

## Research Article

# A Fault Diagnosis Approach for Gears Based on IMF AR Model and SVM

Junsheng Cheng, Dejie Yu, and Yu Yang

*The State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body, Hunan University, Changsha 410082, China*

Correspondence should be addressed to Junsheng Cheng, signalp@tom.com

Received 24 July 2007; Revised 28 February 2008; Accepted 15 April 2008

Recommended by Nii Attoh-Okine

An accurate autoregressive (AR) model can reflect the characteristics of a dynamic system based on which the fault feature of gear vibration signal can be extracted without constructing mathematical model and studying the fault mechanism of gear vibration system, which are experienced by the time-frequency analysis methods. However, AR model can only be applied to stationary signals, while the gear fault vibration signals usually present nonstationary characteristics. Therefore, empirical mode decomposition (EMD), which can decompose the vibration signal into a finite number of intrinsic mode functions (IMFs), is introduced into feature extraction of gear vibration signals as a preprocessor before AR models are generated. On the other hand, by targeting the difficulties of obtaining sufficient fault samples in practice, support vector machine (SVM) is introduced into gear fault pattern recognition. In the proposed method in this paper, firstly, vibration signals are decomposed into a finite number of intrinsic mode functions, then the AR model of each IMF component is established; finally, the corresponding autoregressive parameters and the variance of remnant are regarded as the fault characteristic vectors and used as input parameters of SVM classifier to classify the working condition of gears. The experimental analysis results show that the proposed approach, in which IMF AR model and SVM are combined, can identify working condition of gears with a success rate of 100% even in the case of smaller number of samples.

Copyright © 2008 Junsheng Cheng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## 1. INTRODUCTION

The process of gear fault diagnosis includes the acquisition of information, extracting feature, and recognizing conditions, in which the last two are the prior.

Signal processing methods have been widely used to extract fault feature of gear vibration signals [1, 2]. Fourier transform (FT), which has been the dominating analysis tool for feature extraction of stationary signals, could produce the statistical average characteristics over the entire duration of the data. However, it fails to provide the whole and local features of the signal in time and frequency domain. Unfortunately, the gear fault vibration signals exactly present nonstationary characteristics. On the other hand, the time-frequency analysis methods can generate both time and frequency information of a signal simultaneously. Therefore, in the most recent studies, the time-frequency analysis methods are used in gear fault feature extraction [3–5]. Among all the available time-frequency analysis methods,

the wavelet transform may be the best one [6, 7], however, it still has some inevitable deficiencies [8]. Firstly, energy leakage will occur when wavelet transform is used to process signals due to the fact that wavelet transform is essentially an adjustable windowed Fourier transform. Secondly, the appropriate base function needs to be selected in advance. Moreover, once the decomposition scales are determined, the results of wavelet transform would be the signal under a certain frequency band. Therefore, wavelet transform is not a self-adaptive signal processing method in nature. In addition, the mathematical model needs to be established or the fault mechanism of the gear vibration system needs to be studied before the feature extraction in above-mentioned methods, which usually are quite difficult to be fulfilled in practice. Autoregressive (AR) model, which has no requirements of constructing mathematical model and studying the fault mechanism of a complex gear vibration system in advance, is a time sequence analysis method whose parameters comprise significant information of the system

condition; more importantly, an accurate AR model can reflect the characteristics of a dynamic system. Additionally, it is indicated that the autoregression parameters of AR model are very sensitive to the condition variation [9, 10]. The gear fault vibration signals own shock characteristics, whereas AR model can model transients and its frequency response function can be calculated from autoregression parameters of AR model. Therefore, the autoregression parameters can be used to analyze the condition variation of dynamic systems. However, when the AR model is applied to nonstationary signals, it is difficult to estimate autoregression parameters by the least square method or Yule-Walker equation method. The time-dependent autoregressive and moving average (ARMA) model, on the other hand, can be applied to nonstationary signals, but the more computation time is needed. Furthermore, only when the time-dependent ARMA model is applied to the commonly linear frequency and amplitude modulated signals, can the satisfactory results be obtained [11]. Therefore, it is necessary to preprocess the vibration signals before the AR model is generated. Empirical mode decomposition (EMD) is a new time-frequency analysis method proposed by Huang et al. [12, 13], which is based on the local characteristic time scale of signal and decomposes the complicated signal into a number of intrinsic mode functions (IMFs). By analyzing each IMF component that involves the local characteristic of the signal, the features of the original signal could be extracted more accurately and effectively. In addition, the frequency components involved in each IMF not only relates to sampling frequency but also changes with the signal itself, therefore EMD is a self-adaptive time frequency analysis method that is perfectly applicable to nonlinear and nonstationary processing. Now EMD method has been widely applied to the mechanical fault diagnosis and condition monitoring. In [14], EMD method is combined with smoothed nonlinear energy operator to detect flute breakage. The results demonstrate that this method can efficiently monitor the conditions of the endmill under varying cutting conditions. In [15], a fault diagnosis method for sheet metal stamping process based on EMD and learning vector quantization is proposed. The results show that this method could successfully detect the artificially created defects. In this paper, targeting the nonstationary characteristics of gear vibration signal and disadvantage of AR model, a fault feature extraction method in which IMF and AR model are combined is proposed.

After the feature extraction, the pattern recognition is another point of gears fault diagnosis [16–18]. Conventional statistical pattern recognition methods and artificial neural networks (ANNs) classifiers are studied based on the premise that the sufficient samples are available, which is not always true in practice [19]. In recent years, support vector machines (SVMs) have been found to be remarkably effective in many real-world applications [20–23]. They are based on statistical learning theories that are of specialties for a smaller sample number and have better generalization than ANNs and guarantee that the extremum and global optimal solution are exactly the same. Meantime, SVMs can solve the learning problem of a smaller number of samples [24, 25].

Due to the fact that it is difficult to obtain sufficient fault samples in practice, SVMs are introduced into gears fault diagnosis due to their high accuracy and good generalization for a smaller sample number in this paper.

## 2. EMD METHOD

EMD method is developed from the simple assumption that any signal consists of different simple intrinsic modes of oscillations. Each linear or nonlinear mode will have the same number of extrema and zero-crossings. There is only one extremum between successive zero-crossings. Each mode should be independent of the others. In this way, each signal could be decomposed into a number of intrinsic mode functions (IMFs), each of which must satisfy the following definition [12, 13].

- (1) In the whole dataset, the number of extrema and the number of zero-crossings must either equal or differ at most by one.
- (2) At any point, the mean value of the envelope defined by local maxima and the envelope defined by the local minima is zero.

An IMF represents a simple oscillatory mode compared with the simple harmonic function. With the definition, any signal  $x(t)$  can be decomposed as follows.

- (1) Identify all the local extrema, then connect all the local maxima by a cubic spline line as the upper envelope.
- (2) Repeat the procedure for the local minima to produce the lower envelope. The upper and lower envelopes should cover all the data between them.
- (3) The mean of upper and lower envelope value is designated as  $m_1$ , and the difference between the signal  $x(t)$  and  $m_1$  is the first component,  $h_1$ :

$$x(t) - m_1 = h_1. \quad (1)$$

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

- (4) If  $h_1$  is not an IMF,  $h_1$  is treated as the original signal and repeat (1), (2), (3), then

$$h_1 - m_{11} = h_{11}. \quad (2)$$

After repeated sifting, that is, up to  $k$  times,  $h_{1k}$  becomes an IMF:

$$h_{1(k-1)} - m_{1k} = h_{1k}, \quad (3)$$

then it is designated as

$$c_1 = h_{1k}, \quad (4)$$

the first IMF component from the original data.

- (5) Separate  $c_1$  from  $x(t)$ , we could get

$$r_1 = x(t) - c_1, \quad (5)$$

$r_1$  is treated as the original data and repeat the above processes, therefore the second IMF component  $c_2$  of  $x(t)$

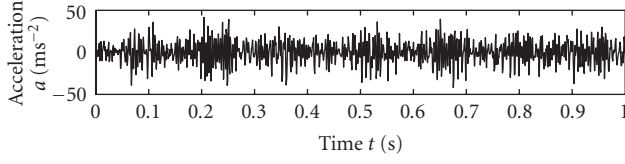


FIGURE 1: Acceleration vibration signal of a gear with a broken tooth.

could be got. Let us repeat the process as described above for  $n$  times, then  $n$ -IMFs of signal  $x(t)$  could be got. Then,

$$\begin{aligned} r_1 - c_2 &= r_2 \\ &\vdots \\ r_{n-1} - c_n &= r_n. \end{aligned} \quad (6)$$

The decomposition process can be stopped when  $r_n$  becomes a monotonic function from which no more IMF can be extracted. By summing up (5) and (6), we finally obtain

$$x(t) = \sum_{j=1}^n c_j + r_n. \quad (7)$$

Thus, one can achieve a decomposition of the signal into  $n$ -empirical modes and a residue  $r_n$ , which is the mean trend of  $x(t)$ . Each of the IMFs  $c_1, c_2, \dots, c_n$  includes different frequency bands ranging from high to low and is stationary.

Figure 1 shows an acceleration vibration signal of a gear with a broken tooth. It is decomposed into 5 IMFs and a remnant  $r_n$  by using EMD method as Figure 2 illustrates. It can be concluded from Figure 2 that each IMF component implies distinct time characteristic scale.

### 3. SUPPORT VECTOR MACHINES (SVMs)

SVM is developed from the optimal separation plane under linearly separable condition. Its basic principle can be illustrated in two-dimensional way as Figure 3 [25]. Figure 3 shows the classification of a series of points for two different classes of data, class A (circles) and class B (stars). The SVM tries to place a linear boundary  $H$  between the two classes and orients it in such way that the margin is maximized, namely, the distance between the boundary and the nearest data point in each class is maximal. The nearest data points are used to define the margin and are known as support vectors.

Suppose there is a given training sample set  $G = \{(x_i, y_i), i = 1 \dots l\}$ , each sample  $x_i \in \mathbb{R}^d$  belongs to a class by  $y \in \{+1, -1\}$ . The boundary can be expressed as follows:

$$\omega \cdot x + b = 0, \quad (8)$$

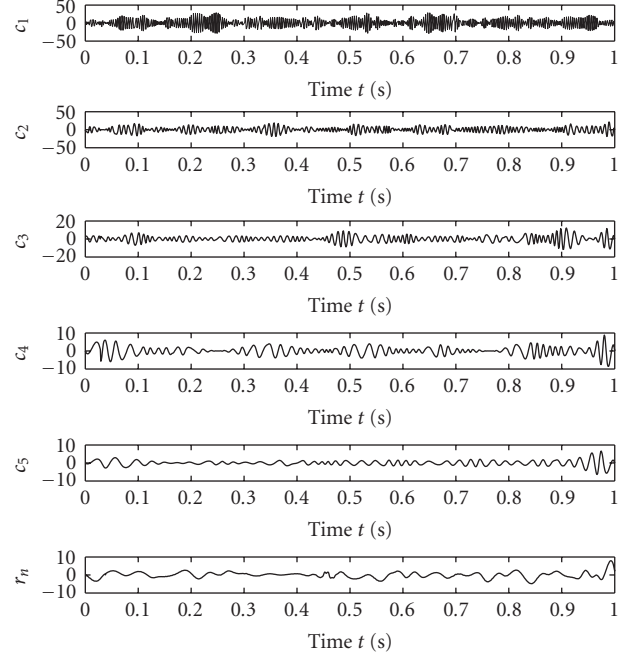


FIGURE 2: The EMD results of a gear vibration signal.

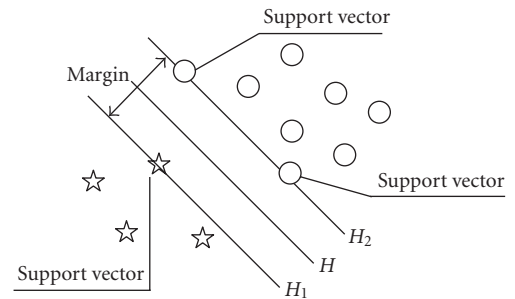


FIGURE 3: Classification of data by SVM.

where  $\omega$  is a weight vector and  $b$  is a bias. So the following decision function can be used to classify any data point in either class A or B:

$$f(x) = \text{sign}(\omega \cdot x + b). \quad (9)$$

The optimal hyperplane separating the data can be obtained as a solution to the following constrained optimization problem:

$$\begin{aligned} &\text{minimize} \quad \frac{1}{2} \|\omega\|^2, \\ &\text{subject to} \quad y_i [(\omega \cdot x_i) + b] - 1 \geq 0, \quad i = 1, \dots, l. \end{aligned} \quad (10)$$

Introducing Lagrange multipliers  $\alpha_i \geq 0$ , the optimization problem can be rewritten as

$$\begin{aligned} \text{minimize } L(\omega, b, \alpha) &= \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j (x_i \cdot x_j), \\ \text{subject to } \alpha_i &\geq 0, \\ \sum_{i=1}^l \alpha_i y_i &= 0. \end{aligned} \quad (11)$$

The decision function can be obtained as follows:

$$f(x) = \text{sign} \left( \sum_{i=1}^l \alpha_i y_i (x_i \cdot x) + b \right). \quad (12)$$

If the linear boundary in the input spaces is not enough to separate into two classes properly, it is possible to create a hyperplane that allows linear separation in the higher dimension. In SVM, it is achieved by using a transformation  $\Phi(x)$  that maps the data from input space to feature space. If a kernel function

$$K(x, y) = \Phi(x) \cdot \Phi(y) \quad (13)$$

is introduced to perform the transformation, the basic form of SVM can be obtained:

$$f(x) = \text{sign} \left( \sum_{i=1}^l \alpha_i y_i K(x, x_i) + b \right). \quad (14)$$

Among the kernel functions in common use are linear functions, polynomials functions, radial basis functions, and sigmoid functions.

#### 4. DIAGNOSIS APPROACH FOR GEARS BASED ON IMF AR MODEL AND SVM

The following autoregressive model  $AR(m)$  could be established for each IMF component  $c_i(t)$  in (7) [26]:

$$c_i(t) + \sum_{k=1}^m \varphi_{ik} c_i(t-k) = e_i(t), \quad (15)$$

where  $\varphi_{ik}$  ( $k = 1, 2, \dots, m$ ),  $m$  are the model parameters and model order of the autoregressive model  $AR(m)$  of  $c_i(t)$ , respectively;  $e_i(t)$  is the remnant of the model and is a white noises sequence whose mean value is zero and variance is  $\sigma_i^2$ . Since the parameters  $\varphi_{ik}$  can reflect the inherent characteristics of a gear vibration system and the variance of the remnant  $\sigma_i^2$  is tightly related with the output characteristics of the system,  $\varphi_{ik}$  and  $\sigma_i^2$  can be chosen as feature vectors  $A_i = [\varphi_{i1}, \varphi_{i2}, \dots, \varphi_{im}, \sigma_i^2]$  to identify the condition of the gears system.

The flow chart of a diagnosis method proposed in this paper is illustrated in Figure 4.

The fault diagnosis approach for gears based on IMF AR model and SVM is represented as follows.

(1) Sample signals  $N$  times at a certain sample frequency  $f_s$  under the circumstance that the gear is normal and the

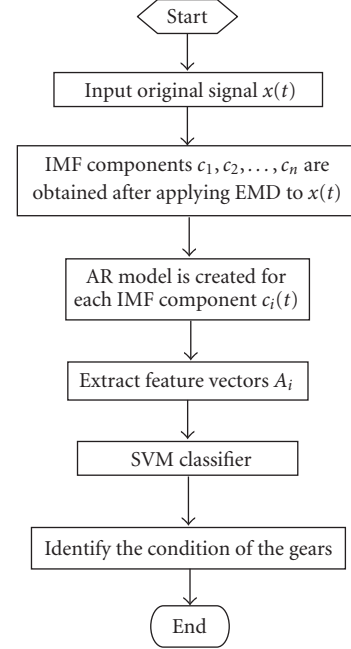


FIGURE 4: The flow chart of the proposed method.

gear has the crack faults. And the  $2N$  signals are taken as samples that are divided into two subsets, the training samples and test samples.

(2) Each signal is decomposed by EMD. Different signal has different amount of the IMFs, denoted by  $n_1, n_2, \dots, n_{2N}$ , and let  $n = \max(n_1, n_2, \dots, n_{2N})$ . If some samples whose amount  $n_k$  ( $k = 1, 2, \dots, 2N$ ) of IMF components is less than  $n$ , it can be padded with zero to  $n$  components  $c_1(t), c_2(t), \dots, c_n(t)$ , that is  $c_i(t) = \{0\}$ ,  $i = n_k + 1, n_k + 2, \dots, n$ .

(3) In order to eliminate the effect of the signal amplitude to the variance of the remnant  $\sigma_i^2$ , normalize each IMF component to achieve a new component:

$$\hat{c}_i(t) = \frac{c_i(t)}{\sqrt{\int_{-\infty}^{\infty} c_i^2(t) dt}}. \quad (16)$$

(4) Establish AR model for the normalized component, determine the order  $m$  of the model and estimate autoregressive parameters  $\varphi_{ik}$  ( $k = 1, 2, \dots, m$ ) and the remnant's variance  $\sigma_i^2$ , where  $\varphi_{ik}$  means the  $k$ th autoregressive parameters of the  $i$ th IMF component. Therefore, the feature vector used as input vector of SVMs is as follows:  $A_i = [\varphi_{i1}, \varphi_{i2}, \dots, \varphi_{im}, \sigma_i^2]$ .

(5) Separate the training set into two classes:  $y = +1$  and  $y = -1$ , which represent two kinds of working condition of the gears, namely, the normal gear and the gear with crack fault. Actually, the decision function  $f(x)$  is determined only by the support vectors, so after the support vectors are obtained the feature vector of test samples can be input into the trained SVM classifier and then the working condition can be classified by the output of the SVMs classifier.

TABLE 1: The identification results based on IMF AR model and SVM.

| Conditions of the signals | IMF   | Feature vectors |                |                |              | Distance           |                    | Results |
|---------------------------|-------|-----------------|----------------|----------------|--------------|--------------------|--------------------|---------|
|                           |       | $\varphi_{i1}$  | $\varphi_{i2}$ | $\varphi_{i3}$ | $\sigma_i^2$ | 6 training samples | 3 training samples |         |
| Normal                    | $c_1$ | 0.4488          | 0.2870         | 0.2498         | 2.1331       |                    |                    |         |
|                           | $c_2$ | -0.7683         | 1.5523         | -1.0823        | 0.9972       | 1.4313             | 0.9421             | +1      |
|                           | $c_3$ | -2.1518         | 2.6944         | -2.0254        | 0.2134       |                    |                    |         |
| Normal                    | $c_1$ | 0.3980          | 0.1908         | 0.2330         | 1.7583       |                    |                    |         |
|                           | $c_2$ | -1.0207         | 1.8408         | -1.6746        | 0.7681       | 1.3609             | 1.0774             | +1      |
|                           | $c_3$ | -2.1360         | 2.7934         | -2.2215        | 0.1856       |                    |                    |         |
| Normal                    | $c_1$ | 0.5110          | 0.2482         | 0.2179         | 2.0377       |                    |                    |         |
|                           | $c_2$ | -0.7941         | 1.5924         | -1.1135        | 0.9576       | 1.7666             | 1.4178             | +1      |
|                           | $c_3$ | -2.0363         | 2.4411         | -1.5479        | 0.2315       |                    |                    |         |
| Crack fault               | $c_1$ | 0.0545          | 6.7798         | 0.1888         | 1.2081       |                    |                    |         |
|                           | $c_2$ | -1.7086         | 2.0489         | -1.3569        | 0.4271       | -1.7755            | -1.5707            | -1      |
|                           | $c_3$ | -2.8216         | 3.9288         | -3.2710        | 0.0439       |                    |                    |         |
| Crack fault               | $c_1$ | 0.0072          | 0.7102         | 0.2035         | 1.0662       |                    |                    |         |
|                           | $c_2$ | -1.7070         | 2.0933         | -1.5511        | 0.3248       | -1.2758            | -1.0311            | -1      |
|                           | $c_3$ | -2.8072         | 3.7685         | -2.9271        | 0.0321       |                    |                    |         |
| Crack fault               | $c_1$ | 0.1515          | 0.5989         | 0.0622         | 1.5854       |                    |                    |         |
|                           | $c_2$ | -1.4817         | 1.8108         | -1.1972        | 0.5092       | -1.5496            | -1.5219            | -1      |
|                           | $c_3$ | -2.8286         | 4.0104         | -3.4727        | 0.0436       |                    |                    |         |

## 5. APPLICATIONS

An experiment has been carried out on the small experiment-rig developed by the Vibration and Test Center of Hunan University itself. The fault is introduced by cutting slot with laser in the root of tooth, and the width of the slot is 0.15–0.25 mm, as well as its depth is 0.1–0.3 mm. The acceleration sensor has been fixed on the cover of the gear box before 30 signals under two circumstances are sampled with sample frequency of 1024 Hz, among which three randomly chosen samples for each condition are taken as training samples, and the remain are test data.

Decompose each vibration signals under different conditions with EMD method into a number of IMFs. The analysis results show that the fault information of gear vibration signals is mainly included in the first three IMF components. Therefore, the AR models of the first three IMF components are established merely. In this paper, the order of the model,  $m$ , is determined with FPE criterion [26]; the autoregressive parameters  $\varphi_{ik}$  ( $k = 1, 2, \dots, m$ ) and the remnant variance  $\sigma_i^2$  of the model are computed with least squares criterion [26]. As, in fact, the system condition is mainly decided by the autoregressive parameters of the first several ones and the remnant variance, those of only the first three ones, that is  $\varphi_{ik}$  ( $k = 1, 2, 3$ ) and  $\sigma_i^2$ , are chosen as feature vectors in this paper for convenience.

Define the normal condition as  $y = +1$  and the one with the crack fault as  $y = -1$ ; choose the linear kernel function to calculate and by formulas (11) we can obtain the parameters of SVM classifier,  $\alpha = [0, 0.1699, 0.6091, 0.7790, 0, 0]^T$ ,  $\|\omega\| = 1.2482$ , and  $b = 2.5942$ . Then, by formula (12) the identification result of each test sample is obtained, part of which are shown in Table 1. Obviously, the identification

results are totally consistent with the fact. For further study of the application of SVMs in the pattern identification with smaller number of samples, the number of training samples decrease to three (one is normal and the others is with crack fault) and the calculation procedure is the same as above. Here, the parameters of the SVM classifier become  $\alpha = [0.5014, 0.5014, 0]^T$ ,  $\|\omega\| = 1.0014$ ,  $b = 2.5485$ . The identification results to the same test samples are shown in Table 1 too.

It can be seen from Table 1 that SVM classifier can still classify the two conditions of gears accurately after the training samples are decreased, which confirm fully that the SVM classifier can be applied successfully to the pattern recognition even in cases where only limited training samples are available. It also can be found, if we compare the distances between test samples with different number of training samples to the optimal separating hyperplane  $H$ , that the distance decreases after the number of training samples become smaller although the gear work states can still be identified by SVM, which shows that in this way the whole performance of the classifier somewhat reduces.

What we discuss above is how to classify two conditions of gears (normal and crack fault), that is, two-class problem. When it comes to the multiple-class problems, that is, how to identify the gears with multiple-class faults (e.g., crack, broken teeth, etc.), generalizing method can be introduced to decompose the multiple-class problems into two-class problems which then can be trained with SVM. In other words, each time take one group of the training samples as one class and therest, which do not belong to the former, can be taken as the other class. Hence, for the  $k$  ( $k \geq 3$ ) classes' problems, the classification of the input space can be achieved by  $k$  decision-functions based on SVM.



TABLE 2: The identification results based on IMF AR model and SVMs.

| Conditions of the signals | SVM classifier |      |      | Identification results |
|---------------------------|----------------|------|------|------------------------|
|                           | SVM1           | SVM2 | SVM3 |                        |
| Normal                    | +1             |      |      | Normal                 |
| Normal                    | +1             |      |      | Normal                 |
| Crack fault               | -1             | +1   |      | Crack fault            |
| Crack fault               | -1             | +1   |      | Crack fault            |
| Broken teeth              | -1             | -1   | +1   | Broken teeth           |
| Broken teeth              | -1             | -1   | +1   | Broken teeth           |

Three SVM classifiers are needed to design if three classes of gear work conditions are to be identified like normal, with crack fault and with broken teeth fault. First of all, define that  $y = +1$  represents the normal condition and  $y = -1$  represents the faults condition, that is, identify the gear whether it has fault or not by SVM1. Secondly, identify the gear whether it has crack fault or not by SVM2, here  $y = +1$  represents crack fault and  $y = -1$  represents other faults. Finally, identify the gear whether it has broken teeth fault or not, here  $y = +1$  represents broken teeth fault and  $y = -1$  represents other faults. The identification approach is the same as above, that is, extract nine samples as training ones at random (three samples with normal condition, three samples with crack fault, and three samples with broken teeth fault); and then calculate the parameters of SVM classifier. The part identification results are shown in Table 2 from which we can see that three SVM classifiers can identify the working conditions and fault patterns of gears accurately.

## 6. CONCLUSIONS

AR model is an information container that contains the characteristics of gear vibration systems, based on which the fault feature of gear vibration signal can be extracted. The most important is that the gear work states can be identified by the parameters of the AR model after the AR model of vibration signals is established without constructing mathematical model and studying the fault mechanism. However, AR model can only be applied to stationary signals, while the gear fault vibration signals always display nonstationary behavior. To target this problem, in this paper before AR model is established, a preprocessing on gear fault vibration signals is carried out with EMD method, which can decompose a signal, in terms of its intrinsic information, into a number of IMFs. The decomposition of EMD is a process of origin signal linearization and stationary in nature, thus AR model can be established for each of the IMF components.

The limitations of the conventional statistical pattern recognition methods and ANNs classifiers are targeted. Support vector machine, which has better generalization than ANNs and can solve the learning problem of smaller number of samples quite well, has been introduced into the pattern recognition.

By the analysis results of three kinds of gears vibration signals among which one is normal and the other two are the gears with crack and gears with broken tooth faults respectively, it has been shown that the gear fault diagnosis

approach based on IMF AR model and SVM can be applied to classify the gear working conditions and fault patterns effectively and accurately even in case of smaller number of samples, which accordingly offers a new approach for the fault diagnosis of gears. However, because it would take more time to determine the parameters of SVM classifier and the AR model, the proposed method cannot be available in real-time. In addition, what is necessary to point out is that the SVM theory is still in its perfecting phase, for example, the problems of kernel functions selection in different condition and so on are still needed to research further.

## ACKNOWLEDGMENT

The support for this research under Chinese National Science Foundation Grant no. 50775068 is gratefully acknowledged.

## REFERENCES

- [1] W. Q. Wang, F. Ismail, and M. F. Golnaraghi, "Assessment of gear damage monitoring techniques using vibration measurements," *Mechanical Systems and Signal Processing*, vol. 15, no. 5, pp. 905–922, 2001.
- [2] D. Brie, M. Tomczak, H. Oehlmann, and A. Richard, "Gear crack detection by adaptive amplitude and phase demodulation," *Mechanical Systems and Signal Processing*, vol. 11, no. 1, pp. 149–167, 1997.
- [3] W. J. Staszewski, K. Worden, and G. R. Tomlinson, "Time-frequency analysis in gearbox fault detection using the Wigner-Ville distribution and pattern recognition," *Mechanical Systems and Signal Processing*, vol. 11, no. 5, pp. 673–692, 1997.
- [4] N. Baydar and A. Ball, "A comparative study of acoustic and vibration signals in detection of gear failures using Wigner-Ville distribution," *Mechanical Systems and Signal Processing*, vol. 15, no. 6, pp. 1091–1107, 2001.
- [5] H. Oehlmann, D. Brie, M. Tomczak, and A. Richard, "A method for analysing gearbox faults using time-frequency representations," *Mechanical Systems and Signal Processing*, vol. 11, no. 4, pp. 529–545, 1997.
- [6] J. Lin and M. J. Zuo, "Gearbox fault diagnosis using adaptive wavelet filter," *Mechanical Systems and Signal Processing*, vol. 17, no. 6, pp. 1259–1269, 2003.
- [7] G. Meltzer and N. P. Dien, "Fault diagnosis in gears operating under non-stationary rotational speed using polar wavelet amplitude maps," *Mechanical Systems and Signal Processing*, vol. 18, no. 5, pp. 985–992, 2004.

- [8] P. W. Tse, W.-X. Yang, and H. Y. Tam, "Machine fault diagnosis through an effective exact wavelet analysis," *Journal of Sound and Vibration*, vol. 277, no. 4-5, pp. 1005–1024, 2004.
- [9] D. Hong, W. Ya, and Y. Shuzi, "Fault diagnosis by time series analysis," in *Applied Time Series Analysis*, World Scientific, River Edge, NJ, USA, 1989.
- [10] W. Ya and Y. Shuzi, "Application of several time series models in prediction," in *Applied Time Series Analysis*, World Scientific, River Edge, NJ, USA, 1989.
- [11] Y. Grenier, "Time-dependent ARMA modeling of nonstationary signals," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 31, no. 4, pp. 899–911, 1983.
- [12] N. E. Huang, Z. Shen, S. R. Long, et al., "The empirical mode decomposition and the Hubert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society A*, vol. 454, no. 1971, pp. 903–995, 1998.
- [13] N. E. Huang, Z. Shen, and S. R. Long, "A new view of nonlinear water waves: the Hilbert spectrum," *Annual Review of Fluid Mechanics*, vol. 31, pp. 417–457, 1999.
- [14] A. M. Bassiuny and X. Li, "Flute breakage detection during end milling using Hilbert-Huang transform and smoothed nonlinear energy operator," *International Journal of Machine Tools and Manufacture*, vol. 47, no. 6, pp. 1011–1020, 2007.
- [15] A. M. Bassiuny, X. Li, and R. Du, "Fault diagnosis of stamping process based on empirical mode decomposition and learning vector quantization," *International Journal of Machine Tools and Manufacture*, vol. 47, no. 15, pp. 2298–2306, 2007.
- [16] L. Wuxing, P. W. Tse, Z. Guicai, and S. Tielin, "Classification of gear faults using cumulants and the radial basis function network," *Mechanical Systems and Signal Processing*, vol. 18, no. 2, pp. 381–389, 2004.
- [17] B. Samanta, "Artificial neural networks and genetic algorithms for gear fault detection," *Mechanical Systems and Signal Processing*, vol. 18, no. 5, pp. 1273–1282, 2004.
- [18] X. Li, S. K. Tso, and J. Wang, "Real-time tool condition monitoring using wavelet transforms and fuzzy techniques," *IEEE Transactions on Systems, Man and Cybernetics C*, vol. 30, no. 3, pp. 352–357, 2000.
- [19] M. Zacksenhouse, S. Braun, M. Feldman, and M. Sidahmed, "Toward helicopter gearbox diagnostics from a small number of examples," *Mechanical Systems and Signal Processing*, vol. 14, no. 4, pp. 523–543, 2000.
- [20] H.-C. Kim, S. Pang, H.-M. Je, D. Kim, and S. Y. Bang, "Constructing support vector machine ensemble," *Pattern Recognition*, vol. 36, no. 12, pp. 2757–2767, 2003.
- [21] G. Guo, S. Z. Li, and K. L. Chan, "Support vector machines for face recognition," *Image and Vision Computing*, vol. 19, no. 9-10, pp. 631–638, 2001.
- [22] O. Barzilay and V. L. Brailovsky, "On domain knowledge and feature selection using a support vector machine," *Pattern Recognition Letters*, vol. 20, no. 5, pp. 475–484, 1999.
- [23] U. Thissen, R. van Brakel, A. P. de Weijer, W. J. Melssen, and L. M. C. Buydens, "Using support vector machines for time series prediction," *Chemometrics and Intelligent Laboratory Systems*, vol. 69, no. 1-2, pp. 35–49, 2003.
- [24] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, NY, USA, 1995.
- [25] V. N. Vapnik, *Statistical Learning Theory*, John Wiley & Sons, New York, NY, USA, 1998.
- [26] G. C. Goodwin and R. L. Payne, *Dynamic System Identification, Experiment Design and Data Analysis*, Academic Press, New York, NY, USA, 1977.