Available online at www.sciencedirect.com
ScienceDirect
Journal of Sound and Vibration 323 (2009) 1077–1089
JOURNAL OF SOUND AND VIBRATION
www.elsevier.com/locate/jsvi

Fault classification of rolling bearing based on reconstructed phase space and Gaussian mixture model

Guo Feng Wang*, Yu Bo Li, Zhi Gao Luo

Mechanical Engineering College, Tianjin University, Tianjin 300072, China

Received 28 July 2007; received in revised form 19 July 2008; accepted 2 January 2009
Handling Editor: L.G. Tham
Available online 17 March 2009

Abstract

Rolling bearings are common and vital elements in rotating machinery and vibration signal is a kind of effective mean to characterize the status of rolling bearing fault and its severity. In this paper, a novel method is introduced to realize classification of fault signal without extracting feature vector preliminarily. By estimating the time delay and embedding dimension of time series, vibration signal is reconstructed into phase space and Gaussian mixture model (GMM) is established for every kind of fault signal in the reconstructed phase space. After these models are built, classification of fault signal is accomplished by computing the conditional likelihoods of the signal under each learned GMM model and selecting the model with the highest likelihood. By testifying of vibration signal under different kinds of bearing status, it is proved that this method is effective for classifying not only fault types but also fault severity. Moreover, all parameters needed in this method could be obtained by analyzing the time series directly so it is very suitable for industry application.

1. Introduction

Rolling bearings are commonly used components in many kinds of machinery. How to identify its fault status and judge its fault type and severity is important in condition monitoring of mechanical equipment. Vibration analysis has been proven to be an effective means for fault diagnosis. For decades, lots of methods have been proposed to extract and analyze features of vibration signal so as to realize fault diagnosis. These methods include autoregressive modeling [1], empirical model decomposition (EMD) [2] and wavelet and wavelet packet methods [3–5], etc. However, mechanical vibration signals demonstrate lots of nonlinear characteristics and the traditional feature extraction methods mentioned above cannot effectively extract these nonlinear fault features [6]. So recently, many researchers try to extract these nonlinear features from vibration signal and judge the fault type according to them. Jingqiu proposes an improved correlation dimension algorithm to realize the diagnosis of rolling bearings [7]. Yan propose that computation of correlation dimension is too complex and Lempel–Ziv complexity [8] and approximate entropy [9] are more suitable for characterizing the nonlinear feature of bearing fault. To realize more accurate fault diagnosis, multiple

*Corresponding author. Tel.: +86 22 81831363.
E-mail address: gfwangmail@tju.edu.cn (G.F. Wang).

0022-460X/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsv.2009.01.003

nonlinear features such as box-counting dimension, multiscale fractal dimension [10,11], and even mel-frequency cepstral coefficients [12] are extracted simultaneously from vibration signal and support vector machine (SVM) [6] and hidden Markov model (HMM) [12] is constructed to realize the mapping between fault type and feature vector. These results show that fault diagnosis based on nonlinear feature extraction is an effective method. But in real application, it is quite a difficult problem on how to select or determine the threshold value of these nonlinear features for finite sample signal. Moreover these features are always not sensitive to change of fault severity, especially when noise exists. So up to now, most nonlinear features are selected only to judge the type of fault and the threshold value is determined according to different equipment which will spend lots of time. In this paper, the periodicity and nonlinearity of vibration signal is characterized by rearranging the vibration signal into reconstructed phase space (RPS) in which the phase trajectory of different kinds of signal will demonstrate different structure. It is worthy of noting that the size of fault could also produce obvious influence on phase trajectory in RPS. To describe the phase trajectory of different signal statistically, Gaussian mixture model (GMM) is utilized to fit the distribution of phase trajectory for every fault signal. After GMM for every kind of signal is learned, maximum likelihood (ML) Bayesian classifier is utilized to realize the classification of vibration signal. The advantage of this method is that all parameters are obtained by analyzing the time series directly. By analyzing the bearing signal of different fault status, it is proven that this method is very effective for classification of not only fault type but also its severity.

The remainder of this paper is organized as follows: Section 2 describes the main principle of method used in this paper which includes method of RPS and GMM. The whole scheme of fault classification is also given in this section. In Section 3, vibration signal from different bearing status is collected and analyzed using methods proposed above. Moreover, selection of time delay, embedding dimension and number of GMM are discussed in this section and the optimal value of these parameters are determined for learning of GMM in RPS and realizing classification of fault type and its severity simultaneously. In Section 4, some useful conclusions are drawn and emphasis of further study is also given.

2. Methods and principles

2.1. Phase space reconstruction

In order to characterize the nonlinear feature of vibration signal, the dynamic system embedded in time series need to be reconstructed. According to Takens's theorem [13,14], a dynamic system can be obtained by reconstructing phase space which is same as original system. RPS technique is founded on underlying principles of dynamical system theory and have been practically applied to a variety of nonlinear signals processing applications [15,16]. The basic idea of the RPS is that a scalar time series $x(t)$ may be used to construct a vector time series that is equivalent to the original dynamics from a topological point of view. Based on time series $x=(x_1, x_2,\ldots,x_N)$, RPS matrix X could be obtained in which row vector is defined as

$$X_n=[x_{n-(d-1)\tau}, x_{n-d\tau}, \ldots, x_{n-\tau}, x_n] \tag{1}$$

where n is index of row vector of matrix X which range from $1+(d-1)\tau$ to N. Symbol N denotes the length of time series. d and τ is embedded dimension and time delay of RPS, respectively. A row vector X_n is a point in the RPS. The sufficient condition for topological equivalence of RPS with original system is that d is greater than twice the box counting dimension of the original system.

When d is not known, as is the case for most real systems, there are many methods to calculate the embedded dimension. Cao' method [14] is adopted in this paper. The advantage of this method is its robustness to noise for vibration signal and there is no need to set the threshold value manually. Principle of Cao' method is described as follows. Suppose the dimension of RPS is d, the ith vector in this d dimensional phase space is given as

$$X_i^d=[x_{i-(d-1)\tau}, x_{i-d\tau}, \ldots, x_{i-\tau}, x_i] \tag{2}$$

In this d dimensional space, one vector $X_{\eta(i)}^d$ is nearest to the vector X_i^d, that is

$$\|X_i^d - X_{\eta(i)}^d\| = \min_{j=n-(d-1)\tau,\ldots,N,\, j\neq i} \|X_i^d - X_j^d\|_\infty \tag{3}$$

where $X_j^d \in R^d$ and $\| \bullet \|_\infty$ is the L_∞ norm which is defined in R^d space. Similarly, if the dimension change from d to $d+1$, the ith vector X_i^{d+1} and its nearest vector $X_{\eta(i)}^{d+1}$ could also be obtained in the phase space of $d+1$ dimension. Then $a(i, d)$ is defined as

$$a(i,d) = \frac{\|X_i^{d+1} - X_{\eta(i)}^{d+1}\|}{\|X_i^d - X_{\eta(i)}^d\|} \tag{4}$$

For all vectors X_i $(i = 1+(d-1)\tau,\ldots,N)$ in m and $m+1$ dimensional phase space, $E(d)$ and $E_1(d)$ are defined as [14]

$$E(d) = \frac{1}{N - d\tau} \sum_{i=(d-1)\tau}^{N} a(i,d) \tag{5}$$

and

$$E_1(d) = \frac{E(d+1)}{E(d)} \tag{6}$$

If signal is deterministic, the embedding dimension exists. That is, when d is larger than a specified value d_0, E_1 tends to a stable value and d_0 is just the correct embedding dimension. But in real application, the length of time series is always finite which results in difficult for judging whether the E_1 is stable or not. So other parameters are proposed which is given as [14]

$$E^*(d) = \frac{1}{N - d\tau} \sum_{i=(d-1)\tau}^{N} |x_{i-d\tau}^d - x_{\eta(i)-d\tau}^d| \tag{7}$$

and

$$E_2(d) = \frac{E^*(d+1)}{E^*(d)} \tag{8}$$

For a stochastic signal, E_2 is equal to one and for a deterministic time series E_2 change with the dimension d. By analyzing the varying of E_1 and E_2 with the dimension d, the embedding dimension of time series could be obtained. By Taken's theory [13], the time lag τ in Eq. (2) could be set to one. However, in practice, it has been found that the appropriate selection of the time lag could reduce the required RPS dimension. A common heuristic for determining time lag is to use the first minimum of the mutual information function. The average mutual information of time series $x = (x_1,x_2,\ldots,x_N)$ is described as

$$I(\tau) = \sum_{i=1}^{N} P(x_i, x_{i+\tau}) \log_2 \left(\frac{P(x_i, x_{i+\tau})}{P(x_i)P(x_{i+\tau})} \right) \tag{9}$$

where $P(x_i)$, $P(x_{i+\tau})$ is the probability density of x_i and $x_{i+\tau}$, respectively. Symbol $P(x_i,x_{i+\tau})$ is joint probability density which could be obtained form two-dimensional histogram. So it is easy to get the curve $I(\tau)$ which represents different mutual information at different lag τ. The time lag of phase space is the first local minimum at the $I(\tau)$ curve. Too small τ value results in weak independency of time series. On the other hand, if the τ is too large, the variable may become independent. Therefore correct selection of time lag is paramount.

2.2. GMM method

Fault of rolling bearing leads to change of dynamic characteristic of mechanical system and RPS is capable of differentiating between signals generated by topologically different systems. By phase space reconstruction of time series, phase trajectory of dynamical systems could be obtained. Generally some geometrical invariants such as correlation dimension, Kolmogorov entropy and Lyapunov exponents are extracted from phase space and adopted as feature vector for classification [8–10]. Because these features are calculated mostly in RPS, so in comparison with these features, phase trajectory could provide more complete description of reconstructed

dynamical system especially in the ability to characterize the periodicity and chaos of time series. Thus in this paper, GMM is used to describe the probability distribution of multidimension vector in phase space. One advantage of this model is its robustness to additive noise components which really exist in mechanical signal [10,12]. Another advantage is that there is no need to concern which geometrical invariants or how many nonlinear features need to be selected. Moreover, these features always need to be determined in advance and usually they are selected empirically [13].

For a class of vibration signals, GMM is used to fit the probability distribution of phase points in RPS. For a group of points in phase space of d dimension, the probability density function is given as

$$p(x) = \sum_{m=1}^{M} \omega_m p_m(x) = \sum_{m=1}^{M} \omega_m N\left(x, \mu_m, \sum_m\right) \tag{10}$$

Symbol $N(x, \mu_m, \sum_m)$ is the Gaussian probability density of d dimension in the phase space. ω_m is the weight value of each Gaussian distribution and M is the number of mixture model. This equation means that probability density of a group of points in d dimensional space could be fit by the weighted sum of several Gaussian probability density functions. Weight value of each Gaussian distribution satisfy

$$\sum_{m=1}^{M} \omega_m = 1 \tag{11}$$

The parameters for the GMM is $(\omega_m, \mu_m, \sum_m)$ which are estimated using the well-known expectation–maximization (EM) algorithm [13]. This iterative method yields a ML estimation of these parameters. For each class of signals, a GMM is built after it is reconstructed in phase space. The only parameter need to be determined in advance is number of mixture model. There are many methods to determine the number of mixture model such as AIC, BIC and some other methods which are referred in Ref. [17]. But it is worthy of noting that this method is instructive and always inaccurate especially when noise exists. So in real application, it is determined manually and it is proven to be feasible by analyzing rolling bearing vibration signal in the next section.

2.3. Bayesian classifier

By building GMM model, a statistics model is constructed for every fault signal which represents the probability distribution of every fault signal in RPS. Thus, for a new signal, Bayesian ML classifier is used here to judge which class it belongs to. That is obtained by computing the conditional likelihoods of the signal under each learned model and selecting the model with the highest likelihood. The likelihoods are computed on a point-by-point basis from the learned GMM models:

$$p(X|c_i) = \prod_{n=1+(d-1)\tau}^{N} p(X_n|c_i) \tag{12}$$

where X is an RPS matrix of dimension d and time lag τ of the signal, X_n is a point in the RPS, and $p(X_n|c_i)$ is the probability of X_n given the ith class calculated using Eq. (12). The classification is realized by calculating

$$\hat{c} = \arg\max_i(p(X|c_i)) \tag{13}$$

where $\hat{c}$ is the class which corresponds to the ML.

2.4. Method of classification based on RPS and GMM

As mentioned above, the RPS could be adopted to depict the nonlinear characteristic in the original time series and GMM could describe the distribution of points in the RPS. So integration of these two methods could be utilized to realize the classification of rolling bearing signal. The algorithm includes two stages.

The first is learning or training stage in which RPS for every kind of fault signal is constructed and GMM is built. This stage is described as the following three steps:

Step 1: Preparing for data used for learning of GMM. The main purpose of this step is to normalize the time series so that the phase trajectory in every RPS has same scope.

Step 2: Calculating the time lag and embedding dimension of RPS. For comparison with each other, time lag and embedding dimension must be determined and the same value for every kind of signal is selected. So in this step time lag and embedding dimension must be calculated for every signal so as to judge whether a uniformed value could be selected or not.

Step 3: Learning the GMM for each class. All samples are constructed in RPS and input into GMM so as to calculate weight value of GMM. This step is done for every kind of fault signal so that all sample data is fitted into GMM.

The second stage is classification. For a new vibration signal, its phase space is reconstructed firstly so as to obtain its phase trajectory and input it into all GMM trained in the previous stage so as obtain the probability of it belongs to each GMM, respectively. The class corresponding to the maximum probability is just the signal belongs to.

3. Experiment and discussions

3.1. Experiment data description

In order to testify the effect of method proposed above, experimental data were collected from ball bearing of an induction motor driven mechanical system which was tested under 1 hp load [18].The whole scheme of test rig is shown in Fig. 1. The test rig consists of four parts which are motor, torque transducer, dynameters, and control electronics, respectively. Two SKF rolling bearings support the motor shaft. An accelerometer was mounted on the motor housing at the drive end and fan end of the motor to acquire the vibration signals from the bearing. Single point faults were introduced to the test bearings using electro-discharge machining and motor speed was 1772 rev/min. The signal was sampled from four kinds of rolling element bearings and each class of data corresponds to the following bearing conditions, respectively: (i) normal bearing; (ii) inner race fault; (iii) ball fault; and (iv) outer race fault. Moreover, in order to testify the ability to classify the fault severity, three kinds of fault sizes are utilized for the same fault type. Here the fault diameter is 0.007, 0.014, and 0.021 in for every fault type. Thus the total number of sample type in this paper is 10.

3.2. Determination of time delay and embedding dimension

Theoretically time delay and embedding dimension value could be obtained for every kind of signal [14]. But it is not practical in engineering application. So it is necessary to select a unique and suitable time delay and embedding dimension for all samples. First variation of time delay for different kinds of vibration signal is analyzed. Four kinds of signals which correspond to normal bearing status, ball fault of 0.007 in, inner race fault of 0.007 in, and outer race fault of 0.007 in are selected, respectively. The maximum time delay is set to 32 and mutual information of each kind of signal is shown in Figs. 2(a–d).

Except for Fig. 2(a) in which the first minimum mutual information appears when delay is equal to four, the first minimum mutual information in Figs. 2(b–d) appears at five. Moreover it shows that the time delay τ in

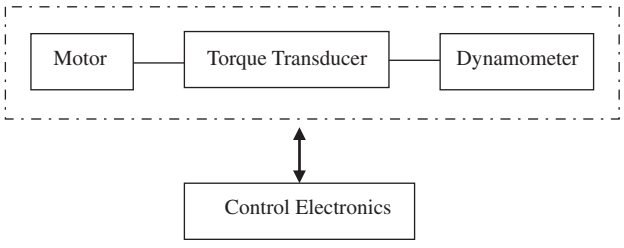

Fig. 1. Scheme of test rig for rolling bearing.

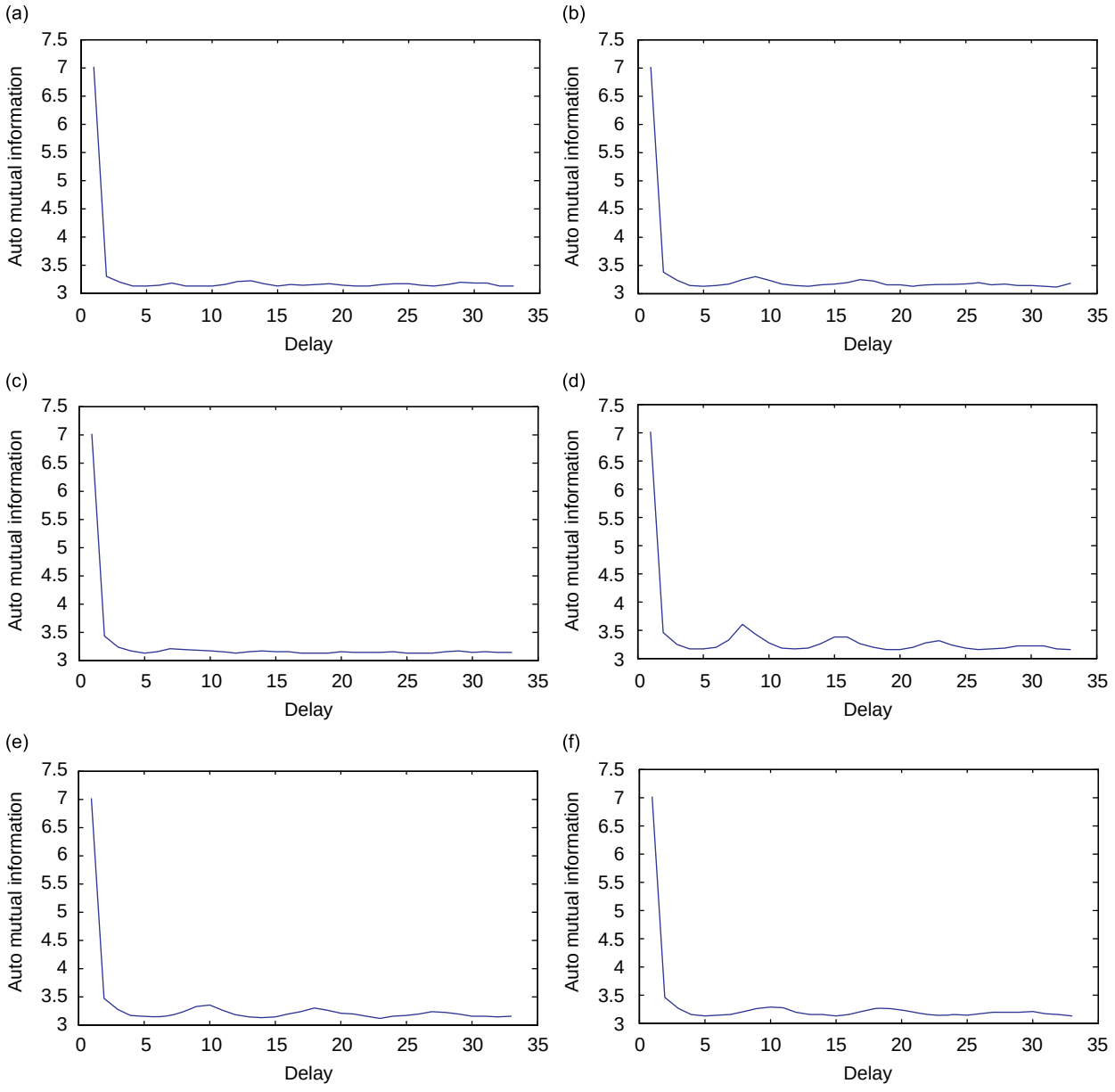

Fig. 2. Mutual information of different kinds of bearing signal: (a) normal status; (b) ball fault with diameter of 0.007 in; (c) inner race fault with diameter of 0.007 in; (d) outer race fault with diameter of 0.007 in; (e) inner race fault with diameter of 0.014 in; and (f) inner race fault with diameter of 0.021 in.

Fig. 2(a) is very close to Figs. 2(b–d). Thus a uniform value of time delay τ could be set to five. For further testifying the correctness, the size of inner race is changed and its mutual information is calculated for vibration signal of bearings whose fault size are 0.014 and 0.021 in so as to obtain the Figs. 2(e–f) in which that the first minimum mutual information is still equal to five. The analysis of other fault signal which is not plotted here could also draw the same conclusion. So finally the time delay τ is set to five.

The next step is to determine the embedding dimension d. The data sampled under different bearing status described above is adopted and Cao'method is taken to calculate $E_1(d)$ and $E_2(d)$ curve which are shown in Figs. 3–8 in which (a) is the $E_1(d)$ curve and (b) is $E_2(d)$ curve.

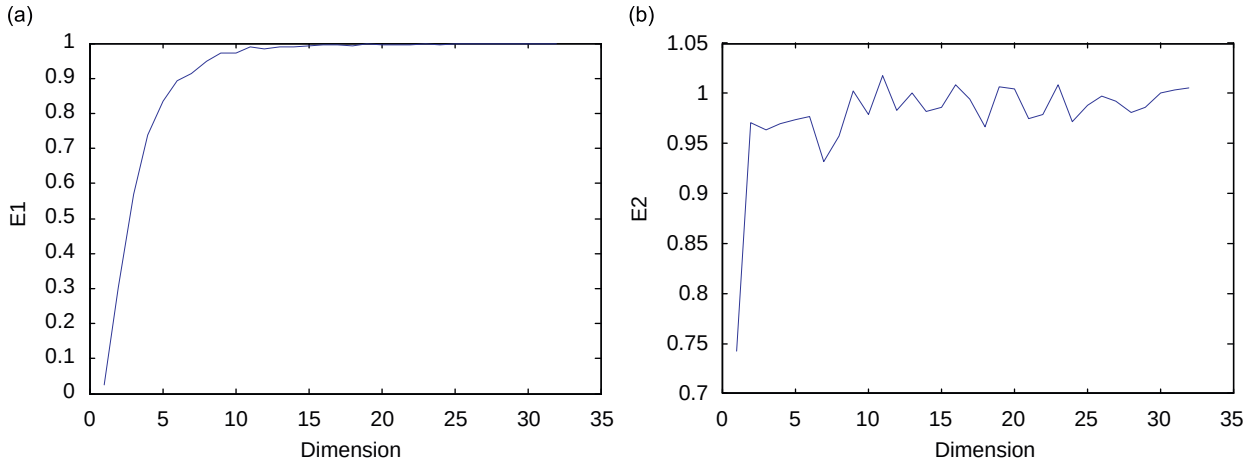

Fig. 3. $E_1(d)$ and $E_2(d)$ curve of normal bearing status.

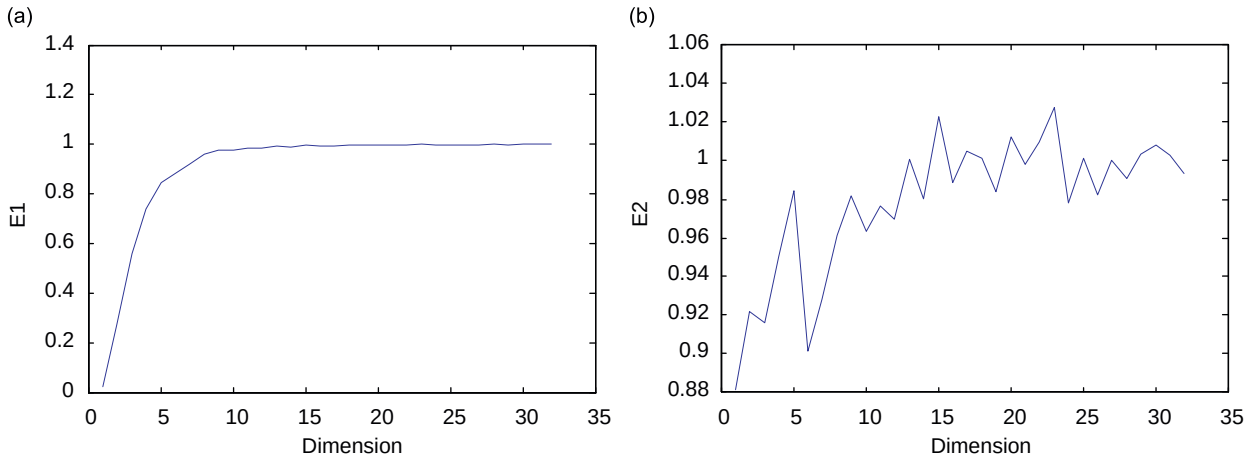

Fig. 4. $E_1(d)$ and $E_2(d)$ curve of ball fault with diameter of 0.007 in.

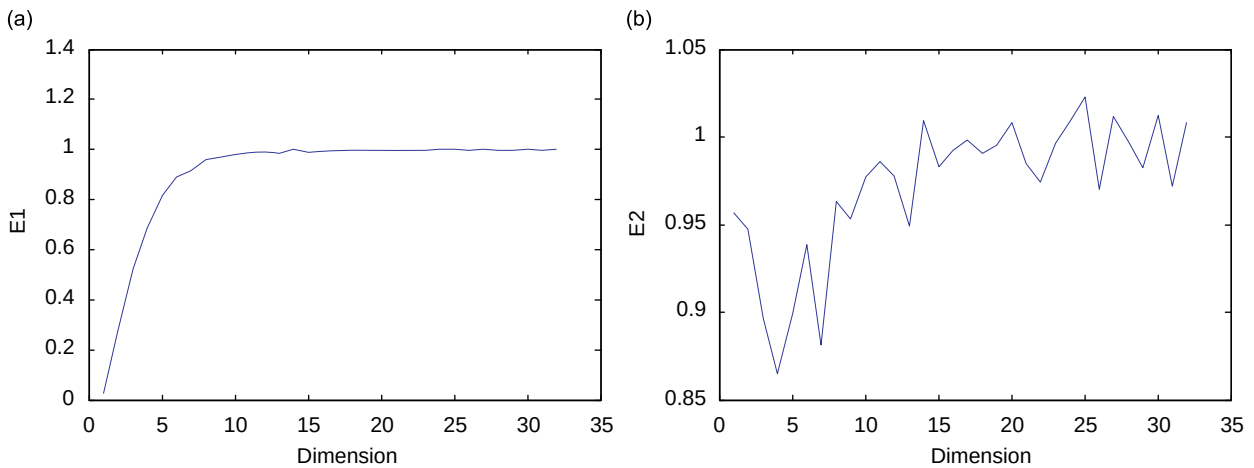

Fig. 5. $E_1(d)$ and $E_2(d)$ curve of inner race fault with diameter of 0.007 in.

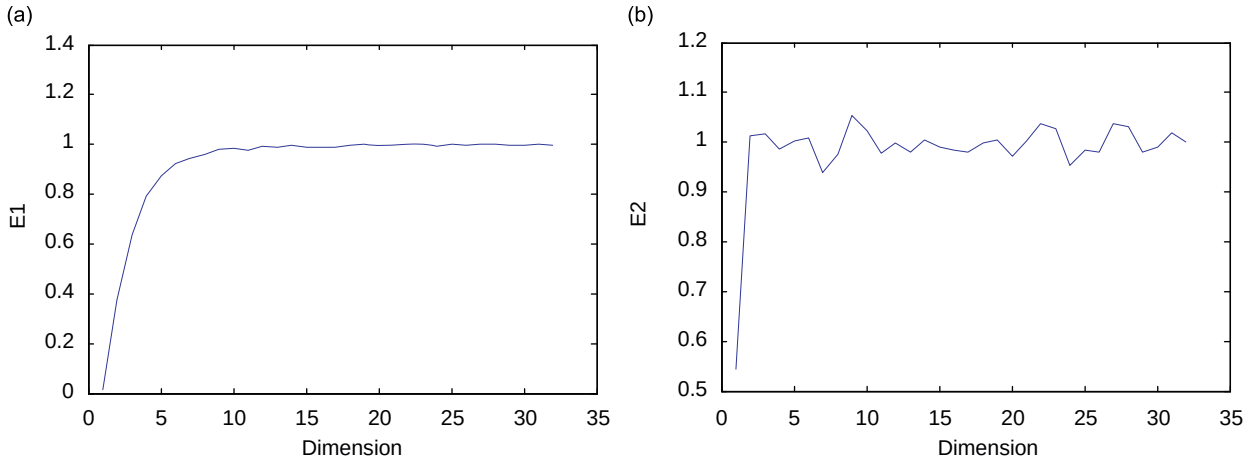

Fig. 6. $E_1(d)$ and $E_2(d)$ curve of outer race fault with diameter of 0.007 in.

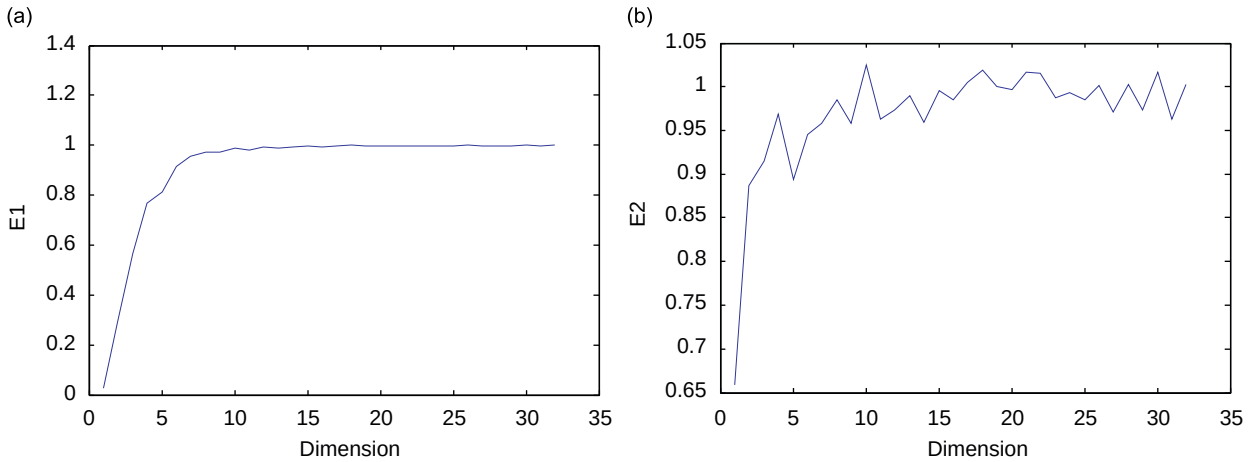

Fig. 7. $E_1(d)$ and $E_2(d)$ curve of inner race fault with diameter of 0.014 in.

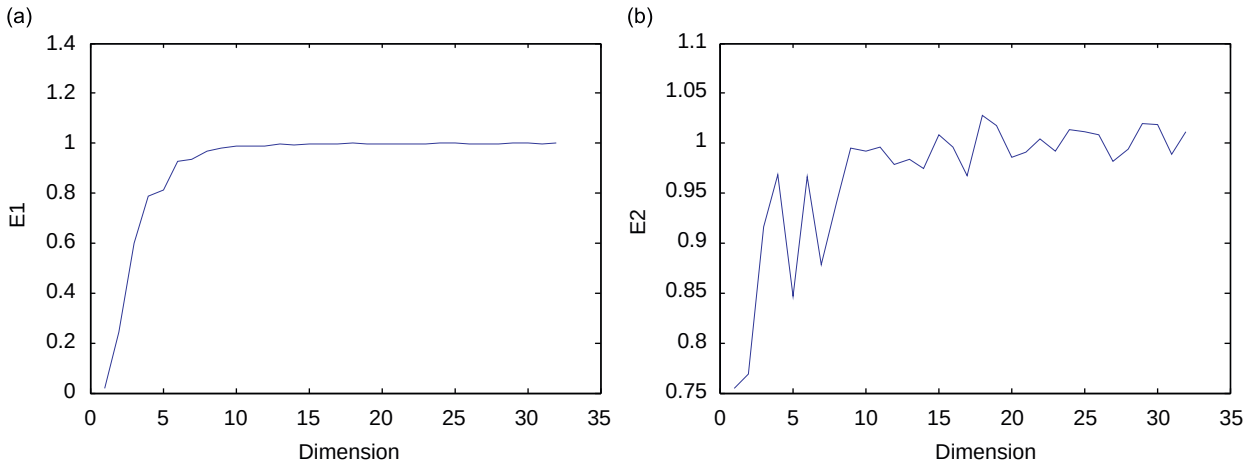

Fig. 8. $E_1(d)$ and $E_2(d)$ curve of inner race fault with diameter of 0.021 in.

From (b) in Figs. 3–8 a conclusion could be drawn that signal is deterministic because $E_2(d)$ varies with the embedding dimension d. Analysis of (a) in Figs. 3–8 show that whatever the type of signal is, the $E_1(d)$ begin to tends to stationary when d is larger than 10 so here the embedding dimension is set to 10. Thus the structure of RPS is time delay τ being equal to five and embedding dimension d being equal to 10.

The third step is to obtain the phase trajectory in RPS for every kind of fault signal. The sample size is 4096 and the phase trajectory of different kinds of fault is shown in Figs. 9(a–f). Because the embedding dimension d is 10 which is larger than three so that it is impossible to obtain a phase trajectory in a figure. Thus multidimensional phase trajectory must be projected on two-dimension phase plane. For comparison and simplification, only $x(n)$ and $x(n-\tau)$ is shown in Fig. 9 and projection on other phase plane is not plotted here, but it is still used for building the GMM described in later section. In Fig. 9, there are obvious differences in

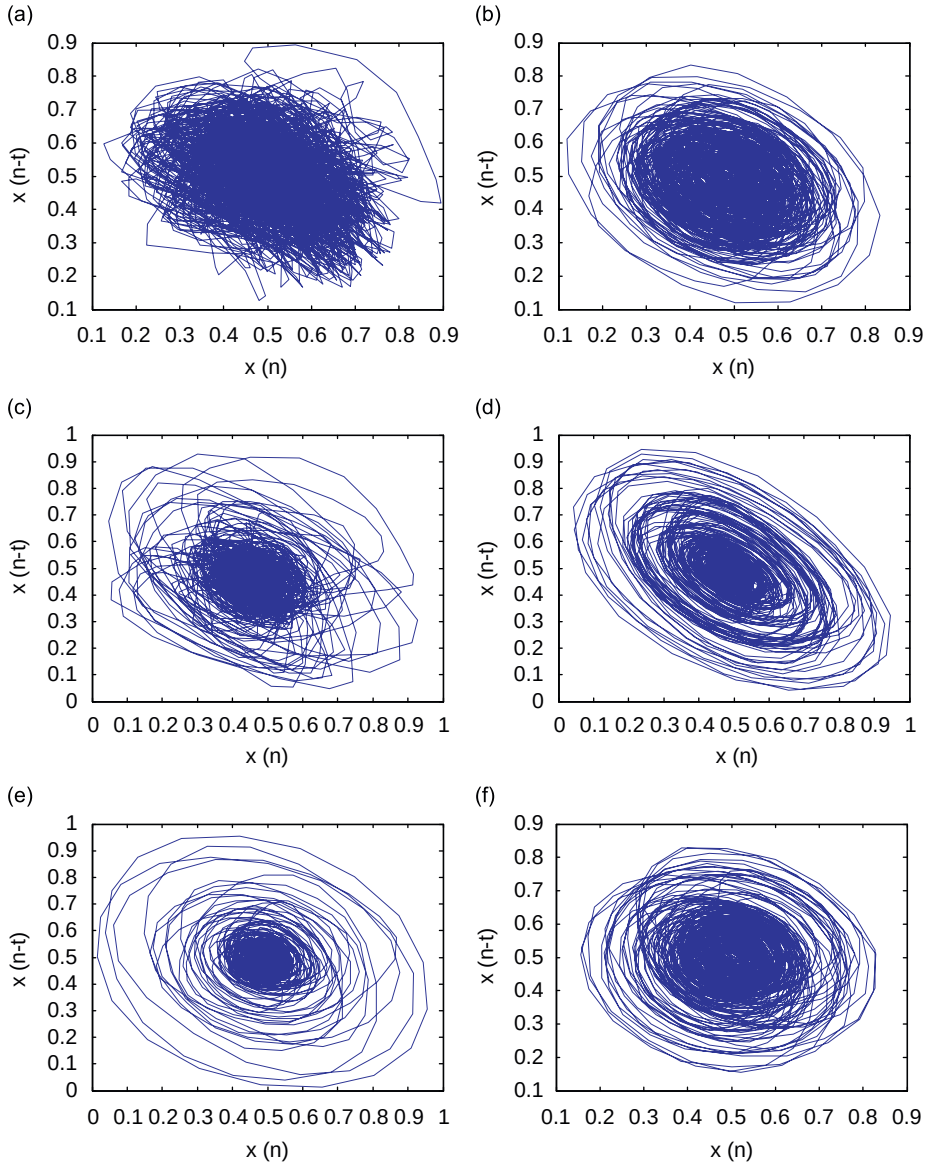

Fig. 9. Graph of phase trajectory projection on $(x(n), x(n-\tau))$ plane for different kinds of bearing signal: (a) normal status; (b) ball fault with diameter of 0.007 in; (c) inner race fault with diameter of 0.007 in; (d) outer race fault with diameter of 0.007 in; (e) inner race fault with diameter of 0.014 in; and (f) inner race fault with diameter of 0.021 in.

phase trajectory between different kinds of bearing status. As depicted in Fig. 9(a), phase trajectory of normal bearing signal demonstrate dense points whose boundary is very irregular which shows that main component of normal bearing signal is noise. However, phase trajectory in Figs. 9(b–f) show that dense orbits points are surrounded by less dense points whose boundary is very regular. So chaos phenomenon does exist for all fault signals [19]. However, the distribution, center, position, and size of phase orbit are very different even for those fault signals which have the same fault type but different fault size. That is just why this method could be used for classification of fault type and fault severity.

3.3. Selection of number of mixture model and sample sizes

After getting the phase trajectory of different kinds of fault, the next work is to build GMM for every kind of bearing signal. An important parameter need to be determined here is number of mixture model. There are many methods used to judge if the number is suitable or not such as AIC, BIC and etc. But these methods may take much more time and the result is always not accurate when noise exists. Another parameter need to be selected is sample size of vibration signal. Because these two parameters all bring influence on accurate rate of classification, the number of mixture model and sample size of each fault signal are used for classification simultaneously. Here a simple method is adopted to determine the value of these two parameters by comparing the accuracy under different parameters. For comparison between different parameters, the same sample number for each kind of fault is selected. Here, sample number used for learning of each GMM model is set to 500 and the sample size used for classification and justification is set to 384.

Here four kinds of fault signal are taken as example. The bearing status of signal are normal bearing status, inner race fault with its fault size equal to 0.007, 0.014 and 0.021 in. The number of GMM is 2, 4, 8, 16, and 32, respectively. The accurate rate curve of every kind of signal with the variation of number of mixture model under sample size being 512 is shown in Fig. 10 in which the accurate rate is almost not influenced by number of mixture model when the number of GMM is larger than four, even the number of sample is reduced to 64. As shown in Fig. 11, increase of number of mixture model also has less influence on accurate rate. So number of mixture model utilized here is four. The variation of accurate rate with the sample size is given in Fig. 12 in which index 1–7 represent the sample size is 128, 256, 512, 1024, 2048, 4096, and 8192, respectively. From Fig. 12, a conclusion could be drawn that the accurate rate increases with the increase of sample size correspondingly. But when sample size is larger than 1024, the increase of accurate rate is not obvious. Considering that it may result in the increase of calculation load, so the size of sample used for classification is set to 1024.

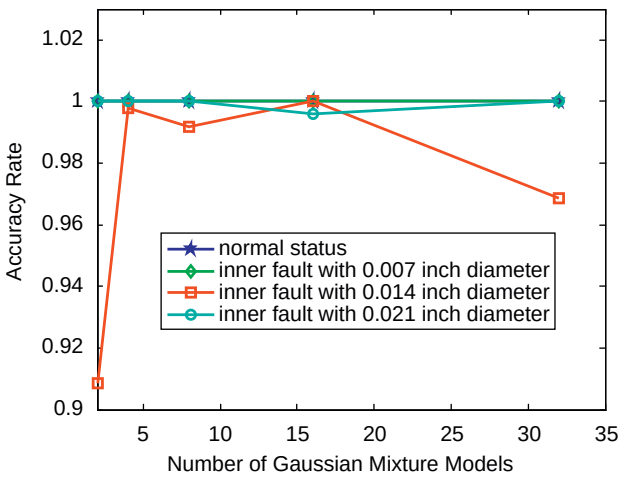

Fig. 10. Variation of accurate rate with increase of number of GMM under sample of 512 points.

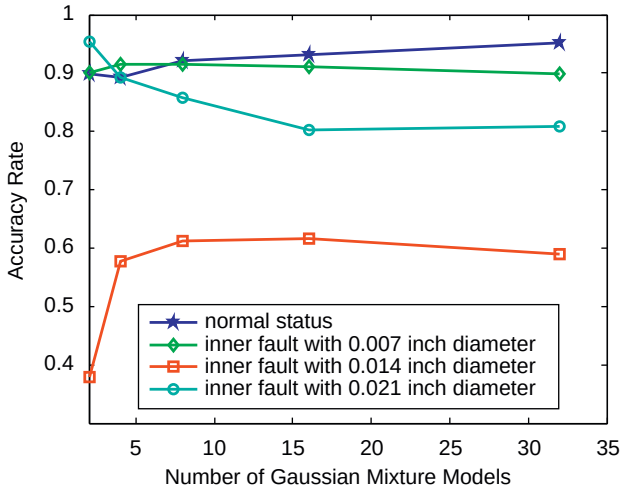

Fig. 11. Variation of accurate rate with increase of number of GMM under sample of 64 points.

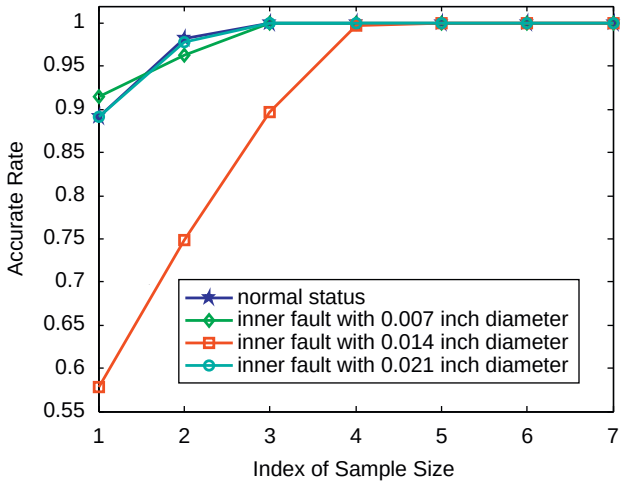

Fig. 12. Variation of accurate rate with increase of sample size under GMM number is four.

Table 1
Hybrid fault classification under different fault type and severity for signal of 1024 points.

Fault	Normal status	0.007 in			0.014 in			0.021 in		
		Ball fault	Inner race fault	Outer race fault	Ball fault	Inner race fault	Outer race fault	Ball fault	Inner race fault	Outer race fault
Accurate rate (%)	100	95.55	100	100	100	81.74	96.44	100	100	85.5

3.4. Classification of fault types and its severity

After these parameters are calculated from time series directly, the whole classification model is built and the results shown in Fig. 12 prove that classification is very accurate. For further testifying the correctness of this

Table 2
Hybrid fault classification under different fault type and severity for signal of 2048 points.

Fault	Normal status	0.007 in			0.014 in			0.021 in		
		Ball fault	Inner race fault	Outer race fault	Ball fault	Inner race fault	Outer race fault	Ball fault	Inner race fault	Outer race fault
Accurate rate (%)	100	100	100	100	100	99.48	100	100	100	100

method, more complex situation is considered. The samples used for classification include different kinds of fault signal under different fault size. Time delay is set to five, embedding dimension is set to 10, number of mixture model is set to four and sample size of vibration signal is set to 1024 and 2048, respectively, the final accurate rate of classifying these 10 kinds of bearing status is given in Tables 1 and 2 in which the accurate rate is very high if the sample size is equal to 1024. Moreover if the sample size is 2048, the accurate rate could almost reach to 100 percent which prove that this method could be used for classification of fault type and fault size of rolling bearings simultaneously.

4. Conclusions

To realize accurate detection and classification of bearing faults, a novel method which integrates RPS and GMM is presented in this paper. By appropriate reconstructing of vibration signal in phase space, the phase trajectory of different kinds of fault signal could be obtained. From these phase trajectory, characteristic of noise is found for vibration signal in normal status and chaos is found for fault signals. For different bearing signal, the distribution, shape, and position of phase trajectory are different obviously. So this method could be adopted for distinguishing between different bearing statuses. To realize the classification of bearing faults, GMM is adopted to realize correct description of phase trajectory distribution in RPS. Comparing between different parameters show that, when number of GMM is larger than four, its variation has slight influence on accurate rate of classification. And the sample size has much influence on result of classification. Therefore, correct selection of sample size could increase the accurate rate greatly. The main advantage of this method is that all parameters, such as time delay, embedding dimension, number of mixture model, and sample size are obtained only by analyzing bearing vibration signal directly so that they are not needed to be determined empirically. Therefore, the classification is much more accurate and it is suitable for industry application. Of course, signals adopted here are only single point fault, so the work in the next step could be focused on classification for multiple points bearing fault. Moreover, the dynamic characteristic of rolling bearing should be studied further to interpret why such phase trajectory is produced which is also a challenging job.

Acknowledgement

The authors would like to thank Case Western Reserve University for their providing free download of the rolling element bearing fault data sets. This project is supported by Chinese National Science Fund (50805100).

References

[1] J. Altmann, J. Mathew, Multiple band-pass autoregressive demodulation for rolling element bearing fault diagnosis, *Mechanical Systems and Signal Processing* 15 (2001) 963–977.
[2] C. Junsheng, Yu. Dejie, Y. Yu, The application of energy operator demodulation approach based on EMD in machinery fault diagnosis, *Mechanical Systems and Signal Processing* 21 (2007) 668–677.
[3] H. Ocak, K.A. Loparo, F.M. Discenzo, Online tracking of bearing wear using wavelet packet decomposition and probabilistic modeling: A method for bearing prognostics, *Journal of Sound and Vibration* 302 (2007) 951–961.

[4] C. Junsheng, Yu. Dejie, et al., Application of an impulse response wavelet to fault diagnosis of rolling bearings, *Mechanical Systems and Signal Processing* 21 (2007) 920–929.
[5] V. Purushotham, S. Narayanan, et al., Multi-fault diagnosis of rolling bearing elements using wavelet analysis and hidden Markov model based fault recognition, *NDT & E International* 38 (2005) 654–664.
[6] J. Yang, Y. Zhang, Y. Zhu, Intelligent fault diagnosis of rolling element bearing based on SVMs and fractal dimension, *Mechanical Systems and Signal Processing* 21 (2007) 2012–2024.
[7] K. Jingqiu, L. Yibing, M. Zhiyong, Y. Keguo, Improved algorithm of correlation dimension estimation and its application in fault diagnosis for industrial fan, *Proceedings of the 25th Chinese Control Conference*, Harbin, China, 2006, pp. 1291–1296.
[8] R. Yan, R.X. Gao, Complexity as a measure for machine fault detection and diagnosis, *Proceeding of IMTC 2003—Instrumentation and Measurement Technology Conference*, Vail, CO, USA, 2003, pp. 65–70.
[9] R. Yan, R.X. Gao, Approximate entropy as a diagnostic tool for machine health monitoring, *Mechanical Systems and Signal Processing* 21 (2007) 824–839.
[10] T. Marwala, U. Mahola, F.V. Nelwamondo, Hidden Markov models and Gaussian mixture models for bearing fault detection using fractals, *International Joint Conference on Neural Networks*, Vancouver, BC, Canada, 2006, pp. 3237–3242.
[11] A.K.S. Jardine, D. Lin, D. Banjevic, A review on machinery diagnostics and prognostics implementing condition-based maintenance, *Mechanical Systems and Signal Processing* 20 (2006) 1483–1510.
[12] F.V. Nelwamondo, T. Marwala, U. Mahola, Early classifications of bearing faults using hidden Markov models, Gaussian mixture models, mel-frequency cepstral coefficients and fractals, *International Journal of Innovative Computing, Information and Control ICIC International* 6 (2006) 1281–1299.
[13] R.J. Povinelli, M.T. Johnson, Time series classification using Gaussian mixture models of reconstructed Phase Spaces, *IEEE Transactions on Knowledge and Data Engineering* 6 (2004) 779–783.
[14] L. Cao, Practical method for determining the minimum embedding dimension of a scalar time series, *Physics D* 110 (1997) 43–50.
[15] J. Kadtke, Classification of highly noisy signals using global dynamical models, *Physics Letters A* 203 (1995) 196–202.
[16] L.F.P. Franca, M.A. Savi, Distinguishing periodic and chaotic time series obtained from an experimental nonlinear pendulum, *Nonlinear Dynamics* 26 (2001) 253–271.
[17] Y. Lee, K.Y. Lee, J. Lee, The estimating optimal number of Gaussian mixtures based on incremental *k*-means for speaker identification, *International Journal of Information Technology* 12 (2006) 13–21.
[18] 〈http://www.eecs.cwru.edu/laboratory/bearing〉, Bearing Data Center Website, Case Western Reserve University.
[19] S.P. Harsha, Nonlinear dynamic analysis of a high-speed rotor supported by rolling element bearings, *Journal of Sound and Vibration* 290 (2006) 65–100.