Prediction Model of Machining Failure Trend Based on Large Data Analysis

Jirong Li
Wuyi University, Jiangmen Guangdong 529020 China

Abstract. The mechanical processing has high complexity, strong coupling, a lot of control factors in the machining process, it is prone to failure, in order to improve the accuracy of fault detection of large mechanical equipment, research on fault trend prediction requires machining, machining fault trend prediction model based on fault data. The characteristics of data processing using genetic algorithm K mean clustering for machining, machining feature extraction which reflects the correlation dimension of fault, spectrum characteristics analysis of abnormal vibration of complex mechanical parts processing process, the extraction method of the abnormal vibration of complex mechanical parts processing process of multi-component spectral decomposition and empirical mode decomposition Hilbert based on feature extraction and the decomposition results, in order to establish the intelligent expert system for the data base, combined with large data analysis method to realize the machining of the Fault trend prediction. The simulation results show that this method of fault trend prediction of mechanical machining accuracy is better, the fault in the mechanical process accurate judgment ability, it has good application value analysis and fault diagnosis in the machining process.

1. Introduction
The structure of complex machining equipment and poor working environment, which belongs to the multiple failures of the high technology of mechanical products, and the current level of technology cannot absolutely guarantee the machining equipment running without fault condition. Including fault processing machinery mainly exists: electrical system fault, rotor bearing fault, digital machine fault, power supply system fault, cooling channel blockage and electromechanical system, transmission system and so on. In order to realize the fault detection of the faults during the machining process, the method of intelligent fault detection and effective fault prediction, mechanical processing, can effectively improve the stability of machining. The fault forecast and diagnosis model of mechanical processing is a huge systematic project that covers a large data processing, artificial intelligence, pattern recognition, fault diagnosis of mechanical processing is based on the real time fault data mining Set on the basis of the analysis of large data processing and feature extraction of fault signals become the key and core steps of fault diagnosis, the fault signal time domain analysis, frequency domain analysis, and statistical analysis, as the data processing and signal processing method, the output fault feature extraction of mechanical processing equipment, the realization of fault trend mechanical processing of prediction[1].
In mechanical processing of the digital processing mode, the transmission of data in large scale, strong disturbance, easy to produce mechanical equipment fault, through accurate data mining on mechanical fault, combined with large data analysis method to realize the trend of machining current, fault prediction, fault trend prediction method of machining of digital processing mode of the main can be divided into fault data mining method based on time domain analysis, fault data mining method based on frequency domain analysis, fault data mining method based on statistical analysis, information theory analysis of the fault data mining algorithm based on pattern recognition and artificial intelligence methods [2]. Combined with the nonlinear time series analysis and signal detection algorithm to achieve the extraction of fault data mining and fault features, fault prediction and data mining to achieve the purpose of fault trend, among them, the reference literature [3] proposed a method based on Chaos The probability analysis of optimization model of trend prediction of mechanical processing digital processing fault pattern classification, combining feature extraction and compression technology, reduces the overhead of fault trend prediction process, improve the accuracy of fault trend prediction, but the computational overhead of the algorithm greatly, bad real-time trend prediction of mechanical processing. The fault algorithm of forecast method is proposed based on normal state feature extraction of fault trend reference in literature [4], coupled with the fault data enhancement, normal state features not fully reflect the digital processing mode under the machining condition information, fault trend prediction accuracy is not high.

Aiming at the above problems, this paper proposes a prediction model of mechanical fault trend based on large data. Fault feature data clustering using genetic algorithm K mean machining, machining feature extraction which reflects the correlation dimension of fault, combined with large data analysis method of fault trend prediction. Finally, the machining simulation analysis is taken and get effective conclusion.

2. Large data analysis model and fault data acquisition of machining failure

2.1. Machining failure big data analysis model

The first step of fault trend prediction of the machining is to construct the fault data model, fault data according to the means of detection and diagnosis principle is divided into mechanical vibration data, noise data processing and other fault data, considering the complicated mechanical components of the input / output parameter model, machining data distribution channel is a big fault the expansion of distribution channel, the MIMO channel propagation model of multi input and multi output fault data information acquisition, processing machinery fault data distribution channel has two main features: one is the distribution channel of big data processing machinery distribution channel model for limited bandwidth, two is constrained mechanical processing fault of large data distribution the distance from the channel, multipath propagation characteristics, easy to be interfered by mechanical vibration and other factors, led to the fault. The array of data acquisition node distribution type structure depends on the data distribution of multipath channel [5]. Assuming data acquisition node is composed of \( N = 2P \) array, data acquisition node element distribution of radial distance is \( d \), data receiving model:

\[
x_m(t) = \sum_{i=1}^{l} s_i(t)e^{j\omega c} + n_m(t), -p + 1 \leq m \leq p
\]

Wherein, \( s_i(t) \) is the vibration data of vibration sensor of mechanical processing equipment, \( s_m(t) \) is received data sequence of heat sensing element \( m \), impulse fault model which building big data processing machinery distribution channel:

\[
h(t) = \sum_{i} a_i(t)e^{j\nu(t)}\delta(t - iT_i)
\]
In the formula, $\theta_i(t)$ is the phase offset fault data distribution channel, extend time window of mechanical equipment data width is $T$, big data fault distribution extends loss equilibrium matrix representation is:

$$
x(t) = [x_{p_1}(t), x_{p_2}(t), \ldots, x_p(t)]^T_{n+1}
$$

(3)

$$
s(t) = [s_1(t), s_2(t), \ldots, s_t(t)]^T_{n+3}
$$

(4)

In the formula, $P$ is a large collection of data bandwidth, $I$ is the number of elements, in the digital processing mode, large data processing machinery fault distribution trend analysis model is expressed as:

$$
c(\tau, t) = \sum_n a_n(t)e^{-2\pi\tau s_n(y)} \delta(t - \tau_n(t))
$$

(5)

Wherein, $a_n(t)$ is the $N$ interference vectors on the path, $\tau_n(t)$ is sensing delay. Through the design, construct the analysis model of the fault trend distribution of large data processing machinery digital processing mode, to provide the basic data for the input processing machinery fault trend prediction.

2.2. Machining failure big data clustering processing

Fault feature data clustering using genetic algorithm K mean machining, machining feature extraction which reflects the correlation dimension of fault, the realization of data clustering, first of all to all the characteristics of data mining from the effective use of big data, big data because there is a lot of interference options, features of data hiding in the original information in the data [6]. Through the information processing algorithm to extract useful features, this paper uses the mining method of extracting the correlation dimension feature, I assume that the failure samples contained in the data set, the sample feature vector for $x_i$, $i = 1, 2, \ldots, n$:

$$
x_i = (x_{i1}, x_{i2}, \ldots, x_{in})^T
$$

(6)

The KNN fuzzy search is carried out in two-dimensional space, and the fuzzy clustering center matrix is obtained:

$$
V = \{v_{ij} | i = 1, 2, \ldots, c, j = 1, 2, \ldots, s\}
$$

(7)

Because the probability of a new category in the clustering process, the clustering center and therefore the need for the central region of the KNN expansion treatment for cross domain data makes a difference. $V$ is the center of clustering which $i$ a vector (Chapter $i$ is a clustering center vector), fuzzy KNN division of matrix representation for:

$$
U = \{u_{ik} | i = 1, 2, \ldots, c, k = 1, 2, \ldots, n\}
$$

(8)

In each K nearest neighbors of the test data, the weights of each class are calculated in turn, and the formula is as follows:

$$
P_{ij} = \sum_{d_i \in \text{KN}} \text{Sim}(x, d_i)y(d_i, C_j)
$$

(9)

The clustering objective function of KNN is defined (defined clustering target function):
\[ J_m(U,V) = \sum_{k=1}^{K} \sum_{i=1}^{N} \mu_m^n (d_{ik})^2 \]  \hspace{1cm} (10) 

In order to determine the K value, an initial value is used for adaptive optimization:

\[ P_{loss} = 1 - \frac{1 - p_0}{\rho} = \frac{p_0 + \rho - 1}{\rho} = \sum_{n=1}^{K} p_{K,n} \]  \hspace{1cm} (11) 

Wherein, \( d_i \) is a test feature vector of fault data, the average amount of packet loss can be expressed as \( \overline{K} \):

\[ \overline{K} = \sum_{k=1}^{K} \sum_{n=1}^{N} k p_{k,n} \]  \hspace{1cm} (12) 

Comparing the second layer in the first class, if \( D_{11} \leq D_0 \), continue to judge, obtained the average time clustering tasks into the queue as:

\[ W = \frac{\overline{K}}{\gamma} = \frac{1}{\gamma} \sum_{k=1}^{K} \sum_{n=1}^{N} k p_{k,n} \]  \hspace{1cm} (13) 

Large packet conversion waiting time is:

\[ W_q = W - \overline{K} = \frac{1}{\gamma} \sum_{k=1}^{K} \sum_{n=1}^{N} k p_{k,n} - \frac{(N-1) \mu + \rho}{\mu r} \]  \hspace{1cm} (14) 

According to the idea of differential evolution and genetic K mean algorithm, the clustering of fault feature data in machining is realized.

3. Design of fault trend prediction algorithm

3.1. Fault feature extraction

This paper presents a prediction model of mechanical fault trend data based on spectral characteristics analysis of abnormal vibration of complex mechanical parts processing process, due to mechanical failure data clustering control parameter \( s_0 \), \( \tau_0 \) often unknown, clustering detection is one of the unknown parameters of multiple hypothesis testing problem, I can use the maximum likelihood estimation value \( \hat{s}_{MLE}, \hat{\tau}_{MLE} \) of \( s_0, \tau_0 \) to tested, i.e.:

\[ l_s(r) = \int r(t) \sqrt{f_s(t)} \left( \hat{s}_{MLE} (t - \hat{\tau}_{MLE}) \right) dt 
= \max_{\hat{s}, \hat{\tau}} \int r(t) \sqrt{sf_s(t - \tau)} dt \]

\[ = \max_{\hat{s}, \hat{\tau}} \left| W_r(a, b) \right| \leq \lambda \]  \hspace{1cm} (15)
In the formula, \( a = 1/s, b = \tau, \lambda \) is the threshold. The expression for the scattered clustering is detection statistics. Because \( s, \tau \) is unknown, defined as the formula for combining decision fault feature extraction:

\[
I_z = \max_b \left| \int r(t) \frac{1}{\sqrt{a}} f^* \left( \frac{t-b}{a} \right) dt \right| \\
= \max_b |W_r(a^*, b)| > \lambda_c \quad \text{if} \quad u_l < u_r
\]

(16)

In the optimal classification plane, nonlinear time series \( x_1, x_2, \ldots, x_n, \ldots \), with total points is \( N \), the time span of the sequence \( \{ x_i \} \) autocorrelation function:

\[
R_{x_x}(f \tau) = \frac{1}{N} \sum_{n=1}^{N} \sum_{j=0}^{N} x_{i+n}^* x_{i+j} 
\]

(17)

Calculation of certain fault state sampling \( l \) dimensional feature vector \( \tilde{X}(l, n_l) \) vibration data of the within class scatter matrix \( \tilde{S}_c \), the optimal setting problem of fault data acquisition solutions can be used to describe the fault \( \chi^* \), big data is modeled as an amplitude modulated signal and array ultrasonic transducer constrained beam directivity characteristics get:

\[
x(t) = \left[ 1 - \cos(2\pi f_{c} t) \right] \left[ 1 + A\cos(2\pi f_{m} t + \phi) \right] \\
\cos[2\pi f_{c} t + B\sin(2\pi f_{m} t + \phi) + \theta] 
\]

(18)

The maximum gradient difference of the prediction of machining failure trend is obtained:

\[
AVG_X = \frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} |G_X(x, y)| 
\]

(19)

Wherein, \( m, n \) are respectively the vector quantization auto correlation coefficients of machining fault data, and then the fault distribution feature extraction is realized.

3.2. Machining failure trend prediction
The spectral characteristics of abnormal vibration in machining process of complex mechanical parts are analyzed:

\[
\rho_{xy} = \frac{\text{Cov}(X, Y)}{\sqrt{D(X)D(Y)}} 
\]

(20)

Wherein, \( \text{Cov}(X, Y) \) is the autocorrelation function of two groups of mechanical fault data sampling, \( D(X) \) and \( D(Y) \) represent the average energy.

By means of empirical mode decomposition (EMD) and Hilbert spectrum extraction method [7], the abnormal vibration of complex machine parts is decomposed into several components, and the time-frequency distribution characteristics of machining fault data are obtained:
\[
 s(t) = \sum_{r=1}^{K} p_r \sin(\omega_r n + \phi_r) + \xi(n) 
\]

Wherein, \(\xi(n)\) is the data set of wide band, \(\phi_r\) is the fault data to interference processing after the phase information, \(\omega_r\) is a recursive feature machining fault data. The machining data transmission equipment fault data distribution channel for the continuous system model assumes that the digital processing, frequency domain model of fault data is expressed as:

\[
x_n = x(t_0 + n\Delta t) + \xi_n
\]

Wherein, \(\xi(.)\) is fault data of time window function, \(\xi_n\) is a measurement error. Analysis model of fault data signal method for machining a group of relatively stable digital processing mode is established using time-frequency analysis, this process the description is shown in Figure 1.

\[\text{Fig. 1} \text{ Model of short time signal with relatively stable fault data in machining under digital processing mode}\]

In the model of relatively stable short time signal given in Fig. 1, the empirical model decomposition and Hilbert spectrum extraction for machining fault data are carried out, and the power spectral density feature is extracted:

\[
y(t) = \frac{1}{\pi} K \int_{t-\tau}^{t+\tau} \frac{x(\tau)}{\tau} d\tau = x(t)^{*} \frac{1}{\pi t}
\]

Wherein, \(K\) is the characteristic coefficient of power spectrum density, \(x(\tau)\) is the fault data in machining in the time-frequency domain of the hyperbolic frequency modulation amplitude, power spectral density data fault feature extraction results:

\[
C_{\tau}(f) = \sum_{k} c_k e^{-j2\pi f t}
\]

The construction of the fault diagnosis and classification in expert system, a training set for \(X = [X_1, X_2, \ldots, X_k, \ldots, X_N]^T\), fault trend of machining failure distribution data \(x(n)\) is decomposed into several IMF components, including any training sample is \(X_k = [x_{1k}, x_{2k}, \ldots, x_{nk}, \ldots, x_{mk}]\), the corresponding output for fault trend prediction is:

\[
Y_k = [y_{1k}, y_{2k}, \ldots, y_{nk}, \ldots, y_{mk}] (k = 1, 2, \ldots, N)
\]

Based on the data analysis, an intelligent expert system is established, and the trend prediction of mechanical machining is realized by combining large data analysis method.
4. Simulation experiment and result analysis

In order to test the performance of the method for predicting mechanical processing trends in the implementation of digital processing fault mode. The experiment is taken based on the Matlab simulation environment, the simulation of the hardware environment: IntelCore3-530 1 G memory, the operating system is Windows 7. Mechanical and large data acquisition frequency is 2~10 KHz, carrier frequency width is 2ms, fault trend prediction of normalized initial frequency were: $f_{i1} = 0.05$, $f_{i2} = 0.15$, the fault data sampling points number is $N = 256$, the signal-to-noise ratio of the interference range is -10~10dB. set according to the simulation environment and parameters, the digital processing model to predict machining fault trend, first time series fault data sampling, time domain processing machinery the data sampling waveform as shown in Figure 2.

![Time domain sampling waveform of machining data](image)

**Fig. 2** Time domain sampling waveform of machining data

The machining data processing model of the digital sampling of the test object, construct the training set and test set, fault trend prediction training simulation and test, reflect the correlation dimension of machining feature extraction fault, the abnormal vibration of complex mechanical parts processing in the process of spectrum analysis, spectrum feature extraction results as shown in Fig. 3.

![Feature extraction results](image)

**Fig.3** Feature extraction results

Figure 3 analysis result showed that this method for feature extraction of data processing machinery fault data, it has a better classification performance, it can accurately identify different fault trend. In order to compare the performance by using this method and the traditional method, get the
accuracy of the results of fault trend prediction is shown in Figure 4. Simulation results show that, the higher prediction accuracy of machining fault trend by using the algorithm of digital processing mode.

![Accuracy comparison of mechanical machining failure trend prediction](image)

**Fig. 4** Accuracy comparison of mechanical machining failure trend prediction

5. Conclusion

The prediction model of machining failure trend is studied in this paper. This paper proposes a prediction model of machining fault trend based on large data. Fault feature data clustering using genetic algorithm K mean machining, machining feature extraction which reflects the correlation dimension of fault, spectrum characteristics analysis of abnormal vibration of complex mechanical parts processing process. The extraction method of the abnormal vibration of complex mechanical parts processing process of multi-component spectral decomposition and empirical mode decomposition Hilbert, according to the feature extraction and decomposition results, combined with large data analysis method of fault trend prediction. The machining results show that this method of fault trend prediction of mechanical machining accuracy is better, the fault of mechanical processing in the process of accurate judgment ability, fault diagnosis and analysis in the process of Good practical value.

Acknowledgments

Fund Project: The Key Project of Department of Education of Guangdong Province (No. 2014 KZD XM055)

References

[1] FU Yue, LI Jing-yi. Finite-time optimal decoupling control for linear multi-variable systems. Control and Decision, 2016, 31(12): 2271-2274.
[2] XIE Hong'an, LI Dong, SU Yang, YANG Kai. Trusted network management model based on clustering analysis [J]. Journal of Computer Applications, 2016, 36(9): 2447-2451.
[3] BARTOLINI N, CALAMONERI T, LA PORTA T F. et al. Autonomous deployment of heterogeneous mobile sensors[J]. IEEE Transactions on Mobile Computing, 2011, 10(6): 753-766.
[4] LIN T Y, SANTOSO H A, and WU K R. Global sensor deployment and local coverage- aware recovery schemes for smart environments [J]. IEEE Transactions on Mobile Computing, 2015, 14(7): 1382-1396.
[5] KASHI S S and SHARIFI M. Coverage rate calculation in wireless sensor networks[J]. Computing, 2012, 94(11): 833-856.
[6] Eldemerdash Y A, Dobre O A, and Liao B J. Blind identification of SM and Alamouti STBC-OFDM signals[J]. IEEE Transactions on Wireless Communications, 2015, 14(2): 972-982.
[7] Karami E and Dobre O A. Identification of SM-OFDM and AL-OFDM signals based on their second-order cyclostationarity[J]. IEEE Transactions on Vehicular Technology, 2015, 64(3): 942-953.