Neurocomputing 157 (2015) 208–222

Contents lists available at ScienceDirect

Neurocomputing

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

Multi-fault diagnosis for rotating machinery based on orthogonal supervised linear local tangent space alignment and least square support vector machine

Zuqiang Su [b], Baoping Tang [a,*], Ziran Liu [a], Yi Qin [b]

[a] School of Mechanical & Electrical Engineering, Henan University of Technology, Zhengzhou 450007, PR China
[b] The State Key Laboratory of Mechanical Transmission, Chongqing University, Chongqing 400030, PR China

ARTICLE INFO

Article history:
Received 1 July 2014
Received in revised form
1 January 2015
Accepted 8 January 2015
Available online 19 January 2015

Keywords:
Vibration signal
Dimension reduction
Fault diagnosis
Supervised manifold learning
Least square support vector machine
Particle swarm optimization

ABSTRACT

In order to improve the accuracy of fault diagnosis, this article proposes a multi-fault diagnosis method for rotating machinery based on orthogonal supervised linear local tangent space alignment (OSLLTSA) and least square support vector machine (LS-SVM). First, the collected vibration signals are decomposed by empirical model decomposition (EMD), and a high-dimensional feature set is constructed by extracting statistical features, autoregressive (AR) coefficients and instantaneous amplitude Shannon entropy from those intrinsic model functions (IMFs) that contain most fault information. Then, the high-dimensional feature set is inputted into OSLLTSA for dimension reduction to yield more sensitive low-dimensional fault features, which not only achieves the combination of intrinsic structure information and class label information of dataset but also improves the discrimination of the low-dimensional fault features. Further, the low-dimensional fault features are inputted to LS-SVM to recognize machinery faults according to the LS-SVM parameters selected by enhanced particle swarm optimization (EPSO). Finally, the performance of the proposed method is verified by a fault diagnosis case in a rolling element bearing, and the results confirm the improved accuracy of fault diagnosis.

1. Introduction

Rotating machinery is an important component in industrial applications and has been widely used in many crucial areas such as electric power generation, mining industry, industrial production and so on. With the development of society and technology, rotating machinery is becoming huger, more complicated and more precise. Additionally, it always operates continuously for extended hours under hostile conditions. As a result, any failure may lead to its fatal breakdown, which brings heavy economic losses and even personal casualties [1]. Consequently, it is very important to develop effective fault diagnosis techniques for rotating machinery to avoid accidents and increase machine reliability. Essentially, fault diagnosis of rotating machinery is a problem of pattern recognition, in which fault feature extraction and intelligent fault recognition are the most important two aspects [2]. Therefore, the accuracy of fault diagnosis can be improved by developing more effective feature extraction methods and classification methods.

Vibration signals contain abundant information about machinery running states and are commonly used in fault diagnosis of rotating machinery [3]. However, the raw vibration signals should be preprocessed before feature extraction, because they are usually non-stationary and non-linear and contain complicated components [4]. Empirical mode decomposition (EMD) [5] is an advanced self-adaptive time–frequency domain signal processing method based on the local characteristic time scale of a signal. It can decompose a multi-component signal into a set of intrinsic mode functions (IMFs) containing different frequency bands, and the fault features extracted from the IMFs are more distinct than those extracted from the raw vibration signals. However, conventional fault features such as statistical features [6], autoregressive model coefficients [7] and Shannon entropy [8] can only characterize the properties of faults from one specific aspect. In view of this, a high-dimensional feature set comprised of all the three features is constructed in this study to obtain more fault information to better characterize the properties of faults. Unfortunately, the high-dimensional fault feature set inevitably contains disturbed and even noise features, and mutual coupling between features also leads to excessive redundant information. Therefore, it is necessary to further extract significantly independent features with low dimension, high sensitivity and good clustering from the

* Corresponding author. Tel.: +86 13658319901; fax: +86 23 65415883.
E-mail addresses: jhonsu1987@126.com (Z. Su), bptang@cqu.edu.cn (B. Tang).

http://dx.doi.org/10.1016/j.neucom.2015.01.016
0925-2312/

high-dimensional fault feature set by an appropriate dimension reduction method.

Dimension reduction methods can be roughly divided into two categories: linear and nonlinear ones. Typical linear dimension reduction methods include principle component analysis (PCA) [9], independent component analysis (ICA) [10], linear discriminant analysis (LDA) [11] and so on. These linear dimension reduction methods are popularly exploited and easy to implement, but they are only applicable to linearly structured datasets with Gaussian distribution and cannot reveal the intrinsic structure information of datasets. In order to overcome the disadvantages of linear dimension reduction methods, some nonlinear dimension reduction methods, known as manifold learning methods, are proposed recently, including ISOMAP [12], locally linear embedding [13], local tangent space alignment [14], etc. These nonlinear dimension reduction methods project a nonlinearly structured high-dimensional dataset into a lower dimensional feature space by preserving local structures in small neighborhoods, and thus successfully acquire the intrinsic features of the high-dimensional dataset. Among these methods, local tangent space alignment has some remarkable advantages such as easy implementation, low computational cost and high capability against noises. However, it can only implement dimension reduction on training dataset but cannot deal with the new coming samples. For this problem, linear local tangent space alignment (LLTSA) is proposed as the linearization of local tangent space alignment [15]. LLTSA inherits the ability of local tangent space alignment in revealing intrinsic structure information of dataset and provides the explicit mapping from high-dimensional input space to low-dimensional feature space. For example, Li et al. [16] have already applied LLTSA to machinery fault diagnosis. Nevertheless, LLTSA is an unsupervised learning method, which neglects the supervisory role of class label information of training samples in classification. This drawback leads to the blindness of dimension reduction and causes the incompleteness of fault decoupling. Although some improvements of local tangent space alignment have been proposed (for example, Zhang et al. [17] proposed an improved local tangent space alignment to deal with the sparse and non-uniformly distributed dataset; Wang [18] proposed a modified local tangent space alignment to solve the problem of outliers; Zhan et al. [19] proposed a robust local tangent space alignment to improve the noise resistance ability), these variants mainly focus on the problem of noise or outliers but neglect the supervisory role of class label information of training samples in classification. Hence, the linearization of these methods is not optimal yet for machinery fault diagnosis. Zhang et al. [20] also proposed a semi-supervised local tangent space alignment, but the employed label information was the low-dimensional presentation rather than class type of training samples. That is to say, the objective of this method is coordinate propagation rather than classification. Although some other improved methods like supervised linear local tangent space alignment [21–23] and orthogonal discriminant linear local tangent space alignment [24] are the supervised variants of LLTSA, they still have some disadvantages as follows: supervised linear local tangent space alignment only introduces class label information within the neighborhood selection phase but keeps local structure information extraction in neighborhood unchanged; orthogonal discriminant linear local tangent space alignment only increases the interclass dissimilarity but does not enhance the intraclass compactness. Therefore, class label information cannot be fully utilized to improve the discrimination power of dimension reduction. In this study, a novel supervised manifold learning method called as orthogonal supervised linear local tangent alignment (OSLLTSA) is proposed for machinery fault diagnosis. The proposed OSLLTSA mainly improves LLTSA from three aspects: (1) class label information is employed to guide the construction of local neighborhood, thus enhancing the intraclass compactness; (2) a novel strategy for local structure information extraction is introduced to increase the interclass dissimilarity; (3) an orthogonal iteration algorithm [25] is applied to compute the optimal orthogonal project matrix, which eliminates the statistical correlations between basis vectors. Through the first two improvements, combination of intrinsic structure information and class label information of dataset can be successfully achieved, so that clustering effect and sensitivity of the extracted low-dimensional fault features are improved. Additionally, since the datasets are orthogonally projected to the low-dimensional feature space, redundant information within the high-dimensional fault feature set can be well removed and more intrinsic structure information is preserved. Thus, the discrimination of the obtained low-dimensional fault features can be further improved.

After fault feature extraction, an effective classification method is necessary to accurately recognize faults. However, commonly used classification methods such as k nearest neighbor classifier (KNNC) [26] and artificial neural networks (ANN) [27] require sufficient training samples for model training, and ANN even suffers overfitting and local optimal solution owing to empirical risk minimization principle. By contrast, support vector machine (SVM) [28] based on structural risk minimization principle can minimize an upper bound on the expected risk, so that it can solve the problem of overfitting and local optimal solution. Moreover, SVM implements classification by using a separating hyper-plane that is determined by a few samples called as support vectors, and thus has great generalization capability for small sample size problem. Least square support vector machine (LS-SVM) [29] is an improvement of SVM, and it can greatly simplify the training process by transforming the quadratic programming problem in SVM to a linear problem. Currently, LS-SVM has already been successfully employed to diagnose faults in machinery. For example, Xu et al. [30] applied LS-SVM to diagnose the faults in bearings; Zhou et al. [31] used LS-SVM to diagnose the vibration faults of centrifugal pump; Zhang et al. [32] proposed a new method based on LS-SVM for roller bearing safety region estimation and state identification. However, the performance of LS-SVM heavily depends upon the selection of its parameters, whereas the theoretical basis of selecting appropriate parameters for LS-SVM is still lacking. Particle swarm algorithm (PSO) [33] is a global optimization method that simulates the social behavior of bird flock foraging, and it is simple to implement and easy to compute. However, PSO suffers the problem of being trapped into local minimum. Worse still, the convergence rate of PSO is very slow in practical applications. In order to overcome the disadvantages, some improvements of PSO have already been proposed. Among these methods, He et al. [34] proposed a new PSO method (NPSO) and pointed out that sub-optimal experience of particle swarm should be introduced within the optimization process to overcome premature; He et al. [35] proposed a modified PSO and pointed out that the parameters of PSO should be adjusted along with the evolution in order to accelerate convergence rate of the optimization process; Ding et al. [36] proposed a new PSO variant on the basis of local stochastic search strategy(LSSPSO) and pointed out that selecting the optimal local position for each particle could improve the performance of the optimization process; Zhai et al. [30] proposed an improved PSO (IPSO) and pointed out that it was the searching velocity rather than the increase of iterations that reflected the real state of the convergence of the particle swarm. However, these methods only focused on one aspect of the problem, but ignored the combination of their advantages to achieve better performance. Inspired by the research results above, this study proposes a novel parameter optimization method called as enhanced PSO (EPSO) to select the parameters of LS-SVM. In the proposed EPSO, searching velocity is redefined to overcome premature in PSO and to indicate the convergence state of the particle swarm; local searching ability is