



Rolling element bearing fault detection in industrial environments based on a *K*-means clustering approach

C.T. Yiakopoulos*, K.C. Gryllias, I.A. Antoniadis

Dynamics and Structures Laboratory, Machine Design and Control Systems Section, School of Mechanical Engineering, National Technical University of Athens, Iroon Polytechniou 9, Athens 15773, Greece

ARTICLE INFO

Keywords:

Condition monitoring
Fault detection
K-means clustering
Vibration analysis
Rolling element bearings

ABSTRACT

A *K*-means clustering approach is proposed for the automated diagnosis of defective rolling element bearings. Since *K*-means clustering is an unsupervised learning procedure, the method can be directly implemented to measured vibration data. Thus, the need for training the method with data measured on the specific machine under defective bearing conditions is eliminated. This fact consists the major advantage of the method, especially in industrial environments. Critical to the success of the method is the feature set used, which consists of a set of appropriately selected frequency-domain parameters, extracted both from the raw signal, as well as from the signal envelope, as a result of the engineering expertise, gained from the understanding of the physical behavior of defective rolling element bearings. Other advantages of the method are its ease of programming, simplicity and robustness. In order to overcome the sensitivity of the method to the choice of the initial cluster centers, the initial centers are selected using features extracted from simulated signals, resulting from a well established model for the dynamic behavior of defective rolling element bearings. Then, the method is implemented as a two-stage procedure. At the first step, the method decides whether a bearing fault exists or not. At the second step, the type of the defect (e.g. inner or outer race) is identified. The effectiveness of the method is tested in one literature established laboratory test case and in three different industrial test cases. Each test case includes successive measurements from bearings under different types of defects. In all cases, the method presents a 100% classification success. Contrarily, a *K*-means clustering approach, which is based on typical statistical time domain based features, presents an unstable classification behavior.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Rolling element bearings consist one of the most widely used industrial machine elements, being the interface between the stationary and the rotating part of the machine. The capability to detect fast, accurately and easily the existence and severity of a fault in an installation during operation is very important as an unexpected failure of machine can drive to unacceptably long time maintenance stops. Thus, due to the importance of bearings, a plethora of monitoring methods and fault diagnosis procedures has been developed, in order to reduce maintenance costs, improve productivity, and prevent malfunctions and failures during operation which could lead to the downtime of the machine.

A large number of them are based on a physical and engineering concept, as established long ago by the term “high-frequency resonance technique (HFRT)” (McFadden & Smith, 1984b). This method is alternatively known as demodulated resonance analysis or envelope power spectral density analysis (Randall, 1987), since it

consists in band-pass filtering a frequency range around a machine structural resonance, which is excited by the periodic impulsive excitation forces, generated on bearings with localized defects. Further demodulating this band generates a “signal envelope” which presents a more clear impulsive nature than the signal itself, able to identify and characterize the nature of the defect.

Parallel, due to the fact that the traditional engineering approaches require a significant degree of engineering expertise, an impressive number of methods have been presented for automated defective rolling element bearing fault detection – and more generally for machine condition monitoring – based on intelligent and/or expert systems. These methods include artificial neural networks (ANN), fuzzy C-means (FCM), hierarchical and partitional clustering, and support vector machines (SVM).

Fuzzy C-means (FCM) clustering has been widely studied and applied to automatic fault diagnosis (Elaine, Wai, Michael, & Joshua, 2004; Saravanan, Cholaiajan, & Ramachandran, 2009; Wang, Wang, & Wang, 2004). A fuzzy clustering algorithm employs a fuzzy mathematics method to partition a set of data into several homogenous groups using procedures that attempt to find natural partitions of patterns, being thus an unsupervised classifier. The

* Corresponding author. Tel.: +30 210 7722332; fax: +30 210 7721525.

E-mail address: chryiako@central.ntua.gr (C.T. Yiakopoulos).

FCM algorithm does not take account the different importance degrees of the features. Therefore, all the features have uniform contributions to clustering, leading to inexact clustering results. There exist several improved clustering algorithms, in which the importance of the features in clustering is considered. Yaqou, Zhengjia, Yanyang, and Xuefeng (2008) introduced a two-stage feature selection and weighting technique based on the compensation distance evaluation technique (CDET), which is incorporated into the FCM algorithm. This clustering algorithm was applied to fault diagnosis of locomotive roller bearings using statistical time-domain and frequency-domain feature parameters. This algorithm overcomes the existing shortcoming in the FCM algorithm and shows an ability to identify the different fault categories.

Artificial neural networks (ANN) have been applied in automated detection of machine conditions (McCormick & Nandi, 1997). However, the traditional neural network approaches have limitations on generalization giving rise to models that can overfit to the training data. This deficiency is due to the optimization algorithms used in ANNs for selection of parameters and the statistical measures used to select the model. Recently, Samanta (2004) used the extracted features from raw and preprocessed signals as inputs to ANNs and support vector machines (SVMs) for two-class recognition. The selection of time-domain input features and the classifier parameters were optimized using genetic algorithms (GA). For most of the cases, the classification accuracy of SVM was better than ANN, without GA. The training time of SVMs was less compared to ANNs. Finally, the performance of both classifiers with GA-based selection was comparable at a success rate of nearly 100%. The main difference between this work and that of Jack and Nandi (2002) is the process of feature extraction from the time-domain signal.

Qiao, Zhengjia, Zhousuo, and Yanyang (2007) proposed a novel method for diagnosis of rotating machinery, based on an improved wavelet package transform (IWPT), a distance evaluation technique and the support vector machines (SVMs) ensemble. The wavelet package transform was based on a biorthogonal wavelet with an impact property that is constructed via a lifting scheme. First, the wavelet package coefficients were processed through Hilbert transform resulting to the calculation of the characteristic frequencies of defective rolling element bearings. Then, the distance evaluation technique was used to select the optimal features, by removing the irrelevant information. Finally, in order to diagnose the type of the fault, the SVMs ensemble with AdaBoost approach was adopted. The average fault diagnosis accuracy has proved much higher than that of single SVMs.

In order to enhance the ability of traditional linear system based feature extraction methods, non-linear fault features such as fractal dimensions can be additionally taken into account for fault diagnosis (Logan & Mathew, 1996a; Logan & Mathew, 1996b). Junyan, Youyun, and Yongsheng (2007) classified various fault types using the capacity, the information and the correlation dimension. The experimental results showed that the classification performance of the support vector machines (SVMs) was improved significantly using in the training procedure eleven time-domain statistical features in tandem with three fractal dimensions.

Genetic programming (GP) is used by Zhang, Jack, and Nandi (2005) to detect faults in rotating machinery. For ANNs and SVMs, genetic algorithms (GA) have been used for feature selection, which is an inherent function of GP. In all cases, the GP has selected fewer features to perform the classification tasks than GA/ANN and GA/SVM approaches. Parallel, the GP demonstrates performance which equals or betters that of the previous approaches.

A combination of independent component analysis (ICA) or principal component analysis (PCA) and SVMs are proposed by Widodo, Yang, and Han (2007) for intelligent diagnosis of induction motor. Initially, total 78 features of time domain and three

parameters of frequency domain are selected. Then, ICA and PCA are applied for feature extraction reducing the parameters to 24 independent and uncorrelated components. After feature extraction, a feature selection process is performed using the distance evaluation technique to remove irrelevant and useless parameters. Next, the SVMs-based classification is trained on data input from original and extracted features followed by Kernel parameters selection. Finally, the SVM is applied to perform the classification process. According to the results, the ICA based feature extraction combining Kernel parameters selection achieves a successful classification of the faults.

Komgom, Mureithi, Aouni, and Marc (2007) introduced time synchronous averaging (TSA) combined with SVM for bearing fault diagnosis. First, TSA and envelope detection of the raw signals by convolving them with a repeating sequence (BPFR, BPFI and BPFO) is performed. Thus, the feature vector contains time and frequency-domain parameters evaluated by the TSA signals obtained. Then, the independent components are determined by the Fast ICA algorithm. The SVMs-based classification is trained on data input from TSA signals. Finally, the SVM is applied to perform the classification process. The performance of the proposed method gives good results when the testing signals are picked up on the same side than the training signals.

A decision tree for bearing diagnosis was established by Lee, Nguyen, and Kwon (2007). Time-domain features are calculated from a three dimension vibration signal and extracted by PCA to reduce training data dimension and remove irrelevant data. The decision has high accuracy but cannot provide the severity level of the bearing fault and it is sensitive to noise.

The very popular *K*-means (MacQueen, 1967) clustering is proposed in this paper, because it is an unsupervised pattern classification method, being thus able to be applied directly to industrial environments without the need of being trained by data measured on a machine under a fault condition. Further advantages of the method are its ease of programming and the accomplishment of a good trade-off between achieved performance and computational complexity.

The first step in the application of *K*-means clustering is to find a set of initial centers. The algorithm is significantly sensitive to the initial randomly selected cluster centers. Varying the starting conditions can produce different stable cluster solutions. In order to overcome this limitation, a starting strategy is implemented. The initial centers are calculated from simulated signals, resulting from a well established model for the dynamic behavior of defective rolling element bearings. This technique allows the production of only one final cluster solution. Besides, this approach achieves to categorize the produced clusters against the unsupervised structure of *K*-means clustering.

After the selection of the initial centers, the proposed method is directly applicable to signals measured in industrial or laboratory environments, leading to a two-stage fault classification procedure: In the first stage, the method leads to a classification if a fault exists or not. In case of a fault, the method further proceeds at the second stage to its classification as an inner or outer race defect.

Quite critical to the success of the method is the selection of proper feature inputs, since many traditionally used features are irrelevant and redundant in fault diagnosis of rotating machinery, resulting to low diagnosis accuracy. Thus, a set of features based on vibration energies in the frequency domain is proposed, which clearly characterize the bearing behavior under various operation conditions.

The method is evaluated in three different industrial cases and one experimental case, each one containing a set of successive measurements which follow the evolution of an outer or an inner race defect. As compared to an alternative *K*-means clustering approach, based on traditional time-domain statistical features, the

proposed K -means clustering approach achieves 100% classification success in all the cases.

The rest of the paper is organized as follows. A brief overview of the K -means clustering approach is performed in Section 2. An overview of a well established dynamic model characterizing the vibration response of defective rolling element bearings is performed in Section 3. Section 4 presents the approach for the implementation of the method, with special emphasis to the proposed frequency domain based feature set. In Section 5 the effectiveness of the method is tested in one literature established laboratory test case and in three different industrial test cases.

2. Overview of the K -means clustering approach

K -means clustering is one of the simplest and most popular unsupervised learning algorithms to solve the clustering problem. K -means clustering generates a specific number of disjoint, flat (non-hierarchical) clusters. That is, the K -means function partitions the observations extracted from the data into k mutually exclusive clusters, and returns a vector of indices, indicating to which of the k clusters each feature set has been assigned.

The M -clustering problem aims at partitioning a data set $X = \{x_1, \dots, x_N\}$, $x_n \in R^d$ into M disjoint subsets C_1, \dots, C_M , such that a function F , referred as clustering criterion, is optimized. The most intuitive and frequently used criterion function is the sum of the squared Euclidean distances between each data point x_i and the cluster center m_k of the subset C_k which contains x_i . This criterion is called clustering error and depends on the centroids m_1, \dots, m_M :

$$F(m_1, \dots, m_M) = \sum_{i=1}^M \sum_{j=1}^{M_i} \|x_{ij} - m_i\|^2, \quad (1)$$

where M_i is the number of objects of the cluster C_i , x_{ij} is the j th object of the i th cluster and m_i is the center of the i th cluster which is defined as

$$m_i = \frac{1}{M_i} \sum_{j=1}^{M_i} x_{ij}, \quad i = 1, \dots, M, \quad (2)$$

A pseudo-code for the implementation of K -means clustering is as follows:

- (A) The algorithm starts with an initial partition of the database in M disjoint subsets $\{C_1, \dots, C_M\}$ and the centroids m_i of these initial clusters are calculated.
- (B) Then, the instances of the database are relocated to the cluster represented by the nearest centroid in an attempt to reduce the square-error. An instance $x_i \in C_s$ in the relocation step can change its cluster membership $x_i \in C_t$ if $\|x_i - m_t\| \leq \|x_i - m_s\|$ for all $j = 1, \dots, M$, $j \neq s$. Then the centroids of the clusters C_s and C_t , and the square-error should be recomputed.
- (C) This process is repeated until the square-error cannot be further reduced, which means that no instance is able to further change its cluster membership.

It has been shown that the performance of the method is strongly related to the metric distance used. A series of different distance measures can be chosen, depending on the kind of data that are clustered.

The Euclidean distance function measures the straight line distance between two points. Each centroid is the mean of the points in that cluster. Given two genes $A_i = \{x_{ik}\}$ and $A_j = \{x_{jk}\}$, $k = 1, \dots, n$ and $i, j \in \{1, \dots, n\}$, $i \neq j$, the Euclidean distance between A_i and A_j is formulated as

$$r(A_i, A_j) = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2}, \quad (3)$$

where $x_i, x_j \in R$ are the measured expression levels. The Euclidean distance r measures the difference in the individual magnitudes of each gene. The Euclidean distance takes the difference between the two genes directly. It should therefore only be used for expression data that are suitably normalized. The Euclidean distance contains information on both direction and magnitude of each vector from the cluster mean. Hence, the genes may be regarded as similar by Euclidean distance though they are very dissimilar in terms of their shapes or vice versa.

Pearson's correlation measures the similarity between two profiles. Each centroid is the component-wise mean of the points in that cluster, after centering and normalizing those points to zero mean and unit standard deviation. Given two series of numbers $A_i = \{x_{ik}\}$ and $A_j = \{x_{jk}\}$, $k = 1, \dots, n$ and $i, j \in \{1, \dots, n\}$, $i \neq j$, the Pearson's correlation (centered) is defined as

$$r(A_i, A_j) = \frac{\sum_{k=1}^n (x_{ik} - m_i)(x_{jk} - m_j)}{\sqrt{\sum_{k=1}^n (x_{ik} - m_i)^2} \sqrt{\sum_{k=1}^n (x_{jk} - m_j)^2}} \quad (4)$$

where m_i and m_j are the means of x_{ik} and x_{jk} , $k = 1, \dots, n$, respectively. The Pearson correlation coefficient is always between -1 and 1 , with 1 denoting that the two series are identical, 0 denoting that they are completely uncorrelated, and -1 denoting that they are perfect opposites.

The correlation distance in clustering considers each gene as a random variable with n observations and measures the similarity between the two genes by calculating the linear relationship between the distributions of the two corresponding random variables. The correlation coefficient is invariant under linear transformation of the data. The correlation distance is insensitive to the amplitude of expression (ignores differences in magnitude) and takes into account the trends of the change. The Pearson correlation coefficient can be used for interval-scaled and ratio-scaled data. The only disadvantage of using correlation is that it is slightly slower than the Euclidean distance.

Cosine metric distance is defined as one minus the cosine of the included angle between points. Each centroid is the mean of the points in that cluster, after normalizing those points to unit Euclidean length. Given two series of numbers A_i and A_j , $i, j \in \{1, \dots, n\}$, $i \neq j$, the cosine metric distance is defined as

$$r(A_i, A_j) = 1 - \frac{\sum_{k=1}^n x_{ik} x_{jk}}{\sqrt{\sum_{k=1}^n x_{ik}^2} \sqrt{\sum_{k=1}^n x_{jk}^2}}, \quad (5)$$

where

$$\frac{\sum_{k=1}^n x_{ik} x_{jk}}{\sqrt{\sum_{k=1}^n x_{ik}^2} \sqrt{\sum_{k=1}^n x_{jk}^2}} \quad (6)$$

is the uncentered Pearson's correlation which is equal to the cosine of the angle of the two n -dimensional vectors A_i and A_j . Eq. (6) is the same function with Eq. (4), except that it assumes that the mean is zero, even when it is not. The uncentered Pearson's correlation distance measures the similarity in shape between two profiles, but it can also capture inverse relationships. In particular, combining K -means with metric distances of Eq. (6) can lead to non-intuitive centroid plots since it can group anti-correlated objects.

Cityblock distance, also known as Manhattan distance, examines the absolute differences between coordinates of a pair of objects. Each centroid is the component-wise median of the points in that cluster. Given two genes A_i and A_j , $i, j \in \{1, \dots, n\}$, $i \neq j$, the city block distance is defined as

$$r(A_i, A_j) = \frac{1}{n} \sqrt{\sum_{k=1}^n |x_{ik} - x_{jk}|} \quad (7)$$

The reasons behind the popularity of the K -means algorithm are

- (a) It is a simple and unsupervised learning algorithm. The K -means is not limited to hypothesis testing based on some prior knowledge. It explores the structure of the data on the basis of similarities that are presented in it.
- (b) The time required is of $O(nML)$, where n is the number of patterns, M is the number of clusters and L is number of iterations taken by the algorithm to converge. The parameters M and L are fixed in advance. Hence, the K -means approach has linear time complexity in the size of the data set.
- (c) Its space complexity is $O(m+n)$. Only the data points and the centroids are stored. The algorithm requires an additional space to store the data matrix.
- (d) It is order independent. Thus, for the same initial conditions, it generates the same partition of the data irrespective of the order in which the patterns are presented in the algorithm.
- (e) With a large number of variables, K -means may be computationally faster.
- (f) It may produce tighter clusters.
- (g) It can be used for a wide variety of data types.
- (h) It is able to find pure subclusters if a large enough number of clusters is specified.
- (i) Although there is no guarantee of achieving a global minimum, the convergence of the algorithm is ensured.

The major disadvantage of K -means clustering is the sensitivity of the algorithm to the choice of the initial conditions. One technique used to address the problem of choosing initial centers is to perform multiple runs, due to the given high speed of the algorithm, each one with a different set of randomly chosen initial centroids, and then select the set of clusters with the minimum clustering error. This strategy may not work very well, depending on the dataset and the number of clusters sought.

Due to the problem of using randomly selected initial centroids, which even repeated runs may not overcome, other techniques are employed for initialization. In the literature, several techniques have been proposed for selecting the best initial positions of the centroids, with the aim of inducing a high quality solution at the

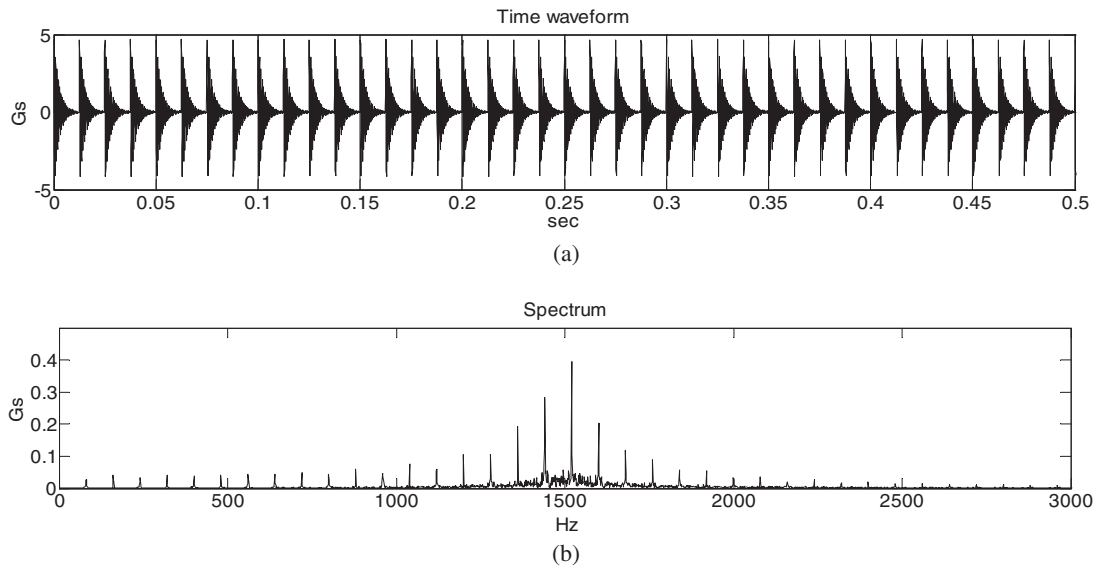


Fig. 1. Vibration response of a rolling element bearing under an outer race defect, as generated by the model of Section 3, Eq. (13), for BPFO = 80 Hz, $f_{\text{shaft}} = 20$ Hz and impact amplitude $A = 5$, (a) time waveform of the signal, (b) frequency spectra of the signal.

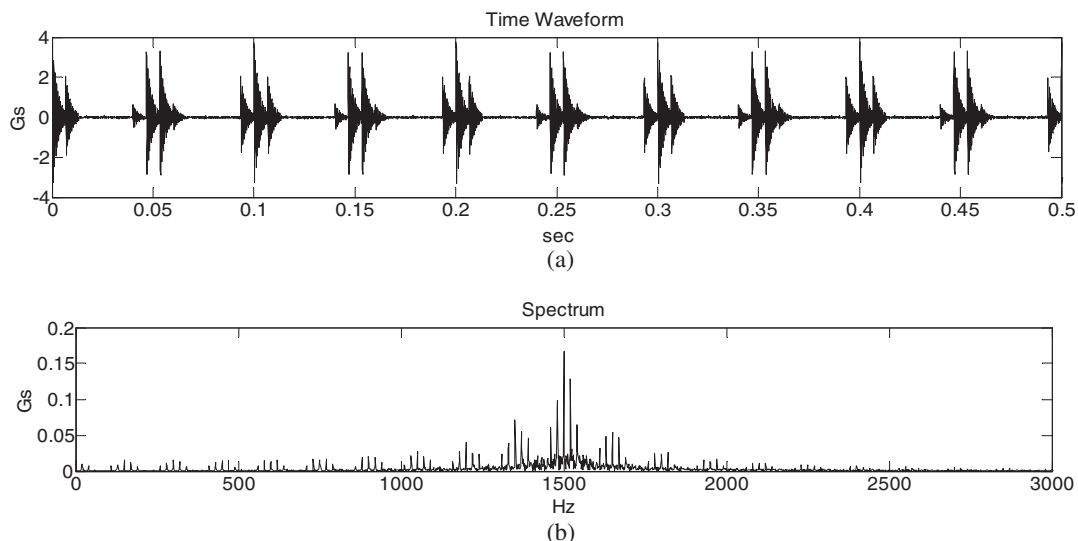


Fig. 2. Vibration response of a rolling element bearing under an inner race defect, as generated by the model of Section 3, Eq. (13), for BPFI = 150 Hz, $f_{\text{shaft}} = 20$ Hz and impact amplitude $A = 10$, (a) time waveform of the signal, (b) frequency spectra of the signal.

end of the K -means run. Milligan (1980) testified the strong dependence of the K -means on the initial conditions and suggested that good final cluster structures can be obtained using Ward's hierarchical method (Ward, 1963). Likas et al. proposed an incremental approach to clustering, called global K -means (Likas, Vlassis, & Verbeek, 2003). The initial centers are obtained using a tree method. The algorithm is deterministic, does not depend on any initial positions for the cluster center and does not contain any empirical parameters. Bradley et al. presented the refinement algorithm (Bradley & Fayyad, 1998) that operates over small subsamples of a given database. The data in each subsample is clustered. Afterwards, all centroids of all the subsamples are clustered together using K -means. The centroids of each subsample are used as initial centers. The final centers that give minimum clustering error are then used as initial conditions in K -means clustering of the original dataset.

The above mentioned initialization algorithms are able to lead to the calculation of good initial centers. They do not constitute just initialization methods. They are clustering methods by themselves and some of them use the K -means algorithm as part of their algorithms. Hence, they suffer from the same problem as the K -means approach.

3. Defective rolling element bearing vibration model

Dynamic behavior of rolling element bearings and machines under localized defects has been a subject of intensive research, leading to a number of well established models (Antoni & Randall, 2002; MacFadden & Smith, 1984a; Randall, 1987). Following actually the same concepts as in Antoni and Randall (2002), MacFadden and Smith (1984a), the basic elements of such a model are as follows.

The repetitive impacts produced by a localized bearing defect can be described by a train of Dirac delta functions $\delta(t)$ of the following form:

$$d(t) = \begin{cases} q(t)d_0 \sum_{k=0}^N \delta(t - kT_d) & \text{for an inner race defect} \\ d_0 \sum_{k=0}^N \delta(t - kT_d) & \text{for an outer race defect} \end{cases}, \quad (8)$$

where d_0 is the amplitude of the impulse force characterizing the severity of the defect. The impulse repetition period T_d is the reciprocal of the characteristic ball pass frequency outer race (BPFO) or ball pass frequency inner race (BPFI) of a rolling element bearing, depending on the type of defect. These two frequencies are proportional to the shaft rotation speed f_{shaft} , and their value depends on the bearing geometric characteristics. The function $q(t)$ is the distribution of the load around the rolling element bearing under radial

load which approximated typically by the well-known Stribeck equation:

$$q(t) = \begin{cases} q_0[1 - (1/2\varepsilon)(1 - \cos \theta)]^n & \text{for } |\theta| < \theta_{\max} \\ 0 & \text{elsewhere} \end{cases}, \quad (9)$$

where q_0 is the maximum load intensity, ε is the load distribution factor, θ_{\max} is the angular extent of the load zone and $n = 3/2$ for ball bearings. The terms q_0 , ε and θ_{\max} are all functions of the diametral clearance of the bearing and the applied load. For a bearing with positive clearance, $\varepsilon < 0.5$ and $\theta_{\max} < \pi/2$.

However, significant variations of the dynamic response of a defective rolling element can exist, when certain parameters of the above equation are assumed to present rather a stochastic than a deterministic behavior. The amplitude of the impacts depends mainly on the load distribution around the circumference of the bearing, as well as on other parameters, such as on the variation in the dynamic stiffness of the assembly, the waviness of the rolling elements and races, and the existence of off-sized balls in the ball complement. Therefore, the series of the impacts should be rather considered as randomly modulated in amplitude.

In parallel, rolling element bearings present a slip motion which introduces nonlinearity effects to the system. As a consequence, the impulses never occur exactly at the same position from one cycle to another.

Thence, a more realistic model for the impact train of a defective rolling element bearing should be of the following form:

$$D(t) = \begin{cases} q(t) \sum_{k=0}^N A_k \delta(t - kT_d - \tau_k) & \text{for an inner race defect} \\ \sum_{k=0}^N A_k \delta(t - kT_d - \tau_k) & \text{for an outer race defect} \end{cases}, \quad (10)$$

Table 2

Typical ranges of the parameters used by the dynamic bearing model.

Model parameters	Outer race defect		Inner race defect	
	Industrial signals	Laboratory signals	Industrial signals	Laboratory signals
Amplitude (Gs)	0.5–40	0.5–20	0.5–20	0.5–10
Slip (%)	0.5–3	0.5–3	0.5–3	0.5–3
Noise level (NL) (%)	0.1–5	0.1–5	0.1–5	0.1–5
Natural frequency (Hz)	1500–5000	1000–7000	1500–5000	1000–7000

Table 1

Energy features extracted from the frequency spectra of the raw signal and of the envelope of the signal. The index K typically takes the values $K = 1, 4$.

Parameter	Symbol
<i>Spectra of the envelope of the signal</i>	
Spectral peak at the K th harmonic of the shaft rotational speed	SKX
Spectral peak at the K th harmonic of the ball pass frequency outer race (BPFO)	SKO
Spectral peak at the K th harmonic of the ball pass frequency inner race (BPFI)	SKI
Sum of shaft harmonic spectral peaks: ($SSHE = \sum_K SKX$)	$SSHE$
Sum of BPFO harmonic spectral peaks: ($SORH = \sum_K SKO$)	$SORH$
Sum of BPFI harmonic spectral peaks: ($SIRH = \sum_K SKI$)	$SIRH$
Sum of BPFI and BPFO harmonic spectral peaks: ($SIORH = SIRH + SORH$)	$SIORH$
<i>Spectra of the raw signal</i>	
Sum of first four shaft harmonic spectral peaks	$SSHR$
High-frequency energy: $HFE = \int_{f_H}^{f_H^{1/2}} f^2 df$	HFE

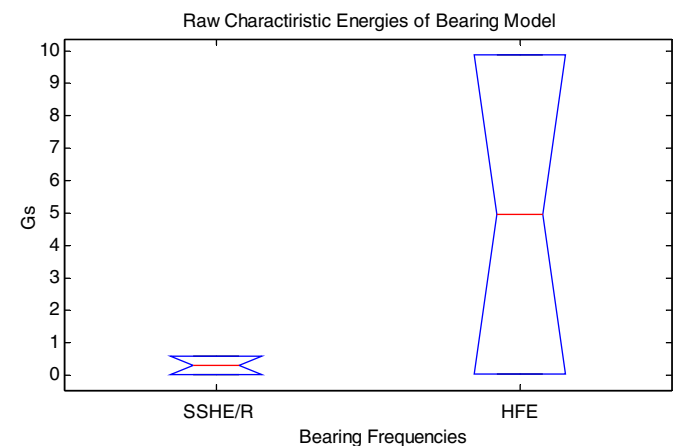


Fig. 3. Energy ranges of: (a) sum of shaft harmonic spectral peaks and (b) high-frequency field, which are calculated in the spectrum of the raw waveform produced by the dynamic bearing model using the parameters of Table 2.

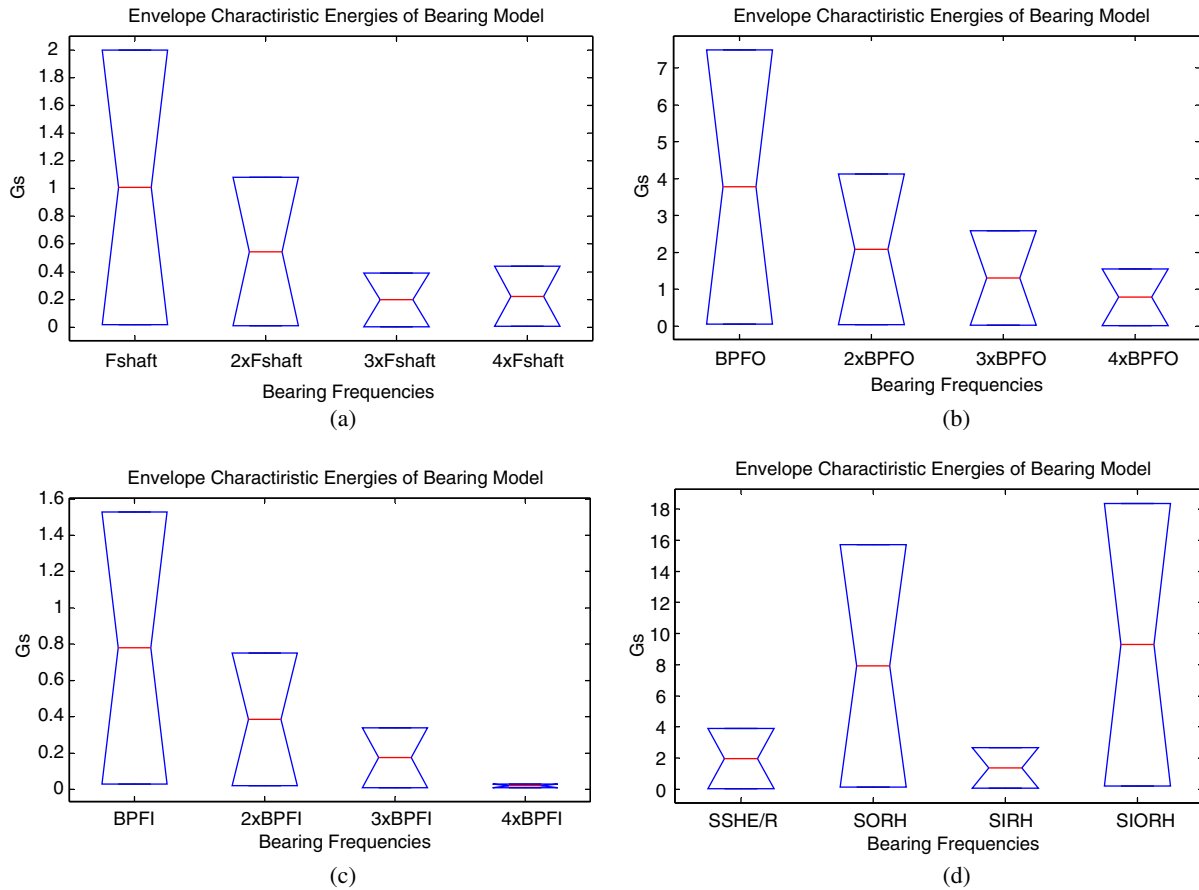


Fig. 4. Energy ranges of: (a) f_{shaft} and its harmonics, (b) BPFO and its harmonics, (c) BPFI and its harmonics and (d) harmonics' sum of each characteristic defect frequency, which are calculated in the envelope spectrum of the waveforms produced by the dynamic bearing model using the parameters of Table 2.

where A_k are random variables for the amplitude of the k th impulse force with a probability density function typically assumed to be normal (Gaussian), τ_k is a random variable for the time lag between two impacts due to the presence of slip, which is typically assumed to be of zero mean and of normal (Gaussian) probability density function.

Table 3

Energy features used in the 1st stage of the proposed K -means based fault classification. For symbol definitions refer to Table 1.

Parameters used in the first stage	
SSHR, SIORH, HFE	

Table 4

Relative energy indices, used in the 2nd stage of the proposed K -means based fault classification. For symbols used in the definitions refer to Table 1. The index K typically takes the values $K = 1, 4$.

Symbol	Definition	Description
$RKOI$	$RKOI = SKO/SKI$	Ratio of the spectral peak at the K th harmonic of the BPFO to the spectral peak at the K th harmonic of the BPFI
$RSOI$	$RSOI = SORH/SIRH$	Ratio of the sum of the BPFO spectral peaks to the sum of the BPFI spectral peaks
$RKIO$	$RKIO = SKI/SKO$	Ratio of the spectral peak at the K th harmonic of the BPFI to the spectral peak at the K th harmonic of the BPFO
$RSIO$	$RSIO = SIRH/SORH$	Ratio of the sum of the BPFI spectral peaks to the sum of the BPFO spectral peaks
$RKXO$	$RKXO = SKX/SKO$	Ratio of the spectral peak at the K th harmonic of the shaft rotation speed to the spectral peak at the K th harmonic of the BPFO
$RSXO$	$RSXO = SSHE/SORH$	Ratio of the sum of the shaft harmonic spectral peaks to the sum of the BPFO spectral peaks
$RKOX$	$RKOX = SKO/SKX$	Ratio of the spectral peak at the K th harmonic of the BPFO to the spectral peak at the K th harmonic of the shaft rotation speed
$RSOX$	$RSOX = SORH/SSHE$	Ratio of the sum of the BPFO spectral peaks to the sum of the shaft harmonic spectral peaks
$RKXI$	$RKXI = SKX/SKI$	Ratio of the spectral peak at the K th harmonic of the shaft rotation speed to the spectral peak at the K th harmonic of the BPFI
$RSXI$	$RSXI = SSHE/SIRH$	Ratio of the sum of the shaft harmonic spectral peaks to the sum of the BPFI spectral peaks
$RKIX$	$RKIX = SKI/SKX$	Ratio of the spectral peak at the K th harmonic of the BPFI to the spectral peak at the K th harmonic of the shaft rotation speed
$RSIX$	$RSIX = SIRH/SSHE$	Ratio of the sum of the BPFI spectral peaks to the sum of the shaft harmonic spectral peaks

When an impulse train is applied to the rolling element bearing, it will excite structural resonances. Assuming for simplicity that the excited structure is a linear MDOF system, the impulsive structural response of such a system to each impulse can be expressed by the following function:

$$s(t) = \sum_{i=1}^M B_i \cdot e^{-2\pi\zeta_i f_{ni} t} \cdot \cos(2 \cdot \pi \cdot f_{0i} \cdot t), \quad (11)$$

$$f_{0i} = f_{ni} \cdot \sqrt{1 - \zeta_i^2}, \quad (12)$$

where $i = 1, \dots, M$ are the excited modes, and for each mode i , f_{ni} is the resonance frequency, and ζ_i is the damping factor. Therefore, the dynamic response $x(t)$ resulting from an induced defect in a bearing can be expressed by

$$d(t) = \begin{cases} \sum_{k=0}^N q(t)A_k\delta(t - kT_d - \tau_k) \otimes \sum_{i=1}^M B_i e^{-2\pi\zeta_i f_{ni}^t} \cos(2\pi f_{oi}t) + n(t) & \text{for an inner race defect} \\ \sum_{k=0}^N A_k\delta(t - kT_d - \tau_k) \otimes \sum_{i=1}^M B_i e^{-2\pi\zeta_i f_{mi}^t} \cos(2\pi f_{oi}t) + n(t) & \text{for an outer race defect} \end{cases}, \quad (13)$$

where the symbol \otimes denotes convolution and $n(t)$ is an additive background noise, which takes into account the effects of the other vibration sources of the system and the external environment. This is the unpredictable measurement noise which is present in any practical measurement system.

An application of the model of Eq. (13) to generate the expected vibration response of a defective rolling element bearing with an

outer and an inner race defect is shown in Figs. 1 and 2, respectively. As it can be observed, a spiky behavior characterizes the time-domain waveforms, while the frequency-domain response is characterized by sidebands around the structural resonance frequency, spaced at distances equal to the corresponding BPFO or BPFI frequency. Moreover in the case of the inner race defect, additional sidebands exist around the inner race defect frequencies, spaced at distances equal to the shaft rotating speed.

4. Proposed K-means clustering classification approach

4.1. Data preprocessing and feature extraction

The dynamic model for the vibration response of defective rolling element bearings described in the previous section foresees

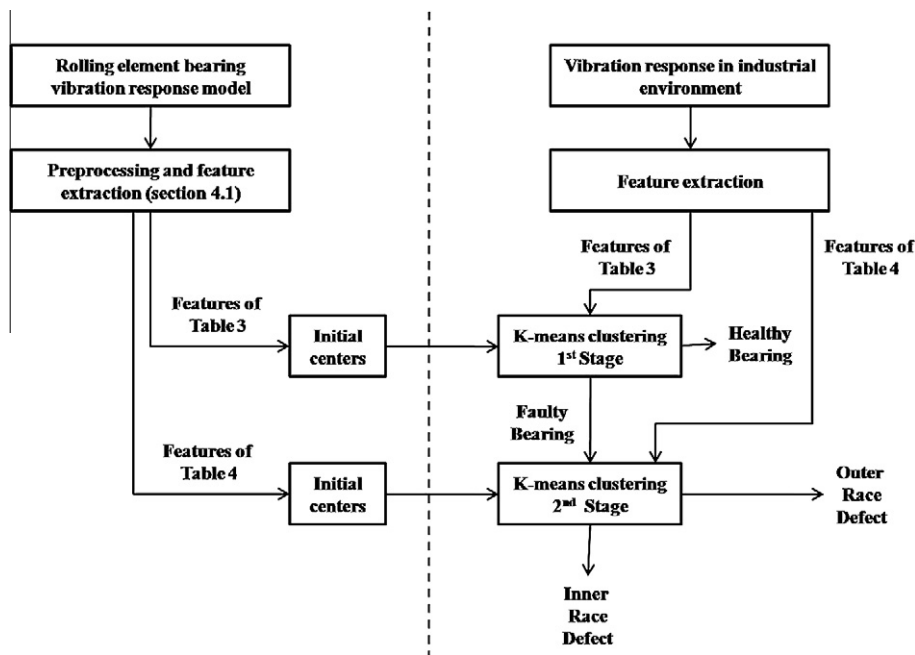


Fig. 5. Flow chart of the proposed two stage K-means clustering based classification algorithm.

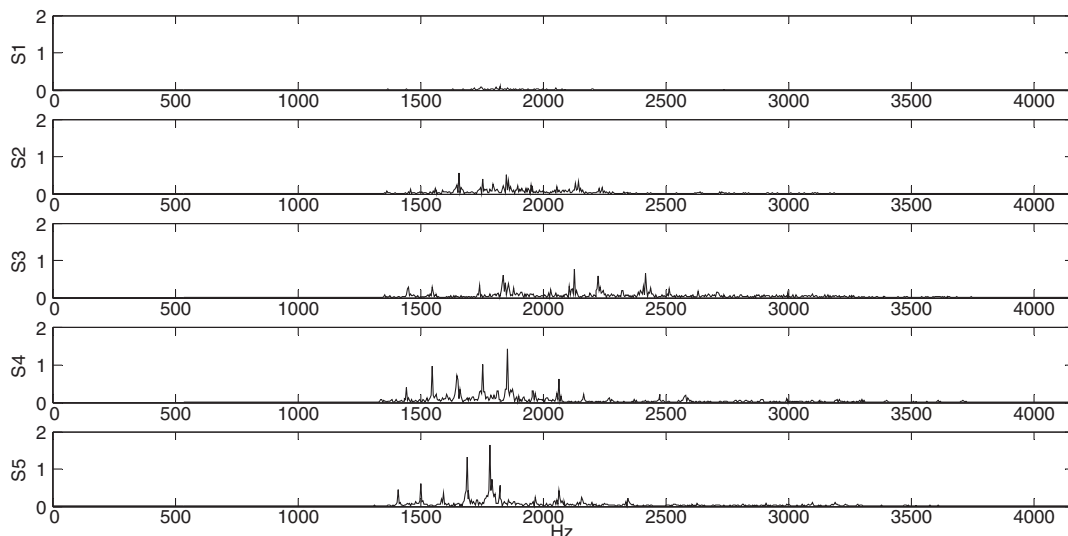


Fig. 6. Test case I: spectra of the raw signals.

that bearing defects cause impacts at characteristic “defect frequencies”, governed by the rotating speed of the machine and the geometry of the bearings, which in turn excite various machine natural frequencies. As a result, a “spiky” behavior is generated in the time-domain signal. The frequency-domain transform of the signal results to entire high-frequency regions around the excited natural frequencies which dominated by characteristic defect frequency sidebands. Due to this characteristic dynamic behavior, different features and feature extraction methods have been proposed, including signal statistical analysis in the time domain, low and high-pass filtering, normalization, time derivation and integration of signals, Fourier transform and wavelet transform.

Some characteristic statistical time-domain feature parameters used are:

Kurtosis, which is a measure of whether the data are peaked or flat relative to a normal distribution. It is defined as

$$KU = \frac{\sum_{i=1}^N (x_i - m)^4}{(N-1)\sigma^4}, \quad (14)$$

where x_i , $i = 1, N$ are N samples of a signal $x(t)$, m is their mean value and σ is their standard deviation.

Skewness, which is a measure of symmetry or the lack of symmetry of the signal, is calculated by

$$SK = \frac{\sum_{i=1}^N (x_i - m)^3}{(N-1)\sigma^3} \quad (15)$$

Variance, which is a measure of dispersion of a waveform about its mean, is defined by

$$VR = \frac{\sum_{i=1}^N (x_i - m)^2}{(N-1)\sigma^2} \quad (16)$$

RMS of a signal, which indicates the severity of a bearing defect:

$$RMS = \sqrt{\frac{\sum_{i=1}^N x_i^2}{N}} \quad (17)$$

Crest factor, which is a measure of how much impacting, is occurring in time waveforms. It is defined as the ratio of the peak value to the RMS value with the DC component removed:

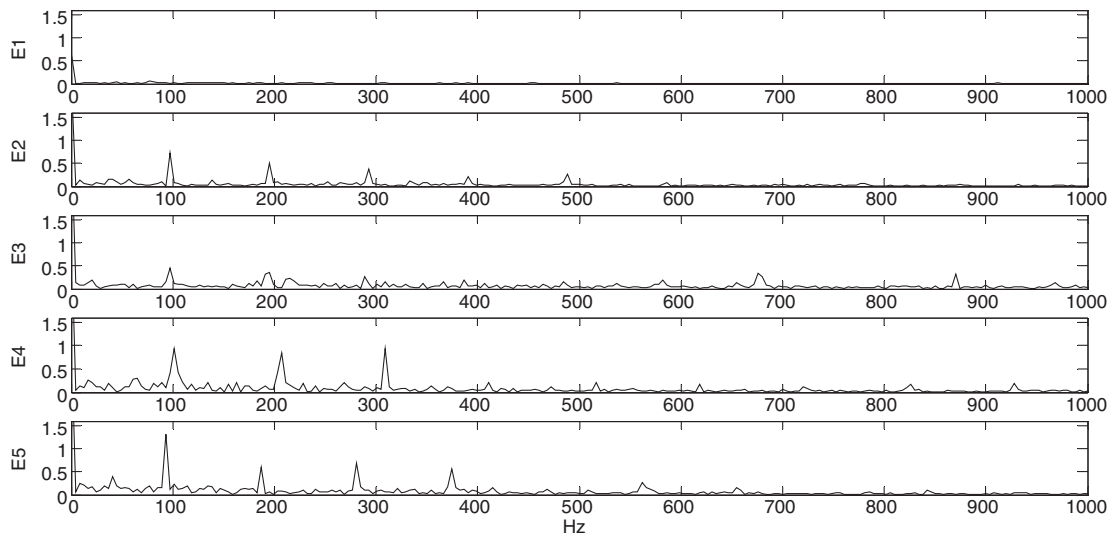


Fig. 7. Test case I: spectra of the demodulated signals.

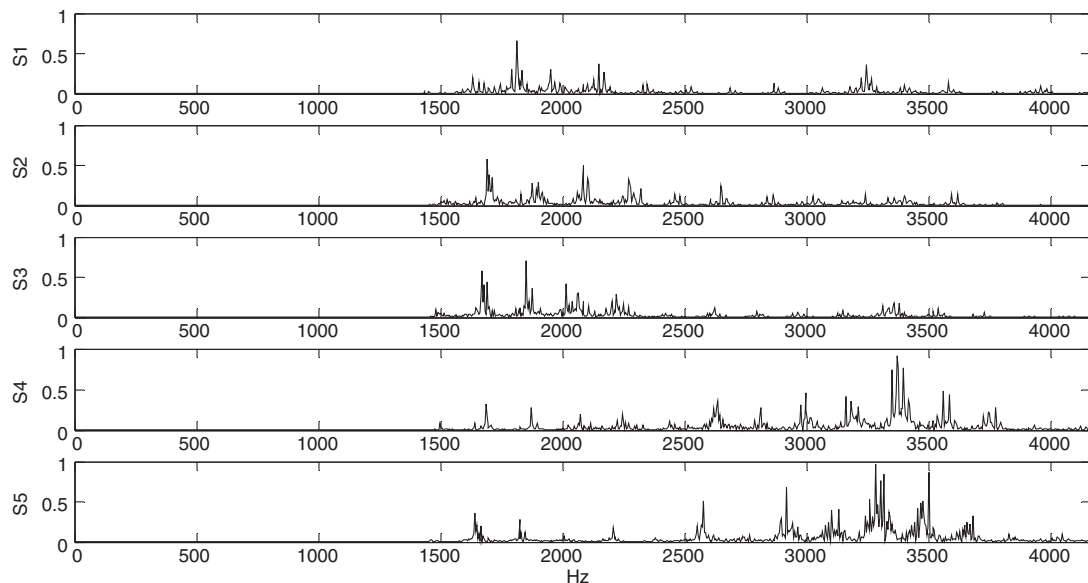


Fig. 8. Test case II: spectra of the raw signals.

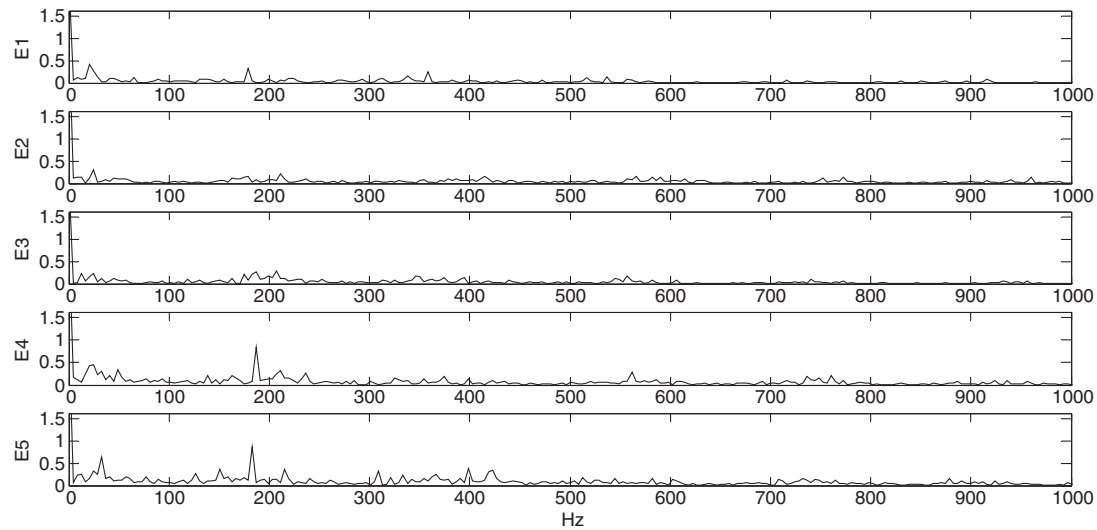


Fig. 9. Test case II: spectra of the demodulated signals.

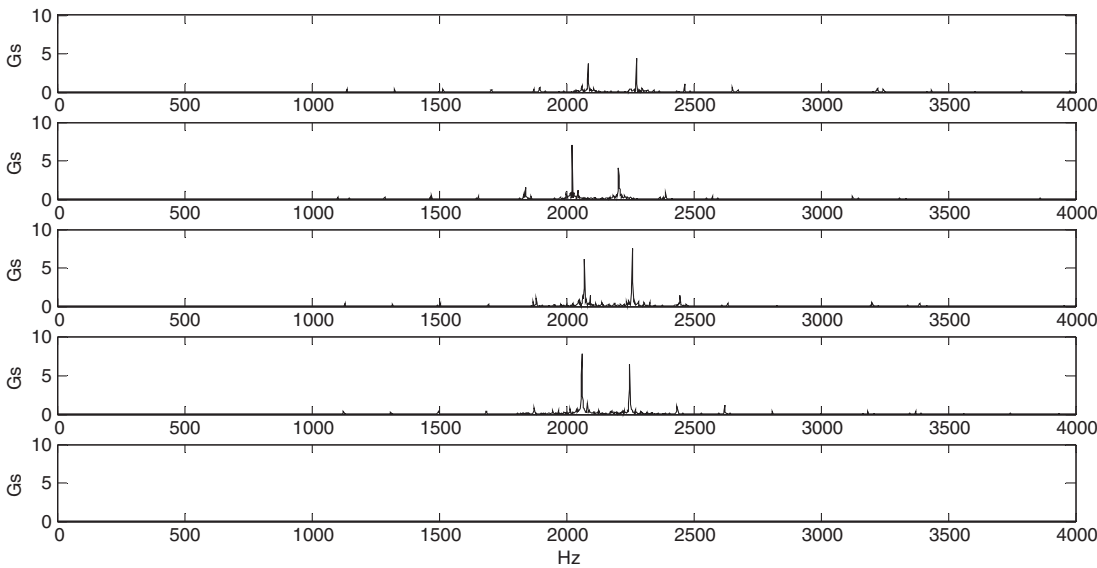


Fig. 10. Test case III: spectra of the raw signals.

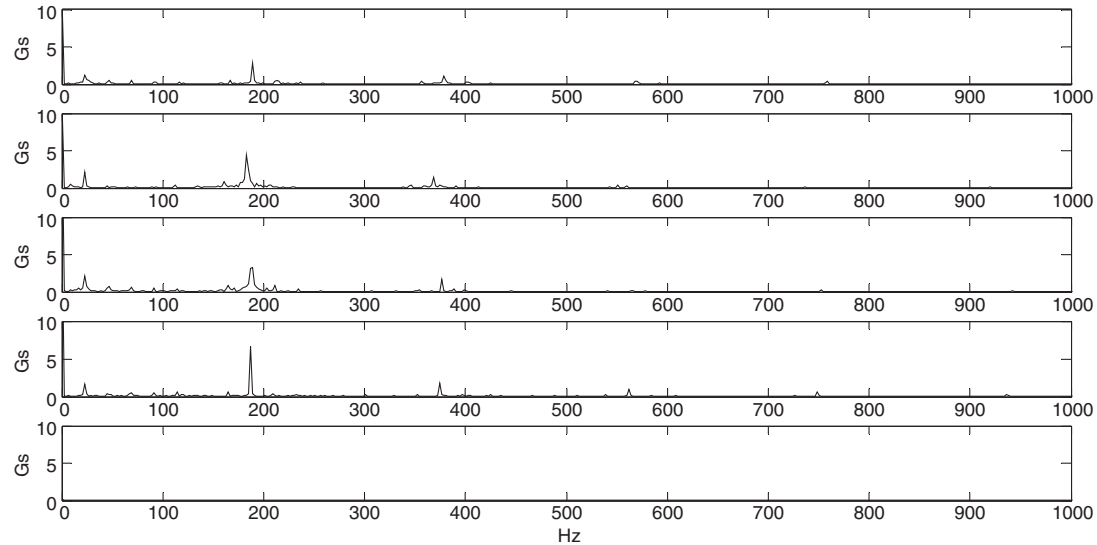


Fig. 11. Test case III: spectra of the demodulated signals.

$$CF = \frac{\max |x_i|}{\sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}} \quad (18)$$

Shape factor, which refers to a value that is affected by an object's shape but is independent of its dimensions. Is defined by the formula:

$$SF = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}}{\frac{1}{N} \sum_{i=1}^N |x_i|} \quad (19)$$

Impulse factor, which is defined by

$$IF = \frac{\max |x_i|}{\frac{1}{N} \sum_{i=1}^N |x_i|} \quad (20)$$

Clearance factor, which is defined by

$$CLF = \frac{\max |x_i|}{\left(\frac{1}{N} \sum_{i=1}^N \sqrt{|x_i|}\right)^2} \quad (21)$$

These parameters -and especially Crest factor and Kurtosis are considered to be especially appropriate for “spiky” signals, like the ones characterizing the behavior of defective rolling element bearings. Additionally to them, a number of other time-domain features may be considered, for example morphological index (MI) (Patargias, Yiakopoulos, & Antoniadis, 2006), standard deviation, mean, etc.

Alternative to the time-domain features, frequency-domain features are considered in this paper, since taking into account the energy at well specified frequency bands can lead to the more

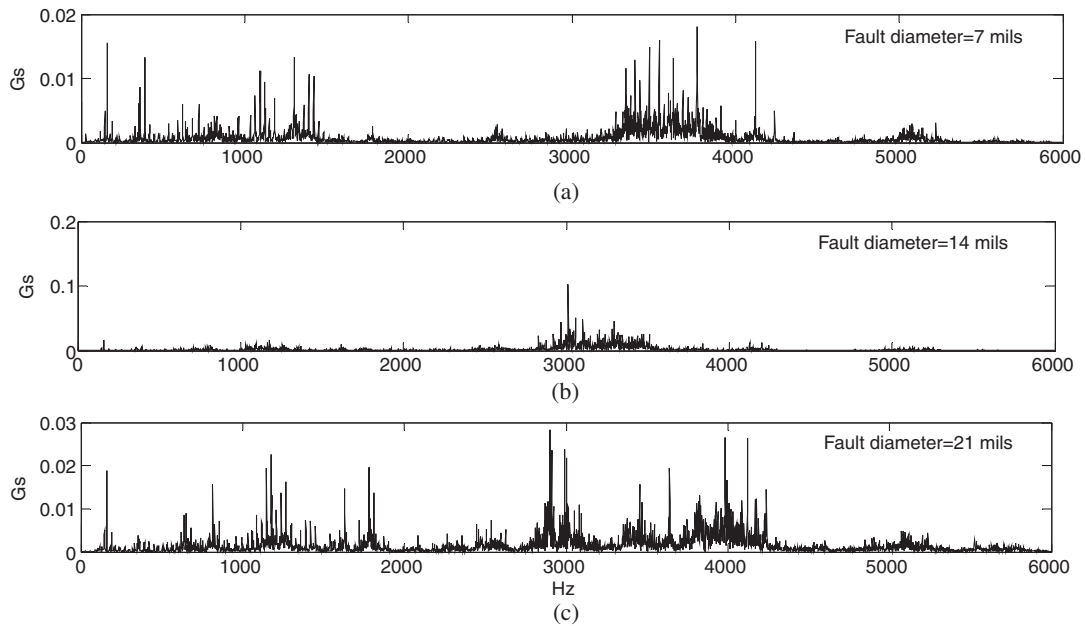


Fig. 12. Test case IV: spectra of the raw signals corresponding to an outer race defect with fault diameter: (a) 7 mils, (b) 14 mils and (c) 21 mils.

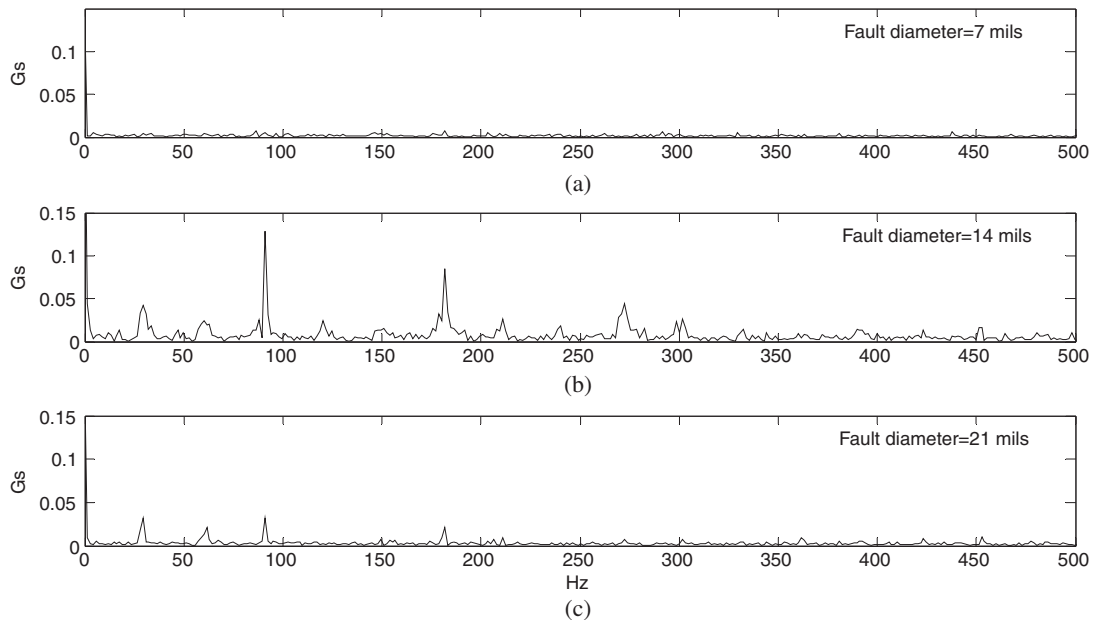


Fig. 13. Test case IV: spectra of the envelope signals corresponding to an outer race defect with fault diameter: (a) 7 mils, (b) 14 mils and (c) 21 mils.

accurate diagnosis and monitoring of the operating condition of rotating machinery.

However, prior to frequency analysis the signals are pre-processed using order analysis. Vibration signals of rotating machinery, especially during run-up or coast down are basically non-stationary signals due to the variations of the shaft rotation speed. As a result, the application of spectrum analysis directly to these kinds of signals will lead to the so-called ‘spectral smearing’. In order to overcome this problem, angular sampling theorem and order analysis method emerged. This kind of non-stationary signals in the time-domain turns to be stationary signals in the angle domain by means of equiangular sampling for reference shaft speed. Applying conventional spectrum analysis to these stationary signals in angle domain, a clear spectrum can be obtained, which is called “order

spectrum”. In the order spectrum, each line represents an order component including frequency of times of reference speed and corresponding amplitude as well.

Especially for rolling element bearings, due to the presence of the characteristic sidebands of the inner or outer race defect frequencies around the structural natural frequencies, demodulation methods are considered.

Demodulation or enveloping based methods offer a stronger and more reliable diagnostic potential, since they are based on a more solid physical background. The corresponding physical mechanism is described in (MacFadden & Smith, 1984a).

The goal of enveloping is first to isolate the measured signal in a relatively narrow frequency band around a specific excited natural frequency using a band-pass filter and then demodulate it to

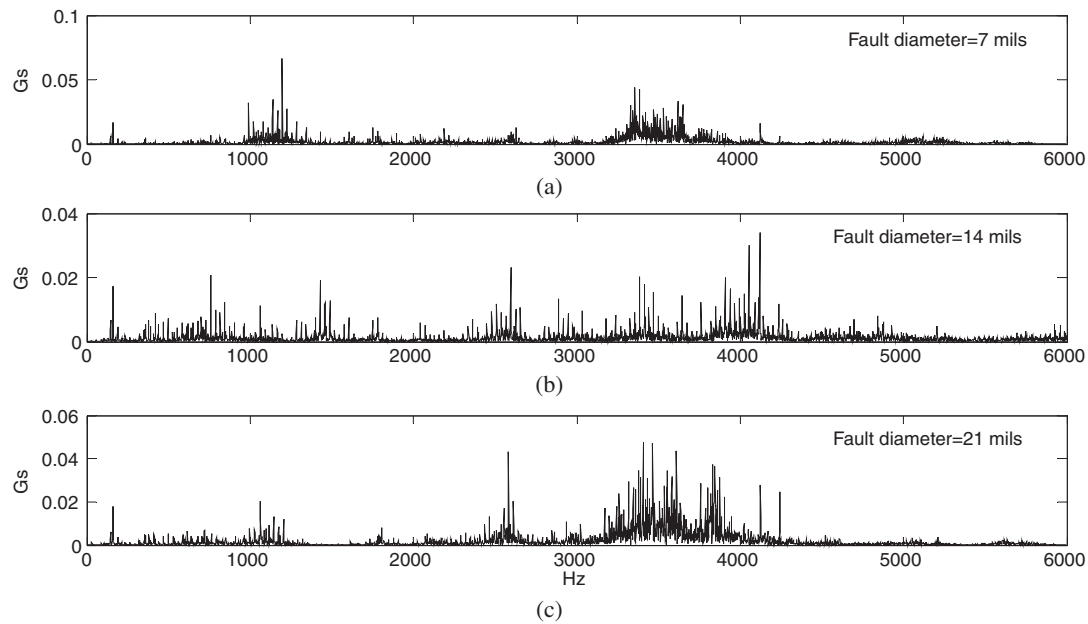


Fig. 14. Test case IV: spectra of the raw signals corresponding to an inner race defect with fault diameter: (a) 7 mils, (b) 14 mils and (c) 21 mils.

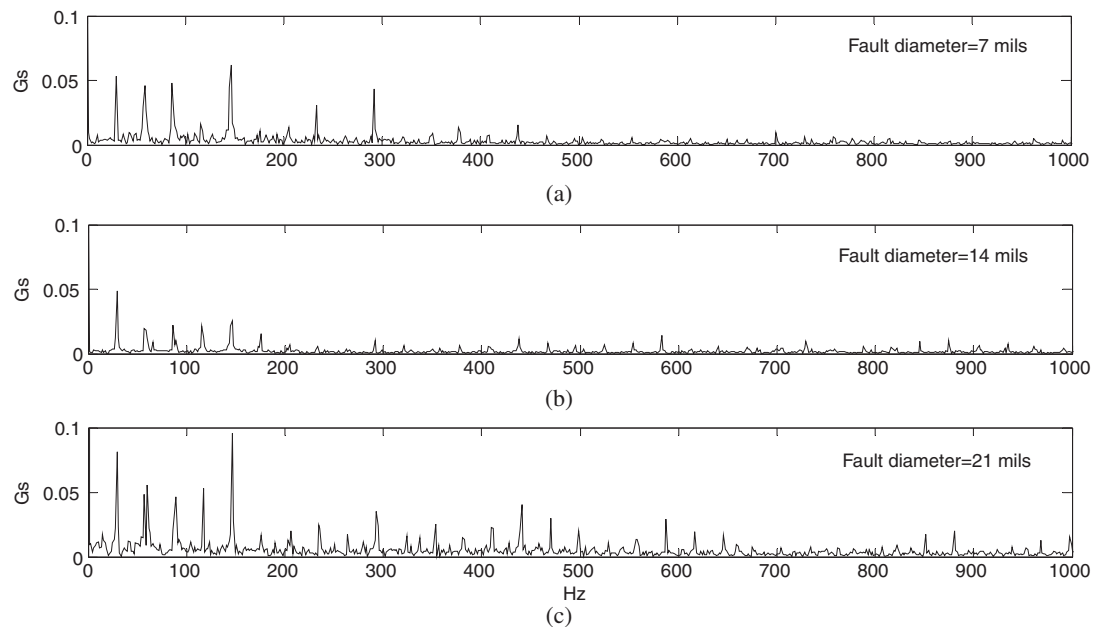


Fig. 15. Test case IV: spectra of the envelope signals corresponding to inner race defect with fault diameter: (a) 7 mils, (b) 14 mils and (c) 21 mils.

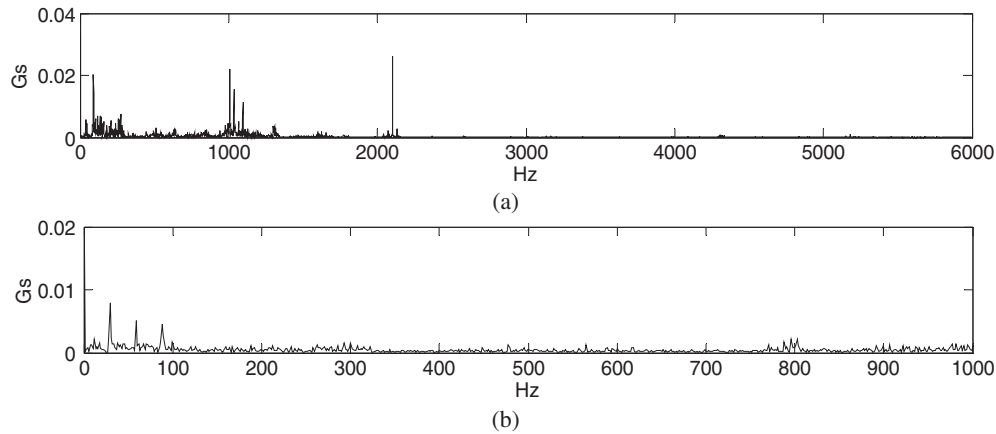


Fig. 16. Test case IV: spectra of the signal of a normal bearing: (a) raw signal and (b) demodulated signal.

produce a low-frequency signal, called the 'envelope'. Carrier signals are removed and this decreases the influence of irrelevant information. In order that the envelope signal is obtained, Hilbert transform is applied. The Hilbert envelope spectrum is given as

$$h(f) = \int_{-\infty}^{+\infty} \sqrt{x^2(t) + H^2[x(t)]} e^{-j2\pi ft} dt, \quad (22)$$

where $H[x(t)]$ is the Hilbert transform of a series $x(t)$.

$$H[x(t)] = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x(\tau)}{t - \tau} d\tau. \quad (23)$$

Afterwards, the vibration energies at characteristic frequency bands of the raw and of the envelope spectrum are properly selected. These vibration energies in the frequency domain, which are presented in Table 1, are the basis for the proposed indices to be used as input features. These frequencies, able to characterize the existence and the type of a bearing defect are: (A) The shaft rotating speed f_{shaft} and its harmonics, (B) The BPFO frequency and its harmonics, (C) The BPFI frequency and its harmonics, and (D) The high-frequency band.

The lowest frequency f_H , as used in the definition of the HFE index in Table 1, is selected equal to 1000 Hz, since this frequency is considered to characterize the typical lower limit of the high high-frequency bands, excited by the bearing impulses, while $f_s/2$ denotes the Nyquist frequency (f_s is the sampling rate). The enveloping procedure takes place in the frequency band range from f_H up to $f_s/2$.

Table 5

Values of the characteristic defect frequencies used in the dynamic bearing model of Section 3, Eq. (13) for all four test cases considered.

Defect frequency	Case I	Case II	Case III	Case IV
f_{shaft}	20	22	22.5	29.5
BPFO	95	136	185	90
BPFI	146	183	243	146

Table 6

Values of the parameters used in the dynamic bearing model of Section 3, Eq. (13) in order to produce the 5 sets ($k = 0, 1, \dots, 4$), of the simulated signals S_d^k used in industrial cases $r = \text{I, II and III}$ and corresponding to an outer ($d = o$) or an inner ($d = i$) race defect.

Model parameters	Outer race defect					Inner race defect				
	S_o^{r0}	S_o^{r1}	S_o^{r2}	S_o^{r3}	S_o^{r4}	S_i^{r0}	S_i^{r1}	S_i^{r2}	S_i^{r3}	S_i^{r4}
Amplitude (Gs)	15	0.5	10	30	40	10	0.5	5	15	20
Slip (%)	1	3	2.5	1.5	0.5	1	3	2.7	2	1
Noise level NL (%)	2	5	2.5	1	0.1	2	5	2.5	1	0.1
Natural frequency (KHz)	1.5	4	3	2	1.5	1.5	4	3	2	1.5

In order to obtain further insight on the behavior of the indices of Table 1, systematic runs of the dynamic bearing model of Eq. (13) are performed in order to generate simulated signals under different bearing faults and under different fault severity levels. The ranges of these parameters used in the simulations are defined in Table 2, and present typical values of these parameters, as expected in the majority of industrial and laboratory applications, involving defective rolling element bearings.

Signals corresponding to vibration response produced by an inner race defect are weaker than signals measured from a bearing with an outer race defect, due to the vibration physical mechanism. Parallel, the vibrations acquired from experimental test rigs are weaker than those in industrial measurements. For this reason, as it is observed in Table 2, the maximum time waveform amplitude for an outer race defect is considered to be twice as high as the corresponding amplitude in the case of an inner race defect. Likewise, the same applies to the simulations that correspond to industrial and laboratory signals.

The natural frequencies which are excited by a bearing fault on a test rig expand in a wide field due to the flexibility of its base structure. Based on the authors' experience, this high-frequency region varies from 1000 to 7000 Hz. On the contrary, as it is

Table 7

Values of the parameters used in the dynamic bearing model of Section 3, Eq. (13) in order to produce the 3 sets ($k = 0, 1, 2$), of the simulated signals S_d^{vk} used in the experimental test case $r = \text{IV}$, corresponding to an outer ($d = o$) or an inner ($d = i$) race defect.

Model parameters	Outer race defect			Inner race defect		
	S_o^{IV0}	S_o^{IV1}	S_o^{IV2}	S_i^{IV0}	S_i^{IV1}	S_i^{IV2}
Amplitude (Gs)	0.5	15	20	0.5	5	10
Slip (%)	3	1.5	1.5	3	2.4	1.2
Noise level NL (%)	5	2.5	3	5	3	2
Natural frequency (KHz)	4	2	3	4	3	2.5

observed in the Table 2, the natural frequencies of the industrial signals lie in a narrow spectrum field due to the rigid base of the machines (1500–5000 Hz). The noise level NL fluctuates from 0.1% to 5% in all the cases. The slip between the impacts is assumed to vary from 0.5% to 3% in all the cases, as well. The standard deviation for the signal amplitude is typically selected up to 5%.

By varying the model parameters of Eq. (13) in the ranges presented in Table 2, the calculated ranges and distributions of the vibration energies of Table 1 is presented in Figs. 3 and 4.

Fig. 3 shows the energy variations of the vibration energy in the frequency bands of: (A) The sum $SSHR/SSHE$ of the shaft harmonic spectral peaks and (B) The energy HFE at the high-frequency field. Due to the excitation of natural frequencies above 1000 Hz, the

spectrum energy is increased in the high-frequency regions. This explains why high-frequency energy (HFE) is higher than the total vibration energy $SSHR/SSHE$ at the shaft harmonics.

Fig. 4 presents the energy variations of the frequency bands of the: (A) the spectral peaks SKX at the shaft rotating speed, (B) the spectral peaks SKO at the BPFO harmonics, (C) the spectral peaks SKI at the BPFI harmonics, (D) their corresponding sums $SSHE$, $SORH$, $SIRH$ and $SIORH$. As expected, the first harmonic (SIX , SIO , SII) of each characteristic defect frequency contains higher amounts of energy than the rest of the harmonic components (SKX , SKO , SKI ; $K = 2, 3, 4$). The maximum total energy of the BPFO harmonics is higher than the relative total energy of the BPFI harmonics. This happens because the vibration response

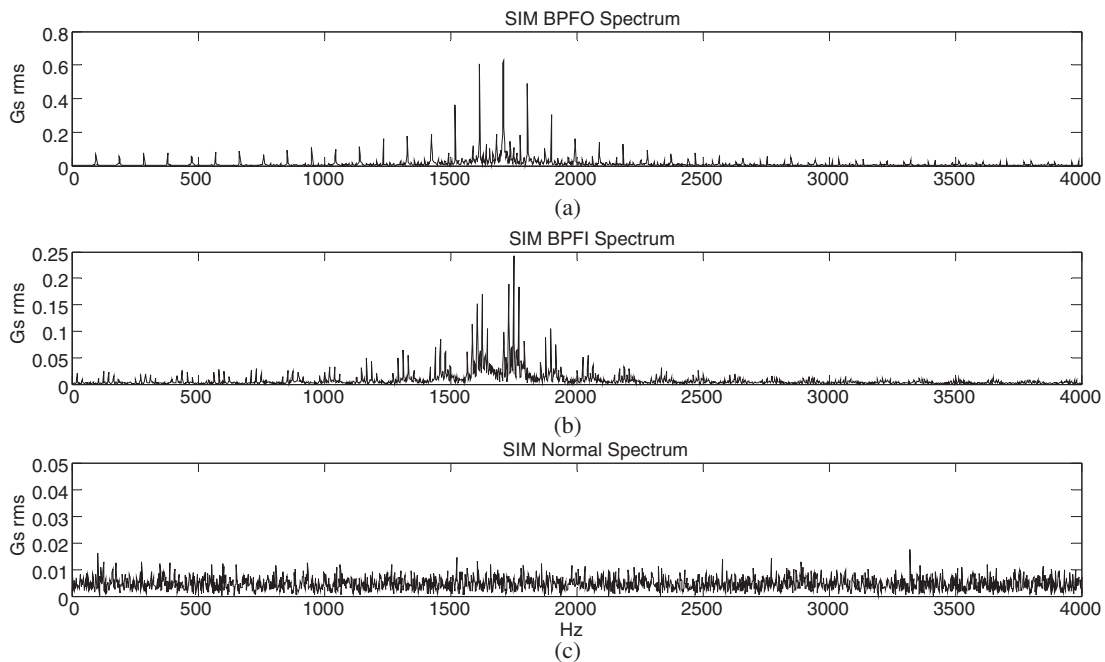


Fig. 17. Test case I: spectra of simulated raw signals: (a) outer race defect S_o^{f0} , (b) inner race defect S_i^{f0} and (c) normal bearing S_n^{f0} .

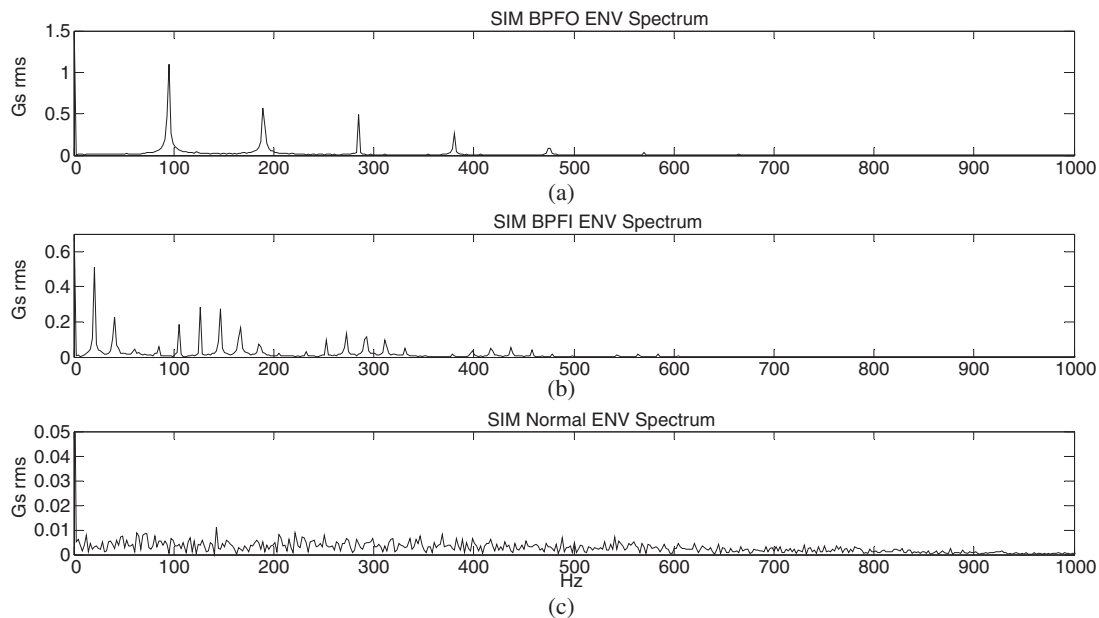


Fig. 18. Test case I: spectra of the envelopes of simulated raw signals: (a) outer race defect S_o^{f0} , (b) inner race defect S_i^{f0} and (c) normal bearing S_n^{f0} .

of a defective bearing with an inner race fault is weaker than the vibration response corresponding to an outer race defect.

In view of the analysis presented in Figs. 3 and 4, it is concluded that for the decision whether a bearing is defective or not, a limited number of indices can be used, as defined in Table 3. Contrarily, for the decision on the type of the defect, an expanded feature set must be used. Such a set of indices is presented in Table 4. The total number of the features presented in the Table 4 depends on the number K of the frequency harmonics to be used. According to the results in Figs. 3 and 4, a number of $K = 4$ harmonics is satisfactory, resulting to a feature vector of a length equal to 30. Especially

for the shaft rotating speed, the number of the shaft rotating speed harmonics is limited to a $(BPFO - f_{\text{shaft}})$ frequency and to a maximum number of $K = 4$.

The purpose of using relative energy indices as integer values in the form of Table 4 is to

- Magnify and further exhibit the differentiation between an outer and an inner race fault.
- Neutralize the effects of the ambivalent vibration energies estimated at characteristic frequency bands. For example, a bearing with an inner race defect is considered. The S4I value is supposed to be equal to 0.95 and the S4O value is

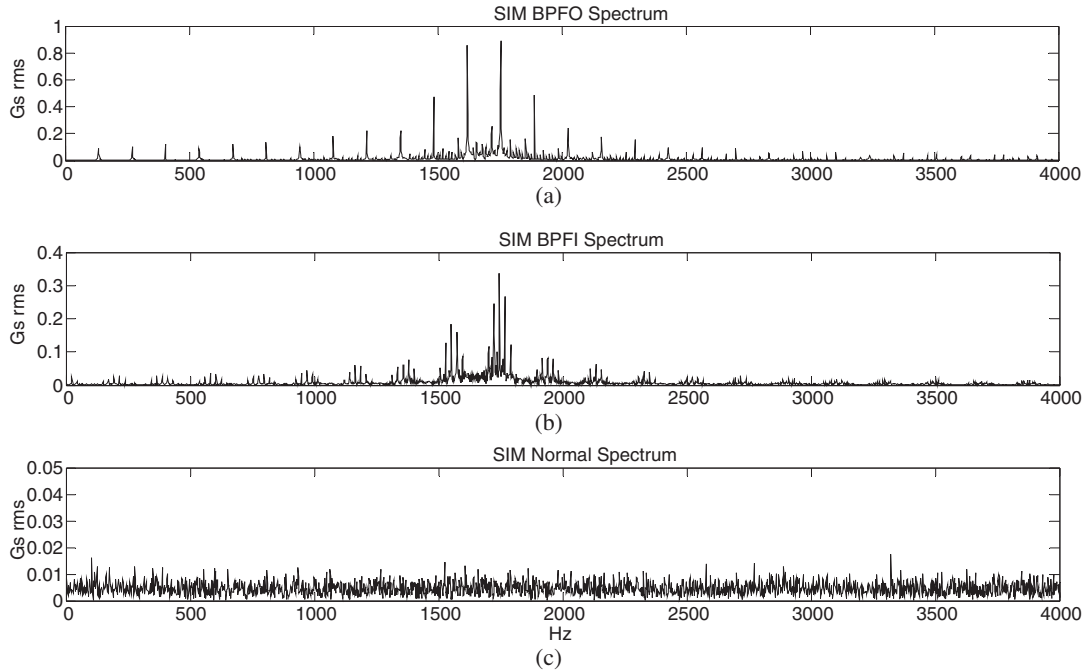


Fig. 19. Test case II: spectra of simulated raw signals: (a) outer race defect S_o^{II0} , (b) inner race defect S_i^{II0} and (c) normal bearing S_n^{II0} .

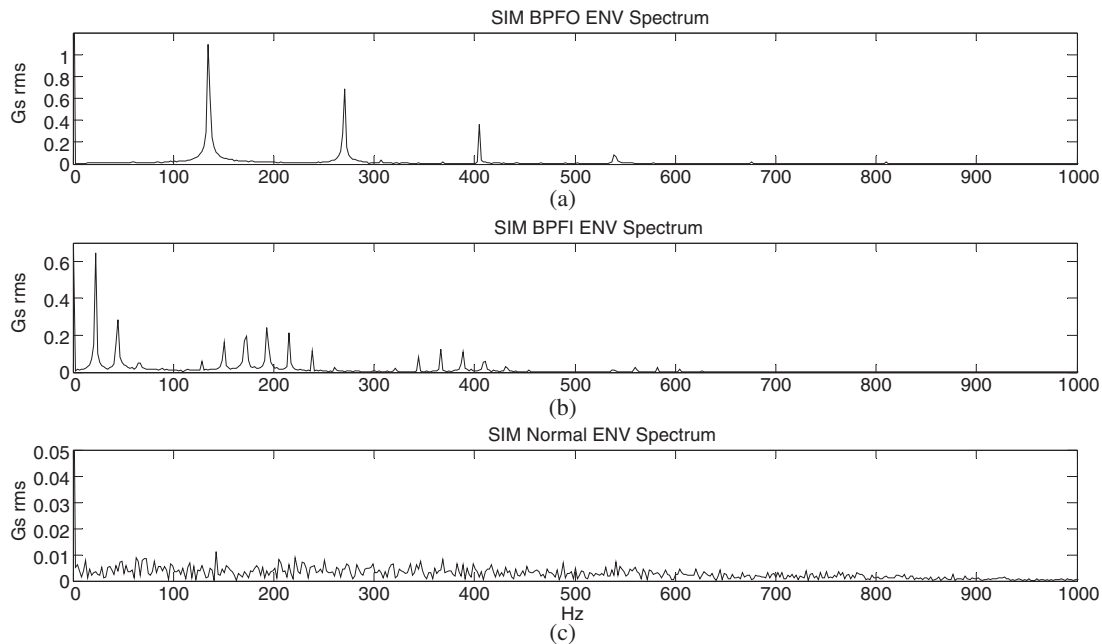


Fig. 20. Test case II: spectra of the envelopes of simulated raw signals: (a) outer race defect S_o^{II0} , (b) inner race defect S_i^{II0} and (c) normal bearing S_n^{II0} .

supposed to be equal to 1.05. The S4O value in many circumstances results to be higher than the S4I because the 4th harmonic of the BPFO coincides in many cases with the 2nd or the 3rd harmonic of the BPFI or even with a sideband of them. Thus, the ratio S4O/S4I is equal to 1.10 and its inverse is equal to 0.90, indicating erroneously an outer race defect. However, the interdependent R4OI and R4IO relative energy integer indices are equal to 1 counteracting each other without affecting any more the outcome of the clustering procedure.

4.2. Proposed classification algorithm

An overall flow chart of the proposed clustering algorithm is presented in Fig. 5. The first step of the algorithm is the selection of proper initial centers, since the performance of *K*-means clustering is considered to be quite sensitive to this choice. Since measurements or experimental data from an actual machine under defective conditions, especially in an industrial environment are quite difficult (if possible at all) to obtain, a key innovation of the proposed approach is that the initial centers of the *K*-means

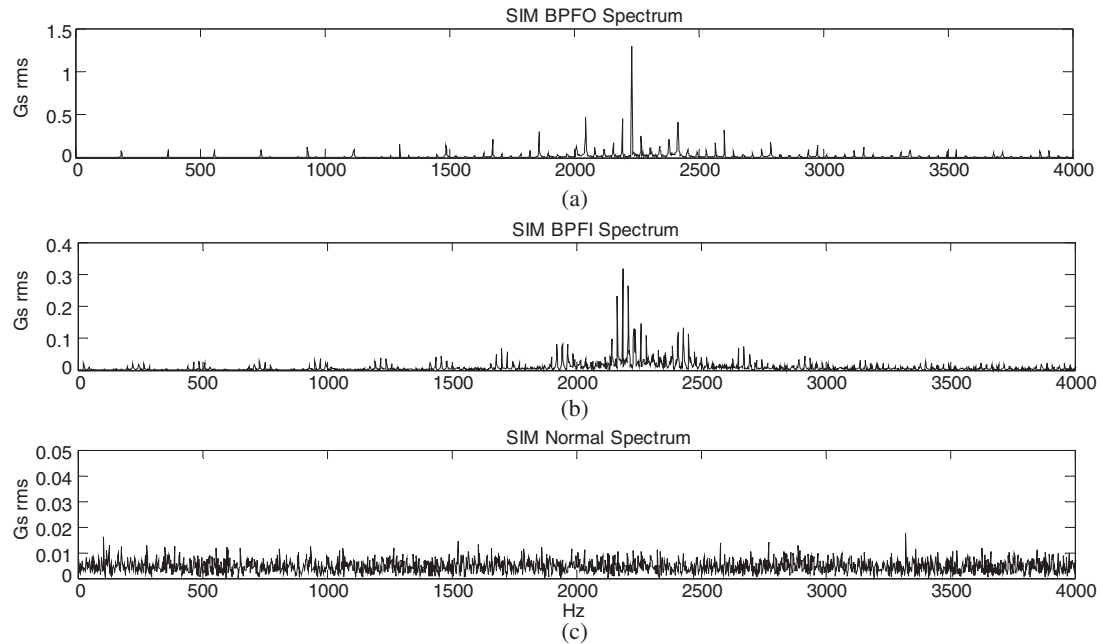


Fig. 21. Test case III: spectra of simulated raw signals: (a) outer race defect S_o^{III} , (b) inner race defect S_i^{III} and (c) normal bearing S_n^{III} .

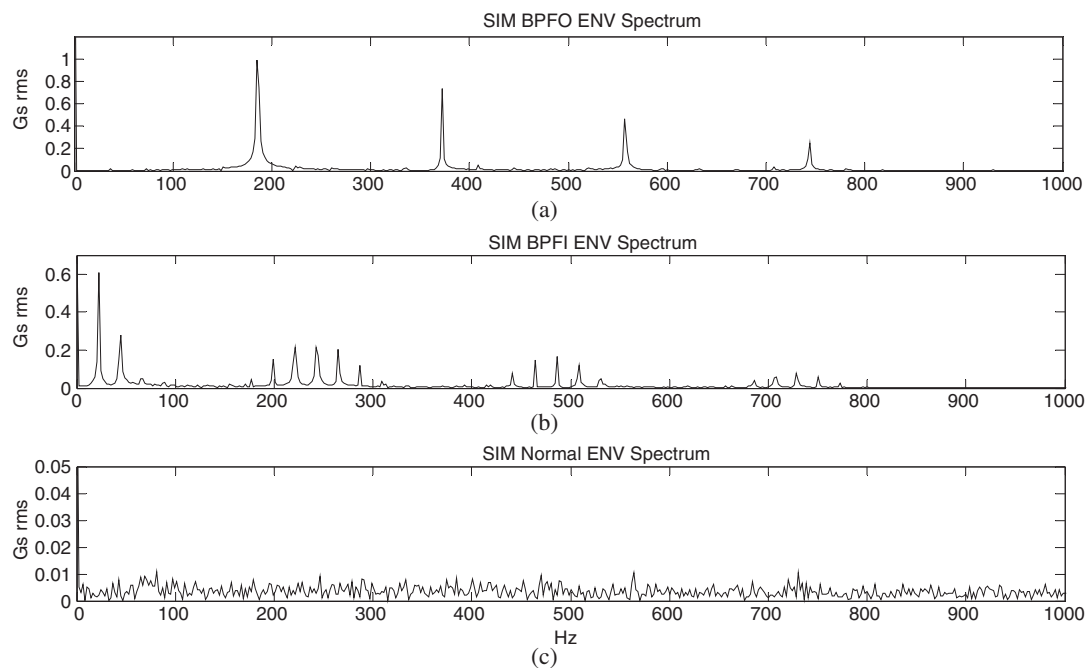


Fig. 22. Test case III: spectra of the envelopes of simulated raw signals: (a) outer race defect S_o^{III} , (b) inner race defect S_i^{III} and (c) normal bearing S_n^{III} .

clustering are produced by simulated signals. These signals, characterizing the vibration response of defective rolling element bearings, are generated by the well established and widely accepted dynamic model, described in Eq. (13). Likewise, the signal presenting a normal condition is simulated by a white noise signal.

This approach allows for the production of only one final cluster solution. Besides, this approach achieves to categorize (label) the produced clusters against the unsupervised structure of K -means clustering. Thus, the proposed approach for the calculation of the initial centers acts in a way as a training session and signifies the object categories.

Then, signals measured on the actual machine on the industrial environment are directly applied, after the proposed preprocessing and feature extraction approach of Section 4.1. The classification then proceeds as a two-stage approach. At the first stage, the objective of the approach is to identify if a fault of any type exists or not. For this reason, only three features are used, as presented in Table 3. At the second stage the type of the fault (inner or outer race bearing defect) is identified. In this case, the integer set of the relative energy features of Table 4 is used. In both classification stages, the correlation metric measure is used, as defined in Eq. (4), since this measure has proven to provide the best results.

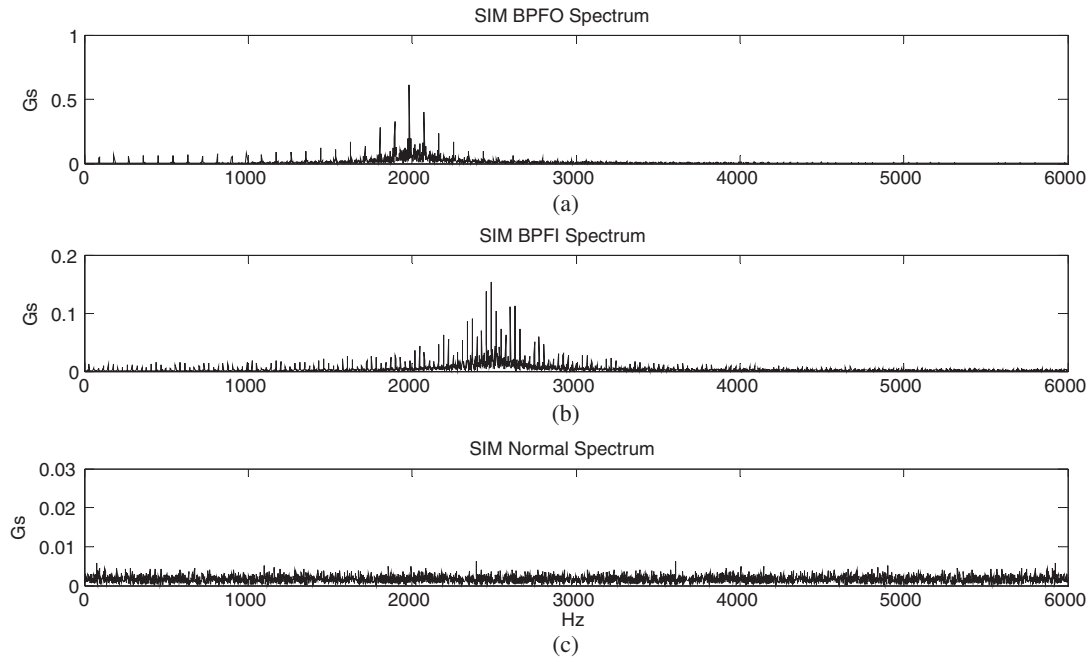


Fig. 23. Test case IV: spectra of simulated raw signals: (a) outer race defect S_o^{IV0} , (b) inner race defect S_i^{IV0} and (c) normal bearing S_n^{IV0} .

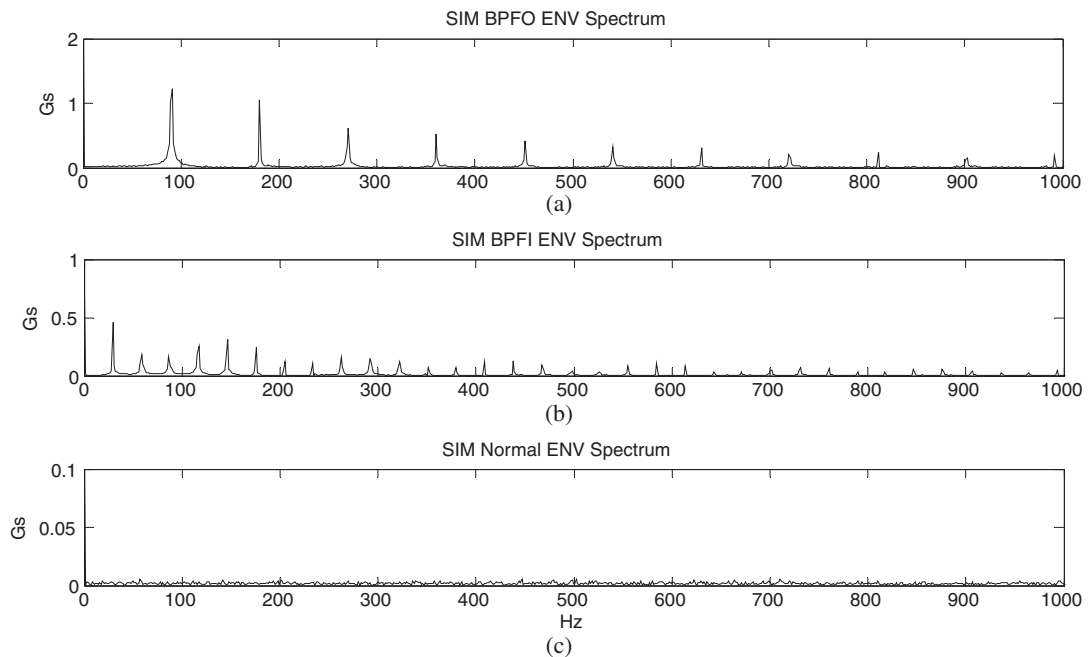


Fig. 24. Test case IV: spectra of the envelopes of simulated raw signals: (a) outer race defect S_o^{IV0} , (b) inner race defect S_i^{IV0} and (c) normal bearing S_n^{IV0} .

5. Applications and evaluation

5.1. Description of the test cases considered

The evaluation of the proposed approach has been performed: (A) in three different test cases from three different industrial installations. Each case includes 5 successive

measurements from the vibration response of defective rolling element bearings in three different industrial installations and (B) in an internationally established bearing data set benchmark, which is obtained from a laboratory test rig under three different operating conditions: (i) normal bearing, (ii) bearing with outer race fault and (iii) bearing with inner race fault (Loparo).

Table 8

Results from the application of the frequency domain based *K*-means clustering for the 1st stage classification using correlation distance (normal or defective bearing).

Case <i>r</i>	Bearing type	Defect type	Result for different centroids				
			C_1^{r0}	C_1^{r1}	C_1^{r2}	C_1^{r4}	C_1^{r5}
I	2312	Outer race defect (5 signals)	✓	✓	✓	✓	✓
II	22205	Inner race defect (5 signals)	✓	✓	✓	✓	✓
III	22228	Inner race defect (4 signals) Normal bearing (1 signal)	✓	✓	✓	✓	✓
IV	6203-2RS JEM	Outer race defect (3 signals) Inner race defect (3 signals) Normal bearing (1 signal)	✓	✓	✓	NA	NA

Table 9

Results from the application of the frequency domain based *K*-means clustering for the 1st stage classification using cosine distance (normal or defective bearing).

Case <i>r</i>	Bearing type	Defect type	Result for different centroids				
			C_1^{r0}	C_1^{r1}	C_1^{r2}	C_1^{r4}	C_1^{r5}
I	2312	Outer race defect (5 signals)	✓	X	X	X	✓
II	22205	Inner race defect (5 signals)	✓	X	X	X	X
III	22228	Inner race defect (4 signals) Normal bearing (1 signal)	✓	✓	✓	✓	✓
IV	6203-2RS JEM	Outer race defect (3 signals) Inner race defect (3 signals) Normal bearing (1 signal)	✓	X	✓	NA	NA

Table 10

Results from the application of the frequency domain based *K*-means clustering for the 1st stage classification using Euclidean or Cityblock distance (normal or defective bearing).

Case <i>r</i>	Bearing type	Defect type	Result for different centroids				
			C_1^{r0}	C_1^{r1}	C_1^{r2}	C_1^{r4}	C_1^{r5}
I	2312	Outer race defect (5 signals)	X	X	X	X	X
II	22205	Inner race defect (5 signals)	X	X	X	X	X
III	22228	Inner race defect (4 signals) Normal bearing (1 signal)	X	X	X	X	X
IV	6203-2RS JEM	Outer race defect (3 signals) Inner race defect (3 signals) Normal bearing (1 signal)	X	X	X	NA	NA

Table 11

Detailed results from the application of the frequency domain based *K*-means clustering for the 1st stage classification using correlation distance for all the *k* sets of centroids (normal or defective bearing). Class 1 denotes a normal bearing and Class 2 denotes a defective bearing. S_n^k denotes a simulated signal of a normal bearing, S_i^k/S_o^k simulated signals from a bearing with an inner/outer race defect and M1–M5/M7 the successive measured signals in each case *r*.

Case <i>r</i>	Classification details									
I	Signal	S_n^k	S_o^k	S_i^k	M1	M2	M3	M4	M5	
	Class	1	2	2	2	2	2	2	2	
II	Signal	S_n^{lk}	S_o^{lk}	S_i^{lk}	M1	M2	M3	M4	M5	
	Class	1	2	2	2	2	2	2	2	
III	Signal	S_n^{llk}	S_o^{llk}	S_i^{llk}	M1	M2	M3	M4	M5	
	Class	1	2	2	2	2	2	2	1	
IV	Signal	S_n^{IVk}	S_o^{IVk}	S_i^{IVk}	M1	M2	M3	M4	M5	M6
	Class	1	2	2	1	2	2	2	2	2

5.1.1. Test case I

The vibration data set in test case I includes signals measured on an industrial fan with a defective rolling element bearing on the outer race. The shaft rotation speed during the measurements was around 1200 rpm (20 Hz). The bearing monitored is of SKF type 2312, with a BPFO frequency equal to 4.72 times the shaft rotation speed, leading to a theoretical estimation of the expected BPFO frequency around 94 Hz. Each signal is 4096 samples long and recorded with a sampling rate of 8.33 KHz. The time interval between two successive measurements is about 2 months. The spectra of the signals are presented in Fig. 6 and the spectra of their envelopes are presented in Fig. 7.

5.1.2. Test case II

The measurements in test case II were performed on a rolling element bearing of an industrial pump, presenting an inner race fault. The shaft rotation speed during the measurements was around 1320 rpm (22 Hz). The bearing type is of 22205 EK type, with a BPFI frequency equal to 8.32 times the shaft frequency, leading to a theoretical estimation of the expected BPFI frequency around 183 Hz. Each signal is 2048 samples long and recorded with a sampling rate of 8.33 KHz. The spectra of the signals are

presented in Fig. 8 and the spectra of their envelopes are presented in Fig. 9.

5.1.3. Test case III

The data set in test case III is obtained from a spherical bearing at an industrial compressor presenting an outer race fault. The shaft rotation speed during the five measurements was around 1350 rpm (22.5 Hz). The bearing monitored is of 22228 cck/w3 type, with a BPFO frequency equal to 8.22 times the shaft frequency, leading to a theoretical estimation of the expected BPFO around 185 Hz. Each signal is 2048 samples long and recorded with a sampling rate of 8.33 KHz. This test case presents a fast progress of the bearing defect. Each of the first four measurements was recorded with a time interval about a week from the previous one. The fifth signal was measured after the shutdown of the machine and the replacement of the defective bearing. The spectra of the signals and their envelopes are presented in Figs. 10 and 11.

5.1.4. Test case IV

The vibration data used at the test case IV have been acquired from a ball bearing test data set obtained from *** Case Western

Table 12

Results from the application of the frequency domain based *K*-means clustering for the 2nd stage classification using correlation distance (outer or inner race defect).

Case <i>r</i>	Bearing type	Defect type	Result for different centroids				
			C_2^0	C_2^1	C_2^2	C_2^4	C_2^5
I	2312	Outer race defect (5 signals)	✓	✓	✓	✓	✓
II	22205	Inner race defect (5 signals)	✓	✓	✓	✓	✓
III	22228	Inner race defect (4 signals) Normal bearing (1 signal)	✓	✓	✓	✓	✓
IV	6203-2RS JEM	Outer race defect (3 signals) Inner race defect (3 signals) Normal bearing (1 signal)	✓	✓	✓	NA	NA

Table 13

Results from the application of the frequency domain based *K*-means clustering for the 2nd stage classification using cosine distance (outer or inner race defect).

Case <i>r</i>	Bearing type	Defect type	Result for different centroids				
			C_2^0	C_2^1	C_2^2	C_2^4	C_2^5
I	2312	Outer race defect (5 signals)	X	✓	✓	✓	✓
II	22205	Inner race defect (5 signals)	✓	✓	✓	✓	✓
III	22228	Inner race defect (4 signals) Normal bearing (1 signal)	✓	✓	✓	✓	✓
IV	6203-2RS JEM	Outer race defect (3 signals) Inner race defect (3 signals) Normal bearing (1 signal)	X	X	X	NA	NA

Table 14

Results from the application of the frequency domain based *K*-means clustering for the 2nd stage classification using Euclidean and Cityblock distance (outer or inner race defect).

Case <i>r</i>	Bearing type	Defect type	Result for different centroids				
			C_2^0	C_2^1	C_2^2	C_2^4	C_2^5
I	2312	Outer race defect (5 signals)	X	X	X	X	X
II	22205	Inner race defect (5 signals)	✓	✓	✓	✓	✓
III	22228	Inner race defect (4 signals) Normal bearing (1 signal)	X	X	X	X	X
IV	6203-2RS JEM	Outer race defect (3 signals) Inner race defect (3 signals) Normal bearing (1 signal)	X	X	X	NA	NA

Table 15
Detailed results from the application of the frequency domain based *K*-means clustering for the 2nd stage classification using correlation distance for all the *k* sets of centroids (outer or inner race defect). Class 1 denotes an outer race defect and Class 2 an inner race defect. S_i^{lk} denotes a simulated signal with inner race defect, S_o^{lk} denotes a simulated signal with an outer race defect, and M1 to M4/M5/M6 the successive measured signals in each case *r*.

Case r	Classification details								
I	Signal	S_o^{lk}	S_i^{lk}	M1	M2	M3	M4	M5	
	Class	1	2	1	1	1	1	1	
II	Signal	S_o^{llk}	S_i^{llk}	M1	M2	M3	M4	M5	
	Class	1	2	2	2	2	2	2	
III	Signal	S_o^{lllk}	S_i^{lllk}	M1	M2	M3	M4		
	Class	1	2	1	1	1	1		
IV	Signal	S_o^{lvk}	S_i^{lvk}	M1	M2	M3	M4	M5	M6
	Class	1	2	1	1	1	2	2	2

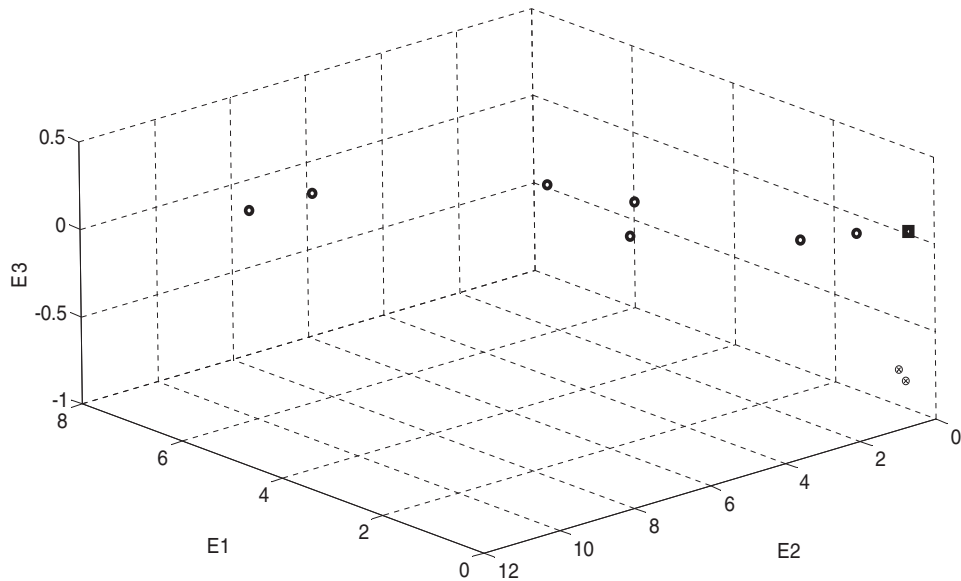


Fig. 25. Test case I – stage 1: classification results for frequency-domain features. The final centers are shown with the marker '⊗'. The outputs characterizing a defective classification are presented with the marker 'o'. The marker '□' represents the simulated normal condition output, which is used as initial center and input to the algorithm.

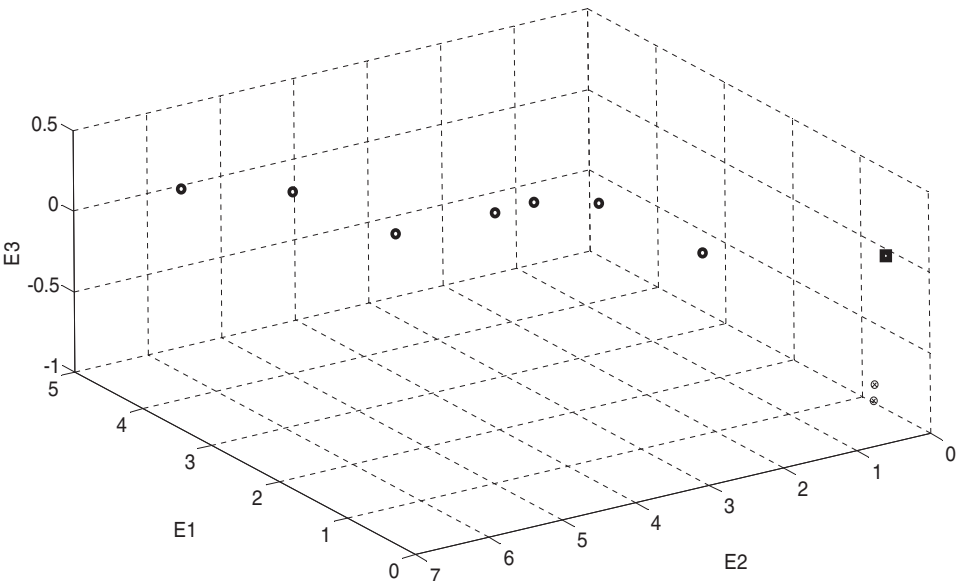


Fig. 26. Test case II – stage 1: classification results for frequency-domain features. The final centers are shown with the marker '⊗'. The outputs characterizing a defective classification are presented with the marker 'o'. The marker '□' represents the simulated normal condition output, which is used as initial center and input to the algorithm.

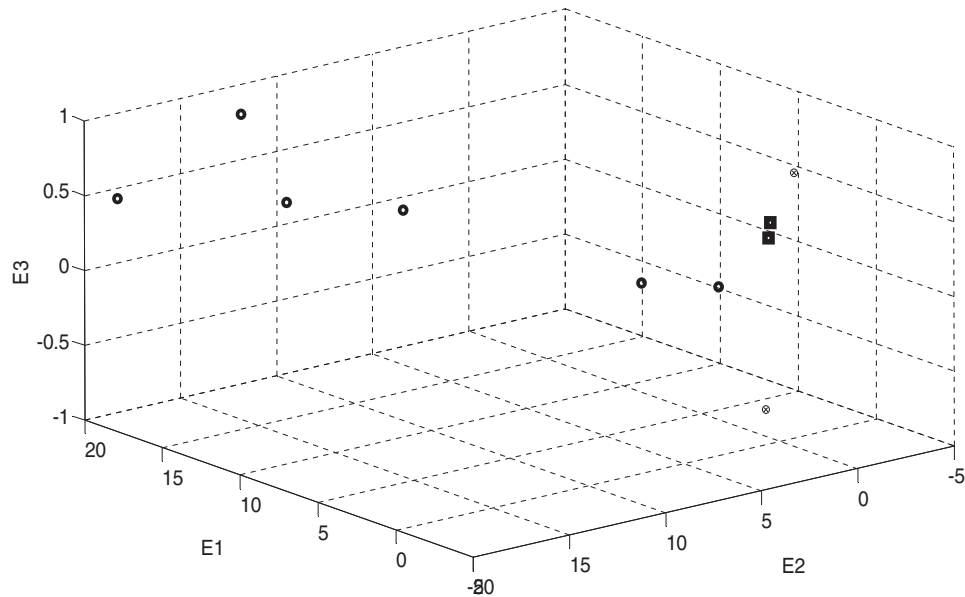


Fig. 27. Test case III – stage 1: classification results for frequency-domain features. The final centers are shown with the marker '⊗'. The outputs characterizing a defective classification are presented with the marker 'o'. The marker '□' represents the simulated normal condition output, which is used as initial center and input to the algorithm.

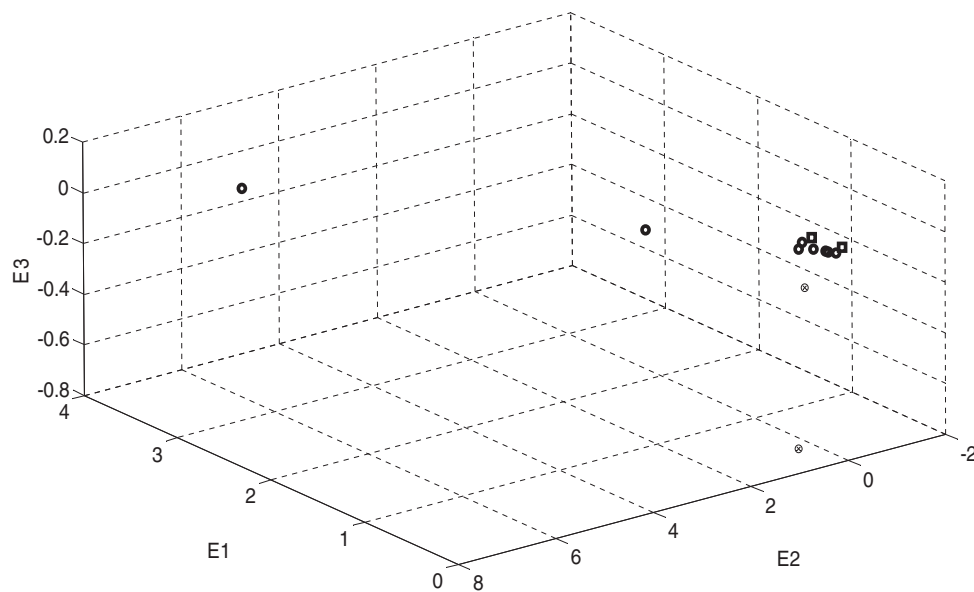


Fig. 28. Test case IV – stage 1: Classification results for frequency-domain features. The final centers are shown with the marker '⊗'. The outputs characterizing a defective classification are presented with the marker 'o'. The marker '□' represents the measured and the simulated normal condition output, which is used as initial center and input to the algorithm.

Table 16

Set A of statistical time-domain feature parameters used for comparison in time domain based *K*-means clustering approach.

Parameter	Symbol
Mean	μ
Root mean square	<i>RMS</i>
Variance	μ_2
Skewness	μ_3
Kurtosis	μ_4
5th Central moment	μ_5
6th Central moment	μ_6
7th Central moment	μ_7
8th Central moment	μ_8
9th Central moment	μ_9

Table 17

Set B of time-domain indices used for comparison in time domain based *K*-means clustering method.

Parameter	Symbol
Skewness	<i>SK</i>
Kurtosis	<i>KU</i>
Crest factor	<i>CF</i>
Clearance factor	<i>CLF</i>
Shape factor	<i>SF</i>
Impulse factor	<i>IF</i>

Reserve University Bearing Data Center Website ([Loparo](#)). The test stand consists of a 1 HP electric motor, a torque transducer, a dynamometer and control electronics. Digital data was collected at

12,000 samples per second. Each signal is 8192 samples long. The motor speed was 1772 rpm (29.53 Hz).

The bearings monitored are deep groove ball bearings manufactured by SKF. The fan end bearing is a 6203-2RS JEM with a BPFI and a BPFO frequency equal to 4.9469 and 3.0530 times the shaft frequency respectively, leading to theoretical estimations of the expected BPFO and BPFI frequencies equal to 90.16 and 146.08 Hz, respectively.

Single point faults are introduced into the bearings using electro-discharge machining with fault diameters of 7, 14 and 21 mils (1 mil = 0.001 inches) at both the inner race and at the outer race. The bearing data set was obtained from the experimental test rig under three different operating conditions: (a) normal condition, (b) outer race fault (3 different fault level cases) and (c) inner race fault (3 different fault level cases). The spectra of the raw signals and of their corresponding envelopes are presented in Figs. 12–16.

5.2. Implementation description

For the implementation of the method the initial centers of the clusters must be first calculated. In order to evaluate the robustness of the method to the selection of the initial centers, a number different set of initial centers are used in all test cases. For the 1st stage of the classification, the first centroid of the set is calculated from a white noise signal, which represents the normal bearing condition and the second centroid is calculated from a simulated signal, which represents a bearing with an outer or inner race defect. For the 2nd stage of the classification, the first centroid of the set is calculated from a simulated signal which corresponds to a bearing with an outer race defect and the second centroid is calculated from a simulated signal which corresponds to a bearing with an inner race defect. In order to compactly describe the implementation procedure and the results obtained, the following notation is used:

S_d^{rk} : Simulated signal, corresponding to a bearing with

defect type d , used in the k th set of initial centers of test case r

E_d^{rk} :

Envelope of the signal S_d^{rk} , obtained by the demodulation procedure described in Section 4.1

F_d^{rk} :

Feature set, derived from the signal S_d^{rk} according to the definitions of Table 3, used in the 1st stage of K -means based classification approach

R_d^{rk} :

Feature set or relative energies, derived from the signal S_d^{rk} according to the definitions of Table 4, used in the 2nd stage of K -means based classification approach

$C_1^{rk} =$

$\{F_n^{rk}, F_d^{rk}\}$:

k th set of initial centers used in the 1st stage of classification of test case r . It consists from a feature set F_n^{rk} derived from a white noise signal corresponding to a normal bearing and a feature set F_d^{rk} derived from a simulated signal, corresponding to a bearing with an inner or outer race defect

$C_2^{rk} =$

$\{R_o^{rk}, R_i^{rk}\}$:

k th set of initial centers used in the 2nd stage of classification of test case r . It consists from a feature set R_o^{rk} derived from a simulated signal, corresponding to a bearing with an outer race defect, and from a feature set R_i^{rk} derived from a simulated signal, corresponding to a bearing with an inner race defect

d :

Subscript indicating the type of defect: $d = n$ indicates a normal bearing, $d = o$ indicates a bearing with an outer race defect, $d = i$ indicates a bearing with an inner race defect

k :

Superscript indicating the number of the set of initial centers used in each test case. A total number of $k = 5$ initial center sets is used in each industrial case (cases I, II, III). A total number of $k = 3$ initial center sets is used in the experimental test case IV

r :

Superscript indicating the test case considered: $r = I, II, III, IV$

Table 18

Results from the application of the set A of time domain based K -means clustering for the 1st stage classification (normal or defective bearing). The features are calculated from the raw and the envelope signals.

Case r	Bearing type	Defect type	Results for different distances			
			Correlation	Cosine	Euclidean	Cityblock
I	2312	Outer race defect (5 signals)	✓	✓	X	X
II	22205	Inner race defect (5 signals)	✓	✓	X	X
III	22228	Inner race defect (4 signals) Normal bearing (1 signal)	X	X	X	X
IV	6203-2RS JEM	Outer race defect (3 signals) Inner race defect (3 signals) Normal bearing (1 signal)	X	X	X	X

Table 19

Results from the application of the set A of time domain based K -means clustering for the 2nd stage classification (outer or inner race defect). The features are calculated from the raw and the envelope signals.

Case r	Bearing type	Defect type	Results for different distances			
			Correlation	Cosine	Euclidean	Cityblock
I	2312	Outer race defect (5 signals)	X	X	X	X
II	22205	Inner race defect (5 signals)	X	X	X	X
III	22228	Inner race defect (4 signals) Normal bearing (1 signal)	X	X	X	X
IV	6203-2RS JEM	Outer race defect (3 signals) Inner race defect (3 signals) Normal bearing (1 signal)	X	X	X	X

Thus, for each test case r ($r = \text{I, II, III and IV}$), systematic applications of the dynamic model of Eq. (13) are performed, in order to generate simulated signals S_d^{rk} with different bearing defects and with different fault severity levels. Different centroid sets C_1^{rk}, C_2^{rk} are then derived after an appropriate preprocessing and feature extraction procedure as in Section 4.1, as derived from the spectrum of the raw signal S_d^{rk} and its envelope E_d^{rk} according to the feature sets of Tables 3 and 4, respectively.

Furthermore, the performance of the K -means clustering approach is evaluated in all four test cases I, II, III and IV using four alternative metric distances: (a) correlation, (b) cosine, (c) Euclidean and (d) Cityblock.

The values of the shaft speeds and of the characteristic defect frequencies, which are used in the implementation of the dynamic bearing model for each test case r , are presented in Table 5. These values are taken directly from the corresponding data of the machine under operation. In the industrial test cases $r = \text{I, II and III}$, each simulation is 8192 samples long and the sampling rate f_s is equal to 16,384 Hz. Likewise, in the experimental test case $r = \text{IV}$, each simulated signal is 8192 samples long and the sampling rate f_s is equal to 12,000 Hz.

The rest of the values of the parameters used for the generation of the signals S_d^{rk} in the $k = 0, 1, 2, 3, 4$ initial center data sets for the industrial test cases $r = \text{I, II, III}$, are presented in Table 6. The rest of the values of the parameters used for the generation of the signals S_d^{rk} in the $k = 0, 1, 2$ initial center data sets for the experimental test case $r = \text{IV}$ are presented in Table 7.

The spectra of the simulated signals S_o^0, S_i^0 and S_n^0 ($r = \text{I, II, III and IV}$) and their corresponding envelopes E_o^0, E_i^0 and E_n^0 are presented in Figs. 17–24. For each test case r , the first simulated signal is a white noise signal that presents the normal bearing condition, the second one presents a bearing with an outer race defect, and the third one presents a bearing with an inner race defect.

Comparing the spectra of the simulated signals in Figs. 17–24 to the corresponding spectra of the measured signals in Figs. 6–16, a similarity in their structure can be observed. This similarity is

further exploited by the selection of the appropriate frequency domain based feature set in Tables 3 and 4, and by the implementation of an appropriate similarity measure in the form of the correlation distance, as further described in Section 5.3.

5.3. Classification performance analysis

The results of the application of the proposed K -means clustering classification are presented in Tables 8–15. As it can be observed, the success rate of the method using the correlation distance is 100% (\checkmark) in both stages of classification, for all data sets and for all the measurements of all four test cases. The reason for the successful usage of the correlation distance is (A) it ignores the differences in magnitude that result for different bearing fault severity levels and (B) it retains its sensitivity to the direction of the change among those data vectors that correspond to the calculated specific spectrum energies, which characterize the type of the bearing defect.

On the contrary, the Euclidean and the Cityblock metric distances fail (X) in both stages of all the measurements in all the cases, except in the 2nd stage of case II. This happens because these metric distances are sensitive to both magnitude and direction of change.

The implementation of the cosine metric distance in most of the cases does not result to successful data classification. Different initial conditions result in different final clusters. The success rate of 100%, achieved by the cosine distance in certain cases, is due to the fact that the cosine formula is defined as one minus the uncentered correlation distance, taking into advantage some of the correlation distance benefits.

Figs. 25–28 present three-dimensional pictures of the first stage classification results. The final centers are shown with the marker points '⊗'. The outputs characterizing a defective classification are presented with the symbol 'o'. The marker '□' represents the measured and simulated normal condition output, which is used as initial center and input to the algorithm.

Table 20

Results from the application of the set B of time domain based K -means clustering for the 1st stage classification (normal or defective bearing). The features are calculated from the raw signals.

Case r	Bearing type	Defect type	Results for different distances			
			Correlation	Cosine	Euclidean	Cityblock
I	2312	Outer race defect (5 signals)	X	X	X	X
II	22205	Inner race defect (5 signals)	\checkmark	X	X	X
III	22228	Inner race defect (4 signals) Normal bearing (1 signal)	\checkmark	X	X	X
IV	6203-2RS JEM	Outer race defect (3 signals) Inner race defect (3 signals) Normal bearing (1 signal)	X	X	X	X

Table 21

Results from the application of the set B of time domain based K -means clustering for the 1st stage classification (normal or defective bearing). The features are calculated from the envelope signals.

Case r	Bearing type	Defect type	Results for different distances			
			Correlation	Cosine	Euclidean	Cityblock
I	2312	Outer race defect (5 signals)	X	X	X	X
II	22205	Inner race defect (5 signals)	\checkmark	\checkmark	X	X
III	22228	Inner race defect (4 signals) Normal bearing (1 signal)	\checkmark	\checkmark	X	X
IV	6203-2RS JEM	Outer race defect (3 signals) Inner race defect (3 signals) Normal bearing (1 signal)	X	X	X	X

Table 22

Results from the application of the set B of time domain based *K*-means clustering for the 2nd stage classification (outer or inner race defect). The features are calculated from the raw and the envelope signals.

Case <i>r</i>	Bearing type	Defect type	Results for different distances			
			Correlation	Cosine	Euclidean	Cityblock
I	2312	Outer race defect (5 signals)	X	X	X	X
II	22205	Inner race defect (5 signals)	X	X	X	X
III	22228	Inner race defect (4 signals) Normal bearing (1 signal)	✓	✓	X	X
IV	6203-2RS JEM	Outer race defect (3 signals) Inner race defect (3 signals) Normal bearing (1 signal)	X	X	X	X

5.4. Comparison to a time domain based features *K*-means clustering approach

In order to evaluate the effectiveness of the proposed frequency-domain feature set, two alternative feature sets are considered in a similar two stage *K*-means clustering approach. These feature sets are now based exclusively to statistical time-domain indices, since these indices have been widely used in the literature for bearing fault detection and classification.

Two alternative feature sets are considered (sets A and B), consisting of statistical time-domain features, as presented in Tables 16 and 17. These time-domain features are calculated for both the raw and the demodulated signal. The results are presented in Tables 18–22.

As it is observed in Tables 18–22, the method results to correct classifications only in a few circumstances using the correlation and cosine metric distances. However, in general the *K*-means clustering method using these two alternative feature sets leads to almost completely erroneous results in both first and the second stage. The statistical time-domain feature parameters confuse the classifier and minimize the accuracy of the classification.

6. Conclusion

The proposed *K*-means clustering approach results to a success rate of 100% at all measurements of all three industrial cases and of the laboratory test case considered. Moreover, the method indicates a quite robust behavior with respect to the choice of the initial centers.

The main reason behind these successful results is that the method takes advantage of already existing engineering knowledge and expertise, concerning the physical behavior of defective rolling element bearings. Based on this knowledge, an effective frequency domain based feature set is proposed, enabling the correlation distance used in *K*-means clustering to indicate the strength and the direction of the linear relationship between two measures. This feature set presents a far superior behavior compared to other traditional feature sets, which are based on statistical time-domain indices. Moreover, the initial centers of the clusters can be calculated using a well established dynamic model, describing the physical behavior of defective rolling element bearings. As proven by the results, the method is quite robust to the choice of these initial centers, eliminating thus one traditional disadvantage of *K*-means clustering.

Taking into account some further advantages of *K*-means clustering, such as its ease of programming and the accomplishment of a good trade-off between achieved performance and computational complexity, an effective automated unsupervised learning procedure results, which can be directly implemented to measured vibration data. The need for training the method with measured data from the specific machine under defective bearing conditions

is eliminated, a fact which consists the major advantage for the implementation of the method, especially in industrial environments.

Acknowledgments

This paper is part of the 03ED_78 research project, implemented within the framework of the “Reinforcement Programme of Human Research Manpower” (PENED) and co-financed by National and Community Funds (25% from the Greek Ministry of Development–General Secretariat of Research and Technology and 75% from E.U. European Social Fund).

References

- Antoni, J., & Randall, R. B. (2002). Differential diagnosis of gear and bearing faults. *Journal of Vibration and Acoustics – Transactions of the ASME*, 124, 165–171.
- Bradley, P. S., & Fayyad, U. M. (1998). Refining initial points for *K*-means clustering. In *Proceedings of the fifteenth international conference on machine learning* (pp. 91–99). San Francisco, CA: Morgan Kaufmann.
- Elaine, Y. C., Wai, W. X., Michael, K. N., & Joshua, Z. H. (2004). An optimization algorithm for clustering using weighted dissimilarity measures. *Pattern Recognition*, 37, 943–952.
- Jack, L. B., & Nandi, A. K. (2002). Fault detection using support vector machines and artificial neural networks, augmented by genetic algorithms. *Mechanical Systems and Signal Processing*, 16, 373–390.
- Junyan, Y., Youyun, Z., & Yongsheng, Z. (2007). Intelligent fault diagnosis of rolling element bearing based on SVMs and fractal dimension. *Mechanical Systems and Signal Processing*, 21, 2012–2024.
- Komgom, N. C., Mureithi, N., Aouni, L., & Marc, T. (2007). On the use of time synchronous averaging, independent component analysis and support vector machines for bearing fault diagnosis. In *First international conference on industrial risk engineering, 17–19 December, 2007, Montreal* (pp. 610–624).
- Lee, H., Nguyen, N., & Kwon, J. (2007). Bearing diagnosis using time-domain features and decision tree. In *Third international conference on intelligent computing, ICIC 2007, August 21–24, 2007, Qingdao, China* (pp. 952–960).
- Likas, A., Vlassis, N., & Verbeek, J. J. (2003). The global *K*-means clustering algorithm. *Pattern Recognition*, 36, 451–461.
- Logan, D., & Mathew, J. (1996a). Using the correlation dimension for vibration fault diagnosis of rolling element bearings – I. Basic concepts. *Mechanical Systems and Signal Processing*, 10, 241–250.
- Logan, D., & Mathew, J. (1996b). Using the correlation dimension for vibration fault diagnosis of rolling element bearings – I. Selection of experimental parameters. *Mechanical Systems and Signal Processing*, 10, 251–264.
- Loparo, K. A. *Bearings vibration data set*. Case Western Reserve University. <<http://www.eecs.case.edu/laboratory/bearing/download.htm>>.
- MacFadden, P. D., & Smith, J. D. (1984a). Model for the vibration produced by a single point defect in a rolling element bearing. *Journal of Sound and Vibration*, 96(1), 69–82.
- MacQueen, J. B. (1967). Some methods for classification and analysis of multivariate observations. *Fifth Berkeley symposium on mathematical statistics and probability* (Vol. 1, pp. 281). Berkeley: University of California Press.
- McCormick, A. C., & Nandi, A. K. (1997). Classification of the rotating machine condition using artificial neural networks. *Proceedings of Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 211, 439–450.
- McFadden, P. D., & Smith, J. D. (1984b). Vibration monitoring of rolling element bearings by the high frequency resonance technique – A Review. *Tribology International*, 11(7/1), 3–10.
- Milligan, G. W. (1980). An examination of the effect of six types of error perturbation on fifteen clustering algorithms. *Psychometrika*, 45, 325–342.

- Patargias, T. I., Yiakopoulos, C. T., & Antoniadis, I. A. (2006). Performance assessment of a morphological index in fault prediction and trending of defective rolling element bearing. *Nondestructive Testing and Evaluation*, 22(1), 39–60.
- Qiao, H., Zhengjia, H., Zhousuo, Z., & Yanyang, Z. (2007). Fault diagnosis of rotating machinery based on improved wavelet package transform and SVMs ensemble. *Mechanical Systems and Signal Processing*, 21, 688–705.
- Randall, R. B. (1987). *Frequency analysis*. Copenhagen: Bruel & Kjaer.
- Samanta, B. (2004). Gear fault detection using artificial neural networks and support vector machines with genetic algorithms. *Mechanical Systems and Signal Processing*, 18, 625–644.
- Saravanan, N., Cholaairajan, S., & Ramachandran, K. I. (2009). Vibration-based fault diagnosis of spur bevel gear box using fuzzy technique. *Expert Systems with Applications*, 36, 3119–3135.
- Wang, X., Wang, Y., & Wang, L. (2004). Improving fuzzy c-means clustering based on feature-weight learning. *Pattern Recognition Letters*, 25, 1123–1132.
- Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of American Statistical Association*, 58, 236–244.
- Widodo, A., Yang, B., & Han, T. (2007). Combination of independent component analysis and support vector machines for intelligent faults diagnosis of induction motor. *Expert Systems with Applications*, 32, 299–312.
- Yaquo, L., Zhengjia, H., Yanyang, Z., & Xuefeng, C. (2008). New clustering algorithm-based fault diagnosis using compensation distance evaluation technique. *Mechanical Systems and Signal Processing*, 22, 419–435.
- Zhang, L., Jack, L. B., & Nandi, A. K. (2005). Fault detection using genetic programming. *Mechanical Systems and Signal Processing*, 19, 271–289.