Journal of Sound and Vibration 400 (2017) 270–287

Contents lists available at ScienceDirect

Journal of Sound and Vibration

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

Weighted low-rank sparse model via nuclear norm minimization for bearing fault detection

Zhaohui Du [a,b], Xuefeng Chen [a,b,*], Han Zhang [a], Boyuan Yang [a], Zhi Zhai [a], Ruqiang Yan [a,c]

[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] School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China

ARTICLE INFO

Article history:
Received 22 January 2017
Received in revised form
9 March 2017
Accepted 31 March 2017
Handling Editor: K. Shin

Keywords:
Sparse singular distribution
Adaptive weighted technique
Nuclear norm regularization
Sparse optimization
Bearing fault detection
Wind turbine

ABSTRACT

It is a fundamental task in the machine fault diagnosis community to detect impulsive signatures generated by the localized faults of bearings. The main goal of this paper is to exploit the low-rank physical structure of periodic impulsive features and further establish a weighted low-rank sparse model for bearing fault detection. The proposed model mainly consists of three basic components: an adaptive partition window, a nuclear norm regularization and a weighted sequence. Firstly, due to the periodic repetition mechanism of impulsive feature, an adaptive partition window could be designed to transform the impulsive feature into a data matrix. The highlight of partition window is to accumulate all local feature information and align them. Then, all columns of the data matrix share similar waveforms and a core physical phenomenon arises, i.e., these singular values of the data matrix demonstrates a sparse distribution pattern. Therefore, a nuclear norm regularization is enforced to capture that sparse prior. However, the nuclear norm regularization treats all singular values equally and thus ignores one basic fact that larger singular values have more information volume of impulsive features and should be preserved as much as possible. Therefore, a weighted sequence with adaptively tuning weights inversely proportional to singular amplitude is adopted to guarantee the distribution consistence of large singular values. On the other hand, the proposed model is difficult to solve due to its non-convexity and thus a new algorithm is developed to search one satisfying stationary solution through alternatively implementing one proximal operator operation and least-square fitting. Moreover, the sensitivity analysis and selection principles of algorithmic parameters are comprehensively investigated through a set of numerical experiments, which shows that the proposed method is robust and only has a few adjustable parameters. Lastly, the proposed model is applied to the wind turbine (WT) bearing fault detection and its effectiveness is sufficiently verified. Compared with the current popular bearing fault diagnosis techniques, wavelet analysis and spectral kurtosis, our model achieves a higher diagnostic accuracy.

* Corresponding author at: State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an 710049, China.
E-mail addresses: duzhaohui3@stu.xjtu.edu.cn (Z. Du), chenxf@mail.xjtu.edu.cn (X. Chen), zhanghan12@stu.xjtu.edu.cn (H. Zhang), yangboyuanxjtu@163.com (B. Yang), zhaizhi@mail.xjtu.edu.cn (Z. Zhai), ruqiang@seu.edu.cn (R. Yan).

http://dx.doi.org/10.1016/j.jsv.2017.03.044
0022-460X/

1. Introduction

The signature detection and accurate diagnosis of faults in bearing systems has become a main requirement for many industrial and on-board applications. When faults do occur and the bearing fails in service, the result could, at best, be the loss of production and income, or, at worst, damage to the industrial process and potentially to the operators and/or the environment. For example, in WT's systems, recent studies show that bearings cause around 70% of gearbox downtime and $21-70\%$ of generator downtime (21% on small generators ($P<1$ MW), 70% on medium generators (1 MW $<P<2$ MW) and 50% on large generators ($P>2$ MW)) [1]. Therefore, fault detection and diagnosis of bearing systems has received increasing attention since the last two decades [1–6]. Detecting the faults reliably, identifying their locations and analyzing their causes are essential to ensure the safety, reliability, efficiency, and performance of industrial applications.

Most of mechanical faults for bearing system produce feature frequency components in the vibration signal, and meanwhile these feature frequencies can be generated depending on which surface is affected by the faults. There is one type of fault frequency, called ball pass frequency of the faulty element, associated with each of the four bearing elements. Accordingly, there are four types of faults: outer-race fault, inner-race fault, and ball and cage fault. From current literatures in bearing fault diagnosis, lots of diagnostic methods based on vibration analysis can be roughly classified into index based monitoring and advanced fault frequency detection. These index based techniques directly calculate statistical parameters, such as Root Mean Square (RMS), Mean, Kurtosis, Crest Factors, and then judge if key indexes are over pre-specific alarm thresholds [7]. These techniques are simple and effective, meanwhile opt to online monitoring, but strong noises and interferences often degrade their performances. Moreover, alarm thresholds are often empirically set and lack of physical interpretations. Thus, many advanced detection techniques have been widely studied and gained great popularity. Two representative methods are multi-resolution wavelet shrinkage and spectral kurtosis filtering. The former is to concentrate the energy of the feature information on a few wavelet coefficients while spreads all strong noises throughout all wavelet coefficients, which makes it suitable to reveal fault-related signals. The latter often decomposes the objective signals by a set of overlapping filter banks and then selects the best band for the filtering step in the process of envelope analysis. There are too many studies and it is a huge work for anyone who attempts to review all the vast achievements. Excellent review articles are recommended as [2,8,9].

One popular data-driven diagnostic strategy, truncated singular value decomposition (TSVD), recently is introduced to perform impulsive feature detection. The main procedure of TSVD consists of the following steps:

(1) A Hankel data matrix is obtained through partitioning the observation signals into a set of local segments.
(2) SVD operation is adopted to decompose the Hankel matrix into a set of principal components, meanwhile, a singular sequence is constructed through extracting all diagonal singular values.
(3) Based on some statistical assumption or expert knowledge, some large singular values in singular sequence are preserved and other smaller values are set as zeros.
(4) Another data matrix is recovered through synthesizing all nonzero singular values and their corresponding singular vectors. Meanwhile, time waveforms of feature signals are extracted by reversing the recovered matrix into 1-dimensional time series.

Therefore, there are two key problems in the TSVD technique: one is how to construct the Hankel data matrix and another is how to select the number of singular values. In [10], the architecture of Hankel matrix is investigated by a set of numerical experiments and recommended as the half of the whole signal length, which often achieves the best results. Meanwhile, difference spectrum of singular sequence is put forward to the selection of effective singular values. Then, Zhao et al further investigates the similarity of Hankel matrix based SVD and wavelet transform in [11]. Meanwhile, singular value decomposition packet technique is also proposed by Zhao for extraction of weak fault features in [12]. In recent, the TSVD is extended to frequency domain in [13] for bearing fault signal de-nosing. On the other hand, many bearing diagnosis techniques utilize SVD operation and related indexes as pre-processing stage to enhance the significance level of fault components or classification accuracy. The short-time matrix series and singular value ratio are introduced in [14] to select optimal filter parameters for rolling bearing fault diagnosis. In [15], bearing fault signatures are directly extracted from a time-varying singular value matrix through performing SVD on a time-varying Hankel data matrix, which illustrates great superiority in improving the signal-to-noise ratio (SNR). In [16], the acquired signals are decomposed into an additive set of principal components through SVD operation, and then some selected singular values are adopted as fault features for an artificial neural network, which is easy to implement and fault feature is noise immune. Similarly, in [17], effective features are constructed through a short-time energy plus SVD in the time-domain or a discrete cosine transform plus SVD in frequency-domain, and the extracted features are subsequently utilized as the inputs of multi-layer support vector machines for identifying faults of the induction motor. SVD-based bearing feature extraction technique combined continuous hidden Markov Model has been also discussed in [18] comprehensively.

Despite achieving many satisfying performance in bearing fault detection results, these SVD-based techniques above have not revealed the underlying physical mechanism and the rationality of SVD operation. Meanwhile, the resulting preserved singular sequence has a significant deviation with respective to the real sequence. More importantly, from one dimensional observation signals to the recovered bearing feature information, the whole algorithmic procedure is often empirically described and lack of rigorously mathematical models, which hinders their engineering applications and theoretical analysis.