions and the later is applied to locate local solutions on the basis of parameters obtained from the former. PSO, proposed by Kennedy and Eberhart [15], performs by iteratively using a population (called a swarm) of candidate solutions (called particles) in the search-space. The swarm consists of m numbers of particles, each of which has own velocity $v_{i,j}(t)$, current position $x_{i,j}(t)$, and local best known position $pbest_j(t)$ ($i = 1, 2, \dots, m; j = 1, 2, \dots, n$). Each particle moves towards own best previous position and the best known positions found by other particles $gbest_j(t)$ in the search-space, which is expected to move the swarm toward the best solutions. The standard PSO can be performed according to the following equations:

$$v_{i,j}(t+1) = w \times v_{i,j}(t) + c_1 \times r1() \times (pbest_j(t) - x_{i,j}(t)) + c_2 \times r2() \times (gbest_j(t) - x_{i,j}(t)) \quad (6)$$

$$x_{i,j}(t+1) = x_{i,j}(t) + v_{i,j}(t+1) \quad (7)$$

where j is the n -th dimension of a particle ($1 \leq j \leq n$), the velocity is restricted to the $[-v_{max}, v_{max}]$ range, $r1()$ and $r2()$ are random numbers in the range of $[0,1]$, $c1$ and $c2$ are positive constants corresponding to personal and social learning factors, and w is the inertia weight. In this study, the initialize parameters with respect to the size of swarm, inertia weight, maximum number of iterations are selected as follows: swarm size $p = 20$, $c1 = 2$, $c2 = 2$, max stall iterations $t_{max} = 6$. The search range of α, β, γ is from $[0, 0, 0]$ to $[200, 200, 200]$.

After global optimization, quasi-Newton minimization algorithm [16], [17] is used to locate minimum solutions around the parameters obtained from PSO, which applies curvature information at each iteration to formulate a quadratic model problem, which has the following form:

$$\min_{x \in R^n} f(x) = \frac{1}{2} x^T H x + b^T x + c \quad (8)$$

where H , the Hessian matrix, is a positive definite symmetric matrix, b is a constant vector, and c is a constant. In this study, Broyden-Fletcher-Goldfarb-Shanno (BFGS) method is applied to efficiently update an approximation to H , which aims to reduce calculation so that a mount of computing time consumption can be saved. For Hessian updating, the BFGS is generally considered as an effective method that can be used for iteratively optimizing the search direction. In BFGS, the formula for generating an approximation to H is described below:

$$H_{k+1} = H_k + \frac{q_k q_k^T}{q_k^T s_k} - \frac{H_k s_k s_k^T H_k^T}{s_k^T H_k s_k} \quad (9)$$

At each major iteration, k , a line search is performed to locate the best solution in the direction:

$$d_k = -H_k^{-1} \cdot \nabla f(x_k) \quad (10)$$

This method has optimal solutions when the partial derivatives of x approach to zero. In this paper, BFGS based quasi-Newton unconstrained minimization serves to locate local optimized parameters after global optimization. The max number of iterations, t_{max} , is set to 80, and max number of function evaluations is 300.

3.3 Fault Classification

Multi-class SVM is adopted for fault classification in this study. A supervisor mode is used for data training and testing using LIBSVM Matlab Toolbox [18, 19]. A grid search method is applied to search best parameters, namely the cost parameter c and the width parameter g in the training phrase. Here, 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 rotating shaft fault diagnostic 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 efficiency of fault classification.

On the basis of above mentioned techniques, the process of the proposed approach is briefly described as follows:

Step 1: collect vibration signals from different conditions of rotating shafts and separate data into training data sets and testing data sets respectively.

Step 2: signals are normalized to make the signals comparable regardless of differences in magnitude using the following equation:

$$X_i = \frac{x - \bar{x}}{\sigma} \quad (11)$$

where x is the i -th element of the signal, \bar{x} and σ are the mean and standard deviation of elements in the given signal respectively.

Step 3: initialize three parameters (α, β and γ), and then decompose vibration signals using this proposed SEW model.

Step 4: extract entropy features using SE from the first m coefficient vectors corresponding to a given signal and construct a feature vector F with the energy as element according to Eqs. (5):

$$F = [H_1, H_2, H_3, \dots, H_m] \quad (12)$$

Step 5: carry out the training and testing procedure of SVM. After that, the testing accuracy of fault classification using current SEW model and SVM can be obtained, the negative value of which can be formed as fitness function. If this is not the best solution, then take next round of optimization. Otherwise, stop optimization and output SEW and SVM models.

Step 6: perform PSO optimization by iterating step 3-5 and updating the global best solutions having the minimum fitness value, which means high

classification accuracy.

Step 7: update SEW model with global best solutions obtained from PSO optimization, and then quasi-Newton minimization algorithm is used to find local minimums of fitness function and obtain local desired solutions by repeatedly perform step 3-5. After PSO and quasi-Newton optimization, SWE model with optimized α, β and γ can be adopted as wavelet kernel with SVM model obtained for fault diagnosis of rotating shafts using the proposed approach.

4 Experimental Study

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

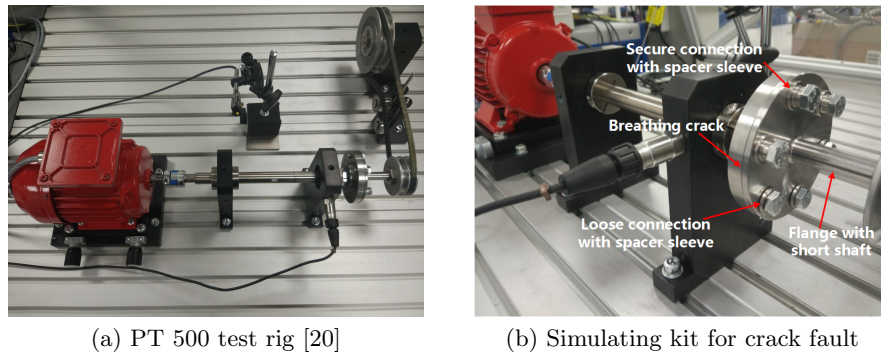


Fig. 2: PT 500 experimental test rig and corresponding crack detection in rotating shaft kit.

4.1 Experimental Setup

In this study, PT 500 machinery diagnostic system [20] is used to collect vibration signals generated from healthy and cracked shafts respectively, as shown in Fig. 2. The fault kit applied is composed of motor assembly, motor control unit, shaft, flange, belt drive kit, and computerised vibration analyser. The control unit is used to collect speed and horsepower data. The piezo-electric sensor and measurement amplifier are used for acceleration measurement. Besides, computerised vibration analyser is used to measure the horizontal vibration signals. This test rig allows to simulate the characteristic behaviour of a shaft with a crack using asymmetrical flange connection. 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 breathing crack in the shaft.

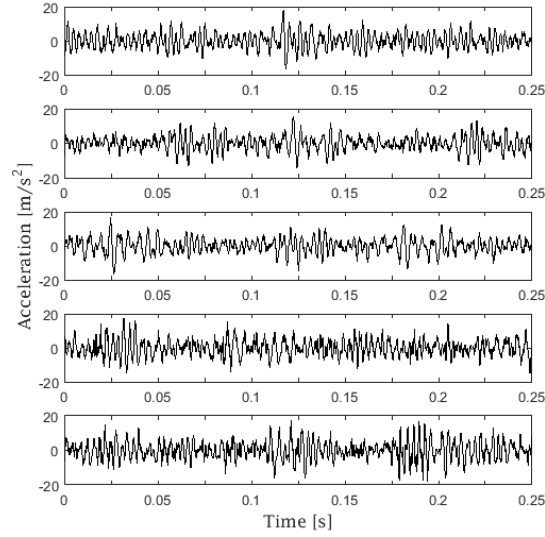


Fig. 3: Typical original vibration signals with five conditions.

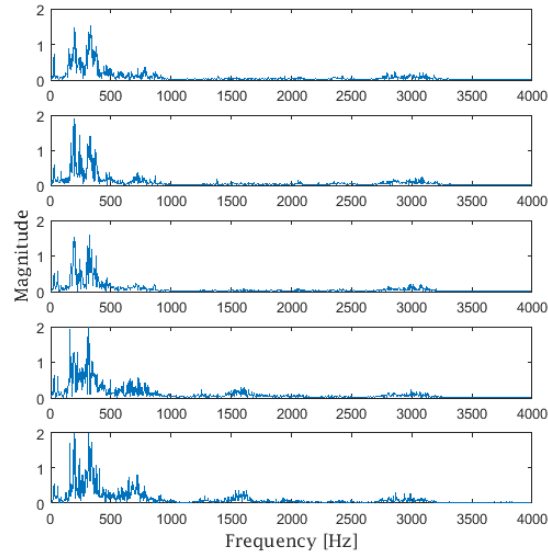


Fig. 4: Frequency spectrums obtained from vibration signals using FFT

In this experiment there are in total five conditions of shaft were considered (denoted as H, C1, C2, C3 and C4) for evaluating and recognizing the health conditions of rotating shafts with breathing crack defects. To be more specific, 0, 1, 2, 3, 4 numbers of loose screws were used respectively to form loose connection and finally closely simulate the breathing cracks in the rotating shafts.

For each condition, 300 data sets were used; therefore the total number of data sets includes 1500 data sets, each of which is a section of vibration signal containing 2000 sampling points. Hence, the entire data set was split into two data sets, namely 750 data sets for training and 750 data sets for testing respectively. Vibration data was captured at a sampling frequency of 8 kHz at 1850 r.p.m. In this experiment, the basic frequency is 30 Hz.

4.2 Results and Analysis

Typical vibration signals with five conditions (i.e., H and C1 - C4) generated from the rotating shaft are presented respectively in Fig. 3, the corresponding frequency spectrum of which using FFT are shown in Fig. 4. As can be seen, resonance frequencies gradually appear with the increasing depth of breathing cracks. For instance, the magnitude of 60 Hz in crack conditions are apparently greater than that in H condition.

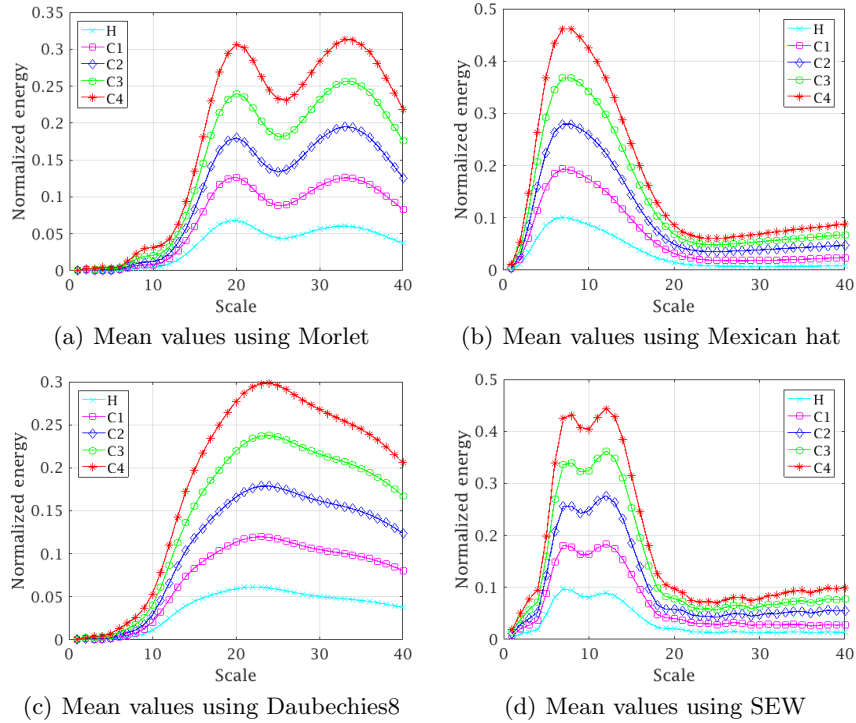


Fig. 5: Mean energy values of coefficients with the first 40 scales using four wavelets.

After vibration data acquisition, the proposed methodology, described in Section II, was implemented to investigate the performance of wavelet kernel and

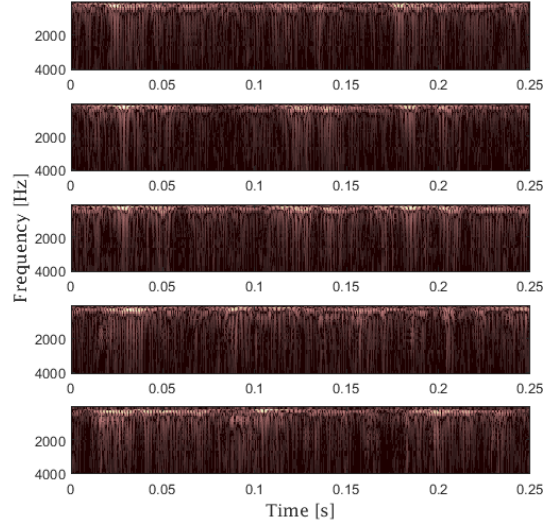


Fig. 6: Time-frequency spectrums of five conditions of rotating shaft using optimized SEW transformation.

feature extraction using different scale numbers. For this purpose, the prediction accuracy using six wavelet kernels on the shelf (e.g., Morlet, Mexican hat, and Daubechies) and proposed SEW wavelet kernel were evaluated and compared in ability of discriminating fault information among various conditions of shafts. Besides, taking optimization time consumption into consideration, the coefficient vectors with the first 40 scales were applied to evaluate the influence of scale numbers on classification accuracy. Firstly, after CWT decomposition, the performance of feature extraction using three existing functions and the proposed SEW function is compared and presented in Fig. 5. It can be seen that mean values of features representing five conditions can be clearly recognized and distinguished by using Shannon entropy. Interestingly, in the right two figures, the mean values of features sharply rise to the submit around 10 scale number, and then gradually decrease and finally keep steady. After 20, the normalized energy of each condition is no more than 0.1 value. In addition, the analysis of time-frequency domains of vibration signals generated from five conditions of rotating shafts using CWT is presented in Fig. 6. From this figure, it can be clearly found that resonance and high frequencies are increasingly apparent, the possible reason of which is that growing depths of crack on the shaft lead to more vibrations that are mainly generated by rotating belt, which is closely connected to the rotating shaft.

Finally, as can be seen in Fig. 7, in comparison with other wavelets, the proposed SEW (represented by red color) is capable of achieving greater accuracy, the average classification accuracy of which is 96.82%. Interestingly, Morlet and Mexican hat enable to obtain more than 90% classification accuracy when

20 and greater number of scales are applied. However, the Daubechies wavelet types achieve minimum accuracies, the best of which is no more than 82% in this study.

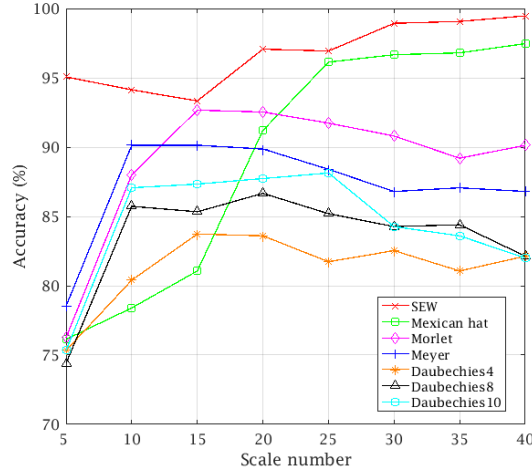


Fig. 7: Comparative prediction accuracies using four kernel functions in SVM for fault classification

To investigate the performance of SVM model used in proposed approach on prediction accuracy, in this study, four different kernel functions (i.e., liner, polynomial, Radial Basis Function (RBF) and sigmod functions) are applied respectively for obtaining optimized parameters in the training phrase for fault classification, the results of which are presented in Fig. 8. As can be seen, the SVM classifier enables to identify crack defects with greater than 90% accuracy when at least 10 scales are applied for fault diagnosis of shaft using the proposed approach. Moreover, it can be notice that the accuracies of fault classification by using four kernel functions gradually ascend following with increasing scale numbers. Interestingly, four lines representing accuracies intersect at 20 that means those four models basically can achieve similar performance when the scale number is 20.

To sum up, through experimental validation, it has proved that the proposed diagnostic approach can be effectively used for incipient crack fault diagnosis of rotating shafts, the best accuracy of which can achieve 99.3%. Moreover, the effectiveness of feature extraction using Shannon entropy was investigated, the result of which presented it can be used for discriminating changes in time and frequency domains. Apart from that, four kernel functions were also evaluated which have showed that RBF function performed better in fault classification than the rest.

Table 2: Classification accuracy using four kernels when 40 decomposition scales are applied using SEW model.

No.	Name	Kernel function	Testing data sets	Misclassified data sets	Testing accuracy (%)
1	Linear	$K(x, y) = x^T y + c$	150	3	98.0%
2	Polynomial	$K(x, y) = (ax^T y + c)^d$	150	3	98.0%
3	Radial basis function	$K(x_i, x_j) = \exp(-\gamma \ x - y\ ^2)$	150	1	99.3%
4	Sigmode	$K(x, y) = \tanh(ax^T + c)$	150	2	98.7%

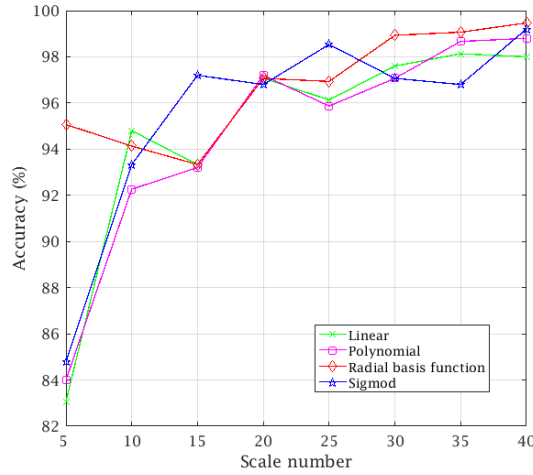


Fig. 8: SVM classification accuracy using different functions

5 Conclusion

In this study, an innovative approach named Self-adaptive Entropy Wavelet (SEW) is proposed to detect incipient transverse crack faults on a rotating shaft based on Continuous Wavelet Transform (CWT), Shannon entropy and SVM classifier. The effectiveness of the proposed approach was investigated by using a test rig, the results of which has proved that the proposed method can be used for crack detection and identification on rotating shafts. The 99.3% classification accuracy can be achieved by using RBF-based SVM. Hence, in the future the proposed method would be further extended to study the fault diagnosis performance of other rotating components in industrial rotating machinery.

Acknowledgement

This work is partially supported by International and Hong Kong, Macao & Taiwan collaborative innovation platform and major international cooperation projects of colleges in Guangdong Province (No.2015KGJHZ026), The Natural Science Foundation of Guangdong Province (No.2016A030307029), and Maoming Engineering Research Center on Industrial Internet of Things (No.517018).

References

1. S. Yin, S. X. Ding, X. Xie, and H. Luo, "A review on basic data-driven approaches for industrial process monitoring," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 11, pp. 6418–6428, 2014.
2. P. Tavner, "Review of condition monitoring of rotating electrical machines," *IET Electric Power Applications*, vol. 2, no. 4, pp. 215–247, 2008.
3. A. Rai and S. Upadhyay, "A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings," *Tribology International*, vol. 96, pp. 289–306, 2016.
4. R. Yan, R. X. Gao, and X. Chen, "Wavelets for fault diagnosis of rotary machines: A review with applications," *Signal Processing*, vol. 96, pp. 1–15, 2014.
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. 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.
7. C. Nagaraju, K. N. Rao, and K. M. Rao, "Application of 3D wavelet transforms for crack detection in rotor systems," *Sadhana*, vol. 34, no. 3, pp. 407–419, 2009.
8. 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.
9. G. Bin, J. Gao, X. Li, and B. Dhillon, "Early fault diagnosis of rotating machinery based on wavelet packets–Empirical mode decomposition feature extraction and neural network," *Mechanical Systems and Signal Processing*, vol. 27, pp. 696–711, 2012.
10. M. Gómez, C. Castejón, E. Corral, and J. García-Prada, "Analysis of the influence of crack location for diagnosis in rotating shafts based on 3 x energy," *Mechanism and Machine Theory*, vol. 103, pp. 167–173, 2016.
11. M. J. Gómez, C. Castejón, and J. C. García-Prada, "Automatic condition monitoring system for crack detection in rotating machinery," *Reliability Engineering & System Safety*, vol. 152, pp. 239–247, 2016.
12. J. Tang, S. Alelyani, and H. Liu, "Feature selection for classification: A review," *Data Classification: Algorithms and Applications*, p. 37, 2014.
13. R. R. Coifman and M. V. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Transactions on information theory*, vol. 38, no. 2, pp. 713–718, 1992.
14. E. Schukin, R. Zamaraev, and L. Schukin, "The optimisation of wavelet transform for the impulse analysis in vibration signals," *Mechanical Systems and Signal Processing*, vol. 18, no. 6, pp. 1315–1333, 2004.

15. J. Kennedy, "Particle swarm optimization," in *Encyclopedia of machine learning*. Springer, 2011, pp. 760–766.
16. D. F. Shanno, "Conditioning of quasi-Newton methods for function minimization," *Mathematics of computation*, vol. 24, no. 111, pp. 647–656, 1970.
17. R. Battiti and F. Masulli, "BFGS optimization for faster and automated supervised learning," in *International neural network conference*. Springer, 1990, pp. 757–760.
18. J. A. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural processing letters*, vol. 9, no. 3, pp. 293–300, 1999.
19. "LIBSVM Matlab Toolbox," (Date last accessed 10-June-2017). [Online]. Available: <https://www.csie.ntu.edu.tw/~cjlin/libsvm/>
20. "PT 500 Machinery Diagnostic System," (Date last accessed 10-June-2017). [Online]. Available: www.gunt.de/static/s3680_1.php