OPEN ACCESS

Prediction of Machine Tool Condition Using Support Vector Machine

To cite this article: Peigong Wang et al 2011 J. Phys.: Conf. Ser. 305 012113

View the article online for updates and enhancements.

You may also like

 Research on Prediction Model of Rock and Soil Layer Information Based on Adjacent Boreholes

Xiang Li, Dingli Su, Jiagao Zhong et al.

- Adaptive prediction of respiratory motion for motion compensation radiotherapy Qing Ren, Seiko Nishioka, Hiroki Shirato et al.
- <u>A multiple model approach to respiratory</u> motion prediction for real-time IGRT Devi Putra, Olivier C L Haas, John A Mills et al.





DISCOVER how sustainability intersects with electrochemistry & solid state science research



This content was downloaded from IP address 3.147.84.71 on 21/05/2024 at 13:17

Prediction of Machine Tool Condition Using Support Vector Machine

Peigong Wang, Qingfeng Meng, Jian Zhao, Junjie Li, Xiufeng Wang

Theory of Lubrication and Bearing Institute , Xi'an Jiaotong University , Xi'an 710049 , China

E-mail:senor@sina.com

Abstract. Condition monitoring and predicting of CNC machine tools are investigated in this paper. Considering the CNC machine tools are often small numbers of samples, a condition predicting method for CNC machine tools based on support vector machines (SVMs) is proposed, then one-step and multi-step condition prediction models are constructed. The support vector machines prediction models are used to predict the trends of working condition of a certain type of CNC worm wheel and gear grinding machine by applying sequence data of vibration signal, which is collected during machine processing. And the relationship between different eigenvalue in CNC vibration signal and machining quality is discussed. The test result shows that the trend of surface roughness value. In trends prediction of working condition, support vector machine has higher prediction accuracy both in the short term ('One-step') and long term (multi-step) prediction compared to autoregressive (AR) model and the RBF neural network. Experimental results show that it is feasible to apply support vector machine to CNC machine tool condition prediction.

1. Introduction

It is a key technology to evaluate and predict the working condition of CNC machine tools for the purpose of ensuring reliable operation of machine tools and improving the performance of machine tools. The CNC machine tools are on a direction of high speed, high accuracy, heavy-load and compound machining, and the mechanical structure is becoming more and more complex that it is difficult to detect potential failures early. Without timely diagnosis and early warning, early failures can cause machine tools to work under worse condition which will result in increasing of rejection rate, fluctuation in the quality and declination of productivity. For the purpose of reducing the defective rate and maintenance costs, it is very important to determine accurately and timely whether there are changes in machine running condition, forecast the development trend of machine running condition, evaluate the working condition of machine tools. Due to abundant working condition information existent in vibration signals of mechanical equipment ^[11], in this paper the time sequence of vibration signals in the working process is picked to predict the trends of machining condition and quality.

Traditional time sequence analysis technique and prediction theory are mainly based on linear autoregressive (AR) model and linear autoregressive moving average model (ARMA), which can obtain better prediction results on the linear system rather than nonlinear system. In recent years, neural network model was successfully applied to non-stationary time sequence prediction^[1], but it can not be widely used because of achieving empirical risk minimization principle only.

Support vector machine (SVM)^[2] is a new machine learning method based on statistic learning theory, which uses new learning mechanism to realize structural risk minimization principle and is suitable for solving the problems such as nonlinear, high dimension and local minimum. Support vector machine can ensure higher accuracy for a long-term prediction compared with traditional regression techniques in many practical applications^[3,4]. Therefore, forecasting model based on

support vector machine becomes one of the research hotspots in artificial intelligence field and has been widely used^[3,4].

2. Theoretical Formulation

SVM was initially developed by Vapnik in 1995^[2]. It is based on linearly separable optimal separating hyperplane and is designed for classification. It can be applied in function approximation problems and also has good performance in regression problems. The core thought_of SVM is the introduction of nonlinear mapping function, which maps original model into high-dimensional space and constructs optimal separating hyper plane in the feature space. It converts a nonlinear problem of low dimension space into a linear one of high dimension space by using kernel function to realize the classification.

2.1. Basic theory of SVM for function regression^[2,5,6]

There are linear regression and nonlinear regression in SVM. Let us consider the following linear regression function. The goal is to estimate unknown real-valued function in the relationship:

$$y = f(\mathbf{x}) = \boldsymbol{\omega} \cdot \mathbf{x} + b \tag{1}$$

where ω is weight vector, b is the 'bias' term, **x** is a multivariate input, and y is a scalar output. Introducing (non-negative) slack variables ξ_i and ξ_i^* , the optimization problem can be expressed as

Minimize
$$\Phi(\omega) = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*), \ C \ge 0$$

Subject to

$$\begin{cases} y_i - \omega \cdot \mathbf{x}_i - b \le \varepsilon + \xi_i, \\ \omega \cdot \mathbf{x}_i + b - y_i \le \varepsilon + \xi_i^*, \quad i = 1, 2, \cdots, l \\ \xi_i, \xi_i^* \ge 0, \end{cases}$$
(2)

where C is a positive constant (regularization parameter), and ε is loss function.

$$f(\mathbf{x}) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(\mathbf{x}_i \cdot \mathbf{x}) + b$$

Lagrange equation is :

$$L(\omega, b, \xi, \xi^{*}) = \frac{1}{2} \|\omega\|^{2} + C \sum_{i=1}^{n} (\xi_{i} + \xi_{i}^{*}) - \sum_{i=1}^{l} \alpha_{i} (\varepsilon_{i} + \xi_{i} - y_{i} + \omega \cdot \mathbf{x}_{i} + b) - \sum_{i=1}^{l} \alpha_{i}^{*} (y_{i} + \varepsilon_{i} + \xi_{i} - \omega \cdot \mathbf{x}_{i} - b) - \sum_{i=1}^{n} (\eta_{i} \xi_{i} + \eta_{i}^{*} \xi_{i}^{*})$$
(4)

 α_i , α_i^* , η_i and η_i^* are Lagrange multiplier.

Dual problem is:

maximize
$$Q(\alpha) = \sum_{i=1}^{n} y_i(\alpha_i - \alpha_i^*) - \varepsilon \sum_{i=1}^{n} (\alpha_i + \alpha_i^*) - \frac{1}{2} \sum_{i,j=1}^{n} (\alpha_i - \alpha_i^*)(\alpha_i - \alpha_i^*)(\mathbf{x}_i \cdot \mathbf{x}_j)$$
 (5)
Subject to

Subject to

$$\begin{cases} \sum_{i=1}^{n} y_{i}(\alpha_{i} - \alpha_{i}^{*}) = 0, \\ 0 \le \alpha_{i} \le 0, \\ 0 \le \alpha_{i}^{*} \le 0, \end{cases}$$
(6)

Regression function is:

$$f(\mathbf{x}) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) (\mathbf{x}_i \cdot \mathbf{x}) + b$$
(7)

Nonlinear regression function is:

$$f(\mathbf{x}) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(\mathbf{x}_i \cdot \mathbf{x}) + b$$
(8)

2.2. Least Square Support Vector Machine (LS-SVM)

The complexity of the SVM algorithm proposed by Vapnik and others is relevant to the number of training samples, the larger the number is, the more complex the corresponding quadratic programming problem is. Low calculating speed can not meet the demand of system dynamic modeling and identification. In order to solve such problems, J.A.K.Suykens proposed an improved algorithm, called Least Square Support Vector Machine (LS-SVM) in 1999^[7], which replaces non-sensitive loss function with quadratic loss function. Quadratic optimization is changed into linear equation by constructing the loss function, so the traditional SVM's quadratic programming problem is conducted by solving a series of linear equations, in which the least square system is used as loss function and inequality constraints in SVM are replaced by equality constraints^{[2, 4].}

On a set of input samples, LS-SVM uses nonlinear mapping $\psi(\mathbf{x})$ to convert the non-linear training data into high dimensional feature space, by which non-linear function estimation problem changes into linear function estimation problem. Now set regression function as $y = \langle \omega \cdot \psi(\mathbf{x}) \rangle + b$. In LS-SVM regression estimation, regression problem is:

Least Square Support Vector Machine (LS-SVM)

$$\min \frac{1}{2} \|\omega\|^{2} + \gamma \frac{1}{2} \sum_{i=1}^{N} e_{i}^{2}$$

s.t $\omega^{T} \psi(s_{i}) + b + e_{i} = y_{i}, \quad i = 1, 2, \dots, N$ (9)

where *e* is error vector and γ is regularization parameter. Introducing Lagrange multiplier λ , $\lambda \in R^{N \times 1}$, (9) can be changed into:

$$\min J_{lssvm} = \min \frac{1}{2} \|\omega\|^2 + \gamma \frac{1}{2} \sum_{i=1}^{N} e_i^2 - \sum_{i=1}^{N} \lambda_i [y_i(\omega^T \psi(\mathbf{x}_i) + b) + e - y_i]$$
(10)

Nonlinear regression estimation function can be expressed as:

$$y = \sum_{i=1}^{N} \lambda_i k(\mathbf{x}_i, \mathbf{x}) + b$$

2.3. Prediction method based on LS- SVM

Time sequence $\{x_t\} \ 1 \le t \le N$ can be predicted by SVM. We take $r \ (1 \le r \le N)$ samples from the whole time sequence as training samples, and the rest as testing samples. For more efficient use of limited data, it is reconstructed, that is to say, transform one-dimensional sequence into a matrix form so that larger amount of information is obtained.

$$\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{r-1} \end{bmatrix} \implies \begin{bmatrix} x_1 & x_2 & \cdots & x_m \\ x_2 & x_3 & \cdots & x_{m+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{r-m} & x_{r-m+1} & \cdots & x_{r-1} \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} x_{m+1} \\ x_{m+2} \\ \vdots \\ x_r \end{bmatrix}$$
(11)

Where m is embedding dimension, maps can be created as: $f : \mathbb{R}^m \to \mathbb{R}$ One step ahead predictive value is expressed as:

$$x_{l+1} = f(x_l, x_{l-1}, \cdots, x_{l-m+1})$$
(12)

Learning samples can be trained by the regression function as following:

$$y_i = \sum_{i=1}^{r-m} (\alpha_i - \alpha_i^*) K(\mathbf{x}_i, \mathbf{x}_i) + b$$
(13)

Here, t=m+1,...,r, $K(\mathbf{x}_i,\mathbf{x}_i)$ is kernel function.

Then the first step prediction is:

$$\hat{x}_{r+2} = \sum_{i=1}^{r-m} (\alpha_i - \alpha_i^*) K(\mathbf{x}_i, \mathbf{x}_{r-m+2}) + b$$
(14)

Here, $\mathbf{x}_{r-m+2} = \{x_{r-m+l}, \cdots, x_{r+1}, \cdots, x_{r+l-1}\}$

3. Experimental details

3.1. Vibration data acquisition



The working condition of a worm wheel grinding CNC machine tools (figure 1) is monitored. After the preparatory work, it begins grinding and collecting the vibration signals at the same time. The vibration signals are stoped collectting and the date are saved after the the grinding is finished.During machine processing of all 8 gears, the vibration signals for each gear are collected. Figure 2 shows the sensors arrangement, accelerometers were mounted on the cover of the grinding wheel spindle by magnetic seat . 0# channel connects to B-axis horizontal (normal direction of grinding surface), 1# channel connects to B-axis vertical, sensitivity is 500mv/g. Sampling frequency is 20000Hz. B-axis rotate speed is 3000r/min in the processing, C-axis rotate speed is 128 r/min. When the grinding is finished, surface roughness Ra (GB1031-1968) of each gear is measured by profile testing instrument as figure 3 shows , randomly selected interval of 180 degrees of two tooth face measured and probe the movement direction of the wheel processing grinding direction perpendicular. The Ra values are listed in table 1.

Ra(um)	0.72	0.52	0.92	1.08	1.32	1.16	0.74	0.82

3.2. Characteristic values

To describe the system condition, the character of the signal should be analyzed, the purpose of characteristic analysis is to transform the original signal into characteristic values and find out the relationship between characteristic values and the system condition. On the basis of characteristic analysis, characteristic values which have good regularity and are sensitive to condition change can be used as pattern vector of fault diagnosis^[8].

The characteristic value is extracted from the vibration signals which are Mentioned in 3.1, such as:

Mean Value,	$Mean = \frac{1}{N} \sum_{i=1}^{N} x_i$
Standard deviation Value,	$Std = \sqrt{\frac{\sum_{i=1}^{n} (X_i - X)^2}{\frac{n-1}{n-1}}}$
Skewness Value,	Skewness = $\frac{1}{N} \sum_{i=1}^{N} x_i^3$
Kurtosis Value,	kurtosis= $\frac{1}{N} \sum_{i=1}^{N} x_i^4$
Peak-to-peak Value,	Pp = max(x) - min(x)
Xr Value,	$Xr = \ [rac{1}{N} \sum_{i=1}^N \sqrt{ig x_iig }]^2$;
Xmean Value,	Xmean= $\frac{1}{N}\sum_{i=1}^{N} x_i $;
Xrms Value,	$Xrms = \sqrt{\frac{1}{N}\sum_{i=1}^{N}x_{i}^{2}};$
Xp Value,	Xp = max(max(x), -min(x)).

In order to eliminate accidental factors, the vibration signals are divided into equal length, each as a sub-sample. For each sub-sample, the average value of characteristic value is calculated, such as \overline{Mean} , \overline{Std} , $\overline{Skewness}$, $\overline{Kurtosis}$, \overline{Pp} , \overline{Xr} , \overline{Xmean} , \overline{Xrms} , and \overline{Xp} . These average value trends and surface roughness trends are compared in Fig2, from which we find that the trend of B-axis horizontal vibration signal Peak-to-peak value fits well with gear surface roughness trend. Thus B-axis horizontal vibration signal Peak-to-peak value is used as characteristic value of processing quality condition prediction in grinding machine.



4. Results and discussion

In the paper Peak-to-peak values are extracted every 2 seconds from the vibration signal in grinding machining process, which form a signal variable time sequence to predict the working condition. For given time series t1, t2, \cdots , tn, time sequence prediction is to estimate the value of n+k time based on the observation value of historical time sequence, which is to find out the relationship between future time value and historical observation value, k is the prediction step, when k=1, it's One-step prediction; when k>1, it's Multi-step prediction. In this paper, the Mean Absolute Percentage Error (MAPE) is applied to asses the prediction ability.

$$MAPE = \frac{100}{N} \sum_{i=1}^{N} \left| \frac{x_i - x_i}{x_i} \right|$$

4.1. The choice of kernel function

The Common kernel function:

1) Linear kernel function: $k(\mathbf{x}, \mathbf{x}') = \mathbf{x} \cdot \mathbf{x}'$;

2) Polynomial kernel function: $k(\mathbf{x}, \mathbf{x}') = (\mathbf{x} \cdot \mathbf{x}' + 1)^d$;

3) Gauss Radial Basis Function (RBF): $k(\mathbf{x}, \mathbf{x}') = \exp(-\|\mathbf{x} - \mathbf{x}'\|^2 / \sigma^2)$

4) Sigmoid function: $k(\mathbf{x}, \mathbf{x}') = \tanh [v (\mathbf{x} \cdot \mathbf{x}_i) + c]$

For a specific problem, which kernel function to choose has not been determined, the general rule is to select by experience, cross-validation method can also be used to select. The feature space of RBF kernel function is infinite dimensional and limited samples in the feature space is linearly separable, so it's a widely used kernel function. The RBF kernel function has little effect on short-term prediction of SVMs, but it shows better ability in long-term prediction^[3]. In the paper RBF kernel function is introduced to make a prediction.

4.2. The choice of LS-SVM parameters and Embedding dimension

The main parameters are the regularization parameter γ and the kernel function width σ^2 for LS-SVM using RBF, these two parameters mainly determine the learning and generalization ability of LS-SVM^[9]. Literature ^[10] investigated practical selection of hyper-parameters for support vector machines (SVM) regression. For the purpose of automatic optimization of parameters in a wide range, Cross-Validation method^[11] and Gridchearch method are chosen in the paper, which are used commonly in model parameters selection. In addition, the embedding dimension m in 'Multi-step' and 'One-step' prediction has a great impact on the prediction accuracy. Figure4 shows the Mean Absolute Percentage Error (MAPE) when m takes different values. γ and σ^2 are the optimal value when chosen by Cross-Validation method and Gridchearch method. Table 2 lists the embedding dimension m, γ and σ^2 , which makes the MAPE minimum in Multi-step'and'One-step' prediction.



 Table 2.
 LS-SVM parameter selection

LS-SVM	γ	σ^{2}	т	MAPE/ %	
parameter					
One-step	5.3425	99.9862	34	4.5825	
Nine-step	16669.1051	1.0364	8	4.9389	

4.3. LS-SVM prediction results

In table 2, m, γ and σ^2 are chosen to predict Peak-to-peak Value sequence of vibration signals in the grinding process. In this sequence, the fist 100 consecutive points are selected for regression training samples, while the rest 59 points are used for the test samples. The prediction results of 'One-step' and 'Nine-step' are respectively shown in Figure5 and Figure6.



4.4. LS-SVM compared with autoregressive (AR) model and the RBF neural network **Table 3** shows the MAPE comparison between LS-SVM and both the autoregressive (AR) model and the RBF neural network in 'One-step' and 'Nine-step' prediction.

Prediction model	One sten	Ning stan	
RBF neural network (MAPE/%)			
Table 3. LS-SVM compared v	th autoregressive	e (AR) model	and the

Prediction model	One-step	Nine-step	
autoregressive (AR)	9.0301	12.0335	
RBF neural network	6.944	8.1897	
LS-SVM	4.5825	4.9389	

5. Conclusions

This paper discusses the relationship between different eigenvalue in CNC vibration signal and machining quality, it also refers to sensitive direction error. In trends prediction of working condition, SVMs is applied to establish 'One-step' and 'Multi-step' prediction model, which makes prediction for grinding machining condition. The test result shows that the vibration signal Pp value in surface normal direction is most relevant to surface roughness value, in addition, SVMs can make a better result both in the short term ('One-step') and long term (multi-step) prediction. SVMs shows better prediction ability in the short term ('One-step') and long term (multi-step) prediction compared with neural network prediction model. Therefore, when the vibration signal Pp value in surface normal direction is used as eigenvalue, it's feasible to monitor and predict CNC machining condition by SVMs.

Acknowledgments

This research work is financially supported by the National Science and Technology Major Special Project of China(No. 2009ZX04014-015). The authors would like to thank the reviewers for their valuable comments on the paper.

References

- [1] Zhang X N. Research on large rotary machinery operation state detection and prediction. PhD thesis. Xi 'an Jiaotong University, Xi'an, China,1998.
- [2] Vapnik V N. The Nature of Statistical Learning theory. [M]. New York: Springer-Verlag. 1995.
- [3] Yang J Y, Zhang Y Y, Zhao R Z. Application of Support Vector Machines in Trend Prediction of Vibration Signal of Mechanical Equipment[J]. Journal of xi'an jiaotong university, 2005,39 (9): 950–953
- [4] Chen B J. Load forecasting using support vector machines: A study on EUN ITE competition 2001 [R]. Taibei: National Taiwan University,2002: 1 11.
- [5] Gunn S. Support vectormachines for classification and regression. ISIS Technical Report, Southamp ton: University of Southamp ton,1998.
- [6] Burges C. A tutorial on support vector machines for pattern recognition. DataMining and Knowledge Discovery (S1384—5810), 1998;2 (2) : 121—167.
- [7] Suykens, J. A. K., & Vandewalle, J. (1999). Least squares support vector machine classifiers. Neural Processing Letters, 9(3), 293–300.
- [8] Li Y. Research on Kernel Based Fault Recognition and Condition Forecasting for Mechanical Power and Transmission Systems . PhD thesis. National University of Defense Technology. Changsha, China ,2007.
- [9] Xiang Z, Zhang T Y, Sun J C.Modelling of nonlinear systems based on recurrent leastsquares support vector machines[J]. Journal of System Simulation, 2006, 18(9): 2684-2687.
- [10] Cherkassky V, Ma Y Q. Practical selection of SVM parameters and noise estimation for SVM regression. Neural Networks 17 (2004) 113–126.
- [11] Michael W. Browne, Cross-Validation Methods [J]. Journal of Mathematical Psychology, 2000. 44: 108-132