ISA Transactions 54 (2015) 193–206

Contents lists available at ScienceDirect

ISA Transactions

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

Application of higher order spectral features and support vector machines for bearing faults classification

Lotfi Saidi *, Jaouher Ben Ali, Farhat Fnaiech

University of Tunis, Tunis National Higher School of Engineering (ENSIT), Laboratory of Signal Image and Energy Mastery (SIME), 5 Avenue Taha Hussein, P.O. Box 56, 1008 Tunis, Tunisia

ARTICLE INFO

Article history:
Received 24 March 2014
Received in revised form
26 July 2014
Accepted 18 August 2014
Available online 3 October 2014
This paper was recommended for publication by Dr. Jeff Pieper

Keywords:
Bi-spectrum analysis
Feature extraction
Bearing defects
Principal component analysis
Support vector machine
Vibration analysis
Receiver operating characteristic

ABSTRACT

Condition monitoring and fault diagnosis of rolling element bearings timely and accurately are very important to ensure the reliability of rotating machinery. This paper presents a novel pattern classification approach for bearings diagnostics, which combines the higher order spectra analysis features and support vector machine classifier. The use of non-linear features motivated by the higher order spectra has been reported to be a promising approach to analyze the non-linear and non-Gaussian characteristics of the mechanical vibration signals. The vibration bi-spectrum (third order spectrum) patterns are extracted as the feature vectors presenting different bearing faults. The extracted bi-spectrum features are subjected to principal component analysis for dimensionality reduction. These principal components were fed to support vector machine to distinguish four kinds of bearing faults covering different levels of severity for each fault type, which were measured in the experimental test bench running under different working conditions. In order to find the optimal parameters for the multi-class support vector machine model, a grid-search method in combination with 10-fold cross-validation has been used. Based on the correct classification of bearing patterns in the test set, in each fold the performance measures are computed. The average of these performance measures is computed to report the overall performance of the support vector machine classifier. In addition, in fault detection problems, the performance of a detection algorithm usually depends on the trade-off between robustness and sensitivity. The sensitivity and robustness of the proposed method are explored by running a series of experiments. A receiver operating characteristic (ROC) curve made the results more convincing. The results indicated that the proposed method can reliably identify different fault patterns of rolling element bearings based on vibration signals.

1. Introduction

As a critical component, rolling element bearings (REBs) are widely used in induction motors (IMs). Bearings are required to run with high reliability, and the occurrences of faults may lead to fatal breakdowns of IMs [1,3]. It is well known that the bearing defects (BDs) represent one of the most common sources of faults in IMs (about 40–50%) [1–4]. Therefore, it is significant to accurately detect and diagnose the existence and severity of the faults occurring in the REB.

The overall procedure for a fault diagnosis scheme can be stated in several steps: data acquisition, signal processing, feature extraction, feature reduction and diagnostics (classifiers) [1,2]. The latter two are the priority. After feature extraction, an intelligent classifier is needed. The critical step in this process is to extract reliable features which are representative of the condition of the REB. Many advanced techniques have been employed to detect and extract such features.

Since REB motion is often rotational, periodicity is regarded as an important characteristic of the vibration signal which arises from bearings. As a result, many techniques are based on or related to periodic analysis, such as spectral analysis [4,5], wavelet analysis [6], time–frequency distribution [7], cyclostationary approach [8,9] and empirical mode decomposition [10]. However, information about bearing condition is often hided by the presence of a strong noise, including random noise and uninteresting vibration associated components [8,9].

For several decades both the first and second-order statistics, such as mean, variance, autocorrelation and power spectrum [11,12] are popular signal processing tools and have been used extensively for mechanical vibration analysis. However, they are

* Corresponding author. Tel.: +216 22 16 9567; fax: +216 71391166.
E-mail addresses: lotfi.saidi@esstt.rnu.tn (L. Saidi), benalijaouher@yahoo.fr (J. Ben Ali), fnaiech@ieee.org (F. Fnaiech).

http://dx.doi.org/10.1016/j.isatra.2014.08.007
0019-0578/

subject to describe linear and Gaussian processes. Or, in practice, most of the situations with nonlinear and non-Gaussian behaviors can be conveniently studied using advanced signal processing techniques, such as higher order statistics (HOS).

Bearing vibration often exhibits nonlinear coupling mechanisms and consequently bi-spectrum analysis offers promise for the analysis of such signal by exploiting two particular characteristics. First, the bi-spectrum is theoretically zero for Gaussian noise and flat for non-Gaussian white noise [8–13], consequently bi-spectrum analysis is insensitive to random noise. Second, the bi-spectrum peaks only at those frequency pairs corresponding to those related components with frequency and phase [11,12].

Motivated by the HOS advantages, that bi-spectrum analysis is an appropriate method for processing complicated vibration signals, this paper presents a novel pattern classification approach for BDs diagnostic, which combines the bi-spectrum analysis features and support vector machine (SVM) classifier. The vibration bi-spectrum patterns are extracted as the feature vectors presenting different BDs. The SVM algorithm is adopted to distinguish four kinds of BDs conditions namely: healthy bearing (HB), inner race fault (IRF), outer race fault (ORF) as well as ball fault (BF) signals which were measured in the experimental test bench running under different working conditions.

However it is expected that a combination of HOS analysis techniques with SVM analysis should provide a very useful tool for condition monitoring and fault diagnosis of IMs. This kind of research based on bearing bi-spectrum features has not been reported in any publications yet.

Relatively, only a few papers discussed the application of HOS analysis (HOSA) in IMs fault diagnosis. Even though many researchers have performed using non-linear features motivated by the HOS but they do not perform the diagnosis and identification of BDs. For these reasons this paper tries to contribute the use of vibration bi-spectrum (third order spectrum) analysis for multi-fault BDs diagnosis.

As mentioned above, for the HOS the stationary and non-stationary Gaussian noise that may be corrupting the signal of interest is removed. It can make up the drawbacks of power spectrum and has strong noise suppression ability. The bi-spectrum, a third-order statistic powerful technique has the advantage of less computing than the other HOSA, but it gives full play of the advantages of the HOSA.

This is due firstly to the difficulty of representing a spectrum of order greater than 5 (result in a space of dimension at least 5) and computation complexity. Comparison made between the bi-spectrum method and benchmark methods using other HOSA can be found in [11,12].

Some literatures about the bi-spectrum diagnosis of BDs have been published, in [8] the twice slice of cyclic bi-spectrum (CBS) is used for the detection of BDs. This method inherits the advantage of the CBS, and is not sensitive to noise. In addition, twice slice of CBS is more direct and clear to show the features for diagnosis purpose.

In [14], an amplitude-modulation detector using bi-coherence (normalized bi-spectrum) is developed to detect incipient defects in the outer race BDs before their characteristic fault frequencies become significant in the power spectrum. In [9], the combination of cyclostationary and HOSA methods is presented. This combination exhibits better results, in terms of bearing status identification.

In [15], Saidi et al. presented two HOS techniques, namely the power spectrum and the slices of bi-spectrum used for the analysis of induction machine stator current leading to the detection of electrical failures. Experimental signals have been analyzed highlighting that bi-spectrum results show their superiority in the accurate detection of rotor broken bars.

Motivated by the HOS advantages, this paper presents a novel pattern classification approach for BDs diagnostic, which combine the bi-spectrum analysis features and support vector machine (SVM) classifier. The vibration bi-spectrum patterns are extracted as the feature vectors presenting different BDs. The SVM algorithm is adopted to distinguish four kinds of bearing condition signals which are measured in the experimental test bench running under different working conditions; with the removal of varying speed and load (or torque) dependence. In fact, a small variation in speed produces deviations that a diagnostic algorithm may erroneously detect as damage, thus providing a false alarm. Encouraging results have been obtained related to the ability of this feature in removing speed and load dependence in order to avoid any bias in data interpretation and identification.

It is generally acknowledged that SVM has a fine performance in solving small sample size, nonlinear and high dimensional pattern recognition problems. It solves satisfactorily the over-fitting and local optimal solution problem of artificial neural networks (ANNs). ANN has fast development in the past several years [17–19], which impels its development in fault diagnosis research and application. Generally, ANN has some disadvantages such as identification ability difference, structure identification difficulty, no standard method to determine the structure of the network, local convergence and other issues. SVM can better solve these problems [16,20], so it achieves better decision accuracy in special cases because of the maximized decision boundary, and it is efficient for large dataset and real time analysis [20].

In IMs fault diagnosis, some researchers have employed SVM as a tool for classification of faults. For example, in [21] a wind turbine gearbox fault diagnosis method based on diagonal spectrum and clustering binary tree SVM has been proposed. The fault features are extracted by the diagonal spectrum. Then the clustering binary tree is used to classify the fault categories.

In [5], an original fault signature based on an improved combination of Hilbert and Park transforms has been proposed to diagnose electrical and mechanical failures in IMs. SVM is used to classify electrical and mechanical faults in IMs based on Hilbret and Park features.

In [22] SVM is used along with continuous wavelet transform. The results obtained are hoped to set up a base for condition monitoring technique of IMs which will be simple, fast and overcome the limitations of traditional techniques. The main drawback of this technique is the problem to determine the optimal parameters for the wavelet filter.

In [23] fault detection of REB using SVM and ANN is presented. The vibration signals are collected to analyze the bearing condition, which has four faults: inner race fault, outer race fault, ball fault and cage fault. Statistical features based on moments and cumulants are used and the optimal features are selected using genetic algorithm. In the classification process, SVM is employed using radial basic kernel function (RBF) with constant kernel parameter.

The remainder of this paper is organized as follows. The basic principles of the bi-spectrum and its derived features are discussed in Section 2. Section 3 presents the proposed method, including the description of the data pre-processing procedure, the selected feature set and the proposed multi-SVM based classifier. Section 4 presents the experimental results including a description of the experimental test bench. Section 5 gives a discussion and comparison with previous works and the paper concludes in Section 6.

2. Bi-spectrum and features derived from the bi-spectrum

2.1. Brief description of bi-spectrum

The bi-spectrum belongs to the class of HOS, used to represent the interaction relationship between frequencies content in a

signal. HOS provides higher order moments and nonlinear combinations of the higher order moments called cumulants [11,12]. Thus, HOS consist of moment and cumulant spectra. An overview of the theory on HOS can be found in [11,12]. Bi-spectral techniques show their effectiveness in quadratic phase coupling peak detection, and because it is a third order moment function, noise background is eliminated in the estimation procedure [12].

The bi-spectrum can be estimated directly from the discrete Fourier transform (DFT) of M realization of sampled version of the non-Gaussian third-order stationary and ergodic IMs vibration signal $x(n)$ as follows [11,12]:

$$\hat{B}(f_1,f_2)=\frac{1}{M}\sum_{i=1}^{M}X_i(f_1)X_i(f_2)X_i^*(f_1+f_2)\approx E\{X(f_1)X(f_2)X^*(f_1+f_2)\} \quad (1)$$

where f_1, f_2 are the frequency indices, X^* denotes the complex conjugate of X and $X(f)$ is the DFT of the discrete signal $x(n)$ and $E\{.\}$ is an average over an ensemble of realizations of a random signal. For deterministic signals, the relationship holds without an expectation operation with the third order correlation being a time-average [12].

The bi-spectrum of real signals contains a lot of symmetries and therefore redundant information. Redundant information already exists in the discrete Fourier transform. Despite these symmetries, here, all the information is displayed. However, focus is concentrated on the positive frequencies and of course only on the principal domain $\Im$ given by [11,12]

$$\Im=\{(f_1,f_2):0\le f_2\le f_1\le f_e/2, f_2\le -2f_1+f_e\} \quad (2)$$

where f_e is the sampling frequency.

In order to minimize the computational cost, the direct method [11,12] has been used to estimate and derive bi-spectrum features from vibration signals. The HOS features are fed to the SVM for automatic BDs identification.

In the next section, the composition of the extracted and selected features dataset is further discussed.

2.2. Features extraction and reduction

2.2.1. Features extraction

In order to characterize the frequency information within REB data, this paper proposes to use the derived bi-spectrum features from the vibration bi-spectrum, which are listed in Table 1 [24]. These features can be useful in discriminating between processes with similar power spectra but different third order statistics [24]. In addition, in order to characterize the regularity and irregularity of the signal from bi-spectrum plots, bi-spectral entropies were also derived (see Table 1) [24]. Then the features vector is expressed as

$$T=[F_1,F_2,F_3,P_1,P_2,P_e,WCOB_1,WCOB_2] \quad (3)$$

In the next section, the principal component analysis (PCA) is introduced to eliminate correlations between features and reduce the dimensionality of the original feature vectors. Note that PCA does not "discard" or "retain" any of pre-defined features. It mixes all of them (by weighted sums) to find orthogonal directions of maximum variance.

2.2.2. Features reduction

Principal component analysis (PCA) involves a mathematical procedure that transforms a number of (possibly) correlated variables into a (smaller) number of uncorrelated variables called principal components (PCs) [25,26]. All the PCs are orthogonal to each other, so there is no redundant information. The PCs as a whole form an orthogonal basis for the space of the data. Thus, the first PC consist of highest variability, the second PC consist of next highest variability and so on for other directions. The first few components are kept and other with less variability are discarded. As shown in Fig. 14, in the proposed work, the first 5 PCs are used to feed to the SVM classifier. The step by step procedure of PCA is provided below:

Let us say we have N observations of M-dimensional data.

- *Step 1:* To perform PCA we should first compute the covariance matrix $\Sigma\in R^M\times^M$; $\Sigma=E\{(x-\bar{x})(x-\bar{x})^T\}$, where x is the given signal, $\bar{x}$ is its mean vector, and $E\{.\}$ is an average over an ensemble of N samples.
- *Step 2:* Find matrix Λ of eigenvectors and diagonal elements of matrix D as eigenvalues of covariance matrix Σ as given by, $\Lambda^{-1}\Sigma\Lambda=D$.
- *Step 3:* Sort the eigenvectors of PCs in the decreasing order of importance of eigenvalues in D.
- *Step 4:* Project the data into the directions of sorted eigenvectors by taking the dot product between the given data and eigenvectors.
- *Step 5:* Choose the first few principal components depending on containment of a given percentage of variability (like 95% or 98% depending on the problem).

Table 1
Derived bi-spectrum features from the bearing vibration signals.

Features	Expression	Notes
Sum of logarithmic amplitudes of the bi-spectrum	$F_1=\sum_{f_1,f_2\in\Im}\log(\lvert B(f_1,f_2)\rvert)$	
Sum of logarithmic amplitudes of diagonal elements in the bi-spectrum	$F_2=\sum_{f_k\in\Im}\log(\lvert B(f_k,f_k)\rvert)$	
First-order spectral moment of amplitudes of diagonal elements in the bi-spectrum	$F_3=\sum_{f_k\in\Im}k\log(\lvert B(f_k,f_k)\rvert)$	
Normalized bi-spectral entropy	$P_1=-\sum_n p_n\log(p_n)$	where $p_n=\frac{\lvert B(f_1,f_2)\rvert}{\sum_{f_1,f_2\in\Im}\lvert B(f_1,f_2)\rvert}$
Normalized bi-spectral squared entropy	$P_2=-\sum_n q_n\log(q_n)$	where $q_n=\frac{\lvert B(f_1,f_2)\rvert^2}{\sum_{f_1,f_2\in\Im}\lvert B(f_1,f_2)\rvert^2}$
Bi-spectrum phase entropy	$P_e=\sum_n p\psi_n\log(p\psi_n)$	where $p\psi_n=\frac{1}{L}\sum_{\Im}1(\phi(B(f_1,f_2))\in\Psi_n$ and $\psi_n=\left\{\phi\middle\vert -\pi+\frac{2\pi n}{N}\le\phi<-\pi+\frac{2\pi(n+1)}{N}\right\}$, $n=0,1,\ldots,N-1$
Weighted center of bi-spectrum (WCOB)	$WCOB_1=\frac{\sum_{\Im}iB(i,j)}{\sum_{\Im}B(i,j)}$ $WCOB_2=\frac{\sum_{\Im}jB(i,j)}{\sum_{\Im}B(i,j)}$	L is the number of points within the non-redundant region given by Eq. (2), ϕ refers to the phase angle of the bi-spectrum, and $1(\bullet)$ is an indicator function which gives a value of 1 when the phase angle ϕ is within the range of bin Ψ_n in Eq. (10). where i and j are the frequency bin indices in the non-redundant region expressed in Eq. (2).

The eigenvalues variable will then be an array where each element describes the amount of variance accounted for by the corresponding PC. We should see that the first eigenvalue will be the largest, and they will rapidly decrease (Fig. 14). PCs with small corresponding eigenvalues are unlikely to be useful, since the variance of the data in those dimensions is so small. However, since we are using the PCs for BDs classification task, we cannot really be sure that any particular number of PCs is optimal; the variance of a feature does not necessarily tell us anything about how useful it will be for classification. An alternative to choosing PCs with the scree plot shown in Fig. 14 is just to try classification with various numbers of PCs and see what the best number is empirically.

3. Bearings multi-fault classification proposed method

3.1. Bearing defects (BDs) signatures

Failure surveys [1,2] by the Electric Power Research Institute (EPRI) indicate that IMs bearing-related faults are about 40% among the most frequent faults in IMs [1–3]. As shown in Fig. 1, the bearings consist mainly of the outer and inner raceways, the balls, and the cage. BDs can be classified into two classes: single-point defects and generalized roughness. Single-point defects are localized and classified into, outer raceway defect; inner raceway defect; as well as ball defect.

Generalized roughness is a type of fault where the condition of a bearing surface has degraded considerably over a large area and become rough, irregular, or deformed.

These faults may enhance vibration and noise level [27]. Moreover, there are internal operating stresses caused by vibration, eccentricity, and bearing current. In addition, bearings can also be affected by other external causes such as:

- contamination and corrosion;
- lack of lubrication causing heating and abrasion;
- defect of bearing's mounting, by improperly forcing the bearing onto the shaft or in the IM's stand.

The single-point defect may be seen by fault frequencies appearing in the machine vibration spectrum record. The frequencies at which these components occur are predictable and depend on the surface on which the bearing contains the fault. Therefore, there are different fault frequency characteristics associated with each component among the four parts of the bearing [27].

Rollers or balls rolling over a local fault in the bearing produce a series of force impacts. If the rotational speed of the races is constant, the repetition rate of the impacts is determined solely by the geometry of the bearing. The repetition rates are denoted bearing frequencies, for example, BPFO (Ball Passing Frequency Outer Race), BPFI (Ball Passing Frequency Inner Race), and BFF (Ball Fault Frequency) are frequently used. Their mathematical equations are as follows [27]:

$$\mathrm{BPFO} = \frac{f_r}{2} N_b \left(1 - \frac{D_b \cos\beta}{D_c}\right) \tag{4}$$

$$\mathrm{BPFI} = \frac{f_r}{2} N_b \left(1 + \frac{D_b \cos\beta}{D_c}\right) \tag{5}$$

$$\mathrm{BFF} = \frac{f_r}{2} \frac{D_c}{D_b} \left[1 - \left(\frac{D_b \cos\beta}{D_c}\right)^2\right] \tag{6}$$

where f_r is the rotor shaft frequency, N_b is the number of rolling elements, D_b is the ball diameter, D_c is the pitch diameter, and β is the ball contact angle.

Table 2 shows the parameters of the bearing used in our experimental test bench (Figs. 5, and 6), taken from the data sheet [28].

3.2. Nature of defective REBs vibration response

3.2.1. The characteristics of rolling bearing signals

A large number of models have been used to describe the dynamic behavior of REBs under different types of defects. According to the traditional approach [27], when rolling elements of bearing pass the defect location, wide band impulses are generated. And those impulses will then excite some of the vibrational modes of the bearing and its supporting structure. The excitation will result in the sensed vibration signals (waveforms) different in either the overall vibration level or the vibration magnitude distribution [27].

Table 2
Drive end bearing information.

Bearing		SKF-6205
Geometry size [mm]	Outside diameter	51.81
	Inside diameter	24.9
	N	9
	D_c	201.9
	D_b	38.86
	$\cos\beta$	0.9
Defect frequencies multiple of running speed [Hz]	Inner ring	5.4152
	Outer ring	3.5848
	Cage train	0.3983
	Rolling element	4.7135

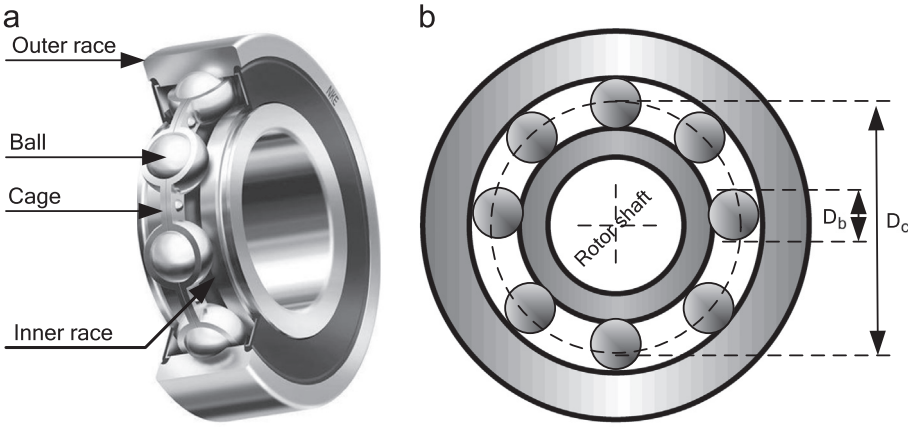

Fig. 1. (a) Exploded view and (b) geometry of REB.

Measured vibration signals consist of two parts: $y(t)=x(t)+n(t)$, where $x(t)$ is the defect-induced impulse responses and $n(t)$ is the background noise, including vibration signals generated by other components, such as rotor unbalance and gear meshing. Because of the structure and the mode of operation of REBs, $x(t)$ has distinct features as follows [13], [27]:

- *Wide frequency range:* BDs usually start as small pits or spalls, and give sharp impulses in the early stages covering a very wide frequency range.
- *Small energy:* The energy created by the BD is very small. A band has to found where the bearing signal dominates over other components.

Because the vibration signals generated by a defective REB have the characteristics mentioned above, it is difficult to identify their faults through a simple classical frequency analysis.

3.2.2. REB signals model

In the paper, the effectiveness of bi-spectrum is illustrated on a synthetic signal of a REB with an ORF. This type of fault was chosen because it is a relatively simple phenomenon to simulate while being often found in rotor machinery condition monitoring. REB with ORF during operation generates a series of periodic shock-pulses whose repetition rate depends on its dimensions and rotational speed of the shaft with which the bearing is assembled. Shock-pulses are generated each time rolling elements strike the defected surface of the bearing outer race and consequently, excite resonances of the structure between the fault location and the vibration sensor [27].

Assuming constant rotational speed and load of the bearing, the vibration signal generated by a REB with defected outer race can be modeled as [8,9,27]

$$x(t)=\sum_i \omega\left(t-i\frac{1}{\mathrm{BPFO}}-\tau_i\right)+n(t), \tag{7}$$

where $\omega(t)$ is the waveform generated by a single impact (related to resonance frequencies of the system), τ_i is an independent and identically distributed random variable and $n(t)$ is an additive background random noise. It now has to be stated that τ_i introduces influence of the rolling elements slips into the model.

In order to exhibit the effect of the stochastic nature of certain critical parameters like slip or varying load, on the vibration spectra, a typical example is considered. The simulated signal generated by Eq. (7), corresponds to the typical response of a bearing with an ORF. The shaft rotation speed f_r is 29.16 Hz.

The characteristic bearing defect frequency BPFO is equal to 3.58 times the shaft rotation speed, leading to an estimation of the BPFO around 104 Hz. The excited natural frequency f_c of the system is assumed to be equal to 4.5 kHz. The signal consists of 2048 samples and the sampling rate is equal to 20 kHz. Fig. 2 illustrates the waveform and spectra of the simulated signals with and without noise effect. In perfect agreement to the expected results of the model defined in Eq. (7), a set of harmonics of the BPFO of 104 Hz appears for zero slip. Their amplitudes are maximum in the frequency band around the resonance frequency of 4.5 kHz. This frequency does not appear in the spectrum, since it does not necessarily coincide to a harmonic of the BPFO frequency.

Fig. 2 displays the simulated signal in time domain (Fig. 2(a), and (b)), it can be found that a serious of spike pulses appear in the time domain waveforms, the interval of the spike pulses of the defect-induced impulses (1/BPFO= 9.6 ms), and in frequency domain (Fig. 2(c), and (d)), with and without additive Gaussian white noise (AGWN), respectively. As seen from the spectrum in Fig. 2(d), no distinct spectral lines can be recognized. Information that could be obtained by examination of the spectrum is only the frequency characteristic of band-pass stationary Gaussian noise used in generated signal. For rotor-machinery vibration signals, natural resonances of the object usually manifest themselves in the same manner. However, for vibration-based condition monitoring purposes, information about high-frequency resonances has relatively limited value. The main interest is in the phenomena that cause excitation of observed resonances. As seen in Fig. 2(d), band-limited noise is amplitude modulated by a component of BPFO frequency. This component contains major information about operation of mechanical source (e.g. REB) required for evaluation of observed objects technical condition. Due to this fact, the usage of an advanced statistical procedure is necessary.

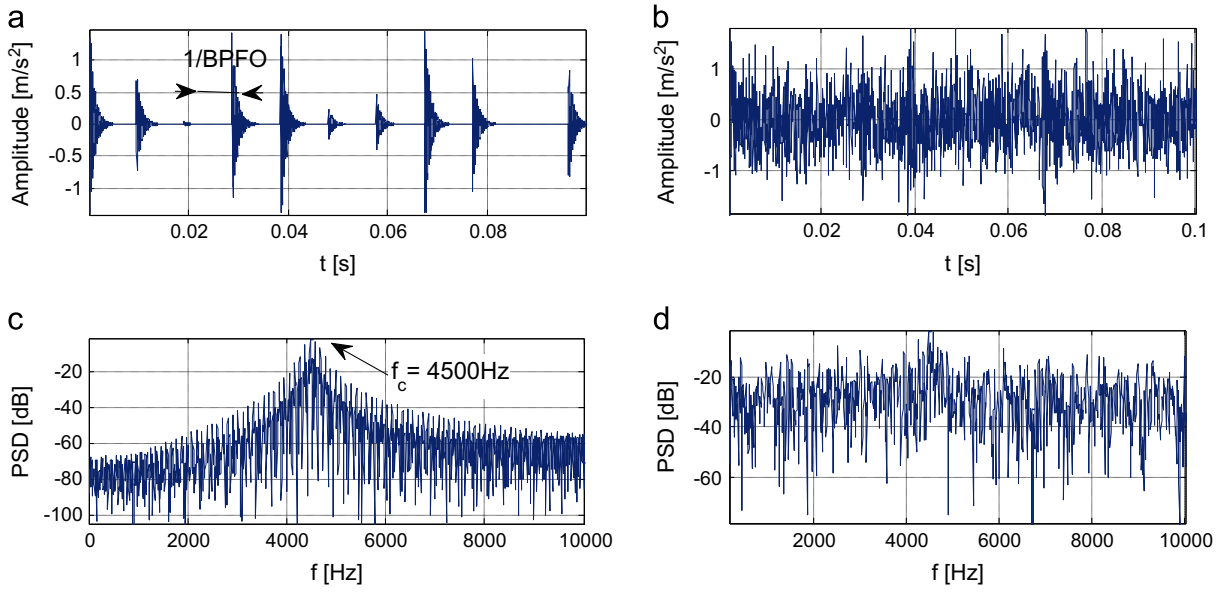

Fig. 2. Simulated short segment of vibration response data showing outer race impacts with and without noise effect: (a, b) waveforms and (c, d) power spectral density (PSD), the magnitude squared of its Fourier transform.

An enlarged presentation of the bi-spectrum vibration of the simulated ORF is shown in Fig. 3. The space between the peaks in the resonant frequency band is about 104 Hz, which is equal to the BPFO. In higher frequency band, the peaks around the highest peak are (f_c, f_c), ($f_c \pm$ BPFO, f_c).

Vertical and horizontal peaks spaced by 104 Hz represent the repetition rate of excitation impulses. Detected modulation components have the strongest magnitude around the (4.5, 4.5) kHz carrier frequency (i.e. center frequency of the random vibration). Such distribution is particularly valuable for mechanical signals as different faults may be observed in different resonant frequency ranges. For example, in the same machinery, advanced gearbox faults might excite different resonances than local bearing faults [27].

3.3. Basic principle of support vector machine (SVM)

3.3.1. Binary SVM

SVM is basically a binary classifier developed by Vapnik [16] in 1995. Unlike other classifiers, SVM minimizes the structural risk rather than empirical risk. During the training of SVM, it maximizes the distance from patterns to the class separating hyperplane [16]. Generally the patterns are not linearly separable; therefore non-linear kernel transformation is performed. There are various kernels that can be used during SVM training. Some of them are quadratic, polynomial and radial basis function (RBF) kernels. In this study RBF kernel is used for kernel transformation. The SVM with a RBF kernel function has two training parameters: regularizing parameter C which controls over fitting of the model and the kernel parameter γ which controls the degree of non-linearity of the model. C and γ are determined using a "grid search" approach [30]. For interested reader, a complete description about SVM is available in [16]. Furthermore, SVMs were originally designed for binary classification. They have then been extended to handle multi-class problems. The idea is to decompose the problem into many binary-class problems and then combine them to obtain the prediction. This is the subject of the next subsection (Section 3.3.2).

3.3.2. Multi-class SVM

As mentioned above SVMs were originally designed for binary (2 classes) classification [20,22–26]. In binary classification, the class labels can take only two values: 1 and −1. In the real problem, however, we deal more than two classes for examples: in condition monitoring of IMs there are several classes such as mechanical unbalance, misalignment, different load conditions, bearing faults, gear faults, etc. Therefore, multi-class SVM is obtained by decomposing the multi-class problem into several number of binary class problems. The multi-class classification methods will be discussed in [20].

Two different approaches are taken into account: one-against-all (OAA) and one-against-one (OAO) [20,29]. In the first one the i-th SVM is trained with all the examples in the j-th class with positive labels and all the other examples with negative labels, while in the latter one each classifier is trained on data from two classes. Here, SVM-OAA is chosen to classify different bearing faults.

We do not make comparisons SVM-OAA strategy with other popular approaches like SVM-OAO in this paper, because of the following reasons:

- Benchmark comparisons on multiclass SVM approaches already exist in literature [29].
- It has been concluded [29] that SVM-OAA is as accurate as any other approach, assuming that all underlying binary SVMs are well tuned.

The training procedure and choice of SVM parameters for training are very important for classification. In this work the process of optimizing the SVM parameters with the cross-validation method (CV) is adopted. CV is one of the most important tools because it gives us an honest assessment of the true accuracy of our system. In other words, the CV process provides a much more accurate picture of our system's true accuracy. Detail information of this strategy has been clearly explained in [20,29,30].

For training and testing SVM-OAA, bi-spectrum features relative to the studied faults are extracted in order to develop the

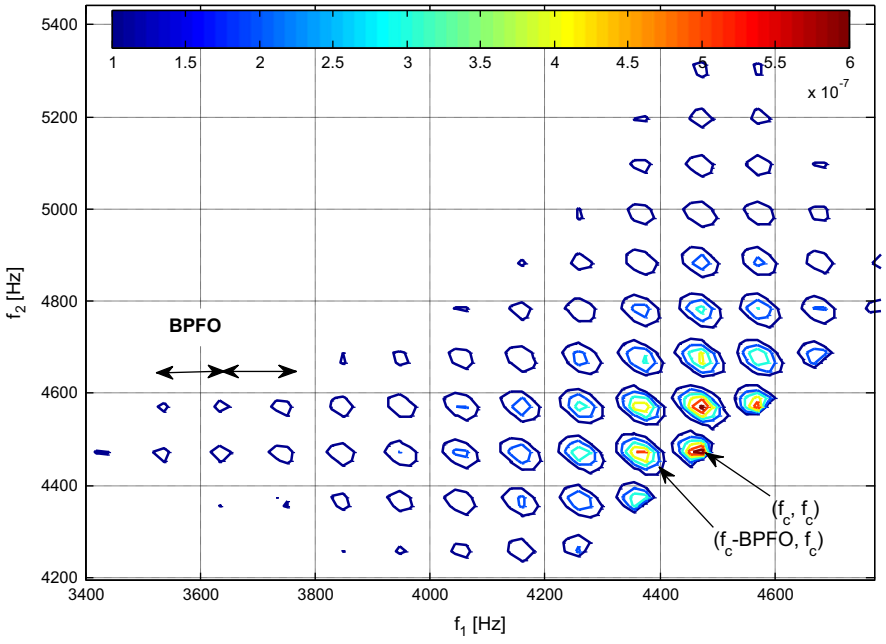

Fig. 3. Enlarged bi-spectrum of simulated outer raceway vibration signal in the bandwidth frequency between 3400 Hz and 4800 Hz.

input vector, which is necessary for the training and the test of the BDs classification. This is the next phase of our research work.

3.4. BDs classification based on SVM

The diagram of fault diagnosis scheme is presented in Fig. 4. The procedure of the proposed method can be summarize as follows:

- *Step 1:* Data acquisition is carried out through the designed test rig.
- *Step 2:* Signal processing is performed using bi-spectrum analysis.
- *Step 3:* Bi-spectrum features calculation from the vibration signals.
- *Step 4:* Bi-spectrum features reduction using the PCA method.
- *Step* 5: Classification process for fault diagnosis by SVM-OAA based on multi-faults classification. Since only four bearing conditions need to be identified in this paper, just three SVM classifiers need to be designed, as indicated in Fig. 4, and Table 3. For SVM1, define healthy bearing (HB) condition as $y=+1$, and the remaining 3 other conditions, as another class, identified as -1; thus the HB could be separated from other conditions by SVM1. Then define the condition with outer race fault (ORF) as $y=+1$ and the other conditions as $y=-1$ for SVM2; thus the ORF could be separated from other conditions by SVM2. Similarly, the ball fault (BF) could be separated from inner race fault (IRF) by SVM3. Then they become a multi-class fault diagnosis system as shown in Fig. 4. Note that all the three SVMs adopt RBF as their kernel function. We choose the RBF for the SVM classifier, since previous studies have shown that it has the best performance in pattern recognition tasks. As well, the RBF kernel is a better choice than other kernels like polynomial kernel because it has lesser hyper-parameters and so the problem becomes less computationally intensive.

SVM-OAA algorithm is used; furthermore, some parameters are predefined for this classifier such as the regularizing parameter C and the kernel parameter γ are set to 100 and 0.5, respectively, which were selected based on the CV method [30]. We always need to cross-validate our experiments in order to guarantee a correct scientific approach. For instance, if we do not cross-validate, the results which we read (such as accuracy) may be highly biased by our test set [16,28]. To avoid such problems just divide our initial dataset in k-fold CV. This will guarantee that any biases to the data are visible to us. As suggested in [28,30], we can perform this with even further division: 10-fold CV, means dividing our data set in 10 parts, then training in 6 and testing in 4 for example, then repeating the process until we have tested in all parts.

4. Experimental results: application of the proposed method in BDs

In this section, the proposed diagnosis approach described above is applied to fault diagnosis of BDs classification.

4.1. CWRU bearing benchmark data: description of experimental set-up and data acquisition

The bearing data center of Case Western Reserve University (CWRU) published bearing signal data on-line for researchers to validate new theories and techniques [28]. All data are annotated with bearing geometric, operating condition and fault information. Fig. 6 shows the test stand from which the test data are collected. As shown in Fig. 5 and 6, the test rig consists of a 2 hp three-phase induction motor (left), a torque transducer (middle), and a dynamometer-load (right). The transducer is used to collect speed and horsepower data. The load is controlled so that the desired torque load levels can be achieved. The test bearing supports the motor shaft at the drive end. Single point faults with fault diameters of 0.1778 mm, 0.3556 mm, and 0.5334 mm were introduced into the test bearing using electro-discharge machining. Accelerometers were placed at the 12 o'clock position when the

Table 3
Binary encoding for each BD.

	SVM1	SVM2	SVM3
HB	+1		
ORF	−1	+1	
IRF	−1	−1	+1
BF	−1	−1	−1

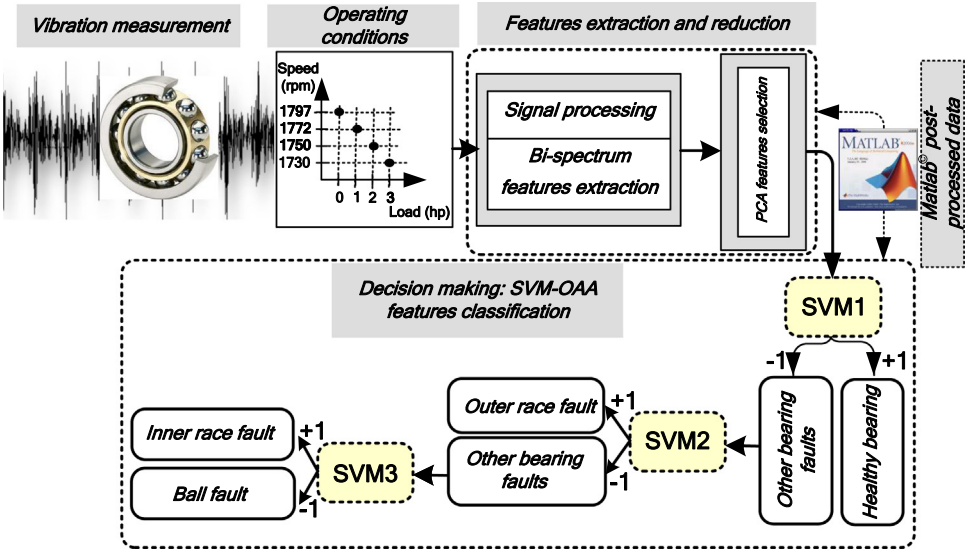

Fig. 4. Flowchart of the proposed diagnostic methodology for multi-faults diagnosis SVM-OAA classifier. The most significant features to distinguish between the predefined classes (HB, IRF, ORF and BF) are selected by means of PCA.

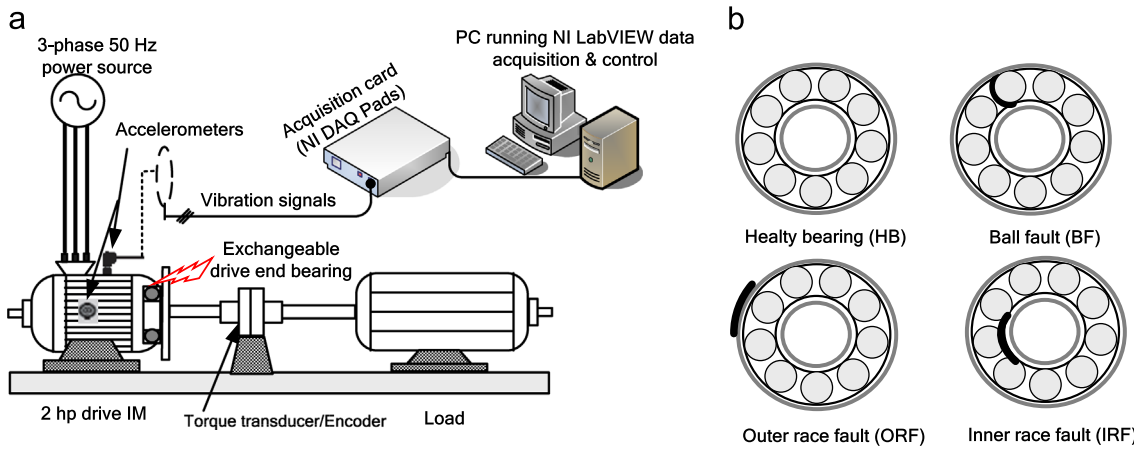

Fig. 5. (a) Schematic of the experimental test rig composed of a 2 hp motor (left), a torque transducer/encoder (center), load (right), and control electronics. The test bearings support the motor shaft. (b) A series of bearing components with faults induced in them indicated in bold line.

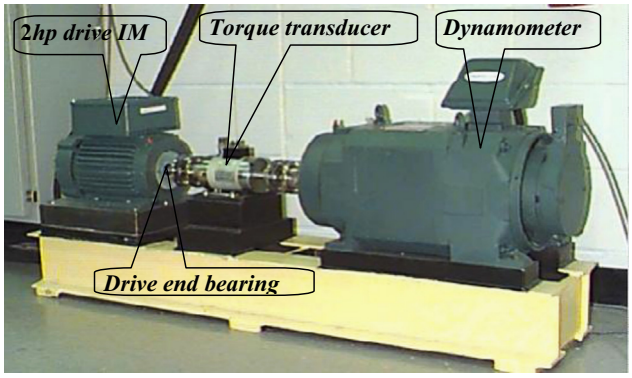

Fig. 6. A photograph of experimental test bench for BDs [28].

Table 4
Description of bearing data set analyzed under rated conditions.

Bearing condition		Fault specifications		Training/testing samples		
		Diameter [mm]	Depth [mm]	Label of classification	The number of training samples	The number of testing samples
HB		0	0	1	20	12
IRF	IRF_{17}	0.1778	0.279	2	78	50
	IRF_{35}	0.3556	0.279			
	IRF_{53}	0.5334	0.279			
	IRF_{71}	0.7112	0.279			
BF	BF_{17}	0.1778	0.279	3	78	50
	BF_{35}	0.3556	0.279			
	BF_{53}	0.5334	0.279			
	BF_{71}	0.7112	0.279			
ORF	ORF_{17}	0.1778	0.279	4	58	38
	ORF_{71}	0.7112	0.279			
	ORF_{53}	0.5334	0.279			

faults were at the inner raceway and the rolling element and at the 6 o'clock position when the fault was at the outer raceway. For better Signal-Noise-Ratio, we use signals from 6 o'clock for analyses in this paper. Vibration data are collected using a 16 channels acquisition system at a sampling frequency of 12 kHz per each one for different bearing conditions, using accelerometers, which are attached to the housing with magnetic bases. A data recorder is equipped with low-pass filters at the input stage to avoid anti-aliasing. The geometry and defect frequencies of the two type bearings are listed in Table 4. Speed and power data were collected using the torque transducer/encoder and were recorded

Table 5
Confusion matrix for the multiclass SVM-OAA resulting from the evaluation of the whole dataset.

		Predicted			
		HB	IRF	ORF	BF
Actual	HB	12	0	1	0
	IRF	1	51	0	1
	ORF	2	1	36	0
	BF	0	0	0	53

Table 6
Confusion matrix for the multiclass SVM-OAA resulting from the evaluation of the reduced dataset.

		Predicted			
		HB	IRF	ORF	BF
Actual	HB	8	0	0	0
	IRF	1	31	0	0
	ORF	1	1	22	0
	BF	0	0	0	32

by hand, the rotating speed of the shaft varied from 1721 to 1729 revolutions per minute (rpm). Vibration signals generated by this test stand are dominated by bearing signals, since there are no gears and shafts are healthy. BDs cover inner and outer races, and ball faults. The deep groove ball bearing 6205-2RS JEM SKF is used in the tests. For interested reader, more equipment and instrumentation details about this test rig can be found from [28].

The bearing data set is obtained from the experimental test rig under four different operating conditions as presented in Fig. 6(b), where 10 identical bearings (SKF-6205) have been used covering the different levels of severity for each fault type: (a) HB (4 measurements), (b) with an ORF (12 measurements), (c) with an IRF (16 measurements), and (d) with a BF (16 measurements).

The SVMs training experiments are conducted on a data set (48 vibration signals) under different fault severities. The classification results are shown in Tables 5, 6 and 7.

The time domain vibration signals of different bearing conditions under the same operating condition (1750 rpm speed and 1 hp load) are presented in Fig. 7. As shown in Fig. 7, the different bearings conditions exhibit different vibration signal characteristics, such as signal variation ratio from the mean, magnitude and others. Therefore, depending on the considered scenarios to distinguish, some statistical frequency features will be more

Table 7
The testing accuracy* for four different bearing conditions using SVM-OAA.

Bearing conditions	SVM including all features Classification accuracy [%]		SVM including selected features Classification accuracy [%]	
	Training	Testing	Training	Testing
HB	99.16	95.00	100	100
IRF	100	96.00	100	96.78
ORF	98.23	89.47	97.78	91.16
BF	100	100	100	100
Average	99.34	95.11	99.44	96.98

* Accuracy is computed based on the confusion matrix provided in the paper.

significant than others. The first few components are kept and other with less variability are discarded (Figs. 11, 12, and 13).

For the healthy and outer raceway defective bearing signals, the power spectral densities (PSDs) are shown in Fig. 8(a) and (b), respectively. The resonance frequency band with the highest energy should be used for bi-spectral analysis, and the central frequency f_c is equal to 3480 Hz, and is obtained from the highest spectral line in Fig. 8(b). Nevertheless, the fault signature is still hardly found in the power spectrum. Moreover, it can be seen from the power spectrum in Fig. 8(b), though the defect-induced frequency, the noise components around BPFO have higher power than it, which may make the diagnosis result be hardly convinced. Thus, we select the frequency band from 3.5 kHz to 4.5 kHz for bispectral analysis.

Vibration bi-spectra are presented in Fig. 9 in non-redundant region of bi-spectrum. Fig. 9 shows vibration bi-spectra in non-redundant region of bi-spectrum, of different BDs types at the same operating condition (1750 rpm speed and 1 hp load). (a) HB. (b) IRF; 0.1778 mm. (c) ORF; 0.1778 mm. (d) BF; 0.1778 mm. With the HB vibration bi-spectrum features vector is given as follows:

$T=$ [0.4769; 0.1657; 0.1080; 3.4928; 0.1006; -4.7530×10^{4}; 52.5810; -54.1058]. Where the bi-spectrum plots are depicted in 2D space of coordinates (f_1, f_2) and the amplitude will be represented in dB scale by a color-bar.

To show the effect of the speed and load on the bi-spectrum features distribution we take as an example the normalized bi-spectrum entropy values for the 16 acquisitions in each condition, which are evaluated, and presented in Fig. 10. This plot shows how this parameter is influenced by the speed and the torque both for healthy and damaged cases and it increases with higher speeds. Moreover, it can be noticed that in low speed cases this parameter value for damaged bearing is almost near to the healthy one when it reaches the highest speed.

The bi-spectrum is computed for each of the four type of bearing conditions including various combinations of speed and load, in the principal domain $\Im$. After feature calculation, normalized bi-spectral entropy (P_1), the normalized bi-spectral squared entropy (P_2) and bi-spectrum phase entropy (P_e) are plotted in Fig. 11 in order to know the structure of the original features. Fig. 11 represents the original features which are not well clustered and have disorder structure. Plotting original feature parameters indicates the necessity of preprocessing of the original features to make them separable and ready for classification. Disorder structure of original features tends to decrease the performance of classifier if it is directly processed in classifier.

To avoid this disadvantage, PCA was proposed to extract and to reduce the feature dimensionality based on eigenvalues of covariance matrix. Therefore, the first five principal components have been selected to replace the original feature vector as shown in Fig. 12.

Fig. 13 shows the feature reduction in component analysis based on eigenvalues of covariance matrix. The projection result is illustrated in Fig. 13. The axes of the projection plane correspond to the maximum variance directions in the initial space. As shown in Fig. 13, the number of features is reduced from 8 to 5.

4.2. Training and test vectors

The experimental vibration data, from four bearing fault types which are HB, with IRF, with ORF and with BF, as well as different levels of severity for each fault type, is listed in Table 1. The training and testing of the SVM model with real time data sets was implemented with the help of LIBSVM software [30]. The total databases comprised 384; (48×8) original features distributed as follows: 32 HB features, 128 IRF features, 96 ORF features, and 128 BF features. Among these 384 data samples, 230 (60%) samples are chosen randomly as training data, and the rest 154 (40%) as testing data. The number of features is reduced to 240; (48×5) selected features were divided into two sets: one for training (containing 60% of the samples), and the other for test (containing 40% of the samples).

After the SVM is trained with 144 features, its performance has been tested with the 96 remaining. The performance of the SVM-OAA is validated by calculating the following performance measure for the train set and test set separately.

The classification accuracy (CA): ratio between total number of correctly classified test samples to total number of test samples is calculated as

$$\text{CA}\,[\%]=\frac{\text{number of correctly classified samples}}{\text{total number of samples in dataset}}\times100 \tag{8}$$

The classification results of all classifiers in the training and testing processes are presented in Table 7. In the training process, all the SVM-OAA classifiers achieve an average accuracy of 99.165% and 99.331% in the whole dataset and the reduced dataset, respectively, some of them without any misclassification out of 230 samples (144 samples) of training data for all bearing features. This indicates that the classifiers are well trained and can be applied for diagnosing BDs. However, in the testing process these classifiers are validated against the test data, the average accuracy is about 96% and 95.416% for the reduced and original features, respectively. The miss-classifications are due to the overlap of machine condition features.

The confusion matrix is a useful tool for analyzing how well the classifier can recognize tuples of different groups, which contains information about actual and predicted classifications done by a classification system.

From the confusion matrix of the SVM in Tables 5 and 6, one can note that SVM finds it difficult to discriminate among ORF, HB, and IRF. Misclassification of 11.76% brings down the diagnostic ability of the SVM-OAA; however, the overall classification accuracy is reasonably good. It can be seen that the BF presented the highest accuracy of 100%.

From the test results shown in Tables 5, 6 and 7 it can be seen that the SVM-OAA classifiers recognize the defect samples effectively, especially for the HB signals, IRF signals and the ORF signals. The recognition results of SVM are ideal because of its high accuracy and a good generalization capability, when the average classification efficiency is close to 96.98%, it is reasonably good.

The results shown in the paper seem to be from a single run of simulation/experiment. The results generated from the data do not have statistical significance and may not be able to generalize. A series of experiments should be conducted to support the author's claims. In addition, in fault detection problems, the performance of a detection algorithm usually depends on the trade-off between robustness and sensitivity. The sensitivity and robustness of the

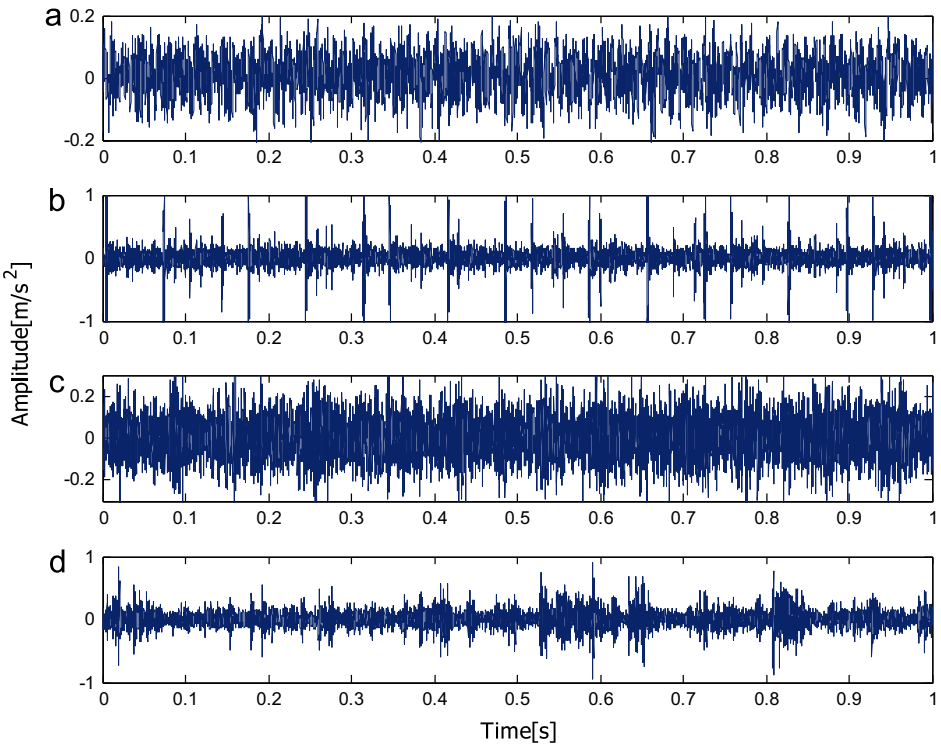

Fig. 7. An example of characteristics vibration signals under different bearing status at the same operating condition (1750 rpm speed and 1 hp load). (a) HB. (b) IRF; 0.1778 mm. (c) ORF; 0.1778 mm. (d) BF; 0.1778 mm.

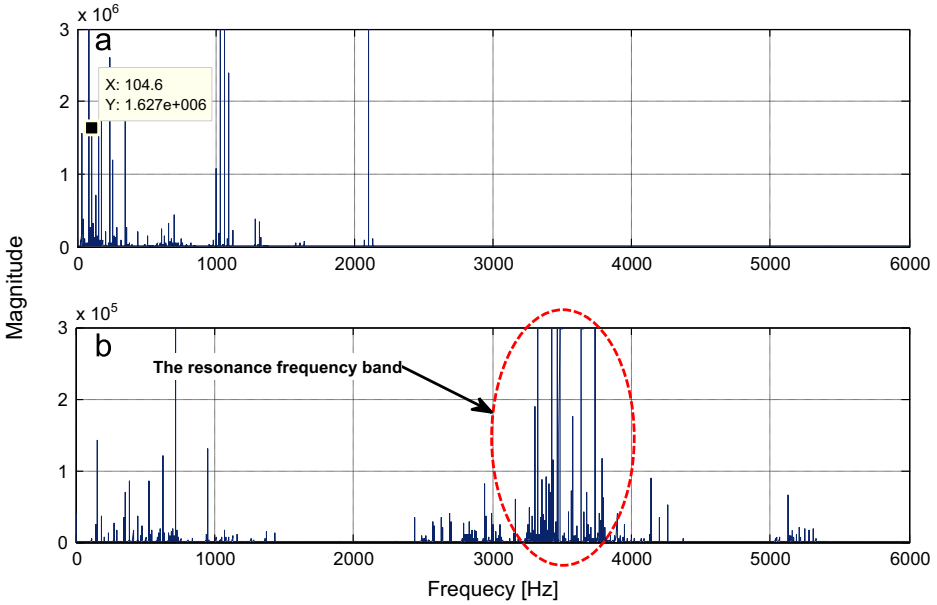

Fig. 8. Frequency band selection for vibration PSD analysis of a: (a) HB and (b) ORF.

proposed method need to be explored by running a series of experiments. A receiver operating characteristic (ROC) curve will make the results more convincing.

4.3. Performance measure of classification: need for ROC analysis

In order to compute the performance of the developed diagnosis methodology, the 10-fold CV method is used. Based on the correct classification of patterns in the test set, in each fold the performance measures are computed. The average of these performance measures is computed to report the overall performance of the SVM-OAA classifier.

In each fold, based on correct classification of test patterns, following parameters are defined:

- *True Positive (TP):* It is the number of actually abnormal test patterns which were classified as abnormal by our automated classifiers.
- *True Negative (TP):* It is the number of normal patterns in the test set which were classified into normal class by the automated classifiers.

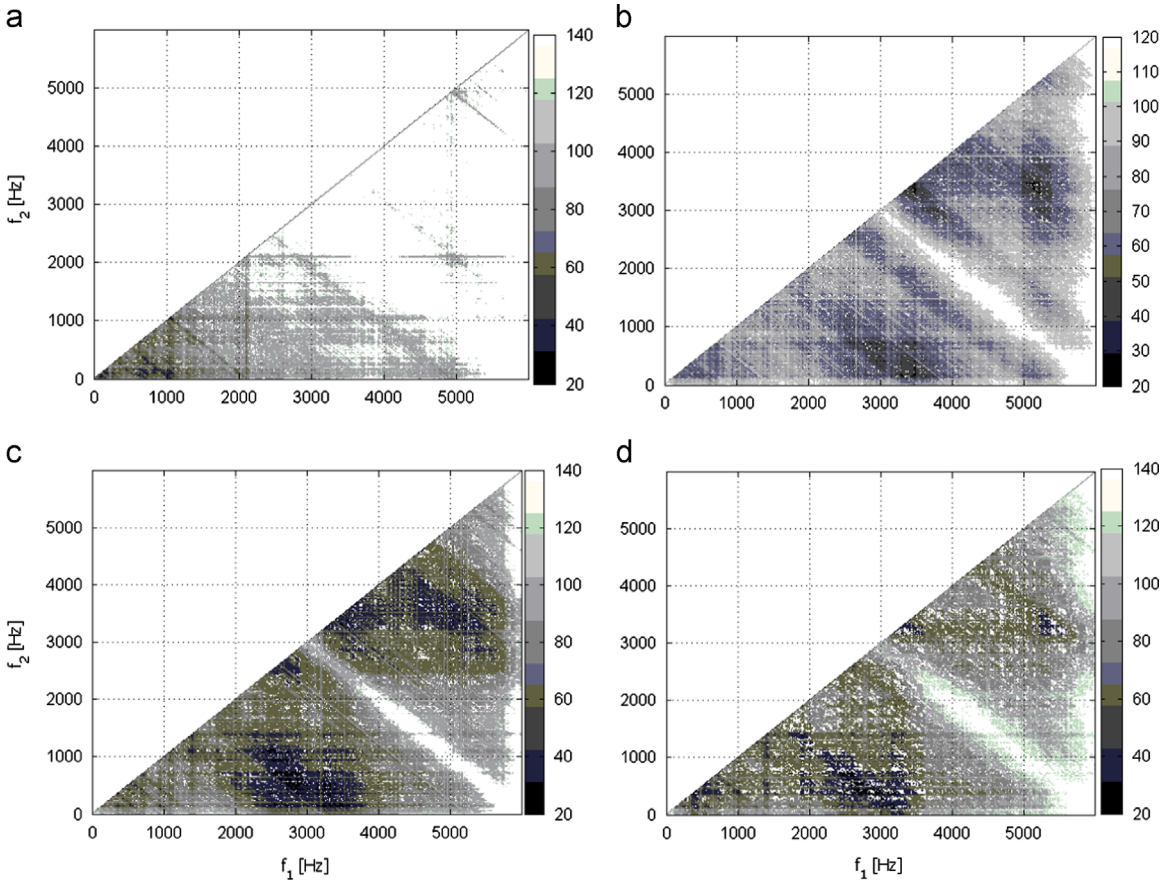

Fig. 9. Vibration bi-spectra of different BDs types at the same operating condition (1750 rpm speed and 1 hp load). (a) HB. (b) IRF; 0.1778 mm. (c) ORF; 0.1778 mm. (d) BF; 0.1778 mm. The magnitude coming out of page and is indicated by a color-bar.

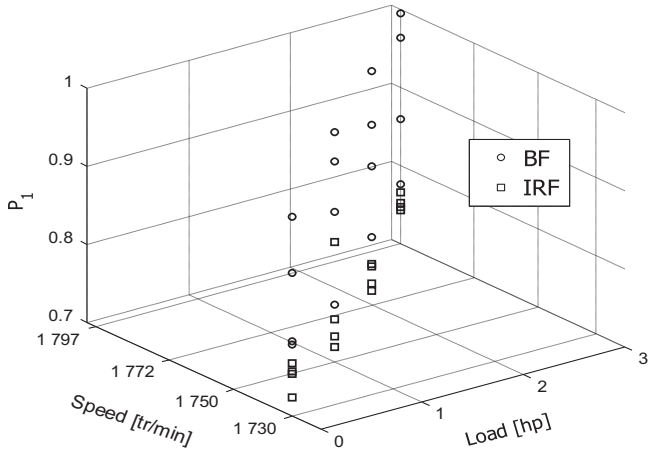

Fig. 10. The effect of speed and load on statistical normalized bi-spectral entropy parameters distribution for two bearing conditions (IRF and BF).

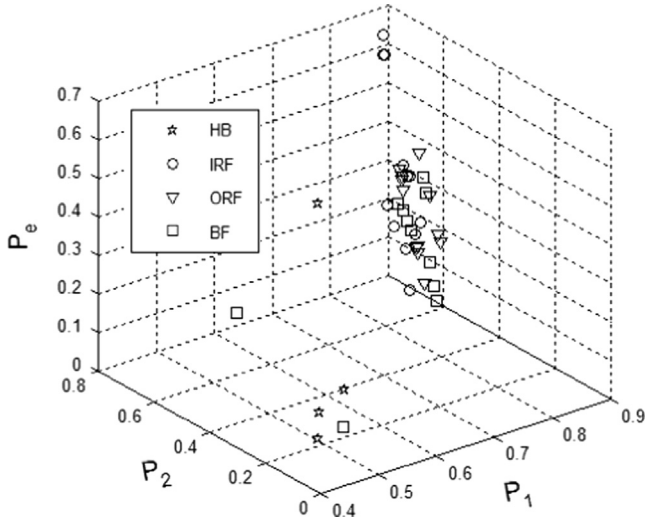

Fig. 11. Distribution of "P_1", "P_2" and "P_e" features of original data (belonging to the four bearing classes).

- *False Positive (FP):* It is the number of normal test patterns classified as abnormal by the automated classifiers.
- *False Negative (FN):* It is the number of abnormal test patterns which are classified as normal by the automated classifiers.

Based on TP, TN, FP and FN, the sensitivity (Sn), and specificity (Sp) are defined as follows:

$$\mathrm{Sn} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \tag{9}$$

$$\mathrm{Sp} = \frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FP}} \tag{10}$$

These performance measures were computed in each fold, and were averaged to compute the average performance.

To examine SVM-OAA classification measure's performance, ROC curves are often the only valid method of evaluation. An ROC curves is a detection performance evaluation methodology,

and demonstrates how effectively a certain detector can separate two groups in a quantitative manner. An ROC curve shows the trade-off between the probability of detection or true positives rate, also called sensitivity and recall versus the probability of false alarm or false positives rate.

Fig. 14 shows average performance of our proposed SVM-OAA classifier. It can be seen that SVM-OAA provides highest accuracy of 99.50%, sensitivity of 100% and specificity of 99.2%.

5. Discussion and comparison with some previous works

Classification accuracy is much improved for all BDs conditions by the use of the bi-spectrum features extraction and selection method. The classification accuracy values are all greater than 92%. Based on these results, it can be concluded that the bi-spectrum featured extraction method effectively improves classification performance for the given BDs classification problem.

Table 8 provides a summary of the studies on automated identification of IMs faults. Tran et al. [26] proposed various experts approaches, such as SVM-OAO, SVM-OAA, adaptive neuro-fuzzy inference system (ANFIS) and relevance vector machine (RVM) for classification of 4 IMs failures (normal, misalignment, bearing fault, mass unbalance) and reported an accuracy of 97.5%, 95%, 82.5% and 97.5%, respectively. They used generalized discriminant components analysis (GDA) of the histogram features (mean, standard deviation, skewness, kurtosis, energy, and entropy) derived to thermal images. Widodo et al.

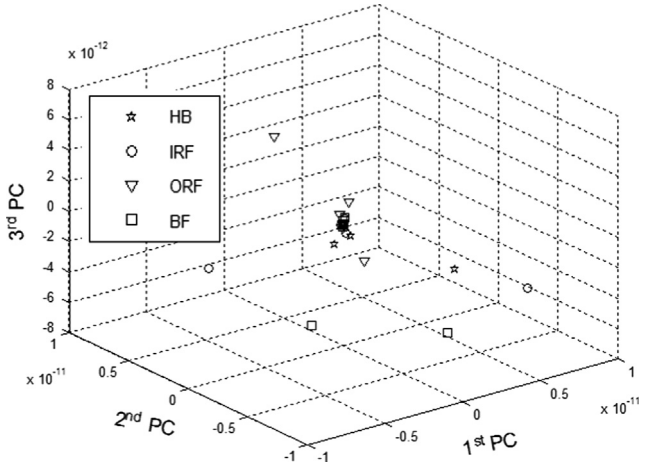

Fig. 12. Original features obtained from PCA.

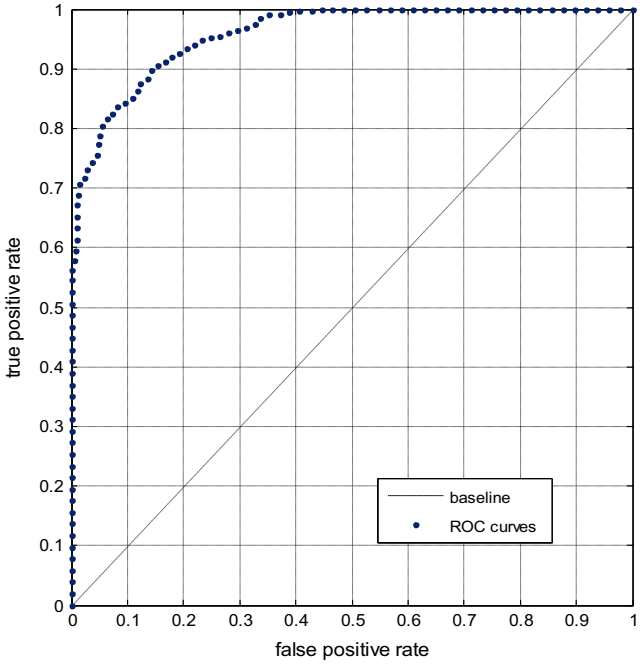

Fig. 14. ROC curve for SVM-OAA performance evaluation classifier.

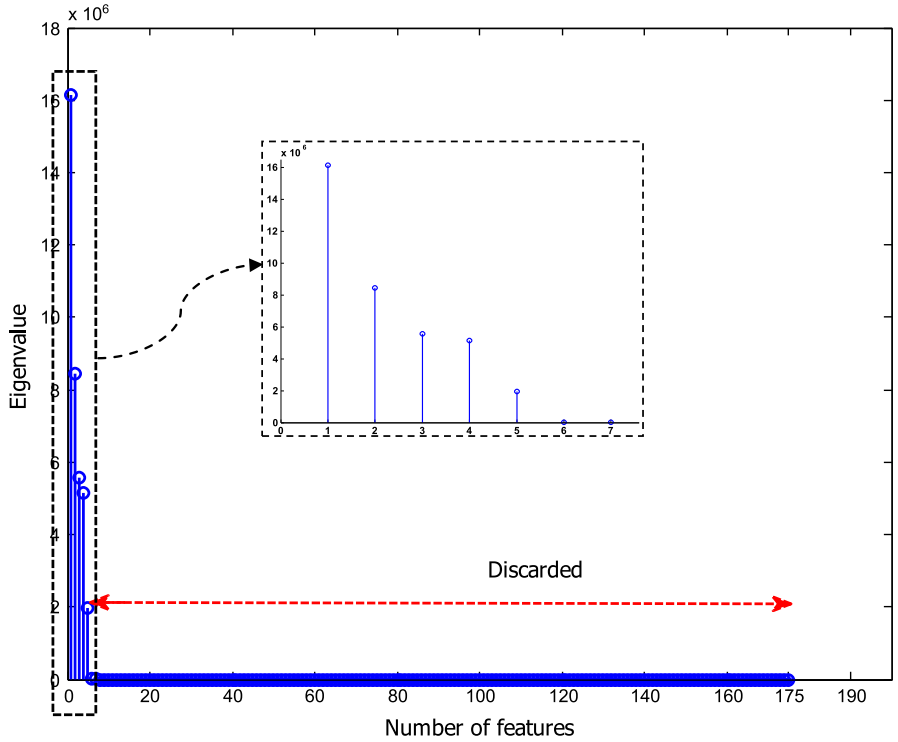

Fig. 13. Eigenvalues of covariance matrix for feature reduction.

Table 8
Comparison of this paper with some previous research.

Literature	Features	Classifier	# of classes	Accuracy[%]
[26]	Histogram features (mean, standard deviation, skewness, kurtosis, energy, and entropy) derived from thermal images + generalized discriminant analysis (GDA).	SVM-OAO, SVM-OAA, adaptive neuro-fuzzy inference system (ANFIS) and relevance vector machine (RVM)	**4**: Normal, misalignment, bearing fault, mass unbalance.	97.5, 95, 82.5 and 97.5, respectively.
[12]	Frequency domain based vibration energy features.	SVM-OAA	**3**: Normal, inner race and outer race BDs	100
[27]	The start-up transient current and discrete wavelet transform with nonlinear feature reduction using kernel PCA and kernel independent component analysis (ICA).	SVM-OAA	**7**: Bowder rotor, broken rotor bars, faulty bearing, eccentricity, phase unbalance, normal condition, mass unbalance.	78.75 using PCA, 80.95 using ICA, 76.19 using kernel PCA, 83.33 using kernel ICA
[28]	Time domain, and frequency domain features derived from vibration and three phases current signals.	ICA and SVM (OAO, OAA).	**7**: Broken rotor bar, bowed rotor, faulty bearing, rotor unbalance, eccentricity, phase unbalance, normal condition	99.97
[7]	Hilbert modulus current space vector (HMCSV) and Hilbert phase current space vector (HPCSV).	SVM-OAA	**4**: Normal condition, electrical fault, air-gap eccentricity fault, outer raceway bearing fault	97.5
[29]	Time domain and frequency domain vibration signals + PCA.	Artificial immunization algorithm SVM, (AIA-SVM)	**14**: Gear damage, structure resonance, rotor radial touch friction, rotor axial touch friction, shaft crack, bearing damage, etc.	97
Proposed methodology	Features derived from HOSA of vibration signals	PCA + SVM-OAA	**6**: HB, IRF, ORF, and BF	96.98

[31] used the start-up transient current and discrete wavelet transform with nonlinear feature reduction using kernel PCA and kernel independent component analysis (ICA) and classified IM faults into seven classes (Bowder rotor, broken rotor bars, faulty bearing, eccentricity, phase unbalance, normal condition, and mass unbalance) with an accuracy of 97.5%. Ben Salem et al. [5] used Hilbert modulus current space vector (HMCSV) and Hilbert phase current space vector (HPCSV) and classified three-phase IM faults into normal, electrical fault, air-gap eccentricity fault, outer raceway bearing fault, using the SVM-OAA method with an accuracy of 97.5%.

In conclusion, evaluation of our method versus previous researches given in Table 6 shows that our experiment demonstrates bi-spectrum features, containing non-Gaussian and nonlinear information of the IM stator current signal, and suggests that HOS has good potential for improving BDs diagnosis. The bi-spectrum feature with the multiclass OAA-SVM classifier produces a good classification performance: the average CA result is 96.98%, and the best individual CA result is 100%. This combination of the feature and classifier is promising for a high-accuracy diagnosis of BDs.

6. Conclusion and future works

In this work, an effort is made to characterize and classify four different REB classes depending on their bi-spectrum features. A flexible test bed was created and vibration signals were taken for four different bearing conditions at different speed and torque conditions. The best CA of four conditions of bearing was 100% on test set for the OAA-SVM. The results show that the OAA-SVM is a strong technique for fault diagnosis of rotating machinery. Also, the results demonstrate the ability of the proposed combination "HOS-SVM" model in diagnosing REB failures.

There are several possible directions for future works based on this approach to fault detection. First, this work has application in fault detection for stochastic processes other than bearing fault data, and should be applied to detect multi-faults in induction motor such as bearing and eccentricities failures. This approach should be applied to other synthetic data sets and simulation models to further test its capacity for fault detection in clean or noisy data. Finally, this method stands to be implemented as real-time fault detection metric in induction machine systems for active monitoring.

References

[1] Heng A, Zhang S, Tan AC, Mathew J. Rotating machinery prognostics: state of the art, challenges and opportunities. Mech Syst Signal Process 2009;23:724–39.
[2] Nandi S, Toliyat HA. Condition monitoring and fault diagnosis of electrical machines – a review. IEEE Trans Energy Convers 2005;20:719–29.
[3] Lee J, Ghaffari M, Elmeligy. S. Self-maintenance and engineering immune systems: towards smarter machines and manufacturing systems. Annu Rev Control 2011;35:111–22.
[4] Frosini L, Bassi E. Stator current and motor efficiency as indicators for different types of bearing faults in induction motors. IEEE Trans. Ind. Electron. 2010;57:244–51.
[5] Ben Salem S, Bacha K, Chaari AK. Support vector machine based decision for mechanical fault condition monitoring in induction motor using an advanced Hilbert–Park transform. ISA Trans 2012;51:566–72.
[6] Konar P, Chattopadhyay P. Bearing fault detection of induction motor using wavelet and support vector machines (SVMs). Appl Soft Comput 2011;11:4203–11.
[7] Cheng J, Yang Y, Yang Y. A rotating machinery fault diagnosis method based on local mean decomposition. Digit Signal Process 2012;22:356–66.
[8] Zhou Y, Chen J, Dong GM, Xiao WB, Wang ZY. Application of the horizontal slice of cyclic bispectrum in rolling element bearings diagnosis. Mech Syst Signal Process 2012;26:229–43.
[9] Yiakopoulos CT, Antoniadis IA. Cyclic bispectrum patterns of defective rolling element bearing vibration response. Forsch Ing 2006;70:90–4.
[10] Xu TL, Lang XZ, Zhang XY, Pei XC. The research based on empirical mode decomposition in bearing fault diagnosis. Appl Mech Mater 2012;103:225–8.
[11] Nikias CL, Petropulu A. Higher-order spectra analysis: a nonlinear signal processing framework. Englewood Cliffs, New Jersey: Prentice-Hall; 1993.
[12] Mendel JM. Tutorial on higher order statistics (spectra) in signal processing and system theory: theoritical results and some applications. Proc. IEEE 1993;79:287–95.
[13] Guoji S, McLaughlin S, Yongcheng X, White P. Theoretical and experimental analysis of bispectrum of vibration signals for fault diagnosis of gears. Mech Syst Signal Process 2014;43:76–89.
[14] Stack JR, Harley RG, Habetler TG. An amplitude modulation detector for fault diagnosis in rolling element bearings. IEEE Trans Ind Electron 2004;51:1097–102.
[15] Saidi L, Henao H, Fnaiech F, Capolino G-A, Cirrincione G. Diagnosis of broken bars fault in induction machines using higher order spectral analysis. ISA Trans 2013;52:140–8.
[16] Vapnik VN. The nature of statistical learning theory. New York: Springer; 1999.
[17] Delgado M, Cirrincione G, Espinosa AG, Ortega JA, Henao H. Bearing faults detection by a novel condition monitoring scheme based on statistical-time features and neural networks. IEEE Trans Ind Electron 2013;30:3398–407.
[18] Jack LB, Nandi AK. Fault detection using support vector machines and artificial neural network, augmented by genetic algorithms. Mech Syst Signal Process 2002;16:373–90.

[19] Lei YG, He ZJ, Zi YY. Fault diagnosis of rotating machinery based on multiple ANFIS combination with Gas. Mech Syst Signal Process 2007;21:2280–94.
[20] Widodo A, Yang BS. Support vector machine in machine condition monitoring and fault diagnosis. Mech Syst Signal Process 2007;21:2560–74.
[21] Wenyi L, Zhenfeng W, Jiguang H, Guangfeng W. Wind turbine fault diagnosis method based on diagonal spectrum and clustering binary tree SVM. Renew Energy 2013;50:1–6.
[22] Gryllias KC, Antoniadis IA. A support vector machine approach based on physical model training for rolling element bearing fault detection in industrial environments. Eng Appl Artif Intell 2012;25:326–44.
[23] Yuan S, Chu F. Fault diagnosis based on support vector machines with parameter optimisation by artificial immunisation algorithm. Mech Syst Signal Process 2007;21:1318–30.
[24] Chua KC, Chandranb V, Acharyaa UR, Lima CM. Application of higher order statistics/spectra in biomedical signals—a review. Med Eng Phys 2010;32:679–89.
[25] Widodo A, Yang BS, Han T. Combination of independent component analysis and support vector machines for intelligent faults diagnosis of induction motors. Expert Syst Appl 2007;32:299–312.
[26] Tran VT, Yang BS, Gua F, Ball A. Thermal image enhancement using bi-dimensional empirical mode decomposition in combination with relevance vector machine for rotating machinery fault diagnosis. Mech Syst Signal Process 2013;38:601–14.
[27] Urbaneka J, Barszcza T, Antoni J. Integrated modulation intensity distribution as a practical tool for condition monitoring. Appl Acoust 2014;77:184–94.
[28] The Case Western Reserve University Bearing Data Center. ⟨http://csegroups.case.edu/bearingdatacenter/pages/download-data-file⟩ [accessed 21.7.14].
[29] Rifkin R, Klautau A. In defence of one-vs.-all classification. J Mach Learn 2004;5:101–41.
[30] LIBSVM – A Library for Support Vector Machines. Available from: ⟨http://www.csie.ntu.edu.tw/∼cjlin/libsvm⟩ [accessed 17.12.13].
[31] Widodo A, Yang BS, Gu DS, Choi BK. Intelligent fault diagnosis system of induction motor based on transient current signal. Mechatronics 2009;19:680–9.