

1641. A novel method for self-adaptive feature extraction using scaling crossover characteristics of signals and combining with LS-SVM for multi-fault diagnosis of gearbox

Xingxing Jiang¹, Shunming Li², Yong Wang³

College of Energy and Power Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

¹Corresponding author

E-mail: ¹jiangxinlong1989@126.com, ²lism0527@163.com, ³wanyong198@126.com

(Received 22 November 2014; received in revised form 18 January 2015; accepted 5 February 2015)

Abstract. Vibration signals of defective gears are usually non-stationary and masked by noise. As a result, the feature extraction of gear fault data is always an intractable problem, especially for multi-fault couple system (two or more fault types simultaneously occur in mechanical systems). Recently, an interesting crossover characteristic of nonlinear data is used to diagnose the different severities of gear faults. Nonetheless, it lacks of self-adaptivity. Consequently, a novel method for self-adaptive feature extraction using scaling crossover characteristics of signals and combining with least square support vector machine (LS-SVM) for multi-fault diagnosis of gearbox is proposed. Firstly, detrended fluctuation analysis (DFA) is introduced to analyze fractal properties and multi-scaling behaviors of vibration signal from multi-fault gearbox. The scale exponents are abrupt changed with the gradual increasing of time scales, which can be observed in the scaling-law curve. Secondly, a criterion based on a Quasi-Monte Carlo algorithm is developed to uncover optimal scaling intervals of scaling-law curve. Several different scaling regions are objectively measured in each of which a single scale exponent can be estimated. Thirdly, a three-dimensional vector, containing three scale exponents which carry definite physical meaning, is used as the feature parameter to describe the underlying dynamic mechanism hidden in gearbox vibration data. Lastly, these vectors are classified by LS-SVM. Moreover, the method of statistical parameters is exploited to classify the multi-fault vibration data which have been investigated by proposed method. The results show that the proposed method is sensitive to multi-fault vibration data of gearbox with similar fault patterns and has a better performance than other methods.

Keywords: multi-fault diagnosis, DFA, crossover characteristic, feature extraction, LS-SVM.

1. Introduction

As a key component of rotating machinery, gears are mostly subjected to progressive deterioration due to severe working conditions [1]. The condition of a gear directly affects the normal running of the machine or even the entire system. The faults caused by gear failures account for 10 % of the malfunctions in rotating machines and 80 % in the transmission machinery, respectively. Therefore, it is necessary and critical that the detection and diagnosis of gear faults are performed to prevent breakdown accidents and to minimize production loss.

Gearbox fault diagnosis has received intensive study for several decades and yet a few lectures focus on multi-fault diagnosis of gearbox. For example, wind turbine gearboxes have multi-gearing and multi-bearing [2]. It is usually difficult to diagnose their potential faults. Especially when multi-faults coexist, vibrations excited by several faults are combined with each other non-linearity and non-stationary. Among the available diagnosis techniques, vibration analysis is the most commonly used and efficient method. However, it is easily influenced by vibration sources from many cases, such as the meshing gears, shafts, and bearings. As a result, the vibration signals are usually quite complex and their analyses are interfered. So far many developed fault detection techniques focus on the evolution of statistical parameters (such as standard deviation, skew and kurtosis, etc.) as a function of time [3], or on frequency analysis [4].

More recently, a series of hybrid time-frequency techniques have been developed, such as wavelet, Wigner-Ville or correlated transforms [5-10]. Both frequency and hybrid techniques rely on the identification of the existing fault frequency. However, when used to analyze the complex gearbox vibration data, the previous methods often produce unsatisfactory results because of their respective drawbacks [11-13] especially to multi-fault gears which have many same feature frequencies. Therefore, the further researches on these methods are in progress, particularly for multi-fault diagnosis.

Recently, De Moura et al. [14, 15] firstly made a constructive attempt to apply DFA to gear and bearing fault diagnosis. In paper [14], DFA was used as a transform tool to compress gearbox vibration data containing 2048 points into the scaling-law curve of 37 values. The extracted the scaling-law curve, by which the relation between the fluctuation function and the time scale can be illustrated graphically in a log-log plot, fully expressed various working conditions contained in the original gearbox data. Afterwards, in paper [15], bearing vibration data containing 4096 points were transformed into the scaling-law curve of 37 values and the extracted the scaling-law curve wholly represented various conditions of severity defects in the original bearing data. Eventually, DFA method was proved to be an excellent tool for data compression. However, only a mono-fractal characteristic was used in [14] and [15], so that it was barely able to expose the underlying nonlinear dynamical mechanism hidden in multi-fractal time series. Lin et al. [16, 17] used DFA to analyze the scaling behavior of gearbox vibration data by a two-dimensional vector containing two scale exponents which carried definite physical meaning and could be used as feature parameters to describe the gearbox vibration data. It showed that DFA method was a powerful tool for uncovering the nonlinear dynamical characteristics buried in non-stationary time series and could capture minor changes of complex system conditions. However, the detections of crossover points and scale regions were relatively subjective since it has been made without rigorous statistical procedures and has generally been determined by eye balling or subjective observation. Crossover points and scale regions determined liked these may be spurious and problematic. It may not reflect the genuine underlying scaling behaviors of a time series.

Multi-scaling behaviors are often found in nature [18]. The identification of these characteristic scales is relevant for a complete understanding of the underlying multi-scale dynamics [19]. As stated by Shao et al. [20]: "there is no consensus on an objective determination approach of the scaling range, which plays a crucial role in the estimation of the scaling exponents". If it has no solid theoretical foundations for the quantitative detection of the crossover time scales, it will produce a subtle issue when dealing with series which have more than one scaling behavior. The detections of crossover points and scale regions have been mainly done by the visualization of the log-log plot and the presence of seemingly different exponent values would be shown in the results [18]. Therefore, it is necessary to develop an independent criterion to estimate the optimal fitting regions of scaling-law curve. Crossover phenomena, i.e. the presence of crossover scales separating regimes with different scaling exponents, may be efficiently unveiled by applying regression analysis and statistical inference. One possibility is to study the local exponents by means of the log derivative plot and look for constant value regions, as suggested by Govindan et al. [21], Bashan et al. [22] and Lopez et al. [23]. However, the method only searched a single crossover point, making it difficult to apply in the cases of multi-crossover points. Additionally, Michalski [24] identified the optimal minimal and maximal scale sizes for persistent processes through a number of extensive Monte Carlo simulations. Moreover, Grech et al. [25] studied the relations between the scale region of artificial correlated and uncorrelated series and series length, Hurst exponent and goodness of fluctuation function linear fit in a given confidence level. Unfortunately, these methods needed to consume much time. In follows, a related approach using linear fit as a way to locate optimal scaling regions was proposed in paper [26]. This criterion based on goodness of linear fit might be naturally extended to other techniques requiring linear fits. It was not computationally prohibitive since it deals with relatively few points. The method will be developed and used in our study for the multi-fault diagnosis. The criterion grounded on a solid statistical foundation can describe multi-scaling behaviors of fractals.

Through the regression analysis and statistical inference, we can identify the crossover time scales that cannot be detected by eye-balling observation and determine the number and locations of the genuine crossover points.

To evaluate the performance of the feature parameters characterizing the multi-fault gearbox, LS-SVM is suitably employed to differentiate the different fault types of gearbox. LS-SVM is an evolution of SVM. The simplicity and inherited advantages over SVM, such as excellent generalization ability and a unique solution, promote the applications of LS-SVM. The most important difference between them is that the loss function of LS-SVM is least square linear system rather than quadratic optimization. LS-SVM has been widely used as classifiers in machine fault diagnosis [27, 28]. LS-SVM will be regarded as classifiers in our study.

The paper is organized as follows. In Section 2, the DFA method is briefly described. Besides, the determinate criterion and identification method of crossover scales are presented. Then, the fame of gearbox multi-fault diagnosis is proposed. Section 3 provides the introduction for all types of vibration signals from multi-fault gearbox. The achieved results of proposed method compared with the analysis results of statistical parameters are investigated in Section 4. An in-depth discussion about the proposed method is given in Section 5. Conclusions of the proposed novel method for multi-fault diagnosis are drawn in the final section.

2. Theoretical backgrounds

2.1. Describe detrended fluctuation analysis (DFA)

Since the seminal study on DFA was proposed by Peng et al. [29], this technique has been widely used in time series analysis. The DFA aims to improve the evaluation of correlations in a time series by eliminating trends in the data. The method is described as follows.

1) Consider a signal $x(i)$, where $i = 1, \dots, N$, and N is the length of the signal. An integrated time series $\tilde{x}(i)$ of $x(i)$ is obtained as:

$$\tilde{x}(i) = \sum_{j=1}^i (x(j) - \langle x \rangle), \quad (1)$$

where $\langle x \rangle$ is the mean of $x(i)$ given as:

$$\langle x \rangle = \frac{1}{N} \sum_{j=1}^N x(j). \quad (2)$$

2) Split series $\tilde{x}(i)$ into $N_s = [N/s]$ non-overlapping windows with data points. $[\cdot]$ denotes the round numbers. If N is not divisible by s , there are some remaining $r = N - N_s s$ values at the end of the series $\tilde{x}(i)$. To solve this problem, the other N_s segments are acquired but starting from $\tilde{x}(r + 1)$. In this way, $2N_s$ windows of s data points are obtained.

3) Let v be the index of the $2N_s$ window. For each one of these windows it takes the polynomial $y_v(i)$ of degree m to fit the data in the window where i is the data index. Then the $2N_s$ local variance is obtained as Eq. (3). If N is divided by s , it simply repeats the N_s values of $F^2(v, s)$:

$$F^2(v, s) = \frac{1}{s} \sum_{i=1}^s \{\tilde{x}[(v-1)s+i] - y_v(i)\}^2, \quad 1 \leq v \leq N_s, \quad (3a)$$

$$F^2(v, s) = \frac{1}{s} \sum_{i=1}^s \{\tilde{x}[(v-N_s-1)s+r+i] - y_v(i)\}^2, \quad N_s + 1 \leq v \leq 2N_s, \quad (3b)$$

where $v = 1, \dots, 2N_s$, must be the same for every step of this technique, determining the detrend polynomial order of analysis.

The averaged root mean square value, i.e. fluctuation function $F(s)$ for the entire time series is given as:

$$F(s) = \sqrt{\frac{1}{2N_s} \sum_{v=1}^{2N_s} (F(v, s))^2} \tag{4}$$

4) Repeated all the above steps 1-3 for several values of s in order to construct the relation between fluctuation function $F(s)$ and time scale s . The relation is presented in the form as:

$$F(s) = s^\alpha \tag{5}$$

By a double logarithmic operation, the parameter α which is called the Hurst exponent or scale exponent can be evaluated. Simultaneously the relation between the fluctuation function and the time scale can be illustrated graphically by the scaling-law curve in a log-log plot. The value of $\alpha = 0.5$ indicates that there are no correlations or only short-term correlations [17]. If $\alpha > 0.5$, the data are long-term correlated. With the increasing of α the persistent long-range correlation of data is enhanced. The case of $\alpha < 0.5$ corresponds to long-term anti-correlations, meaning that large fluctuation is more possible to be followed by small values.

In fact, the complex fractal time series generally do not exhibit mono-scaling behavior characterized by a single scaling exponent. Multi-scaling behaviors are actually very common in natural phenomena. An example for vibration signal of multi-fault gearbox and its scaling-law curve analyzed by DFA is shown in Fig. 1. There are two obvious crossover points which are shown as the square points described in Fig. 1(b). As a result, the two crossover points can be employed to divide the whole scaling-law curve into three different scale regions, in each of which a single scale exponent can be estimated. However, it hardly determines the scale exponent without a theoretical model for multi-scaling behaviors. If the determinations of these regions are subjective, the results derived from different analysis of the same series can not be consistent. Therefore, an independent criterion to estimate the optimal scale region of scaling-law curve is developed in the Subsection 2.2.

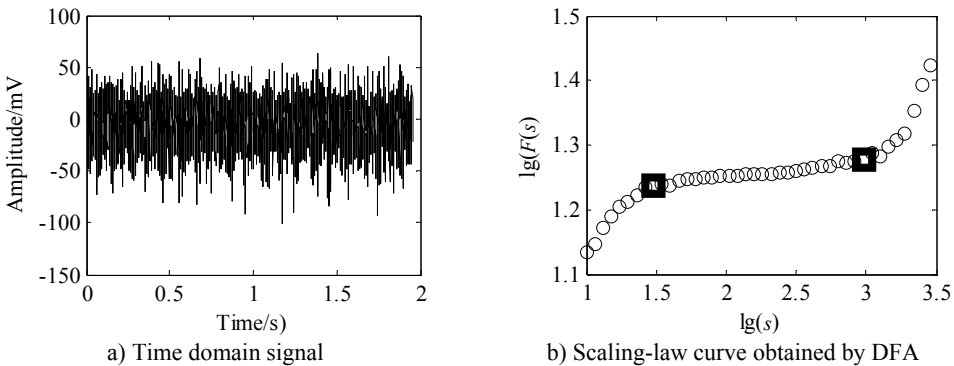


Fig. 1. Example for the vibration data of multi-fault gearbox with multi-scaling behaviors

2.2. The method of uncovering optimal scale intervals

The optimal scale regions of time series with crossover characteristics can be objectively and efficiently estimated by using the method of Quasi-Monte Carlo. The Quasi-Monte Carlo method is described as follows.

- 1) Suppose that $x(i)$ is a time series which has M given scales $s \in \{s_1, s_2, \dots, s_M\}$ where the fluctuation function $F(s) \in \{F(s_k)\}$, ($1 \leq k \leq M$) is calculated.
- 2) Define a number of δ data points. In practice, the parameter δ must be less than the length of minimum scale region. A detailed discussion with the selection of parameter δ is given at the end of this subsection.
- 3) Calculate the logarithmic series $L_s \in \{\log(s_k)\}$ and $L_F \in \{\log(F(s_k))\}$, ($1 \leq k \leq M$).
- 4) Define a matrix $r_{M \times M}$ of default values equal to 0.
- 5) Compute all non-zero elements of $r_{M \times M}$ according as:

$$r_{ij} = R^2(s_i, s_j), \tag{6}$$

where $R^2(s_i, s_j)$ is the coefficient of determination with the linear fit of L_F versus L_s between the logarithmic time scale $\log(s_i)$ and $\log(s_j)$, $1 \leq i \leq (M - \delta + 1)$ and $i + \delta - 1 \leq j \leq M$.

6) Sort all non-zero values r_{ij} in decreasing order while keeping a record of their original subindices. The first element of this list will then provide the first optimal fitting scale interval. If there are repeated values in $r_{M \times M}$, sorting requires an extra criterion that the longest interval length is selected. The slope of linear fit is regard as scale exponent α of the scale interval.

7) Eliminate searched interval in step 6. Then, repeat above steps 2 to 6. The next optimal region can be searched in remaining interval until all linear regions are uncovered.

This algorithm is called as Quasi-Monte Carlo since it actually performs all possible fits at least δ data points between scales s_1 and s_M . In order to more clearly clarify how to operate the method for multi-fault vibration signals, there are two projects used to illustrate the operational process subsequently.

A simple artificial series is performed by this method. The 126 equally spaced points x_i are generated in interval of $[1, 3.5]$. Then, the Eq. (7) is calculated.

$$y_i = \begin{cases} 0.55x_i + \sigma_{1i}, & 1 \leq x_i < 1.5, \\ 0.525 + 0.2x_i + \sigma_{2i}, & 1.5 \leq x_i < 3, \\ -1.575 + 0.9x_i + \sigma_{3i}, & 3 \leq x_i < 3.5, \end{cases} \tag{7}$$

where $\sigma_{1i}, \sigma_{2i}, \sigma_{3i}$ are generated as a set of uniform random numbers of variances $\sigma_1, \sigma_2, \sigma_3$ used as the additive noise. Then the artificial fluctuation functions $F(s_i)$ and scales s_i are generated with series x_i and y_i according to:

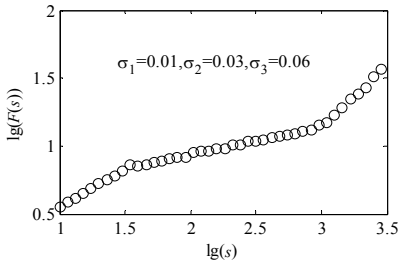
$$s_i = 10^{x_i}, \quad F(s_i) = 10^{y_i}. \tag{8}$$

Consider two kinds of situations. One is that the relationship of variances for random numbers $\sigma_{1i}, \sigma_{2i}, \sigma_{3i}$, is $\sigma_1 \leq \sigma_2 \leq \sigma_3$ and the other one is $\sigma_2 \leq \sigma_1 \leq \sigma_3$. The δ data points are set as 18. The analysis results of artificial series by the method of uncovering optimal fitting scale intervals are shown in Fig. 2 and Fig. 3 respectively. It can be seen that that the fitted curves are perfectly coincident with original data depicted in Fig. 2(b) and Fig. 3(b). The slopes of two examples are (0.5464, 0.2021, 0.9377) and (0.2002, 0.5505, 0.9228) which approximate to the genuine value (0.55, 0.20, 0.90). Moreover, the order of the obtained linear region is in accordance with the sizes of $\sigma_1, \sigma_2, \sigma_3$. As seen in Fig. 2(c) and Fig. 3(c), the coefficients of determination approach to one. It demonstrates that the method of uncovering optimal scale intervals is reliable.

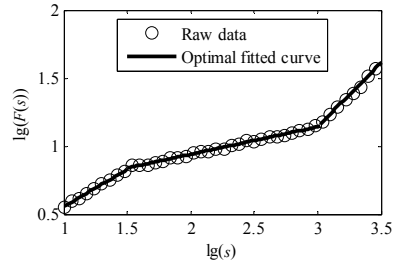
It should be remarked that the key of the method for uncovering optimal scale intervals is to set up the parameter δ . As mentioned in paper [26], the computing time t for this implementation might be approximated by:

$$t \approx k \cdot \frac{(M - \delta + 1)^2}{\#C}, \tag{9}$$

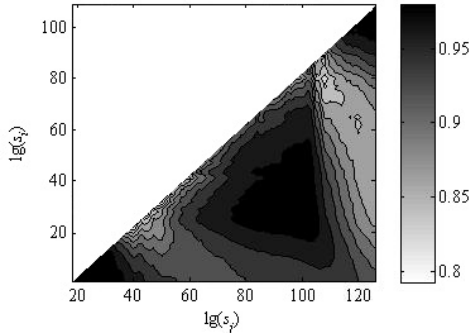
where k is constant. $\#C$ is the number of active cores. This inverse proportionality is due to the fact that elements in the matrix $r_{M \times M}$ are independent of each other, therefore enabling the possibility to split the processing task.



a) Series with noise $\sigma_1 \leq \sigma_2 \leq \sigma_3$

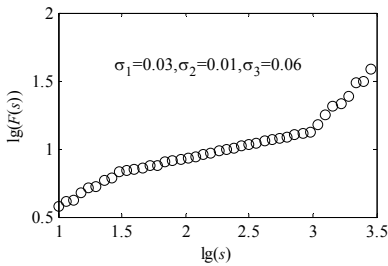


b) Fitting result by the method of uncovering optimal fitting scale intervals

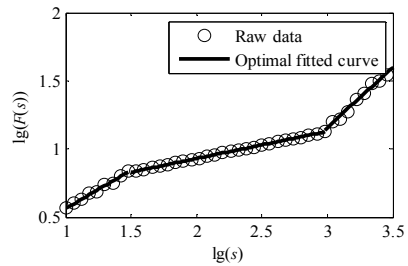


c) Coefficient of determination of the linear fit

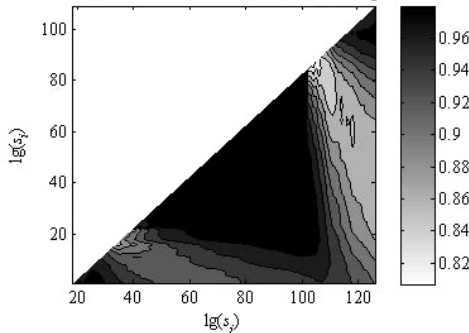
Fig. 2. Analysis results of series with noise



a) Series with noise $\sigma_2 \leq \sigma_1 \leq \sigma_3$



b) Fitting result by the method of uncovering optimal fitting scale intervals



c) Coefficient of determination of the linear fit

Fig. 3. Analysis result of series with noise

According to Eq. (9), the large parameter δ is conducive to reducing computing time. Moreover, the reasonable large parameter δ makes more obvious statistical sense for the linear fit. However, the parameter δ should be less than the data points of the shortest scale regions. Because the parameter δ is too large it could cause a poor statistical sense. As delineated in Eq. (7), the data points of the scale region are 42. Therefore, the δ data points are set as 18 which is a tradeoff between computing time and statistical sense.

2.3. The method of LS-SVM

As an improvement of Vapnik's standard SVM [30], LS-SVM classifier [31] leads to solve a linear system instead of a quadratic optimization method. Its objective function is formulated as:

$$\min \Phi(\omega, \xi) = \frac{1}{2} \|\omega\|^2 + \frac{C}{2} \sum_{i=1}^n \xi_i^2 \text{ s. t. } z_i [(\omega^T \cdot \varphi(y_i)) + \omega_0] = 1 - \xi_i, \quad i = 1, \dots, n, \quad (10)$$

where $z_i \in \{-1, 1\}$ stands for the i th desired output. $\varphi(y_i)$ denotes a nonlinear mapping of the i th input sample of the training data set $\{y_i, z_i\}_{i=1}^n$ from the primal space to the feature space. C is viewed as a form of a "regularization" parameter which controls the tradeoff between the complexity of the machine and the number of non-separable points. ω is a vector which determines the orientation of a separating hyperplane. ξ_i is the slack variables.

In general, ω may potentially become infinite dimensional so that a separating hyper-plane does not exist. Hence the constrain is relaxed by introducing slack variables $\xi_i \geq 0$. The corresponding Lagrangian equation is built as:

$$Q(w, b, \beta, \xi) = \frac{1}{2} \|\omega\|^2 + \frac{C}{2} \sum_{i=1}^n \xi_i^2 - \sum_{i=1}^n \beta_i \{z_i [(\omega \cdot \varphi(y_i)) + \omega_0] - 1 + \xi_i\}, \quad (11)$$

where β_i ($i = 1, \dots, n$) is the Lagrange multiplier.

Then support vectors are got by Karush-Kuhn-Tucker (KKT) condition. The classifier in the dual space takes the form as Eq. (12). The detailed procedure can consult in paper [32]:

$$z(y) = \text{sign} \left[\sum_{i=1}^n \beta_i z_i K(y, y_i) + b \right], \quad (12)$$

where $K(y, y_i)$ is called the kernel function that must satisfy Mercer condition.

2.4. Frame of the novel method for feature extraction using crossover characteristics

As a result, the method of uncovering optimal fitting scale intervals can be employed to divide the whole scaling-law curve into several different scaling regions, in each of which a single scale exponent can be estimated. Taking advantage of the natural crossover characteristics of the scaling-law curve, a novel method for self-adaptive feature extraction using crossover characteristics is proposed in this study. The whole procedure of the proposed novel method can be decomposed as the following four steps:

- 1) Analyze the vibration signal of multi-fault gearbox by DFA and obtain the scaling-law curve described in Subsection 2.1.
- 2) Segment the entire scaling-law curve into several different scaling regions by the method of uncovering optimal scale intervals demonstrated in Subsection 2.2.
- 3) Extract the scale exponents in each scale region as the feature parameters and utilize them to characterize the fault types.

4) Use the method of LS-SVM to classify the feature parameters obtained in the third step.

3. Capture vibration signals of multi-fault gearbox

Gear failures including the localized faults (pit, broken) and distributed fault (worn), as well as coupled fault in power train perhaps cause catastrophic accidents. Therefore, an early recognition of the gear faults is critical for normal operation of a gearbox [33]. Our paper focuses on investigating the multi-fault gearbox. As illustrated in Fig. 4, all trials are performed on a specially designed bench which is composed of a one phase input and three-phase output motor (with nominal power of 0.75 kW and nominal rotation frequency of 1500 rpm), the shaft supporting seats, a flexible coupling, a gearbox and a magnetic powder brake. The rotation of motor is controlled by a frequency inverter. The maximum braking torque of magnetic powder brake is 5 N·m. The gearbox contains two gears (pinion and wheel gear). The gear parameters are displayed in Table 1.

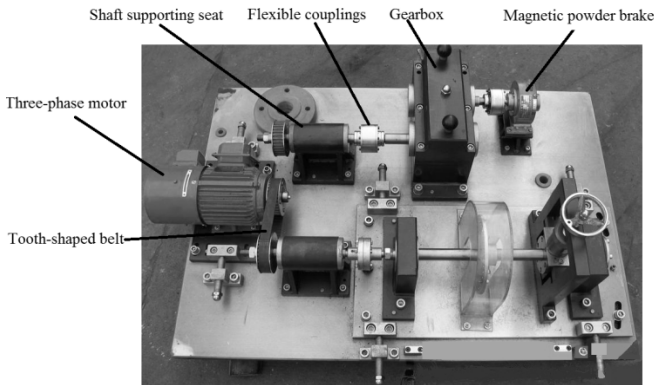


Fig. 4. Bench of multi-fault gearbox

There are six fault types of gearbox: normal, a single broken tooth of wheel, a single pit of wheel, a single worn of pinion, coupled fault of wheel pit and pinion worn, coupled fault of wheel broken and pinion worn which are considered in this experimental case. For brevity, the six typical fault types of gearbox are named as Type-1, Type-2, Type-3, Type-4, Type-5 and Type-6 respectively. Two kinds of rotating conditions (880 rpm and 1500 rpm) are employed for these six fault types of gearbox. When the rotational speed of the motor is 880 rpm, there are four kinds of loads which are the absence of external load and loads of three different currents of magnetic powder brake respectively. Hereby, load-1, load-2 and load-3 represent 0.2 ampere (A), 0.1 A and 0.05 A current of magnetic powder brake, respectively. When the rotational speed is 1500 rpm, a unique absence of external load is set up for all fault types of gearbox. Therefore, each fault type of gear contains five running conditions. 44 data samples are collected for each working condition of one fault type in this experiment. So a total of 1320 data samples are obtained on the designed bench. The sensor used is a piezoelectric accelerometer (DH131E) mounted on the flat surface of gearbox. The sampling frequency is 5120 Hz for all conditions. Each data sample is composed of 10000 data points.

A group of vibration signals for six fault types of gearbox under load-1 with rotational speed of 880 rpm are expressed in Fig. 5.

Table 1. Gear parameters

Gear	Teeth	Module (mm)	Pressure angle (deg.)	Materials
Pinion	55	2	20	S45C
Wheel	75	2	20	S45C

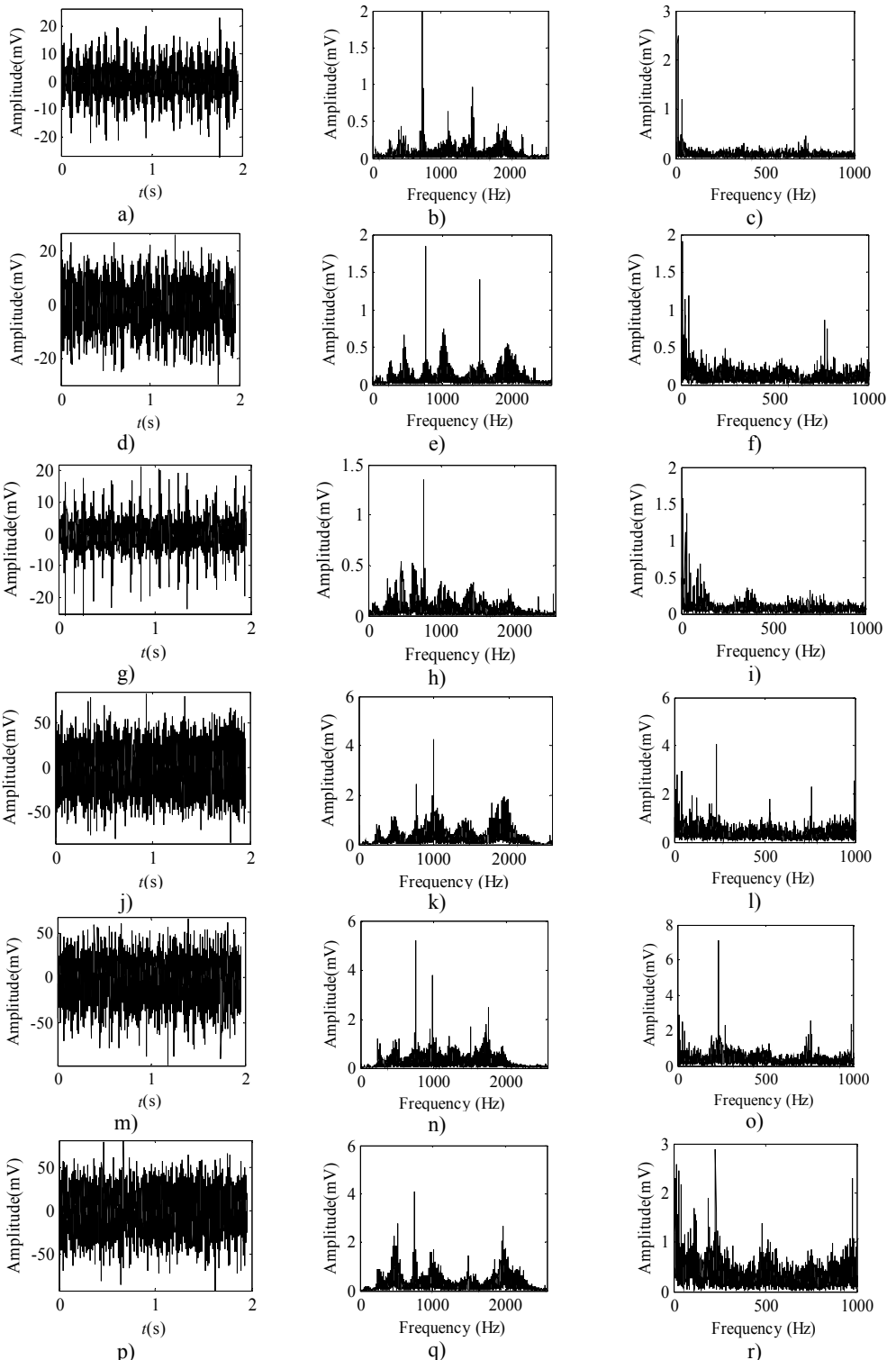


Fig. 5. Acquired raw data, FFT spectrum and envelope spectrum under load-1 with rotational speed of 880 rpm: a), b), c) Type-1; d), e), f) Type-2; g), h), i) Type-3; j), k), l) Type-4; m), n), o) Type-5; p), q), r) Type-6

It can be seen from the first column of Fig. 5 that the raw vibration signals of all fault types of gearbox are heavily contaminated by noise. In general, the feature component of periodic impulse is often encountered when faults occur on gear teeth. However, this characteristic is not obvious in the time waveforms of the raw data except the pit of wheel shown in Fig. 5(g). In addition, the Fast Fourier Transform (FFT) spectral for gear faults should be marked by both the rotational frequency and gear meshing frequency as well as their harmonic based on the study of mechanism of gear faults [34]. As demonstrated in the second column of Fig. 5, this phenomenon cannot be observed well in FFT spectra of all fault types of gearbox. Moreover, the vibration signals for the different fault types of gearbox have almost the same representation in frequency domain. Similarly, the distinct features cannot be obtained from the envelope spectra of all fault types described in the third column of Fig. 5, even if Type-1 can be differentiated from the other types. As a result, it is infeasible and unreliable to recognize the gear states through the waveforms of FFT spectrum and envelope spectrum. Therefore, the proposed method described in Subsection 2.4 is used to diagnose the multi-fault gearbox in Section 4.

4. Experiment results for multi-fault diagnosis based on the proposed novel method

This section is made of two subsections. According to scaling crossover characteristics of multi-fault data, the feature parameters obtained from optimal fitting scale intervals are displayed in Subsection 4.1. The performances of the proposed novel method for the multi-fault diagnosis of gearbox are evaluated in Subsection 4.2. Furthermore, compared with the proposed method, some statistical parameters (pulse factor, kurtosis and form factor) in the time domain and the combination of scale exponents and statistical parameters are also explored to characterize the multi-fault data in Subsection 4.2.

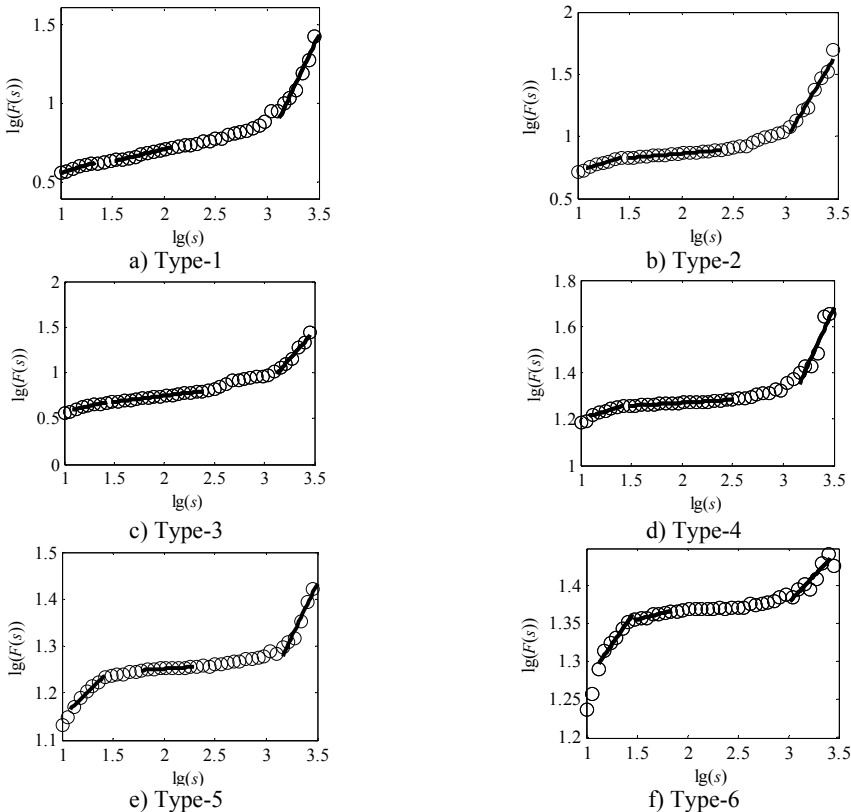


Fig. 6. Analysis result of uncovering optimal scale intervals under load-1 with rotational speed 880 rpm

4.1. Extract the feature parameters from optimal fitting scale intervals

The vital step of the proposed method is the extraction of feature parameters. In this subsection, the method of uncovering optimal scale intervals in Subsection 2.2 is used to analyze the vibration data of multi-fault gearbox. As suggested in paper [16], the degree of DFA introduced in Subsection 2.1 which can be obtained by attempting different orders is finally set as one. Additionally, the range for data points of non-overlapping windows is from 10 to 3000, i.e. the interval of logarithmic s ($\log(s)$) is set to range from 1 to 3, described in paper [17]. In our study, under the same assumptions, the degree of DFA is 1 and the range for the data points of non-overlapping windows is from 10 to 3162, i.e. the interval of $\log(s)$ is uniformly set to range from 1 to 3.5 with the spacing of 0.02 to observe the analysis results in the larger windows. As depicted in Fig. 1(b) and Fig. 6(a)-(f), the scaling-law curves obtained by DFA for all fault types of gearbox clearly exhibit two crossover points. The scaling-law curves can be roughly divided into three scaling regions by using these crossover points where the x -coordinates of two crossover points are close to $\log(s_{25})$ and $\log(s_{100}) = 3$ respectively. Therefore, the ranges of three scale regions approximate to $[s_1, s_{25}]$, $[s_{25}, s_{100}]$, and $[s_{101}, s_{126}]$ respectively. As displayed in Fig. 6(a)-(f), the first and third linear regions have a shorter length than the second region. The shortest length of these linear intervals approximates to 25. To uncovering the optimal scale intervals, the δ data points are set as 18 which is a tradeoff between computing time and statistical sense as explained in Subsection 2.2.

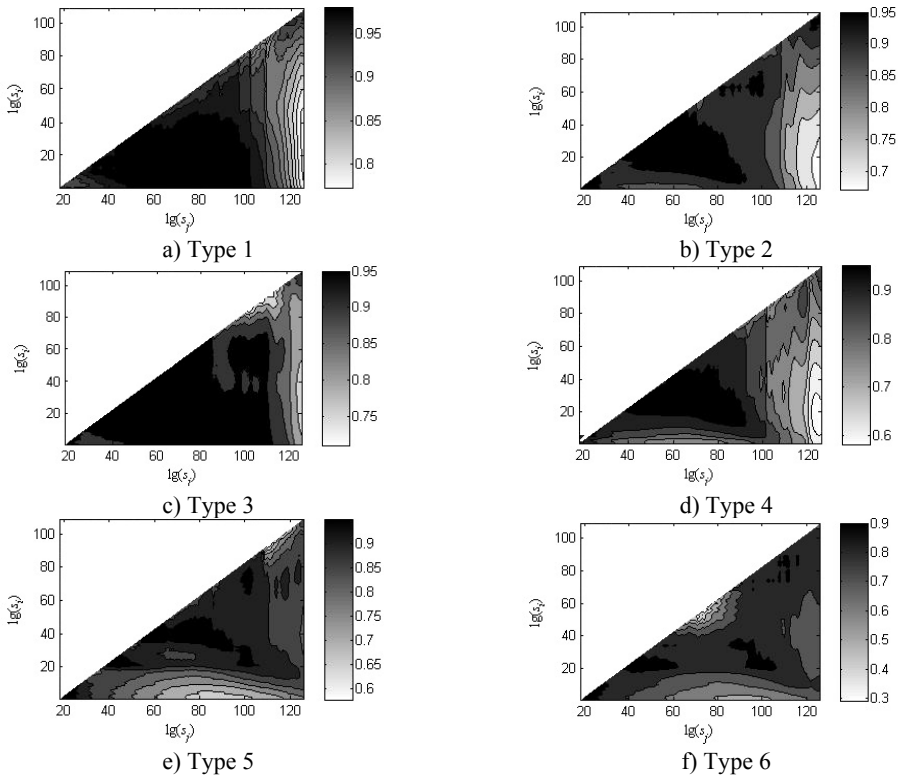


Fig. 7. Coefficients of determination (R^2) of the linear fitting under load-1 with rotational speed 880 rpm

As shown in Fig. 6, the black heavy lines represent the optimal scale intervals fitted by the criterion based on a Quasi-Monte Carlo algorithm. Fig. 6(a)-(f) manifest that the optimal fitted lines can completely depict the outline of scaling-law curves. It can be seen from Fig. 7(a)-(f) that the coefficients of determination before the non-overlapping window of $\log(s_{100})$ are more

apparent than the remaining parts. That is, the linearity of first and second linear regions $[s_1, s_{25}]$, $[s_{25}, s_{100}]$ has a better performance than the one of the third region $[s_{101}, s_{126}]$. As a consequence, the order of the obtained linear regions may be $[s_1, s_{25}]$, $[s_{25}, s_{100}]$ and $[s_{101}, s_{126}]$ or $[s_{25}, s_{100}]$, $[s_1, s_{25}]$ and $[s_{101}, s_{126}]$ which are consistent with analysis in Subsection 2.2. Lastly, three scale exponents, in turn, are estimated from these linear regions by least square fitting. Afterwards three scale exponents are determined, the sequence of scale exponents should be rearranged to conform the original order $[s_1, s_{25}]$, $[s_{25}, s_{100}]$ and $[s_{101}, s_{126}]$ for uniformly forming a standard vector employed in Subsection 4.2 as the input samples. Table 2 reports the three scale exponents for six fault types of gearbox under load-1 with rotational speed of 880 rpm. As reported in Table 2, the scale exponents in the first two scaling regions are less than 0.5 and hence the corresponding data exhibit extremely strong anti-persistent long-range correlations while the opposite situation is shown in third linear region. In next subsection, we will present the discriminability of the features vector V_1 which is expressed as:

$$V_1 = [SE1, SE2, SE3], \tag{13}$$

where SE1, SE2 and SE3 mean the three scale exponents in sequence.

Table 2. Three scale exponents for six fault types of gearbox under load-1 with speed of 880 rpm

Fault type	SE1	SE2	SE3	Fault type	SE1	SE2	SE3
Type 1	0.1798	0.1938	1.6275	Type 4	0.1189	0.0187	1.3390
Type 2	0.2319	0.0707	1.6368	Type 5	0.2107	0.0141	0.4584
Type 3	0.2413	0.1327	1.2454	Type 6	0.1717	0.0364	0.1584

4.2. Compare the classified performance of proposed method with statistical parameters

In general, it is difficult to identify fault types by directly observing features. In this section, LS-SVM are used to realize the recognition for multi-fault gearbox.

As described in Subsection 2.3, the primary principle of LS-SVM which is an intelligent learning algorithm and derived from the work of Vapnik et al. [30, 35] is to map the input from the primal space to the feature space by kernel function shown in Eq. (12). The selection of kernel function $K(y, y_i)$ will affect the learning ability and generalization of LS-SVM. Since the Gaussian Radial Basis Function (RBF) kernel always have a superior performance than the other kinds of kernel functions in many practical applications [36], the RBF kernel which is shown in Eq. (13) is adopted in our research. As well, the RBF kernel is a better choice than other kernels like polynomial kernel because it has lesser hyper-parameters and so the problem becomes less computationally intensive. In addition, the LS-SVM is designed intrinsically for binary classification which is not suitable for multi-class of multi-fault diagnosis. However, in our paper, a six-class classification problem should be implemented by extending the binary classification. Many approaches have been considered for multi-class classification issue, such as “one-against-one”, “one-against-all”, and directed acyclic graph SVM [37]. “one-against-one” is to break down the multi-class problem into a number of smaller binary problems. “one-against-all” considers all classes at once and solves the multi-class problem in one step. In general, “one-against-all” is computationally more expensive to solve a multi-class problem than “one-against-one”. Hsu and Lin [38] discussed the characteristics of these methods and pointed out that the “one-against-one” method is more suitable for practical application than other methods. In this study, we select the “one-against-one” to identify the different fault types:

$$K(y, y_i) = \exp\left(-\frac{\|y - y_i\|^2}{\sigma^2}\right), \tag{14}$$

where σ is the bandwidth of the Gaussian RBF kernel function which controls the degree of

non-linearity of the model.

In order to make fair and objective evaluation for each type of multi-fault gearbox, two key parameters in LS-SVM, C and σ , which play an important role in the classification performance for LS-SVM are optimized by grid searching [39]. The search scopes of parameters C and σ are between 0 and 2 with spacing of 0.1. As mentioned in Section 3, 44 data samples for each operating condition of one fault type are collected from the surface of gearbox. 44 samples are divided into training samples and testing samples, respectively. 15 samples are selected as training samples and the remaining 29 samples are regarded as testing samples. Therefore, a total number of training samples are 90 and testing samples are 174 for six fault types of gearbox under one working condition. The objective function of grid searching method is the average percentage of correct classification, i.e. the ratio of the total number of training samples correctly classified to the total number of training samples.

As displayed in Fig. 8, five of the training samples of V_1 are misclassified, yielding the maximum average percentage of classification accuracy (CA) which is calculated as Eq. (15) is 94.4 % under load-1 with rotational speed of 880 rpm when C and σ are equal to 1.1 and 1.6 respectively:

$$CA (\%) = \frac{\text{Number of correctly classified samples}}{\text{Total number of samples in dataset}} \times 100. \tag{15}$$

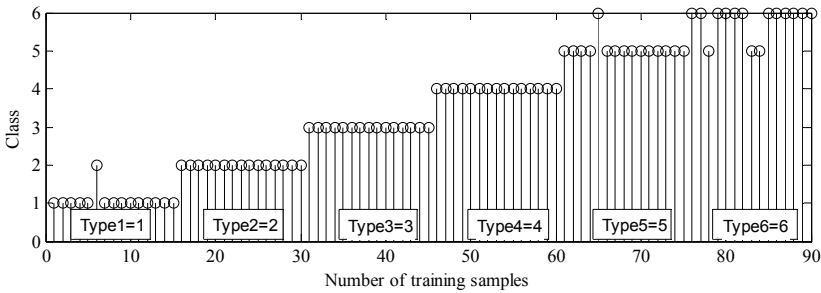


Fig. 8. Optimal training results are obtained by grid searching in LS-SVM

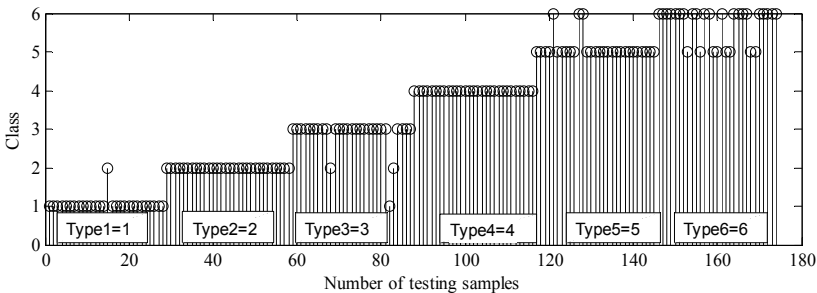


Fig. 9. Classification results in testing samples by LS-SVM

To illustrate the process concretely, the classification results of V_1 in testing phases are drawn in Fig. 9. It can be seen from Fig. 9 that 16 of the testing samples are misclassified, yielding the success rates 90.8 %. Among them, 2 samples of type-1, 3 samples of type-3, 3 samples of type-5 and 8 samples of type-6 are misclassified respectively. For brevity, the CA of testing samples obtained by the proposed method for all types of multi-fault data are listed in Tables 3-4.

Additionally, compared with the proposed method, the vector V_2 of statistical parameters expressed as Eq. (16) is employed to examine the same multi-fault data which have been investigated by proposed method:

$$V_2 = [P, K, F], \tag{16}$$

where P , K and F denote pulse factor, kurtosis and form factor, respectively.

The main parameters in LS-SVM are set in the same way as ones in our proposed method. The analysis results of testing samples about vector V_2 are demonstrated in Table 5 and Table 6.

As illustrated in Table 3 and Table 5, the classified performances of our proposed method are obviously over the performances of statistical parameters in all fault types of gearbox except the slightly less CA for the load-1 of type-5 and load-2 of type-6. Similarly, the comparisons of Table 4 and Table 6 clearly indicate that the performances of proposed method are also over the performances of statistical parameters at the absence of external load. Especially, zero correct classification crops up at type-4 and type-6 of 880 rpm shown in Table 6. As observed in Table 3, the classified performance of heavy load (load-1) is better than light load (load-2). Table 4 presents that the working condition with highly rotational speed has a slightly helpful for improving CA. The contrasts of Table 3 and Table 4 imply that the performances of the absence of external load are slightly poor. The classified results of statistical parameters are unstable and lowly accurate depicted in Table 6.

Table 3. Classification accuracy (CA) of testing samples under load-1, -2, -3 with 880 rpm by proposed method

Fault type	Load-1 (CA %)	Load-2 (CA %)	Load-3 (CA %)	Fault type	Load-1 (CA %)	Load-2 (CA %)	Load-3 (CA %)
Type 1	93.10	86.21	100.00	Type 4	100.00	100.00	96.55
Type 2	100.00	96.55	82.76	Type 5	89.66	79.31	100.00
Type 3	89.66	79.31	72.41	Type 6	72.41	65.52	86.21

Table 4. Classification accuracy (CA) of testing samples under unload with rotational speed of 880 and 1470 rpm respectively by proposed method

Fault type	880 rpm unload (CA %)	1470 rpm unload (CA %)	Fault type	880 rpm unload (CA %)	1470 rpm unload (CA %)
Type 1	86.21	89.66	Type 4	93.10	100.00
Type 2	82.76	96.55	Type 5	96.55	55.17
Type 3	72.41	86.21	Type 6	68.96	72.41

Table 5. Classification accuracy (CA) of testing samples in statistical features (pulse factor, kurtosis and form factor) under load-1, -2, -3 with 880 rpm

Fault type	Load-1 (CA %)	Load-2 (CA %)	Load-3 (CA %)	Fault type	Load-1 (CA %)	Load-2 (CA %)	Load-3 (CA %)
Type 1	86.21	58.62	100.00	Type 4	58.62	96.55	58.62
Type 2	37.93	62.07	79.31	Type 5	100.00	72.41	68.96
Type 3	44.83	62.07	72.41	Type 6	44.83	100.00	55.17

Table 6. Classification accuracy (CA) of testing samples in statistical features under absence of external load with rotational speed of 880 and 1470 rpm respectively

Fault type	880 rpm unload (CA %)	1470 rpm unload (CA %)	Fault type	880 rpm unload (CA %)	1470 rpm unload (CA %)
Type 1	100.00	44.83	Type 4	0	62.07
Type 2	86.21	13.79	Type 5	100.00	48.28
Type 3	20.69	89.66	Type 6	0	13.79

Successively, to further improve the classified performance of proposed method, one idea that vector V_3 which is form as Eq. (17) (the combination of three scale exponents V_1 and three statistical parameters V_2) is attempted to used to characterize the types of multi-fault gearbox:

$$V_3 = [SE1, SE2, SE3, P, K, F], \tag{17}$$

where V_3 represent the combination of vector V_1 and vector V_2 .

The main parameters in LS-SVM are set in the same way as ones in the proposed method. The classified performances of vector V_3 are presented in Table 7 and Table 8. From the comparison of Table 3 and Table 7, the combination of these parameters has a better performance than the single vector V_1 . Identically, Table 4 and Table 8 state that the performances of novel idea are superior to the single vector V_1 under working condition of the absence external load. As reported in Table 7, it must be point out that the CA is 100 % for all fault types under load-1 with rotational speed of 880 rpm.

Table 7. Classified performances of the combination of scale exponents and time domain features at condition of load-1, -2, -3 with 880 rpm

Fault type	Load-1 (CA %)	Load-2 (CA %)	Load-3 (CA %)	Fault type	Load-1 (CA %)	Load-2 (CA %)	Load-3 (CA %)
Type 1	100.00	100.00	100.00	Type 4	100.00	100.00	93.10
Type 2	100.00	93.10	100.00	Type 5	100.00	100.00	86.21
Type 3	100.00	86.21	96.55	Type 6	100.00	100.00	100.00

Table 8. Classified performances of the combination of scale exponents and time domain features at condition of 880, 1470 rpm and unload

Fault type	880 rpm unload (CA %)	1470 rpm unload (CA %)	Fault type	880 rpm unload (CA %)	1470 rpm unload (CA %)
Type 1	93.10	96.55	Type 4	86.21	100.00
Type 2	96.55	96.55	Type 5	100.00	62.07
Type 3	86.21	89.66	Type 6	72.41	68.96

5. Discussions

In this paper, the crossover characteristics of the scaling-law curve was introduced to explore complex multi-fault data and a novel method for multi-fault diagnosis of gearbox was proposed based on DFA, a Quasi-Monte Carlo algorithm and LS-SVM. Applied to multi-fault diagnosis of realistic gearbox, the proposed method in this paper can produce satisfactory performances in distinguishing different fault types of gearbox. Moreover, compared with statistical parameters, the combination of DFA and a Quasi-Monte Carlo algorithm perform better in feature extraction of gearbox vibration data.

Despite being a powerful tool for detecting the different fault types of gearbox, DFA still encounters some problems including discontinuities between trends of two neighboring data segments and practical difficulties in determining the type of the fitting polynomial [20]. Therefore, considerable efforts to resolve the problems existing in the original DFA are still required. This work lies beyond the scope of this study and will be carried out in future. In addition, δ data points have an effect on extracting precision of scale exponents. As a consequence, too large data size may lead to a poor result, whereas too small size may cause an increase in time cost. In our paper, δ data points which are set as 18 are reasonable and yet may not be optimal. Therefore, another problem which requires some research in future is how to determine a more appropriate δ data size for the method of uncovering optimal scale intervals.

In general, the overall procedure for a fault diagnosis scheme can be stated in four steps: data acquisition, signal processing, feature extraction, feature reduction and diagnostics (classifiers). The critical step in this process is to extract reliable features which are representative of the condition of gearbox. In our study, the feature extraction is chosen as the key research and LS-SVM is only regarded as a tool for evaluating the performances of feature extraction. In fact, as reported in Table 3 and Table 7, CA of testing samples approximate to 90 % for all types of multi-fault gearbox when the searching scope of parameters C and σ are limited between 0 and 2. It is proved that the proposed method of feature extraction is very effective. Certainly, it is without doubt that CA of testing samples can be further enhanced when the other optimization methods

such as genetic algorithm and particle swarm optimization [27] are used to tune parameters C and σ on a larger scale, which is a direction for further studies.

Furthermore, although the real multi-fault data collected from a laboratory test were explored in this paper, the proposed method which uncovers the underlying nonlinear dynamical mechanism hidden in time series from the view of data compression and multi-fractal could avoid the effect of noise in some extent and weaken the interference of non-stationary factors. Therefore, the fault diagnosis results in our study can provide good references for improving the layout of rotating machinery. Of course, modern rotating machinery such as wind turbine gearboxes has multi-gearing and multi-bearing. The dynamic responses of these components are complex and interfering with each other. When multi-fault of gears and bearing coexist in real world industrial situation, it makes the observed vibration signals rather complex, which makes it difficult to identify each fault. More research is needed to be done in the future.

6. Conclusions

1) A novel method of uncovering optimal scale intervals which can avoid the situations of spurious and problematic crossover points and scale regions and make the scale exponents self-adaptively extracted is proposed.

2) The classification results show that the proposed novel method is more effective than statistical parameters for multi-fault diagnosis. The load and rotational speed are the influenced factors of correct classification. The heavy load and high rotating speed are more conducive to increasing diagnostic reliability.

3) The combinations of scale exponents and statistical parameters further enhance the CA. The CA is up to 100 % under load-1 with rotational speed 880 rpm for all fault types. It is proved that the united idea could be further developed in multi-fault diagnosis of gearbox.

Furthermore, the multi-fault conditions with different fault types are more general situation in practice. Thus, it is of considerable practical significance to analyze these situations. Also, the proposed method can provide a reference for the fault diagnosis of other rotating machinery.

Acknowledgements

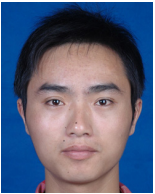
This research was supported by Aviation Science Foundation of China (Grant Number 2012ZD52054) and also supported by Science Combined Project of Nanjing University of Aeronautics and Astronautics (NZ2015103).

References

- [1] **Fakhfakh T., Chaari F., Haddar M.** Numerical and experimental analysis of a gear system with teeth defects. The International Journal of Advanced Manufacturing Technology, Vol. 25, Issue 5, 2005, p. 542-550.
- [2] **Wang Z., Han Z., Gu F., et al.** A novel procedure for diagnosing multiple faults in rotating machinery. ISA Transactions, 2014.
- [3] **Li C. J., Limmer J. D.** Model-based condition index for tracking gear wear and fatigue damage. Wear, Vol. 241, Issue 1, 2000, p. 26-32.
- [4] **Randall R. B.** A new method of modeling gear faults. Journal of Mechanical Design, Vol. 104, Issue 2, 1982, p. 259-267.
- [5] **Wang W. J., McFadden P. D.** Early detection of gear failure by vibration analysis I. Calculation of the time-frequency distribution. Mechanical Systems and Signal Processing, Vol. 7, Issue 3, 1993, p. 193-203.
- [6] **Wang W. J., McFadden P. D.** Early detection of gear failure by vibration analysis II. Interpretation of the time-frequency distribution using image-processing techniques. Mechanical Systems and Signal Processing, Vol. 7, Issue 3, 1993, p. 205-215.
- [7] **Wang W. J., McFadden P. D.** Application of wavelets to gearbox vibration signals for fault detection. Journal of Sound and Vibration, Vol. 192, Issue 5, 1996, p. 927-939.

- [8] **Lin J., Zuo M. J.** Gearbox fault diagnosis using adaptive wavelet filter. *Mechanical Systems and Signal Processing*, Vol. 17, Issue 6, 2003, p. 1259-1269.
- [9] **Padovese L. R.** Hybrid time-frequency methods for non-stationary mechanical signal analysis. *Mechanical Systems and Signal Processing*, Vol. 18, Issue 5, 2004, p. 1047-1064.
- [10] **Fan X., Zuo M. J.** Gearbox fault detection using Hilbert and wavelet packet transform. *Mechanical Systems and Signal Processing*, Vol. 20, Issue 4, 2006, p. 966-982.
- [11] **Huang N. E., Shen Z., Long S. R., et al.** The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, Vol. 454, 1998, p. 903-995.
- [12] **Wu Z., Huang N. E.** Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in Adaptive Data Analysis*, Vol. 1, Issue 1, 2009, p. 1-41.
- [13] **Frei M. G., Osorio I.** Intrinsic time-scale decomposition: time-frequency-energy analysis and real-time filtering of non-stationary signals. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, Vol. 463, Issue 2078, 2007, p. 321-342.
- [14] **De Moura E. P., Vieira A. P., Irmão M. A. S., et al.** Applications of detrended-fluctuation analysis to gearbox fault diagnosis. *Mechanical Systems and Signal Processing*, Vol. 23, Issue 3, 2009, p. 682-689.
- [15] **De Moura E. P., Souto C. R., Silva A. A., et al.** Evaluation of principal component analysis and neural network performance for bearing fault diagnosis from vibration signal processed by RS and DF analyses. *Mechanical Systems and Signal Processing*, Vol. 25, Issue 5, 2011, p. 1765-1772.
- [16] **Lin J., Chen Q.** Fault diagnosis of gearboxes based on the double-scaling-exponent characteristic of nonstationary time series. *Chinese Journal of Mechanical Engineering*, Vol. 48, Issue 13, 2012, p. 108-114.
- [17] **Lin J., Chen Q.** A novel method for feature extraction using crossover characteristics of nonlinear data and its application to fault diagnosis of rotary machinery. *Mechanical Systems and Signal Processing*, Vol. 48, Issue 1, 2014, p. 174-187.
- [18] **Ge E., Leung Y.** Detection of crossover time scales in multifractal detrended fluctuation analysis. *Journal of Geographical Systems*, Vol. 15, Issue 2, 2013, p. 115-147.
- [19] **Du W., Tao J., Li Y., et al.** Wavelet leaders multifractal features based fault diagnosis of rotating mechanism. *Mechanical Systems and Signal Processing*, Vol. 43, Issue 1, 2014, p. 57-75.
- [20] **Shao Y. H., Gu G. F., Jiang Z. Q., et al.** Comparing the performance of FA, DFA and DMA using different synthetic long-range correlated time series. *Scientific Reports*, 2012.
- [21] **Govindan R. B., Wilson J. D., Preißl H., et al.** Detrended fluctuation analysis of short datasets: an application to fetal cardiac data. *Physica D: Nonlinear Phenomena*, Vol. 226, Issue 1, 2007, p. 23-31.
- [22] **Bashan A., Bartsch R., Kantelhardt J. W., et al.** Comparison of detrending methods for fluctuation analysis. *Physica A: Statistical Mechanics and its Applications*, Vol. 387, Issue 21, 2008, p. 5080-5090.
- [23] **López J. L., Contreras J. G.** Performance of multifractal detrended fluctuation analysis on short time series. *Physical Review E*, Vol. 87, Issue 2, 2013, p. 1-9.
- [24] **Michalski S.** Blocks adjustment-reduction of bias and variance of detrended fluctuation analysis using Monte Carlo simulation. *Physica A: Statistical Mechanics and its Applications*, Vol. 387, Issue 1, 2008, p. 217-242.
- [25] **Grech D., Mazur Z.** On the scaling ranges of detrended fluctuation analysis for long-term memory correlated short series of data. *Physica A: Statistical Mechanics and its Applications*, Vol. 392, Issue 10, 2013, p. 2384-2397.
- [26] **Gulich D., Zunino L.** A criterion for the determination of optimal scaling ranges in DFA and MF-DFA. *Physica A: Statistical Mechanics and its Applications*, Vol. 397, Issue 1, 2014, p. 17-30.
- [27] **Zheng H. B., Liao R. J., Grzybowski S., et al.** Fault diagnosis of power transformers using multi-class least square support vector machines classifiers with particle swarm optimisation. *Electric Power Applications*, Vol. 5, Issue 9, 2011, p. 691-696.
- [28] **Fernández F. D., Martínez R. D., Fontenla R. O., et al.** Automatic bearing fault diagnosis based on one-class v-SVM. *Computers & Industrial Engineering*, Vol. 64, Issue 1, 2013, p. 357-365.
- [29] **Peng C. K., Buldyrev S. V., Havlin S., et al.** Mosaic organization of DNA nucleotides. *Physical Review E*, Vol. 49, Issue 2, 1994, p. 1685-1689.
- [30] **Vapnik V. N.** *Statistical Learning Theory*, John Wiley and Sons, New York, 1998.
- [31] **Chang C. C., Lin C. J.** Libsvm: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, Vol. 2, Issue 3, 2011, p. 1-27.

- [32] **Widodo A., Yang B. S.** Support vector machine in machine condition monitoring and fault diagnosis. *Mechanical Systems and Signal Processing*, Vol. 21, Issue 6, 2007, p. 2560-2574.
- [33] **Baydar N., Ball A.** A comparative study of acoustic and vibration signals in detection of gear failures using Wigner-Ville distributions. *Mechanical Systems and Signal Processing*, Vol. 15, Issue 6, 2001, p. 1091-1107.
- [34] **Li Z., Yan X., Yuan C., et al.** Virtual prototype and experimental research on gear multi-fault diagnosis using wavelet-autoregressive model and principal component analysis method. *Mechanical Systems and Signal Processing*, Vol. 25, Issue 7, 2011, p. 2589-2607.
- [35] **Boser B. E., Guyon I. M., Vapnik V. N.** A training algorithm for optimal margin classifiers. *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, 1992, p. 144-152.
- [36] **Yang J., Zhang Y.** Application research of support vector machines in condition trend prediction of mechanical equipment. *Advances in Neural Networks*, Vol. 3498, Issue 3, 2005, p. 857-864.
- [37] **Cui J., Wang Y.** A novel approach of analog circuit fault diagnosis using support vector machines classifier. *Measurement*, Vol. 44, Issue 1, 2011, p. 281-289.
- [38] **Hsu C. W., Lin C. J.** A comparison of methods for multiclass support vector machines. *IEEE Transactions on Neural Networks*, Vol. 13, Issue 2, 2002, p. 415-425.
- [39] **Saidi L., Ali J. B., Fnaiech F.** Application of higher order spectral features and support vector machines for bearing faults classification. *ISA Transactions*, 2014.



Xingxing Jiang received the B.S. degree in Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2012. Now he is a Ph.D. student with College of Energy and Power Engineering, Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing, China. His current research interests include Rotating Machinery Fault Diagnosis and Mechanical Signal and Information Processing.



Shunming Li received Ph.D. degree in mechanics from Xi'an Jiaotong University, China, in 1988. He is a Professor in Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing, China. His current research interests include noise and vibration analysis and control, signal processing, machine fault diagnosis, sensing and measurement technology, intelligent vehicles.



Yong Wang received the B.S. degree in Nanjing Forestry University, Nanjing, China, in 2010. Now he is a Ph.D. student with College of Energy and Power Engineering, Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing, China. His current research interests include vibration analysis and control.