Conference Proceedings Paper – Sensors and Applications

Online condition monitoring of bearings for improved reliability in packaging materials industry

Jovan Gligorijevic[1, 2], Dragoljub Gajic[2, 3, 4, *], Aleksandar Brkovic[1, 2], Ivana Savic-Gajic[5, 6], Olga Georgieva[4], Stefano Di Gennaro[6]

[1] Faculty of Engineering, University of Kragujevac, Serbia
[2] Tetra Pak, Gornji Milanovac Packaging Materials Plant, Serbia
[3] School of Electrical Engineering, University of Belgrade, Serbia
[4] Faculty of Mathematics and Informatics, University of Sofia, Bulgaria
[5] Faculty of Technology, University of Nis, Serbia
[6] Center of Excellence DEWS, University of L'Aquila, Italy

* Author to whom correspondence should be addressed; E-Mail: dragoljubgajic@gmail.com.

Published: 9 November 2015

Abstract: The production processes in the packaging materials industry has to be very efficient and cost-effective. These processes usually take place under extreme conditions and high speeds that requires a high level of reliability and safety. Rollers including their motors and support bearings are the most common components of production machines in the packaging materials industry. Bearing faults, which often occur gradually, represent one of the foremost causes of failures. Therefore detection of their faults in early stage is quite important to assure safe and efficient operation. We present a new automated technique for early fault detection and diagnosis in rolling-element bearings based on vibration signal analysis, wavelet transform and statistical pattern recognition. Accuracy of the technique has been tested on four classes of the recorded vibrations signals, i.e. normal, with the fault of inner race, outer race and balls operation. The overall accuracy of 98.9% has been achieved. The new technique can be used to increase reliability and efficiency in the industry by preventing unexpected faulty operation of bearings.

Keywords: reliability; bearings; fault diagnosis; wavelet transform; statistical pattern recognition.

1. Introduction

In order to further increase its competitiveness the packaging materials industry needs to deploy advanced maintenance strategies and solutions for improved reliability of its production machines as well as safety of its production processes. Reliability-centered maintenance (RCM) is one of such strategies that is usually performed at plant level. As such it plays an increasingly important role in order to achieve a reliable, save and efficient operation of production machines. RCM aims to optimize the entire maintenance program and define optimal measures to be implemented on each of the machines. Manufacturing of paper-based packaging materials usually includes three key production processes which are printing, laminating and slitting, in that order. All these processes take place under extreme conditions and high speeds that requires a high level of reliability and safety. For example, the running speed of a printing machine and a laminator is around 600 m/min while in the case of a slitting machine it reaches 1000 m/min. Apart from a high running speed the laminating process is also characterized by high temperatures since polyethylene is at first melted and then adhered to the packaging material. Rollers together with their supporting bearings and electric motors, as shown in Figures 1 and 2, are the most common components of these production machines. Bearings operate under high loading and severe conditions and their faults represent one of the foremost causes of failures.

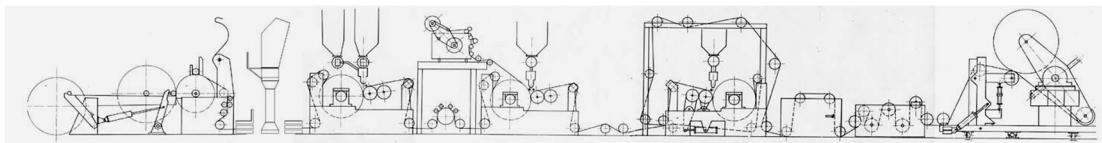**Figure 1.** Roller (left), bearing (middle) and electric motor (right).

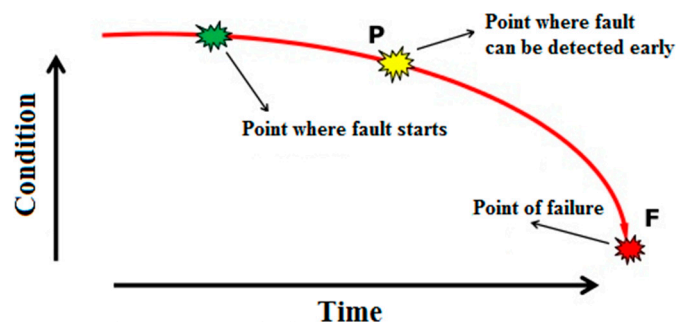**Figure 2.** Scheme of a laminator containing around 50 rollers.

Figure 3. Predictive maintenance Potential to Functional Failure (P-F) curve.

As shown in Figure 3 bearing faults often occur gradually. Defective bearings generate various forces causing high amplitude of vibration. Therefore it is very important to avoid deteriorating condition, degraded reliability and unexpected failures of bearings. Based on an RCM analysis performed on the laminator at Tetra Pak packaging materials plant in Gornji Milanovac, Serbia it has been concluded that for most of bearings a regular check of vibration level every two weeks by a portable, handheld data

collector and analyzer can ensure their reliable and safe operation as well as an early detection of bearing faults. However at a few critical points it has been proposed to do more frequent checks that becomes resource-demanding and not so optimal anymore. Therefore installation of an online vibration monitoring system including deployment of an automated technique for early FDD in bearings. Such a system saves time of maintenance technicians and enables objective, reliable and faster detection and diagnosis of bearing faults, and thus making the entire production process safer. Since vibration sensors and data acquisition systems are already available in the market the Tetra Pak plant has designed an online vibration monitoring system as presented in Figure 4. At the same time the plant also initiated development of a new automated technique for early FDD in bearings to be tested and potentially deployed in this plant once all the equipment is installed. This paper focuses on development of such a technique.
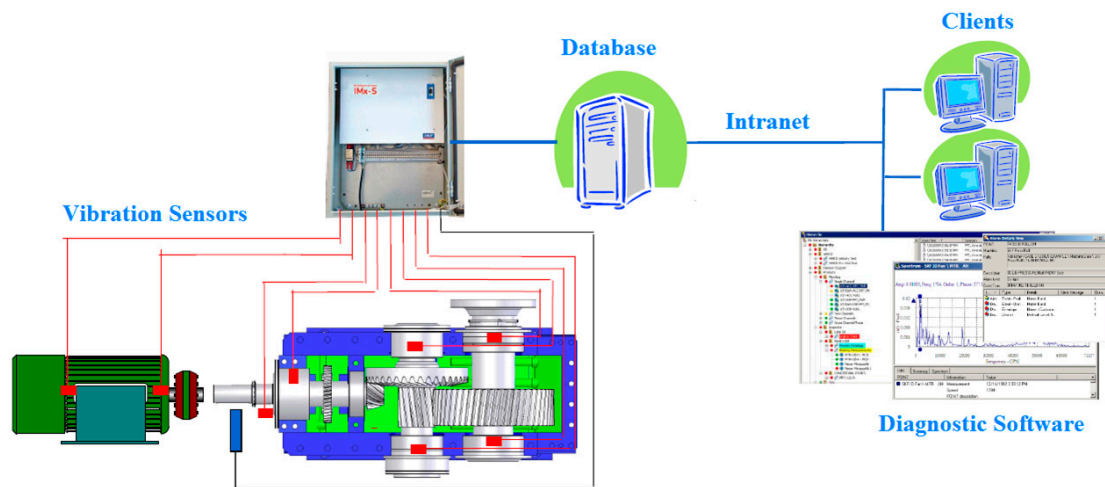

Figure 4. Online vibration monitoring system.

Many techniques for FDD in bearings based on vibration signal analysis have emerged in recent years. Generally, an FDD can be decomposed into three steps: data acquisition, feature extraction, and classification. An effective feature extraction as the key step represents a mapping of vibration signals from their original measured space to the feature space which contains more valuable information for FDD. Even though time-domain features, e.g. peak, mean, root mean square, variance, have also been employed as input features to train a bearing FDD classifier the fast Fourier transform (FFT) is the most widely applied and established feature extraction methods [1]. However, the techniques based on FFT are not suitable for analysis of non-stationary signals. Since vibration signals often contain non-stationary components, for a successful FDD it is very important to reveal such information as well. Thus, a supplementary technique for non-stationary signal analysis is necessary. Time-frequency techniques, e.g. the Wigner–Ville distribution (WVD) [2] and the short-time Fourier transform (STFT) [3] also have their own disadvantages. The WVD bilinear characteristic causes interference terms in the time-frequency domain while the STFT results in a constant resolution for all frequencies having in mind that it uses the same window size for the analysis. The wavelet transform very accurately resolves all these deficiencies. It ensures a good frequency resolution and low time resolution for low-frequency components while for high-frequency components it provides low frequency resolution and good time resolution. Therefore the wavelet transform is widely applied in the vibration signal analysis and feature

extraction for bearing FDD [4, 5]. A precise classification as the next step directly depends on previously extracted features, i.e. there is no a classifier which can make up for the information lost during the feature extraction. As in the case of the feature extraction, we can come across a wide range of classifiers used for FDD in bearings. The classifiers based on artificial neural networks [6-8] and fuzzy logic [9, 10] demonstrated a very reliable classification. However, the disadvantage of the mentioned classification techniques is that they require the availability of a very big training set. They also have a large number of parameters to be selected and adjusted in order to obtain acceptable results [11]. Therefore, there is a strong need to make the classification process simpler, faster and accurate using the minimum number of features and parameters.
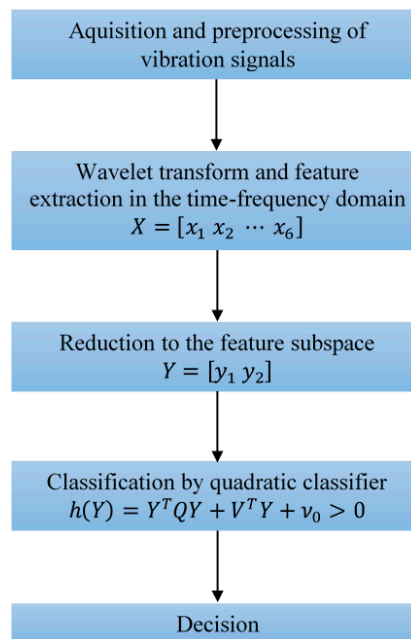
Aquisition and preprocessing of
vibration signals

Wavelet transform and feature
extraction in the time-frequency domain
$X = [x_1 \; x_2 \; \cdots \; x_6]$

Reduction to the feature subspace
$Y = [y_1 \; y_2]$

Classification by quadratic classifier
$h(Y) = Y^T Q Y + V^T Y + v_0 > 0$

Decision

Figure 5. Flowchart of the new technique for early FDD.

In this paper a new technique for early FDD is proposed. As shown in Figure 5 it has a few steps. The first one is acquisition of signals as well as their preprocessing which includes normalization and segmentation. In the next step and after the wavelet transform of vibration signals 6 representative features, i.e. the standard deviation of the wavelet coefficients in six sub-bands, are extracted in the time-frequency domain. In the third step the feature space dimension is optimally reduced to two using scatter matrices while in the final step in total three quadratic classifiers are designed [12], the one for detection and the other two for diagnosis of bearing faults. Using this new approach, the overall complexity of FDD is decreased and at the same time a very high accuracy maintained compared with already available techniques which employ more complex training algorithms.

2. Materials and methods

2.1 *Acquisition and preprocessing of the vibration signals*

Before testing and deployment of the new technique into a real production environment, its capability was tested using the vibration data obtained from the CWRU Bearing Data Center [11] since it has

become a standard reference in the field of FDD in bearings. A ball bearing as one shown in Figure 7 was installed in a motor-driven system presented in Figure 6. An accelerometer with a bandwidth up to 5000 Hz and a 1 V/g output was used to acquire the vibration signals from the bearing.
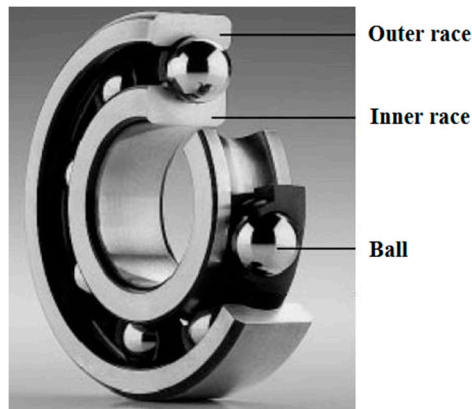
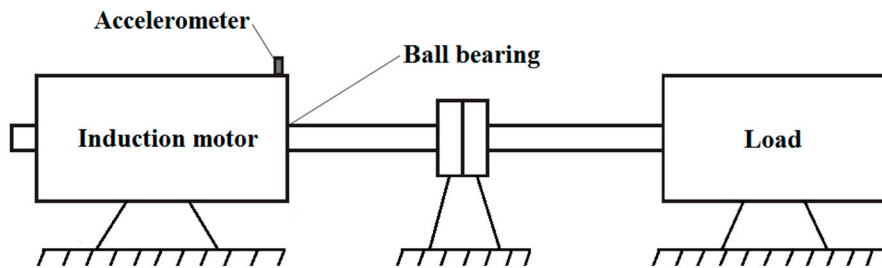

Figure 6. Ball bearing.

Figure 7. Experimental system.

The sampling rate of 12000 Hz is ample having in mind that the frequency content of interest does not exceed 5000 Hz. In total four sets of data were obtained and used:

1. Under normal conditions;
2. With inner race faults;
3. With ball faults;
4. With outer race faults.

The faults with diameter from 0.007 to 0.40 inches and depth of 0.011 inches were introduced separately at the bearing elements. The bearing was tested under different loads, i.e. 0, 1, 2 and 3 hp, while the shaft rotating speeds were 1730, 1750, 1772 and 1797 rpm. Only the smallest fault diameter was selected for this study since we were interested in an early FDD. In order to make the entire technique more robust and less dependent on the vibration signal magnitude the recorded vibration signals were normalized to zero mean and unit variance. The vibration signals collected from each of four different conditions are divided into 256 segments of 1024 sample each, as shown in Figure 8. In total 1024 segments were used, 512 for design and 512 for testing of the new technique for early FDD in bearings.
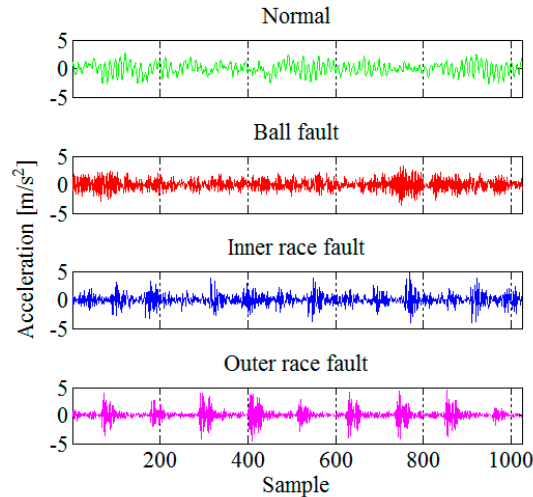
Figure 8. Segments of the vibration signals collected from four different conditions of the ball bearing.

2.2 Wavelet transform

As we already know a signal can be represented as linear combination of basic functions. A unit impulse function with limited power is limited and non-zero mean is the basic function in the time domain. In the frequency domain, the role of basic function is assigned to the sinusoidal function with infinite power and zero mean. When using the wavelet transform to transform the signal from the time domain to the time-frequency domain, the basic function is the wavelet. The wavelet is a function of limited power and zero mean [13], and for which the following is valid

$$\sum_{n=-\infty}^{\infty} |\psi[n]|^2 < \infty, \ \sum_{n=-\infty}^{\infty} \psi[n] = 0. \tag{1}$$

The wavelet can be moved in time for b samples and scaled by the so-called dilation parameter a. In such a case it is given by

$$\psi_{ab}[n] = \frac{1}{\sqrt{a}} \psi\left[\frac{n-b}{a}\right]. \tag{2}$$

If the dilation parameter changes, the basic wavelet ($a = 1$) changes its width, and thus spreads ($a > 1$) or contracts ($0 \le a < 1$) in the time domain as shown in Figure 9. In the analysis of non-stationary signals, such a possibility represents a significant advantage, considering the fact that wider wavelets are useful for extraction of slower changes, i.e. low frequency components, while narrower wavelets are useful for extraction of faster changes, i.e. high frequency components. Following the selection of a and b it is possible to transform segments of the signal $x[k]$ of N samples, and calculate the wavelet transform coefficients in the following way

$$w_{ab}[n] = \sum_{\tau=1}^{N} x[\tau]\psi_{ab}[n-\tau], 1 \le n \le N \tag{3}$$

Only those frequencies which are within the wavelet frequency band $\psi_{ab}[n]$ are extracted, i.e. the signal is filtrated by the wavelet $\psi_{ab}[n]$. Using the obtained wavelet coefficients, the original signal can

be reconstructed that is inverse wavelet transform. It is also possible to independently reconstruct both filtered and rejected part of the signal by the wavelet $\psi_{ab}[n]$ using the so-called detail and approximation coefficients respectively, which are of course a function of the transformation coefficients $\psi_{ab}[n]$.

Figure 9. Sinusoid and two wavelets with different width.

Parameters a and b can be continuous. However it is not so much practical since the signal can be transformed and reconstructed by using smaller number of wavelets, i.e. a limited number of discrete values of a and b. It is known as the discrete wavelet transform (DWT) where parameters a and b are the powers of 2 and in that case frequency bands do not overlap each other. The dilation parameter a, as the power of 2, at each subsequent higher level of transformation, doubles in value in comparison to the value from the previous level. Thus, the signal frequency band from the previous level is split into two halves at every next level, into a higher band which contains finer changes, or details, and a lower band which is an approximation of the signal from the previous level. This technique is known as the wavelet decomposition.

2.3   Dimension reduction in the feature space

Let $X = [x_1 \, x_2 \, \cdots \, x_n]^T$ be transformed into $Y = [y_1 \, y_2 \, \cdots \, y_n]^T = A^T X$ where A is an n-dimensional transformation matrix. Mapping of X into the space which is made up by the eigenvectors $\Phi_1, \Phi_2, \dots, \Phi_n$ of its covariance matrix Σ_X is known as the principal component analysis (PCA) [14]. When reducing the feature space dimension using the PCA [14] the performance of each feature $x_1, x_2, \dots, x_n$ is characterized by its eigenvalue $\lambda_1, \lambda_2, \dots, \lambda_n$, respectively. Thus, by rejecting features we should first reject those with the smallest eigenvalue. In other words, the rejected features have the smallest variance in the new feature space. In the case shown in Figure 10 where the dimension is reduced from two to one, the mapped feature y_2 would be rejected based on the PCA as less informative even though it has better discriminatory potential than y_1.
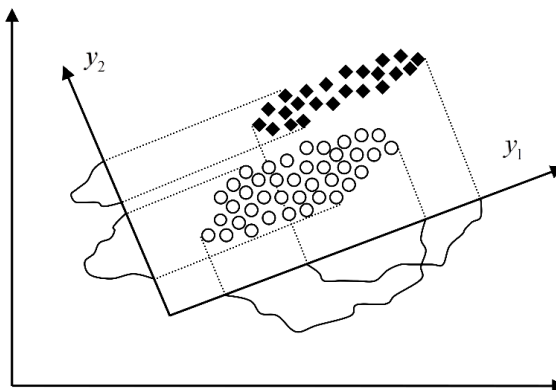

Figure 10. Dimension reduction in the feature space.

Unlike the PCA, the dimension reduction based on scatter matrices [12] is more interesting in this work since it takes into consideration also classification as one of purposes of the dimension reduction. Let L be the number of classes to be classified and M_i and Σ_i, $i = 1 \cdots L$ their mean vectors and covariance matrices. Then the within-class scatter matrix is given as:

$$S_W = \sum_{i=1}^{L} P_i \, E\{(X - M_i)(X - M_i)^T / \omega_i\} = \sum_{i=1}^{L} P_i \, \Sigma_i \tag{4}$$

while is the between-class scatter matrix is defined as:

$$S_B = \sum_{i=1}^{L} P_i \, (M_i - M_0)(M_i - M_0)^T. \tag{5}$$

M_0 is the joint vector of mathematical expectation for all the classes together, i.e.

$$M_0 = E\{X\} = \sum_{i=1}^{L} P_i \, M_i. \tag{6}$$

In addition the mixed scatter matrix can be given as:

$$S_M = E\{(X - M_0)(X - M_0)^T\} = S_W + S_B. \tag{7}$$

Then the problem of dimension reduction is reduced to the identification of the $n \times m$ transformation matrix A which maps the n-dimensional vector X into the m-dimesnional random vector $Y = A^T X$ and also maximizes the criteria $J = tr(S_W^{-1} S_B)$. This criteria is invariant to non-singular linear transformations and results into transformation matrix that takes the following form:

$$A = [\Psi_1 \ \Psi_2 \ \cdots \ \Psi_m] \tag{8}$$

where Ψ_i, $i = 1, \dots, m$ are the eigenvectors of $S_W^{-1} S_B$ which correspond to its m greatest eigenvalues. The dimension reduction based on scatter matrices applied to the case shown in Figure 10 would result into selection of the mapped feature y_2. Obviously it is much better choice than y_1 selected by the PCA in terms of more accurate classification as the main goal of the dimension reduction.

2.4   Design of quadratic classifiers

Quadratic classifiers are known as very robust solutions to the classification problems whose statistical features are change over time [12]. In addition they also provide visual insight into the classification problem. A quadratic classifier to be designed in the two-dimensional feature subspace $Y = [y_1 \ y_2]^T$ can be defined by the following equation:

$$h(Y) = Y^T Q Y + V^T Y + v_0 = [y_1 \ y_2] \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + [v_1 \ v_2] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + v_0. \tag{9}$$

The matrix Q, vector V and scalar v_0 are the unknowns to be optimally determined. Eq. (9) can be represented in a linear form as:

$$h(Y) = [q_{11} \ q_{12} \ q_{22} \ v_1 \ v_2] \begin{bmatrix} y_1^2 \\ 2y_1 y_2 \\ y_2^2 \\ y_1 \\ y_2 \end{bmatrix} + v_0 = V_z^T Z + v_0. \tag{10}$$

In order to achieve as large as possible between-class and as short as possible within-class scattering we have selected the following function as the optimization criterion [12]:

$$f = \frac{P_1 \eta_1{}^2 + P_2 \eta_2{}^2}{P_1 \sigma_1{}^2 + P_2 \sigma_2{}^2} \tag{11}$$

where P_1 and P_2 are probabilities and

$$\eta_l = E\{h(Z)/\omega_l\} = E\{V_z^T Z + v_0/\omega_l\} = V_z^T M_l + v_0 \tag{12}$$

$$\sigma_l{}^2 = var\{h(Z)/\omega_l\} = var\{V_z^T Z + v_0/\omega_l\} = V_z^T \Sigma_l V_z. \tag{13}$$

M_l and Σ_l are the mean vectors and covariance matrices, respectively, of the vector Z for each of classes which should be classified. After optimization of the function f, the optimal vector V_z, and thus the optimal matrix Q and vector V, gets the following form

$$V_z = \begin{bmatrix} q_{11} \\ q_{12} \\ q_{22} \\ v_1 \\ v_2 \end{bmatrix} = [P_1 \Sigma_1 + P_2 \Sigma_2]^{-1} (M_2 - M_1) \tag{14}$$

while the optimal scalar is

$$v_0 = -V_z^T (P_1 M_1 + P_2 M_2) \tag{15}$$

which finishes the design of the quadratic classifier.

3. Results and discussion

In order to extract representative features and before we apply the wavelet transform it is necessary to choose both the basic wavelet type and the number of resolution levels into which the vibration signal segments will be decomposed. After analysis of several types of the basic wavelets, the fourth-order Daubechies wavelet was selected since it demonstrated better discriminatory potential and also has good localizing properties both in the time-frequency domain [13]. The five-level wavelet decomposition of the vibration signals recorded with a sampling frequency of 12000 Hz resulted into the following six frequency sub-bands A_5 0-187.5 Hz, D_5 187.5-375 Hz, D_4 375-750 Hz, D_3 750-1500 Hz, D_2 1500-3000 Hz and D_1 3000-6000 Hz.

Following the wavelet decomposition, the standard deviation of the obtained wavelet coefficients in each sub-band were extracted as representative features in the time-frequency domain. Note that the standard deviation is used here to quantify the average and relative energy in each sub-band because the vibration signals were previously normalized to zero mean and unit variance. In total 6 features were extracted for each of 1024 analyzed segments. Now each original segment recorded in the time domain is represented by its feature vector $X = [x_1 x_2 \cdots x_6]^T$. The extracted features for all four classes are given in Table 1.

Table 1. Standard deviation (SD) of the wavelet coefficients in different sub-bands.

Index	Feature	Sub-band	Normal		Ball fault		Inner race fault		Outer race fault	
			μ	σ	μ	σ	μ	σ	μ	σ
x_1	SD	A_5	2.705	0.585	0.795	0.383	0.728	0.467	0.473	0.484
x_2	SD	D_5	1.694	0.193	0.545	0.093	0.431	0.126	0.247	0.116
x_3	SD	D_4	1.288	0.106	0.636	0.092	0.750	0.058	0.302	0.063
x_4	SD	D_3	1.847	0.199	0.525	0.087	0.919	0.044	0.243	0.056
x_5	SD	D_2	0.800	0.033	1.239	0.052	1.267	0.057	1.255	0.097
x_6	SD	D_1	0.220	0.025	1.041	0.035	0.951	0.037	1.083	0.055
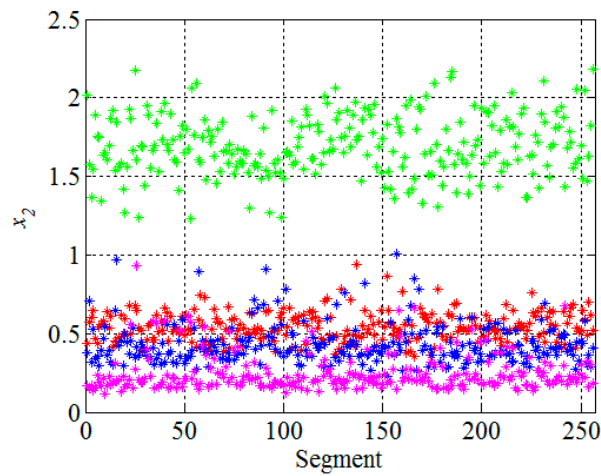
Figure 11. Standard deviation of the wavelet coefficients in the sub-band D_5, normal (green), ball fault (red), inner race fault (blue) and outer race fault (magenta)

Obviously in presence of a bearing fault there is a shift in energy in the vibration signals from lower to higher sub-bands. In Table I and Figure 11, it can also be noticed that most of the extracted features have a certain potential for the fault detection but not for the fault diagnosis. Therefore it is necessary to find their optimal combination in order to achieve a better separability between different classes of bearing condition. That is usually done by a mapping of the existing feature space into a new one whose dimension can be reduced without any significant loss of information that makes the classification process much simpler. Although the PCA is one of the most widely used techniques for reduction of the feature space dimension [14] in this paper we apply the technique based on scatter matrices [12] since it is more suitable for classification problems as described in Section 2.3. At first, we reduce the feature space dimension to enable the fault detection and then repeat the same procedure in order to diagnose the detected fault as shown in Figures 12 and 13. Obviously in this way separabilty between different classes is increased compared with Figures 11. After the dimension reduction we designed suitable quadratic classifiers following the procedure described in Section 2.4, that is also the last step in design of the new technique for early FDD. The first quadratic classifiers shown in Figure 12 separates normal from faulty condition, i.e. performs the fault detection, while the other two quadratic classifiers shown in Figure 13 are able to separate all three different bearing faults from each other, and thus perform the fault diagnosis.
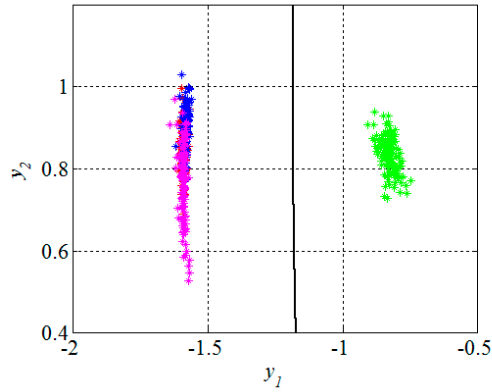
Figure 12. Dimension reduction and classification for the fault detection, normal (green) and faulty (red, blue and magenta) segments of the design set.
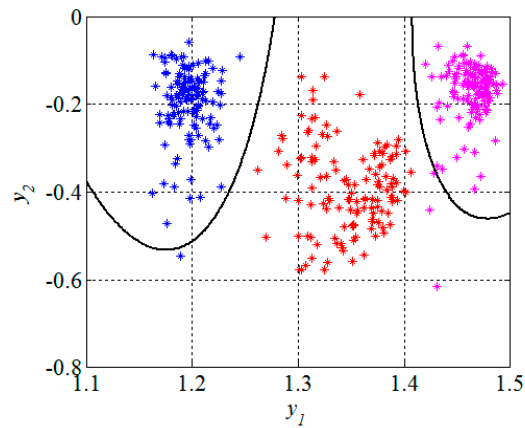

Figure 13. Dimension reduction and classification for the fault diagnosis, segments of the design set with ball fault (red), inner race fault (blue) and outer race fault (magenta).

Unlike in the previous two figures where the design set of 512 segments is shown, Figures 14 and 15 show the remaining set of 512 segments used to test the performance of the new technique for early FDD as well. Statistical performances such as sensitivity, specificity and accuracy are estimated [14].
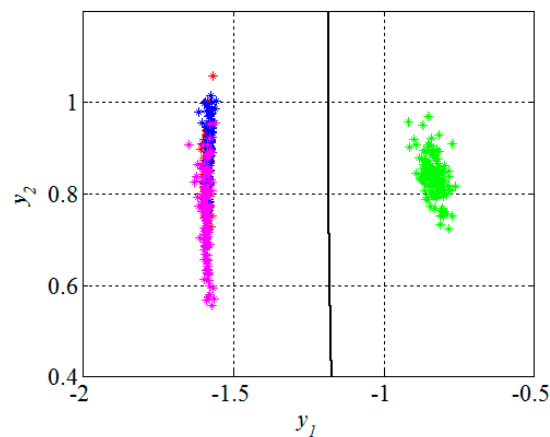

Figure 14. Dimension reduction and classification for the fault detection, normal (green) and faulty (red, blue and magenta) segments of the testing set.
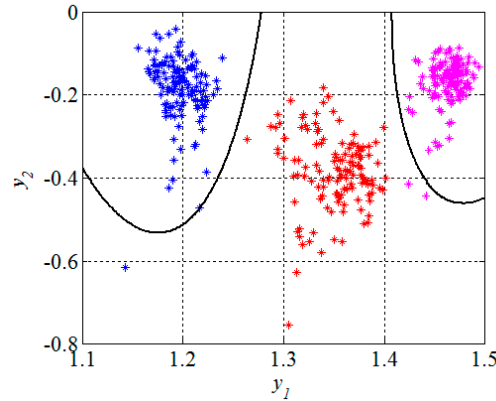
Figure 15. Dimension reduction and classification for the fault diagnosis, segments of the testing set with ball fault (red), inner race fault (blue) and outer race fault (magenta).

The classification results can also begiven by a confusion matrix shown in Table 2, where its each cell contains number of classified features for each combination of three classes of the vibration signals. Based on Figure 15 and the confusion matrix, we can conclude that all the segments from the ball fault class were correctly classified. However, the remaining two classes contained in total three segments which were incorrectly classified, i.e. classified as they belong to the ball fault class. The statistical performances are given in Table 3.The total accuracy of the new technique for early FDD in bearings is 98.9%. Usually, quadratic classifiers are robust and do not result in overtraining when the number of estimated parameters is much less than the number of analyzed samples.

Table 2. Confusion matrix

Fault type (input/output)	Ball fault	Inner race fault	Outer race fault
Ball fault	128	2	2
Inner race fault	0	126	0
Outer race fault	0	0	126

Table 3. Statistical performances

Fault type	Statistical performances [%]		
	Sensitivity	Specificity	Accuracy
Ball fault	100.0	98.4	
Inner race fault	98.4	99.2	98.9
Outer race fault	98.4	99.2	

Taking into account the results of other techniques tested on the same vibrations signals, e.g. 41 papers published in Mechanical Systems and Signal Processing between 2004 and early 2015 [16], the new technique demonstrated a very good performance which is either better or comparable with other available techniques which usually deploy much more complex algorithms. In addition, in this work only the segments with the smallest fault diameter were used because we were interested in incipient FDD. It should also be emphasized that the vibration signals were normalized before further processing. In that way we managed to overcome one of main disadvantages of other techniques in terms of application in a real production environment since most of them also depend on the amplitude of the

vibration signals. However, the amplitude has been found as unreliable in real applications since it varies even with healthy bearings, e.g. depending on their load.

4. Conclusion

In order to further increase reliability and safety of production machines in the packaging materials industry it is necessary to deploy an advanced techniques for automated early fault detection and diagnosis. In this paper we described such a technique to be used in rotating-element bearings as the most common components of production machines in the industry. The new technique based on the wavelet transform and statistical pattern recognition demonstrated a very high accuracy that is either better or comparable with other available techniques which in most cases require a big training set and a large number of parameters necessary to select and adjust in order to obtain acceptable results. A special attention has been paid to robustness of the new technique not only during the feature extraction and the reduction of the feature space dimension but also during the classification process that resulted in the choice of quadratic classifiers known for both their simplicity and a high level of robustness in the applications of this type. Quadratic classifiers have also possibility to visualize the classification results in two-dimensional space. As future work we plan to test the new technique in a real production environment at Tetra Pak.

Acknowledgements

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Tandon, N.; Choudhury, A. A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings. *Tribology International* **2007**, 32, 469-480.

[2] Russell, P.C.; et al. Extraction of information from acoustic vibration signals using Gabor transform type devices. *Measurement Science and Technology* **1998**, 9, 1282–1290.

[3] Koo, I.S.; Kim, W.W. The development of reactor coolant pump vibration monitoring and a diagnostic system in the nuclear power plant. *ISA Transactions* **2000**, 39, 309–316.

[4] Peng, Z.K.; Chu, F.L. Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography. *Mechanical Systems and Signal Processing* **2004**, 18, 199-221.

[5] Kankar, P.K.; et al. Fault diagnosis of ball bearings using continuous wavelet transform. *Applied Soft Computing* **2011**, 11, 2300-2312.

[6] Lazzerini, B.; Volpi, S. Classifier ensembles to improve the robustness to noise of bearing fault diagnosis. *Pattern Analysis and Applications* **2013**, 16, 235-251.

[7] Samanta, B.; Al-Balushi, K.R. Artificial neural network based fault diagnostics of rolling element bearings using time-domain features. *Mechanical Systems and Signal Processing* **2003**, 17, 317-328

[8] Samanta, B.; et al. Artificial neural networks and genetic algorithm for bearing fault detection. *Soft Computing* 2006, 10, 264-271.

[9] Lou, X.; Loparo, K.A. Bearing fault diagnosis based on wavelet transform and fuzzy inference. *Mechanical Systems and Signal Processing* **2004**, 18, 1077-1095.

[10] Lei, Y.; et al. Fault diagnosis of rotating machinery based on a new hybrid clustering algorithm. *The International Journal of Advanced Manufacturing Technology* **2008**, 35, 968-977.

[11] Stepanic, P.; et al. A new approach to detection of defects in rolling element bearings based on statistical pattern recognition. *The International Journal of Advanced Manufacturing Technology* **2009**, 45, 91-100.

[12] Fukunaga, K. *Introduction to Statistical Pattern Recognition*, Academic Press, 1990.

[13] Rao, R.; Bopardikar, A. *Wavelet transforms: Introduction to theory and applications*, Addison-Wesley Longman, 1998.

[14] Sun, W.; et al. Decision tree and PCA-based fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing* **2007**, 21, 1300-1317.

[15] Gajic, D.; Detection of epileptiform activity in EEG signals based on time-frequency and non-linear analysis. *Frontiers in computational neuroscience* **2015**, 9, 38.

[16] Wade, A.; et al. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study. *Mechanical Systems and Signal Processing* **2015**, 64–65, 100-131.