Measurement 48 (2014) 136–148

Contents lists available at ScienceDirect

Measurement

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

Fault diagnosis method based on incremental enhanced supervised locally linear embedding and adaptive nearest neighbor classifier

Zuqiang Su, Baoping Tang *, Jinghua Ma, Lei Deng

The State Key Laboratory of Mechanical Transmission, Chongqing University, Chongqing 400030, PR China

ARTICLE INFO

Article history:
Received 3 July 2013
Received in revised form 24 September 2013
Accepted 31 October 2013
Available online 9 November 2013

Keywords:
Fault diagnosis
Vibration signal
Manifold learning
Incremental enhanced locally linear embedding
Adaptive nearest neighbor classifier

ABSTRACT

A novel fault diagnosis method based on incremental enhanced supervised locally linear embedding (I-ESLLE) and adaptive nearest neighbor classifier (ANNC) is proposed to improve the accuracy of machinery fault diagnosis. Firstly, I-ESLLE is proposed for the non-linear dimensionality reduction of high-dimensional fault samples obtained from vibration signals. I-ESLLE can not only acquire the low-dimensional intrinsic manifold structure embedded in the high-dimensional input space, but also can deal with new fault samples in an iterative and batch model. Then, the low-dimensional fault samples are fed into the proposed ANNC for fault type identification. ANNC exploits "representation-based distance" to select the nearest training samples of new fault sample and identifies fault type in a weighting strategy. Moreover, the number of nearest training samples of each new fault sample is adaptively determined according to the density of the local distribution of the new fault sample. To verify the validity of the proposed fault diagnosis method, a fault diagnosis experiment of gearbox is performed, and the results indicate that the proposed fault diagnosis method outperforms the traditional methods and achieves higher diagnostic accuracy.

1. Introduction

Fault diagnosis is now a very important research area in machinery engineering [1]. Accurate and rapid diagnosis of machinery faults can not only monitor machinery operating condition to avoid fatal breakdown, but also can reduce the maintenance costs and human labors, especially for the huge and long running equipment [2]. Fault diagnosis is essentially a problem of pattern recognition, so more effective feature extraction method and more accurate classifier are needed to obtain higher diagnostic accuracy [3].

Machinery operating vibration signals contain rich fault information, and many fault diagnosis methods are proposed based on the features extracted from vibration signals. For example, Zamanian and Ohadi [4] proposed a fault diagnosis method for gearbox based on Gaussian correlation of vibration signals and wavelet coefficients, and Jena and Panigrahi [5] developed a fault diagnosis method for bearing and gear using adaptive wavelet transform of vibration signal, etc. To obtain more fault information to better characterize the faults and improve the performance of fault diagnosis methods, a large number of features are extracted from vibration signals in different domains [6]. However, there are always non-linear correlations and even redundant and disturbed information existing in the original feature set, besides of the high dimensionality [7]. If the original features are directly inputted into the classifier to identify the fault types, the accuracy of the classifier will decrease while the computational cost will increase. Therefore, an appropriate dimensionality reduction method has to be performed upon the original feature

* Corresponding author. Tel.: +86 13658319901; fax: +86 023 65415883.

E-mail addresses: jhonsu1987@126.com (Z. Su), bptang@cqu.edu.cn (B. Tang).

0263-2241/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.measurement.2013.10.041

set to extract the intrinsic independent features and remove the redundant and disturbed information. The classical dimensionality reduction methods, such as principle component analysis (PCA) [8], multi-dimensional scaling (MDS) [9] and independent component analysis (ICA) [10], can achieve satisfying effects for datasets with linear structures, but their performances degenerate when facing datasets with non-linear structures [11]. The recently proposed manifold learning is a kind of non-linear dimensionality reduction methods, which can effectively approximate the low-dimensional intrinsic manifold structure embedded in the high-dimensional input space [12,13], and has already been successfully applied in facial expression recognition [14], and gene expression [15], etc. However, manifold learning is unsupervised, which cannot utilize the class label information to guide the dimensionality reduction to be more suitable for pattern recognition, namely to better separate samples from different classes. To extend manifold learning into supervised learning method, several supervised manifold learning methods have been proposed, including supervised locally linear embedding (SLLE) [16], supervised Isomap [17,18], and the recently proposed enhanced supervised locally linear embedding (ESLLE) [19], etc. Among them, ESLLE is very suitable for pattern recognition for that it maximize the interclass dissimilarity while minimizing the intraclass dissimilarity by redefining the original distances between samples. But like other manifold learning methods, ESLLE does not provide an explicit mapping from high-dimensional input space to low-dimensional feature space, so it cannot tackle the new fault samples, which is the so-called "out of sample problem". In this paper, a new incremental supervised manifold learning method named I-ESLLE is proposed by integrating the iterative new sample embedding algorithm [20] into ESLLE. I-ESLLE can properly deal with new fault samples in an iterative and batch model, which is more precise than obtaining the low-dimensional representations of new samples by directly extending their local neighborhoods [3]. And the new samples are also added into the dimensionality reduction model incrementally to make the model more flexible. After the dimensionality reduction, a proper classifier is needed to finally identify the fault types of fault samples. Conventional k nearest neighbor classifier (CKNNC) is one of the simplest and most commonly used classifiers, which has already been utilized in fault diagnosis [21,22]. However, CKNNC has several drawbacks, and Xu et al. [23] proposed a variant of CKNNC named coarse to fine k nearest neighbor classifier (CFKNNC) to improve the performance of CKNNC, which exploit "representation-based distance" rather than Euclidean distance to select the nearest training samples of new sample. Nevertheless, CFKNNC still needs to set the number of the nearest training samples of new sample and treats all the nearest training samples equally while identifying the class label of the new sample, which may weaken the classification capacity of CFKNNC. This paper proposed a new extended version of CFKNNC named ANNC. ANNC assigns each training sample a weight according to its contribution in representing the new sample and identifies the class label of the new sample in a weighting strategy. Moreover, the number of nearest training samples of each new fault sample is adaptively determined according to the density of the local distribution of the new sample. By doing this, the selected nearest training samples can well represent the new sample, and those nearest training samples, which are more similar to the new sample, can get higher weight in decision of the class label of the new sample. As a result, ANNC can get higher accuracy than CFKNNC. At last, a novel fault diagnosis method is proposed based on I-ESLLE and ANNC, and the procedure of this method is as follows: Firstly, features are extracted from the original vibration signal to obtain high-dimensional fault samples. Then, the high-dimensional fault samples are processed by I-ESLLE to acquire the low-dimensional representations of the fault samples. Finally, the low-dimensional fault samples are fed into ANNC for classifier training and fault type identification.

The remainder of this paper is organized as follows. In Section 2, locally linear embedding (LLE), SLLE and ESLLE are reviewed briefly, and then I-ESLLE proposed in this paper is described in detail. In Section 3, we briefly review CKNNC and CFKNNC, and ANNC is described at the end of this section. In Section 4, the novel fault diagnosis strategy based on I-ESLLE and ANNC is discussed. A fault simulation experiment of gearbox is performed in Section 5 to verify the validity of the proposed fault diagnosis method. Finally, Section 6 provides the conclusion.

2. Incremental enhanced supervised locally linear embedding

2.1. Locally linear embedding

Given a data set $\mathbf{X} = \{x_i \in R^D, i = 1, \ldots, N\}$ in the high-dimensional input space, the objective of LLE [12] is regaining the low-dimensional representation of $\mathbf{X}$, which is denoted by $\mathbf{Y} = \{y_i \in R^d, i = 1, \ldots, N\}(d \ll D)$. The central idea of LLE is keeping the local neighborhood structure of the high-dimensional data set $\mathbf{X}$ unchanged while mapping $\mathbf{X}$ into the low-dimensional feature space. LLE mainly contains three steps as follows:

(1) Find the k-nearest neighbors $\mathbf{X}_i = \{x_j^i,\ j = 1, \ldots, k\}$ of each sample $x_i \in \mathbf{X}$.
(2) Compute the reconstruction weights W_{ij} of x_j^i that minimize the reconstruction error of x_j^i to x_i.
(3) Compute the low-dimensional representation $\mathbf{Y}$ of $\mathbf{X}$ that keeps the reconstruction weights W_{ij} unchanged.

In step (1), Euclidean distance is the most commonly used criteria to select the k-nearest neighbors $\mathbf{X}_i$ of x_i. Then in step (2), x_i is reconstructed by its k-nearest neighbors $\mathbf{X}_i$, and the optimal reconstruction weight W_{ij} is obtained by solving the constrained least-squares problem as follows:

$$W_i = \arg\min_{W_i}\left(\sum_{j=1}^{k}\|x_i - W_{ij}x_j^i\|^2\right) \qquad \text{Subject to } \sum_{j=1}^{k} W_{ij} = 1 \tag{1}$$