

Incipient mechanical fault detection based on multifractal and MTS methods

Hu Jinqiu*, Zhang Laibin, Liang Wei and Wang Zhaohui

School of Mechanical and Electronic Engineering, China University of Petroleum, Beijing 102249, China

Abstract: An incipient mechanical fault detection method, combining multifractal theory and Mahalanobis-Taguchi system (MTS), which is based on statistical technology, is proposed in this paper. Multifractal features of vibration signals obtained from machine state monitoring are extracted by multifractal spectrum analysis and generalized fractal dimensions. Considering the situation of mass samples of normal mechanical running state and few fault states, the feature parameters corresponding to different mechanical running states are further optimized by a statistical method, based on which incipient faults are subsequently identified and diagnosed accurately. Experimental results proved that the method combining multifractal theory and MTS can be used for incipient fault state recognition effectively during the mechanical running process, and the accuracy of fault state identification is improved.

Key words: Incipient fault, fault detection, multifractal, Mahalanobis-Taguchi system (MTS)

1 Introduction

Incipient fault detection is one of the most important issues to be investigated in mechanical fault diagnosis. Quick and accurate incipient fault identification and diagnosis can help to minimize the quality and productivity offsets, and reduce the risk of hazardous consequences in abnormal situations. Due to incipient fault, some fault features based on vibration analysis are not obvious enough to detect in time normally, which may lead to secondary failures, malfunctions, and even breakdown during operation. This is because some vibration features, which are indicative of machine health, can be obscured by vibration from other machine parts, multiplicity of transmission paths, and ambient noise.

In the last decade, several methodologies were developed for fault detection, such as neural network (Chen, 2008; Wu et al, 2008), time series modeling method (Marseguerra et al, 1992; Lv et al, 2007; Zhao et al, 2007), wavelet analysis (Chen et al, 2007; Rubini and Meneghetti, 2001; Wu and Liu, 2008), empirical mode decomposition (Gao et al, 2008; Wu and Qu, 2008), and time-frequency analysis (Qin et al, 2008). Although these methods produced satisfying experimental and field results, there are still two important drawbacks: they have an inferior ability of fault structure analysis and are insensitive to incipient faults, which result in low accuracy of diagnosis, and failing to track hidden damage processes.

Wang and Zhang (2003) proposed a fractal method to diagnose the incipient faults of engines based on the fractal dimension of signals undergoing condition monitoring, with the emphasis on reflecting the fault essence. However, single fractal approach only can reflect the overall irregularity

of signals, but cannot describe the local scaling properties (Peng et al, 2002). Different local conditions and fluctuation of parameters are important indicators of incipient faults in feature extraction.

Simultaneously intelligent fault identification can be achieved by classification of signals based on the information, and/or features extracted from the signals. Conventional methods (Siddique et al, 2003) rely on mass training data and historical data, especially fault samples for fault pattern recognition. In practice, however, it is not easy to apply these techniques due to the lack of efficient procedures to obtain training data (nonlinear or singular fault samples) and specific knowledge, which are required to develop a diagnosis model, since machines are usually in normal running states during most of their operational lifetime (Wang et al, 2008).

Another important challenge in improving fault identification accuracy is optimizing the fault features. Jade et al (2003) and Lee et al (2004) investigated the use of Kernel Principal Component Analysis (KPCA) for feature extraction, optimization, and denoising of nonlinear systems. Tamura and Tsujita (2007) used the PCA method to determine the number of principal components (PCs) in feature space during the fault detection process, and established relationships between the number of PCs and sensitivities for fault detection. In respect of optimization, Sun et al (2007) developed an Evolving Kernel Principal Component Analysis (EKPCA) method based on an evolution algorithm to select kernel functions and extract nonlinear fault features effectively. Though PCA-based methods partially address feature optimization, some disadvantages still exist. PCA-based methods do not decrease the dimension of feature space in respect of physical mechanism, because every PC is a linear or nonlinear combination of original variables. Secondly, it is difficult to give a physical explanation of abnormal

* Corresponding author. email: hjqcup@yahoo.com.cn

observations based on PCA method.

Aiming at the above three challenges, an incipient fault detection method combining multifractal theory and Mahalanobis-Taguchi system (MTS) is proposed in this paper. Based on multifractal theory, the local conditions of fractal objects are described more precisely; the relationship of local scaling properties and the overall characteristics is established, and the incipient fault features are successfully extracted from the multifractal spectrum and generalized dimensions. Fault identification procedures are developed further for the statistical optimization of feature parameters by optimal Mahalanobis space, and automatic fault classification using the MTS method.

Experimental results of incipient fault detection achieved from the process of field data acquisition, diagnostic model parameter determination, multifractal feature extraction, Mahalanobis distance calculation, feature optimization based on two-level orthogonal array, and final fault classification indicate that the combinatorial method with multifractal theory and MTS for incipient fault detection is effective, feasible, and of remarkable higher accuracy.

2 Incipient fault feature extraction by multifractal theory

2.1 Generalized dimensions

Fractals refer to objects that are either self-similar or self-affine. The whole fractal structure can be regarded as the buildup of local fractals with different local fractal dimensions. This structure composed of different local fractal structures is called Multifractal. In other words, multifractal is defined as a set constituted by singular measures with multi-scaling exponents in fractal structure. For any fractal object, a set of generalized dimensions can reflect the spatial distribution information of the fractal structure.

Based on the Whitney embedding theorem and Packard phase space reconstruction theory, Grassberger and Procaccia (1984) proposed an algorithm, using time series data, to calculate the correlation dimension of strange attractors in dynamical systems. Generalizing the above theory, the correlation integral method to estimate the generalized dimensions of mechanical vibration signals is provided as follows:

Set signal series as $\{x_k, k = 1, 2, \dots, N\}$, and embed them into m -dimensional Euclidean space R^m , then a point set $J(m)$ is obtained, in which the elements are denoted by:

$$X_n(m, \tau) = (x_n, x_{n+\tau}, \dots, x_{n+(m-1)\tau}) \quad n = 1, 2, \dots, N_m \quad (1)$$

where τ is delay time, and $N_m = N - (m - 1)\tau$ is the number of vectors in the point set $J(m)$.

The q -order correlation integral for a discrete fractal set is defined as:

$$c(q, r) = \left\{ \frac{1}{N_m} \sum_{i=1}^{N_m} \left[\frac{1}{N_m} \sum_{j=1}^{N_m} H(r - r_{ij}) \right]^{q-1} \right\}^{\frac{1}{q-1}} \quad (2)$$

where r_{ij} denotes the distance of X_i and X_j in the point set

$J(m)$, that is

$$r_{ij} = d(X_i, X_j) = \left[\sum_{l=0}^{m-1} (x_{i+l\tau} - x_{j+l\tau})^2 \right]^{\frac{1}{2}} \quad (3)$$

$H(x)$ is Heaviside function, that is

$$H(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (4)$$

Then the generalized dimensions can be represented as:

$$D(q) = \lim_{r \rightarrow 0} \frac{\ln(c(q, r))}{\ln(r)} \quad (5)$$

Based on Eq. (5), generalized dimensions $D(q)$ can be obtained by the least square linear-fitting method (Yu et al, 2004). The non-stationary vibration signals can be characterized by a set of fractal dimension values, which provides a potential for signal feature extraction in terms of complexity of the signal.

2.2 Multifractal spectrum

The multifractal spectrum $f(\alpha)$ is also named as singularity spectrum, where α , singularity index, indicates the local dimension of a certain small area of the fractal object. $f(\alpha)$ represents the fractal dimensions of certain subsets which have the same value of α , and describes the characteristics of the fractal object on different levels as well as the distribution of fractal measures in the support set. In other words, $f(\alpha)$ is used to research the overall fractal object in respect of its local scaling properties. α and $f(\alpha)$ are provided from the following Legendre transforms:

$$\tau(q) = (q - 1)D(q) \quad (6)$$

$$\alpha(q) = \frac{d\tau(q)}{dq} \quad (7)$$

$$f(\alpha(q)) = q\alpha(q) + \tau(q) \quad (8)$$

Three important characteristic parameters of a multifractal spectrum can be obtained from Eqs. (6)-(8), which are the width of the multifractal spectrum $\Delta\alpha = \alpha_{\max} - \alpha_{\min}$, the difference of the fractal dimensions between the maximum probability and minimum probability subsets $\Delta f = f(\alpha_{\max}) - f(\alpha_{\min})$, and the maximum of multifractal spectrum f_{\max} . The value of $\Delta\alpha$ describes the degree of inhomogeneity of the distribution of probability measures on the overall fractal structure, which refers to the fluctuating degree of mechanical vibration signals. The value of Δf describes the proportion of the number of elements at the maximum and minimum in the subset developed by related physical variables, which refers to the proportion of the large and small peaks of vibration signals. f_{\max} , the largest height on the multifractal spectrum, represents the irregularity of the number of units having the same probability as the scale ε changes, which refers to the irregularity of the large and small peaks of vibration signals. The proportion of the large and small peaks and the irregularity represent the degree

of severity of the mechanical vibration. Therefore, the multifractal spectrum reflects the proportion and the degree of inhomogeneity of the distribution of probability measures on the overall fractal structure, that is to say, it comprehensively describes the fluctuating degree and the degree of severity of mechanical vibration signals. From this point of view, the features of different incipient fault states can be extracted in this way.

2.3 Multifractal spectrum entropy

Information entropy, which is used to describe the degree of information uncertainty in information theory, is a measure for the homogeneous degree of the probability distribution. Kolmogorov and Chaitin (Li and Vitanyi, 1997) extended the concept of entropy, and proposed the entropy definition based on the system complexity. Based on the calculation theory of entropy, the homogeneity of singularity distribution of fractal structure is described quantitatively, and the characteristic indices of mechanical vibration signals are developed in terms of multifractal spectrum entropy. The multifractal spectrum entropy can describe the singular condition of the signal energy distribution and its probability of the geometric feature distribution. It can be used for precise description and accurate extraction of the intrinsic feature information of non-stationary signals, which is conducive to discovering and identifying the incipient fault signal characteristics.

Set $f(\alpha)$ as the multifractal spectrum of a mechanical vibration signal, and based on the measuring formula of entropy, the calculation of multifractal spectrum entropy H_m is provided as follows:

$$H_m = -\sum_{i=1}^k p_i \log p_i \tag{9}$$

$$p_i = \frac{f(\alpha_i)}{\sum_{i=1}^k f(\alpha_i)} \tag{10}$$

where p_i represents the proportion between $f(\alpha_i)$ and the whole multifractal spectrum. Multifractal spectrum entropy H_m reflects the complexity of non-stationary signal in respect of multifractal spectra with different measures.

3 Optimization and classification of fault features based on MTS

The Mahalanobis-Taguchi system (MTS) is a pattern information technology, which has been used in different applications to make quantitative decisions by constructing a multivariate measurement scale using a data analytic method (Das and Datta, 2007; He and Han, 2007; Pedro et al, 2004). The design model of MTS is shown in Fig. 1, which includes two key steps: (1) constructing Mahalanobis distance (MD) scale; (2) using the signal-to-noise ratio (SNR) to evaluate the quality of measurement, and optimizing all of the information to improve the SNR with an orthogonal array. For the feature recognition based on MTS, mass normal feature samples are utilized to develop Mahalanobis distance space, whereas only a few fault feature samples are used for feature optimization.

3.1 Mahalanobis distance scale

In the MTS, the Mahalanobis space (reference group) is obtained using the standardized variables of healthy or normal data.

- Let
- m = the number of features obtained from original signals;
- n = the number of signal samples;
- n_{ij} = standardized i th observation (sample) of normal

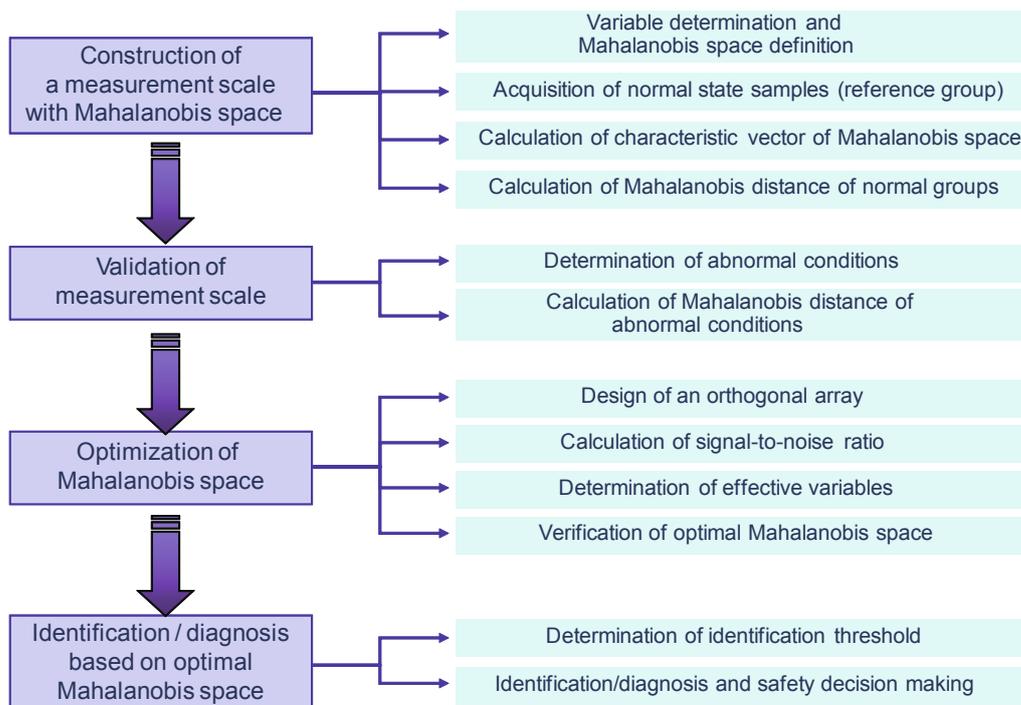


Fig. 1 Design model of MTS

group for j th variable (feature), $i = 1, \dots, n, j = 1, \dots, m$;
 $R_{m \times m}$ = correlation matrix for m standardized variables (feature);

$$R_{m \times m}^{-1} = [a_{ij}]_{m \times m};$$

$Y = (y_1, y_2, \dots, y_m)$ = an observation of test group with unknown condition.

Then the Mahalanobis distance D^2 calculated for the test group observation is given by

$$D^2 = \left[\sum_{i=1}^m \sum_{j=1}^m (a_{ij} y_i y_j) \right] / m \tag{11}$$

When an observation belongs to the reference group, $D^2 \approx 1$, otherwise D^2 should be considerably greater than one and increase as the difference between the observation and the reference group increases. According to the difference degrees between each of the fault states of the mechanical machine and the normal states, different Mahalanobis distance intervals can be mapped to different operational states of the mechanical machine.

3.2 Feature optimization based on two-level orthogonal array

There are various state features of a mechanical machine existing in the Mahalanobis space, whereas parts of them are not relevant to recognition. Hence, selecting the effective features can improve the recognition ability of Mahalanobis space, and save the calculation cost of feature extraction and diagnosis. According to the number of features participating in the optimization, an appropriate two-level orthogonal array is selected. In MTS, an orthogonal array is mainly used to identify useful features, and to assign them different lines through the fewest experiments. Each feature variable has two levels - level 1 and level 2, where level 1 represents using such feature variables to calculate MD. Level 2 represents not using such feature variables to calculate MD, and simultaneously, the appointed features on each row are used to construct the Mahalanobis space.

The estimation value of Mahalanobis distance of practical machine states is defined as the signal factor level, and d state feature samples are extracted from d signal factor levels M_i ($i=1, 2, \dots, d$). That is to say, one of the feature samples is selected corresponding to one certain machine state. Then, the Mahalanobis distances $D_1^2, D_2^2, \dots, D_d^2$ between each feature vector and the Mahalanobis space are calculated based on Eq. (11). The identification performance of the Mahalanobis space is evaluated by signal-to-noise ratio (SNR) based on Eq. (12).

$$\eta = 10 \ln [V(S - V) / r] \tag{12}$$

where

$$r = \sum_{i=1}^d M_i^2, \quad S = \frac{\left[\sum_{i=1}^d (M_i D_i^2) \right]^2}{r}, \quad V = \frac{\sum_{i=1}^d D_i^2 - S}{d - 1}$$

The fault identification performance is improved as the SNR of the Mahalanobis space increases. Considering the SNR of each Mahalanobis space, the average signal-to-noise ratios of features relative to each machine state on level 1 and level 2 are calculated, respectively. If the average SNR of a certain feature on level 1 is higher than level 2, this feature contributes to better fault identification, and should be advised to be reserved, otherwise be eliminated. Finally, the Mahalanobis space newly constructed by reserved features can obtain the theoretically optimal identification performance.

4 Fault identification model based on multifractal theory and MTS

A mechanical fault detection model is developed based on multifractal theory and MTS in this study. Nine multifractal features $X = (x_1, x_2, \dots, x_9) = (D(-2), D(-1), D(0), D(1), D(2), \Delta\alpha, f_{\max}, \Delta f, H_m)$ of mechanical vibration signals are extracted using multifractal theory. Secondly, MTS is used to optimize the above features, i.e. selecting effective features with greater contribution to the SNR of Mahalanobis space. Finally, based on MTS, Mahalanobis distance is used to discriminate mechanical fault states and identify the pattern of fault data from multiple-dimension data. The overall work flow of the detection model is shown in Fig. 2, and the main parameters are selected as follows:

- (1) Selection of data length N . Related literature (Xie and Xie, 1997) shows that during the process of fractal dimension evaluation of the discrete time series signal, the length of series should be larger than 500. Considering the characteristics of mechanical vibration signals and the sampling frequency, the length of series in this model is set as 1,024.
- (2) Selection of delay time τ . Based on the pseudo-phase portrait, delay time τ is set step-by-step from a small value to a large one. When the pseudo-phase portrait extends to its maximal degree, the self-correlation of data reaches minimum, and moreover the motion state information reflected by the original data may not be lost. At this moment, the value of τ is the optimum.
- (3) Selection of embedding dimension m . Trial and error is used to determine m . If the embedding dimension m increases gradually during calculation, it will commonly converge to a lower dimensional space for a deterministic system. That is to say, the fractal dimension will tend to be a constant with increasing embedding dimension. When the fractal dimension begins to converge, the value of embedding dimension m at this moment is the optimum in the model.
- (4) Selection of two-level orthogonal array. Considering the original multifractal feature vector, the factor number c is 9, and the value t of factor level is 2, so the common orthogonal array $L_{12}(2^{11})$ is selected. The variable arrangement of $L_{12}(2^{11})$ is shown in Table 1, including 12 orthogonal experiments. Subsequently the signal-to-noise ratios are computed according to Eq. (12) for feature optimization.

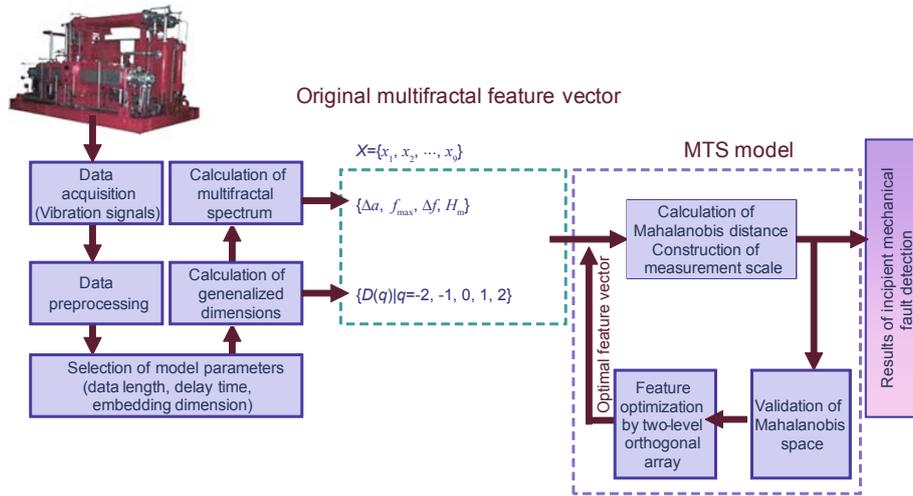


Fig. 2 Fault detection model based on multifractal theory and MTS

Table 1 Orthogonal array $L_{12}(2^{11})$ with the variable arrangement

No.	1	2	3	4	5	6	7	8	9	10	11
Variable	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}
Experiment 1	1	1	1	1	1	1	1	1	1	1	1
Experiment 2	1	1	1	1	1	2	2	2	2	2	2
Experiment 3	1	1	2	2	2	1	1	1	2	2	2
Experiment 4	1	2	1	2	2	1	2	2	1	1	2
Experiment 5	1	2	2	1	2	2	1	2	1	2	1
Experiment 6	1	2	2	2	1	2	2	1	2	1	1
Experiment 7	2	1	2	2	1	1	2	2	1	2	1
Experiment 8	2	1	2	1	2	2	2	1	1	1	2
Experiment 9	2	1	1	2	2	2	1	2	2	1	1
Experiment 10	2	2	2	1	1	1	1	2	2	1	2
Experiment 11	2	2	1	2	1	2	1	1	1	2	2
Experiment 12	2	2	1	1	2	1	2	1	2	2	1

5 Case studies

5.1 Preparation of experimental data

The method presented above was applied to incipient fault detection of oil pumps in the Tarim Oilfield, China. Signal samples of the normal running state, vibration induced by air clearance in a sliding bearing, and race wear of a rolling bearing were acquired from condition monitoring with a sampling frequency of 4,000 Hz. The waveforms in

time domain are shown in Figs. 3-5, respectively, with the data length N of 1,024. Original signals were preprocessed by “db8” wavelet for de-noising. Although there were some differences between the two fault states and the normal state in time domain figures, it was still difficult to identify and distinguish these different states with some quantitative indices. Therefore, it was necessary to make use of multifractal theory to extract the quantitative features of incipient faults.

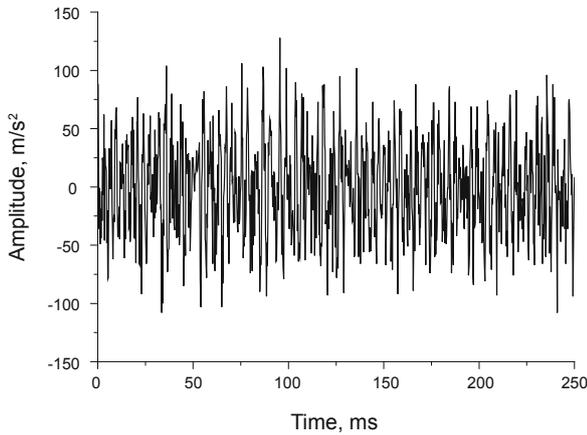


Fig. 3 Normal operational condition

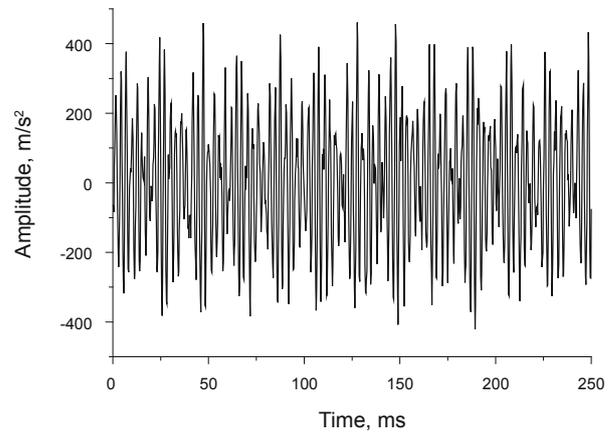


Fig. 5 Race wear of a rolling bearing

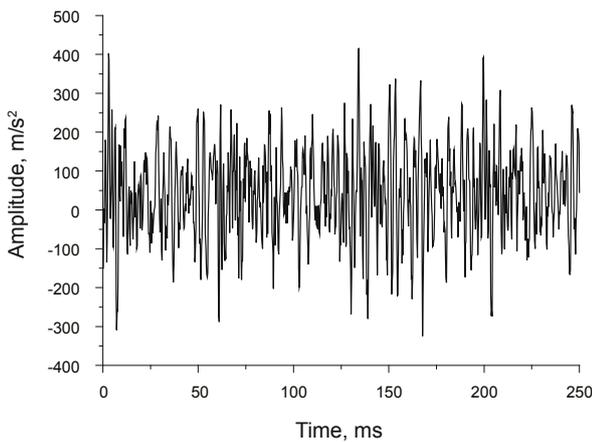


Fig. 4 Air clearance in a sliding bearing

5.2 Fault feature extraction of oil pumps based on multifractal theory

5.2.1 Main parameter selection of the detection model

(1) Selection of delay time τ . The normal data sample was analyzed for phase space reconstruction with different delay times τ , and the evolution process of the phase space trajectory could explain the relationship between delay time and phase space extension, which is shown in Fig. 6. When $\tau = 3, 10$, the phase trajectory extends to its maximal degree; when $\tau = 6, 7$, the phase trajectory shrinks to its minimal degree along the auxiliary diagonal; when $\tau = 1, 12$, the phase trajectory shrinks to its minimal degree along the main diagonal. According to the principle of pseudo-phase portrait method, the delay time τ was set as 3.

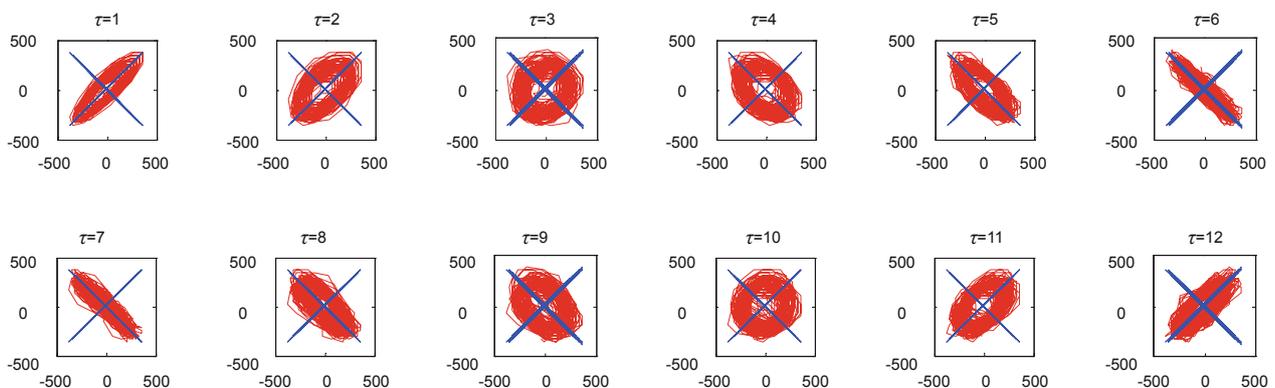


Fig. 6 Pseudo-phase portrait of vibration signal with $\tau = 1$ to $\tau = 12$

(2) Selection of embedding dimension m . According to the Takens theorem and using the trial and error, when weight factor was firstly fixed as $q = 2$, and m was set from 4 to 20, the corresponding double logarithmic plot $\ln(r) - \ln(c(q, r))$ is shown in Fig. 7, and the corresponding fractal dimensions $D(2)$ with different m were calculated

respectively, which are shown in Table 2. The double logarithmic plot $\ln(r) - \ln(c(q, r))$ tends to converge gradually as the embedding dimension m increases from 4. When $m = 10$, the system becomes steady, and the fractal dimension reaches a stable value 3.54. Therefore, the embedding dimension was set as $m = 10$.

Table 2 Fractal dimension with different embedding dimensions

m	4	5	6	7	8	9	10	11	12
$D(2)$	2.3564	2.5387	2.7568	3.0233	3.1959	3.2541	3.5398	3.5452	3.5479

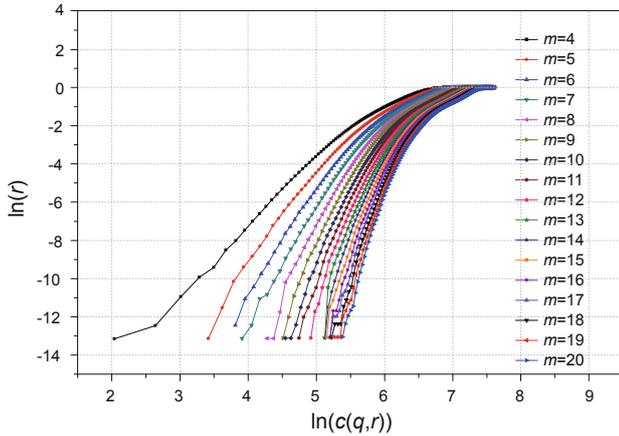


Fig. 7 $\ln(r) - \ln(c(q, r))$ double logarithmic plot

5.2.2 Multifractal feature extraction

Considering the data samples of the above three mechanical states, based on multifractal theory, the weight

factor was set as $q \in [-20, 20]$, and then the generalized dimensions were calculated by Eqs. (1)- (5). When the weight factor q was set as $-2, -1, 0, 1, 2$ respectively, the calculated generalized dimensions of three mechanical running states were prominently different, which could represent the characteristics of mechanical states. Hence, the corresponding generalized dimensions as $D(-2), D(-1), D(0), D(1),$ and $D(2)$ served as five multifractal features. Secondly, multifractal spectrum was calculated based on Eqs. (6)-(8), and then the width of the multifractal spectrum $\Delta\alpha$, the difference of the fractal dimensions Δf between the maximum probability and minimum probability subsets, the maximum of the multifractal spectrum f_m , and the multifractal spectrum entropy H_m were extracted as another four multifractal features. Finally, the calculated multifractal feature vectors ($D(-2), D(-1), D(0), D(1),$ and $D(2), \Delta\alpha, f_m, \Delta f, H_m$) were considered as incipient fault features, as shown in Table 3, corresponding to the above three different mechanical running states.

Table 3 Calculated multifractal feature vectors of three mechanical running states

Multifractal features	$D(-2)$	$D(-1)$	$D(0)$	$D(1)$	$D(2)$	$\Delta\alpha$	f_m	Δf	H_m
Normal state	3.02	2.86	2.61	2.43	2.25	6.04	36.74	-4.29	4.19
Air clearance in a sliding bearing	5.04	4.93	4.79	4.63	4.47	11.58	57.83	-7.40	4.81
Race wear of a rolling bearing	4.37	4.25	4.03	3.82	3.61	9.56	50.71	-6.37	4.54

5.3 Feature optimization and fault identification by MTS

The statistical rule of multifractal characteristic distribution was analyzed by the Mahalanobis distance discrimination method. Nine kinds of multifractal features from 30 normal samples were utilized as original data to construct the Mahalanobis space. In the fault identification experiment, 44 observational signals (20 samples of normal state, 16 samples of race wear of a rolling bearing, and 8 samples of air clearance in a sliding bearing) were used to calculate the Mahalanobis distances, which are shown in Fig. 8. The Mahalanobis distance interval $(0, 1.3), [1.3, 4),$ and $[4, +\infty)$ were separated to represent the normal state, race wear of a rolling bearing, and air clearance in a sliding bearing respectively. Preliminary identification results shown in Table 4 indicate that there are five groups in misjudgment according to the actual working conditions, and the identification accuracy is 88.6%.

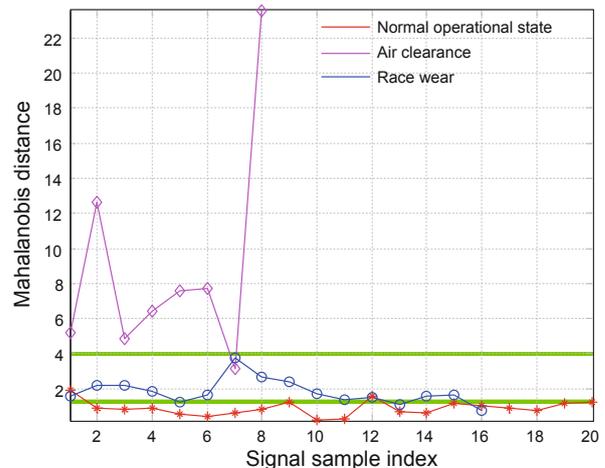


Fig. 8 Mahalanobis distances of observational signals with original features

Table 4 Preliminary identification result of MTS

		Practical running states of oil pump		
		Normal state	Race wear	Air clearance
Mahalanobis distance interval	(0, 1.3)	18	2	0
	[1.3, 4)	2	14	1
	[4, +∞)	0	0	7
Number of misjudgment		2	2	1
Number of signal samples		20	16	8

In order to improve the identification performance, the above features were optimized to select the most effective ones. The SNR of each feature on level 1 and level 2 were computed respectively, based on the orthogonal array $L_{12}(2^{11})$ with the variable arrangement shown in Table 1. The SNR of level 1 related to each feature was subtracted from the

SNR of level 2 to obtain the results of difference shown in Table 5. From Table 5, the SNR of features $D(0)$ and $D(1)$ decrease when they participate in fault identification, which means that these two features play a negative role in the fault identification process, and should be eliminated from the reference group.

Table 5 SNR of multifractal features

Multifractal features	$D(-2)$	$D(-1)$	$D(0)$	$D(1)$	$D(2)$	$\Delta\alpha$	f_m	Δf	H_m
Difference of SNR	0.66	1.44	-0.82	-0.06	0.09	0.66	2.57	0.43	1.23

Finally, the optimal feature vectors ($D(-2)$, $D(-1)$, $D(2)$, $\Delta\alpha$, f_m , Δf , H_m) were utilized to construct the Mahalanobis space. Mahalanobis distances of 44 groups of observational signals were calculated over again, and the results are shown in Fig. 9, from which Mahalanobis distance intervals (0, 3.3), [3.3, 16) and [16,+∞) are mapped to normal state, race wear of a rolling bearing, and air clearance in a sliding bearing respectively. There is no misjudgment or false alarm, and the accuracy of final identification is 100%.

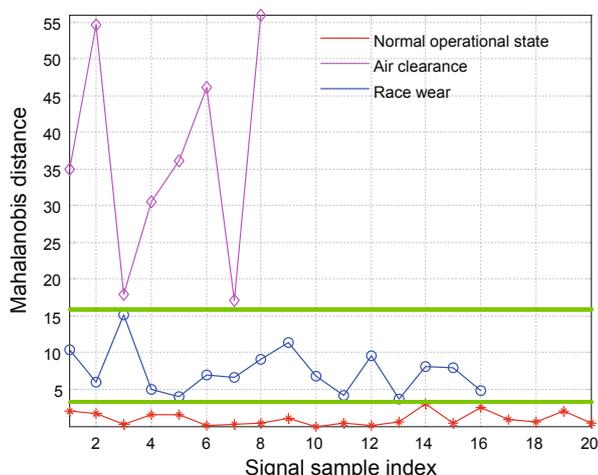


Fig. 9 Mahalanobis distances of monitoring signals with optimal features

6 Conclusions

1) A incipient mechanical fault detection method combining a multifractal method and Mahalanobis-Taguchi system (MTS) is proposed and used in incipient fault detection of oil pumps in the Tarim Oilfield, Xinjiang, China. The multifractal features of nonlinear signals are extracted based on generalized dimensions and multifractal spectrum through machine monitoring. According to MTS, the mechanical running states are distinguished effectively and the incipient faults are identified accurately by Mahalanobis distances. As the number of samples of the vibration spectra of pumps running with mechanical faults in low, the MTS feature optimization method is used to select the essential feature variables with greater contribution to fault identification based on SNR, and further improve the mechanical fault identification accuracy.

2) The case study indicates that the combination of a multifractal theory and MTS and its feature optimization scheme with a few fault samples can effectively identify the incipient faults of oil pumps with high accuracy.

Acknowledgements

The project is supported by the National High Technology Research and Development Program of China (Grant No. 2008AA06Z209); CNPC Innovation Fund (Grant No. 2006-A);

Special Items Fund of Beijing Municipal Commission of Education; and Program for New Century Excellent Talents, Ministry of Education (Grant No. NCET-05-0110).

References

- Chen G. Rough set-genetic algorithm-neural network composite classifier and its application in rotor faults diagnosis. *China Mechanical Engineering*. 2008. 19(1): 61-63 (in Chinese)
- Chen Z X, Xu J W and Yang D B. Denoising method of block thresholding based on DT-CWT and its application in mechanical fault diagnosis. *Chinese Journal of Mechanical Engineering*. 2007. 43(6): 200-204 (in Chinese)
- Das P and Datta S. Exploring the effects of chemical composition in hot rolled steel product using Mahalanobis distance scale under Mahalanobis-Taguchi system. *Computational Materials Science*. 2007. 38(4): 671-677
- Gao Q, Duan C, Fan H, et al. Rotating machine fault diagnosis using empirical mode decomposition. *Mechanical Systems and Signal Processing*. 2008. 22(5): 1072-1081
- Grassberger P and Procaccia I. Dimensions and entropies of strange attractors from a fluctuating dynamics approach. *Physica D: Nonlinear Phenomena*. 1984. 13(1-2): 34-54
- He Z and Han Y J. A study on the multivariate system diagnosis and analysis using Mahalanobis-Taghchi system. *Application of Statistics and Management*. 2007. 26(5): 830-839 (in Chinese)
- Jade A M, Srikanth B, Jayaraman V K, et al. Feature extraction and denoising using kernel PCA. *Chemical Engineering Science*. 2003. 58(19): 4441-4448
- Lee J M, Yoo C, Choi S W, et al. Nonlinear process monitoring using kernel principal component analysis. *Chemical Engineering Science*. 2004. 59(1): 223-234
- Li M and Vitanyi P M B. *A Introduction to Kolmogorov Complexity and its Applications*. New York: Springer-Verlag. 1997
- Lv Y, Li Y R and Wang Z G. New time series decomposition method and its application on machinery fault diagnosis. *Chinese Journal of Mechanical Engineering*. 2007. 43(8): 171-174 (in Chinese)
- Marseguerra M, Minoggio S, Rossi A, et al. Neural networks prediction and fault diagnosis applied to stationary and non-stationary ARMA modeled time series. *Progress in Nuclear Energy*. 1992. 27(1): 25-36
- Pedro S, Nuno F, Raquel C, et al. Fault identification in chemical processes through a modified Mahalanobis-Taguchi strategy. *Computer Aided Chemical Engineering*. 2004. 18: 799-804
- Peng Z K, He Y Y, Lu Q, et al. Wavelet multifractal spectrum: application to analysis vibration signals. *Chinese Journal of Mechanical Engineering*. 2002. 38(8): 59-63 (in Chinese)
- Qin Y, Qin S R and Mao Y F. Research on iterated Hilbert transform and its application in mechanical fault diagnosis. *Mechanical Systems and Signal Processing*. 2008. 22(8): 1967-1980
- Rubini R and Meneghetti U. Application of the envelope and wavelet transform analysis for the diagnosis of incipient faults in ball bearings. *Mechanical Systems and Signal Processing*. 2001. 15(2): 287-302
- Siddique A, Yadava G S and Singh B. Applications of artificial intelligence techniques for induction machine stator fault diagnostics: Review. In: *Proceeding of the IEEE International Symposium on Diagnostics for Electric Machines, Power Electronics and Drives*. New York. 2003. 29-34
- Sun R X, Tsung F G and Qu L S. Evolving kernel principal component analysis for fault diagnosis. *Computers & Industrial Engineering*. 2007. 53(2): 361-371
- Tamura M and Tsujita S. A study on the number of principal components and sensitivity of fault detection using PCA. *Computers & Chemical Engineering*. 2007. 31(9): 1035-1046
- Wang X, Ma J J and Wang S. Statistical state recognition of mechanical equipment using empirical mode decomposition. *China Mechanical Engineering*. 2008. 19(6): 704-708 (in Chinese)
- Wang Z H and Zhang L B. Diagnosis of incipient fault of engine with a fractal method. *Journal of the University of Petroleum, China (Edition of Natural Science)*. 2003. 27(1): 54-56 (in Chinese)
- Wu D J, Chiang P H, Chang Y W, et al. An expert system for fault diagnosis in internal combustion engines using probability neural network. *Expert Systems with Applications*. 2008. 34(4): 2704-2713
- Wu F J and Qu L S. An improved method for restraining the end effect in empirical mode decomposition and its applications to the fault diagnosis of large rotating machinery. *Journal of Sound and Vibration*. 2008. 314(3-5): 586-602
- Wu J D and Liu C H. Investigation of engine fault diagnosis using discrete wavelet transform and neural network. *Expert Systems with Applications*. 2008. 35(3): 1200-1213
- Xie W L and Xie W X. A fractal based analysis of time series data and feature extraction. *Journal of Xidian University*. 1997. 24(2): 172-179 (in China)
- Yu Y L, Xie S L, Cai H T, et al. *Introduction of New Signal Processing Methods*. Beijing: Tsinghua University Press. 2004. 184-190 (in Chinese)
- Zhao J, Xu L M and Liu L. Equipment fault forecasting based on ARMA model. *International Conference on Mechatronics and Automation*. 2007, 5-8 Aug. 3514 - 3518

(Edited by Sun Yanhua)