Mechanical Systems and Signal Processing 94 (2017) 499–524

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

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

Sparsity-aware tight frame learning with adaptive subspace recognition for multiple fault diagnosis

Han Zhang [a,b], Xuefeng Chen [a,b,*], Zhaohui Du [a,c], Boyuan Yang [a,b]

[a] State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an 710049, China
[b] Collaborative Innovation Center of High-End Manufacturing Equipment, Xi'an Jiaotong University, Xi'an 710054, China
[c] Department of Mathematics, University of California, Los Angeles, CA 90095, USA

ARTICLE INFO

Article history:
Received 6 May 2016
Received in revised form 21 November 2016
Accepted 26 February 2017

Keywords:
Sparse representation
Dictionary learning
Transform sparse coding
Discriminative indexes
Nonconvex optimization
Multiple feature decoupling

ABSTRACT

It is a challenging problem to design excellent dictionaries to sparsely represent diverse fault information and simultaneously discriminate different fault sources. Therefore, this paper describes and analyzes a novel multiple feature recognition framework which incorporates the tight frame learning technique with an adaptive subspace recognition strategy. The proposed framework consists of four stages. Firstly, by introducing the tight frame constraint into the popular dictionary learning model, the proposed tight frame learning model could be formulated as a nonconvex optimization problem which can be solved by alternatively implementing hard thresholding operation and singular value decomposition. Secondly, the noises are effectively eliminated through transform sparse coding techniques. Thirdly, the denoised signal is decoupled into discriminative feature subspaces by each tight frame filter. Finally, in guidance of elaborately designed fault related sensitive indexes, latent fault feature subspaces can be adaptively recognized and multiple faults are diagnosed simultaneously. Extensive numerical experiments are sequently implemented to investigate the sparsifying capability of the learned tight frame as well as its comprehensive denoising performance. Most importantly, the feasibility and superiority of the proposed framework is verified through performing multiple fault diagnosis of motor bearings. Compared with the state-of-the-art fault detection techniques, some important advantages have been observed: firstly, the proposed framework incorporates the physical prior with the data-driven strategy and naturally multiple fault feature with similar oscillation morphology can be adaptively decoupled. Secondly, the tight frame dictionary directly learned from the noisy observation can significantly promote the sparsity of fault features compared to analytical tight frames. Thirdly, a satisfactory complete signal space description property is guaranteed and thus weak feature leakage problem is avoided compared to typical learning methods.

1. Introduction

Rotating machinery covers a broad range of engineering equipment and plays an important role in modern industrial systems, such as aircraft engines, wind turbines and power plants. Accordingly, fault diagnosis of rotary machines has taken on

---

* Corresponding author at: State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an 710049, China.

E-mail addresses: zhanghan12@stu.xjtu.edu.cn (H. Zhang), chenxf@mail.xjtu.edu.cn (X. Chen), duzhaohui3@stu.xjtu.edu.cn (Z. Du), yangboyuanxjtu@163.com (B. Yang).

http://dx.doi.org/10.1016/j.ymssp.2017.02.043
0888-3270/© 2017 Elsevier Ltd. All rights reserved.

Nomenclature

y	observed signal
Y	training matrix
Ω	sparse representation dictionary
$\Omega^{\dagger}$	pseudo-inverse of Ω
a	sparse representation coefficient
A	sparse representation coefficient matrix
$\mathcal{S}_f$	linear convolution operator
$\{f_i\}_{i=1}^{r}$	tight frame filter banks
$\mathrm{HT}(\cdot)$	hard-thresholding operator
SRT	sparse representation theory
SATF	sparsity-aware tight frame
WTSC	wavelet tight frame based sparse coding
TQWT	tunable-Q wavelet transform
SVD	singular value decomposition
MAD	median of absolute deviation technique
SNR	signal to noise ratio
CC	correlation coefficient
RRMSE	relative root-mean-square error
ISNR	improved signal to noise ratio
SDFF	significance degree of fault feature

new significance. It is due to its ability to reduce maintenance costs and prevent the harmful and even devastating consequences of failures [1]. The diagnosis can be summarized as the procedure of identifying the cause and effect relationship between multiple signatures and multiple sources. The key of diagnosis lies in extracting and identifying the fault signatures from the measured vibration signals. However, fault features are usually weak and coupling due to severe background noises and feature subspace coupling phenomenon, which make the multiple fault separation and detection problem very challenging, especially at the early stage of incipient faults [2,3].

During the past two decades, numerous efforts have been devoted to the development of vibration-based feature detection techniques in various industrial processes and systems [4–6]. Among which, traditional transform-based methods, such as frequency-domain analysis, time-frequency analysis and time-scale analysis, have been widely studied and their superiority is that fault information could be dramatically enhanced through transforming the collected signal into another appropriate space [7–10]. The success of these methods lies in the matching degree between the transform basis and the target fault feature. During the last decades, these transformations that satisfy the tight frame condition are more attractive as they are not only fast in implementation procedure, but also enjoy the full signal subspace description property, which is very beneficial in weak fault feature identification. Therefore, there are enduring efforts on designing and seeking various tight frame basis which are more suitable to the focused feature signals. However, the oscillation morphology of fault features varies greatly in practice, and one transform working well for one type of fault feature may not work for another. Therefore, the effectiveness of predefined tight frame based methods is undermined seriously due to lack of adaptivity, especially confronting with multiple fault feature coupling problem.

Aiming at designing adaptive transformations, the sparse representation theory (SRT) has been one of the most hot topics in the signal processing society and aroused extensive interests. In recent years, many works are devoted to the improvement of the sparse recovery algorithms [11–13], and its extensive application fields [14–16]. Especially, SRT has obtained huge success in image processing [17], wireless communication and sensor networks [18]. As for machine fault diagnosis community, only a few studies based on SRT have been done [19–22]. The core of sparse approximation based diagnosis is to exploit the low dimensional intrinsic subspace of the feature information through an appropriate sparse representation dictionary which is elaborately designed to the sparse structure of target signals [23]. Compared with traditional transform-based methods, sparse representation dictionary is not predefined, but adaptively optimized to best match the desired feature information, which can be viewed as a feature-orientated strategy.

Main literatures on the feature-orientated dictionary design can be divided into three categories: mathematical model-based approach, dynamic model-based approach and data-driven learning approach. In the first approach, the atoms are constructed by the union of known transform bases to approximately match fault features, such as redundant Fourier dictionary, over-complete Gabor dictionary [24,25,23,26,27]. It is effective for simple fault detection tasks and meanwhile shows great advantage in tight frame design due to its low computational complexity with $O(n)$ or $O(n\log(n))$ floating point operations. However, their performance degrades as these dictionary atoms must satisfy some mathematical property, but not tailored to the fault features. Therefore, the dynamic model based approaches suggest inferring the dictionary from the dynamic response of mechanical faults and discard those stringent mathematical property to meet the engineering requirements. An adaptive impulsive dictionary incorporating the rotational speed and the spall size of the bearing is established by