This article was published in an Elsevier journal. The attached copy is furnished to the author for non-commercial research and education use, including for instruction at the author's institution, sharing with colleagues and providing to institution administration.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier’s archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright
A joint resonance frequency estimation and in-band noise reduction method for enhancing the detectability of bearing fault signals

I. Soltani Bozchalooi, Ming Liang*

Department of Mechanical Engineering, University of Ottawa, 770 King Edward Avenue, Ottawa, Ontario, Canada K1N 6N5

Received 17 April 2007; received in revised form 11 October 2007; accepted 13 October 2007

Available online 22 October 2007

Abstract

The vibration signal measured from a bearing contains vital information for the prognostic and health assessment purposes. However, when bearings are installed as part of a complex mechanical system, the measured signal is often heavily clouded by various noises due to the compounded effect of interferences of other machine elements and background noises present in the measuring device. As such, reliable condition monitoring would not be possible without proper de-noising. This is particularly true for incipient bearing faults with very weak signature signals. A new de-noising scheme is proposed in this paper to enhance the vibration signals acquired from faulty bearings. This de-noising scheme features a spectral subtraction to trim down the in-band noise prior to wavelet filtering. The Gabor wavelet is used in the wavelet transform and its parameters, i.e., scale and shape factor are selected in separate steps. The proper scale is found based on a novel resonance estimation algorithm. This algorithm makes use of the information derived from the variable shaft rotational speed though such variation is highly undesirable in fault detection since it complicates the process substantially. The shape factor value is then selected by minimizing a smoothness index. This index is defined as the ratio of the geometric mean to the arithmetic mean of the wavelet coefficient moduli. De-noising results are presented for simulated signals and experimental data acquired from both normal and faulty bearings with defective outer race, inner race, and rolling element.

Keywords: Resonance frequency estimation; De-noising; Wavelet filtering; Wavelet parameter selection; Smoothness index; Spectral subtraction

1. Introduction

As the failure of machinery elements may lead to catastrophic results or unwanted production delays, the monitoring and diagnosis of the health of machinery elements has attracted much attention over the past few decades [1–4]. One of the major approaches in this area of research is based on the analysis of the vibration data measured through the accelerometers mounted on or near the critical mechanical components. Such data
are information-rich. If processed properly, it can indicate not only the existence of certain faults but also the location of the fault based on the frequency characteristics. However, after all, the vibration signal is an indirect source of information and its effectiveness in fault diagnosis largely relies on the availability of proper signal processing techniques. The main problem affecting the performance of the fault diagnosis methods is the corrupting noise. Measured vibration signal is often severely tainted by various noises, e.g., the background noise present in the measurement device as well as the vibrations generated by other mechanical components not of any significance for condition monitoring. An effective de-noising method would be necessary to remove such corrupting noise and interferences.

Two major de-noising approaches namely wavelet threshold-based (also known as decomposition-based) [5,6] and wavelet filter-based [7–10] have been used to purify the vibration signals measured from faulty bearings or gears. The former applies some thresholding rule to select or shrink the wavelet coefficients to be used in the reconstruction process [11,12] and the latter, uses the filtering characteristic of the wavelet transform at a fixed scale [13,14]. The studies reported in [7,8] are in favor of the wavelet filter-based de-noising. However, the performance of this de-noising scheme is greatly affected by the wavelet parameter selection strategy. In this study a combination of wavelet filter-based de-noising and spectral subtraction is used to enhance the vibration signals measured from faulty bearings. We propose a new approach for the selection of the Gabor wavelet parameters namely scale and shape factor as they define the center frequency and the bandwidth of a Gaussian filter, respectively. The first step of the selection process consists of a resonance frequency estimation algorithm. The estimation result leads us to a desirable scale value. This process is based on the fact that the high signal-to-noise ratio (SNR) frequency band of the vibration signal measured from faulty bearings corresponds to the resonance frequency excited by the fault impacts. Shape factor value associated with the selected scale is then found by minimizing a smoothness index (SI). This index is defined as the ratio of the geometric mean to the arithmetic mean of the wavelet coefficient moduli corresponding to the chosen scale.

Though the filtering process, if performed properly, can increase the SNR of the measured vibration, the in-band noise with frequency content in the range covered by the bandpass filter is not eliminated. Consequently, if the strength of such a noise is high, the wavelet filter-based de-noising approach would no longer be able to provide appreciable quality enhancement for the signal. For this reason, we propose to apply spectral subtraction prior to bandpass filtering such that certain in-band noise can be removed. Spectral subtraction has been used to enhance the speech signals [15]. It has also been used in mechanical fault diagnosis by Dron et al. [16] to improve the sensitivity of scalar indicators (e.g., crest factor and kurtosis). Dron et al. used the spectral subtraction technique to remove the corrupting noise including the vibration as well as electrical interferences. To estimate the power spectral density (PSD) of such noises, the short-time Fourier transform and an averaging procedure is applied. In this paper the spectral subtraction technique is used to reduce the intensity of the in-band noise contributed by the noises present in the measurement device due to, e.g., wiring flaws and electrical interferences. In other words, in our work the spectral subtraction technique is tailored to remove the uncorrelated electrical interferences over a narrow frequency band. As explained later, this approach would justify the use of certain assumptions for simplification purpose, such as uniform noise intensity over the frequency band of interest and time invariance. Instead of using the averaging procedure [16], in this study the PSD of the electrical interferences is estimated by minimizing the ratio of the geometric mean to the arithmetic mean of the envelope of the spectral subtracted signal. The partly purified signal obtained through the spectral subtraction is then wavelet transformed to eliminate the interferences due to the vibrations of the other sources such as shaft imbalance and gear meshing. The proposed spectral subtraction method will improve both the quality of de-noising result and the capability of the filter-based de-noising method to enhance the signals with lower SNR.

This paper hereafter is organized as follows. Section 2 provides a brief overview of the wavelet transform and the wavelet filter-based de-noising method. The proposed wavelet parameter selection methods are presented in Section 3. Section 4 details the spectral subtraction technique as a mean to improve the performance of the wavelet filter-based de-noising method. The proposed algorithm is then experimentally evaluated using both a normal bearing and faulty bearings with different defects (damaged outer race, inner race and rolling element) in Section 5. To demonstrate the efficiency of the proposed method, kurtosis values
are also calculated and compared with the associated smoothness indices for each experimental data set in Section 5. Section 6 concludes the paper.

2. Signal enhancement using an adaptive wavelet filter

One major de-noising approach is bandpass filtering. In this method the high SNR frequency band of the signal is passed through the filter and other frequency components are discarded. This is applicable when the frequency content of the corrupting noise mainly resides outside the frequency region where the energy of the signal is concentrated. This phenomenon leads to formation of a high SNR band in the frequency domain. Due to the resonance excited by the fault impacts, such a high SNR band also exists for the vibration signal measured from faulty bearings. By bandpass filtering the measured vibration signal around this frequency band the impulsive features of these signals can be more clearly identified.

One filtering method would be wavelet transforming the signal at a fixed scale. The continuous wavelet transform (CWT) of \( f(t) \) with respect to a mother wavelet \( \psi(t) \) is obtained by

\[
W_f(s,u) = \int_{-\infty}^{+\infty} f(t)\psi^*_{s,u}(t)dt,
\]

where

\[
\psi_{s,u}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-u}{s}\right),
\]

\( s \) and \( u \) are real and asterisk stands for complex conjugate.

A close look at Eq. (1) shows the resemblance of this equation to the convolution process. Consequently, we can write

\[
F_{a}[W_f(s,u)] = F_{a}[f(u)]F_{a}[\psi^*_{s,0}(-u)],
\]

where \( F_{a} \) denotes Fourier transform of a function with respect to variable \( x \). According to Eq. (3) the wavelet transform \( W_f(s,u) \) for a fixed scale can be interpreted as a filtering process using a filter with an impulse response equal to \( \psi_{s,0}(t) \). [14]

Due to its optimum time and frequency resolutions, we choose Gabor wavelet as the mother wavelet. The Gabor wavelet is defined as

\[
\psi(t) = ce^{-\sigma^2 t^2}e^{i2\pi f_0 t}.
\]

Eqs. (3) and (4) lead to

\[
F_{a}[W_f(s,u)] = \sqrt{s}F_{a}[f(u)]\hat{\psi}(sf),
\]

where \( \hat{\psi}(f) \) is the Fourier transform of \( \psi(t) \) and can be written as

\[
\hat{\psi}(f) = ce^{\frac{-\pi^2}{\sigma^2}(f-f_0)^2}.
\]

The last two equations imply that for a constant scale, the Gabor wavelet transform acts as a bandpass filtering process with a Gaussian filter. The bandwidth and center frequency of this filter can be adjusted by changing shape factor \( \sigma \) and scale \( s \), respectively. The selection of these parameters is addressed in the next section.

3. Wavelet parameter selection

3.1. Scale selection

As explained in the previous section, the shape factor value defines the bandwidth of the bandpass filter. Logically this parameter should be adjusted in accordance with the center frequency of the filter. As a result the process is initiated with the selection of the scale value.
It is well known that the frequency band corresponding to the resonance excited by the fault impacts forms the high SNR band of the vibrations measured from faulty bearings. Accordingly, a resonance frequency estimation method should provide us with the proper scale value. This section presents a novel scale selection method. In this method, prior knowledge of the fault generated impacts and the corresponding resonance phenomenon is exploited. Fig. 1(a) shows a portion of the simulated faulty bearing vibrations. Such vibrations form a periodic time signal. Each period contains two ranges: A resonance active range (RAR) and a resonance inactive range (RIR) (Fig. 1(a)). When the shaft speed goes up, the RAR length remains unchanged or generally increases due to the increased impact intensity. This observation suggests that the RAR will comprise an increased portion of a period when the shaft speed increases (Fig. 1(a) and (b)). On the other hand, for \( f(t) = a(t)\cos(\omega_0 t) \), the instantaneous frequency of \( f(t) \) is given by \( \varphi'(t) \) [17]. Considering the fault generated impulse as modeled below

\[
S(t) = Ae^{-\beta t} \cos(\omega_0 t)u(t),
\]

where \( u(t) \) is the units step function, the instantaneous frequency found for the signal interval associated with the RAR length would be equal to the excited resonance frequency \( \omega_0 \). Hence, when the shaft speed increases an instantaneous frequency that corresponds to the largest proportion increase can be identified as the resonance frequency. To implement the above idea, we form the empirical probability density of the instantaneous frequency for two vibration datasets measured from the same bearing at two different rotational frequencies. With an increase in the rotational speed, we expect a larger probability increase for the resonance frequency compared to the other frequencies.

To elaborate, we define \( \varphi^T_m(n) \) as the instantaneous frequency of the vibration signal measured at sampling frequency \( 1/T \), rotational speed \( \omega_m \) and at sampling time \( nT \). We further denote the empirical probability density function of \( \varphi^T_m(n) (n = 1, \ldots, N) \) by \( Pr^T_m(f) \) given as

\[
Pr^T_m(f) = \frac{N_f}{N},
\]

where \( N_f \) is the number of samples with instantaneous frequency of \( \varphi^T_m(n) = f \), and \( N \) is the total number of samples. Following a rise in the shaft rotational speed, the probability density change for each frequency will be

\[
\Delta Pr^T_{mj}(f) = Pr^T_{m+1}(f) - Pr^T_{m}(f) \quad \text{for} \quad \omega_j > \omega_i.
\]
Then the resonance frequency can be identified as
\[ f_{\text{res}} = \arg \max \{ \Delta P_{r,i}^T(f) \}, \]
where \( f_{\text{res}} \) is the estimated resonance frequency.

Though the Hilbert transform is a simple approach for instantaneous frequency calculation [17], it results in a considerable error in discrete time domain and with the presence of noise. As such, we calculate the instantaneous frequencies as follows. In doing so, we denote the mother wavelet as
\[ \psi(t) = g(t)e^{j\omega t}. \]
It has been shown [13] that the wavelet transform of a signal defined as \( f(t) = a(t) \cos \varphi(t) \) can be written as
\[ W_f(s,u) = \sqrt{\frac{s}{2}}a(u)e^{j\omega u}(\hat{g}(s[\xi - \varphi'(u)]) + a(u, \xi)), \]
where \( \hat{g}(f) \) is the Fourier transform of \( g(t) \), \( \xi = \theta/s \), and \( \varphi'(t) \) is the derivative of \( \varphi(t) \). The corrective term \( v(u, \xi) \) is negligible if \( a(t) \) and \( \varphi'(t) \) have small variations over the support of \( \psi_{s,u}(t) \) defined in Eq. (2) and if \( \varphi''(t) \geq \Delta \omega/s \) (\( \Delta \omega \) is the bandwidth of the mother wavelet). The latter is the condition for the wavelet transformed signal to be analytic. With the Gabor wavelet defined in Eq. (4) and assuming a proper shape factor, both of the above conditions would be satisfied. Neglecting the corrective term in Eq. (8) leads to
\[ \left| \frac{W_f(s,u)}{s} \right|^2 = \frac{1}{4} \sigma^2(u)(\hat{g}(s[\xi - \varphi'(u)])^2. \]
On the other hand, from Eq. (6), we have
\[ \hat{g}(f) = c \sqrt{\frac{\pi}{\sigma^2}}e^{-x^2/\sigma^2}. \]
According to Eq. (10), Eq. (9) reaches the maximum when \( \xi = \varphi'(u) \). In other words, it is possible to find the scale associated with the instantaneous frequency at time \( u \) by finding the scale corresponding to the maximum value of Eq. (9). This process assigns a scale value to each sampling point of the measured vibration. By applying this method to two vibration data sets measured at two different rotational frequencies and finding the empirical probability density for two sets of associated scales, the proper scale value can be selected by
\[ S_p = \arg \max \{ \Delta P_{r,i}^T(S) \}. \]
Fig. 2 shows a flowchart of the proposed de-noising algorithm.

3.2. Shape factor selection

Following the scale selection, shape factor value should also be adjusted so that the fault generated impulses can be clearly identified from the de-noising result. In this section the level of impulsiveness of the bandpass filtered signal is used as a criterion to find the proper shape factor value. An SI, defined as the ratio of the geometric mean to the arithmetic mean of the wavelet coefficient moduli associated with the selected scale, is used to quantify the impulsiveness of the bandpass filtered signal. This is elaborated in the following.

3.2.1. Wavelet coefficient moduli

Eq. (6) indicates that the filtered signal given by the wavelet transform at a fixed scale will be analytic if \( (\sigma f)^2/\sigma^2 \gg 1 \) [13]. Therefore, the modulus of this analytic result would provide the envelope of the bandpass filtered signal [10,13]. Denoting the Gabor wavelet transform of a signal \( V(t) \) at scale \( s \) and shape factor \( \sigma \) by \( W_{V,s}^{\sigma}(u) \), the envelope of the bandpass filtered \( V(t) \) is obtained by
\[ |W_{V,s}^{\sigma}(u)| = \sqrt{\text{Re}(W_{V,s}^{\sigma}(u))^2 + \text{Im}(W_{V,s}^{\sigma}(u))^2}. \]
Since the bandpass filtering is usually treated as a preprocessing step for envelope spectrum analysis [2,4], we can use the characterizing nature of this envelope to adjust the shape factor or the bandwidth of the corresponding bandpass filter.

### 3.2.2. Characterizing the envelope

Though smaller filter bandwidth may lead to a signal with higher SNR, it should be noted that narrowing the filter may result in excessive loss of time resolution for the daughter Gabor wavelet. This should be avoided as it might smear the major featuring characteristic of the faulty bearing vibrations, i.e., the consecutive impulses. On the contrary, for a filter wider than necessary low SNR frequency components or noise content of the signal would also pass through the filter and no de-noising could take place. As a result, a good trade off between the noise elimination and feature preservation should be determined by adjusting the shape factor. As explained later, a reasonable criterion for this adjustment task is the smoothness of the

---

**Fig. 2. Flowchart of the proposed de-noising method.**

1. Calculate the PSD of $V(t)$ denoted by $|\hat{W}(\omega)|^2$
2. Subtract pre-selected values from $|\hat{W}(\omega)|^2$ and reconstruct the signal using new PSD and $\angle V(\omega)$
3. Find the envelope signal using Hilbert transform for the signals reconstructed in previous step
4. Find the time signal corresponding to the minimum smoothness index calculated for the envelope signal found in previous step
5. Wavelet transform the signal found in previous step using the selected scale $S_p$ and for a range of values chosen for the shape factor
6. Find the real part of the wavelet coefficients corresponding to the minimum smoothness index calculated for the wavelet coefficient moduli
7. De-noised Signal

---

Since the bandpass filtering is usually treated as a preprocessing step for envelope spectrum analysis [2,4], we can use the characterizing nature of this envelope to adjust the shape factor or the bandwidth of the corresponding bandpass filter.
envelope of the bandpass filtered result. A robust index quantifying such a characteristic is the ratio of the geometric mean to the arithmetic mean. This index has been used as a measure of spectral flatness in speech signal processing [18]. The geometric mean of a series is defined as

$$G_s = \left( \prod_{n=1}^{N} S(n) \right)^{1/N}$$

for a positive time series \(S(n)\) \((n = 1, 2, \ldots N)\).

Similarly the arithmetic mean of the series is

$$A_s = \frac{1}{N} \sum_{n=1}^{N} S(n).$$

We denote the ratio of the two as \(r_{G/A}\), i.e.,

$$r_{G/A} = \frac{\left( \prod_{n=1}^{N} S(n) \right)^{1/N}}{(1/N) \sum_{n=1}^{N} S(n)}.$$  \hspace{1cm} (13)

Since \(A_s \geq G_s\) [19] (equality holds only when \(S(n) = S(m)\) for all \(m\) and \(n\)), \(r_{G/A}\) is always in the range of \([0,1]\) for a positive time series. A useful property of \(r_{G/A}\) is that it approaches unity for smooth time series and drops to zero for a highly impulsive series.

When the shape factor value is properly adjusted and consequently, the fault generated impulses can be clearly detected, we expect the envelope of the bandpass filtered signal to form a more impulsive time series. Replacing \(S(n)\) in Eq. (13) by the expression given in Eq. (12), \(Z(s)\) can be written as

$$Z(s) = \frac{\left( \prod_{n=1}^{N} W_{Y^{S,p}}(nT) \right)^{1/N}}{(1/N) \sum_{n=1}^{N} W_{Y^{S,p}}(nT)}.$$ \hspace{1cm} (14)

where \(S_p\) is the selected scale, \(T\) the sampling period and \(N\) the number of samples. Accordingly, the proper shape factor value can be found by minimizing \(\eta(\sigma)\).

3.2.3. Measure of impulsiveness: Kurtosis vs. smoothness index

Kurtosis is a measurement of non-Gaussianity [20] and is often used as an indicator of impulsiveness [21]. It is defined as follows:

$$Kurt(Y) = \frac{E((Y - E(Y))^4)}{E((Y - E(Y))^2)^2},$$

where \(Y\) is a random variable.

In the context of machine fault detection, a high kurtosis value is treated as a sign of the presence of faults in a rotating mechanical system. As the smoothness index is proposed for the same purpose, it would be useful to compare the performance of the two indices. In this paper, three typical cases are compared, i.e., (i) white Gaussian noise shown in Fig. 3(a), (ii) white Gaussian noise plus a single outlier (Fig. 3(b)), and (iii) pure impulsive signals at different shaft rotating speeds (Figs. 3(c) and (d)).

As shown in Table 1, the kurtosis value is not always in line with the signal impulsiveness whereas the smoothness index is in good agreement with such a characteristic. The addition of a single outlier causes a large kurtosis surge but only a minor drop of the SI value. As expected, for the pure impulsive signal, the smoothness index reacts correctly leading to, e.g., a quick drop to 0.052 when the simulated shaft speed is set at 20 Hz. Although the kurtosis is also high (25.74) for the same impulsive signal, it is considerably lower than the value obtained for the simulated white Gaussian noise with a single outlier (43.60). Therefore, a high kurtosis level could be merely a result of a random event such as a knock on the machine and may not necessarily mean a high level of impulsivity, the concept adopted for rotating machinery fault detection. Therefore, the tendency of overreaction to a few outliers of random nature would undermine the effectiveness of kurtosis as an impulsiveness measure in the context of rotational machinery condition monitoring.
It is also observed that the kurtosis value is considerably affected by shaft rotational frequency. As shown in Table 1, when the shaft speed increases from 20 to 30 Hz, the kurtosis value exhibits a sharp plunge from 25.74 to 16.89 whereas the SI value has only a slight change from 0.052 to 0.070. Furthermore, unlike the SI which clearly signals a high level of impulsiveness when approaching zero, there seems to be no commonly accepted kurtosis benchmark to be used to distinguish the level of impulsivity. Hence, the dependence of kurtosis on

<table>
<thead>
<tr>
<th>Impulsiveness measure</th>
<th>(i) White Gaussian noise</th>
<th>(ii) White Gaussian noise with a single outlier</th>
<th>(iii) Pure impulsive signal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simulated at 20 Hz shaft rotational speed</td>
<td>Simulated at 30 Hz shaft rotational speed</td>
<td></td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0.245</td>
<td>43.60</td>
<td>25.74</td>
</tr>
<tr>
<td>Smoothness index</td>
<td>0.8426</td>
<td>0.8396</td>
<td>0.052</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.070</td>
</tr>
</tbody>
</table>
rotational speed and the lack of a meaningful benchmark, along with its over-vigilance to a few outliers, would make it difficult to interpret the implication of a certain kurtosis value obtained from a certain test.

3.3. Simulation results

Figs. 4(a) and (b) show the vibrations of a faulty bearing simulated at two different rotational frequencies. A considerable amount of noise is added to both simulated data sets. The signal and noise mixtures (shown in Figs. 5(a) and (b)) are fed to the resonance frequency estimation algorithm to select the scale value. To illustrate, the proposed wavelet filter-based de-noising method is evaluated using the data mixture shown in Fig. 5(a). Following scale selection, the SI found according to Eq. (14) is minimized by using the selected scale $S_p$ given by Eq. (11) and searching the squared shape factor values considered over the range $0.01 \leq \sigma^2 \leq 0.2$ with the step of 0.01. The resulting de-noised signal is displayed in Fig. 6(a). The scale value is selected as 15.9468. As expected this value corresponds to the frequency of 3010 Hz which is very close to the resonance frequency of the simulated signal chosen as 3000 Hz. Fig. 6(b) shows the SI versus squared shape factor values.

Fig. 4. Faulty bearing vibration signals (fault characteristic frequency $= 3.5 \times$ rotational speed in (Hz) and resonance frequency of 3000 Hz corresponding to scale 16, simulated at (a) 780 rpm (13 Hz), and (b) 1020 rpm (17 Hz).

Fig. 5. (a) Simulated signal shown in Fig. 4(a) with noise added and (b) simulated signal shown in Fig. 4(b) with noise added. SNR of the resulting noisy signal is $-10$ in both cases.
The minimum SI (0.6493) is found at $\sigma^2 = 0.06$. Though the simulated fault impulses can be clearly identified from the de-noising result, some spurious artifacts are also formed in between the consecutive impulses. This is mainly caused by the in-band noise that the wavelet filter-based de-noising method is unable to eliminate. This issue is addressed in the following section.

4. Spectral subtraction

In the previous section an adaptive filtering approach was proposed. In this method the bandwidth and the center frequency of a bandpass filter were adjusted so that the high SNR band of the measured signal can pass through the filter and the rest of the frequency components dominated by noise and interferences be eliminated. However, through this method the noise components contained in this frequency band would not be removed. Therefore, the quality of the de-noising result may still be affected by such noises. The level of such influence depends on the intensity of the in-band noise. For vibration signals of very low SNR, especially when the corrupting noise is diffused across all frequencies, bandpass filtering alone would not bring about a substantial enhancement in the quality of the vibration signal. To offset such a drawback of the wavelet filter-based de-noising method, spectral subtraction can be used prior to wavelet transform. Spectral subtraction has been widely used in speech signal processing [15]. In this method the PSD of the pure signal is estimated. This estimate is then used to reconstruct the enhanced signal. To elaborate, consider:

$$Z(t) = S(t) + N(t),$$

where $Z(t)$ is the signal and noise mixture, $S(t)$ is the signal and $N(t)$ is the corrupting noise. Accordingly, we have

$$|\hat{Z}(\omega)|^2 = |\hat{S}(\omega)|^2 + |\hat{N}(\omega)|^2 + \hat{S}(\omega)\hat{N}^*(\omega) + \hat{S}^*(\omega)\hat{N}(\omega).$$ (15)

To find an estimate of the PSD of the signal $|\hat{S}(\omega)|^2$, an approximation of the last three components in the above equation is needed. In speech signal processing, noise is uncorrelated with the speech. Consequently, the last two terms in Eq. (15) can be ignored. Furthermore, a good estimate of $|\hat{N}(\omega)|^2$ can be found by analyzing the intervals where speech is not present [15]. However, none of these is applicable to the vibration signal measured from faulty bearings.

In the context of machinery fault detection, the corrupting noise also contains the vibration from other sources (e.g. shaft imbalance, gear meshing, etc.) which is not necessarily uncorrelated with bearing vibration due to mechanical engagements. However, it should be noted that the purpose of spectral subtraction here is to combat the in-band noise. As explained before the frequency band that would pass through the filter...
corresponds to the resonance excited by the fault impacts. Such impacts usually excite a resonance in the system at much higher frequency than the vibration generated by other machine elements [2]. Therefore, it is reasonable to assume that the major part of the in-band noise consists of the background noise present in the measurement device and mainly caused by the wiring flaws and electrical interferences. To elaborate, we consider:

\[ V(t) = v(t) + N_1(t) + N_2(t), \]  

where \( V(t) \) is the measured vibration, \( v(t) \) is faulty bearing vibration, \( N_1(t) \) is the measured noise caused by other vibration (e.g., shaft imbalance, gear meshing, etc.) and \( N_2(t) \) is the background noise due to wiring flaws and electrical interferences, uncorrelated with \( v(t) \) and \( N_1(t) \). From (16), we have

\[ |\hat{V}(\omega)|^2 = |\hat{v}(\omega) + \hat{N}_1(\omega)|^2 + |\hat{N}_2(\omega)|^2. \]

According to this equation, with an estimate of \( |\hat{N}_2(\omega)|^2 \), \( |\hat{G}(\omega)|^2 = |\hat{v}(\omega) + \hat{N}_1(\omega)|^2 \) can be estimated by

\[ |\hat{G}(\omega)|^2 = |\hat{V}(\omega)|^2 - |\hat{N}_2(\omega)|^2, \]  

where \( \hat{N}_2(\omega) \) and \( \hat{G}(\omega) \) are estimates of \( \tilde{N}_2(\omega) \) and \( \tilde{G}(\omega) \), respectively. Given \( |\hat{G}(\omega)| \), an estimate of \( |\hat{v}(\omega) + \hat{N}_1(\omega)| \), to estimate \( v(t) + N_1(t) \) the phase information is also required. We assume that the phase information is relatively unimportant so that we can approximate \( \angle(\hat{v}(\omega) + \hat{N}_1(\omega)) \) by \( \angle(\hat{V}(\omega)) \).

In Eq. (17), \( |\hat{N}_2(\omega)|^2 \) could be estimated by analyzing the output of the sensors when the mechanical system is idle. This approach however may not always be reliable since the nature of such noises may vary over time, due to the changing working conditions. Accordingly, an on-line estimation process is required to provide the algorithm with the timely estimate of \( |\hat{N}_2(\omega)|^2 \).

As the spectral subtraction is concerned with the in-band noise which would usually correspond to a narrow frequency band, it is reasonable to assume that the energy of the background noise \( N_2(t) \) is uniformly distributed over such a narrow band. Consequently, we can further assume \( N_2(t) \) as white. This assumption will simplify the analysis since the PSD of such noise is flat over the frequency domain and can be represented with a single value for all frequencies. As a result, subtraction of this constant value from the PSD of the measured signal would provide us with an estimate of \( \tilde{G}(\omega) \). Furthermore, we can apply the smoothness minimization criterion introduced earlier to find the proper subtraction value. Although this approach may also reduce the intensity of frequency components other than the white noise, it helps the hidden impulses to manifest in the final de-noised signal, which is the main goal of this spectral subtraction algorithm.

A discretized range of values \([N_{\text{min}}, N_{\text{max}}]\) are considered for \( |\hat{N}_2(\omega)|^2 \). The value on this range which minimizes the SI calculated for the envelope of the spectral subtracted signal is chosen as the best estimate for \( |\hat{N}_2(\omega)|^2 \). As Eq. (17) also yields negative values, it is modified as follows to ensure a non-negative result:

\[ |\hat{G}(\omega)|^2 = \begin{cases} |\hat{V}(\omega)|^2 - |\hat{N}_2(\omega)|^2 & \text{for } |\hat{V}(\omega)|^2 > |\hat{N}_2(\omega)|^2, \\ 0 & \text{otherwise.} \end{cases} \]  

Finally \( v(t) + N_1(t) \) is estimated using

\[ \tilde{G}(\omega) = |\hat{G}(\omega)| \exp[i \angle \hat{V}(\omega)] \quad \text{and} \quad \tilde{G}(t) = F^{-1}[\tilde{G}(\omega)]. \]  

To examine the effect of approximating \( \angle(\hat{v}(\omega) + \hat{N}_1(\omega)) \) by \( \angle(\hat{V}(\omega)) \), the noisy signal shown in Fig. 5(a) is spectral subtracted using the proposed method. The signal is then reconstructed by applying the above phase approximation approach and also using the phase information of the original simulated signal without noise (Fig. 4(a)). The de-noising results are shown in Figs. 7(a) and (b). As one can see the difference is unnoticeable. This indicates that the spectral subtraction result is reasonably insensitive to the error in estimating the phase information. As the phase information would not be available in reality, the approximated phase information can be applied without causing noticeable problem in de-noising. Its worth mentioning that the results presented in Figs. 7(a) and (b) are obtained by the spectral subtraction without using wavelet filtering. A comparison between Figs. 7(a) and 6(a) shows superior de-noising performance of the proposed spectral subtraction method in comparison to the wavelet filter-based de-noising. However, it should be noted that in this case the corrupting noise only contains the additive white Gaussian noise. In the presence of interferences
from other machinery components, wavelet filter-based de-noising is also required as spectral subtraction alone would not be able to de-noise the signal. Hence, the combination of wavelet filtering and spectral subtraction provides a robust de-noising technique.

Following the spectral subtraction step, the reconstructed time signal given by Eq. (19) is wavelet transformed using the scale and shape factor values found through the wavelet parameter selection algorithms explained in previous sections to eliminate $N_1(t)$.

5. Experimental evaluation

The proposed method is evaluated by de-noising signals of four bearings associated with four different conditions: a normal bearing (NB), a bearing with outer race fault (OFB), a bearing with inner race fault...
(IFB), and a bearing with a faulty ball (BFB). The NB, OFB and BFB experimental data are acquired in our lab. To test the robustness of the method, the OFB experiment is set up differently from the NB and BFB cases and for IFB, the data is adopted from [22] with permission. The experiments and de-noising results are presented below.

Fig. 9. (a) Vibration signal measured at 630 rpm rotational speed and the frequency domain representation. (b) De-noising result using spectral subtraction followed by wavelet filter de-noising and its frequency domain representation. (c) De-noising result obtained using wavelet filtering alone and its frequency domain representation.
5.1. Experiment and result for the bearing with an outer race fault

The experiment is carried out using a spectraquest machinery fault simulator (MFK-PK5 M) as shown in Fig. 8. Two well balanced mass rotors (2\textquoteright thick, 4\textquoteright in diameter and 11.11 lbs each) are installed on a 5/8\textquoteright steel shaft and supported by two bearings of type ER10 K (inside, outside, pitch and ball diameters are 0.6250\textquoteright, 1.8500\textquoteright, 1.3190\textquoteright and 0.3125\textquoteright, respectively) with eight rolling elements (balls). The simulator is powered by a 3-hp AC motor which is controlled by a Hitachi drive (SJ200-022NFU). The shaft speed was set at 630 rpm (10.5 Hz). The left bearing has a pre-seeded fault (it was created by the manufacturer and the dimensions of the fault are unknown) on the outer race with a characteristic frequency of 32 Hz ($f_r = \frac{3.052}{10.5}$, specified in the simulator user’s manual). An accelerometer (Montronix VS100-100) with 100 mV/g sensitivity and 1–12 kHz sensitivity range is used to collect the vibration signal. The signal is fed to an NI AT-MIO-16DE-10 DAQ card and then collected through LabVIEW. The signal processing is done using MATLAB on a Pentium® 4/2.52 GHz PC.

To create additional vibration interference, a gearbox is also connected to the driving shaft using a belt as shown in Fig. 8. The meshing frequency of the gearbox, considering the effect of belt connection, is $7.03 \times$ shaft rotational frequency (Hz). For this test, the meshing frequency is 73.8 Hz ($f_m = 7.03 \times 10.5$ Hz). The accelerometer is installed on the simulator base at a location away from the faulty bearing but closer to the belt and the driven gearbox (Fig. 8). The vibration signal is sampled at 20000 samples/s. Part of the raw data and the corresponding frequency domain representation are plotted in Fig. 9(a). The signal is de-noised using the proposed joint wavelet filtering and spectral subtraction method. To select the scale using the resonance frequency estimation algorithm, an additional vibration signal is acquired at 780 rpm (13 Hz) rotational speed. Following the spectral subtraction, SI is minimized using the selected scale and for a range of values chosen for $\sigma^2$.

The proper ($s-\sigma^2$) combination is found as (50, 0.32) and the minimum SI is 0.7398. The de-noising result associated with this ($s-\sigma^2$) combination and the corresponding frequency domain representation are displayed in Fig. 9(b). As shown in the figure the time interval between two consecutive impulses is about 0.0313 s which precisely reflects the fault characteristic frequency of 32 Hz. As shown in the frequency domain representation of the de-noised signal, weak traces of 4th and 5th harmonics (295 and 369 Hz, respectively) of the gearbox meshing frequency can still be noticed. This is because such interference frequencies (295 and 369 Hz) are very close to the frequency band of interest. Similar phenomena can be observed in the other tests and will not be
discussed hereafter. It is however worth mentioning that when the interference frequencies are too close to the frequency components of interest as shown in Fig. 9(b), the de-noising tasks become more challenging. Fig. 9(c) shows the de-noising result with spectral subtraction step eliminated from the algorithm. Though impulsive features can still be detected from the result, comparison of Figs. 9(b) and (c) indicate that a better de-noising performance can be achieved by adding the spectral subtraction component to the bandpass filtering algorithm.

5.2. Experiment and result for ball fault case

In this test, one of the rotors is removed and a bearing (also type ER10 K) with a faulty ball is installed on the belt-end of the shaft. The sensor (accelerometer) is installed on the structure as shown in Fig. 10. The shaft rotational frequency is set at 930 rpm (15.5 Hz). A portion of the measured raw data and the associated frequency domain representation are shown in Fig. 11(a). Following the spectral subtraction, the reconstructed signal is bandpass filtered using the proposed method. To select the scale value an additional vibration data set is acquired at 1054 rpm (17.5 Hz). The smoothness index is then minimized to find the proper shape factor. The de-noising result and its frequency domain representation are shown in Fig. 11(b). In the de-noised result, the fault generated impulses can be clearly identified. The time interval between two
consecutive impulses in Fig. 11(b) is about 0.0323 s which is obviously associated with the fault characteristic frequency of 30.9 Hz.

5.3. Experiment and result for normal bearing

For this test, the faulty bearing is replaced with a normal bearing using the same setup as that in the BFB case and the de-noising algorithm is applied to the acquired vibration data. The vibration data are measured at 11.1 and 12.8 Hz shaft rotational speed to estimate the excited resonance frequency. In the wavelet filtering step the smoothness index drops continuously until reaching a point where the condition for analytic filtered signal \( (\sigma_f^2/\sigma^2 > 1) \) is no longer valid while the final SI is slightly lower than that obtained for the original measured signal. Figs. 12(a) and (b) show the measured signal and the de-noising result, respectively. This signal mainly reflects the gearbox vibrations and as shown in the frequency domain, the peaks correspond to the 2nd (180 Hz), 3rd (270 Hz) and 4th (360 Hz) harmonics of the gearbox meshing frequency.

5.4. Result for a bearing with an inner race fault

To further examine the performance of the method, it is applied to the IFB case using the data acquired from a different type of bearing in a different test condition. The data is downloaded from the Case Western

![Graphs](image-url)
Reserve University bearing data center [22]. A 6203-2RS JEM SKF deep groove ball bearing is used in this case. It contains nine balls and its inside, outside, ball and pitch diameters are 0.6693", 1.5748", 0.2656" and 1.122", respectively. The inner fault size is 0.021" in diameter and 0.011" deep. The inner race fault frequency is 148 Hz ( = 4.9469 × rotational frequency, according to [22]). The shaft rotational frequency is 29.93 Hz (1796 rpm). The vibration data is measured using an accelerometer mounted on the motor supporting base plate of a 2-hp electric motor. The sampling frequency is 12000 Hz. The faulty bearing is installed on the fan end. Further details of the experimental setup can be found in [22]. An additional vibration data set measured from the same bearing rotating at 28.86 Hz (1732 rpm) is used to estimate the excited resonance frequency. The raw data and the de-noised signal are plotted in Figs. 13(a) and (b), respectively. As shown in Fig. 13(b), the period of the impulse is 0.0067 s and corresponds to the inner race fault frequency (148 Hz).

5.5. Discussion: Responses of kurtosis and smoothness index to impulsiveness changes

The kurtosis and the SI values calculated for all the above tests (OFB, BFB, NB and IFB) are summarized in Table 2. As shown in the table, the kurtosis and the SI values for the normal bearing are very close to the expected values, i.e., 3 and 1, respectively, both before and after de-noising. The data in Table 2 also indicate that the SI responds to the impulsiveness changes very well. Following the spectral subtraction and wavelet filtering, the SI is reduced for all three faulty bearing cases, which is anticipated. The improvement in...
impulsiveness can be easily observed by comparing parts (a) and (b) of Figs. 9–13 for OFB, BFB, NB and IFB cases, respectively. However, this fact has not been reflected by the associated kurtosis values. Intuitively, the kurtosis value should increase with the improved impulsiveness but the opposite is observed: The actual kurtosis values have dropped after the impulsivity increase for all the four cases (from 3.60 to 1.60, 8.42 to 3.30, 8.14 to 5.97, and 3.17 to 2.86 for OFB, BFB, IFB and NB, respectively). The above observations indicate that the proposed smoothness index is a reliable measure of impulsiveness of the bearing signals while kurtosis does not seem to be as effective for the same purpose.

A systematic comparison between kurtosis and smoothness index in characterizing impulsive signals of different nature could be an interesting research topic but is beyond the scope of this paper.

6. Conclusion

In this paper, a new de-noising scheme was proposed to enhance the vibration signals measured from faulty bearings. The algorithm involves a spectral subtraction step prior to wavelet filtering. The main purpose of spectral subtraction is to eliminate or at least to reduce the in-band noise with frequency content that is the same as the frequency band covered by the daughter Gabor wavelet. This step is based on the assumption of white uncorrelated in-band noise. The power spectral density of such corrupting noise is estimated by minimizing a smoothness index defined as the ratio of the geometric mean to the arithmetic mean of the envelope of the spectral subtracted signal. The resulting time signal is then wavelet transformed to eliminate the interfering vibration signal resulting from other sources, e.g., shaft imbalance, gear meshing, etc. To find the desirable scale value, a resonance frequency estimation algorithm is developed. Corresponding to the selected scale, a proper shape factor is then determined in a separate step by minimizing the smoothness index (SI) calculated for the wavelet coefficient moduli. The proposed method has been applied to de-noise signals for bearings of different conditions (normal, faulty inner and outer races, and faulty rolling element). The results have shown that the proposed scale and shape factor selection algorithms are very effective in enhancing the detectability of fault signals. It also illustrated that the performance of the wavelet filter-based de-noising method can be improved considerably by preprocessing the signal using spectral subtraction.

Acknowledgments

We wish to thank Professor K.A. Loparo of Case Western Reserve University for his kind permission to use the IFB data. The major part of this study was supported by Natural Science and Engineering Research Council of Canada. This work was also partly supported recently by Ontario Centers of Excellence in partnership with GasTOPS, Ltd. Their support is greatly appreciated.

References


Iman Soltani Bozchalooi is a research associate in the Department of Mechanical Engineering at the University of Ottawa, Canada. He received his M.A.Sc. degree in Mechanical Engineering from the same department in 2007 and his B.A.Sc. degree in Mechanical Engineering from Amir-Kabir University of Technology (Tehran Polytechnic), Iran in 2003. His research interests are in machinery fault diagnosis and signal processing.

Ming Liang is a professor of Mechanical Engineering at the University of Ottawa, Canada since 2000. Dr. Liang worked as an associate professor (1996–2000) and assistant professor (1991–1996) in the same department. He received his Ph.D. in Industrial Engineering from University of Windsor, Canada in 1991 and his M.Sc. and B.Sc. from Northeastern University, China in 1984 and 1982, respectively. His research interests include machine and process fault detection and diagnosis, machining control and monitoring, and manufacturing planning. Dr. Liang is an Editorial Board Member of the IIE Transactions and a Registered Professional Engineer in the Province of Ontario, Canada.