Chatter detection in milling process based on the combination of wavelet packet transform and PSO-SVM

Qingzhen Zheng · Guangsheng Chen · Anling Jiao

Received: 28 May 2021 / Accepted: 1 February 2022 / Published online: 11 February 2022
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2022

Abstract
Chatter is one of the biggest unfavorable factors during the high speed machining process of a machine tool. It severely affects the surface finish and geometric accuracy of the workpiece. To address this obstacle and improve the quality and efficiency of products, it is significantly essential to detect chatter during machining. Therefore, a multi-feature recognition system for chatter detection on the basis of the fusion technology of wavelet packet transform (WPT) and particle swarm optimization support vector machine (PSO-SVM) was proposed in this paper. Firstly, the original vibration signals collected from the acceleration sensor were processed through wavelet packet transform (WPT). The noise and the irrelevant information were remarkably decreased. In addition, the wavelet packets containing chatter-emerging information were chosen and reconstructed. The fourteen time–frequency domain characteristics of the reconstructed vibration signal were calculated and chosen as the multi-feature vectors of chatter detection. Finally, to obtain the optimal radial basis function parameter $g$ and penalty parameter $C$ of the SVM prediction model, the optimization algorithms of $k$-fold cross-validation ($k$-CV), genetic algorithm (GA), and particle swarm optimization (PSO) were employed in optimizing the model parameters of SVM. It was indicated that the PSO-SVM improved obviously the accuracy of chatter recognition than the others. In addition, we applied the optimized SVM prediction model by PSO for detecting chatter state in end milling machining. Chatter recognition results indicated that the model accurately predicted the slight chatter state in advance.

Keywords Chatter detection · Wavelet packet transform · Multi-feature vectors · PSO-SVM · Chatter recognition

1 Introduction
Chatter is one of the biggest unfavorable factors in achieving high performance machining, which is a self-excited vibration that happened between workpieces and cutting tools [1]. It occurs in any machine tooling process and directly affects the surface finish and geometry accuracy of the workpiece, seriously damaging the tool and reducing the life of the machine tools. Timely chatter is detected, which is a prerequisite for improving production efficiency and reducing manufacturing costs. However, the cutting process in the milling is non-stationary due to machine tool spindle wear, the change of operating temperature, workpiece stiffness, and other non-linear factors [2]. Therefore, with the cutting environment changing, the chatter detection and identification methods in the machining process have always been significantly critical problems.

Over recent decades, many researchers paid attention to chatter detection, which has been a research hotspot. In order to detect the phenomenon of chatter, some sensors were generally applied to obtain chatter signals, such as acceleration sensor, acoustic emission, current sensor, and microphone [3–6]. No matter which sensor is chosen, it is of great importance to guarantee that the extracted chatter indicators are sensitive and the designed chatter indicators are closely relevant. Ye et al. [7] extracted the root mean square (RMS) sequence of the real-time acceleration signals, and its coefficient of standard deviation to mean was designed as an indicator to distinguish the machining state. Tangjiticharoen [8] calculated the cumulative power spectrum density (PSD) by the collected three dynamic cutting forces and used its ratios for detecting in-process chatter states during NC turning. In addition, a multi-sensor fusion technique was utilized for extracting chatter features, ensuring that chatter detection is robust and reliable regardless of
variable cutting conditions. Kuljanic et al. [9] investigated the sensibility of chatter onset of several sensors and found that three or four sensors were the most promising solution for reliable and robust chatter identification. Pan et al. [10] used the multi-sensor fusion technique and manifold learning for chatter detection during boring and found that the extracted multi-features can improve the recognition rate. However, some sensors, such as force sensor, maybe be not suitable for the practical machining process. For example, to ensure the reliability and accuracy of measurements, the acoustic emission needs to be close to the machining area between the cutting tool and workpiece [11]. Moreover, the installation and cost of the displacement and force sensor may be difficult and huge, respectively. Wan et al. [12] selected 8 time–frequency domain characteristics and 8 automatic characteristics extracted by stacked-denoising autoencoder as chatter indicators, and they found that the accuracy and reliability of chatter detection in milling was greatly improved.

During the actual cutting process, the acquired signals contain a lot of noise. For extracting the chatter-sensitive characteristics, the signal processing technique is particularly crucial. The proper method for processing time-varying non-stationary signal, including Wigner–Ville distribution, Hilbert–Huang transform, short-time Fourier transform (STFT), wavelet transform, and wavelet packet transform, effectively reduces the content of the noise. These methods effectively enhance the signal-to-noise ratio (SNR). Fu et al. [13] decomposed the collected acceleration signals into a sequence of intrinsic mode functions (IMFs) with using empirical mode decomposition (EMD) to quantize the spectrum characteristic for an online detection system. Ji et al. [14] investigated that ensemble empirical mode decomposition (EEMD) was adopted to treat the acceleration signals and selected the IMFs with chatter information of the milling process to detect milling chatter timely. Although EEMD could address the issue of mode mixing, the application of EMD and EEMD is still restricted owing to lacking theoretical foundation [15]. Wavelet packet transform (WPT) is an effective signal processing method which is especially used to deal with the non-stationary signal. Compared to short-time Fourier transform (STFT), WPT overcomes the drawbacks in high frequency signals, synchronously obtaining the high resolution at both low and high frequency signals. Hence, the measured signals in the milling process were preprocessed by WPT, which effectively extracted the certain frequency band with rich chatter information. Cao et al. [16] applied WPT as a preprocessor to eliminate noise in measurement signals. The performance of the Hilbert–Huang transform (HHT) was enhanced, and its mean and standard deviation were adopted to identify the chatter state at the end milling process. Yao et al. [17] combined wavelet analysis and wavelet packet analysis for processing the measured acceleration signals, and standard deviation and energy ratio were correspondingly extracted as chatter indicators for detection.

Although extracting features in signals with preprocessing presents the stability degree of the machining condition, the threshold methods are usually applied to detect chatter state [18]. However, the threshold approaches required individual experience and inherent mechanism of the dynamic system of machine tools in advance, which are not suitable for industrial applications. Therefore, additional efforts are required to implement intelligent monitoring systems of chatter detection. Through a learning process, the stable and unstable states in the cutting process are significantly discriminated. Several recognition techniques such as neural network, hidden Markov models (HMM), and fuzzy logic have been utilized to detect variable machining conditions. Teti et al. [19] investigated that the neural network and fuzzy logic approaches were extensively applied for monitoring the cutting state. Zhang et al. [20] proposed a hybrid approach of combining hidden Markov model (HMM) and artificial neural network for detecting cutting chatter and found that the cutting chatter was detected timely. However, these recognition techniques need a large amount of samples to ensure the recognition accuracy of chatter.

Alternatively, support vector machine (SVM), a popular supervised machine learning approach, owns the greatest generalization ability and minimizes the classification error. This algorithm can solve the problems of classification and regression for a small sample. Therefore, SVM has been extensively applied for identifying chatter in the milling process [21–23]. However, SVM has the disadvantage of limiting a wide range of industrial applications. For example, the selection of kernel function parameter and penalty parameter of SVM seriously impacts on the recognition rate of the SVM classifier. Therefore, for achieving the optimal kernel function parameter and penalty parameter, several optimization algorithms are employed for SVM. Peng et al. [24], used the k-fold cross-validation method to optimize the gamma parameter of radial basis functions and penalty parameter C. But this approach is a local search strategy which leads to that the SVM classifier is prone to falling into the local minimum [25, 26]. SVM parameters were generally optimized by a genetic algorithm (GA) for improving classification accuracy for monitoring wheel wear, which achieved a great performance [27]. Jia et al. [28] adopted the genetic support vector machine (GA-SVM) to monitor the trend of tool wear during the deep-hole drilling. In addition, particle swarm optimization (PSO) is another optimization algorithm for optimizing SVM parameters, successfully applying for the bearing fault diagnosis [4, 29]. Even though both GA and PSO have
Figure 1 describes the scheme of the proposed approach of chatter detection in this study. The collected vibration signals were decomposed and reconstructed by using WPT, and the energy ratio was chosen as the choice criterion for selecting a characteristic wavelet packet containing chatter-emerging frequency. Subsequently, fourteen time–frequency features were selected as recognition parameters of chatter by calculating the reconstructed characteristic wavelet packets, which formed feature vectors. Finally, parameters of the SVM prediction model were optimized by PSO for identifying test data in end milling machining. Chatter recognition results have indicated that the model accurately predicted the slight chatter state in advance.

2 Proposed methodology of chatter detection

The distribution and amplitude of the collected vibration signals will change as the cutting state changes. Therefore, we can determine whether chatter occurs through detecting the amplitude and distribution of the signal. We supposed that there is a vibration signal \( x_i (i = 1, 2, \ldots, n) \), where \( n \) represents the number of collected data points. Ten features in the time-domain were chosen for identifying chatter state [12, 33], as displayed in Table 1. From Table 1, \( x_m, x_p, \) and \( x_{rms} \) stood for the energy and amplitude of vibration signals in the time-domain. \( x_{std}, x_{ske}, x_{kur}, CF, CLF, SF, \) and \( IF \) reflected the distribution of vibration signals in the time-domain.

Furthermore, on account of time-domain features extracted above cannot directly acquire some potential chatter information, this paper also introduces some frequency-domain features for flutter detection [12] because the amplitude and distribution of frequency components of vibration signal may change with the occurrence of chatter when chatter occurs [13]. In this regard, we selected four frequency-domain parameters as feature indicators of chatter [33, 34], as described in the following.

1. Mean square frequency: \( MSF = \frac{\sum_{m=1}^{n} f_{j}^2 S(j)}{\sum_{m=1}^{n} S(j)} \)
2. One-step autocorrelation function: \( \rho = \frac{\sum_{i=1}^{m} \cos(2\pi f_{j}\Delta t)S(f_{j})}{\sum_{i=1}^{m} S(f_{j})} \)

3. Frequency center: \( FC = \frac{\sum_{i=1}^{m} \cos(f_{j}-FC)S(f_{j})}{\sum_{i=1}^{m} S(f_{j})} \)

4. Standard frequency: \( FV = \frac{\sum_{i=1}^{m} (f_{j}-FC)^{2}S(f_{j})}{\sum_{i=1}^{m} S(f_{j})} \)

where \( S(f_{j}) \) is the power spectrum amplitude that is analyzed through FFT, and \( f_{j} = 1, \ldots, m \) described the \( j \)-th frequency of vibration signal. \( MSF \) is the energy of the frequency-domain signal. \( \rho \) and \( FC \) described the location of the dominant frequency band. \( FV \) represented the energy distribution of the frequency-domain signal.

However, to obtain these frequency-domain features, an amount of time need to spend on fast Fourier transform (FFT). To solve this issue, a fast calculation criterion was employed [35, 36], and four frequency-domain features can be replaced, as shown in Table 2. \( \dot{x}_{i} \) is the first-order difference equation, and \( \dot{x}_{i} = (x_{i} - x_{i-1})/\Delta t \), and \( \Delta t \) is the sampling interval. Hence, fourteen time–frequency features were selected as recognition parameters of chatter in this study, which formed feature vectors.

### 2.2 Wavelet packet transform

However, if the collected vibration signal contains noise, the recognition rate of chatter will be seriously affected. Therefore, the elimination and suppression of noise is significantly important for extracting chatter indicators. Wavelet packet transform (WPT) is an optimal candidate to solve this issue [16]. WPT is performed by using the basic two-channel filter bank that is iterated on the low-pass branch or the high-pass branch. Hence, it simultaneously decomposes the low frequencies and the high frequencies to improve the time–frequency resolution. The collected vibration signals are pretreated by using WPT, and a series of narrow bands are available. At this moment, the energy of broadband noise will be uniformly distributed in these narrow bands. The specific frequency bands containing chatter-emerging frequency were chosen, resulting in the enhancement of the signal-to-noise ratio to a certain extent.

The wavelet method uses a prototype function, which is called the mother wavelet, to shift and scale the signal. A time-domain signal is decomposed into a time–frequency scale. Supposing that there is a mother wavelet function \( \Psi(t) \in L^{2} \), and its shifted and scaled functions family is \( \Psi_{s,u}(t) = \Psi(t-u,s) \). The entire family of functions is produced through the dilatations or contractions of a modulated window \( \Psi(t) \), and the time translations of the family of functions are described by Eq. (1) [37, 38]:

\[
\Psi_{s,u}(t) = \frac{1}{\sqrt{s}} \Psi\left(\frac{t-u}{s}\right) \tag{1}
\]

Here \( s \) and \( u \) represent the scaling and position parameters, respectively. In the case of a continuous time signal \( x(t) \), continuous wavelet transform (CWT) could be presented in Eq. (2) [37, 38]:

\[
CWTx(s,u) = \langle x(t), \Psi_{s,u}(t) \rangle = \frac{1}{\sqrt{s}} \int_{-\infty}^{+\infty} x(t)\Psi^{*}\left(\frac{t-u}{s}\right)dt \tag{2}
\]

where \( CWTx(s,u) \) represents the inner product of signal \( x(t) \) and a family of shifted and scaled wavelets. CWT has a great ability to deal with various kinds of stationary and non-stationary signals, determining the temporal position of each frequency component. However, due to the low computational efficiency of CWT, it takes a lot of time to calculate. Therefore, CWT is not suitable for off-line applications.

| Feature                  | Expression                                      | Feature                  | Expression                                      |
|--------------------------|-------------------------------------------------|--------------------------|-------------------------------------------------|
| Mean                     | \( x_{m} = \frac{1}{n} \sum_{i=1}^{n} x_{i} \)  | Kurtosis                 | \( x_{kur} = \frac{\sum_{i=1}^{n} (x_{i} - \bar{x})^{4}}{(n-1)\bar{x}^{4}} \) |
| Standard deviation       | \( x_{std} = \sqrt{\sum_{i=1}^{n} (x_{i} - x_{m})^{2}} \) | Crest factor             | \( CF = \frac{x_{p}}{\bar{x}} \) |
| Root mean square         | \( x_{rms} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} x_{i}^{2}} \) | Clearance factor        | \( CLF = \frac{x_{p}}{\left( \sqrt{\sum_{i=1}^{n} |x_{i}|} \right)} \) |
| Peak                     | \( x_{p} = \max(|x_{i}|) \)                     | Shape factor             | \( SF = \frac{\sum_{i=1}^{n} |x_{i}|}{\left( \sqrt{\sum_{i=1}^{n} |x_{i}|} \right)} \) |
| Skewness                 | \( x_{skw} = \frac{\sum_{i=1}^{n} (x_{i} - \bar{x})^{3}}{(n-1)\bar{x}^{3}} \) | Impulse factor           | \( IF = \frac{\sum_{i=1}^{n} |x_{i}|}{\left( \sqrt{\sum_{i=1}^{n} |x_{i}|} \right)} \) |

| Feature                  | Expression                                      |
|--------------------------|-------------------------------------------------|
| Mean square frequency    | \( \text{MSF} = \frac{\sum_{i=1}^{m} x_{i}^{2}}{4\text{std}^{2} \sum_{i=1}^{m} x_{i}^{2}} \)  |
| One-step autocorrelation function | \( \rho = \frac{\sum_{i=1}^{m} x_{i} x_{i-1}}{\sum_{i=1}^{m} x_{i}^{2}} \)  |
| Frequency center         | \( FC = \frac{\sum_{i=1}^{m} x_{i} \Delta t}{4\text{std} \sum_{i=1}^{m} x_{i}^{2}} \)  |
| Standard frequency       | \( FV = \text{MSF} - 4\pi^{2} FC^{2} \) |

---

Table 1 Time-domain feature parameters [12, 33]

Table 2 Frequency-domain feature parameters [35, 36]
According to conjugate quadratic filters, Mallat developed a discrete wavelet transform (DWT) with a fast algorithm [39]. However, the high-frequency information is lost by using DWT. To overcome this drawback, wavelet packet transform (WPT) can decompose a discrete signal into a detail signal and approximation signal via high-pass and low-pass filters. Hence, the basic wavelet packet function is presented in the following [39]:

\[
\begin{align*}
  u_{2n}(t) &= \sqrt{2} \sum_{k \in \mathbb{Z}} h(k)u_n(2t - k) \\
  u_{2n+1}(t) &= \sqrt{2} \sum_{k \in \mathbb{Z}} g(k)u_n(2t - k)
\end{align*}
\]

Here, \( g \) and \( h \) and \( g \) represent high-pass and low-pass filter coefficients of wavelet decomposition, respectively. The relationship between them is orthogonal: \( g(k) = (-1)^j h(1 - k) \). \( u_0(t) \) and \( u_1(t) \) correspondingly represent the scaling function and wavelet function. The wavelet packet decomposition coefficients can be acquired by iterative calculation:

\[
\begin{align*}
  x_{2n,j+1}^k &= \sum_{l} h(l - 2k)x_{n,j} \\
  x_{2n+1,j+1}^k &= \sum_{l} g(l - 2k)x_{n,j}
\end{align*}
\]

where \( x_{n,j} \) is wavelet coefficient at the transformation level of \( j(j = 1, 2, \ldots) \); \( n \) stands for \( n \) sub-band; \( m \) stands for the number of wavelet coefficients. Supposing there is a discrete vibration signal \( x(t) \), a three-level decomposition of \( x(t) \) by the WPT method is described in Fig. 2. From Fig. 2, \( x_{i,j}(t) \) represents the \( j \)-th frequency band signal at level \( i \). And \( j(j = 1, 2, \ldots, J, \text{ and } J = 2^J) \) is the number of the decomposed frequency band.

Before the extraction of chatter features, it is necessary to accurately identify the frequency bands containing abundant chatter information of the decomposed wavelet packets. Therefore, the energy ratio of the wavelet packets in end milling machining is used to find out the chatter-emerging frequency band [40]. During the stable milling process, the energy of the wavelet packet node distributes over the entire frequency bands. While energy ratio of a certain wavelet packet will dramatically increase owing to the occurrence of chatter. So the relatively large energy ratio of wavelet packets was chosen. Chatter features are available by calculating the reconstructed wavelet packets.

### 2.3 Chatter recognition based on PSO-SVM

#### 2.3.1 The theory of SVM

Support vector machine (SVM) is a supervised machine learning approach according to statistical learning theory. Through finding the minimum structural risk, the generalization ability of machine learning is improved, decreasing confidence interval and empirical risk [26, 41]. SVM is utilized to find the optimal separating hyperplane in a feature space. This hyperplane could classify a training set. Therefore, the object of the optimization problem changes to discover an optimal hyperplane to minimize the margins from the hyperplane to the nearest training data. Two types of classification cases of linear SVM are displayed in Fig. 3.

![Fig. 3 Selected optimal hyperplane for linearly separable two-class data set](image_url)

Considering that two classes of data points in a set of samples: \( \{x_i, y_i\}, i = 1, 2, \ldots, n, x_i \in \mathbb{R}^n, y_i \in \{+1, -1\} \), \( x_i \) stands for the input vectors; \( y_i \) stands for class labels of \( x_i \); \( n \) represents the total number of samples. Figure 3 described that there are data points composed of different classes of feature vectors. Squares and circles denote class A and class B, respectively. Linear boundary \( B \) separates two data sets to guarantee the maximization of the margins between the nearest data points and the boundary of two classes. The nearest data points are defined as support vectors, which lie on the boundary \( H1 \) and \( H2 \). Generally, the larger margin can reduce the generalization error of the classifier.
The linear boundary H is expressed as follows:
\[ \omega \cdot x_i + b = 0 \]  
(7)

where \( b \) denotes the bias term, and \( \omega \) stands for weight vector that is perpendicular to the optimal hyperplane. The combination of \( \omega \) and \( b \) can determine the position of separating hyperplane. Therefore, the decision functions for classifying samples as class A or class B are as follows:
\[ f(x_i) = \text{sign}(\omega \cdot x_i + b) \]  
(8)

To achieve an accurate classification, the margin \( 2/\|\omega\|^2 \) between the two hyperplanes H1 and H2 should be maximized, which determines the generalization ability of the hyperplane. The constrained optimization problem changes to obtain the optimal hyperplane of data point of classification \([26]\):
\[ \text{Min} \ L(\omega, b, \xi) = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^{n} \xi_i \]  
Subject to \[ y_i(\omega^T \cdot g\varphi(x_i) + b) - 1 + \xi_i \geq 0, \ i = 1, 2, \ldots, n \]  
(9)

where \( C \) stands for penalty parameter that imposes a trade-off between training error and generalization, and \( \xi_i \) denotes slack variables \([25]\), and \( \xi_i > 0 \). \( \varphi(x_i) \) maps \( x_i \) from input space to the higher-feature space, which allows linear classification in higher dimensions \([42]\).

To address it, this optimization problem usually changes into the following quadratic programming (QP) problem by using Lagrange multipliers \( \alpha_i \). Therefore, the optimization problem of Eq. (9) could be rewritten as follows:
\[ \text{Min} \ L = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \]  
Subject to \[ \sum_{i=1}^{n} y_i \alpha_i = 0, 0 \leq \alpha_i \leq C, \ i = 1, 2, \ldots, n \]  
(10)

where \( K(x_i, x_j) \) is called kernel function which should meet the theory of Mercer, and its corresponding dot product of higher-feature space can be defined by
\[ K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j) \]  
(11)

When the quadratic programming problem is settled, the values of the weight vector \( \omega \) and \( b \) should meet:
\[ \begin{cases} \omega = \sum_{i=1}^{n} \alpha_i y_i \varphi(x_i) \\ b = -\frac{1}{2} \left( \max_{i,j} \omega^T \varphi(x_i) + \min_{i,j} \omega^T \varphi(x_j) \right) \end{cases} \]  
(12)

So, the final decision function of SVM is given as follows:
\[ f(x) = \text{sign}(\omega^T \varphi(x) + b) = \text{sign}\left( \sum_{i=1}^{n} \alpha_i y_i K(x_i, x) + b \right) \]  
(13)

In addition, with the help of the kernel functions, the classification of SVM is extended to nonlinearity. Among the kernel functions, the usual functions mainly include linear function, polynomials function, radial basis function (RBF), and sigmoid function \([22, 43]\), as shown in Table 3. The whole performance of the above kernel functions is similar. However, RBF has the advantage of classifying multidimensional data and fewer parameters in comparison to the linear, polynomial, and sigmoid kernel functions \([44]\). Hence, the RBF kernel function is employed in SVM to optimize the solution in this study.

### 2.3.2 Parameters selection of SVM with PSO

In this study, it is meaningfully to choose reasonably RBF parameter \( g \) and penalty parameter \( C \) of SVM because these parameters have a remarkable effect on the recognition performance of SVM classifier. Particle swarm optimization (PSO), proposed by Kennedy and Eberhart in 1995, is an intelligent bionic algorithm. The inspiration of PSO is derived from social behavior such as birds foraging and swarm theory. In terms of theory background, PSO, as a global search algorithm, is a sample without crossover and variation compared to GA \([29]\). Therefore, the parameter setting of PSO is easy to achieve. The execution of PSO begins with random initialization of a swarm of particles for finding out the optimal solution in the search space \([45]\).

Assume that the position and velocity of the \( i \)-th particle are set to \( x_i = [x_{i1}, x_{i2}, \ldots, x_{id}] \) and \( v_i = [v_{i1}, v_{i2}, \ldots, v_{id}] \) in \( d \)-dimensional search space, respectively, where \( i = 1, 2, \ldots, m \), \( m \) represents the scale of particles in the swarm, and \( j = 1, 2, \ldots, d \). According to the PSO algorithm, the direction of movement of each particle is toward its previous best position, finding a global best position of any particle in the swarm. Through evaluating the fitness of each particle, the previous best position (called \( p_{best} \)) of the \( i \)-th particle is computed and then the global best position (\( g_{best} \)) of all particle groups is found. For each iteration, the velocities of all the particles are computed and their positions are updated \([46]\), and the global best value is ultimately obtained.

| Table 3 | Four kernel function and formulas \([22, 43]\) |
|---|---|
| Kernel function type | Mathematical formula |
| Linear | \( K(x_i, x_j) = x_i^T x_j \) |
| Radial basis function (RBF) | \( K(x_i, x_j) = \exp(-\|x_i - x_j\|^2/2), g > 0 \) |
| Polynomial | \( K(x_i, x_j) = (-g x_i^T x_j + r)^d, g > 0 \) |
| Sigmoid | \( K(x_i, x_j) = \tanh(-g x_i^T x_j + r), g > 0 \) |
Let $k$ denote the current generation. To seek out the optimal solution, the current $d$-th dimension’s position and velocity of the $i$-th particle at time $k$ is described as follows [29]:

$$
v_{ij}(k + 1) = \omega v_{ij}(k) + c_{1} r_{1} [p_{ij} - x_{ij}(k)] + c_{2} r_{2} [p_{gij} - x_{ij}(k)] \tag{14}$$

$$
x_{ij}(k + 1) = x_{ij}(k) + v_{ij}(k + 1) \quad i = 1, 2, \ldots, m, j = 1, 2, \ldots, d \tag{15}$$

where $c_{1}$ and $c_{2}$ denote as accelerating constants, and $c_{1}$, $c_{2} > 0$. $r_{1}$ and $r_{2}$ represent random numbers, and its range is $[0, 1]$. $p_{ij}$ represents the best position of $i$-th particle in $j$-dimensional search space, and $p_{gij}$ the best position of the entire swarm. $v_{ij}$ represents the current $i$-th particle velocity in $j$-dimensional search space, $v_{ij} \in [-V_{\text{max}}, V_{\text{max}}]$, and $V_{\text{max}}$ represents the maximum limited velocity. $\omega$ is an inertial weight which is capable to balance local exploration and global exploration. A popularly used inertial weight is linearly decreasing weights (LDW) [30], which is defined as

$$
\omega = \omega_{\text{max}} - \frac{\omega_{\text{max}} - \omega_{\text{min}}}{k_{\text{max}}} \tag{16}
$$

where $\omega_{\text{max}}$ and $\omega_{\text{min}}$ represent the maximal and minimal inertia weight, respectively. $k$ is the number of iterations of controlling process, and $k_{\text{max}}$ is the maximal iteration of PSO.

In this paper, we chose the radial basis function (RBF) as the kernel function. Therefore, the performance of the SVM classifier has been influenced by two user-determined parameters (i.e., penalty parameter $C$ and RBF parameter $g$).

In a PSO system, the particle is composed of penalty parameters (i.e., penalty parameter $C$ and RBF parameter $g$) of SVM. The optimizing procedure of the SVM parameters by PSO is presented in Fig. 4, which is displayed as follows:

1. Input data. Training and testing sets are represented.
2. Particle swarm initialization. The accelerating constants are set to $c_{1}$ and $c_{2}$, $k_{\text{max}}$ is defined as the maximum iteration, and the current iterations $k$ is set to 1. The $m$ particles in the $d$-dimensional space are randomly generated, and the $i$-th particle position and velocity are denoted as $X_{i} = [x_{i1}, x_{i2}, \ldots, x_{id}]$ and $V_{i} = [v_{i1}, v_{i2}, \ldots, v_{id}]$, respectively;
3. Evaluating the fitness of all particles. The optimization function value for each particle is calculated in the corresponding search space, respectively.
4. Comparing the historical best position ($p_{\text{best}}$) of each particle with its current fitness. If the particle current fitness of particle is larger than its $p_{\text{best}}$, then the current fitness replaces $p_{\text{best}}$ to become the current position. Otherwise, $p_{\text{best}}$ remains the same.
5. Comparing the historical best positions ($g_{\text{best}}$) of the whole swarm with its current positions of all particles. If $g_{\text{best}}$ is smaller than the current position of the particle, then the current positions of the whole swarm replace to become the current positions of all particles. Otherwise, $g_{\text{best}}$ remains the same.
6. Updating velocity and position of particles. Velocities and positions of all particles based on Eqs. (11) and (12) are updated, and the next new swarm of the particle is formed, going to step 3.
7. Judging the stopping criterion. The stopping criterion is the maximal iteration, or that the fitness of the particles is smaller than a given required precision. If the stopping criterion is met, end the iteration operation. If the stopping criterion is not met, go to step 3.
8. Obtaining optimized SVM parameters $C$ and $g$.

3 Experimental setup of end milling

As displayed in Fig. 5, to demonstrate the accuracy of the proposed approach of chatter recognition, we had conducted the cutting experiments of aluminum 6061 on a three-axis milling machine center of VMC1165B. The cutting tool was a two-edged carbide end milling cutter, and the corresponding overhang and diameter were 44 mm and 8 mm, respectively. An acceleration sensor was mounted on the
spindle housing, and a data acquisition card of NI USB-6341 was utilized to obtain the vibration signals during end milling machining. The signal sampling frequency was set to 12,000 Hz. In addition, the end milling experiments were carried out under the dry condition without coolant.

It is well known that the occurrence of chatter is closely connected with spindle speed, cutting depth, and feed rate [47]. In this study, for the same spindle speed, the milling depth was set to 0.2 mm each time until the occurrence of chatter. The specific milling parameters were described in Table 4.

In addition, to acquire the inherent frequency of tool and workpiece systems, a hammer experiment was carried out before milling. The position struck by the Hammer was on the tool tip. The transfer function of tool-workpiece systems was acquired through the approach of single-point impaction and response. Therefore, the first-, second-, and third-order natural frequencies are 1494, 2041, and 4160 Hz, respectively.

4 Results and discussion

4.1 The collected vibration signal analysis of milling processes

Figure 6a, b shows the three typical processing states (e.g., stable, transition, and chatter) of the vibration signal. Figure 6c, f describes the partially enlarged vibration signals and FFT of stable cutting state in Fig. 6a, respectively. Figure 6d, h, e, i are the partially enlarged vibration signals and FFT of transition and chatter cutting state in Fig. 6b, respectively. From Fig. 6c, f in the stable cutting, it is obviously found that the vibration signal amplitude is small, and the distribution of frequency components is dispersal, the main frequency peaks mainly concentrating on the 1301 Hz, 2039 Hz, 2801 Hz, 4078 Hz, 5145 Hz. These frequency components nearly correspond with the first-order natural frequency, second-order natural frequency, twice first-order natural frequency, third-order natural frequency, and four-fold first natural frequency, respectively. When increasing axial milling depth up to 0.8 mm and spindle speed up to 6000 r/min, the transition state of chatter appears. In this transition state, the increase of signal amplitude is slight, but the distribution of the frequency components changes drastically (see Fig. 6d, h). Consequently, other frequency components are suppressed, and the frequency focuses on around 2824 Hz. This characteristic frequency is almost twice as large as the first-order natural frequency [48]. This is owing to the serious instability of the helix angle of the milling tool, repeatedly impacting on the workpiece to drive chatter [49]. In addition, the occurrence of chatter inhibits the production of other frequencies. Subsequently, severe chatter occurs with the cumulative effect of energy (see Fig. 6b). This is reflected in the increase in the amplitude of the signal, and the chatter frequency of 2824 Hz is further enhanced.

This paper aims to effectively identify the infantile chatter state for avoiding the undesirable influence on the material surface quality and milling cutter. Therefore, the effective and accurate recognition of the transition stage of chatter is significantly important. In this regard, this paper emphasized on investigating the feature extraction of the transition stage of chatter. In this study, the wavelet basis function chose db10, owing to better orthogonality. WPT was utilized to decompose the measured vibration signals in four layers in terms of stable state and chatter state during end milling machining. The sixteen wavelet packets were obtained correspondingly. Figure 6 was the description of the four-layer WPT of the reconstructed signal in the transition state of chatter and their corresponding FFT in each frequency band, which was corresponding to Fig. 6d. It was shown that the acceleration signal amplitude in the frequency bands $x_{4,7}$ (2625–3000 Hz)

### Table 4 Cutting conditions

| Workpiece  | AL6061             |
|------------|---------------------|
| Tool       | Carbide End Mills Φ8 mm |
| Spindle speed (rpm) | 3000, 4000, 5000, 6000, 7000 |
| Feed per tooth (mm/z) | 0.02 |
and $x_{4,8}$ (3000–3375 Hz) was larger than that in other frequency bands. Therefore, chatter frequency may appear in the frequency bands $x_{4,7}$ and $x_{4,8}$. The energy ratio of each frequency band was calculated, as shown in Table 4. We found that the energy ratio of stable, transition, and chatter state in wavelet packets of $x_{4,7}$ and $x_{4,8}$ was 0.1137, 0.53, and 0.8467, respectively (Table 5). Therefore, the vibration energy increased sharply and mainly concentrated in the wavelet packets of $x_{4,7}$ and $x_{4,8}$ when chatter occurred. Due to the presence of rich chatter information, the wavelet packets of $x_{4,7}$ and $x_{4,8}$ were chosen and reconstructed. The reconstructed vibration signal of characteristic wavelet packets and corresponding to FFT were shown in Fig. 7. It is found that the time-domain features of the reconstructed vibration signal become more obvious, and the frequency spectrum retains complete characteristic information of the transition state of chatter. This demonstrated that the redundant noise and irrelevant information is effectively removed through selecting the characteristic wavelet packets preprocessed by WPT (Fig. 8).

### 4.2 Extraction of chatter features

In this paper, we have mainly focused on the infantile chatter identification (e.g., transition state) based on the vibration signal, as displayed in Fig. 6d. Therefore, for extracting chatter features, the vibration signal is pretreated with four-layer WPT, obtaining sixteen wavelet packets in the time–frequency domain. Then, the characteristic wavelet packets of $x_{4,7}$ and $x_{4,8}$ in stable and chatter transition state, corresponding to Fig. 6c, d, were chosen and reconstructed according to the analysis in Sect 4.1. Subsequently, fourteen time–frequency features were gained through calculating the reconstructed vibration signal. However, the difference between the values of fourteen time–frequency features is too large. This may lead to the failure of convergence and

![Figure 6](https://via.placeholder.com/150)

**Fig. 6** (a) The stable cutting state of vibration signals under cutting conditions of spindle speed 5000 r/min, axial depth of cut 0.2 mm, and feed rate 0.2 mm/z. (b) The transition and chatter cutting states of vibration signals under cutting conditions of spindle speed 6000 r/min, axial depth of cut 0.8 mm, and feed rate 0.2 mm/z. (c) and (f) were the partially enlarged vibration signals and FFT of stable cutting state in (a), respectively. (d) and (h), (e) and (i) were the partially enlarged vibration signals and FFT of transition and chatter cutting state in (b), respectively.
increase the training time. Hence, it is meaningfully necessary to preprocess the chatter feature recognition parameters before training in order to avoid too large a difference between the characteristic parameters, the mean square frequency, frequency center, and standard frequency divided by parameters before training in order to avoid too large a difference. The results were described in Table 6. The fourteen time–frequency features constructed the feature set T = [xm xp x rms x std x skew x kur CF CLF SF IF MSF p FC FV].

In addition, the transition state during milling belonged to the chatter state in this study. Based on the experiment parameter in Table 5, we obtained 30 samples of a stable state and 30 samples of a chatter state. The training set was randomly composed of 20 stable samples and 20 chatter samples, and the testing set consisted of the remaining samples.

### 4.3 Chatter recognition based on time–frequency characteristics of WPT and PSO-SVM

According to the previous analysis in this paper, inappropriate penalty parameter C and RBF parameter g may cause the over-fitting and under-fitting of the SVM classifier, which severely affects the accuracy of SVM prediction classification. But in the process of practical application, it is very difficult to determine the optimal value of C and g. Therefore, the SVM parameters should be set in advance before the application of the SVM prediction model. In this study, the algorithm of PSO was employed to optimize the penalty parameter C and RBF parameter g of the SVM classifier. The initialized parameter values of PSO-SVM were set as follows. The swarm size was set to 20 particles, and the maximum iterations were set to 200. The accelerating constant c1 and c2 were set to 1.5 and 1.7, respectively. The searching range of parameter C of SVM was between 0.1 and 100, while the searching range of parameter g of SVM was between 0.01 and 1000. The fitness curves for PSO to find the optimal parameter of the SVM classification model were shown in Fig. 9. From Fig. 9, the optimum fitness reached 95% and the optimal values of penalty parameter C and RBF parameter g of SVM classifier were C = 74.89 and g = 0.01, respectively. The chatter recognition accuracy of the testing set is 95%, as displayed in Table 7.

In addition, for demonstrating the benefits of the developed PSO-SVM in chatter detection, three other methods of standard support vector machine (SSVM) and k-fold cross-validation support vector machine (k-CV-SVM) were selected and compared for chatter identification. The standard SVM parameters were set to C = 2 and g = 1, which was a reference. And the initialized parameter values of k-CV-SVM were set to threefold cross-validation, 2^−8 ≤ C ≤ 2^8 and 2^−8 ≤ g ≤ 2^8. The initialized parameter values of GA-SVM were set as follows. The maximum evolutionary algebra and population were set to 100 and 20, respectively. And the crossover probability and mutation probability were set to 0.4 and 0.01, respectively. The searching range of parameters C and g of SVM was the same as the PSO algorithm. The classification accuracy of the k-CV method was shown in Fig. 10, and the prediction accuracy was shown in Table 7. From Table 7, it was seen that the chatter identification accuracy rate of training data was 95% under k-CV-SVM. The identification performance of k-CV-SVM was equal to that of SSVM, while the prediction accuracy rate was lower in testing data than that of SSVM. This is due to that the approach of k-CV-SVM belongs to the local search method. In addition, this algorithm is vulnerable to local optima [44, 50]. The fitness curves of the evolutionary algebra of GA were shown in Fig. 11. It was found that the optimum fitness reached 95%. And the chatter identification accuracy rate of testing data was 85%. Therefore, compared to the three chatter identification methods of SVM, k-CV-SVM, and GA-SVM, the PSO-SVM enhanced the accuracy of chatter identification for training data and testing data. Through the method proposed in this study, the phenomenon of chatter in the milling process could be identified and predicted.

### 4.4 Validation of effectiveness of chatter recognition based on WPT and PSO-SVM

For proving the effectiveness of chatter recognition and prediction on the basis of WPT and PSO-SVM, we used the measured vibration signals under milling conditions of

### Table 5 Energy ratio in the frequency bands of wavelet packet

| Energy ratio in wavelet packet node | Cutting state | Stable state | Transition state | Chatter state |
|-----------------------------------|--------------|-------------|------------------|--------------|
| x4,0 (0–375 Hz)                   | 0.0537       | 0.0809      | 0.0127           |
| x4,1 (375–750 Hz)                 | 0.0568       | 0.0168      | 0.0055           |
| x4,2 (750–1125 Hz)                | 0.0522       | 0.0588      | 0.0055           |
| x4,3 (1125–1500 Hz)               | 0.0738       | 0.0146      | 0.0049           |
| x4,4 (1500–1875 Hz)               | 0.0789       | 0.0246      | 0.0069           |
| x4,5 (1875–2250 Hz)               | 0.2125       | 0.0398      | 0.0132           |
| x4,6 (2250–2625 Hz)               | 0.0648       | 0.1036      | 0.0577           |
| x4,7 (2625–3000 Hz)               | 0.0581       | 0.3301      | 0.5617           |
| x4,8 (3000–3375 Hz)               | 0.0556       | 0.1999      | 0.2850           |
| x4,9 (3375–3750 Hz)               | 0.0672       | 0.0619      | 0.0167           |
| x4,10 (3750–4125 Hz)              | 0.0468       | 0.0106      | 0.0059           |
| x4,11 (4125–4500 Hz)              | 0.0270       | 0.0036      | 0.0032           |
| x4,12 (4500–4875 Hz)              | 0.0175       | 0.0095      | 0.0045           |
| x4,13 (4875–5250 Hz)              | 0.0748       | 0.0104      | 0.0037           |
| x4,14 (5250–5625 Hz)              | 0.0254       | 0.0173      | 0.0067           |
| x4,15 (5625–6000 Hz)              | 0.0350       | 0.0176      | 0.0063           |
axial cutting depth of 0.6 mm, spindle speed of 6000 r/min, and feed rate of 0.2 mm/z to identify chatter state. Figure 12 described the time-domain spectrum of the collected vibration signals. For the better identification of cutting state, every 1024 data points were selected as an identification sample of cutting state. And 122 identification samples were obtained. Each sample was processed with WPT, and then fourteen time–frequency features were extracted as feature indicators for chatter recognition. The optimized SVM prediction model by PSO in the Sect. 4.3 was adopted to detect the milling machining state. The identification result was shown by the red asterisk mark in Fig. 12. From Fig. 12, it
The reconstructed vibration signal and FFT of characteristic wavelet packets of $x_{4,7}$ and $x_{4,8}$ in the transition state.

Table 6: The values of time–frequency features of the reconstructed vibration signal

| Time–frequency characteristics | Feature value | Stable  | Transition |
|---------------------------------|---------------|---------|------------|
| Mean ($x_m$)                    | 0.0945        | 0.5718  |
| Standard deviation ($x_{std}$)  | 0.1188        | 0.7215  |
| Root mean square ($x_{rms}$)    | 0.1187        | 0.7211  |
| Peak ($x_p$)                    | 0.3459        | 2.2943  |
| Skewness ($x_{skw}$)            | -2.8949       | -2.8825 |
| Kurtosis ($x_{kur}$)            | 3.8513        | 4.2113  |
| Crest factor ($CF$)             | 2.9121        | 3.1815  |
| Clearance factor ($CLF$)        | 4.3818        | 4.7464  |
| Shape factor ($SF$)             | 1.2570        | 1.2611  |
| Impulse factor ($IF$)           | 0.3459        | 2.2943  |
| Mean square frequency ($MSF/10^7$) | 0.8654       | 0.8291  |
| One-step autocorrelation function ($\rho$) | 0.0339    | 0.0655  |
| Frequency center ($FC/10^5$)    | 0.2935        | 0.2874  |
| Standard frequency ($FV/10^5$)  | 0.3870        | 0.2867  |

Table 7: Chatter recognition accuracy with the different optimization algorithms

| Number of features | $C$  | $g$  | Accuracy/% |
|--------------------|------|------|-------------|
|                    |      |      | Training set | Testing set |
| WPT + SSVM         | 14   | 2    | 1            | 95          | 90          |
| WPT + k-CV-SVM     | 14   | 0.758| 1.32         | 95          | 85          |
| WPT + GA-SVM       | 14   | 3.94 | 0.40         | 95          | 85          |
| WPT + PSO-SVM      | 14   | 74.89| 0.01         | 95          | 95          |

Fig. 9: The adaptive evolutionary curves of PSO

Fig. 10: The classification accuracy of the $k$-CV method

Fig. 11: The adaptive evolutionary curves of GA
was visibly seen that the collected vibration signal amplitude did not change dramatically before the time of 3.891 s. The optimized model identified the early occurrence of chatter at 3.667 s, while the amplitude of the collected vibration signal had a slight fluctuation. It meant that the optimized SVM model accurately identified the light chatter state 0.224 s in advance. In this regard, if some measures to suppress chatter may be executed by this time, the adverse effects of severe chatter on the workpiece and milling cutter could be prevented. Therefore, the proposed method with an excellent and robust performance in this paper accurately identified the infantile chatter state during the milling machining.

5 Conclusion

In this paper, the multi-feature recognition approach for chatter detection combining WPT and PSO-SVM was proposed for improving the accuracy of chatter identification under the milling machining. The chatter detection is essentially a problem of signal processing, feature extraction, and chatter identification. This study emphasizes on the selection of the frequency band containing chatter information of vibration signals with WPT for signal processing. Chatter-emerging frequency bands of $x_{4.7}$ and $x_{4.8}$ were chosen and reconstructed, which effectively removed the redundant noise and useless information to increase the signal-to-noise ratio of vibration signals. Subsequently, fourteen time–frequency features are extracted through calculating the reconstructed vibration signals. The fourteen time–frequency features as the multi-feature indicators of chatter recognition could reflect the change of amplitude and distribution in the time and frequency domain when chatter occurs. In addition, this paper studied that the optimization algorithms of $k$-CV, GA, and PSO were adopted to optimize the penalty parameter $C$ and RBF parameter $g$ of the SVM model. Compared to the two optimization algorithms of $k$-CV and GA, PSO-SVM obviously improved the accuracy of chatter recognition, as shown in Table 7. Moreover, the optimized SVM prediction model by PSO was applied to detect the state of milling machining. The recognition results indicated that the model may accurately predict the slight chatter state in advance. In this regard, chatter suppression will be investigated in future. Chatter may be suppressed through adjusting the milling depth, spindle speed, and other processing parameters.

Author contribution Qingzhen Zheng: methodology, software, writing (original draft preparation), experiment, and data processing; Guangsheng Chen: conceptualization, methodology, writing (review and editing), supervision, project administration; Anling Jiao: format checking, modification of figures and text.

Funding This study was supported by the National Key Research and Development Program of China (Grant Number: 2017YFB1104600).

Data availability Data will be available upon reasonable request.

Declarations

Ethical approval Not applicable.

Consent to participate Not applicable.

Consent for publication All co-authors consent to the publication of this work.

Competing interests The authors declare no competing interests.
References

1. Merritt H (1965) Theory of self-excited machine-tool chatter: contribution to machine-tool chatter research—1. J Eng Ind 87:447–454
2. Qian S, Sun Y, Xiong Z (2015) Support vector machine based online intelligent chatter detection. J Mech E 51:1–8
3. Li X, Wong Y, Nee A (1997) Tool wear and chatter detection using the coherence function of two crossed accelerations. Int J Mach Tool Manuf 37:425–435
4. Thelaidji T, Chenikher S (2013) A new approach of preprocessing with SVM optimization based on PSO for bearing fault diagnosis. Int Conf Hybrid Intell Syst 319–324
5. Soliman E, Ismail F (1997) Chatter detection by monitoring spindle drive current. Int J Adv Manuf Tech 13:27–34
6. Griffin JM, Chen X (2009) Multiple classification of the acoustic emission signals extracted during burn and chatter anomalies using genetic programming. Int J Adv Manuf Tech 45:1152–1168
7. Ye J, Feng P, Xu C et al (2018) A novel approach for chatter online monitoring using coefficient of variation in machining process. Int J Adv Manuf Tech 96:287–297
8. Tangjitstcharoen S (2009) In-process monitoring and detection of chip formation and chatter for CNC turning. J Mater Process Technol 209:4682–4688
9. Kaljanic E, Sortino M, Totti G (2008) Multisensor approaches for chatter detection in milling. J Sound Vib 312:672–693
10. Pan J, Liu Z, Wang X et al (2020) Boring chatter identification by multi-sensor feature fusion and manifold learning. Int J Adv Manuf Tech 109:1137–1151
11. Delio T (1992) Use of audio signals for chatter detection and control. J Manu Sci Eng 114:146
12. Wan S, Li X, Yin Y et al (2021) Milling chatter detection by multi-feature fusion and AdaBoost-SVM. Mech Syst Signal Process 156:107671
13. Fu Y, Zhang Y, Zhou H et al (2016) Timely online chatter detection in end milling process. Mech Syst Signal Pr 75:668–688
14. Ji Y, Wang X, Liu Z et al (2017) EEMD-based online milling chatter detection by fractal dimension and power spectral entropy. Int J Adv Manuf Tech 92:1185–1200
15. Lei Y, Lin J, He Z et al (2013) A review on empirical mode decomposition in fault diagnosis of rotating machinery. Mech Syst Signal Process 35:108–126
16. Cao H, Lei Y, He Z (2013) Chatter identification in end milling process using wavelet packets and Hilbert-Huang transform. Int J Mach Tool Manuf 69:11–19
17. Yao Z, Mei D, Chen Z (2010) On-line chatter detection and identification based on wavelet and support vector machine. J Mater Process Technol 210:713–719
18. Sun Y, Xiong Z (2016) An optimal weighted wavelet packet entropy method with application to real-time chatter detection. IEEE/ASME Trans Mechatron 21:2004–2014
19. Teti R, Jemielaik A, O’Donnell G et al (2010) Advanced monitoring of machining operations. CIRP Ann 59:717–739
20. Zhang CL, Yue X, Jiang YT et al (2010) A hybrid approach of ANN and HMM for cutting chatter monitoring. Adv Mater Res 3225–3232
21. Chen Y, Li H, Jing X et al (2019) Intelligent chatter detection using image features and support vector machine. Int J Adv Manuf Technol 102:1433–1442
22. Chen Y, Li H, Hou L et al (2018) An intelligent chatter detection method based on EEMD and feature selection with multi-channel vibration signals. Measurement 127:356–365
23. Mcy A, Fak A, Ao B (2020) On transfer learning for chatter detection in turning using wavelet packet transform and ensemble empirical mode decomposition. CIRP J Manuf Sci Technol 28:118–135
24. Peng C, Wang L, Liao TW (2015) A new method for the prediction of chatter stability lobes based on dynamic cutting force simulation model and support vector machine. J Sound Vib 354:118–131
25. Zhang X, Liang Y, Zhou J (2015) A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM. Measurement 69:164–179
26. Hsu C-W, Lin C-J (2002) A simple decomposition method for support vector machines. Mach Learn 46:291–314
27. Yang Z, Yu Z (2012) Gridding wheel wear monitoring based on wavelet analysis and support vector machine. Int J Adv Manuf Technol 62:107–121
28. Jia L, Jian-Ming Z, Xiao-Jing B et al (2011) Research on GA-SVM tool wear monitoring method using HHT characteristics of drilling noise signals. Int Conf Consum Electron Commun Netw (CECNet) 635–638
29. Liu Z, Cao H, Chen X et al (2013) Multi-fault classification based on wavelet SVM with PSO algorithm to analyze vibration signals from rolling element bearings. Neurocomputing 99:399–410
30. Hassan R, Cohanim B, De Weck O et al (2005) A comparison of particle swarm optimization and the genetic algorithm. In 46th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 1897
31. Bahrie S, Pothar V (2014) Optimization of milling conditions by using particle swarm optimization technique: a review. Int J Eng Trends and Techn 18:248–251
32. Wang E, Yan P, Liu J (2020) A hybrid chatter detection method based on WPD, SSA, and SVM-PSO. Shock Vib 2020
33. Lei Y, He Z, Zi Y et al (2008) New clustering algorithm-based fault diagnosis using compensation distance evaluation technique. Mech Syst Signal Pr 22:419–435
34. Jia G, Wu B, Hu Y et al (2013) A synaptic criterion for early recognition of cutting chatter. Sci China Technol Sc 56:2870–2876
35. Shi H (2003) Metal cutting theory and practice—a new perspective. Wuhan Huazhong Univ Sci Technol 338–343
36. Zhijian M, Shuzi Y, Hannim S (1988) Early stage diagnosis and monitored control for the chatter of machine tools. J Vib Eng 3
37. Zhu K, San Wong Y, Hong GS (2009) Wavelet analysis of sensor signals for tool condition monitoring: a review and some new results. Int J Mach Tool Manuf 49:537–553
38. Plaza EG, López PN (2018) Analysis of cutting force signals by wavelet packet transform for surface roughness monitoring in CNC turning. Mech Syst Signal Process 98:634–651
39. Mallat SG (1989) A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans Pattern Anal Mach Intell 11:674–693
40. Liu C, Zhu L, Ni C (2018) Chatter detection in milling process based on VMD and energy entropy. Mech Syst Signal Process 105:169–182
41. Vapnik V (1998) Statistical learning theory. Wiley, New York
42. Wang Y, Bo Q, Liu H et al (2018) Mirror milling chatter identification using Q-factor and SVM. Int J Adv Manuf Technol 98:1163–1177
43. Burges CJ (1998) A tutorial on support vector machines for pattern recognition. Data Min Knowl Disc 2:121–167
44. Lin S-W, Ying K-C, Chen S-C et al (2008) Particle swarm optimization for parameter determination and feature selection of support vector machines. Expert Syst Appl 35:1817–1824
45. Long B, Xian W, Li M et al (2014) Improved diagnostics for the incipient faults in analog circuits using LSSVM based on PSO algorithm with Mahalanobis distance. Neurocomputing 133:237–248
46. Huqang C-L, Dun J-F (2008) A distributed PSO–SVM hybrid system with feature selection and parameter optimization. Appl Soft Comput 8:1381–1391
47. Oleaga I, Pardo C, Zulaika JJ et al (2018) A machine-learning based solution for chatter prediction in heavy-duty milling machines. Measurement 128:34–44
48. Suh C, Khurjekar P, Yang B (2002) Characterisation and identification of dynamic instability in milling operation. Mech Syst Signal Pr 16:853–872
49. Zatarain M, Munoa J, Peigné G et al (2006) Analysis of the influence of mill helix angle on chatter stability. CIRP Ann 55:365–368
50. Lin C-J (2003) A practical guide to support vector classification. Available at: http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.