

# A hybrid feature extraction methodology for gear pitting fault detection using motor stator current signal

Qun He, Xiaoru Ren, Guoqian Jiang and Ping Xie

Submitted 24.02.14  
Accepted 10.05.14

*Vibration signals have been widely used to detect gear faults in machines in various industrial applications. However, stator current signals from the drive motor have recently become an alternative and promising tool for detecting and diagnosing the existence and occurrence of gear faults, due to their higher reliability, lower cost and easier distant monitoring ability than traditional vibration-based fault detection methods. Therefore, this paper combines empirical mode decomposition (EMD), fast independent component analysis (FastICA) and a sample entropy measure to propose a hybrid feature extraction methodology for gear fault detection, which is used to analyse the stator current signals of the drive motor. First, the stator current signals obtained from an induction motor are decomposed by EMD into several intrinsic mode functions (IMFs). The signal-to-noise ratio of the stator current signals can be enhanced by removing the IMFs in high-frequency bands. Second, in order to eliminate information redundancy among the IMFs and improve the accuracy of fault detection, the FastICA approach is applied to the selected IMFs to extract independent components. Finally, the sample entropy of the independent components is calculated to quantitatively characterise the differences between healthy and faulty gears and then identify a gear pitting fault. The proposed method is verified by experiments on a real gearbox under different motor rotating speeds. The results demonstrate that the proposed hybrid feature extraction method can provide a more effective and efficient approach to gear pitting fault detection.*

**Keywords:** gear pitting fault, stator current signal, empirical mode decomposition (EMD), fast independent component analysis (FastICA), sample entropy.

## 1. Introduction

Widely used in various mechanical equipment, the gear drive is a power and motion transmission device and its dynamic performance has a significant impact on the whole machine. Gear fault diagnosis is valuable for decreasing economic costs and increasing the operational safety of gearboxes<sup>[1]</sup>. Hence, it is highly desirable to detect gear faults at an early stage and repair the faulty gear in a timely manner, in order to reduce downtime and prevent catastrophic damage to the whole machine<sup>[2]</sup>.

Traditionally, vibration signals are preferred for detecting gear faults<sup>[3]</sup>. However, the useful fault information is usually masked by a large amount of strong background noise due to external excitation and their invasive measurement in nature<sup>[4]</sup>, and the

installation and maintenance of vibration sensors is also relatively difficult<sup>[5]</sup>. An alternative solution to overcome these drawbacks can be the use of stator current signals, which are reliable and easily accessible from the ground without the need for additional sensors or data acquisition devices.

Because the torsional vibration of a gear through the rotary shaft causes the fluctuation of motor air-gap torque, the change of air-gap torque through stator flux causes a change of stator current<sup>[6]</sup>. Therefore, extracting the fault characteristic information from the stator current signal is crucial to fault diagnosis. Some publications related to gear fault detection using the current signal are presented in<sup>[7-10]</sup>. Classical spectral analysis is used to detect a tooth breakage fault of a multistage gearbox and results show that the rotating frequency and mesh frequency components can be detected in the stator current signals<sup>[7]</sup>. In<sup>[8]</sup>, an original, integrated electromechanical model aims at testing the possibility and the interest of tooth fault detection based on electrical measurements on the motor stator. Some works apply discrete wavelet transform and a corrected multi-resolution Fourier transform to investigate the vibration and current transients for gearbox fault detection<sup>[9]</sup>. In<sup>[10]</sup>, the authors have studied that analysing the stator current signals can effectively diagnose gearbox faults under different load conditions.

While the stator current signals collected from the induction motor end generally present non-stationary characteristics, traditional signal processing methods show certain limitations. On the other hand, the stator current signals are usually noisy and fault information hidden in the original current signals is very weak and even buried by fundamental-frequency components. Thus, it is a huge challenge to use stator current signals for gear fault detection. Empirical mode decomposition (EMD) is a newly-developed signal processing tool especially for non-linear and non-stationary signals, which can adaptively decompose a complicated signal into several intrinsic mode components containing relatively independent information from high frequency to low frequency. This paper attempts to use the EMD method to process the measured stator current signal into a series of intrinsic mode functions (IMFs)<sup>[11]</sup> and remove the high-frequency and fundamental-frequency components. However, there is redundant and related information between the intrinsic mode components, which is unfavourable for signal feature extraction. Thus, fast independent component analysis technology is introduced to remove the redundancy between the intrinsic mode components, which is a new signal processing method, along with the blind source separation problem. Introduced in recent years, it is mainly used in mechanical feature extraction, image processing, and so on<sup>[12-13]</sup>. It can decompose the IMFs into a series of independent components.

In order to realise the detection of a gear pitting fault, a hybrid feature extraction methodology combining the EMD, FastICA method and the sample entropy measure is presented in this paper to implement the signal preprocessing, feature extraction and feature quantification. In this method, the EMD is applied to decompose the signal and eliminate high-frequency noise and fundamental-frequency interference. FastICA is used to

Qun He, Xiaoru Ren, Guoqian Jiang\* and Ping Xie are with the Key Lab of Measurement Technology and Instrumentation of Hebei Province, Yanshan University, Qinhuangdao, PR China.

\*Corresponding author. Email: jgq870706@126.com (G Jiang) and pingx@ysu.edu.cn (P Xie).

extract feature components and the sample entropy measure can quantitatively describe the fault characteristics for gear pitting fault detection. Experimental studies are carried out on a real gearbox under different motor rotating speeds to validate the proposed hybrid feature extraction method.

## 2. Methods

### 2.1 Empirical mode decomposition

Empirical mode decomposition (EMD) is a new non-linear and non-stationary signal processing technique proposed by Huang *et al*<sup>[14]</sup>. This method developed from the assumption that any signal consists of different simple intrinsic modes of oscillation. In this way, each signal could be decomposed into a number of intrinsic mode functions (IMFs). Each IMF must satisfy the following definitions: (1) In the whole dataset, the number of zero-crossings and the number of extrema must either be equal or differ at most by one; (2) At any point, the mean value of the envelope defined by the local minima is zero. The main steps of empirical mode decomposition are as follows<sup>[15]</sup>:

1. Assume a random time series  $x(t)$ , identify all local maxima and then get the upper envelope by interpolating between maxima and similarly get the lower envelope. The mean of the upper and lower envelope values is designated as  $m_1$  and the difference between the signal  $x(t)$  and  $m_1$  is the first component  $h_1$ , as:

$$x(t) - m_1 = h_1 \dots\dots\dots(1)$$

Ideally, if  $h_1$  satisfies all the requirements of an IMF, then  $h_1$  is the first IMF component of  $x(t)$ .

2. If  $h_1$  is not an IMF, treat  $h_1$  as the original signal and repeat step 1 until  $h_1$  is an IMF. Then, it is designated as  $c_1 = h_1$ , the first IMF component from the original data.
3. After getting the first component, remove the first component from the original signal and obtain the residual  $r_1$ , as follows:

$$r_1 = x(t) - c_1 \dots\dots\dots(2)$$

Then, treat  $r_1$  as the original signal and repeat the above processes. The second IMF component  $c_2$  of  $x(t)$  will be obtained. Repeat the process as described above  $n$  times. Then, all the IMFs of the signal  $x(t)$  can be obtained. When  $r_n$  is a constant or monotone function, it cannot extract condition signal components and this is the end of the cycle. Finally, the inspected signal  $x(t)$  may be expressed as:

$$\begin{cases} x(t) - c_1 = r_1 \\ r_1 - c_2 = r_2 \\ \dots \\ r_{n-1} - c_n = r_n \end{cases} \dots\dots\dots(3)$$

The decomposition process can be stopped when  $r_n$  becomes a monotonic function or a constant from which no more IMF components can be extracted. Summing up both sides of Equation (3) accordingly, we obtain:

$$x(t) = \sum_{i=1}^n c_i + r_n \dots\dots\dots(4)$$

Thus, one can achieve a decomposition of the signal into  $n$  empirical modes and a residual  $r_n(t)$ , which is the mean trend of  $x(t)$ .

### 2.2 Fast independent component analysis

The method of independent component analysis (ICA) can separate the mixed signal into independent components without any prior knowledge. The basic model of ICA can be expressed as<sup>[16]</sup>:

$$X = AS \dots\dots\dots(5)$$

where  $X = [x_1, x_2, x_3, \dots, x_n]^T$  is random data, the  $x_i$  ( $i = 1, 2, 3, \dots, n$ ) expresses  $n$  observation points,  $S = [s_1, s_2, s_3, \dots, s_n]^T$  is the mixed

signal and  $A$  is the  $m \times n$  mixing matrix.

According to the ICA model, it is easy to find that there are the phenomena of energy ambiguity and order uncertainty.

FastICA is such an algorithm, which is based on a fixed-point iteration scheme for finding a maximum non-Gaussianity of  $WX$ . There are different measures of non-Gaussianity, such as kurtosis (fourth-order cumulant) and negentropy. In this paper, we make use of negentropy as the maximum non-Gaussianity measurement.

The basic steps of the FastICA algorithm are as follows<sup>[17]</sup>:

To preprocess the observed signal: standardisation, centralisation and whitening,  $X$  turns into  $Z$ .

1. Assume the number of independent components needing to be separated is  $N$  and set the number of iterations  $t = 1$ .
2. Randomly initialise a full vector  $W_t$ .
3. Let  $W_t = E\{Zg(W_t^T Z)\} - E\{g'(W_t^T Z)\}W_t$ ;  $g'$  is the derivative of  $g$ .
4.  $W_t = W_t - \sum_{j=1}^{t-1} (W_t^T W_j) W_j$ .
5. Set  $W_t = W_t / \|W_t\|$ , if it is not convergent return to step 4.

Through the above calculation and measurement of mixed signal  $X$ , one can obtain the independent component  $Y$ , mixed matrix  $A$  and the separation of matrix  $W$ .

### 2.3 Sample entropy measure

Given  $N$  data points from a time series  $\{x(n)\} = x(1), x(2), \dots, x(N)$ , take  $m$  vectors  $x_m(1), \dots, x_m(N-m+1)$  defined as  $X_m(i) = [x(i), x(i+1), \dots, x(i+m-1)]$ , for  $1 \leq i \leq N-m+1$ . These vectors stand for  $m$  consecutive  $x$  values, starting at the  $i$ th sample<sup>[18]</sup>:

1. Let  $r$  denote the noise filter level, which is defined as:

$$r = g \times SD, \text{ for } g = 0.1, 0.2, \dots, 0.9 \dots\dots\dots(6)$$

where  $SD$  represents the standard deviation of the data sequence  $X$ .

2. The distance between vector  $X_m(i)$  and  $X_m(j)$ ,  $d[X_m(i), X_m(j)]$  is defined as the maximum absolute difference between their scalar components:

$$d[X_m(i), X_m(j)] = \max_{k=0, \dots, m-1} |x(i+k) - x(j+k)| \dots\dots\dots(7)$$

3. For a given  $X_m(i)$ , count the number of  $j$  ( $1 \leq j \leq N-m, j \neq i$ ), such that  $d[X_m(i), X_m(j)] \leq r$ . This number is represented as  $B_i$ . Then, for  $1 \leq i \leq N-m$ :

$$B_i^m(r) = \frac{1}{N-m+1} B_i \dots\dots\dots(8)$$

Here, note that only the first  $N-m$  vectors of length  $m$  are considered in order to ensure that for  $1 \leq j \leq N-m$ , the vector  $X_{m+1}(i)$  is also defined. The  $B^m(r)$  is defined as:

$$B^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} B_i^m(r) \dots\dots\dots(9)$$

4. Increment the dimension to  $m = m+1$  and compute  $A_i$  as the number of  $X_{m+1}(i)$  within  $r$  of  $X_{m+1}(j)$ , where  $j$  ranges from 1 to  $N-m$ . Then define  $A_i^m(r)$  as:

$$A_i^m(r) = \frac{1}{N-m-1} A_i \dots\dots\dots(10)$$

5. Define  $A^m(r)$  as:

$$A^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} A_i^m(r) \dots\dots\dots(11)$$

Thus,  $B^m(r)$  represents the probability that two sequences will match for  $m$  points, whereas  $A^m(r)$  represents the probability that two sequences will match for  $m+1$  points.

The sample entropy (SampEn) is defined as:

$$SampEn(m, r) = \ln \left[ \frac{B^m(r)}{A^m(r)} \right] \dots\dots\dots(12)$$

### 3. Hybrid feature procedure for gear pitting fault detection

When gear faults occur, the torsional vibration of the gear through the rotating shaft causes a fluctuation of motor air-gap torque. Meanwhile, the change of air-gap torque through the stator flux causes a change of stator current. Therefore, the stator current signals can be used to detect a gear pitting fault.

Given that the stator current signals collected from the drive induction motor end are generally non-stationary and noisy, it is difficult to obtain the ideal detection results directly using traditional stationary signal analysis methods. Therefore, this present study combines the EMD method, FastICA and a sample entropy measure to propose a hybrid feature extraction methodology to analyse the stator current signals in order to detect gear pitting faults. The detailed detection procedure is illustrated in Figure 1. There are three phases in this procedure. First, gear stator current signals are decomposed with the EMD method into several intrinsic mode functions (IMFs) and the high-frequency and fundamental-frequency components are removed to obtain useful IMFs. Second, due to aliasing phenomenon and correlation among selective IMFs, the FastICA method is used to eliminate the redundant information and turn the intrinsic mode functions into independent components. Lastly, the sample entropies of the independent components are calculated to quantitatively characterise useful fault information. Through the above feature extraction procedure, the differences between healthy and faulty gears can be obtained and thus a pitting fault can be also identified.

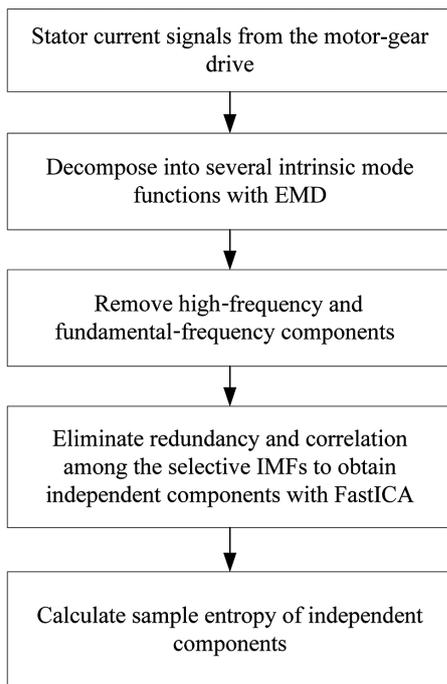


Figure 1. The proposed hybrid feature extraction procedure

### 4. Experiment results

#### 4.1 Test-rig

The test-rig schematic diagram and physical diagram for machine diagnostic purposes in our laboratory are shown in Figures 2(a) and 2(b), which consists of a three-phase asynchronous motor (2.2 kW), a gearbox (type PM250, the reduction ratio is 10.35, two stages, gear ratios are 30/69 and 18/81) with the testing bearings and gears and the load and data acquisition and analysis system. The speed of the motor is adjusted using the integrated speed regulator. The motor directly drives the gearbox through a flexible coupling.

In this paper, a helical gear is applied in the gearbox for the fault

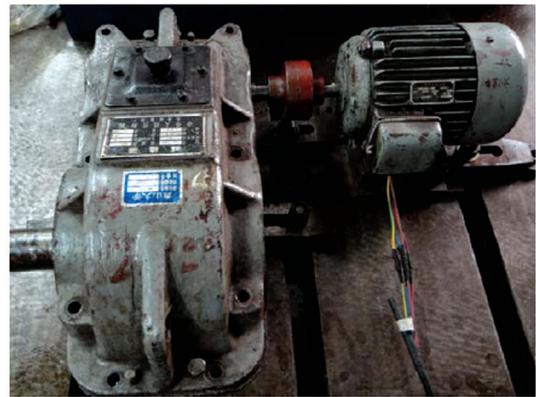
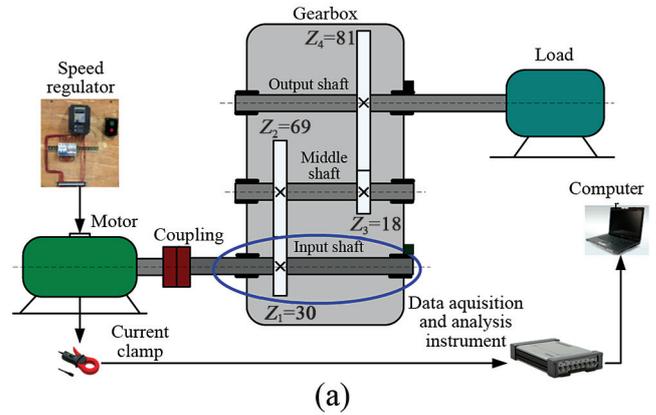


Figure 2. Gear simulation test-rig: (a) the test-rig schematic diagram; (b) the test-rig physical diagram

simulation experiment. Specific parameters of the helical gear are: the number of teeth is 30, normal modulus is 2 mm, the tooth width is 45 mm, the pressure angle is 20°, the helix angle is 8°6'34" and the headspace coefficient is 0.25, a coefficient of the addendum. A gear pitting fault is artificially produced, with some 2 mm-deep pits on the gear surface for the fault simulation experiment, while the gear in normal condition experiment aims to adjust the motor rotating speed with the healthy gearbox for data acquisition. Figure 3 shows the pitting faulty gear for the simulation experiment.



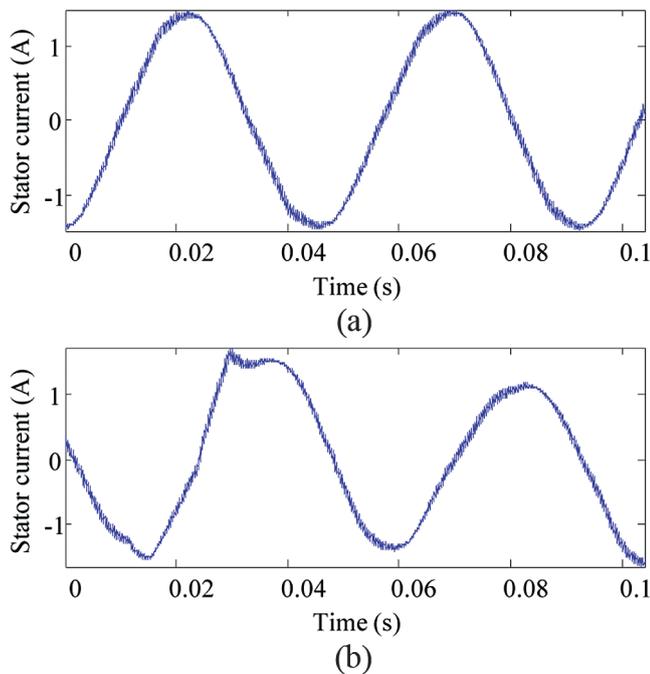
Figure 3. The pitting faulty gear

In the experiment, two different types of signal are collected from the test-rig at the same time: one is the stator current signal, the other is the vibration signal. The current clamp was used to acquire stator current data from the motor. The data sampling rate was 20 kHz. The data was collected using a high-accuracy

signal acquisition device (type INV3018C) that had a built-in anti-aliasing filter to improve the signal-to-noise ratio of the measured current signals. A gear current signal dataset containing two different conditions, (a) healthy and (b) pitting fault, are obtained from the above experimental system, respectively. Each condition corresponds to a subset consisting of 30 samples, each of which contains 8192 sampling points. All the experiments were repeated under the motor rotating speeds of 400, 600 and 800 r/min. The vibration signal acquisition system was composed of a data acquisition card (type NI USB-6215), sensors (type SD1405) and a charge amplifier (type SD1436). The sensors were placed on the gearbox in horizontal and vertical positions to measure the vibration signal in both directions and through the data acquisition card and charge amplifier instrument to the computer. The gear vibration signal acquisition process is the same as the gear current signal acquisition process.

#### 4.2 Signal preprocessing with EMD and FastICA

One sample of stator current signals is selected from a healthy gear and pitting faulty gear, respectively, which contain 2048 sampling points, as shown in Figures 4(a) and 4(b). The EMD is employed to decompose the two samples. According to the Formulae (1)-(3), each sample is decomposed into a series of IMFs, as shown in Figures 5(a) and 5(b). It can be seen from Figure 5 that the residual component of each sample is the fundamental-frequency component.

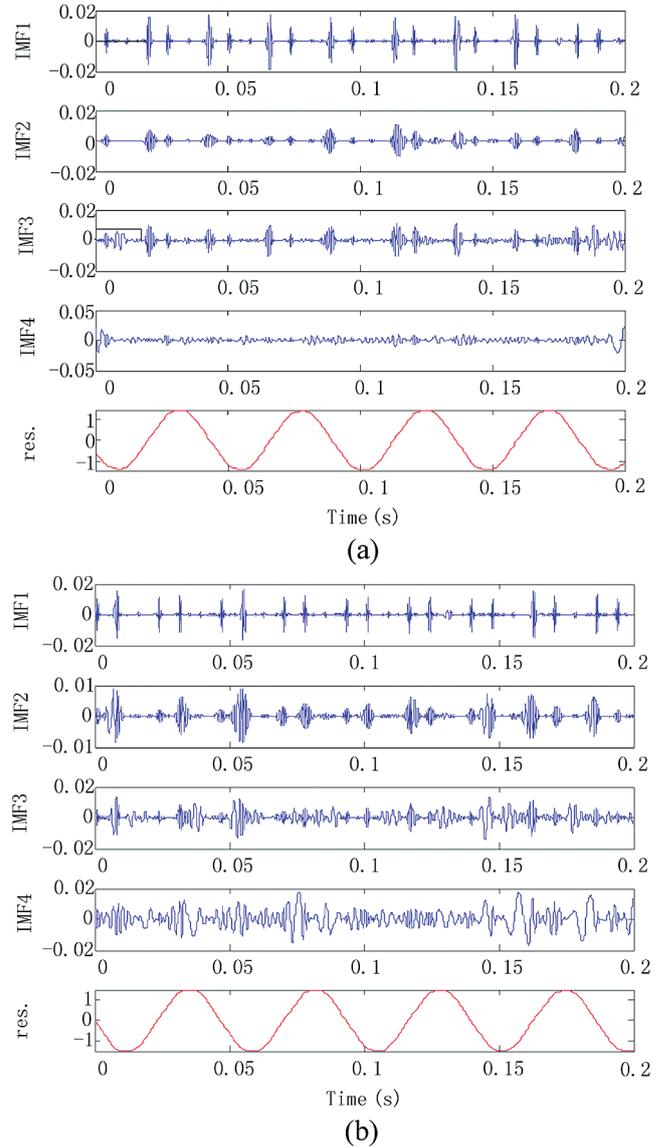


**Figure 4. Stator current signal profiles of (a) the gear in healthy condition (b) the gear in pitting fault condition**

In order to remove the high-frequency components, the Hilbert time-frequency analysis is proposed to handle the IMF components of the two samples. High-frequency noise in a low-voltage distribution network mainly concentrates on above 10 kHz<sup>[19]</sup>. Figure 6 demonstrates the instantaneous frequency of the IMF1 and IMF2 components of the sample from a healthy gear.

It can be seen from Figure 6(a) that the instantaneous frequency of the IMF1 reaches about 10 kHz. So, it can be considered as high-frequency noise. Figure 6(b) shows that the instantaneous frequency of the IMF2 does not reach the high-frequency range. So, the IMF2 component affected by high-frequency noise can be neglected. Similarly, the sample of a faulty gear is treated with the Hilbert time-frequency analysis.

Through the above analysis, the IMF2, IMF3 and IMF4 components of each sample are mixed as the input signal of the FastICA algorithm, respectively. According to the FastICA



**Figure 5. EMD decomposition results of stator current signals of (a) the gear in healthy condition and (b) the gear in pitting fault condition**

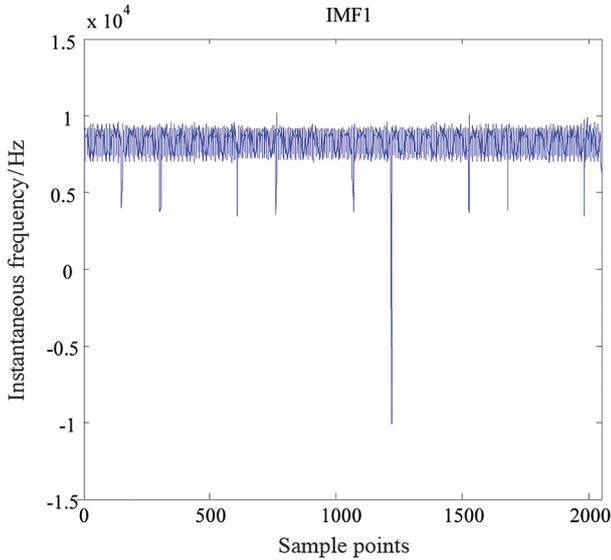
algorithm formula, the selected IMFs are turned into a series of ICs. Figure 7 shows the series of ICs of the two samples.

It can be seen from Figures 6(a) and 6(b) that the number of ICs is the same as the input components. The ICs of each sample are the feature components for gear pitting fault detection.

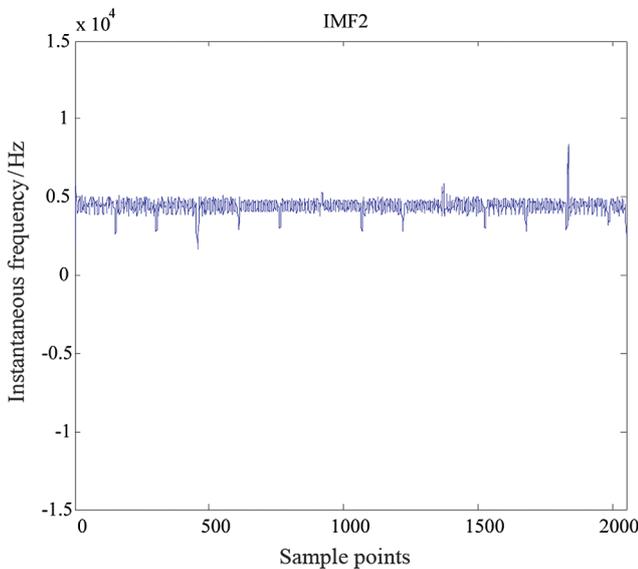
#### 4.3 Characteristic quantification with sample entropy measure

10 samples of stator current signals were selected from a healthy gear and a pitting faulty gear under motor rotating speeds of 600 r/min, respectively. The first 10 samples are the current signals of the gear pitting fault and the following 10 samples are the current signals of a gear in healthy condition. The number of sample points is 2048. The selected 20 samples are preprocessed by the EMD-FastICA method. Consequently, the independent components of each sample are obtained for gear pitting fault detection.

The sample entropy measure is applied to quantitatively describe the independent components. According to Formulae (6)-(12), the sample entropy of the feature components of 20 samples can be obtained. Figure 8 demonstrates the difference in the sample entropy (SampEn) of the ICs in two gear working states. Table 1 shows the sample entropy mean and standard deviation of gear stator current signals in two working states based on two different analysis methods.



(a)

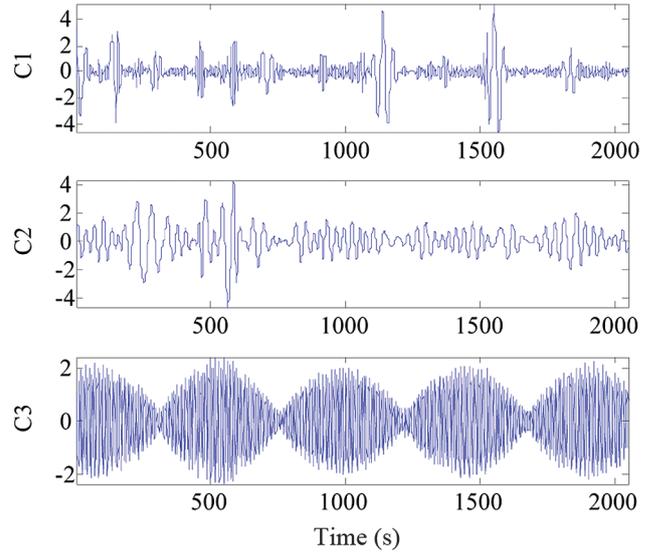


(b)

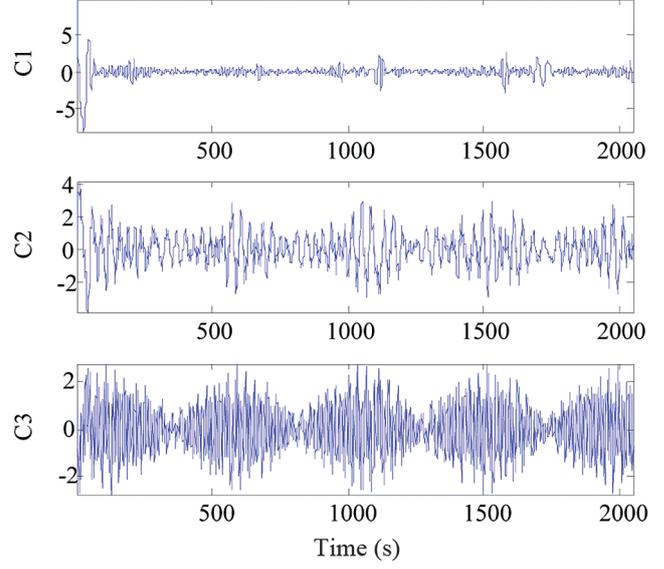
Figure 6. Instantaneous frequency of one sample of the gear in healthy condition of (a) IMF1 and (b) IMF2

It can be seen from Figure 8 and Table 1 that the sample entropy mean of the ICs of the gear pitting fault is higher than the sample entropy mean of the ICs of the gear in normal condition. It can be further seen that the sample entropy mean of IC2 and IC3 has an obvious discrimination on the two gear working states. For the faulty condition, the sample entropy mean of IC2 and IC3 is 0.668 and 0.762, respectively; for the healthy condition, the sample entropy mean of IC2 and IC3 is 0.388 and 0.391, respectively. But the sample entropy values of IC1 have a little discrimination in the two gear working states. This is because the signal characteristic information of the two gear working states concentrated in the IC2 and IC3 components after processing by the EMD-FastICA method. According to the analysis result, the sample entropy values of the gear pitting fault show a generally increasing trend compared with the sample entropy values of a gear in healthy condition. This can be explained that, due to the deterioration of the gear mechanism, the number of frequency components contained in the current signal increases, with an increase in its corresponding complexity value.

To compare the discrimination of the original EMD feature extraction method and the proposed method in the detection of gear pitting fault, Figure 9 shows the differences between the IMFs' sample entropy (SampEn) values in the two gear working states.



(a)



(b)

Figure 7. Decomposed ICs of (a) the gear in healthy condition and (b) the gear in pitting fault condition

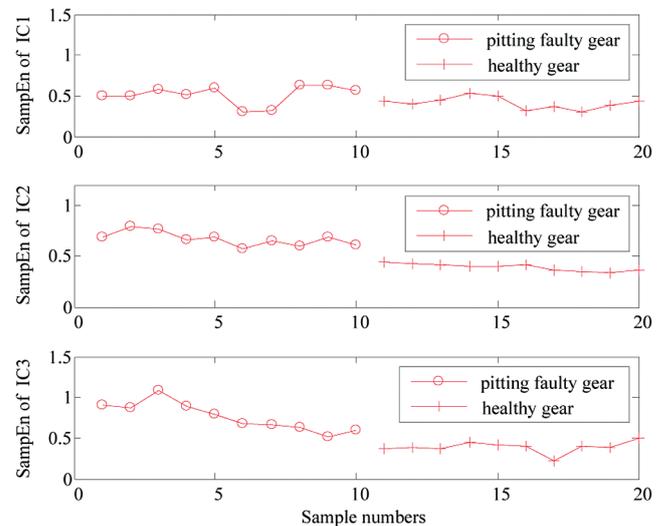
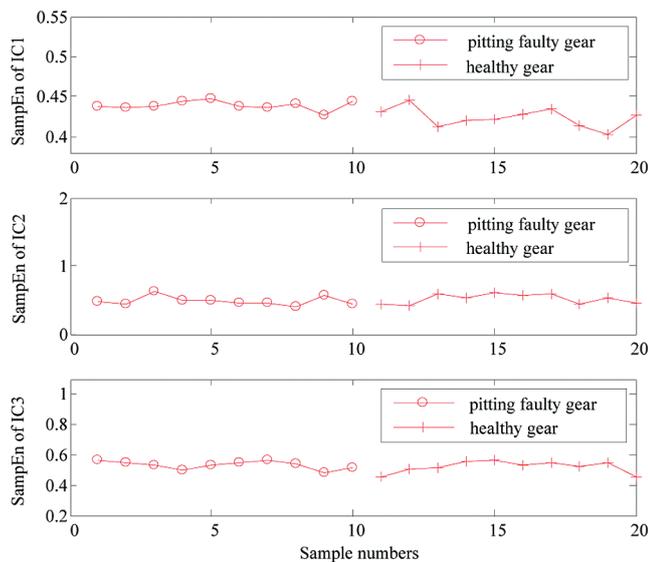


Figure 8. The difference in sample entropy of the ICs in two gear working states

**Table 1. The sample entropy mean and standard deviation of gear stator current signals in two working states based on two different analysis methods**

			Sample number										Sample entropy mean and standard deviation
			1	2	3	4	5	6	7	8	9	10	
EMD-FastICA	Pitting faulty gear	IC1	0.491	0.487	0.576	0.520	0.571	0.394	0.311	0.524	0.547	0.468	0.489±0.082
		IC2	0.685	0.785	0.758	0.656	0.685	0.569	0.648	0.597	0.685	0.607	0.668±0.068
		IC3	0.851	0.870	0.787	0.894	0.791	0.784	0.656	0.629	0.691	0.652	0.762±0.097
	Healthy gear	IC1	0.430	0.406	0.452	0.366	0.491	0.324	0.366	0.526	0.380	0.440	0.418±0.062
		IC2	0.441	0.428	0.411	0.397	0.403	0.413	0.360	0.343	0.329	0.356	0.388±0.038
		IC3	0.367	0.390	0.375	0.446	0.421	0.403	0.215	0.406	0.391	0.493	0.391±0.072
EMD	Pitting faulty gear	IC1	0.438	0.437	0.439	0.448	0.438	0.441	0.426	0.445	0.436	0.445	0.439±0.006
		IC2	0.480	0.438	0.629	0.503	0.500	0.451	0.451	0.408	0.572	0.444	0.488±0.067
		IC3	0.565	0.548	0.529	0.496	0.529	0.548	0.563	0.538	0.485	0.519	0.532±0.026
	Healthy gear	IC1	0.432	0.446	0.412	0.421	0.422	0.429	0.435	0.414	0.403	0.427	0.424±0.013
		IC2	0.444	0.443	0.423	0.583	0.527	0.610	0.574	0.443	0.531	0.461	0.504±0.069
		IC3	0.461	0.505	0.519	0.561	0.562	0.531	0.547	0.523	0.553	0.456	0.522±0.038

It can be seen from Figure 9 and Table 1 that the sample entropy mean of the ICs of two kinds of gear state based on the traditional EMD method are almost equal. As a result, based on EMD-FastICA and sample entropy measure, the gear pitting fault can be detected effectively.

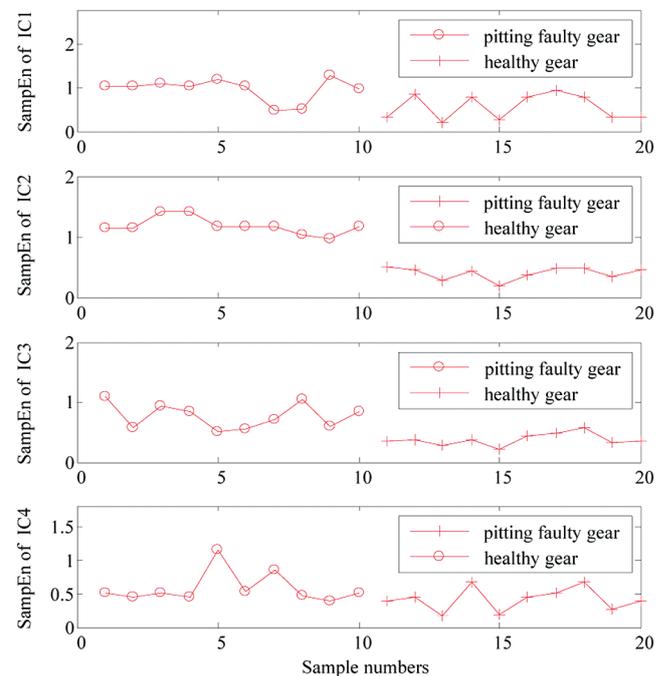


**Figure 9. The difference in sample entropy of the ICs in two gear working states**

Next, 10 samples of vibration signals from the healthy gear and pitting faulty gear under the motor rotating speed of 600 r/min were selected, respectively. The first 10 samples are the vibration signals of the gear pitting fault and the next 10 samples are the vibration signals of the gear in a healthy condition. The number of sample points is 2048. The selected 20 samples are preprocessed by the EMD method, taking the first four IMFs of each sample as the research object, and the FastICA method is then applied to turn the first four IMFs of each sample into feature components. The sample entropy measure is applied to quantitatively describe the feature components. Figure 10 demonstrates the difference in the sample entropy (SampEn) of feature components in the two gear working states. Table 2 shows the sample entropy mean and standard deviation of gear vibration signals in two working states based on EMD-FastICA analysis methods.

It can be seen from Figure 10 and Table 2 that the sample entropy mean of the ICs of the gear pitting fault are generally higher than the sample entropy mean of the ICs of the gear in a

healthy condition. It can be further seen that the sample entropy mean of IC2 has an obvious discrimination in the two gear working states and the sample entropy mean of IC2 in the former 10 samples is 1.173 and the sample entropy mean of IC2 in the next 10 samples is 0.397. According to the result, the sample entropy values of the gear pitting fault are higher than the sample entropy values of the gear in a healthy condition. It can be explained that the complexity of the vibration signals in a gear pitting state is higher than in the normal gear operation state. So, based on the vibration signals, the proposed method can also detect a gear pitting fault effectively.



**Figure 10. The difference in sample entropy of the ICs in two gear working conditions**

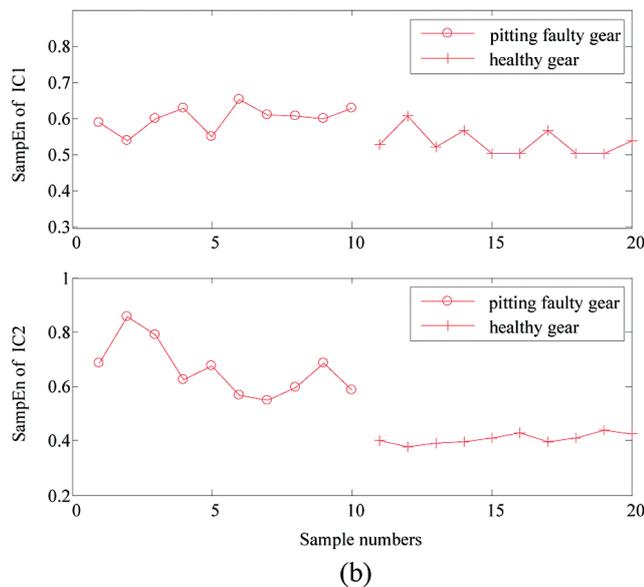
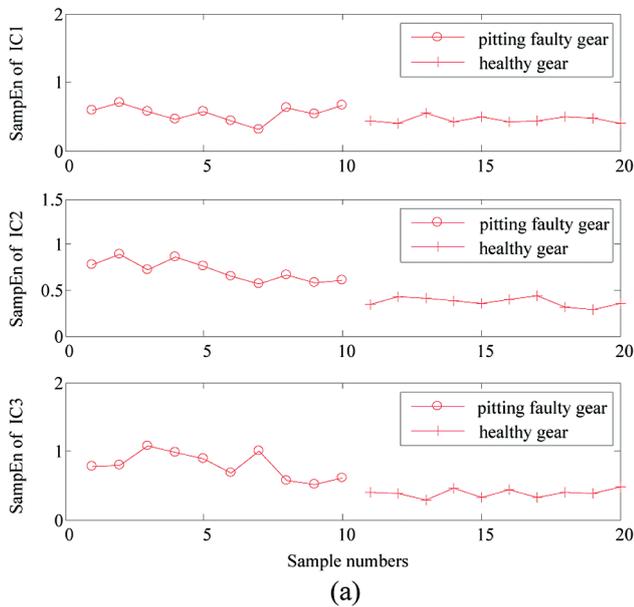
To compare the discrimination of the current signals and the vibration signals in the detection of a gear pitting fault, it can be seen from Figures 7 and 9 that the proposed method based on two types of signal can effectively detect the gear pitting fault. But, from the perspective of signal acquisition, the stator current signal analysis method has three major advantages over the vibration signal analysis method:

1. It is a non-invasive measurement and it does not affect the running of the motor.

**Table 2. The sample entropy mean and standard deviation of gear vibration signals in two working states based on EMD-FastICA analysis methods**

			Sample number										Sample entropy mean and standard deviation
			1	2	3	4	5	6	7	8	9	10	
EMD-FastICA	Pitting faulty gear	IC1	1.012	1.114	1.143	1.208	0.987	1.268	1.145	0.437	0.469	1.432	1.022±0.325
		IC2	1.124	1.116	1.309	1.313	1.151	1.176	1.164	1.102	1.041	1.237	1.173±0.089
		IC3	1.034	0.656	0.908	0.833	0.521	0.744	0.493	0.669	0.637	0.836	0.733±0.171
		IC4	0.518	0.487	0.511	0.507	1.171	0.645	0.929	0.567	0.514	0.532	0.638±0.229
	Healthy gear	IC1	0.323	0.746	0.308	0.745	0.471	0.800	1.039	0.892	0.382	0.504	0.621±0.256
		IC2	0.493	0.405	0.394	0.315	0.308	0.366	0.470	0.561	0.329	0.332	0.397±0.086
		IC3	0.326	0.345	0.301	0.423	0.201	0.487	0.501	0.657	0.338	0.394	0.397±0.127
		IC4	0.479	0.513	0.256	0.923	0.267	0.522	0.541	0.918	0.367	0.485	0.527±0.231

- The overall vibration of the motor does not affect the stator current, so the stator current analysis is less susceptible to outside interference.
- The current sensor price is low, it is easy to operate and can realise remote and continuous fault signal monitoring.



**Figure 11. The differences in sample entropy of the ICs in two gear working states under the motor rotating speed of (a) 400 r/min and (b) 800 r/min**

Therefore, stator current analysis is selected in this study for gear pitting fault detection.

#### 4.4 Gear pitting fault detection under different motor rotating speeds

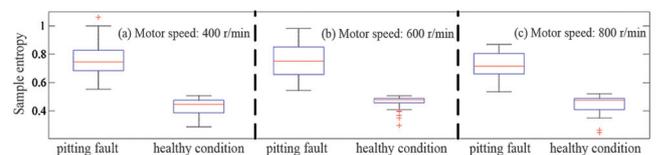
The proposed method is further verified under two different motor rotating speeds of 400 r/min and 800 r/min. Stator current signals from a gear in healthy condition and pitting fault condition were collected. Figure 11 shows the differences in sample entropy of the ICs in two gear working states under the motor rotating speeds of 400 r/min and 800 r/min, respectively.

It can be seen from Figure 11 that there are three independent components (ICs) in Figure 11(a), but only two ICs in Figure 11(b). This is because, after the EMD signal preprocessing, the selected IMFs for feature extraction are not determined. Figure 11(a) shows that the second sub-bands' sample entropy (SampEn) have a better detection quality for a gear pitting fault, and the sample entropy mean of the former 10 samples and the next 10 samples of the second sub-bands is 0.714 and 0.422, respectively. Figure 11(b) demonstrates that the second sub-bands' sample entropy can distinguish the gear pitting fault and the sample entropy mean of the former 10 samples and next 10 samples is 0.698 and 0.413, respectively. Through the above analysis, the proposed method is convenient and effective for the detection of a gear pitting fault under different motor rotating speeds.

#### 4.5 Statistical analysis of fault detection results under different motor rotating speeds

From the stator current signals of two gear working conditions under different motor rotating speeds, 2000 effective sample entropy values were extracted for feature analysis. The statistical results are shown in Figure 12, which is plotted with the Matlab function boxplot in the statistical analysis toolbox. The boxes have lines at the lower quartile, median and upper quartile values. The lower and upper lines of the 'box' are the 25th and 75th percentiles of the sample, respectively. It can be seen from Figure 12 that the sample entropy of a healthy gear is lower than the sample entropy of a pitting faulty gear. This is because the current signals have a certain regularity and complexity when the gear is working under abnormal conditions.

The research results are consistent with the characteristics of sample entropy and theoretical analysis. Hence, the method of EMD-FastICA with sample entropy can be used to detect a gear pitting fault.



**Figure 12. Boxplot of sample entropy values between gear pitting faulty condition and healthy condition under different motor speeds of (a) 400 r/min, (b) 600 r/min and (c) 800 r/min**

## 5. Conclusions

In this paper, a hybrid feature extraction methodology was proposed, where EMD, FastICA and a sample entropy measure are employed to enhance the quality of stator current feature extraction and enable gear pitting fault detection. Experimental studies have been conducted on a real test-rig to collect stator current signals and vibration signals from healthy and pitting fault gears. The experiment results show that the proposed method, compared to the traditional method of EMD and sample entropy measure, has a better detection quality for a gear pitting fault. Next, experimental results demonstrate that, based on vibration signal analysis, the proposed method can also effectively distinguish the gear pitting fault. However, from the perspective of signal acquisition, the stator current signal analysis method has a greater advantage than vibration signal analysis. Moreover, statistical results under different motor rotating speeds also further showed that the proposed method is feasible and effective for the detection of a gear pitting fault. In our future work, the proposed method will be used to detect other fault modes or healthy conditions, including other industrial application areas.

## Acknowledgements

This work is supported by the Natural Science Foundation of Hebei Province (F2011203149) and Scientific Research Project of the Higher Education Institutions of Hebei Province (ZD20131080). We also thank the support of Scientific Technology Research and the Development Programme of Qinhuangdao City (201302A218).

## References

1. Z Liu, J Qu, M J Zuo and H B Xu, 'Fault level diagnosis for planetary gearboxes using hybrid kernel feature selection and kernel Fisher discriminant analysis', *The International Journal of Advanced Manufacturing Technology*, Vol 67, No 5-8, pp 1217-1230, July 2013.
2. D G Lu, G Xiang and Q Wei, 'Current-based diagnosis for gear tooth breaks in wind turbine gearboxes', *Energy Conversion Congress and Exposition (ECCE), 2012 IEEE*, pp 3780-3786, 2012.
3. M Khazaei, H Ahmadi, M Omid, A Banakar and A Moosavian, 'A feature-level fusion based on wavelet transform and artificial neural network for fault diagnosis of planetary gearbox using acoustic and vibration signals', *Insight – Non-Destructive Testing and Condition Monitoring*, Vol 55, No 6, pp 323-330, June 2013.
4. A R Mohanty and C Kar, 'Fault detection in a multistage gearbox by demodulation of motor current waveform', *IEEE Transactions on Industrial Electronics*, Vol 53, No 4, pp 1285-1297, June 2006.
5. N Feki, G Clerc and Ph Velex, 'Gear and motor fault modelling and detection based on motor current analysis', *Electric Power Systems Research*, Vol 95, pp 28-37, February 2013.
6. S H Kia, H Henao and G A Capolino, 'Analytical and experimental study of gearbox mechanical effect on the induction machine stator current signature', *IEEE Transactions on Industry Applications*, Vol 45, No 4, pp 1405-1415, 2009.
7. A R Mohanty and C Kar, 'Multistage gearbox condition monitoring using motor current signature analysis and Kolmogorov-Smirnov test', *Journal of Sound and Vibration*, Vol 290, No 1-2, pp 337-368, February 2006.
8. N Feki, G Clerc and Ph Velex, 'An integrated electro-mechanical model of motor-gear units – applications to tooth fault detection by electric measurements', *Mechanical Systems and Signal Processing*, Vol 29, pp 377-390, May 2012.
9. C Kar and A R Mohanty, 'Vibration and current transient monitoring for gearbox fault detection using multiresolution Fourier transform', *Journal of Sound and Vibration*, Vol 311, No 1-2, pp 109-132, March 2008.
10. C Kar and A R Mohanty, 'Monitoring gear vibrations through motor current signature analysis and wavelet transform', *Mechanical Systems and Signal Processing*, Vol 20, No 1, pp 158-187, January 2006.
11. Z J Shen, X F Chen, X L Zhang and Z J He, 'A novel intelligent gear fault diagnosis model based on EMD and multi-class TSVM', *Measurement*, Vol 45, No 1, pp 30-40, January 2012.
12. S Li, D S Huang, J X Du and C H Zheng, 'Palmprint recognition using FastICA algorithm and radial basis probabilistic neural network', *Neurocomputing*, Vol 69, No 13-15, pp 1782-1786, August 2006.
13. Shih-Hsuan Chiu and C P Lu, 'Noise separation of the yarn tension signal on twister using FastICA', *Mechanical Systems and Signal Processing*, Vol 19, No 6, pp 1326-1336, November 2005.
14. T Wang, M C Zhang, Q H Yu and H Y Zhang, 'Comparing the applications of EMD and EEMD on time-frequency analysis of seismic signal', *Journal of Applied Geophysics*, Vol 83, pp 29-34, August 2012.
15. Y Gan, L F Sui and J F Wu, 'An EMD threshold de-noising method for inertial sensors', *Measurement*, Vol 49, pp 34-41, March 2014.
16. A Ayough, M Zandieh and H Farsijani, 'GA and ICA approaches to job rotation scheduling problem: considering employee's boredom', *The International Journal of Advanced Manufacturing Technology*, Vol 60, No 5-8, pp 651-666, May 2012.
17. Kuo-Kai Shyu, Ming-Huan Lee, Yu-Te Wu and Po-Lei Lee, 'Implementation of pipelined FastICA on FPGA for real-time blind source separation', *IEEE Transactions on Neural Networks*, Vol 1, No 6, pp 958-970, June 2008.
18. Y D Song, J Crowcroft and J X Zhang, 'Automatic epileptic seizure detection in EEGs based on optimised sample entropy and extreme learning machine', *Journal of Neuroscience Methods*, Vol 210, No 2, pp 132-146, September 2012.
19. N Li, Z X Wang and M Bai, 'Fault detection of induction motors based on EMD and ICA', *Journal of Tianjin University of Technology*, Vol 26, No 5, pp 69-70, October 2000.