Int J Adv Manuf Technol (2017) 92:2069–2080
DOI 10.1007/s00170-017-0252-y

ORIGINAL ARTICLE

A hybrid health condition monitoring method in milling operations

Jie Liu[1] · Youmin Hu[1] · Bo Wu[1] · Chao Jin[2]

Received: 3 November 2016 / Accepted: 6 March 2017 / Published online: 25 March 2017

Abstract Cutting chatter has been a very important issue in the milling operations due to its unexpected and uncontrollable characteristics. Developing an effective healthy condition monitoring method is critical to identify cutting chatter exactly. In this paper, a hybrid healthy condition monitoring (HHCM) method, that combines variational mode decomposition (VMD) with genetic algorithm-based back propagation neural network (BPNN) model, is developed for cutting chatter detection and state classification in complex and non-stationary milling operations. First, cutting chatter vibration signal is decomposed into multiple mode components by the VMD. Then, Shannon power spectral entropy is adopted to extract features from decomposed vibration signals. Furthermore, BPNN model is optimized by traditional genetic algorithm to identify and classify machine states in milling operations. Last, impeller milling experiments are conducted and results show that the proposed HHCM method can effectively realize cutting chatter detection and state classification during milling operations.

Keywords Milling operations · Health condition monitoring · Cutting chatter detection · State classification · Variational mode decomposition

✉ Youmin Hu
youmhwh@mail.hust.edu.cn

[1] School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, People's Republic of China

[2] Research Institute of Zhejiang University-Taizhou, Taizhou 318000, People's Republic of China

1 Introduction

Cutting chatter is one of the most unexpected and uncontrollable phenomenon in achieving high-performance machining operations as shown in Fig. 1. It usually arises in combinations of cutting parameters, such as spindle speed, feed rate, and cutting depth. Due to its complex and non-stationary properties, the cutting chatter affects productivity in milling operations, causing harsh noises, tool wear, poor surface quality, lower dimensional accuracy, materials waste, and shortened machine life [4, 37, 39], etc.

By now, the cutting chatter has been analytically studied by many researchers [2, 32, 40]. However, analytic method still can not identically model practical milling systems [1, 16]. So some process-based cutting chatter issues, including cutting chatter detection, cutting chatter prediction, cutting chatter suppression [34], and cutting chatter stability [4, 26], have been focused on. Among these issues, process-based cutting chatter detection helps to identify the state of cutting chatter and provides a foundation for other issues. Typically, skilled industrial users identify, predict, and change the state of cutting chatter in milling operations mainly by their experience and prior knowledge. But labor costs for skilled industrial users increases and false judgement happens.

Thus, it is meaningful to develop healthy condition monitoring systems to monitor milling operations in real time. Typically, the following aspects can be achieved: (a) avoid modeling of complex milling operations; (b) monitor process parameters and various sensor signals of milling operations; (c) detect, identify, and predict the current or incoming state of milling operations based on monitoring data; and (d) establish a milling operation changing strategy to extend machine life, reduce material waste, compensate machine error, and ensure surface quality, machining accuracy and productivity improvement. Therefore,

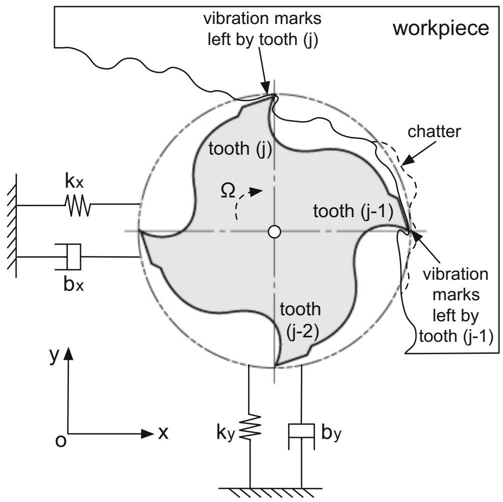

Fig. 1 Dynamic model of milling with two degree of freedom [2]

establishing healthy condition monitoring systems is essential and crucial in practical milling operations.

Extensive and significant research has been conducted on chatter detection of the healthy condition monitoring systems in recent years. Sensor selection, signal processing, feature extraction, and state classification are critical steps in chatter detection, and all these steps have been studied extensively. Various sensor signals have been investigated to monitor chatter, such as acceleration signal [6, 14], acoustic emission [9, 19, 35], cutting force [3, 15, 24], servo current [31], sound [41], etc. These sensors can generate certain features correlated with machining conditions. Due to its easy measurement and high correlation with machine dynamics, vibration signal analysis is the most common approach among sensor signals above.

Generally, signal processing method is much more important. Different signal processing methods have been employed especially in time-frequency domain, to extract relevant and sensitive features from measured nonstationary vibration signals. For example, wavelet transform decomposes the signal into a series of sub-signals and the original signal can be perfectly reconstructed using the wavelet coefficients [11]. These sub-signals then can be used for feature extraction. However, the selection of wavelet basis function and the number of decomposition levels restrict its usage. Different wavelets would lead to quite different performance. Besides, the time-adaptive empirical mode decomposition (EMD) method is used by Hilbert-Huang transform to decompose the signal into a set of orthogonal intrinsic mode functions (IMFs), and then Hilbert spectral analysis is applied to obtain instantaneous frequency array [23]. It does not involve too much manual intervention rather than the wavelet transform. However, due to lack of theoretical basis, some drawbacks, such as mode mixing and end effects, still exist. In addition, ensemble empirical mode decomposition, as an extended version of EMD, has some improvement in solving model mixing of EMD and is sensitive to the strong background noise [44]. Thus, the algorithm should be conducted many times to reduce errors caused by white noise.

Lately, a new variational mode decomposition (VMD) method with advanced multiresolution is proposed [13]. A series of iterative updating process help to minimize the constrained variational model; therefore, calculus of variation decompose the vibration signal into various modes or IMFs. It overcame the disadvantage of lacking theoretical basis and noise sensitivity of EMD. Moreover, the VMD method could adaptively determine the relevant frequency band and estimated the corresponding model. Based on the above-mentioned advantage, it was applied to the applications of rolling bearing, gearbox, wind generation, and economics field, etc.

After time-frequency preprocessing, it needs to model the relationship between vibration signals and cutting chatter state. Before the system modeling, some feature extracting methods should be explored to improve the classification accuracy and reduce computation time, such as normalized energy ratio [49], statistical features [45], wavelet packet coefficients [46], synthetic criterion [27], etc.

It is obvious that states of cutting chatter are affected by milling operation condition, such as spindle speed, feed rate, cutting depth, etc. Extracted features are classified using classification tools to identify cutting chatter in milling operations. Regarding to the nonlinear and the nonstationary characteristics of the milling operations, several classification approaches, such as back propagation neural network (BPNN) [22], generalized Markov chain model [45], generalized hidden Markov model [46], and support vectors machine [29], had been used in previous work. Among these algorithms, BPNN model usually suffers from the problem of multiple local minima and over-fitting. In order to deal with these problems, genetic algorithm (GA) is usually used to optimize the weight and threshold of the BPNN model. Therefore, genetic algorithm-based back propagation neural network (BPNN-GA) model can be considered as an effective method identifying state of cutting chatter in milling operations.

In this paper, an effective hybrid healthy condition monitoring (HHCM) method based on VMD and BPNN-GA model for cutting chatter detection and state classification in milling operations is proposed. The proposed method can be used to identify and classify the vibration signals related to stable, transition, and chatter states. The VMD decomposes the vibration signal under different milling operations and overcomes the disadvantage of lacking theoretical basis and noise sensitivity of traditional EMD. Then, decomposed signals is processed using Shannon power spectral entropy to calculate the frequency and energy distribution, which are closely related to machine states of milling operations. Furthermore, a set of impeller experiments under different milling operations are conducted. Final results show that the performance of the BPNN-GA model based on feature extraction is reliable and improved.

The paper is organized as follows: Section 2 introduces setup of impeller milling experiments under different cutting conditions. Architecture of the HHCM process is illustrated in Section 3, including signal preprocessing, features extraction, and system modeling. Experimental results are analyzed to demonstrate the performance of the proposed method in Section 4. Finally, this paper is concluded with a brief summary of the proposed HHCM method.

2 Experimental setup

To validate the effectiveness of the proposed method, a series of impeller milling experiments were conducted under different cutting conditions. An online chatter monitoring system was set up based on a numerical control milling machine Mikron HSM-600U. The tool is a hard allay milling cutter with six teeth. The workpiece is a rough titanium alloy Ti6Al4V impeller. The vibration signals of the spindle were obtained by the accelerometer (356A16, PCB, USA) which were mounted on the spindle. A data acquisition unit (PXI-1042 National Instruments Corporation, USA) was used to acquire vibration signals. All tests were conducted without coolant. The photograph of the experimental setup is shown in Fig. 2.

Vibration signals of accelerate sensors under a milling operation condition with spindle speed 12,000 rpm, feed rate 0.3 mm/rev, radial depth of cut 0.2 mm, and axial depth of cut 1 mm in time domain are illustrated in Fig. 3. It is noted that the test shows a non-stationary transition from

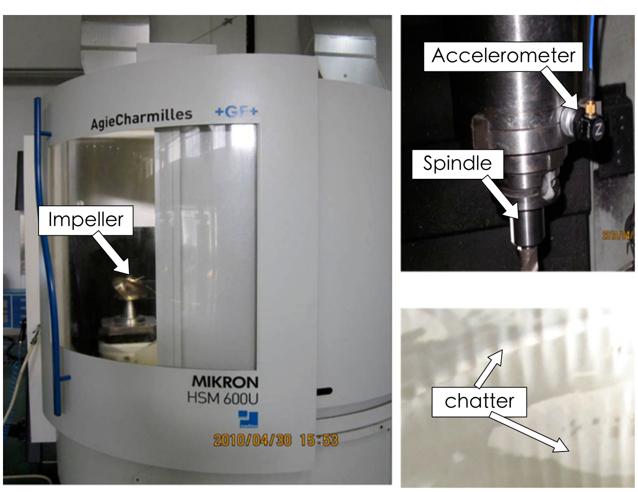

Fig. 2 Experimental setup for impeller milling experiments. **a** Mikron HSM-600U. **b** Installment of accelerometer PCB 356A16. **c** Cutting chatter impacts

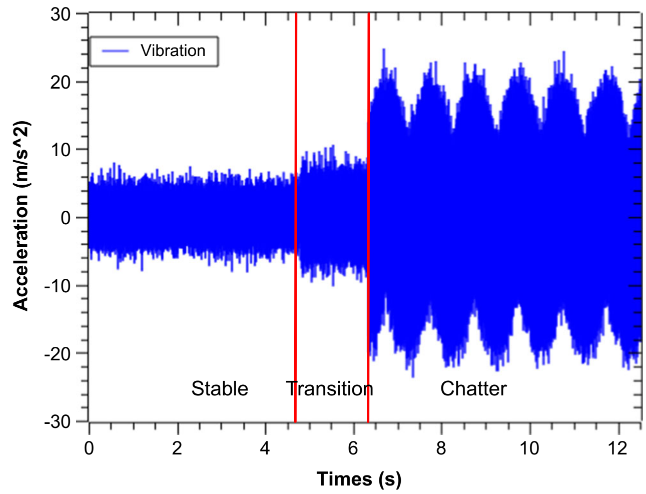

Fig. 3 Non-stationary vibration signals in impeller milling experiments

stable state, via transition state, to chatter state eventually in the time domain. For example, the milling operation condition is stable in the beginning, then transits to transition state around 4.7 s, and reaches to the chatter state around 7.3 s in the end.

Another 27 groups of experimental tests with representative results were also chosen and the relevant milling operation conditions are listed in Table 1. Parameters of impeller milling operations were set with the spindle speed from 3000 to 12,000 rpm, the feed rate from 0.1 to 0.3 mm/rev, the radial depth of cut 0.2 mm, axial depth of cut from 0.5 to 1 mm, and sampling rate from 2000 to 2500 Hz. All algorithms were processed by MATLAB 8.4.0 (2014b) in a laptop with an Intel Core i5 CPU and 8 G RAM.

3 Methodology

Figure 4 shows basic flow chart of healthy condition monitoring method. First, the VMD is used to decompose vibration signals and obtain a set of sub-signals. Then, based on decomposed sub-signals, Shannon power spectral entropies are obtained to realize feature extraction of the sub-signals. Next, the BPNN-GA model is used to build system model. Finally, cutting chatter state in milling operations is identified and classified by proposed HHCM method.

3.1 Signal preprocessing

The VMD is a newly developed methodology for adaptive and quasi-orthogonal signal decomposition by [13]. In the VMD framework, the signal is decomposed into k discrete number of sub-signals, and each component is considered compact around a corresponding center frequency.

Table 1 Experimental settings of impeller milling operations

No.	Spindle speed (rpm)	Feed rate (mm/rev)	Radial depth of cut (mm)	Axial depth of cut (mm)	Sampling rate (Hz)
1	3000	0.1	0.2	0.5	2000
2	3000	0.2	0.2	0.75	2000
3	3000	0.3	0.2	1	2000
4	4500	0.1	0.2	0.75	2000
5	4500	0.2	0.2	1	2000
6	4500	0.3	0.2	0.5	2000
7	6000	0.1	0.2	1	2000
8	6000	0.2	0.2	0.5	2000
9	6000	0.3	0.2	0.75	2000
10	6000	0.1	0.2	0.75	2000
11	6000	0.2	0.2	1	2000
12	6000	0.3	0.2	0.5	2000
13	7500	0.1	0.2	0.5	2000
14	7500	0.2	0.2	0.75	2000
15	7500	0.3	0.2	1	2000
16	7500	0.1	0.2	0.75	2000
17	7500	0.2	0.2	1	2000
18	7500	0.3	0.2	0.5	2000
19	9000	0.1	0.2	0.75	2000
20	9000	0.2	0.2	1	2000
21	9000	0.3	0.2	0.5	2000
22	9000	0.1	0.2	0.5	2000
23	9000	0.2	0.2	0.75	2000
24	9000	0.3	0.2	1	2000
25	12000	0.1	0.2	1	2500
26	12000	0.2	0.2	0.5	2500
27	12000	0.3	0.2	0.75	2500

The process of VMD can be considered as a constrained variational problem, while the formulation of the constrained variational problem is written by

$$\min_{\{u_k\},\{\omega_k\}} \left\{ \sum_{k=1}^{K} \left\| \partial_t \quad \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) e^{-j\omega_k t} \right] \right\|_2^2 \right\}, \quad \text{subject to} \sum_{k=1}^{K} u_k(t) = f(t), \tag{1}$$

where each mode u_k is almost compact around a matching center frequency ω_k, and its bandwidth is assessed by means of H^1 Gaussian smoothness.

In order to deal with unstrained optimization problem in Eq. 1, it can be easily achieved via the augmented Lagrangian method by

$$\mathcal{L}\left(\{u_k\}, \{\omega_k\}, \lambda\right) = \alpha \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * e^{-j\omega_k t} \right] \right\|_2^2 + \left\| f(t) - \sum_{k=1}^{K} u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_{k=1}^{K} u_k(t) \right\rangle. \tag{2}$$

Alternating direct multipliers are adopted to solve (2). The estimated modes u_k and the corresponding updated center frequency ω_k in the frequency domain can be written as follows:

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2}, \tag{3}$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega \left|\hat{u}_k(\omega)\right|^2 d\omega}{\int_0^\infty \left|\hat{u}_k(\omega)\right|^2 d\omega}. \tag{4}$$

where α is the balancing parameter of the data-fidelity constraint. More specically, the process of VMD can be summarized in Fig. 5.

More details of the VMD algorithm can be found in [13]. In addition, initialization and input parameters (balancing parameter α and number of modes K) are key parameters of the VMD. Compared with uniformly spaced distribution, the VMD can get much more reliable and meaningful results with zero initial for the detecting transient signatures [42, 43, 49].

3.2 Features extraction

Frequency and energy distribution are closely related to the cutting chatter state. Normally, the energy of the milling system is mostly dominated by its rotation frequency and harmonic frequencies in stable milling operation condition. When milling operation condition changes, or cutting chatter occurs, the energy in rotation frequency and harmonic frequencies is absorbed to the chatter frequency gradually.

Shannon power spectral entropy, as a function of the probability distribution, is a quantitative uncertainty index of the irregularity and complexity from decomposed sub-signals. In other words, it is linked to the distribution of energy components [7, 48, 50].

Using the VMD, the vibration signals can be decomposed into a set of sub-signals, $u_1(t), u_2(t), ..., u_n(t)$. It is noted that these sub-signals include different frequency bands ranging from low to high. The power spectrum of each component can be calculated as:

$$s_k(f) = \frac{1}{2\pi N}|U_k(\omega)|^2, \tag{5}$$

where N is the length of the time series, $U_k(w)$ is the Fourier transformation of decomposed sub-signal $u_k(j)$, $j = 1, 2, ..., N$.

Probability density function can be estimated by normalization over all frequency components

$$P_{kj} = s_k(f_j)/\sum_{j=1}^{n} s_k(j),\ j = 1, 2, \cdots, N, \tag{6}$$

where $s_k(f_j)$ is the spectral energy for the frequency component f_j of kth decomposed sub-signal $u_k(t)$, P_{kj} is the corresponding probability density, and N is the total number of frequency components in FFT. Then, the power spectral entropy is defined as

$$H_k = -\sum_{j=1}^{N} P_{kj} \cdot \ln P_{kj}. \tag{7}$$

For the sake of convenient comparison in different milling operations, the result is usually normalized by

$$E_k = \frac{H_k}{\ln N} = \frac{-\sum_{j=1}^{N} P_{kj} \cdot \ln P_{kj}}{\ln N}. \tag{8}$$

Then, the power spectral entropy is a non-dimensional indicator in the range of [0, 1].

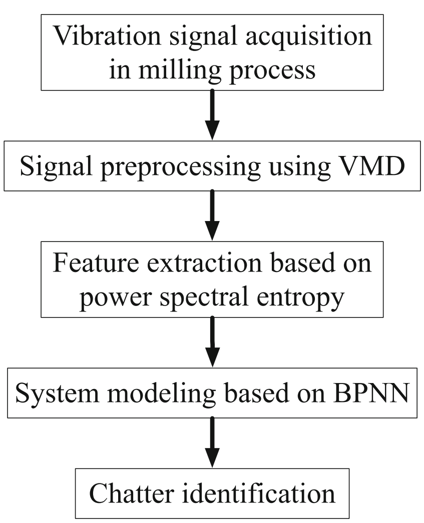

Fig. 4 Basic flow of proposed healthy condition monitoring method

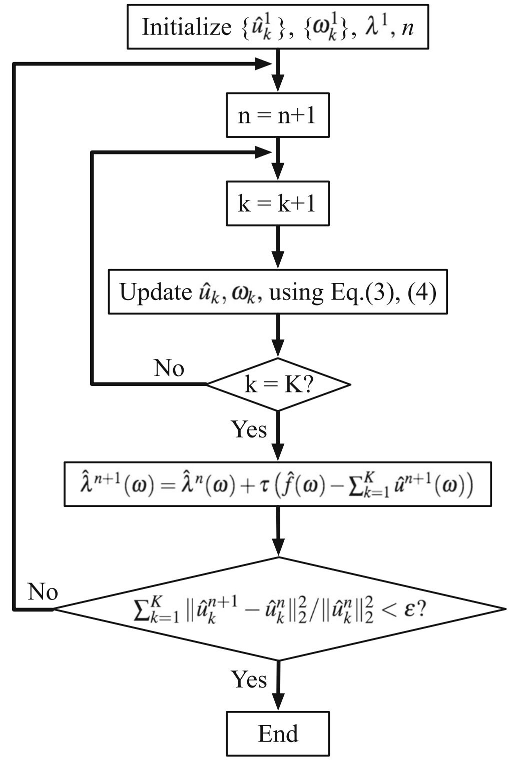

Fig. 5 Basic flow of variational mode decomposition

3.3 System modeling

3.3.1 Back propagation neural network

The BPNN model with an error back propagation arithmetic was proposed in 1986 by McClelland and Rumelhart [38]. It has been widely used in pattern identification of machine condition monitoring fields such as cutting chatter [22], rolling element bearings [17], cutting tool wear [18], surface grinder [25], surface roughness [36], etc. The general structure of BPNN model is shown in Fig. 6. The BPNN consists of three layers, namely input, hidden, and output layers. In addition, the variables n, k, and m mean the

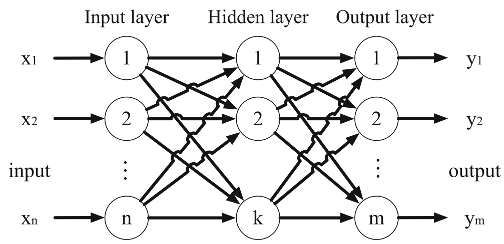

Fig. 6 Structure of back propagation neural network model

total neuron numbers in the input, hidden, and output layers, respectively.

$$v_i = \sum_{j=1}^{n} w_{ij} x_j + \theta_i, \tag{9}$$

where $v_i(i = 1, \cdots, k)$ is the weighted sum of the input to the ith processing element in the hidden layer, w_{ij} is the weight of the connections between the jth element in the input layer and ith element in the hidden layer, x_j is the output of the jth element in the input layer, and θ_i is the weight of the biases between the input and hidden layers. It is noted that number k of the hidden layers in a three layers network model can be described as follows [5]:

$$k = 2 * n + j, j = 0, \cdots, 8. \tag{10}$$

The logistic $sigmoid$ transfer function is used as the activation function. By modifying the connection weight and valve weight in training process of the network model, the the anticipated output can be regulated. In impeller milling operations, the relationship between cutting chatter state and vibration signals under different milling operations is generally nonlinear, so the BPNN can be used to identify and recognize the cutting chatter state.

3.3.2 Genetic algorithm

The GA in particular became popular through the work of John Holland [21]. Inspired by natural selection process, GAs are commonly used to generate high-quality solutions to constrained and unconstrained optimization and search problems by relying on bio-inspired operators such as mutation, crossover, and selection [33]. During these operations, the GA repeatedly modifies a population of individual solutions. At each step, the GA selects individuals at random from the current population to be parents and uses them to produce the children for the next generation. Over successive generations, the optimal population is evolved.

In past years, the GA has been widely used in many engineering applications, such as condition monitoring [8], control system [20], multi-objective optimization [12, 28, 51], and assembly sequence planning [47], etc.

The BPNN-GA model is proposed to identify cutting chatter states. The BPNN model has the capacity to solve complex nonlinear problems, prediction, and optimization, while the GA has the ability to achieve multi-input-parameters optimization in many fields. In this paper, the GA is employed to optimize the input parameters and results by choosing the BPNN model to prediction results as the fitness function. A flow chart shows the optimal process of identifying cutting chatter states in Fig. 7.

4 Results and discussion

4.1 Study of milling operations

In impeller milling operations, with the operating conditions and cutting parameters such as spindle speed, feed rate, and cutting depth adjustment, cutting chatter state can be categorized into three types: stable state, transition state, and chatter state. Typically, this phenomenon can be illustrated using its time-frequency characteristics. On the one hand, the amplitude of vibration signal in time domain increases as shown in Fig. 3. The vibration energy increases gradually with time. On the other hand, the main frequency band of the vibration signal in milling operations changes in frequency domain as shown in Fig. 8.

The power spectrum of results in Fig. 8 represent three states in Fig. 3 respectively. It is found that the main frequency band of the vibration signal moves from the high-frequency domain to the low-frequency domain.

The frequency of vibration signal in stable state can be calculated by

$$\omega = (n \times N)/60, \tag{11}$$

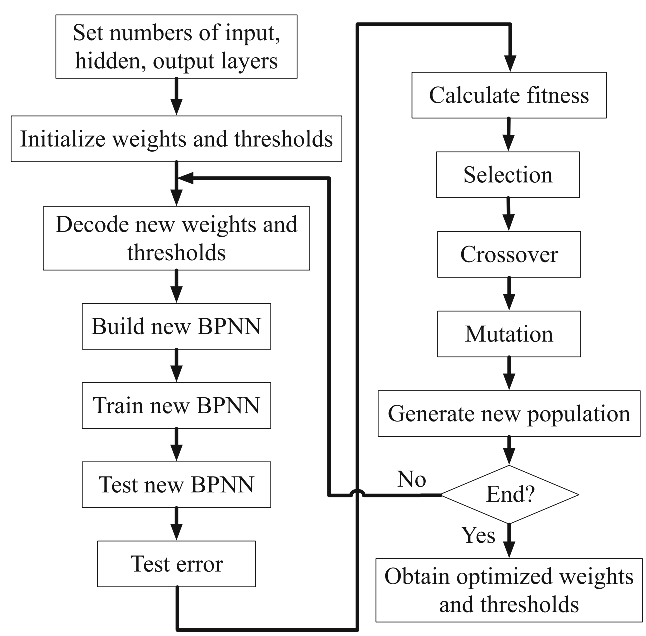

Fig. 7 Structure of back propagation neural network model

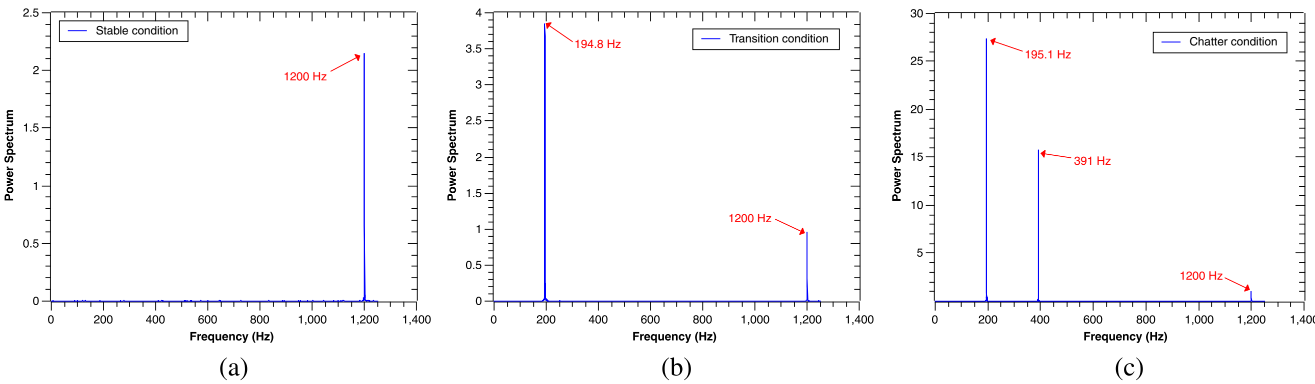

Fig. 8 Vibration signals in **a** stable state, **b** transition state, and **c** chatter state

where n is the spindle speed, N is the tool teeth. Based on the operation condition in Fig. 3, result of Eq. 11 equal to $\omega = (n \times N) / 6 = (120{,}000 \times 6) / 60 = 1200$ Hz, exactly the same as the power spectrum result in Fig. 8a. In the transition state of milling operation, the frequency of the vibration signals begins to appear at about 194.8 Hz and its amplitude also increased obviously as shown in Fig. 8b. In the chatter state of milling operation, another frequency component of 391 Hz occurs with its amplitude increasing sharply as shown in Fig. 8c. Specifically, the frequency components of 194.8 and 391 Hz correspond to the first and second natural frequencies of the milling system. That is, energy of the vibration signal shifts from the main frequency (1200 Hz) to natural frequencies (194.8 and 391 Hz). So this phenomenon provides a reliable basis for proposed HHCM method.

It should be mentioned that if the spindle speed is around 2000 rpm, the corresponding main frequency is about 200 Hz based on Eq. 11. It is close to the natural frequencies (194.8 Hz, 391 Hz). Obviously, this setting may cause resonance vibration, which is a severe fatigue damage to the system. Generally, it is need to figure out the normal operating range of the milling process system before the experiments. In this work, the cutting frequency is set to be far away from the natural frequencies.

4.2 Parameter setting of the VMD

Output of the VMD is affected by typical selection of K and α. Generally, based on the power spectrum under the machine state, the number of interest frequencies helps to determine the parameter K. Using prior knowledge, parameter K cannot be less than the number of interest frequencies. At the same time, parameter α determines the noise level of the decomposed component in frequency domain. It also affects bandwidth of the matching center frequency. For example, one or several additional modes would greatly contain noise for small α and large K, or for large α and small K. While, significant parts of the spectrum are shared by two or more different modes and their center frequencies overlap for small α and small K, or for large α and large K [49]. Specifically, a small value of parameter α will be used to detect impacts [42].

Based on the discussion in Section 4.1, spindle speed determines the main frequency of vibration signal in stable state. However, for a certain milling process, the natural frequencies of vibration signals in transition and chatter states are unchanged. Thus, to some extent, the spindle speed affects the frequency distributions in transition and chatter states. In turn, the frequency distributions will affect the selection of balancing parameter α. For example, if the main frequency is close to the natural frequencies (194.8 Hz, 391 Hz), it will limit the bandwidth of the estimated mode, which is related to the balancing parameter α.

Based on the power spectrum of the chatter state, there are three interest frequencies, that is, the natural frequencies (194.8 Hz, 391 Hz) related with the certain milling process, and the main frequency impacted by the spindle speed. Thus, the modes K is greater than or equal to three. In this work, exact decomposition or reconstruction is not the goal. The estimated modes u_k is used for calculating the power spectrum, which is the inputs of system model. When parameter K is set by three, four, or five, the calculated power spectrum also can be used to build the system model. That is, this method still works if the K is selected as other values. But it does not seem to be the suitable value. Generally, a certain number of the inputs will help to build the system model exactly based on the simulation results and experience. It is noted that too many modes will cause over-segmentation phenomenon.

Based on the frequency component of vibration signals and prior knowledge, quite suitable parameters K and α were set by 6 and 2000 in this paper. The Lagrangian multiplier was effectively shut off. The parameter ω was set to be one, which suggests center frequencies of all the modes were initialized in the uniform distribution. And no DC part was imposed.

4.3 Learning of vibration signals

The vibration signals in chatter state shown in Fig. 8c was decomposed by the VMD. It is convenient to identify the frequency components of the vibration signal. The results of each component and corresponding Fourier spectra are shown in Fig. 9. Six modes were recovered by the VMD from low frequency to high frequency. Similarly, the sub-signals in stable and transition state shown in Fig. 8a, b were also processed by the VMD.

Besides, the normalized energy ration of Shannon power spectral entropy of the sub-signals in different states are shown in Fig. 10. In the stable state, energy of the vibration signal mostly concentrates on 6th mode component, which is exactly the frequency of stable vibration signal. As cutting chatter occurs, the energy entropies in 2nd and 3rd mode components increase dramatically compared with other modes. That is, energy of the 6th mode component is converted to 2nd and 3rd mode in transition and chatter states, respectively. Simulation results above are consistent with the mentioned phenomenon that energy of the vibration signal shifts from the main frequency (1200 Hz) to natural frequencies (194.8 and 391 Hz) of milling system.

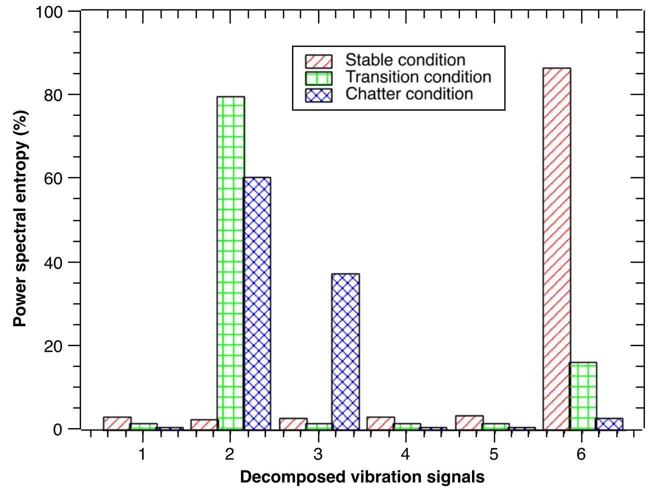

Fig. 10 Normalized energy ration of decomposed vibration signals in stable, transition, and chatter states

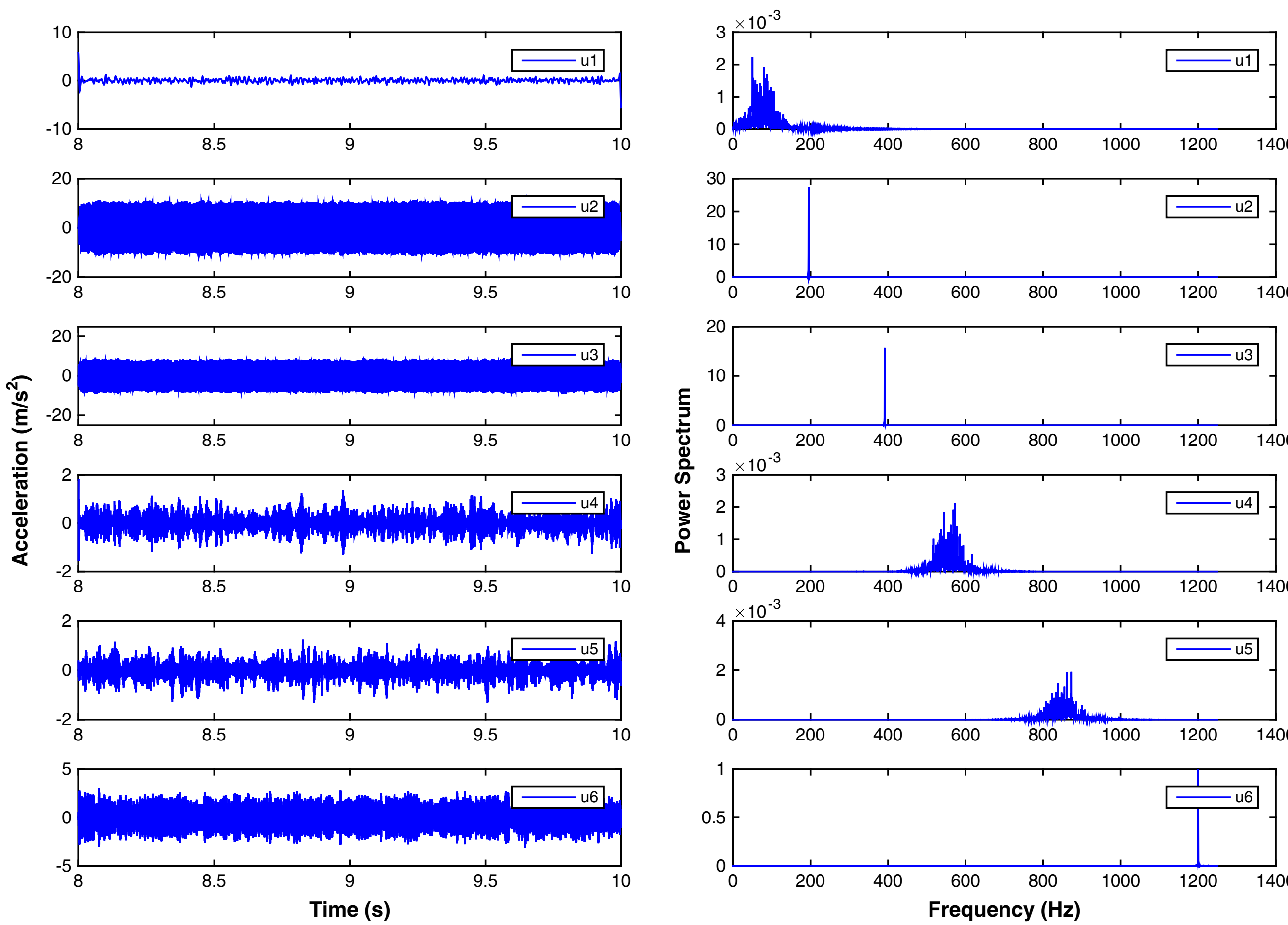

Fig. 9 Result of variational mode decomposition

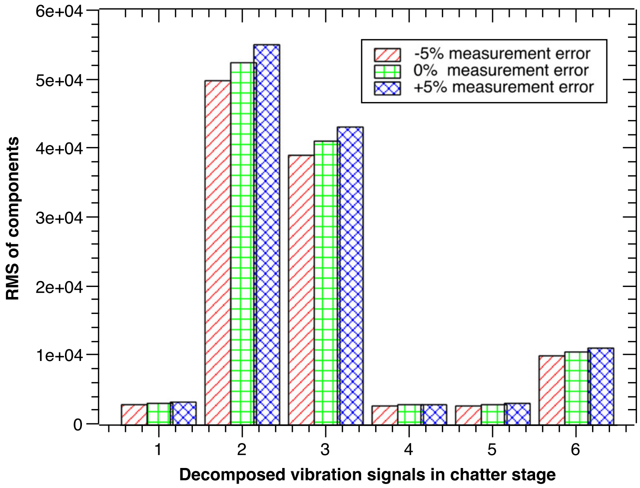

Fig. 11 RMS of decomposed vibration signals in chatter state

Typically, in order to overcome measurement errors as the source of epistemic uncertainty, each value of experimental data is considered with an error of ±5 or ±10%, and then the RMS of decomposed components at different states are taken as the feature observations vector [22, 46] as shown in Fig. 11. The decomposed components is usually calculated by this feature extraction method twice. Compared with methods used in previous work, the Shannon power spectral entropy can deal with uncertainty problem in measurement conveniently and get stable corresponding results, with less calculation time of feature extraction.

Furthermore, if the spindle speed is at some other value that does not exist in the training data, energy of the vibration signal will be also converted to the natural frequencies in the chatter state. That is, the energy entropies in 2nd and 3rd mode components are dominant in the chatter state. Thus, the proposed method will still work if chatter occurs.

The above results show that the feature extraction method based on the VMD and Shannon power spectral entropy can give an alternative option to milling condition monitoring.

4.4 BPNN setting and optimization

The proposed feature extraction method based on the VMD and Shannon power spectral entropy can be used to calculate inputs of back propagation neural network (BPNN) model. So number n of input layers was 6. The number k of the hidden layers was set by Eq. 9, whereas choosing $j = 1$. So $k = 2 \times n + j = 2 \times 6 + 1 = 13$. As milling operation conditions can be categorized by three states, outputs of BPNN were set by [1, 0, 0], [0, 1, 0], [0, 0, 1] corresponding to stable state, transition state, and chatter state, respectively. Based on the structure of BPNN model, the total number of weights was set by $n \times k + k \times m = 6 \times 13 + 13 \times 3 = 117$, the total number of thresholds was $k + m = 13 + 3 = 16$. Specific settings of the BPNN model is depicted by Table 2. While the Levenberg-Marquardt algorithm was chosen to training the BPNN.

Table 2 Parameter setting of BPNN

Type	Input layers	Hidden layers	Output layers	Weights	Threshold
Number	6	13	3	117	16

Table 3 Parameter settings of GA

Type	Population size	Maximum generation	Crossover probability	Mutation probability
Value	40	50	0.7	0.01

Generally, initial weights and thresholds of neural network model are set by the random number in the interval $[-0.5, 0.5]$. These initial values of weights and thresholds have a great impact on training of BPNN, but also can not be obtained accurately. As a commonly used optimization method, the GA was selected to optimize the initial values of weights and thresholds (Table 3).

4.5 Performance analysis

In order to verify the effectiveness of proposed HHCM method, Kennard and Stone algorithm [10, 30] can be used to choose training and testing date sets from 27 groups of raw data sets without loss of generality as shown in the Table 4.

Comparing results in 3rd column with 4th column in Table 5, prediction results in 4th column are much closed to proposed results in 2nd column. That is, the proposed HHCM method using optimized initial weights and thresholds has an enhanced ability to identify and classify cutting chatter states. Furthermore, as shown in results in 2nd column and 3rd column in Table 6, simulation errors of training data and testing data decreases from 0.3401, 0.6267 to 0.1423, 0.0335, respectively. It also verifies that with the optimized BPNN model, the ability of state classification is improved.

To demonstrate the improvement of proposed method compared to previous work [46], a three-level wavelet packet decomposition (WPD) was used to decompose vibration signals and then the root mean square (RMS) of

Table 4 Result of Kennard and Stone algorithm

Type	Experiment No.
Training samples	1, 2, 5, 6, 12, 14, 15, 17, 18, 19, 21, 23, 25
Testing samples	3, 4, 7, 8, 9, 10, 11, 13, 16, 20, 22, 24, 26, 27

Table 5 Prediction results using different algorithms

No.	Proposed output	VMD-BPNN	VMD-BPNN-GA	WPD-BPNN [46]	WPD-BPNN-GA [46]
3	[0 0 1]	[0.0571 0.1035 0.7851]	[0.0002 0.0030 0.9999]	[0.0040 0.1121 0.8970]	[0.0010 0.0158 0.9713]
4	[0 0 1]	[0.0587 0.1182 0.7116]	[0.0024 0.0217 0.9955]	[0.0068 0.1146 0.8595]	[0.0017 0.0127 0.9493]
7	[0 0 1]	[0.0699 0.0626 0.8129]	[0.0012 0.0013 0.9998]	[0.0078 0.1690 0.8316]	[0.0013 0.0131 0.9640]
8	[1 0 0]	[0.9876 0.0105 0.0026]	[0.9989 0.0057 0.0074]	[0.9969 0.0035 0.0006]	[0.9828 0.0091 0.0017]
9	[0 0 1]	[0.0670 0.0921 0.7754]	[0.0006 0.0018 0.9999]	[0.0032 0.2043 0.8425]	[0.0004 0.0317 0.9725]
10	[0 0 1]	[0.0857 0.0624 0.7850]	[0.0025 0.0007 0.9998]	[0.0039 0.1800 0.8574]	[0.0005 0.0269 0.9693]
11	[0 0 1]	[0.0952 0.0228 0.8702]	[0.0075 0.0003 0.9997]	[0.0053 0.1134 0.9109]	[0.0008 0.0076 0.9867]
13	[1 0 0]	[0.9881 0.0096 0.0027]	[0.9990 0.0062 0.0064]	[0.9979 0.0009 0.0001]	[0.9527 0.0302 0.0005]
16	[0 1 0]	[0.0433 0.9956 0.0094]	[0.0130 0.9993 0.0016]	[0.0013 0.9868 0.0256]	[0.0025 0.9487 0.0067]
20	[0 0 1]	[0.0760 0.0380 0.8517]	[0.0010 0.0008 0.9999]	[0.0049 0.2692 0.6639]	[0.0011 0.0173 0.9617]
22	[1 0 0]	[0.9712 0.0243 0.0023]	[0.9970 0.0161 0.0032]	[0.9905 0.0389 0.0001]	[0.9741 0.0339 0.0004]
24	[0 0 1]	[0.0805 0.0252 0.8828]	[0.0018 0.0004 0.9999]	[0.0052 0.1856 0.7747]	[0.0010 0.0094 0.9786]
26	[1 0 0]	[0.9732 0.0182 0.0027]	[0.9974 0.0172 0.0026]	[0.9940 0.0248 0.0001]	[0.9890 0.0032 0.0092]
27	[0 1 0]	[0.0275 0.9958 0.0083]	[0.0080 0.9993 0.0026]	[0.0018 0.0026 0.3584]	[0.0017 0.9548 0.0668]

the wavelet coefficients was used to extract feature from decomposed vibration signals. The eight RMS output were set as inputs of neural network model. Similarly, settings of the BPNN and GA can be referred to the last subsection.

Simulation results of the wavelet packet decomposition-based back propagation model, namely WPD-BPNN, have been displayed in 5th column of Table 5. At the same time, the GA has also been used to optimize WPD-BPNN model used in previous work and corresponding results of the WPD-BPNN-GA method are shown in 6th column. And simulation error of training data and testing data using previous method have been displayed in the 4th and 5th column of Table 6. For example, simulation output of No. 27 experiment using the WPD-BPNN method is not close to the proposed output. At same time, simulation error of testing samples using the WPD-BPNN is much bigger than the VMD-BPNN.

There are two reasons for the bad result. First, the decomposition levels of the wavelet transform limit the extraction of the interest frequencies. The WPD method, as an expansion of classical wavelet decomposition, can split the signal into same-length frequency bands both in low-pass and high-pass bands. In the experiments, the spindle speed affects the distribution of the main frequency and natural frequencies. In this work, different spindle speeds had been used. The distribution of the frequency domain is quite different. Using a fixed decomposition level, the WPD method may not extract the interest frequencies at all. It affects the result of the state classification. For three levels of decomposition, the WPD produces eight different set of coefficients. The interest frequencies locate into the decomposed bands by the three-level WPD. Obviously, four-level WPD have a better decomposition result. But the computation time increases.

Second, system model used in the WPD-BPNN method is not optimized. For example, using the three-level WPD, the total number of the initial weights and thresholds of neural network model is greater. In addition, these initial values have a great impact on training of the BPNN model. Without the optimization process, these initial values were only set by the random number in the interval [−0.5,0.5]. Thus, the BPNN model used in the WPD-BPNN method cannot identify and classify the machine states.

So traditional WPD-BPNN method has a poor pattern identification performance than proposed VMD-BPNN method. That is, ability of feature extraction using the VMD in this paper are quite effective than the previous work. Furthermore, based on comparison of the simulation errors between the WPD-BPNN-GA and the VMD-BPNN-GA as shown in the 3rd and 5th column of Table 6, prediction results using the VMD-BPNN-GA is much smaller than the WPD-BPNN-GA. So the proposed HHCM method is verified to be a better choice than traditional method.

Table 6 The two-norm difference between prediction results and proposed results using different algorithms

Type	VMD-BPNN	VMD-BPNN-GA	WPD-BPNN [46]	WPD-BPNN-GA [46]
Simulation error of training samples	0.3401	0.1423	0.3106	0.34404
Simulation error of testing samples	0.6267	0.0335	1.2527	0.13851

Based on the simulation results above, cutting chatter states of the milling operations in Table 1 have been recognized correctly by proposed HHCM method. As a widely used neural network, the BPNN-GA model has a good performance in pattern identification. The proposed HHCM method can satisfy the requirements of milling system exactly and deal with problems of pattern identification and state classification. Further the basic frame of the proposed HHCM method can also be widely used in other fields.

Though the proposed HHCM method are more accurate and relative reliable than previous method, the optimization process cost too much computation time. For example, it usually takes about several minutes to build the system model in the VMD-BPNN-GA method. Thus, the efficiency of the proposed method should be to be raised. In the future, much emphasis will be focused on the realization of in-situ real-time process monitoring and process control of milling operations. On the one hand, calculation time used in preprocessing, feature extraction, and model optimization can be decreased. A better optimization method or classification model can be adopted to realize and improve ability of pattern identification. On the other hand, the proposed method should be sensitive and robust to the abnormal working range of the milling process. For example, when the cutting frequency is close to the chatter frequency, the proposed method should output an indication to void the resonance vibration, which is a severe fatigue damage to the milling system.

5 Conclusions

In this paper, we proposed a HHCM method using the VMD and the BPNN-GA model for cutting chatter detection and state classification in milling operations. Vibration signal is decomposed into multiple mode components by the VMD, which overcomes the disadvantage of lacking theoretical basis and noise sensitivity of traditional decomposition methods. Then, Shannon power spectral entropy is adopted to extract features from decomposed vibration signals. These features are closely related to states of cutting chatter. Furthermore, states of cutting chatter in impeller milling operations is identified and classified by the BPNN-GA model. Experimental results have showed that the proposed HHCM method are more accurate and relative reliable than previous methods.

The proposed HHCM method is the premise and foundation for the realization of in-situ real-time process monitoring and process control of milling operations in the future. It can be applied to post-analysis in milling operations, such as cutting chatter prediction and cutting chatter suppression. In addition, it also provides a basic frame to deal with similar problems in other machine operations and may be attractive for other application fields, such as rolling bearing, gearbox, wind generation, etc.

Acknowledgments The work here is supported in part by China Scholarship Council with a Scholarship (No. 201606160048), the National Science and Technology Supporting Plan (No. 2015BAF01B04), the National Natural Science Foundation of China (No. 51175208), and the State Key Basic Research Program of China (No. 2011CB706903).

References

1. Abellan-Nebot JV, Subirón FR (2010) A review of machining monitoring systems based on artificial intelligence process models. Int J Adv Manuf Technol 47(1-4):237–257
2. Altintaş Y., Budak E (1995) Analytical prediction of stability lobes in milling. CIRP Ann-Manuf Technol 44(1):357–362
3. Altintas Y, Eynian M, Onozuka H (2008) Identification of dynamic cutting force coefficients and chatter stability with process damping. CIRP Ann-Manuf Technol 57(1):371–374
4. Altintas Y, Weck M (2004) Chatter stability of metal cutting and grinding. CIRP Ann-Manuf Technol 53(2):619–642
5. Bin G, Gao J, Li X, Dhillon B (2012) Early fault diagnosis of rotating machinery based on wavelet packets—empirical mode decomposition feature extraction and neural network. Mech Syst Signal Process 27:696–711
6. Cao H, Lei Y, He Z (2013) Chatter identification in end milling process using wavelet packets and hilbert–huang transform. Int J Mach Tools Manuf 69:11–19
7. Cao H, Zhou K, Chen X (2015) Chatter identification in end milling process based on eemd and nonlinear dimensionless indicators. Int J Mach Tools Manuf 92:52–59
8. Cerrada M, Zurita G, Cabrera D, Sánchez R. V., Artés M., Li C (2016) Fault diagnosis in spur gears based on genetic algorithm and random forest. Mech Syst Signal Process 70:87–103
9. Chiou RY, Liang SY (2000) Analysis of acoustic emission in chatter vibration with tool wear effect in turning. Int J Mach Tools Manuf 40(7):927–941
10. Daszykowski M, Walczak B, Massart D (2002) Representative subset selection. Anal Chim Acta 468(1):91–103
11. Daubechies I (1990) The wavelet transform, time-frequency localization and signal analysis. IEEE Trans Inf Theory 36(5):961–1005
12. Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: Nsga-ii. IEEE Trans Evol Comput 6(2):182–197
13. Dragomiretskiy K, Zosso D (2014) Variational mode decomposition. IEEE Trans Signal Process 62(3):531–544
14. Eynian M (2015) Vibration frequencies in stable and unstable milling. Int J Mach Tools Manuf 90:44–49
15. Feng J, Sun Z, Jiang Z, Yang L (2016) Identification of chatter in milling of ti-6al-4v titanium alloy thin-walled workpieces based on cutting force signals and surface topography. Int J Adv Manuf Technol 82(9-12):1909–1920
16. Fu Y, Zhang Y, Zhou H, Li D, Liu H, Qiao H, Wang X (2016) Timely online chatter detection in end milling process. Mech Syst Signal Process 75:668–688
17. Gan M, Wang C, et al. (2016) Construction of hierarchical diagnosis network based on deep learning and its application in the fault pattern recognition of rolling element bearings. Mech Syst Signal Process 72:92–104
18. Ghosh N, Ravi Y, Patra A, Mukhopadhyay S, Paul S, Mohanty A, Chattopadhyay A (2007) Estimation of tool wear during cnc

milling using neural network-based sensor fusion. Mech Syst Signal Process 21(1):466–479
19. Govekar E, Gradišek J, Grabec I (2000) Analysis of acoustic emission signals and monitoring of machining processes. Ultrasonics 38(1):598–603
20. Grefenstette JJ (1986) Optimization of control parameters for genetic algorithms. IEEE Trans syst, Man, Cybern 16(1):122–128
21. Holland JH (1975) Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence. U Michigan Press
22. Hu Y, Xie F, Wu B, Wang Y (2013) An uncertainty quantification method based on generalized interval. In: Artificial Intelligence (MICAI), 2013 12th Mexican International Conference on IEEE, pp 145–150
23. Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, Yen NC, Tung CC, Liu HH (1998) 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. The Royal Society, pp 903–995
24. Huang P, Li J, Sun J, Zhou J (2013) Vibration analysis in milling titanium alloy based on signal processing of cutting force. Int J Adv Manuf Technol 64(5-8):613–621
25. Hwang R, Cha K (2008) Construction of a prediction model for the structural stability of a surface grinder using backpropagation neural network. Int J Adv Manuf Technol 37(11-12):1093–1104
26. Insperger T, Mann BP, Stépán G, Bayly PV (2003) Stability of up-milling and down-milling, part 1: alternative analytical methods. Int J Mach Tools Manuf 43(1):25–34
27. Jia G, Wu B, Hu Y, Xie F, Liu A (2013) A synthetic criterion for early recognition of cutting chatter. Sci China Technol Sci 56(11):2870–2876
28. Jiang P, Cao L, Zhou Q, Gao Z, Rong Y, Shao X (2016) Optimization of welding process parameters by combining kriging surrogate with particle swarm optimization algorithm. Int J Adv Manuf Technol 86(9-12):2473–2483
29. Jin C, Wu B, Hu Y, Cheng Y (2012) Identification of thermal error in a feed system based on multi-class ls-svm. Front Mech Eng 7(1):47–54
30. Kennard RW, Stone LA (1969) Computer aided design of experiments. Technometrics 11(1):137–148
31. Mannan M, Broms S, Lindström B. (1989) Monitoring and adaptive control of cutting process by means of motor power and current measurements. CIRP Ann-Manuf Technol 38(1):347–350
32. Minis I, Yanushevsky R (1993) A new theoretical approach for the prediction of machine tool chatter in milling. J Eng Indust 115(1):1–8
33. Mitchell M (1998) An introduction to genetic algorithms. MIT press
34. Munoa J, Beudaert X, Dombovari Z, Altintas Y, Budak E, Brecher C, Stepan G (2016) Chatter suppression techniques in metal cutting. CIRP Ann-Manuf Technol 65(2):785–808
35. Nair U, Krishna BM, Namboothiri V, Nampoori V (2010) Permutation entropy based real-time chatter detection using audio signal in turning process. Int J Adv Manuf Technol 46(1-4):61–68
36. Pal SK, Chakraborty D (2005) Surface roughness prediction in turning using artificial neural network. Neural Comput Appl 14(4):319–324
37. Quintana G, Ciurana J (2011) Chatter in machining processes: A review. Int J Mach Tools Manuf 51(5):363–376
38. Rumelhart DE, McClelland JL, PDP Research Group C (eds) (1986) Parallel Distributed Processing: Explorations in the Microstructure of Cognition, vol 1. Foundations. MIT Press, MA, USA
39. Siddhpura M, Paurobally R (2012) A review of chatter vibration research in turning. Int J Mach Tools Manuf 61:27–47
40. Tlusty J, Polacek M (1963) The stability of machine tools against self-excited vibrations in machining. Int Res Prod Eng 1(1):465–474
41. Tsai NC, Chen DC, Lee RM (2010) Chatter prevention for milling process by acoustic signal feedback. Int J Adv Manuf Technol 47(9-12):1013–1021
42. Wang Y, Markert R (2016) Filter bank property of variational mode decomposition and its applications. Signal Process 120:509–521
43. Wang Y, Markert R, Xiang J, Zheng W (2015) Research on variational mode decomposition and its application in detecting rub-impact fault of the rotor system. Mech Syst Signal Process 60:243–251
44. Wu Z, Huang NE (2009) Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv Adapt Data Anal 1(01):1–41
45. Xie F, Wu B, Hu Y, Wang Y (2013) A generalized markov chain model based on generalized interval probability. Sci China Technol Sci 56(9):2132–2136
46. Xie F, Wu B, Hu Y, Wang Y, Jia G, Cheng Y (2014) A generalized interval probability-based optimization method for training generalized hidden markov model. Signal Process 94(1): 319–329
47. Xin L, Jianzhong S, Yujun C (2016) An efficient method of automatic assembly sequence planning for aerospace industry based on genetic algorithm. Int J Adv Manuf Technol, pp 1–9
48. Yu Y, Junsheng C et al. (2006) A roller bearing fault diagnosis method based on emd energy entropy and ann. J Sound Vib 294(1):269–277
49. Zhang S, Wang Y, He S, Jiang Z (2016) Bearing fault diagnosis based on variational mode decomposition and total variation denoising. Measur Sci Technol 27(7):075–101
50. Zhang Z, Li H, Meng G, Tu X, Cheng C (2016) Chatter detection in milling process based on the energy entropy of vmd and wpd. Int J Mach Tools Manuf
51. Zhou Q, Rong Y, Shao X, Jiang P, Gao Z, Cao L (2016) Optimization of laser brazing onto galvanized steel based on ensemble of metamodels. J Intell Manuf, pp 1–15