

Degradation assessment and fault diagnosis for roller bearing based on AR model and fuzzy cluster analysis

Lingli Jiang^{a,b,*}, Yilun Liu^a, Xuejun Li^b and Anhua Chen^b

^a*College of Mechanical and Electrical Engineering, Central South University, Changsha 410083, China*

^b*Hunan Provincial Key Laboratory of Health Maintenance for Mechanical Equipment, Hunan University of Science and Technology, Xiangtan 411201, China*

Received 5 February 2010

Revised 5 May 2010

Abstract. This paper proposes a new approach combining autoregressive (AR) model and fuzzy cluster analysis for bearing fault diagnosis and degradation assessment. AR model is an effective approach to extract the fault feature, and is generally applied to stationary signals. However, the fault vibration signals of a roller bearing are non-stationary and non-Gaussian. Aiming at this problem, the set of parameters of the AR model is estimated based on higher-order cumulants. Consequently, the AR parameters are taken as the feature vectors, and fuzzy cluster analysis is applied to perform classification and pattern recognition. Experiments analysis results show that the proposed method can be used to identify various types and severities of fault bearings. This study is significant for non-stationary and non-Gaussian signal analysis, fault diagnosis and degradation assessment.

Keywords: Degradation assessment, Fault diagnosis, AR mode, Higher-order statistics, Fuzzy cluster analysis

1. Introduction

In modern times, rotating machinery covers a broad range of mechanical equipment and plays an important role in industrial application. Bearing is not only the most important but also a common failure unit in rotating machinery. Fault diagnosis and degradation assessment for bearings are critical to maintain the normal operation of equipment, whose purpose is to analyze the relevant external information to judge the condition of the inaccessible internal components so as to decide if the bearing needs to be replaced or not [1,4]. Vibration analysis-based method is the most principal and effective method for bearing monitoring, and consists of two most important aspects (1) feature extraction (2) condition identification and fault diagnosis.

Effective feature extraction for fault diagnosis is still an ongoing research issue. The common techniques perform feature extraction from the waveforms in the time, frequency or time-frequency domain, and the extracted features include statistical quantities (e.g., mean value, root mean square value, skewness, kurtosis, crest factor, etc.) [8,22], energy or entropy of different frequency bands obtained via wavelet/wavelet packet decomposition or empirical mode decomposition [15,16,19,20], fractal dimension [9], and so on. These waveform-based feature extraction methods always need to combine with failure mechanism to implement fault diagnosis. However, failure mechanism is a complicated process. At the same time, these features are sensitive to different faults and degradation stages, but none of them consistent enough to be used as a sole indicator of spall size.

*Corresponding author: Lingli Jiang, Tel.: +86 15573216060; Fax: +86 0732 5829480; E-mail: linlyjiang@163.com.

Autoregressive (AR) model is a time sequence analysis method whose parameters comprise important information of the system condition. An accurate AR model can reflect the characteristics of the dynamic system. Additionally, it is indicated that the AR parameters of AR model are very sensitive to the condition variation [5,21]. In many cases, the AR parameters are estimated by using the least squares method or the Yule-Walker equation method based on correlation function for the stationary signals. When the AR model is applied to the non-stationary signals, its parameters are difficult to estimate by these approaches. However, bearing fault vibration signals composed of a series of impulse signals and intermixed various noise are typical non-stationary and non-Gaussian signals. Aiming at this problem, Ref. [11] proposed an approach combining empirical mode decomposition (EMD) method and AR model. The EMD method was used as a pretreatment to decompose the non-stationary vibration signal of a roller bearing into a number of intrinsic mode function (IMF) components which are stationary, then the AR model of each IMF component was established, and the means of AR parameters of the IMF components were regarded as the feature vectors. This process is effective but complicated. This paper presents a simpler and effectiver approach that the AR model of the raw bearing vibration signal is established directly and the AR parameters are estimated based on higher-order cumulant. Higher-order cumulant is easy to extract information due to deviation from Gaussianity, to recover the true phase character of the signals, and to detect and quantifying nonlinearities in the time series [7, 14]. These characters make it well suitable to analyze non-stationary and non-Gaussian signals. The AR parameters estimation based on higher-order cumulants are used as feature vectors to reflect different fault types and degradation phases of roller bearings in this paper.

Condition identification and fault diagnosis can be considered as a pattern recognition problem. Fuzzy cluster analysis is one of the major techniques in pattern recognition, which discovers the substructure of a date set by dividing it into several clusters under unsupervised learning [3]. Fuzzy cluster analysis based on fuzzy equivalence relation is applied to perform classification and pattern recognition in this paper.

This paper is organized as follows. In Section 2, the feature extraction approach – the AR parameters estimation based on higher-order cumulants is presented. Fuzzy cluster analysis based on fuzzy equivalence relation is introduced for classification and pattern recognition in Section 3. In Section 4, the experiments and the obtained date sets are introduced. In Section 5, the vibration signals of bearing on the date sets are analyzed and the classification performance based on the proposed approach are discussed. Conclusions of this paper are given in Section 6.

2. AR parameter estimation based on higher-order cumulant

2.1. Moments and cumulant

Dealing with non-Gaussian random processes, the notions of higher-order moments, cumulants, and their polyspectra called higher order statistics are of paramount importance in signal processing [17].

Let $x(k)$ be a real, discrete time, and n th-order stationary random process. Moreover, let $\varpi = [\varpi_1, \varpi_2, \dots, \varpi_n]^T$, $X = [x(k), x(k + \tau_1), \dots, x(k + \tau_{n-1})]^T$. Then the n th-order moment $m_{nx}(\tau_1, \tau_2, \dots, \tau_{n-1})$ is defined as the coefficient in the Taylor expansion of the moment generating function:

$$\Phi_x(\varpi) = E[\exp(j\varpi^T X)] \quad (1)$$

In practice, the n th-order moment $m_{nx}(\tau_1, \tau_2, \dots, \tau_{n-1})$ can be equivalently calculated by taking an expectation over the process multiplied by $(n-1)$ lagged version of itself.

$$\begin{aligned} m_{1x} &= E[x(k)] \\ m_{2x}(\tau) &= E[x(k)x(k + \tau)] \\ m_{3x}(\tau_1, \tau_2) &= E[x(k)x(k + \tau_1)x(k + \tau_2)] \\ &\dots \end{aligned} \quad (2)$$

Similarly, the n th-order cumulant $c_{nx}(\tau_1, \tau_2, \dots, \tau_{n-1})$ is defined as the coefficients in the Taylor expansion of the cumulant generating function, also known as the second characteristic function.

$$\Psi_x(\varpi) = \ln \Phi_x(\varpi) = \ln E[\exp(j\varpi^T X)] \quad (3)$$

Combining Eqs (1) and (3), it is obvious that the cumulants can be expressed in terms of the moments and vice versa. The 1th-, 2th-, 3th- and 4th-order cumulant of $x(k)$ can be written as follows:

$$\begin{aligned} c_{1x} &= m_{1x} \\ c_{2x}(\tau) &= m_{2x}(\tau) - m_{1x}^2 \\ c_{3x}(\tau_1, \tau_2) &= m_{3x}(\tau_1, \tau_2) - m_{1x}[m_{2x}(\tau_1) + m_{2x}(\tau_2) + m_{2x}(\tau_2 - \tau_1)] + 2m_{1x}^3 \\ c_{4x}(\tau_1, \tau_2, \tau_3) &= m_{4x}(\tau_1, \tau_2, \tau_3) - m_{2x}(\tau_1)m_{2x}(\tau_3 - \tau_2) - m_{2x}(\tau_2)m_{2x}(\tau_3 - \tau_1) - m_{2x}(\tau_3)m_{2x}(\tau_2 - \tau_1) \\ &\quad - m_{1x}(\tau_3)m_{2x}(\tau_2 - \tau_1) - m_{1x}[m_{3x}(\tau_2 - \tau_1, \tau_3 - \tau_1) + m_{3x}(\tau_2, \tau_3) + m_{3x}(\tau_2, \tau_4) + m_{3x}(\tau_1, \tau_2)] \\ &\quad + (m_{1x})^2[m_{2x}(\tau_1) + m_{2x}(\tau_2) + m_{2x}(\tau_3) + m_{2x}(\tau_3 - \tau_1) + m_{2x}(\tau_3 - \tau_2) + m_{2x}(\tau_2 - \tau_1)] - 6(m_{1x})^4 \\ &\dots \end{aligned} \quad (4)$$

Higher-than-2th-order moment and cumulant is defined as higher-order moment and cumulant, respectively. With zero-mean assumption, the 2th- and 3th- cumulant are the same as the 2th- and 3th- order moment. The higher-order cumulants are blind to any kind of Gaussian process.

Theorem[2]: Let $x(t) = y(t) + s(t) + v(t)$, where $y(t)$ is a zero-mean stationary non-Gaussian process, $s(t)$ is a deterministic signal, and $v(t)$ is a zero-mean process independent of $x(t)$ with n th-order cumulant identically zero for some $k > 2$ (e.g., $v(t)$ is Gaussian). Then, for $k > 2$

$$c_{kx}(\tau_1, \tau_2, \dots, \tau_{n-1}) = c_{ky}(\tau_1, \tau_2, \dots, \tau_{n-1}) \quad (5)$$

This theorem is only applied to higher-order cumulants. It is obvious that the higher-order cumulant of the signal $x(t)$ is determined by the non-Gaussian process, and can avoid the impact of Gaussian and constant.

In fact, the vibration signal of a roller bearing is composed of Gaussian and non-Gaussian process and constant. The higher-order cumulant provides a novel method to analyze this type of signals.

2.2. AR model and AR parameters estimation

AR(p) model with time-domain vibration signals for fault diagnosis can be depicted as a difference equation [12]

$$\sum_{k=0}^p a_k x(t-k) = e_i(t) \quad (6)$$

where, $a_k (k = 0, 1, \dots, p)$ is the AR parameter, $a_0 = 1$, p is the AR order, and $e_i(t)$ is the remnant of the AR model. The AR parameters can be determined by using cumulants. Any 1-D cumulant slice satisfies the following AR recursion of maximum order p .

$$\sum_{k=0}^p a_k C_{mx}(\tau - k, k_0, \dots, 0) = 0, \text{ for } \tau > 0 \quad (7)$$

where k_0 is a parameter whose choice is discussed hereinafter.

Concatenating Eq. (7) for $\tau = 1, 2, \dots, p+1+M$, where $M \geq 0$, and k_0 is arbitrary, we obtain the cumulant-based normal equations

$$C(k_0)A = 0 \quad (8)$$

where, $A = [1, a_1, a_2, \dots, a_p]^T$, and the entries in Toeplitz matrix $C(k_0)$ are easily deduced as follows:

$$C(k_0) = \begin{pmatrix} C_{mx}(1-0, k_0, 0, \dots, 0) & C_{mx}(1-1, k_0, 0, \dots, 0) \cdots & C_{mx}(1-p, k_0, 0, \dots, 0) \\ C_{mx}(2-0, k_0, 0, \dots, 0) & C_{mx}(2-1, k_0, 0, \dots, 0) \cdots & C_{mx}(2-p, k_0, 0, \dots, 0) \\ \vdots & \vdots & \vdots \\ C_{mx}(p+1+M-0, k_0, 0, \dots, 0) & \cdots & \cdots C_{mx}(p+1+M-p, k_0, 0, \dots, 0) \end{pmatrix}$$

The parameter k_0 should be determined to make $C(k_0)$ has rank p . Thus, the corresponding 1-D slice of the m th-order cumulant is a full rank slice, and the AR(p) parameters can be solved for from Eq. (8).

3. Fuzzy cluster analyses for pattern recognition

Cluster analysis divides data into groups or clusters such that similar data objects belong to the same cluster and dissimilar data objects to different clusters. The resulting data partition improves data understanding and reveals its internal structure. In real applications there is very often no sharp boundary between clusters so that fuzzy clustering is often better suited for the data. Membership degrees between zero and one are used in fuzzy clustering instead of crisp assignments of the data to clusters [18]. Introducing fuzziness to clustering gives us the flexible representations of substructures of the data set.

The most common fuzzy clustering categories are fuzzy clustering based on fuzzy equivalence relation, fuzzy clustering based on objective functions, and the fuzzy generalized k-nearest neighbor rule [6,10,13]. This paper uses fuzzy clustering based on fuzzy equivalence relation for pattern recognition.

3.1. Construction of the fuzzy relation matrix

When fuzzy clustering used as a classifier for pattern recognition, determining the features is an important issue to obtain better clustering performance. The AR parameters solved for from Section 2 are used as features for the following clustering analysis.

Suppose that a sample set contains n samples marked as $\{s_1, s_2, \dots, s_n\}$, where $s_i = \{a_{i0}, a_{i1}, \dots, a_{ip}\}$ is a feature vector, a_{ij} is the j th AR(p) parameter of individual s_i . In order to make the changes of the samples consistent, the features need to be equalized and normalized, i.e. according to the following formula.

$$a'_{ij} = |a_{ij}| / (\text{Max}_{j=1}^p |a_{ij}|) (i = 1, 2, \dots, n; j = 0, 1, \dots, p) \tag{9}$$

The fuzzy relation matrix can be represents as fowllows:

$$R = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix} = \begin{pmatrix} a'_{10} & a'_{11} & \dots & a'_{1p} \\ a'_{20} & a'_{21} & \dots & a'_{2p} \\ \vdots & \vdots & \dots & \vdots \\ a'_{n0} & a'_{n1} & \dots & a'_{np} \end{pmatrix} \tag{10}$$

3.2. Construction of the fuzzy similar matrix

The fuzzy relation matrix R should be further transformed to the fuzzy similar matrix marked as $\overset{R}{\sim}$. A calculation procedure should be done to determine the fuzzy similarity relation between sample $s_i (i = 1, 2, \dots, n)$ and $s_j (j = 1, 2, \dots, n)$, represented as the similar coefficient r_{ij} . r_{ij} can be calculated by many methods, such as the distance method, the intersecting angle cosine method, the maximum or minimum method, the correlation coefficient method, and so on. By considering the easy implementation and intuition, the Euclidean distance method is used here.

$$r_{ij} = 1 - \sqrt{\frac{1}{p} \sum_{k=0}^p (a'_{ik} - a'_{jk})^2} (i, j = 1, 2, \dots, n) \tag{11}$$

Therefore, the obtained matrix $\overset{R}{\sim} = (r_{ij})_{n \times n}$ can be defined as the fuzzy similar matrix.

3.3. Construction of the fuzzy equivalent matrix

The fuzzy similar matrix can not be used for clustering unless it has a fuzzy equivalence relation which should follow three conditions:

- (1) Reflexive, i.e., $r_{ii} = 1$;
- (2) Symmetric, i.e., $r_{ij} = r_{ji}$;
- (3) Transitive, i.e., $\overset{R}{\sim} \circ \overset{R}{\sim} \subset \overset{R}{\sim}$, where “ \circ ” is a fuzzy operational sign.

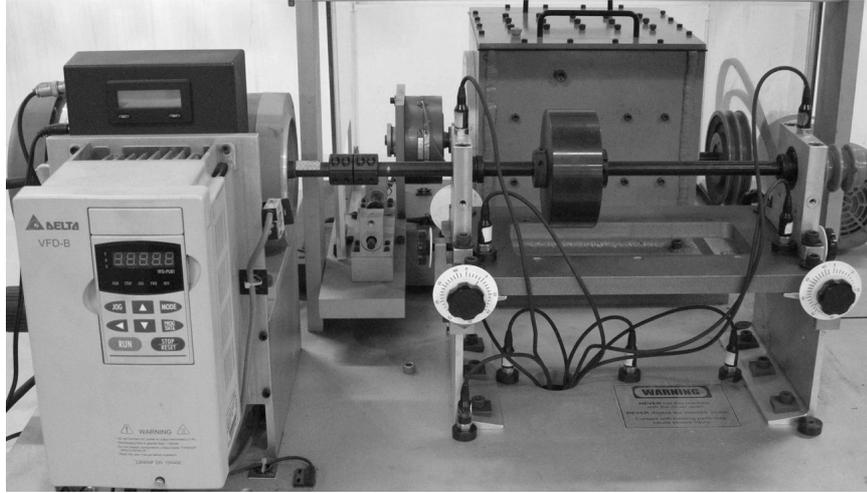


Fig. 1. The machinery fault simulator.

Obviously the fuzzy similar matrix is reflexive and symmetric but not transitive. The transitive closure method is used to obtain a fuzzy equivalent matrix, described as follows:

$$\tilde{R}^2 = \tilde{R} \circ \tilde{R} = \left(\bigvee_{w=1}^n (r_{uw} \wedge r_{wv}) \right)_{n \times n} \tag{12}$$

where, “ \vee ” and “ \wedge ” represent “taking larger” and “taking smaller” operation, respectively.

Perform the same operation $\tilde{R}^4 = \tilde{R}^2 \circ \tilde{R}^2$, $\tilde{R}^8 = \tilde{R}^4 \circ \tilde{R}^4$, ..., until $\tilde{R}^{2k} = \tilde{R}^k$, then the fuzzy equivalent matrix $\tilde{R}^* = \tilde{R}^k$ is obtained.

3.4. λ -cut matrix for clustering

Suppose \tilde{R}^* is the fuzzy equivalent matrix, the λ -cut matrix \tilde{R}^*_λ is denoted as follow:

$$\tilde{R}^*_\lambda = \{r^*_{ij} = 1, if r^*_{ij} \geq \lambda; r^*_{ij} = 0, if r^*_{ij} < \lambda; \lambda \in [0, 1]\} \tag{13}$$

We can get different matrix \tilde{R}^*_λ related to different cutting level λ . If $r^*_{\lambda ij} = 1$, it means that the i th and j th sample belong to the same class. The algorithm is obviously different from the conventional pattern recognition as the classes of clustering are unknown in advance.

4. Experiment and date sets

Experiments were performed on the machinery fault simulator (MFS) from SpectraQuest, Inc. shown in Fig. 1. It can simulate most of faults that commonly occur in rotating machinery, such as misalignment, unbalance, resonance, roller bearing faults, gearbox faults, and so on. The simulator has a range of operating speeds up to 6000 rpm. In this work the simulator is constituted by a motor, a coupling, a testing roller bearing fitted on the left of the shaft near the motor, a working roller bearing on the other side, a bearing load and a shaft. The MFS provides a bearing fault kit including one inner race defect, one outer race defect, one with ball defect, and one combination of defects for performing experiments and studying bearing fault diagnosis. In order to have a detailed investigation and advance studying, single point faults were introduced to the test bearings by using the electro-discharge machining with defect diameters of 7, 14, 28 mils (1mil = 0.001inchs), and the fault depth is 11, 11, 50 mils, respectively, from slight defect to serious defect.

Table 1
Data sets statistics

Data set		Fault type				Fault diameter(mil)			
DAll	DAll_07	N	I	B	O	0	07	07	07
	DAll_14	N	I	B	O	0	14	14	14
	DAll_28	N	I	B	O	0	28	28	28
DSingle	DInner	N	I	I	I				
	DBall	N	B	B	B	0	07	14	28
	DOuter	N	O	O	O				

DAll – the category of different fault types;
 DSingle – the category of different fault severities;
 N – normal; B – ball fault; I – inner race fault; O – outer race fault.

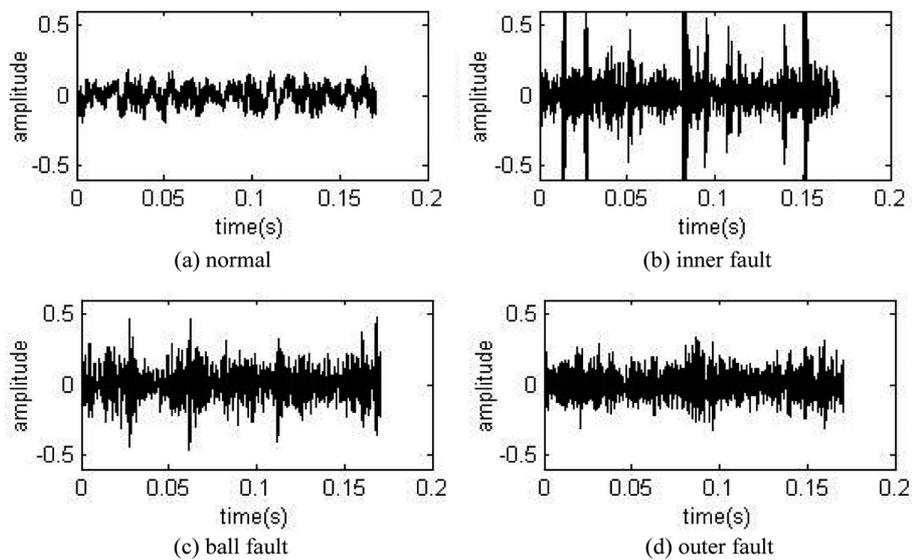


Fig. 2. The raw vibration signals of bearing with different fault type.

The shaft rotating speed was about 30 Hz obtained by a laser speedometer. Acceleration signals were measured by the Dewetron 16 channel data acquisition system and IMI 603C01 accelerometers with 12 kHz acquisition frequency rate. The data was stored in .mat format for further Matlab operation. The vibration signals of the accelerometer located on the top of the left test bearing housing are taken for further analysis.

The experiment data sets of 2 main categories including total of 6 fault data sets are presented in Table 1. For the category of different fault types DAll, each data set includes normal and three different faulty bearings. For the category of different fault severities DSingle, each data set includes a certain fault type bearing with various severity of defect. Each sample in the data sets includes 2048 points. Samples of the raw vibration signals are shown in Fig. 2.

5. Analysis and discussion

The process of the proposed approach combining AR model and fuzzy cluster analysis is shown as Fig. 3. Firstly, the AR mode is established and the AR parameters as the feature vectors are used to construct the fuzzy relation matrix. Then, after a series of deduction and calculation, the fuzzy similar matrix and the fuzzy equivalent matrix are constructed one after another. Finally, by choosing proper λ , the λ -cut matrix is obtained for clustering and pattern recognition.

Table 2
The AR parameters of the data set DA11_07

Sample		a_0	a_1	a_2	a_3	a_4	a_5
s_1	N	1.0000	-2.0548	1.8154	-0.5866	-0.2943	0.2515
s_2	N	1.0000	-1.9376	1.5669	-0.3458	-0.4054	0.2681
s_3	N	1.0000	-2.0438	1.8013	-0.5976	-0.2777	0.2378
s_4	I07	1.0000	-0.7876	1.2450	-0.9504	0.6628	-0.2845
s_5	I07	1.0000	-0.8133	1.2630	-0.9805	0.6782	-0.2903
s_6	I07	1.0000	-0.7843	1.2538	-0.9652	0.6724	-0.3039
s_7	B07	1.0000	-0.7838	1.3212	-1.2079	0.5723	-0.4055
s_8	B07	1.0000	-0.7925	1.3604	-1.2041	0.6074	-0.4003
s_9	B07	1.0000	-0.8234	1.3815	-1.2719	0.6276	-0.4510
s_{10}	O07	1.0000	-0.4966	1.1571	-0.9398	0.4069	-0.3281
s_{11}	O07	1.0000	-0.4354	1.1474	-0.9230	0.3678	-0.3681
s_{12}	O07	1.0000	-0.3976	1.1026	-0.8950	0.3169	-0.3426

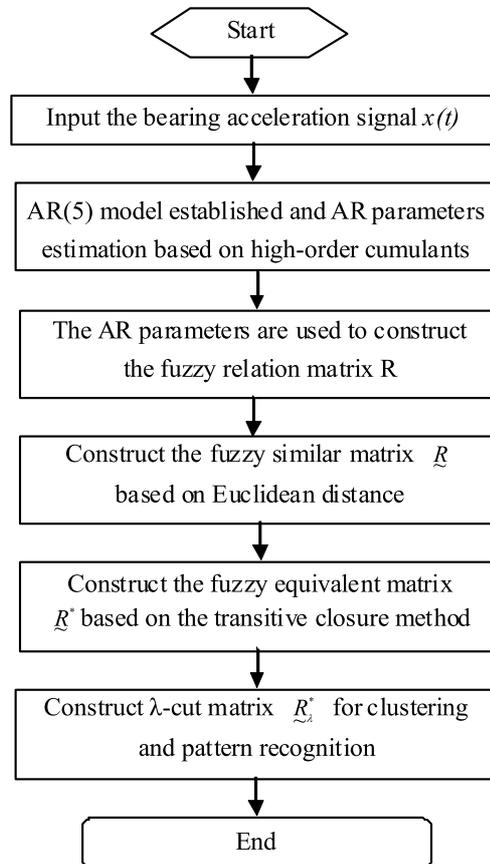


Fig. 3. The flow chart of the analysis process with the proposed method.

5.1. Clustering performance analysis of different fault types

The AR parameters $a_j(j = 0, 1, 2, 3, 4, 5)$ of the samples on the data set DA11_07 are listed in Table 2. There are so many samples but limited space, so only three samples are selected randomly from each of the four fault types. From Table 2, we can see, there is always $a_0 = 1$. In fact, after equalized and normalized by Eq. (9), it also can reflect to different fault types, and the test results showed that there is a better effect by using a_0 as a character to construct a 6 dimension feature vector. According to the flow illustrated in Fig. 3, after a series of deduction and

Table 3
The range of λ to ensure the relevant strategic clusters of the data set DAI1

Data set	Range of λ	$\Delta\lambda = \lambda \max - \lambda \min$
DAI1	DAI1_07 (0.9295,0.9387]	0.0092
	DAI1_14 (0.7811,0.8903]	0.1092
	DAI1_28 (0.7389,0.9387]	0.1998

calculation, the fuzzy equivalent matrix $\tilde{R}_{DAI1_07}^*$ is obtained and shown as follows.

$$\tilde{R}_{DAI1_07}^* = \begin{bmatrix} 1.0000 & 0.9387 & 0.9950 & 0.5847 & 0.5847 & 0.5847 & 0.5847 & 0.5847 & 0.5847 & 0.5847 & 0.5847 & 0.5847 & 0.5847 \\ 0.9387 & 1.0000 & 0.9387 & 0.5847 & 0.5847 & 0.5847 & 0.5847 & 0.5847 & 0.5847 & 0.5847 & 0.5847 & 0.5847 & 0.5847 \\ 0.9950 & 0.9387 & 1.0000 & 0.5847 & 0.5847 & 0.5847 & 0.5847 & 0.5847 & 0.5847 & 0.5847 & 0.5847 & 0.5847 & 0.5847 \\ 0.5847 & 0.5847 & 0.5847 & 1.0000 & 0.9913 & 0.9925 & 0.9295 & 0.9295 & 0.9295 & 0.9046 & 0.9046 & 0.9046 & 0.9046 \\ 0.5847 & 0.5847 & 0.5847 & 0.9913 & 1.0000 & 0.9913 & 0.9295 & 0.9295 & 0.9295 & 0.9046 & 0.9046 & 0.9046 & 0.9046 \\ 0.5847 & 0.5847 & 0.5847 & 0.9925 & 0.9913 & 1.0000 & 0.9295 & 0.9295 & 0.9295 & 0.9046 & 0.9046 & 0.9046 & 0.9046 \\ 0.5847 & 0.5847 & 0.5847 & 0.9295 & 0.9295 & 0.9295 & 1.0000 & 0.9828 & 0.9819 & 0.9046 & 0.9046 & 0.9046 & 0.9046 \\ 0.5847 & 0.5847 & 0.5847 & 0.9295 & 0.9295 & 0.9295 & 0.9828 & 1.0000 & 0.9819 & 0.9046 & 0.9046 & 0.9046 & 0.9046 \\ 0.5847 & 0.5847 & 0.5847 & 0.9295 & 0.9295 & 0.9295 & 0.9819 & 0.9819 & 1.0000 & 0.9046 & 0.9046 & 0.9046 & 0.9046 \\ 0.5847 & 0.5847 & 0.5847 & 0.9046 & 0.9046 & 0.9046 & 0.9046 & 0.9046 & 0.9046 & 1.0000 & 0.9713 & 0.9713 & 0.9713 \\ 0.5847 & 0.5847 & 0.5847 & 0.9046 & 0.9046 & 0.9046 & 0.9046 & 0.9046 & 0.9046 & 0.9713 & 1.0000 & 0.9782 & 0.9782 \\ 0.5847 & 0.5847 & 0.5847 & 0.9046 & 0.9046 & 0.9046 & 0.9046 & 0.9046 & 0.9046 & 0.9713 & 0.9782 & 1.0000 & 1.0000 \end{bmatrix}$$

Based on the fuzzy equivalent matrix $\tilde{R}_{DAI1_07}^*$, we would be obtained different classification results by taking suitable $\lambda \in [0, 1]$. For example, by taking $\lambda_1 \in (0.5847, 0.9046]$, the λ_1 -cut matrix $\tilde{R}_{\lambda_1}^*$ is obtained. All of the samples have been divided into two clusters – $\{s_1, s_2, s_3\}$ and $\{s_4, s_5, s_6, s_7, s_8, s_9, s_{10}, s_{11}, s_{12}\}$. By taking $\lambda_2 \in (0.9295, 0.9387]$, the λ_2 -cut matrix $\tilde{R}_{\lambda_2}^*$ is obtained. All of the samples have been divided into four relevantly strategic clusters – $\{s_1, s_2, s_3\}, \{s_4, s_5, s_6\}, \{s_7, s_8, s_9\}$ and $\{s_{10}, s_{11}, s_{12}\}$.

$$\tilde{R}_{\lambda_1-DAI1_07}^* = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

$$\tilde{R}_{\lambda_2-DAI1_07}^* = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Perform the same analysis process on the data set DAI1_14 and DAI1_28, and the range of λ to ensure the relevant strategic clusters – $\{s_1, s_2, s_3\}, \{s_4, s_5, s_6\}, \{s_7, s_8, s_9\}$ and $\{s_{10}, s_{11}, s_{12}\}$, is listed in Table 3. There is $\Delta\lambda_{DAI1_07} < \Delta\lambda_{DAI1_14} < \Delta\lambda_{DAI1_28}$. Suppose the larger of $\Delta\lambda$, the better performance of the clustering. It is found that the graver severity of the bearing, the better performance of the clustering. That means the bearing degradation can be qualitative assessed in terms of the clustering performance.

To show the efficiency of the selected features and the classifier algorithm, a comparison is made between the fuzzy cluster and BP neural network (BPN) with the AR order 4, 5, 6, 7, 8, respectively. The BPN has one hidden-layer. It is trained using fast back-propagation method. The diagnostic results using both fuzzy cluster and BPN are summarized in Table 4.

From this table, it is observed that among different AR order, AR(5) gives the best accuracy, therefore, it has been chosen for establishing AR model and extracting feature vector. Furthermore, the results show that the diagnostic

Table 4
Comparison between the fuzzy cluster and BPN using different AR order

Method	Diagnostic accuracy (%)				
	AR(4)	AR(5)	AR(6)	AR(7)	AR(8)
Fuzzy cluster	95.83	96.67	90.83	93.33	94.16
BPN	89.17	92.50	87.50	90.83	88.33

Table 5
The AR parameters of the data set DInner

Sample		a0	a1	a2	a3	a4	a5
s1	N	1.0000	-2.0548	1.8154	-0.5866	-0.2943	0.2515
s2	N	1.0000	-1.9376	1.5669	-0.3458	-0.4054	0.2681
s3	N	1.0000	-2.0438	1.8013	-0.5976	-0.2777	0.2378
s4	L07	1.0000	-0.7876	1.2450	-0.9504	0.6628	-0.2845
s5	L07	1.0000	-0.8133	1.2630	-0.9805	0.6782	-0.2903
s6	L07	1.0000	-0.7843	1.2538	-0.9652	0.6724	-0.3039
s7	L14	1.0000	-0.5228	1.2714	-1.0362	0.5294	-0.5762
s8	L14	1.0000	-0.2785	1.0934	-0.6947	0.3649	-0.4149
s9	L14	1.0000	-0.4912	1.2008	-0.9827	0.4708	-0.5385
s10	L28	1.0000	-1.5562	1.9957	-1.6226	0.8730	-0.2062
s11	L28	1.0000	-1.7020	2.2074	-1.8610	1.0392	-0.2633
s12	L28	1.0000	-1.7129	2.2092	-1.8601	1.0277	-0.2579

accuracy of fuzzy cluster analysis is better than those of BPN. This demonstrates the proposed procedure is an efficiency method for fault classification of roller bearings.

5.2. Clustering performance analysis of different fault severities

The AR parameters $a_j(j = 0, 1, 2, 3, 4, 5)$ of the samples on the data set DAll_07 are listed in Table 5. Three samples are selected randomly from each of the four failure severity phase. Then the fuzzy equivalent matrix \tilde{R}_{DInner}^* is shown as follows.

$$\tilde{R}_{DInner}^* = \begin{bmatrix} 1.0000 & 0.9387 & 0.9950 & 0.7339 & 0.7339 & 0.7339 & 0.7339 & 0.7339 & 0.7339 & 0.7339 & 0.7339 & 0.7339 \\ 0.9387 & 1.0000 & 0.9387 & 0.7339 & 0.7339 & 0.7339 & 0.7339 & 0.7339 & 0.7339 & 0.7339 & 0.7339 & 0.7339 \\ 0.9950 & 0.9387 & 1.0000 & 0.7339 & 0.7339 & 0.7339 & 0.7339 & 0.7339 & 0.7339 & 0.7339 & 0.7339 & 0.7339 \\ 0.7339 & 0.7339 & 0.7339 & 1.0000 & 0.9913 & 0.9925 & 0.8665 & 0.8665 & 0.8665 & 0.8527 & 0.8527 & 0.8527 \\ 0.7339 & 0.7339 & 0.7339 & 0.9913 & 1.0000 & 0.9913 & 0.8665 & 0.8665 & 0.8665 & 0.8527 & 0.8527 & 0.8527 \\ 0.7339 & 0.7339 & 0.7339 & 0.9925 & 0.9913 & 1.0000 & 0.8665 & 0.8665 & 0.8665 & 0.8527 & 0.8527 & 0.8527 \\ 0.7339 & 0.7339 & 0.7339 & 0.8665 & 0.8665 & 0.8665 & 1.0000 & 0.8903 & 0.9785 & 0.8527 & 0.8527 & 0.8527 \\ 0.7339 & 0.7339 & 0.7339 & 0.8665 & 0.8665 & 0.8665 & 0.8903 & 1.0000 & 0.8903 & 0.8527 & 0.8527 & 0.8527 \\ 0.7339 & 0.7339 & 0.7339 & 0.8665 & 0.8665 & 0.8665 & 0.9785 & 0.8903 & 1.0000 & 0.8527 & 0.8527 & 0.8527 \\ 0.7339 & 0.7339 & 0.7339 & 0.8527 & 0.8527 & 0.8527 & 0.8527 & 0.8527 & 0.8527 & 1.0000 & 0.9737 & 0.9737 \\ 0.7339 & 0.7339 & 0.7339 & 0.8527 & 0.8527 & 0.8527 & 0.8527 & 0.8527 & 0.8527 & 0.9737 & 1.0000 & 0.9969 \\ 0.7339 & 0.7339 & 0.7339 & 0.8527 & 0.8527 & 0.8527 & 0.8527 & 0.8527 & 0.8527 & 0.9737 & 0.9969 & 1.0000 \end{bmatrix}$$

Based on the fuzzy equivalent matrix \tilde{R}_{DInner}^* , by taking $\lambda_1 \in (0.8665, 0.8903]$, the relevant strategic clusters $\{s_1, s_2, s_3\}, \{s_4, s_5, s_6\}, \{s_7, s_8, s_9\}$ and $\{s_{10}, s_{11}, s_{12}\}$ are obtained.

Perform the same analysis process on the data set DBall and DOuter, and the range of λ to ensure the relevant strategic clusters $\{s_1, s_2, s_3\}, \{s_4, s_5, s_6\}, \{s_7, s_8, s_9\}$ and $\{s_{10}, s_{11}, s_{12}\}$, is listed in Table 6. From this table, there is $\Delta\lambda_{DInner} < \Delta\lambda_{DBall} < \Delta\lambda_{DOuter}$. That means the bearings with outer fault have the best distinguishing for failure severities.

Table 6
The range of λ to ensure the relevant strategic clusters of the data set DSingle

Data set		Value range of λ	$\Delta\lambda = \lambda_{\max} - \lambda_{\min}$
DSingle	DInner	(0.8665,0.8903]	0.0238
	DBall	(0.8793,0.9387]	0.0594
	DOuter	(0.6847,0.9387]	0.2540

6. Conclusion

In this paper, the AR parameters of the raw bearing signals estimated in terms of the higher-order cumulants are applied to classify various fault types and evaluate various fault severities of the roller bearing. This application overcomes AR model can only be applied to stationary signals by using routine methods, and is propping for non-stationary and non-Gaussian signals. Fuzzy cluster analysis based on fuzzy equivalence relation is used as a classifier for pattern recognition.

In the bearing fault simulation experiments, classification results on the category of different fault types showed that different fault types can be identified by choosing proper cutting level, and the graver severity of the bearing, the better performance of the clustering. That means the bearing degradation can be qualitative assessed in terms of the clustering performance. At the same time, classification results on the category of different fault severities showed that various fault severities also can be evaluated. Those investigations demonstrate that the proposed approach can be applied to bearing fault diagnosis and degradation assessment, and the AR parameters of the raw bearing vibration signals estimated in terms of the higher-order cumulants can be used as feature vectors. More over it can reflect not only bearing fault types but also fault severities.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (50775070), the Program for New Century Excellent Talents in University of China (NCET-08-0677), and the Key Program of the Hunan Natural Science Foundation (09JJ8005).

References

- [1] A.K.S. Jardine, D. Lin and D. Banjevic, A review on machinery diagnostics and prognostics implementing condition-based maintenance, *Mechanical Systems and Signal Processing* **20** (2006), 1483–1450.
- [2] B.M. Sadler, G.B. Giannakis and K.S. Lii, Estimation and detection in non-Gaussian noise using higher order statistics, *IEEE Transactions on Signal Processing* **42**(10) (1994), 2729–2741.
- [3] F. Hoppner, F. Klawonn, R. Kruse et al., *Fuzzy cluster analysis*, Wiley, Chichester, 1999.
- [4] H. Qiu, J. Lee, J. Lin and G. Yu, Robust performance degradation assessment methods for enhanced rolling element bearing prognostics, *Advanced Engineering Informatics* **17**(3–4) (2003), 127–140.
- [5] H. Ding, Y. Wu and S.Z. Yang, *Fault diagnosis by time series analysis, applied time series analysis*, World Scientific Publishing Co, Singapore, 1989.
- [6] H. Chih and R. Chen, A New Fault Diagnosis Method of Rotating Machinery, *Shock and Vibration* **15** (2008), 585–598.
- [7] H. Akaike, Fitting autoregressive models for prediction, *Annals of the Institute of Statistical Mathematics* **21** (1969), 1572–9052.
- [8] J.Y. Yang and Y.Y. Zhang, Intelligent fault diagnosis of rolling element bearing based on SVMs and statistical characteristics, *Proceeding of the International Manufacturing Science and Engineering Conference* (2007), 1–12.
- [9] J.Y. Yang, Y.Y. Zhang and Y.S. Zhu, Intelligent fault diagnosis of rolling element bearing based on SVMs and fractal dimension, *Mechanical Systems and Signal Processing* **21** (2007), 2012–2024.
- [10] J.S. Cheng, D.J. Yu, J. SH. Tang et al., Application of SVM and SVD technique based on EMD to the fault diagnosis of the rotating machinery, *Shock and Vibration* **16** (2009), 89–98.
- [11] J.S. Cheng, D.J. Yu and Y. Yu, A fault diagnosis approach for roller bearings based on EMD method and AR model, *Mechanical Systems and Signal Processing* **20** (2006), 350–362.
- [12] J.M. Mendel, Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some application, *Proceedings of the IEEE* **79** (1991), 278–305.
- [13] M.S. Yang and H.M. Shih, Cluster analysis based on fuzzy relations, *Fuzzy Sets and Systems* **120**(2) (2001), 197–212.

- [14] M.A.A.S. Choudhury, S.L. Shah and N.F. Thornhill, Diagnosis of poor control-loop performance using higher-order statistics, *Automatica* **40** (2004), 1719–1728.
- [15] N.G. Nikolaou and A. Antoniadisi, Rolling element bearing fault diagnosis using wavelet packets, *NDT&E International* **35** (2002), 197–205.
- [16] P. Stepanic, L.V. Latinovic and Z. Djurovic, A new approach to detection of defects in rolling element bearings based on statistical pattern recognition, *Int J Adv Manuf Technol* **45** (2009), 91–100.
- [17] T.S. Qiu and X.X. Zhang, Statistical signal processing – non-Gaussian signal processing and its application, *Publishing Housing of Electronics Industry*, Beijing, 2004.
- [18] X.B. Gao, *Fuzzy Cluster Analysis and its Application*, Xi'an Electronic Science and Technology University Press, 2004.
- [19] Y.N. Pan, J. Chen and X.L. Li, Bearing performance degradation assessment based on lifting wavelet packet decomposition and fuzzy c-means, *Mechanical Systems and Signal Processing* **24** (2010), 559–566.
- [20] Y. Yang and J.S. Cheng, A roller bearing fault diagnosis method based on EMD energy entropy and ANN, *Journal of Sound and Vibration* **294** (2006), 269–277.
- [21] Y. Wu and S.Z. Yang, *Application of Several Time Series Models in Prediction*, applied time series analysis, World Scientific Publishing Co, Singapore, 1989.
- [22] Z.J. He, Y.Y. Zi and Q.F. Meng, Fault diagnosis principle of non-stationary signal and application to mechanical equipment, Higher Education Press, Beijing, 2001.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

