Int J Adv Manuf Technol (2017) 89:175–190
DOI 10.1007/s00170-016-9070-x

ORIGINAL ARTICLE

Tool wear monitoring based on kernel principal component analysis and *v*-support vector regression

Dongdong Kong[1] · **Yongjie Chen**[1] · **Ning Li**[1] · **Shenglin Tan**[1]

Received: 1 January 2016 / Accepted: 15 June 2016 / Published online: 27 June 2016

Abstract Machined surface quality and dimensional accuracy are significantly affected by tool wear in machining process, and severe tool wear may even lead to failing of the workpieces being processed. Tool wear monitoring is highly desirable to realize automated or unmanned machining process, which can get rid of the dependence on skilled workers. This paper mainly studies on the methods and techniques of on-line tool wear monitoring through static and dynamic cutting force signals. Sensitive signals related to tool wear are preliminarily selected by using correlation coefficient method. Kernel principal component analysis (KPCA) technique is adopted to fuse these sensitive features for improving training speed and prediction accuracy. Then, the tool wear predictive model based on *v*-support vector regression (*v*-SVR) is constructed through learning correlation between the fused features and actual tool wear. The obtained result shows that the prediction accuracy of the proposed tool wear model is proved effective beyond expectation. Besides, the proposed model still has better generalization ability even in small sample size.

Keywords Tool wear monitoring · Cutting forces · Signal features · KPCA · SVR

✉ Dongdong Kong
kodon007@163.com

[1] School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan, China

1 Introduction

Tool wear is an inevitable phenomenon in actual machining process due to adhesion, abrasion, diffusion, thermal fatigue, and elastic or plastic deformation. The state of tool wear is a vital factor directly affecting surface quality and dimensional accuracy of the workpieces being processed. It is obvious that the usage of worn tools will lead to fringe costs in terms of tool breakage, scrapped components, and unscheduled downtime. More seriously, the manufacturing system may suffer from chatter vibration caused by severe tool wear. Before 1990s, tool changing mainly depends on skilled workers to judge tool wear state through personal experience, which is not economical with the increase of labor costs. Besides, false judgment of tool wear state occurs sometimes because tool wear is a complex phenomenon that indicates itself through non-linear manners. It is meaningful to develop on-line tool condition monitoring (TCM) systems to monitor tool wear state in real time. The following aspects can be achieved by using TCM system: (a) establishing a more reasonable and effective tool changing strategy to extend tool life as long as possible, (b) reducing downtime caused by sudden tool breakage to ensure smooth machining process and productivity improvement, and (c) carrying out the corresponding error compensation to ensure machining accuracy and surface quality of the workpieces being processed.

The late twentieth and early twenty-first centuries have witnessed a rapid development of direct online TCM systems based on machine vision [1, 2]. However, due to the existence of radiation, influence of cutting fluid, and high monitoring costs, it severely limits the promotion of direct method in industry. Indirect method [3–5] has been the main methodologies for online TCM through constructing the correlation model between tool wear and related signals, such as cutting forces [6–8], vibration [9], and acoustic emission [10], which

are acquired from machining processes. An exclusive review of sensor signals for indirect TCM is presented in [11]. In recent years, scholars have put forward many methods of pattern recognition for online TCM, which can be summarized into the following two types. Firstly, the most common methods in research are establishing theoretical models of tool wear. However, machining is a rather complex process subjects to many unknown factors, which makes it unpractical or too expensive to establish accurate tool wear model. Secondly, supervised learning is receiving more and more attentions in TCM with the improvement of data-processing capability of computer. Supervised learning consists of two processes, i.e., learning of the existing training data and classification or regression of the test data. An optimal model that has first-class treatment effect in some certain evaluation criteria is generated once the training stage is finished. All input data can be mapped to corresponding output through this optimal model so as to identify the unknown data. Kaya et al. [12] applied artificial neural networks (ANN) to estimate tool flank wear for CNC milling of Inconel 718. The results showed that the proposed ANN model manifested a better statistical performance with a high correlation between the actual and predicted values of flank wear. Fish et al. [13] applied Hidden Markov models to predict tool wear in milling process by vibratory features extracted in frequency domain. Based on dynamic bayesian networks (DBN), Tobon-Mejia et al. [14] constructed an effective and efficient model to identify current health state of the cutting tools, predicted its remaining useful life, and associated confidence bounds. Dimla D E and Lister P M [15] applied multi-layer perceptron neural networks to realize the classification of tool wear state. The results showed that the correct classification rate of this approach could reach 95.81 %. Chungchoo C and Saini D [16] developed an on-line fuzzy neural network (FNN) model to predict rake face and flank wear in turning with less computational cost.

Although some achievements have been made in the research of TCM systems, the following aspects are still needed to take seriously. Firstly, feature selection is of vital importance before feeding the related features into a TCM system. Previous studies for TCM systems mostly rely on only small amounts of sensitive features or even individual feature. Kious M et al. [17] analyzed the correlation between tool wear rate and average value, root mean square (RMS), and variance of cutting force signals in high-speed milling. They noticed that variance, extracted in time domain, was more appropriate to reflect the varying trend of tool wear. Kamarthi and Pittner [18] used force and vibration signals for flank wear estimation in turning processes. They recommended FFT for vibration signals, and WT seemed more suitable for force signals. This paper extracts large amounts of features related to tool wear from the original static and dynamic cutting force signals, with a view to reflect tool wear process comprehensively. The sensitive features related to tool wear are preliminarily selected by using correlation coefficient method. However, there are plenty of noise and redundant information in these sensitive features, which will lead to higher memory usage and longer computing time. KPCA technique is proposed to address these problems by fusing these sensitive features so as to remove irrelevant information and reduce dimensionality. KPCA is an effective analytical method in dealing with higher-order data and has broad application prospects in feature extraction due to its ability to process not only non-gaussian signals but also non-linear and non-stationary signals. By implementation of KPCA, the sensitive features related to tool wear will be reduced to new low-dimensional feature vectors. This makes contribution to improving training speed and prediction accuracy. Secondly, neural network system has been widely used in tool wear monitoring. However, the prediction accuracy is restricted to a certain degree due to the inherent weakness of neural network, such as local minimum value, over-fitting, and poor generalization especially when there are few samples. Support vector machines (SVM) will be used to construct tool wear model in this paper. Based on statistical learning theory and minimization of structural risk, SVM is a great achievement in data processing and especially fit for pattern classification and non-linear function regression. D. Shi and N. Gindy [19] employed LS-SVM in the prediction of tool wear for broaching process and achieved an effective prediction model. In addition, SVM has better generalization ability even in small samples since SVM has overcome the inherent weakness of neural network in a way.

This paper presents an effective online TCM system, and final result shows that the prediction accuracy of the proposed tool wear model based on *v*-SVR is superior. The paper is organized as follows: Architecture of the on-line tool wear monitoring system is presented in Section 2. KPCA technique is introduced to extract principal features from the sensitive features related to tool wear in Section 3. The *v*-SVR is further introduced to construct the model between extracted features and actual tool wear in Section 4. Prediction results of the proposed tool wear model are presented in Section 5. In Section 6, this paper is concluded with a summary of the proposed tool wear model.

2 Architecture of the on-line tool wear monitoring system

The architecture of the on-line tool wear monitoring system has been designed to predict tool wear values as accurately as possible for hard turning as shown in Fig. 1, and its components are listed in Table 1. The details of this measuring chain will be introduced as follows.

The piezoelectric dynamometer has a great rigidity and consequently a high natural frequency. Moreover, its high resolution enables the smallest dynamic changes in large

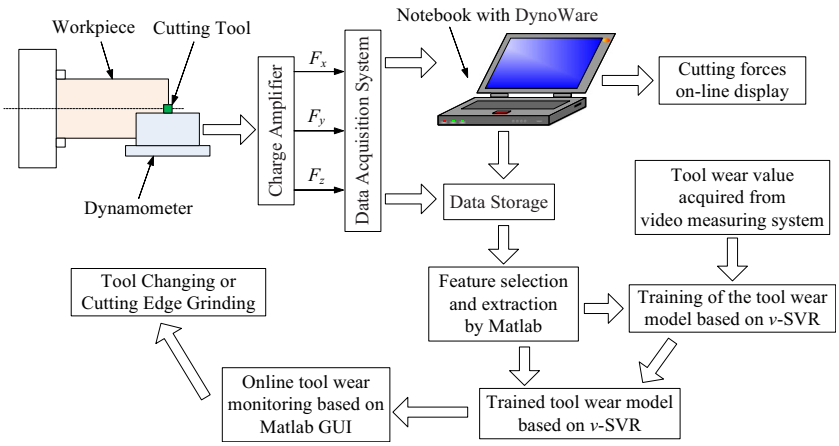


Fig. 1 Architecture of the tool wear monitoring system

forces to be measured. So dynamometer is the most accurate and dedicated sensor to measure cutting forces in machining process. Dynamometer has been widely used in the research of machine manufacturing. The charge amplifier that lays in front of the DQA board is ideal for the measurement of cutting forces with piezoelectric dynamometers. The piezoelectric force sensor produces electric charge which varies proportionally with the mechanical quantities such as force, torque, and acceleration acting on the sensor. The charge amplifier then converts the electric charge into a proportional voltage. The measurands will be read out via interface RS-232C.

DAQ system has been developed specifically for the piezoelectric measuring systems and selected as the hardware platform connecting charge amplifier with the DynoWare package. High resolution of the DAQ system and its high sampling frequency of up to 125 kS/s with 8 measuring channels enable researchers to acquire as accurate information as possible from highly dynamic cutting processes and make it have a very wide range of applications, such as manufacturing, military, and aerospace fields. The DAQ system is connected via a USB 2.0 port to the PC and controlled by the DynoWare program. The DynoWare package has been selected as the software platform due to its powerful function in signal acquisition, data processing, and data management.

Table 1 Experimental components

Components	
Engine lathe	CW6163E
Indexable inserts	Sandvik CNMG120408-PM
Tool holder	Sandvik PCLNR 2525 M 12
Workpieces	Normalizing 50# Steel (HB 160 ~ 197)
Dynamometer	Kistler 9257A
Charge amplifier	Kistler 5070A
Data acquisition system	Kistler 5697A
Notebook with DynoWare	Kistler 2825A
Video measuring system	VMS-1510G (QIM1008)

Moreover, DynoWare offers a powerful online visualization environment for signal analysis. According to practical analysis, DynoWare can conduct signal drift compensation, signal filtering, and signal smoothing in time domain through the selection of different icons. Once the whole tool wear monitoring runs, the DynoWare program will acquire the signals of static and dynamic cutting forces from the dynamometer, output the results in a graphical environment, and finally save the measured data in ASCII format (.txt).

The length of tool flank wear, namely VB, in half of the cutting depth is used to quantify tool wear as shown in Fig. 2. In each set of turning experiments, the actual flank wear will be measured by video measuring system at certain intervals of the cutting process. The video measuring system in terms of VMS-1510G consists of a series of components, such as high-resolution charge-coupled device (CCD) color camera, high-precision working table, and 2D measuring software QIM1008. The measuring accuracy of VMS-1510G can reach micron level. The tool flank wear can be easily measured with the measuring system.

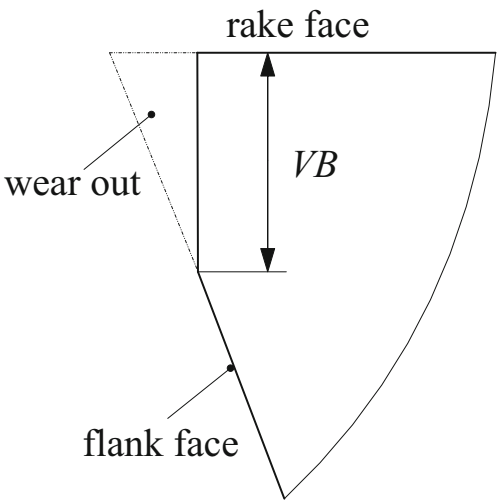


Fig. 2 Tool flank wear in terms of VB

Besides, the output images can be read out by QIM1008 and showed on the display as illustrated in Fig. 3.

The correlations between feature vector and the corresponding VB will be investigated. Once one cutting process is finished, the customized program compiled by Matlab will extract time and time-frequency domain signals related to tool wear from the original cutting forces. Then, the sensitive signals related to tool wear will be preliminarily selected by means of correlation coefficient method. As soon as the sensitive signals are acquired, KPCA will be used to further extract features related to tool wear. Finally, the model between extracted features and actual tool wear is constructed by *v*-SVR.

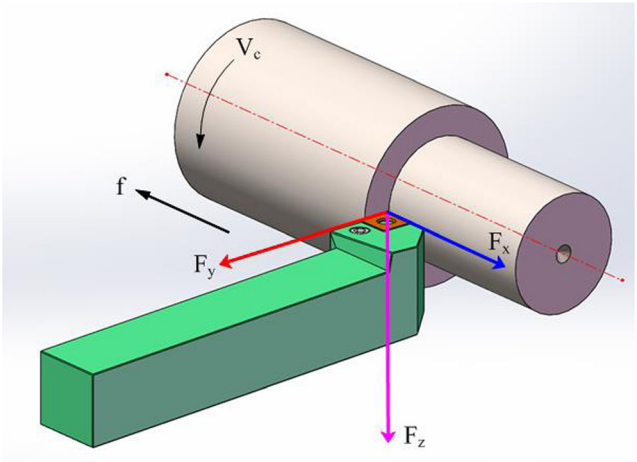


Fig. 4 The three perpendicular cutting forces

3 Feature extraction based on KPCA

The cutting force signals consist of low-frequency content that is an indication of the static cutting forces and higher frequency range that is the natural frequencies of the tool-holder resulted from the excitation of the cutting operation [20]. In this section, large amounts of features related to tool wear are extracted from the original static and dynamic cutting force signals in order to reflect tool wear process comprehensively.

3.1 Signal features derived from time and time-frequency domains

There are three perpendicular sensory signals, i.e., cutting force signals, could be acquired from the dynamometer as shown in Fig. 4. They are feed force F_x, radial force F_y, and main cutting force F_z, respectively. The general purpose time domain features computed from the three cutting forces are (i) mean value, (ii) ratios of the signals, (iii) root mean square (RMS), (iv) variance, (v) skewness, (vi) kurtosis, and (vii) form factor. There are 21 signal features in all derived from the cutting forces through time domain statistical analysis.

Fig. 3 The Video Measuring System (VMS-1510G)

In addition, wavelet analysis has been the most popular time-frequency analysis tool since its time and frequency resolution changes adaptively depending on the frequency of the signal and, hence, is more appropriate for the analysis of non-stationary signals. By Mallat algorithms, the original signal $f(t)$ is firstly decomposed into two parts: the approximation coefficients and the detail coefficients. The new approximation coefficients will be recursively decomposed in the next step. Scale of decomposition relies on the actual requirement.

It is worth noting that multi-resolution analysis decomposes the original signal into different frequency domains effectively. The energy of time domain force signals in different frequency domains will change gradually with the tool wear. An important feature could be acquired from this process, i.e., ratio of the energy of different frequency domains to total energy of the signal. According to the conservation of energy in wavelet transform, the energy of the approximation coefficients and detail coefficients at different scales are employed to represent the energy of the corresponding frequency domain in this paper. By a 5-level MRA analysis as shown in Fig. 5, there are 18 signal features in all could be acquired from the cutting forces.

The details of preliminary selection of sensitive signals related to tool wear by correlation coefficient method will be introduced in Section 5.

3.2 Feature extraction based on KPCA

The main objective of feature extraction is to abstract statistical information on feature category effectively from the signal features scattered in input space so as to improve prediction accuracy and reduce computational complexity. Principal component analysis (PCA), also known as the Karhunen-Loeve transformation, gives the optimal representation of the

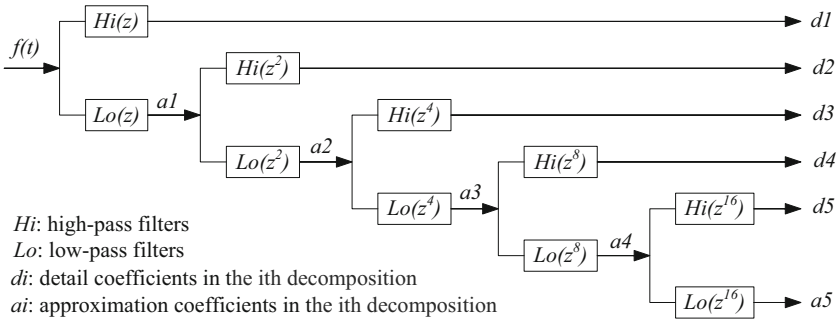


Fig. 5 5-level wavelet decomposition of the original cutting forces

disperse signal features by means of maximum variance. It is an efficient approach to realize dimensionality reduction by removing the correlation between signal features and has been widely used in feature extraction for fault diagnosis as the fore-end of a classifier [21]. However, PCA is not a suitable technique for dealing with non-linear classification problem due to its linear property in essence. Only in high-dimensional non-linear space more effective information with respect to feature selection can be extracted from tool wear which is a complex process manifests itself in nonlinear manners.

Kernel-PCA is a new method for performing a non-linear form of principal component analysis through introducing kernel methods (KMs) [22]. The multi-dimensional signal features acquired from the dynamometer sensor could be expressed through a matrix $X_{n\times m}$ with m samples and n dimensional features. Firstly, we map the data in input space $\boldsymbol{R^n}$ non-linearly into a feature space $\boldsymbol{F}$ by $\Phi : x_i \rightarrow \Phi(x_i)$. Suppose that the data mapped into feature space has been centralized, i.e., $\sum_{i=1}^{m}\Phi(x_i) = 0$. According to the solution procedure of normal PCA, the covariance matrix C in feature space $\boldsymbol{F}$ can be expressed by

$$\overline{C} = \frac{1}{m}\sum_{i=1}^{m}\Phi(x_i)\Phi^T(x_i). \tag{1}$$

The eigenvalues λ and eigenvectors V satisfying $\lambda V = \overline{C}\ V$ will be acquired after the eigenvalue decomposition with respect to $\overline{C}$. However, it is almost impossible to obtain the covariance matrix $\overline{C}$ since the non-linear mapping Φ is not explicit. Fortunately, it is not necessary to calculate the image from input space to feature space explicitly since the vector dot product in feature space can be calculated by using a kernel function in input space. In feature space $\boldsymbol{F}$, the eigenvectors V can be linearly represented by $\Phi(x_i)$,

$$V = \sum_{i=1}^{m}\alpha_i\Phi(x_i). \tag{2}$$

The following formula can be acquired from the inner product of $\Phi(x_k)$ and $\lambda V = \overline{C}\ V$,

$$\lambda\langle\Phi(x_k), V\rangle = \left\langle\Phi(x_k), \overline{C}\ V\right\rangle \quad k = 1, 2\cdots, m. \tag{3}$$

Defining a $m\times m$ matrix $\mathbf{K} = (K_{ij})$, the kernel function $K_{ij} = K(x_i, x_j) = \langle\Phi(x_i), \Phi(x_j)\rangle$, substituting (1) and (2) into (3), we arrive at,

$$m\lambda\boldsymbol{\alpha} = \mathbf{K}\boldsymbol{\alpha}, \tag{4}$$

where $\alpha = (\alpha_1, \alpha_2, \cdots\alpha_m)^T$, $\lambda \neq 0$, $\lambda^* = m\lambda$ are the eigenvalues of kernel matrix **K,** α is the corresponding eigenvectors. Moreover, the eigenvectors α need to be normalized according to the requirement that V belong to unit vectors, i.e., $\langle V, V\rangle = 1$. The normalized α' should meet the following translation,

$$\langle V, V\rangle = \sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_i\alpha_j K_{ij} = \langle\boldsymbol{\alpha}', \mathbf{K}\boldsymbol{\alpha}'\rangle = \lambda^*\langle\boldsymbol{\alpha}', \boldsymbol{\alpha}'\rangle = 1 \tag{5}$$

In Kernel-PCA, principal components corresponding to sample points x_i are obtained from the projections of the mapped $\Phi(x_i)$ onto the eigenvectors V in feature space $\boldsymbol{F}$ as follows:

$$\langle V, \Phi(x_i)\rangle = \sum_{k=1}^{m}\alpha_k'\langle\Phi(x_i), \Phi(x_k)\rangle = \sum_{k=1}^{m}\alpha'\alpha_k' K_{ik}. \tag{6}$$

Noting that the KPCA algorithm is derived under the assumption that $\sum_{i=1}^{m}\Phi(x_i) = 0$, we need to center the mapped data by $\overline{\Phi}(x_i) = \Phi(x_i) - (1/m)\sum_{k=1}^{m}\Phi(x_k)$. Thus, the centered kernel matrix $\mathbf{K}_{center}$ can be expressed in terms of $\mathbf{K}$ as Eq. (7). As for the new samples x_t, a new kernel matrix $\mathbf{K}_{\text{test}}^{C}$ needs to be reconstructed in terms of $K(x_t, x_i)$ and $K(x_i, x_j)$ as Eq. (8).

$$\mathbf{K}_{\text{center}} = \mathbf{K} - 1_{m\times m}\mathbf{K} - \mathbf{K}1_{m\times m} + 1_{m\times m}\mathbf{K}1_{m\times m}, \tag{7}$$

$$\mathbf{K}_{\text{test}}^{C} = \mathbf{K}_{\text{test}} - \mathbf{K}_{\text{test}}1_{m\times m} - 1_{t\times m}\mathbf{K} + 1_{t\times m}\mathbf{K}1_{m\times m}, \tag{8}$$

where $1_{m\times m}$ represents a $m\times m$ ones matrix with the coefficient $1/m$. $1_{t\times m}$ represents a $t\times m$ ones matrix with the coefficient $1/m$. As for the selection of kernel functions, radial basis function is a better choice in case that there is no prior knowledge of the given data. To reduce the computational complexity in KPCA, radial basis function (RBF) is selected as kernel function

$$K(x_i,x_j)=exp\left(-\frac{\left\|x_i-x_j\right\|^2}{2\sigma^2}\right) \tag{9}$$

where σ is the width parameter for optimization.

The compression ratio of signal features directly affects the computation cost and prediction accuracy of the subsequent tool wear predictive model. In this work, the cumulative contribution rate of principal components variance is employed in determining the total amount of principal components as shown in Eq. (10). Thus, the eigenvectors corresponding to the first k largest eigenvalues are retained to extract features from the sensory signal

$$\sum_{i=1}^{k}\lambda_i^*\Big/\sum_{i=1}^{m}\lambda_i^*\geq 90\%, \tag{10}$$

where λ_i^* is the ith eigenvalue of the centered kernel matrix $\mathbf{K}_{\text{center}}$, $\lambda_1^*\geq\lambda_2^*\geq\cdots\geq\lambda_m^*$. The $m\times k$ dimensional subspace constructed by the first k eigenvectors is namely considered as the required projection direction of K_{center} to extract optimal principal components.

The preliminarily selected signal features from Section 3.1 by correlation coefficient method, together with the three cutting parameters, constitutes the feature set that is finally used to predict tool wear value quantitatively. The feature set is randomly divided into two groups, i.e., training sample and test sample. Considering the elements of feature set have different orders of magnitude, all the elements should be normalized before sending to KPCA in order to eliminate the possibility that large order of magnitude weaken the effect of small one. Firstly, the elements of training sample have to be normalized with zero mean and unit standard deviation to improve training efficiency. Then, the elements of test sample need to be normalized by using the mean value x and standard deviation σ_x of training sample as shown in Eq. (11),

$$x'=\frac{x-\overline{x}}{\sigma_x}. \tag{11}$$

Finally, normalized elements of the feature set are sent to KPCA for compressing the sensitive features related to tool wear, and the extracted features will be sent to the tool wear predictive model in Section 4.

4 Tool wear predictive model based on *v*-SVR

SVM is a powerful tool of solving some problems in data mining by means of optimization method and somehow overcomes the traditional difficulties originating from neural network, such as curse of dimensionality and over-fitting. This novel machine-learning tool based on statistical learning theory and structural risk minimization leads to optimal solution of the original problem by solving the corresponding dual problem characterized by the use of kernel functions and the sparseness of the solution. SVM contains support vector classification (SVC) and support vector regression (SVR). The essential difference between SVC and SVR is that SVR search for optimal hyper plane to minimize the total deviation of the distance between the whole sample points and the optimal hyper plane. Considering the existence of outlier points, ε-SVR is proposed by introducing slack variable ξ and penalty parameter C in order to avoid over-fitting and improve generalization ability. However, rational determination of the parameter ε is a tough proposition in practical application sometimes, and ε-SVR cannot control the number of support vectors. Hence, to improve the solution speed of the model, a modified version, v-SVR is proposed by substituting

$v\in(0,1]$ which has explicit physical meaning for the parameter ε [23]. The deduction of v-SVR is introduced as follows. Given a training sample $\{(x_i,y_i);i=1,2,3\ldots N\}$ with feature vectors $x_i\in\boldsymbol{R^n}$ and actual values $y_i\in\boldsymbol{R}$, the regression function can be reconstructed by using nonlinear transformation $\mathrm{x}=\Phi(x)$:

$$f(x)=\omega^T\Phi(x)+b, \tag{12}$$

where $f(x)$ is predictive value of the model. The linear non-separable problem in input space becomes linearly separable in high-dimensional space through non-linear mapping function $\Phi(\cdot)$. The original problem in v-SVR leads to convex quadratic programming with inequality constraints

$$\min_{\omega,\boldsymbol{b},\boldsymbol{\xi},\boldsymbol{\xi}^*,\varepsilon}\quad\frac{1}{2}\|\omega\|^2+C\left(v\varepsilon+\frac{1}{N}\sum_{N=1}^{N}\left(\xi_i+\xi_i^*\right)\right), \tag{13}$$

subject to inequality constraints

$$\begin{aligned}&\omega^T\Phi(x_i)+b-y_i\leq\varepsilon+\xi_i,\\&y_i-\omega^T\Phi(x_i)-b\leq\varepsilon+\xi_i^*,\\&\xi_i,\xi_i^*\geq 0,i=1,2,\cdots,l,\varepsilon\geq 0,\end{aligned}$$

where the insensitive loss function ε becomes a variable that needs to be optimized, slack variables ξ_i and ξ_i^* weak the effect of outlier points on model fitting accuracy via penalty parameter C. In order to solve this optimization

problem with constraints, Lagrange function is introduced as follows:

$$L(\omega,b,\varepsilon,\xi,\xi^*)=\frac{1}{2}\|\omega\|^2+C\left(v\varepsilon+\frac{1}{N}\sum_{N=1}^{N}(\xi_i+\xi_i^*)\right)-\sum_{i=1}^{N}\beta_i\xi_i-\sum_{i=1}^{N}\beta_i^*\xi_i^*-\gamma\varepsilon + \sum_{i=1}^{N}\alpha_i(\omega^T\Phi(x_i)+b-y_i-\varepsilon-\xi_i)+\sum_{i=1}^{N}\alpha_i^*(y_i-\omega^T\Phi(x_i)-b-\varepsilon-\xi_i^*) \tag{14}$$

where $\alpha_i,\alpha_i^*,\beta_i,\beta_i^*,\gamma\geq 0$ are Lagrange multipliers. The solution of Eq. (14) can be achieved by partially differentiating with respect to ξ_i ω, b, ε, and ξ_i^*

$$\begin{cases}\frac{\partial L}{\partial\omega}=\omega+\sum_{i=1}^{N}\alpha_i x_i-\sum_{i=1}^{N}\alpha_i^* x_i=0\\ \frac{\partial L}{\partial b}=\sum_{i=1}^{N}\alpha_i-\sum_{i=1}^{N}\alpha_i^*=0\\ \frac{\partial L}{\partial\varepsilon}=\frac{C}{N}\sum_{i=1}^{N}\nu-\gamma-\sum_{i=1}^{N}(\alpha_i+\alpha_i^*)=0\\ \frac{\partial L}{\partial\xi}=\sum_{i=1}^{N}\frac{C}{N}-\sum_{i=1}^{N}\alpha_i-\sum_{i=1}^{N}\beta_i=0\\ \frac{\partial L}{\partial\xi^*}=\sum_{i=1}^{N}\frac{C}{N}-\sum_{i=1}^{N}\alpha_i^*-\sum_{i=1}^{N}\beta_i^*=0\end{cases}\Rightarrow\begin{cases}\omega=\sum_{i=1}^{N}(\alpha_i^*-\alpha_i)x_i\\ \sum_{i=1}^{N}(\alpha_i^*-\alpha_i)=0\\ \sum_{i=1}^{N}(\alpha_i^*+\alpha_i)=C\nu-\gamma\leq C\nu\\ \alpha_i=\frac{C}{N}-\beta_i\leq\frac{C}{N}\\ \alpha_i^*=\frac{C}{N}-\beta_i^*\leq\frac{C}{N}\end{cases} \tag{15}$$

Substituting (15) into (14), the Lagrange function can be rewritten as follows:

$$L=-\frac{1}{2}\sum_{i,j=1}^{N}(\alpha_i^*-\alpha_i)\left(\alpha_j^*-\alpha_j\right)K(x_i,x_j)+\sum_{i=1}^{N}(\alpha_i^*-\alpha_i)y_i. \tag{16}$$

According to Karush-Kuhn-Tucker conditions, the optimization problem of the objective function expressed by Eq. (16) can be achieved via the optimal solution of the corresponding dual problem

$$\min_{\alpha,\alpha^*}\quad\frac{1}{2}\sum_{i,j=1}^{N}(\alpha_i^*-\alpha_i)\left(\alpha_j^*-\alpha_j\right)K(x_i,x_j)-\sum_{i=1}^{N}(\alpha_i^*-\alpha_i)y_i \tag{17}$$

subject to constraints

$$\sum_{i=1}^{N}(\alpha_i^*-\alpha_i)=0,$$
$$\sum_{i=1}^{N}(\alpha_i^*+\alpha_i)\leq C\cdot\nu,$$
$$0\leq\alpha_i,\alpha_i^*\leq C/N,\quad i=1,2,\cdots,N,$$

Finally, the optimal solution of the dual problem, i.e., the solution of the original problem, can be achieved by means of SMO algorithm [24, 25],

$$\overline{\omega}=\sum_{i=1}^{N}(\alpha_i^*-\alpha_i)\Phi(x_i), \tag{18}$$

$$\overline{b}=\frac{1}{2}\left[y_s+y_t-\sum_{i=1}^{N}(\alpha_i^*-\alpha_i)K(x_i,x_s)-\sum_{i=1}^{N}(\alpha_i^*-\alpha_i)K(x_i,x_t)\right]. \tag{19}$$

The decision function of the v-SVR model can be expressed by

$$f(x)=\sum_{i=1}^{N}(\alpha_i^*-\alpha_i)K(x_i,x)+b. \tag{20}$$

In order to facilitate the calculation of training process, a modified radial basis function is selected as the kernel function in v-SVR model

$$K(x,x_i)=exp\left(-\tau\cdot\|x-x_i\|^2\right), \tag{21}$$

where τ is the adjustment parameter for optimization.

In order to improve the performance of v-SVR model, some parameters need to be reasonably determined, such as

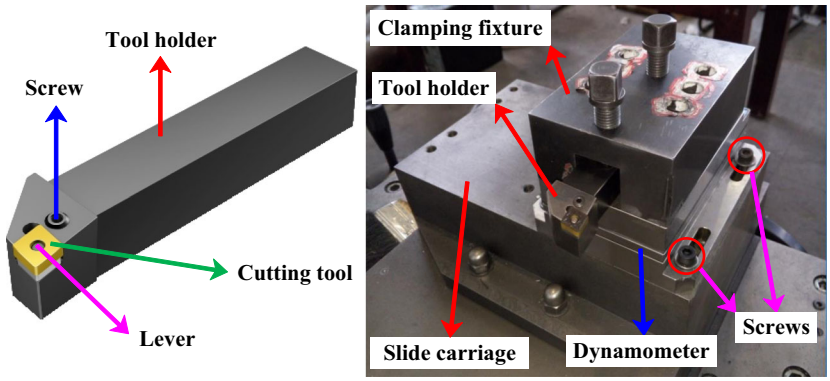


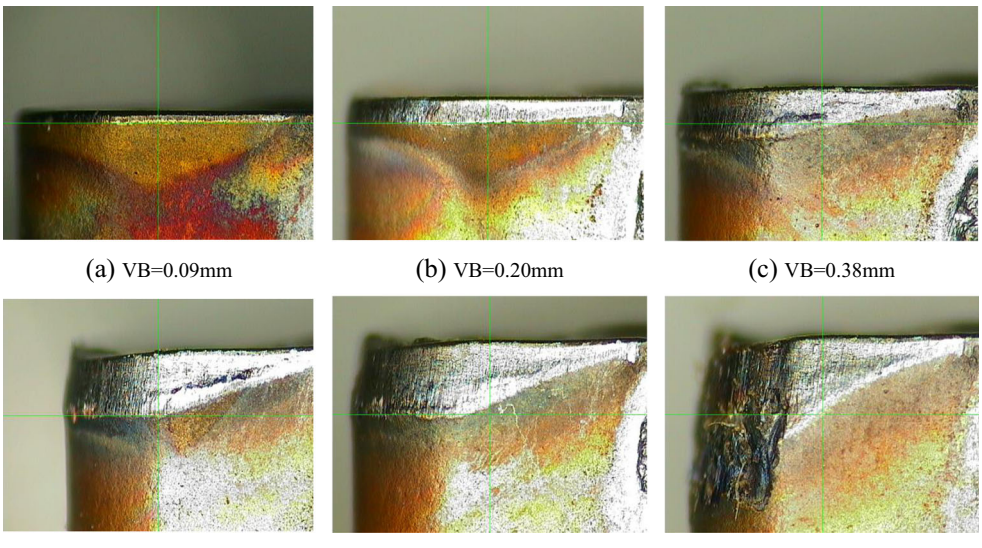
Fig. 6 The installation of tool insert, tool holder, and dynamameter

model parameter v, penalty parameter C, and kernel function parameter τ. The features extracted in Section 3 and actual tool wear value measured by video measuring system can be used to train v-SVR model. The correlation between extracted features and actual tool wear is learned by v-SVR in the training stage. Once the training stage is finished, an optimal v-SVR model that has first-class treatment effect in some certain evaluation criteria is generated and can be used to predict tool wear via input features extracted from KPCA.

5 Experimental results and analysis

This work is mainly focused on the variation of cutting forces under different tool wear values. The emphasis of this study is on making tool inserts go through various wear stages and extracting the desired cutting force signals in a relatively short period of time. Thus, hard turning is conducted for cost reduction given that the presence of cutting fluid will reduce tool wear rate so as to cause tool life extension and mass consumption of workpieces.

All experiments are carried out on an engine lathe which can achieve stepless speed regulation with the combination of a vector converter. The tool insert, fixed by a set of lever and screw, is mounted on the Sandvik turning tool holder which is clamped onto the dynamometer by utilizing a clamping fixture as shown in Fig. 6. Besides, the dynamometer is mounted with four screws on the slide carriage of engine lathe. The selected diamond-shape indexable insert produced by Sandvik has corner radius $R = 0.8$mm, rake angle $\theta = 0°$, and clearance angle $\varnothing = 0°$. The Sandvik tool holder was designed to make the actual installation angle of the indexable insert become rake angle $\theta' = -6°$, clearance angle $\varnothing' = 6°$, tool cutting edge angle $\alpha = 90°$, and tool minor cutting edge angle $\beta = 10°$. The normalizing 50 steel (HB160 ~ 197) cylindrical blocks with diameter 200 mm and length 400 mm are mounted on the spindle via three-jaw chuck and live center and rotate at a varied speed in every cutting process to keep the cutting speed constant. The distance between the dynamometer and cutting



(a) VB=0.09mm (b) VB=0.20mm (c) VB=0.38mm

(d) VB=0.50mm (e) VB=0.558mm (f) Tool breakage

Fig. 7 Wear process of the cutting tool in major flank face ($V_c = 300$mm/min, $a_p = 2$mm, $f = 0.3$mm/r)

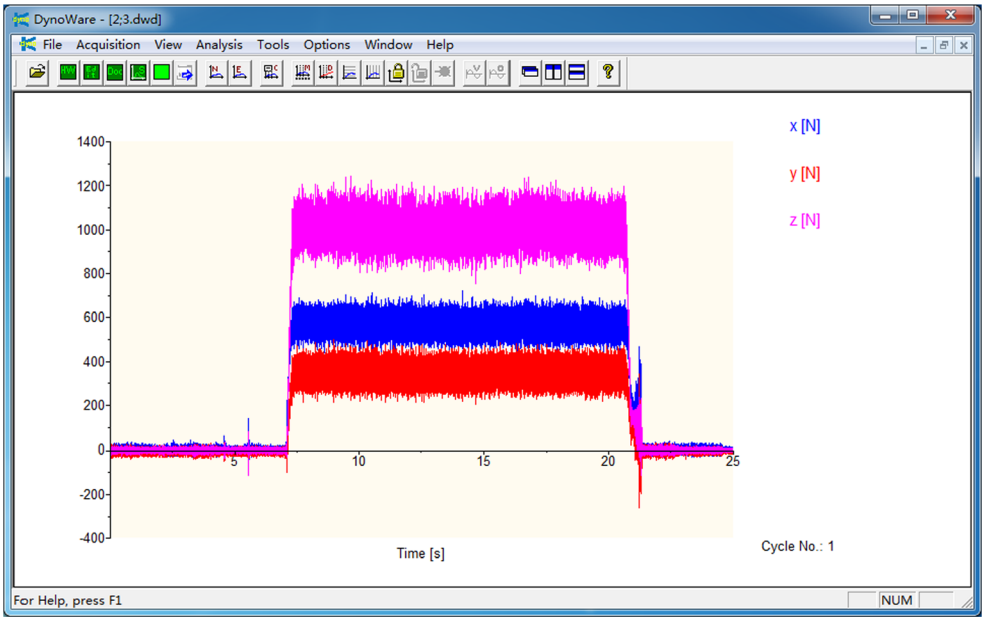


Fig. 8 Graphical user interface of the DynoWare package

position, i.e., the cantilever length of the tool holder, should be as short as possible so as to reduce signal loss and remain constant throughout the experiment to fix the influence of cantilever length on experimental results.

The measured VB in half of the cutting depth on tool flank face is employed as quantitative index to quantify tool wear. The machine tool will have a larger vibration and even chatter when the tool flank wear value exceeds 0.5 mm. Besides, tool breakage may occur suddenly when the flank wear value is more than 0.5 mm. In order to protect the machine tool and workpiece surface, 0.5 mm is used as the threshold to determine whether the tool is worn out or not. In the turning experiment, indexable insert is taken down to measure the actual flank wear by video measuring system once each cutting process is finished. The tool wear process is presented in Fig. 7. The force signals collected via the dynamometer are amplified through charge amplifier and then directly sent to DAQ system. The DynoWare package has been selected as the software to acquire sensory signals from DAQ system in the cutting process. The graphical user interface of the DynoWare is presented in Fig. 8. The measuring time, sampling rate (20 kHz), cycles, and trigger delay time can be set up by the user. Once one sampling process is finished, the acquired data will be exported to a .txt file with filename generated by user settings. Afterwards, the customized program compiled by Matlab will automatically read the new.txt file, extract time, and time-frequency domain signal features related to tool wear and then save the extracted features to build a feature set.

This paper focuses on the variation of static and dynamic cutting forces with tool wear under different cutting parameters. Comprehensive experimental method is employed in this paper. The experimental cutting parameters determined according to the recommended cutting parameters for indexable insert are listed in Table 2. Quite a few dry turning experiments have been carried out to investigate the correlation between sensory signals and actual tool wear. In each set of the trials, the flank wear of the indexable insert will be measured after each cutting process until the insert is broken.

Twenty-one signal features related to tool wear can be obtained via time domain statistical analysis of the acquired cutting force signals. Table 3 shows the correlation coefficients of the obtained signal features with tool wear in terms of VB. The correlation of F_x/F_z and skewness with the variation of VB intuitively reveal the linear dependence between signal features and VB as shown in Figs. 9 and 10. It is evident that skewness should be eliminated due to its poor linear relationship with VB. It is also explained that correlation coefficient

Table 2 Experimental cutting parameters

Test number	Cutting speed V_c (m/min)	Cutting depth a_p (mm)	Feed f (mm/r)
1	300	1	0.2
2	300	1	0.3
3	300	1	0.4
4	300	2	0.2
5	300	2	0.3
6	300	2	0.4
7	350	1	0.2
8	350	1	0.3
9	350	1	0.4

Table 3 Correlation coefficients of the time domain features and VB

Signal features	F_x	F_x/F_z	RMS_F_x	variance_F_x	skewness_F_x	kurtosis_F_x	form factor_F_x
Coefficients	*0.95*	*0.95*	*0.95*	0.004	−0.31	−0.12	*−0.84*
Signal features	F_y	F_y/F_z	RMS_F_y	variance_F_y	skewness_F_y	kurtosis_F_y	form factor_F_y
Coefficients	*0.95*	*0.95*	*0.95*	0.16	−0.09	−0.23	*−0.77*
Signal features	F_z	F_x/F_y	RMS_F_z	variance_F_z	skewness_F_z	kurtosis_F_z	form factor_F_z
Coefficients	*0.90*	0.002	*0.90*	0.15	0.03	−0.13	−0.26

method can effectively select out sensitive features related to tool wear and eliminate non-sensitive features. Therefore, the features corresponding to the correlation coefficients in italic should be selected as sensitive features since that they have an obvious linear relationship with VB. It can be seen that the selected 10 sensitive features change gradually with the increase of VB as shown in Fig. 11. However, the changing trends of $F_i(i=x,y,z)$ and RMS_F_i with the increase of VB are almost identical. In addition, considering that the correlation coefficients of F_i and RMS_F_i are equal with each other, RMS_F_i should be regarded as redundant features with regard to F_i and can be deleted from the 10 sensitive features.

In addition, wavelet multi-resolution analysis (MRA) is employed to get more information related to tool wear in this paper. The cutting force signals- can be decompose into different frequency domains effectively via MRA analysis. With a 5-level MRA analysis of the cutting force signals sampled at 20 kHz, the corresponding frequency domain information is decomposed into different frequency domain ranges as shown in Fig. 12. The energy of each frequency band will change gradually with the tool wear. Ratio of energy of each frequency band to total energy of the signal can be acquired. The energy of each frequency band is represented by the corresponding wavelet coefficients according to the conservation of energy. For simplicity, "Ca" and "Cd" represent the energy ratio of approximation coefficients and detail coefficients, respectively. The correlation coefficients of energy ratio in different frequency bands and tool wear in terms of VB are shown in Table 4. In the table, Cd1_F_x represents the ratio of energy in detail sections to total energy after the first layer wavelet decomposition of F_x, and so on. The features corresponding to the correlation coefficients in italic are sensitive to tool wear and should be selected as sensitive features to predict tool wear. The correlation of the 12 sensitive features relevant to tool wear with the variation of VB is shown in Fig. 13.

The 7 sensitive features selected from time domain and 12 sensitive features selected from MRA analysis, together with three cutting parameters, constitute the 22-dimension feature vector in each sample. There are a total of 120 sample feature vectors acquired from the 9 sets of cutting experiments. The feature set containing sensitive features related to tool wear is shown in Table 5. Only candidate-sensitive features are shown in this paper due to space constraints. It is worth noting that correlation coefficient method cannot effectively identify non-linear features related to tool wear which are extremely abundant in the cutting processes. Thus, mapping the signal features acquired by correlation coefficient method into high-dimensional non-linear space is more helpful to improving identification accuracy of the predictive model as introduced in Section 4. Besides, computational complexity will be substantially increased if the sensitive features are directly sent to

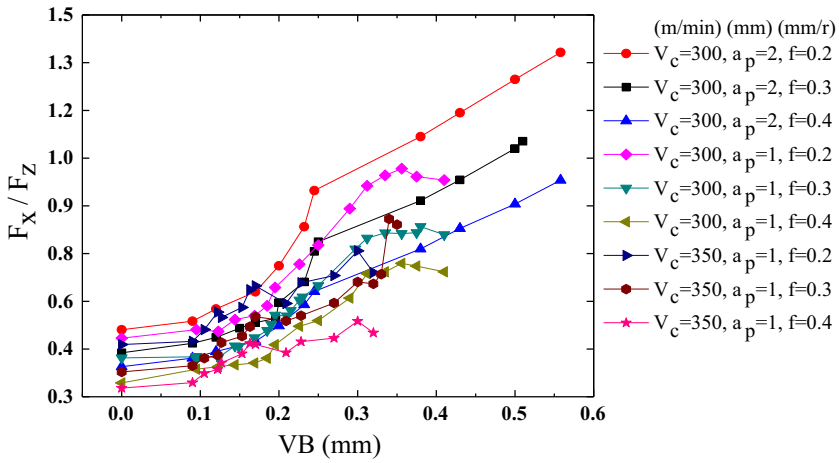


Fig. 9 The correlation of F_x/F_z with the variation of VB

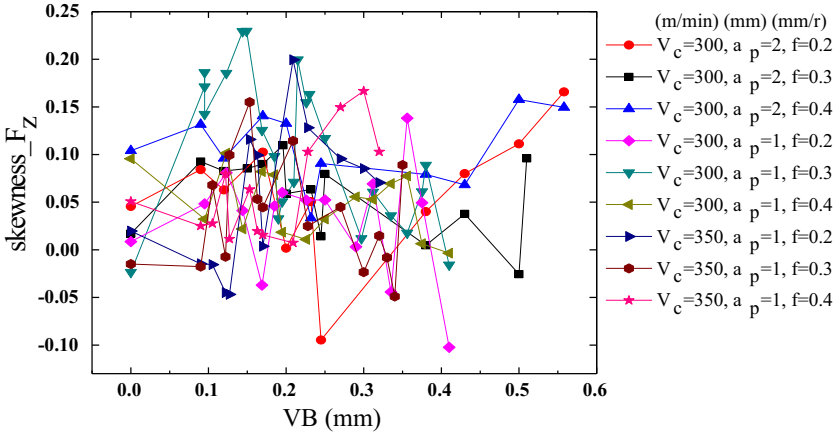
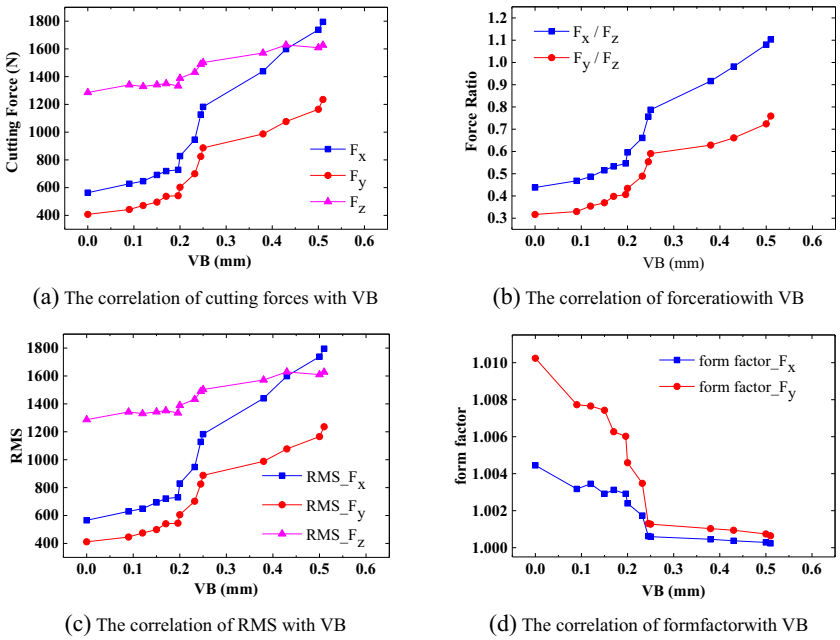


Fig. 10 The correlation of skewness_F_z with the variation of VB



(a) The correlation of cutting forces with VB

(b) The correlation of forceratiowith VB

(c) The correlation of RMS with VB

(d) The correlation of formfactorwith VB

Fig. 11 The correlation of sensitive features with the variation of VB. ($V_c = 300$mm/min, $a_p = 2$mm, $f = 0.3$mm/r)

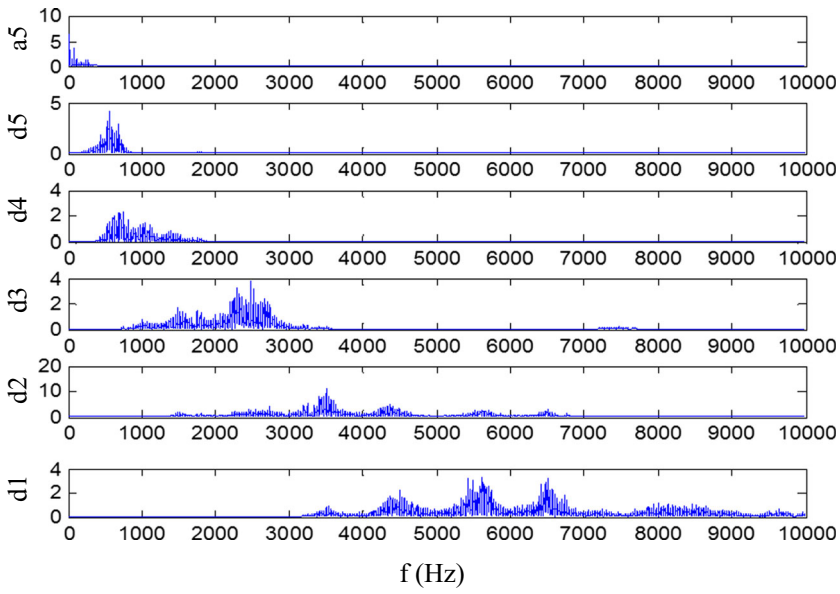
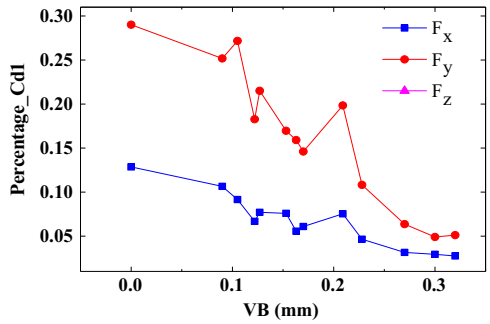


Fig. 12 The frequency graph of 5-level MRA analysis of the cutting force signals. ($V_c = 300$mm/min, $a_p = 2$mm, $f = 0.4$mm/r, VB = 0.232 mm)

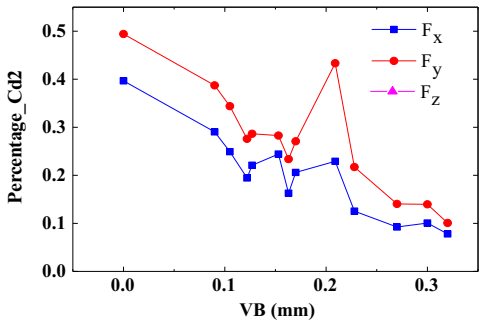
Table 4 Correlation coefficients of the time-frequency domain features and VB

Signal features	Cd1_F_x	Cd1_F_y	Cd1_F_z	Cd2_F_x	Cd2_F_y	Cd2_F_z	Cd3_F_x	Cd3_F_y	Cd3_F_z
Coefficients	*−0.83*	*−0.83*	−0.54	*−0.80*	*−0.77*	−0.33	*−0.81*	*−0.82*	−0.24
Signal features	Cd4_F_x	Cd4_F_y	Cd4_F_z	Cd5_F_x	Cd5_F_y	Cd5_F_z	Ca5_F_x	Ca5_F_y	Ca5_F_z
Coefficients	*−0.85*	*−0.82*	−0.04	*−0.85*	*−0.83*	−0.46	*0.84*	*0.81*	0.40

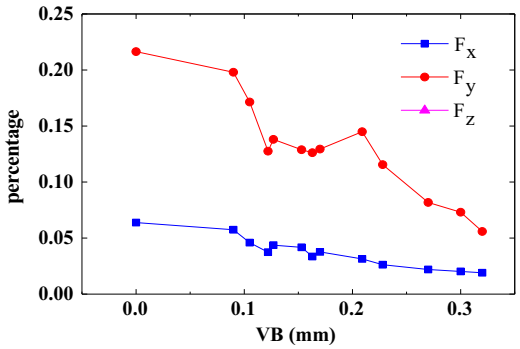
predictive model without dimensionality reduction. All these questions will be solved by KPCA technique as introduced in Section 3.2. The feature set can be randomly divided into two sub-sets, i.e., training sample and test sample. The training



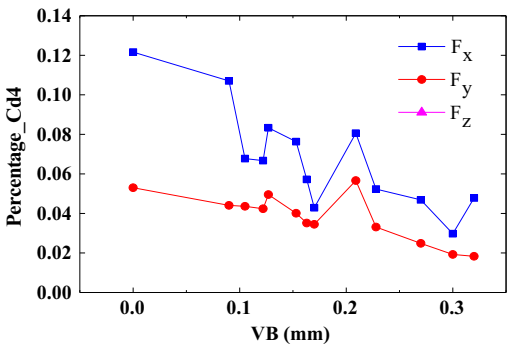
(a) Energy ratio of details in first layer decomposition



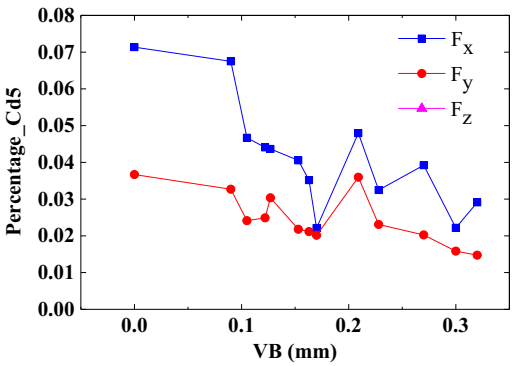
(b) Energy ratio of detailsin second layer decomposition



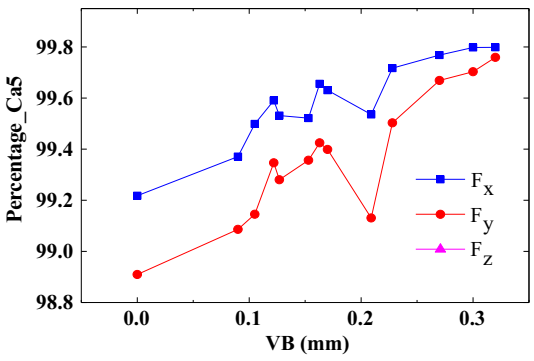
(c) Energy ratio of detailsin third layer decomposition



(d) Energy ratio of detailsin forth layer decomposition



(e) Energy ratio of detailsin fifth layer decomposition



(f) Energy ratio of approximations in fifth layer decomposition

Fig. 13 The correlation of energy ratio in different frequency bands with the variation of VB (V_c = 300mm/min, a_p = 2mm, f = 0.4mm/r)

Table 5 Feature set: total sample features acquired from cutting processes

No.	VB	V_c	f	a_p	F_x	F_y	F_z	F_x/F_z	F_y/F_z	form factor_F_x	form factor_F_y	Cd1_F_x	Cd1_F_y	Cd2_F_x	…
1	0	300	0.3	2	563.3974	407.4174	1285.838	0.438156	0.31685	1.00445	1.010241	0.135519	0.394712	0.533766	…
2	0.09	300	0.3	2	627.6973	442.4423	1340.969	0.468092	0.329942	1.003174	1.007728	0.090986	0.253465	0.361388	…
3	0.12	300	0.3	2	646.7556	471.0351	1329.193	0.486578	0.354377	1.003452	1.007657	0.106832	0.273913	0.406745	…
4	0.15	300	0.3	2	691.5883	495.5315	1341.571	0.515506	0.369366	1.002913	1.007428	0.0827	0.207441	0.33127	…
5	0.17	300	0.3	2	719.411	537.6161	1349.853	0.532955	0.398277	1.003124	1.006266	0.086483	0.184559	0.371878	…
6	0.196	300	0.3	2	728.1585	541.8142	1333.074	0.546225	0.40644	1.002914	1.00602	0.084178	0.179918	0.350858	…
7	0.2	300	0.3	2	827.3809	602.7857	1388.427	0.595912	0.43415	1.002397	1.004593	0.069666	0.137533	0.287355	…
8	0.232	300	0.3	2	945.9244	699.9071	1431.502	0.660792	0.488932	1.001733	1.003478	0.055538	0.096877	0.207702	…
9	0.245	300	0.3	2	1127.039	824.5501	1489.295	0.75676	0.553651	1.000624	1.001304	0.015919	0.030142	0.065674	…
10	0.38	300	0.3	2	1439.203	987.1875	1570.667	0.916301	0.628515	1.000453	1.001031	0.013692	0.027302	0.050276	…
11	0.43	300	0.3	2	1598.55	1076.281	1628.501	0.981608	0.660903	1.000372	1.00094	0.011753	0.026818	0.043198	…
12	0	300	0.2	2	470.3819	277.1479	921.066	0.510693	0.300899	1.001827	1.006313	0.058991	0.207694	0.207402	…
13	0.09	300	0.2	2	522.4769	310.261	972.0051	0.537525	0.319197	1.001491	1.005045	0.050864	0.168312	0.164268	…
14	0.12	300	0.2	2	569.0307	353.8873	987.7723	0.576075	0.358268	1.001244	1.004104	0.042619	0.136742	0.138994	…
15	0.17	300	0.2	2	654.0504	428.2648	1038.54	0.629779	0.412372	1.001063	1.002973	0.032446	0.08905	0.114591	…
16	0.2	300	0.2	2	746.862	490.9908	1049.529	0.711616	0.46782	1.000904	1.00253	0.026659	0.073089	0.096196	…
17	0.232	300	0.2	2	889.0709	597.7792	1065.689	0.834269	0.560932	1.000636	1.001345	0.016558	0.045061	0.063765	…
18	0.245	300	0.2	2	1061.286	701.7991	1119.501	0.947999	0.626885	1.00023	1.000766	0.004726	0.011634	0.017744	…
19	0.38	300	0.2	2	1328.088	850.3037	1188.428	1.117516	0.715486	1.000173	1.002008	0.003745	0.010607	0.015345	…
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
120	0.32	350	0.4	1	481.9003	530.1426	962.0265	0.500922	0.551069	1.007357	1.009448	0.136328	0.136764	0.774603	…

Table 6 Training sample produced by KPCA process

No.	VB	V_c	f	a_p	KPCA1	KPCA2	KPCA3	KPCA4	KPCA5
1	0	300	0.3	2	−0.00324	0.016538	−0.03345	−0.00824	−0.01025
2	0.32	350	0.4	1	−0.01613	0.006723	0.023187	−0.0133	−0.00132
3	0.335	300	0.4	1	0.029464	0.008056	0.02194	0.00667	0.001976
4	0.17	300	0.2	2	0.034232	−0.00497	−0.01245	−0.00468	0.004283
…	…	…	…	…	…	…	…	…	…
60	0.15	300	0.3	2	0.010824	0.016531	−0.02129	−0.00904	0.005836

Table 7 Test sample produced by KPCA process

No.	VB	V_c	f	a_p	KPCA1	KPCA2	KPCA3	KPCA4	KPCA5
1	0.09	300	0.2	2	0.019355	−0.00944	−0.02419	−0.00596	0.00525
2	0	350	0.2	1	−0.00839	−0.03424	−0.01854	−0.01541	−0.01009
3	0.127	350	0.2	1	0.007498	−0.03516	−0.00852	−0.01261	−0.00724
4	0.195	300	0.3	1	−0.00392	−0.0116	0.000214	0.006234	0.009388
…	…	…	…	…	…	…	…	…	…
60	0.123	300	0.2	1	−0.00283	−0.03372	−0.02275	0.012522	0.0203

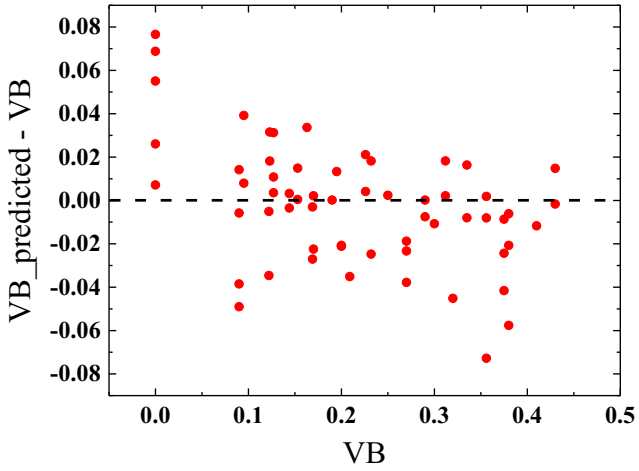


Fig. 14 Absolute error between the predicted and actual measured VB

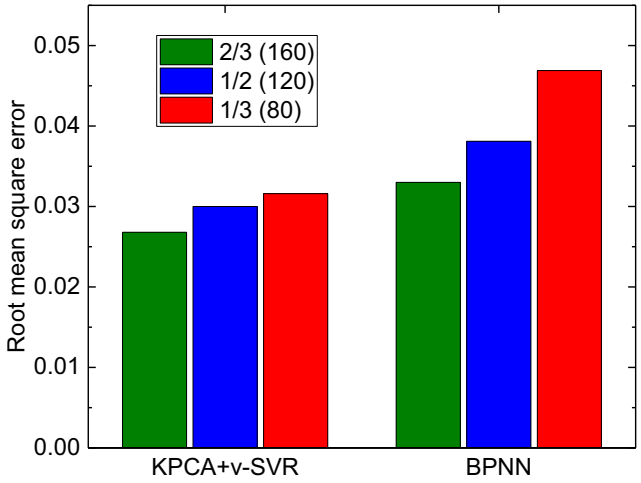


Fig. 16 Root mean square error

sample contains 60 feature vectors, the rest 60 feature vectors make up the test sample. The elements of feature vectors in training sample need to be normalized against their respective mean values and standard deviations before sending to KPCA. And the test sample will be normalized according to Eq. (11). The width parameter σ in kernel function expressed by Eq. (9) is selected as 100. The original 22-dimension feature in each sample is compressed into 5-dimension through the treatment of KPCA technique. The ultimate training sample and test sample produced by KPCA are listed in Tables 6 and 7, respectively. Then, the training sample and the corresponding tool wear measured by video measuring system will be employed to train the SVR-based tool wear predictive model, and the model parameter v is selected as 0.5. In addition, penalty parameter C and adjustment parameter τ are determined by means of grid-search and cross-validation method and ultimately selected as 16 and 64, respectively.

Once the training stage is finished, the $(a_i^*-a_i)$ and b will be acquired and used to construct the SVR-based tool wear model which will be validated by the test sample. The output of the SVR-based tool wear model represents tool wear state in terms of VB. Absolute error between the predicted value of the proposed predictive model and actual measured VB is shown in Fig. 14. Square correlation coefficient between the predicted and actual tool wear is 0.9485. Root mean square error between the predicted and actual tool wear is 0.028. Fortunately, maximum absolute error appears in the point of VB = 0 which does not need to be predicted. Thus, a good agreement between the predicted and actual measured VB can be found at different levels of tool wear. And absolute error between the predicted and actual measured tool wear follows an approximately normal distribution. Finally, the experimental results show that the proposed SVR-based predictive model is proven effective to predict tool wear by using the features extracted from KPCA.

6 Comparison with BPNN model

Neural networks are complicated network systems consisting of large numbers of simple neurons which are connected to each other as illustrated in Fig. 15. The neurons in each layer receive information from several other neurons in the previous layer. All the input information will

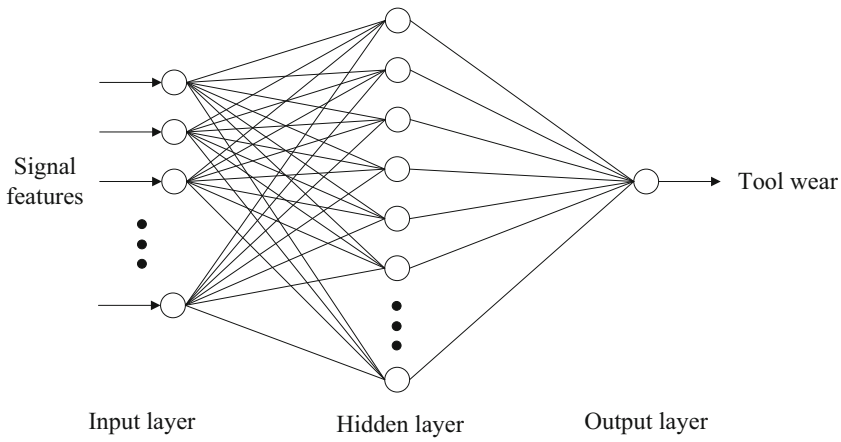


Fig. 15 Structure of the BPNN model for tool condition monitoring

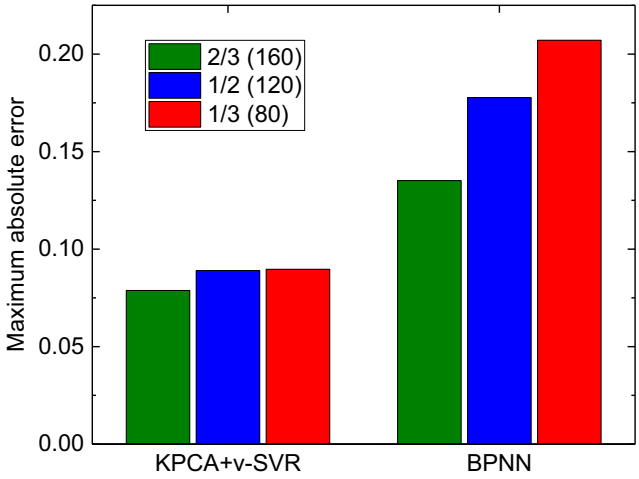


Fig. 17 Maximum absolute error

pass through artificial neuron models which are composed of weight, bias, and activation function. The produced neuron output will be transmitted to the next layer. The back-propagation neural network (BPNN), which consists of forward calculation and error back propagation processes, is one of the most widely used neural network models. When desired outputs are not appeared in the output layer during training process, the error signal will return along the original path and modify the weights of each layer so as to minimize itself until the expectation error is reached.

Network structure is determined as following. Fifteen nodes are selected in the hidden layer which is determined by trail. The bipolar sigmoid and sigmoid activation functions are selected for the hidden layer and output layer, respectively. The training algorithm is selected as variable learning rate momentum gradient descent algorithm in which the learning rate is set as 0.05. The training target error and maximum training times are selected as 0.001 and 10,000, respectively.

Note that sample features need to be normalized to cancel the difference in order of magnitude between each dimension data so as to avoid larger network prediction error. Data normalization method is as follows:

$$x_i = \frac{x_i - \max(x_i)}{\max(x_i) - \min(x_i)}, i = 1, 2, 3 \ldots N \tag{22}$$

where N is the dimension of sample features, $x_i \in \boldsymbol{R}^n$.

A total of 240 sample feature vectors, which consist of selected sensitive features from Section 3.1 and three cutting parameters, are acquired from the 9 sets of cutting experiments. The feature set is randomly divided into two groups, i.e., training sample and test sample. Training samples of different sizes are utilized to train the proposed model and BPNN-based predictive model. The rest of total sample are utilized to test the performance of the trained model, respectively. Prediction results of the two models under different training samples are shown in Figs. 16 and 17, respectively. 2/3, 1/2, and 1/3 represent the ratio of training sample size to total sample. Contents in brackets are the number of feature vectors in training sample.

The root mean square error (RMSE) of the proposed model is still less than 0.032 with the significant decrease of training sample size. The maximum absolute error (MAE) of the proposed model is less than 0.09 even in less-scale training sample. However, the RMSE and MAE of the BPNN-based model increase rapidly with the reduction of training sample size. The comparison result shows that the proposed model is more suitable for tool wear monitoring. It is obvious that SVM has better generalization ability even in small sample size. Analysis of the prediction results reveals that the proposed model by the combination of KPCA and v-SVR still has better prediction accuracy even if the size of training sample is sharply reduced.

7 Conclusions

This paper presents a tool wear predictive model based on the combination of KPCA and v-SVR. To test the effectiveness of the proposed model, 7 time domain features and 12 wavelet domain features are selected as discriminative features to characterize different tool wear stages. Firstly, KPCA technique is successfully applied to extract features from multi-dimensional signals acquired from machining processes. Compared with PCA, KPCA technique can map data samples into a high-dimensional space and further extract sensitive features related to tool wear. Secondly, v-SVR has been proved effective to predict tool wear by using the features extracted from KPCA, and the predictive accuracy is much beyond expectation. There is a good agreement between the predicted value of the proposed model and the actual tool wear measured by video measuring system. Besides, the proposed model has better generalization ability even in small samples.

Acknowledgments The authors are grateful to the financial sponsorship from 863 National High-Tech Research and Development Program of China (no. 2013AA041108).

References

1. Jurkovic J, Korosec M, Kopac J (2005) New approach in tool wear measuring technique using CCD vision system. Int J Mach Tools Manuf 45:1023–1030
2. Castejón M, Alegre E, Barreiro J, Hernández LK (2007) On-line tool wear monitoring using geometric descriptors from digital images. Int J Mach Tools Manuf 47:1847–1853
3. Dimla DE Sr, Lister PM (2000) On-line metal cutting tool condition monitoring. I: force and vibration analyses. Int J Mach Tools Manuf 40:739–768

4. Salgado DR, Alonso FJ (2007) An approach based on current and sound signals for in-process tool wear monitoring. Int J Mach Tools Manuf 47:2140–2152
5. Choudhury SK, Rath S (2000) In-process tool wear estimation in milling using cutting force model. J Mater Process Technol 99(1): 113–119
6. Ghosh N, Ravi YB, Patra A, Mukhopadhyay S, Paul S, Mohanty AR, et al. (2007) Estimation of tool wear during CNC milling using neural network-based sensor fusion. Mech Syst Signal Process 21: 466–479
7. Dong J, Subrahmanyam KVR, Wong YS, Hong GS, Mohanty AR (2005) Bayesian-inference-based neural networks for tool wear estimation. Int J Adv Manuf Technol 30:797–807
8. Eiji S, Rei H, Masayuki U, et al. (2003) Intelligent recognition of end milling conditions based on cutting force model. Trans Jpn Soc Mech Eng C 69(7):1927–1932
9. Kilundu B, Dehombreux P, Chiementin X (2011) Tool wear monitoring by machine learning techniques and singular spectrum analysis. Mech Syst Signal Process 25:400–415
10. Kima JS, Kangb MC, Ryuc BJ, Jic YK (1999) Development of an on-line tool-life monitoring system using acoustic emission signals in gear shaping. Int J Mach Tools Manuf 39:1761–1777
11. Dimla DE Sr (2000) Sensor signals for tool-wear monitoring in metal cutting operations—a review of methods. Int J Mach Tools Manuf 40:1073–1098
12. Kaya B, Oysu C, Ertunc HM (2011) Force-torque based on-line tool wear estimation system for CNC milling of Inconel 718 using neural networks. Adv Eng Softw 42:76–84
13. Fish RK, Ostendorf M, Bernard GD, Castanon DA (2003) Multilevel classification of milling tool wear with confidence estimation. IEEE Trans Pattern Anal Mach Intell 25:75–85
14. Tobon-Mejia DA, Medjaher K, Zerhouni N (2012) CNC machine tool's wear diagnostic and prognostic by using dynamic Bayesian networks. Mech Syst Signal Process 28:167–182
15. Dimla DE Sr, Lister PM (2000) On-line metal cutting tool condition monitoring. II: tool-state classification using multi-layer perceptron neural networks. Int J Mach Tools Manuf 40:769–781
16. Chungchoo C, Saini D (2002) On-line tool wear estimation in CNC turning operations using fuzzy neural network model. Int J Mach Tools Manuf 42:29–40
17. Kious M, Ouahabi A, Boudraa M, Serra R, Cheknane A (2010) Detection process approach of tool wear in high speed milling. Measurement 43:1439–1446
18. Kamarthi SV, Pittner S (1997) Fourier and wavelet transform for flank wear estimation—a comparison. Mech Syst Signal Process 11:791–809
19. Shi D, Gindy NN (2007) Tool wear predictive model based on least squares support vector machines. Mech Syst Signal Process 21: 1799–1814
20. Scheffer C, Heyns PS (2004) An industrial tool wear monitoring system for interrupted turning. Mech Syst Signal Process 18:1219–1242
21. Shi D, Tsung F (2003) Modelling and diagnosis of feedback-controlled processes using dynamic PCA and neural networks. Int J Prod Res 41:365–379
22. Schölkopf B, Smola A, Müller K (1998) Nonlinear component analysis as a kernel eigenvalue problem. Neural Comput 10(5): 1299–1319
23. Schölkopf B, Smola AJ, Williamson RC, Bartlett PL (2000) New support vector algorithms. Neural Comput 12(5):1207–1245
24. J C Platt (1998) Sequential minimal optimization: a fast algorithm for training support vector machines. Microsoft Research, Technical Report MSR-TR-98-14
25. Platt JC (1999) Fast training of support vector machines using sequential minimal optimization," advances in kernel methods: support vector learning. MIT Press, Cambridge