Mechanical Systems and Signal Processing 156 (2021) 107671

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

Milling chatter detection by multi-feature fusion and Adaboost-SVM

Shaoke Wan [a,b,*], Xiaohu Li [a,b], Yanjing Yin [a,b,c], Jun Hong [a,b]

[a] Key Laboratory of Education Ministry for Modern Design & Rotor-Bearing System, Xi'an Jiaotong University, Xi'an, Shaanxi, China
[b] School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi, China
[c] Luoyang Bearing Research Institute Co., Ltd., Luoyang, Henan, China

A R T I C L E I N F O

Article history:
Received 4 October 2020
Received in revised form 15 December 2020
Accepted 19 January 2021
Available online 4 February 2021

Keywords:
Milling chatter detection
Multi-feature fusion
Strong classifier
Adaptive boosting
Support vector machine

A B S T R A C T

Unstable chatter vibration in the milling process significantly affect the machining quality and efficiency. In order to suppress or avoid the chatter vibration in the cutting operation, detection of chatter onset is highly needed. Until now, most of the existing chatter detection methods designed chatter indicators by extracting signal features, and the threshold of designed chatter indicator is usually needed, which is difficult to determine and might not be applicable in different cutting conditions. In fact, chatter detection is essentially a typical classification problem, hence milling chatter detection based on machine learning method is presented in this paper. In order to obtain the needed data set, milling experiments under different cutting conditions were performed. Multi-features are utilized for the chatter detection, including the dimensionless features in time domain and frequency domain, and the automatic features extracted by stacked-denoising autoencoder (SDAE). In order to improve the accuracy of chatter classification and avoid the negative effects of possible samples with wrong labels, adaptive boosting (Adaboost) algorithm that consists of a series of weak classifiers by support vector machine (SVM) is utilized and further improved. Experimental verification and performance analysis are also performed, and the results show that the presented method can detect the chatter with a high accuracy and is applicable in different milling conditions.

© 2021 Elsevier Ltd. All rights reserved.

1. Introduction

Chatter occurs as undesired vibration in the machining process, leading to poor surface finishes, tool damage and severe noise, and conservative cutting parameters are usually selected to guarantee a stable cutting operation, which significantly decrease the machining efficiency [1]. In order to guarantee the machining quality of workpiece and improve machining efficiency, the chatter detection and suppression methods have always been important issues in the past several decades [2]. Due to the fact that it has been desired to apply certain actions which can suppress or avoid the chatter during machining operation [3], the chatter detection is quite important and highly needed.

Until now, different methods and technologies have been proposed to detect the chatter in cutting process [2], and it can be found that the conventional chatter detection process mainly includes the monitoring signal acquisition, feature extrac-

* Corresponding author at: Key Laboratory of Education Ministry for Modern Design & Rotor-Bearing System, Xi'an Jiaotong University, Xi'an, Shaanxi, China.

E-mail address: skwan2020@xjtu.edu.cn (S. Wan).

https://doi.org/10.1016/j.ymssp.2021.107671
0888-3270/© 2021 Elsevier Ltd. All rights reserved.

tion and indicator design for chatter identification. As the chatter vibration occurs between the cutting tool and workpiece, it certainly reflects in the variation of vibration state, sound, and the cutting load. Hence, different kinds of monitoring signals have been utilized for chatter detection, including vibration signal [4–8], motor current signal [9–11], cutting force or torque signal [12–17], sound signal [18–20] and so on, and different kinds of sensors are needed. In order to improve the robustness and reliability of chatter detection method under variable cutting conditions, researchers also utilized multi-sensors to extract more features for the chatter detection. Kuljanic et al. [21] compared sensitivity of different sensors to the chatter onset in milling, and found that the accuracy and robustness of chatter identification were improved with combination of force and vibration signals. Pan et al. [22] investigated the boring chatter identification with multi-sensors, and found that the recognition results can be improved with multi-features extracted from different kinds of sensors. However, though more features can be extracted with different sensors and the chatter identification results can be improved, the installation of sensors (such as force sensor and displacement sensor) might be costly and difficult, which is not applicable in practical application of milling process.

In order to extract the chatter-sensitive features and design the related chatter indicators for chatter detection, different signal processing methods have been utilized, mainly including time-domain method, frequency-domain method and time–frequency domain method. Time domain methods usually extracts the time domain features of the monitoring signal, and design corresponding indicators for chatter detection. Schmitz T L [23] utilized the variance of synchronously audio signal as the chatter indicator for chatter detection in milling. Ye et al. [24] extracted the standard deviation and mean the of root mean square sequence, respectively, and determined their ratio as the indicator to identify the onset of chatter. Wan [25] proposed an adaptive filter to extract the chatter sensitive components, which can process the monitoring signal on-line, and designed a special dimensionless indicator in time domain to detect the onset of milling chatter at early stage. Considering that the distribution of frequency components in the monitoring signals changes when chatter occurs in the machining process, some frequency domain methods were also proposed. Kondo et al. [26] introduced a new criterion for detecting regenerative chatter by application of spectral analysis. Thaler et al. [27] extracted features in frequency space through pre-processing signals, with which the chatter can be successfully detected with a carefully selected threshold. Tangjitsitcharoen [28] utilized the power spectrum density of dynamic cutting force signals to identify the chatter onset in turning. Considering that the process when cutting state becomes unstable (chatter) from stable is non-stationary, the time-frequency domain methods for chatter detection were investigated extensively, with which the instantaneous and sudden occurrence of chatter are easily detected. Due to the excellent ability of extracting features from non-stationary signal, wave packet decomposition (WPD) has been widely utilized for the chatter detection. Liu et al. [29] utilized WPD to reconstruct the signals based on wavelet packet node with the maximum energy, and then the chatter features were extracted. Chen et al. [27] applied the WPD to preprocess the monitoring signals, and then different kinds of feature were selected and compared for the chatter identification. Considering that each wave packet's energy in WPD changes during the development of chatter, Yao et al. [30] designed two indicators for chatter detection based on the standard deviation of wavelet transform and the wavelet packet energy ratio in the chatter-emerging frequency band, and the results showed that the method is applicable in different cutting operations. Empirical mode decomposition (EMD) and ensemble empirical mode decomposition (EEMD) were also utilized for the chatter detection. Ji et al. [31,32] proposed online milling chatter detection methods based on EEMD and improved EMD, respectively, with which the signals were decomposed into a series of intrinsic mode functions (IMFs), and then the chatter-sensitive IMFs were utilized to identify the chatter, in addition, multi-indicator synthetic evaluation for chatter detection was also proposed. Fu et al. [33] proposed an energy aggregation characteristic-based chatter detection method, in which the measured vibration signals were decomposed to a series of IMFs with EEMD, and then chatter-sensitive components were extracted based on the majority energy rule. With the development of time–frequency analysis technology, a new non-recursive signal processing method, which is called as variational mode decomposition (VMD), has also been utilized in the chatter detection. Zhang et al. [8] decomposed the non-stationary chatter signal based on VMD and WPD, and utilized the energy entropy to evaluate the onset of chatter. Liu et al. [34] presented a chatter detection method based on VMD, in which the parameters of VMD was selected automatically, and a chatter indicator based on the energy entropy was also proposed. Over all, it can be found that the conventional chatter detection technologies usually utilize different signal processing methods to extract features that can reflect the chatter state, and then different chatter indicators are designed to identify the chatter onset, and it should be noted that the threshold of indicator for chatter detection is very essential. Though different indicators were proposed in the above mentioned works, the selection of threshold for the chatter detection has always been an important issue, and a fixed threshold is usually determined based on cutting experiments or personal experience. However, whether the preselected threshold is applicable in other cutting conditions still cannot be guaranteed.

Due to the fact that chatter detection is essentially a classification problem, hence some intelligent classification methods based on machine learning were utilized for the chatter detection. Tarng et al. [35] presented an on-line drilling chatter identification method with an adaptive neural network, and a spindle regulation method was utilized to suppress the chatter once the chatter was detected. Zhang et al. [36] combined the advantages of artificial neural network (ANN) and hidden markov model (HMM) together to monitor the cutting process, and experimental results showed that the cutting chatter can be detected efficiently. Kwak et al. [37] also applied the neural network as a diagnostic technique of grinding state to evaluate the grinding chatter. Lamraoui et al. [6] proposed a milling chatter detection method with neural network classification, in which the selected features were classified into stable and chatter, and the selection of threshold was not required. Considering that massive data is usually needed with neural network, support vector machine (SVM) was utilized to detect the machining chatter, in which relatively small size of samples is needed. Peng et al. [15] proposed a new chatter stability prediction method, in which the SVM was utilized to find the chatter stability boundary based on the simulated dynamic cutting force. Shao et al. [38] presented a chatter recognition method based on principal component analysis and SVM, and the feature obtained with fast Fourier transform (FFT) was utilized. These intelligent classification methods can avoid the selection of threshold, and usually take high applicability in different machining conditions. However, the accuracy of chatter detection is highly determined by the selected features for classification, and some extracted features are still sensitive to cutting conditions, which easily affect the generalization ability of proposed methods. Meanwhile, the accuracy of designed classifier might not be sufficient for chatter detection and should be improved.

In summary, the chatter detection methods have been investigated extensively. However, in the most of existing conventional chatter detection methods, thresholds of designed chatter indicators are usually needed and quite difficult to determine, and a fixed threshold might not be applicable in other cutting conditions. Though some intelligent classification methods based on ANN or SVM show advantage of unneeded threshold selection, the selected features for the training of chatter identification directly affect the performance of classification, and appropriate features are usually needed. In addition, the chatter detection is essentially implemented with the classifier designed by these intelligent classification algorithms, hence the classifier with high accuracy is highly needed.

In order to improve the performance of chatter detection in milling, a new chatter detection method based on multi-feature fusion and Adaboost-SVM is presented in this paper. To simplify the process of data set production and enrich the needed training data set, milling experiments with different cutting parameters, workpiece shape and even the workpiece material, are performed. Multi-features including the time domain features, frequency domain features are utilized. In addition, features extracted by stacked-denoising autoencoder (SDAE) are also considered, and different combination of features are then compared. In order to improve the accuracy of chatter classification, adaptive boosting (Adaboost) algorithm which consists of a series SVM classifiers (Adaboost-SVM) is utilized to train a strong classifier for chatter detection, which can increase the classification accuracy for chatter detection. In addition, considering the inevitable mislabeled samples in the training data set, an improved Adaboost-SVM algorithm is then proposed. After that, the performance of presented chatter detection method is analyzed and verified.

The rest of this paper is organized as follows. Section 2 presents the signal acquisition and features extraction from the monitoring signals. In Section 3, the proposed methodology is presented. The analysis and verification of proposed chatter detection method are both shown in Section 4. Conclusions are drawn in Section 5.

2. Signal acquisition and feature extraction

In this section, the signal acquisition is firstly performed to obtain the needed data set. Then the conventional features including time domain features and frequency domain features are carefully selected based on Student's *t* test (*t*-test). In addition, the automatic feature extraction based on SDAE is also presented. In order to improve the feature extraction quality, different types of feature combination are also analyzed and compared.

2.1. Signal acquisition

As shown in Fig. 1, milling experiments are performed on a three-axis milling machine tool. Due to the fact the vibration signal measured with acceleration sensor is more sensitive to chatter state [21] and the sensor's installation is simple, hence the vibration signals are utilized for chatter detection in this paper. During the experiments, an B&K4525 type accelerometer

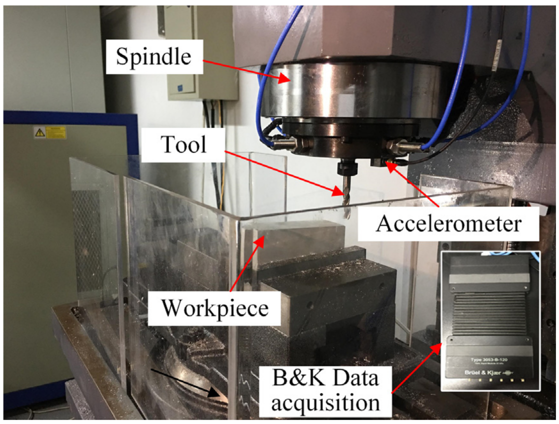

Fig. 1. Experimental setup for chatter detection.

sensor is attached to the front end of spindle system, and a B&K data acquisition system is used to record the data, with a sampling frequency 4096 Hz.

Actually, the chatter detection with intelligent learning algorithm is essentially a typical two-classification problem, which classifies the monitoring signal into stable or chatter, and the supervised learning method is usually needed. For a supervised learning method, the data set directly affects the performance of classification model. In order to improve the quality of data set for the chatter classification model, milling experiments under different cutting conditions were performed. During the milling experiments, an end mill cutter with two flutes, diameter 6 mm is utilized. In addition, two types of workpiece with different shapes are used, including the wedge-shaped workpiece and step-shaped workpiece (shown in Fig. 2), for the convenience of expression, the two types of workpiece are denoted as #1 and #2 respectively. With the workpiece #1, the cutting depth changes from 0 mm to 8 mm with a cutting length 100 mm, and the cutting depth varies with 0.5, 1, ···, 5 mm when the workpiece #2 is selected. To enrich the information of data set, a series of milling experiments with different cutting conditions were performed, as listed in Table 1. In the milling experiments performed on the workpiece #1, the axial cutting depth increases along the feed direction, and the cutting state changes from stable to chatter gradually. Fig. 3 shows the surface finishes of machined wedge-shaped workpiece, it can be found that the surface finishes varies gradually with the increment of axial cutting depth, which indicates the milling process changes from stable to unstable. Detailed surface finishes at three stages with a length of cutting time 5 s are also presented in Fig. 3. During stage A (17–22 s), there is no chatter marks and the milling process is stable. With the increment of cutting depth, slight chatter marks can be found (shown in stage B (38–53 s)), and significant chatter marks emerge during the stage C (65–70 s), which means the occurrence of serious chatter. With workpiece #2, the monitoring signals under stable and chatter state can also be obtained.

All the monitoring vibration signals in the above milling experiments are utilized as the needed data set for classifier design, including the training data set and testing data set. Due to the fact that the labeling of samples in the data set is needed and it is usually time-consuming to label the onset of chatter manually by spectrum analysis, hence, it is desired to label the data set based on the related surface finishes. In each milling experiment, 100 samples with a length of 512 sampling signal points are extracted from the stable and chatter conditions respectively, and 1800 samples are obtained in total. When the chatter marks exist on the machined surface finishes, the related samples are labeled as chatter, if not, the samples are labeled as stable. All the samples are divided into training data set and testing data set randomly, and the training data set include 1400 samples and the testing data set contain 400 samples. The ratio of chatter samples and stable samples keep equal in both training data set and testing data set. It can be found that the data set utilized in this paper takes the advantage of comparable proportions between the chatter samples and stable samples (with same ratio), simple labeling process and including different cutting conditions.

2.2. Conventional feature extraction

Similar with the conventional chatter detection methods by extracting all kinds of features from the monitoring signals, features extracted from the samples in the data set are also needed and also directly affect the performance of chatter identification model. When chatter occurs, the monitoring signals change in both time domain and frequency domain [21], hence, time-domain features and frequency-domain features are usually utilized.

Generally, there are a series of typical time-domain statistical indicators (listed in Table 2, where x_k denotes the sampling point of time series) utilized to represent the changes of time series, which can be used as the features for chatter detection. However, some time-domain statistical indicators, such as average amplitude and variance, are easily affected by the cutting conditions and not conducive to the chatter detection. Hence, in this paper, six dimensionless time-domain statistical indicators are selected as the time-domain features used for chatter detection, including waveform index S_f, peak index C_f, pulse index CL_f, skewness index S_v, kurtosis index K_v and margin index I_f.

In addition, considering that some hidden information cannot be obtained directly from the above selected time-domain features, some frequency-domain features are also introduced to the chatter detection. Due to the fact that the frequency components change with the occurrence of chatter [33], hence three conventional frequency-domain features are utilized, including the gravity frequency index FC, mean square frequency index MSF and frequency variance index VF, and the expressions are as follows:

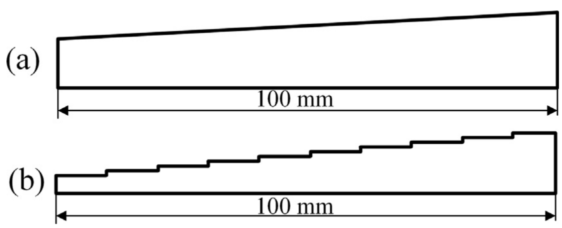

Fig. 2. Schematic diagram of workpiece in milling experiments: (a) wedge-shaped; (b) step-shaped.

Table 1
Parameters of milling experiments under different cutting conditions.

Test No.	Shape of workpiece	Workpiece material	Spindle speed (rpm)	Axial cutting depth (mm)	Radial cutting depth (mm)	Feed rate (mm. s^{-1})
1	#1	AL7075	3600	0–8	2	1
2	#1	AL6061	5100	0–8	2	1
3	#2	AL6061	2700	0.5–5	2	1
4	#2	AL6061	2850	0.5–5	1	1
5	#2	AL7075	3700	0.5–5	2	1
6	#2	AL6061	4200	0.5–5	2	0.5
7	#2	AL6061	5400	0.5–5	1	1
8	#2	AL7075	5500	0.5–5	2	0.5
9	#2	AL6061	5700	0.5–5	2	0.75

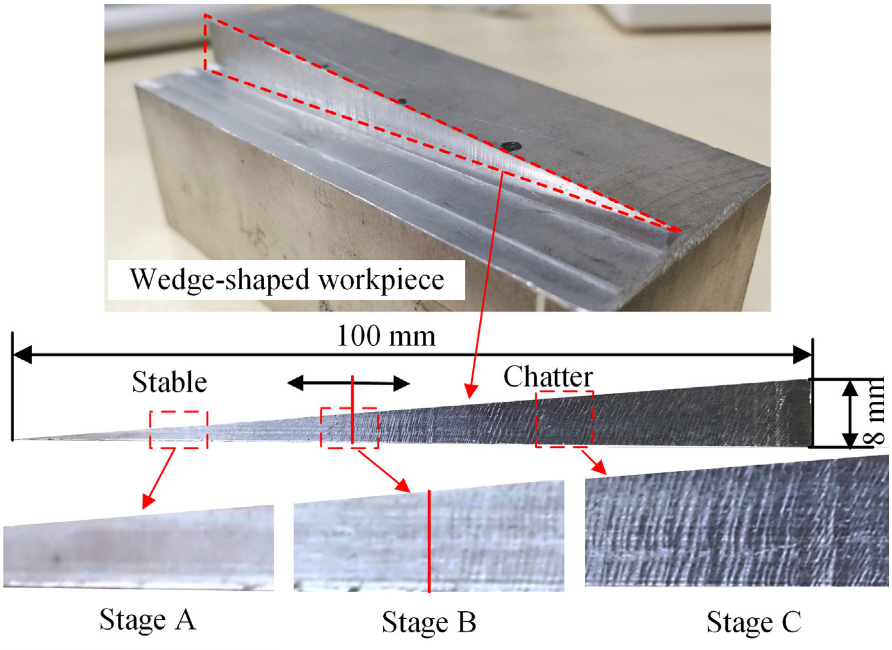

Fig. 3. Surface finishes of machined workpiece (wedge-shaped).

Table 2
Typical time-domain statistical indicators.

Root mean square	Square root amplitude	Average amplitude	Skewness	Kurtosis	Variance
$x_{rms} = \sqrt{\frac{1}{N}\sum_{k=1}^{N} x_k^2}$	$x_s = \left(\frac{1}{N}\sum_{k=1}^{N}\sqrt{\lvert x_k\rvert}\right)^2$	$x_a = \frac{1}{N}\sum_{k=1}^{N}\lvert x_k\rvert$	$S = \frac{1}{N}\sum_{k=1}^{N}\lvert x_k\rvert^3$	$K = \frac{1}{N}\sum_{k=1}^{N} x_k^4$	$\sigma^2 = \frac{1}{N}\sum_{k=1}^{N} x_k^2$
Waveform index	Peak index	Pulse index	Skewness index	Kurtosis index	Margin index
$S_f = \frac{x_{rms}}{x_a}$	$C_f = \frac{x_{max}}{x_{rms}}$	$CL_f = \frac{x_{max}}{x_a}$	$S_v = \frac{S}{\sigma^3}$	$K_v = \frac{K}{\sigma^4}$	$I_f = \frac{x_{max}}{x_s}$

$$\mathrm{FC} = \sum_{i=0}^{n} \omega_i x(\omega_i) / \sum_{i=0}^{n} x(\omega_i) \tag{1}$$

$$\mathrm{MSF} = \sum_{i=0}^{n} \omega_i^2 x(\omega_i) / \sum_{i=0}^{n} x(\omega_i) \tag{2}$$

$$\mathrm{VF} = \sum_{i=0}^{n} (\omega_i - \mathrm{FC})^2 x(\omega_i) / \sum_{i=0}^{n} x(\omega_i) \tag{3}$$

where $x(\omega_i)$ denotes the Fourier transform of time-domain signals $x(t)$, and ω_i is the related frequency state, n denotes the length of Fourier transform series.

Next, the selected features' sensitivities to chatter are examined. The selected nine features are calculated and shown in Fig. 4. For the chatter detection, the probability distribution of the selected feature in stable condition should be different with the one in chatter condition, which means the feature should be sensitive to cutting state. Through the results in Fig. 4, though overlap between the two types samples (chatter and stable) exist to some extent, the differences in distribution still can be found, except for the feature of skewness index S_v. To further test the selected features' differences in terms of chatter and stable state, Student's t test (t-test) method is utilized to test whether obvious differences exist in the feature's distribution between chatter and stable, and only the feature of skewness index S_v doesn't satisfy the significant difference level when the level value is set as 10%, hence S_v is removed from the above selected features.

2.3. Automatic feature extraction based on stacked-denoising autoencoder

In Section 2.2, it can be found that the selection of conventional features is highly dependent on the operator's experience. Hence, a novel feature extraction strategy is proposed based on autoencoder (AE). AE is developed by Rumelhart et al. [39] and has been widely utilized for the representation learning of inputted information. The AE is essentially one kind of special ANN, and its structure is shown in Fig. 5. The input layer and hidden layer are called as encoder, which compresses the input signal $\boldsymbol{Y}_i$ to the output of encoder $\boldsymbol{H}_i$ with smaller dimensions:

$$\boldsymbol{H}_i = f(\boldsymbol{Y}_i\boldsymbol{W}) \tag{4}$$

where $\boldsymbol{Y}_i$ is the i th input vector, $\boldsymbol{W}$ denotes the weighting matrix, and $f(\cdot)$ is the activation function.

Then, the hidden layer and output layer are consisted as decoder, and $\boldsymbol{H}_i$ is reconstructed as $\boldsymbol{X}_i$ with the output layer:

$$X_i = f(H_iW') \tag{5}$$

where $\boldsymbol{X}_i$ is the i th output vector, W' denotes the weighting matrix.

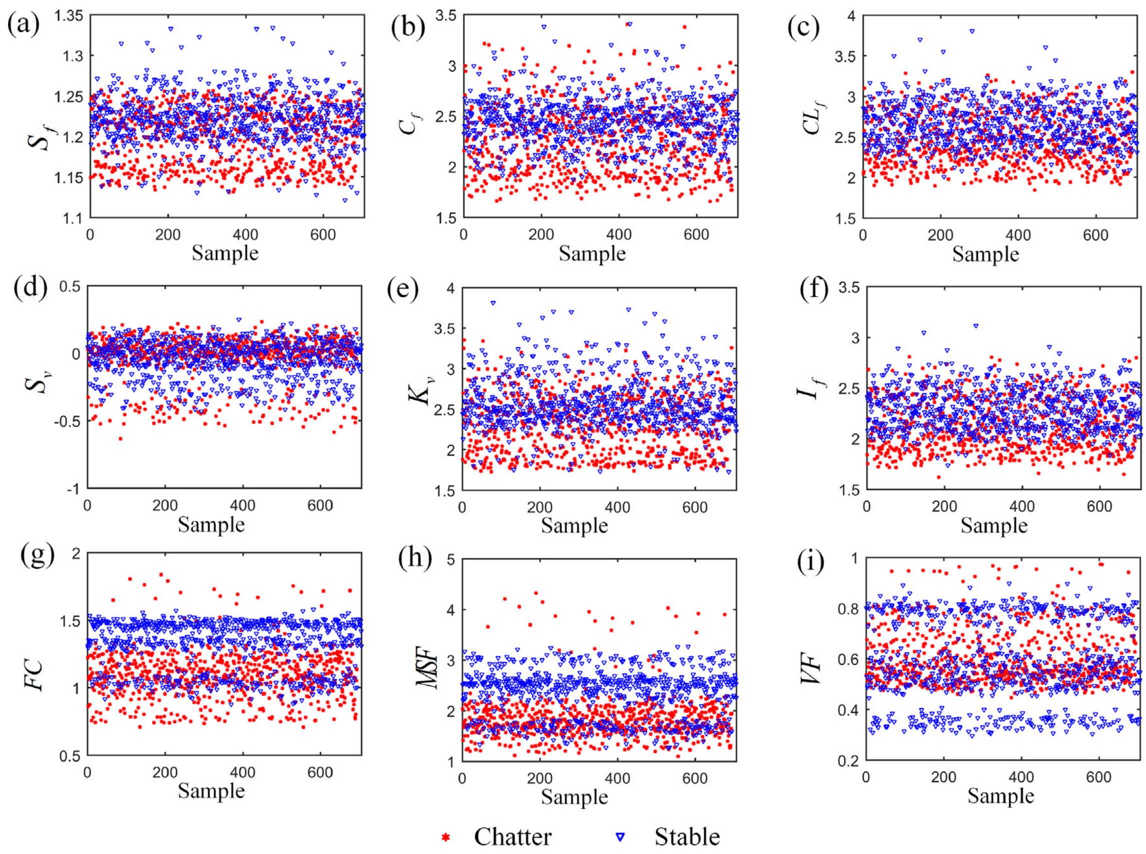

Fig. 4. Distribution of selected features with the training data set (with 1400 samples).

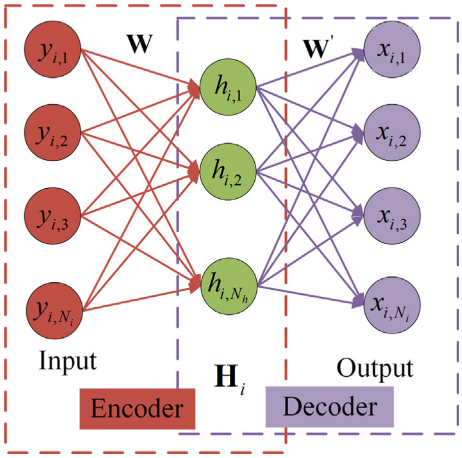

Fig. 5. Typical structure of the autoencoder.

Through the basic working principle, it can be found that the output of encoder $\boldsymbol{H}_i$ actually represents the major characteristics hidden in the original data, hence the output of encoder $\boldsymbol{H}_i$ is proposed to be utilized as the features extracted from the monitoring signals for chatter detection.

In order to improve the quality of extracted features with AE, the stacked-denoising autoencoder (SDAE) [40] with multi denoising autoencoder (DAE) is utilized in this paper, and the structure is shown with Fig. 6. Here, a three-stacked SDAE is utilized, and the dimension of data is compressed to 1/4 of the input data with each DAE, hence an output data with length of 8 is obtained for a samples with length of 512 data points, and the elements in the final output data are utilized as the features for chatter detection.

2.4. Multi-features combination and analysis

In order to improve the quality of features for chatter detection, different types of multi-features are compared, including the manually selected features (shown in Section 2.2), automatic features extracted by SDAE and their combination. Fig. 7 shows the combination features with manually selected features and the features extracted by SDAE.

Similar with the distribution analysis in Section 2.2, the distribution of different types of multi-features are also performed. Considering that the data set with multi-features take high-dimensional space and are difficult to be illustrated directly, the t-distributed stochastic neighbor embedding (t-SNE) algorithm proposed by Laurens et al. [41] is utilized here to reduce the dimensionality of data set, and the results are shown in Fig. 8. Through Fig. 8(a), it can be found that the chatter and stable state can be distinguished roughly with the manually selected features. However, significant overlap between the samples with different labels exist. Meanwhile, the samples with same state (chatter or stable) distributes with multiple clusters, which means these manually selected features still have certain relationships with the cutting conditions. Though the results of automatic features extracted by SDAE (shown in Fig. 8(b)) illustrate a better performance than the manually selected features, overlap and multiple clusters still can be found. Fig. 8(c) shows the results of combination features with

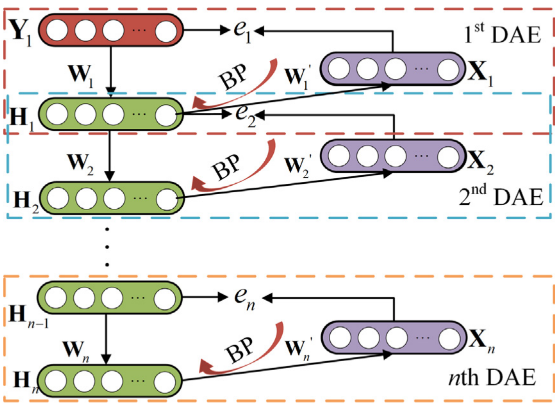

Fig. 6. Structure of the SDAE.

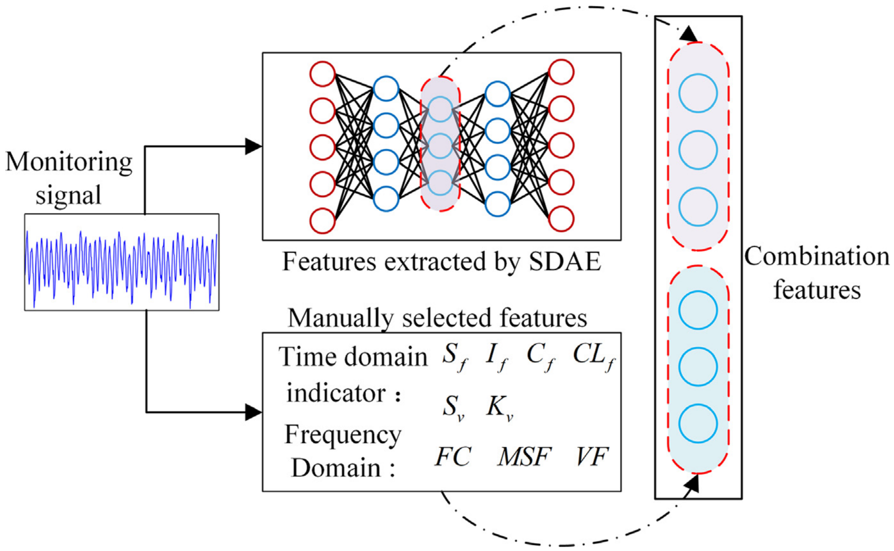

Fig. 7. Combination features with manually selected features and features extracted by SDAE.

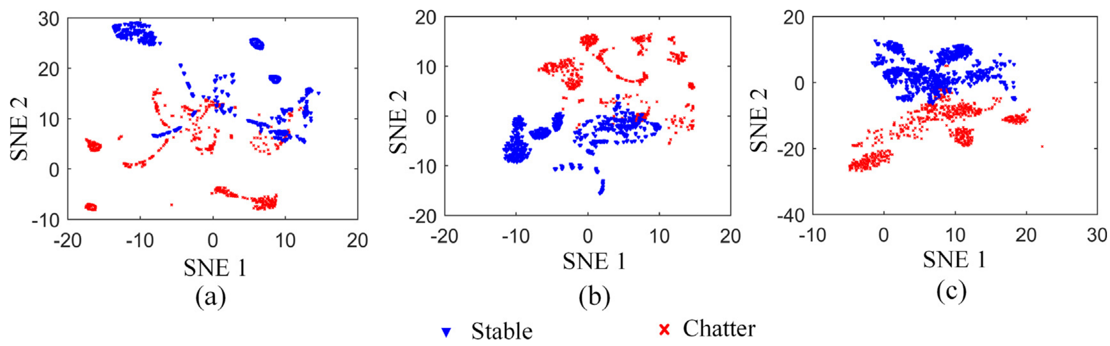

Fig. 8. Distribution of different features by t-SNE: (a) manually selected features; (b) features extracted by SDAE; (c) combination features.

manually selected features and automatic features by SDAE, and it can be found that the degree of overlap between the samples with different labels is further reduced, in addition, the distributions of samples with same label are more gathered, which is quite beneficial for the chatter classification.

3. Chatter identification with improved Adaboost-SVM

Next, the needed classifier for chatter identification is designed in this section. Due to the fact that SVM takes advantages of excellent performance in the generalization ability, high accuracy in classification and strong applicability in the condition with small data set, it has been widely utilized for the machine condition monitoring [42]. The accuracy of SVM for the chatter detection with composed features was tested in this paper, and the results showed that the accuracy was lower for the chatter detection (less than 90%). Hence, it is desired to construct a strong classifier with the Ensemble Learning method [43], with which the accuracy of chatter classification can be significantly improved. In this paper, adaptive boosting (Adaboost) algorithm developed by Freund et al. [44] is utilized to construct a strong classifier with a series of weak classifiers, and the SVM is selected as the needed weak classifier.

Fig. 9 illustrates the schematic structure of Adaboost-SVM. The working principle of Adaboost-SVM can be described as follows: firstly, a subset is obtained by sampling the points in the training data set with an initially same sampling weight D_1; then, the SVM is utilized as the first weak classifier, and the classification error of the first weak classifier in the training data set is calculated, and the sampling weight of subset for the next weak classifier D_2 is further updated, in addition, the weight of 1st weak classifier in the final strong classifier α_1 is also obtained. Next, iterate the above process with an updated sampling weight until the T th classifier. Finally, a strong classifier can be obtained by combining these weak classifiers with respective weight α_m , $m = 1, 2, \cdots, T$.

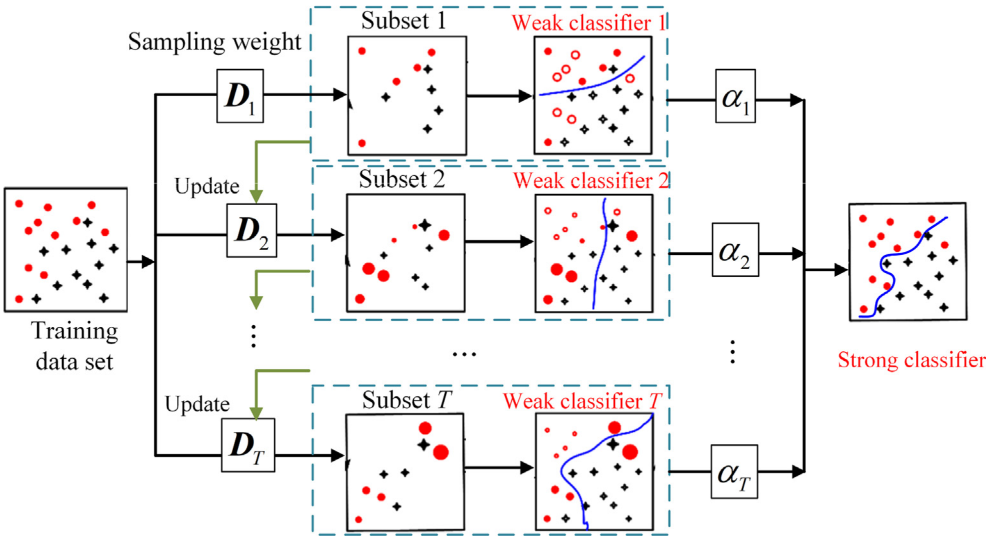

Fig. 9. Schematic structure of the Adaboost-SVM.

The main idea of Adaboost algorithm is to continuously increase the sampling weight of incorrectly classified samples and reduce the sampling weight of correctly classified samples simultaneously. Hence the classification model can always focu on the samples that are difficult to classify. However, the labeling of samples in the training data set might be wrong in some conditions, and these samples with wrong labels easily affect the performance of final strong classifier, as the sampling weight increases significantly once these mentioned samples were incorrectly classified due to the wrong labels. Unfortunately, in order to quickly make the needed data set for chatter identification model, the labels of samples are roughly determined based on the surface finishes with and without chatter marks (as descripted in Section 2.1), and the samples are easily labeled incorrectly, as chatter might have actually happed even no obvious chatter marks exist. Hence, in this section, an improved Adaboost algorithm is proposed to reduce the negative effects of samples with wrong labels.

Table 3 shows the improved Adaboost algorithm, the samples in the training data set are labeled with -1 and 1, which represents chatter and stable state, and weak classifier algorithm can be selected manually. Compared with the original algorithm, two parameters are introduced in the improved Adaboost algorithm: consecutive misclassification times of samples $\beta_{m,i}$ (for the m th weak classifier and i th samples), and inhibitory factor r. Different with the conventional Adaboost algorithm, the updating sampling weight is replaced by

$$D_{m+1,i} = \frac{D_{m,i}}{Z_m} e_m^{\alpha_m h_m(x_i)\cdot y_i \frac{2}{1 + e^{r\cdot\beta_{m,i}}}} \tag{6}$$

and when inhibitory factor $r = 0$, the improved Adaboost algorithm becomes conventional one. Z_m is the normalization factor, with the following expression:

$$Z_m = \sum_{i=1}^{n} D_{m,i} \exp(-\alpha_m y_i h_i(x_i)) \tag{7}$$

In order to further test the performance of improved Adaboost algorithm, simulations are performed to compare the actual classification accuracy with conventional Adaboost-SVM and improved Adaboost-SVM algorithm. Suppose there are 100 chatter samples and 100 stable samples that satisfy the two-dimensional normal distribution $N(-2, 2, 4, 4, 0)$ and $N(2, -2, 4, 4, 0)$ respectively, and the chatter and stable state are labeled with -1 and 1 respectively. The simulation data set is shown in Fig. 10, 20% of the chatter samples and stable samples are selected randomly and labeled with wrong labels. A conventional Adaboost-SVM and improved Adaboost-SVM are utilized to obtain the needed strong classifier for the data set shown in Fig. 10. During the simulations, the number of weak classifiers is 5 and the size of the subset for weak classifier is 20, and the inhibitory factor in the improved Adaboost algorithm is set as $r = 1$. The simulation results with conventional Adaboost-SVM and improved Adaboost-SVM are listed in Tables 4 and 5 respectively.

Through the simulation results listed in Tables 4 and 5, it can be found that the actual classification accuracy reaches to 96% with the improved Adaboost-SVM, and only 84.5% can be reached by the conventional Adaboost-SVM. The results indicate that the samples with wrong label significantly affect the classification accuracy of strong classifier, and the improved Adaboost can well guarantee the classification accuracy even the samples with wrong labels possibly exist, which is quite important for the proposed chatter detection method in this paper.

Table 3
The pseudo-code of improved Adaboost algorithm.

Input :	Training data set $S = \{(x_1, y_1), (x_2, y_2) \ldots (x_n, y_n)\}, y_i \in \{-1, 1\}$
	Training subset size: M
	Weak classifier algorithm: ς
	Numbers of weak classifier : T
	Inhibitory factor：r
Output :	Classification results: $H(x_i) \in \{-1, 1\}$

Initialize sampling weight: $D_{1,i} = n^{-1}$, $i = 1, 2, \ldots, n$
Initialize consecutive misclassification times of the sample：$\beta_{1,i} = 0, i = 1, 2, \ldots, n$
For m=1：T
Training subset sampling: S_m
Classification results with weak classifier: $h_t = \varsigma(S_m)$
Weighted classification error：$e_m = \sum_{i=1}^{n} D_{m,i} \delta(h_m(x_i) \neq y_i)$
if $e_m > 0.5$ **then**
$\alpha_m = 0$ **, continue**
end if
$\alpha_m = \frac{1}{2} \ln\left(\frac{1 - e_m}{e_m}\right)$
Update sampling weight：$D_{m+1,i} = \frac{D_{m,i}}{Z_m} e_m^{\alpha_m h_m(x_i) \cdot y_i \frac{2}{1+e^{r \cdot \beta_{m,i}}}}$
if $h_m(x_i) \neq y_i$ **then** $\beta_{m+1,i} = \beta_{m,i} + 1$
else $\beta_{m+1,i} = 0$ **end if**
end for
The final classification results：$H(\boldsymbol{x}) = \sum_{m=1}^{T} \alpha_m h_m(\boldsymbol{x})$

4. Verification and performance analysis

In this section, the performance of proposed method is analyzed and verified. Considering that the selected features directly affect the accuracy of chatter classification, hence the performance of chatter classification with different types of multi-features are compared. The actual performance of improved Adaboost-SVM under the condition that parts of samples are incorrectly labeled is also investigated. After that, more milling experiments with different cutting conditions are also performed to test the generalization ability of developed strong classifier for chatter detection.

4.1. Chatter classification performance with different types of multi-features

In Section2, different features for chatter classification have been presented, including the manually selected features in time-domain and frequency-domain, automatic features extraction by SDAE and combination features by the two types of multi-features. The distribution of features by t-SNE (shown in Fig. 8) indicates that the composed features show better clustering performance, in order to further investigate the actual chatter classification performance with different types of multi-feature, the classification results with the improved Adaboost-SVM are compared. With different types of multi-features, different classifiers are trained, and then the classification accuracy for the testing data set is also counted. During the tests, training data set in Section 2 is utilized for the training of strong classifier, with subset size 100, weak classifier number 50 and inhibitory factor $r = 1$. Fig. 11 shows the classification accuracy with different types of multi-features for chatter detection. It is found that, with the manually selected features, only about 90% accuracy can be achieved in both training data set and testing data set. With the automatic feature extraction based SDAE, the classification accuracy increases to 93% both in training data set and testing data set. However, with the combination features by manually selected features and features extracted with SDAE, the classification accuracy increased significantly to about 97%, which indicates that the combination features are more applicable for the chatter detection. Hence, combination features are selected for the chatter detection in this paper.

4.2. Performance analysis considering samples with wrong labels

In Section 3, an improved Adaboost algorithm has been proposed considering possible samples with wrong labels. In order to further verify the performance of proposed method, the training data set and testing data set are utilized in this section, and part of the samples in the training data set are intentionally marked with wrong labels to test the actual accuracy in both training data set and testing data set. Different ratios of samples with wrong labels are considered, and the comparisons between the Adaboost-SVM and improved Adaboost-SVM are also performed, and the results are shown in Fig. 12. During the test, the subset size and weak classifier number are same in the two methods, which are set as 100 and 50 respec-

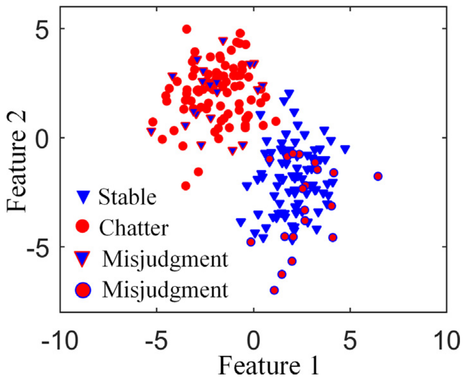

Fig. 10. Distribution of data set for simulation.

tively, and inhibitory factor $r = 1$ for the improved Adaboost-SVM, and maximum 30% of samples with wrong labels is considered. Through the results, it can be found that the samples with wrong labels exactly affect the classification accuracy of developed classifier, and the accuracy decreases with a higher ratio of samples with wrong labels for both conventional Adaboost-SVM and improved Adaboost-SVM. In addition, for the conventional Adaboost-SVM algorithm, the samples with wrong labels in the training data set significant affect the classification accuracy in the testing data set. However, with the improved Adaboost-SVM that is proposed in this paper, the accuracy in the testing data set can be achieved to almost same level with the training data set, which indicates that the proposed improved-Adaboost-SVM can well reduce the decline of classification accuracy caused by the possible samples with wrong labels in the training data set. It should also be noted that, though the data set selection method by labeling the samples roughly based on the surface finishes significantly shorten the needed time, some samples with wrong labels easily exist in data set, hence the proposed improved-Adaboost-SVM is quite useful for the actual application of chatter detection with classification method.

4.3. Generalization ability analysis

Though the proposed method shows high classification accuracy in the testing data set, the chatter detection performance of developed strong classifier under the other milling conditions is also desired to be verified, which can indicate the actual generalization ability of developed classifier with the improved Adaboost-SVM.

Two more milling experiments are performed with the experimental setup shown in Section2, and same wedge-shaped workpiece is utilized and the material is AL7075. Fig. 13 shows the chatter identification results with spindle speed 4800 rpm, radial cutting depth 2 mm and feed rate 90 mm/min. The length of each sampling data is 512, and combination features of the sampling data are extracted and then inputted to the trained model. For the convenience of chatter classification, the stable state and chatter state are noted as 1 and −1 respectively. Fig. 14 shows the results with spindle speed 5400 rpm, and the other cutting parameters are same with the ones in Fig. 13. Through the results, it can be found that the stable state and chatter state in the milling can be classified generally, which indicates the ability of chatter detection of proposed method with multi-features and an improved Adaboost-SVM. It is also found that misclassification exists in both classification results, especially around the stage that chatter is identified. However, these misclassifications can be explained as follows: it can be found that these misclassification occurs at the transition state from stable to chatter, in which the cutting state and corresponding signal characteristics are difficult to be recognized [33]. Hence, difference between the stable and chatter is not very obvious (the surface finishes shown in Fig. 2 also indicate this point). In summary, the results indicate that the proposed method can identify the chatter with an enough accuracy and is applicable for milling chatter detection in different milling conditions.

Table 4
Simulation results with conventional Adaboost-SVM.

Classifier	Classification accuracy	Weight of weak classifier α_m	Sampling weight of samples with wrong label
$h_1(\boldsymbol{x})$	82%	0.4	0.2
$h_2(\boldsymbol{x})$	86%	0.2416	0.2915
$h_3(\boldsymbol{x})$	79%	0.1001	0.3253
$h_4(\boldsymbol{x})$	50%	0.0718	0.3470
$h_5(\boldsymbol{x})$	85.5%	≈0	0.3486
$H(\boldsymbol{x})$	84.5%	–	–

Table 5
Simulation results with improved Adaboost-SVM.

Classifier	Classification accuracy	Weight of weak classifier α_m	Sampling weight of samples with wrong label
$h_1(\boldsymbol{x})$	69%	0.182	0.2
$h_2(\boldsymbol{x})$	64%	0.2327	0.2253
$h_3(\boldsymbol{x})$	58.5%	0.1139	0.2310
$h_4(\boldsymbol{x})$	95.5%	0.5277	0.2346
$h_5(\boldsymbol{x})$	99%	0,3331	0.3273
$H(\boldsymbol{x})$	96%	–	–

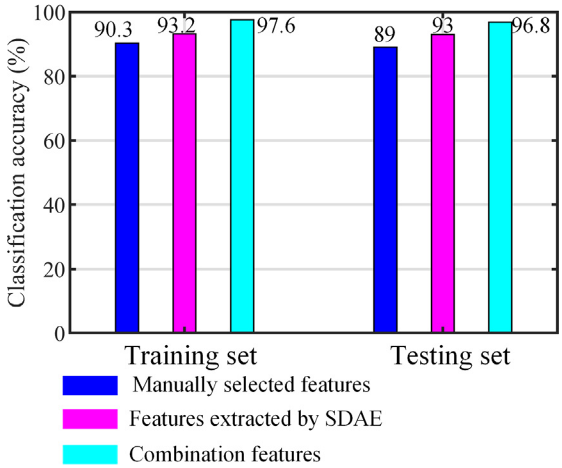

Fig. 11. Classification accuracy with different types of multi-features.

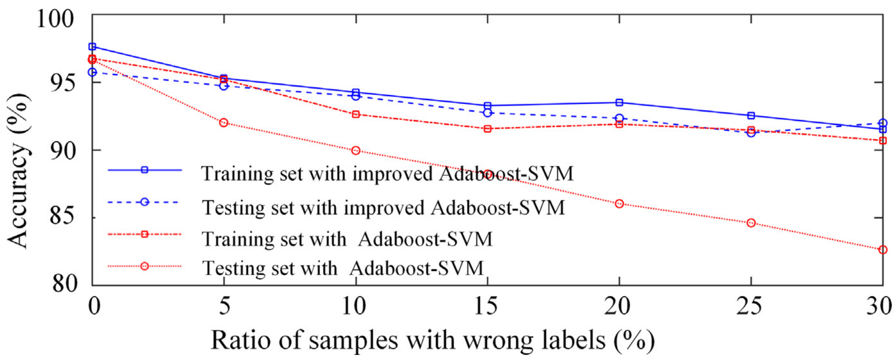

Fig. 12. Classification accuracy under different ratios of samples with wrong labels.

5. Conclusions

In the most of existing conventional chatter detection methods, thresholds of chatter indicators are needed and quite difficult to be determined, and the preselected threshold based on experiments might not be applicable in other conditions. In order to improve the performance of chatter detection in milling, a new chatter detection method based on multi-feature fusion and Adaboost-SVM is presented in this paper, which treats the chatter detection as a typical classification problem in machine learning. The following conclusions are derived from this research:

1. The combined 16 features, which consists of 8 manually selected features in time domain and frequency domain and 8 features automatically extracted by SDAE, highly improves the accuracy and reliability of milling chatter identification.
2. With the improved Adaboost-SVM algorithm presented in this paper, the decline of chatter classification accuracy caused by the possible samples with wrong labels in the training data set, which easily exist during the preparation of data set for chatter detection, can be avoided to a large extent.

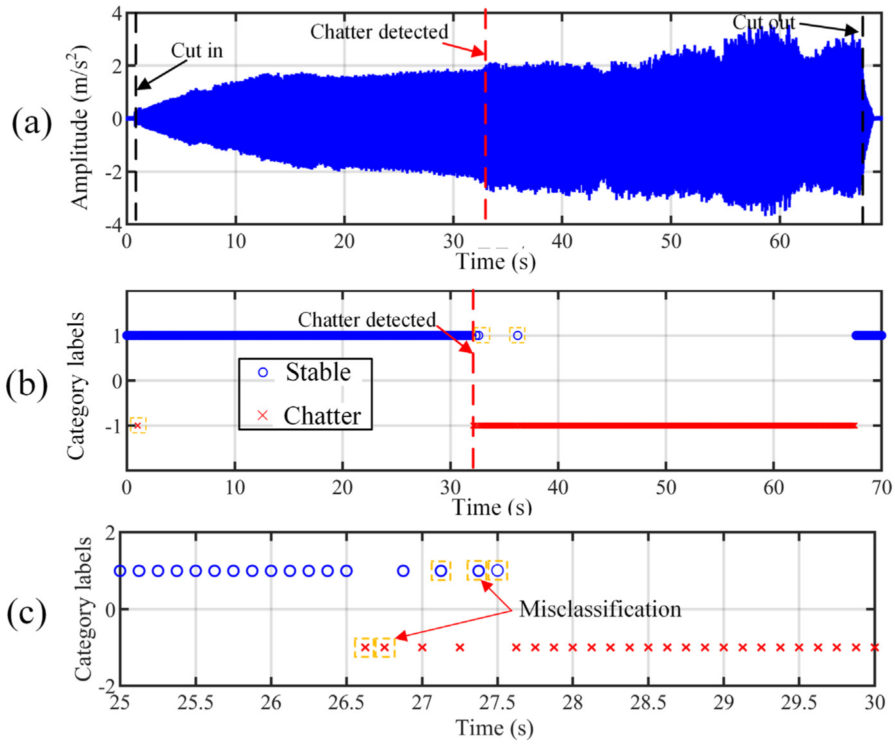

Fig. 13. Chatter identification results with proposed method (spindle speed 4800 rpm, radial cutting depth 2 mm and feed rate 90 mm/min).

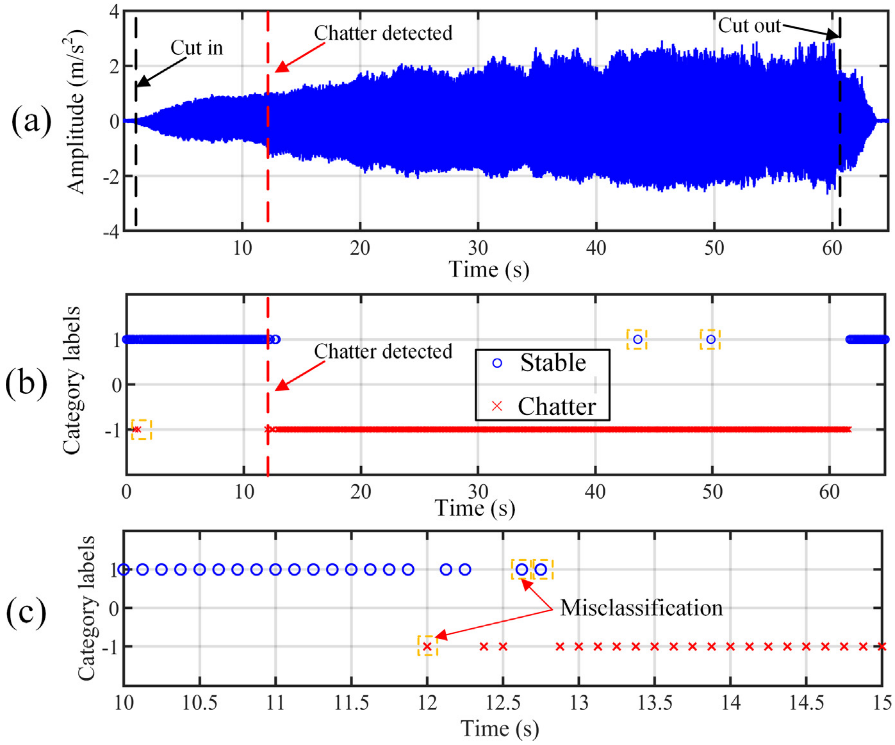

Fig. 14. Chatter identification results with proposed method (spindle speed 5400 rpm, radial cutting depth 2 mm and feed rate 90 mm/min).

3. With the proposed method, chatter can be well identified under more other milling conditions with the trained strong classifier by improved Adaboost-SVM, which avoid the needed threshold selection in conventional chatter detection methods.

CRediT authorship contribution statement

Shaoke Wan: Conceptualization, Methodology, Investigation. **Xiaohu Li:** Supervision, Conceptualization, Investigation, Funding acquisition. **Yanjing Yin:** Investigation, Visualization, Resources. **Jun Hong:** Supervision, Conceptualization, Project administration.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (No. 2018YFB2000504), Major Science and technology projects of Shaanxi Province of China (No. 2018zdzx01-02-01HZ01) and Open funded project of Henan Key Laboratory of high-performance bearing technology (No. 2020ZCKF04). The authors express their gratitude for their support. Our deepest gratitude goes to the anonymous reviewers for their careful work and thoughtful suggestions that have helped improve this paper substantially.

References

[1] Y. Altintas, M. Weck, Chatter stability of metal cutting and grinding, CIRP Ann. 53 (2) (2004) 619–642, https://doi.org/10.1016/S0007-8506(07)60032-8.
[2] L. Zhu, C. Liu, Recent progress of chatter prediction, detection and suppression in milling, Mech. Syst. Sig. Process. 143 (2020) 106840, https://doi.org/10.1016/j.ymssp.2020.106840.
[3] J. Munoa, X. Beudaert, Z. Dombovari, Y. Altintas, E. Budak, C. Brecher, G. Stepan, Chatter suppression techniques in metal cutting, CIRP Ann. 65 (2) (2016) 785–808, https://doi.org/10.1016/j.cirp.2016.06.004.
[4] X. Dong, W. Zhang, Chatter identification in milling of the thin-walled part based on complexity index, Int. J. Adv. Manuf. Technol. 91 (9-12) (2017) 3327–3337, https://doi.org/10.1007/s00170-016-9912-6.
[5] L. Zhu, C. Liu, C. Ju, M. Guo, Vibration recognition for peripheral milling thin-walled workpieces using sample entropy and energy entropy, Int. J. Adv. Manuf. Technol. 108 (9-10) (2020) 3251–3266, https://doi.org/10.1007/s00170-020-05476-7.
[6] M. Lamraoui, M. Barakat, M. Thomas, M.E. Badaoui, Chatter detection in milling machines by neural network classification and feature selection, J. Vib. Control 21 (7) (2015) 1251–1266, https://doi.org/10.1177/1077546313493919.
[7] H. Cao, Y. Yue, X. Chen, X. Zhang, Chatter detection based on synchrosqueezing transform and statistical indicators in milling process, Int. J. Adv. Manuf. Technol. 95 (1-4) (2018) 961–972, https://doi.org/10.1007/s00170-017-1283-0.
[8] Z. Zhang, H. Li, G. Meng, X. Tu, C. Cheng, Chatter detection in milling process based on the energy entropy of VMD and WPD, Int. J. Mach. Tools Manuf. 108 (2016) 106–112, https://doi.org/10.1016/j.ijmachtools.2016.06.002.
[9] D. Aslan, Y. Altintas, On-line chatter detection in milling using drive motor current commands extracted from CNC, Int. J. Mach. Tools Manuf. 132 (2018) 64–80, https://doi.org/10.1016/j.ijmachtools.2018.04.007.
[10] Y. Li, S. Zhou, J. Lin, X. Wang, Regenerative chatter identification in grinding using instantaneous nonlinearity indicator of servomotor current signal, Int. J. Adv. Manuf. Technol. 89 (1-4) (2017) 779–790, https://doi.org/10.1007/s00170-016-9117-z.
[11] Y. Liu, X. Wang, J. Lin, W. Zhao, Early chatter detection in gear grinding process using servo feed motor current, Int. J. Adv. Manuf. Technol. 83 (9-12) (2016) 1801–1810, https://doi.org/10.1007/s00170-015-7687-9.
[12] Y. Mei, R. Mo, H. Sun, K. Bu, Chatter detection in milling based on singular spectrum analysis, Int. J. Adv. Manuf. Technol. 95 (9-12) (2018) 3475–3486, https://doi.org/10.1007/s00170-017-1366-y.
[13] M. Lamraoui, M. Thomas, M. El Badaoui, F. Girardin, Indicators for monitoring chatter in milling based on instantaneous angular speeds, Mech. Syst. Sig. Process. 44 (1-2) (2014) 72–85, https://doi.org/10.1016/j.ymssp.2013.05.002.
[14] P. Huang, J. Li, J. Sun, J. Zhou, Vibration analysis in milling titanium alloy based on signal processing of cutting force, Int. J. Adv. Manuf. Technol. 64 (5-8) (2013) 613–621, https://doi.org/10.1007/s00170-012-4039-x.
[15] C. Peng, L. Wang, T.W. Liao, 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 (2015) 118–131, https://doi.org/10.1016/j.jsv.2015.06.011.
[16] M. Luo, H. Luo, D. Axinte, D. Liu, J. Mei, Z. Liao, A wireless instrumented milling cutter system with embedded PVDF sensors, Mech. Syst. Sig. Process. 110 (2018) 556–568, https://doi.org/10.1016/j.ymssp.2018.03.040.
[17] W. Sun, M. Luo, D. Zhang, Machining vibration monitoring based on dynamic clamping force measuring in thin-walled components milling, Int. J. Adv. Manuf. Technol. 107 (5-6) (2020) 2211–2226, https://doi.org/10.1007/s00170-020-05153-9.
[18] T.L. Schmitz, K. Medicus, B. Dutterer, Exploring once-per-revolution audio signal variance as a chatter indicator, Mach. Sci. Technol. 6 (2) (2002) 215–233, https://doi.org/10.1081/MST-120005957.
[19] Y. Kakinuma, Y. Sudo, T. Aoyama, Detection of chatter vibration in end milling applying disturbance observer, CIRP Ann. 60 (1) (2011) 109–112, https://doi.org/10.1016/j.cirp.2011.03.080.
[20] N.-C. Tsai, D.-C. Chen, R.-M. Lee, Chatter prevention for milling process by acoustic signal feedback, Int. J. Adv. Manuf. Technol. 47 (9-12) (2010) 1013–1021, https://doi.org/10.1007/s00170-009-2245-y.
[21] E. Kuljanic, M. Sortino, G. Totis, Multisensor approaches for chatter detection in milling, J. Sound Vib. 312 (4-5) (2008) 672–693, https://doi.org/10.1016/j.jsv.2007.11.006.
[22] J. Pan, Z. Liu, X. Wang, C. Chen, X. Pan, Boring chatter identification by multi-sensor feature fusion and manifold learning, Int. J. Adv. Manuf. Technol. 109 (3-4) (2020) 1137–1151, https://doi.org/10.1007/s00170-020-05611-4.
[23] T.L. Schmitz, Chatter recognition by a statistical evaluation of the synchronously sampled audio signal, J. Sound Vib. 262 (3) (2003) 721–730, https://doi.org/10.1016/S0022-460X(03)00119-6.

[24] J. Ye, P. Feng, C. Xu, Y. Ma, S. Huang, A novel approach for chatter online monitoring using coefficient of variation in machining process, Int. J. Adv. Manuf. Technol. 96 (1-4) (2018) 287–297, https://doi.org/10.1007/s00170-017-1544-y.
[25] S. Wan, X. Li, W. Chen, J. Hong, Investigation on milling chatter identification at early stage with variance ratio and Hilbert–Huang transform, Int. J. Adv. Manuf. Technol. 95 (9-12) (2018) 3563–3573, https://doi.org/10.1007/s00170-017-1410-y.
[26] E. Kondo, H. Ota, T. Kawai, A new method to detect regenerative chatter using spectral analysis, part 1: Basic study on criteria for detection of chatter, J. Manuf. Ence Eng. 119 (1997) 461–466.
[27] T. Thaler, P. Potočnik, I. Bric, E. Govekar, Chatter detection in band sawing based on discriminant analysis of sound features, Appl. Acoust. 77 (2014) 114–121, https://doi.org/10.1016/j.apacoust.2012.12.004.
[28] S. Tangjitsitcharoen, In-process monitoring and detection of chip formation and chatter for CNC turning, J. Mater. Process. Technol. 209 (10) (2009) 4682–4688, https://doi.org/10.1016/j.jmatprotec.2008.10.054.
[29] C. Liu, L. Zhu, C. Ni, The chatter identification in end milling based on combining EMD and WPD, Int. J. Adv. Manuf. Technol. 91 (9-12) (2017) 3339–3348, https://doi.org/10.1007/s00170-017-0024-8.
[30] Z. Yao, D. Mei, Z. Chen, On-line chatter detection and identification based on wavelet and support vector machine, J. Mater. Process. Technol. 210 (5) (2010) 713–719, https://doi.org/10.1016/j.jmatprotec.2009.11.007.
[31] Y. Ji, X. Wang, Z. Liu, Z. Yan, L.i. Jiao, D. Wang, J. Wang, EEMD-based online milling chatter detection by fractal dimension and power spectral entropy, Int. J. Adv. Manuf. Technol. 92 (1-4) (2017) 1185–1200, https://doi.org/10.1007/s00170-017-0183-7.
[32] Y. Ji, X. Wang, Z. Liu, H. Wang, L. Jiao, D. Wang, S. Leng, Early milling chatter identification by improved empirical mode decomposition and multi-indicator synthetic evaluation, J. Sound Vib. 433 (2018) 138–159. https://doi.org/10.1016/j.jsv.2018.07.019.
[33] Y. Fu, Y. Zhang, H. Zhou, D. Li, H. Liu, H. Qiao, X. Wang, Timely online chatter detection in end milling process, Mech. Syst. Sig. Process. 75 (2016) 668–688, https://doi.org/10.1016/j.ymssp.2016.01.003.
[34] C. Liu, L. Zhu, C. Ni, Chatter detection in milling process based on VMD and energy entropy, Mech. Syst. Signal Process. 105 (2018) 169–182. https://doi.org/10.1016/j.ymssp.2017.11.046.
[35] Y.S. Tarng, T.C. Li, M.C. Chen, On-line drilling chatter recognition and avoidance using an ART2 – A neural network, Int. J. Mach. Tools Manuf. 34 (7) (1994) 949–957, https://doi.org/10.1016/0890-6955(94)90027-2.
[36] C.L. Zhang, X. Yue, Y.T. Jiang, W. Zheng, A Hybrid Approach of ANN and HMM for Cutting Chatter Monitoring, Adv. Mater. Res. 97–101 (2010) 3225–3232.
[37] J.-S. Kwak, M.-K. Ha, Intelligent diagnostic technique of machining state for grinding, Int. J. Adv. Manuf. Technol. 23 (5-6) (2004) 436–443, https://doi.org/10.1007/s00170-003-1899-0.
[38] Q. Shao, C.J. Feng, Pattern Recognition of Chatter Gestation Based on Hybrid PCA-SVM, Appl. Mech. Mater. 120 (2012).
[39] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors, Nature 323 (6088) (1986) 533–536, https://doi.org/10.1038/323533a0.
[40] Y. Bengio, P. Lamblin, D. Popovici, Larochelle. Greedy layer-wise training of deep networks[C]//Advances in neural information processing systems. (2007): 153-160.
[41] V.D.M. Laurens, G. Hinton, Visualizing data using t-SNE, J. Mach. Learn. Res. 9 (2008) 2579–2605.
[42] W.B.S. Yang, Support vector machine in machine condition monitoring and fault diagnosis, Mech. Syst. Signal Process. (2007).
[43] Y. Liu, X. Yao, Ensemble learning via negative correlation, Neural Networks 12 (10) (1999) 1399–1404, https://doi.org/10.1016/S0893-6080(99)00073-8.
[44] Y. Freund, R.E. Schapire, A decision-theoretic generalization of on-line learning and an application to boosting, J. Comput. Syst. Sci. 55 (1) (1997) 119–139, https://doi.org/10.1006/jcss.1997.1504.