Damage identification and external effects removal for roller bearing diagnostics

Pirra M.[1], Fasana A.[2], Garibaldi L.[3], and Marchesiello S.[4]

[1,2,3,4] *Dynamics & Identification Research Group, DIMEAS, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy*
miriam.pirra@polito.it
luigi.garibaldi@polito.it

ABSTRACT

In this paper we introduce a method to identify if a bearing is damaged by removing the effects of speed and load. In fact, such conditions influence vibration data during acquisitions in rotating machinery and may lead to biased results when diagnostic techniques are applied. This method combines Empirical Mode Decomposition (EMD) and Support Vector Machine classification method. The vibration signal acquired is decomposed into a finite number of Intrinsic Mode Functions (IMFs) and their energy is evaluated. These features are then used to train a particular type of SVM, namely One-Class Support Vector Machine (OCSVM), where only one class of data is known. Data acquisition is done both for a healthy bearing and for one whose rolling element presents a 450 μm damage. We consider three speeds and three different radial loads for both bearings, so nine conditions are acquired for each type of bearing overall. Feature evaluation is done using EMD and then healthy data belonging to the various conditions are taken into account to train the OCSVM. The remaining data are analysed by the classifier as test object. The real class each element belongs to is known, so the efficiency of the method can be measured by counting the errors made by the labelling procedure. These evaluations are performed by applying different kinds of SVM kernel.

1. INTRODUCTION

Rolling bearings are among the most widely used components in machinery. Their condition monitoring and fault diagnosis are then very important in order to prevent the occurrence of breakdowns. A wide range of different methods has been proposed since the Seventies to get proper fault diagnosis techniques. Signal analysis is an important topic in mechanical fault diagnosis research and applications thanks to its ability to extract the fault features and identify the fault patterns. Methods such as Fourier analysis and time-domain analysis take into account the acquired signal and are based on the assumption that the process generating the signal itself is stationary and linear. Unluckily, the faults are time localised transient events, so this kind of techniques could provide a wrong information.

Some possible ways to overcame these aspects are presented in Randall and Antoni (2011). They develop an interesting review of diagnostic analysis of acceleration signals from rolling element bearings, especially when a strong masking noise is present due to other machine components such as gears. They show industrial applications that confirm the reliability of their methods. Another interesting method that could be efficiently used in the vibration-based condition monitoring of rotating machines is presented in Antoni (2006). He shows how the Spectral Kurtosis (SK), in contrast to classical kurtosis analysis, provides a robust way of detecting incipient faults even in the presence of strong masking noise. The other appealing aspect is that it allows to design optimal filters efficiently to filter out the mechanical signature of faults.

A useful tool to analyse non-stationary signals such as those related to bearing vibrations is wavelet transform. Its strength comes from the simultaneous interpretation of the signal in both time and frequency domain that allows local, transient or intermittent components to be exposed. As drawback there is the dependence on the choice of the wavelet basis function. An example of wavelet-based analysis technique for the diagnosis of faults in rotating machinery from its vibrating signature is Chebil, Noel, Mesbah, and Deriche (2009).

An innovative technique in the time–frequency domain is the Empirical Mode Decomposition (EMD) (Huang et al., 1998). It allows any complicated signal to be decomposed into a collection of Intrinsic Mode Functions (IMFs) based on the local characteristic time scale of the signal. It is self-adaptive because the IMFs, working as the basis functions, are determined by the signal itself rather than being pre-determined. Hence, EMD is highly efficient in non-stationary data analysis. It has been applied to a wide variety of problems, going from geophysics to structural health monitoring (Huang & Shen, 2005). Lots of authors apply EMD to rotating machines and bearings with diagnostic intents, usually in association with other techniques. Some examples are Gao, Duan, Fan, and Meng (2008), where combined mode functions are introduced, Junsheng, Deije, and Yu (2006), that use EMD jointly with an AutoRegressive model and Yu, Deije, and Junsheng (2006), that train an Artificial Neural Network (ANN) classifier with the EMD energy entropies.

Another worth of interest aspect is the search for methods able to remove effects produced in vibrations by external factors, such as environmental temperature or test rig assemblies. Some examples are presented in Pirra, Gandino, Torri, Garibaldi, and Machorro-López (2011) and in Machorro-López, Bellino, Garibaldi, and Adams (2011), where the multi-variate statistical technique named Principal Component Analysis (PCA) is used successfully in bearing fault detection and rotating shaft. Other factors influencing vibrations related to rotating elements are varying load and speed. In fact a variation in these factors produces some difficulties in recognising the presence of fault in a signal. Bartelmus and Zimroz (2009) show how in condition monitoring of planetary gearboxes is important to identify the external varying load condition. In particular, they analyse in detail how many factors influence the vibration signals generated by a system in which a planetary gearbox is included and show how the load has a consistent contribution. As far as bearings are concerned, instead, some works are presented in Cocconcelli, Rubini, Zimroz, and Bartelmus (2011) and Cocconcelli and Rubini (2011). They inspect the continuous change of rotational speed of the motor, that represent a substantial drawback in terms of diagnostics of the ball bearing. In fact, the large part of algorithms proposed in the literature needs a constant rotation frequency of the motor to identify fault frequencies in the spectrum. They tackle the problem with encouraging results aided by ANN and Support Vector Machine (SVM).

These two last techniques could be grouped under the terms of *soft* or *natural* computing. They are well developed in Worden, Staszewski, and Hensman (2011), an exhaustive tutorial overview of their basic theory and their applications in the context of mechanical systems research. SVM in particular, is widely used for condition monitoring and damage classification (Widodo & Yang, 2006), (Rojas & Nandi, 2006). It is based on the concept of separating data objects into different classes through an hyperplane. However, this method assumes that all types of instances are known before applying it. A particular case of SVM is the One-Class SVM (OCSVM), that is well suited for a diagnostic technique purpose. In fact, it allows the creation of the separating hyperplane starting from the knowledge of only one class, that is what usually happens in damage detection. Shin, Eom, and Kim (2005) adopt this method for machine fault detection and classification in electro-mechanical machinery from vibration measurements.

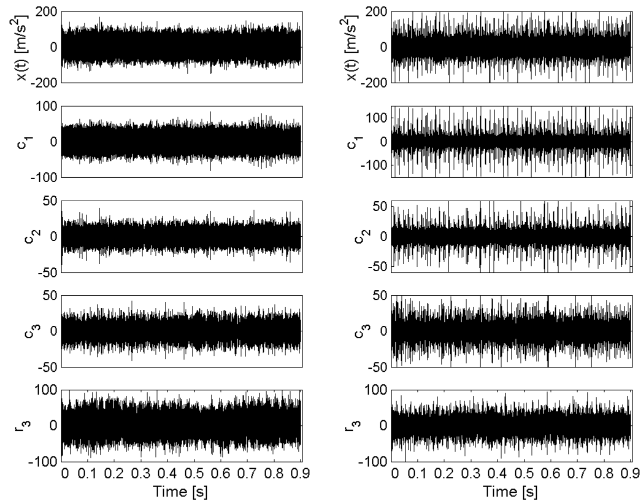

Figure 1. Acceleration signal and its decomposition for a healthy bearing (left) and for a faulty rolling element one (right).

The intent of our work is to find a parameter able to remove the influence of various external conditions in order to detect properly a damage in a roller bearing. This paper is organised as follows. In next two sections EMD method and OCSVM are presented with some theoretical background. Our algorithm is explained in Section 4 and then its application on a test rig is developed in the following session.

2. Empirical Mode Decomposition

Empirical Mode Decomposition is a method presented by Huang et al. (1998) and based on the local characteristic time scales of a signal. This approach could be seen as a self-adaptive signal processing method that can be applied to non-linear and non-stationary process. In particular, it allows a complex signal function to be decomposed into a number of intrinsic mode functions (IMFs). Each one of these components contains frequencies changing with the signal itself and it has to satisfy the following definition:

- In the entire data set, the number of extrema and the number of zero crossings must either be equal or differ at most by one.
- At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

Thanks to this definition, each IMF represents the simple oscillation mode involved in the signal. According to Huang et al. (1998) a *sifting* process is used in order to extract the IMFs from a given signal $x(t)$. It consists of different steps:

1. Identify all the extrema of the signal, and connect all the local maxima by a cubic spline line as the upper envelope. Repeat the same procedure on the local minima to produce the lower envelope.
2. Designate the mean of the two envelopes as m_1, and the difference between the signals $x(t)$ and m_1 as the first component, h_1, i.e.

$$x(t) - m_1 = h_1. \tag{1}$$

   Ideally, if h_1 is an IMF, then take it as the first IMF component of $x(t)$. Otherwise, consider h_1 as the original signal and repeat the first two step obtaining

$$h_1 - m_{11} = h_{11}. \tag{2}$$

   Repeat the sifting process up to k times when h_{1k} becomes an IMF, that is

$$h_{1(k-1)} - m_{1k} = h_{1k}. \tag{3}$$

   The first IMF component is then designated as

$$c_1 = h_{1k}. \tag{4}$$

3. Separate c_1 from the original signal $x(t)$ to obtain the residue r_1:

$$r_1 = x(t) - c_1. \tag{5}$$

4. Consider r_1 as the original signal and repeat the above process n times, obtaining the other IMFs $c_2, c_3, \ldots, c_n$ satisfying

$$\begin{aligned} r_1 - c_2 &= r_2 \\ &\vdots \\ r_{n-1} - c_n &= r_n \end{aligned} \tag{6}$$

5. Stop the decomposition process when r_n becomes a monotonic function from which no more IMFs can be extracted. The sum of Eq. (5) and Eq. (6) gives

$$x(t) = \sum_{i=1}^{n} c_i + r_n. \tag{7}$$

From Eq. (7) we can see how the signal $x(t)$ can be decomposed into n empirical modes and a residue r_n, that could be interpreted as the mean trend of the signal. Each IMF c_i includes different frequency bands ranging from high to low and is stationary.
Figure 1 shows two signals, a healthy and a damaged one. The last one refers to a 450 μm fault on a rolling element. In both cases, the original signal and 3-IMFs decomposition of the signal itself are presented.

3. One-Class Support Vector Machine

Support vector machine (SVM) is a computational learning method developed during the 80s, based on the statistical learning theory (Vapnik, 1982). It is well suited for classification, because given some data points which belong to a certain class it is able to state the class a new data point would

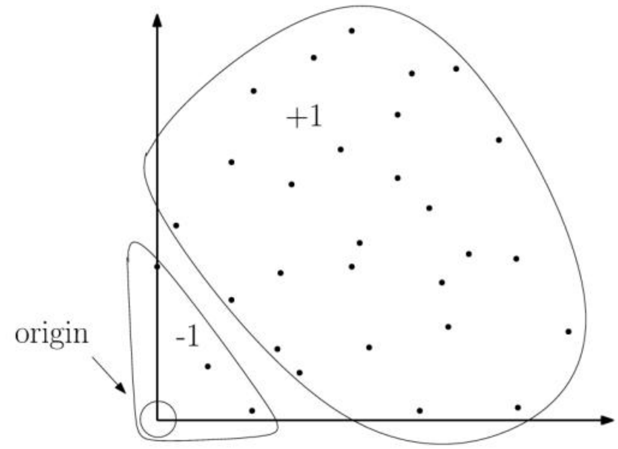

Figure 2. One-Class SVM classifier where the origin is the only member of one class.

be in. If we consider an n-dimensional input data made up of a number of samples belonging to a class, namely positive or negative, SVM constructs a hyperplane that separates the two classes. Moreover, this boundary would satisfy the condition that the distance from the nearest data points in each class is maximal. In this way, an optimal separating hyperplane is created, namely the maximum margin. The points in both classes nearest to this margin are called support vectors and, once selected, they contain all the information necessary to define the classifier. Every time a new element appears, it could be classified according to where it places respect to the separating hyperplane.
SVM could also be applied in case of non-linear classification using a function $\phi(x)$ that maps the data onto a high-dimensional feature space, where the linear classification is then possible. Furthermore, if a kernel function $K(x_i, x_j) = (\phi^T(x_i) \cdot \phi(x_i))$ is applied, it is not necessary to evaluate explicitly $\phi(x_i)$ in the feature space. Various kernel function could be used, such as linear, polynomial or Gaussian RBF. This property enables SVM to be used in case of very large feature spaces because the dimension of classified vectors does not influence directly the SVM performance.
When more than two classes are present, a Multi-class SVM could be adopted. Two different approaches are taken into account: One-against-all (OAA) and One-against-one (OAO). In the first one the i-th SVM is trained with all the examples in the j-th class with positive labels and all the other examples with negative labels, while in the latter one each classifier is trained on data from two classes.
It is clear that in the previous cases, two or more classes of data are given since the beginning of the analysis. In more general diagnostic applications, instead, only one type of data objects is usually acquired: the healthy one. This could be seen as the detection of patterns in data that do not conform to

a well defined notion of normal behaviour, so we could refer to anomaly detection. One-Class SVM is the application of the SVM approach to the general concept of anomaly detection, as presented by Schlkopf et al. in Schlkopf, Williamson, Smola, Taylor, and Platt (2000). In their method they construct a hyper-plane around the data, such that this is maximally distant from the origin and can separate the regions that contain no data. They propose to use a binary function that returns +1 in region containing the data and -1 elsewhere. For a hyper-plane w which separates the data x_i from the origin with maximal margin ρ, the following quadratic program has to be solved:

$$\min_{w\in F,\xi\in\mathbb{R}^n,\rho\in\mathbb{R}} \quad \frac{1}{2}||w||^2 + \frac{1}{\nu n}\sum_i \xi_i - \rho \tag{8}$$

$$\text{subject to} \quad (w \cdot \Phi(x_i)) \geq \rho - \xi_i, \quad \xi_i \geq 0 \tag{9}$$

where ξ represents the slack variable and ν is a variable taking values between 0 and 1 that monitors the effect of outliers (hardness and softness of the boundary around data).
If w and ρ solve the minimisation problem presented in Eq. (8) - (9), the decision function

$$f(x) = \text{sign}((w \cdot \Phi(x_i)) - \rho) \tag{10}$$

is positive for most instances representing the majority of data.
Figure 2 shows graphically the idea presented here, with only few points around the origin that are negatively labelled.

4. Methodology

The previous sections introduced the background and the theoretical aspects of the two methods that now we want to use jointly. The goal of this study is the search for a method able to identify a damage in a rotating element of a roller bearing by removing the effect of external conditions influencing vibrations.
The diagnosis method consists of different steps:

1. Collect vibration signals under various condition of speed and radial load applied, both for a healthy and a damaged bearing.
2. Apply EMD and decompose the original signal into some IMFs; then choose the first n to extract the features used during the analysis.
3. Evaluate the total energy for the n selected IMFs:
$$E_j = \int_{-\infty}^{+\infty} |c_j(t)|^2 dt \qquad j = 1, \ldots, n. \tag{11}$$
4. Create a feature vector with the energies of the n selected IMFs:
$$F = [E_1, \ldots, E_n]. \tag{12}$$
5. Normalise the feature vector dividing F for this value:
$$E_N = \sqrt{\sum_{j=1}^{n} |E_j|^2} \tag{13}$$

(a)

(b)

Figure 3. DIRG test rig (a) and roller bearing used during the tests with the damaged roller in the white circle (b).

6. Obtain the n-dimension normalised feature vector:
$$F' = [E_1/E_N, \ldots, E_n/E_N]. \tag{14}$$
7. Consider 75% of healthy data as training and the remaining 25% as test together with damaged data. All loads and speeds are analysed together.
8. Train the one-class SVM classifier on training data and evaluate the label assigned by the classifier to test data. The real class is known so mistakes in labelling could be computed.
9. Repeat point 7. and 8. 30 times permuting healthy data order to give statistical significance to the analysis and evaluate the error percentage in labels assignment.

5. Application to bearing data

Several conditions can influence data during acquisitions in our test rig analysis: speed, external load, temperature variations. Detecting and removing the effects of these factors is important to avoid any bias during the application of diagnostic techniques. In fact, a small variation in speed or in the temperature of the oil circulating in a system produces deviations that a diagnostic algorithm may erroneously detect as a damage, thus providing a false alarm. In this paper we try to introduce a method able to identify a damage in a rotating element of a roller bearing by removing the effect of speed and external load.
Accelerations are acquired on a test rig assembled by Dynamics & Identification Research Group (DIRG) at Department of Mechanical and Aerospace Engineering (Figure 3 a).

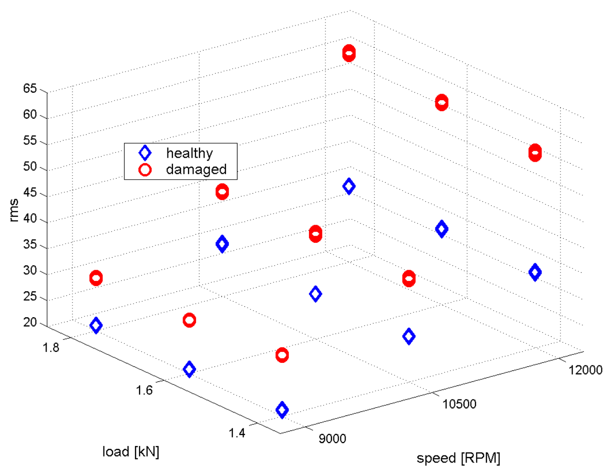

Figure 4. RMS value.

This bearing test rig is designed to perform accurate testing of bearings different levels of damage in a controlled laboratory conditions, especially regarding the minimization of spurious signals coming from the mechanical sounds of other bearings, rotating shafts, gear wheels meshing and other vibrating elements. Hence, we are sure that the only variations in accelerations are given by speed and load that can be properly changed and monitored.

We consider three different speed values (9000, 10500 and 12000 RPM) and three radial loads (1.4, 1.6 and 1.8 $\times$ 10^3 N) and we acquire data for each combination. In particular, 10 acquisitions registering 1 second of vibrating signal at sampling frequency 102.4 kHz are collected for each of the nine cases. This is done both for a healthy bearing and for a damaged one. In the last case, we analyse a bearing with a greater than 450 μm fault on a rolling element (Figure 3 b). Notice that the temperature of the oil circulating is almost constant between the different acquisitions, so we are certain that the only variations detected through vibrations are caused by load and speed changing.

In Figure 4 Root Mean Square values for the 10 acquisitions in each condition are evaluated. This plot shows how this parameter is influenced by the speed both for healthy and damaged case and it increases with higher speeds. Moreover, it can be noticed that in low speed cases this parameter value for damaged bearing is almost near to the healthy one when it reaches the highest speed. For example, RMS value for a damaged bearing at 9000 RPM for the three loads is around 30. If we consider the healthy case at the highest speed evaluated (12000 RPM) RMS is around 34, so it can be noticed that the undamaged bearing at higher speed has a parameter value greater than the faulty one at lower speed. It means that if we consider the RMS parameter taking into account all nine conditions together, the difference between healty and faulty

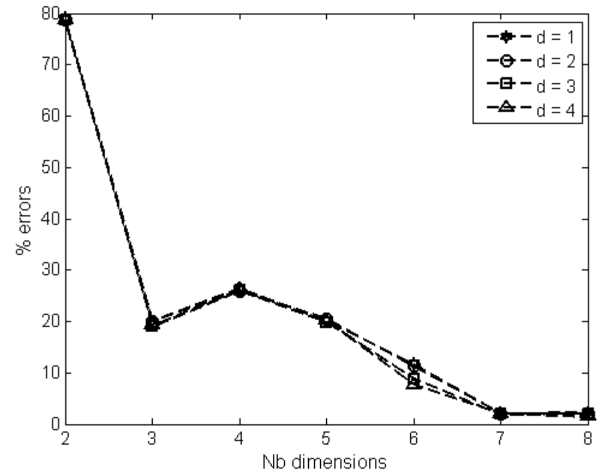

Figure 5. Error percentage for linear kernel.

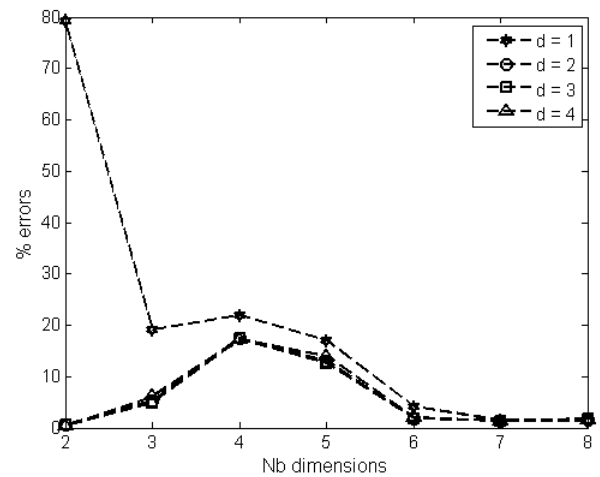

Figure 6. Error percentage for polynomial kernel.

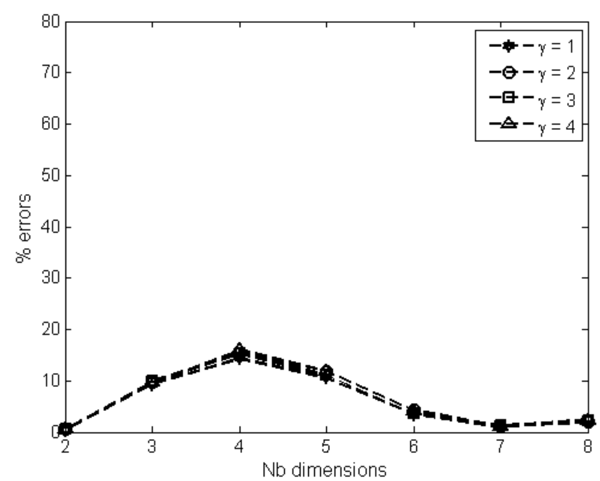

Figure 7. Error percentage for gaussian kernel.

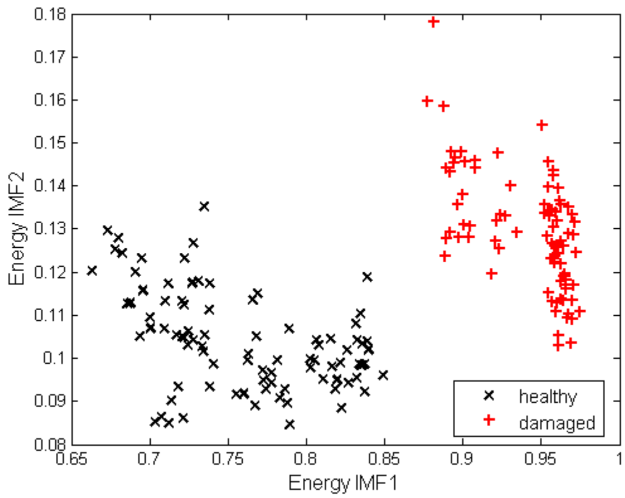

Figure 8. 2-dimension feature vector F' representation.

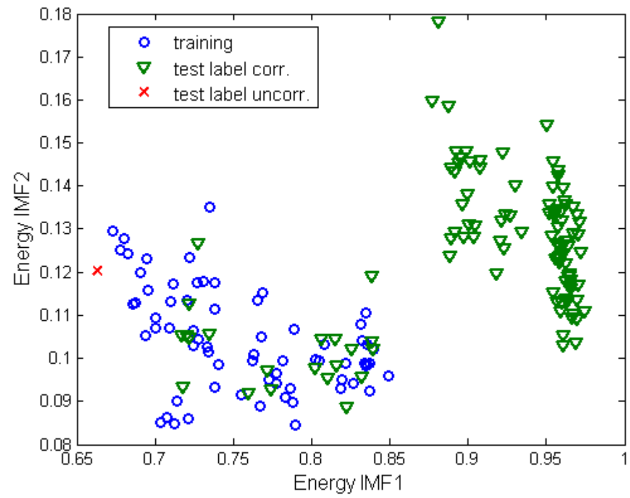

Figure 9. 2-dimension feature vector F' representation after OCSVM.

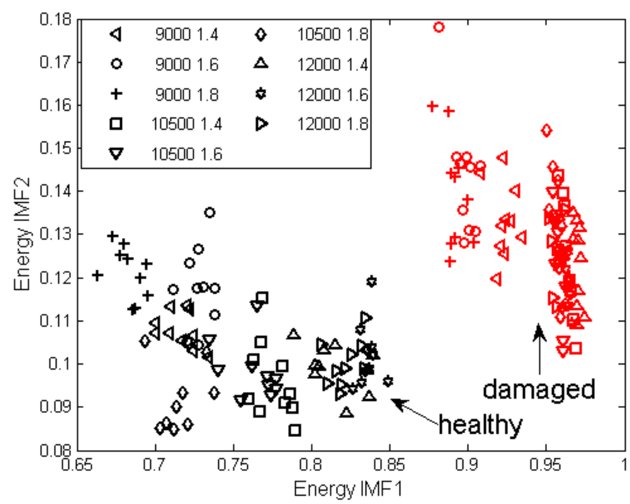

Figure 10. 2-dimension feature vector F' representation with different conditions: the first number is the speed expressed in RPM, the second is the load expressed in kN.

bearings may be strongly biased. This observation leads to the need of a parameter that could avoid such kind of problems.

According to the methodology presented in Section 4, we obtain a normalised feature vector F'. We decide to take into account the first 8 IMFs which include the most dominant fault information, so this vector is in a 8-dimensions space. The analysis through OCSVM is done starting from the first two dimensions of the feature vector. Then we add a new dimension each time until the whole feature vector F' is used. We choose to include the feature from the beginning according to the fact that EMD operates in form of collection of filters organised in a filter bank structure. In particular, the first mode could be considered similar to a highpass filter while the other modes are characterised by a set of overlapping bandpass filters (Flandrin & Rilling, 2004). In such way, taking the feature starting from the beginning of the vector, we move from higher frequency contents to lower ones.

As stated in Section 4, the 75% of healthy data are used to train the classifier, while the 25% of them are added to damaged data as testing instances. Since the exact belonging is known, it is interesting to evaluate the errors in labelling made by the OCSVM classifier. In this way, an evaluation of the relation between the number of dimensions and a proper identification procedure could be done. Moreover, three different SVM kernels are compared through the application to the acquired data:

- linear: $K(x_i, x_j) = (x_i^T x_j)^d$
- polynomial: $K(x_i, x_j) = (x_i^T x_j + 1)^d$
- Gaussian: $K(x_i, x_j) = \exp(-\gamma ||x_i - x_j||^2)$

For each kernel, parameters d and γ take values going from 1 to 4 and labelling mistakes are evaluated in percentage. Figures 5, 6 and 7 present the different behaviours of the three kernels when the number of feature and the parameters values increase. The error percentage for the linear kernel tends to decrease when the dimensions go from 1 to 8. Hence, in order to provide a good detection ability a greater number of features should be considered. The same behaviour is observed for polynomial kernel when $d = 1$, while for the other values of the parameter less errors are present for 2, 6, 7 and 8 dimensions. The error trend in the case of a gaussian kernel does not seem to be conditioned by parameter γ, while the minimum number of labelling errors are found when the feature vector has 2 and 7 dimensions. On the whole, a gaussian kernel or a polynomial one with parameter $d > 1$ give successful results in detecting the damage regardless of speed

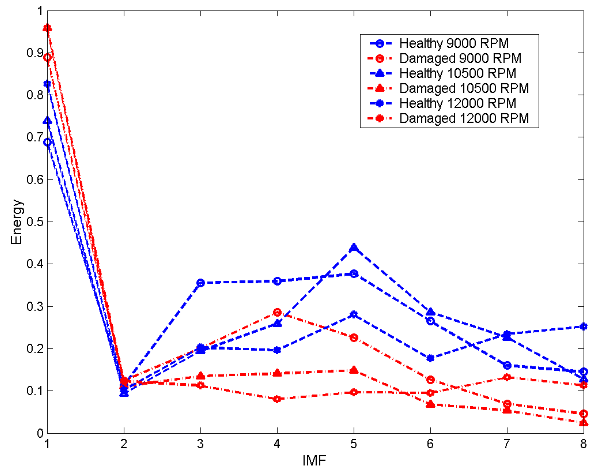

Figure 11. Normalised feature vector F' values for three speeds for both undamaged and damaged case.

and load influence.
To emphasise this fact, we can concentrate on the 2-dimensions feature vector F' thanks to the fact that it gives interesting results and because it is easier to visualise. If we consider the totality of 180 values computed using our methodology for both healthy and damaged bearing, we obtain the plot in Figure 8. In this picture, it is clear how data divide into two groups according to their state rather then depending on their condition of load and speed. This explains the great efficiency of the classifier in damage identification, due to the perfect distinction between the two classes of data. It could be seen in Figure 9 how OCSVM with Gaussian kernel and $\gamma = 1$ works. The testing data are well classified (green triangles) and only one belonging to the faulty class is labelled as healthy producing an error (red cross).
Furthermore, any dependence on different loads and speeds seems to be removed as pointed out in Figure 10. The nine symbols represent the various conditions for the undamaged and damaged bearing and, on the whole, no particular division based on the rotational speed or on the load applied is noticed.
Figure 11 could help to explain the ability of the method in the speed and load influence removal. Values of one acquistion feature vector F' are plotted for each of the speeds considered, both for the healthy and for the damaged bearing. Firstly, the vector normalisation presented at step 5 and 6 in the Methodology Section helps to remove the contribution of highest energies and, so, to mitigate the various conditions influence on the features. Moreover, as it could be noticed in the Figure, this aspect is particular observable for the 'frequency content' represented by c_2. The normalised values of the energies here, in fact, tends to be very similar independently of the speed considered, giving a great contribution in the removal of this parameter influence.

6. Conclusion

In this paper we proposed a method for the detection of damages in roller bearings with the removal of speed and load dependence. This methodology combines Empirical Mode Decomposition, used to produce a proper feature vector, with the One-Class Support Vector Machine technique, exploited to classify the data. Since the original class belonging was known, different SVM-kernels have been tested in order to find those with lower error rate. Encouraging results have been obtained related to the ability of this feature in removing speed and load dependence in order to avoid any bias in data interpretation and identification. Further applications could deal with various damage entity comparisons and with other damage type, such as sandblasted inner ring. Moreover, other factors influence removal, such as temperature, and the comparison of this method with other techniques used to obtain the feature vector, such as wavelet decomposition, could be developed.

References

Antoni, J. (2006). The spectral kurtosis: a useful tool for characterising non-stationary signals. *Mechanical System and Signal Processing*, *20*, 282307.

Bartelmus, W., & Zimroz, R. (2009). Vibration condition monitoring of planetary gearbox under varying external load. *Mechanical Systems and Signal Processing*, *23*, 246257.

Chebil, J., Noel, G., Mesbah, M., & Deriche, M. (2009). Wavelet decomposition for the detection and diagnosis of fault in rolling element bearings. *Jordan Journal of Mechanical and Industrial Engineering*, *3*, 260-267.

Cocconcelli, M., & Rubini. (2011). Support Vector Machines for condition monitoring of bearings in a varying-speed machinery. In *Proceeding International Conference on Condition Monitoring, Cardiff, UK.*

Cocconcelli, M., Rubini, R., Zimroz, R., & Bartelmus, W. (2011). Diagnostics of ball bearings in varying-speed motors by means of Artificial Neural Networks. In *Proceeding International Conference on Condition Monitoring, Cardiff, UK.*

Flandrin, P., & Rilling, G. (2004). Empirical Mode Decomposition as a filter bank. *IEEE Signal Processing Letters*, *11*(2), 112-114.

Gao, Q., Duan, C., Fan, H., & Meng, Q. (2008). Rotating machine fault diagnosis using empirical mode decomposition. *Mechanical System and Signal Processing*, *22*, 1072-1081.

Huang, N. E., & Shen, S. (Eds.). (2005). *Hilbert-Huang Transform and Its Applications*. World Scientific, Singapore.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., et al. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. In *Proceedings of the Royal Society* (p. 903-995).

Junsheng, C., Deije, Y., & Yu, Y. (2006). A fault diagnosis approach for roller bearings based on EMD method anda AR model. *Mechanical System and Signal Processing*, *20*, 350-362.

Machorro-López, J., Bellino, A., Garibaldi, L., & Adams, D. (2011). PCA-based techniques for detecting cracked rotating shafts including the effects of temperature variations. In *Proceeding 6th International Conference on Acoustical and Vibratory Surveillance Methods and Diagnostic Techniques, Compigne, France.*

Pirra, M., Gandino, E., Torri, A., Garibaldi, L., & Machorro-López, J. M. (2011). PCA algorithm for detection, localisation and evolution of damages in gearbox bearings. *Journal of Physics. Conference series*, *305*(1).

Randall, R. B., & Antoni, J. (2011). Rolling element bearing diagnostics - A tutorial. *Mechanical System and Signal Processing*, *25*, 485-520.

Rojas, A., & Nandi, A. B. (2006). Practical scheme for fast detection and classification of rolling-element bearing faults using support vector machines. *Mechanical Systems and Signal Processing*, *20*, 1523-1536.

Schlkopf, B., Williamson, R. C., Smola, A. J., Taylor, J. S., & Platt, J. C. (2000). Support vector method for novelty detection. *Advances in Neural Information Processing Systems*, *12*, 582-586.

Shin, H. J., Eom, D.-H., & Kim, S.-S. (2005). One-class support vector machines - an application in machine fault detection and classification. *Computer & Industrial Engineering*, *48*, 395-408.

Vapnik, V. N. (Ed.). (1982). *Estimation of dependences based on empirical data*. Springer-Verlag, New York.

Widodo, A., & Yang, B. (2006). Support vector machine in machine condition monitoring and fault diagnosis. *Mechanical Systems and Signal Processing*, *21*, 2560-2574.

Worden, K., Staszewski, W. J., & Hensman, J. J. (2011). Natural computing for mechanical systems research: A tutorial overview. *Mechanical Systems and Signal Processing*, *25*, 4-111.

Yu, Y., Deije, Y., & Junsheng, C. (2006). A roller bearing fault diagnosis method based on EMD energy entropy and ANN. *Journal of Sound and Vibration*, *294*, 269-277.