New method of bearing fault diagnosis based on mmemd and DE_ELM

Yongtao Hu1, Mingru Dong1, Guozhu Wang1, Zheng Fan1, Shuqing Zhang2

1School of Electrical Engineering and Automation, Henan Institute of Technology, Xinxiang, People’s Republic of China
2Institute of Electrical Engineering, Yanshan University, Qinhuangdao, People’s Republic of China
E-mail: zhshq-yd@163.com

Abstract: A new method of bearing fault diagnosis based on multi-masking empirical mode decomposition (MMEMD) and extreme machine learning optimised by differential evolution algorithm (DE_ELM) is proposed in this study. MMEMD is an improvement of empirical mode decomposition (EMD). By adding masking signals to the signals to be decomposed in different levels, MMEMD can restrain low-frequency components effectively in the sifting process and then suppress the mode mixing. Differential evolution algorithm is applied to determine the parameters of ELM for improving the classification accuracy. The four parameters are determined at one time by uniformly coded as the individuals of the differential evolution algorithm. To achieve the bearing fault diagnosis, the fault signals are first decomposed into different intrinsic mode functions (IMFs) and the sample entropy of each IMF was calculated as the fault feature. Then the fault feature extraction and fault classification are progressed. The results show that the method can identify different faults with high reliability and accuracy.

1 Introduction

Bearings are essential components of mechanical equipment [1, 2]. As the reason of long-term running in atrocious conditions such as variable speeds, alternating and heavy loads, bearings are damaged inevitably. If the equipment keeps running with a damaged bearing, the mechanical system may be broken down eventually [3, 4], which result in enormous economic losses and serious casualties. Therefore, an effective method of bearing fault diagnosis is extremely valuable. Two key steps of bearing fault diagnosis are feature extraction and fault classification [5].

Empirical mode decomposition (EMD) is a kind of nonlinear, non-stationary signal processing method and applied widely in fault feature extraction [6–8]. EMD can decompose the bearing vibration signal into a series of intrinsic mode functions (IMFs) [9]. However, EMD has the disadvantage of mode mixing, which affects the effectiveness of fault feature. To overcome the shortcomings of EMD, a new improved method named multi-masking empirical mode decomposition (MMEMD) is proposed in this paper. By adding multi-masking signals to the signals to be decomposed in different levels, MMEMD can restrain effectively low-frequency component of mixing in high-frequency component in the sifting process, and then suppress the mode mixing.

Extreme learning machine (ELM) is designed on the basis of single hidden layer feedforward networks and widely used in various applications such as classification, regression, unsupervised learning, feature selection and so on [10]. There are three parameters that may affect the performance of ELM, including hidden node number, input weight and hidden node bias [11]. Differential evolution (DE) algorithm is adopted to select the optimal parameters for the advantages of simple principle, less parameters, strong optimisation capability and the fast convergence [12, 13].

This paper proposes a bearing fault diagnosis method based on the MMEMD and DE-ELM. Inter order to obtained the fault feature set, the vibration signals are decomposed into a sum of IMFs by MMEMD, and then sample entropies are calculated as the fault features for fault classification. DE algorithm is used to optimise the ELM to improve the accuracy of classification. To verify the performance of the method, bearing fault diagnosis experiment and actual wind turbine high speed bearing fault diagnosis are progressed. The results show that the method can identify different faults with high reliability and accuracy.

2 Sample entropy extraction based on MMEMD

2.1 Algorithm of MMEMD

Sampling signals are decomposed to a series of IMFs with orthogonality and completeness [14] by MMEMD. For the original signal $x(t)$, the algorithm of MMEMD is as follows.

Step 1: Decompose $x(t)$ into IMFs by EMD and select the most relevant $j$ IMFs according to the correlation between $x(t)$ and IMFs.

Step 2: Let $i = 1$, $x_i(t) = x(t)$.

Step 3: Determine the frequency of IMF$i$ and IMF$(i + 1)$ according to Step 1. Take the mean of them as the masking signal frequency $f_m$, and the amplitude of IMF$i$ as the masking signal amplitude $a_m$.

The masking signal $s_m(t)$ is obtained as follows:

$$s_m(t) = a_m \sin(2\pi f_m t)$$

Step 4: Construct $x_{ii}(t)$ and $x_{i+1}(t)$ according to (2) and (3)

$$x_i(t) = x(t) + s_m(t)$$

$$x_{i+1}(t) = x(t) - s_m(t)$$

Step 5: Obtain the $r$th order IMF of $x_{ii}(t)$ and $x_{i+1}(t)$ by means of sifting as EMD. Take average of them and get the IMF$i$, $c_i(t)$.

Step 6: Subtract $c_i(t)$ from $x_i(t)$ to obtain the residue, $r(t)$

$$r(t) = x(t) - c_i(t)$$

Step 7: Let $x_j(t) = r(t)$ and $i = i + 1$, repeat steps above from Step 3 until $i = j$, the all IMFs can be obtained.
2.2 Feature extraction based on sample entropy

After signals are decomposed by MMEMD into different frequency bands perfectly, sample entropies can be calculated and be taken as the fault features for classification. Sample entropy is an improved algorithm to approximate entropy, which could measure time sequence complexity and effectively reflect the bearing conditions [15].

Sample entropy is a modification of approximate entropy calculated as follows:

\[ S = \ln \left( \frac{1}{n-m+1} \sum_{i=m}^{n-m+1} c^n(i) - \frac{1}{n-m} \sum_{i=m}^{n-m} c^{m+1}(i) \right) \]  

where \( n \) is the signal length, \( m \) is the match length, \( r \) is the noise filter parameter and \( c^n(i) \) is the probability that any two signals match each other. According to experiment, \( m = 4 \) and \( r \) is set to the 20% of the standard deviation of the signal. After the vibration signal decomposed into IMFs, the sample entropies of all IMFs are calculated as the fault feature as follows:

\[ x = [S_1, S_2, ..., S_J] \]  

3 Differential evolution optimised extreme learning machine

3.1 Extreme learning machine classifier

ELM is a new kind of single hidden layer neural network, which overcomes the shortcomings of traditional neural network and has been used in mechanical fault diagnosis recently. The principle ideal of the ELM classifier can be described as follows:

Suppose \( T = \{ (x_i, y_i) \} (i = 1, 2, ..., N) \) is a training set. Where \( x_i \) is the input features vector, \( y_i \in \{ 1, 2, ..., M \} \) is the label of \( x_i, N \) is the size of training set and \( M \) is the number of classes. The \( j \)th output of ELM classifier is given by

\[ f(x) = h(x)w_j, \quad j = 1, 2, ..., M \]  

where \( w_j \) is the output weight vector and \( h(x) \) is the output vector with respect to the input \( x \). It is a constrained optimisation problem which constrained items are as follows:

\[ \begin{align*}
\min L_c &= \frac{1}{2} || w ||^2 + \frac{C}{2} \sum_{i=1}^{N} || \xi_i ||^2 \\
\text{s.t.} & \quad h(x_i) - w \cdot \tilde{y}_i - \xi_i, \quad i = 1, 2, ..., N \\
& \quad h(x_i) = g(\omega_i x_i + b) 
\end{align*} \]  

where \( w = [w_1, w_2, ..., w_M] \) is the output weight matrix, \( C \) is the regularised parameter, \( \tilde{y}_i = [\tilde{y}_1, \tilde{y}_2, ..., \tilde{y}_M] \) is the target output vector of \( x_i, \xi_i = [\xi_1, \xi_2, ..., \xi_M] \) is the training error vector, \( g \) is the activation function, \( \omega_i \) is the input weight vector and \( b \) is the hidden node bias vector.

According to the KKT theorem, (32) can be converted to its dual optimisation problem as follows:

\[ \begin{align*}
L_{\text{dual}} &= \frac{1}{2} || w ||^2 + \frac{C}{2} \sum_{i=1}^{N} || \xi_i ||^2 \\
& - \sum_{i=1}^{N} M \sum_{j=1}^{M} a_{ij}(h(x_i)w_j - \tilde{y}_j + \xi_i) 
\end{align*} \]  

where \( a_{ij} \) is the Lagrange multipliers. By partial differential method, the solution can be obtained as follows:

\[ w = H(H^T + HH^T)^{-1} \tilde{y} \]  

where \( H = [h(x_1), h(x_2), ..., h(x_N)] \) is the output matrix, \( I \) is the identity matrix and \( \tilde{y} = [\tilde{y}_1, \tilde{y}_2, ..., \tilde{y}_N]^T \) is the target output matrix. Hence the output of ELM classifier is

\[ f(x) = h(x)w = h(x)H(H^T + HH^T)^{-1} \tilde{y} \]  

It is necessary to mention that the output \( f(x) \) is not the class label, but a vector that is related to the class label. The index of the maximal element is also the class label that is easy to be obtained. The parameters that need users to set are the hidden node number, input weight and hidden node bias. Normally, the hidden node number can be determined easily by experience. Also the input weight and the hidden node bias are selected by differential evolution.

3.2 DE_ELM algorithm

Differential evolution algorithm is an intelligent optimisation algorithm, which has the advantages of simple principle, less parameters, strong optimisation capability and fast convergence. The purpose of differential evolution is to improve the recognition accuracy by searching the optimal parameters. So the parameters input weight and hidden node bias are taken as the individual. Moreover the classification error is set as the fitness, the \( j \)th individual fitness function can be expressed as follows:

\[ \text{fit}(j) = \frac{N_{\text{err}}}{N} \]  

where \( N_{\text{err}} \) is the fault recognition number and \( N \) is the total number of input samples. The steps of DE_ELM algorithm are as follows.

(i) Initialise parameters, set the maximum iteration number to 100 and make it as the terminal condition, set the iteration number \( t = 0 \).
(ii) Encode the optimise parameters and initialise the individuals randomly.
(iii) Input the training samples to be recognised and save the individual fitness.
(iv) Individual variation, crossover and select from the generation before to produce the next generation of population, \( t = t + 1 \).
(v) Check the terminal condition, if satisfied, go to step (vi), otherwise return step (iii).
(vi) Output the optimal parameters and obtain the optimised ELM.

4 Bearing fault diagnosis model based on MMEMD and DE_ELM

According to the discussion above, the bearing fault diagnosis model based on MMEMD and DE_ELM is proposed as shown in Fig. 1.

The process of bearing fault diagnosis can be described as follows.

(i) Sample: \( n \) points are sampled from bearing vibration signal.
(ii) Decomposition: the signal is decomposed into a sum of IMFs for feature extraction.
(iii) Feature extraction: the sample entropies of all IMFs are calculated and normalised as the feature vectors.
(iv) Training: the feature vector and the labels of the training set are put into DE_ELM for training.
(v) Testing: the feature vector of testing set is put into the well trained DE_ELM and the fault diagnosis of bearing are achieved.

5 Experiment and application

In order to verify the high performance of MMEMD and the feasibility of the proposed fault diagnosis method, bearing fault diagnosis experiments and application are conducted.
5.1 Experiment of bearing fault diagnosis

Bearing vibration data was collected from the bearing fault diagnosis experiment rig. It consists of a 1.47 kW motor, a torque transducer/encoder, a dynamometer and control electronics. The type of bearing is SFK6205. Single point faults with diameters of 0.007 inches in ball, inner raceway and outer raceway were introduced to the test bearings by electro-discharge. The motor speed was 1750 r/min, load was 1.47 kW and the sampling frequency was 12 kHz. The signal length is 0.1 s, covering 1200 points. Take 100 samples from each bearing for diagnosis experiment, of which 50 for training and the other for testing.

The signals were decomposed by MMEMD and the decomposition result of a sample is shown in Fig. 2. It can be seen that the vibration characteristics are obvious and there is no mode mixing. Hilbert envelope is employed to analysis the IMFs as shown in Fig. 3. It is easy to find the fault characteristic frequency and its fold frequency when the bearing is in trouble. So MMEMD can decompose the bearing vibration signal into different frequency bands effectively, which help to achieve good features for fault classification.

The normalised sample entropies of different bearings are listed in Table 1. The sample entropies are different greatly among different bearings. Fig. 4 is the fuzzy c-means (FCM) clustering result of the training set. It shows that the most bearings are classified correctly. So MMEMD sample entropy can be taken as the features for classification and a 400 × 3 feature matrix is obtained.

The feature matrix was taken apart into a training set and testing set and each set comprised 200 samples. The training set and labels were input to DE_ELM and a classifier model was obtained. 12 samples randomly selected from the testing set were put into the model for testing, all the testing samples are classified correctly.

In order to further validate the effectiveness of the proposed method, the testing samples were increased the other method were taken as a comparison. Fig. 5 is the training fitness of ELM.

**Fig. 1** Bearing fault diagnosis model based on MMEMD and DE_ELM

**Fig. 2** Decomposition result of a sample
(a) Waveform, (b) Amplitude-frequency characteristic

**5.1 Experiment of bearing fault diagnosis**

Bearing vibration data was collected from the bearing fault diagnosis experiment rig. It consists of a 1.47 kW motor, a torque transducer/encoder, a dynamometer and control electronics. The type of bearing is SFK6205. Single point faults with diameters of 0.007 inches in ball, inner raceway and outer raceway were introduced to the test bearings by electro-discharge. The motor speed was 1750 r/min, load was 1.47 kW and the sampling frequency was 12 kHz. The signal length is 0.1 s, covering 1200 points. Take 100 samples from each bearing for diagnosis experiment, of which 50 for training and the other for testing.

The signals were decomposed by MMEMD and the decomposition result of a sample is shown in Fig. 2. It can be seen that the vibration characteristics are obvious and there is no mode mixing. Hilbert envelope is employed to analysis the IMFs as shown in Fig. 3. It is easy to find the fault characteristic frequency and its fold frequency when the bearing is in trouble. So MMEMD can decompose the bearing vibration signal into different frequency bands effectively, which help to achieve good features for fault classification.

The normalised sample entropies of different bearings are listed in Table 1. The sample entropies are different greatly among different bearings. Fig. 4 is the fuzzy c-means (FCM) clustering result of the training set. It shows that the most bearings are classified correctly. So MMEMD sample entropy can be taken as the features for classification and a 400 × 3 feature matrix is obtained.

The feature matrix was taken apart into a training set and testing set and each set comprised 200 samples. The training set and labels were input to DE_ELM and a classifier model was obtained. 12 samples randomly selected from the testing set were put into the model for testing, all the testing samples are classified correctly.

In order to further validate the effectiveness of the proposed method, the testing samples were increased the other method were taken as a comparison. Fig. 5 is the training fitness of ELM.
optimised by DE, GA and PSO. From Fig. 5, it can be seen that DE has the best performance of optimisation. The classification result is listed in Table 2. It shows that no matter which classifier is used, the classification accuracy is better when MMEMD is used for feature extraction. Also for different feature extraction method, the DE_ELM achieves the best classification result. That means the MMEMD-DE_ELM method more effective.

5.2 Application to wind turbine bearing fault diagnosis

The proposed method is applied to wind turbine bearing fault diagnosis to verify its effectiveness in practical applications. The high-speed bearing vibration signals were collected from a 2 MW wind turbine by condition monitoring system. Data were collected at 10-minute intervals with sample rate 97.656 kHz and sample time of 6 s. The data were transferred and stored in the server once a day. The high-speed shaft is driven by a 20 tooth gear and the rated speed is 1800 r/min. The time domain wave and amplitude-frequency characteristic of the high-speed bearing in a condition of normal and fault are shown in Fig. 6. The unit is Gs, where 1G is the earth standard gravitational acceleration.

Around 100 samples were taken from bearing in condition of normal and fault for diagnosis. Take 50 samples as the training set and the other 50 samples as testing set. The signals were decomposed by MMEMD and the sample entropy features are obtained. The FCM clustering result of the training data set is given in Fig. 7, which shows that both the bearings are classified correctly.

The training set and labels ware put into DE_ELM and a DE_ELM classifier model was obtained. The testing set was put into the model to realise the fault diagnosis. The diagnosis results with the input of 5, 15 and 25 samples are shown in Table 3. It shows that all the testing samples are classified correctly.

6 Conclusion

A new method combining MMEMD sample entropy and ELM optimised by differential evolution is proposed for bearing fault diagnosis. MMEMD can restrain the mode mixing effectively, which is important for bearing signal decomposition and fault feature extraction. Combined with sample entropy, the fault features can be extracted quickly and accurately. Differential evolution is used for optimising the parameters of ELM classifier, which could improve the classification accuracy. The bearing fault diagnosis experiment and wind turbine bearing fault diagnosis application results indicated that the proposed method is effective in fault diagnosis.

7 Acknowledgment

This work was supported by the Science and Technology Research Project of Henan Province (Grant No. 182102210033,
Table 2  Classification result of different testing samples and method

| Method        | Feature extraction | Classifier | 5, %       | 25, %       | 50, %       |
|---------------|--------------------|------------|------------|------------|------------|
| EEMD          |                    | BP         | 80         | 87%        | 89.5       |
|               |                    | SVM        | 85         | 90%        | 92.5       |
|               |                    | GA_ELM     | 90         | 92%        | 92         |
|               |                    | PSO_ELM    | 90         | 93%        | 91         |
|               |                    | DE_ELM     | 95         | 98%        | 99.5       |
| MMEMD         |                    | BP         | 90         | 95%        | 95.5       |
|               |                    | SVM        | 95         | 96%        | 96         |
|               |                    | GA_ELM     | 95         | 98%        | 97         |
|               |                    | PSO_ELM    | 100        | 99%        | 99         |
|               |                    | DE_ELM     | 100        | 100        | 100        |

Table 3  Diagnosis results with input of 5, 15 and 25 samples

| Numbers | Condition | Wrong number | Accuracy, % |
|---------|-----------|--------------|-------------|
| 5       | normal    | 0            | 100         |
|         | fault     | 0            | 100         |
| 15      | normal    | 0            | 100         |
|         | fault     | 0            | 100         |
| 25      | normal    | 0            | 100         |
|         | fault     | 0            | 100         |

8 References

[1] Wang, Y., Xu, G., Liang, L., et al.: ‘Detection of weak transient signals based on wavelet packet transform and manifold learning for rolling element bearing fault diagnosis’, Mech. Syst. Signal Process., 2015, 54–55, (3), pp. 259–276.

[2] Amar, M., Gondal, I., Wilson, C.: ‘Vibration spectrum imaging: a novel bearing fault classification approach’, IEEE Trans. Ind. Electron., 2015, 62, (1), pp. 494–502.

[3] Wang, T., Liang, M., Li, J., et al.: ‘Rolling element bearing fault diagnosis via fault characteristic order (FCO) analysis’, Mech. Syst. Signal Process., 2014, 45, (1), pp. 139–153.

[4] Shao, R., Hu, W., Huan, X., et al.: ‘Multi-damage feature extraction and diagnosis of a gear system based on higher order cumulant and empirical mode decomposition’, J. Vib. Control, 2013, 21, (4), pp. 736–754.

[5] Kankar, P., Sharma, S., Harsha, S.: ‘Rolling element bearing fault diagnosis using wavelet transform’, Neurocomputing, 2011, 74, (10), pp. 1638–1645.

[6] Li, Y., Xu, M., Liang, X., et al.: ‘Application of bandwidth EMD and adaptive multiscale morphology analysis for incipient fault diagnosis of rolling bearings’, IEEE Trans. Ind. Electron., 2017, 64, (8), pp. 6506–6517.

[7] Chen, Y., Zhang, P.: ‘Bearing fault detection based on SVD and EMD’, Appl. Mech. Mater., 2012, 184–185, (6), pp. 70–74.

[8] Kopsinis, V., Mclaughlin, S.: ‘Development of EMD-based denoising methods inspired by wavelet thresholding’, IEEE Trans. Signal Process., 2009, 57, (4), pp. 1351–1362.

[9] Pei, Y., Wu, Y., Jia, D.: ‘Research on PD signals denoising based on EMD method’, Proc. Electrotechniczn., 2012, 88, (1), pp. 137–140.

[10] Flandrin, P., Rilling, G., Gonçalves, P.: ‘Empirical mode decomposition as a filter bank’, IEEE Signal Process. Lett., 2004, 11, (2), pp. 112–114.

[11] Jerri, A.: ‘The Shannon sampling theorem – its various extensions and applications: a tutorial review’, Proc. IEEE, 2005, 65, (11), pp. 1565–1596.

[12] Han, M., Pan, J.: ‘A fault diagnosis method combined with LMD, sample entropy and energy ratio for roller bearings’, Measurement, 2015, 76, (12), pp. 7–19.

[13] Zhao, Y., Wang, G., Yin, Y., et al.: ‘Improving ELM-based microarray data classification by diversified sequence features selection’, Neural Comput. Appl., 2016, 27, (1), pp. 155–166.

[14] Yin, X., Hadjiloucas, S., He, J., et al.: ‘Application of complex extreme learning machine to multiclass classification problems with high dimensionality: A THz spectra classification problem’, Digit. Signal Process., 2015, 40, (C), pp. 40–52.

[15] Zheng, Y., Xu, X., Ling, H., et al.: ‘A hybrid fireworks optimization method with differential evolution operators’, Neurocomputing, 2015, 148, (19), pp. 75–82.

Fig. 6  High-speed bearing signal
(a) Waveform, (b) Amplitude-frequency characteristic

Fig. 7  FCM clustering result of the training set

182102210261), Natural Science Foundation of Hebei Province (Grant No. E2018203439, E2018203339).