Journal of Intelligent Manufacturing (2022) 33:1483–1502
https://doi.org/10.1007/s10845-021-01736-9

A kMap optimized VMD-SVM model for milling chatter detection with an industrial robot

Yu Wang[1] · Mingkai Zhang[1] · Xiaowei Tang[1] · Fangyu Peng[1] · Rong Yan[1]

Received: 2 March 2020 / Accepted: 2 January 2021 / Published online: 29 January 2021

Abstract

Industrial robots play an important role in the milling of large complex parts. However, the robot is less rigid and prone to vibration-related problems; chatter, which affects machining quality and efficiency, is more complex and difficult to monitor. In this paper, a variational mode decomposition-support vector machine (VMD-SVM) model based on information entropy (IE) is built to detect chatter in robotic milling. Significantly, the vibration signals are classified into four states for the first time: stable, transition, regular chatter, and irregular chatter. To improve the accuracy of the identification model based on VMD-SVM, a novel hyper-parameter optimization strategy—the kMap method—is proposed in this paper for optimizing three-dimensional hyper-parameters in the VMD-SVM model. The hyper-parameters of VMD-SVM are jointly optimized by the kMap method, with constant step sizes. As an improved grid search (GS), kMap reduces the operation time to the same order of magnitude as the heuristic algorithm (HA) [comprising particle swarm optimization (PSO) and genetic algorithm (GA)]. The VMD-SVM model with the hyper-parameters optimized by kMap exhibits higher accuracy and better stability than the hyper-parameters optimized by PSO and GA. The results of the validation experiments show that the kMap-optimized identification model is effective in industrial robotic milling.

Keywords VMD-SVM · Industrial robot · Chatter identification · Hyper-parameter optimization

Introduction

Industrial robots are used for milling large and complex parts owing to their advantages of low cost, wide workspace range, and high flexibility. However, compared with CNC machine tools, the low rigidity of the robot makes chatter more likely, influencing machining accuracy, quality, and efficiency. Extensive research has been conducted on the prediction of chatter and various chatter suppression methods have been summarized. Generally, chatter in robotic milling is composed of both regenerative and mode coupling chatter. However, differing from the typical issues concerning regenerative chatter in conventional CNC machining, mode coupling chatter was identified as the dominant source of vibrations in robotic machining at low cutting speeds and regenerative chatter was the dominant source at high speeds (Pan et al. 2006; Gienke et al. 2019). The stiffness matrix depends on the current configuration in terms of the robot. The damping effect will always increase the stability of the system but is difficult to accurately identify. In addition, the main chatter mechanism of robotic milling is different for different machining cases; thus, chatter analysis cannot yield a confident result using their stability criteria. Except for chatter prediction, detecting the vibration state of milling timely is an efficient method to improve the performance of machining equipment, which reduces the frequency and time of chatter occurrence and contributes to further research on the chatter mechanism. Therefore, it is essential to study chatter identification in robotic milling.

Traditional chatter identification is mainly achieved by observing the processed surface and analysis of the physical signal spectrum. Recently, some scholars have identified chatter through feature extraction and setting thresholds for machined surface images (Lei and Soshi 2017), cutting force signals (Tangjitsitcharoen et al. 2015), vibration signals (Tao et al. 2019; Musselman et al. 2019) and current signals (Aslan and Altintas 2018). The value of the characteristic threshold

✉ Xiaowei Tang
txwysxf@126.com

[1] National NC System Engineering Research Center, School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, China

is important and significantly influences the accuracy of the identification model. By introducing a machine learning-based chatter identification model, with physical vibration signals as the input, the correspondence between the input and output requires establishing. Generally, the chatter identification process is realized by feature extraction and classification. Deep learning such as deep belief networks (DBN) is used to extract feature automatically and classify chatter simultaneously (Fu et al. 2019b). However, the DBN needs more time to complete the model training. In addition, machine learning is more suitable for timely chatter monitoring than deep learning because of its small computation time. Feature selection and extraction of signals in machine learning have a significant impact on model accuracy. In addition, external disturbances, modeling errors, and uncertainties are common when measuring in practical applications (Stojanovic and Prsic 2020; Nannapaneni et al. 2020). These noises and uncertainties affect the performance of feature extraction and classification. Generally, non-Gaussian noise is eliminated by designing appropriate filters. When chatter occurs during machining, the energy is concentrated near the mode frequency of the machining system, and chatter frequency bands will be observed, while the frequency band centers are not fixed. Signal analysis methods based on modal decomposition, such as VMD and empirical mode decomposition (EMD) (Zhao et al. 2020), and empirical wavelet transformation (EWT) are usually used to extract features of the signals with these characteristics. VMD can accurately separate the harmonic components of non-stationary signals, regardless of how close their frequency components are. Compared with VMD, EMD does not have a strong mathematical foundation. In addition, VMD can be used to reduce the impact of non-Gaussian impulsive noise (Dutta et al. 2016). Aneesh et al. (2015) compared the performance of VMD (Dragomiretskiy and Zosso 2013) with EWT in feature extraction. The classification results of SVM show that VMD feature extraction performs better than EWT. VMD, an adaptive signal machining method, is conducive to the extraction of chatter features, effectively dealing with the characteristic that chatter frequency bands shift during machining. Liu et al. (2017) used the chatter of milling on an NC machine as the analysis object, extracted sample features by VMD and Shannon power spectral entropy, and classified and predicted samples using a probabilistic neural network (PNN). However, the parameters that affect the accuracy of the model are selected based on prior knowledge, without further optimization processes. After the data features are extracted, SVM/support vector classification (SVC) is usually used in chatter detection and prediction. Chen et al. (2020) used a multivariate filter method to select the p-leader multifractal features and adopted SVM for chatter classification. The stability lobe diagram of milling was predicted using extended SVC and an artificial neural network (ANN) (Friedrich et al. 2017), and these continuously learning algorithms can be applied to higher-dimensional problems with arbitrary input dimensions. The dominant frequency bands are identified by localizing the frequency bands at which the energy is high in the average fast Fourier transform (FFT) plot to highlight the chatter-related characteristics. Furthermore, the combination of VMD and SVM (Abdoos et al. 2016) facilitated mechanical fault diagnosis. However, the VMD-SVM model is rarely used in chatter detection. The optimization of hyper-parameters can improve model accuracy. Mutation sine and cosine algorithm-particle swarm optimization algorithm (SCA-PSO) (Fu et al. 2019a), chaos sine and cosine algorithm (CSCA) (Fu et al. 2018), quantum chaos fly optimization algorithm (QCFOA) (Xu et al. 2019) and other optimization methods have been used to optimize the hyper-parameters of the VMD-SVM model, where VMD and SVM are optimized respectively.

The VMD-SVM model in the above literature is mainly used in the classification of power quality events, fault diagnosis of variable load-bearing, fault diagnosis for rolling bearings, vibration trend measurement for a hydropower generator, etc. Moreover, there are limited studies on the chatter identification method of robotic milling based on the VMD-SVM model. This study constructs a VMD-SVM model to identify robotic milling chatter. Generally, the hyper-parametric optimization of VMD and SVM is conducted separately, which is insufficient to obtain the optimal performance of the VMD-SVM model. For better performance of the chatter identification model in robot milling based on VMD-SVM, the kMap method is proposed in this paper to jointly optimize the hyper-parameters in the identification model. The remainder of this paper is organized as follows. In section “Dataset construction based on novel chatter classification”, the experiment settings and data augmentation methods for constructing the dataset are presented. An identification model based on VMD-SVM is constructed and the details of the proposed kMap optimization method are described in section "Chatter identification model based on VMD-SVM". In section "Identification accuracy and experiment analysis", the identification models based on raw cutting data and the data containing non-Gaussian noise optimized by different methods are analyzed. The effectiveness of the proposed approach was validated through 12 robot milling experiments. Finally, the conclusions are presented in section "Conclusions".

Dataset construction based on novel chatter classification

The full-discretization method (Tang et al. 2017) was used to obtain the stability lobe diagram in this paper. The ranges of spindle speeds and cutting depths were selected for the cutting experiment according to the case and experience. The vibration data of the experiments were collected to build vibration sample sets under different vibration conditions. Then, the chatter characteristics in robotic milling were analyzed and its identification model was trained.

Experimental platform and parameter setting

The planar milling experiments were conducted on the robotic milling platform with ABB IRB6660, as shown in Fig. 1. The vibration signals of the spindle and the vibration texture of the machined surface in the milling process were recorded. A SANDVIK Φ 25 face milling cutter (600-025A25-10H) with 3 10 mm blade-diameter cutter teeth (600-1045M-ML 1030) was used to cut the Ni–Al bronze workpiece. Vibration signals were collected via an NI data acquisition system and a 3-dimensional acceleration sensor (DYTRAN 3263A2T). Simultaneously, an industrial vision measurement system (Camera: Basler aca2440-20gc, lens: OPT c1614-5m) was used to take pictures of the surface of the workpiece to record the vibration texture of the surface.

According to the stability lobe diagram (Fig. 2) and actual machining experience, a rotation speed range of 3000–6000 rpm and a cutting depth range of 0.2–2 mm were selected for the cutting experiments. In the milling process, the feed per tooth and cutting width remained constant. The feed per tooth was fz = 0.05 mm/ft and the cutting width was 8 mm. The sampling frequency of the vibration data was 10 kHz. A total of 160 sets of different machining parameters were selected for cutting. Up milling was performed along the positive direction of the Y-axis of the robotic coordinate system. The machining parameters are shown in Table 1, represented visually by four types of differently shaped/colored symbols in the stability lobe diagram. Specifically, the red point and the green '×' express the stable and transition machining parameter states, respectively. The black '#' and the blue '*' both denote the chatter state; the black '#' signifies regular chatter and the blue '*' represents irregular chatter. In the stability lobe diagram, above the curve is unstable and below is stable. Stable data are in the stable region and chatter (regular chatter and irregular chatter) data are in the unstable region. In addition, the transition cutting parameter points are around the curve. The accuracy of prediction is not sufficiently high but the cutting data of the four vibrations are similar in size. The practical meanings of these four types used to distinguish different machining vibration states are expounded in section "Classification of vibration states in robotic milling".

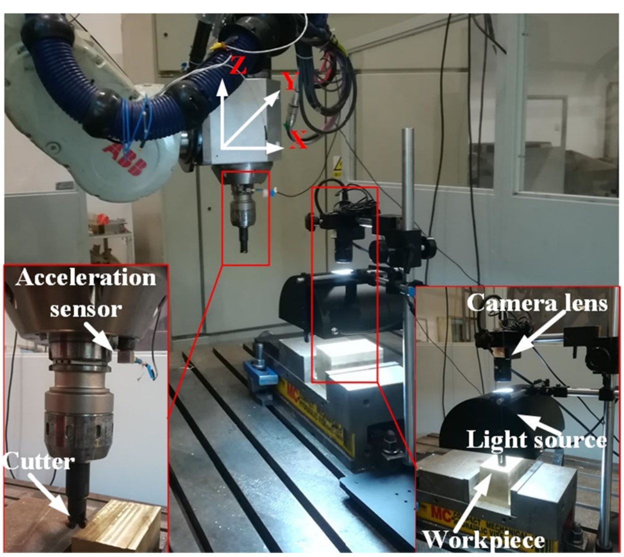


Fig. 1 Planar milling experiment setup on the robotic milling platform

Construction of vibration dataset based on the sliding window method

Generally, the original vibration data are collected in three stages: the feeding, cutting, and relieving stages, as shown in Fig. 3. The spectrum features of the three stages are not completely consistent (Fig. 4). To ensure the consistency of the sample characteristics in the sample splicing and sliding window operation, the samples in the feeding and relieving stages were not considered. By intercepting the raw vibration data of each section, the data of the cutting stage can be obtained as samples. Simultaneously, multiple cutting during the experiment will be conducted to collect sufficient vibration samples for analysis under each set of cutting parameters.

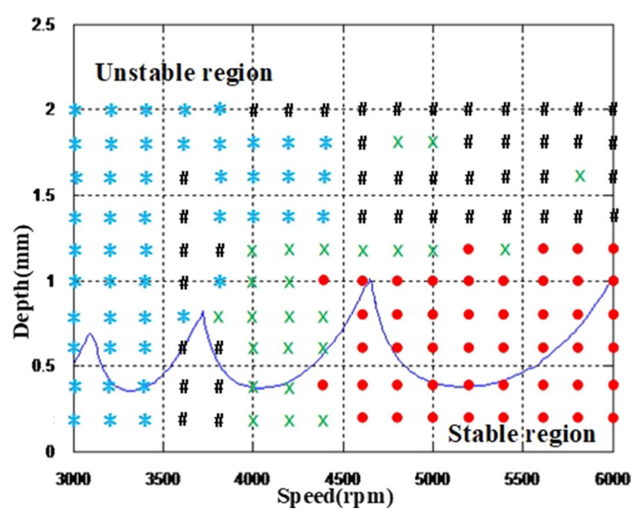


Fig. 2 Stability lobe diagram

Table 1 Milling parameters table

Speed n (rpm)	Feed f (mm/s)	Depth d (mm)									
		0.2	0.4	0.6	0.8	1	1.2	1.4	1.6	1.8	2
3000	7.5	1	17	33	49	65	81	97	113	129	145
3200	8	2	18	34	50	66	82	98	114	130	146
3400	8.5	3	19	35	51	67	83	99	115	131	147
3600	9	4	20	36	52	68	84	100	116	132	148
3800	9.5	5	21	37	53	69	85	101	117	133	149
4000	10	6	22	38	54	70	86	102	118	134	150
4200	10.5	7	23	39	55	71	87	103	119	135	151
4400	11	8	24	40	56	72	88	104	120	136	152
4600	11.5	9	25	41	57	73	89	105	121	137	153
4800	12	10	26	42	58	74	90	106	122	138	154
5000	12.5	11	27	43	59	75	91	107	123	139	155
5200	13	12	28	44	60	76	92	108	124	140	156
5400	13.5	13	29	45	61	77	93	109	125	141	157
5600	14	14	30	46	62	78	94	110	126	142	158
5800	14.5	15	31	47	63	79	95	111	127	143	159
6000	15	16	32	48	64	80	96	112	128	144	160

In this study, the machined plane size of the workpiece in the experiments is 70 × 120 mm and the single cutting range is 8 × 70 mm. In the cutting area, it can be seen from Fig. 5 that the features of the vibration signal under the same machining parameters maintain good consistency. Therefore, it is assumed that the robot maintains the same vibration state under the same processing parameters when the cutting area is small. In this way, for the same parameters, vibration data from the cutting stage of multiple data files can be directly collected to represent the vibration state. To facilitate training the identification model in the following chapters, sliding window sampling is conducted for the vibration signals under the same processing parameters to construct standard vibration data sample sets (Fig. 6), according to the following:

$$S_{ij} = S_i[1 + j * s\_len,\ fs + j * s\_len]\ ,\quad \left(j = 0, 1, 2, \ldots, INT(\frac{Ni - fs}{s\_len}) \right) \tag{1}$$

where S_{ij} is the jth vibration sample obtained from the vibration signal of the ith parameter group. j is the number of sliding windows, and fs is the sampling frequency. fs is uniformly set to 10 k. s_len represents the sliding step size of the window, and $INT()$ is the integer component of the value in parentheses. The process is shown in Fig. 7.

Classification of vibration states in robotic milling

Previous studies conducted FFT on vibration signals and analyzed their frequency components to determine whether chatter occurs. In addition to the tool passing frequency and its frequency multiplication, the data of the chatter state contains other obvious frequency components. However, it is found that the frequency distribution in the spectrum diagram of the data in the chatter state is not completely similar in the experimental process. The frequency of some chatter data is regular in the spectrum diagram, while the frequency of others is chaotic. When the time–frequency information of the vibration data is input into the image through a continuous wavelet transform (CWT), it can be found that the

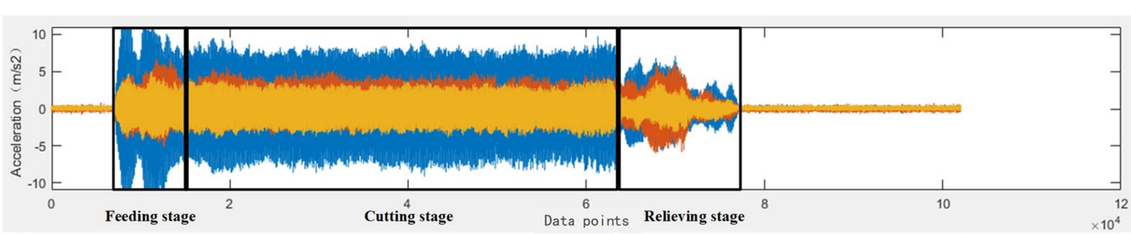


Fig. 3 Original milling vibration data

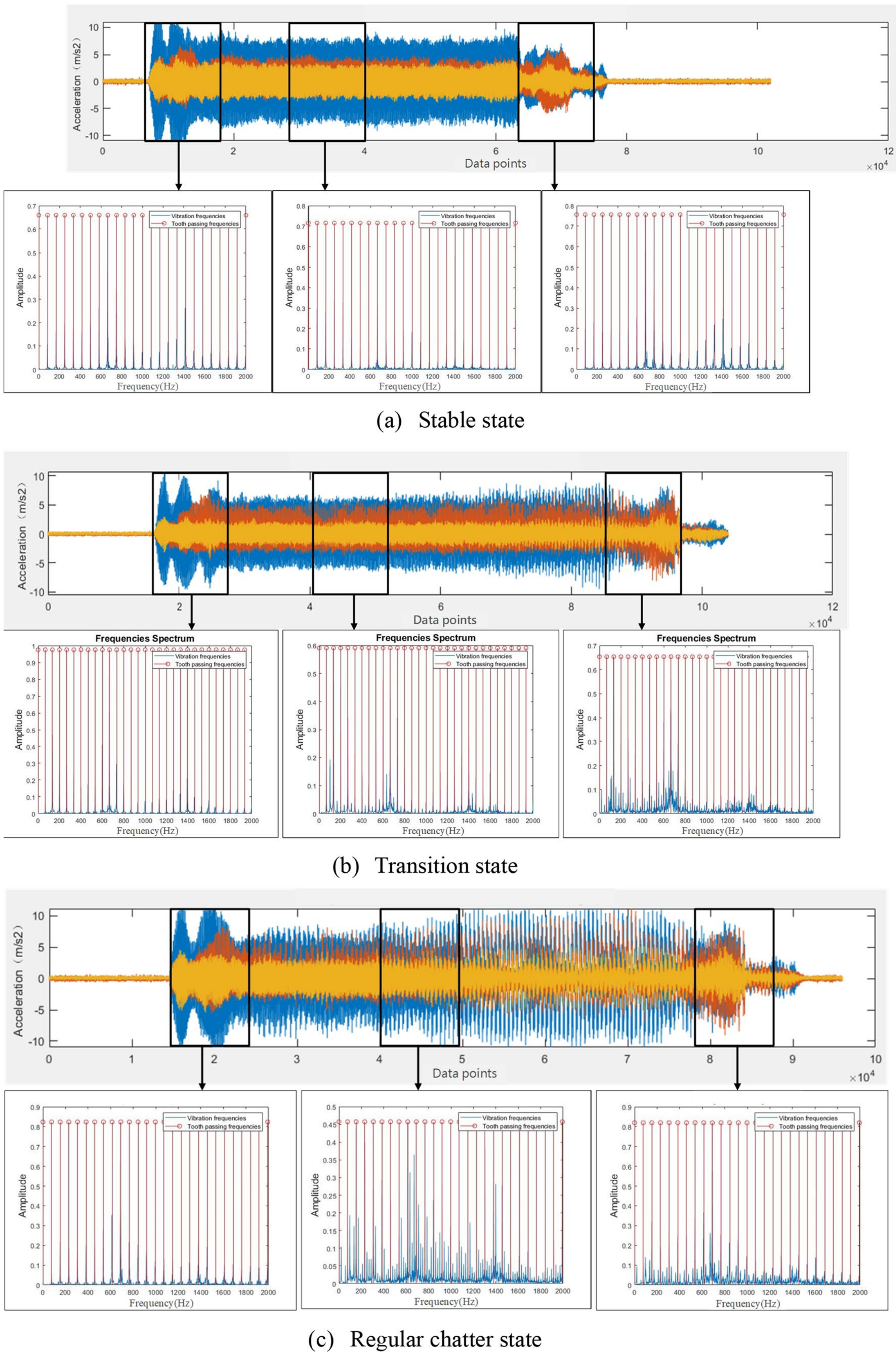
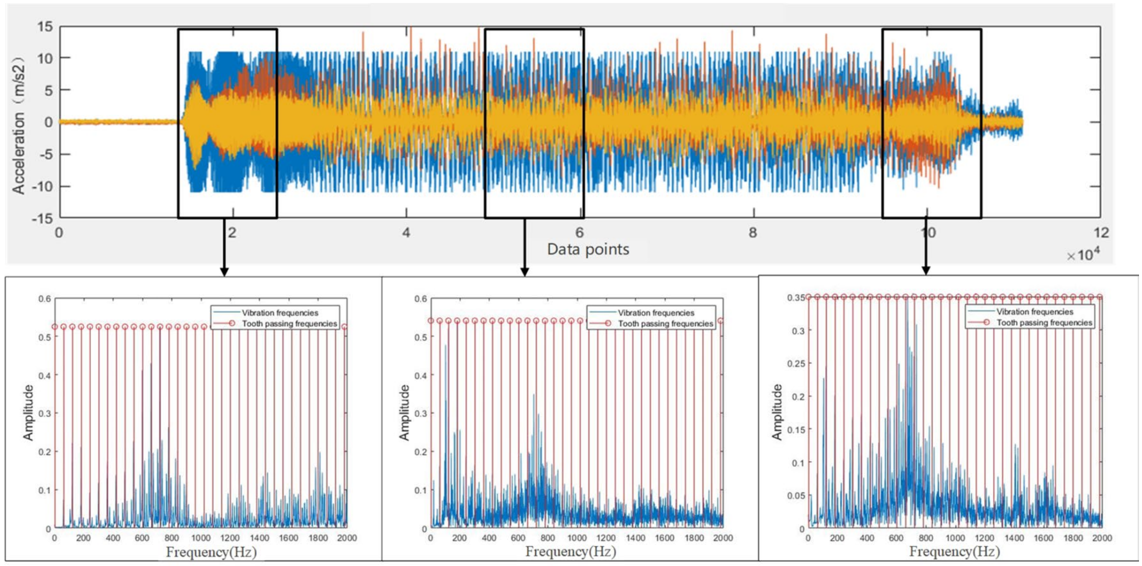


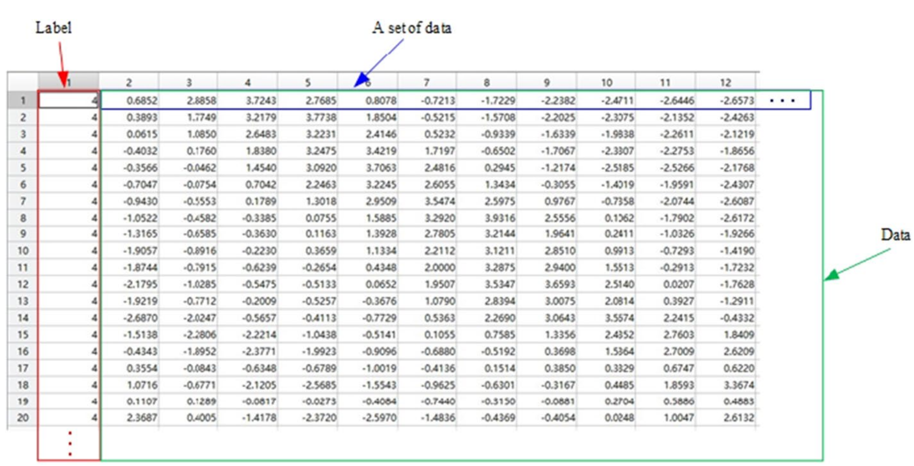
Fig. 4 FFT spectra of samples with different vibration states at different cutting stages



(d) Irregular chatter state

Fig. 4 (continued)

Fig. 5 The vibration data in the same machining parameters



	1	2	3	4	5	6	7	8	9	10	11	12
1	4	0.6852	2.8858	3.7243	2.7685	0.8078	-0.7213	-1.7229	-2.2382	-2.4711	-2.6446	-2.6573
2	4	0.3893	1.7749	3.2179	3.7738	1.8504	-0.5215	-1.5708	-2.2025	-2.3075	-2.1352	-2.4263
3	4	0.0615	1.0850	2.6483	3.2231	2.4146	0.5232	-0.9339	-1.6339	-1.9838	-2.2611	-2.1219
4	4	-0.4032	0.1760	1.8380	3.2475	3.4219	1.7197	-0.6502	-1.7067	-2.3307	-2.2753	-1.8656
5	4	-0.3566	-0.0462	1.4540	3.0920	3.7063	2.4816	0.2945	-1.2174	-2.5185	-2.5266	-2.1768
6	4	-0.7047	-0.0754	0.7042	2.2463	3.2245	2.6055	1.3434	-0.3055	-1.4019	-1.9591	-2.4307
7	4	-0.9430	-0.5553	0.1789	1.3018	2.9509	3.5474	2.5975	0.9767	-0.7358	-2.0744	-2.6087
8	4	-1.0522	-0.4582	-0.3385	0.0755	1.5885	3.2920	3.9316	2.5556	0.1062	-1.7902	-2.6172
9	4	-1.3165	-0.6585	-0.3630	0.1163	1.3928	2.7805	3.2144	1.9641	0.2411	-1.0326	-1.9266
10	4	-1.9057	-0.8916	-0.2230	0.3659	1.1334	2.2112	3.1211	2.8510	0.9913	-0.7293	-1.4190
11	4	-1.8744	-0.7915	-0.6239	-0.2654	0.4348	2.0000	3.2875	2.9400	1.5513	-0.2913	-1.7232
12	4	-2.1795	-1.0285	-0.5475	-0.5133	0.0652	1.9507	3.5347	3.6593	2.5140	0.0207	-1.7628
13	4	-1.9219	-0.7712	-0.2009	-0.5257	-0.3676	1.0790	2.8394	3.0075	2.0814	0.3927	-1.2911
14	4	-2.6870	-2.0247	-0.5657	-0.4113	-0.7729	0.5363	2.2690	3.0643	3.5574	2.2415	-0.4332
15	4	-1.5138	-2.2806	-2.2214	-1.0438	-0.5141	0.1055	0.7585	1.3356	2.4352	2.7603	1.8409
16	4	-0.4343	-1.8952	-2.3771	-1.9923	-0.9096	-0.6880	-0.5192	0.3698	1.5364	2.7009	2.6209
17	4	0.3554	-0.0843	-0.6348	-0.6789	-1.0019	-0.4136	0.1514	0.3850	0.3329	0.6747	0.6220
18	4	1.0716	-0.6771	-2.1205	-2.5685	-1.5543	-0.9625	-0.6301	-0.3167	0.4485	1.8593	3.3674
19	4	0.1107	0.1289	-0.0817	-0.0273	-0.4084	-0.7440	-0.3150	-0.0881	0.2704	0.5886	0.4883
20	4	2.3687	0.4005	-1.4178	-2.3720	-2.5970	-1.4836	-0.4369	-0.4054	0.0248	1.0047	2.6132

Fig. 6 Standard vibration data sample sets

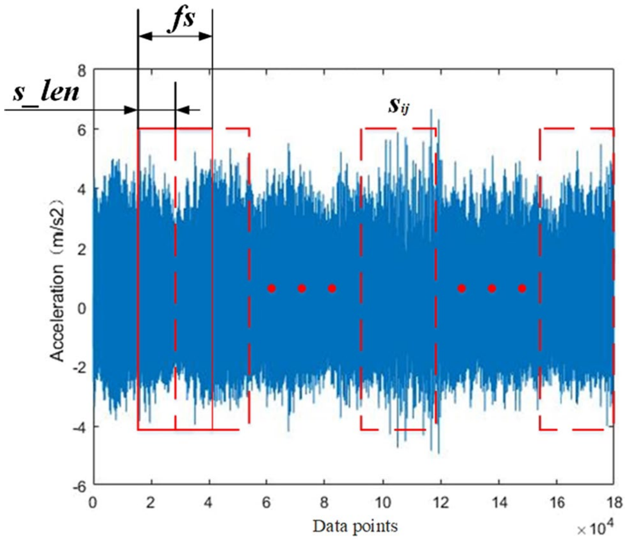


Fig. 7 Sliding window sampling process under the same processing parameters

frequency components of the data with regular frequency have little change with time, while the frequency components of the data with a disordered frequency vary with time. The vibration data in the chatter state is divided into regular chatter (time-invariant chatter) and irregular chatter (time-varying chatter). All vibration data can be divided into four types: stable, transition, regular chatter, and irregular chatter. FFT is conducted for the vibration signals of the different vibration types with their frequency components being analyzed. The results of the FFT are shown in Fig. 8 and the CWT results are shown in Fig. 9.

According to the wavelet time–frequency spectrum diagram, its frequency component is constant in the stable state and the frequency bands are continuous. In the transition state, most of the frequency components remain unchanged and some frequency bands are discontinuous. In the regular chatter state, there is a small number of time-varying frequency components with discontinuous frequency bands. In the irregular chatter state, most frequency components are time-varying and each frequency band is almost completely discontinuous. Therefore, the time-varying chatter components in robotic milling can be reflected and captured by wavelet analysis.

Obvious differences can be found by observing the vibration patterns of the machined surface (Fig. 10) under four states:

1. Stable state
   There are sometimes normal visible gear marks on the surface but no obvious vibration marks. When directly observed by the naked eye, the surface is very smooth and clean, as shown in Fig. 10a.
2. Transition state
   Compared with the stable state, there is a slight vibration texture on the surface, as shown in Fig. 10b.
3. Regular chatter state
   Compared with the stable state, there is an obvious regular vibration texture on the surface, as shown in Fig. 10c.
4. Irregular chatter state

Compared with the previous states, the surface has obvious irregular vibration textures and the surface quality is the worst, as shown in Fig. 10d.

According to the above standards, vibration textures generated using all the processing parameters are observed. Among them, the stable state and transitional state accounted for 28.75 and 15%, respectively. The percentages of regular chatter and irregular chatter states are 26.875 and 29.375%, respectively.

Chatter identification model based on VMD-SVM

The entire framework of the identification model is shown in Fig. 11. After obtaining the vibration signal datasets, VMD and IE are used in the preprocessing of raw signals. VMD is used to decompose the original signal in the frequency domain. The chaotic characteristics of the vibration signal are further extracted and the dimension of the data is reduced by computing IE. Then, an SVM-based classification model is trained to divide the vibration signals into four types. Finally, the hyper-parameters of the VMD-SVM model are optimized by kMap to improve the accuracy of the identification model.

VMD-SVM model based on IE

In view of higher decomposition accuracy, VMD is used to process signals adaptively, which extracts features of vibration better by considering the characteristic that vibration frequency bands are variable. The solution target of the VMD is shown below:

$$\begin{aligned} &\min_{\{u_k(t)\},\{\omega_{k(t)}\}} \sum_k l_k \\ &s.t. \quad \sum_k u_k(t) = f(t) \end{aligned} \tag{2}$$

where $\{u_k(t)\} = \{u_1(t), u_2(t), \ldots, u_k(t)\}$ represents the K intrinsic mode function (IMF) components to be preset and estimated. $\{\omega_k(t)\} = \{\omega_1(t), \omega_2(t), \ldots, \omega_k(t)\}$ is the relevant

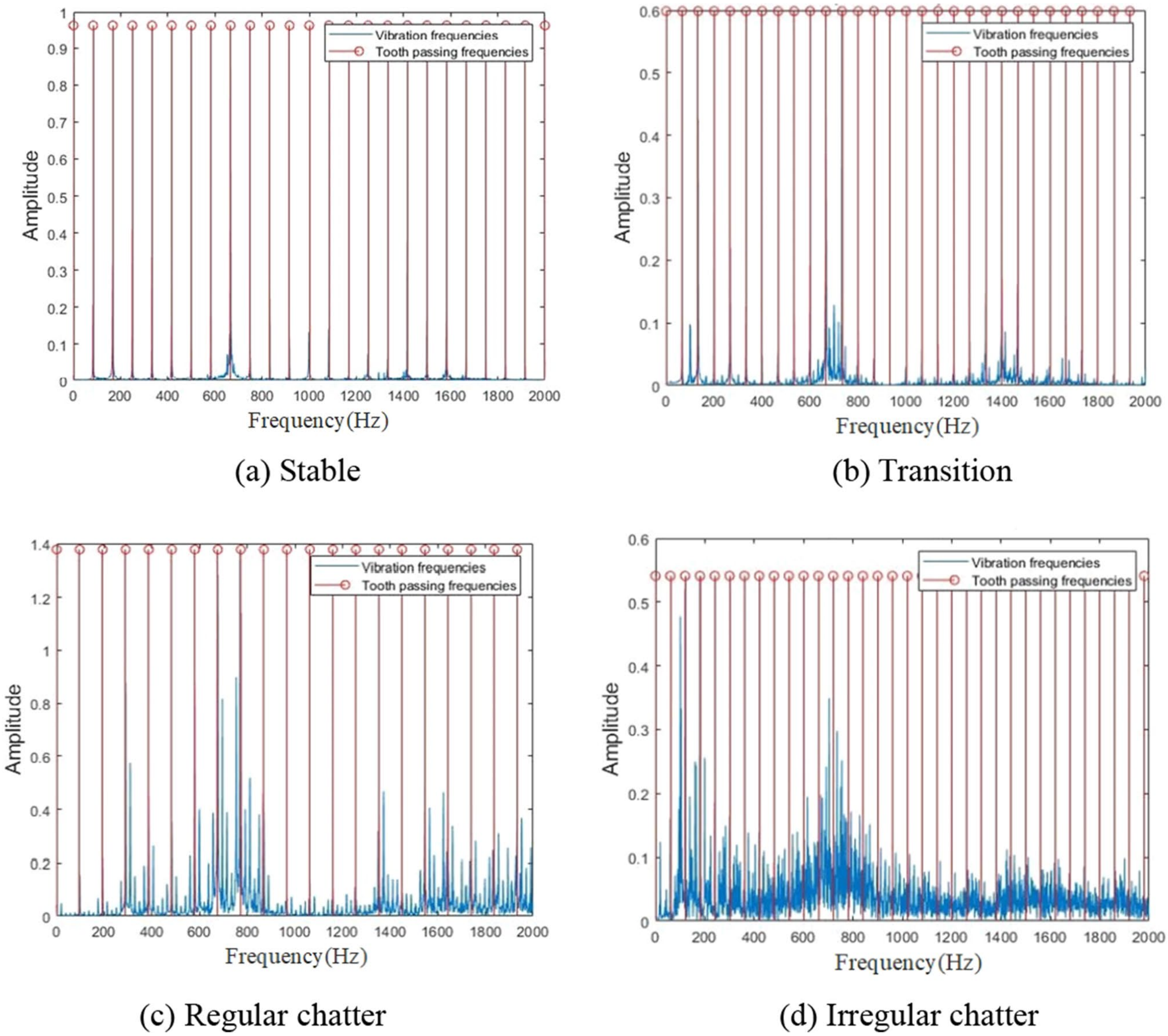


Fig. 8 Frequency spectra of 4 vibration states based on FFT

center frequency and $f(t)$ is the raw signal. u_k, ω_k, and f are simplified to represent $u_k(t)$, $\omega_k(t)$, and $f(t)$.

The Lagrangian multiplier method (LMM) is utilized to solve the optimization problem in Eq. (3). The secondary penalty factor α and Lagrangian multiplier $\lambda(t)$ are introduced to obtain the Lagrangian equation and solution, as follows:

$$L(\{u_k\},\{\omega_k\},\lambda) = \alpha \sum_k l_k + \left\| f - \sum_k u_k \right\|_2^2 + \left\langle \lambda(t), f - \sum_k u_k \right\rangle \tag{3}$$

Then, the alternate direction method of multipliers (ADMM) is used to solve Eq. (3), and the extreme point is found by alternately updating $u_k^{n+1}, \omega_k^{n+1}$, and λ^{n+1}, obtaining k signal sub-sequences (Liu et al. 2017).

The IE of the sub-signals can be obtained as the characteristic variable according to the characteristics that the frequency components of vibration signals become increasingly complex when chatter occurs. The calculation equation for IE is:

$$E_i = -\mathbf{x}_i \log \mathbf{x}_i, \quad i = 1, 2, \ldots, n \tag{4}$$

where E_i is the IE of the ith signal subsequence, and $\boldsymbol{x}_i$ represents the signal subsequence extracted with VMD. The feature vector obtained from IE is taken as the input and the SVM classification model based on the radial basis function (RBF) is used to realize identification of the vibration state

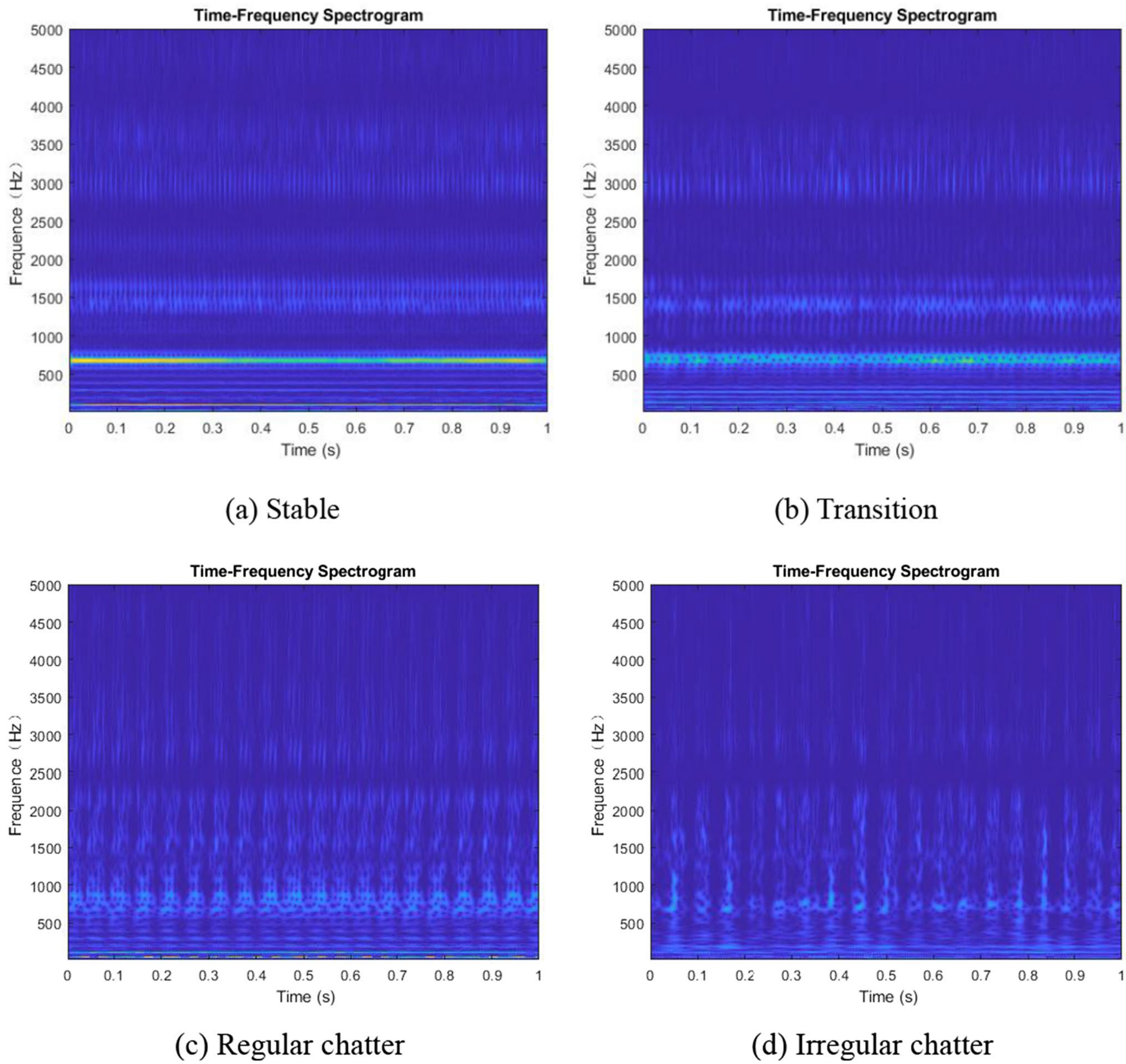


Fig. 9 Time–frequency spectra of 4 vibration states based on CWT

in the process of robotic milling (the vibration signal collected in this paper is the signal sequence of sampling points with a sampling frequency of 10 kHz).

kMap identification model optimization method

Generally, the SVM algorithm is used for classification, which involves the hyper-parameters σ and C. The optimization of hyper-parameters is a two-dimensional hyper-parameter optimization problem. With the use of the VMD algorithm, the modal decomposition number K and the second penalty factor α in VMD will also become hyper-parameters of the model. The optimization of hyper-parameters becomes a 4-dimensional hyper-parameter optimization problem. According to the influence of K and α on the reconstruction characteristics of simulation signals, the value of α should be greater than or equal to half of the sampling frequency (Lv et al. 2016). The GS is used to optimize multi-dimensional hyper-parameters but its operation time is long. This paper sets $\alpha = 5000$ and proposes a GS-based method, the kMap method, to solve the VMD-SVM three-dimensional hyper-parameter optimization problem.

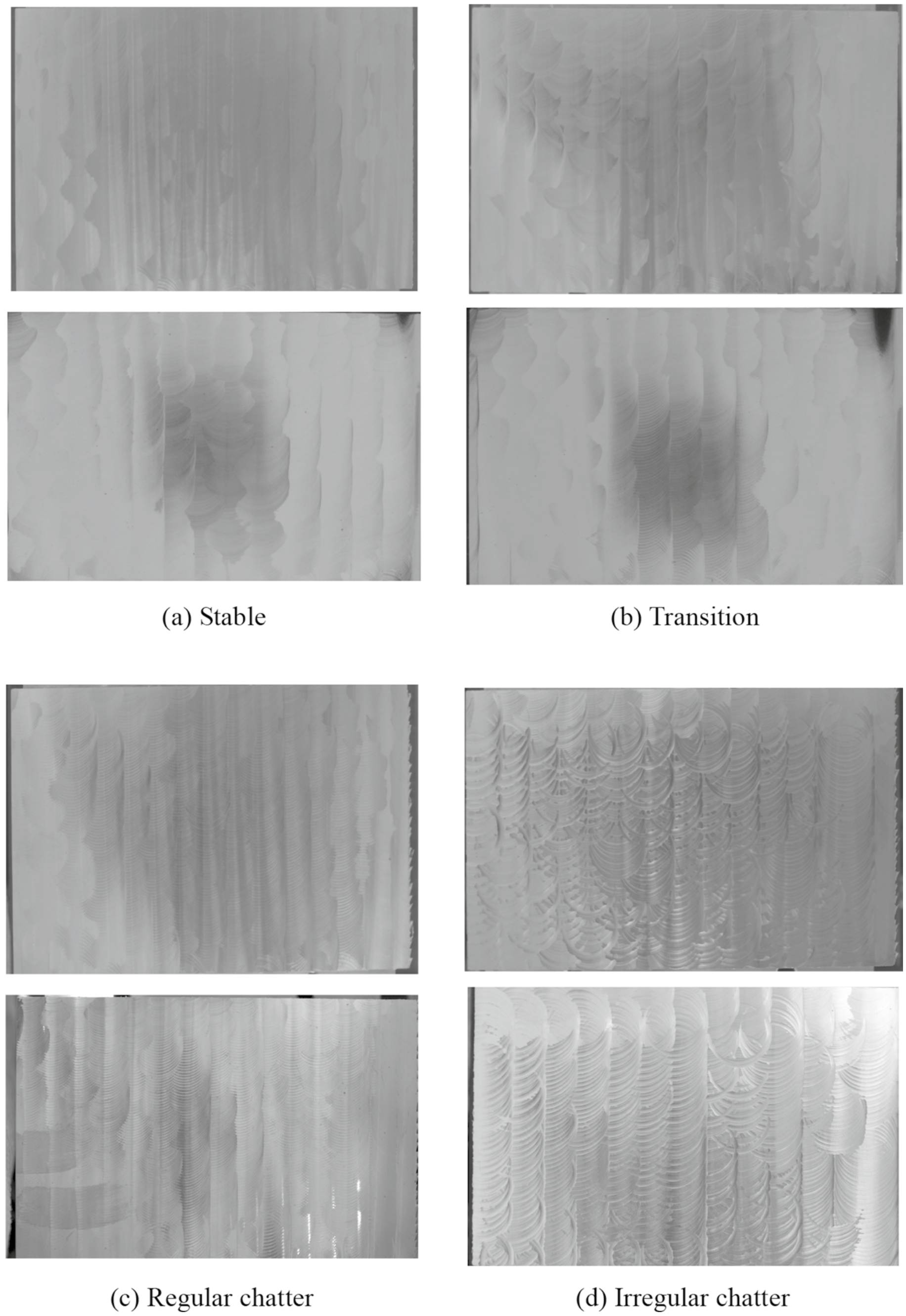


Fig. 10 Vibration texture of workpiece surface under 4 vibration states

First, the range of hyper-parameters and step length are set:

$$\begin{cases} C \in [2^{C\_\min}, 2^{C\_\max}], \quad C\_n = INT((C\_\max - C\_\min)/C\_step) \\ \sigma \in [2^{\sigma\_\min}, 2^{\sigma\_\max}], \quad \sigma\_n = INT((\sigma\_\max - \sigma\_\min)/\sigma\_step) \\ K \in \{1, 2, ..., n\} \end{cases} \tag{5}$$

where the value of C and σ are discretized with 2 as the basis number. *C_step* and *σ_step* represent the discrete step lengths of the corresponding hyper-parameters. K is the number of modal decompositions, so its value is generally a discrete integer. Within this range, the optimal hyper-parameter combination, (C, σ, K), is determined to ensure that the VMD-SVM model achieves higher training accuracy.

Then, a group value range of hyper-parameters is set with a small degree of dispersion. The influence trend of these 3 hyper-parameters on the model precision is first analyzed using the GS. The accuracy distributions in different parameters obtained by setting $C\_\min = \sigma\_\min = -5$, $C\_\max = \sigma\_\max = 5$, $C\_step = \sigma\_step = 0.2$, and $K \in \{2, 3...10\}$, as shown in Fig. 12. Where the surfaces of different colors

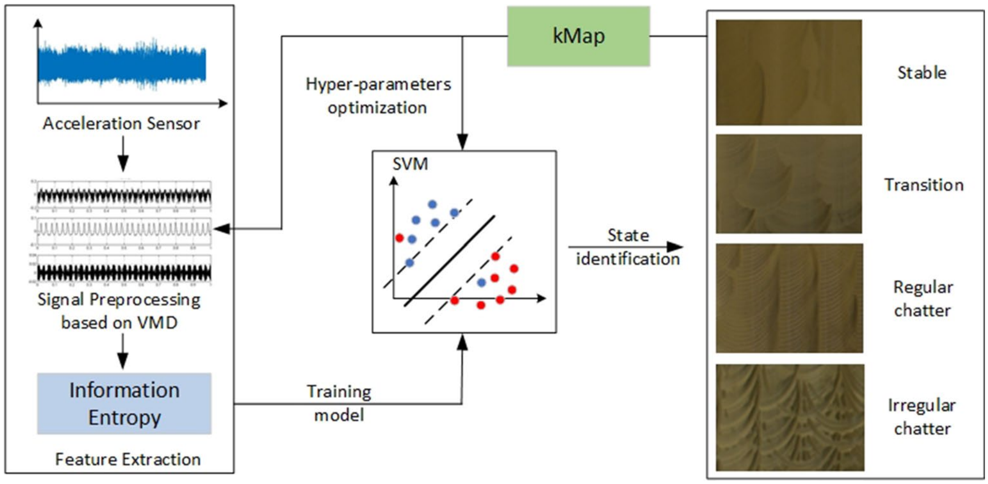


Fig. 11 Chatter identification model framework for robotic milling

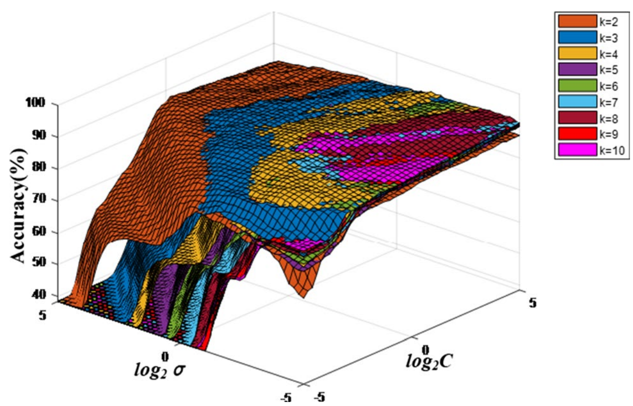
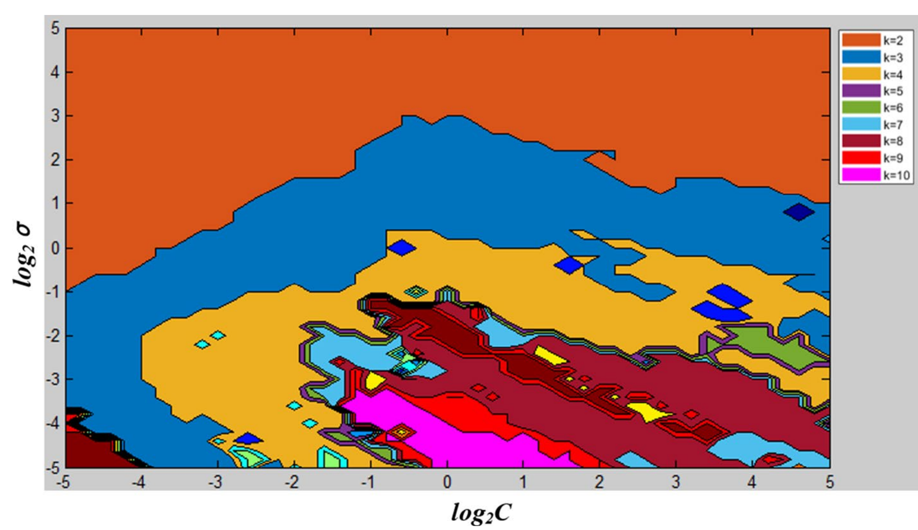


Fig. 12 Training accuracy distribution under different parameter combinations

represent the model training accuracy under combinations of C and σ when K is set as different values.

It can be seen from Fig. 12 that the influence of the 3 hyper-parameters on the model precision studied in this paper is not a linear relationship. Especially, when the model precision reaches a higher value range, the influence relationship becomes more complex. Convert Fig. 12 to a plane perspective of $C-\sigma$, as shown in Fig. 13. The differently colored regions in the figure correspond to different values of C and σ, representing the value of K that can maintain the highest accuracy of the model in the current region. According to the figure, the optimal value of K is not fixed. However, it is not difficult to find from the figure that these colored areas are divided into continuous blocks in most cases. Therefore, as long as the value of K in each $C-\sigma$ region can be determined to ensure that the model has the highest accuracy, the 3-dimensional

Fig. 13 Training accuracy distribution diagram on the $C-\sigma$ plane



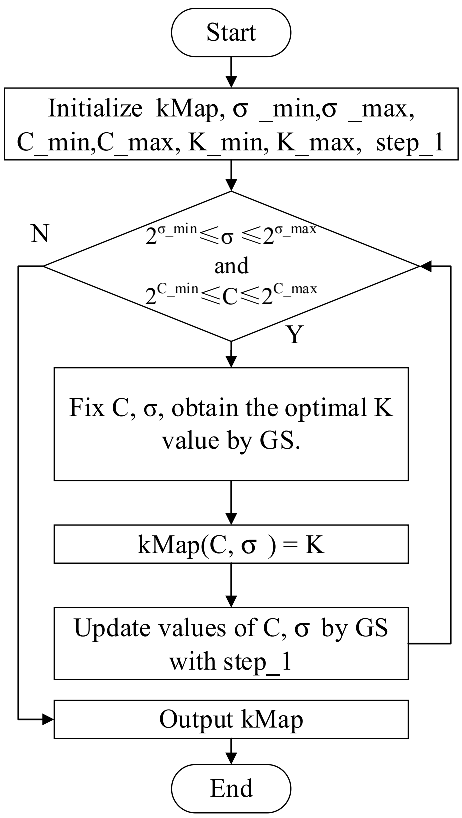


Fig. 14 First step: obtain the K-contour map (kMap) under the large step

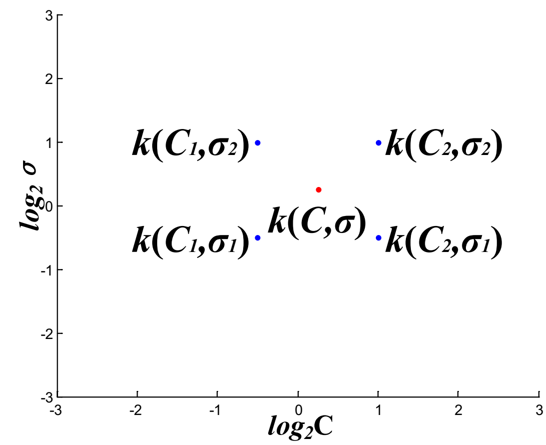


Fig. 16 Bilinear Interpolation

hyper-parameter optimization problem can be transformed into the $C-\sigma$ two-dimensional optimization problem.

The continuous region can be identified by determining the boundaries of different continuous regions. Specifically, a large step size is given and the values of C and are σ updated by GS. The optimal value of K under the current values of C is σ, obtained as follows:

$$\begin{aligned} & k(C,\sigma) = \underset{K\in\{1,2,3,\ldots,n\}}{\arg\max}\ acc(C,\sigma,K) \\ & s.t.\ k(C,\sigma) = i-1,\left(\left|acc_i - acc_{i-1}\right| \le \varepsilon\right) \end{aligned} \tag{6}$$

where *acc* represents the training accuracy of the model. At the same time, the higher the value of K, the higher the computational complexity of the VMD algorithm. Therefore, with the constraint condition set, a smaller value of K is selected as the optimal K when the difference between the model accuracies corresponding to the current K and the previous K is less than the convergence accuracy. The

Fig. 15 Optimal K distribution on the $C-\sigma$ plane

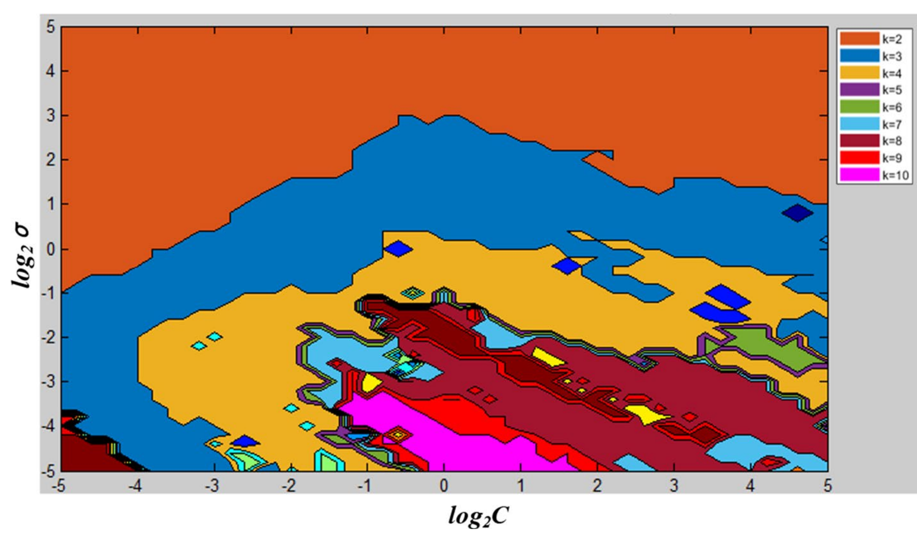
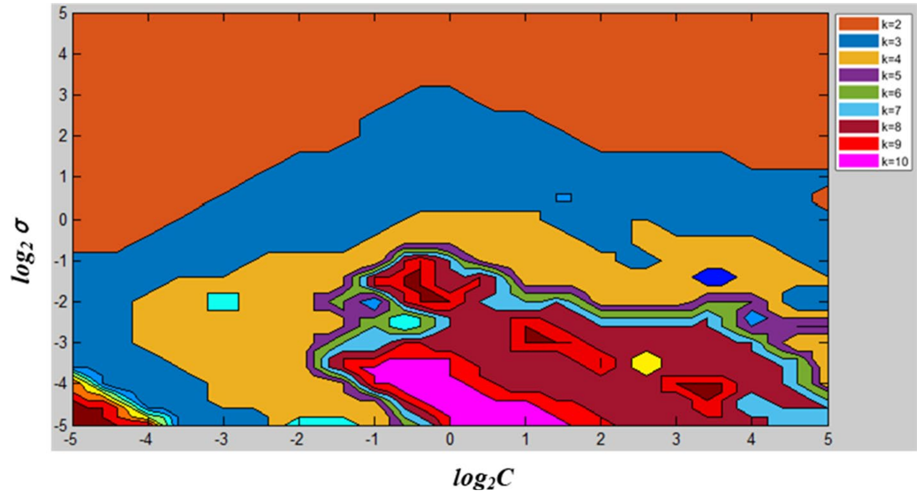
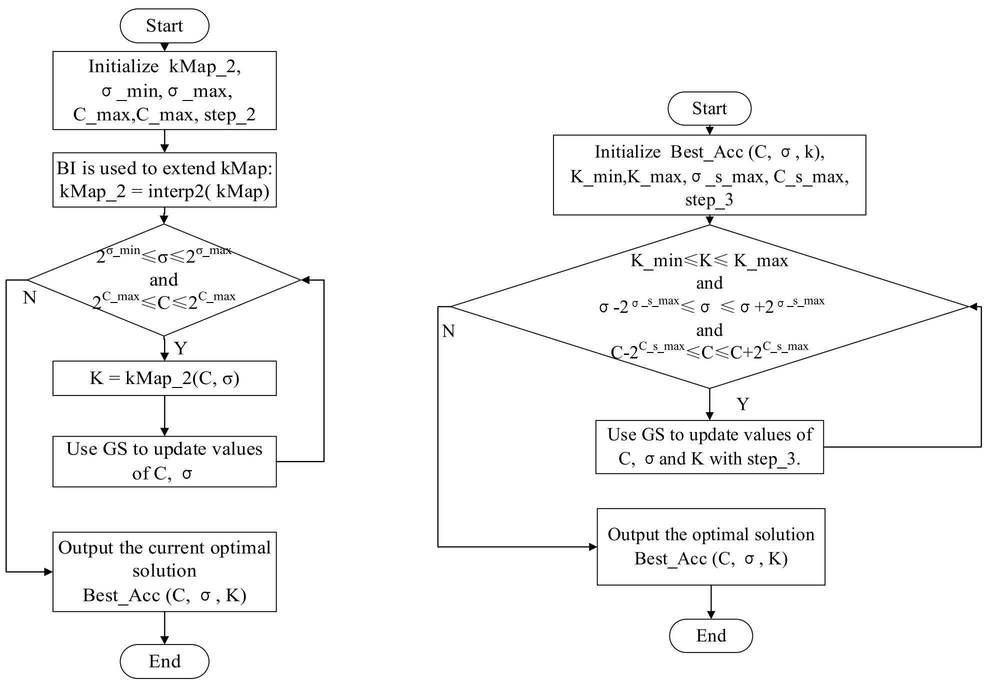


Fig. 17 kMap obtained from BI





(a) (b)

(a) Second step: update and expand kMap based on BI with the middle step size

(b) Third step: obtain optimal hyper-parameters with the small step size

Fig. 18 Flow chart with the middle and small step size

optimal K matrix corresponding to all grid points at the current large step size can be obtained when all values of C and σ are traversed. The specific process is as follows:

1. The value ranges and step sizes of C, σ, K are set.
2. Execute the loop body, update C and σ with GS until the search is completed. Then execute (4).
3. According to Eq. (6), the optimal K of the current C and σ is calculated.
4. Obtain the optimal K distribution diagram for all values of C and σ, referred to as kMap.

The process of obtaining a contour map of K is shown in Fig. 14:

Taking the parameter settings in Fig. 14 as an example. When the optimal matrix K is calculated with $C\_step=\sigma\_step=0.5$ rather than $C\_step=\sigma\_step=0.2$, the contour map of the K matrix is drawn in the $C-\sigma$ plane, as shown in Fig. 15. Different colors represent the K values in different regions corresponding to the highest model accuracy.

As shown in Fig. 15, each optimal K region is a continuous block; its range is similar to that of when $C\_step=\sigma\_step=0.2$. Therefore, kMap can be used to further search for the optimal solution with the two-dimensional $C-\sigma$ grid method under the middle step size, until the target step size is reached.

Bilinear interpolation (BI) is required to interpolate kMap to obtain the same optimal K values as the discrete grid points with a middle step size. The interpolation formula is shown below:

$$k(C,\sigma)=\begin{bmatrix}\frac{C_2-C}{C_2-C_1} & \frac{C-C_1}{C_2-C_1}\end{bmatrix}\begin{bmatrix}k\left(C_1,\sigma_1\right) & k\left(C_1,\sigma_2\right)\\ k\left(C_2,\sigma_1\right) & k\left(C_2,\sigma_2\right)\end{bmatrix}\begin{bmatrix}\frac{\sigma_2-\sigma}{\sigma_2-\sigma_1}\\ \frac{\sigma-\sigma_1}{\sigma_2-\sigma_1}\end{bmatrix} \tag{7}$$

where $k(C_1, \sigma_1)$, $k(C_1, \sigma_2)$, $k(C_2, \sigma_1)$, and $k(C_2, \sigma_2)$, are the 4 known points closest to the current interpolation point, as shown in Fig. 16. According to Eq. (7), the kMap corresponding to the middle step size can be obtained. Setting $C\_step=\sigma\_step=0.2$, the kMap obtained by the BI of the kMap generated by $C\_step=\sigma\_step=0.5$ is shown in Fig. 17. Its continuous region and boundary are well preserved.

After obtaining kMap with a middle step size, the optimal solution can be further searched according to the GS. The specific process of the algorithm is shown in Fig. 18a. The optimal combination (C, σ, K) generated from the discrete step length can be taken as the value range center. A small value range and search distance are given and the optimal solution can be searched using the GS. If the optimal (C, σ, K) is not updated, the optimal solution can be considered found. The specific process is shown in Fig. 18b.

For the optimization of the values of the 3 hyper-parameters, the overall algorithm flow based on the kMap method is as follows:

1. Initialize parameters.
2. Search for all K values with the highest accuracy under large step sizes of C, σ and obtain kMap.
3. According to kMap with the large step size, the BI method is used to update and expand kMap. The GS is used to search for the current optimal (C, σ, K) with the middle step size.

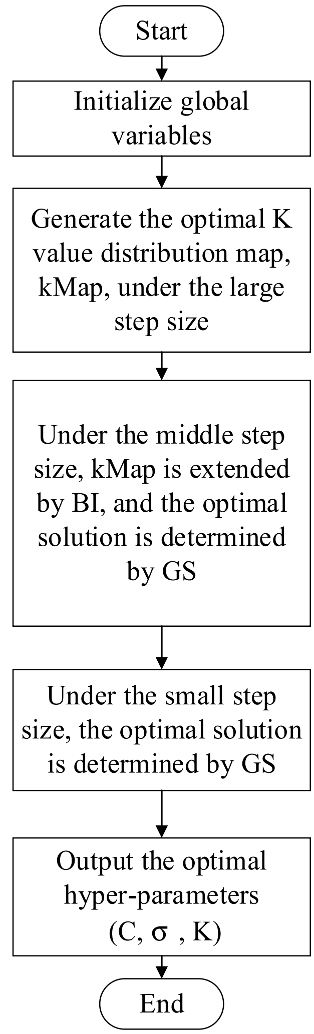


Fig. 19 Overall flow chart of kMap algorithm

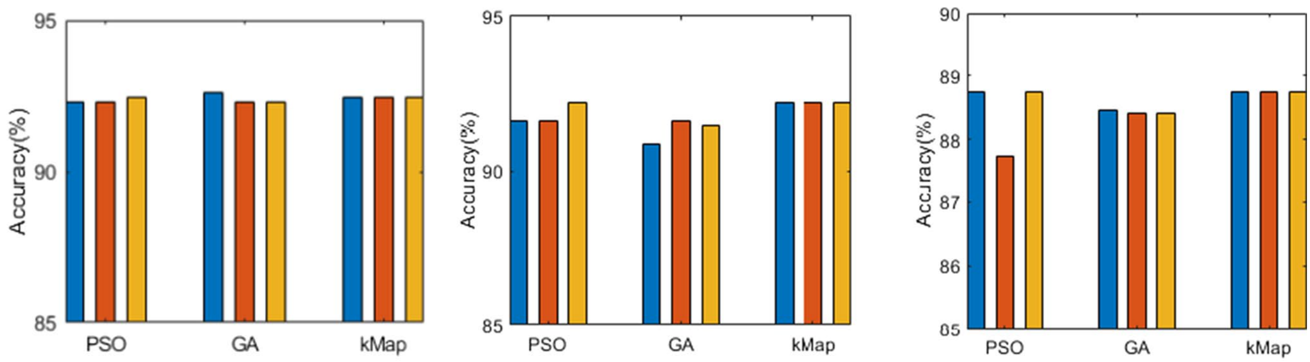
4. Taking the optimal hyper-parameter combination generated in (3) as the center point, update the optimal (C, K) by GS with a small step size.
5. Output the optimal (C, K), end.

The overall flow chart is shown in Fig. 19.

To select optimal hyper-parameters, assuming that n C, m σ and g K values are used to identify the model for optimal parameter selection, the time complexity of kMap is O(nmg). The time complexity of kMap is the same as that of the GS algorithm proposed in the following section. However, the number of parameters taken by kMap is far less than that of GS and part of the calculation process is greatly simplified, with the computational burden greatly reduced.

Table 2 The results optimized with GS and kMap

	Accuracy (%)	Time	C	σ	K
GS	92.59	37 h	1.35	0.43	4
kMap	92.45	261 s	7.2602	0.1303	8



(a) Raw data (b) $R_1 = 0.4, R_2 = 0.1, \varepsilon = 0.7$ (c) $R_1 = 0.8, R_2 = 0.6, \varepsilon = 0.7$

Fig. 20 Identification model accuracies

Identification accuracy and experiment analysis

Optimization result and analysis of kMap

As an improved GS algorithm, the operation time of kMap is far less than that of GS. The operation results are shown in Table 2. In the optimizing process of GS, $C\_min = \sigma\_min = -5$, $C\_max = \sigma\_max = 5$, $C\_step = \sigma\_step = 0.02$, and $K \in \{2, 3...10\}$. The obtained optimal model accuracy was 92.59% and the time used was approximately 37 h. The hyper-parameters of the VMD-SVM model are optimized by kMap with the same parameter ranges. The sizes of the large step, middle step, and small step are set as 0.5, 0.2, and 0.02, respectively. The accuracy of the model optimized by kMap was 92.45%. kMap saves large computational expense with a small accuracy cost.

The heuristic algorithm (HA) [e.g., PSO (Kennedy and Eberhart 1995) and GA (Holland 1973)] is an effective optimization method. Compare these methods in kMap, with the following optimization settings of PSO, GA, and kMap: optimization ranges of C, σ, K: [− 5, 5], [− 5, 5), [2, 10], evolution time number is 200, and population size is 20. In addition, the maximum velocity of the particle and termination error value are set as 0.02 and 1e-25, respectively, in PSO. The individual length of the gene and the termination error value are 11×2 and 1e-25 in GA. The step sizes of kMap are taken as three different sets of values; (0.5, 0.2, 0.02), (0.5, 0.1, 0.02), and (0.4, 0.12, 0.024).

It is assumed that the stochastic disturbance has a non-Gaussian distribution in the case of industrial robotic milling. The measured acceleration data $X(t)$ is written as:

$$X(t) = X'(t) + e(t) \tag{8}$$

where $X'(t)$ and $e(t)$ are the real acceleration data and stochastic noise, respectively. It is assumed that the measurement noise $e(t)$ has a non-Gaussian distribution, with approximately normal distribution classes (Stojanovic and Prsic 2020):

$$P_\varepsilon = \{p(e) : p(e) = (1 - \varepsilon)p_1(e) + \varepsilon p_2(e)\} \tag{9}$$

where the probability density $p(e)$ represents a mixture of primary probability density $p_1(e)$: $N(0, R_1)$ and contaminating probability density $p_2(e)$: $N(0, R_2)$. The degree ε is in range $0 < \varepsilon < 1$, while R_1 and R_2 are covariance matrices of primary and contaminating terms in a non-Gaussian distribution. Non-Gaussian noises with different R_1, R_2, and ε are added to the cutting data. These datasets containing noise are used to train the identification model, and PSO, GA, and kMap are used to obtain optimal hyper-parameters. The results are shown in Fig. 20.

The operation time of kMap is of the same order of magnitude as PSO and GA, i.e., hundreds of seconds.

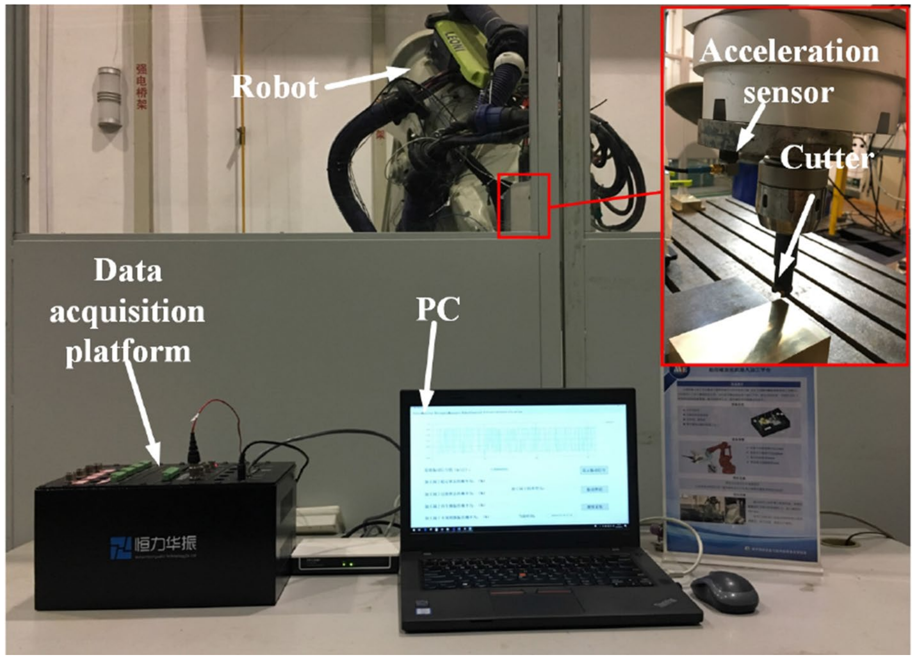


Fig. 21 Robotic milling vibration monitoring experiment

However, the average optimization performance of kMap is slightly better than that of PSO and GA. It is worth noting that kMap inherits the advantages of stable optimization of the GS; this means that in the same optimization range, the optimization result of kMap is almost unchanged when the lengths of the searching steps are changed. It is obvious that non-Gaussian noise affects the accuracy of the identification model so it is important to appropriately design a cutting signal filter.

Chatter identification experiment

The chatter identification model optimized by kMap was used in the experiments of robotic milling. The hardware is a Henrywaltz data acquisition module (Fig. 21). The cutting tools and the material used in the experiment are the same as section "Experimental platform and parameter setting" and are machined by up-milling. Different machining parameters were set for the robotic milling experiments. The experimental machining parameters are listed in Table 3. The results of the software used to monitor vibration states are shown in Fig. 22. The results are presented in the form of four types of vibrations in robotic milling.

Take experiments 3, 4, 11, and 12 as examples for analysis. The machined surfaces under the four sets of processing parameters are shown in Fig. 23. In experiment 3, an irregular vibration pattern appears on the machined surface, which is judged as irregular chatter according to experience. In experiment 4, there are slight vibration marks on the processed surface except for the gear mark, which is judged as a transition state according to experience. In experiment 11, the machined surface is smooth and clean, except for the gear mark, and is judged as a stable state according to experience. In experiment 12, the processed surface shows an obvious regular vibration pattern, which is determined as regular chatter according to experience.

Table 3 Processing parameters of chatter identification experiment

Number	Speed (rpm)	Feed (mm/s)	Depth (mm)	Width (mm)
1	3200	8	1.1	8
2	3200	8	2	8
3	3600	9	2.1	8
4	4000	10	0.5	8
5	4000	10	1.2	8
6	4400	11	0.3	8
7	4400	11	0.7	8
8	5000	12.5	1.2	8
9	5200	13	0.5	8
10	5200	13	0.9	8
11	5800	14.5	0.5	8
12	5800	14.5	1.7	8

On average, it takes 1.256 s to identify the vibration state, with the vibration data of 5000 sampling points, and the time for data feature extraction is relatively large. Notably, the identification model can be used for chatter identification when more than 5,000 data are collected in the identification experiment, i.e., after 0.5 s of sampling. The identification results of the 4 groups of experiments

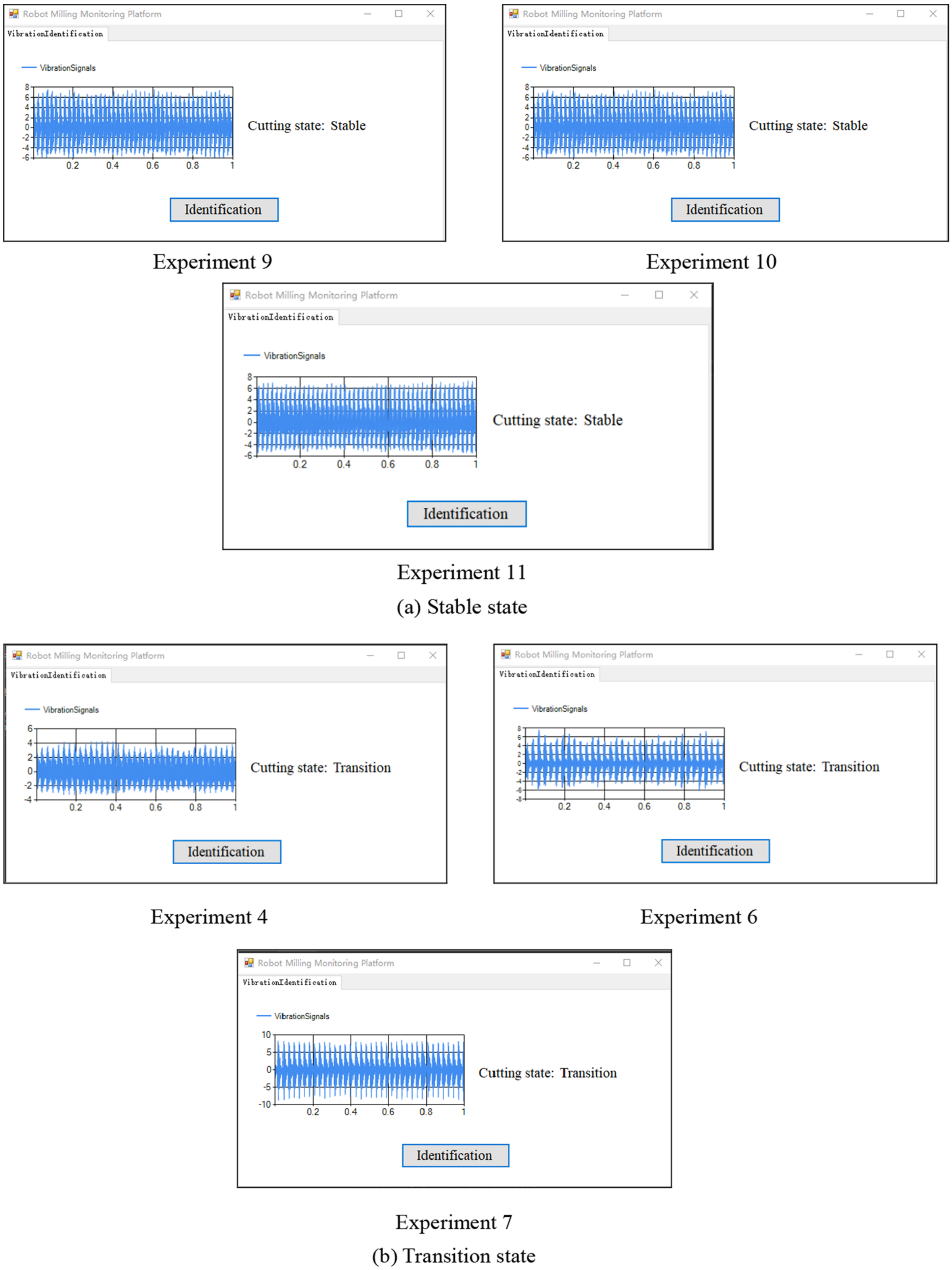


Fig. 22 Robotic milling chatter monitoring

are irregular chatter, transition, stable, and regular chatter states. The validity of the model in chatter identification is verified by the same results as those obtained from machining surfaces and experiment results.

Conclusions

To better fit the actual situation of robotic milling and provide the basis for further study of chatter mechanisms,

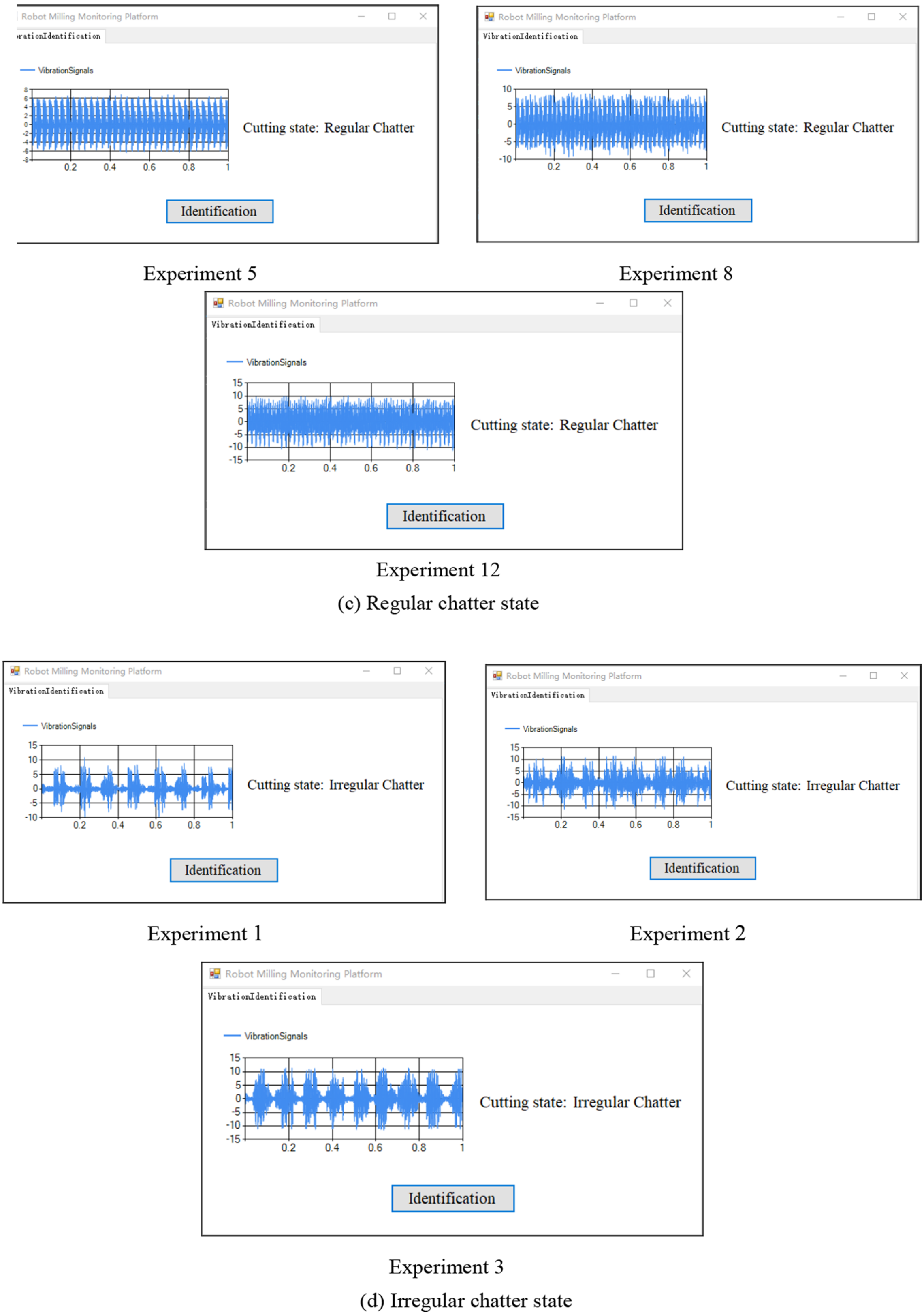
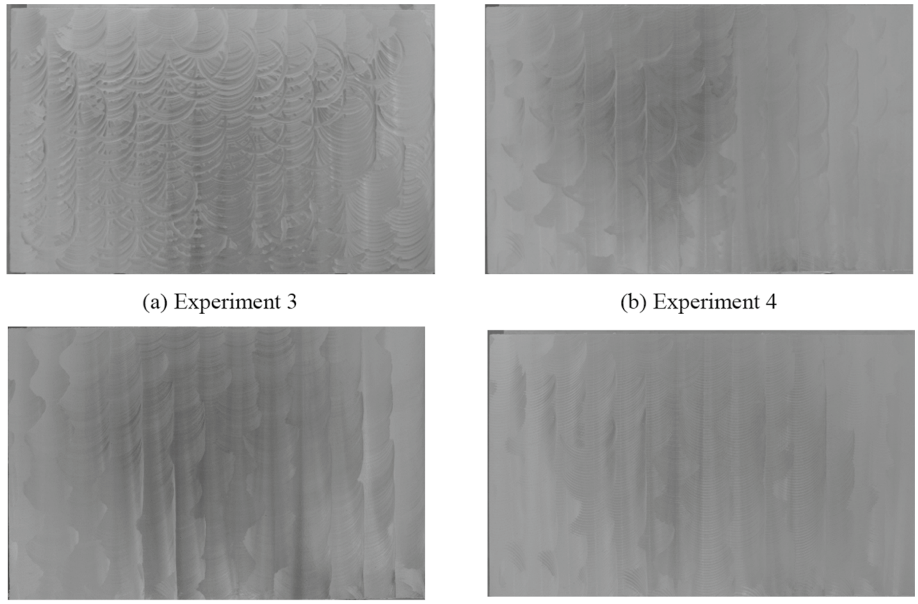


Fig. 22 (continued)

Fig. 23 Machined surface



(a) Experiment 3 (b) Experiment 4 (c) Experiment 11 (d) Experiment 12

this paper divides the vibration states of industrial robotic milling into four types: stable, transitional, regular chatter, and irregular chatter. The VMD-SVM model is applied for chatter detection of robotic milling for the first time, which is trained using 160 robotic milling experiments. To improve the accuracy of the chatter identification model, a novel optimization method—the kMap algorithm—is proposed in this paper for the optimization of 3-dimensional hyper-parameters of the VMD-SVM model. Finally, validation experiments in 12 parameter sets are used to confirm the effectiveness of the chatter identification model. Through the research of this paper, the efficiency and accurate identification of chatter in robotic milling can be realized. Based on the above research on chatter detection in robotic milling, some conclusions are summarized as follows:

1. The vibration data is identified by classifying the vibration data promptly measured using a VMD-SVM model based on IE where VMD adaptively decomposes the vibration signals and IE quantifies the clutter degree of vibration data. VMD and IE can effectively extract the features of vibration data. The experimental results show that the frequency and time–frequency characteristics of the 4 types of vibration signals proposed in this paper (stable, transitional, regular chatter, and irregular chatter) correspond to the textures on the machined surfaces in robotic milling. The identification accuracy of VMD-SVM model is 92.59% and can effectively identify chatter in robot processing.
2. Compared to the optimization performance of GS, GA, and PSO, the kMap proposed in this paper shows comprehensive advantages in terms of optimization time, accuracy, and stability. The optimization time of kMap was 261 s; this is far less than that of GA (37 h). The average optimization performance of kMap is slightly better than that of PSO and GA. Furthermore, kMap inherits the advantages of stable optimization of GS, i.e., the optimization result of kMap is almost unchanged within the same optimization range. In addition to the application case in this study, the kMap algorithm can be applied to other multidimensional hyper-parameter optimization. To apply kMap to other applications, it is necessary to adjust the optimization range and step size.

Acknowledgements This research is supported by National Natural Science Foundation of China under Grant No. 51805189, National Natural Science Foundation of China under Grant No. 51625502, and Science Fund for Creative Research Groups of the National Natural Science Foundation of China under Grant No. 51721092.

References

Abdoos, A. A., Mianaei, P. K., & Ghadikolaei, M. R. (2016). Combined VMD-SVM based feature selection method for classification of power quality events. *Applied Soft Computing, 38,* 637–646.

Aneesh, C., Kumar, S., Hisham, P. M., & Soman, K. P. (2015). Performance comparison of variational mode decomposition over empirical wavelet transform for the classification of power quality disturbances using support vector machine. *Procedia Computer Science, 46,* 372–380.

Aslan, D., & Altintas, Y. (2018). On-line chatter detection in milling using drive motor current commands extracted from CNC. *International Journal of Machine Tools and Manufacture, 132,* 64–80.

Chen, Y., Li, H., Hou, L., Bu, X., Ye, S., & Chen, D. (2020). Chatter detection for milling using novel p-leader multifractal features. *Journal of Intelligent Manufacturing*. https://doi.org/10.1007/s10845-020-01651-5.

Dragomiretskiy, K., & Zosso, D. (2013). Variational mode decomposition. *IEEE Transactions on Signal Processing, 62*(3), 531–544.

Dutta, T., Satija, U., Ramkumar, B., & Manikandan, M.S. (2016). A novel method for automatic modulation classification under non-Gaussian noise based on variational mode decomposition. In: *2016 twenty second national conference on communication (NCC)* (pp. 1–6).

Friedrich, J., Hinze, C., Renner, A., Verl, A., & Lechler, A. (2017). Estimation of stability lobe diagrams in milling with continuous learning algorithms. *Robotics and Computer-Integrated Manufacturing, 43,* 124–134.

Fu, W., Tan, J., Xu, Y., Wang, K., & Chen, T. (2019a). Fault diagnosis for rolling bearings based on fine-sorted dispersion entropy and SVM optimized with mutation SCA-PSO. *Entropy, 21*(4), 1–23.

Fu, W., Wang, K., Li, C., Li, X., Li, Y., & Zhong, H. (2018). Vibration trend measurement for a hydropower generator based on optimal variational mode decomposition and an LSSVM improved with chaotic sine cosine algorithm optimization. *Measurement Science and Technology, 30*(1), 1–15.

Fu, Y., Zhang, Y., Gao, H., Mao, T., Zhou, H., Sun, R., & Li, D. (2019b). Automatic feature constructing from vibration signals for machining state monitoring. *Journal of Intelligent Manufacturing, 30*(3), 995–1008.

Gienke, O., Pan, Z., Yuan, L., Lepper, T., & Van Duin, S. (2019). Mode coupling chatter prediction and avoidance in robotic machining process. *International Journal of Advanced Manufacturing Technology, 104*(5–8), 2103–2116.

Holland, J. H. (1973). Genetic algorithms and the optimal allocation of trials. *SIAM Journal on Computing, 2*(2), 88–105. https://doi.org/10.1137/0202009.

Kennedy, J., & Eberhart, R. C. (1995). Particle swarm optimization. *Proceedings of IEEE International Conference on Neural Networks, 4,* 1942–1948.

Lei, N., & Soshi, M. (2017). Vision-based system for chatter identification and process optimization in high-speed milling. *International Journal of Advanced Manufacturing Technology, 89*(9–12), 2757–2769.

Liu, J., Wu, B., Wang, Y., & Hu, Y. (2017). An integrated condition-monitoring method for a milling process using reduced decomposition features. *Measurement Science and Technology, 28*(8), 1–13.

Lv, Z., Tang, B., Zhou, Y., & Zhou, C. (2016). A novel method for mechanical fault diagnosis based on variational mode decomposition and multikernel support vector machine. *Shock and Vibration, 2016,* 1–11.

Musselman, M., Xie, H., & Djurdjanovic, D. (2019). Nonstationary signal analysis and support vector machine based classification for vibration based characterization and monitoring of slit valves in semiconductor manufacturing. *Journal of Intelligent Manufacturing, 30*(3), 1099–1110.

Nannapaneni, S., Mahadevan, S., Dubey, A., & Lee, Y. T. (2020). Online monitoring and control of a cyber-physical manufacturing process under uncertainty. *Journal of Intelligent Manufacturing*. https://doi.org/10.1007/s10845-020-01609-7.

Pan, Z., Zhang, H., Zhu, Z., & Wang, J. (2006). Chatter analysis of robotic machining process. *Journal of Materials Processing Technology, 173*(3), 301–309.

Stojanovic, V., & Prsic, D. (2020). Robust identification for fault detection in the presence of non-Gaussian noises: application to hydraulic servo drives. *Nonlinear Dynamics, 100*(3), 2299–2313.

Tang, X., Peng, F., Yan, R., Gong, Y., Li, Y., & Jiang, L. (2017). Accurate and efficient prediction of milling stability with updated full-discretization method. *International Journal of Advanced Manufacturing Technology, 88*(9–12), 2357–2368.

Tangjitsitcharoen, S., Saksri, T., & Ratanakuakangwan, S. (2015). Advance in chatter detection in ball end milling process by utilizing wavelet transform. *Journal of Intelligent Manufacturing, 26,* 485–499.

Tao, J., Qin, C., Xiao, D., Shi, H., Ling, X., Li, B., & Liu, C. (2019). Timely chatter identification for robotic drilling using a local maximum synchrosqueezing-based method. *Journal of Intelligent Manufacturing, 31,* 1243–1255.

Xu, B., Li, H., Zhou, F., Yan, B., Liu, Y., & Ma, Y. (2019). Fault Diagnosis of Variable Load Bearing Based on Quantum Chaotic Fruit Fly VMD and Variational RVM. *Shock and Vibration, 2019,* 1–20.

Zhao, X., Qin, Y., He, C., & Jia, L. (2020). Underdetermined blind source extraction of early vehicle bearing faults based on EMD and kernelized correlation maximization. *Journal of Intelligent Manufacturing*. https://doi.org/10.1007/s10845-020-01655-1.