Fault Diagnosis of Ball-bearings Using Principal Component Analysis and Support-Vector Machine

O.R. Seryasat [1]*, H. Ghayoumi Zadeh [2], M. Ghane [3], Z. Abooalizadeh [4], A. Taherkhani [5], F. Maleki [6]

[1, 2, 3] Department of Electrical Engineering, Takestan branch, Islamic Azad University, Takestan, Iran
[4, 5, 6] Department of Physics, Faculty of Science, Takestan branch, Islamic Azad University, Takestan, Iran.
Corresponding author: *Omid.Seryasat@gmail.com

Abstract: Due to the importance of rolling bearings as one of the most commonly used industrial machinery elements, it is necessary to develop proper monitoring and fault diagnosis procedure to suppress malfunctioning and failure of these elements during operation. For rolling bearing fault detection, it is expected that a desired time domain analysis method has good computational efficiency. In this paper, first, the features in time and frequency domain such as Mean Square, Moments, Cumulant, kurtosis, Skeweness, Zero Crossing Rate, Peak Rate, standard deviation, maximum value, Crest factor, Clearance factor, Shape factor and Impulse factor which are widely used in fault diagnostics, have been extracted from the vibration signal. Indeed, the numbers of 12 features have been extracted from the time domain signal. Then they are going to PCA algorithm, After PCA processing, the redundant features can be removed effectively. In this work, 12 features decrease to six efficient features. Although most of the features are reduced, the average diagnosis accuracy does not decrease. For some states, the diagnosis accuracy arises a little for the information fusion performance of PCA. Then, the features that extracted have been classified successfully using MSVM classifier.
[O.R. Seryasat, H. Ghayoumi Zadeh, M. Ghane, Z. Abooalizadeh, A. Taherkhani, F. Maleki. **Fault Diagnosis of Ball-bearings Using PCA and Support-Vector Machine**. *Life Sci J* 2013;10(1s):393-397] (ISSN:1097-8135). http://www.lifesciencesite.com. 64

Keywords: Fault diagnosis; Roller bearing; SVM; Principal component analysis (PCA)

1. Introduction

Roller bearings are the important and frequently encountered components in the rotating machines that find widespread industrial applications. Therefore, fault diagnosis of the roller bearings has been the subject of extensive research. Rolling bearing faults can have many reasons, e.g. wrong design, improper mounting, acid corrosion, bad lubrication and plastic deformation [1, 2]. The process of roller bearing fault diagnosis includes the acquisition of information, extraction of features and recognition of conditions. The latter two have priority to the first one. Different methods are used for the acquisition of information; they may be broadly, classified as vibration and acoustic measurements, temperature measurements and wear debris analysis. Among these, vibration measurements are commonly used in the condition monitoring and diagnostics of the rotating machinery [3]. The vibration measurement of the roller bearing can be made using some accelerating sensors that are placed on the bearing house. When faults occur in the roller bearing, the vibration signal of the roller bearing would be different from the signal under the normal condition [4-6]. So far, many conventional vibration-signal-analysis-based methods have been applied to rotating machine fault diagnosis. Quite a few works have been done in this field, e.g. by Wang and McFadden [7], Shiroishi et al. [8], Scholkopf [9], Dellomo [10], Li et al. [11], Jack and Nandi [12], Nikolaou and Antoniadis [13], Samanta et al. [14], Al-Ghamd and Mba [15], and Purushotham et al. [16]. The possibilities of using support vector machines (SVMs) in machine condition monitoring applications are being considered only in recent years. For example, Nandi [17], and then, Jack and Nandi [18] have provided a procedure for condition monitoring of rolling element bearing. Then they improved their work by using GAs for automatic feature selection in machine condition monitoring [12, 19-20]. Samanta et al. developed a procedure similar to that of Jack and Nandi but different in processing time-domain signal [14], where only two cases were studied which are false and normal conditions. Finally, Rojas and Nandi [20] have worked on the training of SVMs by using the sequential minimal optimization (SMO) algorithm. But, multi-class Support vector machines (MSVMs), based on statistical learning theory that are of specialties for a smaller sample number have better generalization than ANNs and guarantee the local and global optimal solution are exactly the same [21]. Meantime, the learning problem of a smaller number of samples can be solved by SVM.

Recently, it has been found that SVMs can be effectively applied to many applications [22-25]. Due to the fact that it is practically difficult to obtain sufficient fault samples, SVMs are introduced into rotating machinery fault diagnosis due to their high accuracy and good generalization for a smaller sample number. In this paper, the interesting point of this investigation is the introduction of an effective method for fault detection and diagnosis in such systems through features in optioned from vibration signals by Principal component analysis (PCA) and support vector machines (SMVs) that used for classification of rolling-element bearing faults.

The extracted features from original and preprocessed signals by using PCA are used as inputs to the classifiers for two-class (normal or fault) recognition.

The classifier parameters this features are classified successfully using SVM classifier, the classifiers are trained with a subset of the experimental data for known machine conditions and are tested using the remaining set of data. The procedure is illustrated using the experimental vibration data of a rotating machine.

1. **SVM**

In order to calculate decision surfaces directly instead of modeling a probability distribution across training data, SVM makes use of a hypothetic space of linear functions in a high dimensional feature space. A support vector (SV) kernel is utilized for mapping the data from input space to a high-dimensional feature space; this makes easy the process of the problem in linear form.

SVs are samples that have [28]. SVM always finds a global minimum because it usually tries to minimize a bound on the structural risk, rather than the empirical risk. Empirical risk is defined as the measured mean error rate on the training set as below:

$$R_{emp}(\alpha)=\frac{1}{2l}\sum_{i=1}^{l}|y_i-f(x_i,\alpha)| \qquad (1)$$

, where l is the number of observations, y_i is the class label and x_i is the sample vector. The structural risks, defined as a structure derived from the inner class of the function in the nested subset, find the subset of the function that minimizes the bound on the actual risk. SVM achieves this goal by minimizing the following Lagrangian formulation:

$$L_p=\frac{1}{2}\|w\|^2-\sum_{i=1}^{l}\alpha_i y_i(x_i.w+b)+\sum_{i=1}^{l}\alpha_i \qquad (2)$$

Where α_i is positive Lagrange multiplier [27, 28]. SVM uses some kernels to map the data from the input space to a high-dimensional feature space which facilitates the problem to be processed in linear form. In this paper, linear and radial basis function (RBF), quadratic and polynomial kernels have been used.

2. **Principal component analysis (PCA)**

The PCA is an unsupervised feature reduction method.

Principal component analysis (PCA) is a statistical technique that can be constructed by several ways, one commonly cited of which is stated in this appendix. By stating a few directly useful properties of PCA for radar signal analysis, we by no means, tend to give an even superficial survey of this ever-growing topic. For deeper and more complete coverage of PCA and its applications, [26] are also nice and shorter materials to explain some general properties of PCA. For simplifying the presentation, all the following properties of PCA are proved under the assumption that all eigenvalues of whichever covariance matrix concerned are positive and distinct. One PCA construction: Assume a random vector X, taking values in $\Re^m$, has a mean and covariance matrix of μ_X and $\sum_X$, respectively. $\lambda_1>\lambda_2>\cdots>\lambda_m>0$ are ordered eigenvalues of $\sum_X$, such that the i-th eigenvalue of $\sum_X$ means the i-th largest of them. Similarly, a vector α_i is the i-th eigenvector of $\sum_X$ when it corresponds to the i-th eigenvalue of $\sum_X$. To derive the form of principal components (PCs), consider the optimization problem of maximizing $\mathbf{var}[\alpha_1^T X]=\alpha_1^T\sum_X\alpha_1$, subject to $\alpha_1^T\alpha_1=1$. The Lagrange multiplier method is used to solve this question.

$$L(\alpha_1,\phi_1)=\alpha_1^T\sum_X\alpha_1+\phi_1(\alpha_1^T\alpha_1-1)$$

$$\frac{\partial L}{\partial\alpha_1}=2\sum_X\alpha_1+2\phi_1\alpha_1=0 \quad\Rightarrow$$

$$\sum_X\alpha_1=-\phi_1\alpha_1 \quad\Rightarrow$$

$$\mathrm{var}[\alpha_1^T X]=-\phi_1\alpha_1^T\alpha_1=-\phi_1.$$

Because $-\varphi_1$ is the eigenvalue of $\sum_X$, with α_1 being the corresponding normalized eigenvector, $\mathbf{var}[\alpha_1^T X]$ is maximized by choosing α_1 to be the first eigenvector of $\sum_X$. In this case, $z_1=\alpha_1^T X$ is named the first PC of X, α_1 is the vector of coefficients for z_1, and $\mathbf{var}(z_1)=\lambda_1$.

To find the second PC, $z_2=\alpha_2^T X$ we need to maximize $\mathbf{var}[\alpha_2^T X]=\alpha_2^T\sum_X\alpha_2$ subject to z_2 being uncorrelated with z_1.

Because $\mathbf{cov}(\alpha_1^T X,\alpha_2^T X)=0\Rightarrow\ \alpha_1^T\sum_X\alpha_2=0$ $\Rightarrow\ \alpha_1^T\alpha_2=0$, this problem is equivalently set as maximizing $\alpha_2^T\sum_X\alpha_2$, subject to $\alpha_1^T\alpha_2=0$, and $\alpha_2^T\alpha_2=1$. We still make use of the Lagrange multiplier method.

$$L(\alpha_2,\phi_1,\phi_2)=\alpha_2^T\sum_X\alpha_2+\phi_1\alpha_1^T\alpha_2+\phi_2(\alpha_2^T\alpha_2-1)$$

$$\frac{\partial L}{\partial \alpha_2} = 2\Sigma_X \alpha_2 + \phi_1 \alpha_1 + 2\phi_2 \alpha_2 = 0$$

$$\Rightarrow \quad \alpha_1^T (2\Sigma_X \alpha_2 + \phi_1 \alpha_1 + 2\phi_2 \alpha_2) = 0 \quad \Rightarrow \phi_1 = 0$$

$$\Rightarrow \Sigma_X \alpha_2 = -\phi_2 \alpha_2 \Rightarrow \alpha_2^T \Sigma_X \alpha_2 = -\phi_2 .$$

Because $-\varphi_2$ is the eigenvalue of Σ_X, with α_2 being the corresponding normalized eigenvector, $\mathbf{var}[\alpha_2^T X]$ is maximized by choosing α_2 to be the second eigenvector of Σ_X. In this case, $z_2 = \alpha_2^T X$ is named the second PC of X, α_2 is the vector of coefficients for z_2, and $\mathbf{var}(z_2) = \lambda_2$. Continuing in this way, it can be shown that the i-th PC $z_i = \alpha_i^T X$ is constructed by selecting α_i to be the i-th eigenvector of Σ_X, and has variance of λ_i. The key result in regards to PCA is that the principal components are the only set of linear functions of original data that are uncorrelated and have orthogonal vectors of coefficients.

Experimental Procedure Two data sets, each containing twenty data files, were collected from two bearings which are the same but with different faults. The first data file was collected from each test bearing when the loading was zero, and the bearing was running at the highest speed (3000 rpm).

The load was then increased step by step, the speed was kept at 3000rpm, and four other data files were collected. The load was then brought back to zero, and speed was decreased by 500 rpm; then, the next five data files were collected under five different loads similar to the first five data files. This procedure was continued until all twenty five sets of data were collected.

The sampling frequency was chosen as 41.67 kHz; this sampling frequency along with the data record size of 4098 guarantees that the sampling procedure covers at least 1.6 revolutions of shaft at the lowest speed.
The diagram block of detection of the type of faults in bearings has been illustrated in Table (1).

3. **Test bearings**

An impact impulse is generated every time a ball hits a defect in the raceway or every time a defect in a ball hits the raceway. Each of such impulses excites a short transient vibration in the bearings at its natural frequencies. Each time this defect is rolled over, an impact is produced whose energy depends on the severity of the defect. Many failure modes of a rolling element bearing produce such a discontinuity in the path of the rolling elements. Moreover the majority of rolling element bearing failure cases begin with a defect on one of the raceways. In this research, defects on inner raceway (IRD) and normal Bearing (GBR) are used.

Conclusion Due to the importance of rolling bearings as one of the most populous used industrial machinery elements, development of proper monitoring and fault diagnosis procedure to suppression malfunctioning and failure of these elements during operation is necessary.

For rolling bearing fault detection, it is expected that a desired time domain analysis method has good computational efficiency.

A procedure is presented for diagnosis of bearing condition using one classifiers, namely, SVMs with feature reduction from time-domain vibration signals by PCA.

4. **Performance evaluation**

The performance of a classifying factor can be evaluated with the use of the criteria of sensitivity, accuracy and specificity.

$$\text{sensitivity} = \frac{TP}{TP + FN} \qquad (16)$$

$$\text{specificity} = \frac{TN}{TN + FP} \qquad (17)$$

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \qquad (18)$$

In the above formulas: TP is the number of correct positive classifications (the learning machine classified correctly); TN is the number of correct negative classifications (the learning machine classifies correctly); FP is the number of incorrect positive classifications (the learning machine classified incorrectly), and FN is the number of incorrect negative classifications (the learning machine does not classify correctly).

5. **Figures and Tables**

5.1. Figures

In this section, the diagram block of detection of the type of faults (Fig.1), the original acceleration vibration signal for two types of faults at 3000rpm speed and 500N load have been shown (Fig.2).

5.2 Table

In this section, the roller bearing fault diagnosis for two type faults at 3000rpm speed and 1000N load have been shown in Table 1.

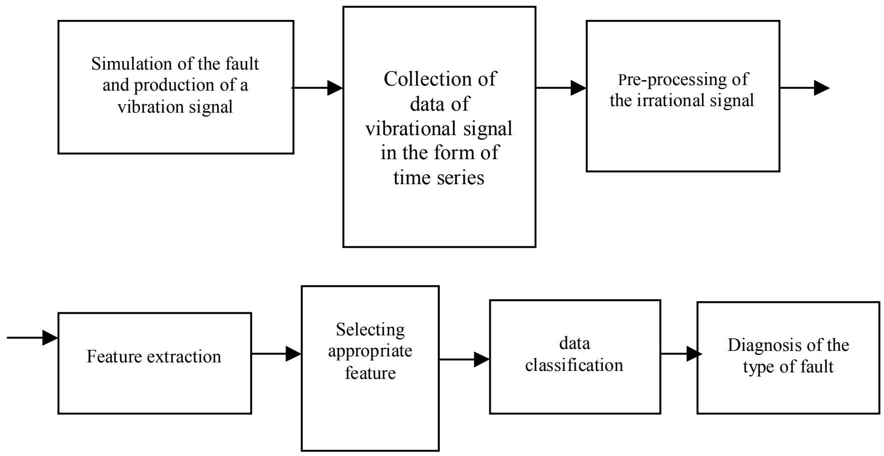

Fig.1: the diagram block of detection of the type of faults.

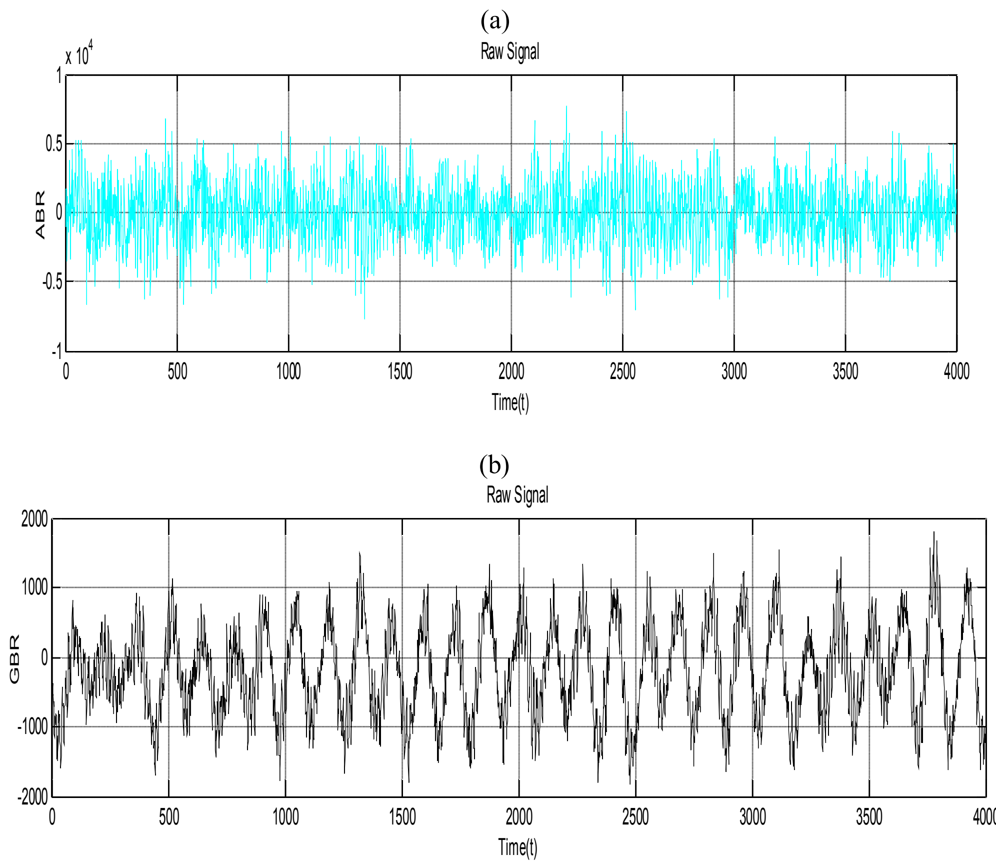

Fig.2: Original acceleration vibration of the signal for two different faults, (a): Inner race way fault, (b) Good bearing

TABLE 1: The result of the two set of data to SVM for feature extraction using PCA algorithm

DATA	PERCENT TEST EDUCATION	PERCENT TEST ACCURACY	SENCITIVE PERCENT	ACCURACY PERCENT
NORMAL BEARING	100	100	100	100
FAULTY BEARING	92.49	92.75	92.58	92.76

1. Conclusion

In this paper, first, the features in time and frequency domain such as Mean Square, Moments, Cumulant, kurtosis, Skeweness, Zero Crossing Rate, Peak Rate, standard deviation, maximum value, Crest factor, Clearance factor, Shape factor and Impulse factor which are widely used in fault diagnostics, have been extracted from the vibration signal. Indeed, the numbers of 12 features have been extracted from the time domain signal. Then they are going to PCA algorithm, After PCA processing, the redundant features can be removed effectively. In this work, 12 features decrease to six efficient features. Although most of the features are reduced, the average diagnosis accuracy does not decrease. For some states, the diagnosis accuracy arises a little for the information fusion performance of PCA. Then, the features that extracted have been classified successfully using MSVM classifier.

References

1. N.G. Nikolaou, I.A. Antoniadis, "Rolling element bearing fault diagnosis using wavelet packets," *NDT & International* 35 (2002) 197–205.
2. V. Purushotham, S. Narayanan, S.A.N. Prasad, "Multi-fault diagnosis of rolling bearing elements using wavelet analysis and hidden Markov model based fault recognition," *NDT&E International* (2005) 1–11.
3. J. Shiroishi, Y. Li, S. Liang, T. Kurfess, S. Danyluk, "Bearing condition diagnostics via vibration and acoustics emission measurement," *Mechanical Systems and Signal Processing* 11 (1997) 693–705.
4. P.D. Mcfadden, J.D. Smith, "Model for the vibration produced by a single point defect in a rolling element bearing," *Journal of Sound and Vibration* 96 (1) (1984) 69–82.
5. A. Barkov, N. Barkova, J. Mitchell, "Condition assessment and life prediction of rolling element bearings," *Sound and Vibration* 38 (1995) 10–17, +27-31.
6. V. Purushotham, S. Narayanan, S.A.N. Prasad, "Multi-fault diagnosis of rolling bearing elements using wavelet analysis and hidden Markov model based fault recognition," *NDT&E International* (2005) 1–11.
7. W.J. Wang, P.D. McFadden, "Application of wavelets to gearbox vibration signals for fault
8. detection," *Journal of Sound and Vibration* 192 (1996) 927–939.
9. J. Shiroishi, Y. Li, S. Liang, T. Kurfess, S. Danyluk, "Bearing condition diagnostics via
10. vibration and acoustic emission measurements," Mechanical Systems and Signal Processing 11 (1997) 693–705.
11. B. Scholkopf, "SVMs—A practical consequence of learning theory," *IEEE Intelligent systems* 13 (1998) 18–19
12. M.R. Dellomo, "Helicopter gearbox fault detection: A neural network-based approach," *Transaction of the ASME, Journal of Vibration and Acoustic* 121 (1999) 265–272.
13. B. Li, M.Y. Chow, Y. Tipsuwan, J.C. Hung, "Neural-network-based motor rolling bearing fault diagnosis," *IEEE Transactions on Industrial Electronics* 47 (5) (2000) 1010–1060.
14. L.B. Jack, A.K. Nandi, "Support vector machines for detection and characterization of rolling element bearing faults," *Proceedings of Institution of Mechanical Engineers, Part C, Journal of Mechanical Engineering Science* 215 (2001) 1065–1074.
15. N.G. Nikolaou, I.A. Antoniadis, "Rolling element bearing fault diagnosis using wavelet packets," *NDT&E International* 35 (2002) 197–205.
16. B. Samanta, K.R. Al-Balushi, S.A. Al-Araimi, "Artificial neural networks and support vector machines with genetic algorithm for
17. bearing fault detection," *Engineering Application of Artificial Intelligent* 16 (2003) 657–665.
18. A.K. Nandi, "Advanced digital vibration signal processing for condition monitoring," *Proceedings of COMADEM 2000, Houston, TX, USA*, 2000, pp. 129–143.
19. L.B. Jack, A.K. Nandi, "Comparison of neural networks and support vector machines in condition monitoring applications," *Proceeding of COMADEM 2000, Houston, TX, USA*, 2000, pp. 721–730.
20. L.B. Jack, A.K. Nandi, "Genetic algorithm for feature extraction in machine condition monitoring with vibration signals," *IEE Proceeding of Vision & Image Signal Processing* 147 (2000) 205–212.
21. L.B. Jack, A.K. Nandi, "Fault detection using support vector machines and artificial neural networks, augmented by genetic algorithms," *Journal of Mechanical Systems and Signal Processing* 16 (2002) 373–390.
22. A. Rojas, A.K. Nandi, "Practical scheme for fast detection and classification of rolling-element bearing faults using support vector machines," *Mechanical Systems and Signal Processing* 20 (7) (2006) 1523–1536.
23. V.N. Vapnik, "The Nature of Statistical Learning Theory," *Springer Verlag*, New York, 1995.
24. H.-C. Kim, S. Pang, H.-M. Je, et al., "Constructing support vector machine ensemble," *Pattern Recognition* 36 (12) (2003) 2757–2767.
25. G. Guo, S.Z. Li, K.L. Chan, "Support vector machines for face recognition," *Image and Vision Computing* 19 (2001) 631–638.
26. O. Barzilay, V.L. Brailovsky, "On domain knowledge and feature selection using a support vector machine," *Pattern Recognition Letters* 20 (5) (1999) 475–484.
27. U. Thissen, R.V. Brakel, A.P. De Weijer, et al., "Using support vector machines for time series prediction," *Chemometrics and Intelligent Laboratory Systems* 69 (1) (2003) 35– 49.
28. I. Golightly and D. Jones, Senior Member, IEEE "Visual control of an unmanned aerial vehicle for power line inspection".
29. Jolliffe, Theodoridis and Koutroumbas, 2006; Chapter 8, Mardia et al., 1979; Chapter 11, Anderson, 2003, p266~p272.
30. C.J.C. Burges, "A Tutorial on Suport Vector Machines for Pattern Recognition," *Knowledge Discovery Data Mining*, vol. 2, no.2, pp. 121- 167, 1998.
31. C. J. C. Burges and B. Schoclkopf, "Improving speed and accuracy of Support Vector Learning Machines" *Advances in Neural information processing System*, vol. 9, pp. 375-381, 1997.

1/5/2013