

A new data mining approach for gear crack level identification based on manifold learning

Zhixiong Li*, Xinping Yan*, Yu Jiang**, Li Qin*, Jingping Wu***

*Key Laboratory of Marine Power Engineering and Technology (Ministry of Transportation), Wuhan University of Technology, Wuhan 430063, P. R. China, E-mail: jylys@hsu.edu.cn

**School of Information Engineering, Huangshan University, Huangshan 245021, P. R. China

***School of Mechanical and Manufacturing Engineering, The University of New South Wales, UNSW Sydney, NSW 2052, Australia

crossref <http://dx.doi.org/10.5755/j01.mech.18.1.1276>

1. Introduction

Gear transmission systems are widely served in industrial application. Generally working in severe conditions, gears are subjected to progressive deterioration of their state [1]. The breakdowns of the transmission machinery resulted from the gear failures account for 80%. One of the most frequently occurred fault modes is the crack [2]. A crack fault may lead to the cause of gear tooth broken, which could bring serious damage to the machinery. Therefore, the identification of gear crack level is crucial to prevent the system from malfunction.

Up to date fault diagnosis of industrial gearboxes has received intensive study for several decades, and vibration signal analysis is manifestly the most commonly used method for detecting gear failures. However, the nonlinear property of vibration signals made the diagnosis of gear fault very difficult, especially for the early gear cracks [2]. Classical methods, including spectral analysis, time domain averaging and envelope detection, etc., are ineffective for the feature extraction of nonstationary signal [3-5]. This is because most of the conventional techniques are based on the assumption that the vibration signals are stationary [3]. Hence, advanced techniques, including Wigner-Ville distribution (WVD) [6], empirical mode decomposition (EMD) [7, 8] and wavelet transform (WT) [9, 10], etc., are introduced into the analysis of nonstationary signals. These methods can easily handle a large number of variables and are very powerful for fault detection [6-10]. In general, the original fault characteristics obtained from these advanced signal processing methods contain some redundant ones. If use them to reveal the working states of the machine, the detection rate may be low. Hence, it is imperative to use the data mining techniques to eliminate the useless features before diagnosis, i.e. feature selection. The challenge is that it is not easy to determine the most distinguished features. One of the most popular data mining methods, principal component analysis (PCA) and its derivative algorithms, have been proved to be a useful tool for feature reduction and extraction [11]. However, their main limitation lies in their ability to capture the nonlinear properties of the original data [12-14]. The same problems are also found in other methods [15], including multi-dimensional scaling (MDS), linear discriminate analysis (LDA) and independent component analysis (ICA). For this reason, the manifold learning algorithms are developed for the nonlinear feature selection. Compared with the linear ones, the purpose of manifold learning methods is to

project the original high-dimensional data into a lower dimensional feature space by preserving the local topology of the original data. Hence the intrinsic structure of the data of interest can be extracted effectively. The representative methods include Isomap [12], Laplacian eigenmap [13] and locally linear embedding (LLE) [14], etc. Successful applications of these new nonlinear feature selection methodologies can be found in the field of image processing, speech spectrograms, EEG and ECG signals for medical diagnose [15]. Furthermore, manifold learning is seldom researched in condition monitoring and fault diagnosis field, especially for the gear fault detection. Yang et al. [16] proposed a method of nonlinear time series noise reduction based on principal manifold learning applied to the analysis of gearbox vibration signal with tooth broken, but only for signal denoising. Jiang et al. [15] proposed the supervised manifold learning algorithm (S-LapEig) for feature extraction. The gear pitting and gear bear faults were investigated in their study and high recognition rate was obtained. Hence, it is worth investigating the feature extraction for different gear crack levels using the new nonlinear feature selection methods.

This paper aims to tackle gear crack severity identification. Due to the good ability of local geometry structure information preservation, the LLE have become a hot research topic in the field of image processing [15]. Improve methods, such as Hessian LLE (HLLE), supervised locally linear embedding and LLE-LDA (ULLELDA), etc., have been proposed for the image feature extraction [15]. In this paper, a new method is proposed based on the empirical mode decomposition (EMD) and (SLLE) for gear crack diagnosis. The SLLE was extended to the gear fault diagnosis for feature reduction and extraction. Even though many researchers have performed fault detections of gearbox using EMD, they did not employ nonlinear feature selection technique to support their works [7, 8]. To verify the efficacy of the proposed scheme, the experimental tests were carried out in the present work, and the analysis results demonstrate that the proposed method based on the EMD-SLLE is effective and efficient for the gear crack level identification.

2. Description of new data mining algorithm

As discussed, the new method based on the combination of empirical mode decomposition (EMD) and supervised locally linear embedding (SLLE) is presented to the gear crack identification. We review EMD first, and then SLLE.

2.1. Empirical mode decomposition (EMD)

EMD [17] is useful advanced signal processing technique for the analysis of the vibration signals. EMD has the ability to decompose a signal into a number of monocomponent signals, named as intrinsic mode functions (IMFs) [17]. IMFs represent simple oscillatory modes embedded in the signal [18]. An IMF is a function that satisfies the following definitions [17].

1. In the whole analysis dataset, the number of extrema and the number of zero-crossings must either equal or differ at most by one.

2. At any point, the mean value of the envelope defined by local maxima and the envelope defined by the local minima is zero.

To extract IMFs from a vibration signal x , all the local extrema are firstly identified. Then a cubic spline line connects all the local maxima as upper envelope and all the minima as lower envelope. The mean of upper and lower envelope is subtracted from x to obtain h_1 . Check h_1 for the IMF conditions. If it satisfies the conditions it is an IMF, otherwise upper and lower envelopes are found for the h_1 and the process is repeated till the first IMF c_1 is got. Subtract c_1 from x and the result is now treated as new original signal and the above process is repeated to get the second IMF. Keep continuing the process till no more IMF can be extracted. Thus, at the end of the EMD decomposition we obtain

$$x = \sum_{i=1}^N c_i + r_N \quad (1)$$

where r_N is the final residue and c_i ($i = 1, 2, \dots, N$) is the i th IMF.

2.2. Supervised locally linear embedding (SLLE)

The effective feature extraction is important for the pattern recognition of gear fault. The linear eigenvector-based methods have achieved significant successes in the feature extraction. These algorithms include principal component analysis (PCA) and linear discriminate analysis (LDA), etc. However, the linear solutions may lead to losing the nonlinear properties of the original data [10-12]. Unlike the linear eigenvector-based feature extraction algorithms, LLE preserves local topology of high-dimensional data in the reduced space. This advantage is essential to maintain the nonlinear properties of the input data and thus benefits characteristic information extraction.

Locally linear embedding proposed by Roweis and Saul [14] aims to discover distinct low dimensional manifold embedded in the nonlinear high-dimensional data. Due to its good ability of local geometry structure information preservation, LLE has become a hot research topic in the image processing and EEG signal analysis [15], etc. Improve methods, such as Hessian LLE (HLLE), supervised LLE (SLLE) and LLE-LDA (ULLELDA), etc., have been proposed for the feature extraction problem [15]. Given a high dimensional feature space $F = [f_1, f_2, \dots, f_n] \in R^p$ (n is the total sample number and p the dimensionality of each sample), the objective of LLE is to reconstruct a nonlinear mapping to project F

into a reduced manifold space $H = [h_1, h_2, \dots, h_n] \in R_q$ ($q \ll p$). A brief description of LLE algorithm is given as follows [14]:

Step 1: Compute k neighbours of every sample.

Step 2: Compute the local reconstruction weight matrix W by minimizing the following cost function:

$$\min \varepsilon(W) = \left| \sum_{i=1}^n w_j^i (f_i - f_{ij}) \right|^2 \quad (2)$$

where k is the number of nearest neighbours used for reconstructing each data point and w_j^i is the weight values.

If f_i and f_j are not neighbours, $w_j^i = 0$ and $\sum_{j=1}^k w_j^i = 1$.

The local covariance matrix $Q_{ja}^i \in R^{k \times k}$ is introduced to calculate the weight values, and

$$Q_{ja}^i = (f_i - f_{ij})^T (f_i - f_{ia}) \quad (3)$$

where f_i and f_{ia} are the neighbours of f_i . Hence, by the means of Lagrange multiplier method, the local reconstruction weight matrix can be obtained as

$$w_j^i = \frac{\sum_{a=1}^k (Q^i)_{ja}^{-1}}{\left(\sum_{b=1}^k \sum_{c=1}^k (Q^i)_{bc}^{-1} \right)} \quad (4)$$

Step 3: Map the original dataset to the embedded coordinates. Compute the reconstructed q -dimensional manifold space S , by minimizing the following constraint

$$\min \varepsilon(H) = \sum_{i=1}^n \left| h_i - \sum_{j=1}^k w_j^i h_{ij} \right|^2 \quad (5)$$

where h_i is the projection vector of f_i in the embedded coordinates, and h_{ij} are the neighbours of h_i . Eq. (5) can be rewritten as

$$\min \varepsilon(H) = \sum_{i=1}^l \sum_{j=1}^k m_j^i h_i^T h_j = \text{tr}(HMH^T) \quad (6)$$

where the cost matrix M can be expressed as

$$M = (I_{l \times l} - W)^T (I_{l \times l} - W) \quad (7)$$

Hence, the minimization of Eq. (6) can be reduced to an eigenvalue problem, and H could be determined by the q smallest nonzero eigenvectors of M .

However, LLE is an unsupervised learning method, and can not use the category information of original data efficiently [15]. To overcome this problem, Ridder [19] proposed a supervised LLE method (SLLE) for classification. SLLE finds low-dimensional representation of the high-dimensional data using the same steps as in LLE excepting the distance calculation of k neighbours. The distance introduced in SLLE is defined by

$$D(f_i, f_j) = \|f_i - f_j\| + \alpha \max \|f_i - f_j\| \Delta \quad (8)$$

where $\alpha \in [0, 1]$ is a tuning parameter, and Δ is the character function defined below.

$$\Delta = \begin{cases} 0 & \text{if } f_i \text{ and } f_j \text{ belong to same class,} \\ 1 & \text{else.} \end{cases} \quad (9)$$

3. Experimental setup and tests

A crack in a gear tooth may cause serious damage to the gear. Thus an early recognition of the gear crack is essential for the regular operation of a gearbox. Experimental tests of a two-stage gear transmission have been

carried out using an experimental setup. The gearbox in our experiment is illustrated in Fig. 1 and the fault simulator setup with sensors is shown in Figs. 2-4. Gear #Z40 is the tested gear. The gearbox is driven at a set input speed using a nominal power 0.4 kW, nominal speed 1500 rpm DC drive motor, and the torque is applied by a nominal power 4 kW, nominal speed 1500 rpm DC generator. Two piezoelectric accelerometers (CA-YD-106) mounted on the flat surface were used for measuring gearbox body acceleration. An optical pick-up sensor was used as a tachometer for the measurement of speed signal on the input shaft (Fig. 5). The software DASP is used for recording the signals.

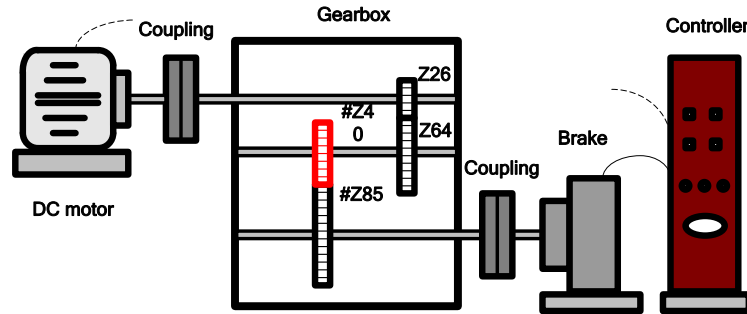


Fig. 1 Gear transmission diagram of the experimental system



Fig. 2 Overview of the gearbox test setup

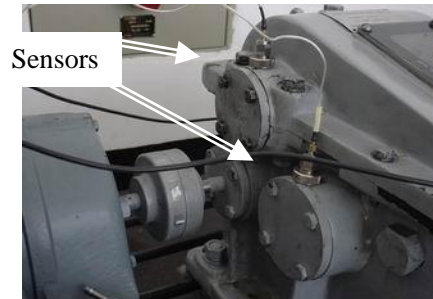


Fig 3 Piezoelectric sensor arrangement

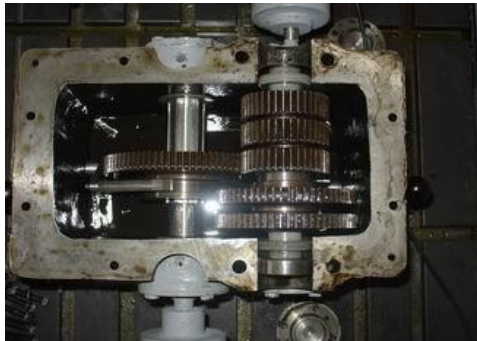


Fig. 4 Insight of the gearbox test setup



Fig. 5 Optical pick-up sensor for speed measurement

In this paper, the slight and serious cracks were processed in the root of gear #40. The slight crack length was up to 5 mm, about 10 percent of the gear face width, and the serious crack length was up to 15 mm, about 30 percent of the gear face width. The vibration signals of normal and faulty gears were collected under 750 rpm of the drive speed and full load. The sampling frequency was 10,000 Hz, and sample length was 19.456 for all conditions. Therefore, 20 data samples were obtained for each gear condition and there were altogether 60 data samples.

Fig. 6 shows the time and FFT frequency spectrum of one sample of the three gear conditions. From the spectrum waveforms we could learn that the vibration signals of different working conditions have been corrupted by heavy noise and have almost the same representation in time and frequency domain. Therefore, it is infeasible and unreliable to recognize the gear states directly using FFT frequency spectrum. For this reason, the new approach based on EMD-SLLE is applied to the gear crack diagnosis in this work.

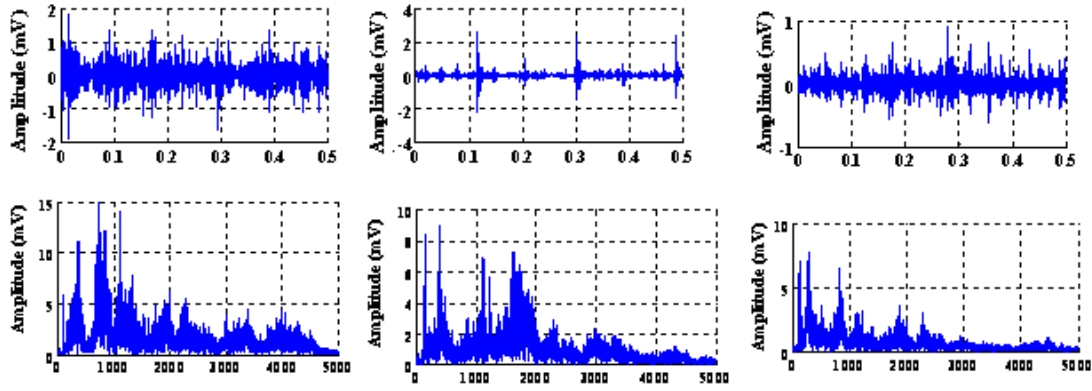


Fig. 6 Time and FFT frequency spectrum of the gear vibration signals: (left) normal, (middle) slight crack and (right) serious crack

4. Application of proposed diagnosis method

As mentioned above, the EMD-SLLE diagnostic approach is proposed for the gear crack detection. EMD is capable for dealing with noised signal, and the intrinsic characteristics could be presented as IMFs. Additionally,

SLLE can preserve distinct nonlinear features of the data of interest in a low dimensional space, which could benefit the pattern recognition of the data. The experimental test results validate that the proposed method is able to recognize different gear crack modes. A flow chart of the proposed diagnosis method is illustrated in Fig. 7.

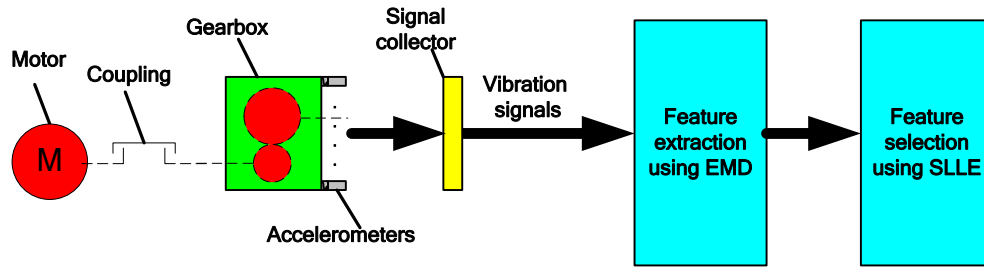


Fig. 7 The flow chart of the diagnostic process

The vibration signals were decomposed into 8 IMFs with EMD in this study. The energy distribution [7] and the kurtosis [18] of IMFs could be good indicators for detection and characterization of early gear damage. Hence, the energy and kurtosis values of each IMF signal were calculated as important features for the detection of incipient gear crack. Moreover, the root mean square (RMS), crest factor (CF), skewness, frequency center (FC), root mean square frequency (RMSF) and standard deviation

frequency (STDF) of each IMF were extracted as additional fault characteristic information. After normalization processing the original feature space $F_{64 \times 60}$ was obtained. The energy distribution of different gear states is shown in Fig. 8. It can be seen that distinguished changes appear in the energy values when crack damage occurs, and vary with the crack severities. Similar changes are also appeared in the rest of statistic features.

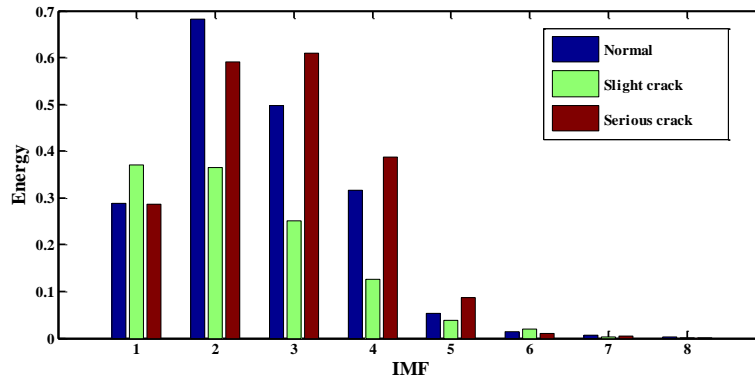


Fig. 8 The energy distribution of different gear operating conditions

As mentioned in section 2, the SLLE was employed to eliminate redundant features. The performance of SLLE and the classical PCA was discussed. Fig. 9 shows the results of redundance reduction using SLLE and PCA, respectively. It is evident from Fig. 9 that three different

gear working states can be identified correctly by SLLE; however, the pattern recognition performance of PCA is unsatisfied. The analysis results indicate that the extraction of nonlinear features can enhance the gear crack identification significantly.

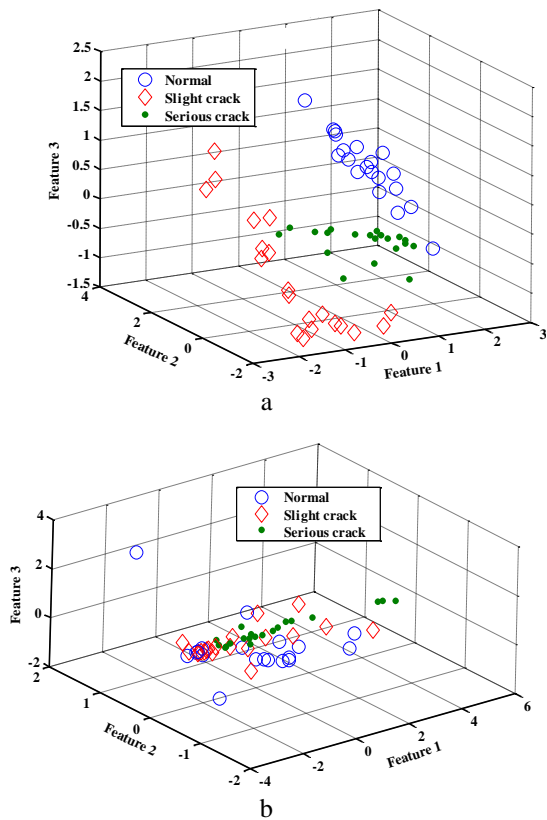


Fig. 9 The performance of feature extraction: a - SLLE; b - classical PCA

The radial basis function (RBF) neural network classifier was used to verify the efficiency of the proposed diagnosis technique and provide automated decision for the detection of rotor multiple faults. In the pattern recognition procedure, there are 30 pieces of data for training and the other 30 pieces for testing the recognition rate. The classification rates of three methods are shown in Table.

Table
Fault diagnosis results of each method with RBF classifier

Feature extraction method	Training rate (%)	Testing rate (%)
EMD	76.67	73.33
EMD-PCA	90.0	86.67
EMD-SLLE	100.0	96.67

From Table we can see, the fault detection rate of EMD-SLLE is higher than that with EMD-PCA or just EMD. For the three patterns, the classification errors of EMD-PCA and EMD are 13.33% and 26.67%, respectively. Contrast with them, the classification error of EMD-SLLE is 1.33%. As a result, we can see that the EMD-SLLE algorithm has better performance than EMD-PCA and EMD.

5. Conclusions

The incipient fault signal often has nonstationary features and is usually heavily corrupted by noise. It is difficult to obtain high-quality features through linear based signal processing methods. The nonlinear approach based on EMD-SLLE is therefore presented for the gear crack detection and diagnosis in this paper. The effectiveness of the proposed method is evaluated and compared

with the linear based scheme in the experimental investigation. The analysis results on experimental data demonstrate that the presented diagnostic approach is feasible and efficient for feature extraction and fault identification of gear cracks. The proposed diagnosis system in this work may provide practical utilities for gear crack diagnosis. Further research is to provide the proposed method to the industrial data of gear fault vibration.

Acknowledgment

This project is sponsored by the grants from the National Natural Sciences Foundation of China (No. 50975213 and No. 50705070) and the Program of Introducing Talents of Discipline to Universities (No. B08031).

References

1. **Fakhfakh, T.; Chaari, F.; Haddar, M.** 2005. Numerical and experimental analysis of a gear system with teeth defects, *International Journal of Advanced Manufacturing Technolog* 25(5-6): 542-550. <http://dx.doi.org/10.1007/s00170-003-1830-8>
2. **Lei, Y.; Zuo, M.J.** 2009. Gear crack level identification based on weighted K nearest neighbor classification algorithm, *Mechanical Systems and Signal Processing* 23(5): 1535-1547. <http://dx.doi.org/10.1016/j.ymssp.2009.01.009>
3. **Zheng, H.; Li, Z.; Chen, X.** 2002. Gear fault diagnosis based on continuous wavelet transform, *Mechanical Systems and Signal Processing* 16(2-3): 447-457. <http://dx.doi.org/10.1006/mssp.2002.1482>
4. **Li, Z.; Yan, X.; Yuan, C.; Peng, Z.; Li, L.** 2011. Virtual prototype and experimental research gear multi-fault diagnosis using wavelet-autoregressive model and principal component analysis method, *Mechanical Systems and Signal Processing* 25(7): 2589-2607. <http://dx.doi.org/10.1016/j.ymssp.2011.02.017>
5. **Saravanan, N.; Ramachandran, K. I.** 2009. Fault diagnosis of spur bevel gear box using discrete wavelet features and Decision Tree classification, *Expert Systems with Applications* 36(5): 9564-9573. <http://dx.doi.org/10.1016/j.eswa.2008.07.089>
6. **Choy, F. K.; Mugler, D. H.; Zhou, J.** 2003. Damage identification of a gear transmission using vibration signatures, *Journal of Mechanical Design, Transactions of the ASME* 125(2): 394-403. <http://dx.doi.org/10.1115/1.1564571>
7. **Lei, Y.; Zuo, M.J.; He, Z.; Zi, Y.** 2010. A multidimensional hybrid intelligent method for gear fault diagnosis, *Expert Systems with Applications* 37(2): 1419-1430. <http://dx.doi.org/10.1016/j.eswa.2009.06.060>
8. **Cheng, J.; Yu, D.; Yang, Y.** 2008. A fault diagnosis approach for gears based on IMF AR model and SVM, *Eurasip. J. Adv. Sign. Process.*, 2008: 1-7.
9. **Lin, J.; Zuo, M.J.** 2003. Gearbox fault diagnosis using adaptive wavelet filter, *Mechanical Systems and Signal Processing* 17(6): 1259-1269. <http://dx.doi.org/10.1006/mssp.2002.1507>
10. **Wu, J.; Chan, J.** 2009. Faulted gear identification of a rotating machinery based on wavelet transform and artificial neural network, *Expert Systems with Applications* 36(5): 8862-8875. <http://dx.doi.org/10.1016/j.eswa.2008.11.020>

11. **Kim, B.; Kim, G.; Lee, H.** 2008. Ex situ plasma diagnosis by recognition of X-ray photoelectron spectroscopy data using a neural network, *Materials and Manufacturing Processes* 23(5): 528-532.
<http://dx.doi.org/10.1080/10426910802104310>
12. **Belkin, M.; Niyogi, P.** 2003. Laplacian eigenmaps for dimensionality reduction and data representation, *Neural Computation* 15(6): 1373-1396.
<http://dx.doi.org/10.1162/089976603321780317>
13. **Tenenbaum, J. B.; Silva, V.; Langford, J.C.** 2000. A global geometric framework for nonlinear dimensionality reduction, *Science* 290: 2319-2323.
<http://dx.doi.org/10.1126/science.290.5500.2319>
14. **Roweis, S.; Saul, L.** 2000. Nonlinear dimensionality reduction by locally linear embedding, *Science* 290: 2323-2326.
<http://dx.doi.org/10.1126/science.290.5500.2323>
15. **Jiang, Q.; Jia, M.; Hu, J.; Xu, F.** 2009. Machinery fault diagnosis using supervised manifold learning, *Mechanical Systems and Signal Processing* 23(7): 2301-2311.
<http://dx.doi.org/10.1016/j.ymsp.2009.02.006>
16. **Yang, J.; Xu, J.; Yang, D.** 2006. Noise reduction method for nonlinear time series based on principal manifold learning and its application to fault diagnosis, *Chinese Journal of Mechanical Engineering* 42(8): 154-158.
<http://dx.doi.org/10.3901/JME.2006.08.154>
17. **Huang, N.E.; Shen, Z.; Long, S. R. et al.** 1998. The empirical mode decomposition and the Hubert spectrum for nonlinear and non-stationary time series analysis, *Proc. R. Soc. A Math. Phys. Eng. Sci.* 454(1971): 903-995.
<http://dx.doi.org/10.1098/rspa.1998.0193>
18. **Parey, A. and Tandon, N.** 2007. Impact velocity modelling and signal processing of spur gear vibration for the estimation of defect size, *Mechanical Systems and Signal Processing* 21(1): 234-243.
<http://dx.doi.org/10.1016/j.ymsp.2005.12.011>
19. **Ridder, D.; Kouropteva, O.; Okun, O. et al.** 2003. Supervised locally linear embedding, *Lecture Notes in Artificial Intelligence* 2714: 333-341.

Zhixiong Li, Xiping Yan, Yu Jiang, Li Qin, Jingping Wu

NAUJAS DUOMENŲ PARINKIMO TRAKTAVIMAS
PLYŠIO LYGIUI KRUMPLIARATYJE NUSTATYTI
NAUDOJANT DAUGIARIOPĄ TYRIMĄ

R e z i ū m ė

Plyšys krumpliaratyje yra įprastas krumpliaratinių mechanizmų pažeidimo tipas, nes netikėtai atsiradęs didelis plyšys gali suardyti pavara ir padaryti nemažą ekonominių nuostolių. Efektyvus prasidedančio pažeidimo nustatymas ir diagnozė yra svarbus normaliam mechanizmo darbui. Vienas iš svarbiausių momentų diagnozuojant pažeidimą yra jo tipo nustatymas ir identifikavimas. Literatūros apžvalga rodo kad ribotas tyrimas nustatant netiesines kontūro savybes pagal daugiariopus tyrimo algoritmus mechaniniams pažeidimams aptikti ir netiesinis tipo nustatymas plyšiu krumpliaratyje aptikti yra nepakankami. Šiame straipsnyje aprašomas naujas duomenų parinkimo

metodas, paremtas empiriniu dekompozicijos metodu ir lokaliu tiesiniu duomenų panaudojimu plyšio lygiui krumpliaratyje nustatyti. Empirinis dekompozicijos metodas taikomas svyravimų signalams išskirti į atskiras būdingas funkcijas pažeidimo tipui nustatyti, o lokalius tiesinius duomenis panaudoti defekto savybėms nustatyti. Eksperimentiniai svyravimų duomenys gauti naudojant krumpliaratinių pažeidimų bandymo įrangą buvo apdoroti ir panaudoti būklei įvertinti pasiūlytuju metodu. Tyrimo rezultatai parodė, kad taikant pasiūlytą metodiką gali būti nustatytos įvairių krumpliaratinių plyšio tikslios svarbiausių svyravimų signalų tarpusavio charakteristikos. Energijos pasiskirstymas ir statistinės IMF charakteristikos kinta keičiantis krumpliaratinių eksploatacijos sąlygoms, ir daugiausia išsiskiriantys atvejai gali būti nustatyti taikant pasiūlytąjį lokalinį tiesinį duomenų nustatymo metodą. Taisyti straipsnyje pasiūlytą charakteristikų nustatymo metodą yra geriau nei svarbiausių komponentų analizės metodą.

Zhixiong Li, Xiping Yan, Yu Jiang, Li Qin, Jingping Wu

A NEW DATA MINING APPROACH FOR GEAR
CRACK LEVEL IDENTIFICATION BASED ON
MANIFOLD LEARNING

S u m m a r y

Gear crack is a common damage model in the gear mechanisms, and an unexpected serious crack may break the transmission system down, leading to significant economic losses. Efficient incipient fault detection and diagnosis are therefore critical to machinery normal running. One of the key points of the fault diagnosis is feature extraction and selection. Literature review indicates that only limited research considered the nonlinear property of the feature space by the use of manifold learning algorithms in the field of mechanic fault diagnosis, and nonlinear feature extraction for gear crack detection are scarce. This paper reports a novel data mining method based on the empirical mode decomposition (EMD) and supervised locally linear embedding (SLLE) applied to gear crack level identification. The EMD was used to decompose the vibration signals into a number of intrinsic mode functions (IMFs) for feature extraction, whilst the SLLE for nonlinear feature selection. The experimental vibration data acquired from the gear fault test-bed were processed for feature reduction and extraction using the proposed method. Study results show that the sensitive characteristics between different gear crack severity vibration signals can be revealed effectively by EMD-SLLE. The energy distribution and the statistic features of IMFs vary with the change of the gear operation conditions, and the most distinguished features can be extracted by nonlinear method of SLLE. In addition, the performance of feature extraction of SLLE is better than that of the linear method of principal component analysis (PCA).

Keywords: gear, crack level, data mining.

Received February, 17, 2011
Accepted January 11, 2012