Information Reconstruction Method for Improved Clustering and Diagnosis of Generic Gearbox Signals

Fangji Wu[1, 2], Jay Lee[2]

[1] *State Key Laboratory for Manufacturing Systems Engineering, Research Institute of Diagnostics and Cybernetics, Xi'an Jiaotong University, Xi'an, Shaanxi, 710049, China*
wfjridc@gmail.com
[2] *NSF I/UCR Center for Intelligent Maintenance System, University of Cincinnati, OH, 45221, USA*
jay.lee@uc.edu

ABSTRACT

Gearbox is a very complex mechanical system that can generate vibrations from its various elements such as gears, shafts, and bearings. Transmission path effect, signal coupling, and noise contamination can further induce difficulties to the development of diagnostic system for a gearbox. This paper introduces a novel information reconstruction approach to clustering and diagnosis of gearbox signals in varying operating conditions. First, vibration signal is transformed from time domain to frequency domain with Fast Fourier Transform (FFT). Then, reconstruction filters are employed to sift the frequency components in FFT spectrum to retain the information of interest. Features are further extracted to calculate the coefficients of the reconstructed energy expression. Then, correlation analysis (CA) and distance measurement (DM) techniques are utilized to cluster signals under diverse shaft speeds and loads. Finally, energy coefficients are used as health indicators for the purpose of fault diagnosis of the rotating elements in the gearbox. The proposed method was used to solve the gearbox problem of the 2009 PHM Conference Data Analysis Competition and won with the best score in both professional and student categories.*

1. INTRODUCTION

Gearbox is one of the most widespread and crucial rotating mechanical systems in modern industry. It provides a speed-torque conversion from a higher speed motor to a slower but more forceful output or vice-versa. A gearbox usually consists of rotating elements such as gears, shafts, and bearings and static elements such as box body and bearing caps. During operation, a gearbox system can suffer the following: gear failures such as wear, scoring, interference, surface fatigue, plastic flow and fracture; bearing failures such as wear, scoring, surface fatigue and brinelling; and shaft failures such as fatigue cracking and overload (Forrester 1996). All these defects can worsen the operating condition and excite excess vibration, and potentially cause major unexpected breakdowns and safety issues. Condition monitoring and fault prognostics of gearbox system have been used for many applications to some degree of success (Peng and Chu 2004, Suh et al. 1999, Wang et al. 2007, Byington et al. 2004). The major challenge is to effectively and accurately identify abnormal patterns early with a sound estimation of the remaining useful life (RUL).

The 2009 PHM Conference Data Analysis Competition is focused on the detection and magnitude estimation of mechanical faults from a generic gearbox using accelerometer data and information about bearing geometry. Participants are scored based on their ability to correctly identify fault type, location, magnitude and damage in the gear system. Data were collected at 30, 35, 40, 45 and 50 Hz shaft speed while being subjected to either high or low loading. Additionally, repeated runs are included in the data, although the run time and load were not sufficient to induce significant fault progression. There are a total of 560 vibration data files to be classified and diagnosed. Details of the Data Analysis Competition are provided on the website http://www.phmsociety.org/competition/09.

This paper introduces a novel information reconstruction approach for clustering and diagnosis of gearbox signals in varying operating conditions.

Fig. 1 is a schematic diagram of the proposed approach. First, vibration signal is transformed from time domain to frequency domain with Fast Fourier Transform (FFT). Second, reconstruction filters are employed to sift the frequency components in FFT spectrum to retain the information of interest and eventually obtain the reconstructed FFT spectrum. Features are further extracted from the modified spectrum to calculate the coefficients of the reconstructed energy expression (energy fitting model). Then, correlation analysis (CA) and distance measurement (DM) techniques are used for clustering signals under diverse shaft speeds and loads. Finally, energy coefficients are used as health indicators for fault diagnosis of the rotating elements in the gearbox. Basically, this approach is a hybrid of data-driven and model-driven schemes. It can be applied as a systematic method for gearbox health assessment system.

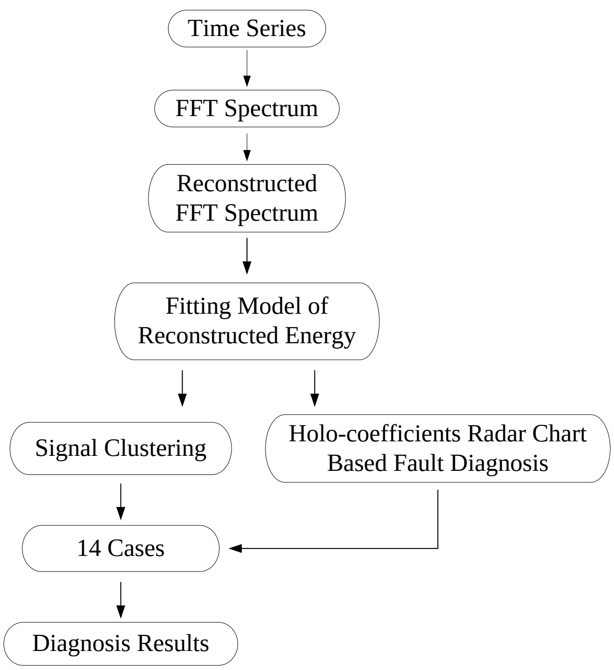

Fig.1. Overview of information reconstruction method

This paper is organized as follows. In Sec. 2, the scheme of reconstructing FFT spectrum is introduced. The feature extraction and reconstructed energy are presented in Sec. 3. Sec. 4 shows the signal clustering process and result of accelerometer data. Sec. 5 introduces holo-coefficients map for gearbox fault diagnosis. The generalization and improvement of the information reconstruction method is discussed in Sec. 6. Finally, conclusions are presented in Sec. 7.

2. RECONSTRUCTED FFT SPECTRUM

To gain further understanding of the gearbox signals, many tools have been developed. These tools consist of time synchronous average (Dempsey 2004) and autoregressive moving average (Wang and Wong 2002) model for time domain analysis; FFT (Lin et al. 1993), power spectrum (Baydar and Ball 2000), and cepstrum (Badaoui et al. 2001) for frequency domain analysis; short-time Fourier transform (Pinnegar and Mansinha 2003), Wigner-Ville distribution (Baydar and Ball 2000), wavelet transform (Sung et al. 2000), and Hilbert-Huang Transform (Huang et al. 1998) for time-frequency analysis, among others. For the 2009 PHM competition case, vibration data were collected using accelerometers mounted on both the input and output shaft retaining plates. The signal can be described as a complicated measurement with a wide-range energy distribution. However, only some parts of signal are related to specific machine conditions. The main idea of spectrum analysis is to either look at the whole spectrum or look closely at certain frequency components of interest and then extract features from the signal.

To remove or reduce noise and effects from other unexpected sources and further enhance signal components of interest, a reconstruction approach is used to filter and assemble the frequency components to reconstruct signal without loss of information of interest. The scheme of reconstruction method is illustrated in Fig. 2. Each signal is transformed to FFT spectrum. Then, eighteen band-pass filters are applied to select specific frequency bands within the signal. Finally, all the eighteen frequency segments are reassembled together to reconstruct a new signal. The functions of these eighteen band-pass filters are listed in Table 1, which shows the criteria for defining these filters. In this table, frequency components are obtained by calculating corresponding vibration characteristic frequencies of shafts, gears and bearings. Frequency order is the ratio of the characteristic frequency to the shaft rotating frequency.

For shaft, defects such as unbalance and bend will excite harmonic frequency components of shaft rotating frequency. For gear, characteristic frequencies are gear meshing frequency (GMF) and its side band frequencies. GMF is equal to the number of teeth multiplied by the rotational frequency of the gear. It is the periodic signal at the tooth-meshing rate due to deviations from the ideal tooth profile. Side band signals are induced by amplitude modulation effects due to variations in tooth loading; frequency modulation effects due to rotational speed fluctuations and non-uniform tooth spacing; and additive impulses associated with tooth faults. For bearing, a defect on the inner or outer race will cause an impulse each time a rolling element contacts the defect. For an inner race defect this occurs at the inner race ball pass frequency (BPFI), and for an outer race defect this occurs at outer race ball pass frequency (BPFO). A defect on rolling element will cause an impulse each time the defect surface contacts the inner or outer races, which will excite the ball spin frequency (BSF). These characteristic frequencies can be expressed as:

$$\begin{cases} BPFI = \dfrac{N}{2}(f_o - f_i)(1 + \dfrac{d}{D}\cos(\alpha)) \\ BPFO = \dfrac{N}{2}(f_o - f_i)(1 - \dfrac{d}{D}\cos(\alpha)) \\ BSF = (f_o - f_i)(\dfrac{D}{d} - \dfrac{d}{D}\cos^2(\alpha)) \end{cases} \tag{1}$$

where N is the number of rolling elements, f_o is the rotational frequency of the outer race, f_i is the rotational frequency of the inner race, d is the diameter of the rolling elements, D is the pitch circle diameter, α is the contact angle.

Table 2 lists the corresponding meaning of these eighteen filters and shows why these filters are defined. The i-X GMF means i-th harmonic frequency of gear meshing frequency. To cite an example, Fig. 3 shows the FFT spectrum of input side signal of File-29 and Fig. 4 shows its reconstructed FFT spectrum.

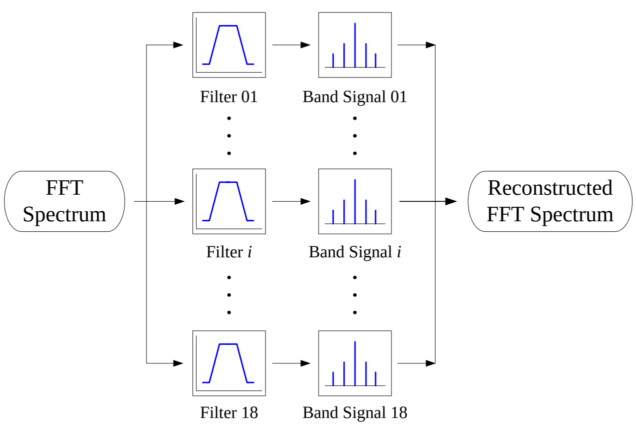

Fig.2. FFT spectrum reconstruction

Table 1. Functions of reconstruction filters

Filter 01	Retaining 01X order component
Filter 02	Retaining 02X order component
Filter 03	Retaining 03X order component
Filter 04	Retaining 04X order component
Filter 05	Retaining 05X order component
Filter 06	Retaining 45X order component
Filter 07	Retaining 06X-10X order component
Filter 08	Retaining 14X-18X order component
Filter 09	Retaining 22X-26X order component
Filter 10	Retaining 30X-34X order component
Filter 11	Retaining 38X-42X order component
Filter 12	Retaining 46X-50X order component
Filter 13	Retaining 54X-58X order component
Filter 14	Retaining 62X-66X order component
Filter 15	Retaining 78X-82X order component
Filter 16	Retaining 94X-98X order component
Filter 17	Retaining 110X-114X order component
Filter 18	Retaining 126X-130X order component

Table 2. Corresponding meaning of filter functions

Filter 01	Characteristic frequency component of input shaft unbalance
Filter 02	Characteristic frequency component of bent input shaft
Filter 03	Characteristic frequency component of outer race defect of input-shaft bearing
Filter 04	Characteristic frequency component of ball defect of input-shaft bearing
Filter 05	Characteristic frequency component of inner race defect of input-shaft bearing
Filter 06	Natural frequency of rotating element or gear ghost frequency component
Filter 07	Output-shaft helical 1X GMF
Filter 08	Input-shaft helical 1X GMF Output-shaft helical 2X GMF Output-shaft spur 1X GMF
Filter 09	Output-shaft helical 3X GMF
Filter 10	Input-shaft helical 2X GMF Output-shaft helical 4X GMF Input-shaft spur 1X GMF Output-shaft spur 2X GMF
Filter 11	Output-shaft helical 5X GMF
Filter 12	Input-shaft helical 3X GMF Output-shaft helical 6X GMF Output-shaft spur 3X GMF
Filter 13	Output-shaft helical 7X GMF
Filter 14	Input-shaft helical 4X GMF Output-shaft helical 8X GMF Input-shaft spur 2X GMF Output-shaft spur 4X GMF
Filter 15	Input-shaft helical 5X GMF Output-shaft spur 5X GMF
Filter 16	Input-shaft helical 6X GMF Input-shaft spur 3X GMF Output-shaft spur 6X GMF
Filter 17	Input-shaft helical 7X GMF Output-shaft spur 7X GMF
Filter 18	Input-shaft helical 8X GMF Input-shaft spur 4X GMF Output-shaft spur 8X GMF

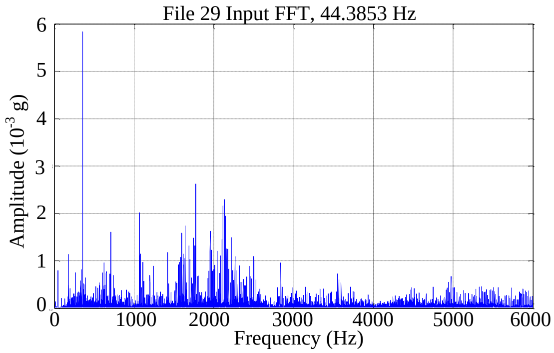

Fig.3. FFT spectrum of File-29

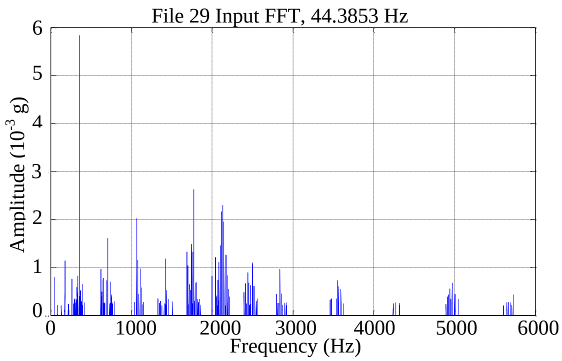

Fig.4. Reconstructed FFT spectrum of File-29

3. FEATURE EXTRACTION AND RECONSTRUCTED ENERGY

Based on the reconstructed FFT spectrum, eighteen features are extracted and they serve as coefficients in the reconstructed energy model. The reconstructed energy can be expressed as:

$$\begin{cases} f_E = f_{EI} + f_{EO} \\ f_{EI} = (\alpha_1 + \cdots + \alpha_6) \times E_{Imax} + (\alpha_7 + \cdots + \alpha_{18}) \times E_{Iall} \\ f_{EO} = (\beta_1 + \cdots + \beta_6) \times E_{Omax} + (\beta_7 + \cdots + \beta_{18}) \times E_{Oall} \end{cases} \quad (2)$$

where f_E is the total energy index of input and output side signals, f_{EI} is the energy index of input side signal, f_{EO} is the energy index of output side signal, E_{Imax} and E_{Omax} are the maximum energy components of input and output side signals, E_{Iall} and E_{Oall} are the full energy values of input and output side signals, α_1 to α_6 are derived by dividing the energy of the first six band signals of input side signal by E_{Imax}, β_1 to β_6 results from dividing energy of first six band signals of output side signal by E_{Omax}, α_7 to α_{18} are computed when energy of last twelve band signals of input side signal is divided by E_{Iall} , and finally, β_7 to β_{18} are determined by dividing the energy of last twelve band signals of output side signal by E_{Oall}.

In the reconstructed energy expression, energy coefficients are selected to have certain classification power. The basic idea is to identify and further classify the data with similar attributes to a specific group. For example, α_1 is supposed to classify the data either to unbalance group or normal group. Moreover, energy coefficients are also supposed to be comprehensible for user or have physical meaning. This is necessary whenever the classified pattern is to be used for supporting a decision to be made. If the classified pattern is a group without explanation, the user may not trust it. In this paper, knowledge comprehensibility can be achieved by using high-level knowledge representations described in the previous section.

4. SIGNAL CLUSTERING

Given a set of data items, partitioning this set into subsets, such that items with similar characteristics or features are grouped together, is the general idea of signal clustering (Goebel and Gruenwald 1999). A natural way of signal clustering is based on certain similarity measure or distance measure between two signals. In this section, CA and DM on energy coefficients are introduced and evaluated for clustering signals under diverse shaft speeds and loads. Vector of energy coefficients can be constructed as

$$C_E = [\alpha_1, ..., \alpha_{18}, \beta_1, ..., \beta_{18}]^T \quad (3)$$

Then, CA on two signals is defined as

$$CA = (C_{Ei} \cdot C_{Ej}) / (|C_{Ei}| * |C_{Ej}|) \quad (4)$$

where $\cdot$ means dot product, $|\cdot|$ means the largest singular value of a vector. The result of CA ranges between zero and one, with higher CA signifying a higher correlation. DM on two signals is

$$DM = \| C_{Ei} - C_{Ej} \| \quad (5)$$

where $\|\cdot\|$ is the Euclidean distance, with lower DM signifying a higher similarity.

4.1 Determination of Repeated Runs

Using the tachometer signal, rotating speed can be calculated as shown in Fig. 5. There are five distinct groups corresponding to the 5 shaft speeds and each group contains exactly 112 data points. Repeated runs identification was then applied to each speed regime. Consider 50 Hz speed regime, CA for File-157 on these 112 files is illustrated in Fig. 6, while DM for the same scenario is shown in Fig. 7. CA shows that File-183, File-227 and File-498 have the largest correlation value to File-157 and they can be considered as its repeated runs. DM also shows that these three files have the smallest distance value to File-157 and confirms that they are its repeated runs.

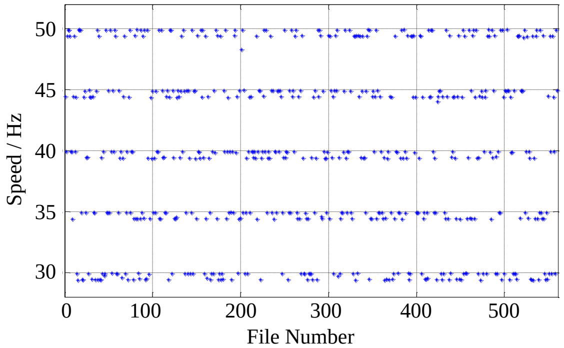

Fig.5. Input shaft speeds

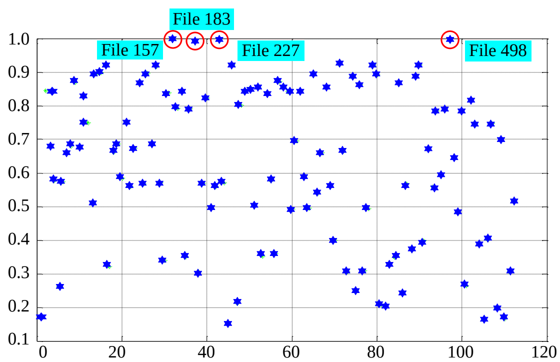

Fig.6. CA for File-157

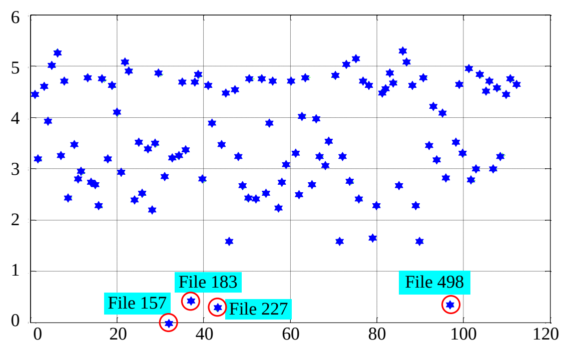

Fig.7. DM for File-157

4.2 Identification of Diverse Loading Runs

After identifying the 4 repeated runs, the 112 files in 50Hz regime are now clustered into 28 groups. CA for File-157 on 28 files from these 28 groups, one file from each group, is illustrated in Fig. 8. DM for File-157 on these 28 files is illustrated in Fig. 9. CA shows that File-250 has the largest correlation value to File-157 and they are from the same pattern. DM shows that File-250 has the smallest distance value to File-157 and they are from the same pattern, one with high load and the other with low load. After identifying the high and low loading runs, 112 files in each speed regime are reduced into 14 groups.

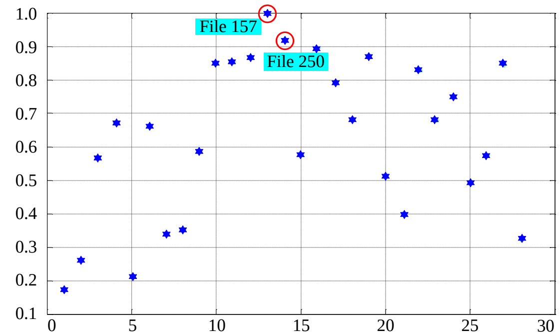

Fig.8. CA for File-157

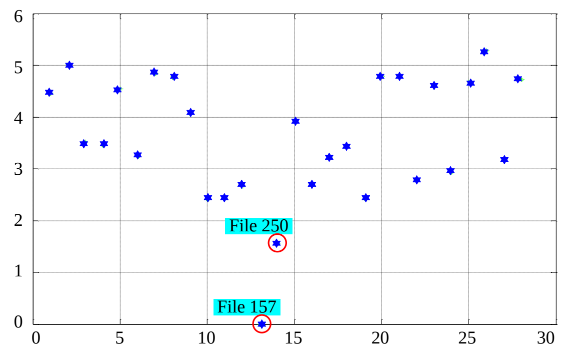

Fig.9. DM for File-157

4.3 Identification of Diverse-Speed Runs

At this point, each speed regime has 14 groups (replications and loading, considered). This section will then describe how the 14 unique patterns are identified across the 5 speed regimes. Consider File 157 (with File-250 as its load pair) in 50Hz regime, its CA and DM with 28 files (one from each of the 28 groups in the same speed regime after identifying replications) in 45Hz regime, are illustrated in Fig. 10 and 11, respectively. Both figures show that File-157 and File-55 (File-62 was its load pair as determined in a previous step) share the same pattern. By doing the same process for the other 3 speed regimes, it was found that File-59, File-69 in 30Hz, File-34, File-88 in 35Hz, File-56, File-213 in 40Hz, File-55, File-62 in 45Hz, and, File-157, File-250 in 50Hz can be clustered as one pattern (Pattern A).

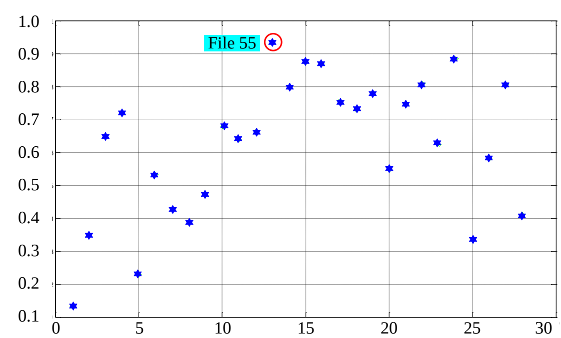

Fig.10. CA for File-157

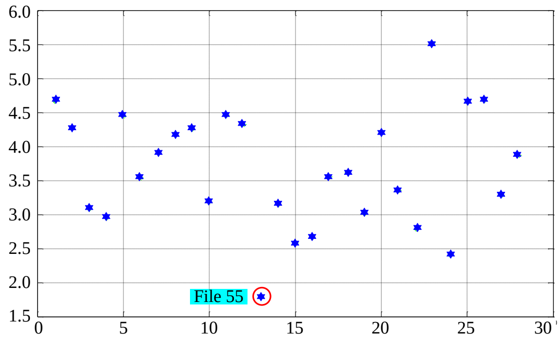

Fig.11. DM for File-157

5. HOLO-COEFFICIENTS MAP/RADAR CHART AND FAULT DIAGNOSIS

The fault diagnosis of rotating elements in the gearbox is performed using energy coefficients as health indicators. A holo-coefficients map comprises of all the energy coefficients. In the map (e.g. Fig. 12 and Fig. 14), the contribution rate of each coefficient can be revealed very clearly along with operating conditions. A more advanced format of holo-coefficients map is holo-coefficients radar chart. The multivariate data in holo-coefficients map are displayed in holo-coefficients radar chart starting from the same point and in different equi-angular spokes, with each spoke representing one of the variables. The data length of a spoke is proportional to the magnitude of the variable for the data point. In the chart (e.g. Fig. 13 and Fig. 15), radial 1 to 18 correspond to α_1 to α_{18}, and radial 19 to 36 correspond to β_1 to β_{18}. The map and chart can be treated as qualitative tools for fault diagnosis. The rules that authors used for qualitative diagnosis are: 1) energy coefficient of a defect should be higher than normal case; the threshold of faulty case depends highly on gearbox set and its dynamic characteristics; usually an energy coefficient larger than 0.4 should trigger a warning, 2) bearing defect may excite lower energy coefficient compared to shaft and gear defect; 3) a high energy coefficient in hard working condition such as high loading and high speed is more reliable for fault detection. Moreover, holo-coefficients map can be updated for quantitative diagnosis. This will be further discussed in next section for generalization of the proposed approach.

Fig. 12 shows the holo-coefficients map of files of Pattern A. Fig. 13 is the transformed radar chart format of Fig. 12. From the figure, input shaft unbalance (radials 1 and 19) and bearing outer defect at input shaft output side (radial 21) are diagnosed. The unbalance excites 1X frequency component as measured from the input side signal and this component is also distinct in output side signal due to the transmission effect of the rigid gearbox housing. The contribution rate of coefficient 3 in 40 Hz is also considerable. However, with the increase in speed, its contribution decreases. Fig. 14 shows the holo-coefficients map of Pattern B (File-60 in 30Hz, File-19 in 35Hz, File-185 in 40Hz, File-36 in 45Hz, and File-258 in 50Hz). Fig. 15 is the transformed radar chart format of Fig. 14. It is determined that this pattern contains gear error defect at idler shaft 2 location (radials 8 and 26).

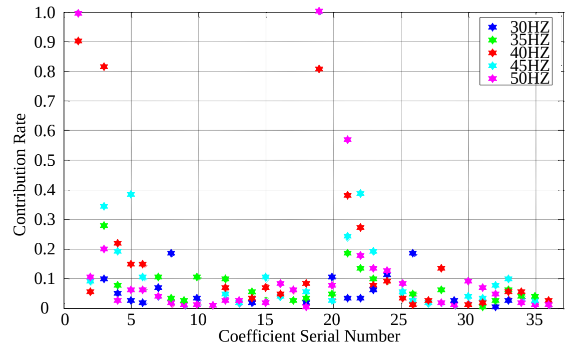

Fig.12. Holo-coefficients map of pattern A

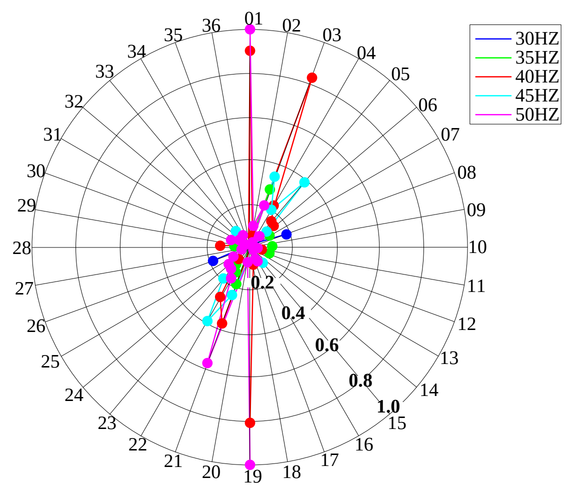

Fig.13. Holo-coefficients radar chart of pattern A

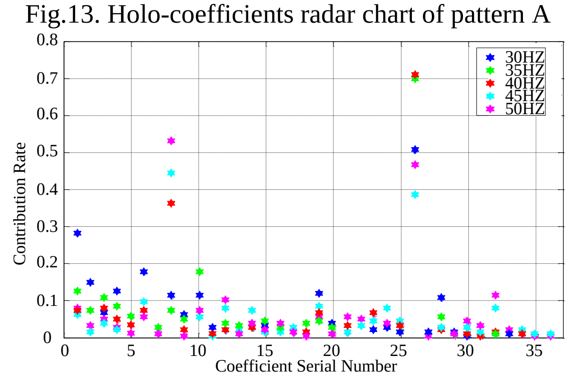

Fig.14. Holo-coefficients map of pattern B

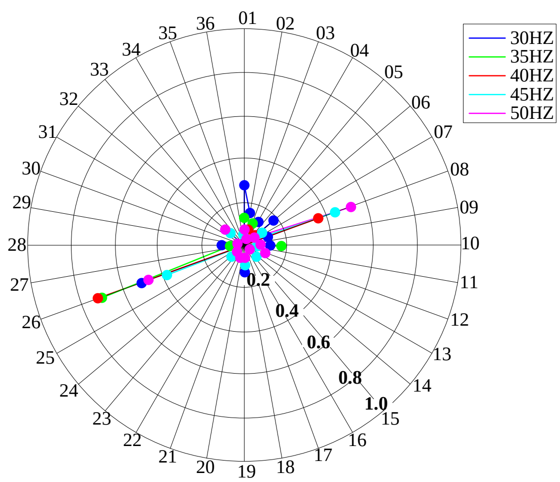

Fig.15. Holo-coefficients radar chart of pattern B

6. GENERALIZATION AND IMPROVEMENT OF INFORMATION RECONSTRUCTION METHOD

The information selection and feature extraction are the crucial steps of the proposed information reconstruction method. The effect of feature selection are (1) to improve classification and diagnosis performance; (2) to visualize the data for model construction; (3) to reduce dimensionality and (4) to remove noise. Improper selection of information of interest and poor extraction of features can lead to under-fitting and over-fitting issues during model creation of the PHM activity of a gearbox system. In developing the energy expression, there is a risk of generating too many energy coefficients which is called over-fitting. Over-fitting will decrease the efficiency and accuracy of the classification since irrelevant attributes can confuse the data mining algorithm. On the contrary, under-fitting means energy coefficients are not enough to support the decision making process.

For over-fitting, it is desirable to have a procedure to prune the ensemble of energy coefficients while keeping the expected classification performance and avoiding the risks in feature selection. The method for selection of energy coefficients that was discussed in this paper relied on expert knowledge which is user-driven and domain-dependent. Had the data files been labeled *a priori,* original files can then be taken as training data, therefore, objective methods, which are data-driven and domain-independent, can be employed to optimize the energy coefficients. The principal component analysis (PCA) can be used to prune the energy coefficients. Because of its ability to discriminate directions with the largest variance in a data set, it is suitable to use PCA for identifying the most representative features. One can first classify data files by pattern; then, apply PCA to feature vectors of data files in each pattern to find the most representative features for the corresponding pattern; finally, assemble retained features from each pattern to obtain the final feature set. Fisher criterion can be another approach for pruning the energy coefficients. Suppose that we have a set of features in the pattern labeled ω_1 and another set of features in the pattern labeled ω_2. Fisher criterion method actually tries to find the feature set to maximize the distance between two patterns and minimize the deviation within each pattern. A Fisher criterion score can be expressed as:

$$SF_i = \left| \frac{m(\omega_1)_i - m(\omega_2)_i}{\sigma^2(\omega_1)_i + \sigma^2(\omega_2)_i} \right| \tag{6}$$

where $m(\omega_1)_i$ and $m(\omega_2)_i$ are the mean value for the *i*-th feature in ω_1 and ω_2 pattern, $\sigma^2(\omega_1)_i$ and $\sigma^2(\omega_2)_i$ are the standard deviation. By deleting features with small Fisher criterion score, one can exclude irrelevant features from original feature set. Moreover, other advanced feature selection methods such as support vector machine (SVM) and genetic algorithm (GA) based approaches can also be applied (Bradley and Mangasarian 1998, Yang and Honavar 1997).

For under-fitting, more efficient signal processing methods are needed to extract more distinguishable features or more information about the gearbox set is needed to define specified attributes such as natural frequency of gears and bearings. In the current energy expression, the weighting coefficients reflecting the relative importance of energy coefficients are same. If there is evidence proving one energy coefficient is more distinguishable than others, the energy expression can be improved further to have more efficient performance and a more accurate diagnosis. Finally, holo-coefficients radar chart is capable for quantitative diagnosis. However, in order to achieve this goal, there are three sub-tasks need to be considered. First, experiment should be carried out in detail to record the relationship between single energy coefficient and single defect. Second, experiment should be carried out in detail to record the relationship between whole energy coefficients and multi-defects. Third, a model need to be established to represent the relationship between energy coefficients and defects, and then a quantitative reference system and thresholds for quantitative diagnosis can be obtained.

7. CONCLUSION

This paper addressed the information reconstruction method for solving the challenging problem of the 2009 PHM Conference Data Analysis Competition. With this method, raw data can be represented by a reconstructed energy model. Then, based on the energy coefficient of this model, signal clustering can be performed for determination of repeated runs, identification of diverse loading runs, and identification of diverse speed runs. Thus, 560 vibration data files can be classified into 14 patterns. For fault diagnosis of rotating elements in the gearbox, holo-coefficients map and radar chart are used. In the map and chart, the contribution rate of each energy coefficient can be revealed very clearly along with operating conditions. Finally, in order to further apply the information reconstruction method to other gearbox sets besides the one used for PHM competition and to further improve the current approach, four issues are discussed as 1) over-fitting issue, 2) under-fitting issue, 3) weighting coefficient, and 4) quantitative diagnosis. The proposed information reconstruction method can further be applied to the gearbox set working in varying working condition such as helicopter gearbox and wind turbine gearbox for signal clustering and fault diagnosis.

For development of gearbox diagnostic system, extraction of features that are less sensitive or not sensitive to working conditions is critical to accuracy; simulation of the problem-solving process of experts to get diagnosis results with computer is critical to efficiency. In the future, solving problems without interference of experts or performing computer-aided pre-diagnosis before resorting to experts could be expected with the further development of intelligent diagnostic systems.

ACKNOWLEDGMENT

The authors would like to thank Prof. Yudi Shen and other researchers at Research Institute of Diagnostics and Cybernetics, Xi'an Jiaotong University for their kind support during the work of the competition.

REFERENCES

B. D. Forrester. (1996). *Advanced Vibration Analysis Techniques for Fault Detection and Diagnosis in Geared Transmission Systems.* Ph. D. Thesis, Swinburne University of Technology, Melbourne, Australia.

Z. Peng and F. Chu. (2004). Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography, *Mechanical Systems and Signal Processing*, vol. 18, pp. 199-221.

J.H. Suh et al. (1999). Machinery fault diagnosis and prognosis: application of advanced signal processing techniques, *CIRP Annals – Manufacturing Technology*, vol. 48, issue 1, pp. 317-320.

J.-Z. Wang, et al. (2007). Gearbox fault diagnosis and prediction based on empirical mode decomposition scheme, *Proceedings of the Sixth International Conference on Machine Learning and Cybernetics*, pp. 1072-1075.

C.S. Byington, et al. (2004). Data-driven neural network methodology to remaining life predictions for aircraft actuator components, *IEEE Aerospace Conference Proceedings*, vol. 6, pp. 3581-3589.

P. J. Dempsey (2004). A Comparison of Vibration and Oil Debris Gear Damage Detection Methods Applied to Pitting Damage, *in Proceedings of COMADEM 2000, 13th International Congress on Condition Monitoring and Diagnostic Engineering Management*, Houston, TX.

W. Wang and A. K. Wong. (2002). Autoregressive Model-Based Gear Fault Diagnosis, *Journal of Vibration and Acoustics*, vol. 124, pp. 172-179.

H. H. Lin, D. P. Townsend, F. B. Oswald (1993). Prediction of Gear Dynamics Using Fast Fourier Transform of Static Transmission Error, *Mechanics Based Design of Structures and Machines: An International Journal*, vol. 21, pp. 237-260.

N. Baydar, A. Ball (2000). Detection of gear deterioration under varying load conditions by using the instantaneous power spectrum, *Mechanical Systems and Signal Processing*, vol. 14, pp. 907-921.

M. E. Badaoui et al. (2001). Use of the moving cepstrum integral to detect and localise tooth spalls in gears, *Mechanical Systems and Signal Processing*, vol. 15, pp. 873-885.

C. R. Pinnegar, L. Mansinha (2003). Time-local spectral analysis for non-stationary time series: The S-Transform for noisy signals, *Fluctuation and Noise Letters*, vol. 3, pp. 357-364.

N. Baydar, A. Ball (2000). A comparative study of acoustic and vibration signals in detection of gear failures using Wigner-Ville distribution, *Mechanical Systems and Signal Processing*, vol. 15, pp. 1091-1107.

C. K. Sung, H. M. Tai and C. W. Chen. (2000). Locating defects of a gear system by the technique of wavelet transform, *Mechanism and Machine Theory*, vol. 35, pp. 1169-1182.

N.E. Huang, et al. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proceedings of the Royal Society of London*, Series A 454 pp. 903–995.

M. Goebel, L. Gruenwald (1999). A survey of data mining and knowledge discovery software tools, *ACM SIGKDD Explorations Newsletter*, vol. 1, pp. 20-33.

P. S. Bradley, O. L. Mangasarian. Feature selection via concave minimization and support vector machines, *in Proceedings of 13th International Conference on Machine Learning*, San Francisco, CA.

J. Yang and V. Honavar. Feature subset selection using a genetic algorithm, Feature Extraction, Construction and Selection: A Data Mining Perspective, pp. 117-136, 1998, second printing, 2001.

Fangji Wu is currently a Ph.D. student in State Key Laboratory for Manufacturing Systems Engineering, Research Institute of Diagnostics and Cybernetics, Xi'an Jiaotong University in China, and a visiting scholar in IMS Center, Department of Mechanical Engineering at University of Cincinnati in US. His current research focuses on component-level and system-level CBM and PHM; health management system design and industrial applications; data-driven and model-based methods for fault detection, diagnosis, and prognosis.

Jay Lee is Ohio Eminent Scholar and L.W. Scott Alter Chair Professor at the Univ. of Cincinnati and is founding director of National Science Foundation (NSF) Industry/University Cooperative Research Center (I/UCRC) on Intelligent Maintenance Systems which is a multi-campus NSF Center of Excellence between the Univ. of Cincinnati (lead institution), the Univ. of Michigan, and the Univ. of Missouri-Rolla. His current research focuses on autonomic computing, embedded IT and smart prognostics technologies for industrial and healthcare systems, design of self-maintenance and self-healing, systems, and dominant design tools for product and service innovation. He is also a Fellow of ASME, SME, as well as International Society of Engineering Asset Management (ISEAM).