Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Sound and Vibration 287 (2005) 614–624
JOURNAL OF SOUND AND VIBRATION
www.elsevier.com/locate/jsvi

Short Communication

Feature separation using ICA for a one-dimensional time series and its application in fault detection

Ming J. Zuo*, Jing Lin, Xianfeng Fan

Department of Mechanical Engineering, University of Alberta, Edmonton, Alta., Canada T6G 2G8

Received 27 August 2002; received in revised form 24 January 2005; accepted 6 February 2005
Available online 7 April 2005

Abstract

Principal component analysis (PCA) is a method that transforms multiple data series into uncorrelated data series. Independent component analysis (ICA) is a method that separates multiple data series into independent data series. Both methods have been used in fault detection. However, both require signals from at least two separate sensors. To overcome this requirement and utilize the fault detection capability of ICA and PCA, we propose to use wavelet transform to pre-process the data collected from a single sensor and then use the coefficients of the wavelet transforms at different scales as input to ICA and PCA. The effectiveness of this method is demonstrated by applying it to both a simulated signal series and a vibration signal series collected from a gearbox. The results show that the method of combining wavelet transform and ICA works better than the method of combining wavelet transform and PCA for impulse detection based on a one-dimensional vibration data series.

1. Introduction

Vibration signals from a gearbox carry abundant information about its health. To detect possible faults in meshing gears, bearings, or other components, it is necessary to isolate the signals generated by the faults from the collected composite vibration signals.

*Corresponding author. Tel.: +1 780 492 4466; fax: +1 780 492 2200.
E-mail address: ming.zuo@ualberta.ca (M.J. Zuo).

0022-460X/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsv.2005.02.005

Independent components analysis (ICA) was proposed a few years ago for feature separation. It requires little prior knowledge about the components to be isolated; however, at least two sensors must be available for signal collection and the number of sensors must be at least equal to the number of sources to be separated. ICA has been successfully used in acoustic signal separation [1,2] and image de-noising [3–5]. Recently, it has been introduced to dynamic signal analysis of mechanical systems [6,7]. However, this method cannot be applied directly when there is only one sensor collecting signals.

Principal component analysis (PCA) is a multivariate data analysis technique that transforms a set of correlated variables into a set of uncorrelated variables [8]. Each member of the resulting set of uncorrelated variables is called a principal component. We are interested in determining its suitability for fault detection because one of the identified principal components may reveal the signature of a hidden fault. As with ICA, however, this method cannot be applied directly when only a single variable is observed.

Wavelet transform may be considered as a series of band pass filters when applied to the data collected from a single sensor. The results of the transform, which exist in different frequency regions, say N regions ($N>1$), may be considered as different mixtures of the sources that have generated the collected signals. These N groups of data can then be used as input to ICA or PCA for identification of the hidden sources. In this paper, we investigate the potential of wavelet analysis as a data preprocessing method that may enable application of ICA or PCA in fault detection when only one sensor is available for signal collection.

The organization of this paper is as follows. The principles of ICA and PCA are briefly reviewed in Section 2. In Section 3, we outline the method of using wavelet transform to generate data that can then be used as input for ICA analysis in fault detection. In Section 4, we use both a simulated signal series and a real vibration signal series collected from a gearbox to test the performance of the method. The conclusions of this paper are given in Section 5.

2. Principles of ICA and PCA

ICA is a technique for separating independent sources linearly mixed in signals. Suppose that there are N independent sources of vibration, and N sensors at different locations are used to record vibration signals. The signals recorded by each sensor come from different sources with different mixing ratios. Let $s_1(t), s_2(t), \ldots, s_N(t)$ be the signals produced by the N independent sources and $x_1(t), x_2(t), \ldots, x_N(t)$ be the observations from the N sensors. The sensors record these signals simultaneously. The task of ICA is to estimate the mixing ratios of the source signals in the collected signals and obtain the independent source signals.

To identify the independent components successfully, we need a rule for evaluating the independency of the identified components. According to the Central Limit Theorem, the distribution of the sum of a large number of independent random variables tends to a Gaussian distribution. Since the collected signals are weighted sums of the independent sources, the sources to be isolated must have less Gaussianity than the collected signals. Thus, non-Gaussianity can be used for separating independent components. Hyvarinen and Oja [5] proposed to use negentropy to evaluate the non-Gaussianity of the separated components so as to evaluate separation performance. With this concept, we can seek the separation that provides the least Gaussianness of the separated components. The popular FastICA algorithm proposed by Havarinen and Oja [9] is often used to carry out the ICA procedure.

PCA is a technique that obtains linear transformations of a group of correlated variables such that the transformed variables are uncorrelated [8]. For example, consider two variables, x_1 and x_2. For each variable, we have obtained the following N observations:

$$x_{11}, x_{12}, \ldots, x_{1N}; x_{21}, x_{22}, \ldots, x_{2N},$$

where x_{1i} and x_{2j} denote the ith and the jth observations of variables x_1 and x_2, respectively. The PCA method seeks two new axes, D_1 and D_2, that make the projections of the collected data onto D_1 have the largest variability and at the same time, the projections of the collected data onto D_2 have the smallest variability. This way, we have expressed the collected data as their two principal components. Most of the variation in the original data is explained by the first principal component, D_1, and the remaining variation in the original data is explained by the second principal component, D_2.

ICA renders the separated components independent of one another while PCA renders the separated components uncorrelated with one another. PCA separates the components based only on the second-order cumulant while ICA separates the components on high-order cumulants. Therefore, ICA can be considered a generalization of PCA.

3. Feature extraction for one-dimensional series

Whenever there are multiple sensors collecting data, both ICA and PCA may be used directly for feature extraction, however, when there is only one sensor collecting data, neither can be used directly for feature extraction. In this section, we propose a method for preprocessing the available one-dimensional data so that ICA and PCA can be used for feature extraction.

In developing the proposed method, we make the following assumptions:

(1) There is only a single sensor collecting data. The data collected by this sensor is a time series or a sample, which contains information on possible multiple sources.
(2) We are interested in identifying a single source.

This is the situation when we are interested in detecting whether a gearbox contains a gear tooth fault, and we have vibration data from a single sensor mounted on the gearbox. When analyzing this one-dimensional data, we are not at all concerned whether there are faults other than the gear tooth fault; however, after we finish answering the question of whether there is a gear tooth fault, we may use the same one-dimensional vibration data to answer the question of whether there is a bearing fault in this gearbox. The key point is that we are interested in identifying one source at a time. This is different from the original idea behind ICA and PCA, which attempt to identify multiple sources simultaneously.

The available data is a single time series. To apply ICA or PCA for feature extraction, we need to have more than one time series. In the following, we propose a method to generate multiple time series from the single available time series.

When we perform a time-invariant linear transform on a time series, the phase relationships among the independent sources do not vary. If we perform several time-invariant linear transforms one by one on a single time series, we obtain several data series that may be used as input for the ICA or PCA model. Because all are obtained from the same time series, they do not contain enough information for isolation of multiple sources. We will show, however, that they are useful in identifying a specific source.

If the sources hidden in the single time series do not have the same frequency band, and the linear transform applied is a band pass filter, then the transformed signal may cover only a small interval of the frequency band. The mixing ratios for different sources are different for different linear transforms, as shown in Fig. 1. Suppose that distributions $s1$ and $s2$ are the frequency spectra of two independent sources contained in the observed time series. Areas T1 and T2 in Fig. 1 stand for the frequency distributions of two different time-invariant linear transform functions. The two areas have different mixing ratios of $s1$ and $s2$. The independent sources will remain independent after a linear transform is performed. For the same source, the results obtained from different linear transforms are correlated if those linear transforms are correlated. The correlation coefficient of the two time series obtained can be measured by the correlation coefficient of the two transform functions; this, in turn, can be measured by the correlation coefficient of the two filtering series.

As stated earlier, the separated sources from the collected signals will be made independent as much as possible under the ICA iteration procedure. Assume that s_1 and s_2 are two independent components hidden in a single signal series, x. Assume that x_a and x_b are the resulting time series obtained by performing linear transforms T_1 and T_2, respectively (refer to Fig. 1). Let s_{1a} and s_{1b} represent the contributions from s_1 in x_a and x_b, respectively, and s_{2a} and s_{2b} the contributions from s_2 in x_a and x_b, respectively. Then, s_{1a} and s_{1b} are still independent of s_{2a} and s_{2b}, respectively. We also know that s_{1a} and s_{1b} are correlated and s_{2a} and s_{2b} are correlated if these two linear transforms are correlated. The two signal series obtained, x_a and x_b, can then be fed into ICA or PCA for identification of a specific fault source.

Wavelet transform decomposes a signal series in the time domain into a two-dimensional function in the time-scale (frequency) plane. The wavelet coefficients measure the time-scale

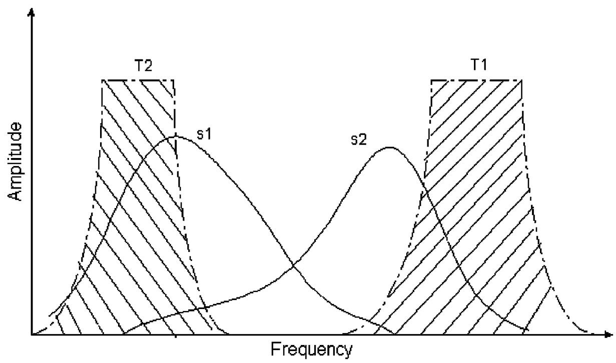

Fig. 1. Illustration of different frequency band occupations after different linear transforms.

(frequency) content in a signal indexed by the scale parameter and the translation parameter. Let $\varphi(t)$ be the mother wavelet. The wavelet family consists of a series of daughter wavelets that are generated by dilation and translation from the mother wavelet $\varphi(t)$:

$$\varphi_{a,b}(t) = \sqrt{|a|}\varphi[(t-b)/a],$$

where a is the scale parameter, b is the location parameter, and $\sqrt{|a|}$ is used to guarantee energy preservation. The wavelet transform of signal $x(t)$ is defined as the inner product of $\varphi_{a,b}(t)$ and $x(t)$ in the Hilbert space of L^2 norm defined as:

$$W(a,b) = \langle \varphi_{a,b}(t), x(t) \rangle = \int x(t)\varphi^*_{a,b}(t)\,\mathrm{d}t,$$

where the symbol * stands for the complex conjugate.

Wavelet transform can be thought of as a series of band pass filters. The results of the transform, which exist in different frequency regions, may be thought of as different mixtures of

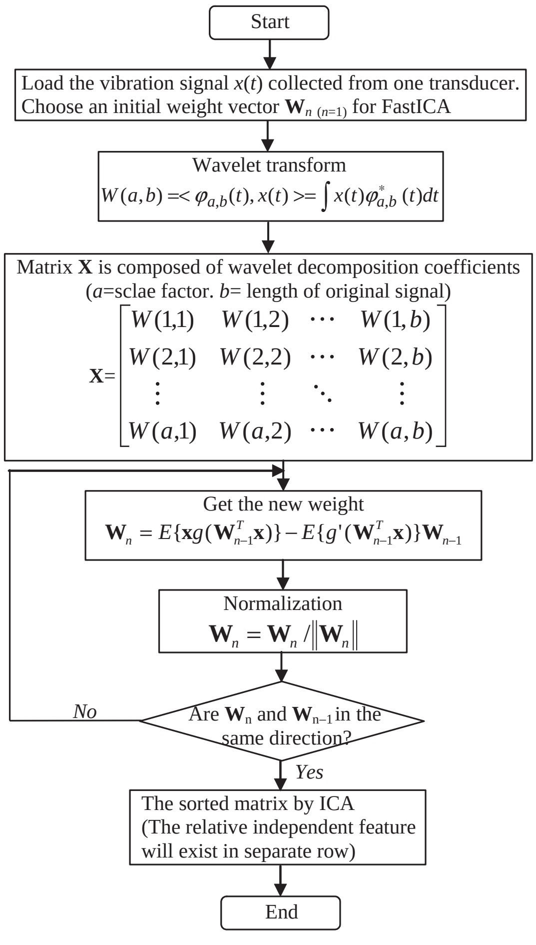

Fig. 2. Flowchart of the signal processing scheme of the proposed method.

the independent sources. These different mixtures may be considered to be signals collected at different "locations", or more accurately, through different "sensors" with different frequency ranges. This way, the one-dimensional vibration signal is transformed into multidimensional data that satisfy the requirements of ICA and PCA. The preprocessing of the one-dimensional vibration data with wavelet transform makes ICA and PCA usable for identification of a hidden source.

Gear tooth fault is a commonly seen gearbox fault. When a tooth is chipped or cracked, an impulse is generated whenever this tooth meshes [10]. The impulses generated by a faulty tooth are different from impulses generated by normal teeth in meshing. When we want to detect gear tooth fault, we look for these impulses in the collected vibration signal. Morlet wavelet is often used in the wavelet transform because it matches individual impulses well. This helps to capture impulses, as illustrated in [11]. FastICA algorithm [9] is used to perform ICA. The procedure for combining wavelet transform and ICA is shown in Fig. 2.

4. Experimental studies

To verify the feasibility of the proposed method for feature detection, we first use a simulated signal. The simulated signal is composed of periodic impulse signals $e^{-100t}\sin(200t)$ with a period of 500 s and a chirp signal $\sin(50\pi t^2)$. These two independent sources are shown in Fig. 3(a) and (b). The mixed signal of these two independent sources is shown in Fig. 3(c). The mixing ratio between the impulses and the chirp signal is 0.2:1. Because periodic impulses often represent fault signatures in machinery, our task is to detect the existence of periodic impulses in the mixed signals.

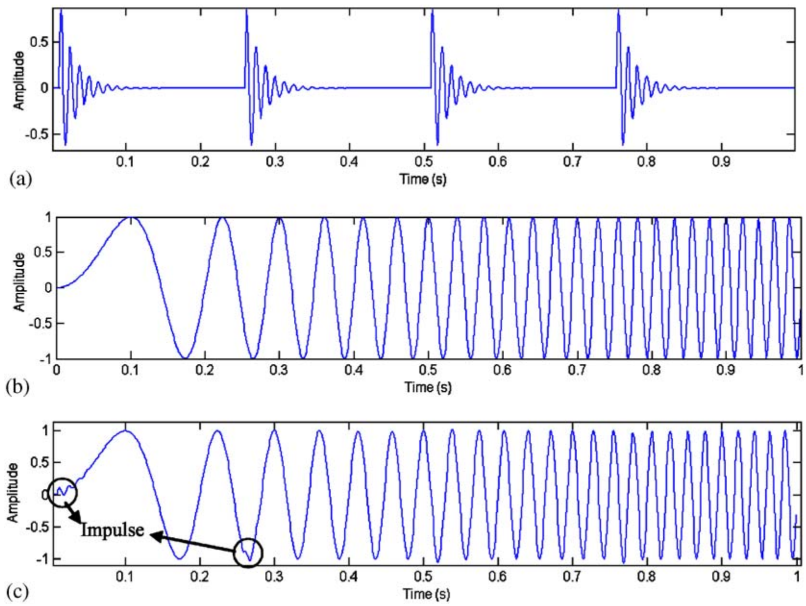

Fig. 3. Waveforms of the simulated signal: (a) the impulse signal; (b) the chirp signal; (c) the mixture signal.

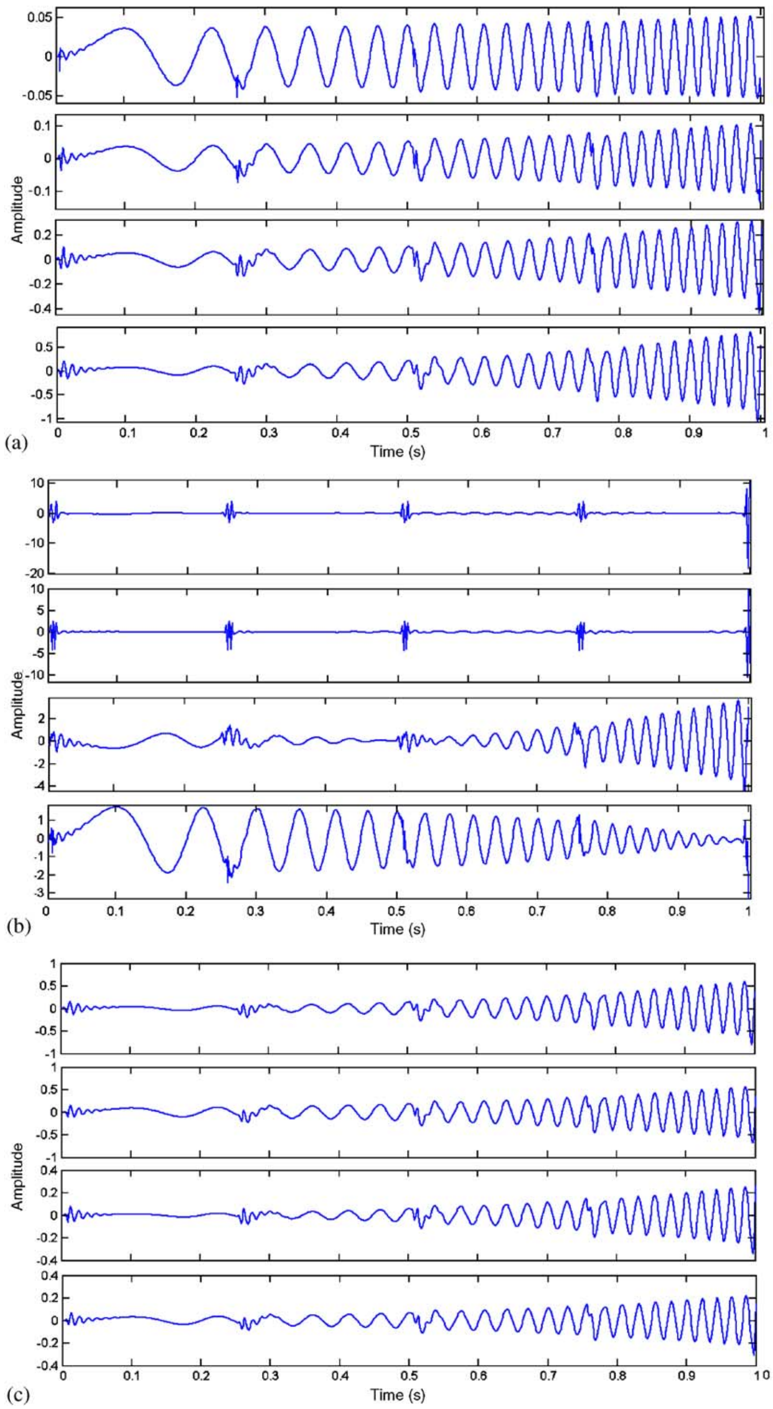

Fig. 4. Analysis results of the simulated signal: (a) the wavelet decompositions; (b) the separation results with ICA; (c) the separation results with PCA.

Fig. 4(a) is the wavelet decomposition of the mixed signal with the scale parameter having values from 1 to 4. There are no obvious periodic impulses at any scale value in Fig. 4(a). ICA is then used to process the obtained matrix of wavelet coefficients, and the results are shown in Fig. 4(b). Fig. 4(b) clearly shows that periodic impulses are present in the first two subplots. This experiment proves that the proposed methods combining ICA and wavelet transform is effective for impulse detection. For comparison, we also used PCA to process the obtained wavelet coefficient matrix. The obtained results are shown in Fig. 4(c). The identified principal components are shown in descending order. Unfortunately, the periodic impulses are not obvious in Fig. 4(c). It would appear that PCA is not as useful as ICA for identification of the impulses in this example.

We next analyzed a signal series collected from a gearbox. The experiment was conducted on a gearbox dynamic simulator produced by SpectraQuest. Spur gears were used in the gearbox. One tooth was broken on gear ♯4. The signal was collected through an accelerometer. The experimental system is illustrated in Fig. 5. The rotating frequencies of the input, middle, and output shafts are 14.9, 4.96, and 8.28 Hz, respectively, and the meshing frequencies of the pairs of gears 1&2 and 3&4 are 238.4 and 198.67 Hz, respectively.

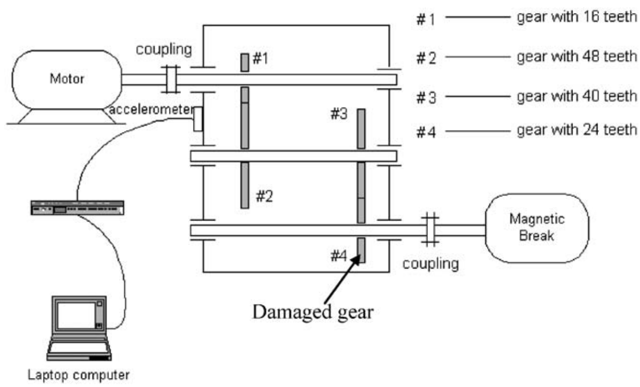

Fig. 5. Illustration of the experimental system.

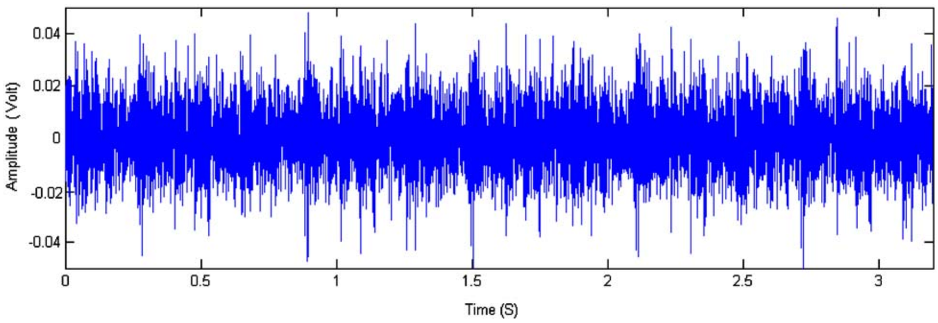

Fig. 6. Waveform of signals from the gearbox with one broken tooth on gear #4.

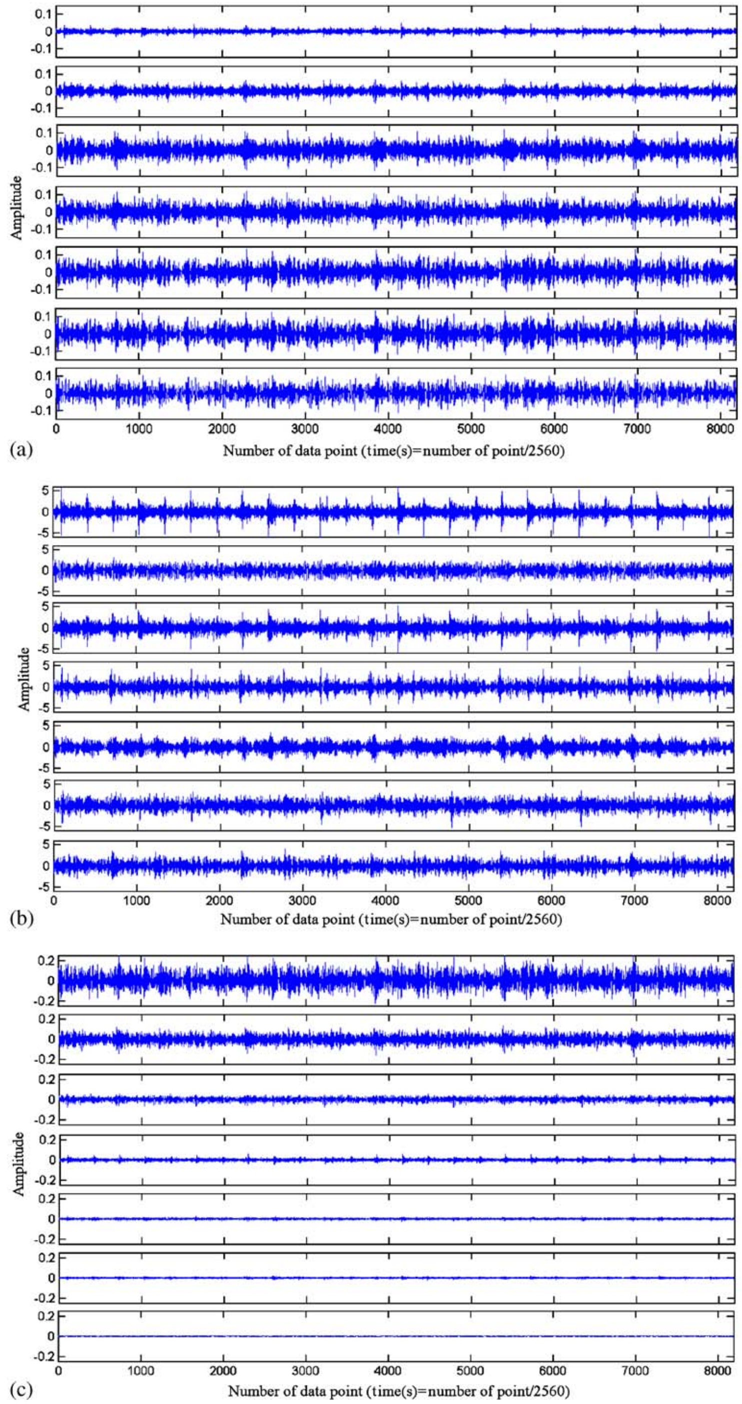

Fig. 7. Analysis results of the gearbox signal: (a) the wavelet decompositions; (b) the separation results with ICA; (c) the separation results with PCA.

Because the damaged gear was on the output shaft, it was expected that periodic impulses would be found in the collected signals with a period of $1/8.28\,\text{Hz} = 0.1208\,\text{s}$. The collected vibration signals are plotted in Fig. 6. The data was sampled at 2560 Hz with low pass filtering at 1000 Hz. Morlet wavelet was employed to decompose the signals from scales 1 to 7 and the results obtained are shown in Fig. 7(a). We then used ICA to process those decomposition coefficients and obtained the results shown in Fig. 7(b). In Fig. 7(b), it is easy to find the periodic impulses which appear in the first subplot. The period of about 0.1208 s is the same as the rotating period of gear ♯4. PCA was also used to process the same data given in Fig. 7(a). The results of processing it using PCA are shown in Fig. 7(c). The corresponding eigenvalues are [38.537, 10.840, 2.5213, 0.7788, 0.2413, 0.1256, 0.0124]. Though some impulses can be found in Fig. 7(c), the time intervals of the impulses are not as obvious as those shown in Fig. 7(b).

5. Conclusions

ICA has been proven to be an effective tool for simultaneous separation of multiple independent sources that are hidden in the collected signals provided by multiple sensors. When, however, there is only a single sensor available for data collection, ICA cannot be applied directly. In this paper, we have proposed to use wavelet transform to preprocess the vibration data collected from a single sensor. With the proposed method, we are able to obtain multiple data series at different scales of the wavelet transform. These multiple data series can then be used as input to ICA or PCA for detection of a single independent source. The proposed method was applied to impulse detection using a simulated signal series and a real signal series collected from a gearbox with a gear tooth fault. It was demonstrated that ICA in combination with wavelet transform worked better for impulse identification than PCA in combination with wavelet transform.

Acknowledgement

This research was supported by the Natural Sciences and Engineering Research Council of Canada. Critical comments from reviewers have been very helpful in improving previous versions of this paper.

References

[1] J.H. Lee, H.J. Jung, T.W. Lee, S.Y. Lee, Speech feature extraction using independent component analysis, *IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. III, 2000, pp. 1631–1634.
[2] H.M. Park, H.Y. Jung, T.W. Lee, S.Y. Lee, On subband-based blind signal separation for noisy speech recognition, *Electronic Letters* 35 (23) (1999) 2011–2012.
[3] T.W. Lee, M.S. Lewicki, Unsupervised image classification, segmentation, and enhancement using ICA mixture models, *IEEE Transactions on Image Processing* 11 (3) (2002) 270–279.
[4] P.O. Hoyer, A. Hyvärinen, A multi-layer sparse coding network learns contour coding from natural images, *Vision Research* 42 (12) (2002) 1593–1605.

[5] A. Hyvärinen, E. Oja, Independent component analysis: algorithms and applications, *Neural Networks* 13 (4–5) (2000) 411–430.
[6] W. Li, F. Gu, A.D. Ball, A.Y.T. Leung, C.E. Phipps, A study of the noise from diesel engines using the independent component analysis, *Mechanical Systems and Signal Processing* 15 (6) (2001) 1165–1184.
[7] G. Gelle, M. Colas, C. Serviere, Blind source separation: a tool for rotating machine monitoring by vibrations analysis, *Journal of Sound and Vibration* 248 (5) (2001) 865–885.
[8] J.E. Jackson, *A User's Guide to Principal Components*, Wiley, New York, 1991.
[9] A. Hyvärinen, E. Oja, A fast fixed-point algorithm for independent component analysis, *Neural Computation* 9 (7) (1997) 1483–1492.
[10] W.J. Stazewskin, G.R. Tomlinson, Local tooth fault detection in gearboxes using a moving windows procedure, *Mechanical Systems and Signal Processing* 11 (3) (1997) 331–350.
[11] J. Lin, M.J. Zuo, Gearbox fault diagnosis using adaptive wavelet filter, *Mechanical Systems and Signal Processing* 17 (6) (2003) 1259–1269.