ARTICLE IN PRESS

Mechatronics xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Mechatronics

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

Sparse classification of rotating machinery faults based on compressive sensing strategy

Gang Tang [a], Qin Yang [a], Hua-Qing Wang [a,*], Gang-gang Luo [a], Jian-wei Ma [b]

[a] School of Mechanical and Electrical Engineering, Beijing University of Chemical Technology, Beijing 100029, China
[b] Department of Mathematics, Harbin Institute of Technology, Harbin 150001, China

A R T I C L E I N F O

Article history:
Received 30 October 2014
Revised 29 March 2015
Accepted 7 April 2015
Available online xxxx

Keywords:
Rotating machinery
Fault diagnosis
Compressive sensing
Random dimensionality reduction mapping
Sparse representation

A B S T R A C T

Rotating machinery is integral to production processes and is composed of several components susceptible to malfunction. Safety and stable operation necessitate early identification of potential faults. Fault identification and classifications for rotating machinery are investigated in this study utilizing expanded monitoring data. A representation classification strategy for rotating machinery faults is developed based on a newly developed compressive sensing theory focusing on extraction and classification of fault features through sparse representation combined with random dimensionality reduction mapping. Original characteristics of a vibration signal are sampled and preserved by applying a small number of random projections and a learning redundant dictionary then constructed to sparsely represent the vibration signal. Fault signal impulse information is determined through an optimization strategy with sparsity promoting. A rolling bearing is utilized as an example in the study with simulations and experiments indicating that fault characteristics may be directly extracted from a small number of random projections without ever reconstructing the vibration signal completely and with minimal prior fault characteristic frequency knowledge required.

1. Introduction

Rotating machinery equipment has become increasingly complex with the development of industrial automation necessitating potential fault early warning actions for industrial equipment management to ensure safe and stable operations. Fault diagnosis is often confronted with many difficulties as sources of complications and manifestations of faulty rotating machinery components, especially rolling bearings have grown more complicated. Bearings subject to dynamic load variety during operation produce vibrations reflecting the status of operation states for both the bearings and the rotating machinery. Fault classification and diagnosis of rotating machinery commonly apply vibration signal analysis for quality and safety assurance.

Non-stationary and nonlinear vibration signals intensify difficulty in fault feature extraction during the industrial production, most notably during early stages. Various methods have been designed to analyze these signal types, *e.g.*, Hilbert–Huang transform (HHT) [1], wavelet analysis [2] and mathematical morphological filtering [3]. Accuracy of the signal mode decomposition, however, may be affected by the end effects [4] and mode mixing problem [5] of HHT. Energy leakage also occurs with a finite length of wavelet basis functions while the selection criterion of wavelet bases remains an issue for further study [6]. Mathematical morphological filtering characteristics lack an exact analysis process with results dependent on filter choices [7]. On the other hand, fault pattern recognition has been conducted to sufficiently reach reasonable classifications based on fault features extractions. Artificial neural network (ANN) algorithm realizes nonlinear mapping between symptoms and faults [8], pursuing superior accuracy at a cost of sample size and convergent rate [9]. Rough sets (RS) method demonstrates positive effects on redundancy removal by attributes reduction, but is ineffective in interference resistance [10]. Support vector machine (SVM) is a machine learning method based on statistical learning theory, and produces a favorable generalization performance, yet challenges present with processing larger samples and multiclass problems as increases are required for training time and computational complexity [11]. Data monitoring for fault classifications is also often cumbersome as high volumes of data are retrieved. Development of new suitable methods for processing the complex data more efficiently and extracting fault features more accurately is necessary.

Compressive sensing [12,13] is a newly developed theory that may contribute to solutions as it emerges as a potential approach for perfect reconstruction where expected detection may be

* Corresponding author.
E-mail address: wanghq_buct@hotmail.com (H.-Q. Wang).

http://dx.doi.org/10.1016/j.mechatronics.2015.04.006
0957-4158/

achieved even with under-sampled signals. Compressive sensing may lower storage costs with transmission and processing capabilities for high-dimensional data and is currently successfully applied to signal processing and imaging in fields such as medical imaging [14], geological exploration [15], radar systems [16] and radio communication [17]. Davenport [18] and Haupt [19] have proved that the under-sampled data in the majority of applications actually retain the original data structure and related information, allowing extraction of required characteristic from sampled values via appropriate detection algorithm, and the signal detection to be realized without ever reconstructing the original signal completely. Further research has progressed as Duarte [20] proposed a detection algorithm based on matching pursuit while Meng [21] applied compressive sensing to the detection of sparse events in wireless sensor networks.

Researchers have preliminarily explored the compressive sensing effectiveness in the field of mechanical diagnosis [22–24], but further studies are needed to attain effective sparse representation and sampling methods for machine failure diagnosis. Compressive sensing theory fundamentals mainly relate to sparse representation and random dimensionality reduction mapping, thus providing promising potential for complex rotating machinery signals with sparse impulse features. A sparse classification strategy for rotating machinery faults will be developed in this paper based on this to resolve fault classification issues as related to large-volume monitoring data for rotating machinery.

Presentation of this study will be in the following order: A brief introduction to compressive sensing theory, Section 2; Description for sparse representation of fault vibration signals based on over-complete dictionary, Section 3; Introduction for the proposed method of random dimensionality reduction mapping and sparse representation classification strategy, Section 4; Simulations and experimental results, Section 5; summary and discussions, Section 6; conclusions, Section 7.

2. Basic idea of compressive sensing theory

Compressive sensing theory dictates that a signal may be optimally reconstructed from a limited number of samples beyond requirements of the Nyquist–Shannon sampling theorem. Recovery of the sampled data relies upon the structure and essential information of the signal and is no longer dependent on bandwidth of the signal from compressive sensing theory perspective. Compressive sensing basics can be described as [25]:

If an m-dimensional signal X can be sparsely represented by $\Psi = [\varphi_1, \varphi_2, \ldots, \varphi_n]$, denoting a pre-chosen transform base or a learning sparse dictionary, then it can be possibly reconstructed accurately by d linear measurements ($d \ll m$) of another non-coherent base $\Phi = [\phi_1^T, \phi_2^T, \ldots, \phi_d^T]$, as:

$$y = \Phi X = \Phi \Psi \alpha = \widetilde{\Psi} \alpha \in \mathbb{R}^d \quad (1)$$

The high-dimensional signal X can be projected onto a low-dimensional measurement y, which is commonly known as a data dimension reduction with the aid of a measurement matrix Φ, and $\widetilde{\Psi}$ as the named sensing matrix and $\mathbb{R}$ as the set of real numbers. The projections of signal X on each vector φ_i of Ψ are described by the coefficients $\alpha_i = \langle X, \varphi_i \rangle$. If a signal is highly approximated by a linear combination of K non-zero coefficients in bases Ψ, then x is K-sparse ($K \ll m$). Each different K sparse vector projected onto the same d-dimensional observation vector is avoided as a "principle of restricted isometry property" (RIP) is developed to ensure a sparse solution and unique reconstruction [26].

Dimension d of y is much smaller than the total number of samples m for $d \ll m$, indicating the problem described by Eq. (1) is underdetermined, and several feasible theoretical solutions for α may exist. If RIP is satisfied [12], according to compressive sensing theory, an accurate sparse solution of α may be pursued by ℓ_1-norm optimization of this under-determined problem. [27]

$$\hat{\alpha}_1 = \arg\min \|\alpha\|_1 \;\; s.t. \quad X = \Psi \alpha \quad (2)$$

where $\|\alpha\|_1$ denotes the ℓ_1-norm, since the practical environment has changed dramatically, the following mathematical model is applied to describe the noisy vibration signal of rotating machinery [26]:

$$X = \Psi \alpha + w \quad (3)$$

where w denotes the random noise or interferences in practical conditions, the minimization problems with sparse constraints based on the compressive sensing theory and sparse representation can be described as [28]:

$$\hat{\alpha}_1 = \arg\min \|\alpha\|_1 \;\; s.t. \quad \|X - \Psi \alpha\|_2 \leqslant \varepsilon \quad (4)$$

where $\| \bullet \|_2$ denotes the Euclidean norm and ε^2 is a likely upper limit of the noise power $\|w\|_2^2$.

3. Sparse representation of fault vibration signals based on over-complete dictionary

Bearing fault signal is utilized as an example to explain the sparse representation of rotating machinery fault signals. Impact forces are generated in each collision resulting in fault impulse signals when a local fault exists in a rolling bearing. Complexity and non-stationary properties of the rolling bearing signals result in acquired vibration signals with high sampling frequencies and large data volume for processing Thus it is urgent and significant to study how to represent the huge vibration data flexibly.

Recent development of compressive sensing theory has generated focused attention on sparse representation. Mallat and Zhang [29] stated that a signal can be represented by an over-complete dictionary with elements of atoms, indicating that sparse representation of a signal can be conducted through an optimal linear combination of several selected atoms in a mapping process. A signal can be represented by or sufficiently approximated to a linear combination of predefined atoms according to the theory. If the atomic bases are well defined, the represented coefficients would be extremely sparse, then the compression efficiency may be greatly improved and processing cost largely reduced [30].

A rolling bearing signal is often so complex that it is not always sparse in common orthogonal transform bases. To achieve a sparse representation for k categories of rolling bearing faults, a target training sample set E is defined.

Provided there are n_i samples in the ith category training set E_i:

$$E_i = [v_{i1}, v_{i2}, \ldots, v_{in_i}] \in \mathbb{R}^{m \times n_i}$$

where $v_{ij} \in \mathbb{R}^{m \times 1}$ indicates the jth sample of the ith category fault, $i = 1, 2, \ldots, k;\; j = 1, 2, \ldots, n_i$.

The training sample set E composed of all the k categories target training samples can be expressed as:

$$E = [E_1, E_2, \ldots, E_i, \ldots, E_k]$$

$$= \left[v_{11}, \ldots, v_{1n_1}, v_{21}, \ldots, v_{2n_2}, \ldots, v_{i1}, \ldots, v_{in_i}, v_{k1}, \ldots, v_{kn_k}\right] \in \mathbb{R}^{m \times n}$$

$$n = n_1 + n_2 + n_3 + \cdots + n_k$$

Another m-dimensional signal x belonging to ith category fault that is not the component of E can be expressed as a linear combination of the training set E:

$$x = E\alpha + w \in \mathbb{R}^m \quad (5)$$

The signals x of each category are defined as testing datasets, i.e. monitoring data of practical conditions and ensured to be different for each test, where $\alpha = [0, \ldots, 0, \alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in_i}, 0, \ldots, 0]^T \in \mathbb{R}^n$ is