Estimation the wear state of milling tools using a combined ensemble empirical mode decomposition and support vector machine method

Chuangwen XU*, Yuzhen CHAI *, Huaiyuan LI* and Zhicheng SHI*
*Lanzhou Institute of Technology, Provincial Key Laboratory for Green Cutting Technology and Application of Gansu Province
Gong Ping Road No. 1, Qilihegu, Lanzhou, China
E-mail: xuchuangwen@126.com

Received: 4 April 2018; Revised: 8 May 2018; Accepted: 3 June 2018

Abstract
Vibrational signals resulting from tool wear have non-linear and non-stationary features. It is also difficult to acquire large numbers of typically worn samples in practice. In this work, a method of predicting the wear of milling tools is proposed based on ensemble empirical mode decomposition (EEMD) and the use of a support vector machine (SVM). The EEMD method is used to decompose the original non-stationary vibration acceleration signals into several stationary intrinsic mode functions (IMFs). The energies of the signals in these different frequency bands change when the tool is worn. Thus, the tool wear state can be identified by calculating the EEMD energies and energy entropies of the different vibrational signals. The correlation coefficients between the IMF components and original signal were calculated and wear-sensitive IMFs chosen. A SVM is then established by considering the energy features extracted from a number of wear-sensitive IMFs that contain primary information on tool wear. These are considered as the inputs to judge the wear state of the tool. The results show that the method is capable of predicting the wear state of the milling tool to good effect. Furthermore, the predictions made using an LS-SVM based on EEMD method are more accurate than those made using FFT, Wavelet analysis and EMD methods.

Keywords: Cutting vibration, Ensemble empirical mode decomposition, Intrinsic mode function, Energy entropy, Support vector machine, Tool wear prediction

1. Introduction

The tooling system is an extremely important functional part of high-grade numerical-control (NC) machine tools. As such, monitoring the running state of NC tooling systems is a key aspect of the NC system, and one that has a decisive effect on the manufacturing process (Li et al., 2017). Therefore, as manufacturing processes continue to develop, it is essential that the tool state is monitored in real-time so as to provide a basis for warning operators of the need for tool radius compensation and cutting parameter optimization (Kannatey-Asibu et al., 2017). As a result of such action, environmental pollution and processing costs can be reduced, and productivity and product quality improved (Ghani et al., 2011).

Milling machines are more complicated than other mechanical machining methods for a variety of reasons: the diverse nature of the milling conditions encountered, variability of the cutting parameters, random nature of tool wear, and complexity and incomplete controllability of the cutting process. There are also a large number of factors that influence the wear of the milling tool. Thus, monitoring models are faced with a great challenge: adaptability. Subsequently, adaptability modeling and extraction of wear features during milling are difficult to accomplish and constitute bottleneck problems in research on tool-wear monitoring.

At present, obtaining satisfactory milling machine results is still primarily dependent on the skill and experience of the workers. This strongly influences the stability of milling quality and delays the intelligentization and digitization of the milling process. As a result, this problem has become one that needs to be solved urgently. In this context, studying the theory, technology, and intelligentization of on-line monitoring of milling tool-wear is of great significance in realizing precise control over the milling process and promoting intelligentization and digitization of mechanical manufacturing. Moreover, it can greatly enrich the basic theory and methods used in intelligent milling machining. It
will also bring tremendous social and economic benefits.

Wear-monitoring technology is generally divided into indirect and direct monitoring methods. Indirect methods involve inferring tool wear indirectly by use of various sensors. This approach to identifying the state of tool wear is one that has been accepted by scholars all over the world. Based on signal processing and analysis technology, these methods extract features of the acquired signals (cutting force, vibration, acoustic emission, cutting temperature and power, etc.) using time domain, frequency domain, time–frequency domain, and time series analysis methods. Moreover, features that reflect tool wear can be used to make a multi-dimensional eigenvector. Then, mode identification methods can be used to study and identify the state of tool wear.

Wang et al. (Wang et al., 2014) introduced heterogeneous ensemble learning to realize tool condition monitoring in which the support vector machine, hidden Markov model and radius basis function are selected as base classifiers and a stacking ensemble strategy is further used to reflect the relationship between the outputs of these base classifiers and tool wear states. Kilundu et al. (Bovic et al., 2011) used data mining technology in their tool-wear monitoring research while Chia, Liang, Yen et al. (Yen et al., 2013) monitored tool wear occurring during milling by using a self-organizing feature mapping neural network model. Qun, Ren et al. (Qun et al., 2010) integrated multi-sensor and tool wear information using Takagi–Sugeno–Kang fuzzy modeling and thus provided a highly-precise and highly-reliable method of tool wear prediction. Tomas Kalvoda et al. (Tomas and Hwang, 2010) carried out research on tool wear monitoring using a method based on Hilbert–Huang transformation. Geramifar et al. (Geramifar et al., 2012) introduced a method to continuously monitor the state of tool wear during milling by employing a hidden Markov model. Jun, Hong, Zhou et al. (Zhou et al., 2011) used features of the acoustic emission signal to identify tool wear using an autoregressive integrated moving average model based on a few external features as input. Liu Lu et al. (Liu et al., 2011) automatically identified the tool wear state by employing a hypersphere SVM. Wang Guofeng et al. (Wang et al., 2011) analyzed multi-sensor information (e.g. multi-directional cutting force and vibration signal) and acquired information on time, frequency, and wavelet domains to use as features to classify wear. Nie Peng et al. (Nie and Chen, 2011) established a system to predict tool wear based on PCA and an improved back propagation (BP) neural network using the Levenberg–Marquardt algorithm. Guan Shan (Guan, 2011) proposed a method of multi-feature analysis of acoustic emission signals based on LS-SVM which was aimed at classifying tool wear and predicting the extent of wear under a variety of cutting conditions.

The above discussion summarizes the various methods used to monitor wear. In the early stages, the methods were mainly aimed at analysis in the time domain. In the middle stages, they moved more towards time–frequency analysis and employed Fourier transformations. In more recent years, wavelet analysis and neural networks have begun to feature. It is apparent, therefore, that developments in signal processing technology have positively affected the promotion of tool wear prediction. To date, however, no good methods have been found to predict the state of tool wear during milling. The reasons for this are two-fold: Firstly, changes in the cutting parameters affect the stability of the signal processing technology during milling. Secondly, the monitoring signals observed during milling exhibit strong non-linearity and non-stationarity. This causes the frequency-domain features of the signals to vary with time and so the milling signals cannot be described using fixed frequency-domain features.

The signal frequency after FFT is mixed, so it cannot identify from high frequency signal and the high frequency noise. The wavelet transform maps the signal into the wavelet domain and divides the signal scale at the same time, so the details are concerned from the internal signal. The disadvantage is that wavelet transformation need to determine the wavelet base also threshold and hierarchical number according to the different characteristics of the object, and the selection of the parameters has a great influence on the filtering accuracy, that is, The wavelet transform is non-adaptive. The empirical mode decomposition (EMD) method can decompose the signal adaptively according to different frequencies, and achieve noise reduction by component high pass, low pass, band pass or threshold filtering. But when the high frequency signal denoising, there is modal aliasing, which affects the filtering effect of non-stationary signals this includes abnormal events. In order to solve the modal aliasing problem, the ensemble mean empirical mode decomposition is proposed. The EEMD method makes use of the zero mean value of Gauss white noise and the uniform distribution characteristics of the frequency to make the signal feature scale evenly distributed and the abnormal events can be processed smoothly. Thus the high frequency denoising effect of the EMD is better than that of the high frequency denoising, and the processing process is adaptive and the wavelet filtering is simple. The EEMD method can obtain a complete and accurate time-frequency distribution of the signal energy, which has higher accuracy and resolution than the traditional FFT spectrum, and does not produce energy leakage at the same time. When the tool...
is worn, the natural frequency of the cutting system will be changed. At this time, the energy of the tool wear signal will change in the natural frequency section.

Empirical mode decomposition (EMD), however, is a newly developed signal-processing method that is suitable for non-linear and non-stationary signals. EMD is a method of time–frequency analysis that was proposed by Huang in 1998. The EMD method involves decomposing a signal into the sum of numerous intrinsic mode functions (IMFs) based on the timescales of the local features of the signal (Huang et al., 1998). As EMD is a local method of time–frequency analysis which uses an adaptive basis, it avoids the effects associated with poor decomposition which generally occur due to improper choice of basis function during wavelet decomposition. Moreover, each of the decomposed IMFs represent information specific to a certain frequency-band and have specific significance. This favors the extraction of weak abnormal information from the signal. However, as a result of the fitting of envelope lines and mean-value curves, the zero-mean problem and end effects of the algorithm mean that the IMF components obtained using EMD from the measured signals lack physical significance in mechanical failure applications. This has a heavy impact on the application of EMD processing in signal analysis. This is especially the case when the timescales of the signals decomposed via EMD exhibit ‘jumpy’ changes. As a result, IMF components containing features of different timescales are likely to occur, that is, we end up with modal mixtures, (Peng et al., 2005 and Chen et al., 2014).

To solve this problem, Wu et al. proposed the use of EEMD. Using EEMD, a non-stationary signal can be decomposed into a finite number of approximately stationary IMFs. Of crucial importance is that the degree of modal mixing in the IMFs obtained via EEMD is lower than that via EMD (Wu and Huang, 2004, 2009). The signal energies in different frequency bands change when tools become worn. Therefore, the state of wear can be identified by calculating the EEMD energy entropy of the different vibration signals (Yang and Shao, 2013). On this basis, a new method of predicting the wear of milling tools based on EEMD and SVM is proposed in this study. In the method, the energy features extracted from the IMF components that contain primary information on tool wear are regarded as the input to the SVM to identify the state of the tool wear.

2. Experimental materials and procedures
2.1 Workpiece materials
0Cr18Ni9 stainless steel with high strength, good rigidity, strong corrosion resistance, good wear resistance, low price and excellent comprehensive performance, and is widely used in aerospace, chemical, petroleum, building and food in all walks of life, which has extensive application. 0Cr18Ni9 stainless steel is a material that is typically difficult to machine: it shows a high tenacity and thermal intensity as well as a low thermal conductivity. Furthermore, it is liable to undergo significant plastic deformation, severe work hardening, and generates much heat during cutting, and the heat is hard to dissipate, so the tool tip operates at a high temperature and chips are liable to adhere to the cutting edges. As a result, built-up edges are easily to produce, so as to accelerate tool wear. In view of the wide application of 0Cr18Ni9 stainless steel and its difficult to process particularity, 0Cr18Ni9 stainless steel is chosen as the experimental object.

In the experiment, 0Cr18Ni9 stainless steel square stock with a length, width, and height of 420 mm, 320mm, and 100 mm respectively was used. The mechanical properties were as follows: tensile strength $\sigma_s \geq 520$ MPa, yield strength $\sigma_{0.2} \geq 205$ MPa, elongation rate $\delta \geq 40\%$, area reduction $\psi \geq 60\%$, elastic modulus $E > 115$ GPa, and a hardness of 159 HB. In addition, the chemical components are listed in Table 1.

<table>
<thead>
<tr>
<th>Element</th>
<th>C</th>
<th>Mn</th>
<th>Mo</th>
<th>Co</th>
<th>P</th>
<th>Cr</th>
<th>Ni</th>
<th>Cu</th>
<th>N2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass fraction</td>
<td>0.02</td>
<td>1.71</td>
<td>2.10</td>
<td>0.12</td>
<td>0.03</td>
<td>16.31</td>
<td>10.14</td>
<td>0.24</td>
<td>0.04</td>
</tr>
</tbody>
</table>

2.2 Cutting tool materials and geometrical parameters
TiAlN_ carbide tool with 10 μm of composite coating produced by Lamina Company, Switzerland was used to process the 0Cr18Ni9 stainless steel. In addition, it had two teeth and the types of arbor and insert were 90W25-3K13 and APMT 1135 PDTR LT30 respectively. The geometrical parameters of the insert were as follows: front angle 9°, back angle 16°, inclination angle 4°, minor back angle 17°, end-cutting edge angle 4°, corner radius 0.5 mm, and cutting edge angle 90° after installing a insert (see Figure 1). The diameter of the tool used in the experiments is 16mm.
2.3 Cutting experiments

Cutting experiments were conducted on an NC milling machine (type XKA 714) and the cutting rate adjusted by changing the spindle speed. During the experiments, the vibrations in the x, y, and z directions were measured by using a DH5922N testing system. A microscope with a super-high depth of field (Smartzoom 5) was employed to observe the wear mode and measure the extent of wear of the flank. The vibration acceleration test system used is illustrated in Fig. 2. The acceleration transducer is located on the workpiece. In our experiments, the peripheral milling and radial cutting depths were both 2 mm, and the acceleration sampling frequency was set to 10 kHz (in dry cutting mode).

Cutting conditions were chosen according to an orthogonal experiment table. Under the conditions determined, the experiments proceeded as follows:

1) Two blades were installed on the cutter bar for the cutting experiments and a single-cutting length of 420 mm was chosen. Cutting was continuously performed 5 times using NC programming. The vibrational signals were collected during cutting and then the cutter bar and blades were taken down.

2) The extent of wear of the cutter flank, $VB$, was measured.

3) The cutter bar and blades were reinstalled and then the NC milling machine was turned on. Cutting was continuously performed 5 times under the control of the NC programmer. The cutting process was continued until a reasonable amount of wear had been caused to the blades ($VB \geq 0.3$ mm). In this context, the cutting experiment for this set of conditions was ended.

4) Another set of cutting conditions were chosen and then steps 1–3 repeated. When all the different cutting conditions had been completed, the experiment was ended.

In the experiments made using the same cutting conditions, the new blades used last for some time before they reach a certain amount of wear (which is consistent with what happens in actual cutting jobs). Tool wear is a process that only varies gradually. Hence, only the experimental data obtained in the last 3 s of each cutting process will truly reflect the state of tool wear corresponding to the measured $VB$. This greatly reduces the amount of data that needs to be analyzed and makes subsequent data processing much more convenient. A method that cutting time is gradually increased is used in each cutting experiment to facilitate the calibration of cutting time. The wear on the tools was measured using the following procedure:

1) The tool shank and positions of tool bits 1 and 2 were marked. A black spot was marked where the generatrix of the shank was vertical to the normal of the primary tool flank. An infrared (IR) emitter was fixed onto the Smartzoom 5 so that IR ray emitted was parallel to the normal of the primary tool flank and coincided with the black spot (Figs. 3a and 3b).

2) To guarantee that different tools were in the same position under the lens, a white rectangular block was pasted onto the display screen. During each observation process, the end of the tool bit was checked to ensure that it coincided with the block (Fig. 3c).
3) To ensure the same measuring position was used, the coordinate system in the Smartzoom 5 was employed (the coordinate system was determined according to special points in the tool image taken before wear was incurred). The tool nose was taken as the origin. A distance equal to two-thirds of the cutting depth used in the experiment was measured along the blade line. Then, the horizontal and vertical coordinates of the point corresponding to this two-thirds depth were ascertained. The straight line passing through this horizontal coordinate was employed as the line used to monitor the tool wear. Clearly, the tool nose could not be determined once the tool had started to wear. Therefore, the tool root was used as the standard location to which the relative amount of tool wear was recorded. Subsequently, all the coordinates used were recorded and used to calculate the degree of tool wear (Figs. 3d and 3e).

4) Pictures were taken of the tools after putting them into the set reference position. Using the coordinates determined before wear started, the wear-monitoring line could be obtained from the tool image. The initial wear point on the line was manually moved to the bottom edge of the tool wear area and the vertical coordinate recorded. The extent of tool wear could then be acquired by finding the difference between the vertical coordinates before and after the movement. Finally, the data was recorded (Fig. 3f).

5) The extent of tool wear grows with time (i.e. the wear value measured each time is larger than that previously measured). Thus, any decline in the extent of wear suggests that errors are present in the experimental results (this may be caused by various uncertain factors, e.g. local non-uniformity in the tool texture, inaccurate monitoring of the wear line, etc.). Therefore, the whole of the wear area was analyzed macroscopically during data acquisition to help eliminate (as far as possible) the effects of accidental factors on the experimental data. In addition, measurements were conducted three times in order to avoid errors during the measurement process. These measurements were then averaged and the mean taken as the extent of wear of the tool flank.

![Fig. 3 The tool wear offline measurement](image)

Experiments were conducted using orthogonal experiments based on three factors at four levels (the experimental variables include the cutting rate $v$, feed per tooth $f$, and cutting depth $a_c$). The factors and levels used in the milling experiments are given in Table 2, and Table 3 constitutes the orthogonal experiment table employed.

<table>
<thead>
<tr>
<th>Table 2 The factors and levels of milling experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factors</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>$v$ (m/min)</td>
</tr>
<tr>
<td>$f$ (mm/tooth)</td>
</tr>
<tr>
<td>$a_c$ (mm)</td>
</tr>
</tbody>
</table>

| Table 3 Orthogonal experiment table of milling experiment |
## 2.4 Variation in tool wear and extraction of sample data

By using a microscope with a super-high depth of field, the extent of wear of the tool flanks could be observed and measured. The wear mode and variation in the extent of wear \( VB \) with cutting length \( L \) is shown in Figs. 4 and 5 (\( v = 100 \text{ m/min} \), \( f = 0.05 \text{ mm/tooth} \), and \( a_p = 3 \text{ mm} \)). Fig. 4 shows that when the stainless steel (0Cr18Ni9) was milled, wear primarily occurred to the rake face, flank, and boundary. Also, multi-mode wear probably appears simultaneously due to the complex nature of the cutting process. As stainless steel is a hard material to machine, serious friction occurs between the rake face and sheared metal, as well as between the flank and machined surface. As a result, high contact pressure and temperature is generated which causes the rake face, flank, and boundary of the tool to wear.

### Table 1: Cutting Parameters

<table>
<thead>
<tr>
<th>Group number</th>
<th>( v ) (m/min)</th>
<th>( f ) (mm/tooth)</th>
<th>( a_p ) (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>0.05</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>0.06</td>
<td>3.5</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>0.07</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>0.08</td>
<td>2.5</td>
</tr>
<tr>
<td>5</td>
<td>70</td>
<td>0.05</td>
<td>3.5</td>
</tr>
<tr>
<td>6</td>
<td>70</td>
<td>0.06</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>70</td>
<td>0.07</td>
<td>2.5</td>
</tr>
<tr>
<td>8</td>
<td>70</td>
<td>0.08</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>90</td>
<td>0.05</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>90</td>
<td>0.06</td>
<td>2.5</td>
</tr>
<tr>
<td>11</td>
<td>90</td>
<td>0.07</td>
<td>4</td>
</tr>
<tr>
<td>12</td>
<td>90</td>
<td>0.08</td>
<td>3.5</td>
</tr>
<tr>
<td>13</td>
<td>110</td>
<td>0.05</td>
<td>2.5</td>
</tr>
<tr>
<td>14</td>
<td>110</td>
<td>0.06</td>
<td>3</td>
</tr>
<tr>
<td>15</td>
<td>110</td>
<td>0.07</td>
<td>3.5</td>
</tr>
<tr>
<td>16</td>
<td>110</td>
<td>0.08</td>
<td>4</td>
</tr>
</tbody>
</table>

Fig. 4 shows the wear mode and variation in the extent of flank face wear \( VB \) with cutting length \( L \). The wear process in milling tools can be divided into three stages: primary, normal, and dramatic. In the primary wear stage (0–3 m), the wear rate is rapid. When the extent of wear increases to 0.1 mm, the wear rate declines to a small value. We then enter the normal wear stage (3–30 m), during which the extent of wear slowly grows from 0.1 to 0.3 mm. Finally, in the

---

dramatic wear stage, the extent of wear of the tool flank increases significantly. Fig. 5 shows that the extent of wear of two tool flanks which were started simultaneously vary in the same way with cutting length. The samples used in the subsequent analysis were extracted based on the overall change in tool wear. The tool wear can be divided into six states, according to the extent of wear of the tool flanks (Table 4).

![Fig. 5 The variation of the flank face wear VB with cutting length L](image)

Table 4 Wear states of the tool flanks

<table>
<thead>
<tr>
<th>States</th>
<th>Extent of wear /mm</th>
<th>State number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$0 \leq VB \leq 0.10$</td>
<td>VB1</td>
</tr>
<tr>
<td>2</td>
<td>$0.10 &lt; VB \leq 0.15$</td>
<td>VB2</td>
</tr>
<tr>
<td>3</td>
<td>$0.15 &lt; VB \leq 0.20$</td>
<td>VB3</td>
</tr>
<tr>
<td>4</td>
<td>$0.20 &lt; VB \leq 0.25$</td>
<td>VB4</td>
</tr>
<tr>
<td>5</td>
<td>$0.25 &lt; VB \leq 0.30$</td>
<td>VB5</td>
</tr>
<tr>
<td>6</td>
<td>$0.30 &lt; VB$</td>
<td>VB6</td>
</tr>
</tbody>
</table>

3. EEMD and feature extraction

3.1 The EEMD method

The EEMD method reduces the degree of modal mixing by adding white noise to the recorded signal. This is done so that the signal exhibits a certain continuity on different scales based on the statistical nature (uniformly distributed frequency) of the Gaussian white noise. The specific decomposition steps used and principles involved are as follows (Zhang et al., 2012):

1) The ensemble average time $M$ is initialized and numerical amplitudes assigned to the newly added white noise. Moreover, $i = 1$.

2) White noise $n_i(t)$ of given amplitude is added to the original signal $x(t)$ to form a new signal $x_i(t)$. Where, $x_i(t)$ and $n_i(t)$ refer to the $i$th additional noise signal and the $i$th white noise series, respectively. The Gaussian white noise directly affects the decomposition of the signal using EEMD and avoids modal mixing occurring.

$$x_i(t) = x(t) + n_i(t) \quad i = 1, 2, 3 \cdots M$$  \hspace{1cm} (1)

3) The new signal $x_i(t)$ is decomposed to a sum of IMFs using the EMD algorithm.

$$x_i(t) = \sum_{j=1}^{S} c_{i,j}(t) + r_{i,j}(t) \quad j = 1, 2, 3 \cdots S$$ \hspace{1cm} (2)

The number of IMFs $c_{i,j}(t)$ is $S$, and $r_{i,j}(t)$ (residual function) represents the average trend of the signal. The IMFs are labeled ($c_{1,1}, c_{1,2}, \ldots, c_{S,1}$) in order of frequency (from high to low).

4) Steps 2 and 3 are repeated $M$ times. The IMF sets thus obtained from the signals have white noise of different given amplitude added to them, as follows:

$$[e_{1,j}(t), e_{2,j}(t), \ldots, e_{M,j}(t)] \quad j = 1, 2, 3 \cdots S$$
5) The mean value of the decomposed IMF sets is calculated as the final result and $c_j(t)$ is taken to be the IMF as decomposed by EEMD.

$$ c_j(t) = \frac{1}{M} \sum_{i=1}^{M} c_{ij}(t) \quad i = 1, 2, 3 \ldots M \quad j = 1, 2, 3 \ldots S $$

(3)

The larger the value $M$, the closer the sum of the corresponding IMF with white noise approaches to zero. In this context, the resulting EEMD decomposition is:

$$ x(t) = \sum_{j} c_j(t) + r(t) $$

(4)

where, $r(t)$ refers to the final residual component, which represents the average trend of the signal. Each signal, $x(t)$, can be decomposed into a sum of a finite number of IMFs and a residual component using the EEMD method. The intrinsic mode components $c_j(t)$ represent the component signal bands from high to low frequency. Each frequency band contains different frequency components, which also vary as the signal $x(t)$ changes.

Fig. 6 displays an original vibration acceleration signal for a tool in a worn state ($v = 70$ m/min, $f = 0.07$ mm/z, $a_p = 2.5$ mm, and $VB = 0.079$ mm). The signal can be decomposed into 13 IMF components and one residual component using the EMD method, as shown in Fig. 7. It can be seen from this figure that the non-stationary vibration acceleration signal arising from the tool’s wear state is decomposed into a finite number of stationary IMF components by the EMD method. In addition, the different IMF components correspond to different timescales.

Fig. 8 shows the decomposition results obtained under the same conditions using EEMD. It can be seen from a comparison of Figs. 8 and 7 that the latter has a lower degree of modal mixing when the amplitudes change in the same range. Obviously, if there is basically no difference in the waves while the amplitude range varies, then the degree of modal mixing with the smaller change in amplitude range is less than that with larger changes. Thus, it can be concluded that the degree of modal mixing using EEMD is lower than that found using EMD.

3.2 Sensitive IMFs extraction

The next step is to identify the most sensitive IMFs. As well as information on the state of tool wear, the original vibrational signals also contain various irrelevant information (state of the machine, etc.) and background noise. The original signals have been decomposed into a series of IMF components representing the intrinsic signals using EEMD. By eliminating the IMFs containing noise and information irrelevant to the wear state of the tool, we can improve the accuracy with which the tool wear is identified. Depending on the characteristics of the tool-wear-induced vibrational signals, the steps involved in this process are as follows:

1) Signal preprocessing: Apart from useful information, the vibrational signals collected from the experimental tool-wear identification process also contain much interference. The signals tend to exhibit troublesome characteristics including non-linearity, non-stationarity, and lack smoothness. They therefore deviate from ideal (true) behavior which causes errors and even mistakes to be made in the subsequent analysis. Therefore, the preprocessing of the vibrational signals is a very important part of their analysis.
The noise present can be reduced using a nonlinear method of wavelet thresholding. Generally speaking, after wavelet transformation, the wavelet coefficients with large amplitudes mainly correspond to genuine signals, while those of low amplitude are primarily noise. The threshold noise-reduction method refers to when noisy signals are processed via wavelet decomposition and then a threshold used to discard the information related to noise. Wavelets whose coefficients are larger than the set threshold are retained and those that are not have their coefficient set to zero. The processed wavelet coefficients can then be used to reconstruct the (noise-reduced) signal. The white noise in the signal can completely constrained by using the threshold method. The threshold value employed is generally calculated using one of four methods: rigrsure, heursure, sqtwolog, or minimax. The most suitable thresholds to use to reduce noise in the data used to determine tool wear are based on the signal-to-noise ratio (SNR). For threshold selection, there are 4 threshold selection rules: rigrsure, heursure, sqtwolog, and minimax. When choosing threshold to reduce noise, there is always a contradiction between noise removal and retention of high frequency components of useful signals. The rigrsure rule is an adaptive threshold selection based on Stein's unbiased likelihood estimation principle. Given a threshold lambda $\lambda(j)$, its likelihood estimate is obtained, and then the likelihood $\lambda(j)$ is minimized to obtain the selected threshold. This is a software threshold estimate. Sqtwolog is a fixed threshold form. It produces a
threshold of $\lambda$. The best predictor variable threshold is chosen by heursure. If the signal-to-noise ratio is very small, the signal processed by the unbiased likelihood estimation principle has very large noise, and this threshold type is needed. Minimax uses the maximum and minimum principle to generate the threshold, and uses the minimum mean square error as the objective function to produce an extremum instead of no error. It is a threshold selection method based on the statistical extremum estimator principle. This threshold estimator minimizes the maximum mean square error in a given function (Su and Jun, 2008).

![Graphs showing decomposition results obtained under the same conditions using EEMD](image)

2) Calculation of correlation coefficients: The correlation coefficients between the IMF components and original signal are determined according to (Nie et al., 2012):

$$r = \frac{\text{cov}(X, Y)}{\sqrt{D(X)}\sqrt{D(Y)}} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}}$$

(5)

where $\text{cov}(X, Y)$ is the covariance between two sequences $X$ and $Y$ whose variances are $D(X)$ and $D(Y)$.
3) The correlation coefficient $\alpha_n$ between $x(t)$ and $c_i(t)$ is calculated. The correlation coefficient $\beta_n$ between $c_i(t)$ and the signal $x_f(t)$ of the machine tool under no-load operation is also found.

4) The difference between the two correlation coefficients is calculated to obtain the total correlation coefficient.

$$\chi_n = \alpha_n - \beta_n$$

5) The wear-sensitive factors of the IMF components of the signal $x(t)$ are defined and calculated (Lei, 2011).

$$\delta_n = \frac{\chi_n - \min(\chi)}{\max(\chi) - \min(\chi)} \quad \chi \in \chi_n \quad n = 1, 2, 3 \ldots N$$

6) The wear-sensitive IMFs are chosen according to the wear-sensitive factors and displayed in descending order of magnitude.

7) The differences between pairs of adjacent wear-sensitive factors are calculated.

8) The index $n$ corresponding to the maximum difference calculated above is found.

The algorithm used to select the most wear-sensitive IMFs is based on the following principles. IMFs that reproduce many features of the vibration signal $x(t)$ will be similar to the signal itself and have large correlation coefficients. Therefore, the correlation coefficients ($\alpha_n$) can be considered as factors that can be used to evaluate the sensitivity of the IMFs. The signal $x_f(t)$ of the machine tool operating without load is taken as reference. The correlation coefficients $\beta_n$ between $x_f(t)$ and vibration signal $x(t)$ are calculated to evaluate the non-wear information that is irrelevant to tool wear and contained in the IMFs of the signal $x(t)$. As well as the similarity between the IMFs and signal itself, the calculation of the IMF sensitive factors also takes the similarity between the IMFs and non-wear signals into account. Therefore, the selection algorithm for the sensitive IMFs can enrich the information regarding the potential components relating to tool wear, and reduce the information of the machine tool under no-load operation that is irrelevant to wear, so as to obtain more distinct wear features.

### 3.3 Calculation of IMF energy and energy entropy

The frequency distribution of the vibrational signal changes when the tool is in a state of wear. As a result, the energy distribution of the vibrational signal also undergoes corresponding changes in the worn state. The vibration acceleration signal containing information on tool wear is decomposed into a group of IMF components using EEMD and then the energy $E_i$ of each IMF component is calculated (Xia et al., 2014).

$$E_i = \int |x_i(t)|^2 dt \quad i = 1, 2, 3 \ldots M$$

The energy vector of the IMFs is then obtained.

$$T = [E_1, E_2, E_3 \ldots E_n]$$

The above vectors are normalized so as to obtain the following result.

$$T' = [p_1, p_2, p_3 \ldots p_n]$$

$$p_i = \frac{E_i}{E}$$

$$E = \sum_{i=1}^{n} E_i$$

Based on this, the definition of EEMD energy entropy is acquired (Yang et al., 2006).

$$H_{EN} = -\sum_{i=1}^{n} p_i \log p_i$$

The quantity $p_i$ refers to the normalization energy of each IMF component and represents the weight of each IMF in the total wear energy. The components of the energy vector $T'$ decrease in frequency from $p_1$ to $p_n$ successively. For example, $p_1$ represents the energy of the first IMF component $c_1$ (the energy component with the highest frequency) while $p_n$ refers to the energy of $c_n$, namely, the energy component with the lowest frequency.
4. Principles of LS-SVM

LS-SVM is a method that improves on the traditional SVM method. In LS-SVM, inequality constraints appearing in traditional SVM are transformed into equality constraints, and the square error and loss function are considered to be the empirical loss of the training set. In this way, a quadratic programming problem is converted into a problem involving the solution of a set of linear equations. This makes solution of the problem faster and improves convergence precision. The basic idea of the LS-SVM classification algorithm is to transform the input space to a higher space using a non-linear transformation defined by a kernel function in order to make the samples linearly separable. The optimal classification face is constructed in the higher space to form the rule used to perform sample classification. The basic LS-SVM theory is mainly used to deal with two-class classification problems. Multi-class classification problems are transformed into two-class ones by using combined classifiers according to the basic idea of the LS-SVM. Moreover, one of two modes are used in the combined classifier: one-to-one (as employed in this study) and one-to-many. All the possible two-class classifiers are constructed in the $N$ training samples using the algorithm (so that there are $N(N–1)/2$ LS-SVM sub-classifiers). In addition, the two-class classifiers are combined and the voting method is used to choose the class with maximum votes as the class of points to be measured.

LS-SVM algorithm: We are given a training sample \( \{x_i, y_i\}_{i=1}^{l} \in \mathbb{R}^n \), where \( x_i \in \mathbb{R}^l \) and \( y_i \in \mathbb{R} \) refer to the input and output of the training sample, respectively. The objective optimization function in the LS-SVM algorithm is (Dai, 2011):

$$
\min \frac{1}{2} (\omega \cdot \omega) + c \sum_{i=1}^{l} \xi_i^2

\phi(x_i) \cdot \omega + b + \xi_i \quad i = 1, \ldots, l
$$

where, \( \phi(x) \), \( \omega \), \( \xi \), \( b \) and \( c \) represent the mapping function of the kernel space, weight vector, error variable, offset value, and adjustable parameter, respectively. To find the minimum value of the optimization function, a Lagrange function is established.

$$
\min \frac{1}{2} (\omega \cdot \omega) + c \sum_{i=1}^{l} \xi_i^2 - \sum_{i=1}^{l} \alpha_i (\phi(x_i) \cdot \omega + b + \xi_i - y_i)
$$

(10)

where, \( \alpha_i \) refers to Lagrangian multiplier. Eq. (10) is further transformed to obtain \( \alpha \) and \( b \), so the LS-SVM for function estimation is:

$$
y(x) = \sum_{i=1}^{l} \alpha_i K(x_i, x) + b
$$

(11)

The kernel function \( K(x_i, x_j) \) can be freely chosen. To sum up, the original optimization problem to be solved is transformed into a set of linear equations using an LS-valued function. The constraints are also converted to equalities which greatly reduces the complexity of the problem.

5. Experimental verification

According to the theory outlined above, the procedure used to predict tool wear based on EEMD energy entropy and LS-SVM is shown in Fig. 9.

![Fig. 9 The flow chart of tool wear state recognition](image)

1) Signal preprocessing

The experiments conducted were based on Table 3. For each cutting condition employed, 30 samples were
collected under six tool-wear states, producing 480 vibration signal samples in total. The degree of tool wear was such that the measured \( VB \) value changed only slightly if 420 mm per feed was used to generate the data used to determine tool wear. Thus, 2100 mm (that is, cutting 420 mm for 5 times) was used. The changes associated with 2100 mm (cutting length) are considered to satisfy a linear relation. Table 5 displays the correlation found between the denoised signals (using four wavelet bases) and the original signals, and the SNRs. The table indicates that there are good correlations and high SNRs after implementing noise reduction using the heursure and rigrsure wavelet bases. Therefore, the heursure wavelet basis was used to reduce the noise in the rest of this study. Fig. 10 shows the vibration signals obtained from the cutting tools that have been denoised using the heursure wavelet basis.

Table 5  The correlation found between the denoised signals (using four wavelet bases) and the original signals, and the SNRs

<table>
<thead>
<tr>
<th></th>
<th>Heursure</th>
<th>Rigrsure</th>
<th>Sqtwolog</th>
<th>Minimax</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R )</td>
<td>0.9950</td>
<td>0.9950</td>
<td>0.9526</td>
<td>0.9720</td>
</tr>
<tr>
<td>( \text{SNR} )</td>
<td>19.15</td>
<td>19.15</td>
<td>8.86</td>
<td>11.21</td>
</tr>
</tbody>
</table>

![Fig. 10 The noise reduction processing using heursure wavelet basis.](image)

2) EEMD

Each signal was decomposed using the EEMD method to obtain a finite number of IMF components. The different signals produce different numbers of IMF components and so the first \( M \) IMFs containing wear information were taken to use as research objects. The correlation coefficients and wear-sensitive factors were calculated according to Eqs. (5)–(7), Determining the IMF component closely related to the original signal and contain the primary information of the original signal. The other IMF components are poorly correlated with the original signal and contain noise from the original signal. Therefore, the other IMFs are abandoned to achieve signal denoising, dimension reduction, and normalization. vibration of high frequency component is likely to occur. If the high frequency component in the IMF component is the multiple of milling the intermittent period of spindle, consideration is caused by intermittent vibration.

The correlation coefficients and wear-sensitive factors were calculated according to Eqs. (5)–(7), as shown in Fig.11. In Fig.11. \( \alpha_n, \beta_n, \chi_n \) and \( \delta_n \) represent the correlation coefficient between \( x(t) \) and \( c_n(t) \), the correlation coefficient \( \beta_n \) between \( c_n(t) \) and the signal \( x(t) \), the total correlation coefficient and the wear-sensitive factors, respectively. \( n \) means that the collected signals can be decomposed by EEMD and get \( n \) IMF components.

It can be seen from the Fig.11 that the first four IMF components exhibit a favorable correlation with the original signal and contain the primary information of the original signal. The other IMF components are poorly correlated with the original signal and contain noise from the original signal. Therefore, the other IMFs are abandoned to achieve signal denoising, dimension reduction, and normalization. Fig. 12 displays the IMF sensitive factors calculated using the sensitivity algorithm and shows them displayed in descending order. It can be seen that the maximum difference occurs between \( c_2 \) and \( c_3 \). Thus, \( c_1 \) and \( c_2 \) for the first two IMFs are regarded as the wear-sensitive factors, and are used to construct the energy of the tool-wear signal. The selection algorithm has effectively selected the IMFs relevant to tool wear and can thus highlight the characteristics of the tool wear. In order to strengthen the dispersion features of the different energies of each vibration signal under different tool wear states, however, the first four IMF components are adopted to calculate the energy entropies.
3) Calculation of energies and energy entropies of the first four IMFs

The percentage energies of the first four IMFs of the tool-wear vibration data collected in experiment No. 7 (under the sixth state) were calculated according to the methods introduced in Eqs. (8) and (9). The results are shown in Fig. 13. It can be seen from the figure that the energy distributions corresponding to different extents of wear change after EEMD decomposition. That is, the percentage energy gives a good indication of the state of wear of the tools and fully exhibits the disparities between selected eigenvalues. Fig. 14 displays the eigenvalues of the data from different groups under the same state ($0.01 < V_B < 0.02$). The eigenvalues exhibit the same trends. That is, the selected eigenvalues have favorable repeatability. Only those eigenvalues with favorable disparities and repeatability can result in accurate identification. Therefore, these two points determine the quality of the selected eigenvalues. It can be seen from the above analysis that the eigenvalues chosen in this study satisfy the above points, so they are satisfactory eigenvalues.

![Fig. 11 Correlation coefficients calculated according to Eqs. (5)-(7)](image1)

![Fig. 12 IMF sensitive factors calculated using sensitivity algorithm](image2)

![Fig. 13 Characteristic value differences of percentage energies of the first four IMFs of the tool-wear vibration data collected in experiment No. 7](image3)
Fig. 14 The normalization energy repetitive, $E_i/E_c$, of each IMF $cn$, for various cases with $0.01< VB<0.02$

4) All the test signals are collected to form eigenvectors according to the above Steps (1)–(3). Then, the eigenvectors are normalized to make up an energy eigenvector matrix $T(480, 5)$ to use as the input to the LS-SVM. A description of the data set is given in Table 6. The eigenvectors of the training samples in just experiment No.7 (under all wear states) are displayed in Table 7. In Table 7, these numbers of the rightmost column represent tool wear states of training samples according to the actual extent of wear of the measured tool and the wear condition classification in Table 4.

5) The eigenvectors of the energy and energy entropy are established.

6) The classifier for the tool wear states composed of SVMs is then established. Then, the energy eigenvector $T$ of the IMFs related to tool wear is input into the SVM to train the SVM.

7) Using, in turn, FFT (Guan et al., 2014), Wavelet analysis (Guan et al., 2014), EED and EEMD method, the characteristics vectors of the samples are input into multiple classifiers composed of SVMs for training. The times of the sigmoid kernel function is $d = 3$, kernel parameter of the Gaussian RBF, and error penalty factor are, $\gamma = 0.5$, and $C = 1$, respectively. The computational accuracy for ending iteration is set at $\varepsilon = 0.001$. Finally, the eigenvectors of the test samples in 20 groups under 6 tool wear states (shown in Table 6) are input into the trained SVMs to identify the tool wear modes. The eigenvectors of the test samples in experiment No.4 under 6 tool wear states (shown in Table 6) are input into the trained SVMs to identify the tool wear modes. The prediction classification results are displayed in Table 8. The classification results of test samples using different feature vector extraction methods are displayed in Table 9.
Table 8 The eigenvectors of the test samples and prediction classification result in just experiment No.4 (Part Samples)

<table>
<thead>
<tr>
<th>Wear states</th>
<th>Wear extent /mm</th>
<th>Eigenvector</th>
<th>Prediction classification result</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$E_1/E$</td>
<td>$E_2/E$</td>
</tr>
<tr>
<td><strong>VB1</strong></td>
<td></td>
<td>0.013</td>
<td>0.4876</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.025</td>
<td>0.6414</td>
</tr>
<tr>
<td><strong>VB2</strong></td>
<td></td>
<td>0.019</td>
<td>0.5495</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.134</td>
<td>0.6420</td>
</tr>
<tr>
<td><strong>VB3</strong></td>
<td></td>
<td>0.159</td>
<td>0.5856</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.166</td>
<td>0.5292</td>
</tr>
<tr>
<td><strong>VB4</strong></td>
<td></td>
<td>0.228</td>
<td>0.6719</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.243</td>
<td>0.6423</td>
</tr>
<tr>
<td><strong>VB5</strong></td>
<td></td>
<td>0.264</td>
<td>0.6762</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.269</td>
<td>0.7105</td>
</tr>
<tr>
<td><strong>VB6</strong></td>
<td></td>
<td>0.314</td>
<td>0.7092</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.325</td>
<td>0.6379</td>
</tr>
</tbody>
</table>

Table 9 Classification results of test samples using different feature vector extraction methods

<table>
<thead>
<tr>
<th>Extraction methods</th>
<th>Wear states</th>
<th>Total recognition rate(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>VB1</strong></td>
<td><strong>VB2</strong></td>
</tr>
<tr>
<td>FFT</td>
<td>79</td>
<td>59</td>
</tr>
<tr>
<td>Wavelet analysis</td>
<td>82</td>
<td>100</td>
</tr>
<tr>
<td>EMD</td>
<td>83</td>
<td>100</td>
</tr>
<tr>
<td>EEMD</td>
<td>86</td>
<td>100</td>
</tr>
</tbody>
</table>
Table 9 shows that the identification rate of the LS-SVM classifier based on the EEMD is clearly superior to those of the other three classifiers. The LS-SVMs based on Gaussian RBF can efficiently identify the state of tool wear in the test samples. This shows that our method based on EEMD energy entropy and use of LS-SVM is an effective one to identify the tool-wear state. At the same time, the recognition rate of the VB6 sample is the lowest in Table 9 because of the edge gap emerged easily and the rapid wear of the cutting tool in the dramatic wear stage. The samples collected are less, and the amount of wear is distributed among the 0.30-0.35mm, and the coverage of the sample is narrow, which leads to insufficient training. It affects the overall recognition rate. However, the error identification of such samples does not affect the performance of the tool.

6. Conclusions

The vibrational signals from milling tools subject to wear were processed using an EEMD method to explore the energies of the principal components that contain information about the state of tool wear. The energies obtained were normalized to generate energy eigenvectors to use as input to an LS-SVM for training and testing purposes. The following conclusions can be made after comparing the predictions made using three different functions:

1) EEMD is a self-adaptive signal processing method which can be used to precisely process non-linear and non-stationary signals. The degree of modal mixing using EEMD is lower than that achieved using the EMD method.

2) The energies of the signals in the different frequency bands change when the tools are worn. Hence, the state of the tool wear can be identified by calculating the EEMD energies and energy entropies of the different vibrational signals.

3) The correlation coefficients between each IMF component and original signal can be calculated and sensitive IMFs chosen according to wear-sensitive factors. Moreover, the tool-wear state can be successfully identified by combining the energy features extracted from the wear-sensitive IMFs containing primary tool-wear information with the LS-SVM.

4) The predictions made using an LS-SVM based on EEMD method are more accurate than those made using FFT, Wavelet analysis and EMD methods.

Acknowledgement

This project is supported by National Natural Science Foundation of China (Grant No. 51465029). The authors thank to Gansu province university collaborative innovation team construction plan for their support (Grant No. 2016C-07) and Lanzhou talent innovation and entrepreneurship Project(Grant No. 2015-RC-4).

References


Guo, S., Classification and prediction of tool wear based on multi feature analysis and fusion of acoustic emission signals, Doctoral Dissertation of Jilin University, (2011).


Huang, N. E., Shen, Z. and Long, S. R., The empirical mode decomposition and the Hilbert spectrum for nonlinear and


