Bearing fault diagnosis based on adaptive variational mode decomposition

Jun Zhou Xue, Tian Ran Lin*, Jin Peng Xing, Chao Ni
Center for Structural Acoustics and Machine Fault Diagnosis,
School of Mechanical and Automotive Engineering, Qingdao University of Technology
Qingdao, China
*trlin@qut.edu.cn

***Abstract*—Variational mode decomposition (VMD) is a novel signal processing technique which decomposes original signals into several intrinsic mode functions (IMFs). Nevertheless, the mode number (*K*) and the quadratic penalty term (*α*) to be preset have a significant effect on the decomposition result. This paper proposed an adaptive VMD technique based on particle swarm optimization (PSO) algorithm where *K* and *α* can be determined adaptively based on the signal characteristic. Firstly, a weighted kurtosis index composing of kurtosis index and a correlation coefficient is chosen and used as the fitness function of PSO optimization. Secondly, the maximum weighted kurtosis index is served as the optimized fitness function of PSO to determine the number of modes (*K*) and the quadratic penalty term (*α*) adaptively . Finally, the related IMF from VMD decomposition is used in the envelope analysis to obtain the defect signal components for fault diagnosis. The effectively of the proposed technique in extracting the fault characteristic frequency (FCF) from noise contaminated signals is evaluated using a bearing experimental data. The advantage of the proposed method is further illustrated by comparing the diagnosis result with that using empirical mode decomposition (EMD).**

Keywords-Fault diagnosis, Variational mode decomposition, Particle swarm optimization, Weighted kurtosis index, Rolling bearing

I. INTRODUCTION

Rolling element bearings are the most critical but sensitive mechanical component in a rotating machine. When a bearing failure occurs, it can result in unexpected shutdown or devastating functional failures of a machine and even human casualty [1-2]. Therefore, bearing condition monitoring (CM) is important to ensure a continuous operation of a machine. The operation condition of a rolling element bearing can be reflected in the vibration CM signal. A bearing defect under the influence of changing load, changing speed, friction or other factors can lead to amplitude-modulated and frequency-modulated (AM-FM) CM signals, and the FCF can be extracted from the signal for an accurate fault diagnosis.

Many methods have been developed for signal analysis and fault diagnosis of rolling element bearings. For instance, Tang et al. [3] proposed a fan bearing health evaluation based on wavelet transform. Wang et al. [4] proposed a technique based on wavelet packet decomposition to extract the weak transient signals produced by a defect bearing for bearing fault diagnosis. Because the mother wavelet function used in the signal analysis needs to be selected manually, signal analysis using wavelet transform and wavelet packet transform cannot be made adaptive to suit the signal characteristic. In contrast, EMD can adaptively decompose a nonlinear, non-stationary signal into multiple IMFs [5]. However, the drawbacks of EMD such as mode mixing and end effect are also well known [6-8]. VMD is an adaptive signal decomposition technique originally proposed by Dragomiretskiy et al.[9]. VMD can outperform EMD in tonal signal detection, tonal signal separation, and is much more robust in the analysis of noise contaminated signals. VMD has been widely employed in mechanical fault diagnosis. Mohanty et al. [10] used VMD analysis to achieve an accurate bearing fault diagnosis. Wang et al. [11] detected multiple signal signatures produced by rotor-to-stator rubbing by using VMD technique. Markert et al. [12] applied an optimized VMD technique for wind turbine condition monitoring. Nevertheless, the number of decomposed modes (K) and the quadratic penalty term (α) used in VMD are either chosen manually or are optimized independently by neglecting the interaction between the two parameters which can easily be trapped in a local optimal. Yan et al. [13] proposed an improved VMD technique based on the genetic algorithm which can optimize the two parameters simultaneously. Nevertheless, the fitness function used in the optimization considers only the impact properties of the decomposed IMFs but ignores their correlation with the original signal which can result in information loss.

Inspired by the success of the previous work and taking into consideration the limitation of existing works, this paper presents an adaptive VMD technique based on PSO where the weighted kurtosis index similar to that used in Ref. [14] is used as the objective function. The structure of this paper is arranged as follows: The background theories of VMD, PSO and the weighted kurtosis index are introduced in Section II. Section III elaborates the adaptive VMD method proposed in this work. The validity of the proposed adaptive VMD technique is evaluated utilizing an experimental signal acquired from a bearing test rig which can be found in Section IV. Conclusion is drawn in Section V.

II. THEORETICAL BACKGROUNDS

A. Variation Mode Decomposition

VMD is a non-recursive signal decomposition technique which decomposes a real signal $x(t)$ into several IMFs (u_k)

The financial support from the Shandong Provincial Government key research project grant (Grant No 2018GGX109011) for this work is gratefully acknowledged

with specific sparsity properties [9]. Each IMF (u_k) is compact around the center frequency ω_k, and the bandwidth is estimated through the squared L^2 norm of the gradient. The solution of constrained variational problem is given as [9].

$$L_{\text{norm}}^2 = \underset{\{u_k(t)\},\{\omega_k\}}{\operatorname{argmin}} \sum_k \left\| \frac{\partial}{\partial t} \left\{ \left[\delta(t) + \mathrm{j}\frac{1}{\pi t} \right] * u_k \right\} \exp(-\mathrm{j}\omega_k t) \right\|_2^2 \tag{1}$$

$$\text{subject to } \sum_{k=1}^{K} u_k(t) = x(t)$$

where $\delta()$ is the Dirac delta function, j is the imaginary number, k is the number of IMFs, the symbol $*$ denotes convolution.

Due to the difficulty in solving the constrained problem, the quadratic penalty term α and Lagrangian multipliers λ are utilized to convert the above constrained issue to an unconstrained variational problem. The augmented Lagrangian is described below:

$$L\{\{u_k(t)\},\{w_k\},\lambda(t)\} = \alpha \sum_k \left\| \frac{\partial}{\partial t} \left\{ \left[\delta(t) + \mathrm{j}\frac{1}{\pi t} \right] * u_k \right\} exp(-\mathrm{j}\omega_k t) \right\|_2^2 + \left\| x(t) - \sum_k u_k(t) \right\|_2^2 + \langle \lambda(t), x(t) - \sum_k u_k(t) \rangle \tag{2}$$

The saddle point of Eq. (2) is found using alternating direction multiplier method (ADMM).

The processing steps of VMD are described below:

1. Initializes $\{\bar{u}_k^1(\omega)\}$, $\{\omega_k^1\}$, $\{\bar{\lambda}^1(\omega)\}$ where the overhead hat indicates the Fourier transform, presets the initial iteration number n to 0, and presets the quadratic penalty term α and the number of modes K in the decomposition.
2. For $k = 1{:}K$, updates $\bar{u}_k(\omega)$, ω_k and $\bar{\lambda}(\omega)$, for modes having the center frequency $\omega \geq 0$.

$$\bar{u}_k^{n+1}(\omega) = \frac{\bar{x}(\omega) - \sum_{k \neq K} \bar{u}_k(\omega) + \frac{1}{2}\bar{\lambda}^n(\omega)}{1 + 2\alpha(\omega - \omega_k^n)^2} \tag{3}$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\bar{u}_k^{n+1}(\omega)|^2 \, d\omega}{\int_0^\infty |\bar{u}_k^{n+1}(\omega)|^2 \, d\omega} \tag{4}$$

$$\bar{\lambda}^{n+1}(\omega) = \bar{\lambda}^n(\omega) + \tau \left[\bar{x}(\omega) - \sum_{k=1}^{k} \bar{u}_k^{n+1}(\omega) \right] \tag{5}$$

3. Repeats the iteration process until the following convergence condition is met.

$$\sum_{k=1}^{k} \frac{\|\bar{u}_k^{n+1}(\omega) - \bar{u}_k^n(\omega)\|_2^2}{\|\bar{u}_k^n(\omega)\|_2^2} < \varepsilon \tag{6}$$

4. The IMF $u_k(t)$ in time domain can be obtained by taking the inverse Fourier transforming of Eq. (3) and keeping the real part of the inversed signal.

A major drawback of VMD is that the parameters used in the decomposition needs to be preset, an inappropriate setting the mode number K can have an impact on the characteristics of decomposed modes. Furthermore, an incorrected selection of the quadratic penalty term α will affect the performance of noise suppression in the signal decomposition. Therefore, the ability to select the parameters K and α adaptively to the signal characteristics in the decomposition is key for the application of VMD technique.

B. Particle Swarm Optimization

Particle swarm optimization (PSO) is an optimization algorithm based on the swarm intelligent inspired from the foraging behaviors of a bird flock which was proposed originally by Kennedy in 1995 [15]. The algorithm has been widely used in parameter optimization for machine learning and artificial intelligent neural networks these days.

A particle is a potential optimal solution of PSO which can be described by three features: fitness value, velocity, and position [15]. During the optimization process, the particles continually refresh their velocities and positions according to the best individual value (P) and the best group value (G) in each iteration. The updated particles continue to search the optimal values in the search space until reaching the preset maximum number of iterations.

PSO algorithm is as follows:

1. Initializes the swarm X_i (i = 1, 2, ⋯, M), the maximum number of iterations (N), initial positions and velocities.
2. Selects the fitness function according to the actual problem.
3. Calculates the optimal local extremum value P and the optimal global extremum value of the population G, retains the better result.
4. Updates the positions and velocities of all particles according to

$$v_{id}^{n+1} = \theta v_{id}^n + c_1 \eta (p_{id}^n - x_{id}^n) + c_2 \eta (g_d^n - x_{id}^n) \tag{7}$$

$$x_{id}^{n+1} = x_{id}^n + v_{id}^{n+1} \tag{8}$$

where θ is the inertia weight; $d=1, 2, \cdots, D$, D is the dimensional space; c_1 and c_2 are the learning rates; η is a random value between $[0,1]$, x_{id}^n and v_{id}^n the position and velocity of i^{th} particle in n^{th} iteration of d dimension.

5. Repeats steps 2 and 3 until reaching the preset maximum number of iterations.

Comparing with other optimization algorithms such as grey wolf optimizer (GWO) and firefly algorithm (FA), PSO is more appropriate to the optimize parameter choice due to its simple mechanism, faster convergence speed, and excellent performance in global optimization search [15]. A key criterion for a successful application of PSO is the selection of the fitness function.

C. Weighted Kurtosis Index

The correlation coefficient and the kurtosis index are two important parameters in vibration-based machine fault diagnosis. The correlation coefficient can characterize the correlation between two signals, but it is sensitive to noise when detecting the impact signal. The kurtosis index is a dimensionless index which is sensitive to impacts presented in a vibration signal. It can be used to detect the impulses produced by a defect of a rotating component such as bearings for fault diagnosis of a rotating machine. However, it is susceptible to the distribution density of the impacts. Selecting only the maximum kurtosis index as the optimization fitness function for impact signal detection can cause information loss problem. By taking consideration of the advantages and drawbacks of these two parameters, Miao et al. [14] constructed a so-called weighted kurtosis index as the fitness function by binding the two parameters into one objective function. The weighted kurtosis index is adopted as the objective function of grey wolf optimizer in optimizing a time varying filtering based empirical mode decomposition method for machinery fault diagnosis. As a summary. The weighted kurtosis index can be expressed as follows:

$$WKI = KI * |C| \tag{9}$$

$$KI = \frac{\frac{1}{N}\sum_{n=0}^{N-1} x^4(n)}{\left(\frac{1}{N}\sum_{n=1}^{N-1} x^2(n)\right)^2} \tag{10}$$

$$C = \frac{E[(x-\bar{x})(y-\bar{y})]}{E[(x-\bar{x})^2]E[(y-\bar{y})^2]} \tag{11}$$

where WKI represents the weighted kurtosis index, KI is the kurtosis index of the discrete signal $x(n)$, N is the length of the signal, C represents the correlation coefficient between signals x and y; $E[\,]$ is the mathematical expectation. The absolute value of the correlation coefficient $|C| \leq 1$ can be obtained from the Schwarz inequality.

III. THE PROPOSED METHOD

Selecting a wrong value for the parameter K and α in VMD can affect the end result. In this paper, K and α are selected adaptively by PSO using the maximum weighted kurtosis index as fitness function as

$$\begin{cases} f = \underset{\beta=(K,\alpha)}{\operatorname{argmax}}\{WKI_i\} \\ \text{subject to } K = 2,3,\cdots,10 \\ \qquad \alpha\epsilon[1000, 10000] \end{cases} \tag{12}$$

where f is the fitness function, WKI_i $(i = 1, 2, 3 \cdots K)$ represents the weighted kurtosis index value of the IMFs from VMD decomposition, $\beta = (K, \alpha)$ is the parameter set to be optimized, where $K\epsilon[2,10]$, $\alpha\epsilon[1000,10000]$ are the preset range of the two parameters. The flow diagram of the proposed adaptive VMD technique is described in Fig. 1.

The corresponding signal processing steps are elaborated below:

1. Initializes the parameters of PSO and sets the range of VMD parameters to be optimized.
2. Decomposes the vibration signal using VDM, calculates the KCI of each IMF and saves the maximum KCI after each iteration.
3. Terminates the iteration process when $n \geq N$
4. Saves the optimal parameters and the maximum WKI
5. Decomposes the input signal once again using VMD with optimal parameters, calculates the WKI of each IMF.
6. Selects the IMF with the maximum WKI for envelope analysis to extract the related FCF for bearing fault diagnosis.

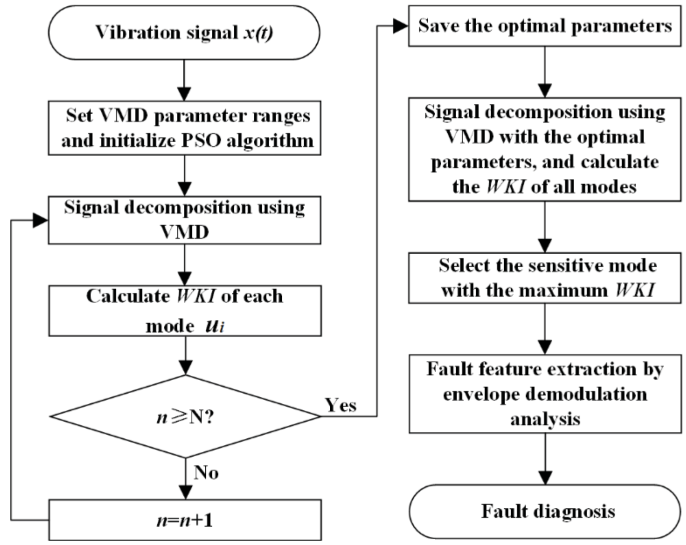

Figure 1. The flow diagram of the proposed method

IV. EXPERIMENTAL VALIDATIONS

A bearing experimental data is used in this section to verify the validity of the present adaptive VMD technique. The data is acquired from a Spectra Quest's Machinery Fault Simulator shown in Fig. 2. The test bearing type and the related parameters are listed in Table I. Multiple faults are simulated for the test bearing and the bearing FCFs are displayed in Table II. The experiment was conducted at a constant shaft rotating frequency (f_r) of 21 Hz. Vibration condition monitoring signals were acquired using a B&K4370 accelerometer attached on top of the test bearing house. The sampling frequency used in the experiment is 10 kHz, and the signal length is 1s.

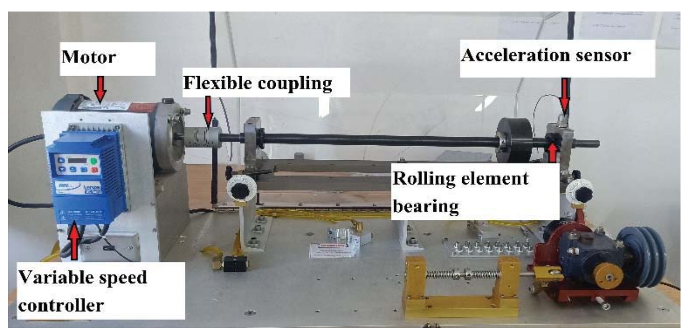

Figure 2. Spectra Quest's Machinery Fault Simulator

Table I ER-16K single row rolling element bearing and its parameters

Parameter	Value
Inner ring diameter (mm)	25.4
Pitch diameter (mm)	38.5064
Rolling element diameter (mm)	7.9375
Number of rolling element (n)	9
Contact angle (°)	9.08

Table II Bearing fault characteristic frequencies

Inner ring	Outer ring	Rolling element
$5.43f_r$	$3.572f_r$	$2.322f_r$

The time-domain waveform and the frequency spectrum of the vibration condition monitoring signal are shown in Fig. 3 and Fig. 4 respectively. As result of the presence of noise in the experiment and multiple faults of the bearing, it is difficult to detect the periodic impacts produced by each individual defect from Fig. 3. It is observed from the frequency spectrum in Fig. 4 that there is a resonance band between 2000-3000Hz in the spectrum. The identified structural resonant band is used to bandpass filter the original signal and an envelope demodulation analysis is applied on the filtered signal as shown in Fig. 5. It is demonstrated that the inner race FCF can be clearly identified from the envelope spectrum. Whilst other two FCFs are largely submerged by the high noise floor.

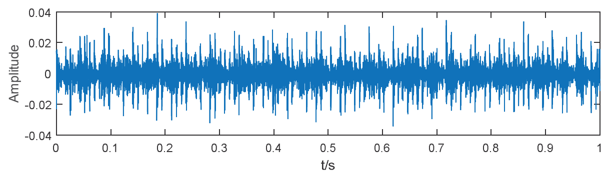

Figure 3. The time-domain waveform of the vibration signal

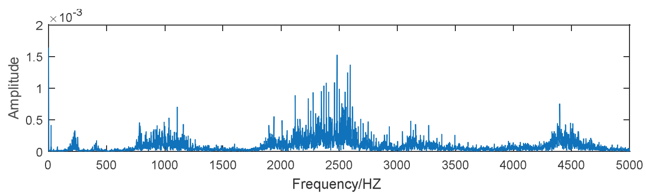

Figure 4. The corresponding frequency spectrum

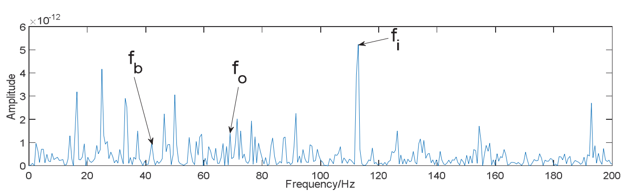

Figure 5. The envelope spectrum of the filtered signal

To reduce the effect of noise interference, the vibration condition monitoring signal is processed using the adaptive VMD technique presented in this paper. The parameters used in the PSO algorithm are population size $M = 20$, maximum number of iterations $N = 50$, the learning rates $c_1 = c_2 = 1.5$. The optimal convergence curve of the maximum *WKI* value is shown in Fig. 6, and the corresponding optimal combination of $[K, \alpha]$ is found to be $[6, 3034]$. These two optimal parameters are then used in VMD to analyze the vibration condition monitoring signal, and the IMFs are exhibited in Fig. 7. The *WKIs* corresponding to the 6 IMFs from the decomposition are demonstrated in Fig. 8. It is demonstrated that IMF4 has the largest *WKI* value and is the most sensitive to the bearing fault signal components. The envelope spectrum of the IMF4 is shown in Fig. 9. It is shown that the noise floor shown in Fig. 5 has been effectively subdued and each of the FCFs of the compounded bearing defects can be clearly identified from the analysis.

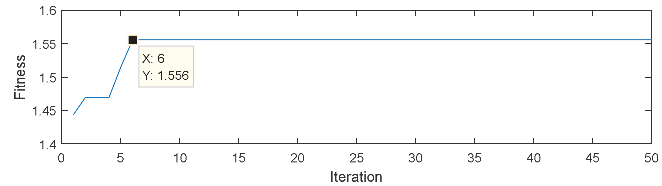

Figure 6. The convergence curve of the PSO optimization

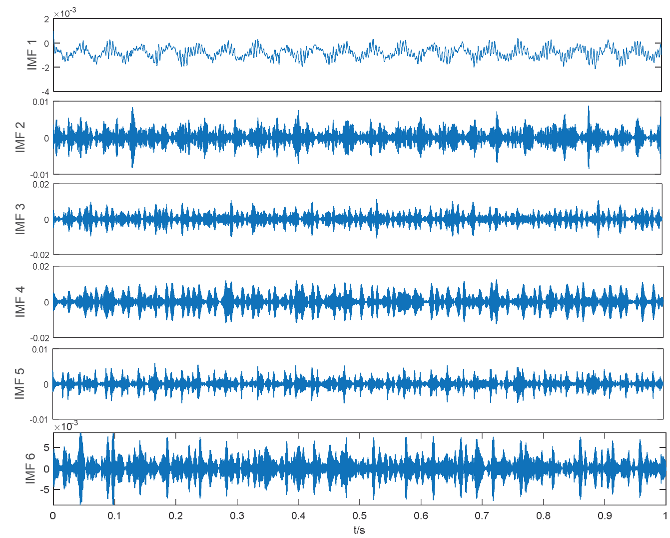

Figure. 7 The IMFs decomposed by the adaptive VDM technique

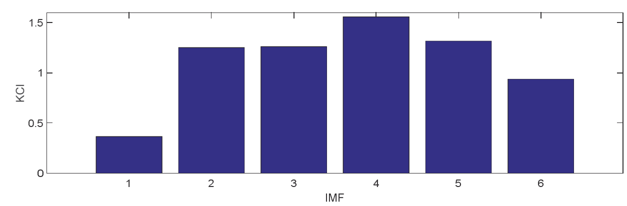

Figure 8. The *WKI*s of the corresponding IMFs

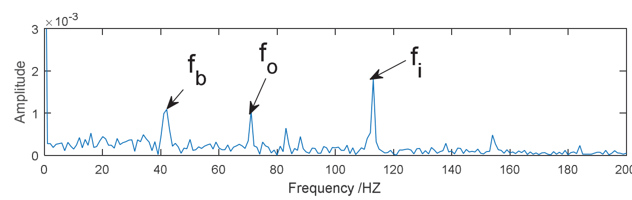

Figure 9. The envelope spectrum of IMF4

To evaluate the validity of the presented method, the vibration condition monitoring signal is processed by EMD and the IMFs and the results are shown in Fig. 10. The *WKI* value of each IMF is calculated and the IMF with the largest *WKI* value (IMF2) is used in the following envelope analysis. The envelope spectrum of IMF2 is exhibited in Fig. 11. Comparing the result shown in Fig. 11 with that shown in Fig. 9, it is concluded that the proposed adaptive VDM technique can produce better signal analysis result than that of EMD technique.

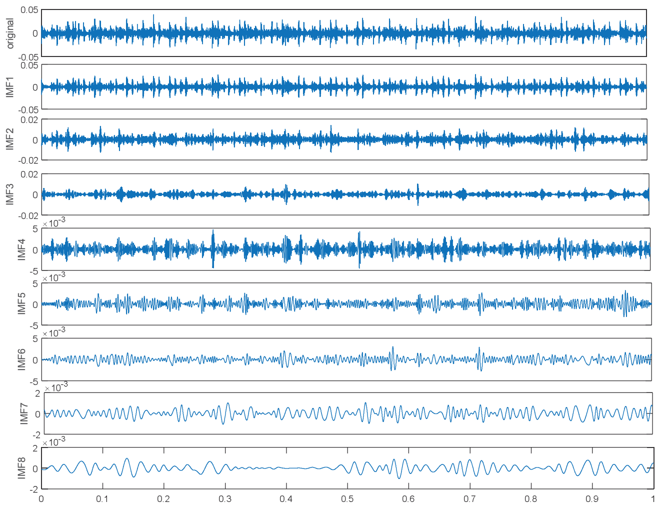

Figure 10. The IMFs decomposed using EMD

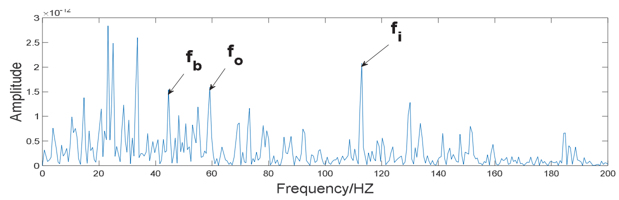

Figure 11. The envelope spectrum of IMF2 in Fig. 10

V. Conclusion

In the study, an adaptive VMD technique utilizing PSO for parameter optimization was presented for bearing fault diagnosis. The presented technique effectively resolves the problem of the original VMD technique in the selection of parameters K and α to suit the signal characteristics. The validity of the presented technique is evaluated using a bearing experimental data with compounded bearing defect signals. It is shown that the proposed technique can effectively subdue the noise interference in the signal to produce a clean signal spectrum for an accurate bearing fault diagnosis. A comparison study also showed that the proposed technique can outperform that of the EMD technique.

References

[1] R. B. Randall. "Vibration-based condition monitoring: industrial, aerospace and automotive applications". Condition monitoring December 19, 2010.

[2] T. R. Lin, E. Kim, and A. C. C. Tan. "A practical signal processing approach for condition monitoring of low speed machinery using Peak-Hold-Down-Sample algorithm." Mechanical Systems and Signal Processing . vol.36, no.2, pp. 256-270, 2013.

[3] Q. Miao, C. Tang, W. Liang, and P. Michael. "Health Assessment of Cooling Fan Bearings Using Wavelet-Based Filtering." Sensors . vol.13, no.1, pp. 274-291, 2013

[4] W. Yi, G. H. Xu, L. Liang, and K. S. Jiang. "Detection of weak transient signals based on wavelet packet transform and manifold learning for rolling element bearing fault diagnosis." Mechanical Systems and Signal Processing. vol.54, pp. 259-276, 2014.

[5] N. E. Huang et al. "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis." Proceedings: Mathematical, Physical and Engineering Sciences. vol.454, no.1971, pp. 903-995, 1998.

[6] Z. H. Wu, 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.

[7] G. Rilling, P. Flandrin. "On the Influence of Sampling on the Empirical Mode Decomposition." IEEE International Conference on Acoustics IEEE, 2006.

[8] Y. G. Lei, J. Lin, Z. J. 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. 1-2, pp. 108–126, 2013.

[9] K. Dragomiretskiy, D. Zosso. "Variational Mode Decomposition." IEEE Transactions on Signal Processing. vol.62, no.3, pp. 531-544, 2014.

[10] G. K. K. Mohanty, R. K. Solomon. "Bearing Fault Analysis using Variational Mode decomposition." International Conference on Industrial & Information Systems 2014.

[11] Y. X. Wang, R. Markert. "Filter bank property of variational mode decomposition and its applications." Signal Process. vol.120, pp. 531-544, 2016.

[12] Y. X. Wang, R. Markert, J. W. Xiang, W. G Zheng. "Research on variational mode decomposition and its application in detecting rub-impact fault of the rotor system." Mechanical Systems and Signal Processing. vol.60-61, pp. 243-251, 2015.

[13] X. A. Yan, M. P. Jia, L. Xiang. "Compound fault diagnosis of rotating machinery based on OVMD and a 1.5-dimension envelope spectrum." Measurement Science and Technology. vol.27, no.7, pp. 1361-6501, 2016.

[14] X. Zhao, Z. Liu, Q. Miao, and L.Wang. "An optimized time varying filtering based empirical mode decomposition method with grey wolf optimizer for machinery fault diagnosis." Journal of Sound and Vibration . vol.418,pp. 52-78, 2018.

[15] J. Kennedy, R. Eberhart. "Particle swarm optimization." Proceedings of ICNN'95 International Conference on Neural Networks IEEE, 2002.