

EEMD Method and TWSVM for Fault Diagnosis of Roller Bearings

Guo Xiaoxuan

School of Mathematics and Systems Science
Beihang University
Beijing, China
e-mail: gxx930121@163.com

Guo Xiaoxuan *

School of Mathematics and Systems Science
Beihang University
Beijing, China
e-mail: gxx930121@163.com

* Corresponding Author

Abstract— The ensemble empirical mode decomposition (EEMD) is a self-adaptive signal processing technique for nonlinear and non-stationary signals, which can alleviate the mode mixing problem occurring in empirical mode decomposition (EMD). As a improved support vector machine (SVM) method, Twin support vector machine (TWSVM) is a powerful tool for supervised learning, which are successfully applied to classification and regression problems. This paper proposes an effective fault diagnosis method for roller bearings based on EEMD and TWSVM. First, the vibration signals collected from the roller bearings are decomposed using EEMD and intrinsic mode functions (IMF) are produced. Second, the sample entropy of the most IMFs are calculated as the feature of initial signal. At last, these features, as training and recognition samples, are fed into TWSVM to identify the bearing fault conditions. The experiment results show that the proposed method can accurately recognize the bearing normal, inner race, outer race and ball fault under small samples.

Keywords—EEMD; EMD; sample entropy; TWSVM; fault diagnosis.

I. INTRODUCTION

Roller bearings are essential components of rotating machinery and the health condition of roller bearings is always a topic issue. In rotating machinery, the failure of roller bearings can result in the deterioration of mechanical condition and safety. Therefore, it is significant to find a method which can precisely and automatically diagnose the faults occurring in the bearings. The vibration-based signal processing technique is one of major tools for diagnosing bearing faults. Then, the fault characteristic information can be extracted from the vibration signals.

Recently, various methods have been proposed for bearings fault diagnosis in [1-4]. Traditional signal processing techniques cover time-domain method and Fourier transform, applied to linear and stationary data. Since the vibration signals captured from the faulty bearings are strongly nonlinear and non-stationary, the conventional signal processing techniques inevitably cause serious errors. Empirical mode decomposition (EMD), a novel time-frequency analysis technique, has been introduced by Huang et al[5]. to deal with the nonlinear and non-stationary signals. EMD is based on the local characteristic time scales of a signal and can self-

adaptively decompose the original signal into a set of intrinsic mode functions (IMFs). The IMFs have actually physical meaning. Nonetheless, one of the major shortcomings of EMD is the mode mixing problem, which includes either a single IMF consisting of components of widely disparate scales or a similar scale signal remains in different IMFs. To remedy the deficiency of the EMD, the ensemble empirical mode decomposition (EEMD), a new noise-assisted data analysis method, is presented recently by Wu and Huang[6]. The EEMD method can eliminate the mode mixing problem automatically by adding finite white noise to the investigated signal. Therefore, EEMD will be employed as the tool to process the vibration signals of roller bearings in this paper.

In recent years, support vector machine (SVM), replacing the artificial neural network method, has been successfully applied in fault diagnosis.[7] The SVM approach is systematic and motivated by statistical learning theory (SLT) and Bayesian arguments. In the case of small samples, SVM has a better performance than neural network method and, to a large extent, overcomes the problems such as curse of dimensionality and over-learning. Twin Support Vector Machine (TWSVM) is developed by Jayadeva et al. [8], which obtains non-parallel hyperplanes by solving two novel formulations of quadratic programming problems (QPP). The generalization of TWSVM has been shown to be significantly better than standard SVM for both linear and nonlinear kernels. TWSVM possesses more excellent processing speed and recognition rate than traditional SVM, thus it has a vast potential for practical application.

This paper combined the merits of EEMD and TWSVM, then a new fault diagnosis method to automatically and effectively diagnosis faults of roller bearings is presented. Because sample entropy can show the complexity of signal stably, compute the sample entropy of principal IMF components as feature extracted from the vibration signals. As training and recognition samples of machine, sample entropy are input into TWSVM to design a roller bearings fault classification recognizer. This paper uses one-against-rest TWSVM to classify fault conditions. The remainder of this paper is organized as follows. Section 2 introduces EMD method and EEMD method, respectively. Section 3 briefly discusses the sample entropy formulation. The new fault

diagnosis method based on EEMD and TWSVM is presented in Section 4. Section 5 uses experiments to illustrate the effectiveness of our methods. Finally, some concluding remarks are given in Section 6.

II. EMPIRICAL MODE DECOMPOSITION AND ENSEMBLE EMPIRICAL MODE DECOMPOSITION

A. EMD algorithm and mode mixing

Based on the time scale of the local signal characteristics, EMD decomposed signals into a set of complete and almost orthogonal components named several intrinsic mode functions (IMF). Each IMF must satisfy the two conditions: (1) the number of extreme points and the corresponding number of zero crossing points in entire data segment must either be equal or have a difference of no more than one; (2) at any time the mean of upper envelope curves formed by local maximum points and lower envelope curves formed by local minimum points should be zero. The IMFs represent the natural oscillatory mode embedded in the signal and work as the basic functions, which are determined by the signal itself. Thus, EMD is a effective self-adaptive signal processing method for nonlinear and non-stationary signals. However, one drawbacks of EMD is the mode mixing problem, which is defined as a single IMF including oscillation of dramatically disparate scales, or a component of a similar scale residing in different IMFs. In this case the IMF components would lose physical meaning.

B. EEMD algorithm

To overcome the problem of mode mixing in EMD, a new noise-assisted data analysis method named EEMD is proposed. It defines the true IMF components as the mean of an ensemble of trials and every trial consists of the decomposition results of the signal plus a white noise of finite amplitude.

EEMD is a substantial improvement of EMD, and its procedures are as follows:

(1) Initialize the number of ensemble M , the amplitude of the added white noise, and $m=1$.

(2) Add a random white noise signal to the original signal

$$x_m(t) = x(t) + n_m(t) \quad (1)$$

where $n_m(t)$ indicates the m th added white noise series, and $x_m(t)$ is the noise-added signal of the m th trial.

(3) Decompose $x_m(t)$ into a series of IMFs $c_{i,m}$ utilizing EMD as follows:

$$x_m(t) = \sum_{i=1}^{N_m} c_{i,m} + r_{N_m} \quad (2)$$

where $c_{i,m}$ denotes the i th IMF of the m th trail, r_{N_m} denotes the residue of the m th trail and N_m is the IMFs number of the m th trail.

(4) If $m < M$, then repeat steps (2) and (3), and add different random white noise signals each time.

(5) Obtain $I = \min(N_1, N_2, \dots, N_M)$ and calculate the ensemble means of the M trials for each IMF

$$c_i = \frac{1}{M} \sum_{m=1}^M c_{i,m} \quad (3)$$

where $i = 1, 2, \dots, I$

(6) c_i is the ensemble mean of corresponding IMF of the decompositions.

III. FEATURE EXTRACTION WITH SAMPLE ENTROPY

A. Sample entropy

Entropy, a nonlinear dynamics parameter measuring the occurrence probability of new information in time series, has been applied in many scientific fields. Recently, a number of different entropy estimators have been applied to quantify the complexity of the signal. Approximate entropy (ApEn) is proposed by Pincus in 1991 from the perspective of time series complexity and successfully applied to the physiological signal analysis. Approximate entropy possess many advantages such as good anti-interference abilities. However, the results of approximate entropy were somewhat biased and inconstant.

In order to overcome these defects of ApEn, Richman and Moorman presented an improved method of sample entropy in 2000[9], which not only maintained the merits of approximate entropy but also simplified the calculation and reduced the computation time by half[10]. Sample Entropy (SampEn) examines time series for similar epochs and assigns a non-negative number to the sequence, with larger values corresponding to more complexity or irregularity in the data [8]. Sample entropy could be expressed as $\text{SampEn}(m, r, N)$, where N is data length, r is similar margin and m is embedding dimension. For given N data points from a time series $\{x(n)\} = x(1), x(2), \dots, x(N)$, the sample entropy algorithm is as follows.

(1) Form a set of m -dimensional vectors according to the order of serial number.

$$X(i) = [x(i), x(i+1), \dots, x(i+m-1)], \quad (4)$$

$$i = 1, 2, \dots, N-m+1$$

(2) Define the maximum distance between vector $X(i)$ and vector $X(j)$ as $d[X(i), X(j)]$.

$$d[X(i), X(j)] = \max_k [x(i+k) - x(j+k)] \quad (5)$$

$$k = 0, 1, \dots, m-1; i, j = 1, 2, \dots, N-m+1, j \neq i$$

(3) For a given $X(i)$, count the number of j , denoted as B_i , such that $d[X(i), X(j)]$ is less than or equal to r .

$$B_i^m(r) = \frac{1}{N-m+1} B_i \quad (6)$$

(4) Then set

$$B^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} B_i^m(r) \quad (7)$$

(5) Increase the dimension to get a $(m+1)$ dimensional vector, then repeat the steps of (1)~(4) to obtain $B^{m+1}(r)$. When N is a finite value, the sample entropy can be defined as

$$\text{SampEn}(m, r) = -\ln \frac{B^{m+1}(r)}{B^m(r)} \quad (8)$$

B. Feature extraction

In general, the larger energy the IMF component had, the more information of the original signal this IMF component would contain. In this paper, the sample entropy can be used to describe the complexity or irregularity of IMF component. Thus, the sample entropy of IMF components with major energy would be extracted as the feature of the original signal. Then, a set of eigenvectors composed of these sample entropies are used as the input data of fault pattern recognition. As it is known to all, most energy of signal is contained in the first several IMFs. To simplify the calculation process, only count the sample entropy values of first few IMFs.

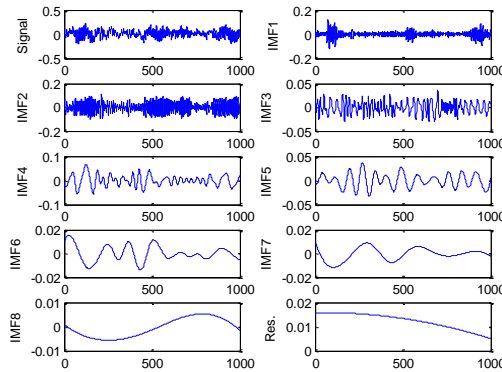


Figure 1. The decomposition result of normal bearing signal with EEMD

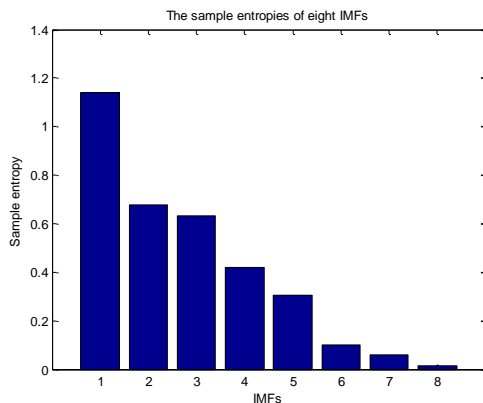


Figure 2. The sample entropy values of eight IMFs

To better illustrate, a normal bearings signal, which has 1797 rpm motor speeds is decomposed by EEMD. The decomposition results are showed in Fig. 1. It can be seen that the signal is divided into eight IMFs. Then calculating the sample entropy of each IMF and the corresponding values are given in Fig. 2. As displayed in Fig. 2, the major information of original signal is embodied in the first five IMFs. Therefore only extract the first five sample entropies to consist the eigenvector as the signal feature.

IV. THE PROPOSED BEARING FAULT DIAGNOSIS METHOD

A. Brief review of TWSVM

For the last decade, support vector machine (SVM) as powerful tools for pattern classification and regression, have already widely applied in a variety of fields. SVM

represent novel learning techniques that have been introduced in the framework of structural risk minimization (SRM) and in the theory of VC bounds. The central idea of SVM is to find the optimal separating hyperplane between the positive and negative examples. The optimal hyperplane is defined as the one giving maximum margin between the training examples that are closest to the hyperplane. As a improved method of SVM, Twin Support vector machine (TWSVM) is first proposed by Professor Jayadeva in 2007. This algorithm mainly aim at generating two nonparallel planes such that each plane is closer to one of the two classes and is as far as possible from the other. As a novel support vector machine, the main difference between the classic support vector machine is that solving a pair of quadratic programming problems (QPPs) instead of a single QPP. Meanwhile, there are two discriminate curves in TWSVM.

Thus, the training speed of TWSVM is four times faster than that of SVM. TWSVM can also be extended to nonlinear cases by considering the two kernel-generated surfaces.

TWSVM possesses more excellent processing speed and recognition rate than traditional SVM. Therefore it has a vast potential for practical application, such as pattern recognition, data classification, function fitting and so on.

B. The proposed diagnosis method

The proposed bearing fault diagnosis method based on EEMD and TWSVM is shown in Fig. 3. It includes the following four procedures. First of all, the vibration signals captured from the roller bearings is decomposed into a set of IMFs by EEMD. Secondly, calculating the sample entropy values of the first several IMFs as the feature extracted from the original signal. Finally, input these features into TWSVM network, as the training data, to built the fault classifiers. Thereby, recognize the fault conditions of the roller bearings and final diagnosis result can be produced.

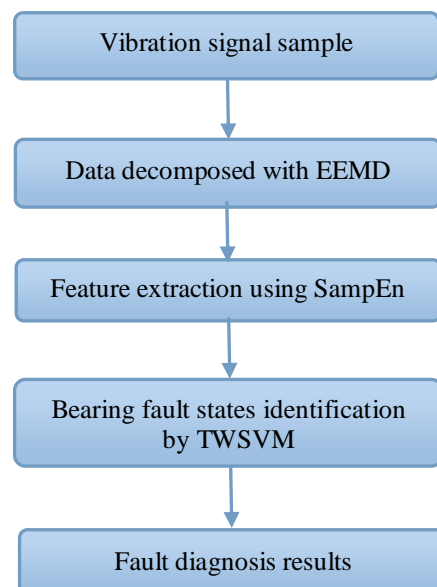


Figure 3. The processing procedures of the proposed method

V. EXPERIMENTS AND RESULTS

TWSVM is initially developed to solve two-class or binary pattern recognition problems. To classify roller bearings states, which are normal, the inter race fault, the outer race fault and the ball fault, multi-class TWSVM classifiers must be built. First of all, combine several binary classifiers to form a multi-class classifier, then train the TWSVM network, lastly, identify fault conditions. In this paper, it is needed four binary classifier (TWSVM 1, TWSVM 2, TWSVM 3, TWSVM 4) to differentiate four types state of roller bearing. Recognition method is showed as the following table:

TABLE I. RECOGNITION METHOD

The output of classifier	The state of bearing
TWSVM 1 = 1	Normal state
TWSVM 1 = -1	Other states
TWSVM 2 = 1	Inter race fault
TWSVM 2 = -1	Not inter race fault
TWSVM 3 = 1	Outer race fault
TWSVM 3 = -1	Not outer race fault
TWSVM 4 = 1	Ball fault
TWSVM 4 = -1	Not ball fault

Experiments are conducted using a 2 hp Reliance Electric motor, and acceleration data is measured at location near to and remote from the motor bearings. Through the experiments, collect a bearing data set consisting of four data subsets under normal state, inter race fault, outer race fault and ball fault. Each of the four data subsets contains 30 groups vibration signals. In this four TWSVM input data, separately select 40 groups as training sample data, the rest 20 groups as testing sample data. The results and prediction accuracy are shown in Fig. 4, Fig. 5, Fig. 6, Fig. 7 and Table 2. In this experiment, the TWSVM for four distinct conditions produce the classification accuracy of 92.5%, 97.75%, 98.5% and 90%, respectively. This result implies that the proposed method could diagnosis the roller bearings fault effectively and stably.

TABLE II. PREDICTION ACCURACY OF FOUR STATES

States	Normal	Inter race fault	Outer race fault	Ball fault
Accuracy	92.5%	97.75%	98.5%	90%

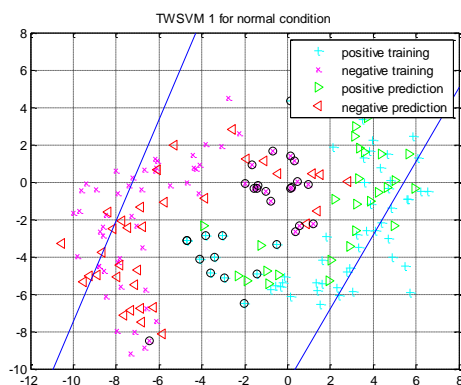


Figure 4. TWSVM 1 classifier for normal condition

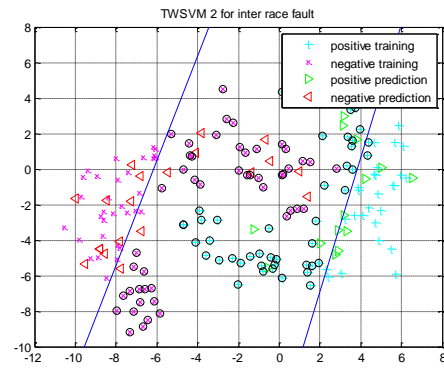


Figure 5. TWSVM 2 classifier for inter race fault

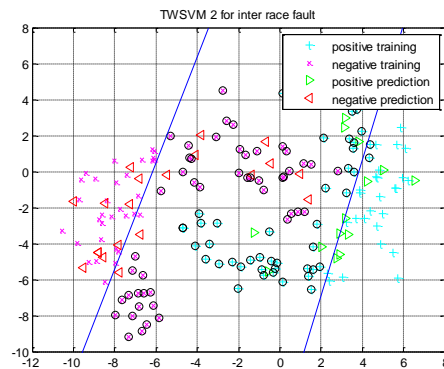


Figure 6. TWSVM3 classifier for outer race fault

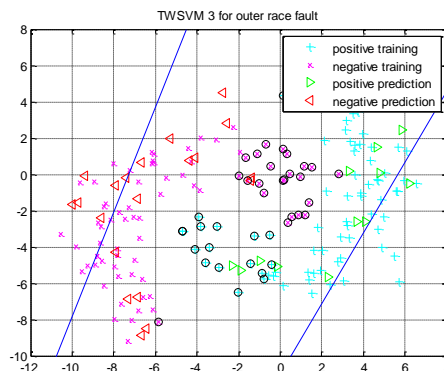


Figure 7. TWSVM 4 classifier for ball fault

VI. CONCLUSIONS

Based on ensemble empirical mode decomposition (EEMD) and Twin support vector machine (TWSVM), present an effective and automated fault diagnosis method for roller bearings. The detailed steps of the method are as follows: (1) Pre-process the investigated signals captured from the roller bearings with EEMD; (2) Compute the sample entropy of the most IMFs for each vibration signal; (3) Use the values of the sample entropy as the train data of TWSVM to built recognition classifier. The results of experiment demonstrate our method is a promising method for the fault diagnosis of roller bearings.

ACKNOWLEDGEMENT

The research work is supported by National Natural Science Foundation of China under Grant No. 61379001 and Key Project of Beihang University under Grant No. 201412.

REFERENCES

- [1] Y.G. Lei, Z.J.He and Y.Y. Zi, Fault diagnosis of rotating machinery based on multiple ANFIS combination with Gas [J]. *Mechanical Systems and Signal Processing*, 2007, 21(5): 2280-2294.
- [2] Y. G. Lei, Z. J. He, and Y. Y. Zi, A new approach to intelligent fault diagnosis of rotating machinery [J]. *Expert Systems with Applications*, 2008, 35(4): 1593-1600.
- [3] Y. G. Lei, Z. J. He, and Y. Y. Zi, Application of the EEMD method to rotor fault diagnosis of rotating machinery [J]. *Mechanical Systems and Signal Processing*, 2009, 23(4): 1327-1338.
- [4] Y. G. Lei, Z. J. He, and Y. Y. Zi, EEMD method and WNN for fault diagnosis of locomotive roller bearings [J]. *Expert Systems with Applications*, 2011, 38(6): 7334-7341.
- [5] N.E. Huang et al, Huang N E, Shen Z, Long S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[C]//*Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. The Royal Society, 1998, 454(1971): 903-995.
- [6] Z.H. Wu, N.E. Huang, Ensemble empirical mode decomposition: a noise-assisted data analysis method [J]. *Advances in adaptive data analysis*, 2009, 1(01): 1-41.
- [7] C. J. Burges, A tutorial on support vector machines for pattern recognition [J], *Data Mining Knowledge Discovery*, 1998, 2(2) :955-974.
- [8] Jayadeva et al., Twin support vector machines for pattern classification [J]. *IEEE Trans. Pattern Anal. Machine Intell.*, 2007, 29(5) :905-910.
- [9] J. S. Richman, J. R. Moorman, Physiological time series analysis using approximate entropy and sample entropy [J]. *Am.J.Physiol.Heart Physio*, 2000, 278:2039-2049.
- [10] H. Liu, W. X. He, X. P. Chen, Approximate entropy and sample entropy for physiological time-series [J],*Chinese Journal of Scientific Instrument*, 2004, 25(4):806~812.
- [11] X. P. Chen, C. F. Yin, W. X. He., Feature extraction of gearbox vibration signals based on EEMD and sample entropy [J], *Fuzzy Systems and Knowledge Discovery*, 2013, 10:811-815.
- [12] X. P. Chen, C. F. Yin, W. X. He., Feature extraction of gearbox vibration signals based on EEMD and sample entropy [J], *Fuzzy Systems and Knowledge Discovery*, 2013, 10:811-815.
- [13] W. H. Hu et al., Fault diagnosis of gas turbine based on support vector machine [J], *Chinese Control and Decision Conference*, 2014, 26:2853-2856.
- [14] J. W. Lou, C. B. Hu, J. L. Zhao., Research of EEMD sample entropy in bearing fault recognition with SVM [J], *Journal of Mechanical Transmission*, 2014, 38(3):41-45.
- [15] Case Western Reserve University. Bearing Data Center. Available: <http://csegroups.case.edu/bearingdatacenter/pages/welcome-case-western-reserve-university-bearing-data-center-website>.