www.czasopisma.pan.pl PAN POLSKA AKADEMIA NAUK www.journals.pan.pl

PAN ARCHIVES OF ACOUSTICS
Vol. **45**, No. 2, pp. 283–295 (2020)
DOI: 10.24425/aoa.2020.133149

Research Paper

Enhancement in Bearing Fault Classification Parameters Using Gaussian Mixture Models and Mel Frequency Cepstral Coefficients Features

Youcef ATMANI[(1),(2)], Said RECHAK[(1)], Ammar MESLOUB[(3)], Larbi HEMMOUCHE[(3)*]

[(1)] *Ecole Nationale Polytechnique – ENP, El Harrach*
Algiers, Algeria; e-mail: youcef.atmani@g.enp.edu.dz, Said.rechak@g.enp.edu.dz

[(2)] *Ecole Nationale Supérieure de Technologie – ENST, Dergana*
Algiers, Algeria; e-mail: youcef.atmani@enst.dz

[(3)] *Ecole Militaire Polytechnique – EMP, Bordj El Bahri*
Algiers, Algeria; e-mail: mesloub.a@gmail.com
*Corresponding Author e-mail: l.hemmouche@gmail.com

(*received July 28, 2019; accepted January 30, 2020*)

Last decades, rolling bearing faults assessment and their evolution with time have been receiving much interest due to their crucial role as part of the Conditional Based Maintenance (CBM) of rotating machinery. This paper investigates bearing faults diagnosis based on classification approach using Gaussian Mixture Model (GMM) and the Mel Frequency Cepstral Coefficients (MFCC) features. Throughout, only one criterion is defined for the evaluation of the performance during all the cycle of the classification process. This is the Average Classification Rate (ACR) obtained from the confusion matrix. In every test performed, the generated features vectors are considered along to discriminate between four fault conditions as normal bearings, bearings with inner and outer race faults and ball faults. Many configurations were tested in order to determinate the optimal values of input parameters, as the frame analysis length, the order of model, and others. The experimental application of the proposed method was based on vibration signals taken from the bearing datacenter website of Case Western Reserve University (CWRU). Results show that proposed method can reliably classify different fault conditions and have a highest classification performance under some conditions.

Keywords: bearing faults; Gaussian mixture models; Mel frequency cepstral coefficients; feature extraction; diagnosis.

1. Introduction

The vibrations analysis technique is the most popular method used for diagnosis rotating machinery defects, in condition based maintenance. The development of resources allocated to the analysis of failures in these systems has generated improvements in the techniques and methodologies used. Many obtained results are used by the maintenance community, to prevent recurrence of failures and saving precious financial and technological resources. Consequently, the field failure analysis has become an integral part of the product life cycle.

In this way, the condition monitoring of bearing failures and their prediction are implemented from data of various sources. The causes of their deterioration can be: incorrect design, faulty installation, brine ling, corrosion, poor lubrication, fatigue, wear, or plastic deformation. The traditional used techniques can be classified into three principal domains namely, time domain analysis, frequency domain analysis, and time-frequency domain analysis. In addition, the fault signature pattern can be solved through a classification problem based on several statistical or machine learning emerging approaches.

As a part of learning-based techniques group, the Gaussian Mixture Model (GMM) approach was successful in the speech recognition domain. It is a probabilistic model designed for data density estimation, and mostly used in the tasks of speaker identification

and verification. However, in literature only few works are dedicated to the use of GMM as a classifier or MFCCs as features in the diagnosis by classification-based approach of mechanical defects. The number of coefficients is often fixed in advance (12, 13 or 14), without proof of its optimality, nor its link with the size of the signal frames, nor its relationship with the parameters of the classifier. This is why this paper is particularly focused on this aspect and tries to provide answers to these questions by associating the GMMs and the MFCCs in a supervised classification problem, intended for the diagnosis of bearing defects.

Randall and Antoni (2011) investigated the effectiveness of conventional signal processing techniques through the diagnosis of accelerometer-type vibratory signals from rolling bearings, in an environment marked by interference from other signals of other machine components such as gears. Different separation and denoising techniques (self-adaptive noise cancellation, discrete random separation, time synchronous averaging, spectral kurtosis etc.) were tested. Once filtering was optimized, the envelope analysis method was used in the final diagnosis.

Ericsson *et al.* (2004) compared different vibration analysis techniques for automatic detection of local defects in bearings. Several analysis tools for bearing condition monitoring were presented. About 103 laboratory and industrial environment test signals were used for describing a large-scale evaluation of automatic bearing monitoring methods. Results show that two wavelet-based and two based on envelope and periodization techniques are the best-performed methods.

Cerrada *et al.* (2018) reviewed the recent methods and techniques used to achieve the fault severity evaluation in the main components of the rolling bearings, such as inner race, outer race, and ball. The paper is mainly focused on two approaches. The first one is data-driven based approach such as signal processing for extracting the proper fault signatures associated with the damage degradation. The second concerns learning approaches that are used to identify degradation patterns with regards to health conditions.

Barszcz and Sawalhi (2012) introduced Minimum Entropy Deconvolution (MED) rolling element bearings fault detection. Two cases are presented to show its applications. The first one was taken from a fan bladed test rig. The second case was taken from a wind turbine with an inner race fault. The work particularly focuses on selecting the optimal parameters for the MED filter.

Girondin *et al.* (2013) proposed a vibration-based automated framework to readjust the fault frequencies from the theoretical frequencies of bearings in the transmission of a helicopter. The method provides the confidence index of the readjusted frequency. The algorithms were then tested with data from two test benches and from flight conditions. According to authors, results in flight conditions, frequency readjustment demonstrates good performances when applied on the spectrum.

Purushotham *et al.* (2005) presented a method for detecting localized bearing defects based on discrete wavelet transform (DWT) and Hidden Markov Models (HMM). Single, multiple and combined point defects on inner race, outer race and ball faults were considered for analysis. The variable time-frequency resolution detects well periodicity due to repetitive force impulses generated by each pass of the rolling element over the defect. The results were compared with others from spectrum analysis.

Narendiranath Babu *et al.* (2017) investigated fault diagnosis on journal bearing using Debauchies Wavelet-02 (DB-02). An experimental set up was mounted to acquire vibration signal from accelerometer. Many kinds of faults simulating oil losses cases were tested. Then a process of faults classification based on the Artificial Neural Networks (ANN) was used. The results show that the classification rate reached 85.7%.

Mannepalli *et al.* (2016) investigated identification of the accent in the speech recognition systems. The samples of speeches were collected from the native speakers of different accents of Telugu language for both training and testing. In this study, MFCC features were associated with GMM classifier for classification of the speech based on accent. According to authors, estimated efficiency of the proposed system to recognize the speaker, and the region he belongs, based on accent, is 91 %.

Shen *et al.* (2014) proposed a two-layer structure consisting of support vector regression machines (SVRMs) to recognize bearing fault patterns and track the fault sizes. Statistical parameters were first extracted to track the fault evolutions. The extracted features were then used to train the proposed two-layer SVRMs structure. Once the SVRM trained, features extracted from other vibration signals can be used to predict the unknown bearing health conditions. The effectiveness of the proposed method was validated by experimental datasets collected from a test rig.

Narendiranath Babu *et al.* (2018) applied Empirical Mode Decomposition (EMD) Artificial Neural Network (ANN) and Deep Neural Network (DNN) for self-aligning bearing fault diagnosis. Time domain and time-frequency domain features were extracted and introduced in the neural networks with the pattern recognition tool in MATLAB. The results show that ANN and DNN reached a high classification rate.

Yu *et al.* (2010) presented a diagnosis approach based on Gaussian mixture model (GMM). Discrete wavelet transform (DWT) feature vectors that represent different machine conditions, were extracted. After introducing the features into the classifier, the mod-

els were built. Diagnosis can be accomplished through finding out the GMM whose posteriori probability for a given testing feature vector is the maximum of all. Results were compared with those obtained by Multi-Layer Perceptron (MLP) neural network, in this work.

Aye *et al.* (2015) modelled a quantification index called Degradation Assessment Index (DAI) from Acoustic Emission (AE), for the detection of slow speed bearing faults, using an integrated approach. Incipient damage was detected under changing operating conditions. The model developed from the integration of Polynomial Kernel Principal Component Analysis (PKPCA), a Gaussian Mixture Model (GMM) and an Exponentially Weighted Moving Average (EWMA), is useful in the detection of slow speed bearings faults under variable operating conditions.

Yu (2011) proposed a Locality Preserving Projections (LPP)-based Feature Extraction (FE) approach. The effectiveness of the proposed approach for bearing defect and severity classification was evaluated experimentally on bearing test-beds. In a similar approach, GMM-based negative log likelihood probability, was evaluated for bearing defect and severity classification. The input features to the GMM were obtained from the vibration signal by using a so-called LPP.

Nelwamondo *et al.* (2006) used linear and nonlinear features extracted from time-domain vibration signals of a rotating machine with normal and defective bearings. The Multi-Scale Fractal Dimension (MFD), MFCC and kurtosis were used as features to classify bearing faults using GMM and HMM. Results of the used classifiers were compared in terms of accuracy and computational time.

Benkedjouh *et al.* (2018) investigated a method for machinery condition monitoring based on MFCC and Support Vector Machine (SVM). The approach is a supervised classification using extracted features such as temporal indicators and MFCC coefficients. The diagnosis accuracy assessment is carried out by conducting various experiments on acceleration signals collected from a rotating machinery. That is an automatic detection system for mechanical components defects.

2. Problem overview

The subject matter of this study, is related to the detection of anomalies in rotating machines as part of the predictive maintenance. In general, the diagnosis of defects can be done in three ways as shown in Fig. 1. The first one is based on the human (expert) auditory apparatus in its ability to distinguish between sounds emanating from different machine defects in operation. This was the original method of the maintenance concept. Inspired from this first concept, the second way is based on measured sounds. In the literature there are many works where acoustic waves are obtained and treated consequently by many signal processing methods. Some deemed reliable classification approaches are based on GMM and MFCCs as the features extraction vector. The result leads to the identification of fault classes. That is the same concept what we are going to develop in this paper. However, it is based on accelerometers signals. In other words, we operate in a step upstream of the acoustic acquisition, because we measure displacement, speed or acceleration of oscillatory movement generated by defects in organs. This is the third and the last way mentioned above.

As the bearing represents a vital piece in a rotating machine, the diagnosis of its defects will be the subject of this study. The classification process based on GMM

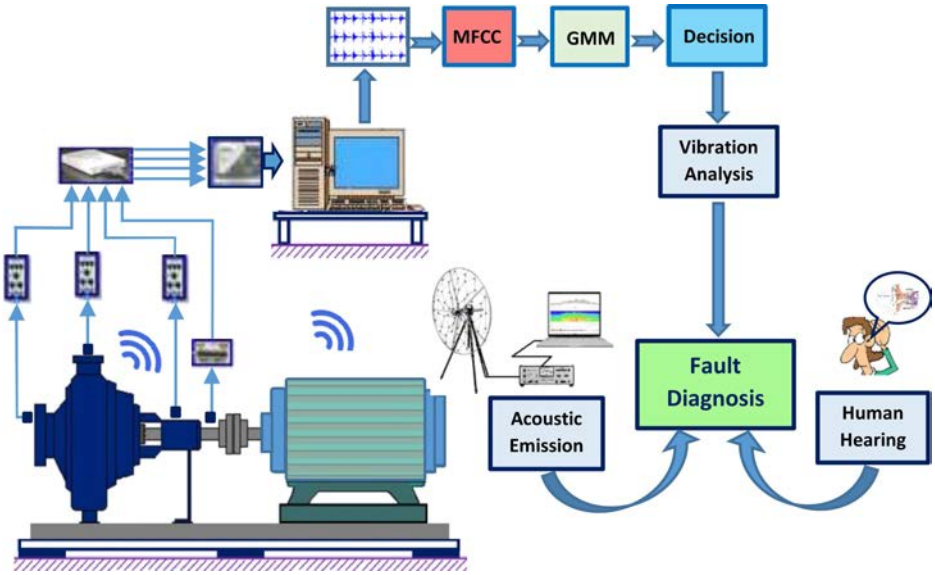

Fig. 1. The main possibilities offered to the diagnosis of defects in rotating machines.

and MFCC features extraction is applied to the fault identification related to rolling elements in the bearing shown in Fig. 2. Therefore, they present characteristic frequencies depending on the localization of the defect. Theoretically, five characteristic frequencies of fault locations can be calculated using the following equations (Randall, Antoni, 2011; Ocak, Loparo, 2004):

- Ball Pass Frequency, Inner race (BPFI):

$$\text{BPFI} = \frac{nf_r}{2}\left(1 + \frac{d}{D}\cos\phi\right), \tag{1}$$

- Ball Spin Frequency (BSF):

$$\text{BSF} = \frac{Df_r}{2d}\left(1 - \left(\frac{d}{D}\cos\phi\right)^2\right), \tag{2}$$

- Ball Pass Frequency, Outer race (BPFO):

$$\text{BPFO} = \frac{nf_r}{2}\left(1 - \frac{d}{D}\cos\phi\right), \tag{3}$$

- Fundamental Train Frequency (FTF):

$$\text{FTF} = \frac{f_r}{2}\left(1 - \frac{d}{D}\cos\phi\right), \tag{4}$$

where n is the number of rolling elements, d and D are the rolling element diameter and the pitch diameter of the bearing respectively, f_r is the rotational speed of the shaft (inner race speed) and ϕ is the contact angle (from the radial).

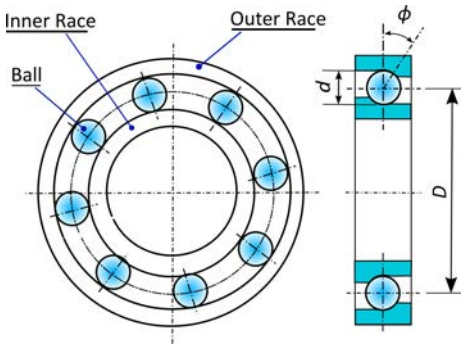

Fig. 2. Ball bearing elements and geometric parameters.

In others terms, localized faults in a rolling element bearing may occur in the outer race, the inner race, the cage, or a rolling element. High frequency resonances between the bearing and the response transducer are excited when the rolling elements strike a local fault on one of these components. The problem is how to detect and identify the various types of faults.

3. Proposed method description

This section describes the general principle of the proposed method as well as basic theoretical elements about cepstral coefficients and the classifier.

3.1. Classification scheme

The bearing fault diagnosis problem can be modelled as a classification problem, where one wants to get the best degree of separability between classes representing the fault conditions. In order to apply that concept to build a bearing fault diagnosis system capable of identifying the fault condition, a methodology based on GMM has been developed. This classifier has been proven in various classification problems, like radars (Mesloub *et al.*, 2018), speech recognition (Mannepalli *et al.*, 2016), and bioinformatics. Based on the well-known Expectation-Maximization (EM) algorithm for parameters estimation, it can approximate any probability density function (PDF) by a finite Gaussian mixture (McLachlan, Peel, 2000). It can also be completely represented with three parameters: mean vectors, covariance matrices, and the mixture weights.

Hence, the architecture of the proposed framework, as shown in Fig. 3, consists of four major stages. The first step is importing vibration signals (Loparo, 2012), and building the preprocessing data sets. Later, these ones are divided in training and test subsets. The second step is explained in Subsec. 3.3 below and clarified in Sec. 5, to show the relevancy of proposed feature. The third one is the model construction, which is in our case the estimation of GMM parameters, of each class. Finally, the last step is reserved to performance evaluation (confusion matrix and probability of correct classification). The chosen performance criterion is the higher Average Classification Rate (ACR).

The Gaussian mixture builds models for all possible faults types and the normal condition. Diagnosis of the bearing fault is achieved by calculating the probability of the feature vector, given the entire previously constructed fault model. GMM with maximum probability then determines the bearing condition. When relevant features are extracted, reference models that will be used to classify faults are built and are used to

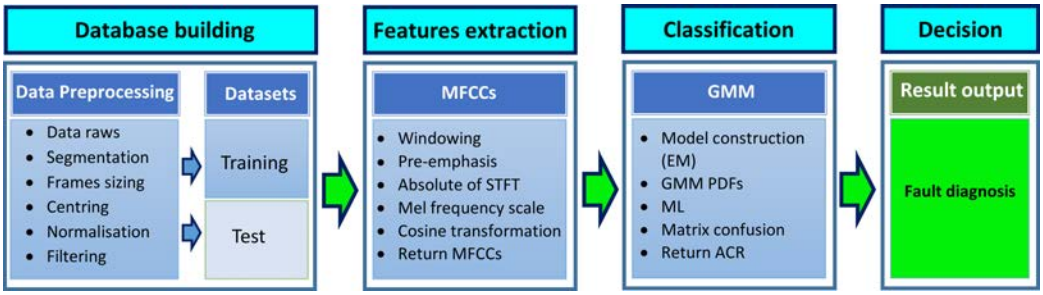

Fig. 3. The flow chart of the proposed bearing fault diagnostic procedure.

classify the fault conditions. In this case, we used the feature vectors in each frame as such without any manipulation. We tested the GMM with varying number of Gaussians, varying analysing frame lengths, varying MFCCs numbers and varying frame test lengths. The number of replicates is fixed to 100, in order to avoid the problem of EM algorithm, which can be trapped in one of the many local maxima of the likelihood function.

3.2. Features extraction

The complex cepstral coefficients known as MFCC have been widely and successfully used in the field of speech recognition. They can be defined as a wavelet in which frequency scales are placed on a linear scale for frequencies less than 1 kHz and on a log scale for frequencies above 1 kHz. It is well known too, that vibration signals contain both linear and non-linear features (Nelwamondo *et al.*, 2006). Therefore, the MFCCs with their time and frequency information of the signal, makes them more useful for the feature extraction.

The process consists of a transformation of the signal from the time domain to frequency domain and mapping the transformed signal in hertz, onto Mel-scale (Huang *et al.*, 2001; Wang *et al.*, 2002). For the calculation of MFCCs, it is necessary to go through the following steps:

- pre-emphasis filtering,
- take the absolute value of the short time Fourier transformation using windowing, in Fig. 4,
- Warp to auditory frequency scale (Mel-scale),
- take the discrete cosine transformation of the log-auditory-spectrum,
- return the firsts q MFCCs.

Pre-emphasis filtering, which is a special kind of finite impulse response (FIR), can be used to compensate the fluctuation of the spectrum energy between low frequencies and higher frequencies.

Let $x[n]$ be the raw signal at sample n, and $s[n]$ the signal after the high-pass filtering

$$s[n] = x[n] - \alpha x[n-1], \qquad n = 1, 2, ..., N, \tag{5}$$

where α is a parameter controlling how much it is filtered and is often set between 0.95 and 1 in practice.

The next step is to transform the signal from time domain to frequency domain by applying Short Time Fourier Transformation (STFT) together with a window function. It assumes that the signal over a very short time period is at least nearly stationary, thus able to be transformed to frequency domain. This can be done by:

$$\begin{aligned} X_a[k] &= \sum_{n=0}^{N-1} s[n] \cdot w_a[n] \cdot e^{-\frac{i2\pi kn}{N}} \\ &= \sum_{n=0}^{N-1} s[n] \cdot w_a[n] \cdot e^{-i\omega k}, \quad 0 \le k < N, \end{aligned} \tag{6}$$

where $w_a[n]$ is the window function, which is a zero valued function everywhere except inside the window, and i is the imaginary unit.

To keep the frames continuous, a Hamming window is preferred

$$\begin{aligned} &w_a[n] = \alpha - \beta \cos\left(\frac{2\pi n}{N-1}\right), \\ &0 \le n < N, \qquad \alpha = 0.54, \qquad \beta = 1 - \alpha = 0.46. \end{aligned} \tag{7}$$

A fact of human hearing ability is that we are more sensitive to sounds between 20 and 1000 Hz. Thus, it is less efficient to assign a signal the same scale at high frequencies as at lower frequencies. An adjustment can be made by mapping the data from Hertz-scale onto Mel-scale

$$\text{Mel} = \begin{cases} f, & f \le 1000, \\ 2595 \log_{10}\left(1 + \dfrac{f}{700}\right), & f > 1000. \end{cases} \tag{8}$$

In addition, its inverse is given by

$$f = \begin{cases} \text{Mel}, & \text{Mel} \le 1000, \\ 700\left(e^{\text{Mel}/2595} - 1\right), & \text{Mel} > 1000. \end{cases} \tag{9}$$

Given the STFT of a input window frame $x_a[k]$, we define a filterbank with M filters ($m = 1, 2, ..., M$) that are linear on Mel scale but nonlinear on Hertz scale, where m is triangular filter given by

$$M_m[k] = 1 - \left| \frac{k - \frac{N-1}{2}}{\frac{N-1}{2}} \right|, \tag{10}$$

where N is the length of the filter. Notice again that these filters are linear on Mel scale and they need to be transformed back to Hertz scale. Thus, we can then compute the log-energy of each filter as

$$S[m] = \ln\left[\sum_{k=0}^{N-1} |X_a[k]|^2 M_m[k]\right], \quad 0 < m \le M. \tag{11}$$

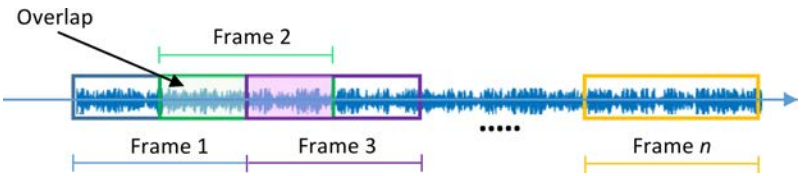

Fig. 4. Windowing is applied to minimize the discontinuities at the edge of each frame.

The Mel-frequency cepstrum coefficients are then the discrete cosine transform of the M filter outputs:

$$c[q] = \sum_{m=0}^{M-1} \left[S[m] \cos\left(\frac{\pi q \left(m - \frac{1}{2}\right)}{M} \right) \right], \quad 0 < m \leq M. \quad (12)$$

3.3. Gaussian mixture models

The GMMs are widely used in data mining, pattern recognition, machine learning, and statistical analysis. The empirical probability distribution of sampled data can be estimated by a GMM using a linear combination of Gaussian distributions (Duda *et al.*, 1995; Dempster *et al.*, 1977). It is a statistical model using a weighted sum of probability density functions of multiple Gaussian distributions to approach the empirical distribution of sampled data. In our work, as shown in Fig. 5, this data is formed by MFCC concatenated for different frames.

A complete Gaussian mixture model θ is parameterized by mixture weights w_i, mean vectors μ_i, and the covariance matrices Σ_i from all the M mixture components (Bishop, 2006)

$$\theta = \{\mathbf{w}, \boldsymbol{\mu}, \boldsymbol{\Sigma}\}, \quad (13)$$

where

$$\mathbf{w} = \{w_i\}, \qquad i = 1, 2, ..., M,$$
$$\boldsymbol{\mu} = \{\mu_i\}, \qquad i = 1, 2, ..., M,$$
$$\boldsymbol{\Sigma} = \{\Sigma_i\}, \qquad i = 1, 2, ..., M.$$

The Gaussian mixture density is

$$p(\mathbf{x}|\theta) = \sum_{i=1}^{M} w_i p_i(\mathbf{x}). \quad (14)$$

For a N-dimensional vector $\mathbf{x}$, the multivariate Gaussian distribution takes the form :

$$\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{N/2}} \frac{1}{|\boldsymbol{\Sigma}|^{1/2}} \cdot \exp\left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\}, \quad (15)$$

where N is the MFCC number.

Hence, the density function is:

$$p_i(\mathbf{x}) = \mathcal{N}(\mathbf{x}|\mu_i, \Sigma_i). \quad (16)$$

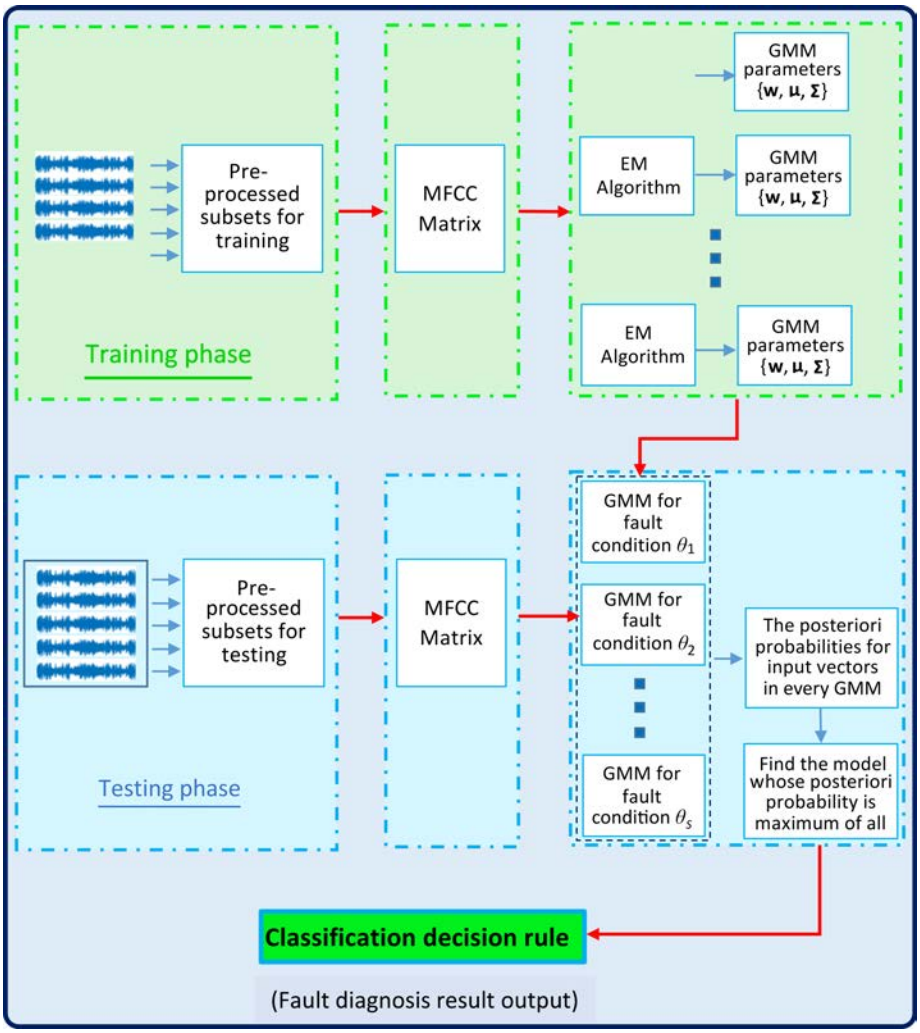

Fig. 5. The flow chart of the classification of data using GMM.

The description of the GMM is a linear superposition of Gaussians,

$$p(\mathbf{x}) = \sum_{i=1}^{M} w_i \mathcal{N}(\mathbf{x}|\mu_i, \Sigma_i). \tag{17}$$

Normalization and positivity require

$$\sum_{i=1}^{M} w_i = 1, \qquad 0 \le w_i \le 1. \tag{18}$$

For fault diagnosis, each fault condition (class) is represented by a GMM. Given training fault features from each fault condition, the goal of training model is to estimate the GMM parameters. It can be done using the Expectation Maximization (EM) algorithm, which yields a maximum likelihood estimate (Dempster *et al.*, 1977; Mc Lachlan, Krishnan, 2008)

$$\widehat{s} = \arg \sum_{t=1}^{T} \log p(x_t|\theta_k), \tag{19}$$

where k represents the index of the type of fault, whereas C is the total number of known fault conditions (classes) and $\mathbf{x} = \{x_1, x_2, ..., x_T\}$ is the unknown fault vibration segment and $p(x_t|\theta_k)$ is given in Eq. (14).

Once the GMMs for C fault conditions have been trained, diagnose a testing fault feature vector is straightforward. We just need to find out the model θ, which has the maximum posteriori probability for a given testing fault feature vector.

4. Application case of the proposed method

The proposed method of classification is applied to bearing faults data collection obtained from the CWRU bearing data centre. The vibration signals are recorded with names "xxx.mat" in Matlab files format. The "xxx" means three numbers, as given in Table 1 for the selected ones. Every file contains two signals DE and FE (Drive End and Fan End). More details on experimental set up are described in (Loparo, 2012). The signals selected for our classification contain healthy and faulty bearing at the inner, outer (at centred position) and ball. The sampling frequency for all the selected vibration signals is 48 kHz.

Table 1. The 48k bearing fault data files selected.

Fault width 0.007 in (0.18 mm)			Normal data
IR	Ball	OR centred	
109	122	135	97
110	123	136	98
111	124	137	99
112	125	138	100

Overall, the selected dataset covers 16 scenarios for the four considered fault conditions (04 classes) because each of the four ones represents four loading cases (four different speeds). In addition, by considering the two signals relating to the measuring points (Drive End and Fan End of each scenario), we get a total of 32 signals to be processed in our classification scheme.

The procedure of pre-processing began by reshaping the all dataset vibration signals because their lengths in their original forms are not the same. Thereby, this first operation generated 14 blocks of uniform sizes of 229,376 samples, for each class. After, these created blocks have undergone the necessary pretreatments (centering, normalization etc.). An example of the pre-processed signals corresponding to each of the four classes is shown in Fig. 6. Every block is divided in analysis frames (or segments) with various lengths to perform which ones give the higher classification rate. Then, six frame lengths, as shown in Table 2, were explored.

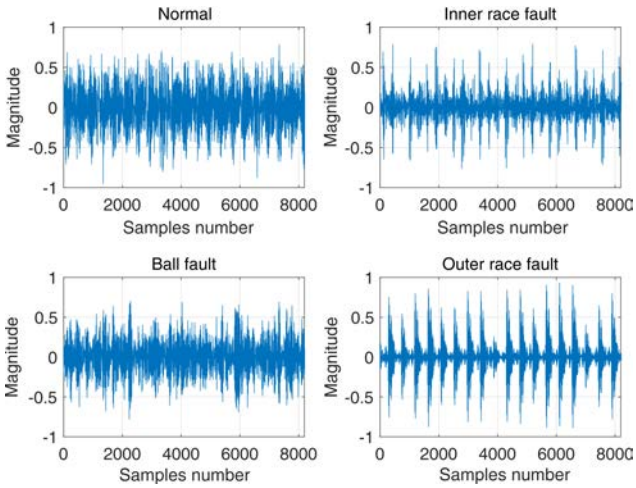

Fig. 6. The four classes pre-processed vibration signals.

Table 2. Number of used frames for each frame length.

Frame length	Training data	Testing data	All data
512	3584	2688	6272
1024	1792	1344	3136
2048	896	672	1568
4096	448	336	784
8192	224	168	392
16384	112	84	196

In order to create a training and testing subsets from the dataset, first eight blocks were devoted to training and modelling (57.14%), and the rest were kept for testing the proposed method (42.86 %).

The feature extraction is applied for both training and testing data. This process based on the number of input rows, the Hamming window length, the hop length, and other parameters, computes at its end

the MFCCs for each frame. Features extracted in an arbitrary frame of the vibration signal are shown in Fig. 7. The horizontal axis shows index of each coefficient among the fourteen (14) extracted from one frame. Also, all conditions under investigation (four classes 04) are displayed in the same plot. As the lengths of the windows chosen (time durations) are a function of the size of the analysis frames and the sampling frequency, the result leads to variable time duration. Features Extraction are done with variable MFCC number. This solution will allow us to test different configurations later and to choose the one that maximizes the ACR (the best).

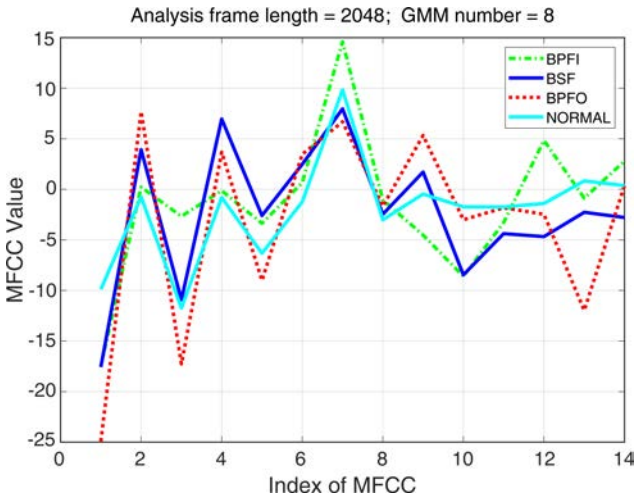

Fig. 7. MFCC values corresponding to different fault conditions.

The last interesting parameter for this methodology, as shown in Fig. 8, is the number of analyzed frames taken in a test sequence (or testing frame). Hence, many tests are reached in order to select the best combination giving the higher ACR.

5. Results and discussion

Along this study, a better classification means a maximum ACR. Therefore, a value of 100% is considered ideal. It is obtained from the average of the diagonal terms of the confusion matrix, at the end of each classification test. This is the reference criterion whose value judges the optimality of the other parameters involved in the classification process, like number of components in the Gaussian mixture, the number of MFCCs, the length of the analysis frame, and the length of the test sequence.

The purpose of this study is to find a combination of the optimal parameters that allows a better classification. To achieve this, different values of each of the classification parameters were tested. The results obtained, and the influences of each parameter on the classification process are presented below.

5.1. Effect of the GMM components number

In the literature, the influence of the number of components of the mixture on the classification process by GMM, is a subject that occupies an important place. Hence, the estimation of optimal number of pa-

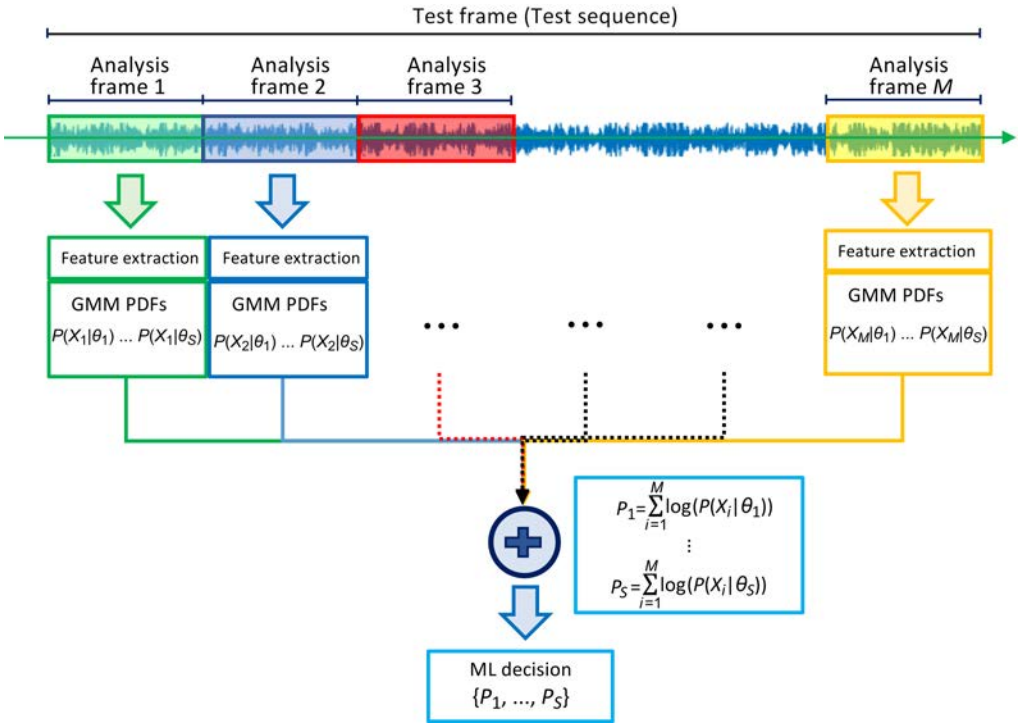

Fig. 8. Feature extraction regarding testing and analysing frames.

rameters is essential because the parameters with inappropriate components may not evaluate the mixture model accurately. It is a challenging attempt to decide the optimum number of components. To estimate both the model order and components together some techniques and criterion were developed (Akaike's information criterion (AIC), Minimum Description Length (MDL), Bayesian Inference Criterion (BIC), among others) (Akaike, 1974; McKenzie, Alder, 1994).

However, in this work, this aspect has not been developed as it represents in itselfan entire dimensional study. The purpose of this work is to find the number of mixture components, which maximises the ACR. Therefore, we try increasing number of mixture components from 2 to 12 (with the step of 2). Two cases were explored. The first, illustrated in Fig. 9, relates to a frame of 2048 samples. The numbers of MFCCs tested are 10, 12, 13, and 14. At this stage, we do not yet know the optimal number needed. We used the most used numbers in the literature. In the end, the number of eight (8) components was obtained and satisfied the highest ACR, regardless of the number of MFCCs. The second case illustrated in Fig. 10, supported the choice of the first one, because the 1024-samples length frame makes the component number dependent on the number of MFCCs. It is also emphasized that the classification rates in this case are relatively lower than the first ones.

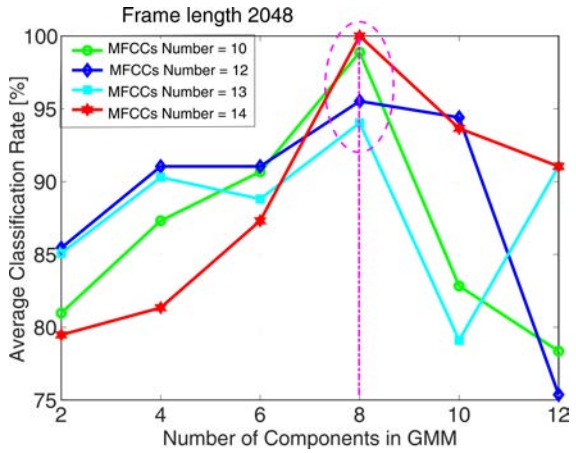

Fig. 9. ACR *versus* GMM components number with frame length of 2048 samples.

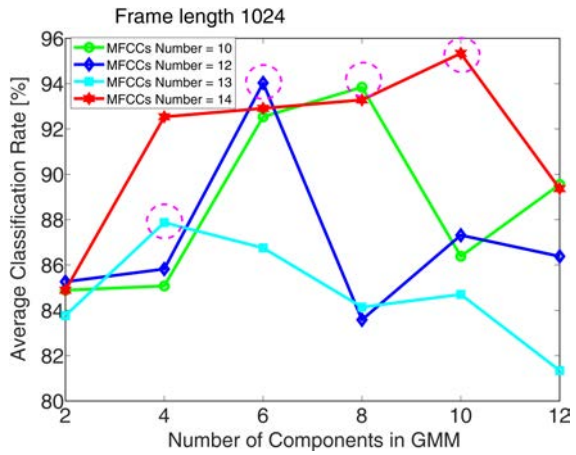

Fig. 10. ACR *versus* GMM components number with frame length of 1024 samples.

On the other hand, it should be noted that the choice of length frames of 2048 and 1024 samples is fortuitous at this stage of the study. It is mostly justified by their average positions in a range of frames included between 512 and 16384 samples. As a consequence and for all these arguments and experimental considerations, the fixed number of components to be considered in all the executions of the EM algorithm by the GMM classifier, is eight (8). This means that with this value, the EM process gives the best estimation of the model parameters, and hence the best classification accuracy of bearing fault conditions.

5.2. Effect of the MFCC number

The choice of the number of MFCCs to include in the classification process is largely empirical (Rabiner, Schafer, 2011). Therefore, the method was tested with two parameter combinations. The features extraction process used frame lengths of 1024 samples and 2048 samples. The time durations are equal respectively to 21.3 ms and 42.7 ms with 50% overlap. The number of MFCCs tested was between 2 and 20. Trying to improve the classification performance, we investigate the effect of changing the number of MFCC extracted from each frame and compare the classification results as shown in Fig. 11. It quickly became apparent that the very high cepstral coefficients were not helpful for classification, neither were the very low ones. For more than 14 coefficients, the performance became rapidly degraded. Indeed, too many coefficients kill the classification.

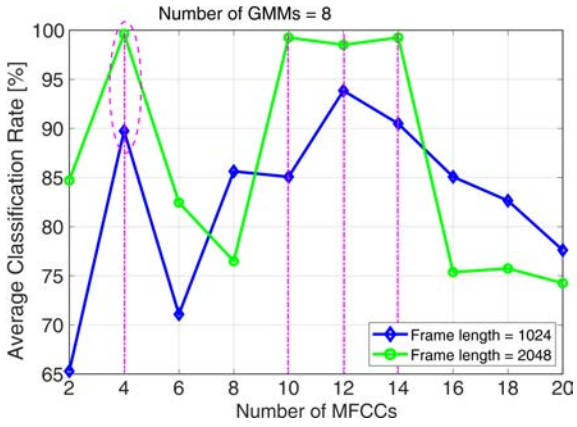

Fig. 11. ACR *versus* number of MFCCs (for N GMMs = 8).

Two areas are particularly interesting. The first one is the number of MFCCs equal to 4 with the highest ACR from the curves. The second is a range of MFCC number between 10 and 14. For the latter the ACR

is also high. It fits well with literature results. The number of 12 or 13 are the most cited and most used.

On the one hand, the frame of 2048 samples seems once again to have a higher classification rate compared to that of 1024 samples. On the other hand, the latter has a high ACR with the number of MFCCs of 4 in the first zone, but relatively fluctuating in the range of MFCCs between 10–14.

It follows that the choice of the number of MFCCs will be four (4), for obvious reasons of saving in computing time. It remains to be determined whether the frame of 2048 will always remain the best one or another may eventually replace it. The answer to this question will be discussed in the following sub section, which examines the influence of the length of the analysis frame on the classification rate. In all cases, the number of MFCCs is now fixed to the value of four (4) and its confusion matrix is shown in Table 3.

Table 3. Confusion Matrixes in case of: $N_{\mathrm{MFCC}} = 4$; NGMM = 8; frame length = 2048.

Test sequence (Frame test) = 10 analysis frames				
	BPFI	BSF	BPFO	Normal
BPFI	1.0000	0	0	0
BSF	0	0.9851	0	0.0149
BPFO	0	0	1.0000	0
Normal	0	0.0299	0	0.9701

In any way, the optimal number of coefficients depends on the quantity of training data, the details of the training algorithm (in particular how well the PDFs can be modelled as the dimensionality of the feature space increases), the Gaussian mixture, the background noise characteristics, and maybe the available computing resources. If we call up the limitations of MFCC, a serious one of the original MFCC feature extraction technique is that the filter bandwidth is not an independent design parameter but instead is determined by the frequency range of the filter bank, the number of filters, and type of window used.

5.3. Effect of the analysing frame length

In this work, the initial goal was to segment into frames having the lengths: 64, 128, 256, 512, 1024, 2048, 4096, 8192, and 16384 points. The observation is that for the frames less than 512 (256, 128, and 64), the process does not converge with the parameters initially set for the automatic fault classification. As a result, these three frames have been eliminated because they require special precautions. The frames studied vary between 512 and 16384 samples. Each frame is then multiplied by a Hamming window of the same length.

As shown in Fig. 12, the higher ACR is obtained for the frame length with 2048 samples. If we also consider the number of MFCCs, four (4) is best than the

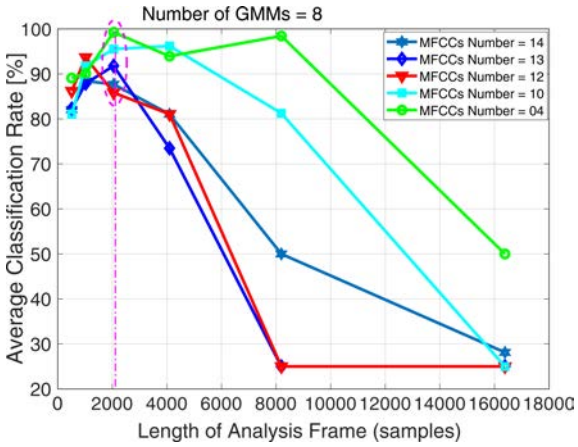

Fig. 12. ACR *versus* frame analysis length (GMM Number = 8).

rest shown in curves. This result confirms the one that was obtained previously in the subsection above. With regard to this evolution of the curves, the optimal value of the length frame is 2048 samples.

5.4. Effect of the number of analysing frames

The block diagram illustrating the analysis frames numbers present in each test sequence, was explained in Fig. 8. In all the tests carried out before, the test sequence was always fixed at ten (10) analysis frames. In this subsection, we evaluated three other possibilities in terms of the number of analysis frames contained in a test sequence (frame test): five (5), two (2), and one (1) frame.

The obvious result is shown in Fig. 13. It is mentioned that the highest ACR is assigned to the longest test sequence (the one that contains ten (10) frames of analysis). It means that the estimation is better with the longer samples then the lower ones. The confusion matrices of each case A, B, C, and D, are shown in Table 4. The curve corresponding to the number

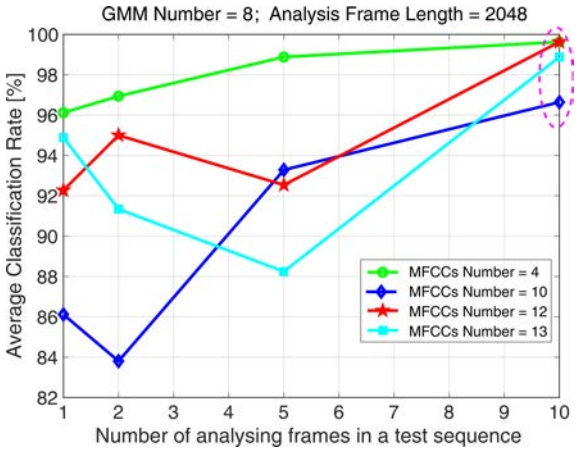

Fig. 13. ACR *versus* the test sequence (frame test).

Table 4. Confusion matrixes in case of: $N_{\mathrm{MFCC}} = 4$; $N_{\mathrm{GMM}} = 8$; frame length = 2048.

Frame test with 1 analysis frames; average accuracy = 94.8881%				
	BPFI	BSF	BPFO	Normal
BPFI	1.0000	0	0	0
BSF	0	0.9030	0	0.0970
BPFO	00045	0	0.9955	0
Normal	0	0.1030	0	0.8970
Frame test with 2 analysis frames; average accuracy = 96.3433%				
	BPFI	BSF	BPFO	Normal
BPFI	1.0000	0	0	0
BSF	0	0.9493	0	0.0507
BPFO	0.0030	0	0.9970	0
Normal	0	0.0925	0	0.9075
Frame test with 5 analysis frames; average classification rate = 96.9449%				
	BPFI	BSF	BPFO	Normal
BPFI	1.0000	0	0	0
BSF	0	0.9552	0	0.0448
BPFO	0	0	1.0000	0
Normal	0	0.0299	0	0.9701
Frame with 10 analysis frames; average classification rate = 99.1604%				
	BPFI	BSF	BPFO	Normal
BPFI	1.0000	0	0	0
BSF	0	0.9851	0	0.0149
BPFO	0	0	1.0000	0
Normal	0	0.0299	0	0.9701

of four (4) MFCCs has the highest classification rate. This again represents a confirmation of the results obtained previously. The other curves were plotted to show these notorious differences between the ACRs related to every MFCC number used.

5.5. Summarization

At the end of this study, the parameters of the classification optimized during the different tests are finally set to their final optimal values. Then, the number of components of the mixture is fixed at 8, the number of MFCCs is set at 4, the size of the analysis frame is set at 2048 samples and the size of the test sequence is fixed at 10 analysis frames.

The results of this last test carried out with these optimal parameters below are represented by the confusion matrix in Table 5, the scatter plot 2D (two dimensions) of Fig. 14 and the scatter plot 3D (three dimensions) of Fig. 15. In these figures, we can easily notice some dispersion in the BPFO and BSF classes. It is probably due to modelled noise or irregularities related to the structures of covariance matrices estimated by the EM algorithm. The average classification rate achieved reached 99.6269%. This is a good result obtained by the methodology developed during this experimental study.

Table 5. Confusion matrix NMFCC = 4; NGMM = 8; frame length = 2048.

Test sequence = 10 analysis frames; average classification rate = 99. 6269%				
	BPFI	BSF	BPFO	Normal
BPFI	1.0000	0	0	0
BSF	0	0.9851	0	0.0149
BPFO	0	0	1.0000	0
Normal	0	0	0	1.0000

The process of optimizing the parameters in the proposed methodology has made the GMM-MFCC duo powerful enough to provide a good classification. However, two aspects that were not addressed during this work require perspectives for exploration.

The first one is related to the extraction of features through an optimization of the proportion of data needed for training and the study of the influence of the number of filters used for the extraction of MFCCs. The latter is related to the windows used, the size of the Fourier transform and the different bands

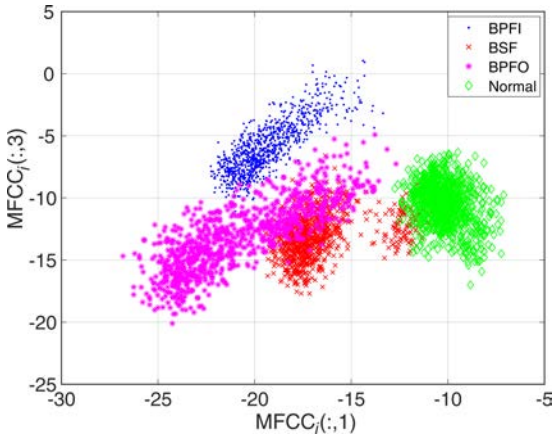

Fig. 14. Scatter plot 2D for the features (MFCCs) of the bearing fault conditions.

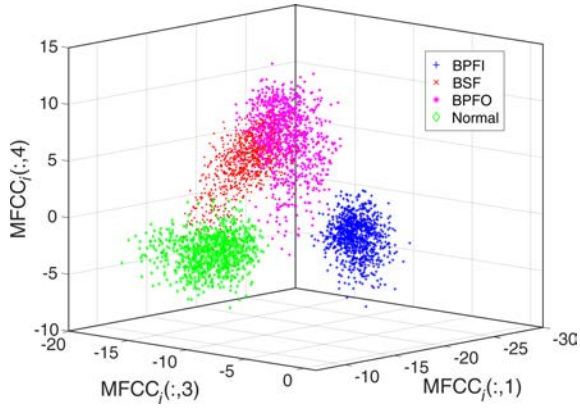

Fig. 15. Scatter plots 3D for the features (MFCCs) of the bearing fault conditions.

of the frequency spectrum. Indeed, different possibilities must be tested and different relationships must be established to optimize this parameter, to increase the range of MFCCs for good classification. Also, Singh *et al.* (2012) prepared a table in which are set the parameters mentioned above and influencing the sensitivity of the MFCCs in a classification process. In addition, Singh *et al.* scored the channel distortion among these.

The second aspect is related to the GMM classifier, which consists of multivariate Gaussian distribution components. Each component is defined by its mean, covariance, and a vector of mixing proportions (weights) defines the mixture. These parameters are estimated by using maximum likelihood (ML) estimation obtained iteratively using expectation-maximization algorithm (EM). Gaussian mixture models require that you specify a number of components before being fit to data. For many cases, it might be difficult to know the appropriate number of components. Akaike Information Criterion (AIC) and Bayes Information Criterion (BIC) are statistic fits and can help choosing of the best fitting Gaussian mixture model over varying numbers of components. They can be used to compare multiple model fit to the same data. As scalars, smaller values of them allow a better model. The negative log likelihood, the number of iterations in the Expectation-Maximization (EM) algorithm, the regularization parameter value and effect of the tolerance for posterior probabilities are the other parameters that can be analyzed for optimization problem in perspectives explorations.

6. Conclusion

This study used a combination of GMM with MFCC features for intelligent fault diagnosis of bearing fault conditions. It is a supervised classification-based approach. The optimization of its input parameters, namely the number of MFCCs, the size of the analysis frame, the number of components of the mixture, and others, made it possible to reach at the end a high Average Classification Rate (ACR). In addition, and although it has been applied to a reputable noisy database with some non-stationary characteristics, the GMM-MFCC classification has proved successful. This can open new possibilities for this duo, in its application to other types of failures present in rotating machinery (gears, misalignment etc.).

Acknowledgements

This work was completed in the GMD laboratory (ENP, El Harrach-Algiers, Algeria). The authors would like to thank the Delegated Ministry for Scientific Research (MDRS) for granting financial support for CNEPRU Research Project No. J03004820140301. The authors would like to thank Professor K.A. Loparo of CaseWestern Reserve University for his permission to use their bearing database.

References

1. Akaike H. (1974), A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, **19**(6): 716–723, doi: 10.1109/tac.1974.1100705.
2. Aye S.A., Heyns P.S., Thiart C.J. (2015), Fault detection of slow speed bearings using an integrated approach, *IFAC-Papers On Line*, **48**(3): 1779–1784, doi: 10.1016/j.ifacol.2015.06.344.
3. Barszcz T., Sawalhi N. (2012), Fault detection enhancement in rolling element bearings using the minimum entropy deconvolution, *Archives of Acoustics*, **37**(2): 131–141, doi: 10.2478/v10168-012-0019-2.
4. Benkedjouh T., Chettibi T., Saadouni Y., Afroun M. (2018), Gearbox Fault Diagnosis Based on Mel-Frequency Cepstral Coefficients and Support Vector Machine, [in:] Amine A., Mouhoub M., Ait Mohamed O., Djebbar B. [Eds], *Computational Intelligence and Its Applications. CIIA 2018. IFIP Ad-*

vances in Information and Communication Technology, Vol. 522, pp. 220–231, Springer, doi: 10.1007/978-3-319-89743-1_20.

5. Bishop C.M. (2006), *Pattern Recognition and Machine Learning*, Springer Science + Business Media LLC.
6. Cerrada M. *et al.* (2018), A review on data-driven fault severity assessment in rolling bearings, *Mechanical Systems and Signal Processing*, **99**, 169–196, doi: 10.1016/j.ymssp.2017.06.012.
7. Dempster A.P., Laird N.M., Rubin D.B. (1977), Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society, Series B (Methodological)*, **39**(1): 1–38, www.jstor.org/stable/2984875.
8. Duda R.O., Hart P.E., Stork D.G. (1995), *Pattern Classification and Scene Analysis. Part 1: Pattern Classification*, 2nd ed., John Wiley and Sons, Inc.
9. Ericsson S., Grip N., Johansson E., Persson L.E., Sjöberg R., Strömberg J.O. (2004), Towards automatic detection of local bearing defects in rotating machines, *Mechanical Systems and Signal Processing*, **19**(3): 509–535, doi: 10.1016/j.ymssp.2003.12.004.
10. Girondin V., Pekpe M.K., Cassar J.P., Morel H. (2013), Bearings fault detection in helicopters using frequency readjustment and cyclostationary analysis, *Mechanical Systems and Signal Processing*, **38**(2): 499–514, doi: 10.1016/j.ymssp.2013.03.015.
11. Huang X., Acero A., Hon H. (2001), *Spoken language processing: a guide to theory, algorithm, and system development*, Prentice Hall PTR.
12. Loparo K.A. (2012), *Bearings Vibration Data Sets*, Case Western Reserve University, http://csegroups.case.edu/bearingdatacenter/home.
13. Mannepalli K., Sastry P.N., Suman M. (2016), MFCC-GMM based accent recognition system for Telugu speech signals, *International Journal Speech Technology*, **19**(1): 87–93, doi: 10.1007/s10772-015-9328-y.
14. McKenzie P., Alder M. (1994), Selecting the optimal number of components for a Gaussian mixture model, *Proceedings of IEEE International Symposium on Information Theory*, doi: 10.1109/ISIT.1994.394626.
15. McLachlan G.J., Krishnan T. (2008), *The EM algorithm and extensions*, 2nd ed., New Jersey: John Wiley & Sons Inc.
16. McLachlan G.J., Peel D. (2000), *Finite mixture models*, New York: John Wiley & Sons.
17. Mesloub A., Abed-Meraim K., Belouchrani A. (2018), Ground moving target classification based on micro-Doppler signature using novel spectral information features, *2017 IEEE European Radar Conference (EURAD)*, pp. 255–258, doi: 10.23919/EURAD.2017.8249195.
18. Narendiranath Babu T., Aravind A., Rakesh A., Jahzan M., Rama Prabha D. (2018), Application of EMD ANN and DNN for self-aligning bearing fault diagnosis, *Archives of Acoustics*, **43**(2): 163–175, doi: 10.24425/122364.
19. Narendiranath Babu T., Himamshu H.S., Prabin Kumar N., Rama Prabha D., Nishant C. (2017), Journal bearing fault detection based on daubechies wavelet, *Archives of Acoustics*, **42**(3): 401–414, doi: 10.1515/aoa-2017-0042.
20. Nelwamondo F.V., Marwala T., Mahola U. (2006), Early classifications of bearing faults using hidden Markov models, Gaussian mixture models, Mel-frequency Cepstral coefficients and fractals, *International Journal of Innovative Computing, Information and Control*, **2**(6): 1281–1299, http://www.ijicic.org/05-077-1.pdf.
21. Ocak H., Loparo K.A. (2004), Estimation of the running speed and bearing defect frequencies of an induction motor from vibration data, *Mechanical Systems and Signal Processing*, **18**(3): 515–533, doi: 10.1016/s0888-3270(03)00052-9.
22. Purushotham V., Narayanan S., Suryanarayana A.N.P. (2005), Multi-fault diagnosis of rolling bearing elements using wavelet analysis and hidden Markov model based fault recognition, *NDT & E International*, **38**(8): 654–664, doi: 10.1016/j.ndteint.2005.04.003.
23. Rabiner L.R., Schafer R.W. (2011), *Theory and applications of digital speech processing*, Pearson Higher Education, Inc.
24. Randall R.B., Antoni J. (2011), Rolling element-bearing diagnostics – A tutorial, *Mechanical Systems and Signal Processing*, **25**(2): 485–520, doi: 10.1016/j.ymssp.2010.07.017.
25. Shen C., Wang D., Liu Y., Kong F., Tse P.W. (2014), Recognition of rolling bearing fault patterns and sizes based on two-layer support vector regression machines, *Smart Structures and Systems*, **13**(3): 000-000, http://dx.doi.org/10.12989/sss.2014.13.3.000.
26. Singh N., Khan R.A., Shree R. (2012), MFCC and prosodic feature extraction techniques: a comparative study, *International Journal of Computer Applications*, **54**(1): 0975–8887, doi: 10.5120/8529-2061.
27. Wang J.C., Wang J.F, Weng Y.S. (2002), Chip design of MFCC extraction for speech recognition, *Integration,* the *VLSI Journal*, **32**(1–2): 111–131, doi: 10.1016/S0167-9260(02)00045-7.
28. Yu G., Li C., Sun J. (2010), Machine fault diagnosis based on Gaussian mixture model and its application, *International Journal of Advanced Manufacturing Technology*, **48**(1–4): 205–212, doi: 10.1007/s00170-009-2283-5.
29. Yu J. (2011), Bearing performance degradation assessment using locality preserving projections and Gaussian mixture models, *Mechanical Systems and Signal Processing*, **25**(7): 2573–2588, doi: 10.1016/j.ymssp.2011.02.006.