G Model
JMSY-562; No. of Pages 7
ARTICLE IN PRESS
Journal of Manufacturing Systems xxx (2017) xxx–xxx
Contents lists available at ScienceDirect
Journal of Manufacturing Systems
journal homepage: www.elsevier.com/locate/jmansys

45th SME North American Manufacturing Research Conference, NAMRC 45, LA, USA

Virtualization and deep recognition for system fault classification☆

Peng Wang, Ananya, Ruqiang Yan, Robert X. Gao*

Department of Mechanical and Aerospace Engineering, Case Western Reserve University, Cleveland, OH, 44106, USA

ARTICLE INFO

Article history:
Received 18 November 2016
Received in revised form 15 February 2017
Accepted 3 March 2017
Available online xxx

Keywords:
Condition monitoring
Deep machine learning
Virtualization

ABSTRACT

Efficient gearbox health monitoring and effective representation of diagnostic results of dynamical systems have remained challenging. In this paper, a new approach to using deep learning for translating diagnostic results of one-dimensional time series analysis into graphical images for fault type and severity illustration is presented, with gearbox as a representative example. Specifically, time sequences are first converted by wavelet analysis to time-frequency images. Next, a deep convolutional neural network (DCNN) learns the underlying features in the time frequency domain from these images and performs fault classification. Experiments on gearbox data demonstrates effectiveness and efficiency of the developed approach with a classification accuracy better than 99.5%.

1. Introduction

The widespread application of gearboxes in manufacturing and transportation has continually motivated research that aims at effective and efficient gearbox health monitoring and fault diagnosis techniques. A majority of the techniques reported in the literature utilize vibration signals acquired by sensors for fault analysis.

The most important components in gearbox vibration spectra are the tooth meshing frequency and its harmonics, together with sidebands caused by modulation phenomena [1]. Increments in the number and amplitude of such sidebands may indicate a fault occurrence and deterioration. Traditional methods to obtain vibration spectra are Fast Fourier Transforms (FFT) and Short-Time Fourier Transforms (STFT). It has been shown that FFT is unable to reveal the dynamic features of non-stationary signals, while STFT cannot achieve satisfactory resolution in both the time and frequency domains at the same time [2]. In recent years, it has been demonstrated that time-frequency distribution of the wavelet transform is more effective in identifying frequency transitions along time than the traditional approaches, owing to its multi-resolution feature localization capabilities [3]. Other spectral analysis methods such as Hilbert transforms are applied as a post-analysis technique to wavelet transform, to identify the frequency components of interest. Fault detection and fault severity classification in this approach are achieved by visual inspection of the numbers and energies of sidebands. This limits the technique to small volume data analysis. With large scale data to be analyzed, it is advantageous to have an automated classification system to interpret the time-frequency distributions of the vibration signals and make the decision (classified gearbox fault type and severity).

There has been ongoing research towards using neural networks in combination with wavelet transforms to achieve good classification accuracy. Use of neural networks for fault classification has been demonstrated previously in the case of helicopter gearbox [4]. However, multi-sensor installation was required to achieve class separability in fault classification [5] and the accuracy was dependent on feature selection. In this paper, an automatic approach for feature identification based on deep learning is established. Deep neural architecture has been used in numerous classification applications such as image recognition [6,7], object classification [8] and handwritten digit data classification [9]. Deep neural networks have also been used in health diagnosis such as in the case of electric power transformers and aircraft engines [10], and electro motors [11]. Moreover, it has been demonstrated that such networks are robust, and can be trained on large scale data [12] and to be not affected by image distortion [13].

In this paper, wavelet analysis and deep convolution neural network (DCNN) are integrated together for gearbox fault severity classification. The advantage of using wavelet transforms with deep neural architecture is two-fold. Wavelet transforms can represent time series data in the time-frequency domain, capturing the fault related frequency components effectively. With converted 2D images based on acquired wavelet coefficients, DCNN, as a translator, can dig and extract the deep features in the images related to fault severity, and finally classify the vibration data.

☆ Peer review under responsibility of the Scientific Committee of NAMRI/SME.
* Corresponding author.
E-mail address: robert.gao@case.edu (R.X. Gao).

http://dx.doi.org/10.1016/j.jmsy.2017.04.012
0278-6125/

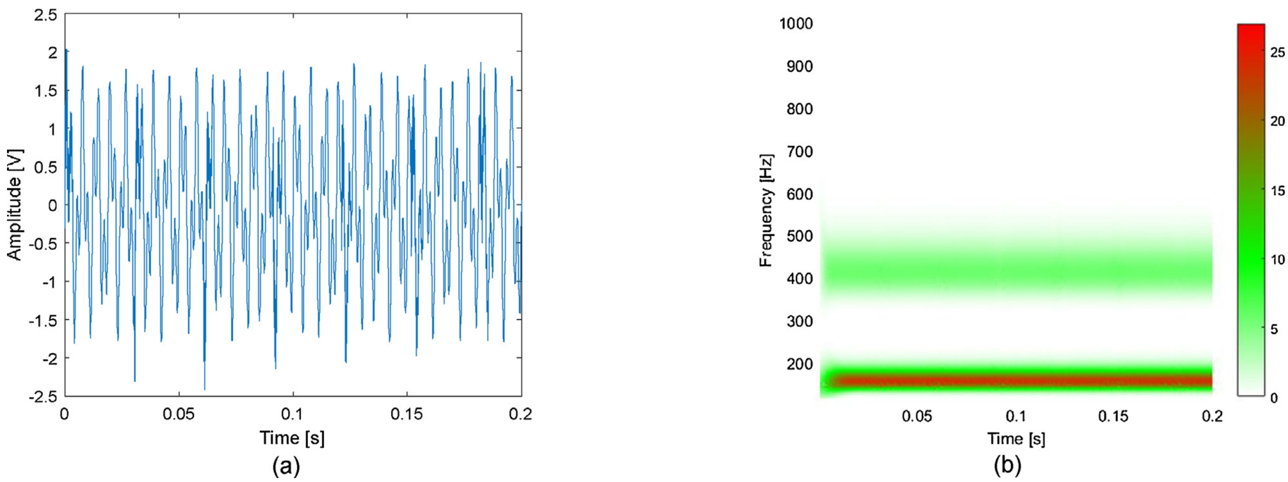

Fig. 1. (a) Vibration Signal and (b) Corresponding Time-Frequency Spectral through Wavelet Transform.

The remainder of the paper is organized as follows. The theoretical background for the proposed method of wavelet analysis and deep learning is discussed in Section 2. Section 3 deals with the experimental setup and the network structure, and also provides the results. Concluding remarks are presented in Section 4.

2. Diagnosis methods

A combination of wavelet transforms and deep learning analysis is proposed for health condition monitoring of gearbox faults. In this section, fundamental theory of wavelet analysis and DCNN are first introduced, followed by the illustration of automated fault severity classification based on time-frequency images through integrated wavelet and DCNN.

2.1. Wavelet analysis

Fault detection directly from the time series, e.g., vibration signals measured from the gearbox, is difficult. Various methods are used to process these time series into a suitable time-frequency domain. A time-frequency representation simultaneously describes the occurrence of a signal component and its frequency development [14]. In traditional methods like STFT, high resolution in both frequency and time domain is not possible. However, such a resolution is necessary in the case of fault detection where time-frequency information of fault occurrence is not available beforehand.

Wavelet transform [15] is a method which is able to represent the signals in multiple resolutions. In early stages of fault initiation, the vibration component associated with the fault is small and requires high sensitivity deduction. It has been shown that time-frequency distributions are effective in detecting such faults [16]. Additionally, different faults produce vibration components of different duration. Since wavelet transform uses varying window sizes, all the components can be displayed simultaneously.

Considering a vibration signal $x(t)$, it can be decomposed into a series of wavelet coefficients at different scales through wavelet transform [3]. Mathematically, it can be expressed as

$$WT_x(t, \quad s) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} x(\tau)\, \psi_{s,t}{}^{*} \left(\frac{\tau\text{-}t}{s} \right) d\tau \tag{1}$$

where s is the scaling factor for dilation, t is the time for translation, $\psi_{s,t}{}^{*}(\cdot)$ denotes the conjugate of the function $\psi_{s,t}(\cdot)$, which represents a wavelet family that is generated by dilating and translating a fixed valued wavelet base $\psi(t)$, and expressed as

$$\psi_{s,t}(\cdot) = \frac{1}{\sqrt{s}} \psi \left(\frac{\tau - t}{s} \right)_{(s,t) \in R^2} \tag{2}$$

where **R** denotes the set of real numbers. By choosing a suitable wavelet base, the vibration signal can be transformed into the time-frequency domain, which is able to reflect the transition of frequency components over time with a high resolution.

Fig. 1(a) shows a simulated vibration signal with 2 gear meshing frequency components (160 Hz and 420 Hz) modulated on a 3600 Hz base and a white noise. Formulation of this vibration signal is $e^{-400*t}\sin(2\pi * 3600 * t) + 0.7 * \sin(2\pi * 420 * t) + \sin(2\pi * 160 * t)$. Fig. 1(b) denotes the time-frequency spectra converted by wavelet transform [17]. The gear meshing frequencies 420 Hz and 160 Hz can be seen clearly.

2.2. Deep convolutional neural network

Deep neural network (DNN) is a type of artificial neural network with multiple hidden layers of units embedded between the input and output layers, and it has been gaining popularity in the recent years, particularly in the field of image processing and classification. Images contain features that reflect the underlying architecture that makes each image a unique representation. These features might not appear to be visually distinct between images. However, deep neural networks can be trained to learn such features at different levels of abstraction [18].

Deep convolutional neural network (DCNN) is one of the DNN structures. It consists of hierarchically arranged trainable stages that “learn” the efficient internal representation for all data [19]. The neurons in a DCNN are arranged in the form of feature maps. In the case of image processing, the input to the network is a 2D array representing the image. A typical convolutional network consists of two to four stages of convolutional layers and pooling layers.

The input to a convolutional layer is an image x of size $m \times n$. The convolutional layer contains k kernels (filters) of size $p \times q$, smaller in dimension than the input image. The output of the convolutional layer is a set of k feature maps of size $(m-p+1) \times (n-q+1)$ by striding over one pixel. For example, as shown in Fig. 2, a kernel with size 5×5 is applied to deliver a pixel in the output map. The kernel, realized by assigning a weight k_{ij} to each pixel in the input image and calculated as weighted sum, extracts certain features (e.g. edge information) contained in the image. The weighted sums are then added by an additive bias and passed through a non-linear