Fault diagnosis of automaton based on local characteristic-scale decomposition and individual feature selection

Wei Du and Liqing Fang

Department One, Mechanical Engineering College, Shijiazhuang, 050003, China

Abstract. In order to accurately use the vibration signal to fault diagnosis, a method of automaton fault diagnosis based on local characteristic-scale decomposition (LCD) and individual feature selection (IFS) was proposed. Firstly, the vibration signal of automaton was decomposed with local characteristic-scale decomposition to obtain intrinsic scale components (ISC) in different scales, and the correlation dimension of these components were calculated as characteristic parameters. Then, the individual optimal feature subset for each pair of fault states was selected using an improved Fisher feature selection method. Finally, the feature matrix was input to multi-class support vector machine for fault classification and recognition. The results of automaton fault test demonstrate that the proposed method is of high accuracy in fault diagnosis.

1. Introduction

Automaton is the core component of the anti-aircraft gun system, which in the process of reciprocating motion accompanied by intense impact, friction, vibration and beating [6], resulting in non-linear and non-stationary short-term shock vibration signal. Feature extraction is the key problem in fault diagnosis. The accuracy of fault diagnosis and the reliability of early prediction are directly related to whether fault information can be extracted accurately. The vibration signal of automaton has fractal characteristics in a certain scale, and it is feasible to quantify its structural characteristics as feature parameters for fault diagnosis. The correlation dimension is widely used in the fractal technology of fault diagnosis, for it can reflect the dynamic structure of the attractor of non-linear system sensitively [7,8].

The existence of noise has a great influence on the calculation of the correlation dimension, and even result in non-convergence of the correlation dimension [3]. Local characteristic-scale decomposition (LCD) algorithm [2] is a new adaptive signal decomposition method, which can decompose the nonstationary signal into a series of intrinsic scale components (ISC), characterizing the different time scale of signal, and the noise reduction can be achieved by eliminating the residual components. LCD method is more effective than empirical mode decomposition (EMD) [5] and local mean decomposition (LMD) [4] in terms of speed of operation, suppression of end-effect and mode aliasing [1].

Support vector machine (SVM) is a widely used intelligent classification method in fault diagnosis, and the multi-class classifier is realized by combining binary classifiers. All of the binary classifiers share a same set of feature sets to distinguish all categories. This feature selection method is called shared feature selection (SFS). In practice, a feature may be easy to distinguish between two categories, but cannot effectively distinguish between all classes. In order to improve the effectiveness of the feature set and improve the diagnostic efficiency in the fault diagnosis model, an individual feature selection (IFS) approach is proposed in this paper. First, an improved individual feature

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

selection approach based on Fisher method is used to select the individual feature subsets for each binary classifier. Then, the multi-class classifier is constructed by using one-against-one (OAO) approach. The fault signals of the measured automaton are analysed to verify the effectiveness of the method.

2. Theoretical background

2.1 Local characteristic-scale decomposition

LCD is based on the Intrinsic time-scale decomposition (ITD) method. Assume that a multicomponent signal can be decomposed into a finite number of intrinsic scale components (ISC) with physical meaning of instantaneous frequency, and any two ISC components are independent of each other [2]. According to the definition of the ISC components, the signal LCD decomposition process is as follows:

(1) Determine all extreme points (τ_k, X_k) of $x(t)$, $k = 1, 2, 3, \cdots, M$, where M is the number of extreme points. The two consecutive extreme points can divide $x(t)$ into several intervals, and linearly transform $x(t)$ between any two adjacent extreme points:

$$L_t^{(k)} = L_k + \frac{L_{k+1} - L_k}{X_{k+1} - X_k}(x_t - X_k),\ t \in (\tau_k, \tau_{k+1}) \qquad (1)$$

Note that $L_t^{(k)}$ represents a baseline signal segment obtained by linear transformation of the k-th section of the original signal, and the base-line signal segments in adjacent regions are connected at the beginning and the end to obtain a baseline signal L_t, in formula (1)

$$L_{k+1} = aA_{k+1} + (1-a)X_{k+1} \qquad (2)$$

where the parameter a generally takes the value 0.5.

(2) The baseline signal L_t is separated from the original signal $x(t)$ to obtain a residual signal $P_1(t)$. If $P_1(t)$ satisfies the ISC component criterion, let $\mathrm{ISC}_1(t) = P_1(t)$. Otherwise, $P_1(t)$ is used as the original signal and repeat steps (1), (2), cycling k times until the ISC component P_k is obtained, that is, $\mathrm{ISC}_1(t)$.

(3) $\mathrm{ISC}_1(t)$ is separated from the original signal $x(t)$ by repeating steps (1) and (2) for n times until the residual component is a monotonic function or less than a preset threshold, and n components satisfying the ISC component criterion are obtained. Then

$$x(t) = \sum_{i=1}^{n} \mathrm{ISC}_i(t) + r_n \qquad (3)$$

The ISC components are separated by LCD method from the original signal step by step and the multi-scale adaptation of the signal data is realized, obtaining the different levels of information in the signal. By eliminating the residual components, the purpose of noise reduction is achieved.

2.2 Correlation dimension

Compared with other fractal dimensions, such as box dimension, information dimension and volume dimension, correlation dimension can be calculated easier, and identify the working state of the system better, reflecting the degree of inhomogeneity of system attractor. The G-P algorithm based on the idea of time-space embedding phase space reconstruction is used to obtain the correlation dimension $C_m(r)$ of the one-dimensional time series with embedding dimension m. $C_m(r)$ reflects the ratio of the points with distance less than r $(r > 0)$ on the attractor in the reconstructed phase space, with $C_m(r) = r^{D_2(m,r)}$ in the scale-free region of r. When $r \to 0$, we get the correlation dimension

$$D_2(m,r)=\lim_{r\to 0}\frac{\ln C_m(r)}{\ln r} \qquad (4)$$

The logarithmic curve $\ln C_m(r)-\ln r$ is plotted, and the least squares linear fitting of its scale-free region is performed. And calculate the slope to get the correlation dimension of the time series. The selection of embedding dimension m and delay time τ has a great influence on the calculation of correlation dimension. The determination methods of parameter τ are autocorrelation function method, C-C method and mutual information method. The parameter m is determined by the correlation method, the observation method and the Cao method. The mutual information method and the Cao method are used to determine the parameters τ and m respectively.

2.3 Improved Fisher feature selection method

We assumed that in the two-class problem (C_1,C_2) gives a sample set $\{x_{i,j}\}$, $i=1,2,\cdots,N$, $j=1,2,\cdots,D$, a total of N samples, and each sample contains D-dimensional features. The fisher criterion value $J_F(d)$ of the d - dimensional feature F_d is defined as the ratio between the between-class variance $S_{B,d}$ and the intra-class variance $S_{W,d}$, that is

$$J_F(d)=\frac{S_{B,d}}{S_{W,d}} \qquad d=1,2,\cdots,D \qquad (5)$$

In (5), between-class variance and intra-class variance are defined as

$$S_{B,d}=(m_{1,d}-m_{2,d})^2 \qquad (6)$$

$$S_{W,d}=\sigma_{1,d}^2+\sigma_{2,d}^2 \qquad (7)$$

where $m_{1,d}$ and $\sigma_{1,d}^2$ are the mean and variance of feature F_d in class C_1, $m_{2,d}$ and $\sigma_{2,d}^2$ are the mean and variance of F_d in class C_2, respectively.

The greater the ability of a feature to discriminate between classes, the smaller the interclass spacing and the larger the interclass spacing, and the larger the Fisher criterion value is. However, when the two features are linearly related and both have higher Fisher criterion values, both of them will be selected, resulting in redundancy [9]. In order to improve the effectiveness of the selected feature set and the diagnostic efficiency, the Fisher feature selection method is improved: Firstly, the Fisher criterion value of each feature is calculated and the uncorrelated features of poor classification ability with Fisher criterion value less than 0.5 are removed. Then, for the remaining features, the correlation coefficient for each of the two features is calculated, and if the correlation coefficient is greater than 0.98, features which have smaller Fisher criterion value are removed, for there is significant redundancy. The remaining M features are ranked by the Fisher criterion from large to small, and the first m features are trained in SVMs and test the recognition rate each time, where the value of m is $(2,3,\cdots,M)$. The feature subset with the highest classification precision and the least number of features is the the optimal feature subset of each fault state.

3 Experiment analysis

3.1 Data acquisition

Measured vibration signals were from a type of air-guided automaton of anti-aircraft gun. According to the common failure mode of the automaton, the three working states of the block wear (F1), transmission spring failure (F2) and normal (N) were set. Among them, the block was set an area of about 7mm × 10mm wear failure, and the failure mode of transmission spring was plastic deformation. Taking into account the practical application of the convenience and safety of operation, manual way was used to complete the processes of latch, unlatch, lock and unlock.

Respectively, vibration signals of three states of the process were measured. The vibration signal acquisition system used in the experiment included three-way piezoelectric accelerometer CA-YD-

193A01, NI9234 data acquisition card and NI cDAQ-9171 chassis data acquisition module and the measurement and control software SignalPad, the acceleration sensor installation position shown in Figure 1. The sampling frequency was 20kHz, each state measured 40 sets of data. In order to reduce the interference of the unrelated signals in the experimental operation, only the signals in the working process of block were intercepted and analysed. A group of data samples consisted of the time series of 1600 sampling points. The vibration signals of the three working states of the automaton are shown in Figure 2. It is suggested from the figure that the time-domain waveform of each state vibration signal difference is not obvious, therefore, only time-domain waveform is not accurate and further analysis and recognition are needed.

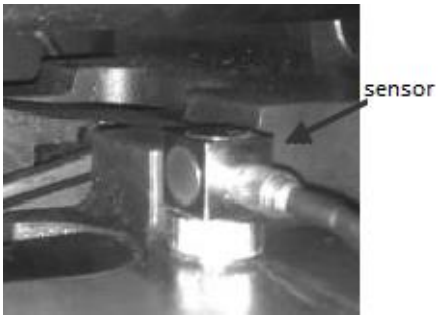


Figure 1. Installation location of sensor

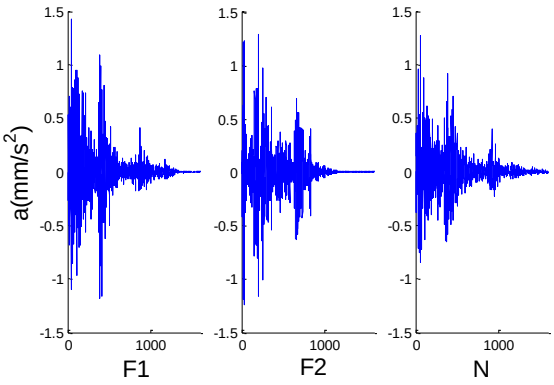


Figure 2. Time domain waveform of automaton in three states

3.2 Feature extraction

The data samples were decomposed by the LCD, and most of the sample data were decomposed into 7 ISC components and 1 residual component. Figure 3 shows the result of the decomposition of a set of sample data in the vibration signal of the block wear. The delay time τ and the embedding dimension m were determined by the mutual information method and the Cao method, and the correlation dimension of the first 7 ISC components were calculated by G-P algorithm, obtaining three 40 × 7 dimensionality matrix of normalized data. Figure 4 shows the correlation dimension errorbar of each ISC component in three states of the automaton, and the result is the statistics of all data. It can be seen that the correlation dimension of some ISC components has certain separability, which indicates that the irregularity and complexity of the signal are different at different scales. Therefore, the features extracted by the combination of LCD and correlation dimension have the ability to characterize the operation information of the automaton status.

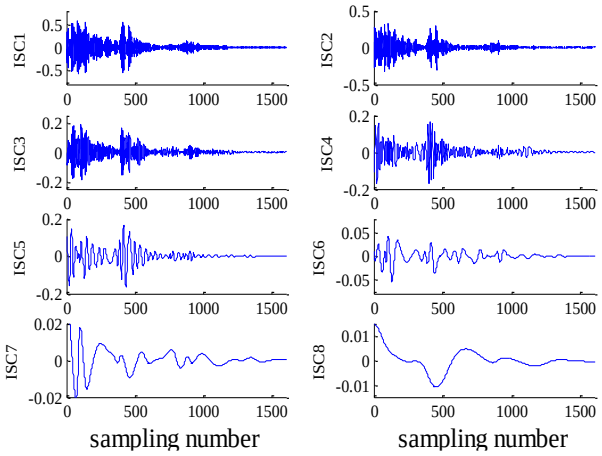


Figure 3. Decomposition of signal F1

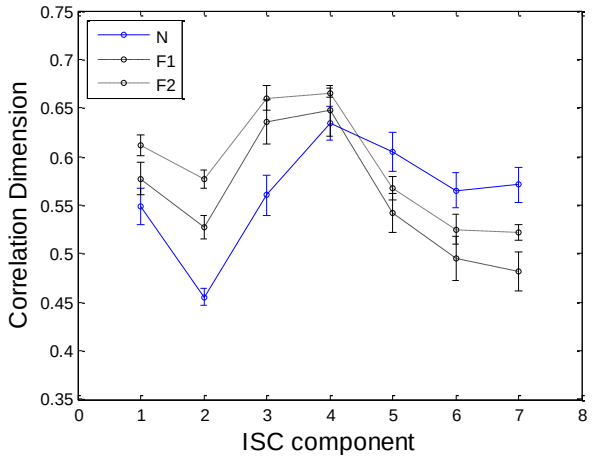


Figure 4. Correlation dimension of ISC component

3.3 Feature selection and fault diagnosis

20 sets of data were randomly selected as training samples, and the remaining 20 groups were used as test samples. The improved Fisher feature selection method was used to select the optimal feature subset for each binary classifier. The feature selection results are shown in Tab.1, where the number 1 to 7 represent the 7 ISC correlation dimensions. The RBF kernel function with the parameter $g=1$ and the penalty parameter $C=1$ were chosen to train three binary SVMs with their respective optimal subset.

Table 1. Results of IFS

State	F2	N
F1	2,3,6	2,3,5,7
F2	-	1,2,7

A one-against-one approach was used to construct a multi-class classifier to form the final fault diagnosis model. The results of the tests and the overall diagnostic results (IFS) for each state are shown in Tab.2. In order to verify the superiority of the proposed method, the recognition rate (SFS) using shared feature selection method is also given. Overall, the proposed method improves the

recognition rate by approximately 8.34% compared with the traditional method, and obtains the ideal diagnosis effect.

Table 2. Recognition results of SVM

State	Error diagnosis	Correct diagnosis	Recognition rate
F1	4	16	80%
F2	1	19	95%
N	3	17	85%
IFS	8	52	86.67%
SFS	13	47	78.33%

4. Conclusion

Aiming at the non-linear and non-stationary characteristics of automaton vibration signals, a fault diagnosis method based on local characteristic-scale decomposition and individual feature selection was proposed. Experimental results indicate that the LCD of automaton vibration signal and the correlation dimension of ISC component provide an effective basis for state recognition. The improved Fisher feature selection method was used to optimize the feature subset, which can improve the validity of the selected feature set and the recognition rate of the automaton fault diagnosis.

References

[1] Cheng Junsheng, Yang Yi, Yang Yu. 2012. “Local characteristic-scale decomposition method and its application to gear fault diagnosis”, *Journal of Mechanical Engineering*, **9**, pp. 64-71.
[2] Cheng Junsheng, Zheng Jinde, Yan Yu. 2012. “A nonstationary signal analysis approach—the local characteristic-scale decomposition method”, *Vibration Engineering*, **2**, pp. 215-220.
[3] Argyris J, Andreadis I, Pavlos G, et al. 1998. “The Influence of noise on the correlation dimension of chaotic attractors”, *Chaos, Solitons & Fractals*, **9**, pp. 343-361.
[4] Smith J S, 2005. “The local mean decomposition and its application to EEG perception data”, *Journal of the Royal Society Interface*, **2**, pp. 443-454.
[5] Huang N E, Shen Zheng, Long S R. 1999. “A new view of nonlinear water waves: the Hilbert spectrum”, *Annual review of fluid mechanics*, **31**, pp. 417-457.
[6] Pan Hongxia, Du Heng, Ma Chunmao. 2015. “High-speed automaton fault diagnosis based on local wave and information entropy”, *Vibration, Measurement & Diagnosis*, **6**, pp. 1159-1164.
[7] Alberto R N, María E, Montesino O. 2005. “A method for the correlation dimension estimation for on-line condition monitoring of large rotating machinery”, *Mechanical Systems & Signal Processing*, **19**, pp. 939-954.
[8] Wang Bingcheng, Ren Zhaohui, Wen Bangchun. 2012. “Fault diagnosis method of rotating machines based on nonlinear multi-parameters”, *Journal of Mechanical Engineering*, **5**, pp. 63-69.
[9] Zhou Yong, He Chuangxin. 2012. “Bearing fault diagnosis under varying load conditions based on individual feature selection and relevance vector machine”, *Journal of Vibration and Shock*, **3**, pp. 157-161.