Mechanism and Machine Theory 98 (2016) 114–132

Contents lists available at ScienceDirect

Mechanism and Machine Theory

journal homepage: www.elsevier.com/locate/mechmt

Hierarchical fuzzy entropy and improved support vector machine based binary tree approach for rolling bearing fault diagnosis

Yongbo Li, Minqiang Xu *, Haiyang Zhao, Wenhu Huang

Department of Astronautical Science and Mechanics, Harbin Institute of Technology (HIT), No.92 West Dazhi Street, Harbin 150001, People's Republic of China

ARTICLE INFO

Article history:
Received 22 August 2014
Received in revised form 17 November 2015
Accepted 19 November 2015
Available online xxxx

Keywords:
Hierarchical fuzzy entropy (HFE)
Laplacian score (LS)
Improved support vector machine based binary tree (ISVM-BT)
Fault feature extraction

ABSTRACT

A novel rolling bearing fault diagnosis method based on hierarchical fuzzy entropy (HFE), Laplacian score (LS) and improved support vector machine based binary tree (ISVM-BT) is proposed in this paper. Focus on the difficulty of extracting fault feature from the non-linear and non-stationary vibration signal under complex operating conditions, HFE method is utilized for fault feature extraction. Compared with multi-scale fuzzy entropy (MFE) method, HFE method considers both the low and high frequency components of the vibration signals, which can provide a much more accurate estimation of entropy. Besides, Laplacian score (LS) method is introduced to refine the fault feature by sorting the scale factors. Subsequently, the obtained features are fed into the multi-fault classifier ISVM-BT to automatically fulfill the fault pattern identifications. The experimental results demonstrate that the proposed method is effective in recognizing the different categories and severities of rolling bearings faults.

1. Introduction

Rolling bearings are important and fragile parts in the industry filed, which are frequently damaged under harsh environment, resulting in huge economic loss if not detected in time [1]. Much research in recent years has focused on the fault diagnosis of rolling bearings. Among them, the vibration analysis method has been widely applied for diagnosing the rolling bearing fault due to its intrinsic merits of revealing bearing failure [2–4]. It is generally accepted that vibration analysis method consists of two aspects: fault feature extraction and fault pattern classification [5]. However, due to the influence of kinds of factors such as friction, clearances, overloading and so on, the measured vibration signal often represents non-linear and non-stationary characteristics, which is a challenge to extract the underlying fault features in complex vibration signals. Therefore, many researchers have suggested ways of fault feature extraction. For example, correlation dimension was proposed by Logan to identify the rolling bearing operating conditions [6]. However, correlation dimension method has been largely unsuccessful to short data. To avoid the drawbacks of correlation dimension, Pincus [7] introduced approximate entropy (ApEn), which was successfully applied to monitor the rolling bearing health conditions by Yan [2]. Unfortunately, the various lengths of data will affect the ApEn calculation significantly, especially for processing short time series. The estimated value is uniformly lower than that expected one as well [8]. In order to overcome the disadvantages of ApEn, Richman and Moorman proposed Sample entropy (SampEn), which had been widely used in a range of applications both in physiological and machinery signals [9]. Although SampEn can improve performance in contrast to ApEn, it measures the complexity of data in a single scale, which may lead to inaccurate results. Costa [10] proposed a multi-scale entropy procedure to calculate the SampEn over a range of scales. Multi-scale entropy (MSE) has been widely applied quantify the time series' complexity. For instance, Wu applied MSE to

* Corresponding author. Tel.: +86 451 86414320.
E-mail addresses: liyongbo0532@126.com (Y. Li), xumqhit@126.com (M. Xu), weiyu1219@126.com (H. Zhao), leobo28@foxmail.com (W. Huang).

http://dx.doi.org/10.1016/j.mechmachtheory.2015.11.010
0094-114X/

diagnose the rolling bearing fault [11], and Liu conducted the bearing fault pattern identifications by the combination of LMD and MSE [5]. However, MSE only takes into account the low frequency components due to the progressive smoothing operations. Besides, MSE is not well adapted to the practical measured bearing fault vibration signals, which contain rich frequencies and the fault information may be embedded in both lower and higher frequency components. Hence, it is not adequate to extract the fault feature by using MSE method. To overcome the disadvantages of MSE, Jiang et al. proposed hierarchical entropy (HE) method to estimate the complexity of data recently [12], which is successfully applied to analyze the cardiac interbeat interval time series. HE was firstly applied to identify different rolling bearing fault patterns by Zhu et al. [13], and the superiority of HE was validated by comparing with MSE method.

Based on the SampEn, a novel approach called fuzzy entropy (FuzzyEn) was proposed by Chen et al. [14], which replaced the Heaviside function with fuzzy membership function with a better continuity. FuzzyEn was applied to rolling bearing fault pattern recognition by Zheng et al. [15]. Recently, the multi-scale fuzzy entropy (MFE) was developed to enhance the physical meanings and statistical sense of FuzzyEn [15]. However, the coarse-grained procedure used in MFE essentially represents a linear smoothing, which only captures the low frequency components, ignoring the high frequency components [16]. Furthermore, the calculation of MFE is time-consuming when analyzing the long time series. Hence, a novel approach called hierarchical fuzzy entropy (HFE) is proposed in this paper to overcome the weakness of MFE. Compared with MFE, HFE has some advantages as follows: firstly, for each scale of time series, it considers a lower frequency component produced by averaging the components in the previous scale, as well as a higher frequency component produced by taking the difference of two consecutive scales; secondly, the hierarchical decomposition has higher computational efficiency than the coarse-graining procedure, so HFE needs fewer time than MFE for the long time series analysis. Therefore, HFE is utilized to extract the fault feature from rolling bearing vibration signals in this paper.

Therefore, HFE is taken as a feature extractor to extract the fault information from the vibration signals in this paper. However, the feature vectors extracted from vibration signals using HFE are high dimension with information redundancy, which will make the diagnosis accuracy decreasing and time consuming. In this paper, we introduce an effective approach called Laplacian Score (LS) to select the first several important scale factors to construct the new fault feature vectors [17]. By virtue of the LS method, the fault feature vectors can be automatically ranked according to their importance and correlations with the main fault information [4], and then we select the first four important scale factors as the new fault feature vectors. By using LS, it can not only reduce the data dimension but also enhance the identification accuracy greatly.

Normally, after obtaining the fault features using HFE, another focus is to achieve the fault pattern recognition by using a multi-fault classifier. Support vector machine (SVM) based on statistical learning theory was put forward by Vapnik [18], which had been demonstrated to be effective in making a reliable decision for a smaller number of datasets. Since SVM has high accuracy and good generalization capabilities, it has been widely applied in fault diagnosis and classification field [19].

SVM was originally designed as a binary classifier, nevertheless, the practical identification problem is mostly the multi-class problem. So far, varieties of techniques for solving the multi-class problem based on SVM have been developed, such as: one against one (OAO) [20] one against all (OAA) [21] decision directed acyclic graph SVM(DDAGSVM) [22] and support vector machines based on binary tree (SVM-BT) [19]. Compared with other classifiers, SVM-BT has the superiorities as follows: fewer sub-classifiers, none of unclassifiable region and good classification performance, which is suitable for practical application of multi-class fault recognitions [19].

A major concern of SVM-BT is the adoption of hierarchical structures for the training of a multi-class SVM, which is vitally important for classification performance. Furthermore, the hierarchical structures should be designed before training each sub-classifiers of SVM, and the much research in recent years has focused on hierarchical structures design for SVM-BT. Among many measures, inter-class Euclidean distance (ED) has been one of the most effective methods to quantify the similarity of two classes, inter-class ED is used to construct the binary tree hierarchy [23], in which the class with bigger distance from other classes will be separated earlier and in the higher node of the binary tree architecture. In addition, intra-class sample distribution was introduced to design the hierarchical structures of binary tree [24,25]. Namely, the class with wider distribution range will be classified earlier. The inter-class ED and intra-class sample distribution are both the approaches to describe the distinguishability from different aspects. Considering the merits of the two approaches, in this paper, we introduce a novel measure to construct the hierarchical structure of binary tree, called improved SVM-BT (ISVM-BT),which takes advantages of the above two approaches and reflects the class separability more comprehensively.

The rest of this paper is organized as follows: in Section 2, we firstly recall SampEn, FuzzyEn and MFE methods. In Section 3, the proposed HFE method is described detailed. The ISVM-BT method is introduced, meanwhile, the superiority of ISVM-BT is validated in Section 4. The proposed method based on HFE and ISVM-BT is introduced briefly and experiment data analysis is presented in Section 5. Finally, conclusions are given in Section 6.

2. Review of SampEn, FuzzyEn and multi-scale fuzzy entropy

2.1. Sample entropy

To overcome the drawbacks of approximate entropy (ApEn), the sample entropy (SampEn) is proposed by Richman [9], which can provide more precise complexity estimation. For a given time series $\{x(i), i = 1, 2, \cdots, N\}$, the main calculation procedures of SampEn can be written as follows.

(1) Construct an m dimension vector according to the Eq. (1):

$$X_i^m = \{x(i), x(i+1), \ldots, x(i+m+1)\}, 1 \le i \le N-m+1 \tag{1}$$