Classification of Bearing Faults Through Time-Frequency Analysis and Image Processing

Damiano Rossetti[1*], Yu Zhang[2], Stefano Squartini[1] and Stefano Collura[3]

[1] Università Politecnica delle Marche, Ancona, Italy, e-mail: *d.rossetti@pm.univpm.it, s.squartini@univpm.it*

[2] University of Lincoln, Lincoln, United Kingdom, e-mail: *yzhang@lincoln.ac.uk*

[3] Loccioni Group, Maiolati Spontini, Ancona, Italy, e-mail: *s.collura@loccioni.com*

Abstract— **The present work proposes a new technique for bearing fault classification that combines time-frequency analysis with image processing. This technique uses vibration signals from bearing housings to detect bearing conditions and classify the faults. By means of Empirical Mode Decomposition (EMD), each vibration signal is decomposed into Intrinsic Mode Functions (IMFs). Principal Components Analysis (PCA) is then performed on the matrix of the decomposed IMFs and the important principal components are chosen. The spectrogram is obtained for each component by means of the Short Time Fourier Transform (STFT) to obtain an image that represents the time-frequency relationship of the main components of the analyzed signal. Furthermore, Image Moments are extracted from the spectrogram images of principal components in order to obtain an array of features for each signal that can be handled by the classification algorithm. 8 images are selected for each signal and 17 moments for each image, so an array of 136 features is associated with every signal. Finally, the classification is performed using a standard machine learning technique, i.e. Support Vector Machine (SVM), in the proposed technique. The dataset used in this work include data collected for various rotating speeds and loads, in order to obtain a set of different operating conditions, by a Roller Bearing Faults Simulator.**

The results have shown that the developed technique provides classification effectively, with a single classifier, of bearing faults characterized by different rotating speeds and different loads.

Keywords: Empirical Mode Decomposition; Principal Component Analysis; Spectral Analysis; Image Processing; Machine Learing

I. Introduction

The detection and the prediction of faults in industrial machines is a crucial aspect to enhance reliability and reduce maintenance cost.

Bearings are widely used in induction motors and rotating machine, so they are common components in many mechanical systems. Due to their diffusion, many of the failures that occur in industrial machines are ascribed to these components, in fact statistical studies show that bearing failures account for around 50% of the total failures of these devices [1, 2, 3]. A bearing failure can accelerate machinery deterioration and bring dangerous consequences for the machine operators present, so bearing condition monitoring becomes an important measure to ensure machine safety.

Due to the importance of this topic, a lot of works and studies were published during past years by providing many different approaches to improve the detection and the classification of the faulty bearings [4]. Fault diagnosis of bearings is usually based on vibration signals, and a set of features is extracted from these signals in order to classify the faults. The features could be in the time domain, frequency domain or time-frequency domain, and the prevalent techniques include Kurtosis [5, 6], Wavelet Transform [7, 8], Cepstrum [9], or Envelope Spectrum [10].

The bearing behavior depend on several parameters and their diagnosis during regular operation will often involve analysis of non-stationary signals as the rotating speed, loads, and environmental conditions vary with time. Therefore, the algorithms for fault detection and classification should allow analysis of non-stationary or quasi-stationary signals.

Zvokelj et al. [11] realized the non-stationarity of the data collected from the monitoring of bearings, and so proposed an approach that combine the Principal Component Analysis (PCA) and the Ensemble Empirical Mode Decomposition (EEMD) methods. PCA reduces the dimensionality of the data and EEMD decomposes the signals into various time scales to allow the extraction of appropriate characteristics from the faulty bearing signals.

Both time and frequency domains provide some features that can be used as characteristics of these non-stationary vibration signals, useful to determinate the healthy state of the bearing. A means to extract these characteristics is by using images, which can provide a comprehensive description of vibration signals, including information about the various bearings faults.

In the work presented by Wei Li et al. [12], the spectrum images are processed with two-dimensional PCA to reduce the dimensions, then a minimum distance method is applied to classify the bearing faults.

Liang Hua et al. [13] presented a method where the acquired mechanical vibration signals are converted into color time-frequency spectrum images by the processing of pseudo Wigner-Ville distribution. Then a feature extraction method based on quaternion invariant moment is used, combining image processing technology and multi weight neural network technology.

Klein et al. [14] provided a method for the bearing fault diagnosis by applying the image processing techniques such as ridge tracking and related algorithms on the time-frequency representation, available by STFT or wavelet, of the non-stationary bearing signals.

All the works in literature deal with the analysis and the diagnosis of the bearing faults by trying to get as close as possible to the real working conditions of a bearing. Therefore, the main problems for this type of analysis is the non-stationarity of the signals or the variation of the operating parameters.

The method that is presented in this paper provides an innovative approach that uses time-frequency analysis and image processing with a machine learning technique for the bearing faults classification; dealing with the aforementioned problems simultaneously. The proposed method involves a feature extraction technique that decomposes the vibration signals with Empirical Mode Decomposition (EMD) and Principal Component Analysis. Then Spectrograms and Images Moments are used to extract characteristics that allow the discrimination between different bearing faults by using a classification algorithm. This approach provides a classification of damaged bearings for different and variable setups in stationary and non-stationary conditions. Due to its characteristics, the method can be used to obtain a single control model suitable to monitor various bearings that operating under different conditions.

This paper is organized as follows: Section II presents the theoretical principles used for this work, in Section III the used data is described, the methodology is presented in Section IV, Section V describes the experiments performed and presents the results obtained in the diverse operating conditions addressed. Finally, conclusions are drawn in Section VI.

II. THEORETICAL PRINCIPLES

A brief overview of the theoretical principles of the techniques used in this paper, are now given.

A. Vibration signals generated by bearing defects

The vibration response of a defective bearing consists of a series of impulses that are generated every time a running roller passes over the surface of the bearing damages. The amplitude of the impulses are a measure of the intensity of the shock. Their size depends on the speed, the spatial extent of the damage as well as on the load conditions on the bearing. The frequency at which these impulses are produced is called the bearing characteristic frequency (BCF), which depends on the shaft speed, the geometry of the bearing, and the site of defects. Generally, four types of BCFs are encountered for ball bearings: bearing pass frequency of outer race, bearing pass frequency of inner race, ball spin frequency, and the fundamental train frequency, which correspond to defects on the outer race, the inner race, the roller, and the cage, respectively.

The amplitude of these BCFs is characteristically a sign of defect severity, and the presence of harmonics of the BCFs is an indication of the defect origin.

B. Empirical Mode Decomposition

Empirical Mode Decomposition [15] is based on the direct detection of local signal extrema at a variety of intrinsic time scales to decompose the signal. The resulting Intrinsic Mode Functions (IMFs) are considered as the most important characteristics of the signal, and since the decomposition is based on localized timescales, is readily applicable to nonlinear and nonstationary signals.

An IMF of a signal from EMD satisfies the following two conditions: (1) the number of extrema and the number of zero crossings are equal, or their difference is no more than 1, and (2) its local mean is zero. Specifically, given a signal *x(t)*, the constituent IMFs, $c_i(t)$, can be obtained and summed such that:

$$x(t) = \sum_{i=1}^{K} c_i(t) + r_K(t) \qquad (1)$$

where K is the number of IMFs, $r_K(t)$ the final residue, which is the mean trend of the signal, and $c_i(t)$ represents the IMFs that are nearly orthogonal to each other and whose mean is close to 0.

C. Principal Components Analysis

Principal Components Analysis [16] is a feature extraction method that reduces the dimensionality J of data $\boldsymbol{X}_{IxJ}$, where I is the number of observations and J the number of features, with a minimum loss of information by projecting the data into a lower dimensional subspace, which contains most of the variance of the original data. PCA thus represents data as the product of the mutually orthogonal data called scores $\boldsymbol{T}_{Ixn} = [t_1, t_2, \dots, t_n]$ and transposed linear transformation matrix $\boldsymbol{P}_{jxn} = [p_1, p_2, \dots, p_n]$ also called the principal component matrix, as shown in Eq.(2):

$$\boldsymbol{X} = \boldsymbol{T}\boldsymbol{P}^T = \sum_{j=1}^{n} t_j {p_j}^T \qquad (2)$$

with $n < J$.

To create the best subspace for our signals, it is crucial ot determine the appropriate number of principal components n to select. Due to the importance of this aspect, many studies have been presented in literature with different techniques proposed. For this study, the selection of the number of principal components has been performed by a comparative test. The procedure is based on the evaluation of classification accuracy provided by each different number of principal components in order to select the number that provides the best performance.

D. Spectral Analysis

Spectrograms are a visual representation of the spectrum of frequencies in a sound or other signal as they vary with time. They provide a way to recognize fault patterns in the time-frequency domain characteristics. In this work, spectral analysis is employed to create images that characterize the signals. A spectrogram is obtained by calculating the short time Fourier transform (STFT) [17] of the signal:

$$STFT\{x(t)\}(\tau,\omega) = \int_{-\infty}^{\infty} x(t)w(t-\tau)e^{-j\omega t}dt \qquad (3)$$

where x(t) is the original signal, and w(t) is the window function, for this work a Hamming window is used, and then:

$$spectrogram(t,\omega) = |STFT(t,\omega)|^2 \qquad (4)$$

The spectrogram provide an image where the darkest colors represent higher amplitude of energy and show the relation between the signal in time domain and frequency domain.

E. Images Moments

In image processing, moments are scalar quantities used to characterize an image and to capture its significant features. Given an function $I(x,y)$, that represent the pixel intensity of an image, if this function is piecewise continuous and has nonzero values only in a finite region of the (x,y) plane, then the moments sequence M_{ij} is uniquely determined by $I(x,y)$. By considering that an image segment has finite area, or at least is piecewise continuous, moments of all orders exist and a complete moment set can be computed and used uniquely to describe the information contained in the image.

In the case of a greyscale image, the image moments M_{ij} are calculated by:

$$M_{ij} = \sum_x \sum_y x^i y^j I(x,y) \tag{5}$$

where x and y are the pixel abscissa and ordinate. The central moments μ_{ij} are invariants with respect to translation and are defined as:

$$\mu_{ij} = \sum_x \sum_y I(x,y)(x-\bar{x})^i (y-\bar{y})^j \tag{6}$$

where $\bar{x} = \frac{M_{10}}{M_{00}}$ and $\bar{y} = \frac{M_{01}}{M_{00}}$ are the components of the centroid.

The 10 central moments up to order 3 are used to calculate the 7 Hu moments [18], that are invariant to translation, changes in scale, and rotating. For this work, all the 10 central moments and the 7 Hu moments are used as features set for the images.

F. Support Vector Machines

In order to obtain a model able to classify the faulty bearings, a Support Vector Machine (SVM) approach has been considered [19]. SVMs are binary classifiers that discriminate whether an input vector x belongs to class +1 or to class -1 based on the following discriminant function:

$$f(x) = \sum_{i=1}^{N} \alpha_i d_i K(x, x_i) + b \tag{7}$$

where $d_i \in \{-1,+1\}$, $0 < \alpha_i < C$ and $\sum_{i=1}^{N} \alpha_i d_i = 0$. The terms x_i are the support vectors and b is a bias term that together with the α_i are determined during the training process of the SVM. The input vector x is classified as +1 if $f(x) \geq 0$ and -1 if $f(x) < 0$. The kernel function $K(\cdot,\cdot)$ can assume different forms. The radial basis function has been chosen as kernel, that is given by $\mathrm{K(u,v)} = \exp(-\gamma \|u-v\|^2)$. The ɣ and C parameters, and the chosen kernel, are crucial to obtain the best performance. The dataset has been tested using the C and ɣ parameters, performing the grid search approach, within the following ranges: $C = \{2^{-5}, 2^{-4}, \ldots, 2^{5}\}$ and $\gamma = \{2^{-5}, 2^{-4}, \ldots, 2^{3}\}$. All the experiments have been performed using LibSVM (a library for Support Vector Machines) [20] that provide a multiclass classification used in this work. A total of 4 classes are used, each one associated with a different bearing used.

III. BEARING FAULTS SIMULATOR

The data used for this study was obtained from a Bearing Faults Simulator (Figure 1), a laboratory device produced by GUNT that mimics the behavior of a standard rotary machine. An electric motor transmits rotation via a shaft to the bearing located in its housing. A micrometer screw generates a radial force that acts on the movably mounted housing with the stationary outer bearing race, thereby exerting a radial load on the bearing. The vibration sensor is a piezoelectric accelerometer with integral electronics mounted on the vertical axis that transforms vibrations into electrical signals collected by the control unit, then transferred into the PC.

Three bearings are used for this work, each one characterized by a fault associates to a single BCF: bearing with fault on the outer race (BOF), bearing with fault on the inner race (BIF) and bearing with fault on the rolling element (BRF). A fourth bearing with no faults (BNF) is used as reference.

Figure 1. Bearing Faults Simulator

Rotating machinery usually works under different loads and speeds, for this reason the presented technique is aimed to be suitable under stationary and non-stationary conditions. Hence

TABLE 1. DESCRIPTION OF BEARING FAULTS SIMULATOR SETUPS

Stationary State	Parameters	
	Rotating Speed (rpm)	*Load (mm)*
Setup 1	1700	4
Setup 2	1700	6
Setup 3	3000	4
Setup 4	3000	6
Nonstationary State	**Parameters**	
	Rotating Speed (rpm)	*Load*
Setup 1	1700	Variable Load from 4 to 6 and from 4 to 6
Setup 2	3000	Variable Load from 4 to 6 and from 4 to 6
Setup 3	Variable Speed from 1700 to 3000 and from 3000 to 1700	4
Setup 4	Variable Speed from 1700 to 3000 and from 3000 to 1700	6

two rotating speeds and two loads are used to create 4 different simulator setups with which demonstrate the effectiveness of the proposed method. A first set of signals was recorded under constant conditions, in order to simulate stationary or quasi-stationary behavior. To simulate a more real situation that involves variable speed and load, a second acquisition was carried out by using simulator setups with the same speeds and loads of the previous case, but by varying one of the two parameters during the acquisition. All acquired signals have duration of 8 seconds and they were acquired with a sample frequency of 32 KHz.

The characteristics of the used setups are listed in TABLE 1, for both the stationary and non-stationary case. The loads 4 and 6 mm, corresponding to a force applied on the cage of approximately 5 and 32 N respectively.

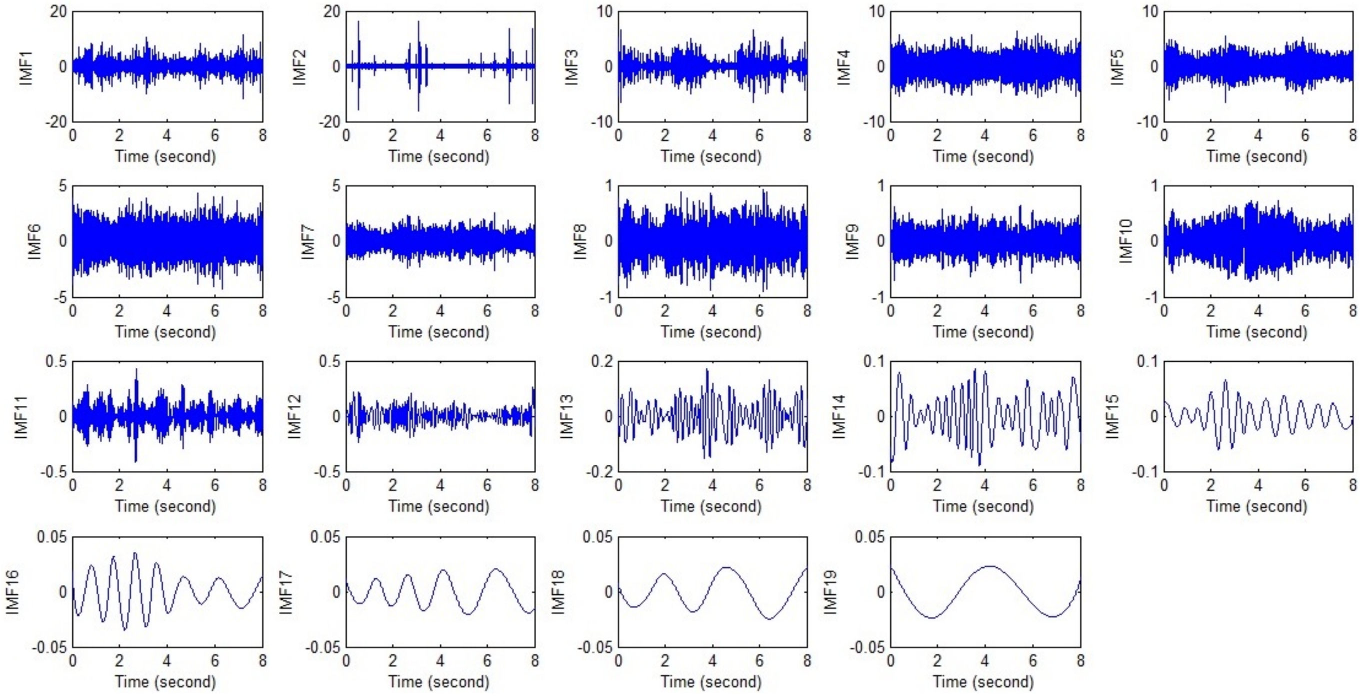

Figure 2. Example of IMFs of bearing vibration signal

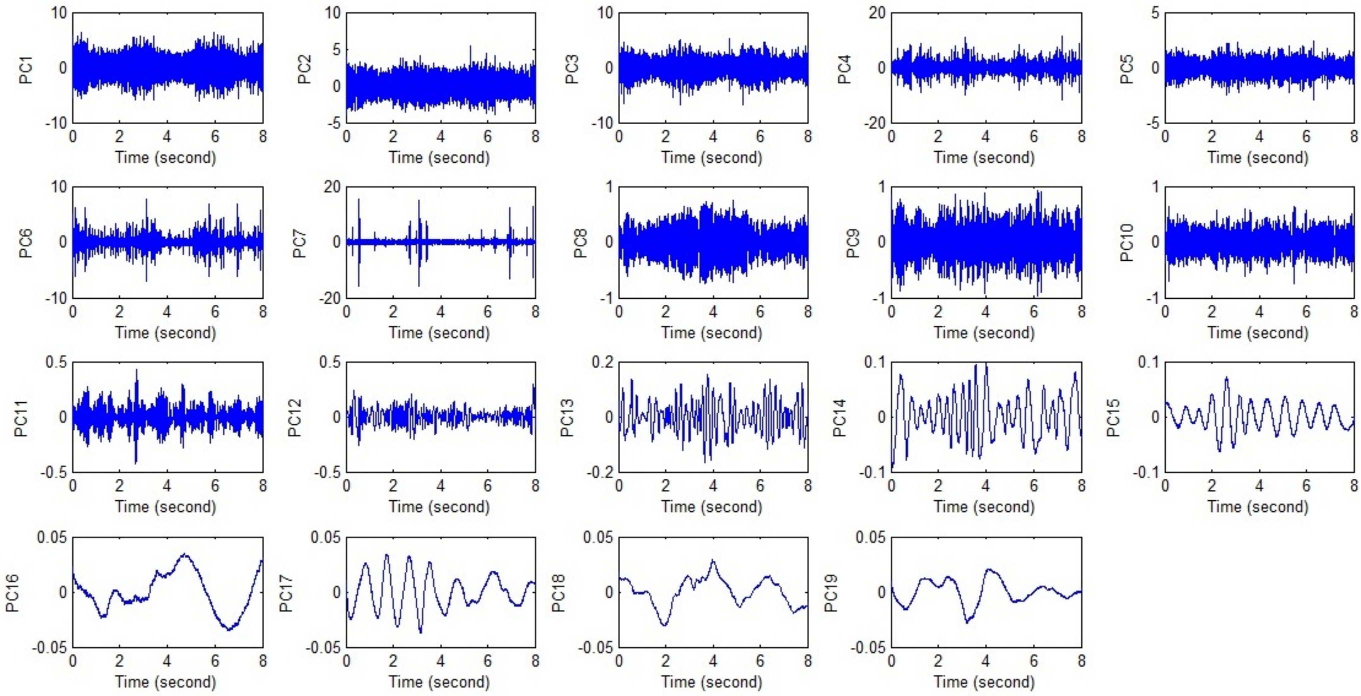

Figure 3. PCs of the IMFs after performing EMD on the original signal

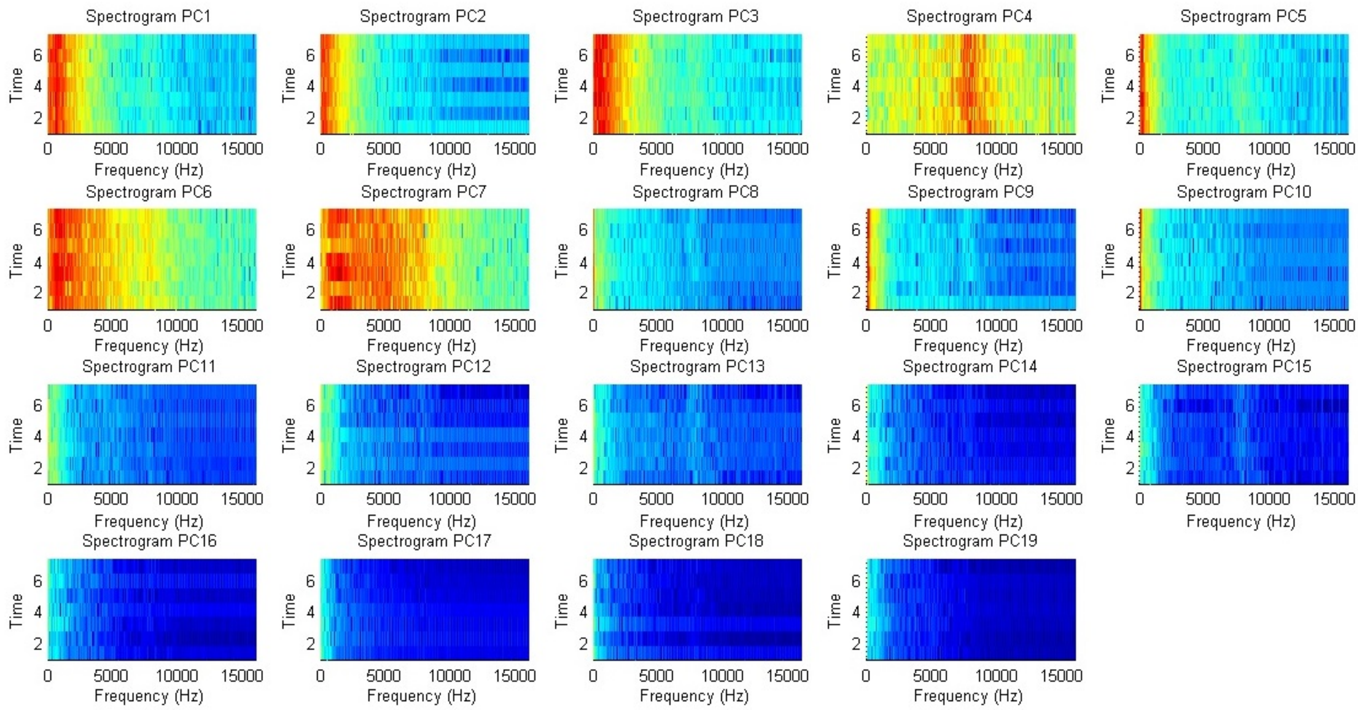

Figure 4. Spectrograms of the PCs

IV. PROPOSED TECHNIQUE

The proposed EMD–PCA method combines EMD, which adaptively decomposes signals into various time scales, with PCA, which provides a mechanism to extract useful information from a signal by reducing the number of involved variables. The EMD technique is suitable for analyzing non-stationary signals recorded from non-linear systems, which is often the case of real systems that use bearings. Furthermore due to its adaptive empirical nature appropriate for the processing of signals exhibiting non-linear characteristics.

In the first step of the presented technique, the vibration signal is decomposed into the IMFs through the EMD algorithm. For example, Figure 2 shows the extracted IMFs for a signal decomposition case; the residual is not shown because it will not be considered in the next step. All IMFs, less the residual, are arranged in a matrix so that each column corresponds with an IMF, and each row represents a different time instant. This matrix is used for the PCA in the following step.

In the second step, PCA gets a new representation of the IMF matrix in a new space created by the Principal Components (PCs). Figure 3 shows the principal components obtained from the IMF from Figure 2.

Each PC is taken as a standalone signal to obtain a representation in the time-frequency domain through the spectrogram. The spectrogram is saved and used as an image to characterize the PC. By calculating the spectrograms of all the PCs, a series of images that characterize the time-frequency relationship of the original signal are obtained; shown in Figure 4. A selection is performed to keep only those images that contain information useful to differentiate the signals exhibiting different faults. The choice of the number of images, thus the number of PCs, used to characterize a signal has been made by evaluating the classification accuracy provided by the classification algorithm that will be introduced later. The algorithm needs the same number of features for each observation, for this reason to find the best number of PCs, the accuracy is evaluated on the overall dataset, and for each signal the same number of images has been chosen.

As expressed earlier, the behavior of a bearing is dependent on many factors, some under the control of the operator; such as the rotating speed and the load, while others are due to internal and external sources to the machine independent of human control. A monitoring system that must provide information about the state of a device must be robust to minimal changes of the system and take into account all factors that may affect the system, including those that are difficult to control. The use of a machine learning algorithm is suitable for a black box approach to create a model that takes into account all the factors that influence the phenomenon, both controllable and uncontrollable.

In order to use the algorithm based on SVM to classify faulty bearings, it is necessary convert the images into a series of values that can be handled by the algorithm. To obtain this new representation, the selected images are converted into greyscale images in order to extract their image moments; 17 for each image. The moments related to the images of a single signal are gathered together to form the feature vector associated with that signal.

V. EXPERIMENTAL TRIALS

To demonstrate the efficacy of the proposed method for identifying different faulty bearings for stationary and nonstationary conditions, two typologies of tests are presented.

The first test shows the accuracy obtained with a dataset composed by stationary signals only. The second tests the possibility to use a model trained on stationary data to classify signals collected in nonstationary conditions.

A. Bearing faults classification with stationary signals

All the observations related to the stationary signals are gathered together to form a unique dataset. The evaluation of the accuracy has been performed using a Cross Validation (CV) K-Folds technique with 6 non-overlapping folds. At each CV iteration, a different combination of Training and Test sets is selected from the dataset. The Validation set is extracted from the selected Train set and used to identify the best values for the parameters C and ɣ, necessary to train the SVM model. The final performance is evaluated in terms of mean accuracy and standard deviation of the Test set.

The accuracy of a model trained with a machine learning algorithm is strongly influenced by the data used for training, because the data must be able to characterize the phenomenon under observation, providing examples of all operating conditions that must be monitored. For this reason, an important parameter to take into account in order to obtain an accurate classification, is the number of observations associated with each class. Therefore, various tests were carried out with an increasing number of observations to see how the accuracy varies and then select the appropriate number of observations.

Along with the study on number of observations, the evaluation of the number of PCs to take into account to obtain the best classification accuracy has been performed.

The graph in Figure 5 shows the evolution of the average accuracy obtained on the test set with CV as a function of the number of observations that each class contains as well as the number of PCs.

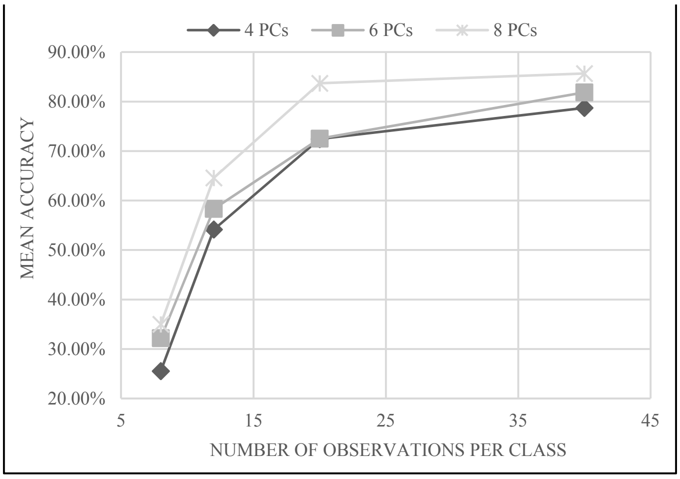

Figure 5. Accuracy evolution in function of the number of the observations and the numer of PCs

It is possible to see that with 8 PCs and 40 observation for each class, so 160 observations for the dataset, a mean accuracy of 85% is obtained for the test set. As referred in the Paragraph IV, 17 moments are extracted from each image so the selection of 8 PCs produces a features vector with 136 elements. TABLE 2 shows in detail the mean accuracies and the corresponding standard deviations for each bearing obtained for the different number of observations per setup with 8 PCs.

TABLE 2. CLASSIFICATION ACCURACY ON TEST SET FOR EACH BAERINGS UNDER STATIONARY CONDITIONS

n^a	BNF	BOF	BIF	BRF	Mean Test Set Accuracy	Standard Deviation
8	37.50%	50.00%	37.50%	12.50%	34.38%	23.83
12	58.33%	50.00%	75.00%	75.00%	64.58%	21.53
20	85.00%	90.00%	85.00%	75.00%	83.75%	10.14
40	87.50%	90.00%	80.00%	85.00%	85.63%	9.99

a. *n* is the number of observations per class in the Dataset.

The detailed results provide a complete view on the performance of the method for each bearing and show how increasing the number of observations per class lead to an increase of accuracy for each bearing. It is observable a common trend on the single bearing results, with an accuracy improvement at each observation increment. Only one case for BIF shows an opposite trend, an accuracy decrease in the step from 20 to 40 observation. This behavior is due to the specific added observations in the last increment that introduced some information that lead to an incorrect classification for this fault.

In parallel with the overall increase of the mean accuracy, the standard deviation decreases, confirming the best performance is obtained with 40 observations per class.

B. Bearing faults classification with nonstationary signals

The proposed bearing faults classification technique is also suitable for providing accurate classification for signals recorded under variable rotating speed and load.

In this case, the classifiers trained with the stationary datasets introduced in the previous test are used to classify the signals measured under nonstationary conditions. Due to the best performance provided in the previous test, 8 PCs are maintained in this test to characterize the signals. For the training phase, the 4 datasets with a different number of observations for each class, respectively 8, 12, 20 and 40 observations, are used. The test set is composed by all those signals measured under variable speed and load, in this case 16 observations are chosen for each class. This test does not involve a CV procedure since the train and test sets are fixed, so the performance is evaluated only in terms of accuracy on the test set, and no standard deviation values are reported as in the stationary case study. TABLE 3 details the results obtained for each bearing.

TABLE 3. CLASSIFICATION ACCURACY ON TEST SET FOR EACH BAERINGS UNDER NON-STATIONARY CONDITIONS

n^b	BNF	BOF	BIF	BRF	Test Set Accuracy
8	50.00%	81.25%	56.25%	37.50%	56.25%
12	87.50%	0.00%	93.75%	56.25%	59.38%
20	68.75%	68.75%	93.75%	25.00%	64.06%
40	75.00%	93.75%	87.50%	31.25%	71.88%

b. *n* is the number of observations per class in the Train Set.

This test confirms the good performance of the proposed technique for the classification of faulty bearings. Obviously, the classification performance using nonstationary data for the test is inferior to that using the stationary data for both train and test, but the dataset with the highest number of observation per class in the train set provide however an accuracy on test set of 71%.

By looking in detail the results for each bearing, it is not possible observe a common trend among the results. For each characteristic bearing fault the accuracy does not follow the increase of the number of observations. This behavior is different from the stationary case where a common trend is visible. The main reason for this difference is with the non-stationary case the CV is not used and therefore the classification accuracies are not averaged. The lack of average causes the fluctuation of the single accuracies but at the same time it is possible to see that the overall accuracy shows an improvement at each increment of observations. It is also important to note that for BNF, BOF and BIF the performance are comparable with those obtained with the test on stationary data. For the BRF instead, the method provides an accuracy remarkably low and this accuracy drags down the overall accuracy. The changing of the operating conditions affects the frequency response of bearings and in the BRF case this leads the classification algorithm to misclassify the fault.

VI. CONCLUSION

In this paper, we present a new technique based on images analysis of the spectrograms that use Empirical Mode Decomposition and Principal Component Analysis to create a time-frequency representation of the vibration signals. The image moments are extracted from the spectrograms and used as features for the classification with an algorithm based on Support Vector Machine. The presented method is suitable for multiple rotating speeds and loads conditions as well as stationary and non-stationary signals. The developed method has provided good results, with accuracies of 85.63% and 71.88%, for stationary and nonstationary respectively with constant and variable parameters.

The presented technique is the first step for the development of a complete methodology for the diagnostic and prognostic of the roller bearings. In particular, the technique will be tested on signals collected from real industrial rotary machine to verify the accuracy obtained with data relative to a standard work cycle of the machine. Further studies will be conducted to test different methods for the feature extraction from the images as well as different algorithms for classification.

ACKNOWLEDGMENTS

The author would like to thank Jarek Grebenik who helped me with the review of this paper, we discussed many aspects of the project together and his points of view were really useful; and Pierre Francq who was helpful analyzing the data and also looked into optimization of the system.

REFERENCES

[1] O. Pat, "Report of Large Motor Reliability Survey of Industrial and Commercial Installations, Part I," *IEEE Transactions on Industry Applications,* Vols. IA-21, no. 4, pp. 853-864, July/August 1985.

[2] V. T. Olav and D. Magnus, "A Survey of Faults on Induction Motors in Offshore Oil Industry, Petrochemical Industry, Gas Terminals, and Oil Refineries," *IEEE Transactions on Industry Applications,* vol. 31, no. 5, pp. 1186-1196, September/October 1995.

[3] V. T. Olav and D. Magnus, "Failure Identification and Analysis for High-Voltage Induction Motors in the Petrochemical Industry," *IEEE Transactions on Industry Aapplications,* vol. 35, no. 4, pp. 810-818, July/August 1999.

[4] R. Akhand and U. S.H., "A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings," *Tribology International,* vol. 96, pp. 289-306, 2016.

[5] W. Dong, T. Peter W. and T. Kwok Leung, "An enhaced Kurtogram method for fault diagonsis of rolling elements bearings," *Mech Syst Signal Process,* vol. 35, pp. 176-199, February 2013.

[6] Y. Wang and M. Liang, "An adaptive SK technique and its application for fault detection of rolling element bearings," *Mechanical Systems and Signal Processing,* vol. 25, no. 5, p. 1750–1764, 2011.

[7] L. Hongkun, X. Fujian, L. Hongyi and Z. Xuefeng, "Incipient fault information determination for rolling element bearing based on synchronous averaging reassigned wavelet scalogram," *Measurement,* vol. 65, pp. 1-10, April 2015.

[8] W. Peter and D. Wang, "The design of a new sparsogram for fast bearing fault diagnosis: Part 2 of the two related manuscripts that have a joint title as "two automatic vibration-based fault diagnostic methods using the novel sparsity measurement - Parts 1 and 2"," *Mech Syst Signal Process,* vol. 40, no. 2, pp. 499-519, 2013.

[9] P. Choon-Su, C. Young-Chul and K. Yang-Hann, "Early fault detection in automotive ball bearings using the minimum variance cepstrum," *Mechanical Systems and Signal Processing,* vol. 38, p. 534–548, July 2013.

[10] S. Yuh-Tay, "An envelope analysis based on the resonance modes of the mechanical system for the bearing defect diagnosis," *Measurement,* vol. 43, p. 912–934, August 2010.

[11] Z. Matej, Z. Samo and P. Ivan, "Multivariate and multiscale monitoring of large-size low-speed bearings using Ensemble Empirical Mode Decomposition method combined with Principal Component Analysis," *Mechanical Systems and Signal Processing,* vol. 24, p. 1049–1067, 2010.

[12] L. Wei, Q. Mingquan, Z. Zhencai, W. Bo and Z. Gongbo, "Bearing fault diagnosis based on spectrum images of vibration signals," *Measurement Science and Technology,* vol. 27, no. 3, February 2016.

[13] H. Liang, Q. Yujian, G. Juping, C. Ling, Z. Xinsong and Z. Hairong, "Mechanical Fault Diagnosis Using Color Image Recognition of Vibration Spectrogram Based on Quaternion Invariable Moment," *Mathematical Problems in Engineering,* vol. 2015, 2015.

[14] K. Renata, M. Eyal, R. Eduard and W. Itai, "Bearing diagnostics using image processing methods," *Mechanical Systems and Signal Processing,* vol. 45, p. 105–113, 2014.

[15] H. Norden E., S. Zheng, L. Steven R., W. Manli C., S. Hsing H., Z. Quanan, Y. Nai-Chyuan, T. Chi Chao and L. Henry H., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences,* vol. 454, no. 1971, pp. 903-995, 1998.

[16] H. Harold, "Analysis of a complex of statistical variables into principal components," *Journal of Educational Psychology,* vol. 24, pp. 417-441, 1933.

[17] O. Alan V. and S. Ronald W., in *Discrete - time signal processing*, Englewood Cliffs, Prentice-Hall, 1989, pp. 713-718.

[18] H. Ming-Kuei, "Visual Pattern Recognition by Moment Invariants," *IRE Transactions on Information Theory,* vol. 8, pp. 179 - 187 , February 1962.

[19] S. Haykin, "Support Vector Machines," in *Neural Networks*, 2005, pp. 341-350.

[20] C. Chih-Chung and L. Chih-Jen, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology,* vol. 2, pp. 1-27, 2011.