Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Sound and Vibration 282 (2005) 899–917
JOURNAL OF SOUND AND VIBRATION
www.elsevier.com/locate/jsvi

Design of mixture de-noising for detecting faulty bearing signals

Yimin Shao*, Kikuo Nezu

Department of Mechanical System Engineering, Gunma University, Tenjin-cho 1-5-1, Kiryu City 376-8515, Japan

Received 18 July 2003; received in revised form 21 January 2004; accepted 15 March 2004

Abstract

Detecting the waveform of a noisy signal is a key problem in the detection of early bearing faults under actual plant conditions. Mixture de-noising is found to be a useful technique for identifying bearing signals and greatly improves the fault diagnostics of the bearings. The mixture de-noising technique consists of an adaptive noise-canceling filter and a wavelet-based de-noise estimator. The mixture de-noising technique can substantially improve the signal-to-noise ratio when the signal is contaminated by noise. The performance of mixture de-noising under different noise ratios, bearing failure sizes and shaft speeds are discussed in this paper. This paper shows that the diagnostic role of failure diagnosis and analysis techniques can be made more effective with the application of mixture de-noising.

1. Introduction

Vibration signature-based diagnostics are mainly concerned with the extraction of features from a diagnostic signal which can indicate whether the vital components of a machine are good or defective [1]. Vibration monitoring is based on the principle that all systems produce vibration. When a machine is operating properly, vibration is small and constant; however, when faults

*Corresponding author. Current address: College of Mechanical Engineering, Chongqing University, Chongqing, China.
E-mail address: ymshao@cqu.edu.cn (Y. Shao).

0022-460X/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsv.2004.03.051

develop and some of the dynamic processes in the machine change, the vibration spectrum also changes.

For the localized defects of rolling element bearings, each time the rolling element passes over the localized defect, an impulse of vibration is generated. The impulses are related to the defect and its severity, and all vibration-monitoring techniques are essentially based on the recording and quantification of these vibration impulses [2]. Numerous methods have been developed for detecting the localized faults of bearings: for example, statistical methods [3], time and frequency domain analyses [4], adaptive noise canceling [1,5–7] (ANC), and neural network identification methods [1,8]. However, there is still no ideal way to diagnose the early failure of bearings from noisy signals. One main reason is that the signal-to-noise ratio is lower, and the impulse signal is more difficult to extract when the bearings are running under actual plant conditions.

The time wave of a typical bearing failure is similar to that shown in Fig. 1. Although there is a defect on the outer race of the bearing in Fig. 2, the fault is not easily identified from the time waveform and enveloping spectrum of the vibration signal. This is because the vibration signal of the bearing includes the heavy background noise of normal bearing vibration and the noise from the vibration of the machine.

Although adaptive filtering can be applied to remove the machine noise [1,5,7], the noise cancellation is not significantly successful because the noise from the normal bearing vibration is stronger than the noise from other sources. Mixture de-noising, with its noise-cancellation capability, can separate the bearing fault signal from noise in actual plant conditions.

Section 2 of this paper formulates the problem of bearing fault signals and noise, and briefly describes the algorithm for mixture de-noising. The performance of mixture de-noising by simulation is discussed in Section 3, and experimental results are presented in Section 4. Conclusions are provided in Section 5.

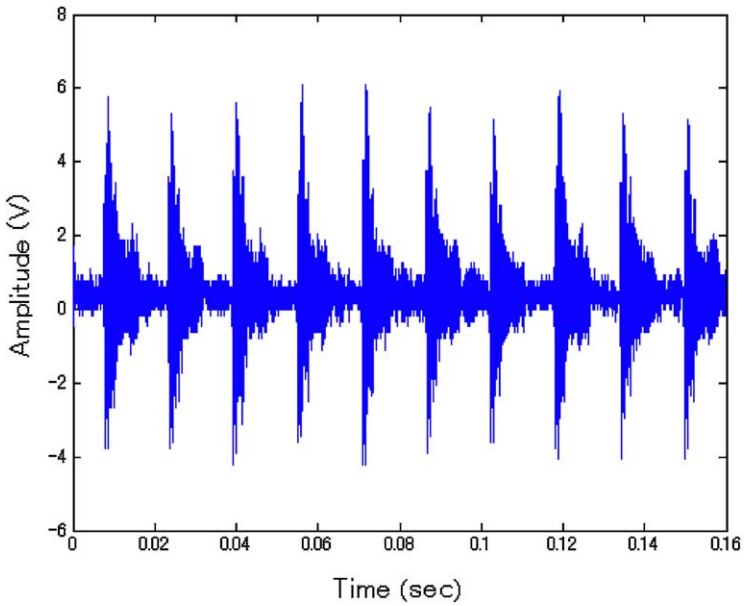

Fig. 1. Typical time wave form of noise signal of faulty bearing.

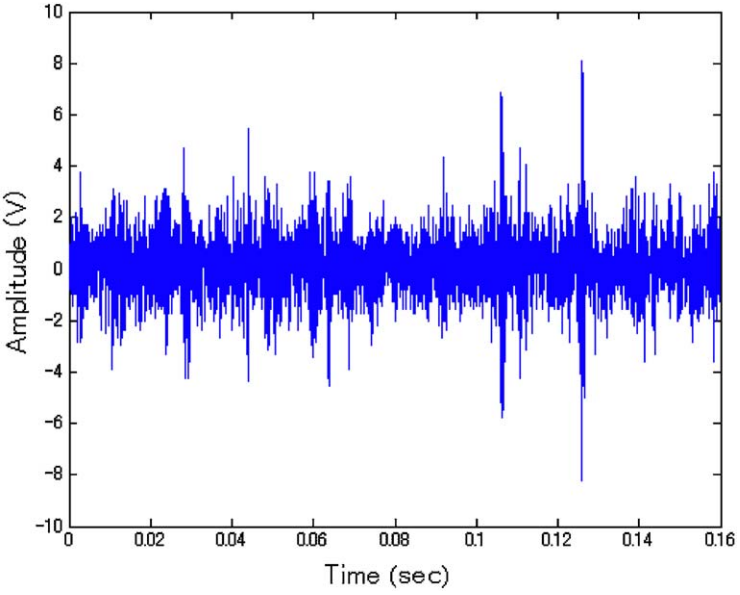

Fig. 2. Noise vibration signal of bearing with local defect.

2. The mixture de-noising principle

The general mixture de-noising concept is shown in Fig. 3. The mixture de-noising technique consists of an adaptive noise-canceling filter and a wavelet-based de-noise estimator. The mixture de-noising technique can substantially improve the signal-to-noise ratio when the signal of interest is contaminated by noise.

2.1. Review of adaptive noise canceling

A signal S corrupted with noise N_0 is received at the primary sensor while the bearing runs under faulty conditions, as shown in Fig. 3. S is the fault signal of the bearing and N_0 is the noise of normal bearing vibrations. A reference noise N_1, which is related to noise N_0 in some unknown way but not correlated with signal S, is received at the reference sensor. The filter output y is then adaptively filtered to match N_0 as closely as possible. Then, the filter output is subtracted from the primary input $S + N_0$ to produce the system output ε:

$$\varepsilon = S + N_0 - y. \tag{1}$$

Then

$$\varepsilon^2 = S^2 + (N_0 - y)^2 + 2S(N_0 - y). \tag{2}$$

Taking the expectation of both sides of Eq. (2)

$$E(\varepsilon^2) = E(S^2) + E((N_0 - y)^2). \tag{3}$$

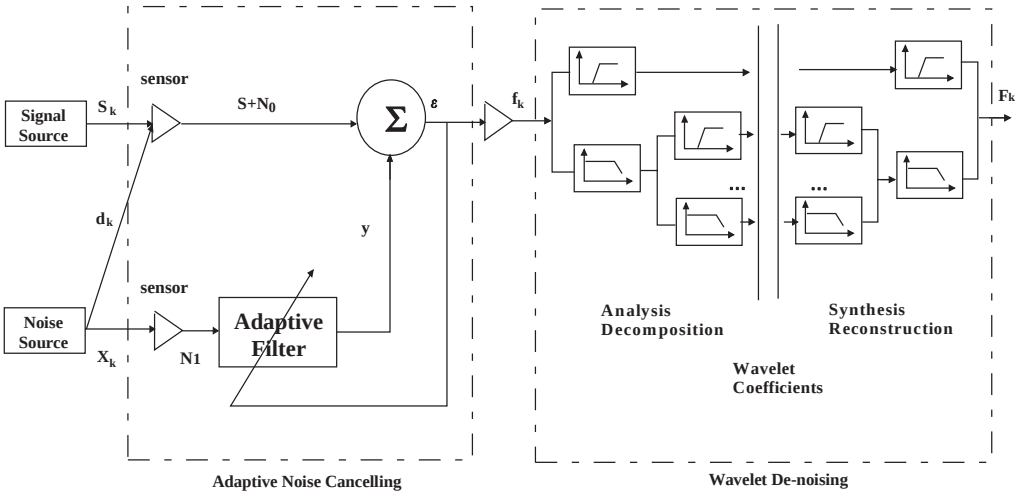

Fig. 3. Mixture de-noising in a machine noise canceling.

The adaptive filter output is

$$y = \mathbf{W}^{\mathrm{T}}\mathbf{X}, \tag{4}$$

$$E(\varepsilon^2) = E(S^2) + E((N_0 - \mathbf{W}^{\mathrm{T}}\mathbf{X})^2), \tag{5}$$

where $\mathbf{W}$ is the weight vector and $\mathbf{X}$ is the input vector. The signal power $\mathbf{E(S^2)}$ will be unaffected as the filter weights are adjusted to minimize $\mathbf{E}(\varepsilon^2)$:

$$\min\ E(\varepsilon^2) = E(S^2) + \min\ E((N_0 - \mathbf{W}^{\mathrm{T}}\mathbf{X})^2). \tag{6}$$

For an optimal set of filter weights, the output y is the best least-squares estimate of primary noise N_0. From the following equation:

$$(\varepsilon - S) = (N_0 - y) \tag{7}$$

the adjustment of the filter weights to minimize the output power causes ε to be the best least-squares estimate of the signal S.

The ANC filter is easily implemented by computer software. The implementation method is the same as that of the LMS (least-mean-squares) algorithm. This algorithm is shown by the following equation [9]:

$$\mathbf{W}_{j+1} = W_j + 2\lambda\varepsilon_j\mathbf{X}_j, \tag{8}$$

where $\mathbf{W}_j$ is the weight vector at the jth instant of time, $\mathbf{X}_j$ is the reference input vector at the jth instant of time, ε_j is the error signal at the jth instant of the output of the noise canceler, and λ is the gain constant that regulates the speed and stability of adaptation. To ensure the convergence

of the adaptive algorithm, the value of λ should be chosen such that [9]

$$0<\lambda<\frac{1}{(L+1)(\text{signal power})}, \tag{9}$$

where L is the index of the last filter weight.

2.2. Review of wavelet de-noising

The classical additive model of a signal $f(x_n)$ corrupted with noise z_n is shown as follows [10]

$$s_n = f(x_n) + \sigma z_n, \quad n = 1, 2, 3, \ldots, N, \tag{10}$$

where $s_n = (s_1, s_2, s_3, \ldots, s_N)$ represents the observed signal, and $f(x_n)$ is the purely deterministic signal. The z_n values are identically and independently distributed Gaussian random variables. It is assumed that f can be well represented by a linear combination of P wavelet basis functions ϕ_p, as follows [11]:

$$f(x) \approx \sum_{p=1}^{N} \alpha_p \phi_p(x), \tag{11}$$

where $\alpha = (\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_p)$ are wavelet coefficients. Orthogonal wavelets are used in this paper.

This principle is also the basis for the popular wavelet threshold method that reduces the noise in signals. Small wavelet coefficients are assumed to be dominated by noise and carry little information. Replacing these coefficients by zero eliminates a major part of the noise without affecting the signal very much. The new wavelet coefficients can be estimated by the hard and the soft function [11] of Eq. (12):

$$\begin{aligned} \eta_\beta^{\text{hard}}(x) &= x1(|x|>\beta), \\ \eta_\beta^{\text{soft}}(x) &= \text{sign}(x)(|x|-\beta), \end{aligned} \tag{12}$$

where β is the smoothing parameter, or the threshold. Several ways of selecting β have been proposed for the Gaussian noise [10–12]. The β of Eq. (13) [10] was selected in this paper:

$$\beta = \sigma\sqrt{2\log M}, \tag{13}$$

where M is the number of samples.

2.3. A simplified machine model for the signal of bearing vibration

The detected signal from the sensor for the defective bearing can be divided into three parts: the signal of a normal bearing, the signal of other rotating components, and the signal of the fault bearing. The signal of the defective bearing is defined as the "signal" in this paper. Fig. 4 shows a simplified model for noise canceling in a machine. It is assumed that there is an equivalent noise source corresponding to the various noise sources, and this noise reaches the primary and

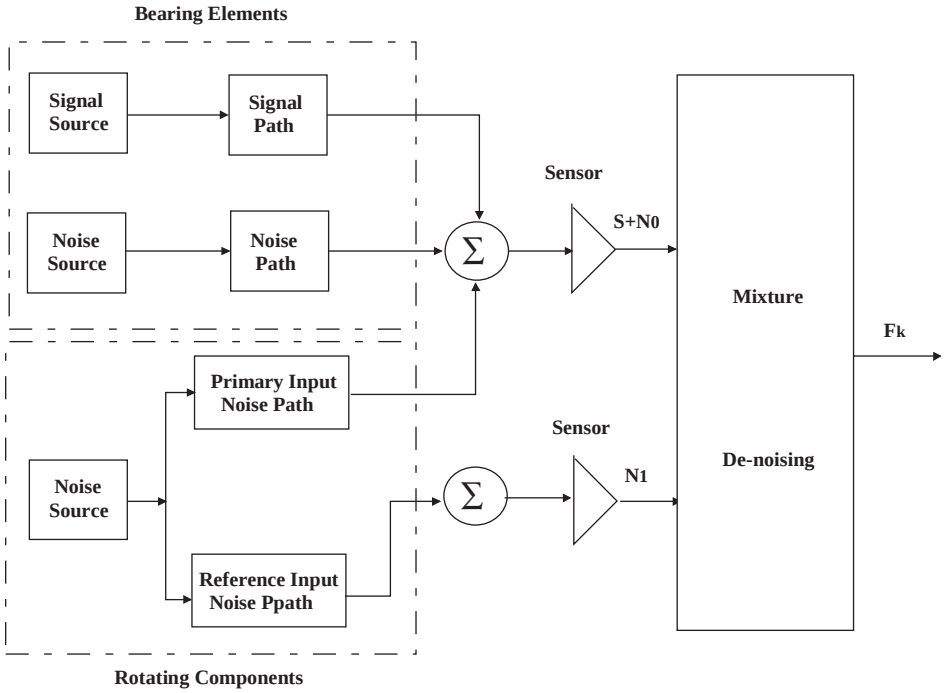

Fig. 4. A simplified model for noise canceling in a machine.

reference sensors through equivalent channels. All propagation paths are assumed to be equivalent to linear time-invariant filters.

For actual test operation, two cases are considered for the fault diagnosis of bearings. In the first case, the noises in the primary and reference inputs are correlated and the signals are relatively uncorrelated or weakly correlated. In the second case, some noises in the two inputs are uncorrelated (these noises are assumed to be random noise).

According to the mixture de-noising principle, the noises from nearby rotating components will be canceled by the adaptive filter, and uncorrelated random noise and normal bearing vibration will be canceled by wavelet de-noising. The defective bearing will be easily detected from the signal by the mixture de-noising.

3. Simulations

The performance of the mixture de-noising estimator was examined by a simulation signal. The simulation was presented with a Gaussian impulse of a periodic signal with a cosine signal (15 Hz) and the random signal in the primary input. The random signal includes a standard deviation of 1 and a standard deviation of 4. The reference input includes a cos signal (15 Hz) and a random signal of standard deviation 4. The relationship between the de-noising ability and the noise ratio

of the signal (from 30% to 90%) is discussed in this section. The noise ratio of the signal is defined as follows:

$$k = \sum_{n=1}^{N} |z_n| \Big/ \sum_{n=1}^{N} |s_n| \times 100\%, \tag{14}$$

where z_n is the Gaussian noise signal, and s_n is the contaminated signal. The noise ratio of the signal in this paper is typical of the ratio of Gaussian noise and the contaminated signal in the primary input shown in Fig. 4; it does not include other types of noise (for example, the noise of the periodic signal).

The kurtosis factor was chosen to evaluate the performance of the mixture de-noising. The kurtosis factor is defined as follows:

$$\text{kurtosis factor} = \int_{-\infty}^{+\infty} |x|^4 p(x)\,\mathrm{d}x \Big/ \left(\int_{-\infty}^{+\infty} |x|^2 p(x)\,\mathrm{d}x \right)^2, \tag{15}$$

where $p(x)$ is the probability density function. The value of the kurtosis factor of the Gaussian impulse is 8.1 in this study. The de-noised signal will contain the noise signal if the value of the kurtosis factor of the signal is smaller than 8.1, and the de-noised signal will be distorted if the value of the kurtosis factor of the signal is greater than 8.1.

3.1. Mixture de-noising applied to simulated data

The noisy signal is shown in Fig. 5. The result of de-noising with the signal of Fig. 4 using an adaptive filter is shown in Fig. 6. Fig. 6 shows that there are some noises in the estimated signal after application of the adaptive filter. The result of de-noising the signal of Fig. 4, using mixture

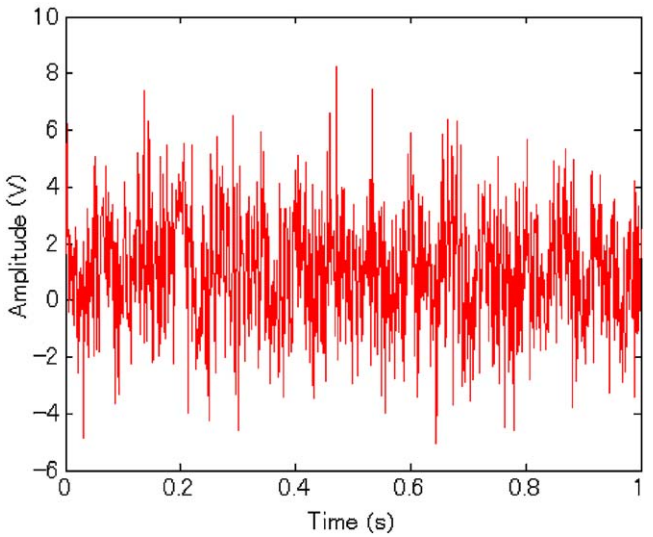

Fig. 5. Noisy signal with a noise ratio of 50%.

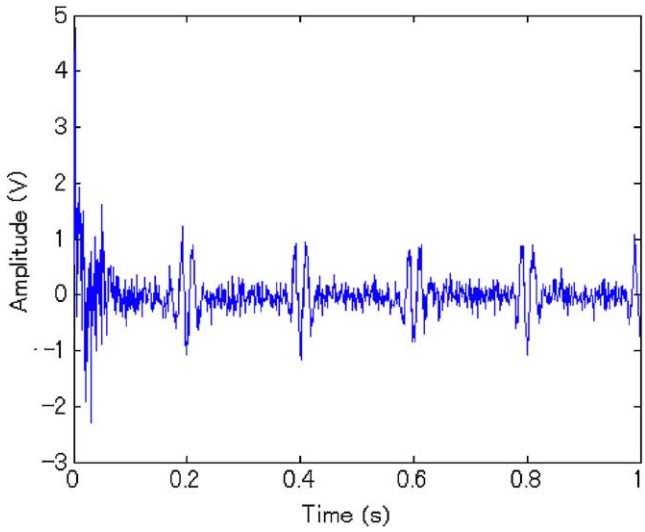

Fig. 6. Output signal of adaptive filter.

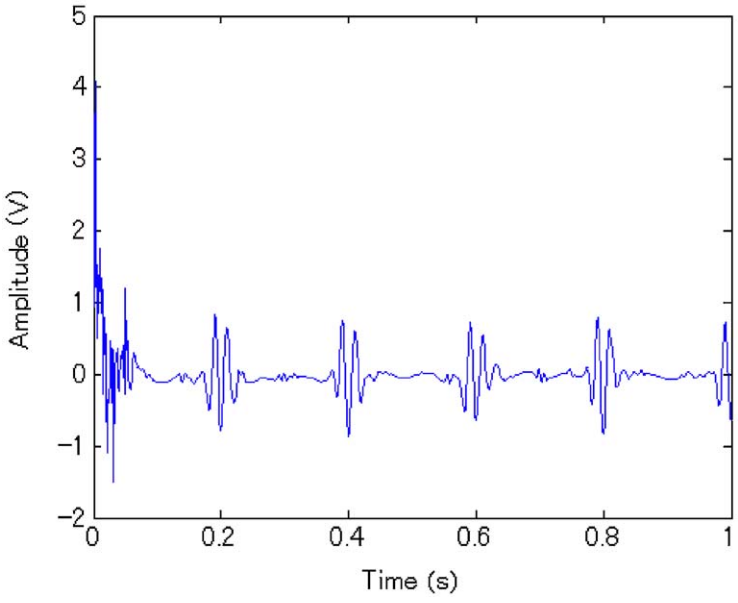

Fig. 7. De-noised signal by mixture de-noising.

de-noising, is shown in Fig. 7. A comparison of the waveform of the de-noised signal in Fig. 7 with the signal in Fig. 6 shows that the two waveforms are very similar, but the noise of the signal has been canceled in Fig. 7. The value of the kurtosis factor is 3.2 in Fig. 4. The value of the kurtosis

factor is 4.9 in Fig. 6 after application of the ANC, but the value of the kurtosis factor reaches 8.0 in Fig. 7. Fig. 7 shows that, when the signal is contaminated by noise, mixture de-noising produces good results.

3.2. The relationship between learning ratios and noise ratios

The adaptive filter is a key step for mixture de-noising because wavelet de-noising is weak when the noise ratio is more than 90% [12], or when the noise signal contains non-Gaussian noise. The relationship between the learning ratios of the adaptive filter and the noise ratios of the signal is discussed as follows.

The relationship between the learning ratios of adaptive filtering and the noise ratios of the signal is shown in Fig. 8. The noise ratios range from 0% to 100%, and the learning ratios of the adaptive filter range from 0.0 to 0.06. Fig. 8 shows that the signal will be distorted after de-noising when the learning ratio of the adaptive filter is over 0.05, because the values of the kurtosis factor are bigger than 8.1.

A zoomed view of the graph of Fig. 8 is shown in Fig. 9. In Fig. 9, there are two phases of increases and decreases as the noise ratio increases. The values of the kurtosis factor increase when the noise ratio ranges from 60% to 80%. The peak ratio is a decreasing process when the noise ratio ranges from 80% to 100%. Fig. 9 shows that the value of the kurtosis factor is almost the same as the values (near 8.0) when the noise ratio is less than 60%, and the signal is hard to reconstruct from the noisy signal when the noise ratio is more than 90%.

The values of the kurtosis factor increase when the noise ratio range is 60–80% because the signal is distorted in the de-noised signal. The de-noised signal for the 70% noise ratio is shown in

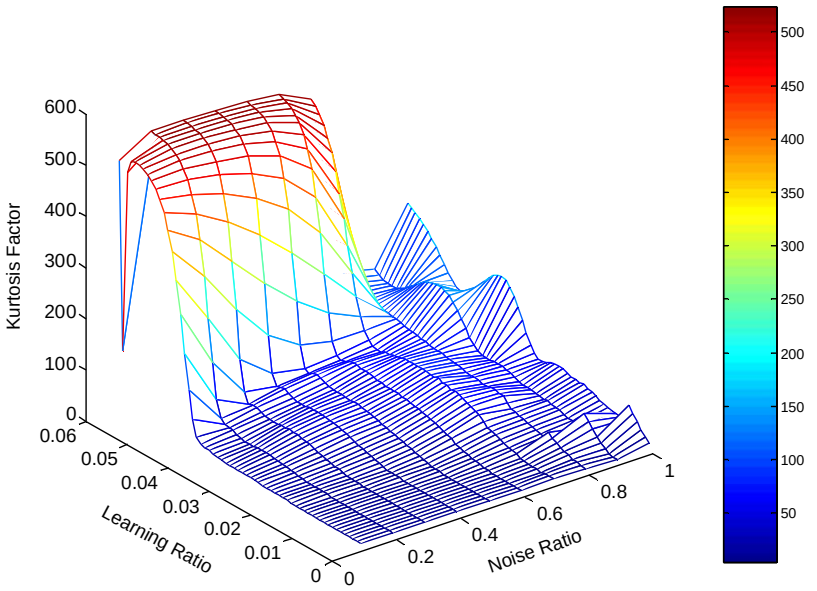

Fig. 8. Relationship between learning ratios and noise ratios.

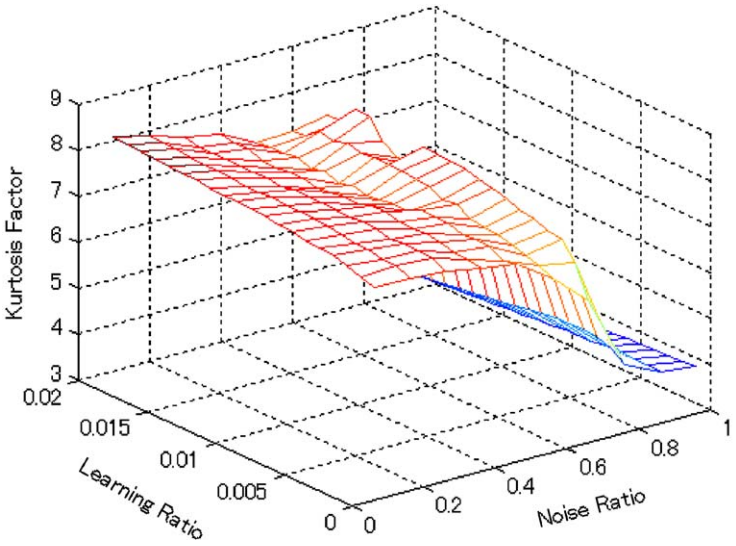

Fig. 9. A zoomed view of Fig. 8.

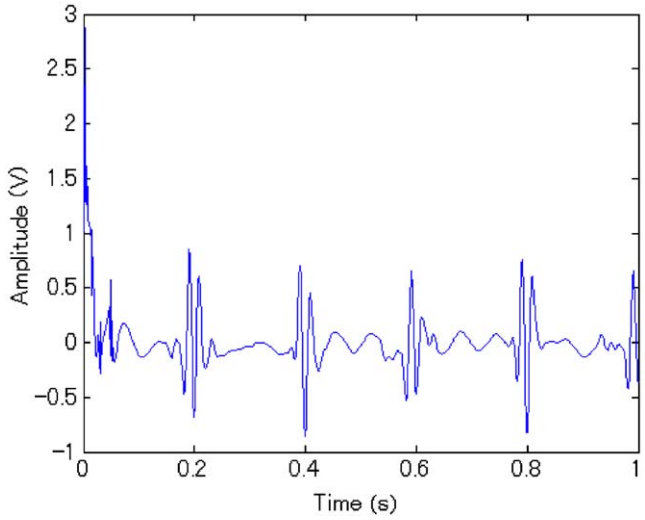

Fig. 10. De-noised signal by mixture de-noise (noise ratio 70%).

Fig. 10. A comparison of the waveforms of the de-noised signal in Fig. 10 and the signal in Fig. 7 shows that the two waveforms are very similar, but the de-noised signal is distorted because the value of the kurtosis factor is more than 8.1 in Fig. 10. These results show the effect of mixture de-noising relative to the noise ratio of the observation signal.

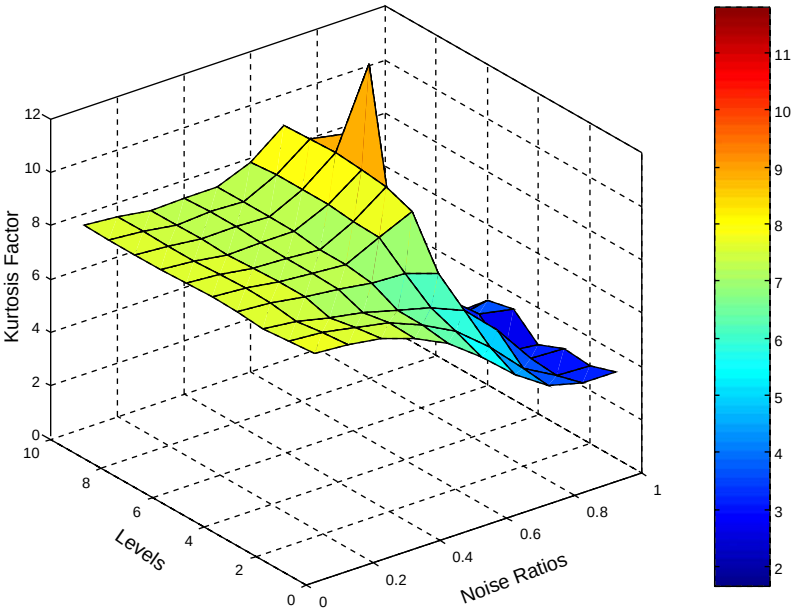

Fig. 11. Relationship between noise ratios and levels of wavelet decomposition.

3.3. The relationship between levels of wavelet decomposition and noise ratios

Fig. 11 shows the relationship between the levels of wavelet decomposition and noise ratios when the learning ratio of the adaptive filter is 0.008. There are three phases of stable, increasing and decreasing values of the kurtosis factor as the noise ratios increase and the levels of wavelet decomposition are more than 5. The values of the kurtosis factor are near 7.8 in the stable phase, when the noise ratios range from 0% to 60%. The values of the kurtosis factor range from 7.8 to 8.7 in the increasing phase, when the noise ratios range from 60% to 80%. The values of the kurtosis factor range from 8.7 to near 3 in the decreasing phase, when the noise ratios range from 80% to 100%. These results show that de-noising works well for different noise ratios when the levels of wavelet decomposition are greater than 5.

3.4. The variation of mixture de-noising with filtering time

The variation of adaptive neural filtering with filtering time is shown in Fig. 12. Fig. 12 shows that the process of mixture de-noising is stable after 2 s when the learning ratio is 0.008 and the level of wavelet decomposition is 8. This result shows that the process of mixture de-noising is not fast, but the result is very good.

4. Mixture de-noising applied to actual machine data

Fig. 13 shows the experimental arrangement, which includes a motor, pulley system, two types of bearings, housing, and a load device. The motor was controlled by an alternating governor

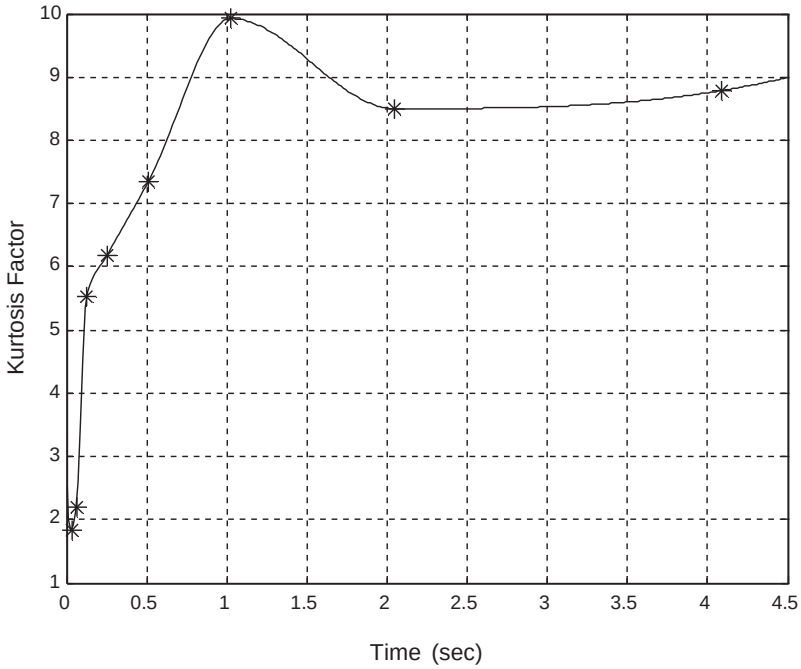

Fig. 12. Performance of mixture de-nosing.

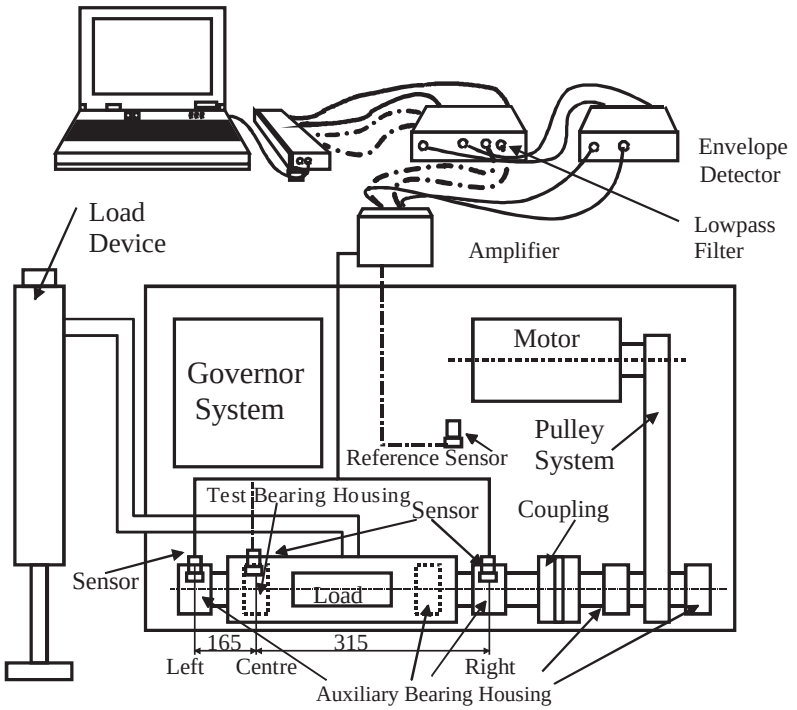

Fig. 13. Experimental arrangement.

system (FUJI Frenic 5000 G9S/P9S), and the bearing load and rotating speed were adjusted according to experimental requirements. Since the motor and bearings were housed in the same foundation, it was possible to simulate the noise conditions of the vibration under different loads and speeds. One defective bearing was installed in the center position. The type of bearings used in the study was NSK N210. The bearing defects were artificially localized in a rectangular shape in the outer race and roller. The failure sizes (length mm × width mm) were as follows: small (8.7×0.3), medium (3.3×0.9), and large (1.9×0.9). The test bearings were subjected to two loads of 5 and 8 kN. The defective bearing signature was recorded by placing the sensor on the bearing housing 165 mm from the center position of the defective bearing at shaft speeds of 200, 500 and 1700 rev/min.

4.1. The effect of mixture de-noising

Fig. 14 shows the time waveform of the noisy signal of the defective bearing. Fig. 15 shows the result of the adaptive filter for the actual machine data, which shows that the failure occurred in the outer race with the following characteristics: failure size small; shaft speed 1700 rev/min; load 8 kN. Generally, the fault is hard to find using spectral or envelope analysis [13]. The impulse of the defective bearing is also hard to identify in Fig. 15, even after the signal was processed using adaptive filtering. However, the impulse of the defective bearing is easy to identify in Fig. 16, when the signal was processed using mixture de-noising. The value of the kurtosis factor is 2.7 in Fig. 14 and 3.3 in Fig. 15, but the value is 4.3 in Fig. 16 after the application of mixture de-noising. The results show that mixture de-noising has the ability to cancel the machine noise.

Fig. 17 shows the time waveform of the normal signal from the bearing (shaft speed 600 rev/min; load 8 kN). The de-noised normal signal by mixture de-nosing is shown in Fig. 18. The

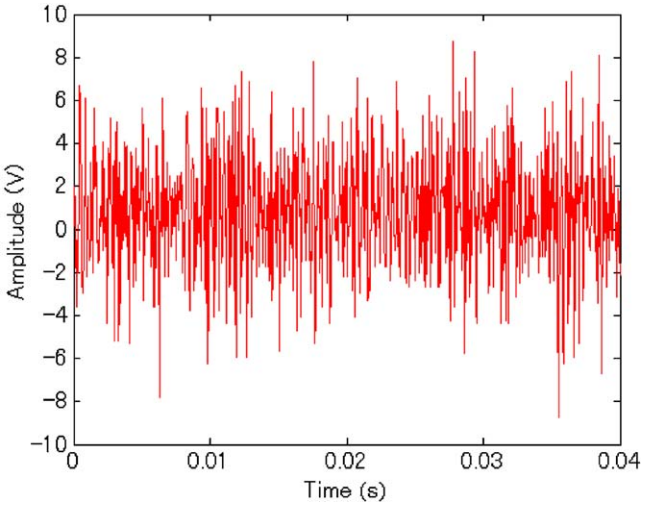

Fig. 14. Time waveform of the noisy signal of the defective bearing.

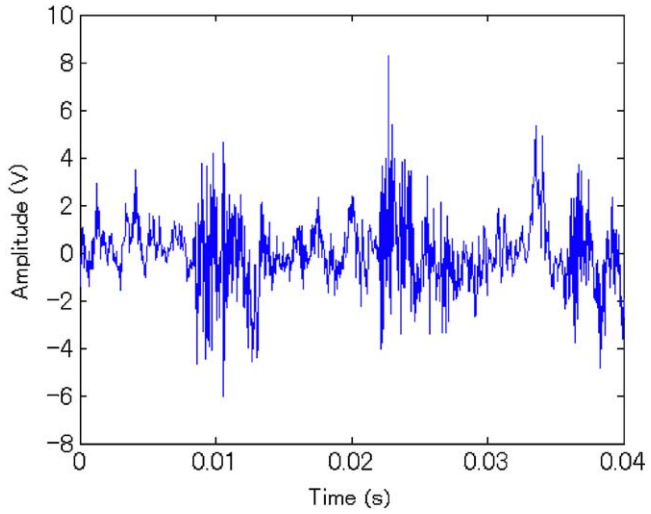

Fig. 15. Output signal of adaptive filtering.

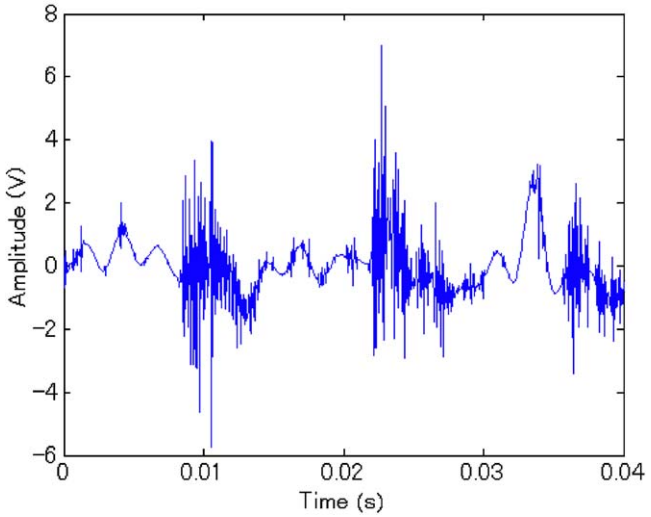

Fig. 16. Output signal of mixture de-noising.

features of the bearing fault cannot be identified from this figure, although the value of the kurtosis factor has increased from 3.02 (in Fig. 17) to 3.63 (in Fig. 18).

4.2. The relationship between kurtosis factor and different experimental conditions

Fig. 19 shows the influence of shaft speeds and failure sizes on the kurtosis factor when the failure of the bearings occurs in the outer race and the measurement distances are 165 mm from

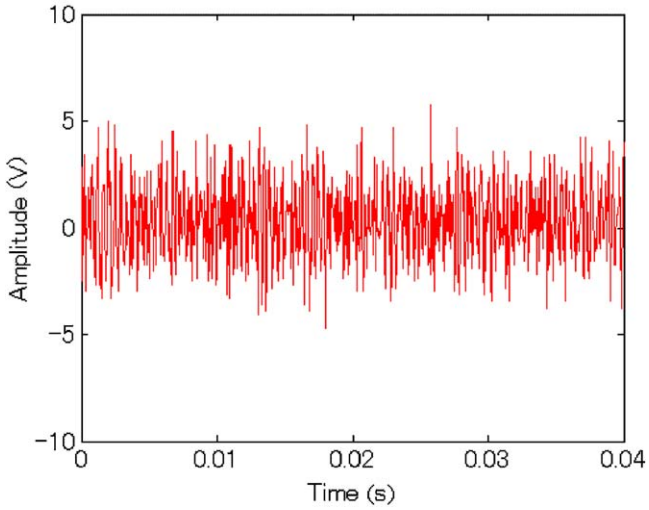

Fig. 17. Normal signal of a bearing.

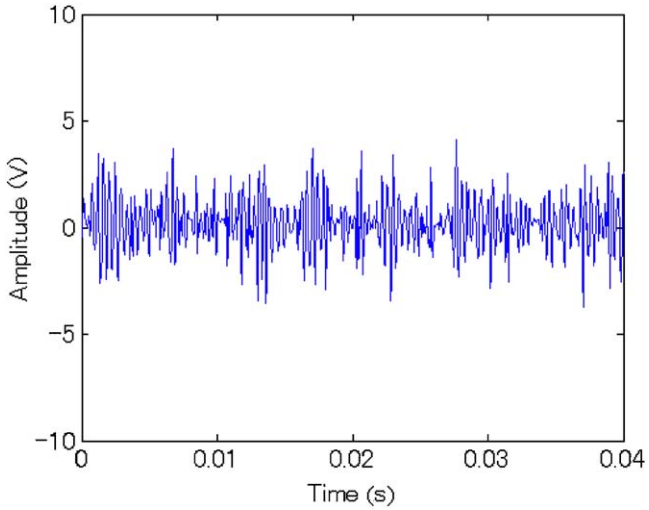

Fig. 18. De-noised normal signal of a bearing.

the sensors to the failure bearings. The shaft speeds of 200, 600, 1100 and 1700 rev/min are shown as 0.1, 0.2, 0.3 and 0.4, respectively, in Fig. 19. The normal, small, medium and large faults of the bearing condition are shown as 1, 2, 3 and 4, respectively, in Fig. 19.

The change in the values of the kurtosis factor for different shaft speeds follows an increasing trend as the shaft speeds increase from low to high, and the range of the increase becomes very

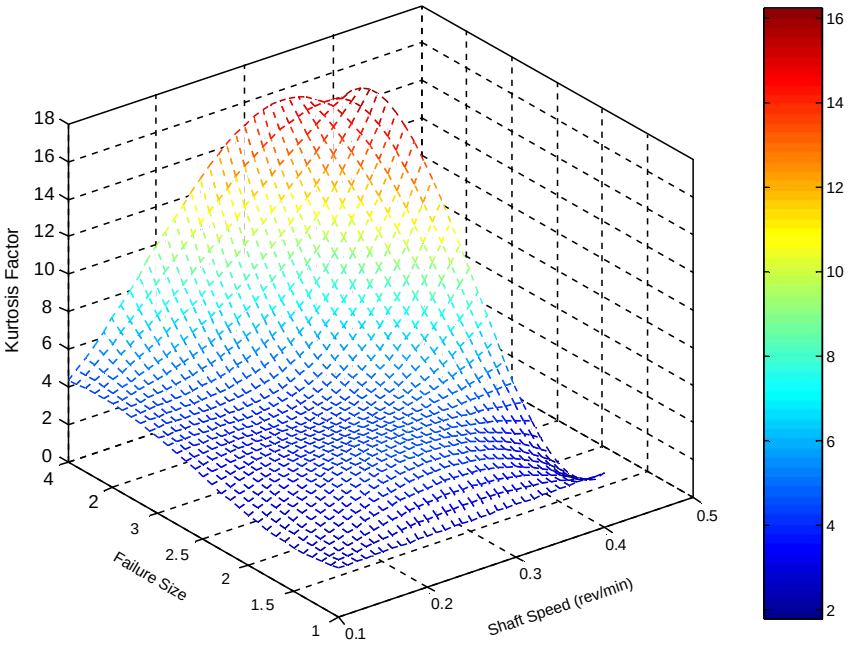

Fig. 19. Relationship between failure sizes and shaft speeds.

small when the shaft speed is more than 1100 rev/min. The change in the values of the Kurtosis factor for different failure sizes also follows an increasing trend with the increase of failure sizes.

The values of the kurtosis factor of the normal bearing are close to 3. The results in Fig. 19 show that the values of the kurtosis factor of the faulty bearing under different test conditions are more than 4.0; therefore, the detection of bearing defects becomes easier after the application of mixture de-noising.

4.3. The relationship between learning ratios and failure sizes

Fig. 20 shows the relationship between learning ratios and failure sizes for the shaft speed of 1100 rev/min. Experimental results show that the learning ratios follow a decreasing trend as the failure size of the bearing increases. The reason may be that the noise ratio of the bearing vibration signal decreases with increasing shaft speed. The experimental result is the same as the simulation result.

4.4. The relationship between learning ratios and shaft speeds

Fig. 21 shows the relationship between learning ratios and shaft speeds for the medium fault. Experimental results show that the learning ratios follow a decreasing trend as the shaft speed increases, but there is a sudden change at 500 rev/min. The reason may be that the signal at 500 rev/min contains the high amplitudes of low frequencies from the increasing amplitudes of the

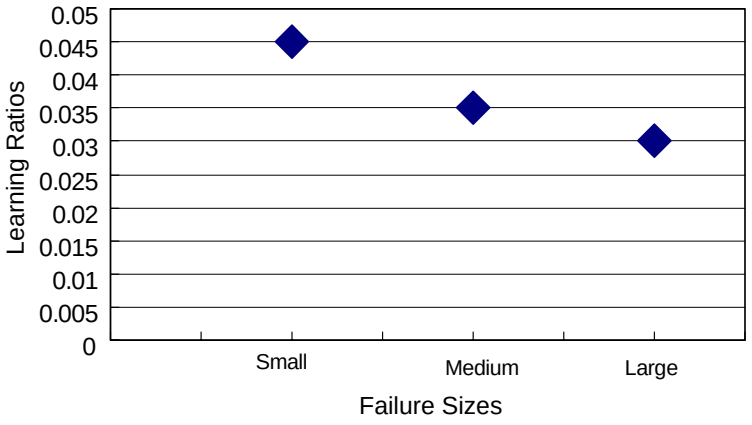

Fig. 20. Relationship between learning ratios and failures.

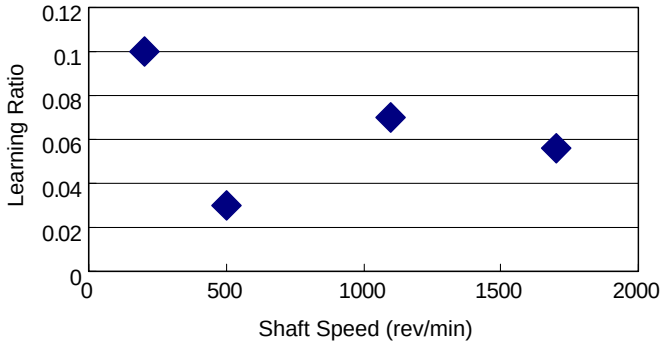

Fig. 21. Relationship between learning ratios and shaft speeds.

vibration of the pulley transmission system, because it is close to the natural frequencies of the transmission system and so the learning ratio must be a small value to trace the change of the signal. The experimental result shows that the learning ratios need to change according to the different experimental conditions.

4.5. Comparison of performance of mixture de-noising and adaptive filtering

The kurtosis factors of the three failure sizes (for failure occurring in the outer race) for the two noise-canceling algorithms are shown in Fig. 22. Generally, the kurtosis factor is 3 when the bearing is running under normal conditions. Fig. 22 shows that the kurtosis factor is greater than 3.5 and over 4.0 for the three sizes of bearing failure, and that the mixture de-noising algorithm is more suitable for identifying bearing faults than is the adaptive filtering. These results show that mixture de-noising is more capable of fault detection when the signal includes heavy noise conditions, but the deference of the kurtosis factor is not big at small sizes compared with

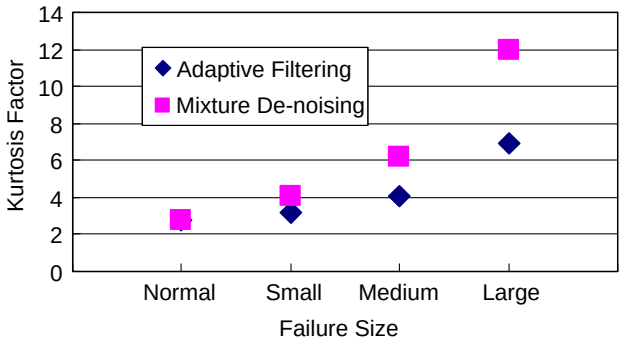

Fig. 22. The kurtosis factors of the three failure sizes for two noise-canceling algorithms.

adaptive filtering, as shown in Fig. 22. The reason may be that the impulse signal is weak at small defect sizes and the background noises from bearing running contain many non-Gaussian signals. So, in order to improve the performance of mixture de-noising, the development of new wavelet de-nosing algorithms for non-Gaussian noise is necessary. This study has begun [14].

5. Conclusions

It has been shown that mixture de-noising is a powerful method for identifying faults in bearings, depending on the speed of the bearing and the size and location of the defect, especially when extracting early fault features of bearings under heavy noise conditions. Experimental results show that the noise-canceling ability of mixture de-noising is stronger than that of adaptive filtering.

The experimental results also show that there is a wide, stable range (0.0001–0.07) of learning rates for mixture de-noising, but the learning ratios change according to different experimental conditions. The selection method of optimum learning for the de-noising problem will be discussed in our next paper.

References

[1] G.K. Chaturvedi, D.W. Thomas, Bearing fault detection using adaptive noise cancelling, *Journal of Mechanical Design* 104 (1982) 280–289.

[2] I.E. Alguindigue, A.L. Buczak, R.E. Uhrig, Monitoring and diagnosis of rolling element bearings using artificial neural networks, *IEEE Transactions on Industrial Electronics* 40 (2) (1993) 209–217.

[3] D. Dyer, R.M. Stewart, Detection of rolling element bearing damage by statistical vibration analysis, *Journal of Mechanical Design* 100 (1978) 229–235.

[4] J. Matthew, R.J. Alfredson, The condition monitoring of rolling element bearings using vibration analysis, *Journal of Vibration, Acoustics, Stress, and Reliability in Design* 106 (1984) 447–453.

[5] Y. Shao, K. Nezu, Detection of self-aligning roller bearing fault using asynchronous adaptive noise cancelling technology, *JSME International Journal, Series C* 42 (1) (1999) 33–43.

[6] Y. Shao, K. Nezu, Extracting symptoms of bearing faults from noise using a non-linear neural filter, *Proceedings of the Institution of Mechanical Engineers Part I, Journal of Systems and Control Engineering* 216 (2002) 169–179.
[7] J. Antoni, R.B. Randall, Optimization of SANC for separating gear and bearing signals, *Proceedings of the 14th International Congress*, Vol. 1, UK, 2001, pp. 89–96.
[8] Y. Shao, K. Nezu, Feature extraction of machinery diagnosis using neural network, *Proceedings of IEEE International Conference on Neural Networks*, Vol. 1, Perth, Australia, 1995, pp. 459–464.
[9] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, Upper Saddle River, NJ, 1996.
[10] D.L. Donoho, De-nosing by soft-thresholding, *IEEE Transactions of Information Theory* 41 (3) (1995) 613–627.
[11] S. Sardy, P. Tseng, A. Bruce, Robust wavelet denosing, *IEEE Transactions on Signal Processing* 49 (6) (2001) 1146–1152.
[12] Y. Shao, H. Kumehara, K. Nezu, Bearing fault diagnostic using wavelet transforms, *Transactions of the Japan Society of Mechanical Engineers, Series C* 69 (687) (2003) 2957–2969.
[13] R.B. Randall, Bearing diagnosis in helicopter gearboxes, *Proceedings of the 14th International Congress*, Vol. 1, UK, 2001, pp. 89–96.
[14] A. Antoniadis, D. Leporini, J.C. Pesquet, Wavelet thresholding for some classes of non–Gaussian noise, *Statistica Neerlandica* 56 (4) (2002) 434–453.