Fault Diagnosis of Low Speed Bearing Using Support Vector Machine

Achmad Widodo[†], Jong-Duk Son[†], Bo-Suk Yang[†]
Dong-Sik Gu[*], Byeong-Keun Choi[*]
Yong-Han Kim[**], Andy C.C Tan[**], Joseph Mathew[**]

ABSTRACT

This study presents fault diagnosis of low speed bearing using support vector machine (SVM). The data used in the experiment was acquired using acoustic emission (AE) sensor and accelerometer. The aim of this study is to compare the performance of fault diagnosis based on AE signal and vibration signal with same load and speed. A low speed test rig was developed to simulate various defects with shaft speeds as low as 10 rpm under several loading conditions. In this study, component analysis was also performed to extract the feature and reduce the dimensionality of original data feature. Moreover, the classification for fault diagnosis was also conducted using original data feature without feature extraction. The result shows that extracted feature from AE sensor gave better performance in faults classification.

Keywords: Fault diagnosis; Low speed bearing; Component analysis; Support vector machines.

1. Introduction

There are industries that equipped by low speed rotating machinery such as rolling machine in paper mill, steel pipe and mining industries. Also, low speed rotating machine can be found in wind turbine power plant. As rotating machinery, bearing is critical component and sometimes it must carry heavy loads and operate at high efficiency and reliability. Therefore, condition monitoring and fault diagnosis of low speed bearing is very useful to guarantee the effectiveness and reliability of this machine. Establishing intelligent system for faults detection of low speed rotating machine is a solution. To face this issue, the research area in machine learning has been applied to perform condition monitoring, faults detection and classification.

Furthermore, machine fault diagnosis based on classification technique such as SVM will perform well if the data input consists of useful features which represent each condition or class. In the present study, we employ component analysis that is aimed to support the data preparation process. Using component analysis, the useful features can be extracted from the original data and high-dimensional of original data can be reduced by removing irrelative features, so that the classifier will reach high accuracy.

2. Methods

2.1. Support vector machine (SVM)

SVMs are a kind of machine learning based on statistical learning theory. The basic idea of applying SVM to pattern classification can be stated as follows: first, map the inputs vectors into one features space, possible in higher space, either linearly or nonlinearly, which is relevant with the kernel function. Then, within the feature space from the first step, seek an optimized linear division, that is, construct a hyperplane which separates two classes. It can be extended to multi-class. SVMs training always seek a global optimized solution and avoid over-fitting, so it has ability to deal with a large number of feature. A complete description about SVMs is available in [1]. In the linear separable case, there exists a separating hyperplane whose function is

$$\mathbf{w}\cdot\mathbf{x} + b = 0 \qquad (1)$$

which implies

$$y_i(\mathbf{w}\cdot\mathbf{x} + b = 0) \geq 1, \quad i = 1,\ldots, N \qquad (2)$$

By minimizing $||\mathbf{w}||$ subject to this constrain, the SVMs approach tries to find a unique separating hyperplane. Here $||\mathbf{w}||$ is the Euclidean norm of $\mathbf{w}$, and the distance between the hyperplane and the nearest

† Pukyong National University, Korea.
E-mail : achmadwid2006@yahoo.com
Tel : (051) 625-1604
* Gyongsang National University, Korea.
** Queensland University of Technology, Australia.

data points of each class is $2/\|\mathbf{w}\|$. By introducing Lagrange multipliers α_i, the SVMs training procedure amounts to solving a convex quadratic problem (QP). The solution is a unique globally optimized result, which has the following properties

$$w = \sum_{i}^{N} \alpha_i y_i \mathbf{x}_i \tag{3}$$

Only if corresponding $\alpha_i > 0$, these $\mathbf{x}_i$ are called support vectors.

When SVMs are trained, the decision function can be written as

$$f(\mathbf{x}) = sign\left(\sum_{i=1}^{N} \alpha_i y_i (\mathbf{x} \cdot \mathbf{x}_i) + b \right) \tag{4}$$

For a linear non-separable case, SVMs perform a nonlinear mapping of the input vector $\mathbf{x}$ from the input space R^n into a higher dimensional Hilbert space, where the mapping is determined by kernel function. According to the different classification problems, the different kernel function can be selected to obtain the optimal classification results.

2.2 Component analysis (CA)

Component analysis is a technique of multivariate statistical analysis that can linearly or nonlinearly transforms an original set variables into a substantially smaller set variables. It can be viewed as a classical method of multivariate statistical analysis for dimensionality reduction. Because of the fact that a small set of uncorrelated or independent variables is much easier to understand and use in further analysis than a larger set of correlated or dependent variables. This technique has been widely applied to virtually every substantive area including cluster analysis, visualization of high-dimensionality data, regression, data compression and pattern recognition. In this research, component analysis is used to extract the sensitive feature from original features and to reduce the dimension of original features by means of principal component analysis (PCA) [2], independent component analysis (ICA) [3], kernel PCA [4] and kernel ICA.

3. Signal process and preparation

The time series signal can be used to perform fault diagnosis by analysing vibration or acoustic signal obtained from experiment. Statistical methods are widely used that can able to present the physical meaning of time data series. For instance, the use of overall root-means-square (RMS) and crest factor (ratio of peak value to RMS) has been applied for detection of localized defects [5]. Moreover, probability density has also been used popularly for bearing defect detection [6].

In this study, statistical method is employed to investigate the characteristic of the system by calculating 14 statistical feature parameters in time and frequency domain presented as follows: mean, RMS, shape factor, skewness, kurtosis, crest factor, entropy error, entropy estimation, histogram lower and upper, peak value, RMS frequency, frequency center and root variance frequency.

For detecting rolling element bearing failures, we also performed envelope analysis to show bearing characteristics frequencies by isolating other unwanted signals. Envelope analysis typically refers to sequence of the following procedures: (1) Band-pass filtering (BPF), (2) Signal rectification, (3) Hilbert transform of low-pass filtering, and (4) Power spectrum. The purpose of BPF is to reject the low-frequency high-amplitude signals to eliminate random-noise outside the pass-band. In the present study, we employed six band-pass setting for enveloping: BPF1: 5-15 kHz, BPF2: 15-25 kHz, BPF3: 25-35 kHz, BPF4: 35-55 kHz, BPF5: 55-75 kHz, BPF6: 75-100 kHz. From the enveloping process, six features called mean-peak ratio [7] are calculated using Eqs. (5)-(7).

$$mPRO = 20\log_{10} \frac{\sum_{j=1}^{n} (P_j - A_s)}{A_s} \text{ (dB)} \tag{5}$$

$$mPRI = 20\log_{10} \frac{\sum_{j=1}^{n} (P_j - A_s) + \sum_{j=1}^{n} (Ps_i - A_s)}{A_s} \text{ (dB)} \tag{6}$$

$$\text{where } A_s = \frac{\sum_{k=a}^{b} S_k}{(b-a)} \tag{7}$$

4. Experiment

The low speed machinery fault simulator was developed to conduct research on low speed condition monitoring and fault diagnosis [8]. This test rig enables modelling of bearing and gearbox faults under different loading conditions and verification of condition monitoring at low speed as low as 10 rpm. At the driving end, the shaft is attached to a reduction gear box (10.1:1) through a coupling. The constant radial load can be applied close to the driven-end support for long period that is measured by load cell. An AE sensor (type R3a from Physical Acoustic Corporations) with frequency range 25-530 kHz was attached on the top of the bearing housing using magnetic holder as shown in Fig. 1a.

The data acquisition process is presented in Fig. 1b. The bearings used in this study are roller bearing: SKF NF307 and N307, with the inner ring and outer ring are separable. The test bearing enables an easy access to the raceway for seeded defects and to observe the surface condition. The faults of crack and spall was simulated by a hair-line scratching using diamond bit and grinding using air-speed grinding tool, respectively. All type seeded defect bearings used in this study (Fig. 2) are listed as follows: inner-race crack (IFC1), inner race spall (IF1), outer-race crack (OFC1), outer-race spall (OF1), small spall on roller (BF1), medium spall on roller (BF2). In addition, the normal bearing was also experimented for benchmarking. Totally, we have 6 classes of faulty bearings and one normal condition for classification.

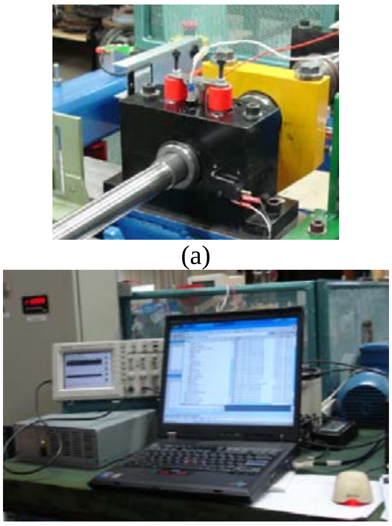

Fig. 1 Location of AE sensor (a) and data acquisition process (b).

5. Training and classification

SVM based multi-class classification is applied to perform the classification process using one-against-one [9]. Sequential minimal optimization (SMO) proposed by Platt [10] is used to solve the SVM classification problem.

In this study, RBF kernel ($K = \exp(-\|\mathbf{x}-\mathbf{x}_j\|^2/2\gamma^2)$) and polynomial ($K = (\gamma\ \mathbf{x}^T\cdot\mathbf{x}_j + r)^d$) are used as the basic kernel function of SVM. To select proper kernel parameters (C, γ and d), we used cross-validation technique [11] to obtain good performance of classification and to avoid overfitting or underfitting problem.

6. Results and discussion

The results of this study can be shown in Tables 1, 2, and 3. In these tables, we listed the kernel function, classification error (%) for training and testing and number of support vector (SV). The classification error is determined by using ratio of correct classification and on the whole of training or testing respectively.

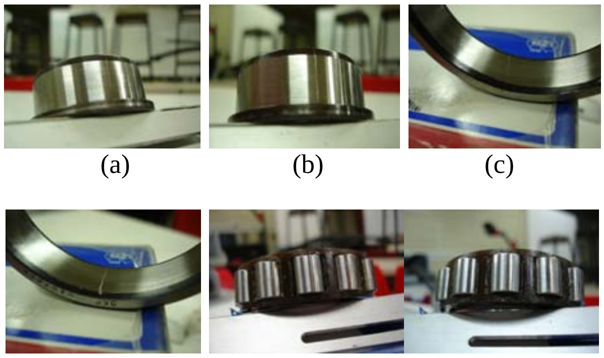

Fig. 2 Seeded defects on the bearing: (a) IFC1, (b) IF1, (c) OFC1, (d) OF1, (e) BF1, (f) BF2.

Table 1 shows the classification performance of original signal without feature extraction via component analysis. The results are not good for both kernel functions: RBF and polynomial. Even though the training process reached good accuracy for RBF kernel, however the testing performance is bad, the error is very high. The lack performance of classification is due to the presence of irrelative features that decrease the accuracies.

Table 2 presents classification accuracy of vibration signal augmented by feature extraction using component analysis. ICA feature extraction gives best performance among others both for RBF and polynomial kernel. This phenomenon can be explained that ICA finds the components not merely uncorrelated but independent. Independent components are more useful for classification rather than uncorrelated components. The reason is the negentropy in ICA could take into account the higher order information of the original inputs better than PCA. However, RBF kernel outperforms polynomial kernel in accuracies. KICA also gives better performance although its accuracies are not as good as ICA does.

In this study, PCA and KPCA are not well performing. The best performance in the present study is given by AE signal augmented by ICA feature extraction (Table 3). ICA can extract the useful feature and reduce the dimensionality. It gives good data input for classification process. Moreover, KICA also performs better than PCA and KPCA in this study.

Table 1 Classification performance of original data

Signal	Kernel	Parameters	Error (Training/Testing), %	SV
Vibration	RBF (C, γ)	(1, 0.125)	0.0/85.7	48
	Polynomial (d, C)	(2, 1000)	82.1/87.8	56
AE	RBF (C, γ)	(1, 0.125)	0.0/87.5	48
	Polynomial (d, C)	(2, 1000)	75.0/85.7	56

Table 2 Classification performance: vibration, 20 rpm, load 5 kN

Component analysis	Kernel parameters		Error (Training/Testing), %		SV	
	RBF (C, γ)	Polynomial (d, C)	RBF	Polynomial	RBF	Polynomial
ICA	(128, 8)	(2, 1000)	9.3/20.4	26.8/30.6	45	26
PCA	(1, 0.125)	(2, 1000)	0.0/85.7	91.1/89.8	48	56
KICA	(128, 1)	(2, 1000)	20.8/73.4	5.4/79.6	47	53
KPCA	(32, 0.125)	(2, 1000)	79.2/87.8	44.6/83.7	48	53

Table 3 Classification performance: AE, 20 rpm, load 5 kN

Component analysis	Kernel parameters		Error (Training/Testing), %		SV	
	RBF (C, γ)	Polynomial (d, C)	RBF	Polynomial	RBF	Polynomial
ICA	(128, 2)	(2, 1000)	14.6/11.2	7.1/18.3	42	35
PCA	(1, 0.125)	(2, 1000)	0.0/85.7	75.0/73.5	48	56
KICA	(16, 1)	(2, 1000)	31.3/75.6	1.7/75.5	44	54
KPCA	(32, 0.125)	(2, 1000)	79.7/85.7	48.2/83.7	48	53

7. Conclusions

In this study, fault diagnosis of low speed bearing based on classification method using SVM has been presented. The data used in the classification process are vibration and AE signals. Statistical features are calculated from time and frequency domain of each signal. SVM based multi-class classification is trained by statistical features with and without feature extraction. In this case, feature extraction is performed by component analysis via ICA, PCA, KICA and KPCA. The results show that ICA outperforms among others feature extraction technique for vibration and AE signals. Moreover, the comparison of classification performance shows that AE signal is better than vibration signal based on experimental work.

References

[1] Vapnik, V.N., 1999, *The Nature of Statistical Learning Theory*, Springer, New York.

[2] Jolliffe, I.J., 1986, *Principal Component Analysis*, Springer, New York.

[3] Hyvarinen, A., Orja, E., 2000, Independent component analysis: algorithm and applications, *Neural Networks* 13, 411-430.

[4] Scholkopf, B., Smola, A., Muller, K.R., 1998, Nonlinear component analysis as a kernel eigenvalue problem, Neural Computation, 10(5), 1299-1319.

[5] Heng, R. B. W., Nor, M. J. M., 1998, Statistical analysis of sound and vibration signals for monitoring rolling element bearing condition, *Applied Acoustic* 53, 211-226.

[6] Shi, D.F., Wang, W. J., Qu L. S., 2004, Defect detection for bearings using envelope spectra of wavelet transform, *Journal of Vibration and Acoustic*, *Trans. ASME* 126(4), 567-573.

[7] Kim Y. H., Tan A. C. C., Mathew, J., Kosse, V., Yang, B. S., 2007, A comparative study on the acoustic emission technique and acceleration measurement for low speed condition monitoring," *APVC 2007*.

[8] Kosse, V., Tan, A. C. C., 2006, Development of testing facilities for verification of machine condition monitoring for low speed machinery, *WCEAM 2006*, 181-186.

[9] Hsu, C.W., Lin, C.J., 2002, A comparison of methods for multiclass support vector machines, *IEEE Transaction on Neural Network*, 13(2), 415-425.

[10] Platt J., 1999, Fast training of support vector machines using sequential minimal optimization, in: B. Scholkopf, et al. *Advances in Kernel Methods-Support Vector Learning*. MIT Press, Cambridge.

[11] Hsu, C. W., Chang, C. C., Lin, C. J., 2005, *A practical guide to support vector classification*. Technical Report, Department of Computer Science and Information Engineering, National Taiwan University, Available at: http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf.