Knowledge-Based Systems 89 (2015) 56–85

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

Intelligent fault diagnosis of roller bearings with multivariable ensemble-based incremental support vector machine

XiaoLi Zhang [a,b], BaoJian Wang [a], XueFeng Chen [a,*]

[a] State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an 710049, PR China
[b] Key Laboratory of Road Construction Technology and Equipment, Ministry of Education, Chang'an University, Xi'an 710064, PR China

A R T I C L E   I N F O

Article history:
Received 17 October 2014
Received in revised form 22 May 2015
Accepted 24 June 2015
Available online 29 June 2015

Keywords:
Multivariable
Ensemble
Incremental
Support vector machine
Intelligent fault diagnosis
Roller bearing

A B S T R A C T

Since roller bearings are the key components in rotating machinery, detecting incipient failure occurring in bearings is an essential attempt to assure machinery operational safety. With a view to design a well intelligent system that can effectively correlate multiple monitored variables with corresponding defect types, a novel intelligent fault diagnosis method with multivariable ensemble-based incremental support vector machine (MEISVM) is proposed, which is testified on a benchmark of roller bearing experiment in comparison with other methods. Moreover, the proposed method is applied in the intelligent fault diagnosis of locomotive roller bearings, which proves the capability of detecting multiple faults including complex compound faults and different severe degrees with the same fault. Both experimental and engineering test results illustrate that the proposed method is effective in intelligent fault diagnosis of roller bearings from vibration signals.

1. Introduction

Roller bearing is the most common component in industrial rotating machinery. Defects in roller bearings can lead to machinery malfunctioning, even to economical loss and human casualties. Therefore, fault diagnosis of roller bearing plays a very important role in machinery maintenance and process automation. Various signal processing and pattern recognition approaches are proposed as available diagnostics tools [1]. It is worthy to note that time–frequency analysis including wavelet analysis [2–5] and Hilbert–Huang transform [6–8] can reveal defect frequency characteristics of the vibration signal. However, there are a number of factors affecting the complexity of the bearing signal [9]. Firstly, it is almost impossible to precisely determine the bearing characteristic frequencies with variation of bearing geometry and assembly. Secondly, different locations of a bearing defect cause different behaviors in transient signal response, which is easily buried in the wide band resonance and noise signals. Thirdly, it is hard to distinguish signals and extract their representative features in different severity stages with a same defect type. Fourthly, operational speed and load of the shaft greatly affect the way and the degree of a machinery vibration, which also affect the measured vibration signals. Finally, it is also impossible to obtain the baseline information about some particular bearing, which makes signal processing methods impractical. All above, there is a great demand of constructing a reliable, intelligent and automated procedure for condition monitoring and diagnosis of roller bearing.

Pattern recognition techniques, such as neural network [10–12], decision tree [13–15], support vector machine [16–18], can offer diagnostic decision by learning the latent rules from the observed signals or measured variables with a fault type. However, there is a bottleneck that machinery fault samples are often very scarce. Support vector machine is of specialties for small samples based on statistical learning theory and therefore exhibits good performance in intelligent fault diagnosis of roller bearings [19,20].

Support vector machine usually require representative training samples to generate an appropriate decision boundary among different classes, so the whole training data is usually attained in prior and trained in one batch [21]. When a new data belonging to a new class occurs, a typical approach to learn new knowledge involves discarding the existing classifier and retraining another classifier with representative samples that have been accumulated so far. This learning style of support vector machine results in loss of previously acquired information, which is known as catastrophic forgetting phenomenon [22]. Moreover, it brings data storage burden and computation waste for repetitious retraining. The terrible case is that the previous training data was lost in the past and the new unseen data is not coming, which is infeasible for support

* Corresponding author. Tel./fax: +86 29 82663689.
E-mail address: chenxf@mail.xjtu.edu.cn (X. Chen).

http://dx.doi.org/10.1016/j.knosys.2015.06.017
0950-7051/

vector machine to learn adequate knowledge so as to recognize the whole patterns. To tackle the problem, incremental based learning methods have been proposed so far. Laskov et al. focuses on the design and analysis of efficient incremental support vector machine learning with the aim of providing a fast, numerically stable and robust implementation [23]. A detailed analysis of convergence and algorithmic complexity of incremental support vector machine learning is carried out in Ref. [23]. Moreover, Parikh et al. develop an ensemble-based incremental learning approach for data fusion application [24]. The experimental results indicate that the ensemble-based incremental learning approach often performs better than an ensemble classification system or a single classifier [24]. Erdem et al. proposed an ensemble of support vector machine for incremental learning on optical character recognition and volatile organic compounds dataset [25], which is proved to be capable of learning new information from subsequent datasets including new instances of previously unseen classes.

Since machinery fault samples are generally attained little by little as machinery condition varies in operation and representative fault samples are very scarce, it is important to learn fault information incrementally so as to make correct machinery condition assessment utilizing previously learned knowledge and newly attained data in an online fashion without requiring access to previous data. With a view to design a well intelligent system that can effectively correlate multiple monitored variables with defect types, a multivariable ensemble-based incremental support vector machine (MEISVM) is proposed to improve classification performance of typical support vector machine and learn new information gradually. The highlights and contributions are as follows:

- Multivariable ensemble-based incremental support vector machine is proposed to improve classification performance of support vector machine, since machinery fault samples are gradually attained as machinery condition varies in operation and the representative fault samples are very scarce.
- Multivariable ensemble-based incremental support vector machine has well capacity of correlating the measured multiple variables with a corresponding fault type of roller bearing from sensor information, which exhibits good performance in comparison with other proposed methods in literature on a benchmark of roller bearing experiment.
- Multivariable ensemble-based incremental support vector machine is applied in the intelligent fault diagnosis of locomotive roller bearings, which can detect multiple faults including complex compound faults and different severe degrees with a same fault type.

The organization of the paper is as follows. The basic theory of support vector machine is briefly reviewed in Section 2. The multivariable ensemble-based incremental support vector machine is explained in Section 3. Experiments are conducted to testify the effectiveness of the proposed methods in contrast to other methods in Section 4. The proposed method is finally conducted to detect complex faults of locomotive roller bearings in Section 5. General conclusions are drawn in Section 6.

2. The basic theory of support vector machine

Support vector machine is a powerful machine learning method for classification and regression problems based on statistical learning theory and structural risk minimization principle [19,26]. Most cases in practical are multi-classes, such as fault diagnosis. Many approaches have been proposed to extend the binary support vector machine to multi-class problems. The most common strategies are called "one-against-one" and "one-against-all". Many comparison studies and arguments show that the one-against-one approach outperforms other approaches in many cases [27–30]. Therefore, classification is done using one-against-one approach, which is briefly introduced as follows.

Given a training sample set in the input space

$$ST = \{(\boldsymbol{x}_t, y_t) | \boldsymbol{x}_t \in H, \quad y_t \in \{1, 2, \cdots, k\}, \quad t = 1, \cdots, l\} \tag{1}$$

where $\boldsymbol{x}_t$ is an input vector, y_t is the corresponding label of $\boldsymbol{x}_t$, l is the number of the training samples, and k is the number of different classes. One-against-one support vector machine constructs $k(k-1)/2$ binary classifiers to recognize different classes. For training data from ith and jth classes, the following binary classification problem can be solved:

$$\text{Minimize} \quad \frac{1}{2}(\boldsymbol{w}^{ij})^T \boldsymbol{w}^{ij} + C\sum_t \xi_t^{ij} (\boldsymbol{w}^{ij})^T \tag{2}$$

$$\text{Subject to} \quad \begin{aligned} &(\boldsymbol{w}^{ij})^T \Phi(\boldsymbol{x}_t) + b^{ij} \geqslant 1 - \xi_t^{ij} \quad \text{if } y_t = i, \\ &(\boldsymbol{w}^{ij})^T \Phi(\boldsymbol{x}_t) + b^{ij} \leqslant 1 - \xi_t^{ij} \quad \text{if } y_t = j, \\ &\xi_t^{ij} \geqslant 0. \end{aligned} \tag{3}$$

where $\Phi(\boldsymbol{x}_i)$ is the vector mapped from the input space of $\boldsymbol{x}_i$, $\boldsymbol{w}^{ij}$ is the weight vector, ξ_t^{ij} is a slack variable, C is a penalty constant. By incorporating kernels and rewriting it in Lagrange multipliers, the above binary classification problem can be transformed into a dual quadratic optimization problem and finally forms decision function.

$$f^{ij}(\boldsymbol{x}) = sign\left(\sum_{t=1}^{l} y_t \alpha^{ij} \cdot K(\boldsymbol{x}, \boldsymbol{x}_t) + b^{ij}\right) \tag{4}$$

where α^{ij} is the Lagrange multiplier, $K(\boldsymbol{x}, \boldsymbol{x}_i)$ is kernel function, b^{ij} is bias value.

After all the $k(k-1)/2$ classifiers are constructed, the classification decision of the one-against-one support vector machine is made using the following strategy: if $f^{ij}(\boldsymbol{x})$ says sample $\boldsymbol{x}$ is in the ith class, then the vote for the ith class is added by one. Otherwise, the jth is increased with one. After being tested with the $k(k-1)/2$ classifiers respectively, $\boldsymbol{x}$ belongs to the class which has maximal votes.

3. Multivariable ensemble-based incremental support vector machine (MEISVM)

The ensemble-based incremental learning approach is firstly proposed by Parikh et al. for data fusion application [24]. Inspired by the approach, multivariable ensemble-based incremental support vector machine is proposed for intelligent fault diagnosis of roller bearing, which mainly contains three parts as illustrated in Fig. 1: multivariable feature extraction, ensemble learning and incremental learning.

3.1. Multivariable feature extraction

Identifying significant variables or extracting fault features form large amount of measured sensory information is challenging and has been a focal point in the fault diagnosis domain. It is desirable that variables and features extracted from sensors are sensitive to machinery faults and robust to the varying machinery running conditions and background noise [9]. So there has been a lot of signal processing approach to obtain desirable features for machinery fault diagnosis, among which Fast Fourier Transform (FFT) is one of the most widely used and well-established methods. When a fault occurs, new frequency components may appear and a change of the convergence of frequency spectrum may take place.