

1280. Health assessment of rotary machinery based on integrated feature selection and Gaussian mixed model

Zhipeng Wang¹, Chen Lu², Zili Wang³, Jian Ma⁴

School of Reliability and Systems Engineering, Beihang University, Beijing, 100191, China

Science and Technology on Reliability and Environmental Engineering Laboratory, Beijing, 100191, China

²Corresponding author

E-mail: ¹wpw104@163.com, ²luchen@buaa.edu.cn, ³wzlj@buaa.edu.cn, ⁴majian3128@126.com

(Received 5 February 2014; received in revised form 7 April 2014; accepted 13 April 2014)

Abstract. Bearing failure is the most common failure mode of all rotary machinery failures, and can interrupt the production in a plant causing unscheduled downtime and production losses. A bearing failure also has the potential to damage machinery causing soaring machinery repair and/or replacement costs. In order to prevent unexpected bearing failure, a health assessment method is proposed in this paper. It employs an integrated feature selection approach and Gaussian mixture model (GMM). Firstly, the integrated feature selection approach, which combines empirical mode decomposition (EMD), singular value decomposition (SVD) and Principal Component Analysis (PCA), processes nonlinear and non-stationary vibration signals of a bearing and extracts features for health assessment. Then, GMM is utilized to evaluate and track the health degradation of the bearing in terms of confidence values (CV). This method, which is notable for bearing health tracking and detect the defect at its incipient stage, can be used without the need for failure datasets in applications. Finally, the feasibility and efficiency of this method was validated by two datasets of different bearing experiments.

Keywords: health assessment, empirical mode decomposition, singular value decomposition, principal component analysis, Gaussian mixture model.

1. Introduction

Currently, traditional maintenance strategy is either reactive or blindly preventive. Both scenarios are extremely wasteful. In order to reduce the maintenance cost, the innovation from reactive maintenance to predictive maintenance has been driven by the technology development from diagnostics to prognostics. That is a transformation from the traditional fail and fix (FAF) maintenance practices to predict and prevent (PAP) paradigm [1].

Bearings are common and vital elements in rotating machinery. With the exception of abrupt catastrophic failures, most of the faults have progression processes to failure. By tracking the health degradation of a bearing, unscheduled machinery outages and costly damage caused by a bearing failure can be avoided. Considering the complexity of rotating machinery and the severe working conditions, the data-driven methods are widely applied for online condition monitoring. Nowadays, a variety of data-driven methods exist, such as: support vector machine [1], fuzzy logic [2], wavelet packet [3], proportional hazard model [4], and artificial neural networks [5-7]. Research on health assessment of bearing has attracted much attention. However, most current health assessment methods require fault data to extract the feature conditions, but such data is rare in real situations [8, 9]. Hence, accurately tracking the health condition of bearings using normal condition data is a hot topic in condition-based maintenance. To this end, this paper proposes a method based on Gaussian mixture model (GMM).

Bayesian inference-based GMM is a data-driven method that is highly efficient in monitoring non-Gaussian processes that consist of multiple modes and significant multi-Gaussianity in monitoring data [10]. Thus, the GMM-based health assessment method is notable for bearing health tracking and detect the defect at its incipient stage. Moreover, GMM can be used without fault datasets in applications.

To employ GMM as a health assessment algorithm involves a significant process: extracting the suitable features from the raw data. These features will be the inputs of the GMM. To ensure

the efficiency of MTS, it is important that this feature extraction is performed appropriately – otherwise, we will suffer from the well-known “garbage in, garbage out” problem. Therefore, feature extraction needs to be investigated.

Traditionally, the Fourier transform (FT) has been the dominant analysis tool for extracting the features of stationary signals, as it could produce the statistical average of characteristics over the duration of the data. However, FTs are not suitable for nonlinear and non-stationary signals. Unfortunately, bearing vibration signals are both nonlinear and non-stationary. In the most recent studies, time-frequency analysis methods have been used to study the rotating machinery vibration signal [11-14]. Over the past decades, the wavelet transform has been a fast-evolving mathematical and signal processing tool. The basic wavelet transform involves the operations of dilation and translation, which lead to a multi-scale analysis of the signal. Wavelet transforms are capable of analyzing non-stationary signals, but they still have some inevitable deficiencies [15]. One of these is that the commonly used wavelet analysis is essentially a kind of FT with adjustable windows, and thus it suffers many of the shortcomings of Fourier spectral analysis [16]. Another disadvantage of wavelet analysis is its non-adaptive nature. Once the wavelet function is selected, it must be used to analyze all the data [15, 17]. Therefore, a self-adaptive method for nonlinear and non-stationary signals is needed to avoid the disadvantages of wavelet transform.

In 1998, Huang et al. proposed the empirical mode decomposition (EMD), which uses local time and scale characteristics of the signal to decompose it into a number of intrinsic mode functions (IMFs). IMFs not only relate to the sampling frequency, but also change within the signal itself, so the method is self-adaptive in decomposing nonlinear and non-stationary signals [17]. Nowadays, it has been broadly used in fault diagnosis [16, 18, 19]. Therefore, EMD can be used for feature extraction in this paper.

Meanwhile, considering that the vibration signal usually exhibits similar features to a periodic impulse when a local fault appears in the bearing, matrix singular value decomposition (SVD) techniques can be used to extract features [18]. According to the matrix theory, singular values are the intrinsic characteristics of a matrix with favorable stability, invariant ratio, and rotation. SVD can be used as a tool for signal regularization, noise reduction, signal detection, and estimation [20]. However, the construction of the initial matrix remains a drawback in the implementation of SVD. Generally, a form of phase space reconstruction from chaos theory is employed [18, 21]. But, this requires reconstruction parameters, such as the delay time and embedding dimension, to be determined, which will involve extensive computations. Thus, this construction method is unavailable in application. The EMD can decompose a signal into a number of IMFs that can be used to construct the original matrix for SVD. Additionally, Principal Component Analysis (PCA) has been extensively applied for data analysis. It is a simple, non-parametric method of extracting relevant information from confusing data sets. Therefore, the current paper proposed an integrated feature selection approach combined with EMD, SVD and PCA to process nonlinear and non-stationary signals.

This paper is organized as follows: Section 2 introduces EMD, SVD, PCA and GMM; Section 3 describes the experimental studies performed to verify this method; and Section 4 gives the conclusions and some ideas for future work.

2. Methodology

2.1. Empirical mode decomposition

The method of EMD is developed from the assumption that any signal consists of different simple intrinsic modes of oscillation. Each signal can be decomposed into a number of IMFs, each of which must satisfy the following definition [17]: (1) In the whole data set, the number of extreme crossings should either be equal or differ by at most one to the number of zero crossings; (2) At any point, zero is the mean value of the envelope defined by the local maxima and the envelope defined by the local minima.

An IMF represents a simple oscillatory mode compared with the simple harmonic function. With the definition, any signal $x(t)$ can be decomposed as follows [18]:

Step 1: Identify all the local extrema, and then connect all the local maxima by a cubic spline line as the upper envelope.

Step 2: Repeat the procedure for the local minima to produce the lower envelope. The upper and lower envelopes should cover all the data between them.

Step 3: The mean of upper and low envelope value is designated as $m_1(t)$; and the difference between the signal $x(t)$ and $m_1(t)$ is the first component, $h_1(t)$, i.e.:

$$x(t) - m_1(t) = h_1(t). \quad (1)$$

Ideally, if $h_1(t)$ is an IMF, then $h_1(t)$ is the first component of $x(t)$.

Step 4: If $h_1(t)$ is not an IMF, $h_1(t)$ is treated as the original signal and repeat step 1, 2, 3; then:

$$h_1(t) - m_{11}(t) = h_{11}(t). \quad (2)$$

After repeated sifting, i.e. up to k times, $h_{1k}(t)$ becomes an IMF, that is:

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

Then, it is designated as:

$$c_1(t) = h_{1k}(t). \quad (4)$$

The first IMF component from the original data $c_1(t)$ should contain the finest scale or the shortest period component of the signal.

Step 5: Separate $c_1(t)$ from $x(t)$, we could get:

$$r_1(t) = x(t) - c_1(t), \quad (5)$$

$r_1(t)$ is treated as the original data and repeat the above processes; therefore, the second IMF component $c_2(t)$ of $x(t)$ could be got. Let's repeat the process as described above for n times, then n -IMFs of signal $x(t)$ could be got. Then:

$$\begin{aligned} r_1(t) - c_2(t) &= r_2(t), \\ \vdots \\ r_{n-1}(t) - c_n(t) &= r_n(t). \end{aligned} \quad (6)$$

The decomposition process can be stopped when $r_n(t)$ becomes a monotonic function from which no more IMFs can be extracted.

With the definition, any signal can be decomposed [19]. We finally obtain:

$$x(t) = \sum_{j=1}^n c_j(t) + r_n(t). \quad (7)$$

Thus, one can achieve a decomposition of the signal into n -empirical modes and a residue $r_n(t)$, which is the mean trend of $x(t)$.

EMD decomposes signal into several IMFs one by one iteratively. The new IMF $c_n(t)$ is decomposed from the previous residue $r_{n-1}(t)$. Each IMF $c_1(t)$, $c_2(t)$, ..., $c_n(t)$ contains lower-frequency oscillations than the prior-extracted one. With the higher frequency parts are separated from signal $x(t)$, each residue $r_1(t)$, $r_2(t)$, ..., $r_n(t)$ is more smooth than the prior one.

Until $r_n(t)$ becomes a monotonic function from which no more IMFs can be extracted, the process of EMD can be stopped. Therefore, the final residue $r_n(t)$ represents the central tendency of signal $x(t)$, and is usually a monotonic trend rather than white noise.

EMD has been proved as a fast, effective, self-adaptive method for nonlinear and non-stationary time series analysis, but there is still a drawback, which is worth noting: the end effect, appearing distortion at the end of signal in the decomposition process. In this paper, the mirror periodic extending method is employed to solve this problem.

2.2. Singular value decomposition

SVD is a powerful and effective feature extraction tool in linear algebra, and has been used for fault diagnosis in rotating machinery [21]. The SVD is defined as follows:

X denotes an $m \times n$ matrix with $m \geq n$. According to the SVD theorem, the matrix X can be decomposed in the form given by [22]:

$$X = U\sigma V^T, \quad (8)$$

where $U(m \times m)$ and $V(n \times n)$ are orthogonal matrices, and σ is an $m \times n$ diagonal matrix of singular values ($\sigma_{ij} \neq 0$, if $i = j$ and $\sigma_{11} \geq \sigma_{22} \geq \dots \geq 0$). The columns of the orthogonal matrix U and V are called the left and right singular vectors, respectively. An important property of U and V is that they are mutually orthogonal. The singular values represent the importance of individual singular vector in the matrix composition. In other words, a singular vector that corresponds to larger singular values contains more information about the structure of the pattern embedded in the matrix than the other singular vector.

2.3. Principal component analysis

Principal Component Analysis (PCA) has been called one of the most valuable results from applied linear algebra [5]. PCA is used abundantly in several forms of analysis, because it is a simple, non-parametric method of extracting relevant information from confusing data sets. With minimal additional effort PCA provides a roadmap for how to reduce a complex data set to a lower dimension to reveal the sometimes hidden, simplified structure that often underlie it [8]. It is a way of identifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences. Since patterns in data can be hard to find in data of high dimension, where the luxury of graphical representation is not available, PCA is a powerful tool for analyzing data.

The purpose of PCA is to identify the structure behind a multivariable stochastic observation in order to obtain a compact description of it. Clusters are more likely to be distinguished by projection in a high-variance direction than in a low variance one. Hence, the dimensionality reduction in the PCA could retain most of the intrinsic information in the data [9]. In PCA, the k th principal component direction is along an eigen vector direction belonging to the k th largest eigen value of the covariance matrix of the input data. An advantage in using PCA to preprocessing data is a reduction in data dimension that eliminates redundant information and allows simplifying the classification process [10].

2.4. Health assessment based on the Gaussian mixed model

This paper regards health assessment as evaluating the overlap between the current observed features and those observed during the normal condition. This overlap adopts Confidence Values (CVs) as an index for evaluating the severity of a health condition. This index was introduced by Jay Lee [9, 23] and has been accepted by some researchers [24-26]. The Confidence Value lies between 0 and 1, with 1 being the normal state. In this paper, the GMM is employed to assess the

health state by using only normal condition datasets.

The GMM, as shown in Fig. 1, is a statistical model that uses a weighted sum of probability density functions of multiple Gaussian distributions to depict the distribution of a vector in the probability space [24]. This method can theoretically be applied to approximate an arbitrary feature distribution within an arbitrary accuracy [23].

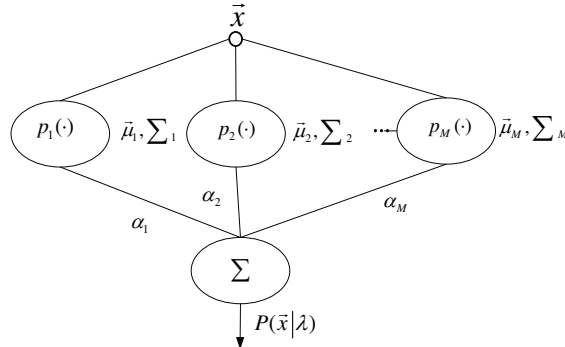


Fig. 1. Structure of a GMM with M mixture components

The mathematic model of GMM is parameterized by mixture weights α , mean vectors $\vec{\mu}$, and the covariance Σ from all M mixture components as follows:

$$\lambda = \left\{ \alpha_i, \vec{\mu}_i, \Sigma_i; i = 1, 2, \dots, M \right\}. \quad (9)$$

GMM can be given as:

$$P(x) = \sum_{i=1}^M \alpha_i p(x|\theta_i), \quad (10)$$

$$p(x|\theta_i) = \frac{1}{(2\pi)^{d/2} |\sigma_i|^{1/2}} \exp\left(-\frac{1}{2} (x - \vec{\mu}_i)' \sigma_i^{-1} (x - \vec{\mu}_i)\right), \quad (11)$$

where $p(x|\theta_i)$ is a single Gaussian function, x is the d -dimensional feature vector, M is the number of mixtures, and θ_i represents the parameters of the i -th Gaussian function, including mean vector $\vec{\mu}_i$ and covariance matrix σ_i .

For health assessment, given the training feature vectors from normal machine conditions, the goal of model training is to estimate the parameters of the GMM built under normal machine conditions [24]. Several methods can be used to estimate these parameters. Maximum likelihood (ML) estimation, which is the most popular and well-established method, is used in the current work. The ML parameters can be estimated iteratively by using the expectation maximization algorithm. The main steps as follows [25-28]:

1) (E step) Using Bayes' theorem, calculate the posterior probability of model i , given a data vector x_n and current model parameter θ :

$$p(i|x_n, \theta) = \frac{p_i p(x_n|\theta_i)}{\sum_{j=1}^M p_j p(x_n|\theta_j)}. \quad (12)$$

2) (M step) Maximum likelihood (ML) re-estimation of model parameters is given by:

$$\hat{p}_i = \frac{1}{N} \sum_{n=1}^N p(i|x_n, \theta), \tag{13}$$

$$\hat{\mu}_i = \frac{\sum_{n=1}^N p(i|x_n, \theta)x_n}{\sum_{n=1}^N p(i|x_n, \theta)}, \tag{14}$$

$$\hat{\sigma}_i = \frac{\sum_{n=1}^N p(i|x_n, \theta) (x - \vec{\mu}_i)'(x - \vec{\mu}_i)}{\sum_{n=1}^N p(i|x_n, \theta)}. \tag{15}$$

Repeating the E and M steps, and the calculation converges to a stable solution that represents the maximum likelihood solution of the problem, thus getting converged mean, covariance matrix and prior vectors.

When using GMM to approximate the feature distribution, how to properly choose the number of the mixture is critical, or the problems of underfitting overfitting will arise. Bayesian Information Criterion (BIC) is introduced to address this issue [26].

BIC defines a log-likelihood function and penalty term, and its mathematical description given GMM is depicted as follows:

$$BIC(H_j|D) = -2 \times \ln(P(D|\hat{\theta}_j, H_j)) + k \times \ln(n), \tag{16}$$

where H_j is the j th candidate model, D is the training feature, $\ln(P(D|\hat{\theta}_j, H_j))$ is the maximized log-likelihood of the j th candidate model, k is the number of the parameters to be estimated and n is the size of the feature. The candidate GMM with the lowest BIC score can be deemed as the best model.

After establishing the GMMs for normal condition and monitoring condition separately, the Confidence Value (CV) can be calculated by the overlap of the GMMs as follows:

$$CV = overlap = \frac{\int g_1(x)g_2(x)dx}{\sqrt{\int (g_1(x))^2 dx} \sqrt{\int (g_2(x))^2 dx}}, \tag{17}$$

where $g_1(x)$ and $g_2(x)$ are GMMs that represent the normal and current monitoring conditions, respectively. This paper uses only normal condition datasets for overlap calculation, where each overlap value between the clustered normal dataset and online measurement data can be considered Confidence Value (CV), as well as for tracking health degradation status.

3. Experimental results

3.1. Experimental setup

In this paper, two datasets from different bearing test rigs were used to verify the proposed method. The first dataset was generated from a test rig in the Case Western Reserve University Bearing Data Center (<http://csegroups.case.edu/bearingdatacenter/pages/download-data-file>). In this experiment, as Table 1 shows, 6205-2RS JEM SKF deep-groove ball bearings were tested. Three common types of bearing faults were tested: inner race fault, outer race fault, and rolling element fault. Under four states, including the normal condition, the vibration signal was acquired with a stabilized motor speed of 1797 r/min and a sampling rate of 12 kHz.

The second dataset was generated from bearing run-to-failure tests performed under a constant load condition by the National Science Foundation Industry and University Cooperative Research Program (NSF I/UCR) Center on Intelligent Maintenance Systems (IMS) [29]. As Table 2 shows,

Rexnord ZA-2115 double row bearing was selected in the run-to-failure test. The rotation speed was kept constant at 2000 rpm, and 6000 lbs (2730 kg) of radial load was placed onto the shaft and bearing via a spring mechanism [29]. Vibration data were collected every 20 min at a sampling rate of 20 kHz. The test was continued until an inner race fault in the bearing was found. Both datasets were analyzed to verify the proposed method.

Table 1. Data of 1# bearing experiment

Type	Rexnord ZA-2115 double row
Failure mode	Inner race fault
	Outer race fault
	Rolling element fault

Table 2. Data of 2# bearing experiment

Subject	Bearing
Type	Rexnord ZA-2115 double row
Failure mode	Inner race fault

3.2. Feature extraction

Firstly, the vibration signal from two datasets was decomposed into n -empirical modes that contain lower-frequency oscillations than the previously extracted ones and a residue $r_n(t)$, as shown in Fig. 2. Information on bearing health condition is considered in the high-frequency bands. Therefore, in this paper, n was designated as 7. The first seven IMFs and a residue $r_7(t)$ were extracted to construct the feature matrix for the following SVD process. Then, the constructed feature matrix was decomposed to obtain an 8-dimensional feature vector of singular values. After that, PCA was employed to reduce a complex data set to a lower dimension to reveal the hidden, simplified information. In this paper, 8 dimensions were reduced to 2 dimensions.

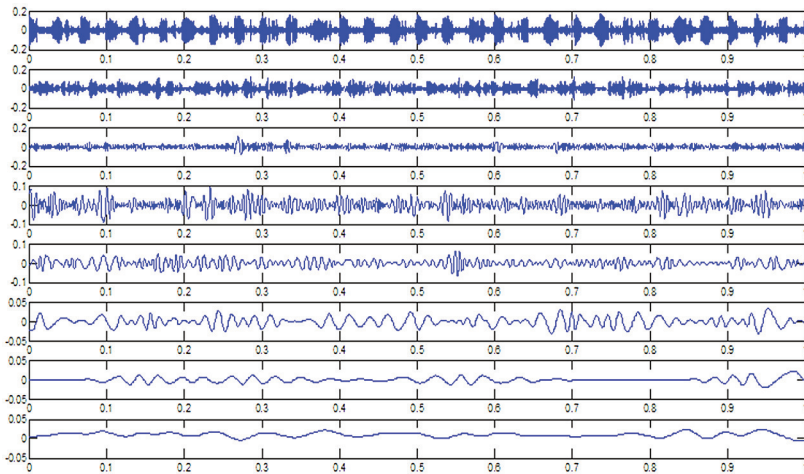


Fig. 2. A sample of EMD decomposition results of normal condition datasets

3.3. Health assessment by using only normal condition data

To verify this proposed method, both datasets from different test rigs were used. From the first dataset, a sample of normal condition data were employed to construct $g_1(x)$ in Eq. (17) and train the GMM model. Then, another 4 samples of normal condition data and 2 samples of data acquired under each fault mode were used to test the trained GMM model.

As shown in Table 3, Confidence Values (CVs) of normal condition from the first dataset were near 1, while Confidence Values (CVs) of all fault modes were close to 0. This result verified that the GMM is very notable for health assessment.

From the second dataset, with the first 400 min collected datasets considered normal condition data, the Confidence Values (CVs) from the beginning to the end of the life testing of the bearing were calculated by using Eq. (17), as shown in Fig. 3.

Table 3. Confidence Values (CVs) of the first dataset

Type	No.	CVs	Means of CVs
Normal condition	1	0.986	0.9445
	2	0.965	
	3	0.873	
	4	0.954	
Inner race fault	1	$3.089e^{-143}$	$1.01e^{-140}$
	2	$2.018e^{-140}$	
Outer race fault	1	$1.040e^{-64}$	$0.65e^{-60}$
	2	$1.291e^{-60}$	
Rolling element fault	1	$2.191e^{-33}$	$1.09e^{-33}$
	2	$4.214e^{-35}$	

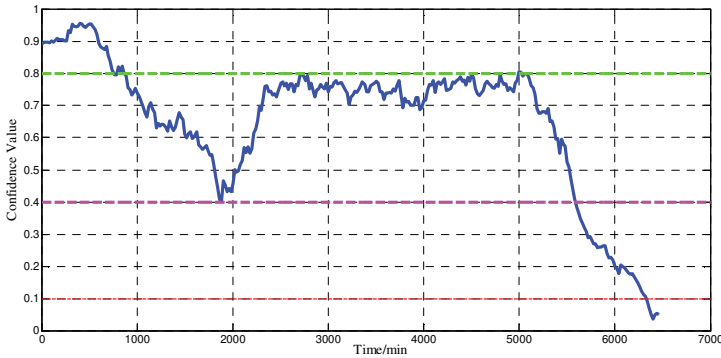


Fig. 3. Confidence Values obtained from the second dataset

To explain the advantages of the proposed method, we compared it with the result of the reference [29], which employed the Self Organizing Map (SOM) to process the same dataset for performance degradation assessment. The result is shown in Fig. 4. It is found that reference [29] uses the value of minimum quantization error (MQE) to reflect the health state. The larger the value of MQE, the worse the health state is. If there is no fault data, it is unavaial to know how large the MQE is means the failure appearance. In this situation, the method of reference [29] is useless for condition-based predictive maintenance. Meanwhile, the proposed method in this paper can be used without fault datasets in real applications, since the Confidence Values (CVs) gradually tend to 0. Besides, by comparison, it is also implied that the proposed method shows more sensitive to small initial changes in health conditions, and can detect defects at the incipient stage.

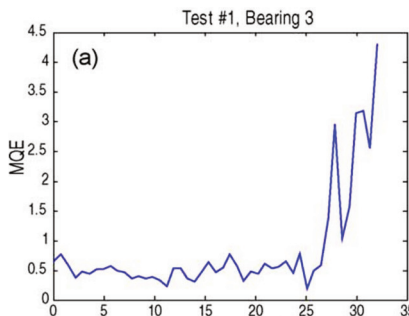


Fig. 4. Health assessment results from reference [29]

The blue curve in Fig. 3 represents the detected health degradation. In the first 900 min, the bearing was in good health and considered to be in the “normal stage”. From approximately

900 min to 5400 min, initial defects appeared and propagated, which decreased the Confidence Values (CVs); this is considered the “defect stage”. From approximately 5400 min to 6400 min, the defects worsened considerably, and the Confidence Values (CVs) decreased rapidly; this condition is considered the “danger stage”. A few minutes after 6400 min, fault occurred, and this condition is considered the “failure stage”. In Fig. 3, several thresholds of 0.8, 0.4, and 0.1 were selected. These thresholds were used to determine the health scale of the bearing, which is significant for condition-based predictive maintenance.

4. Conclusions

Condition-based maintenance has become critical to reducing operating and maintenance costs. Health assessment is one of the key techniques for condition-based maintenance. However, the strong nonlinearity features in vibration signal bring difficulties for health assessment. In this paper, a method based on an integrated feature selection approach and GMM is proposed for health assessment of rotary machinery. For the nonlinear and non-stationary vibration signal, an integrated feature selection approach combined with EMD, SVD and PCA is employed to reveal the health condition. Then, GMM is utilized for health assessment, which has been verified by two different experimental datasets. The proposed method has many benefits. First, it extracts those features that evidently reflect the health state of the bearings, because EMD is self-adaptive in decomposing nonlinear and non-stationary signals and SVD extracts the intrinsic characteristics of a matrix with favorable stability. Second, it is a reliable, real-time multivariate analysis method, because the calculation of GMM values has no expensive computational requirements, which means this method can be employed in real-time. Third, since the GMM can assess the overlap between the current observed features and those observed during the normal condition, therefore the proposed method can be used without fault datasets in real applications. Finally, compared with the existing method by reference [29], the proposed method shows more sensitive to small initial changes in health conditions, and can detect defects at the incipient stage. Nevertheless, additional work is needed to further validate the method in wider applications. Besides, the determination of thresholds for Confidence Values (CVs) is also very crucial for accurate health assessment and needs further works to select thresholds self-adaptively.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant No. 61074083, 50705005 and 51105019), the Technology Foundation Program of National Defense (Grant No. Z132013B002), as well as the Innovation Foundation of BUAA for PhD Graduates.

References

- [1] **Yu J.** A hybrid feature selection scheme and self-organizing map model for machine health assessment. *Applied Soft Computing*, Vol. 11, 2011, p. 4041-4054.
- [2] **Li B., Liu P., Hu R., Mi S., Fu J.** Fuzzy lattice classifier and its application to bearing fault diagnosis. *Applied Soft Computing*, Vol. 12, 2012, p. 1708-1719.
- [3] **Pan Y., Chen J., Guo L.** Robust bearing performance degradation assessment method based on improved wavelet packet-support vector data description. *Mechanical Systems and Signal Processing*, Vol. 23, 2009, p. 669-681.
- [4] **Van Tung Tran, Pham H. T., Yang B., Nguyen T. T.** Machine performance degradation assessment and remaining useful life prediction using proportional hazard model and support vector machine. *Mechanical Systems and Signal Processing*, 2012, p. 320-330.
- [5] **Yang Y., Liao Y., Meng G., Lee J.** A hybrid feature selection scheme for unsupervised learning and its application in bearing fault diagnosis. *Expert Systems with Applications*, Vol. 38, 2011, p. 11311-11320.

- [6] **Pan Y., Chen J., Guo L.** Robust bearing performance degradation assessment method based on improved wavelet packet-support vector data description. *Mechanical Systems and Signal Processing*, Vol. 23, 2009, p. 669-681.
- [7] **Zhu X., Zhang Y., Zhu Y.** Bearing performance degradation assessment based on the rough support vector data description. *Mechanical Systems and Signal Processing*, Vol. 34, 2013, p. 203-217.
- [8] **Djurdjanovic D., Lee J., Ni J.** Watchdog agentan – infotronics-based prognostics approach for product performance degradation assessment and prediction. *Advanced Engineering Informatics*, Vol. 17, 2003, p. 109-125.
- [9] **Liao L., Lee J.** Design of a reconfigurable prognostics platform for machine tools. *Expert Systems with Applications*, Vol. 37, 2010, p. 240-252.
- [10] **Yu J.** A nonlinear kernel Gaussian mixture model based inferential monitoring approach for fault detection and diagnosis of chemical processes. *Chemical Engineering Science*, Vol. 68, 2012, p. 506-519.
- [11] **Oehlmann H., Brie D., Tomczak M., Richard A.** A method for analysing gearbox faults using time – frequency representations. *Mechanical Systems and Signal Processing*, Vol. 11, 1997, p. 529-545.
- [12] **Yan R., Gao R. X.** An efficient approach to machine health diagnosis based on harmonic wavelet packet transform. *Robotics and Computer-Integrated Manufacturing*, Vol. 21, 2005, p. 291-301.
- [13] **Zou J., Chen J.** A comparative study on time – frequency feature of cracked rotor by Wigner-Ville distribution and wavelet transform. *Journal of Sound and Vibration*, Vol. 276, 2004, p. 1-11.
- [14] **Baydar N., Ball A.** A comparative study of acoustic and vibration signals in detection of gear failures using Wigner – Ville distribution. *Mechanical Systems and Signal Processing*, Vol. 15, 2001, p. 1091-1107.
- [15] **Tse P. W., Yang W., Tam H. Y.** Machine fault diagnosis through an effective exact wavelet analysis. *Journal of Sound and Vibration*, Vol. 277, 2004, p. 1005-1024.
- [16] **Junsheng C., Dejie Y., Yu Y.** Research on the intrinsic mode function (IMF) criterion in EMD method. *Mechanical Systems and Signal Processing*, Vol. 20, 2006, p. 817-824.
- [17] **Huang N. E., Shen Z., Long S. R., Wu M. C., Shih H. H., Zheng Q., Yen N., Tung C. C., Liu H. H.** The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Mathematical, Physical and Engineering Sciences*, Vol. 454, 1998, p. 903-995.
- [18] **Cheng J., Yu D., Tang J., Yang Y.** Application of SVM and SVD technique based on EMD to the fault diagnosis of the rotating machinery. *Shock and Vibration*, Vol. 16, 2009, p. 89-98.
- [19] **Yu D., Cheng J., Yang Y.** Application of EMD method and Hilbert spectrum to the fault diagnosis of roller bearings. *Mechanical Systems and Signal Processing*, Vol. 19, 2005, p. 259-270.
- [20] **Van Der Veeken A., Deprettere E. F., Swindlehurst A. L.** Subspace-based signal analysis using singular value decomposition. *Proceedings of the IEEE*, Vol. 81, 1993, p. 1277-1308.
- [21] **Sandy J.** Monitoring and diagnostics for rolling element bearings. *Sound and Vibration*, Vol. 22, 1988, p. 16-20.
- [22] **Wade M. J., Sánchez A., Katebi M. R.** On real-time control and process monitoring of wastewater treatment plants: real-time process monitoring. *Transactions of the Institute of Measurement and Control*, Vol. 27, 2005, p. 173-193.
- [23] **Liao L., Lee J.** A novel method for machine performance degradation assessment based on fixed cycle features test. *Journal of Sound and Vibration*, Vol. 326, 2009, p. 894-908.
- [24] **Yu G., Sun J., Li C.** Machine performance assessment using Gaussian mixture model (GMM). Shenzhen, China, 2008.
- [25] **Yu J.** Bearing performance degradation assessment using locality preserving projections and Gaussian mixture models. *Mechanical Systems and Signal Processing*, Vol. 25, 2011, p. 2573-2588.
- [26] **Liu W., Zhong X., Lee J., Liao L., Zhou M.** Application of a novel method for machine performance degradation assessment based on Gaussian mixture model and logistic regression. *Chinese Journal of Mechanical Engineering*, Vol. 24, 2011, p. 879-884.
- [27] **Zhang B., Zhang C., Yi X.** Competitive EM algorithm for finite mixture models. *Pattern Recognition*, Vol. 37, 2004, p. 131-144.
- [28] **Huda S., Yearwood J., Togneri R.** A stochastic version of Expectation Maximization algorithm for better estimation of Hidden Markov Model. *Pattern Recognition Letters*, Vol. 30, 2009, p. 1301-1309.
- [29] **Qiu H., Lee J., Lin J., Yu G.** Robust performance degradation assessment methods for enhanced rolling element bearing prognostics. *Advanced Engineering Informatics*, Vol. 17, 2003, p. 127-140.