The International Journal of Advanced Manufacturing Technology (2019) 103:2377–2390
https://doi.org/10.1007/s00170-019-03597-2

ORIGINAL ARTICLE

Robust cepstral-based features for anomaly detection in ball bearings

Ricardo Sousa[1] · Joel Antunes[2] · Filipe Coutinho[3] · Emanuel Silva[2] · Joaquim Santos[2] · Hugo Ferreira[4]

Received: 9 October 2018 / Accepted: 12 March 2019 / Published online: 22 April 2019

Abstract

This paper proposes the linear frequency cepstral coefficients as highly discriminative features for anomaly detection in ball bearings using vibration sensor data. These features are based on cepstral analysis and are capable of encoding the patterns of a spectral magnitude profile. Incipient damages on bearings can grow rapidly under normal use resulting in vibration and harsh noise. If left undetected, this damage will worsen, leading to high maintenance costs or even injury. Multiple interferences in an industrial environment contaminate the signal, making it a challenge to correctly identify the bearings' condition. Many studies have attempted to overcome this issue at the signal level. However, the discriminative capacity of the current vibration signal features is still vulnerable to interference, which motivates this work. In order to demonstrate the benefits of these features, we (1) show that they are computationally efficient and suitable for real-time incremental training; (2) conduct discriminative analysis by evaluating the separability performance and comparing it with the state of the art; and (3) test the robustness of the proposed features under noise interference, which is ideal for use in the harsh operating conditions of industrial machinery. The data was obtained from a laboratory workbench setting that reproduces bearing fault scenarios. Results show that the proposed features are fast, competitive when compared to state-of-the-art features, and resilient to high levels of interference. Despite the higher performance when using the quadratic model, the proposed features remain highly discriminative when used with several other discriminant function.

Keywords Anomaly detection · Spectral profile features · Noise resilience

1 Introduction

Rolling elements (such as ball bearings) have a major impact on a rotating machinery's operation condition. Faults due to wear and tear, high temperatures, misalignment, lack of adequate lubrication, or fine particle contamination are a constant in industrial environments. These rolling elements can be found in important components such as motors, compressors, gearboxes, and pumps. Thus, detecting bearings faults in its early stages is essential in reducing maintenance costs and avoiding injury [22].

Predictive maintenance (PdM) is a set of methodologies to detect, predict, and diagnose machine component faults as early and as accurately as possible [8]. Knowledge-based PdM is a non-invasive methodology that resorts to machine learning and data mining methods to build their procedures [19]. Anomaly detection is one of the PdM procedures whose goal is to identify precursor events in the sensor signals that anticipate failure.

More specifically, anomaly detection only requires negative (non-failure) labelled data to generate detection models. This has two significant advantages. Firstly, feature extraction does not depend on the rare positive (failure) labelled data. Second, feature extraction does not require collecting and labelling all the possible failure modes of a component—a formidable if not impossible task.

Vibration signals (displacement signals) obtained from accelerometers attached to bearings' housing structures are considered one of the most reliable means of monitoring the bearings' conditions. These signals contain meaningful information that reflects the bearing's health [5]. For anomaly detection, discriminative features of signals are needed to differentiate the various conditions of the bearings in various machines working modes. However, the signal is distorted by noise interference from the industrial environment, which leads to detection errors (false positives and negatives). It is therefore important to design these

✉ Ricardo Sousa
rtsousa@inesctec.pt

Extended author information available on the last page of the article.

features to be resilient to this interference. This is crucial since it compromises the next stage of anomaly detection—generating the classification model of the health state. The more discriminative and resilient to interferences these features are, the less complex the classification stage will be. Feature extraction is the generation of measurements that describe the signal patterns used to characterize the condition of the bearing element. Some features consist of measures with physical meaning associated to bearings physical model [21]. They are discriminative if they can be used to unequivocally encode particular waveform patterns that can be successfully used to differentiate between the various conditions of this physical process [2].

Data streams are the main abstraction used for collecting monitoring data through IIoT (Industrial Internet of Things) sensors networks in industrial environments [1]. Vast amounts of data are continuously produced in real-time and at many sampling rates, which gives rise to demanding processing time requirements [23]. Fast feature extraction methods are therefore required in this context.

This work proposes the use of the noise-robust and discriminative linear frequency cepstral coefficients (LFCC) features to detect anomalies on bearings with high speed and accuracy so that it can be easily integrated into an anomaly detector. These types of features are generally used in speech processing tasks where signal pattern diversity is high. In fact, these features are closely related to the mel-frequency cepstral coefficients (MFCC), which are the state of the art in speech recognition methods and are used in very complex classification tasks.

The LFCC features describe the spectral magnitude profile. The DCT transform is also used to decorrelate these features with almost the same compaction efficiency of the Karhunen-Loève transform (optimal decorrelation transformation) [17]. We also suggest the use of median filtering over the feature space since it prevents cluster overlapping and significantly increases the feature's discriminant properties.

To verify the benefit of these features, we compared the LFCC to a baseline that consists of the traditional time domain features and to the empirical mode decomposition (EMD) features, which is widely applied in bearing vibration analysis. In this task, we (1) measure the runtime of all features computation; (2) perform discriminative analysis by evaluating the performance with simple discriminant functions; and (3) test the robustness of the features under noise interference (with several noise levels).

The remainder of this paper is organized as follows. Section 2 briefly reviews the traditional time domain and the EMD features that are often used in anomaly detection in bearings elements. Section 3 describes the LFCC and discuss its main advantages. Section 4 explains the evaluation experiments. The results are presented and discussed in Section 5 and the main conclusions are reported in Section 6.

2 Related work

A short literature review is provided in this section referencing the most prominent work on feature generation for bearing vibration signal analysis. In this description, we favour the features that can be computed over small segments (windows or frames) of the signal (short-term features), which are compatible with data streaming anomaly detectors.

We selected just two features types: the traditional time domain statistical features and the state-of-the-art EMD features. The time domain features are the baseline and the EMD features are used as the golden standard to show the competitive benefits of LFCC.

2.1 Time domain features

Time domain features have been used in bearing vibration analysis as a first approach to anomaly detection. In general, these features encode the patterns of variations in signals' amplitude and are based on statistical measures [3, 9, 15]. The most common features are root mean square (RMS), kurtosis (K), crest factor (CrF), clearance factor (ClF), impulse factor (If), and shape factor (SF), which are described in the following expressions:

$$\mathrm{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_n - \bar{x})^2} \tag{1}$$

$$K = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_n - \bar{x})^4}{RMS^4} \tag{2}$$

$$\mathrm{CrF} = \frac{1/2[\max(x_n) - \min(x_n)]}{RMS} \tag{3}$$

$$\mathrm{ClF} = \frac{1/2[\max(x_n) - \min(x_n)]}{\frac{1}{N}(\sum_{i=1}^{N}|x_n|)^2} \tag{4}$$

$$\mathrm{If} = \frac{1/2[\max(x_n) - \min(x_n)]}{\frac{1}{N}\sum_{n=1}^{N}|x_n|} \tag{5}$$

$$\mathrm{SF} = \frac{RMS}{\frac{1}{N}\sum_{n=1}^{N}|x_n|} \tag{6}$$

where n is the frame sample index, x_n is the element of the frame, $\bar{x}$ is the mean of the frame values, and N is the size of the frame. These features capture the spiky and impulsive nature of damaged bearing vibration [11]. As advantages, these features are very easy to implement and are highly

computational efficient. They are also very interpretable since they are directly related to the impulsive events on the signal. As drawbacks, these features present low performance in noisy environments since they are strongly dependent on signals' amplitude and cannot distinguish between noise and vibration impulses. Although no work on short-term measures was found using these features, they can be easily applied in this context.

2.2 Empirical mode decomposition features

EMD is an iterative method that decomposes the signal into its elementary components referred to as the intrinsic mode functions (IMFs) [12, 16, 21, 25]. IMFs form an orthogonal basis of the signal. Hence, these components represent decorrelated information with the potential to discriminate bearing conditions. This method acts like a filter bank where the weak impulses caused by bearings faults are enhanced. Figure 1 displays the spectrum magnitude of the EMD components with a frequency axis in logarithmic scale.

The filter bank is nearly dyadic and its filters are focused on the low frequencies of the spectrum. This characteristic gives more resolution and enhancement to low-frequency information. The pseudo-code of the EMD computation is presented in Algorithm 1.

Algorithm 1 EMD components computation.

Initialization: $x_n -$ *input signal*; $i = 0$
while x_n *not monotonic* **do**
  $1-$ *Compute the maxima and the minima*
  $2-$ *Interpolate a function with all* $maxima - M_n$
  $3-$ *Interpolate a function with all* $minima - m_n$
  $4-$ *Compute* $\mu_n = (M_n + m_n)/2$
  $5 - c_n = x_n - \mu_n$
  if [$c[n]$ *is IMF function*] **then**
    $x_n = c_n$
    $IMF_n^i = c_n$
    $i = i + 1$
  else
    $x_n = c_n$

Basically, the method involves the computation of the frame's upper and lower envelopes through interpolation of the maxima and minima, respectively. Then, the envelope's average is subtracted from the frame. This is an iterative process that repeats itself until the remaining frame's signal is monotonic.

The main advantage is the fact that these components form a nearly orthogonal base that can be used to produce the discriminant features. EMD does not require any domain transformation and can be easily applied to non-stationary signals. The main inconvenience is the long processing time, in particular, the interpolation stages [13]. Since it is an adaptive decomposition, the filter bank is not the same for all frames, even in the number of filters. Therefore, the same IMF_n^i does not reflect exactly the same information in different frames.

It is worth mentioning that the more recent variants of EMD, such as ensemble EMD (EEMD) and complementary EEMD (CEEMD), retain these same advantages and disadvantages [13]. In our work, we used the original EMD for comparison since it is the simpler one.

The most well-known features based on EMD analysis are the normalized IMF component energy EMD(Eng), entropy EMD(Ent) and the singular values EMD(SV) obtained from the singular value decomposition (SVD) [22]. The EMD(Eng) features are computed as the ratio between the IMF_n^i energy E_i and the sum of all component's energy E [25]. The EMD(Ent) are computed simply by applying Eq. 7 to the normalized energies [22].

$$\text{EMD(Ent)}_i = -\frac{E_i}{E}\log_2\left(\frac{E_i}{E}\right) \tag{7}$$

The EMD(SV) features are obtained by performing a tall matrix SVD where each column contains the IMF components. The singular values have the properties of stability and rotation and scale invariance, which is beneficial for encoding patterns. However, the computational cost is high.

3 Linear frequency cepstral coefficients

3.1 Linear frequency cepstral coefficients features

In this section, we describe the LFCC features' computation. The rationale behind the use of LFCC is the fact that the acoustic speech signal is of a vibrational nature. In fact, vibrational and acoustic signals can be described by both the spectral and cepstral analysis since they are the reflection of vibrational physical models. LFCC are also spectral profile descriptors based on cepstral analysis that encode the possible magnitude patterns. They are less affected by additive and noisy interference.

These features are very similar to the MFCC features which are considered the state of the art in speech recognition and regarded as highly discriminative. The difference between LFCC and MFCC is that the latter models the human auditory system, which cannot be applied in this context [14]. To the best of our knowledge, this is the first time that LFCC analysis is employed in bearing vibration analysis.

Figure 2 illustrates the LFCC computation. LFCC are short-term features that are computed from a signal frame where the vibration signal is considered to be stationary

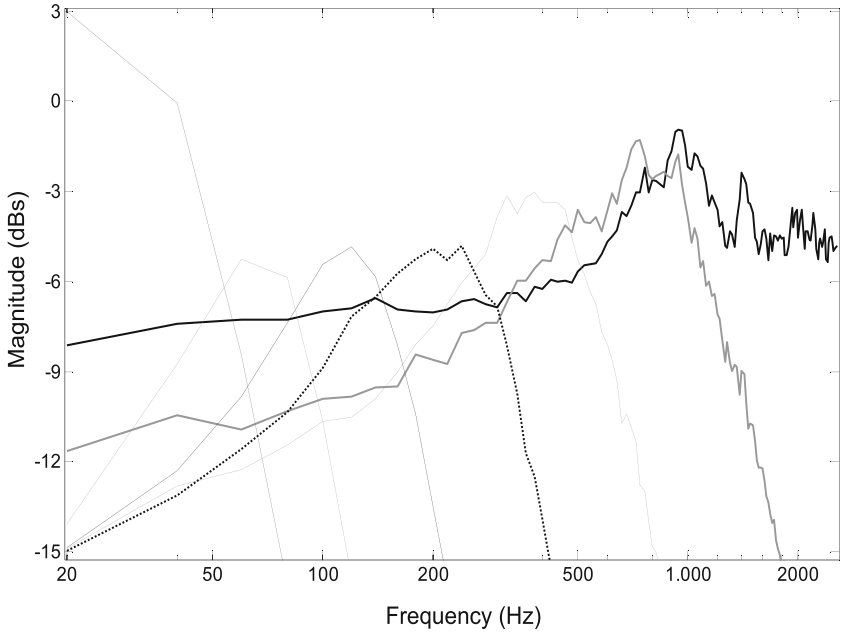

Fig. 1 Spectrum magnitude of the IMF components from a 256 sample frame of bearing vibration signal. The frequency axis uses a logarithmic scale

(e.g., 50 ms). For each signal frame, the Hanning window is applied in order to reduce the "spectral leakage" effect (introduction of new frequency components artefacts) to obtain $x_w[n]$ [7]. Then, the DFT's magnitude is obtained by first using the fast Fourier transform (FFT) and then calculating the power spectrum $S[k]$. The DFT presents already decorrelated components.

Here, the noise is spread throughout the components and its effect is reduced in each one. Next, a triangular-shaped filter bank (weighting masks) is applied to reduce the resolution of the power spectrum and compress the information in a few coefficients. The filter masks are equally spaced and have an overlap of 50% in order to give the same importance to all spectral regions. This step produces separated bands $B[k, i]$.

The energy $E[i]$ of each band i is computed and the logarithmic function $\log_{10}(.)$ is then applied to enhance and capture subtle but discriminative information. Each filter therefore summarizes the information of a spectral region into a single value of energy $E_{\log}[i]$. Finally, the discrete cosine transform (DCT), which is also derived from the DFT, is applied to decorrelate the $E_{\log}[i]$ values and produce the cepstral coefficients $C[i]$. The DCT also has the approximate compaction efficiency of the Karhunen-Loève transform for vibration signals [27].

3.2 Median filter application

Preliminary experiments have shown that the LFCC present high and spiky like variability when the signal's frames are relatively small ($\sim$ 64 to 256 samples). To get more stable values, we used another level of framing applied to the LFCC when the coefficients vary a long time. In each frame, a median was applied (median filters) to remove the effect of outliers on the LFCC coefficient values. This procedure exploits the related consecutive feature examples.

Figure 3 illustrates the effect of removing the outlier examples in two LFCC coefficients.

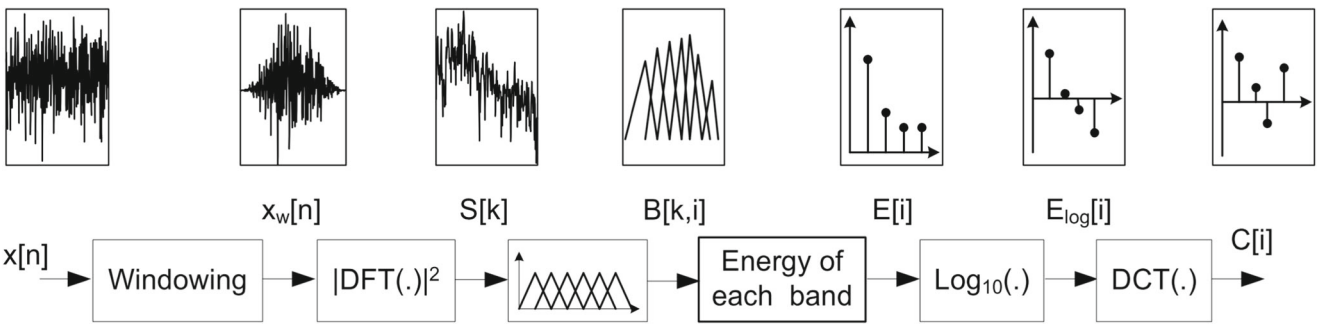

Fig. 2 Diagram for the computation of LFCC features

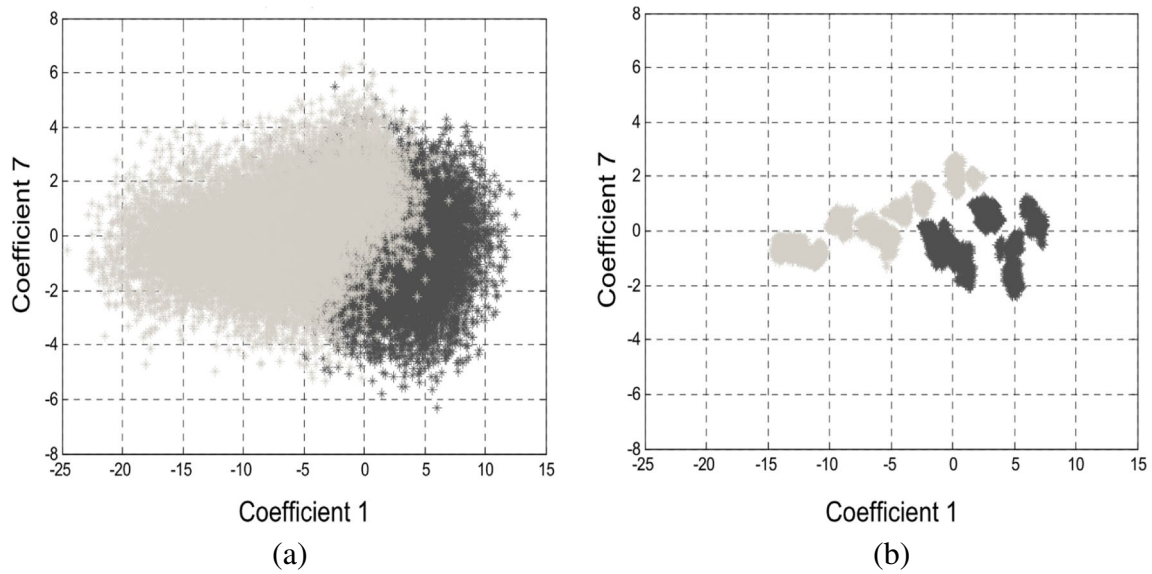

Fig. 3 Effect of the median filter considering the coefficient 1 and 7 of the LFCC with 13 coefficients **a** without median filter and **b** with median filter. Normal and damaged labelled examples are represented in dark and light gray, respectively

These clusters show that the median both reduces the overlap between clusters and increase the distance between the normal and damaged labelled clusters. Moreover, Fig. 3b also suggests that the various operating regimes within the normal and abnormal classes can also be separated.

3.3 Run-time complexity comparison

Table 1 shows the run-time complexity of each method. For the LFCC, the FFT is the operation that most significantly contributes to its run-time complexity, and is of the order of $O(n.\log(n))$, where n is the frame size [20]. The EMD method has a run-time complexity of the order of $O(n.\log(n))$ [26].

When considering the SV, the complexity increases by exactly $O(kn^2)$, where k is the number of features [10]. Considering that k is small compared to n, the complexity for the SVD operation is $O(n^2)$.

Table 1 Run-time complexity of the aforementioned features

Method	Time complexity
LFCC	$O(n.\log(n))$
EMD(SV)	$O(n.\log(n) + n^2)$
EMD(Ent)	$O(n.\log(n))$
EMD(Eng)	$O(n.\log(n))$
Time	$O(n)$

4 Evaluation

This section presents the evaluation of the LFCC in terms of run-time, discriminative capacity, and noise resilience. Since it is currently not possible to obtain vibration data labelled with bearing faults from the application site, a laboratory workbench was used to reproduce several fault scenarios. Semi-synthetic signals were also produced due to the difficulty of creating realistic noise interference on the workbench. The fault scenarios mimic the operation of milling machines that are installed at the application site.

The scientific questions will be contextualized in a supervised anomaly detection framework because of the possibility of labelling all the data. First, we describe the experimental setup that was used to collect the vibration signals and the respective settings. Next, we explain how the semi-synthetic signals were generated. Finally, the discriminative performance evaluation is described.

4.1 Experimental setup

The laboratory workbench is comprised essentially of two bearing housings, a motor flange, several 6305 ball bearings, an AISI 304 stainless steel shaft, an elastomer flexible coupling, a AM8121-0F00 servo-motor controlled by EL7211 driver, and PLC programmed via TwinCAT3 software by Beckhoff (Fig. 4a).

Figure 4b presents a 4353 B triaxial accelerometer from Brüel&Kjaer. This accelerometer was attached to the top of the housing that supported the bearing subjected to analysis. The accelerometer is connected to a C

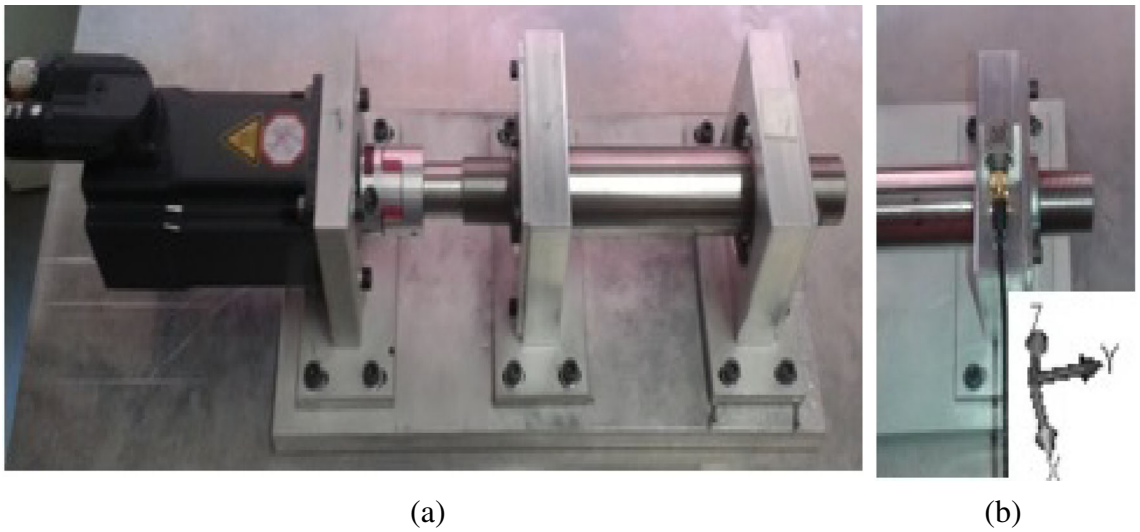

Fig. 4 Laboratorial workbenchs. **a** Experimental setup and **b** the triaxial accelerometer attached to the housing of the bearing subjected to analysis

Series NI 9234 (sound and vibration signal capturing module) from National Instruments, itself connected to a DAQ-9171 chassis, which in turn is connects to a PC. The LabVIEW programming environment from National Instruments provides the accelerometer's signal conditioning and data acquisition.

As an important step, the analysis required the vibration frequencies information, which were provided by the bearings' physical and operating (e.g., number of spheres and speed) specifications.

The sampling frequency was determined considering the necessary information for this work and the overloading limit of the hardware and software. Therefore, knowing the sampling rates of the NI 9234 module, we use

$$f_{\text{smp}} = 2.n.f_{\text{spin}} \tag{8}$$

to compute the sampling frequency, where f_{smp} is the sampling frequency (Hz), f_{spin} is the spindle speed frequency (Hz), and n is the number of teeth. The multiplication factor 2 reflects the Nyquist-Shannon theorem of sampling, which states that the minimum sampling frequency should be twice that of the maximum frequency of the collected information.

In this work, maximum f_{spin} is 100 Hz (6000 rpm) and the maximum n in the available mills is 24; the minimum sampling frequency is therefore 4800 Hz. Since the value of 4800 Hz is not one of the recommended sampling frequencies of the NI 9234 module, the closest higher value of 5120 Hz was used.

The bearing component's health is the central factor in this analysis. Thus, we considered a normal undamaged condition and three types of bearing damage on the various components: outer race, inner race and on the sphere. These fault types were selected because they are the most likely parts of the bearing to become damaged in the application site [6, 24]. Bearing faults were introduced by puncturing the relevant components to produce visible fissures. Figure 5 shows the three types of damage used in this work. One bearing element was used for each damage.

To create signal diversity, we selected 1000, 2000, and 3000 rpm motor speeds according to machine operating modes in the application site. We collected 4 s of signal segments, which are sufficient to provide stability in the sensor readings. Each segment is composed of 3 acceleration values (respectively the x, y, and z accelerometer axes) with 20,480 samples each. Note that the information conveyed by the three axes is orthogonal, which can also contribute to the discriminative capacity of the anomaly detection model.

In the data collection process, no load was attached to the rotation system. During signal collection, we first captured 5 segments (one for each test scenario) in different moments with the bearing in normal condition, while it was in motion. Afterwards, the motor was stopped and the bearing was damaged. Finally, the motor is again set in motion and 5 more segments were captured. The signal segments were stored onto files that were labelled as normal (negatives) and damaged (positives). The data is available online as separate files at the URL.[1]

4.2 Semi-synthetic signals and features datasets

For the semi-synthetic signals' production, we chose to add white noise since all signal frequencies can be equally

[1] https://gitlab.com/cese/adw/tree/master/data/inegi/ensaios_rolamentos_3

Fig. 5 Bearing damage types: **a** outer race, **b** inner race, and **c** on the sphere damages

affected. However, even though this is in general true, in a real scenario, the low frequencies are more often affected due to the existence of other rotational vibration sources. We added white noise to the signal segments (all three axes) and generated three additional data sets with the signal-to-noise ratios (SNRs) of 0, −6, and −12 dBs respectively. As a reference, we also considered the signal without noise addition in our experiments. In the noise addition cases, we assumed that the signals' implicit noise component (e.g., noise from instrumentation) presents little impact on the measurements and is substantially smaller than the added noise.

Each signal segment was divided into small 50-ms frames (256 samples for each axis), which simulates the data stream framing. With an overlap of 50% (equivalent to 128 samples), 849 frames were produced for each axis. Each frame was labelled according to the fault types (these fault labels are stored in the files' name). Then, we combined the five signal segments' frames of normal condition followed by five signal segments frames of damaged condition, which results in a total of 8478 frames for each noise level. Finally, we extract the abovementioned features, which yields four data sets of 8478 examples each. Each example is composed of the three concatenated features vectors from x, y, and z axis values.

4.3 Run-time evaluation

The goal of the run-time analysis is to measure the feature extraction processing time taken for one frame consisting of the three axis values. This run-time evaluation, together with the discriminative analysis, is then used to make the accuracy-efficiency trade-off decision. All the methods were implemented in MATLABb and run on a Intel® Core™ i7-4720HQ 2.60-GHz PC with 8 GB RAM on windows 8.1.

4.4 Discriminative analysis

In order to test the feature's resilience to noise interference, we assessed the discriminative capacity of the LFCC against the time and EMD features. The goal is to find the combination of features and classification models that presents the best performance separating the normal and damage labelled examples for the various noise levels. These results will provide information used to implement an anomaly detection system that may eventually require more complex and highly non-linear classifiers. The following most frequently used functions were therefore selected to test the adequacy of the models for anomaly detection: linear, quadratic, and logistic discriminants.

The linear function is the simplest model that can be easily implemented but is the least flexible (most biased). The logistic function, on the other hand, is the most flexible of the models as it supports more complex separation borders that may be required to differentiate between the normal and damaged labelled examples. However, it could more easily over-fit to the data. Preference is therefore given to the simpler model (easier to implement and less risk of over-fitting). The models' training was conducted using MATLAB's Statistics Toolbox.

To evaluate the model and the features, we used Accuracy (percentage of correctly classified examples) for the global performance analysis and the False Alarm Rate (False Positive Rate), which is of particular interest when these features are incorporated into an alarm system. The best performance is high Accuracy and low False Alarm Rate, simultaneously. However, we consider Accuracy as more important and the False Alarm Rate is used as tiebreaker between equally accurate methods.

Tenfold cross-validation was applied. The normal and damaged labelled examples were equally distributed to

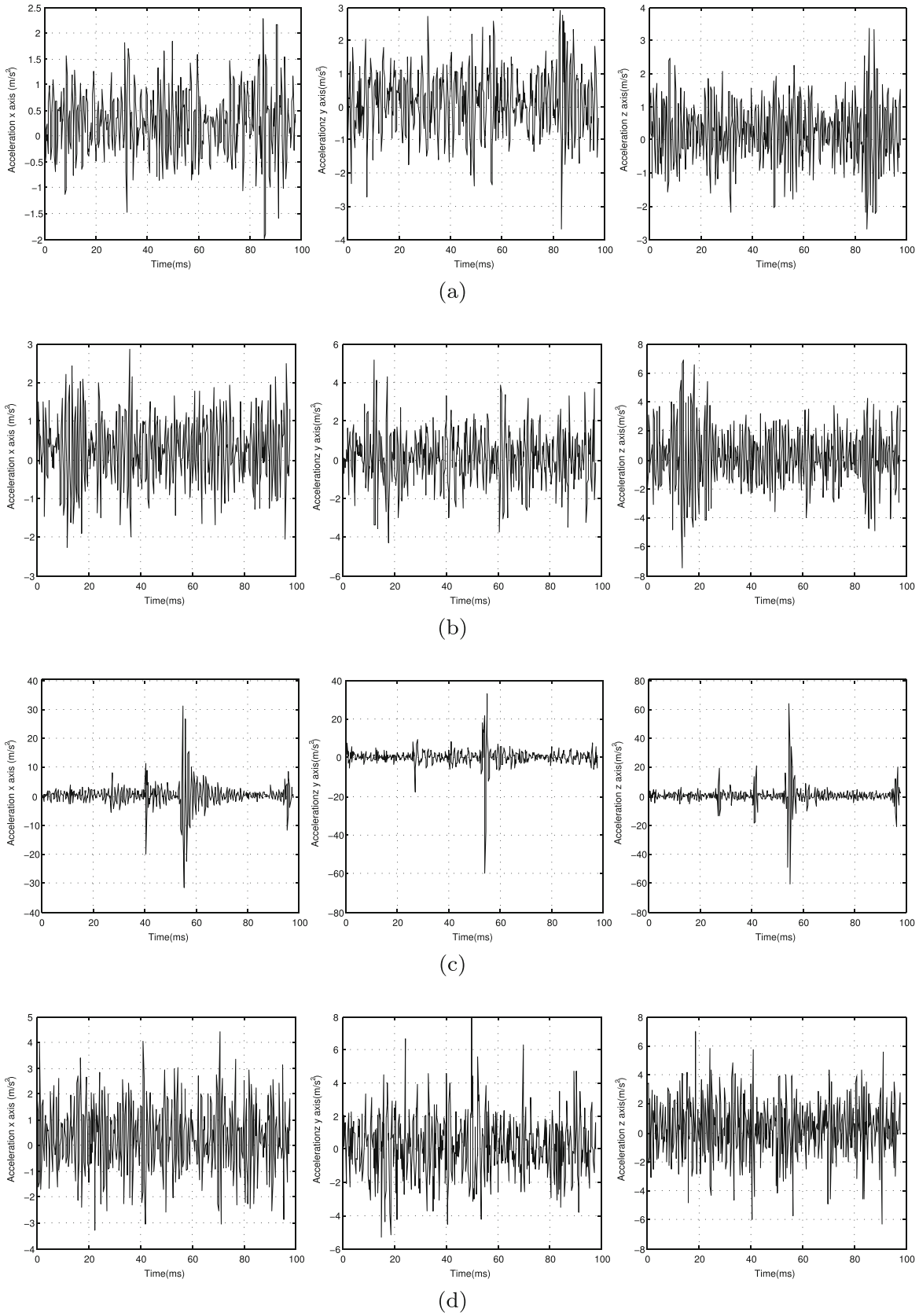

Fig. 6 Visualization of the x, y, and z axis signals of **a** without damage bearing, **b** damage in the outer race, **c** damage in the inner race, and **d**) damage in the sphere

all folders in order to create balanced data sets using a random shuffle. In addition, the cross-validation process was repeated 10 times since the examples are shuffled differently each time. At the end, the final performance was measured by computing the mean value of all runs. This process was performed so that we could have consistent and significant values.

Finally, the Friedman and Nemenyi post hoc tests were applied to the performance results in order to find the groups of methods that differ significantly and rank them accordingly [4]. Both grouping and ranking were performed with a 5% significance level.

Table 2 Mean run-time of features extraction from one three-axis frame

Method	Mean execution time (μs)
LFCC13	1,29
LFCC5	1,17
Time	0,11
EMD(SV)	218
EMD(Ent)	176
EMD(Eng)	170

5 Results

This section presents the evaluation results that answer the research questions. We present the run-time of the methods and the Accuracy and False Alarm Rate measurements for each discriminant function and noise level.

5.1 Acquired vibration signals

This subsection presents a few examples of vibration signals visualizations. Note that these signals do not contain added noise. Figure 6 depicts the x, y, and z accelerometer axis signals from a bearing without damage, with damage in the outer race, in the inner race and in one sphere. Each plot presents 100-ms segments.

The bearings with damage presented pulses and the increase of amplitude caused by the damages. In particular, the damage in the inner race showed larger pulses. In fact, these pulses are very subtle and hardly observed in a direct observation of the signal. These subtle changes reinforces the need of features that can differentiate the several bearings conditions.

5.2 Run-time analysis

Table 2 shows the mean run-time required to extract the features from one signal frame with 256 samples and three axes. The EMD produces a different number of IMF for each frame. We therefore only considered the minimum of 5 IMFs in this experiment (including the residual). This means that the total number of features for the EMD(SV), EMD(Ent), and EMD(Eng) models is 15 each, considering the three axes.

In order to make a fair comparison, we also calibrated the LFCC with five coefficients for each axis, denoted as LFCC5. To prove that these features can reach higher performance, we also calibrate the LFCC with 13 coefficients for each axis. This number of 13 coefficients was selected because it is the typical value used in speech recognition systems [18]. This LFCC calibration is denoted as LFCC13. The six time features are grouped in one vector for each axis, yielding a total of 18 features. These experiments did not involve the median filter since this operation does not contribute significantly to the run-time and can be combined with all the mentioned methods.

As expected, the extraction of time features is the fastest due to its simplicity. The EMD feature extraction are the slowest due to the spline functions. In particular, the SV presents the highest run-time due to the SVD operation. The LFCC are faster than the EMD algorithms since only the FFT operation contributes significantly to the run-time. The number of LFCC coefficients has little effect on the run-time.

5.3 Discriminative and noise robustness analysis

In this section, we present the performance measures of the linear, quadratic, and logistic discriminative models for each noise level, using the abovementioned features. In the tables below `WN` corresponds to the *without noise adding case* and the wmf subscript corresponds to method without applying the median filter. The italicized numbers mark the best values for each noise level. Tables 3 and 4 present the Accuracy and the False Alarm Rate for the linear model, respectively.

The LFCC13 presents the highest Accuracy for all noise levels and are very robust since the Accuracy is higher than 90% for all the noise levels. The time features are the least robust because they are very sensitive to the noise. These tables also reveal that the median filter substantially increases Accuracy in all cases. The False Alarm Rate decreases in most of the cases. Some cases, such as the EMD(Ent)$_{wmf}$, false positives are very high which means that the model is prone to classifying the examples as damaged condition. Effectively, EMD(Ent)$_{wmf}$ presented one of the lowest Accuracy values. For the -12 dB noise level, the time features almost reach a very poor 50% that is close to the performance of a binary random classifier.

Table 3 Accuracy (%) for a linear model

Methods	SNR (dBs)			
	WN	0	−6	−12
LFCC13	**100.0**	**99.9**	**99.4**	**92.6**
$LFCC13_{wmf}$	97.0	90.2	83.6	74.4
LFCC5	98.5	96.8	94.0	87.7
$LFCC5_{wmf}$	91.8	87.2	80.3	72.3
Time	87.0	80.3	78.1	66.8
$Time_{wmf}$	78.5	74.1	67.3	58.8
EMD(SV)	96.3	96.4	93.8	86.6
$EMD(SV)_{wmf}$	89.7	90.0	78.1	62.5
EMD(Eng)	96.9	94.3	89.1	87.0
$EMD(Eng)_{wmf}$	89.7	83.4	82.4	82.1
EMD(Ent)	87.2	84.1	78.0	76.5
$EMD(Ent)_{wmf}$	75.9	75.5	74.2	73.5

Table 5 Accuracy (%) for a quadratic model

Methods	SNR (dBs)			
	WN	0	− 6	− 12
LFCC13	**100.0**	**100.0**	**99.9**	**98.0**
$LFCC13_{wmf}$	99.2	92.1	87.4	85.3
LFCC5	100.0	99.9	99.0	89.7
$LFCC5_{wmf}$	96.2	87.3	80.0	74.0
Time	72.1	71.9	77.9	73.3
$Time_{wmf}$	67.2	64.8	59.5	56.9
EMD(SV)	98.7	99.2	97.6	89.4
$EMD(SV)_{wmf}$	87.7	85.1	81.2	81.8
EMD(Eng)	98.9	96.3	92.1	81.0
$EMD(Eng)_{wmf}$	78.9	75.4	74.2	75.0
EMD(Ent)	99.8	96.8	96.3	87.9
$EMD(Ent)_{wmf}$	87.7	89.0	76.5	60.8

Tables 5 and 6 present the Accuracy and the False Alarm Rate for the quadratic model, respectively.

Similarly, the LFCC13 features present the highest Accuracy values and are very robust for all noise levels. With a quadratic model, the LFCC13 features also obtained the lowest False Alarm Rate values. The median filter also improves the Accuracy and decreases the False Alarm Rate in all cases.

Tables 7 and 8 provide the Accuracy and the False Alarm Rate for the logistic model, respectively.

With the logistic model, the LFCC13 also presented the best results. However, the EMD(SV) features presented the best results for −12 dB noise level. As anticipated, the time features again showed to be the least discriminative. The median filter again improves Accuracy and False Alarm Rates for all methods.

Table 4 False Alarm Rate (%) for a linear model

Methods	SNR (dBs)			
	WN	0	− 6	− 12
LFCC13	0,0	**0.1**	0.9	11.2
$LFCC13_{wmf}$	3.5	11.2	17.8	24.9
LFCC5	0.8	4.2	6.7	14.5
$LFCC5_{wmf}$	9.1	12.9	18.8	25.2
Time	12.0	20.6	23.1	34.6
$Time_{wmf}$	19.2	29.8	37.9	43.8
EMD(SV)	4.2	7.3	11.6	16.4
$EMD(SV)_{wmf}$	7.0	11.5	21.3	36.0
EMD(Eng)	5.3	8.1	9.3	7.2
$EMD(Eng)_{wmf}$	7.0	7.7	3.3	0.6
EMD(Ent)	7.0	8.4	6.0	5.1
$EMD(Ent)_{wmf}$	6.6	4.3	**0.4**	**0.0**

Table 6 False Alarm Rate (%) for a quadratic model

Methods	SNR (dBs)			
	WN	0	− 6	− 12
LFCC13	**0.0**	**0.0**	**0.0**	**2.3**
$LFCC13_{wmf}$	0.8	9.3	12.7	13.5
LFCC5	0.0	**0.0**	0.8	15.8
$LFCC5_{wmf}$	4.4	15.6	19.6	20.6
Time	0.1	0.5	8.7	30.7
$Time_{wmf}$	1.5	2.6	9.7	39.2
EMD(SV)	2.6	1.7	4.4	6.7
$EMD(SV)_{wmf}$	7.7	8.5	10.1	9.8
EMD(Eng)	2.1	1.1	2.6	5.0
$EMD(Eng)_{wmf}$	6.4	10.2	9.5	8.7
EMD(Ent)	0.1	6.4	7.3	15.7
$EMD(Ent)_{wmf}$	7.7	13.8	26.4	40.4

Table 7 Accuracy (%) for a logistic model

Methods	SNR (dBs)			
	WN	0	− 6	− 12
LFCC13	**100.0**	**100.0**	**99.9**	94.5
LFCC13$_{wmf}$	97.9	90.2	83.5	74.4
LFCC5	**100.0**	99.0	96.2	87.7
LFCC5$_{wmf}$	93.0	87.2	80.3	72.3
Time	90.3	81.8	78.0	66.9
Time$_{wmf}$	79.3	74.6	67.6	58.8
EMD(SV)	99.8	99.5	99.0	**97.8**
EMD(SV)$_{wmf}$	95.7	95.1	95.7	96.8
EMD(Eng)	98.4	95.7	95.4	93.2
EMD(Eng)$_{wmf}$	89.5	89.1	90.5	91.9
EMD(Ent)	100.0	99.9	97.6	86.3
EMD(Ent)$_{wmf}$	95.7	90.1	78.2	62.5

5.4 Ranking of discriminative capacity

In this section, we present the critical diagrams of the Accuracy and False Alarm Rate considering the combinations of features and discriminant function pairs. We identify the linear discriminant, quadratic discriminant and the logistic discriminant in the diagrams as `LD`, `QD`, and `LogD` respectively. This analysis only includes the methods with median filtering since it improves the performance for all combinations of features and discriminative functions. The Friedman test revealed that the performance results for each method were different for all measures. Briefly, the diagram is composed by a ranking bar where the best methods are on the right. The black bars define a group of methods where the performance is not statistically different.

Figure 7 displays the critical diagram Accuracy considering the combinations of features and discriminant functions.

The LFCC13 with quadratic discriminant is the best combination to obtain a high detection rate. This combination is included in a distinct group which also includes the LFCC13 features combined with linear and logistic discriminants. This fact means that these features are suitable for any model. EMD(SV) with the logistic discriminant is also included in this best group. These facts reveal that EMD(SV) features are also very discriminative but only with a complex discriminant such as the logistic function. On the other hand, the time features present very low Accuracy for any discriminant.

Figure 8 depicts the critical diagram of the False Alarm Rate considering the combinations of features and discriminant functions.

The diagram reveals that the LFCC13 with quadratic discriminant produces the least number of false alarms. The best group include also other combinations of LFCC with other discriminant functions. Similarly, the EMD(SV) with the logistic discriminant function is in the best group.

Table 8 False Alarm Rate (%) for a logistic model

Methods	SNR (dBs)			
	WN	0	−6	−12
LFCC	**0.0**	**0.0**	**0.1**	5.7
LFCC$_{wmf}$	2.1	10.0	17,2	25.2
LFCC5	0.0	0.7	3.2	13.1
LFCC5$_{wmf}$	6.8	12.9	19.5	26.0
Time	9.6	18.2	22.3	34.1
Time$_{wmf}$	21.6	27.7	34.6	43.4
EMD(SV)	0.2	0.5	1.1	**2.9**
EMD(SV)$_{wmf}$	4.8	5.4	5.5	4.5
EMD(Eng)	1.0	3.2	2.7	5.0
EMD(Eng)$_{wmf}$	6.2	6.7	6.4	6.6
EMD(Ent)	0.0	0.1	2.5	14.9
EMD(Ent)$_{wmf}$	4.8	10.7	21.8	36.2

5.5 General remarks

Considering all the results, we point out the following remarks:

- The median filter improves the Accuracy of all features and improves the False Alarm Rate of the most accurate methods.
- The best feature and model combination is the LFCC13 with quadratic discriminant function for both Accuracy and False Alarm Rate. However, we considered that the EMD(SV) with the logistic discriminant performs equally well since it is also in the best group of combinations.
- The LFCC features present high rates for any discriminant function and the EMD features are better separated with a logistic function. This fact suggests that LFCC presents high separability of the normal and damage labelled examples.
- The LFCC can be used for the fast training of simple models without losing discriminative capacity.

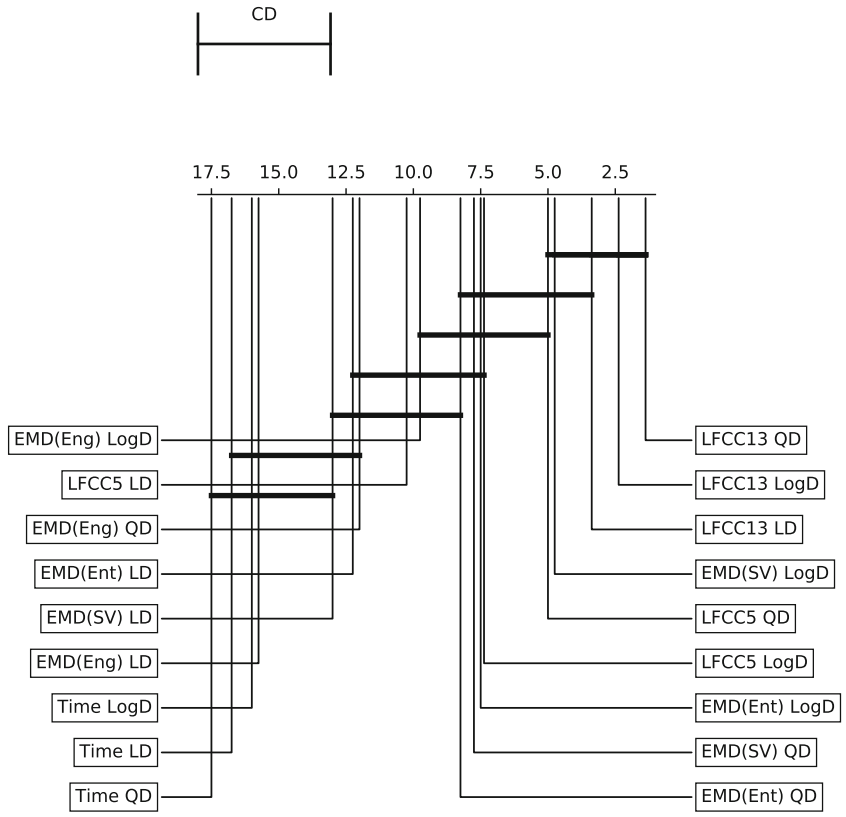

Fig. 7 Critical diagram for Accuracy. Nemenyi post hoc test at 5% significance

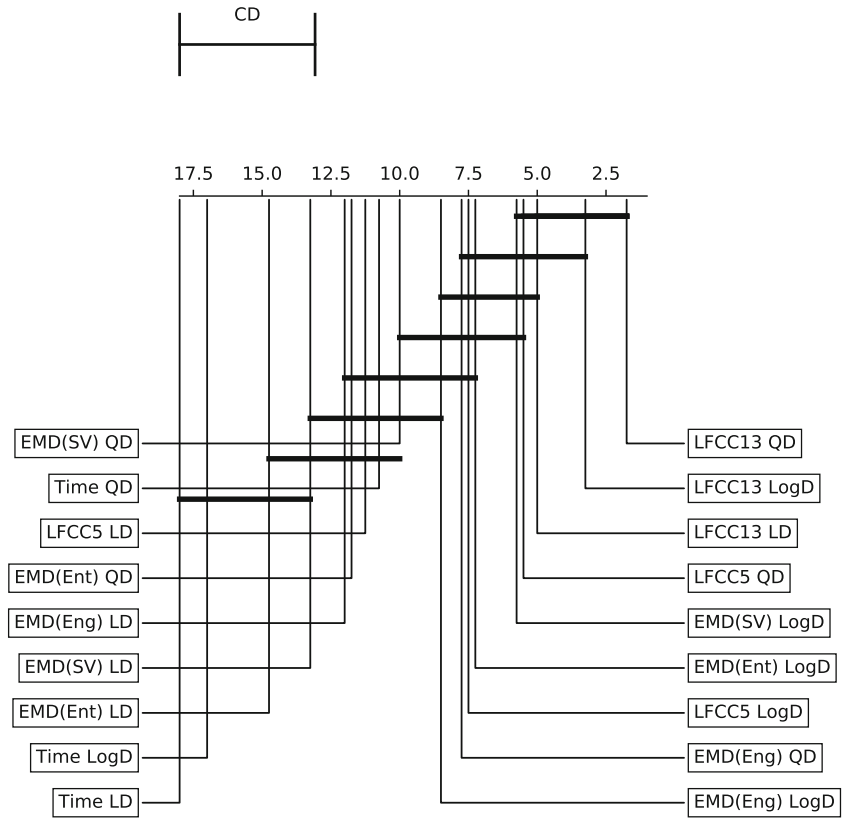

Fig. 8 Critical diagram for False Alarm Rate. Nemenyi post hoc test at 5% significance

6 Conclusion

This work proposes the use of LFCC features for anomaly detection in bearing vibration signals. The results reveal that the LFCC extraction is faster than the EMD features with similar or higher performance. It is therefore better suited for data streams and real-time systems. Moreover, LFCC presents high data separability (even for simple model functions) between the normal and damage conditions almost independently of the function model (although the quadratic function is preferable).

As future work, tuning methods can be applied to find the best parametrization for LFCC features. In addition, the LFCC features can be submitted to feature selection method to improve accuracy and efficiency. Finally, the results of this work will be used in an anomaly detector based on the quadratic model. This work was conducted in a supervised anomaly detection setting. In future works, we aim to extend these results to a semi-supervised (and if possible, unsupervised) context, where very little labelled data is available.

Acknowledgments The authors would like to thank the reviewers for the careful reading and constructive feedback on the material presented in this article.

Funding information This work is co-financed by the ERDF – European Regional Development Fund through the Operational Programme for Competitiveness and Internationalization - COMPETE 2020 under the PORTUGAL 2020 Partnership Agreement, and through the Portuguese National Innovation Agency (ANI) as a part of project ADIRA I4.0 with reference POCI-01-0247-FEDER-017922.

References

1. Bifet A, Holmes G, Pfahringer B, Kirkby R, Gavaldà R (2009) New ensemble methods for evolving data streams. In: Proceedings of the 15th ACM SIGKDD International Conference, KDD '09. ACM, New York, pp 139–148. ISBN 978-1-60558-495-9
2. Bishop CM (2006) Pattern recognition and machine learning (Information Science and Statistics). Springer, Berlin. ISBN 0387310738
3. Cerrada M, Sánchez R-V, Li C, Pacheco F, Cabrera D, de Oliveira JV, Vásquez RE (2018) A review on data-driven fault severity assessment in rolling bearings. Mech Syst Signal Process 99:169–196. ISSN 0888-3270. https://doi.org/10.1016/j.ymssp.2017.06.012. http://www.sciencedirect.com/science/article/pii/S0888327017303242
4. Demšar J (2006) Statistical comparisons of classifiers over multiple data sets. J Mach Learn Res 7:1–30. ISSN 1532-4435
5. El-Thalji I, Jantunen E (2015) A summary of fault modelling and predictive health monitoring of rolling element bearings. Mech Syst Signal Process 60-61:252–272. ISSN 0888-3270. https://doi.org/10.1016/j.ymssp.2015.02.008. http://www.sciencedirect.com/science/article/pii/S0888327015000813
6. Georgoulas G, Nikolakopoulos G (2016) Bearing fault detection and diagnosis by fusing vibration data. In: IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society, pp 6955–6960. https://doi.org/10.1109/IECON.2016.7794118
7. Harris FJ (1978) On the use of windows for harmonic analysis with the discrete Fourier transform. Proc IEEE 66(1):51–83. ISSN 0018-9219. https://doi.org/10.1109/PROC.1978.10837
8. Hashemian HM (2011) State-of-the-art predictive maintenance techniques. IEEE Trans Instrum Meas 60:3480–3492
9. Heng RBW, Nor MJM (1998) Statistical analysis of sound and vibration signals for monitoring rolling element bearing condition. Appl Acoust 53(1):211–226. ISSN 0003-682X. https://doi.org/10.1016/S0003-682X(97)00018-2. http://www.sciencedirect.com/science/article/pii/S0003682X97000182
10. Holmes MP, Gray AG, Isbell Jr CL (2008) QUIC-SVD: fast SVD using cosine trees. In: Advances in Neural Information Processing Systems 21, Proceedings of the Twenty-Second Annual Conference on Neural Information Processing Systems, Vancouver, pp 673–680. http://papers.nips.cc/paper/3473-quic-svd-fast-svd-using-cosine-trees
11. Howard I, Howard IM (1994) Defence Science, Technology Organisation (Australia), Aeronautical, and Maritime Research Laboratory (Australia). A review of rolling element bearing vibration: detection, diagnosis and prognosis / Ian Howard. DSTO Aeronautical and Maritime Research Laboratory Melbourne
12. Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, Yen N-C, Tung CC, Liu HH (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc R Soc A: Math Phys Eng Sci 454:903–995. https://doi.org/10.1098/rspa.1998.0193
13. Imaouchen Y, Kedadouche M, Alkama R, Thomas M (2017) A frequency-weighted energy operator and complementary ensemble empirical mode decomposition for bearing fault detection. Mech Syst Signal Process 82:103 – 116. ISSN 0888-3270. https://doi.org/10.1016/j.ymssp.2016.05.009. http://www.sciencedirect.com/science/article/pii/S0888327016300802
14. Jothilakshmi S, Ramalingam V, Palanivel S (2009) Unsupervised speaker segmentation with residual phase and MFCC features. Expert Syst Appl 36:9799–9804
15. Kumar S, Goyal D, Dang RK, Dhami SS, Pabla BS (2018) Condition based maintenance of bearings and gears for fault detection – a review. Mater Today: Proc 5(2, Part 1):6128–6137. ISSN 2214-7853. https://doi.org/10.1016/j.matpr.2017.12.219. http://www.sciencedirect.com/science/article/pii/S2214785317332078. 7th International Conference of Materials Processing and Characterization
16. Lin J, Chen Q (2013) Fault diagnosis of rolling bearings based on multifractal detrended fluctuation analysis and Mahalanobis distance criterion. Mech Syst Signal Process 38(2):515–533. https://doi.org/10.1016/j.ymssp.2012.12.014
17. Michie D, Spiegelhalter DJ, Taylor CC, Campbell J (eds) (1994) Machine learning, neural and statistical classification. Ellis Horwood, Upper Saddle River. ISBN 0-13-106360-X
18. Muda L, Begam M, Elamvazuthi I (2010) Voice recognition algorithms using mel frequency cepstral coefficient (mfcc) and dynamic time warping (dtw) techniques. CoRR, arXiv:1003.4083. http://dblp.uni-trier.de/db/journals/corr/corr1003.html#abs-1003-4083
19. Okoh C, Roy R, Mehnen J (2017) Predictive maintenance modelling for through-life engineering services. Procedia CIRP 59:196–201. ISSN 2212-8271. https://doi.org/10.1016/j.procir.2016.09.033. http://www.sciencedirect.com/science/article/pii/S2212827116309726. Proceedings of the 5th International

Conference in Through-life Engineering Services Cranfield University, 1st and 2nd November 2016
20. Poularikas AD (2010) Transforms and applications primer for engineers with examples and MATLAB®. Electrical Engineering Primer Series. CRC Press, Boca Raton. ISBN 9781420089325. https://books.google.pt/books?id=fgrdgTm45X4C
21. Rai Akhand, Upadhyay SH (2016) A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings. Tribol Int 96(Complete):289–306. https://doi.org/10.1016/j.triboint.2015.12.037
22. Rai Akhand, Upadhyay SH (2017) Bearing performance degradation assessment based on a combination of empirical mode decomposition and k-medoids clustering. Mech Syst Signal Process 93(Complete):16–29. https://doi.org/10.1016/j.ymssp.2017.02.003
23. Read J, Bifet A, Holmes G, Pfahringer B (2012) Scalable and efficient multi-label classification for evolving data streams. Mach Learn 88(1-2):243–272. ISSN 0885-6125
24. Saruhan H, Sandemir S, Çiçek A, Uygur I (2014) Vibration analysis of rolling element bearings defects. J Appl Res Technol 12(3):384–395. https://doi.org/10.1016/s1665-6423(14)71620-7
25. Tabrizi AA, Garibaldi L, Fasana A, Marchesielo S (2014) Ensemble empirical mode decomposition (EEMD) and Teager-Kaiser Energy Operator (TKEO) based damage identification of roller bearings using one-class support vector machine. In: Cam VLx, Mevel L, Schoefs F (eds) EWSHM - 7th European Workshop on Structural Health Monitoring. IFFSTTAR, Inria, Université de Nantes, Nantes. https://hal.inria.fr/hal-01022990
26. Wang Y-H, Yeh C-H, Young H-WV, Hu K, Lo M-T (2014) On the computational complexity of the empirical mode decomposition algorithm. Physica A: Stat Mech Appl 400(Complete):159–167. https://doi.org/10.1016/j.physa.2014.01.020
27. Zhou X, Garcia-Romero D, Duraiswami R, Espy-Wilson C, Shamma S (2011) Linear versus mel frequency cepstral coefficients for speaker recognition. In: 2011 IEEE Workshop on Automatic Speech Recognition Understanding, pp 559–564. https://doi.org/10.1109/ASRU.2011.6163888

Affiliations

Ricardo Sousa[1] · Joel Antunes[2] · Filipe Coutinho[3] · Emanuel Silva[2] · Joaquim Santos[2] · Hugo Ferreira[4]

Joel Antunes
jantunes@inegi.up.pt

Filipe Coutinho
f.coutinho@adira.com

Emanuel Silva
ejsilva@inegi.up.pt

Joaquim Santos
jsantos@inegi.up.pt

Hugo Ferreira
hmf@inesctec.pt

[1] LIAAD-INESC Porto, University of Porto, Campus of FEUP, Rua Dr. Roberto Frias, s/n, Porto, Portugal

[2] INEGI Porto, University of Porto, Campus of FEUP, Dr. Roberto Frias, s/n, Porto, Portugal

[3] ADIRA Metal Forming Solutions S.A., Canelas, Portugal

[4] CESE-INESC Porto, University of Porto, Campus of FEUP, Dr. Roberto Frias, s/n, Porto, Portugal