Tool Condition Monitoring Based on Second Generation Wavelet Transformation and Hyper-sphere Support Vector Machine

Xiao-qing ZHANG, Ying-chuan WU and Wei-xiong LIU
China Aerodynamics Research and Development Center, Mianyang, China

Keywords: Condition monitoring, Hyper-sphere, SVM, Second generation wavelet transformation, Tool.

Abstract. A method of tool condition monitoring based on the second generation wavelet transformation and hyper-sphere support vector machine was proposed to improve classifying precision of tool condition in the process of machining. Cutting force signal and vibration signal were filtered by the second generation wavelet transformation. Lots of features were extracted by wavelet packet transformation in the time domain and frequency domain analysis. The sensitive features to tool condition as the average value, the mean square root, wavelet coefficient were selected by principal component analysis. The feature vector made up by the features was inputted to hyper-sphere support vector machine, which built the relation between tool condition and features to predicting tool condition automatically. The experimental results show hyper-sphere SVM are of excellent study ability, generalization ability, and of high recognized precision with small training samples.

Introduction

Tool condition online monitoring is the key to automatic and unmanned machining process, operator can change or sharpen worn tool in advance through tool condition monitoring, to ensure the machining accuracy and surface smoothness. Methods of tool condition monitoring are divided into 2 classes: direct monitoring and indirect monitoring [1]. Due to the coolant liquid and the motion of cuttings, accuracy cannot be guaranteed and difficulty is high in direct methods. Through the collection of cutting force signal, vibration signal, acoustical signal, temperature signal and current signal, extracting sensitive features by time-frequency domain analysis, and using theory of artificial neural network, fuzzy recognition, fuzzy clustering and support vector machine to set up relationship between tool condition and signal feature, indirect method can classify the condition accurately [2]-[6]. In the process of machining, inhomogeneity of material caused the change of amplitude under the same condition, so that the feature appears to be instable. In addition, system built by cutting tool, machine tool and work piece is highly nonlinear, so that the system parameters is difficult to determine, this makes tool condition system hard to meet the requirements of industrial application. A large number of studies have shown that the key technic of tool condition monitoring is feature optimization selection and pattern recognition [7]-[10]. Whether the selected feature can reflect the variation tendency of tool wear and meet the requirement of accuracy has great influence on the result. Moreover, model parameters training needs a large number of samples, and experiment cost is larger, that leads the progress slowness. Aimed in these problems, in this paper, the second generation wavelet is used to filter cutting force signal and vibration signal, put the features selected by PCA method into hyper-sphere SVM based on statistics learning theory and describe distribution range of samples by calculating minimal hyper-sphere boundary of different tool condition, then improve the training algorithm to obtain the better learning ability and the classification results. The test results show that under the circumstances of less training samples, the proposed method has increased the accuracy of monitoring system, which has a high industrial value.
Experiment Design

Experimental Device

Tool condition monitoring system is shown in figure 1. Kistler 9257B Three-way dynamometer is installed on the slip board of CA6140 NC lathe. The tool is installed on the tool carrier by bolts. B&K4370 acceleration sensor is installed in the vertical direction of the cutting tool to measuring vibration signal caused by tool wear in machining. After the sensor signal is amplified and processed, input it into YANHUA PCI-1715U high-speed data acquisition card. The sampling frequency is 25000Hz, and sampling length is 4000. Tool type is 45w25-4k16. The material of 41065A blade is YT15 cemented carbide. The geometry angle of blade: normal rake angle is 12°, edge inclination angle is -2°, angle of law is 8°. The material of workpiece is steel No.45, and its diameter is 95mm.

Experimental Scheme

In the process of product processing, artifacts, cutting tools, processing methods and machine tool are fixed factors, the three elements of cutting and tool wear are change factors. Among them, tool wear is the uncontrollable factor. The purpose of the experiment is to study the relationship between tool wear and the signal features, so that tool wear must be turned into controllable factor, the level of distribution is according to Table 1. In order to obtain sufficient data to inspect method, all factor combination method is used in cutting experiment, the wear condition is measured by microscope.

Table 1. Factors and levels table.

<table>
<thead>
<tr>
<th>Factors</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cutting speed V/ m/s</td>
<td>1</td>
</tr>
<tr>
<td>Feed speed f/ mm/r</td>
<td>0.2</td>
</tr>
<tr>
<td>Back engagement of cutting edge ap/ mm</td>
<td>A 0.5</td>
</tr>
<tr>
<td>Tool wear interval</td>
<td>0-0.1</td>
</tr>
</tbody>
</table>

Hyper-sphere Support Vector Machine

Set the training sample for \( \{(x_i, y_i), \cdots, (x_i, y_i)\} \subseteq \mathbb{R}^n \times \mathbb{R}^m\), \(\{(x_i', y_i'), i = 1, 2, \cdots, l_s\}\) is for class s samples, \(l_1 + \cdots + l_m = l\) is the total number of samples. The input vector \(x\), the desired output \(y\), the dimensions of the input vector \(n\), the number of classes \(m\), and radius \(R\) can be found to contain all the samples in this class as many as possible. Introducing the slack variable \(\xi_n\), allows a certain error, the minimum radius square of hyper-sphere \(R\) can be solve through the objective function in Formula. (1) \([11][12]\):
\[
\min[F(R_x, a^t_i, \xi_{st})] = \min(\sum_{i=1}^{m} R_x + C \sum_{s=1}^{m} \sum_{i=1}^{ls} \xi_{st}).
\]

(1)

\[
\begin{aligned}
\text{s.t.} & \quad \|x_i^t - a^t_i\| \leq R_x + \xi_{st} \quad i = 1, \cdots, ls; s = 1, \cdots, m \\
& \quad \xi_{st} \geq 0 \quad i = 1, \cdots, ls; s = 1, \cdots, m
\end{aligned}
\]

Where \( C \) is the penalty parameter which weighs the number of error samples and the size of the hyper-sphere to controlling the penalty degree of error samples.

According to the KKT condition, the minimal hyper-sphere problem is transformed into dual form by Lagrange optimization method:

\[
\begin{aligned}
\max & \quad \sum_{s=1}^{m} \sum_{i=1}^{ls} a^t_i \{x_i^t, x_{j}^t\} - \sum_{s=1}^{m} \sum_{i=1}^{ls} \sum_{j=1}^{ls} a^t_i a^t_j \{x_i^t, x_{j}^t\} \\
\text{ s.t. } & \quad \sum_{i=1}^{ls} a^t_i = 1 \quad s = 1, \cdots, m \\
& \quad 0 \leq a^t_i \leq C \quad i = 1, \cdots, ls
\end{aligned}
\]

(2)

Where \( a^t_i \) is the Lagrange multiplier.

Solving the above quadratic optimization problems, \( m \) hyper-spheres are calculated, that each hyper-sphere represents a class of samples. The point which can be called support vector on the sphere plays a key role in the definition of the sphere, as shown in figure 2.

![Figure 2. Schematic diagram of Hyper-sphere support vector machine.](image)

According to the theory, the kernel function \( K(x, y) \) corresponds to an inner product in transformation space. In Formula (2), the corresponding part is changed into a kernel function, and the quadratic programming of the minimal hyper-sphere is obtained in the feature space:

\[
\begin{aligned}
\max & \quad \sum_{s=1}^{m} \sum_{i=1}^{ls} a^t_i K(x_i^t, x_{j}^t) - \sum_{s=1}^{m} \sum_{i=1}^{ls} \sum_{j=1}^{ls} a^t_i a^t_j K(x_i^t, x_{j}^t).
\end{aligned}
\]

(3)

Formula (3) is an optimal equation is for multi-objective classification. The parameter to be optimized is the total number of samples \( l \).

According to the KKT condition, the sample corresponding to \( 0 \leq a^t_i \leq C \) meet as follows:

\[
R_x^2 - (K(x_i^t, x_i^t) - 2 \sum_{j=1}^{ls} a^t_j K(x_j^t, x_i^t) + a^2_i) = 0.
\]

(4)

By the Formula (4) different types of value of \( R_x \) can be calculated. For the unidentified sample \( z \), the distance to the each center of hyper-sphere can be calculated by decision function to determining which class it belongs to [13]:

\[
fs(z) = K(z, z) - 2 \sum_{i=1}^{ls} a^t_i K(z, x_i^t) + \sum_{i=1}^{ls} \sum_{j=1}^{ls} a^t_i a^t_j K(x_i^t, x_j^t).
\]
The minimal \( f_k(z) \) can be found by comparison, and \( z \) belongs to the \( k \) class. Meantime, the credibility of the classification results is as follows:

\[
B_k = \begin{cases} 
1, & R_k \geq f_k(z). \\
R_k / f_k(z), & \text{otherwise}.
\end{cases}
\]  

Formula (5) shows that when the sample is in the sphere, the credibility is 1, otherwise the credibility is less than 1. Farther away from the center of the sphere, the smaller the credibility.

**Second Generation Wavelet Transformation**

Swelden and Daubechies proved that, the second generation wavelet can be used in all situations Mallat algorithm can be used in. Compared with the first generation, wavelet basis function of the second generation transforms in time domain completely, which is more efficient than Mallat algorithm. Figure 3 shows the process of signal decomposition and reconstruction realized by the second generation wavelet.

**Signal decomposition** includes three steps: Split, make prediction and update.

1. **Split**
   
   Signal \( s_j \) will be split into two: odd number sequence \( s_{2k+1} \) and even number sequence \( s_{2k} \), meet \( s_j = s_{2k+1} + s_{2k} \).

2. **Prediction**
   
   Keep even number sequence \( s_{2k} \) unchanged, and estimate odd number sequence by even number sequence, the different value between original and estimation is called wavelet coefficient, expressed as:

\[
d_{j-1} = s_{2k+1} - P(s_{2k}).
\]  

Where \( P \) is prediction operator, \( P(s_{2k}) \) is the value of estimation of \( s_{2k+1} \) made by \( s_{2k} \) in some method.

3. **Update**
   
   Correction operator \( U \) is introduced, and scale coefficients is calculated by updating even number sequence \( s_{2k} \) using wavelet coefficient \( d_{j-1} \):

\[
s_{j-1} = s_{2k} + U(d_{j-1}).
\]  

Where \( U \) is correction operator \( r \), \( U(d_{j-1}) \) is some combination of wavelet coefficient \( d_{j-1} \).
The three steps above constitute a process of ascension, repeat these steps to input $s_{j-1}$ to build a whole wavelet transform process. From the perspective of the frequency domain, wavelet coefficient $d_{j-1}$ refers to the high frequency components of original data, scale coefficients $s_{j-1}$ refers to the low frequency components.

Signal reconstruction is the reverse transformation of decomposition, it also includes three steps: back prediction, back updating and combination.

1. Back updating
   Restore even number sequence by frequent signal $s_{j-1}$ and detail signal $d_{j-1}$:
   $$ s_{2k} = s_{j-1} - U(d_{j-1}) $$

2. Back prediction
   Restore odd number sequence by even number sequence $s_{2k}$ and detail signal $d_{j-1}$:
   $$ s_{2k+1} = d_{j-1} + P(s_{2k}) $$

3. Combination
   Combine even number sequence $s_{2k}$ with odd number sequence $s_{2k+1}$ to build original signal:
   $$ s_{j} = s_{2k} + s_{2k+1} $$

The second generation wavelet can construct adaptive wavelet based on the local features of the signal, and realize nonlinear wavelet transform by changing prediction operator and modified operator, which is necessary to extract the local sensitive feature of signal.

**Signal Analysis and Feature Optimization**

We used the second generation wavelet based on the linear adaptive algorithm to process vibration signal and cutting force signal, and calculate the time and frequency feature of frequent signal of different level, then extracted time and frequency feature to make a multi-dimensional feature vector $X = [x_1, x_2, \ldots, x_n]$. Compared with the traditional wavelet analysis method, time used for the same signal feature extraction is 9 ms and 21ms. As on-line tool condition monitoring system must meet the requirements of real time operation, excessive feature may cause not only slowness, but also information redundancy, this could lead to error. In this paper, we adopted the method of PCA to reduce dimensionality of vector $X$, respectively for each sensor signal processing, extract the vector which has more than 0.85 contribution rate to build a new vector $Y$ as input vector: average of cutting force $F_x$ and $F_y$, RMS of dynamic cutting force $F_x$, RMS of vibration and its frequency band energy between [1800,2200] Hz.

**Experimental Analysis and Results**

Each experiment was repeated three times, obtaining 768 samples in different tool conditions. 8, 16, 27, 32 and 64 samples were randomly selected from each tool condition respectively. This samples were set as the training samples of the hyper-sphere support vector machine, remaining samples as test samples. The kernel function of support vector machine is based on the radial basis function. The penalty parameter $C$ and radial basis function kernel parameter $\sigma$ were selected by using grid search method. The optimal value was obtained by the combination test of the penalty parameter $C$ and kernel parameter $\sigma$. The predicted results of the test samples were shown in Table 2. As the number of training samples was 32, the classification accuracy was 53.1%; As the number of training samples increased to 64, the classification accuracy reached 75.9%; As the number of training samples reached 108, the classification accuracy reached 88.6%; As the training sample number was 128, the classification accuracy would be 90.6%; As the training number of samples increased to 256, the
classification accuracy was 92.2%. According to the results, under the condition of small samples, the classification accuracy increased with the increase of the number of training samples. However, when the number of samples reached to a certain number, the classification accuracy is stable.

Table 2. Predicted results of tool conditions.

<table>
<thead>
<tr>
<th>Number of samples</th>
<th>Tool conditions</th>
<th>Predicted results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>A</td>
</tr>
<tr>
<td>8</td>
<td>A</td>
<td>101</td>
</tr>
<tr>
<td>8</td>
<td>B</td>
<td>23</td>
</tr>
<tr>
<td>8</td>
<td>C</td>
<td>26</td>
</tr>
<tr>
<td>8</td>
<td>D</td>
<td>22</td>
</tr>
<tr>
<td>16</td>
<td>A</td>
<td>135</td>
</tr>
<tr>
<td>16</td>
<td>B</td>
<td>16</td>
</tr>
<tr>
<td>16</td>
<td>C</td>
<td>9</td>
</tr>
<tr>
<td>16</td>
<td>D</td>
<td>13</td>
</tr>
<tr>
<td>27</td>
<td>A</td>
<td>142</td>
</tr>
<tr>
<td>27</td>
<td>B</td>
<td>11</td>
</tr>
<tr>
<td>27</td>
<td>C</td>
<td>7</td>
</tr>
<tr>
<td>27</td>
<td>D</td>
<td>5</td>
</tr>
<tr>
<td>32</td>
<td>A</td>
<td>146</td>
</tr>
<tr>
<td>32</td>
<td>B</td>
<td>7</td>
</tr>
<tr>
<td>32</td>
<td>C</td>
<td>6</td>
</tr>
<tr>
<td>32</td>
<td>D</td>
<td>3</td>
</tr>
<tr>
<td>64</td>
<td>A</td>
<td>115</td>
</tr>
<tr>
<td>64</td>
<td>B</td>
<td>5</td>
</tr>
<tr>
<td>64</td>
<td>C</td>
<td>3</td>
</tr>
<tr>
<td>64</td>
<td>D</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 3 showed that the comparison of classification results by BP neural network and hyper-sphere support vector machine under the same number of training samples. When the number of samples was small, the classification accuracy of the hyper-sphere support vector machine was higher than that of the BP network. When the classification accuracy of BP neural network was not quite different from that of the hyper-sphere support vector machine, the number of training samples required by BP network was much larger than that of the hyper-sphere support vector machine. The results show that the hyper-sphere support vector machine has a good generalization ability and classification effect, and it is good for online modeling.

Table 3. Comparison of different classification results.

| Classification | Number of samples |                   |                   |                   |                   |                   |
|----------------|-------------------|-------------------|-------------------|-------------------|-------------------|
|                |                   | 32    | 64    | 108   | 128   | 160   | 210   |
| BP neural network |                   | 18.9  | %     | 41.7  | %     | 73.6  | %     | 82.9  | %     | 90.7  | %     | 92.6  | %     |
| Hyper-sphere SVM |                   | 53.1  | %     | 75.9  | %     | 88.6  | %     | 90.6  | %     | 92.2  | %     | 92.8  | %     |

**Conclusion**

The main factors influencing the accuracy of tool wear recognition are nonlinearity and non-stationarity of signal, which cause the wear feature appear to be unstable, for the same wear pattern, features range in a certain amplitude, not a fixed value. As a result of unstable feature, the accuracy of recognition is decreased because the neural network is based on the fitting of features. In this paper, the second generation wavelet filtering technique is used to optimize the feature selection and process threshold, reduced the influence made by feature variation, and increased the accuracy, the main result is as follows:
(1) It can improve the signal processing speed, and it is advantageous to the online real-time monitoring of tool condition by using the second to filter cutting force signal and vibration signal;
(2) PCA method can achieve the optimization of the feature of choice;
(3) Compared with the BP neural network, the hyper-sphere support vector machine based on the least structure risk principle has better generalization ability and recognition accuracy under small samples condition.

References


