Mechanical Systems and Signal Processing 35 (2013) 150–166

Contents lists available at SciVerse ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

Roller element bearing fault diagnosis using singular spectrum analysis

Bubathi Muruganatham*, M.A. Sanjith, B. Krishnakumar, S.A.V. Satya Murty

Electronics and Instrumentation Division, Indira Gandhi Centre for Atomic Research, Kalpakkam, Tamil Nadu 603102, India

ARTICLE INFO

Article history:
Received 7 November 2011
Received in revised form
13 August 2012
Accepted 14 August 2012
Available online 1 September 2012

Keywords:
Singular spectrum analysis
Bearing fault
Vibration analysis
Induction motor
Neural network

ABSTRACT

Most of the existing time series methods of feature extraction involve complex algorithm and the extracted features are affected by sample size and noise. In this paper, a simple time series method for bearing fault feature extraction using singular spectrum analysis (SSA) of the vibration signal is proposed. The method is easy to implement and fault feature is noise immune. SSA is used for the decomposition of the acquired signals into an additive set of principal components. A new approach for the selection of the principal components is also presented. Two methods of feature extraction based on SSA are implemented. In first method, the singular values (SV) of the selected SV number are adopted as the fault features, and in second method, the energy of the principal components corresponding to the selected SV numbers are used as features. An artificial neural network (ANN) is used for fault diagnosis. The algorithms were evaluated using two experimental datasets—one from a motor bearing subjected to different fault severity levels at various loads, with and without noise, and the other with bearing vibration data obtained in the presence of a gearbox. The effect of sample size, fault size and load on the fault feature is studied. The advantages of the proposed method over the exiting time series method are discussed. The experimental results demonstrate that the proposed bearing fault diagnosis method is simple, noise tolerant and efficient.

1. Introduction

Bearing faults are a major problem in rotating machines and they are diagnosed mainly by vibration analysis. The vibration produced by a bearing fault is non-stationary in nature, so it is essential to obtain the fault features accurately. Various methods of feature extraction and fault detection using time series analysis have been implemented so far.

A few approaches use statistical features of the vibration time series for fault detection [1]. Two common methods used in time series analysis are autoregressive (AR) model [2–3,5], autoregressive moving average (ARMA) model [4]. AR analysis uses time variations and vibration amplitude to establish the mathematical model. AR models of different orders for each bearing defect type have been formed for analysis [2]. The AR coefficients are used as fault features which are given as an input to a classifier such as ANN. AR models of the same order for each bearing defect type were established in [3]. In an AR model, the value of the residual squared error should approach zero for the model to be more accurate. It depends on the order selection, which in turn depends on the number of measured points of the signal. If the order of the AR or ARMA model is large, then the ANN inputs will increase correspondingly, which will require more training time. If the order is small, then the residual squared error will be bigger. Also, if the noise level is high, then the performance of

* Corresponding author. Tel.: +91 44 27480500x22624; fax: +91 44 27480228.
E-mail addresses: bubathi@igcar.gov.in, mbubathi@gmail.com, mbubathi@yahoo.co.in (B. Muruganatham).

0888-3270/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ymssp.2012.08.019

the model will be poor. Wanga et al. [5] showed that the approach based on the difference values of the AR coefficients (difference of AR coefficients of ideal signal for healthy condition and AR coefficients for faulty conditions) with an ANN is superior to that using the AR coefficients with an ANN. Recently, methods using the difference histogram [6], zero crossing intervals [7] and multiscale fuzzy sample entropy [8] have been proposed to detect bearing faults. Most of the aforementioned methods are complex in nature. Very few of them use a simple approach, and thus further research towards developing a simple time series method for bearing fault feature extraction is required.

SSA which is based on singular value decomposition (SVD) has been used in many applications such as climatic and meteorological data analysis [9], tool condition monitoring [10] and rotor-to-stator rub failure detection [11]. The first work towards the applicability of singular spectrum analysis for bearing fault detection was done by the present authors by considering only the inner race fault [12].The principal components in which the fault frequency and its harmonics were present were selected and their time domain parameters were used for fault detection. Cascade SSA for multilevel decomposition [13] similar to wavelet decomposition is proposed for bearing damage detection. Based on the proportion of variance assumed, the corresponding principal components were selected from the decomposed signal. The number of principal components used for reconstruction was used as a bearing health degradation indicator.

In this work, a simple time series method for detection and classification of the bearing faults using SSA is proposed. A new fault feature and novel approach for the selection of the principal components are discussed which requires neither knowledge of the fault frequency nor the amount of variance, as used in the earlier two works based on SSA. Two methods are proposed. In the first method, a selected subset of the singular values (SV) is adopted as the fault feature. The correspondence between the singular spectrum (plot of SV) and the power spectrum is also shown. In the second method, the energy of the principal components corresponding to the selected SV numbers is used as the fault feature. An ANN is used for fault classification as it has been shown to be very effective in bearing fault classification [14]. The methods are evaluated using data obtained from a bearing test rig and a gearbox test rig. The block diagram of the proposed method for roller bearing fault diagnosis is given as in Fig. 1.

The paper hereafter is organized as follows. Section 2 introduces the SSA method used for feature extraction. The two sets of experimental data are described in Section 3. First method (SV based fault diagnosis) is explained in Section 4. Section 5 describes the second method (energy feature based fault diagnosis) and comparison of the proposed method with the existing methods is done in Section 6. Finally, the conclusion is drawn in Section 7.

2. Singular spectrum analysis

SSA [15] is a time series analysis technique which decomposes the signal into an additive set of independent time series (principal components). The set of series resulting from the decomposition is interpreted as consisting of a trend representing the signal mean at each instant, a set of periodic series, and an aperiodic noise. SSA is applied to the vibration signals obtained under various bearing conditions considered for the study. The SSA algorithm is described below.

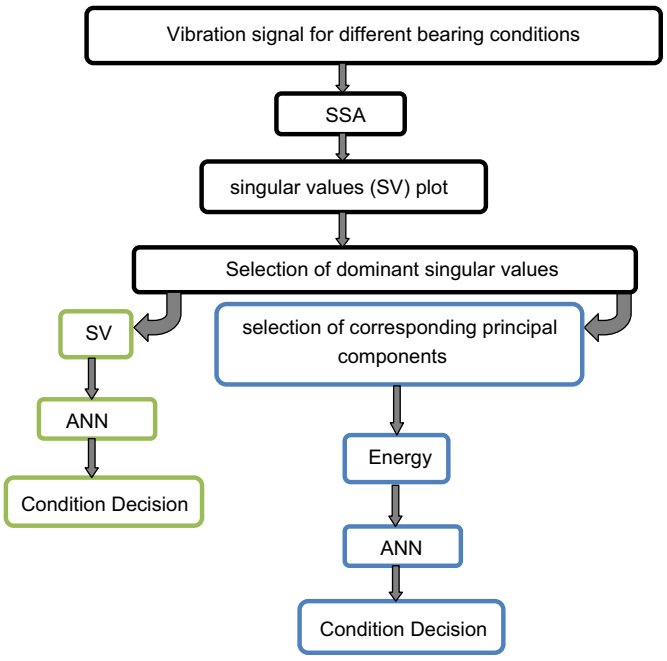

Fig. 1. Block diagram of the proposed roller bearing fault diagnosis method based on SSA.

2.1. Step1—Embedding

In this step, the bearing vibration signal $F=(f_0, \ldots, f_{N-1})$ of N data points is mapped into a sequence of multidimensional lagged vectors. Let L be an integer (window length), $1<L<N$ and K be the number of lagged vectors ($K=N-L+1$), $X_i=(f_{i-1,\ldots\ldots}f_{i+L-2})^T$, where $1\le i\le K$.

The obtained trajectory matrix X is given as (1)

$$X=\begin{pmatrix} f_0 & f_1 & \cdots & \cdots & f_{K-1} \\ f_1 & f_2 & \cdots & \cdots & f_k \\ . & . & \cdots & \cdots & . \\ . & . & . & \cdots & . \\ f_{L-1} & f_{L-2} & \cdots & \cdots & f_{N-1} \end{pmatrix} \tag{1}$$

2.2. Step 2—Singular value decomposition

In this step, the SVD of the matrix X is obtained. It is further decomposed into a sum of mutually orthogonal, unit-rank, elementary matrices. Let $S=XX^T$ be an $L\times L$ matrix. Let $\lambda_1, \lambda_2, \ldots \lambda_d$ be the non-zero eigenvalues of S ($d=L$ if no eigenvalue is zero) arranged in decreasing order, and $U_1, U_2 \ldots U_d$ be the corresponding eigenvectors. The vectors $V_j=X^T(U_j/\sqrt{\lambda_j})$ for $j=1,2,\ldots,d$ is constructed.

$$\text{The SVD of} \quad X=X_1+X_2+\ldots+X_d \tag{2}$$

where $X_j=\sqrt{\lambda_j}U_jV_j^T$

The obtained L numbers of SV's are the square roots of the eigenvalues of matrix S. The plot of the SV in decreasing order is known as singular spectrum (SS), and thus gives the method its name.

2.3. Step 3—Diagonal averaging

Each matrix of decomposition (2) is transformed into a new time series of length N by applying a linear transformation known as diagonal averaging. The matrices $X_1, X_2, \ldots, X_L$ get converted to principal components $C_1, C_2, \ldots, C_L$. The diagonal averaging algorithm is as follows:

Let y be a $L\times K$ matrix with elements y_{ij}, where i, j values range are $1\le i\le L$ and $1\le j\le K$. Diagonal averaging transforms the matrix Y into the series $g_0 \ldots g_{N-1}$ (principal component) by the formula:

$$g_k=\begin{cases} \frac{1}{k+1}\sum_{m=1}^{k+1} y_{m,k-m+2} & \text{for } 0\le k\le L^*-1 \\ \frac{1}{L^*}\sum_{m=1}^{L^*} y_{m,k-m+2} & \text{for } L^*-1\le k\le K^* \\ \frac{1}{N-k}\sum_{m=k-K^*+2}^{N-K^*+1} y_{m,k-m+2} & \text{for } K^*\le k\le N-1 \end{cases} \tag{3}$$

where $L^*=min(L,K)$, $K^*=max(L,K)$

The simplicity and efficiency of the SSA for time series based bearing fault classification will be illustrated using two datasets in the following sections.

3. Experimental data

3.1. Test data1

The Test data1 set used in this work was acquired from the Case Western Reserve University bearing data centre [16]. The experimental setup as depicted in Fig. 2, consists of a 2 HP three-phase induction motor (Reliance Electric IQPreAlert motor) with a torque transducer, accelerometer and a dynamometer. The drive end bearing (series deep grove ball bearing) [Appendix] data are used in this analysis. Single point faults are introduced in the inner raceway, outer raceway and ball elements of different bearings using electron-discharge machining with fault diameters of 0.18, 0.36, 0.53, 0.71 mm and a depth of 0.28 mm. Vibration data are collected at 12 kHz and for different fault sizes with load varying from 0 to 3 HP with 1 HP increments. The speed varies from 1797 rpm (no-load) to 1730 rpm (3 HP load). The data for the outer race fault are taken with the fault position centered at the 6 o'clock position with respect to the load zone. The load zone is centered at 6 o'clock position. Fig. 3 shows the recorded bearing acceleration vibration signals for the healthy and faulty conditions.

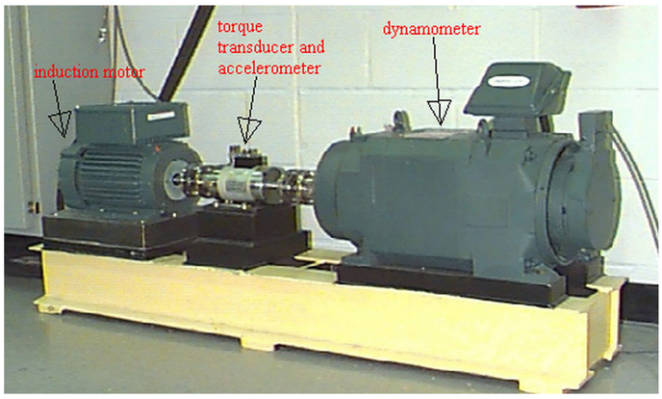

Fig. 2. Experimental setup—test data1 [16].

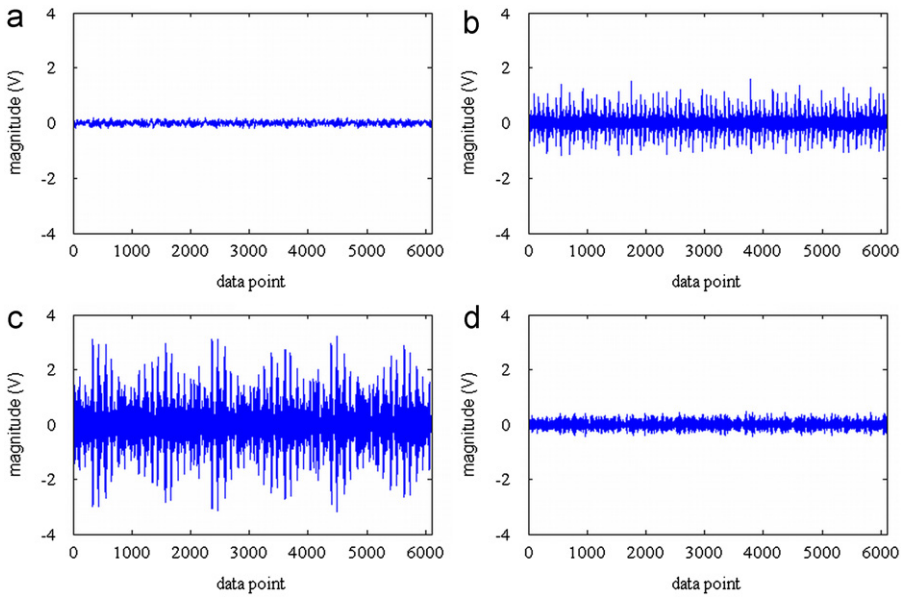

Fig. 3. Bearing vibration signal of test data1 (0.18 mm fault, 1 HP load) under (a) H; (b) IF; (c) OF; (d) BF.

The following four bearing conditions are considered for the study: healthy bearings (H), a bearing with a point defect on inner race (IF), a bearing with a point defect on the outer race (OF) and a bearing with a point defect on a ball (BF).

3.2. Test data2

The test data2 set used was obtained from a spur gear test rig [17,18] at the University of New South Wales (UNSW). The bearing fault data are taken from [18]. Fig. 4 shows the gear test rig used. Fig. 5 gives the schematic representation of the test setup. The test rig consists of a single stage spur gearbox (set with 1:1 ratio and 32 gear teeth) driven by a three phase motor and driving a hydraulic pump/motor set. The shaft of the gearbox has two double row ball bearings (Appendix). The shaft speed was 6 Hz. The sampling frequency of the acceleration signal is 48 kHz. The inner and outer race had a notch fault of width 0.8 mm and depth 0.3 mm, while the ball had a fault with both width and depth of 0.5 mm. Fig. 6 gives the outer race fault details. The total torque load on the gear was 100 N m. The obtained vibration signals had a SNR in the range of 15 to 25 dB [17]. Fig. 7 shows the obtained vibration signal for the bearing under different conditions.

4. SV and ANN based bearing fault diagnosis—1st method

The singular values are the representation of the signal in the singular spectrum domain. At the end of Step 2 of the SSA method (Section 2), we get L number of SV. These SV are plotted with respect to their number (SS plot). The SS plots for the different bearing conditions are compared and then the appropriate SV are selected. They are adopted as the inputs to an

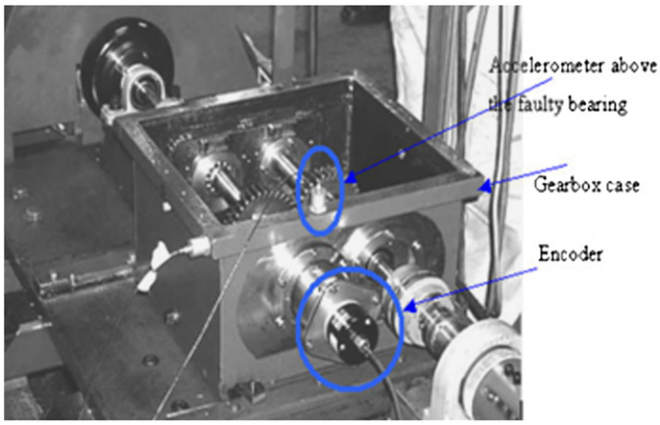

Fig. 4. Spur gear test rig [17]—test data2.

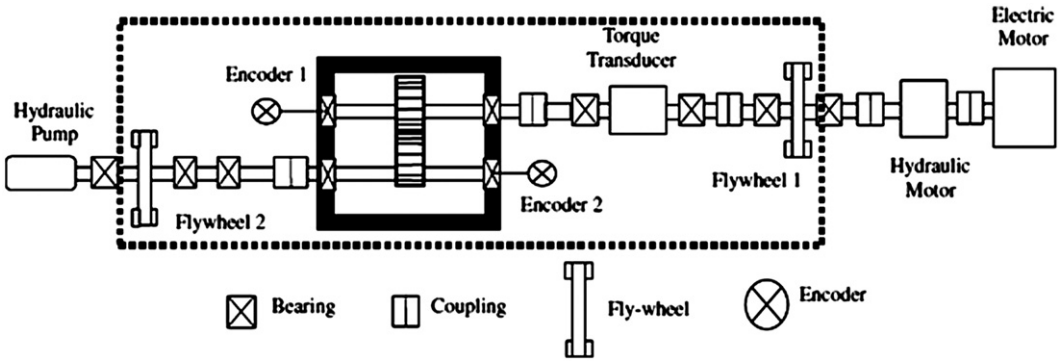

Fig. 5. Test data2 setup schematic representation [17].

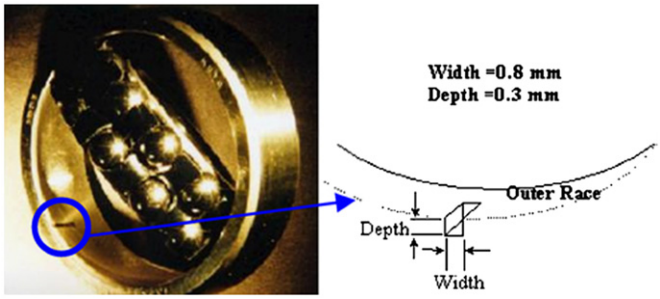

Fig. 6. Outer race fault [17].

ANN for condition diagnosis. This method is tested with the vibration signal of test data1 and test data2, which are described in the subsequent sections.

4.1. Test data1 analysis

4.1.1. Selection of singular values

Choice of window length: Different window lengths ($L=5$, 10 and 15) were used for embedding the signal, and the corresponding L number of SV are obtained based on the procedure as given in Section 2. The SS plots of the bearing vibration signal in healthy and faulty conditions for values of L are shown in Fig. 8. The larger SV in the singular spectrum represent the large amplitude components in the decomposition, and the low-amplitude components of the signal are represented by the smaller SV. The SV should be selected such they differs for all the bearing conditions. For $L=5$ (Fig. 8(a)), all the SV are found to separate the different bearing conditions. If all these SV are used it would result in processing the whole signal. As seen from Figs. 8(b) and 8(c), for $L=10$ and $L=15$, many of the SV differ for all the bearing conditions. With an appropriate L value the fault feature of the signal would be extracted more clearly. As compared to

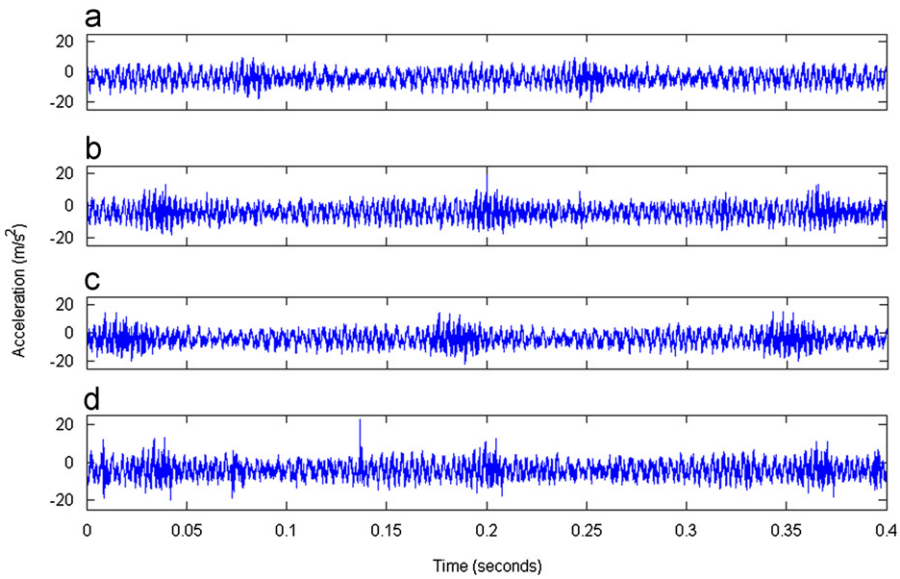

Fig. 7. Bearing vibration signals of test data2: (a) H, (b) IF, (c) OF (d) BF.

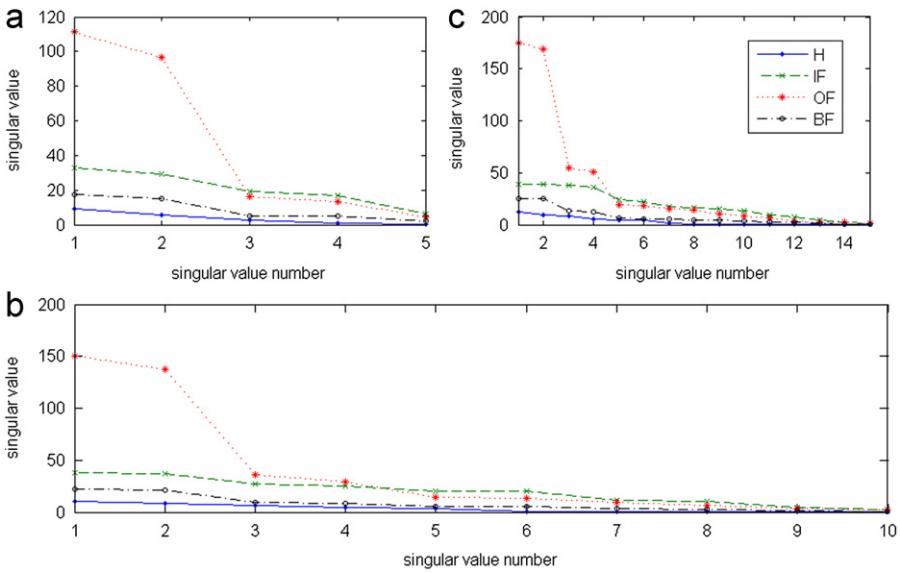

Fig. 8. SS plot of bearing in healthy and faulty conditions (a) $L=5$, (b) $L=10$ and (c) $L=15$.

$L=15$, $L=10$ requires fewer SV to be selected with respect to the total number available. We could also use the singular values from case $L=5$ or $L=15$, but we require fewer SV when $L=10$. Hence, a window length of 10 is selected.

Selection criteria of SV: The SV should have distinct values for healthy, inner race fault, outer race fault and ball fault should be selected so that not only the faulty condition is identified but also the type of fault is known (even for different combination of load and fault sizes).

To select the appropriate SV, the singular spectrum plots were obtained for four different fault sizes (0.18, 0.36, 0.53 and 0.71 mm) taken at 1 HP load, under healthy and various types of faulty conditions, which is shown in Fig. 9. From Fig. 9, it is seen that there are many SV which differentiate the healthy and faulty conditions, i.e., these SV have distinguishable values with respect to each of the conditions. With reference to Fig. 9(a)–(d), SV1 to SV6 are able to classify the faults for 0.18, 0.53 and 0.71 mm. However, for the 0.36 mm case (Fig.9(b)), SV1, 2, 5 and 6 are found to classify the faults. So, we have selected the SV which can satisfy the criteria. Hence, SV1, SV2, SV5 and SV6 are selected for fault classification purpose. Also, we have further verified the selected SV, by considering the vibration signal obtained for no-load, 2 and 3 HP. It was found that the selected SV have different values for healthy and faulty conditions under

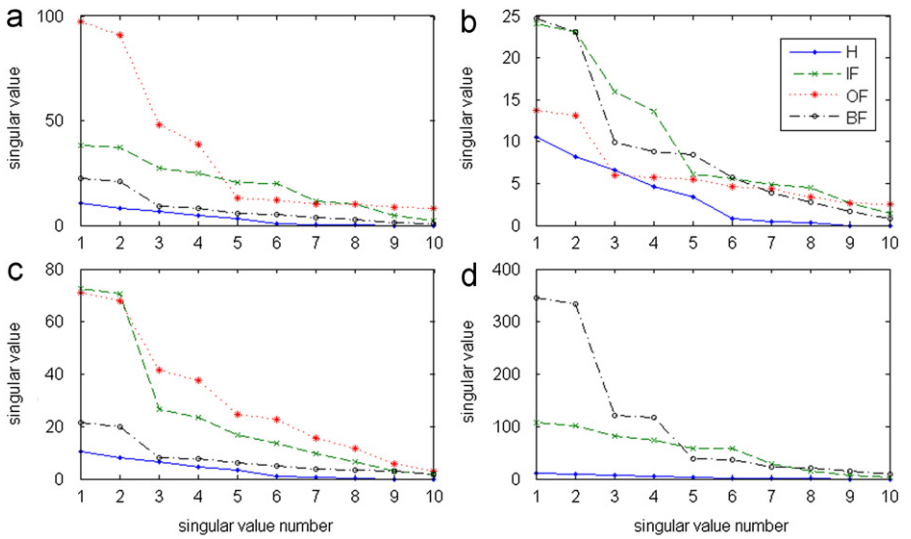

Fig. 9. Shows the SS plot for $L=10$ obtained for bearing under different fault severity conditions, 1 HP load (a) 0.18 mm, (b) 0.36 mm, (c) 0.53 mm and (d) 0.71 mm.

different loads which make them suitable for fault classification. The effect of SV 1, 2, 5 and 6 under varying load and fault sizes are studied in Section 4.1.3. The pattern of the SV remains constant for varying load and fault sizes.

It is seen that SV 1, 2, 5 and 6 are able to classify faults even under different load and fault size. Hence, they are selected.

The SS plot follows the same trend found in the power spectral density plot, as shown in Fig. 16. Hence, it is possible to use the singular spectrum to detect changes in the signal, as it also represents the frequency spectrum of the signal [10]. This correspondence between the singular spectrum and the frequency spectrum is the basis of the processing technique [10], which is shown in Section 4.3.

SV are normalized using the min–max normalization technique. These are then used as an input to an ANN. Based on the output of the ANN; the bearing condition is identified as healthy or faulty including its type.

4.1.2. ANN performance results using test data1

The ANN used in this work is a feed-forward back propagation neural network (BPNN) with a single hidden layer architecture. There are four nodes in the input layer each representing the SV. There are four nodes used in the output layer. The output pattern of the nodes depends on the bearing condition: H [1-1-1], IF [-1 1 -1 -1], OF [-1 -1 1 -1], and BF [-1 -1 -1 1]. The BPNN is created, trained and implemented using the MATLAB neural network toolbox. A total of 480 datasets of 4 normalized features (1st, 2nd, 5th and 6th SV) are obtained for the different bearing conditions, out of which 336 are used to train the network and the remaining are used to test the network. BPNN training parameters used were a mean square error of 10^{-10}, a minimum gradient of 10^{-10} and a maximum iteration number (epoch) of 1000. The training process would stop if any of these conditions are met. Log sigmoid and tan sigmoid activation functions are used for the hidden and outer layers, respectively.

The experimental setup has no masking source elements such as gear vibration. Hence, gaussian white noise was added such that noisy signals of SNR $+10$ and -10 dB are obtained for each recorded vibration signal, respectively. SV values are obtained for three cases—original data, noisy data of $+10$ dB SNR and noisy data of -10 dB SNR. Table 1 shows the BPNN overall classification accuracy. As observed from Table 1, even in the presence of noise, high classification accuracy is obtained. The performance of the fault diagnosis method is highly dependent on two factors: First, how well the fault features contain information about the different fault conditions; second, the ability of the classifier to accurately differentiate the different types of fault with respect to the input fault features. The selected singular values have distinct values for each type of bearing condition, i.e., each condition has a specific value of singular value, which makes the different conditions easily separable. Thus, they have good information about each type of condition. Even for the noisy signal, the SV 1, 2, 5 and 6 has distinguishable values for each of the four conditions. The ANN is trained and tested with a k-fold cross validation technique ($k=10$) which gives much less chance of over training. It is the combination of the distinct SV (even under noisy conditions) and the good generalization capability of the ANN that contribute to the accurate classification of the bearing conditions even under noisy conditions.

The generalization ability of the ANN was investigated. The terms "over training" and "generalization" are interrelated. Over training means over fitting, i.e., the model tends to work only with a particular dataset. Over training an ANN makes it rigid and hinders its ability to handle unseen data, i.e., it has low generalization ability. In our work, the ANN is trained for different numbers of hidden neurons (1 to 30) and the optimum number of hidden neurons which attains the required mean square error (*square of the difference between the target and actual value*) was chosen. Dalianis et al., conducted a

Table 1
BPNN performance for test data1—1st method.

Signal	Hidden neuron	Classification accuracy (%)
Test data1	13	96.53
Noisy test data1, SNR +10 dB	13	98.61
Noisy test data1, SNR −10 dB	13	100

Table 2
Confusion matrix for test data1—1st method.

	H	IF	OF	BF
H	12	0	0	0
IF	0	48	0	0
OF	0	0	35	1
BF	0	1	3	44

Table 3
Confusion matrix for noisy test data1 of +10 dB SNR—1st method.

	H	IF	OF	BF
H	12	0	0	0
IF	0	47	0	1
OF	0	0	35	1
BF	0	0	0	48

Table 4
Confusion matrix for noisy test data1 of −10dB SNR—1st method.

	H	IF	OF	BF
H	12	0	0	0
IF	0	48	0	0
OF	0	0	36	0
BF	0	0	0	48

study to detect the over training of an ANN [19] and it was shown that over training is present if the output error decreases up to a point and increases thereafter. In the implemented back propagation ANN, the error between the estimated output and the target values decreases to its lowest value and no further increase in error was found. Also, to test whether the ANN could classify the unseen data, it was trained with the features obtained for the original signal and tested with an independent dataset (data obtained from a noisy signal with SNR of +10 dB). In this case, a classification accuracy of 90% is obtained. It means that the network is correctly classifying the unseen data. Hence, there is very less chance of over training in the used structure. The above two proofs shows that the ANN has no overtraining, i.e., it has a very good generalization capability.

Tables 2–4 give the confusion matrix of the BPNN obtained in each case. The diagonal elements give the number of datasets correctly classified and the non diagonal elements gives the misclassifications. As seen from the confusion matrices, the method gives a very high detection rate for each condition.

4.1.3. *Effect of load and fault size on SV*

To study how the SV are affected with respect to the changes in load and also with the increase in fault severity, vibration signals from bearings with fault sizes of 0.18, 0.53 and 0.71 mm, recorded at no-load and 3 HP load were considered. Figs. 10 and 11 show the variation of the SV with respect to load and fault size for the inner race fault and the ball fault, respectively. In the case of the inner race fault, as fault severity increases, the SV increase, but the SV decreases with an increase in load. For the ball fault, all the SV increase with increase in fault size, while for increase in load, SV1 and SV2 increase and SV5 and SV6 decrease. With an increase in outer race fault size, SV1 and SV2 decrease while SV5 and SV6 increase. However, for an increase in load, SV1 and SV2 increase while SV5 and SV6 decrease. It is seen that the effect depends on the location of the fault. However, the profile of all the SV remains same for a change in load. It is seen that the ratios SV1/SV2 and SV5/SV6 both have a constant value of 1 for all faulty conditions, while for healthy conditions

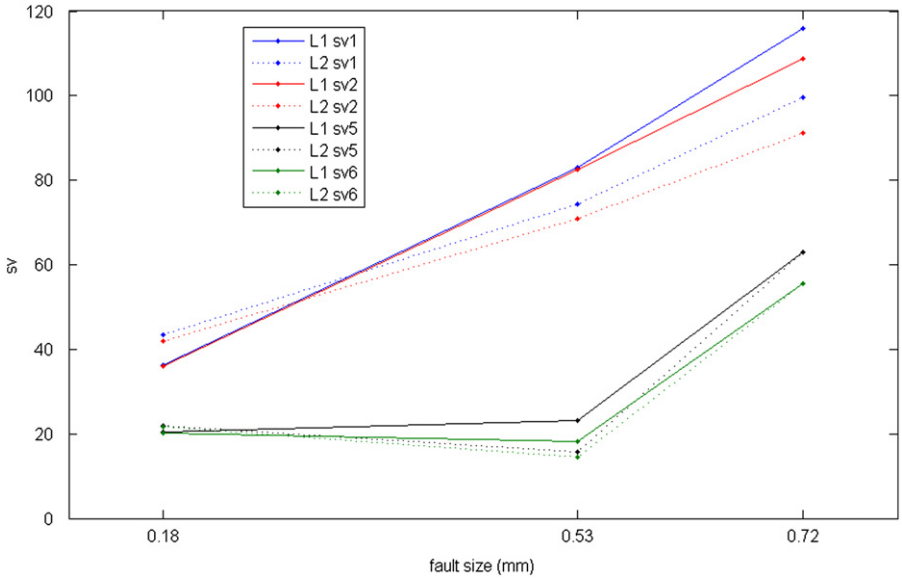

Fig. 10. Effect of load and fault size on SV—inner race fault ($L1=0$ HP, $L2=3$ HP load).

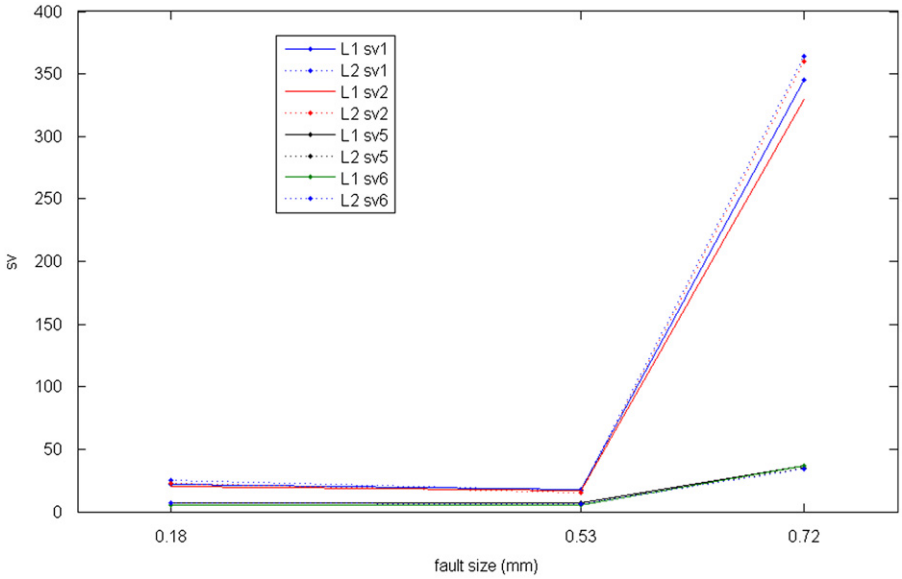

Fig. 11. Effect of load and fault size on SV—ball fault ($L1=0$ HP, $L2=3$ HP load).

$(SV1/SV2) > 1$, $(SV5/SV6) > 3$. Thus, the selected SV are able to classify different load and fault sizes. The fault size effect is more pronounced compared to that of load.

4.1.4. *Effect of sample size on SV*

To investigate the effect of sample size on the SV, the testdata1 obtained for the 0.18 mm fault size taken at 1 HP load is considered. The SV values are obtained for sample sizes of 500, 1000, 2000 and 4000. Fig. 12 shows the variation in the SV with respect to the sample size for healthy and faulty cases. As seen from Fig. 12, as sample size increases, the SV also increase. However, it is found that the ratios (SV1/SV2) and (SV5/SV6) have a constant value of 1 for all faulty conditions, while for healthy conditions the ratio $(SV1/SV2) > 1$ and $(SV5/SV6) > 4$. This shows that trend of the SV for healthy and faulty cases follow the same pattern for different sample sizes. The selected SV have distinct values for each condition even under sample size variation and this easily separates the healthy and different fault conditions. This shows that fault diagnosis can be done accurately and sample size variation has little affect on the classification accuracy.

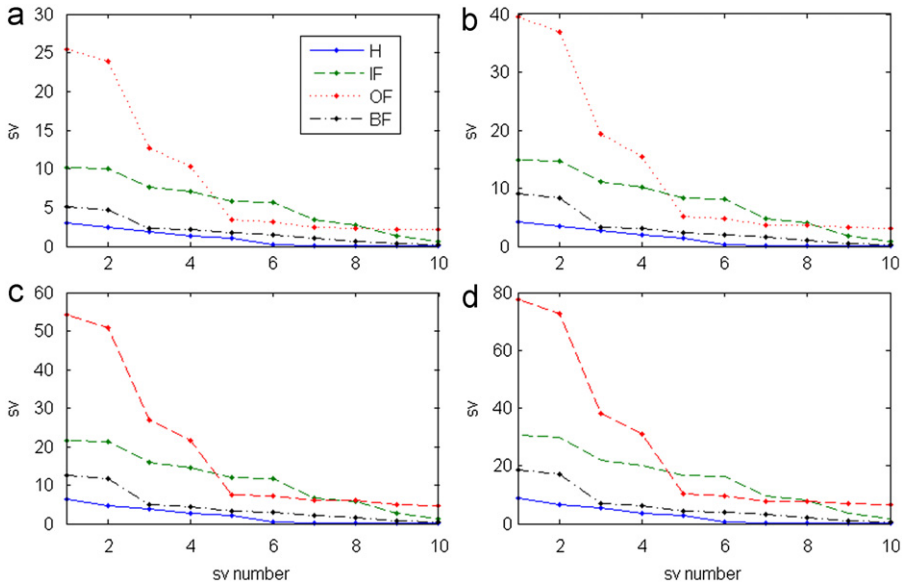

Fig. 12. Effect of sample size (data points) on SV (a) 500, (b) 1000, (c) 2000 and (d) 4000.

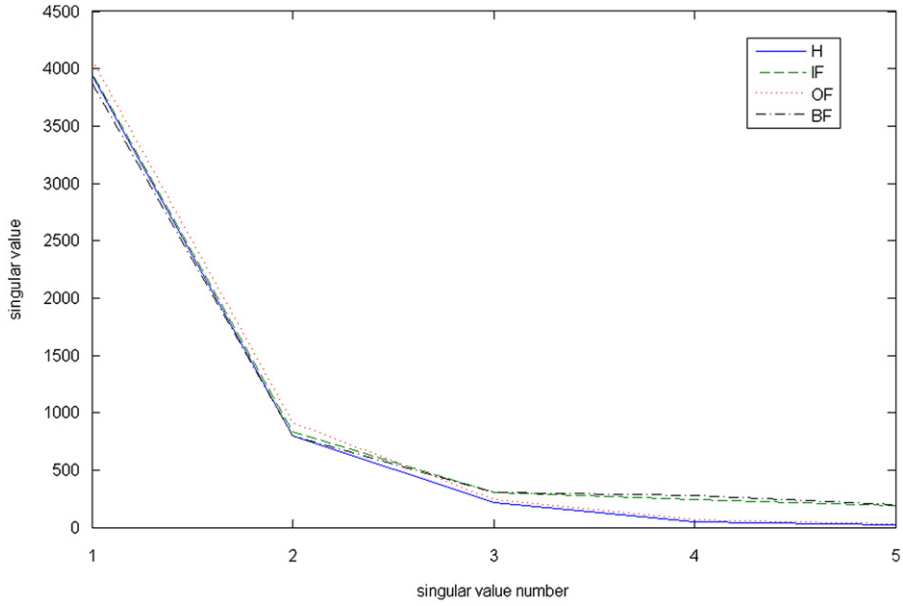

Fig. 13. SS plot of test data2 for $L=5$.

4.2. Test data2 analysis

In the test data1 setup, there was no masking source such as gearbox vibration. To study the method in the presence of a gearbox, signals obtained from the experimental setup of test data2 are used. The procedure followed is similar to that of the test data1 analysis.

4.2.1. Selection of singular values and window length—Test data2

Similar to test data1, different window lengths were used. Figs. 13–15 shows the SS plots obtained for test data2 using $L=5$, 10 and 15, respectively. As observed from Figs. 13 and 15, healthy SV coincide with SV of the outer race fault. Also, SV of the inner race fault and ball fault have similar values. Hence, different bearing condition could not be easily separated using SV's for $L=5$ (Fig. 13) and $L=15$ (Fig. 14). In the case of $L=10$ (Fig. 14), the SV of healthy and outer race fault have close values, while for the inner race fault and ball fault the SV are separated, which is not observed in case of $L=5$ and

$L=10$ plots. Hence, $L=10$ is used as the optimal window length for fault diagnosis. It is seen that SV number 5, 6 and 7 are sufficient to separate the different bearing conditions under study. The SV's of healthy and outer race fault, inner race fault and ball fault have close values which is in consistent with the power spectra density plot shown in Fig. 17.

Mostly, the selection of the principal components obtained from the decomposition of SSA is based on either the first few eigenvalues or those components' eigenvalues which cover a predefined % of the variance of the total eigenvalues [13]. In the proposed method, based on the singular value plot the selection of the SV and principal components are done. As shown in the analysis of test data2, the first few singular values are the same for healthy and faulty conditions (due to the presence of the gearbox) and hence the choice of these components would lead to a wrong diagnosis. Also, the choice of the amount of variance required is unknown and in the case of the gearbox application the amount of variance may include the masked eigenvalues.

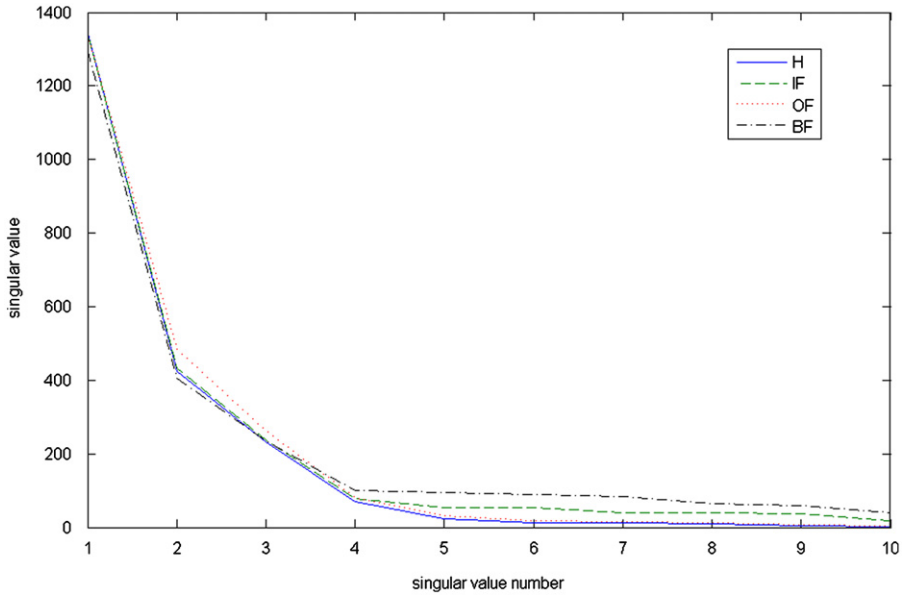

Fig. 14. SS plot of test data2 for $L=10$.

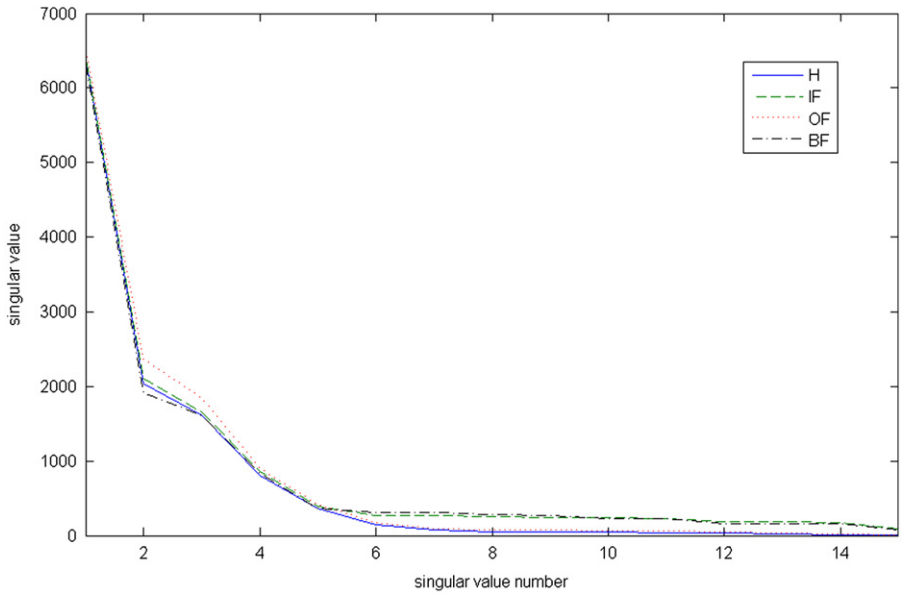

Fig. 15. SS plot of test data2 for $L=15$.

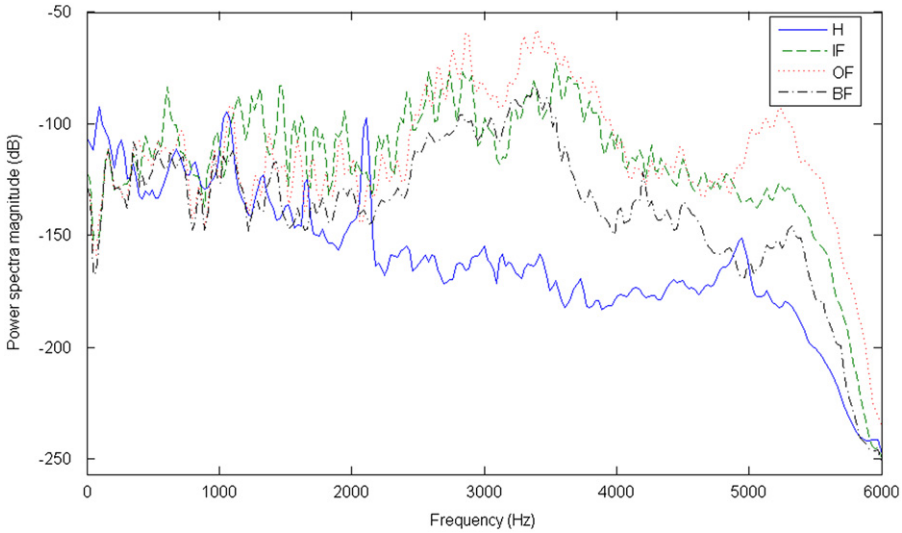

Fig. 16. PSD of test data1 (0.18 mm fault size, 1 HP load).

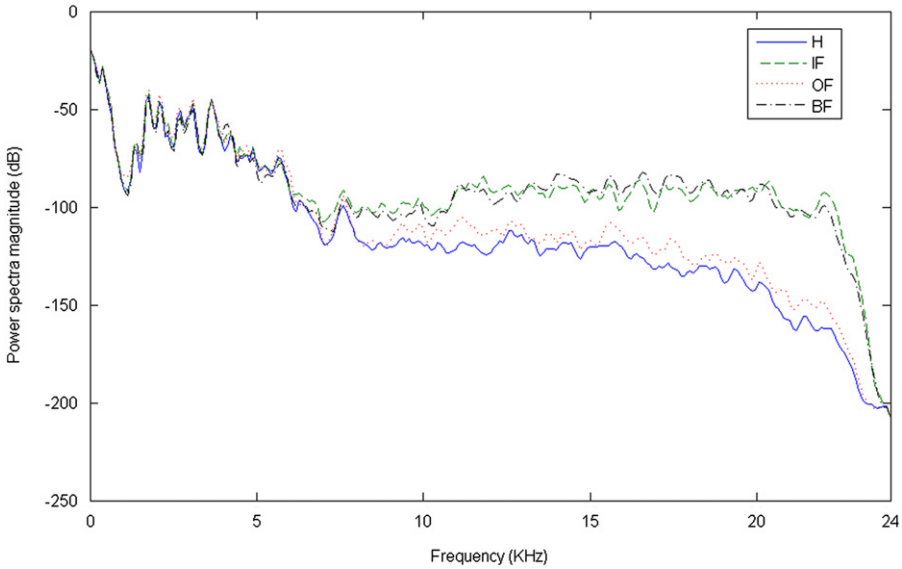

Fig. 17. PSD of test data2.

4.2.2. ANN performance results using test data2

The ANN used for test data2 is same as that of the test data1 analysis except that the network uses 3 inputs and 17 hidden neurons. 48 sets of 3 SV's (SV number 5, 6 and 7) are obtained for each bearing condition. The SV values are normalized (min–max normalization) in [0-1] range. 33 sets of input SV's in each condition were used for training the ANN and rest for testing. The overall classification accuracy of the BPNN obtained is 93.33%. The confusion matrix obtained for BPNN is shown as in Table 5. High overall classification accuracy and an excellent individual class accuracy is obtained (Table 5).

4.3. Comparison between power spectrum and singular spectrum

The power spectral density (PSD) is calculated based on Welch's method using a Hanning window [17] of length 512 samples. Figs. 16 and 17 give the PSD plots obtained for test data1 and test data2, respectively. For test data1(Fig. 16), in the lower frequency range (up to 2200 Hz), the power spectra has very minute masking due to the motor and dynamometer, and at a few places the PSD intersect for healthy and faulty cases. In the region from 2200 to 4000 Hz (Fig. 16), there is a huge increase in the level of power spectral magnitude due to the presence of defect. For test data2

Table 5
Confusion matrix for test data2—1st method.

	H	IF	OF	BF
H	13	0	0	2
IF	0	15	0	0
OF	0	0	15	2
BF	0	2	0	13

(Fig. 17), it is seen that up to 8 kHz, there is no change in the power spectra for the healthy and fault cases. It is known that a large part of the frequency region is masked by the gearbox vibration [17]. Only above 10 kHz, does the healthy power spectra start to separate from the faulty ones. Similar to Fig. 16, there is an increase in power spectra in the higher frequency region (>10 kHz) as a consequence of a defect in the bearing, and the power spectra can be separated for healthy and faulty cases in this region.

From the SS plot of test data1 (Figs. 8 and 9), SV of healthy and faulty conditions are separated in various regions of the singular spectrum. A similar trend is observed in the PSD of test data1. While for test data2, it is seen that the healthy SV and faulty SV coincide with each other in the initial part of the SS plots (Fig. 14). This is due to the masking of the signal due to gear vibration, as observed in Fig. 17. For test data2, the SV's of healthy and outer race fault, inner race fault and ball fault all have close values, which is in agreement with the PSD shown in Fig. 17. Also, for test data1, the SS plots are concurrent with the power spectra shown in Fig. 16. This shows the correspondence between the power spectrum and singular spectrum.

5. Energy feature and ANN based bearing fault diagnosis—2nd method

At the end of step3 of the SSA technique, the vibration signal is decomposed into L components—C_1, C_2, …, C_L. These principal components are time series of different frequency regions. The correlation between these components and the bearing fault is studied in this method. Based on the condition of the bearing, i.e., healthy or faulty, the resonant frequency changes. The energy of the signal also changes in the different frequency regions [20]. In the 1st method described earlier, appropriate SV were obtained. The components corresponding to these SV are selected. From the L components, m number of components are selected ($1<m<L$). The energy of the corresponding principal components is calculated as $E=\sum_{n=1}^{N}[x(n)]^2$, where, x is signal (principal component) and $n=1$ to N data points. Thus, energy $E=\{E_1,E_2,\ldots,E_m\}$, gives the energy distribution in the frequency domain for a bearing in healthy or different faulty conditions. Each energy feature is normalized by the total energy of the m-components. These are then given as inputs to an ANN for fault diagnosis.

5.1. Application to test data1—2nd method

From the analysis done in Section 4.1.1, a window length $L=10$ is found to be suitable. Hence, the signals are decomposed into 10 components. Fig. 18 shows the principal components corresponding to the decomposition of the vibration signal for an inner race fault with a window length $L=10$. The appropriate SV are SV number 1, 2, 5 and 6 (from 1st method). The components (C_1, C_2, C_5 and C_6) corresponding to the selected SV are chosen from which the normalized energy (E_1, E_2, E_5 and E_6) features are obtained, which are then given as input to ANN. Similar to 1st method, gaussian white noise is added to obtain noisy signals having SNR of $+10$ and -10 dB which are separately analyzed.

5.1.1. ANN performance results for test data1—2nd method

The ANN consists of 4 input nodes, 13 hidden nodes and 4 outputs. The architecture, number of training and testing datasets and other parameters are the same as used in the 1st method (Section 4.1.2). Table 6 gives the testing accuracy of the BPNN. Tables 7–9 show the confusion matrix of the output result for the three cases of signals analyzed. For the original and noisy signal of SNR$=+10$ dB, the BPNN gives higher classification. However for SNR$=-10$ dB, the classification drops down from 95% to 65%. This is due to the energy values for the different fault conditions in the three chosen components having close values in the very low SNR case. Also, the chosen value of L was based on the analysis done on the original signal, not the signal with added noise.

5.2. Application to test data2—2nd method

From the analysis done in Section 4.2.1, a suitable window length is 10 and the suitable SV numbers are 5, 6 and 7. The principal components (C_5, C_6 and C_7) corresponding to the selected SV numbers are obtained and their energy features (E_5, E_6 and E_7) are calculated. An ANN is then used for diagnosis with the normalized energy features as an input.

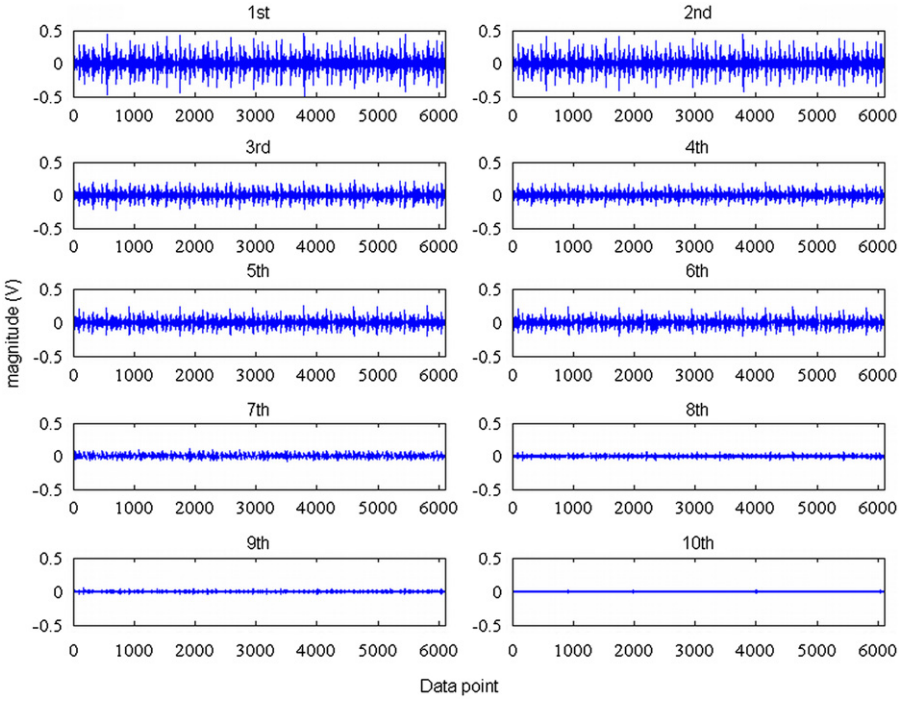

Fig. 18. SSA decomposition of bearing vibration signal for inner race fault (0.18 mm fault size, 1 HP load).

Table 6
BPNN performance for test data1—2nd method.

Signal	Hidden neuron	Classification accuracy (%)
test data1	13	95.14
Noisy test data1, SNR +10 dB	13	96.53
Noisy test data1, SNR −10 dB	13	64.58

Table 7
Confusion matrix for test data1—2nd method.

	H	IF	OF	BF
H	11	0	0	1
IF	0	47	1	0
OF	0	4	31	1
BF	0	0	0	48

Table 8
Confusion matrix for noisy test data1, SNR= +10 dB—2nd method.

	H	IF	OF	BF
H	11	0	0	1
IF	0	47	1	0
OF	0	1	34	1
BF	0	0	0	48

5.2.1. ANN performance results for test data2—2nd method

The ANN used is same as for test data1 (Section 5.1.1) except that 17 hidden neurons are used. Similar to the 1st method applied to test data2, the number of training and test datasets are the same. On testing, the classification accuracy is found to be 100%. Table 10 gives the confusion matrix obtained.

Table 9
Confusion matrix for noisy test data1, SNR= −10 dB—2nd method.

	H	IF	OF	BF
H	12	0	0	0
IF	0	23	13	12
OF	0	6	22	8
BF	0	4	8	36

Table 10
Confusion matrix for testdata2—2nd method.

	H	IF	OF	BF
H	15	0	0	0
IF	0	15	0	0
OF	0	0	15	0
BF	0	0	0	15

6. Comparison with published works

Table 11 shows the comparison of the proposed method with the recent published works [1,6–8,21–24,26] which are based on time series and data based fault diagnosis approaches. All these works uses the same bearing data [16] as used in this paper. The classification rate calculated depends on the number of datasets used for training and testing of the classifier. The works including the present one have used different numbers of datasets ranging from 63 to 940.

In [1,21,22], the algorithm is applied to data consisting of only one fault size at a particular load and used fewer samples for training and testing of the classifier. Time domain features used are affected by noise [1].

The difference histogram (DH) based method [6] is a simple method, but the size of histogram bin influences the classification and it has to be sufficient large for higher accuracy. The DH bins along with an ANN was used to discriminate between the faults, but the healthy condition was not included.

William and Hoffman [7] implemented a simple technique based on a zero crossing (ZC) feature. Similar to the SV feature, the ZC feature also represent the information present in the frequency domain. The feature calculation depends on the signal length which requires knowledge of the longest expected time duration between successive zero crossing (Tq).The classification accuracy was affected by the length and varies from 91.5% to 97.1%. Also, the BPNN when trained only with higher defect sizes and tested for other defect sizes gave good classification. The effects of noise on ZC features are not studied.

In [8], multi scale entropy as a feature was found to be suitable for bearing fault detection. However, in noisy conditions (white noise power is between [0.25, 1] times signal power) the classification accuracy comes down from 97.42% to 73.94%. The proposed method (method 1) works well even in case of noisy conditions.

A hidden Markov model (HMM) and Gaussian mixture model (GMM) were employed for fault detection by Nelwamondo et al. [22]. Sample size influenced the classification and an accuracy of 81% to 100% was obtained for different bearing locations. The average accuracy is still less than the one obtained in the proposed work using BPNN.

Complex calculation is used in [21] and separate GMM was constructed for each bearing fault type along with a Bayesian classifier used for fault detection [23].

In [24], HMM was used for bearing fault detection which was based on comparison of monitoring index with the reference vectors which are large in number. It required more training time and it was not fixed, as random values are used to fix the model parameters which are calculated by the Viterbi algorithm [24]. Over fitting of data occur in HMM due to usage of the Baum–Welch method [25].

Li and Zhang [26] used a non linear dimension reduction method known as supervised locally linear embedding projection (SLLEP), which transforms the original signal (assumed to be in a higher dimension) to a low dimension feature space. The performance of SLLEP is very sensitive to the regularization parameter which represents the generalization capacity of algorithm.

In all the above described methods a high classification rate is achieved, but the proposed algorithm does not suffer from the limitations of the above described methods such as parameter dependent performance, and features affected by noise and sample sizes. Of all the methods discussed so far, the method based on ZC and DH is simple. In the DH method, discriminating between healthy and faulty bearings was found to be trivial using a single histogram bin. The ZC feature calculation depends on the signal length which requires knowledge of longest expected time duration between successive zero crossings. Sample size affects DH and ZC based methods. It was shown in Section 4.1.4 that fault diagnosis using the proposed method is not affected by sample size variation. The effect of noise on classification and the performance in the presence of gearbox vibration is not studied in the ZC and DH methods. The proposed method works well even in presence

Table 11
A study between the proposed method and some of the recent published works which uses bearing dataset [16].

Method	Dataset (single point defect size width; load; signal length)	Training and testing datasets	Condition classified	Average testing accuracy (%)	Fault features
BPNN [1]	0.18 mm, 1 HP load; 6000 data points	Train—48 Test—32	H, IF, OF and BF	100	Kv, normal negative log-likelihood value
DH and BPNN [6]	0.18, 0.36, 0.53 mm; 0–3 HP load; 30000 data points	Both 144	healthy and faulty	92 (IF, OF and BF classification only)	6 histogram bins
ZC and BPNN [7]	0.18, 0.36, 0.53 mm; 0–3 HP load; 1024–4096 data points	–	H, IF, OF and BF	91.5–97.1	ZC count and ZC duration
Multi scale entropy and support vector machine (SVM) [8]	0.18, 0.36, 0.53, 0.71 mm; no-load; 2048 data points	Train—525 Test—237	H, IF, OF and BF	97.42	6 entropies at different time scale
Time domain feature and SVM [8]				75.64	Rms, skewness, kurtosis, max
Wavelet denoising and SVM [21]	0.18 mm, 2 HP load; 12,000 data points	Test—63	H, IF, OF and BF	100	2 weibull negative log-likelihood values
HMM and GMM [22]	0.18 mm	–	H, IF, OF and BF	HMM-99-100 GMM-94-99	Mel-frequency cepstral coefficients, multi-scale fractal dimension, kurtosis (Kv)
Reconstructed phase space, GMM and Bayesian classifier [23]	0.18, 0.36, 0.53 mm; 1 HP load; 1024 data points	Both 384	H, IF, OF and BF	81–100	4 GMM
k-means clustering and HMM [24]	0.18, 0.36, 0.53, 0.71 mm; 0–3 HP load.; 3000 data points	Both 940	H, IF, OF and BF	96	Power index and 17 reference vector's
SLLEP with Minimum distance, k-NN and SVM [26]	0.53 mm; 3 HP load; 1030 data points	Train—150 test—100	H, IF, OF and BF	96.64–99.67	low dimension feature space of original signal
SSA and BPNN [present work]	0.18, 0.36, 0.53, 0.71 mm; 0–3 HP load; 6100 data points	Train—336 test—144	H, IF, OF and BF	96.53–100 95–100%	4 singular values 3 energy features

of noise (test data1) and masking sources such as gears (test data2). Also, the proposed algorithm is evaluated using the data obtained for various fault sizes and at different loads which very few of the mentioned published works have tested.

7. Conclusions

The present work discussed a new technique for bearing fault feature extraction based on SSA. A new fault feature (singular value) was introduced for bearing fault detection. A new approach for selection of the principal components was employed. Two methods were discussed. In first method, singular values extracted from the bearing vibration signal are used as fault features which are given to ANN for automated fault diagnosis. This method is simple, detects faults even in a signal with a low SNR and also with masking sources such as gears with higher classification accuracy. In the second method, the correlation between the principal components corresponding to the selected SV and the bearing fault has been investigated. It shows that the features obtained from the principal components corresponding to SV numbers are useful for fault detection. An energy feature was used for fault identification. The correspondence between the singular spectrum and the power spectrum has also been shown. The proposed method is found to have more advantages over the existing time series and data based methods. The experimental results demonstrate that the proposed bearing fault diagnosis method is simple, noise tolerant and efficient.

Acknowledgments

The authors thank *Case Western Reserve University* for providing the Bearing Fault Datasets freely over the web. Mr. Bubathi Muruganatham thanks *Prof. Robert Bond Randall*, University of New South Wales, Australia for providing the bearing fault data and permission to use the data, figures and pictures related to test data2 used in the work. Research Fellowship provided to Mr. Bubathi Muruganatham by *Dept. of Atomic Energy, India* is gratefully acknowledged. The anonymous reviewers are thanked for their very good suggestions and comments provided. Special thanks to the first reviewer for correctly the writing.

Table A1
Bearing details.

Details	Test data1(SKF 6205)	Test data2 (Koyo 1205)
Ball diameter (mm)	7.94	7.12
Pitch diameter (mm)	39.04	38.5
Number of balls	9	12
Contact angle (°)	0	0

Appendix

See Table A1.

References

[1] B. Sreejith, A.K. Verma, A. Srividya, Fault diagnosis of rolling element bearing using time–domain features and neural networks, Proc. Third ICIIS, Kharagpur, India, December 2008, 1–6.
[2] D.C. Baillie, J. Mathew, A comparison of autoregressive modeling techniques for fault diagnosis of rolling element bearings, Mech. Syst. Sig. Process. 10 (1) (1966) 1–17.
[3] C. Junsheng, Y. Dejie, Y. Yu, A fault diagnosis approach for roller bearings based on EMD method and AR model, Mech. Syst. Sig. Process. 20 (2) (2006) 350–362.
[4] F. Li, L. Ye, G. Zhang, G. Meng, Bearing fault detection using higher-order statistics based ARMA, Key Eng. Mater. 347 (2007) 271–276.
[5] C.C. Wanga, Y. Kang, P.C. Shen, Y.P. Chang, Y.L. Chung, Applications of fault diagnosis in rotating machinery by using time series analysis with neural network, Expert Syst. Appl. 37 (2) (2010) 1696–1702.
[6] B.J.V. Wyk, M.A.V. Wyk, G. Qi, Difference histograms: a new tool for time series analysis applied to bearing fault diagnosis, Pattern Recognit. Lett. 30 (6) (2009) 595–599.
[7] P.E. William, M.W. Hoffman, Identification of bearing faults using time domain zero-crossings, Mech. Syst. Sig. Process. 25 (8) (2011) 3078–3088.
[8] G. Xiong, L. Zhang, H. Liu, et al., A comparative study on ApEn, SampEn and their fuzzy counterparts in a multiscale framework for feature extraction, Zhejiang Univ. Sci. A (Appl. Phys. Eng.) 11 (4) (2010) 270–279.
[9] R. Vautard, M. Ghil, singular spectrum analysis in nonlinear dynamics with applications to paleoclimatic time series, Physica D 35 (3) (1989) 395–424.
[10] F.J. Alonso, D.R. Salgado, Analysis of the structure of vibration signals for tool wear detection, Mech. Syst. Sig. Process. 22 (3) (2008) 735–748.
[11] W.J. Wang, J. Chen, X.K. Wu, Z.T. Wu, The application of some non-linear methods in rotating machinery fault diagnosis, Mech. Syst. Sig. Process. 15 (4) (2001) 697–705.
[12] B. Muruganatham, M.A. Sanjith, B. Krishna Kumar, S.A.V. Satya Murty, P. Swaminathan, Inner race bearing fault detection using singular spectrum analysis, Proceecidings of IEEE International Conference on Computer, Control and Communication Technologies (ICCCCT 2010), India, Oct. 2010, pp. 573–579.
[13] B. Kilundu, X. Chiementin, P. Dehombreux, Singular spectrum analysis for bearing defect detection, J. Vib. Acoust. 133 (5) (2011). 051007-1–051007-7.
[14] B. Li, M.Y. Chow, Y. Tipsuwan, J.C. Hung, Neural-network-based motor rolling bearing fault diagnosis, IEEE Trans. Ind. Electron. 47 (5) (2000) 1060–1069.
[15] N. Golyandina, V. Nekrutkin, A. Zhigljavsky, Analysis of Time Series Structure—SSA and Related Techniques, Chapman & Hall/CRC, Florida, 2001.
[16] Bearing data Centre, Case Western Reserve University, Available: 〈http://csegroups.case.edu/bearingdatacenter/home〉.
[17] N. Sawalhi, R.B. Randall, Simulating gear and bearing interactions in the presence of faults Part I. The combined gear bearing dynamic model and the simulation of localised bearing faults, Mech. Syst. Sig. Process. 22 (8) (2008) 1924–1951.
[18] R.B. Randall, Vibration Based Condition Monitoring: Industrial, Aerospace and Automotive Applications, Wiley, 2011.
[19] P.J. Dalianis, S.G. Tzafestas, G. Anthopoulos, A study of the generalization capability versus training in back propagation neural networks, Proc. Syst. Man Cybern. 4 (1993) 485–490.
[20] Y. Yu, Y. Dejie, C. Junsheng, A roller bearing fault diagnosis method based on EMD energy entropy and BPNN, J. Sound Vib. 294 (1-2) (2006) 269–277.
[21] S. Abbasion, A. Rafsanjani, A. Farshidianfar, N. Irani, Rolling element bearings multi-fault classification based on the wavelet denoising and support vector machine, Mech. Syst. Sig. Process. 21 (7) (2007) 2933–2945.
[22] 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,, Int. J. Innovative Comput. Appl., Inf. Control 2 (6) (2006) 1281–1299.
[23] G.F. Wang, Y.B. Li, Z.G. Luo, Fault classification of rolling bearing based on reconstructed phase space and Gaussian mixture model, J. Sound Vib. 323 (3-5) (2009) 1077–1089.
[24] Tony Boutros, Ming Liang, Detection and diagnosis of bearing and cutting tool faults using hidden Markov models, Mech. Syst. Sig. Process. 25 (6) (2011) 2102–2124.
[25] Z. Garamani, An introduction to hidden Markov models and Bayesian networks, Int. J. Pattern Recognit Artif Intell. 15 (1) (2001) 9–42.
[26] B. Li, Y. Zhang, Supervised locally linear embedding projection (SLLEP) for machinery fault diagnosis, Mech. Syst. Sig. Process. 25 (8) (2011) 3125–3134.