Sensors & Transducers, Vol. 179, Issue 9, September 2014, pp. 93-98

Sensors & Transducers

Discrimination the Roller Bearing Faults with Harmonic Wavelet Package, Kernel Principal Component Analysis and Relevance Vector Machine

[*] Tao Xu, Ailing Pei, Yong Liu
Shenyang Aerospace University, No. 37 Daoyi South Avenue, Daoyi Development District, Shenyang City, 110136, P. R. China
[*] Tel.: +86-024-89724448, fax: +86-024-89723871
[*] E-mail: xutao@sau.edu.cn

Received: 23 May 2014 /Accepted: 29 August 2014 /Published: 30 September 2014

Abstract: For the purpose of discriminating the roller bearing faults, this paper proposes a novel method with harmonic wavelet package, Kernel Principal Component Analysis (KPCA) and Relevance Vector Machine (RVM). Firstly, compute the vector energy with wavelet coefficients after the roller bearing vibration signals have been decomposed with harmonic wavelet package. The feature vectors are available after the vector energy has been standardized. Secondly, this paper uses several kernel principal components to extract nonlinear feature from the former vector according to the cumulative principle component contribution rate while greater than 90 %. Thirdly, a multiple classification model is proposed with RVM to discriminate the roller bearing faults. Finally, vibration data from Case Western Reserve University is used to test the proposed method. The experimental results show the proposed method reflects more roller bearing fault characteristics than the conventional method. Meanwhile, it can identify roller bearing states more accurately and efficiently. *Copyright © 2014 IFSA Publishing, S. L.*

Keywords: Discriminate fault, Roller bearing, Harmonic wavelet package, Kernel principal component analysis, Relevance vector machine.

1. Introduction

Vibration signals are widely used for the roller bearing condition monitoring. Comparing the signals of a machine in normal and faulty conditions, discriminating the roller bearing faults is possible. Because of the non-stationary characteristics of roller bearing vibration, many researchers study time-frequency technology in order to extract feature from vibration signals. Due to the adaptive performance of Hilbert-Huang transformation in time-frequency domain, Empirical Mode Decomposition (EMD) is used to process roller vibration signals to extract faulty feature [1-3]. Wavelet transform also possesses good performance for time-frequency analysis. Morphological wavelet is used to extract energy feature from roller bearing vibration signals [4]. Compared with wavelet transform, harmonic wavelet transform possesses better filter performance when decomposing signals in interested time-frequency domain. Meanwhile, the signals after transformation possesses the same resolution as the original signals, that overcome the limitation of Mallat algorithm in wavelet transform [4]. For harmonic wavelet analysis is effective to extract the singular components in the non-stationary signals, the time-frequency profile of harmonic wavelet is performed to analyze the signals of gear fault [5]. However, with the increase of harmonic wavelet decomposition layer, analysis band will gradually tend to lower frequency or higher frequency without arbitrarily selected band. Harmonic wavelet package overcomes the former limitation because it adaptively subdivides the whole band to extract interested frequency. Therefore, harmonic wavelet package is used to decompose roller bearing signals into different frequency domain and compute energy to prepare the feature vector [6, 7].

After the feature vector is available, a multiple classifier is necessary to discriminate the faults of roller bearing. For the nonlinear mapping performance of neural network, Back Propagation (BP) is used to propose the classification model to classify faults [6]. The fuzzy classifier is established to diagnosis roller bearing faults after the necessary rule set is obtained by intuition and domain knowledge [8, 9]. When a smaller number of samples are available, Support Vector Machine (SVM) possesses better nonlinear mapping performance. Based on decision tree architecture, a multiple SVMs classifier is designed with to diagnose roller bearing faults [3, 7]. For the purpose of efficiently classifying, Proximal Support Vector Machine (PSVM) is used to discriminate different bearing states [10, 11]. Relevance Vector Machine (RVM) converts nonlinear mapping in low-dimensional space into linear mapping in high-dimensional space just as SVM. Yet, its training process is implemented within Bays architecture and it removes the irrelevant points with Automation Relevance Decision (ARD) principle premising a priori parameters to achieve sparse model [12, 13]. RVM overcomes such limitations of SVM as error parameters determination, model sparse and kernel function satisfying Mercer condition. It possesses similar or better accuracy for regression prediction, classification and identification compared with SVM.

This paper proposes a novel discrimination method for the roller bearing faults diagnosis. This paper is organized as follow. Section 2 presents the scheme of discrimination method. Section 3 presents the process of feature extraction with harmonic wavelet package. Section 4 presents the process of feature reduction with KPCA. Section 5 presents the architecture of multi-classification model with RVM. To illustrate the proposed method above, experimental results are given in section 6. Finally, we draw the meaningful conclusion in section 7.

2. Scheme of Discrimination the Roller Bearing Faults

For the purpose of discriminate the roller bearing faults, this paper proposes a novel method with harmonic wavelet package, KPCA and RVM. Four roller bearing faults are considered including good bearing, bearing with inner race fault, bearing with outer race fault and bearing with roller fault. The scheme of the proposed method is shown in Fig. 1.

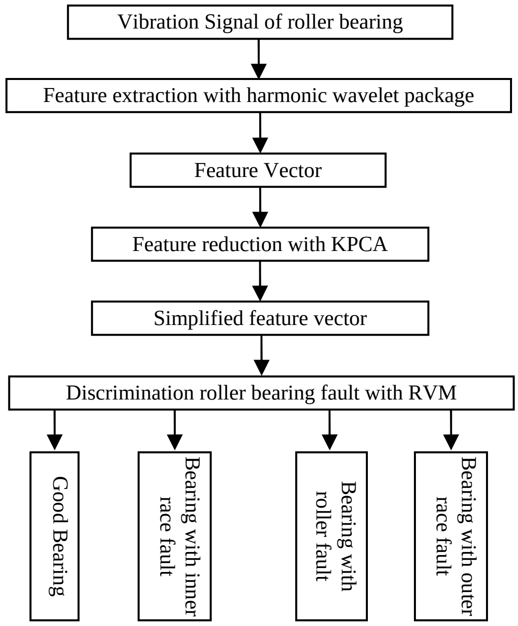

Fig. 1. Scheme of discrimination method.

Vibration signal is collected in order to represent the characteristics of the roller bearing state. After the vibration signal is processed by harmonic wavelet package, feature vector is achieved. Then, KPCA implements feature reduction from the former vectors to achieve the simplified feature vector. A RVM Classifier is proposed with decision tree architecture to discriminate good bearing, bearing with inner race fault, bearing with outer race fault and bearing with roller fault.

3. Feature Extraction with Harmonic Wavelet Package

From the spectrum of wavelet package, Newland proposes harmonic wavelet that is a novel wavelet formation approach [14]. Harmonic wavelet is a kind of complex spectrum wavelet with box-shape. Its expression in frequency domain is

$$\psi_{m,n}(x) = [\exp(in2\pi x) - \exp(im2\pi x)] \big/ [i2\pi(n-m)x], \quad (1)$$

where m, n determine the wavelet transform level just the same as j from 2^{-j} in the binary wavelet. Since harmonic wavelet is a complex wavelet with real part and imaginary part, it possesses good performances of locking phase and filter. The interested frequencies are reserved while other frequencies are shielded after transformation [15]. However, harmonic wavelet analysis can not choose interested frequency arbitrarily without infinite subdivision performance. Anyway, adaptive and infinite subdivision may be implemented based on the idea of binary wavelet package. Suppose interested frequency band is $B = 2^{-j} f_{\mathrm{h}}$, in which f_{h} is the high frequency and

$$\begin{cases} m = sB \\ n = (s+1)B \end{cases}, s = 0, 1, 2, \cdots, 2^j - 1, \tag{2}$$

Therefore, any decomposition layer can be subdivided to achieve the interested frequency band.

Conventional feature extraction method with harmonic wavelet package is to standardize the testing signal for the purpose of eliminating the effects of different variables that are described in equation (3).

$$\overline{X} = [X - sum(X)] / \sigma(X), \tag{3}$$

where the sum function computes the sum of X and σ indicates the standard deviation of X. However, conventional standardization method is not appropriate for vibration signal, whose sum may approach zero because of the contradiction between positive signal and negative signal. So, this paper first computes the sum of square of wavelet coefficients in each frequency domain. Then, compute the square root of the sum. After the square root has been standardized, the feature vector is prepared. The process of feature extraction method with harmonic wavelet package is described in detail as below.

1) Decompose vibration signal into multiple scales with harmonic wavelet package to achieve coefficients of each scale.

2) Compute the energy from wavelet coefficients in each scale with equation (4).

$$E_{H_{N,i}} = \sqrt{\int \left| H_{N,i} \right|^2 dt} = \sqrt{\sum_{j=1}^{M} \left| H_{N,i,j} \right|^2}, \tag{4}$$

where N indicates the number of frequency bands and M indicates the number of wavelet coefficients in each band after harmonic wavelet decomposition.

3) Standardize the energy of wavelet coefficients with equation (5).

$$\overline{E_{H_{N,i}}} = \left[E_{H_{N,i}} - \mathrm{mean}(E_{H_{N,i}}) \right] / D_{\sigma}(E_{H_{N,i}}), \tag{5}$$

where mean function computes the average energy of each band and D_{σ} indicates the standard deviation of the energy of each band.

4) The standard fault feature is described as equation (6).

$$\overline{E_{H_N}} = \left[\overline{E_{H_{N,1}}} \quad \overline{E_{H_{N,2}}} \quad \cdots \quad \overline{E_{H_{N,M}}} \right], \tag{6}$$

4. Feature Reduction with KPCA

Principal Component Analysis (PCA) is a conventional dimension reduction method to deal with linear data. Meanwhile it is unable to deal with nonlinear data. Yet, PCA can be extended to nonlinear space with the projection of nonlinear kernel function. Scholkopf proposes Kernel Principal Component Analysis (KPCA) with the aid of the SVM theory to solve nonlinear problem [16]. With KPCA, those linear inseparable data can be projected to the high dimension space in which it is linear separable for them. Based on PCA, KPCA implements the nonlinear data mapping to the higher dimensional feature space with kernel function and it is convenient to extract better feature in the high dimensional space [17, 18].

By the nonlinear mapping function Φ, the time series of data $x_i (i = 1, ..., M)$ in original space R are projected to a new high dimensional feature space F, in which x_i are denoted by $\phi(x_i)$. The process of KPCA for feature reduction is described in Fig. 2.

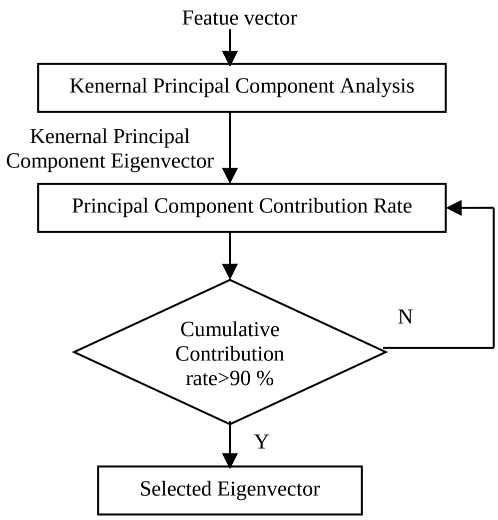

Fig. 2. Process of KPCA for feature reduction.

The process of feature reduction is described in detail as follow.

1) Compute the kernel matrix from feature vector matrix.

2) Centralize the kernel matrix.

3) Decompose the centralized matrix to achieve standardized eigenvectors and determine the number of kernel principal components.

4) Compute the projection of the eigenvector in the feature space.

5. Discrimination Roller Bearing Faults with RVM

Relevance Vector Machine (RVM) is proposed by M. E. Tipping based on Sparse Bayesian learning theory, which ensures sparse performance by assigning eigenvector with zero mean Gaussian prior distribution after introducing ultra-parameter [12, 13]. Meanwhile, estimate ultra-parameter with the maximizing edge likelihood function and automatically adjust rule coefficients in the process of ultra-parameter estimation [19]. Besides the good performances as Support Vector Machine (SVM), RVM overcomes intrinsic limitation of SVM. With computation efficiency, its kernel function is not limited by Mercer condition [20, 21]. Based on the idea of Multi-classification Support Vector Machine, this paper proposes three models with RVM to implement multi-faults discriminating for roller bearing. This paper considers four bearing states including good bearing, bearing with inner race fault, bearing with roller fault and bearing with outer race fault.

Just as the model with SVM in literature [7], this paper proposes decision tree model with RVM. Suppose the classification number is k, the scheme of decision tree is to establish k-1 two-classification models. The target output of i^{th} classification model is 0 for the i^{th} samples, is 1 for the rest $(i+1)^{th}$ samples, $(i+2)^{th}$ samples, …, k^{th} samples. And the target output of $(i+1)^{th}$ classification model is 0 for the $(i+2)^{th}$ samples, is 1 for the rest $(i+3)^{th}$ samples, $(i+4)^{th}$ samples, …, k^{th} samples. After the process of one two-classification RVM model, the number of classes decreases until all classes have been classified. With RVM this paper proposes roller bearing discrimination method with decision tree architecture which is shown in Fig. 3.

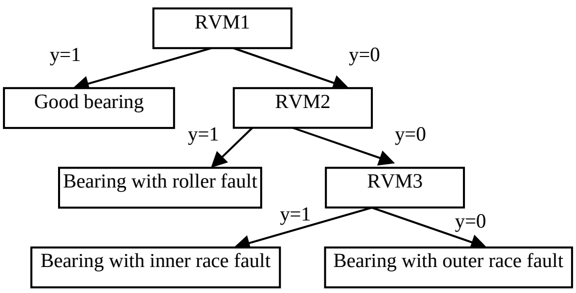

Fig. 3. Decision tree classification architecture with RVM.

In the process of classification, decision tree model discriminates good bearing firstly because it is easy to classify the good bearing and faulty bearing. Then, it discriminates bearing with roller fault which is different from race faults. Finally, it discriminates bearing with inner race fault and bearing with outer race fault. Anyway, if one RVM misclassifies bearing, it will not be corrected, that affects the accuracy of decision tree model.

6. Experimental Illustration

To verify the effectiveness of the discrimination method for rolling bearing fault diagnosis, the test data is acquired from electrical engineering lab of rolling bearing fault simulation in Case Western Reserve University [22]. The test stand is shown in Fig. 4.

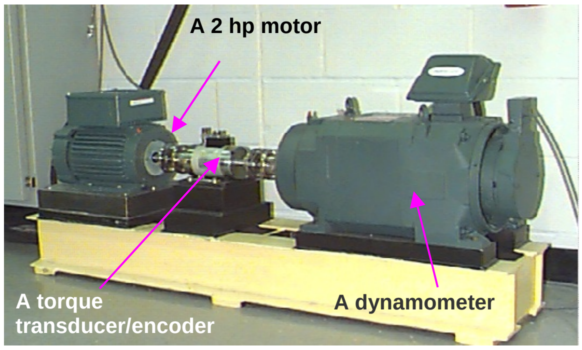

Fig. 4. Test stand of roller bearing.

The test bearings support the motor shaft. Single point faults were introduced to the test bearings using electro-discharge machining with fault diameters of 7 mils, 14 mils, 21 mils, 28 mils, and 40 mils (1 mil=0.001 inches). SKF bearings were used for the 7, 14 and 21 mils diameter faults, and NTN equivalent bearings were used for the 28 mil and 40 mil faults. Vibration data was collected using accelerometers, which were attached to the housing with magnetic bases. Accelerometers were placed at the 12 o'clock position at both the drive end and fan end of the motor housing.

During some experiments, an accelerometer was attached to the motor supporting base plate as well. Vibration signals were collected using a 16 channel DAT recorder, and were post processed in a Matlab environment. All data files are in Matlab (*.mat) format. Digital data was collected at 12,000 samples per second, and data was also collected at 48,000 samples per second for drive end bearing faults. Speed and horsepower data were collected using the torque transducer/encoder.

To illustrate the advantages of the proposed feature extraction and reduction method with harmonic wavelet package and KPCA, this paper compares it with conventional wavelet packet and KPCA. Firstly, decompose the vibration data with harmonic wavelet packet and the wavelet packet respectively. Secondly, extract the eigenvector, implement kernel principal component analysis and compute its principal component contribution rate. Contribution rates of the former seven principal components are shown in Table 1.

Table 1. Contribution rates of principle component.

Eigenvalue number	Wavelet packet-eigenvector λ	Wavelet packet-eigenvector Principal component contribution rate (%)	Wavelet packet-eigenvector Principal component cumulative contribution rate /%	Harmonic wavelet packet-eigenvector λ	Harmonic wavelet packet-eigenvector Principal component contribution rate (%)	Harmonic wavelet packet-eigenvector Principal component cumulative contribution rate (%)
1	8.6511	72.05	72.05	18.306	78.14	78.14
2	1.3106	10.92	82.97	3.3144	14.15	92.29
3	0.8497	7.08	90.05	0.9923	4.24	96.53
4	0.5443	4.53	94.58	0.3743	1.60	98.13
5	0.3034	2.53	97.11	0.2410	1.02	99.16
6	0.1187	0.99	98.10	0.1227	0.52	99.68
7	0.1027	0.86	98.96	0.0754	0.32	100

Table 1 shows that both harmonic wavelet package and wavelet package possess good performance in the process of feature compression with kernel principal component analysis. The first principal component contribution rate exceeds 70 %, the former three principal components cumulative contribution rate exceeds 90 % and the former five principal components cumulative contribution rate exceeds 95 %. Meanwhile, the principal component contribution rate with harmonic wavelet package exceeds the principal component contribution rate with wavelet packet respectively. So, harmonic wavelet package possesses better performance when feature compressing. In order to ensure the good dimension reduction effect and minimal information loss in the process of feature extraction, feature dimension should be large than 3 by experiences. With the proposed feature extraction and reduction method, 200 groups are collected including 50 groups with normal state, 50 groups with inner race fault, 50 groups with roller fault and 50 groups with outer race fault. After these groups have been used to train the MRVM (Multiple Relevance Vector Machine) classifying model that is established with decision tree architecture, the model possesses the capability to discriminate different roller bearing state. Then, collect other 200 groups to test the classifying capability of the model. Meanwhile, we use the same process to verify the method with wavelet package and KPCA. This paper studies four different working states. The fault diameter range includes 0.007 inch, 0.014 inch and 0.021 inch. The rolling speed range includes 1730 r/min, 1750 r/min and 1797 r/min. The accuracy and elapsed time of two methods are compared in Table 2.

Table 2. Comparison results of two methods.

Working state	Wavelet-KPCA-MRVM				Harmonic wavelet-KPCA-MRVM			
	Principal component cumulative contribution rate /%	Principal component number	Accuracy/ %	Elapsed time/s	Principal component cumulative contribution rate /%	Principal component number	Accuracy/ %	Elapsed time/s
0.007 inch 1730 r/min	98	6	**99.5**	**0.328**	97	4	**100**	**0.250**
0.007 inch 1797 r/min	98	6	**100**	**0.282**	98	4	**100**	**0.281**
0.014 inch 1730 r/min	98	9	**99.5**	**0.312**	98	5	**100**	**0.297**
0.021 inch 1750 r/min	95	4	**85.5**	**0.297**	95	3	**95.5**	**0.297**

Table 2 shows that the corresponding fault states with different sizes and different speeds in this paper. The proposed method possesses higher diagnostic accuracy, less elapsed time and few kernel components than the method with wavelet package and KPCA.

7. Conclusions

This paper combines harmonic wavelet packet and KPCA for the purpose of feature extraction and reduction from the rolling bearing vibration signal. Harmonic wavelet packet possesses outstanding performance when extracting feature from time-frequency domain. KPCA reduces dimension of principal components and excludes those components with little contribution. So, the sensitive features with fault characteristics are left for discriminating faults of rolling bearing more efficiently. The classification model with RVM is established to discriminate roller bearing faults. Based on the decision tree architecture, four bearing states can be discriminated with feature vector. After the experimental data is acquired from the rolling bearing test stand in Case Western Reserve University, the feature extraction and discrimination roller bearing faults of the proposed method are verified. Experimental result shows that the proposed method extracts feature more accurately and effectively than the method with wavelet packet and KPCA. Meanwhile, classification with RVM can discriminate roller bearing state more accurately and efficiently.

Acknowledgements

Thanks for the support from Aviation Science Fund (2012ZD54013) and those provide help to this research.

References

[1]. Dejie Yu, Junsheng Cheng, Yu Yang, Application of EMD method and Hilbert spectrum to the fault diagnosis of roller bearings, *Mechanical Systems and Signal Processing*, Vol. 19, Issue 2, 2005, pp. 259-270.

[2]. Cheng Junsheng, Yu Dijie, Yang Yu, A fault diagnosis approach for roller bearings based on EMD method and AR model, *Mechanical Systems and Signal Processing*, Vol. 20, Issue 2, 2006, pp. 350-362.

[3]. Yu Yang, Dejie Yu, Junsheng Cheng, A fault diagnosis approach for roller bearing based on IMF envelope spectrum and SVM, *Measurement*, Vol. 40, Issue 9-10, 2007, pp. 943-950.

[4]. Xing Han, Jingqi Xiong, Rui Sun, et al, Research on the roller bearing fault diagnosis based on morphological wavelet and LSSVM algorithm, in *Proceeding of the International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering*, Mingshan City, China, 15-18 July 2013, pp. 1888-1892.

[5]. Wang Zhigang, Li Yourong, Li Fang, Fault diagnosis method based on harmonic wavelet analysis, *Journal of Vibration and Shock*, Vol. 25, Issue 2, 2006, pp. 125-128.

[6]. Zhao Yuan-Xi, Xu Yong-Gang, Gao Li-Xin, et al, Fault pattern recognition technique for roller bearing acoustic emission based on harmonic wavelet packet and BP neural network, *Journal of Vibration and Shock*, Vol. 29, Issue 10, 2010, pp. 162-165.

[7]. Yu Jintao, Ding Mingli, Meng Fangang, et al, Acoustic emission source identification based on harmonic wavelet packet and support vector machine, *Journal of Southeast University* (*English Edition*), Vol. 27, Issue 3, 2011, pp. 300-304.

[8]. V. Sugumaran, K. I. Ramachandran, Automatic rule learning using decision tree for fuzzy classifier in fault diagnosis of roller bearing, *Mechanical Systems and Signal Processing*, Vol. 21, Issue 5, 2007, pp. 2237-2247.

[9]. V. Sugumaran, K. I. Ramachandran, Fault diagnosis of roller bearing using fuzzy classifier and histogram features with focus on automatic rule learning, *Expert Systems with Applications*, Vol. 38, Issue 5, 2011, pp. 4901-4907.

[10]. V. Sugumaran, V. Muralidharan, K. I. Ramachandran, Feature selection using decision tree and classification through proximal support vector machine for fault diagnostics of roller bearing. *Mechanical Systems and Signal Processing*, Vol. 21, Issue 2, 2007, pp. 930-942.

[11]. V. Sugumaran, G. R. Sabareesh, K. I. Ramachandran, Fault diagnostics of roller bearing using kernel based neighborhood score multi-class support vector machine, *Expert Systems with Applications*, Vol. 34, Issue 4, 2008, pp. 3090-3098.

[12]. M. E. Tipping, Sparse Bayesian learning an the relevance vector machine, *Journal of Machine Learning Research*, Vol. 1, Issue 3, 2001, pp. 211-244.

[13]. M. E. Tipping, A. C. Faul, Fast marginal likelihood maximisation for sparse Bayesian models, in *Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics*, Key West, Florida, America, 3-6 January 2003, pp. 1-13.

[14]. D. E. Newland, Harmonic wavelet analysis, in *Proceeding of Royal Society*, London, Great Britain, Vol. 443, 1993, pp. 203-225.

[15]. D. E. Newland, Wavelet analysis of vibration, part 1: Theory, *Journal of Vibration and Acoustics*, Vol. 116, Issue 4, 1994, pp. 409-416.

[16]. B. Scholkopf, A. Smola, K. R. Muller, Nonlinear component analysis as a kernel eigenvalue problem, *Neural Computation*, Vol. 10, Issue 5, 1998, pp. 1299-1319.

[17]. Yong Xu, David Zhang, Fengxi Song, et al, A method for speeding up feature extraction based on KPCA, in *Proceedings of the International Conference on Intelligent Computing*, Hefei, China, 23-26 August 2005, pp. 1056-1061.

[18]. Zhao Xiaoqiang, Xue Yongfei, Yang Wu, An improved KPCA algorithm of chemical process fault diagnosis based on RVM, in *Proceedings of the 32nd Chinese Control Conference*, Xi'an, China, 26-28 July 2013, pp. 6083-6087.

[19]. J. C. Mackay, The evidence framework applied to classification networks, *Neural Computation*, Vol. 4, Issue 5, 1992, pp. 720-736.

[20]. Begüm Demir, Sarp Ertürk, Hyperspectral image classification using relevance vector machines, *IEEE Geoscience and Remote Sensing Letters*, Vol. 4, Issue 4, 2007, pp. 586-590.

[21]. B. Gholami, W. M. Haddad, A. R. Tannenbaum, Relevance vector machine learning for neonate pain intensity assessment using digital imaging, *IEEE Transactions on Biomedical Engineering*, Vol. 57, Issue 6, 2010, pp. 1457-1466.

[22]. Seeded fault test data, Bearing Data Center, The Case Western Reserve University, (http://csegroups.case.edu/bearingdatacenter/pages/download-data-file).