

A feature extraction method for shearer cutting pattern recognition based on improved local mean decomposition and multi-scale fuzzy entropy

Lei Si^{1,2}, Zhongbin Wang^{1,*}, Chao Tan¹, Xinhua Liu^{1,3} and Xihua Xu¹

¹School of Mechatronic Engineering, and

²School of Information and Electrical Engineering, China University of Mining and Technology, Xuzhou 221116, China

³State Key Laboratory of Robotics and System, Harbin Institute of Technology, Harbin 150080, China

Aiming at accurately identifying shearer cutting patterns, this article proposes a new feature extraction method based on improved local mean decomposition (LMD) and multi-scale fuzzy entropy (MFE). The cubic trigonometric Hermite interpolation was used to calculate local mean and envelope estimate functions to improve LMD decomposition results and a sum of product functions was acquired. Furthermore, MFE, referring to the calculation of fuzzy entropy over a range of scales, was designed to measure the complexity and self-similarity of vibration signals and extract the features from the decomposition results. Subsequently, the obtained feature vectors were fed into two classifiers of support vector machine and back propagation neural network to realize the cutting pattern recognition. The experimental results indicate the applicability and effectiveness of the methodology and demonstrate that the proposed algorithm could perform better in identifying different cutting categories of shearer.

Keywords: Feature extraction, local mean decomposition, multi-scale fuzzy entropy, shearer cutting pattern.

NOWADAYS, vibration signal analysis techniques are being widely used and are effective in the fields of fault diagnosis and pattern recognition^{1,2}. Accurately identifying the cutting patterns of a shearer can improve the efficiency of coal mining, reduce wear on the machine components and prolong its life. Due to the close relationship between shearer rocker vibration and cutting conditions, it is feasible and reasonable to monitor the shearer condition and diagnose the cutting patterns on the basis of vibration signal analysis.

However, the vibration of shearer rocker usually displays strong nonlinear, non-Gaussian and non-stationary characteristics. Hence the state feature information cannot be extracted accurately from the vibration signals only in the time or frequency domain. To extract the

feature information, many signal analysis methods have been developed such as time-domain statistical features, short-time Fourier transformation, wavelet transform, Wigner–Ville distribution, and so on^{3–6}. However, these traditional signal processing methods have their own drawbacks, especially the lack of self-adaptive feature^{7,8}.

Unlike the above methods, empirical mode decomposition (EMD)⁹ and local mean decomposition (LMD)¹⁰ have been applied in many fields^{11–17}. EMD displays some defects of the end effect and mode mixing problems, etc. which limits its applications^{18,19}. LMD can adaptively decompose a complicated multi-component signal into several product functions (PFs). However, there are still some shortcomings that need to be addressed, such as the distortion of decomposition results and lower convergence speed. To overcome the defects of original LMD, many researchers have developed improved measures, such as cubic spline interpolation (CSI) method^{20,21}, cubic Hermite interpolation (CHI) method^{22,23}, rational Hermite interpolation (RHI) method^{24–28} and so on. However, these methods have their own drawbacks. In this article, an improved LMD method based on the cubic trigonometric Hermite interpolation (CTHI) with shape parameters is proposed and the actual experiment data are used to verify the effectiveness of the CTHI–LMD method.

After the original vibration signals are decomposed, a major goal is to extract the features from the PF components with more cutting pattern information. As a statistical measure method, approximate entropy (AppEn) and sample entropy (SamEn) have been proposed^{29,30}, which can only reflect the information of time series in a single scale. Costa *et al.*^{31,32} developed a multi-scale entropy (MSE) method on the basis of SamEn, which has been used in fault diagnosis of rolling bearing^{33,34}. For a shearer, the vibration signals of different cutting patterns possess diverse complexity in various timescales and the entropy values also differ from each other. Multi-scale fuzzy entropy (MFE) can evaluate the self-similarity of original data and provide uncertain and unsatisfactory analysis^{35–37}. Hence, MFE is used as the feature extractor

*For correspondence. (e-mail: wangzbpaper@126.com)

to acquire the feature information from the decomposed PF components of the CTHI-LMD method.

Naturally, the extracted features are used to train a classifier to automatically recognize the categories of shearer cutting pattern. In the fields of fault detection and condition monitoring, there are many pattern recognition techniques based on artificial intelligence algorithms, such as artificial neural networks³⁸, adaptive neuro-fuzzy inference system³⁹, support vector machine (SVM)⁴⁰, and so on. In this article, two classifiers of SVM and back propagation neural network (BPNN) have been constructed based on the extracted feature vectors to automatically identify the cutting patterns of the shearer, and experiments performed to verify the applicability and effectiveness of the proposed feature extraction method to the vibration signals of the shearer rocker.

Improved local mean decomposition

According to the literature on LMD, we can obtain the following decomposition result

$$x(t) = \sum_{i=1}^k PF_i(t) + u_k(t), \tag{1}$$

where $x(t)$ is the original signal, $PF_i(t)$ denotes the i th PF component, k the number of PFs and $u_k(t)$ is a residual signal after decomposition.

However, the original LMD method has many shortcomings. In this article, CTHI with shape parameters is introduced for LMD.

Cubic trigonometric Hermite interpolation algorithm

The basic function of CTHI can be constructed using trigonometric functions.

Definition 1: For a given real number λ and $0 \leq t \leq 1$, the four functions

$$\begin{cases} f_i(t) = \lambda S^2 - \lambda S^3 + C^3, \\ f_{i+1}(t) = 1 - \lambda S^2 + \lambda S^3 - C^3, \\ g_i(t) = \frac{2}{\pi}(-\lambda + S + \lambda S^2 - S^3 + \lambda C^3), \\ g_{i+1}(t) = \frac{2}{\pi}(-C + \lambda S^2 - \lambda S^3 + C^3), \end{cases} \tag{2}$$

can be described as the basis functions of CTHI with parameter λ , where $S = \sin(\pi t/2)$, $C = \cos(\pi t/2)$.

After the calculations, the defined basic functions should satisfy the following conditions

$$\begin{aligned} f_i(0) &= 1, f_{i+1}(0) = 0, g_i(0) = 0, g_{i+1}(0) = 0, \\ f_i(1) &= 0, f_{i+1}(1) = 1, g_i(1) = 0, g_{i+1}(1) = 0, \\ f'_i(0) &= 0, f'_{i+1}(0) = 0, g'_i(0) = 1, g'_{i+1}(0) = 0, \\ f'_i(1) &= 0, f'_{i+1}(1) = 0, g'_i(1) = 0, g'_{i+1}(1) = 1, \\ f_i(t) + f_{i+1}(t) &= 1, g_i(t) = -g_{i+1}(1-t). \end{aligned} \tag{3}$$

It can be summarized that the basic functions of CTHI possess the same properties as CHI. By introducing the parameter λ , we can adjust the shape of the splines by setting different values for it, which can guarantee enough physical sense for the decomposed PFs.

Definition 2: Given a partition of range $[a, b]$: $a = x_0 < x_1 < x_2 < \dots < x_N = b$ and a discrete date (x_i, y_i, d_i) , y_i is the local maximum or minimum at time x_i and d_i is the first-order derivative at time x_i . Then the following formula can be obtained

$$TH_i(x) = f_i(t)y_i + f_{i+1}(t)y_{i+1} + g_i(t)h_i d_i + g_{i+1}(t)h_i d_{i+1}, \tag{4}$$

where $TH_i(x)$ is the cubic trigonometric Hermite interpolation spline of each partition $[x_i, x_{i+1}]$ in the range $[a, b]$, $h_i = x_{i+1} - x_i$, $t = (x - x_i)/h_i$ and $f_i(t)$, $f_{i+1}(t)$, $g_i(t)$ and $g_{i+1}(t)$ are the basic functions of CTHI. It can be proved easily that the CTHI splines are C^1 continuous.

Selection of optimum envelopes

Using CTHI, the envelopes of LMD can be constructed reasonably well. According to the above analysis, the shapes of the envelopes will correspondingly change when the parameter λ is set to different values, which can affect the decomposed PF components. On the basis of calculation time and accuracy, the range λ is set as $[-20, 20]$ in the present study²⁸. In order to select optimum envelop with proper λ , a novel evaluation criterion is proposed to implement the selection of optimum envelopes in each rank.

Grey correlation degree (GCD) is a quantitative indicator to describe the correlation between objects in the grey correlation analysis method. Therefore, the GCD is utilized as the evaluation criterion to select the optimum PF (OPF) component here, and the algorithm process is as follows.

Assume that the original signal $x(t) = \{x_0(t), t = 1, 2, \dots, T\}$ and the decomposed PF components with different λ are denoted as $PF_i^1(t), PF_i^2(t), \dots, PF_i^p(t)$, $PF_i^j(t) = \{y_i^j(t), t = 1, 2, \dots, T; j = 1, 2, \dots, p\}$, where p is the width of range λ . The correlation coefficient ξ between $x(t)$ and $PF_i^j(t)$ can be calculated as follows

$$\xi_i^j(t) = \frac{\min_j \min_t |x_0(t) - y_i^j(t)| - \gamma \max_j \max_t |x_0(t) - y_i^j(t)|}{|x_0(t) - y_i^j(t)| - \gamma \max_j \max_t |x_0(t) - y_i^j(t)|}, \quad (5)$$

where γ is a constant and $\gamma \in (0, 1]$. Here, γ is set as 0.5.

The GCD (R) between $x(t)$ and $PF_i^j(t)$ can be calculated using the mean of all correlation coefficients

$$R_i^j = \frac{1}{T} \sum_{t=1}^T \xi_i^j(t). \quad (6)$$

The PF component with the biggest R can be selected as the optimum PF with the optimum envelopes in each sifting process.

Design implementation of CTHI-LMD

An improved local mean decomposition (LMD) method based on cubic trigonometric Hermite interpolation (CTHI) is proposed here:

(1) Identify all the local extrema n_i of the original signal $x(t)$ and generate the upper envelope $E_u(t)$ and lower envelope $E_l(t)$ using the CTHI algorithm. For different values of the shape controlling parameter λ , the corresponding upper and lower envelopes can be denoted as $E_u^1(t), E_u^2(t), \dots, E_u^p(t)$ and $E_l^1(t), E_l^2(t), \dots, E_l^p(t)$.

(2) Calculate $m_i(t)$ and $a_i(t)$ using $E_u(t)$ and $E_l(t)$, denoted as $m_i^1(t), m_i^2(t), \dots, m_i^p(t)$ and $a_i^1(t), a_i^2(t), \dots, a_i^p(t)$ respectively

$$\begin{cases} m_i^j(t) = \frac{E_u^j(t) + E_l^j(t)}{2}, \\ a_i^j(t) = \frac{|E_u^j(t) - E_l^j(t)|}{2}, j = 1, 2, \dots, p. \end{cases} \quad (7)$$

(3) According to the decomposition steps of LMD, we can obtain p PFs with different values of λ : $PF_i^1(t), PF_i^2(t), \dots, PF_i^p(t)$. The optimum PF (OPF $_i(t)$) should be selected from the p PFs based on the GCD criterion.

(4) The residual $u_i(t) = x(t) - OPF_i(t)$ is then taken as the original signal and the above procedures are repeated until $u_i(t)$ is a constant or monotonic. Considering the orthogonality of each PF component, the orthogonality criterion (OC) is used in the proposed CTHI-LMD method to overcome the drawback of PF criteria. OC can be defined as

$$OC = \frac{\left| \sum_{t=0}^T x(t)m_{in}(t) \right|}{\left| \sum_{t=0}^T [x(t) - m_{in}(t)]m_{in}(t) \right|}, \quad (8)$$

where $m_{ik}(t)$ is the local mean function of the i th PF component at sifting iteration step n .

Using OC as the PF criteria cannot only guarantee the orthogonality of PF components, but also reduce the iteration number and duration of the decomposition process. Figure 1 depicts a flowchart of the proposed CTHI-LMD method.

Feature extraction based on multi-scale fuzzy entropy

Fuzzy entropy

As an improvement of AppEn and SamEn⁴¹, fuzzy entropy (FuzEn) employs a fuzzy function, such as exponential

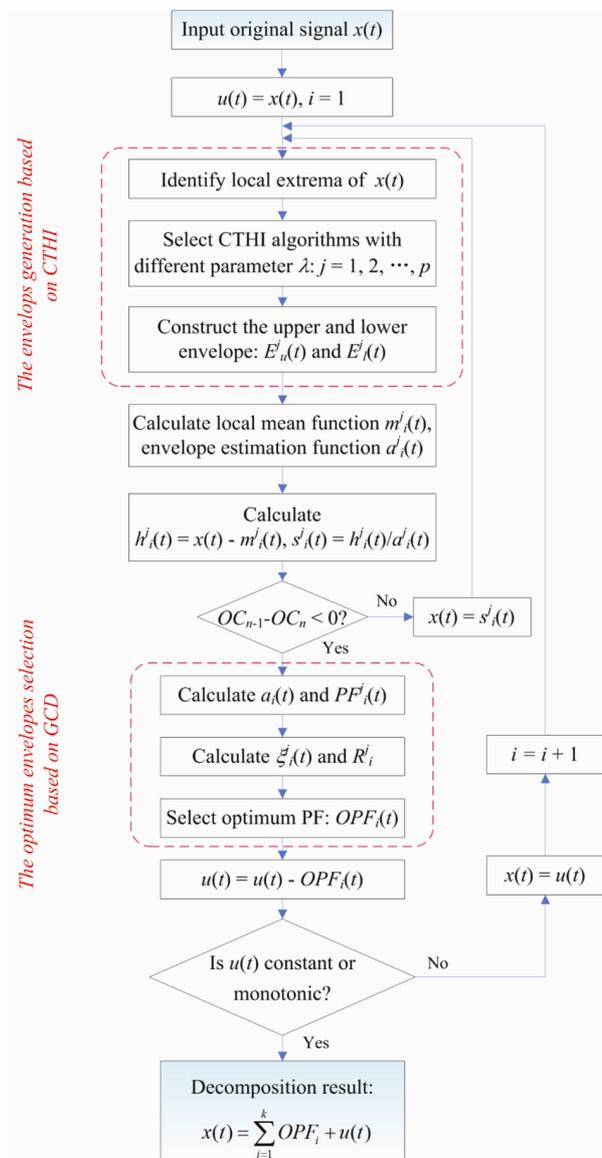


Figure 1. Flowchart of the improved local mean decomposition method.

function $e^{-(d/r)^n}$ to measure the similarity, which can avert the drawbacks of AppEn and SamEn. The specific steps of FuzEn are described as follows.

(1) For a given time series $\{u(i), 1 \leq i \leq N\}$, where N is the length, the m -dimensional vector can be constructed as

$$X_i^m = \{u(i), u(i+1), \dots, u(i+m-1)\} - u_0(i),$$

$$i = 1, 2, \dots, N - m + 1, \tag{9}$$

where $u_0(i)$ is the mean of the m elements, calculated as

$$u_0(i) = \frac{1}{m} \sum_{j=0}^{m-1} u(i+j). \tag{10}$$

(2) The maximum distance between X_i^m and X_j^m is defined as d_{ij}^m

$$d_{ij}^m = \max_{k \in (0, m-1)} \{|[u(i+k) - u_0(i)] - [u(j+k) - u_0(j)]|\},$$

$$i \neq j = 1, 2, \dots, N. \tag{11}$$

(3) The similarity degree D_{ij}^m of X_i^m and X_j^m can be calculated as follows

$$D_{ij}^m = e^{-(d_{ij}^m/r)^n}, \tag{12}$$

where n and r denote the gradient and width of the border respectively.

(4) The following formulas can be defined as

$$\phi^m(n, r) = \frac{1}{N-m} \sum_{i=1}^{N-m} \left(\frac{1}{N-m-1} \sum_{\substack{j=1 \\ i \neq j}}^{N-m} D_{ij}^m \right),$$

$$\phi^{m+1}(n, r) = \frac{1}{N-m} \sum_{i=1}^{N-m} \left(\frac{1}{N-m-1} \sum_{\substack{j=1 \\ i \neq j}}^{N-m} D_{ij}^{m+1} \right). \tag{13}$$

(5) The fuzzy entropy can be calculated as follows

$$\text{FuzEn}(m, n, r) = \lim_{N \rightarrow \infty} [\ln \phi^m(n, r) - \ln \phi^{m+1}(n, r)]. \tag{14}$$

Usually N is finite, $\text{FuzEn}(m, n, r)$ can be expressed as follows

$$\text{FuzEn}(m, n, r) = \phi^m(n, r) - \ln \phi^{m+1}(n, r). \tag{15}$$

Multi-scale fuzzy entropy

In analogy to MSE, the calculation method of MFE mainly includes two procedures, summarized as given below.

(1) The coarse-grained vectors of original time series are constructed as follows

$$y_j(\tau) = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i, \quad 1 \leq j \leq \frac{N}{\tau}, \tag{16}$$

where the scale factor $\tau = 1, 2, \dots$.

(2) FuzEn of each coarse-grained time series can be calculated using eqs (9)–(15) and then plotted as the function of scale factor for MFE analysis.

The parameters of MFE can be selected as follows: $m = 2$, $r = 0.1 \sim 0.25 \cdot \text{SD}$ (SD is the standard deviation the original time series), $r = 0.15 \cdot \text{SD}$, $n = 2$, $\tau_{\max} = 15$.

Experimental validation

Collection of vibration signals

Figure 2 shows a self-designed experimental system for shearer cutting coal. The coal seam was composed of three parts: $f = 2$, $f = 3$ and coal seam with some gangues (f is the Protodikonov's hardness coefficient). Thus, the shearer mainly displayed four cutting patterns, including the idling pattern, and three kinds of coal seams, which could be represented by the symbols of F1, F2, F3 and F4 respectively. We collected 240 data samples with 60 data samples under each cutting pattern to verify the proposed algorithm; each sample contained 5000 data. Figure 3 shows vibration acceleration signals under the four cutting patterns.

Comparative analysis of different decomposed methods

In order to verify the effectiveness of the CTHI-LMD method, three other methods, i.e. original LMD, CHI-LMD and RHI-LMD were employed to decompose the vibration signals. The following four indicators of orthogonal index (OI), energy difference (ED), number of iterations for obtaining each PF component and consuming time were used to evaluate the decomposition performance in detail as follows

(1) The orthogonal index can be calculated as

$$\text{OI} = \frac{\sum_{i=1}^{N_{\text{PF}}} \sum_{j=1}^{j < i} \left| \sum_{k=1}^N \text{PF}_{ik} \times \text{PF}_{jk} \right|}{\sum_{k=1}^N (x_k - u_k)^2}, \tag{17}$$

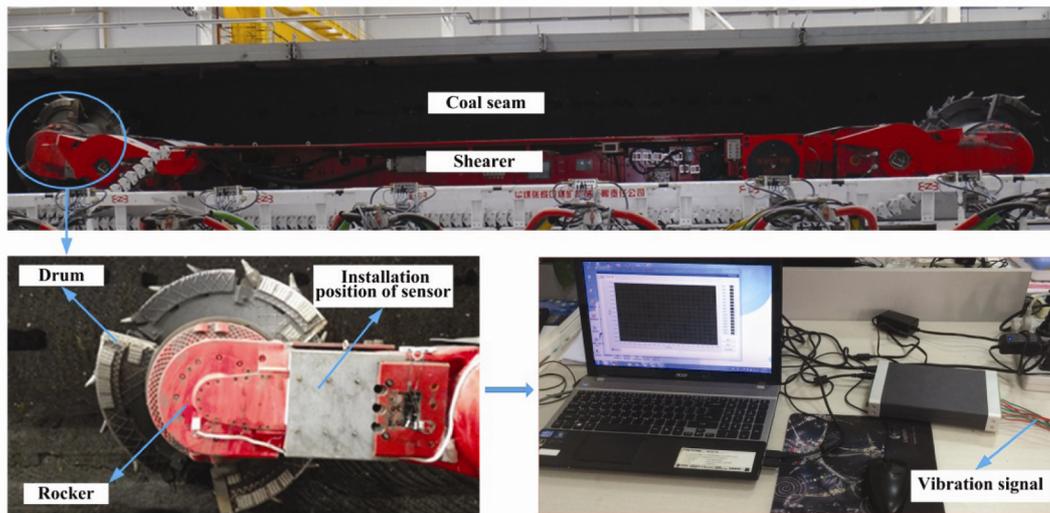


Figure 2. The self-designed experimental system for shearer cutting coal.

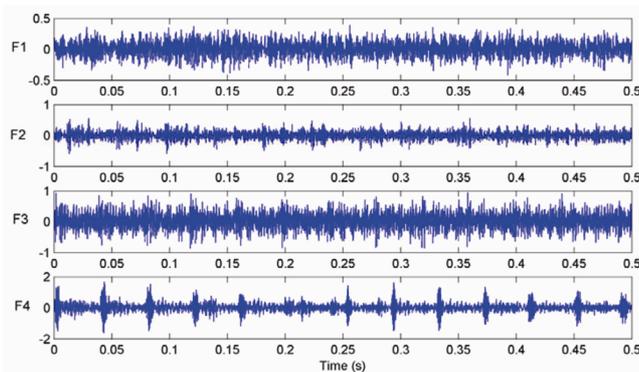


Figure 3. The vibration acceleration signals of each shearer cutting pattern.

where N_{PF} and N denote the number of PF components and length of the signal respectively.

(2) The energy difference can be calculated as

$$ED = \frac{\left(\sum_{i=1}^{N_{PF}} \int_0^{+\infty} |PF_i(t)|^2 dt \right) + \int_0^{+\infty} |u(t)|^2 dt - \int_0^{+\infty} |x(t)|^2 dt}{\int_0^{+\infty} |x(t)|^2 dt} \times 100\% \quad (18)$$

(3) The number of iterations for each PF: In the LMD theory, there is no consistent mathematical model to define the best envelope and smaller iteration time of each PF will shorten the decomposition time and improve calculation efficiency.

(4) Consuming time: For different decomposition methods, the consuming time should be taken into consideration.

In the simulation experiment, the vibration signal of shearer cutting pattern F4 was selected as the analysis object and was decomposed into a sum of PF components (OPF components for CTHI-LMD) based on the four methods respectively. Figures 4–7 show the corresponding decomposition results. It should be noted that since the front PF or OPF components contain the maximum pattern information, here we have plotted only the first four PF or OPF components due to lack of space.

It is clear from Figures 4 to 7 that some difference exists in the PF or OPF components derived from the four methods. This is because different ways are adopted to calculate the local mean value function and envelope estimation function. At the same time, different iterative termination conditions are also used in the four methods. In order to directly show the decomposition performance of the four methods, a future comparison was carried out. The four evaluating indicators were employed to quantify the decomposition performance, and Table 1 provides the results of the comparison.

From Table 1, the following conclusions can be drawn. First, the OI value of CTHI-LMD (0.0643) is smaller than that of RHI-LMD (0.0954), CHI-LMD (0.1246) and LMD (0.1708), which indicates the proposed decomposition method possesses better orthogonality property. Secondly, CTHI-LMD has a smaller energy difference than the other three methods. This indicates that the energy value of CTHI-LMD decomposition results closer to the energy of the original signal. Meanwhile, the iteration number of the first four OPFs obtained by CTHI-LMD (7, 6, 7, 5) is smaller than that of PFs obtained by other three methods, which demonstrates that the envelopes of LMD generated by the CTHI algorithm are superior. Lastly, CTHI-LMD will consume more time (18.7516 sec) than RHI-LMD (14.7567 sec), CHI-LMD (11.4128 sec) and LMD (10.5764 sec). In general, the proposed decomposition

method is demonstrated to be more suitable for vibration signals of the shearer rocker.

Multi-scale fuzzy entropy analysis for vibration signals

The MFE method was utilized to extract features from the original signals and the first components (OPF1) under 15 scales for four cutting patterns. Figures 8 and 9 illustrate the MFE curves of the original signals and OPF1 components over 15 scales of the cutting data under four cutting conditions.

As seen from Figures 8 and 9, it is not easy to identify the four cutting patterns just from the MFE curves, especially the two cutting conditions of F2 and F3. In Figure 8, obviously, the cutting pattern F2 cannot be separated from the cutting pattern F3 based on the corresponding MFE curves. In addition, the FuzEn values of the shearer with idling pattern are distinctly larger than others when

the scale factor $\tau \geq 3$, which indicates that the vibration signals of idling pattern are more complex than those of shearer cutting different coal seams. Meanwhile, the shearer with cutting pattern F4 has larger FuzEn values than the other two cutting patterns (F2 and F3).

Application to shearer cutting pattern recognition

A classifier based on SVM was applied for classification of the four categories. The pattern features extracted by MFE of OPF1 over 15 scales were used to construct the feature vector, and then train and test the SVM classifier. However, not all FuzEn values in 15 scales have a close relationship to the cutting patterns and a large dimension of input vector for SVM would consume more time and decrease the classification accuracy rate. In order to refine the extracted features, the Laplacian score (LS) algorithm⁴² was employed to rank the FuzEn values over different scales according to their importance. Based on

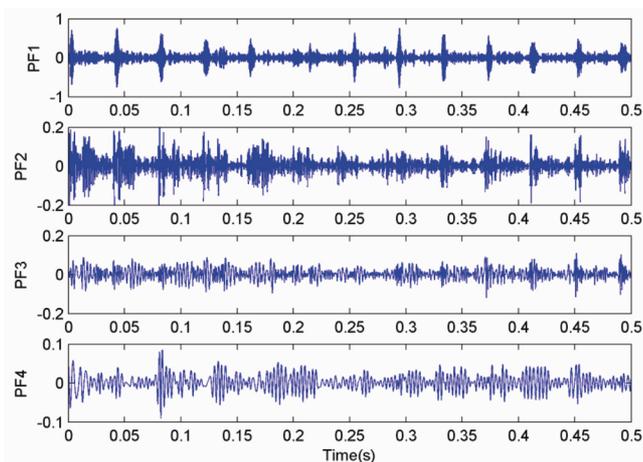


Figure 4. LMD decomposition results of the vibration signal of shearer rocker with cutting pattern F4.

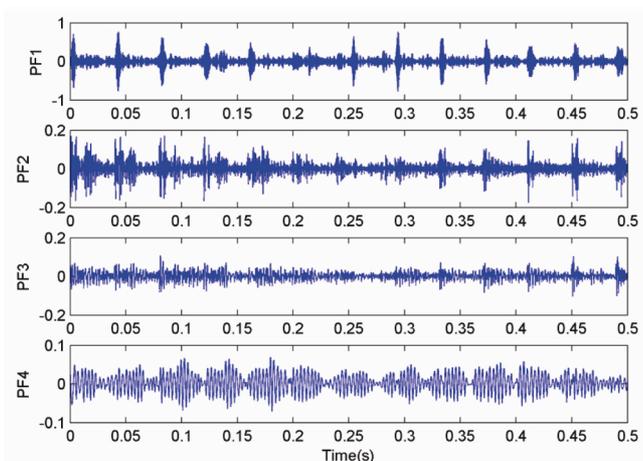


Figure 5. Cubic Hermite interpolation based LMD decomposition results of the vibration signal of shearer rocker with cutting pattern F4.

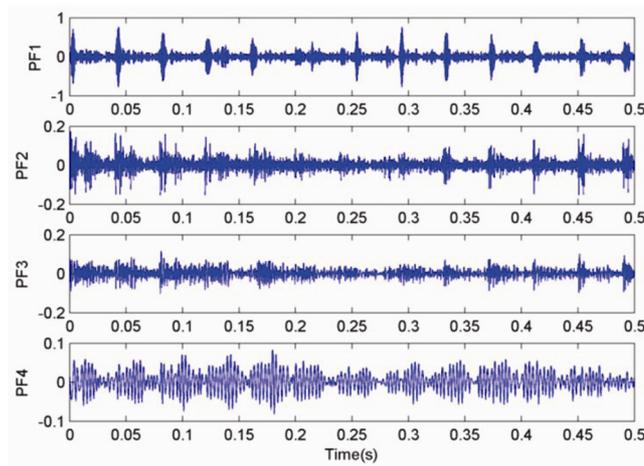


Figure 6. Rational Hermite interpolation based LMD decomposition results of the vibration signal of shearer rocker with cutting pattern F4.

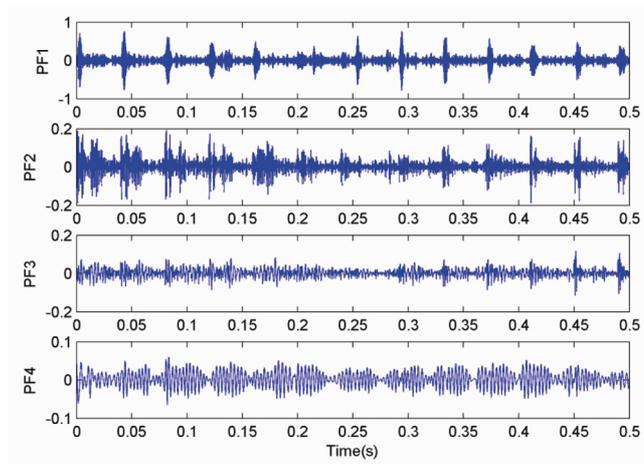


Figure 7. CTHI-LMD decomposition results of the vibration signal of shearer rocker with cutting pattern F4.

Table 1. Comparison between LMD, CHI-LMD, RHI-LMD and CTHI-LMD methods using four evaluating indicators

Method	Orthogonal index	Energy difference (%)	Number of iterations				Consuming time (sec)
			PF1	PF2	PF3	PF4	
LMD	0.1708	9.25	16	14	12	12	10.5764
CHI-LMD	0.1246	6.77	13	9	11	10	11.4128
RHI-LMD	0.0954	3.04	9	6	10	8	14.7467
CTHI-LMD	0.0643	2.74	7	6	7	5	18.7516

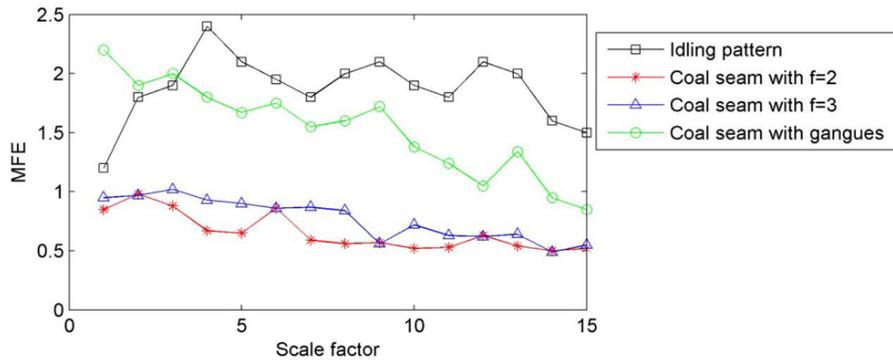


Figure 8. Multi-scale fuzzy entropy (MFE) over 15 scales of original signals.

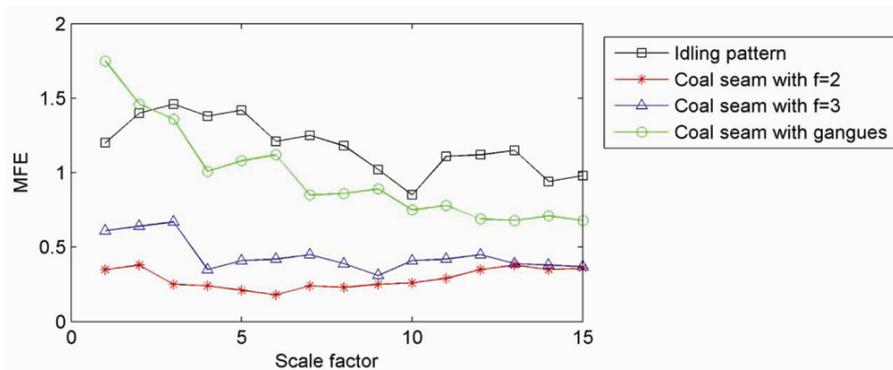


Figure 9. MFE over 15 scales of optimal components decomposed from the CTHI-LMD method.

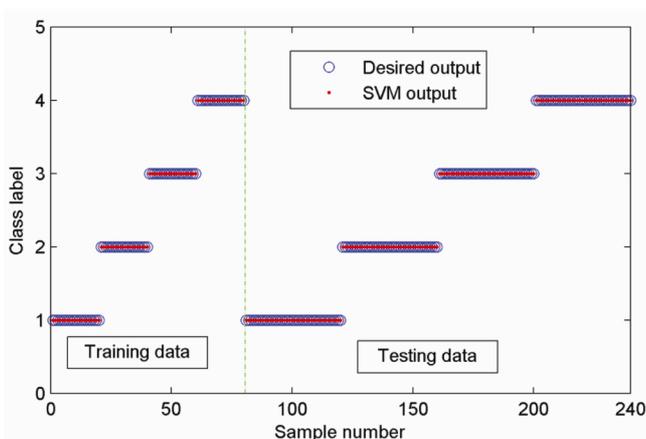


Figure 10. Classification results of support vector machine with the proposed method.

the procedures described in Li *et al.*¹⁵, the LS of each feature could be calculated easily and then ranked as follows

$$LS_{10} < LS_5 < LS_{15} < LS_{11} < LS_7 < LS_{13} < LS_6 < LS_{14} < LS_{12} < LS_8 < LS_9 < LS_3 < LS_1 < LS_4 < LS_2.$$

It should be noted that the features with smaller LS values contain more important information. Thus the first five features in the front ($\tau = 10, 5, 15, 11, 7$) with the most important information were chosen as feature vectors for the SVM model. In addition, 20 samples with different cutting conditions, randomly selected from the dataset, were used to train the SVM. The residual 160 samples were used to test its recognition performance and the class labels of the four cutting patterns were marked

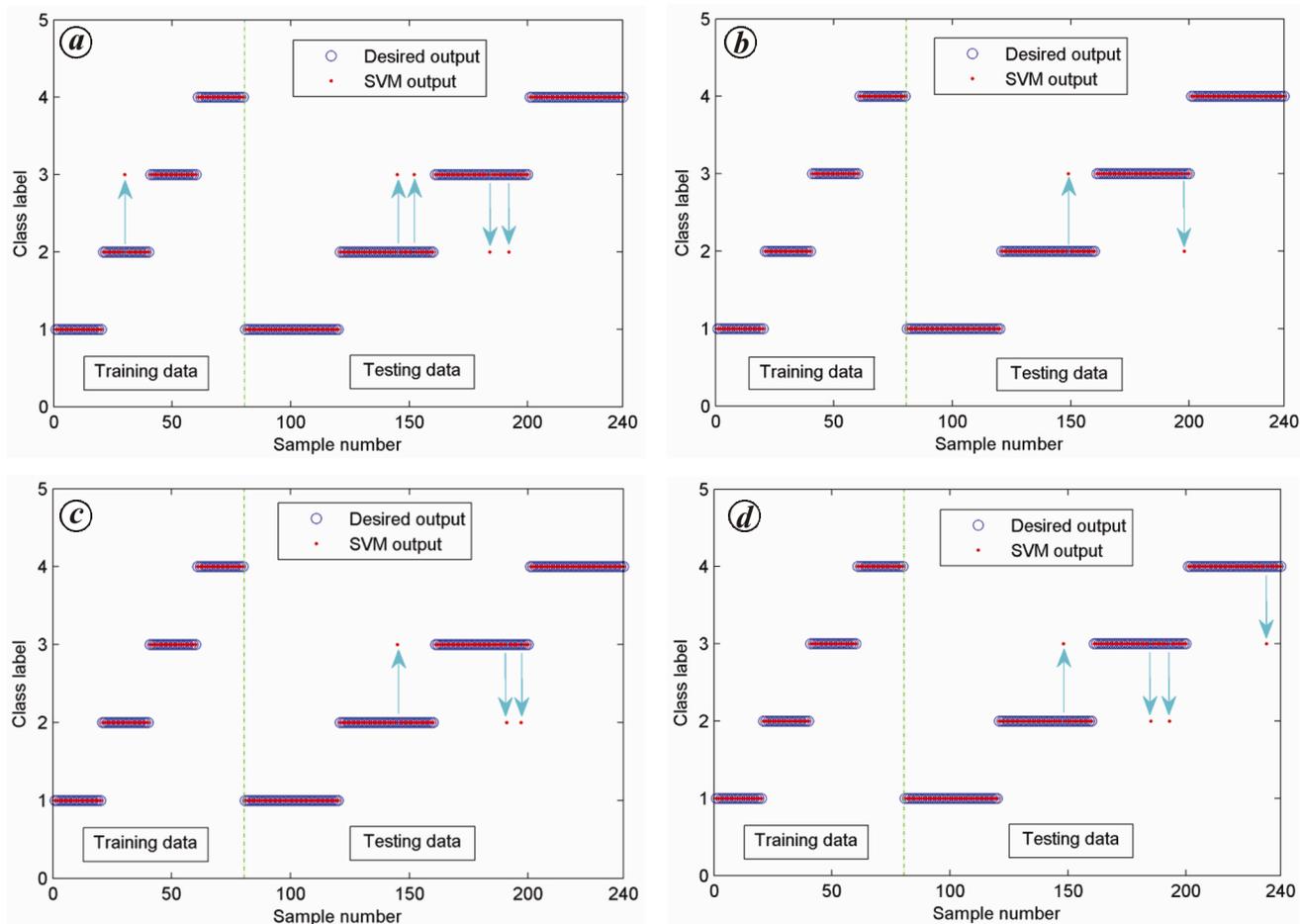


Figure 11. Classification results of SVM with different methods. *a*, SVM without preprocessing by the decomposition methods. *b*, SVM with RHI-LMD and MFE. *c*, SVM with CHI-LMD and MFE. *d*, SVM with LMD and MFE.

as 1–4. Furthermore, the genetic algorithm (GA) was utilized to optimize the parameters of SVM. Finally, the classification results of the proposed method were plotted (Figure 10). It is clear from Figure 10 that no sample was misclassified in the training and testing phase, and the overall recognition accuracy reached 100%. The results indicate that SVM with the proposed method displayed perfect classification performance and was suitable for the recognition of shearer cutting pattern.

In order to verify the necessity of preprocessing vibration signals by decomposition methods, the MFE values of the original signals (Figure 8) were used to construct the feature vectors. Using the same simulation settings mentioned above, the classification results were plotted (Figure 11 *a*). It can be observed from Figure 11 *a* that the overall classification accuracy is 97.92%, which is lower than that of SVM. The reason is that the interference noise in vibration signals affected the recognition results of SVM, which could be retained by the use of CTHI-LMD. The results demonstrate the necessity to decompose the vibration signals before extracting pattern characteristics.

To further validate the superiority of CTHI-LMD in the recognition of shearer cutting pattern, the other three decomposition methods, i.e. RHI-LMD, CHI-LMD and LMD were used to preprocess the vibration signals and the pattern features of the first PF component were extracted by the MFE method. Then the feature vectors were also constructed by the LS method for a SVM classifier to recognize the various cutting patterns of the shearer. The simulation settings were the same as mentioned above. Figure 11 *b–d* shows the classification results of SVM with different preprocessing methods after the training and testing phase. It can be clearly observed from the figures that two, three and four samples are misclassified into the wrong pattern degrees through SVM with RHI-LMD, CHI-LMD and LMD methods respectively, and the corresponding recognition accuracies are 99.17%, 98.75% and 98.33%, which are obviously lower than that of SVM with the proposed method. The results provide compelling evidence that CTHI-LMD can acquire the optimal component with much more pattern information, generating higher classification accuracy than the other decomposition methods.

Table 2. Classification results of back propagation neural network with different preprocessing methods

Preprocessing method	MFE (%)	LMD and MFE (%)	CHI-LMD and MFE	RHI-LMD and MFE (%)	CTHI-LMD and MFE (%)
Training sample	98.75	98.75	100	100	100
Testing sample	96.88	97.5	97.5	98.75	100
Overall accuracy	97.5	97.92	98.33	99.16	100

MFE, Multi-scale fuzzy entropy.

In order to illustrate the universality of the proposed feature extraction method, another common classifier, BPNN was also employed to recognize the shearer cutting patterns. The number of training and testing samples remained the same as those of the SVM model and GA was also used to optimize the parameters of BPNN to achieve better classification ability. The classification results of BPNN with different preprocessing methods are listed in Table 2 after simulations.

As seen from Table 2, BPNN with the proposed method could perfectly recognize the cutting patterns and the overall recognition accuracy reached 100%, which is obviously higher than RHI-LMD and MFE (99.16%), CHI-LMD and MFE (98.33%), LMD and MFE (97.92%) and single MFE (97.5%). The simulation results indicate that the proposed feature extraction method of CTHI-LMD and MFE is well-suited and effective in representing the characteristic information of shearer cutting patterns.

Conclusion

This study proposed a new feature extraction method based on improved LMD and MFE for the recognition of shearer cutting pattern. The CTHI algorithm is employed to construct the envelopes of LMD reasonably well and a sum of PF components can be obtained with different λ values in each rank. Then, the optimal component is selected from the PF components in each sifting process according to the size of grey correlation. Furthermore, MFE is introduced to analyse the complexity of OPF components and the obtained feature vectors are taken as inputs of the SVM and BPNN classifiers. Finally, the simulation results indicate that the proposed approach possesses superior performance in representing the characteristic information of shearer cutting patterns and is applicable to identify the shearer cutting patterns.

According to the results of this study, we can conclude that the proposed method can be used for shearer cutting pattern recognition for vibration signal processing. However, there are some limitations, such as incomplete acquisition of signals in other parts of the shearer. As a major emphasis of future work, more useful vibration signals will be acquired from other key parts of the shearer and some efficient data fusion algorithms will be investigated.

Conflict of interests. The authors declare that they have no competing interests.

- Feng, Z. and Zuo, M. J., Vibration signal models for fault diagnosis of planetary gearboxes. *J. Sound Vib.*, 2012, **331**, 4919–4939.
- Zvokeelj, M., Zupan, S. and Prebil, I., EEMD-based multiscale ICA method for slewing bearing fault detection and diagnosis. *J. Sound Vib.*, 2016, **370**, 394–423.
- Lee, J. H., Kim, J. and Kim, H. J., Development of enhanced Wigner–Ville distribution function. *Mech. Syst. Signal Process.*, 2001, **13**, 367–398.
- Lin, J. and Qu, L. S., Feature extraction based on Morlet wavelet and its application for mechanical fault diagnosis. *J. Sound Vib.*, 2000, **234**, 135–148.
- Sanz, J., Perera, R. and Huerta, C., Fault diagnosis of rotating machinery based on auto-associative neural networks and wavelet transforms. *J. Sound Vib.*, 2007, **302**, 981–999.
- Fu, Y. B., Chui, C. K. and Teo, C. L., Accurate two-dimensional cardiac strain calculation using adaptive windowed Fourier transform and Gabor wavelet transform. *Int. J. Comput. Assist. Radiol.*, 2013, **8**, 135–144.
- Hasson, M., Wavelet-based filters for accurate computation of derivatives. *Math. Comput.*, 2005, **75**, 259–280.
- Heidari, M., Homaei, H., Golestanian, H. and Heidari, A., Fault diagnosis of gearboxes using wavelet support vector machine, least square support vector machine and wavelet packet transform. *J. Vibroeng.*, 2016, **18**, 860–875.
- Huang, N. E. *et al.*, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. London, Ser. A*, 1998, **454**, 903–995.
- Smith, J. S., The local mean decomposition and its application to EEG perception data. *J. R. Soc. Interface*, 2005, **2**, 443–454.
- Do, V. T. and Nguyen, L. C., Adaptive empirical mode decomposition for bearing fault detection. *Stroj. Vestn. – J. Mech. E.*, 2016, **62**, 281–290.
- Wei, L. Y., A hybrid ANFIS model based on empirical mode decomposition for stock time series forecasting. *Appl. Soft Comput.*, 2016, **42**, 368–376.
- Wu, X., Liu, T. B., Liu, R. and Zhao, L., Surge detection methods using empirical mode decomposition and continuous wavelet transform for a centrifugal compressor. *J. Mech. Sci. Technol.*, 2016, **30**, 1533–1536.
- Liu, H. H. and Han, M. H., A fault diagnosis method based on local mean decomposition and multi-scale entropy for roller bearings. *Mech. Mach. Theory*, 2014, **75**, 67–78.
- Li, Y. B., Xu, M. Q., Wang, R. X. and Huang, W. H., A fault diagnosis scheme for rolling bearing based on local mean decomposition and improved multiscale fuzzy entropy. *J. Sound Vib.*, 2016, **360**, 277–299.
- Cheng, J. S., Zhang, K. and Yang, Y., An order tracking technique for the gear fault diagnosis using local mean decomposition method. *Mech. Mach. Theory*, 2012, **55**, 67–76.
- Cheng, J. S., Zhang, K. and Yang, Y., Local mean decomposition method and its application to roller bearing fault diagnosis. *China Mech. Eng.*, 2009, **20**, 2711–2717.

18. Peng, Z. K., Tse, P. W. and Chu, F. L., An improved Hilbert–Huang transform and its application in vibration signal analysis. *J. Sound Vib.*, 2005, **286**, 187–205.
19. Li, Y. B., Xu, M. Q., Wei, Y. and Huang, W. H., An improvement EMD method based on the optimized rational Hermite interpolation approach and its application to gear fault diagnosis. *Measurement*, 2015, **63**, 330–345.
20. Hu, J. S., Yang, S. X. and Ren, D. Q., Spline-based local mean decomposition method for vibration signal. *J. Data Acquisition Process.*, 2009, **24**, 82–86.
21. Wang, M. D., Zhang, L. B. and Liang, W., Local mean decomposition method based on B-spline interpolation. *J. Vib. Shock*, 2010, **29**, 73–78.
22. Deng, L. F. and Zhao, R. Z., Rotor vibration signal analysis based on the CHI-LMD method. *J. Vib. Shock*, 2014, **33**, 58–64.
23. Li, H. L., Guo, L. H., Chen, T., Yang, L. M. and Wang, X. S., Iris recognition based on PCHIP-LMD. *Opt. Precis. Eng.*, 2013, **21**, 197–206.
24. Zhang, K., Cheng, J. S. and Yang, Y., The local mean decomposition method based on rational spline and its application. *J. Vib. Eng.*, 2011, **24**, 96–103.
25. Xie, J., Tan, J. Q. and Li, S. F., Rational cubic Hermite interpolating spline and its approximation properties. *Chin. J. Eng. Math.*, 2011, **28**, 385–392.
26. Li, Y. B., Xu, M. Q., Zhao, H. Y., Yu, W. and Huang, W. H., A new rotating machinery fault diagnosis method based on improved local mean decomposition. *Digit. Signal Process.*, 2011, **46**, 201–214.
27. Zhao, H. Y., Xu, M. Q. and Wang, J. D., Local mean decomposition based on rational Hermite interpolation and its application for fault diagnosis of reciprocating compressor. *J. Mech. Eng.*, 2015, **51**, 83–89.
28. Li, J. C., Zhong, Y. E. and Xie, C., Cubic trigonometric Hermite interpolating splines curves with shape parameters. *Comput. Eng. Appl.*, 2014, **50**, 182–185.
29. Pincus, S. M., Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci. USA*, 1991, **88**, 2297–2301.
30. Richman, J. S. and Moorman, J. R., Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol.–Heart Circ. Phys.*, 2000, **278**, H2039–H2049.
31. Costa, M., Goldberger, A. L. and Peng, C. K., Multiscale entropy analysis of complex physiologic time series. *Phys. Rev. Lett.*, 2002, **89**, 068102.
32. Costa, M., Goldberger, A. L. and Peng, C. K., Multiscale entropy analysis of biological signals. *Phys. Rev. Lett.*, 2005, **71**, 021906.
33. Zhang, L., Xiong, G. and Liu, H., Bearing fault diagnosis using multi-scale entropy and adaptive neuro-fuzzy inference. *Expert Syst. Appl.*, 2010, **37**, 6077–6085.
34. Zheng, J. D., Cheng, J. S. and Yang, Y., A rolling bearing fault diagnosis approach based on multiscale entropy. *J. Hunan University (Natural Sciences)*, 2012, **39**, 38–41.
35. Chen, W., Zhuang, J., Yu, W. and Wang, Z., Measuring complexity using FuzzyEn, ApEn, and SampEn. *Med. Eng. Phys.*, 2009, **31**, 61–68.
36. Zheng, J. D., Cheng, J. S., Yang, Y. and Luo, S. R., A rolling bearing fault diagnosis method based on multi-scale fuzzy entropy and variable predictive model-based class discrimination. *Mech. Mach. Theory*, 2014, **78**, 187–200.
37. Zheng, J. D., Chen, M. J., Cheng, J. S. and Yang, Y., Multiscale fuzzy entropy and its application in rolling bearing fault diagnosis. *J. Vib. Eng.*, 2014, **27**, 145–151.
38. Wu, J. D. and Tsai, Y. J., Speaker identification system using empirical mode decomposition and an artificial neural network. *Expert Syst. Appl.*, 2011, **38**, 6112–6117.
39. Polat, K. and Güneş, S., An expert system approach based on principal component analysis and adaptive neuro-fuzzy inference system to diagnosis of diabetes disease. *Digit. Signal Process.*, 2007, **17**, 702–710.
40. Si, L., Wang, Z. B., Liu, X. H., Tan, C., Liu, Z. and Xu, J., Identification of shearer cutting patterns using vibration signals based on a least squares support vector machine with an improved fruit fly optimization algorithm. *Sensors*, 2016, **16**, 90.
41. Yentes, J. M., Hunt, N. and Schmid, K. K., The appropriate use of approximate entropy and sample entropy with short data sets. *Ann. Biomed. Eng.*, 2013, **41**, 349–365.
42. He, X., Cai, D. and Niyogi, P., Laplacian score for feature selection. *Adv. Neural Inf. Process. Syst.*, 2005, 507–551.

ACKNOWLEDGEMENTS. The support of the National Natural Science Foundation of China (51605477 and U1510117), the National Key Basic Research Program of China (2014CB046301), Open Foundation of State Key Laboratory of Robotics and System (HIT), Harbin Institute of Technology (SKLRS-2016-KF-15) and the Priority Academic Program Development (PAPD) of Jiangsu Higher Education Institutions in carrying out this research is acknowledged. We thank the reviewers for their suggestions that helped improve the original manuscript.

Received 18 October 2016; revised accepted 3 January 2017

doi: 10.18520/cs/v112/i11/2243-2252