uOttawa

L'Université canadienne
Canada's university

**FACULTÉ DES ÉTUDES SUPÉRIEURES
ET POSTOCTORALES**

uOttawa

L'Université canadienne
Canada's university

**FACULTY OF GRADUATE AND
POSDOCTORAL STUDIES**

Iman Soltani Bozchalooi
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.A.Sc. (Mechanical Engineering)
GRADE / DEGREE

Department of Mechanical Engineering
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Bearing Vibration and Oil Debris Signal Enhancement for Machinery Condition Monitoring

TITRE DE LA THÈSE / TITLE OF THESIS

Dr. M. Liang
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Dr. X. Huang

Dr. D. Necsulescu

Gary W. Slater
Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Bearing Vibration and Oil Debris Signal Enhancement for Machinery Condition Monitoring

Iman Soltani Bozchalooi

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements for the degree of

Master of Applied Science
in
Mechanical Engineering

Ottawa-Carleton Institute for Mechanical and Aerospace Engineering
University of Ottawa
Ottawa, Canada

April 2007

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-32479-0
Our file Notre référence
ISBN: 978-0-494-32479-0

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Abstract

Vibration signal and lubricant oil condition are two major sources of information for machine health condition monitoring. Though vibration signal is an indirect indicator of machine conditions, it contains very rich information. On the other hand, the lubricating oil analysis provides a direct indicator of machine health conditions. The joint use of the two sources of information would compensate for their limitations and thus better maintenance actions can be expected. However, this alone is not sufficient since the two sources are often severely contaminated by background and machine interference noises. Using such contaminated data without careful de-noising will inevitably cause misleading maintenance decisions and hence premature machine failure as well as lost productivity. As such, this thesis addresses the de-noising issues for both vibration and oil condition signals. Due to different natures of the vibration signals and signals measured through oil debris monitoring sensors, different approaches will be developed in this study for the enhancement of the two types of signals. In de-noising vibration signals, this research focuses on bearings since they are one of the most vulnerable and frequently used components in rotating machinery. The results obtained based on bearings could be applied to other rotating machine components with some modifications.

Wavelet transform, in particular the Gabor wavelet transform, has been used for de-noising impulsive signals measured from faulty bearings. However, it has been a challenging task to select proper wavelet parameters. This work introduces a method to guide the selection process by a smoothness index (SI). The SI is defined as the ratio of the geometric mean to the arithmetic mean of the wavelet coefficient moduli of the vibration signal. For the signal contaminated by Gaussian white noise, we have shown that the modulus of the wavelet coefficients follows Rician distribution. Based on this observation, we then prove that the SI converges to a constant number (0.8455...) in the absence of mechanical faults or for very low signal to noise ratio. This result provides a dimensionless SI upper bound corresponding to the most undesirable case. We have also shown that the SI value decreases in the presence of impulses with properly selected parameters.

However, this approach is based on the assumption that the most impulsive components of the measured vibration are due to the faults. This assumption may not be valid in general. On the other hand, the proposed method requires a global search for the minimum SI for all combinations of wavelet parameters in the chosen discretized ranges which is a computationally demanding task. In addition, through bandpass filtering the signal, the in-band noise with frequency content in the range covered by the daughter wavelet is not eliminated. As a result, the performance of the wavelet filter based de-noising method deteriorates as the background noise intensity increases.

To mitigate the above difficulties, a novel scale selection method is proposed. In this approach we incorporated our knowledge of the resonance frequency excitation phenomenon in the scale selection algorithm. Furthermore, to improve the efficiency of the method, spectral subtraction is applied prior to wavelet transform. The proposed spectral subtraction method leads to improvements in both the final result of the process and the capability of the wavelet filter based de-noising method for lower SNR vibration signals. The proposed joint spectral subtraction and wavelet de-noising method has been successfully tested using experimental data.

For the oil condition signals, the main issue is that the oil debris sensor is not only sensitive to the metal debris or particles but the structural vibrations as well. The weak signals of small particles are often concealed in the vibration signals. This either causes false alarm (since the shape of a particle signal resembles that of a vibration signal in certain ways) or leaves existent machine faults undetected. Adaptive Line Enhancement technique is proposed to remove such interferences. The method has been tested on both simulated and experimental data.

Acknowledgments

I firstly, would like to express my deep gratitude to my mother whose sacrificing love and encouragements made this work possible. I dedicate this work to my mother Prof. Mahvash Oskoui and to the memory of my father Abbas Soltani.

I feel grateful to my advisor Prof. Ming Liang for his moral as well as financial support and also his invaluable guidance, encouragement and friendship. I have benefited so much from his discussions on every detail of this research, and he has been always willing to make time for me no matter how busy he was. He has been a real mentor and a true friend.

I would like to extend special thanks to those at GasTOPs Ltd for providing the MetalSCAN sensor, Leo Denner of University of Ottawa for his helps on the experimental work, John Perrins, James MacDermid, Doug Taylor for making the parts for the fault simulator and Professor K. A. Loparo for his kind permission for use of their bearing vibration data.

I also wish to thank my lab colleagues and friends Raymond Wang and Hoonbin Hong for all of their support and also Dr. Xianfeng Fan for the experiments on the MetalSCAN sensor.

Finally, I would like to thank my dearest Reyhaneh for her unlimited love and support during the course of this study.

Table of Contents

List of Figures

Nomenclature

$A_s\ (G_s)$	Arithmetic (geometric) mean of a series
A	Amplitude of the simulated impulse
$a(t)$	Local signal to noise ratio (SNR), which is equal to $\frac{\mu(t)}{\nu}$
C_w	Set of all wavelet coefficients
D	A dictionary of orthonormal basis
$Ei(x)$	Exponential integral
$En(X)$	Entropy of a discrete value random variable X
$F_Z(z)$	Cumulative probability distribution of random variable Z
${}_1F_1$	Confluent hypergeometric function
$F_x(\cdot)$	Fourier transform of a function with respect to variable x
$\hat{F}(\omega)$	Fourier transform of $F(t)$
$\tilde{F}$	An estimate of F
$F^{-1}(\ )$	Inverse Fourier transform
$\angle(F)$	Phase of the imaginary value F
$f_{X,Y}(x,y)$	Joint probability density of random variables X and Y
f_0	Modulation frequency of the mother Gabor Wavelet
f_c	Cage rotational frequency
f_{ir}	Inner race rotational frequency
f_{or}	Outer race rotational frequency
f_b	Ball rotational frequency
f_{fo}	Outer race fault characteristic frequency
f_{fi}	Inner race fault characteristic frequency
f_{res}	Estimated resonance frequency
$\langle f, g \rangle$	Inner product of two functions f and g
$G(t)$	Vibration of the faulty bearing plus the vibration of other components
g_m	The m_{th} vector of the orthonormal basis $\mathrm{B} = \{g_m\}_{m \in N}$
I_0	Modified Bessel function of first kind and zeroth order
N_a	Total number of samples in a measured data set
N_f	Number of samples with instantaneous frequency of f

$N_1(t)$	Corrupting noise due to sources of vibration other than faulty bearing
$N_2(t)$	White Gaussian background noise present in all measuring devices
$\Pr_m^s(f)$	Empirical probability density of the instantaneous frequency f for the signal measured at sampling frequency $\frac{1}{T_s}$ and rotational speed ω_m
Re (Im)	Real (imaginary) part of a complex value
$r_{G/A}$	Ratio of the geometric mean to the arithmetic mean
s	Scale
t	Time index
u	Wavelet translation index
$u(t)$	Unit step function
$V(t)$	Measured vibration
$v(t)$	Vibration signature of a faulty bearing in the absence of noise
$W_f(s,u)$	Wavelet transform of the function $f(t)$ associated with scale s and translation index u
$W_f^{s,\sigma}(u)$	Gabor wavelet transform of a signal $f(t)$ associated with scale s and shape factor σ
W	Filter weight vector
x_k	k_{th} sample of data set X
β	Structural damping coefficient
$\Delta \Pr_{i,j}^s(f)$	Probability change at instantaneous frequency f due to an increase in rotational speed from ω_i to ω_j
$\varepsilon^\lambda[M]$	Approximation error using M largest wavelet coefficients and basis $\mathrm{B}^\lambda = \{g_m^\lambda\}_{1 \le m \le N}$
ε_k	Error between the filter output and k_{th} desired response
$\varphi'(t)$	Derivative of $\varphi(t)$
γ	Euler's constant
$\eta(s,\sigma)$	Smoothness index as a function of scale s and shape factor σ
$\mu(t)$	Modulus of the wavelet coefficients calculated for the noise-free vibration signal
ν^2	Noise variance

θ	Thresholding function
σ	Shape factor
ω_0	Excited resonance frequency
$\psi(t)$	Mother wavelet
$\psi_{s,u}(t)$	Daughter wavelet associated with scale s and translation index u
$\Psi(n)$	Euler's psi function

Chapter 1: Introduction

1.1 Background

As the failure of the machinery elements may lead to catastrophic results or unwanted production delays, fault detection and diagnosis has been a hot research topic in the past few decades. With the advances in digital signal processing methods and due to hardware simplicities, there has been an increasingly strong interest in the application of vibration analysis techniques for fault detection and diagnosis of rotating machinery elements. Oil condition data analysis is another widely used fault detection approach in industry. Each of the two types of information sources has its own merits and drawbacks. The vibration signal is 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 indirect and its effectiveness in fault detection largely relies on availability of proper signal processing techniques. The oil condition data, on the other hand, contain information about the size and quantity of metal debris in the oil and thus provide a direct perception of the machine condition. Nonetheless, it would be difficult to determine the location or source of the fault based on oil condition data analysis. The two types of data complement each other and the joint use of them would lead to better fault detection results. However, both sources are often severely tainted by various noises and thus the benefits of joint application of the two may never be materialized without proper de-noising.

The main approach in vibration based condition monitoring technique is based on the analysis of the vibration data acquired through the accelerometers mounted on or near the critical mechanical components. Bearings are among widely used and probably the most critical elements. As such, bearing fault detection and diagnosis has been the subject of extensive research (McFadden and Smith 1984a, McFadden and Smith 1985, Antoni and Randal 2006, McFadden and Smith 1984b). Several designs are available for bearings but all follow the same principles (Harris 1984). Figure 1.1 shows the building blocks and the geometry of a bearing (Braun 1986).

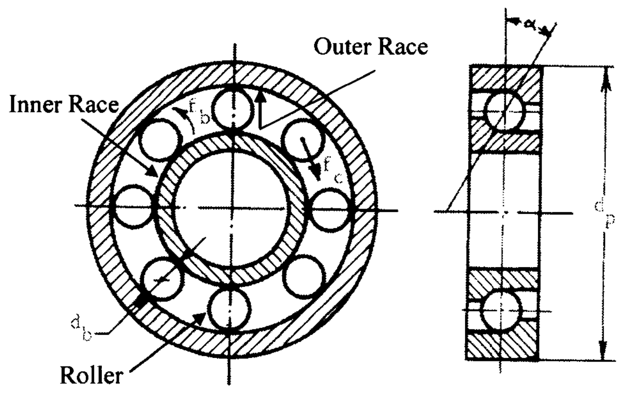
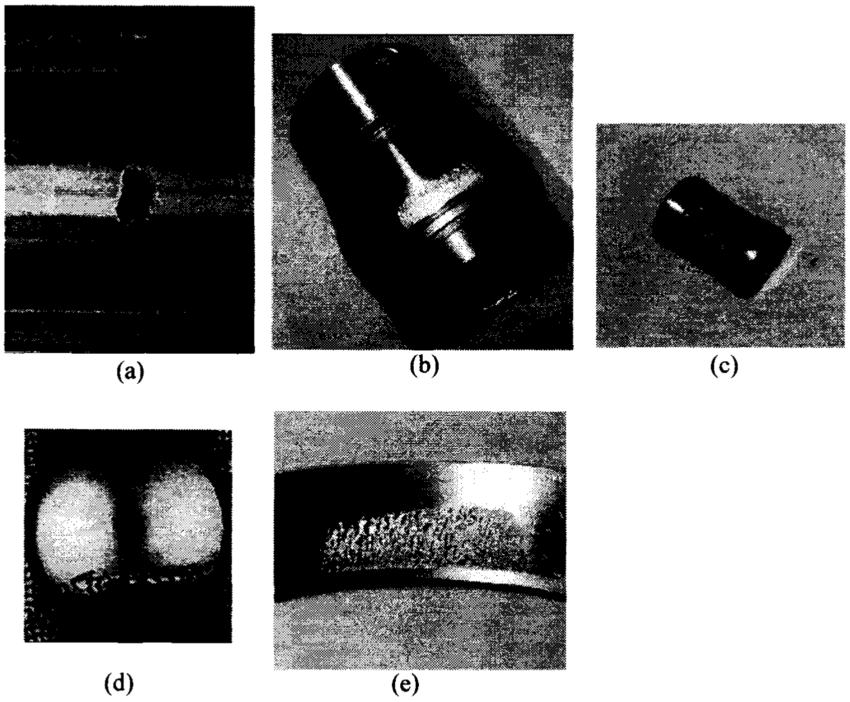


Figure 1.1 Bearing Geometry

Bearing faults usually appear as cracks or spalls on the surfaces of the bearing especially on the roller, outer race or inner race. Figure 1.2 shows examples of such faults (Kotzalas and Harris 2001, Qiu *et al* 2003).


(a) (b) (c)

(d) (e)

Figure 1.2 (a), (b) Ball and roller bearing inner race defects (c),(d) roller and ball defects (e) outer race defect (Kotzalas and Harris 2001, Qiu *et al* 2003)

When these faults come in contact with mating surfaces during the operation of the bearing, they generate impulses which are the main features to be detected through vibration analysis. The frequency of repetition of these impulses called "fault characteristic frequency" can be related to the rotational frequency of the shaft and the geometry of the bearing. These relations are provided below (Braun 1986, Ericsson *et al* 2005).

$$f_c = \frac{f_{ir}}{2}(1-\zeta) + \frac{f_{or}}{2}(1+\zeta)$$

$$f_b = \frac{d_b}{2d_p}(1-\zeta)(1+\zeta)(f_{or}-f_{ir})$$

$$f_{fo} = N_b(f_{or}-f_c)$$

$$f_{fi} = N_b(f_{ir}-f_c)$$

where $\zeta = \frac{d_b}{d_p}\cos\alpha$, f_{or} and f_{ir} are outer and inner race rotational frequencies respectively, α, d_b, d_p, f_b and f_c are defined in Fig. 1.1, f_{fo} and f_{fi} are outer and inner race fault characteristic frequencies respectively.

A comprehensive analytical model for the vibration of a roller bearing with a single fault has been described by McFadden and Smith (1984a). Comparing the analytical model with the measured vibration may lead to an effective method for fault detection. For this purpose, the fault features should be made evident via proper signal processing means as the measured signals are often contaminated by intensive noise, especially when these faults are at their early stages of development. As a result, a major step in fault detection is de-noising. One such de-noising method is filter based. With this approach, high SNR frequency band of the vibration signal passes through the filter whereas other frequency components are eliminated. Consequently the filtered signal should display the features of a faulty bearing more clearly. Bandpass filtering is also applied as a pre-processing step in the high frequency resonance technique (McFadden and Smith 1984b).

Besides the filter based de-noising, decomposition based de-noising such as the "soft-thresholding" method proposed by Donoho and Johnstone (Donoho and Jonestone 1994, Donoho 1995) has also been used for bearing vibration de-noising (Lin and Qu 2000). Though the comparison of the two methods seems in favour of the filter based approach

(Qiu *et al* 2003, Qiu *et al* 2006), its performance largely relies on the proper selection of parameters, i.e., the center frequency and the bandwidth of the filter. Several studies have attempted to address related issues by choosing proper parameters of a daughter Morlet wavelet and by applying the wavelet filter based de-noising method (Nikolaou and Antoniadis 2002, Qiu *et al* 2003, Lin and Zuo 2003, Qiu *et al* 2006).

Oil debris monitoring is another method to assess the health condition of machinery. In off-line oil analysis, oil samples are collected and analysed in laboratories. The analysis results can provide information about the health state of the mechanical components. Another approach uses a chip detector that utilizes magnetic capture to collect the metallic debris. An alarm system notifies the operator when the amount of collected debris passes some predefined threshold (Dempsey 2000, Miller and Kitaljevich 2000).

ODM (Oil Debris Monitor) is another on-line oil condition monitoring system. It is installed on the lubricating oil lines and provides a full flow passage way for the lubricating oil. It can detect metallic particles contained in the oil that pass through it (Industry Canada, 2007). This sensor generates a signature similar to a full period of a sine function with each metal particle passing. By processing the output signal it is possible to estimate the level of wear of the mechanical components.

This sensor has shown superior performance compared to the traditional magnetic chip collector. In contrast to the chip collectors this sensor is also sensitive to non-ferromagnetic particles. In addition, it does not require periodic inspection or cleaning (Miller and Kitaljevich 2000). However, like any other measuring devices the performance of the sensor is also affected by the noise and interferences which mask the signatures of interest. The interferences are due to vibrations of the structure where the sensor is mounted and manifest as addition of a combination of modulated sinusoidal signals to the sensor output. This masking effect leads to malfunction of the wear estimation system, either causing false alarms or leaving faults undetected. Research on this topic, except some studies carried out by a few researchers in our lab, has not been reported in the accessible literature.

1.2 Proposed study

1. *Design of an efficient de-noising algorithm to enhance the vibration signal measured from faulty bearings*

In this study, we propose a new criterion to specify the center frequency and bandwidth of a bandpass filter by adjusting the scale and shape factor of the Gabor wavelet. The Gabor wavelet is used because of optimal time and frequency resolution (Mallat 1998). The proposed criterion is reflected by the smoothness index, SI, which is the ratio of the geometric mean to the arithmetic mean of the wavelet coefficient moduli. It will be demonstrated that the envelope of the bandpass filtered signal or the modulus of the wavelet coefficients follows Rician distribution. This result is later used to prove that the SI converges to a constant number (0.8455...) in the absence of mechanical faults or when the SNR of the vibration signal is very low. It is also shown that the SI value decreases in the presence of fault signatures with properly selected parameters. As our purpose is to detect or reveal the fault signature, the selection of the Gabor wavelet parameters can be guided by the process of minimizing SI.

However, this approach is based on an assumption that the most impulsive components of the measured vibration are due to mechanical faults. This assumption may not be valid in general whereas a coloured noise, electrical interference or any other signal components not of our interest may show similar impulsive features. One example is the EMI (Electro Magnetic Interferences) from the PWM (Pulse Width Modulation) AC controllers which are usually used in CNC machines. On the other hand, the proposed method requires a global search for the minimum SI calculated for all combinations of wavelet parameters on the chosen discretized ranges which is a computationally demanding task. Though the computation time required by the method does not pose a major problem while a sudden breakdown of the bearing without any earlier warning is not expected, reducing the computational complexity of the algorithm can increase its reliability and applicability. For these reasons a novel scale selection method is proposed. In this method we use our knowledge of the fault signature and the consequent resonance frequency excitation phenomenon. The best shape factor is then found in a separate step by minimizing SI at a chosen scale.

Furthermore, through bandpass filtering the signal, the in-band noise with frequency content in the range covered by the bandpass filter is not eliminated. As a result, the performance of the wavelet filter based de-noising method deteriorates as the background noise intensity increases. This deterioration goes to the point where the wavelet filter based de-noising method does not provide considerable enhancement in the quality of the vibration signal. In order to improve the efficiency of the method, spectral subtraction is applied prior to bandpass filtering. Spectral subtraction has been used to enhance the speech signals degraded by additive noise (Lim and Oppenheim 1979). It has also been used in mechanical fault diagnosis to improve the sensitivity of scalar indicators (e.g., crest factor and kurotsis) (Dron, Bolaers and Rasolofondraibe 2004).

In this work the spectral subtraction technique is used to reduce the intensity of the background noise that is uncorrelated with bearing vibrations in order to find an approximation of the bearing vibration plus the vibration generated by other sources like shaft imbalance, gear meshing, etc. This partly purified signal is then bandpass filtered in order to eliminate the noise due to vibrations of the other machinery elements. The proposed method leads to improvements in both the final result of the process and the capability of the wavelet filter based de-noising method for lower SNR vibration signals.

2. *Development of an algorithm to eliminate some interfering vibration signal in oil debris sensors*

In the final part of this work the problem of interferences affecting the output of oil debris sensors is addressed. To the best of our knowledge, this problem is investigated in Manufacturing Systems and Machine Health (MSMH) Lab at the Department of Mechanical Engineering of University of Ottawa for the first time. Adaptive Line Enhancement is applied to remove such noises and clean the signature of the passing metallic particles. This approach uses the correlation between the measured signal and a delayed version of the same data. This correlation is due to such interferences since the passages of metallic particles follow a random pattern with no correlation between the corresponding output signatures. The proposed method removes the interferences due to the vibrations and leaves the signatures of the passing metallic particles intact. Accordingly, more reliable wear estimate can be expected.

1.3 Organization of the thesis

This thesis is organized as follows. Chapter 2 provides a review of different de-noising algorithms used to enhance the vibration signals measured from faulty bearings and gears. This chapter also presents a brief comparison between two major de-noising algorithms namely filter based and decomposition based. This is followed by a review of existing approaches to the selection of the wavelet parameters for wavelet filter based de-noising and their limitations. Chapter 3 presents a new method for wavelet parameter selection. This chapter starts with an explanation of envelope detection. Ratio of the geometric mean to the arithmetic mean is presented and its implication on parameter selection is addressed. The drawbacks of the proposed method are explained in chapter 4. In this chapter a novel scale selection method is proposed to improve the robustness and applicability of the proposed de-noising algorithm. Later the concept of spectral subtraction and its application in performance enhancement of the wavelet filter based de-noising is investigated. The proposed de-noising method is evaluated using experimental data. Chapter 5 deals with the problem of interfering signals that affect the performance of the oil debris sensors. The nature of such interferences is explained and a possible solution through adaptive noise cancellation is explored. Adaptive Line Enhancement (ALE) technique is proposed to mitigate the problem and results are provided for both simulated and experimental data. Chapter 6 concludes this work and presents suggestions for future research.

Chapter 2: Review of De-noising Methods

The vibration generated by a single faulty mechanical element is expected to follow some known patterns and can be modeled. The knowledge of these patterns would help us construct a health condition monitoring system. However, when such components are installed as part of a complex system it would be difficult to use the same health assessment techniques without additional modifications. In other words the fault detection and diagnosis schemes should be tailored for every single application. To develop an effective machine health assessment method, availability of proper signal enhancement techniques is very important. These techniques remove noise and interferences which are defined as any signal component other than the vibration signal generated by the faulty mechanical element. As explained in the introduction, two major de-noising approaches, namely, decomposition based de-noising and filter-based de-noising methods, have been used to enhance the vibration signals measured from faulty bearings or gears. In the following sections, we provide a review of these de-noising schemes which is preceded by a brief introduction to wavelet transform.

2.1 Wavelet transform

The wavelet transform as a specialized field can be traced to the work of Grossman and Morlet (Grossmann and Morlet 1984). Their work was motivated by the possibility of expressing certain seismic signals as a combination of translations and dilations of an oscillatory function called wavelet (Rao and Bopardikar 1998).

2.1.1 Continuous Wavelet transform

The continuous wavelet transform (CWT) of $f(t)$ with respect to a wavelet $\psi(t)$ is defined as (Mallat 1998, Rao and Bopardikar 1998):

$$W_f(s,u) = \int_{-\infty}^{+\infty} f(t)\frac{1}{\sqrt{s}}\psi^*(\frac{t-u}{s})dt \tag{2-1}$$

where s and u are real and asterisk denotes complex conjugate.

Equation (2-1) can be written in a more compact form by using:

$$\psi_{s,u}(t)=\frac{1}{\sqrt{s}}\psi(\frac{t-u}{s}) \tag{2-2}$$

If the mother wavelet $\psi(t)$ satisfies the admissibility condition, such that

$$C\equiv\int_{-\infty}^{\infty}\frac{|\hat{\psi}(\omega)|^2}{|\omega|}d\omega,0<C<\infty$$

then

$$f(t)=\frac{1}{C}\int_{u=-\infty}^{\infty}\int_{s=-\infty}^{\infty}\frac{1}{|s|^2}W(s,u)\psi_{s,u}(t)dsdu \tag{2-3}$$

In other words, the original time signal can be reconstructed using the wavelet coefficients calculated in the transform given by equation (2-1).

2.1.2 Orthogonal Wavelets

Orthogonal wavelet transform provides a non-redundant wavelet representation. The wavelet transform is given by:

$$\langle f,\psi(2^{-j}t-n)\rangle=\frac{1}{\sqrt{2^j}}\int_{-\infty}^{\infty}f(t)\psi^*(2^{-j}t-n)dt \tag{2-4}$$

where $\langle\psi(2^{-k}t-m),\psi(2^{-j}t-n)\rangle=0$ for $m\neq n$ or $j\neq k$.

Then

$$f=\sum_{j=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\langle f,\psi_{j,n}\rangle\psi_{j,n} \tag{2-5}$$

$$\psi_{j,n}=\frac{1}{\sqrt{2^j}}\psi(2^{-j}t-n)$$

As can be seen, a discrete number of wavelet coefficients are required to reconstruct the original function $f(t)$.

2.2 De-noising

2.2.1 Wavelet filter-based de-noising

The convolution of two signals $h(t)$ and $x(t)$ is defined as:

$$h(t)*x(t)=\int_{-\infty}^{\infty}h(\tau)x(t-\tau)d\tau \tag{2-6}$$

According to this relation and considering equation (2-1), we have (Rao and Bopardikar 1998):

$$f(u)*\psi_{s,0}^{*}(-u)=\int_{-\infty}^{+\infty}f(t)\frac{1}{\sqrt{s}}\psi^{*}(\frac{t-u}{s})dt=W_f(s,u) \tag{2-7}$$

Denoting Fourier transform of a function with respect to the variable x by F_x, we obtain:

$$F_u[W_f(s,u)]=F_u[f(u)]F_u[\psi_{s,0}^{*}(-u)] \tag{2-8}$$

According to Eq. (2-8), for a fixed scale we can consider wavelet transform $W_f(s,u)$ as the output of a filter with impulse response $\psi_{s,0}^{*}(-u)$ and input $f(u)$. When proper filter $F_u[\psi_{s,0}^{*}(-u)]$ is used, this filtering process can result in a signal with higher signal to noise ratio.

2.2.2 Wavelet threshold de-noising (Decomposition based)

Equation (2-5) shows the reconstruction formulas for orthogonal wavelets. On the other hand, according to equation (2-4) the wavelet coefficients are the correlations of the function $f(t)$ with the wavelet basis. With this interpretation we expect to see higher wavelet coefficient values on the intervals where $f(t)$ has higher correlation with the daughter wavelet or in other words where the wavelet better describes the features of the function. Bearing this concept in mind, it is possible to construct an approximation of the original function using the larger wavelet coefficients. In this case the approximation function f_m can be considered as a projection of the original function on the M vectors in I_m (a collection of wavelets chosen from $\mathrm{B}=\{g_m\}_{m\in N}$ for the reconstruction purpose Mallat (1998)):

$$f_m = \sum_{m \in I_m} \langle f, g_m \rangle g_m \tag{2-9}$$

where $\mathrm{B} = \{g_m\}_{m \in N}$ is an orthonormal basis. f_m can also be written as an inner product thresholding:

$$f_m = \sum_{m=0}^{\infty} \theta_T\left(\langle f, g_m \rangle\right) g_m \tag{2-10}$$

where

$$\theta_T(x) = \begin{cases} x & if\,|x| \geq T \\ 0 & if\,|x| < T \end{cases} \tag{2-11}$$

T is the threshold value which defines I_m. Considering the orthonormality of the wavelet basis, the approximation error will be the sum of the remaining coefficients:

$$\varepsilon[M] = \|f - f_m\|^2 = \sum_{m \notin I_m} |\langle f, g_m \rangle|^2 \tag{2-12}$$

To minimize this error, the indices in I_m must correspond to the M vectors having the largest inner product amplitude $|\langle f, g_m \rangle|$. As mentioned before, these are the vectors that best correlate f and thus can be interpreted as the main features of f (Mallat 1998). It is obvious that this error approaches zero as M increases.

Following this concept, it is possible to discriminate between the available basis families according to their capability of condensing the major features of f in smaller number of large wavelet coefficients. In other words for properly selected basis family the error given in equation (2-12) goes quickly to zero as M increases, providing that the decreasingly sorted values $|\langle f, g_m \rangle|$ have a fast decay. More specifically, consider a dictionary D that is a union of orthonormal bases in a signal space of dimension N:

$$D = \bigcup_{\lambda \in \Lambda} \mathrm{B}^{\lambda}$$

Each orthonormal basis is a family of N vectors:

$$\mathrm{B}^{\lambda} = \{g_m^{\lambda}\}_{1 \leq m \leq N}$$

As defined before the approximation error is:

$$\varepsilon^{\lambda}[M]=\sum_{m\notin I_M^{\lambda}}\left|\left\langle f,g_M^{\lambda}\right\rangle\right|^2=\|f\|^2-\sum_{m\in I_M^{\lambda}}\left|\left\langle f,g_M^{\lambda}\right\rangle\right|^2 \tag{2-13}$$

The basis family $\mathrm{B}^{\alpha}=\left\{g_m^{\alpha}\right\}_{1\le m\le N}$ is a better basis than $\mathrm{B}^{\gamma}=\left\{g_m^{\gamma}\right\}_{1\le m\le N}$ for approximating f if for all $M\ge 1$:

$$\varepsilon^{\alpha}[M]\le\varepsilon^{\lambda}[M]$$

Using equation (2-13) the above condition can be written as:

$$\forall M\ge 1,\ \sum_{m\in I_M^{\alpha}}\left|\left\langle f,g_M^{\alpha}\right\rangle\right|^2\ge\sum_{m\in I_M^{\gamma}}\left|\left\langle f,g_M^{\gamma}\right\rangle\right|^2$$

Marshal and Olkin (1979) has proved that the above criterion is satisfied if and only if, for all concave functions $\Phi(u)$, the following is satisfied

$$\sum_{m=1}^{N}\Phi\left(\frac{\left|\left\langle f,g_m^{\alpha}\right\rangle\right|^2}{\|f\|^2}\right)\le\sum_{m=1}^{N}\Phi\left(\frac{\left|\left\langle f,g_m^{\gamma}\right\rangle\right|^2}{\|f\|^2}\right)$$

In practice, two bases are compared using a single concave function $\Phi(u)$. In other words the cost of approximating f in a basis B^{λ} is defined by the sum:

$$C(f,\mathrm{B}^{\lambda})=\sum_{m=1}^{N}\Phi\left(\frac{\left|\left\langle f,g_m^{\lambda}\right\rangle\right|^2}{\|f\|^2}\right)$$

If the basis B^{α} is a better basis than B^{γ} for approximating f then:

$$C(f,\mathrm{B}^{\alpha})\le C(f,\mathrm{B}^{\gamma})$$

This condition is necessary but not sufficient since we test a single concave function. Coifman and Wickerhauser (1992) found a best basis B^{α} in D using Shannon entropy as the cost function. The entropy $\Phi(x)=-x\log_e x$ is a concave function for $x>0$ and the corresponding cost is called the entropy of the energy distribution defined as:

$$C(f,\mathrm{B})=-\sum_{m=1}^{N}\frac{\left|\left\langle f,g_m^{\lambda}\right\rangle\right|^2}{\|f\|^2}\log_e\frac{\left|\left\langle f,g_m^{\lambda}\right\rangle\right|^2}{\|f\|^2} \tag{2-14}$$

It is important to note that these results do not hold for non-orthogonal bases in general.

Now we can consider the problem of signal estimation in a noisy environment. The noisy signal $Z[n]$ is given as:

$$Z[n]=Y[n]+W[n]$$

where $Y[n]$ is the signal and $W[n]$ is the noise. In the case of white Gaussian W[n], the energy of the noise is diffused across all frequencies and as a result over all the vectors of any orthogonal basis. With this fact in mind and following the previous discussions, estimation process can thus be optimized by finding a basis that concentrates the signal energy over a few coefficients. In other words it is possible to reconstruct the signal by using larger wavelet coefficients and discarding the remaining coefficients that are mostly related to the contaminating noise. This is the hard threshold de-noising approach where the thresholding function is given by equation (2-11). In the case of coefficient selection approach given in equations (2-10) and (2-11), it is shown that minimum mean square error (MSE) is achieved when the following ideal selection threshold is applied (Mallat 1998):

$$T_m = \sigma_m^{\ 2}$$

$$\sigma_m^{\ 2} = E\left\{ \left| \langle W, g_m \rangle \right|^2 \right\}$$

The thresholding function can be considered in other forms as an attenuation rule where each wavelet coefficient is attenuated according to the noise variance. The following equation shows such a thresholding method:

$$\tilde{Y} = \sum_{m=0}^{N-1} \langle Z, g_m \rangle \theta[m] g_m$$

It can be shown that the following thresholding function minimizes the MSE of the estimation (Mallat 1998).

$$\theta[m] = \frac{\left| \langle f, g_m \rangle \right|^2}{\left| \langle f, g_m \rangle \right|^2 + \sigma_m^{\ 2}}$$

In practice we cannot implement ideal attenuation and selection since we do not know the values $\langle f, g_m \rangle$.

A hard thresholding is a coefficient selection procedure as given by equation (2-11) and produces a MSE that is larger than the ideal coefficient selection. Another approach for thresholding is soft thresholding rule where the threshold value is subtracted from the coefficients that are larger than the threshold T and as before the coefficients smaller than T are set to zero.

Donoho and Johnston (1994) proved that in the case of Gaussian noise, it is possible to have an estimation error that is within a factor of $2\log_2 N$ of an ideal selection error by applying the hard or soft thresholding rule and using an appropriate threshold$T_m = \sigma_m \sqrt{2\log_e N}$. However, in this approach the noise variance σ_m should be known.

When the corrupting noise has white Gaussian nature it is possible to estimate the noise variance. For this purpose we assume that the original function has a relatively smooth behaviour. In this case the function to be estimated does not contain much high frequency information or in other words the wavelet coefficients corresponding to the finest scale are mostly related to the additive white Gaussian noise. Then an estimation of the wavelet coefficients variance can give an estimate of the noise variance σ. An estimator insensitive to large outliers is given below (Mallat 1998):

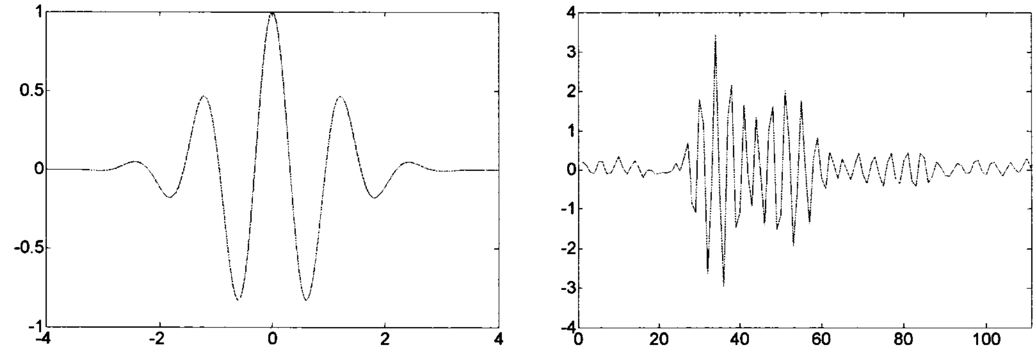
$$\tilde{\sigma} = \frac{1}{0.6745} Med\left(\left|\left\langle Z, \psi_{1,m} \right\rangle\right|\right)_{0 \le m < \frac{N}{2}}$$

where *Med* stands for median.

2.3 Non-orthogonal wavelets for threshold de-noising

For the vibration signal measured from mechanical components like bearings and gears, the main fault features are the impulses generated due to the presence of cracks or spalls. As a result, the main purpose of de-noising is to make the hidden fault generated impulses visible and to eliminate the corrupting noise. Consequently, all the components in the measured signal other than the mentioned impulsive features are considered as noise. To perform the threshold de-noising explained in previous section, orthogonal wavelets should be used. As mentioned before to obtain better de-noising results, vectors of the selected wavelet basis should be well correlated with the features of the signal. In this case the chosen wavelet basis should contain vectors similar to the fault generated impulses. This characteristic is not well met among the orthogonal wavelets (Lin and Qu 2000). However, disregarding the orthogonality constraint provides a wider range of selections of wavelet basis. One such wavelet is Gabor wavelet that is also often referred to as Morlet wavelet. The real part of this wavelet is shown in Figure 2.1a. An example of a fault generated impulse due to the passage of a roller over a crack on the outer race of a roller bearing is displayed in Figure 2.1b.

Lin and his co-investigators (Lin and Qu 2000, Lin, Zuo and Fyfe 2004) used Morlet wavelet as the wavelet basis and applied wavelet threshold de-noising approach to de-noise the vibration signal measured from both faulty bearings and gears. In their studies, a variation of soft thresholding rule proposed by Donoho (1995), called "generalized soft-thresholding" and a thresholding rule based on Hyvarinen (1999) were applied. The results were reportedly better compared to the threshold de-noising results yielded by orthogonal wavelet basis.



(a) (b)

Figure 2.1 (a) Morlet wavelet, (b) Signature of a faulty bearing

2.4 Comparison of threshold de-noising and wavelet filter based de-noising

Qiu et al (2003, 2006) compared wavelet filter based approach with decomposition based de-noising method. They argued that though the threshold de-noising method performs very well on Gaussian noise and can almost achieve optimal noise reduction, there are still some issues which require further investigation. The first is how to select an optimum wavelet for a particular type of signal. The second issue is how to select a threshold and develop a thresholding function for the reconstruction process. Despite the large variety of threshold selection strategies (e.g., Donoho and Jonestone 1994, Donoho 1995, Chang, Bin and Vetterli 2000, Hansen and Bin 2000), threshold selection in cases where prior knowledge about the data is limited is yet to be investigated. Finally the idea of condensing the energy of the signal on a few wavelet coefficients works well only if signal is a smooth curve with little or no abrupt changes. However, for the measured vibration from faulty bearings or gears which usually consists of many impulses, a sparse wavelet representation is very difficult to achieve. In the studies of Qiu et al (2003, 2006), they applied both de-noising algorithms to de-noise simulated signals and measured signals from faulty bearings. Their de-noising results have shown that the wavelet filter based method outperformed the threshold based method.

As explained before, the wavelet threshold de-noising method relies on the basic idea that the main energy of the signal can be concentrated in a few coefficients. The nonlinear thresholding techniques will then retain a few larger coefficients and discard the noise coefficients. It was also explained before that these large wavelet coefficients result from the high correlation between the signal and the daughter wavelet which is chosen so as to resemble the features of interest.

On the other hand, fault impacts usually excite natural frequencies of the bearing, sensor and structure. This excitation generates an interval of ringing or oscillations which damps out according to the damping characteristics of the system. Considering this phenomenon and the fact that a short duration of measured signal is often used for de-noising and fault detection, we do not expect very different excited resonance frequencies or damping characteristics for these fault signatures. As a result, it is not unrealistic to assume that almost all the fault generated impulses in one data set are very similar, affected by the same structural damping coefficient and composed of the same range of excited resonance frequencies.

In this case using a daughter wavelet at a specific scale we expect to see the large wavelet coefficients to repeat at the same time instants where the fault generated impulses are located in the original noisy vibration signal. The wavelet coefficients corresponding to one specific scale are the result of bandpass filtering the signal with the filter formed by the daughter wavelet. This explanation further justifies application of wavelet filter based de-noising method. However, the parameters of the daughter wavelet such as scale should be chosen in a way that the filtering process results in a better realization of existence of impulses and consequently better detection performance.

2.5 Review of parameter selection criteria in wavelet filter based de-noising

2.5.1 Previous studies on shape factor selection and their limitations

As explained earlier, a wavelet similar to the features of our interest is Gabor wavelet. This wavelet is defined as:

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

Using this wavelet, the parameter σ called shape factor may also be adjusted to improve de-noising performance. Nikolaou and Antoniadis (2002) applied wavelet filter based de-noising method with Gabor wavelet. In this work two parameters of this wavelet namely scale and shape factor were selected in order to achieve a better demodulation performance. Three criteria defined below were used to select the shape factor:

Shannon entropy

$$E(\sigma) = -\sum_{k=1}^{M} d_k \log d_k \qquad (2\text{-}15)$$

where $d_k = \dfrac{|C_w(i,j)|}{\sum_{i,j} |C_w(i,j)|}$.

Magnification factor

$$c_r(\sigma) = \frac{MN \max\left(C_w(i,j)\right)}{\sqrt{\sum_{i,j} C_w(i,j)^2}}$$

where C_w is the matrix formed by all wavelet coefficients and MN is the number of the elements in matrix C_w,

and the ratio of the two as a combined index

$$f_{sc}(\sigma) = \frac{c_r(\sigma)}{E(\sigma)}.$$

These indices were calculated using all the wavelet coefficients. The shape factor that minimizes $E(\sigma)$ and maximizes $c_r(\sigma)$ as well as $f_{sc}(\sigma)$ was selected. According to (Nikolaou and Antoniadis 2002), all the indices provided almost the same result. This form of entropy given by equation (2-15) was first used by Lin and Qu (2000).

The magnification factor given above would be very sensitive to outliers in the data as its value is defined by a single maximum wavelet coefficient among all the coefficients. As a result, the presence of a single large wavelet coefficient can lead to wrongful selection of the shape factor. On the other hand, as explained before entropy as a concave function, given in equation (2-14), cannot be used for performance measurement of non-orthogonal bases. However, another interpretation may justify application of the form given in equation (2-15) as a performance measurement criterion. Entropy *En* is defined for a discrete-value random variable X as (Hyvarinen and Karhunen 2001):

$$En(X) = -\sum_i P(X = a_i)\log P(X = a_i)$$

where a_i is a possible value of X and $P(X = a_i)$ is the probability of the random variable X taking value a_i. Considering a function

$$f(p) = -p\log p \quad \text{for} \quad 0 \le p \le 1$$

One has

$$En(X) = -\sum_i f\left(P(X = a_i)\right)$$

Figure 2.2 shows function $f(p)$ plotted on interval [0, 1].

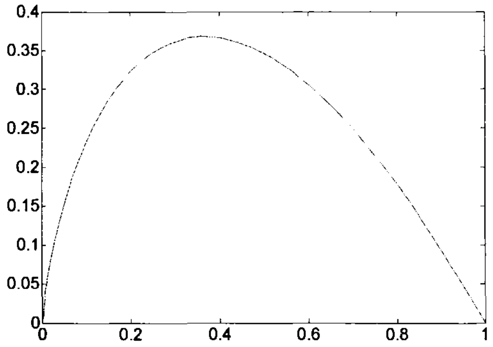


Figure 2.2 The function *f*(*p*) plotted on interval [0,1]

As shown in Figure 2.2, the entropy is small if the probabilities $P\left(X = a_i\right)$ are close to 1 or zero and larger for values in between. The entropy of a random variable can be interpreted as the level of information gained by the observation of the variable. More random and unstructured random variables result in larger entropy values. If all probabilities are close to zero except one which will be closer to one there is little randomness in the variable. In other words the random variable is more structured. This is reflected in small entropy value. On the other hand if all the probabilities are equal, then they are relatively far from zero or one and entropy takes larger values. This means that the variable is random and unstructured.

A similar concept can be applied to a time series. In the case of the wavelet coefficients, a few large wavelet coefficients are more suitable. Hence, the proposed criterion given in equation (2-15) may be used as a measure of sparsity of the time series. Qiu et al. (2003, 2006) also used this criterion for shape factor selection.

However, there is an important problem common to all these approaches. Selection of a basis according to the above criteria matches the concept of decomposition based de-noising method as the final result is formed using the whole frequency content of the signal. This approach finds the best shape factor to match the whole frequency content of the signal covered by the selected scales while finally a specific frequency band will be chosen and other frequency components of the signal are discarded. In simple words, through these selection criteria, the shape factor is chosen utilizing all the wavelet coefficients calculated at all pre-selected scales while finally one scale will be selected and the remaining scales will not be used in the de-noising process.

As we will explain later, the shape factor defines the bandwidth of the bandpass filter. Depending on the chosen center frequency or scale, this bandwidth should be adjusted. In other words, adjusting the bandwidth of the bandpass filter disregarding the choice of center frequency may not be justified.

2.5.2 Previous studies on scale selection and their limitations

Nikolaou and Antoniadis (2002) selected the best scale in a separate step. They introduced two different variations for this purpose. In the first variation the demodulated signal was formed by the maximum value of the absolute wavelet coefficients at each time translation index over all pre-selected scales. This approach was taken based on the assumption that different resonance frequencies might be excited each time a fault comes in contact with the mating surfaces. As a result, consecutive impulses might correspond to different frequency bands. In the second variation, a threshold is defined and at each scale the number of wavelet coefficients exceeding this threshold is calculated. The scale corresponding to the highest number of wavelet coefficients that are larger than the threshold is chosen as the best scale. The experimental results on this work however reported similar outcomes through both approaches. This result indicates that the assumption that similar resonance frequencies are excited due to consecutive impulses is valid. The first approach accentuates the magnitude of wavelet coefficients. This approach may not work well in the presence of other vibration components or when the incipient fault signal is very weak. In the second variation a single threshold value is selected based on the mean and standard deviation of the modulus of all the wavelet coefficients. This approach again leads to misleading results especially for the weak signatures of incipient faults.

In the works by Qiu et al. (2003, 2006) scale selection was performed based on the periodicity criterion defined by the singular values of the matrices formed by the wavelet coefficients. In simple words, a scale resulting in the wavelet coefficients that display the highest periodicity was selected as the best scale. This selection criterion is based on the assumption that the most periodic component of the vibration signal is the vibration resulting from consecutive fault impulses. This approach however may raise questions while vibration signals measured from rotating machinery contains other periodic components like gear meshing or shaft imbalance. Furthermore, roller slippage may occur in practice which affects the level of periodicity of the generated impulses. In this case, building the whole de-noising process based on the periodicity assumption may cause misleading results.

Lin and Zuo (2003) proposed to find the best scale-shape factor combination by maximizing the kurtosis. In their work the vibration signal measured from faulty gears were de-noised. Kurtosis can be considered as a measure of non-Gaussianity (Hyvarinen and Karhunen 2001) of the filtered signal. Hence the most non-Gaussian frequency band of the signal is passed through the filter and other frequency components are eliminated. This approach is based on the assumption that the most non-Gaussian portion of the measured signal corresponds to the vibrations generated by the faulty bearings or gears. This however may not be valid in general while some electrical noise, interference or signal component may also have similar non-Gaussian characteristics. On the other hand, it is reported that kurtosis is very sensitive to outliers in the signal (Hyvarinen and Karhunen 2001). Furthermore, this approach, like other previous methods, requires examination of all available scale-shape factor combinations to select the best pair, which is time-consuming and may not be suitable for on-line applications.

Though the above has revealed some limitations of the existing selection criteria, they do have their merits. It is very difficult to completely reject an existing method and to design a new one that outperforms all the other selection algorithms in all aspects. One approach may not be successful in certain circumstances but may outperform other selection algorithms in other cases. In the following two chapters, a new wavelet parameter selection algorithm is proposed. The main goal is an efficient and exact selection method which keeps its high performance regardless of the working conditions.

2.6 Summary of existing methods for wavelet parameter selection

2.6.1 Shape factor selection

1- Entropy minimization.

This approach finds the best shape factor to match the whole frequency content of the signal covered by the selected scales. Finally a specific frequency band will be chosen and other frequency components of the signal are discarded

2- Kurtosis maximization

This method requires a global search for proper parameters and is very sensitive to outliers in the data.

3- Magnification factor maximization.

4- Maximization of a combined index formed by Shannon entropy and magnification factor.

Criteria (3) and (4) are both very sensitive to outliers in the data.

2.6.2 Scale selection

1- Scale with highest number of wavelet coefficients exceeding some threshold.

This may not work very well for weak signatures of incipient faults.

2- Scale with the highest periodicity of corresponding coefficients.

This approach assumes that the most periodic component of the signal corresponds to the faulty bearing vibration while there are other periodic components in the vibration signal measured from rotating machinery.

3- Kurtosis maximization.

It requires a global search for proper parameters assuming that the most non-Gaussian component of the measured signal corresponds to faulty bearing vibrations which may not be true in general.

Chapter 3: Wavelet Parameter Selection in Signal De-Noising and Fault Detection

As reviewed in Chapter 2, the filter-based de-noising method compares favourably with the threshold-based method (Qiu et al 2003, 2006). However, the performance of this de-noising scheme is greatly affected by the wavelet parameters, i.e., the scale and shape factor. The scale defines the center frequency of the corresponding filter and the shape factor defines the bandwidth. An efficient selection method can boost the performance of the health assessment unit considerably. Therefore, this chapter focuses on the wavelet parameter selection problem in the filter-based de-noising method.

3.1 Envelope Detection

In this work we choose Gabor wavelet as the mother wavelet. This wavelet is obtained through frequency modulation of a Gaussian window. Gaussian windows are used because of their optimal time and frequency energy concentration, proved by the Heisenberg Uncertainty theorem (Mallat 1998). Gabor wavelet is defined as:

$$\psi(t) = ce^{-\sigma^2 t^2} e^{i2\pi f_0 t} \tag{3-1}$$

In this study, constant c is chosen as:

$$c = \sqrt[4]{\frac{8\sigma^2}{\pi}} \tag{3-2a}$$

so that:

$$\int_{-\infty}^{\infty} \mathrm{Re}(\psi(t))^2 dt \approx \int_{-\infty}^{\infty} \mathrm{Im}(\psi(t))^2 dt \approx 1 \tag{3-2b}$$

To elaborate on this, $\int_{-\infty}^{\infty} \mathrm{Re}(\psi(t))^2 dt$ is written as:

$$\begin{aligned} \int_{-\infty}^{\infty} \mathrm{Re}(\Psi(t))^2 dt &= \int_{-\infty}^{\infty} \left[ce^{-\sigma^2 t^2} \cos 2\pi f_0 t\right]^2 dt = c^2 \int_{-\infty}^{\infty} e^{-2\sigma^2 t^2} \frac{1+\cos 4\pi f_0 t}{2} dt \\ &= \frac{c^2}{2}\left[\int_{-\infty}^{\infty} e^{-2\sigma^2 t^2} dt + \int_{-\infty}^{\infty} e^{-2\sigma^2 t^2} \cos 4\pi f_0 t dt\right] \end{aligned} \tag{3-3a}$$

From (Gradshteyn and Ryzhik 2000) the above statement can be expressed as:

$$\frac{c^2}{2}\left[\sqrt{\frac{\pi}{2\sigma^2}}+e^{-\frac{2(\pi f_0)^2}{\sigma^2}}\sqrt{\frac{\pi}{2\sigma^2}}\right] \tag{3-3b}$$

Similarly:

$$\begin{aligned}\int_{-\infty}^{\infty}\mathrm{Im}\left(\Psi(t)\right)^2 dt &= \int_{-\infty}^{\infty}\left[ce^{-\sigma^2t^2}\sin 2\pi f_0 t\right]^2 dt = c^2\int_{-\infty}^{\infty}e^{-2\sigma^2t^2}\frac{1-\cos 4\pi f_0 t}{2}dt\\ &=\frac{c^2}{2}\left[\int_{-\infty}^{\infty}e^{-2\sigma^2t^2}dt-\int_{-\infty}^{\infty}e^{-2\sigma^2t^2}\cos 4\pi f_0 t dt\right]=\frac{c^2}{2}\left[\sqrt{\frac{\pi}{2\sigma^2}}-e^{-\frac{2(\pi f_0)^2}{\sigma^2}}\sqrt{\frac{\pi}{2\sigma^2}}\right]\end{aligned} \tag{3-3c}$$

The last term inside the bracket in Eqs. (3-3b) and (3-3c) is very small for $f_0 \geq 1$ and $\sigma \leq 1$, which can be neglected. Hence, for $c=\sqrt[4]{\frac{8\sigma^2}{\pi}}$ the above leads to:

$$\int_{-\infty}^{\infty}\mathrm{Re}(\psi(t))^2 dt \approx \int_{-\infty}^{\infty}\mathrm{Im}(\psi(t))^2 dt \approx \frac{c^2}{2}\sqrt{\frac{\pi}{2\sigma^2}}=1 \tag{3-3d}$$

According to Eqs. (2-8) and (3-1), one obtains:

$$F_u[W(s,u)]=\sqrt{s}F_u[f(u)]\hat{\psi}^*(sf) \tag{3-4}$$

where $\hat{\psi}(f)$ is the Fourier transform of $\psi(t)$ obtained by:

$$\hat{\psi}(f)=\int_{-\infty}^{\infty}ce^{-\sigma^2t^2}e^{i2\pi f_0 t}e^{-i2\pi f t}dt=c\sqrt{\frac{\pi}{\sigma^2}}e^{-(\frac{\pi^2}{\sigma^2})(f-f_0)^2} \tag{3-5}$$

Equations (3-4) and (3-5) show that the wavelet transform at a constant scale acts like a bandpass filtering process with a Gaussian filter. As explained before, for the properly selected bandwidth and center frequency, this process results in a signal with higher signal to noise ratio. The bandwidth of this filter is adjusted by the shape factor σ and the center frequency is adjusted by the scale s as illustrated in Figure 3.1.

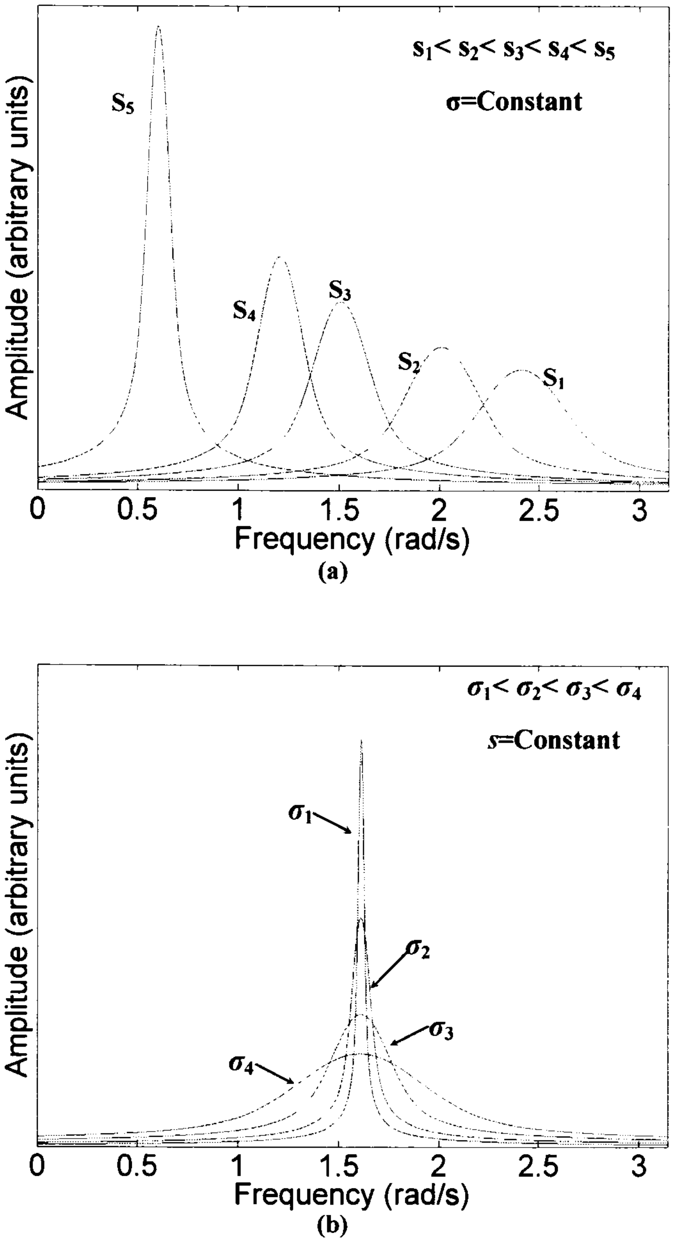


Figure 3.1 Gabor wavelet in frequency domain for (a) fixed shape factor and different scales (b) fixed scale and different shape factor

On the other hand, according to Eq. (3-5) the filtered signal is analytic for $\frac{(\pi f_0)^2}{\sigma^2} >> 1$ (Mallat 1998). Therefore, the modulus of this analytic result provides the envelope of the band pass filtered signal (Mallat 1998, Nikolaou and Antoniadis 2002). Denoting the

Gabor wavelet transform of a signal $f(t)$ at scale s and shape factor σ by $W_f^{s,\sigma}(u)$, the envelope of the bandpass filtered $f(t)$ is obtained by

$$\left|W_f^{s,\sigma}(u)\right| = \sqrt{\operatorname{Re}\left(W_f^{s,\sigma}(u)\right)^2 + \operatorname{Im}\left(W_f^{s,\sigma}(u)\right)^2} \qquad (3\text{-}6)$$

This envelope will later be used in the de-noising process.

3.2 The G/A ratio

The geometric mean of a series is defined as:

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

for a positive time series $S(n)$ (n =1, 2 ,..., N). Similarly the arithmetic mean of the series is:

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

The ratio of the two, $r_{G/A}$, is therefore

$$r_{G/A} = \frac{G_s}{A_s} \qquad (3\text{-}7)$$

As $A_s \geq G_s$ (Abramowitz and Stegun 1972) (Equality only holds when $S[n] = S[m]$ for all m and n), $r_{G/A}$ is always between zero and one for a positive time series. The geometric mean can be rewritten as follows:

$$G_s = \exp\left(\frac{1}{N}\sum_{n=1}^{N} \ln(S(n))\right) \qquad (3\text{-}8)$$

Then for a continuous function of time, one has:

$$r_{G/A} = \frac{\exp\left[\frac{1}{T}\int_0^T \ln(f(t))dt\right]}{\frac{1}{T}\int_0^T f(t)dt} \qquad (3\text{-}9)$$

where $f(t)$ is defined over $[0,T]$.

An important property of G/A ratio is that it approaches unity for flat functions and zero for peaky functions. In an extreme case, $r_{G/A} = 1$ if $f(t)$ = constant and $r_{G/A} \approx 0$ when $f(t)$ is highly impulsive in nature. The G/A ratio has been used as a measure of spectral flatness in speech signal processing (Markel and Gray 1976, Kay 1979).

3.3 The smoothness index and rationale of using it for parameter selection

As mentioned earlier, wavelet transform is to be used as a bandpass filter for noise reduction. The performance of such a filter depends on the selection of scale and shape factor of the daughter wavelet. For properly selected scale and shape factor, wavelet transform will result in higher coefficient values in the time interval where the fault generated impulses are located and lower coefficients elsewhere. This way, we can magnify parts of the signal that are of our interest and suppress the other parts. In other words, it is expected that the modulus of the wavelet transform forms a more peaky function of time when proper parameters are used. This is true when the vibration signal contains fault generated impulses. As such, the smoothness index defined below, can be used as a criterion to select the parameters, namely scale s and shape factor σ. Replacing $f(t)$ in Eq. (3-9) by $\left|W_f^{s,\sigma}(u)\right|$ given in Eq. (3-6) yields

$$\eta(s,\sigma) = \frac{\exp\left[\frac{1}{T}\int_0^T \ln\left(\left|W_f^{s,\sigma}(u)\right|\right)du\right]}{\frac{1}{T}\int_0^T \left|W_f^{s,\sigma}(u)\right|du} \qquad (3\text{-}10)$$

More specifically, the best parameters s and σ are achieved if $\eta(s,\sigma)$ is minimized. Figure 3.2 summarizes the steps of the proposed method.

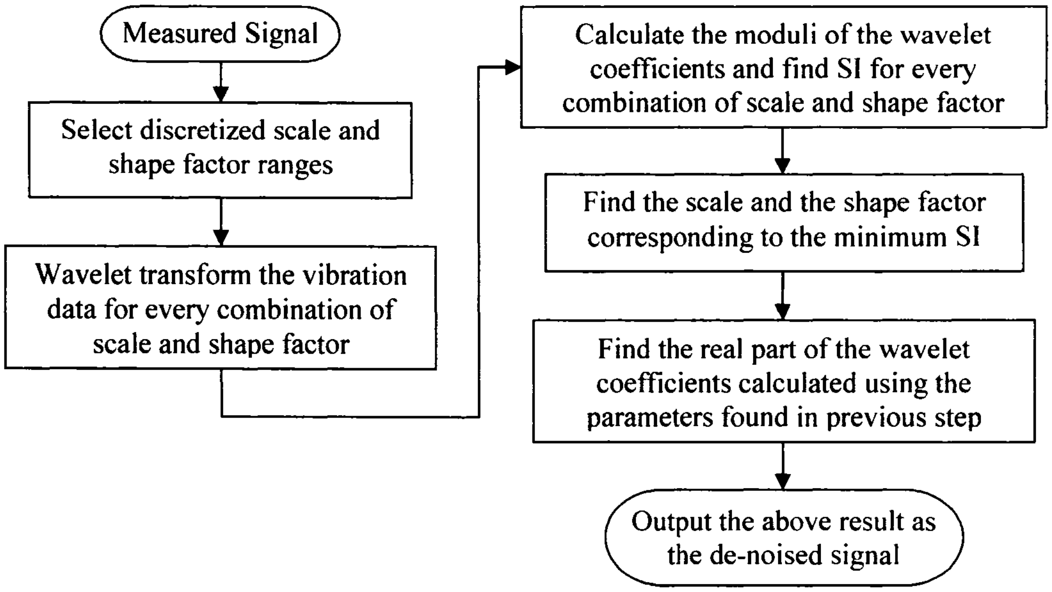


Figure 3.2 Flowchart of the proposed de-noising algorithm

3.4 Probability density function (pdf) of wavelet coefficient modulus

To assess the behavior of the smoothness index, the statistical characteristics of the wavelet coefficient modulus are required, which will be detailed in the following. The fault signature resulting from the fault generated impulse may be expressed as:

$$S(t) = A.e^{-\beta t}\cos(\omega_0 t)u(t) \tag{3-11}$$

where β is the structural damping coefficient, $u(t)$ is a unit step function and ω_0 is the excited resonance frequency. Due to the resemblance to a theoretical impulse, "fault generated impulse" or "impulse" is used interchangeably to refer to such signatures hereafter. Assuming such an impulse is generated each time when a fault comes in contact with the mating surface, we can model the measured vibration from a faulty bearing as a series of impulses with period T_p which corresponds to the characteristic fault period:

$$V(t)=\sum_{m=-M}^{M} A.e^{-\beta(t-mT_p)}\cos\omega_0(t-mT_p)u(t-mT_p)+w(t)=v(t)+w(t) \tag{3-12}$$

where $w(t)$ is continuous white Gaussian noise and $v(t)$ is the impact vibration due to the fault and contains $2M+1$ impulses. Wavelet transforming $V(t)$ using Eqs. (2-1) and (3-1), one obtains:

$$\begin{aligned} W_V^{s,\sigma}(t) &= \int_{-\infty}^{\infty} V(t')\frac{1}{\sqrt{s}}\psi^*\left(\frac{t'-t}{s}\right)dt' \\ &= \int_{-\infty}^{\infty}\sum_{n=-M}^{M} Ae^{-\beta(t'-nT_p)}\cos\omega_0(t'-nT_p)u(t'-nT_p)\frac{c}{\sqrt{s}}e^{-(\frac{\sigma}{s})^2(t'-t)^2-i\frac{2\pi}{s}(t'-t)}dt' \\ &\quad + \int_{-\infty}^{\infty} w(t')\frac{c}{\sqrt{s}}e^{-(\frac{\sigma}{s})^2(t'-t)^2-i\frac{2\pi}{s}(t'-t)}dt' \end{aligned} \tag{3-13}$$

Substituting

$$W_v^{s,\sigma}(t)=\int_{-\infty}^{\infty}\sum_{m=-M}^{M} Ae^{-\beta(t'-mT_p)}\cos\omega_0(t'-mT_p)u(t'-mT_p)\frac{c}{\sqrt{s}}e^{-(\frac{\sigma}{s})^2(t'-t)^2-i\frac{2\pi}{s}(t'-t)}dt' \tag{3-14}$$

into Eq. (3-13) yields:

$$W_V^{s,\sigma}(t)=W_v^{s,\sigma}(t)+\frac{c}{\sqrt{s}}\int_{-\infty}^{\infty} w(t')e^{-(\frac{\sigma}{s})^2(t'-t)^2-i\frac{2\pi}{s}(t'-t)}dt' \tag{3-15}$$

where $W_v^{s,\sigma}(t)$ is the wavelet transform in the absence of noise. Decomposing the noise related component into real and imaginary parts leads to:

$$\begin{aligned} W_w^{s,\sigma}(t) &= \frac{c}{\sqrt{s}}\int_{-\infty}^{\infty} w(t')e^{-(\frac{\sigma}{s})^2(t'-t)^2-i\frac{2\pi}{s}(t'-t)}dt' = \int_{-\infty}^{\infty} w(t')e^{-(\frac{\sigma}{s})^2(t'-t)^2}\frac{c}{\sqrt{s}}\cos\frac{2\pi}{s}(t'-t)dt' \\ &\quad - i\int_{-\infty}^{\infty} w(t')e^{-(\frac{\sigma}{s})^2(t'-t)^2}\frac{c}{\sqrt{s}}\sin\frac{2\pi}{s}(t'-t)dt' \end{aligned} \tag{3-16}$$

As both real and imaginary parts in the above equation are linear transforms of Gaussian random variables, they are Gaussian random variables with zero mean and variance given by:

$$
\begin{aligned}
\nu^2 &= \operatorname{var}\left(\operatorname{Re}\left(W_w^{s,\sigma}(t)\right)\right) \\
&= \operatorname{var}\left(\int_{-\infty}^{\infty} w(t')\frac{c}{\sqrt{s}}e^{-(\frac{\sigma}{s})^2(t'-t)^2}\cos\frac{2\pi}{s}(t'-t)dt'\right) \\
&= E\left\{\left[\int_{-\infty}^{\infty} w(t')\frac{c}{\sqrt{s}}e^{-(\frac{\sigma}{s})^2(t'-t)^2}\cos\frac{2\pi}{s}(t'-t)dt'\right]^2\right\} \\
&= E\left\{\left[\int_{-\infty}^{\infty} w(t')\frac{c}{\sqrt{s}}e^{-(\frac{\sigma}{s})^2(t'-t)^2}\cos\frac{2\pi}{s}(t'-t)dt'\right]\left[\int_{-\infty}^{\infty} w(t'')\frac{c}{\sqrt{s}}e^{-(\frac{\sigma}{s})^2(t''-t)^2}\cos\frac{2\pi}{s}(t''-t)dt''\right]\right\} \\
&= E\left\{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w(t')w(t'')\frac{c}{\sqrt{s}}e^{-(\frac{\sigma}{s})^2(t'-t)^2}\cos\frac{2\pi}{s}(t'-t)\frac{c}{\sqrt{s}}e^{-(\frac{\sigma}{s})^2(t''-t)^2}\cos\frac{2\pi}{s}(t''-t)dt'dt''\right\} \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E\left[w(t')w(t'')\right]\frac{c}{\sqrt{s}}e^{-(\frac{\sigma}{s})^2(t'-t)^2}\cos\frac{2\pi}{s}(t'-t)\frac{c}{\sqrt{s}}e^{-(\frac{\sigma}{s})^2(t''-t)^2}\cos\frac{2\pi}{s}(t''-t)dt'dt''
\end{aligned}
$$

where E is the statistical expectation operator. In the above equation, if $t' \neq t''$, $E[(w(t')w(t'')] = 0$ ($w(t')$ and $w(t'')$ are independent). For $t' = t''$, ν^2 is given as follows:

$$
\begin{aligned}
&\nu^2 = \operatorname{var}\left(\operatorname{Re}\left(W_w^{s,\sigma}(t)\right)\right) = \int_{-\infty}^{\infty} E\{w(t')^2\}[\frac{c}{\sqrt{s}}e^{-(\frac{\sigma}{s})^2(t'-t)^2}\cos\frac{2\pi}{s}(t'-t)]^2 dt' \approx \\
&\operatorname{var}(w(t))\int_{-\infty}^{\infty}[\frac{c}{\sqrt{s}}e^{-(\frac{\sigma}{s})^2(t'-t)^2}\cos\frac{2\pi}{s}(t'-t)]^2 dt'
\end{aligned}
\qquad \text{(3-17a)}
$$

On the other hand, according to Eq. (3-2b),

$$
\int_{-\infty}^{\infty}[\frac{c}{\sqrt{s}}e^{-(\frac{\sigma}{s})^2(t'-t)^2}\cos\frac{2\pi}{s}(t'-t)]^2 dt' \approx \int_{-\infty}^{\infty}[\frac{c}{\sqrt{s}}e^{-(\frac{\sigma}{s})^2(t'-t)^2}\sin\frac{2\pi}{s}(t'-t)]^2 dt' \approx 1
$$

So one can write:

$$
\begin{aligned}
\nu^2 &= \operatorname{var}\left(\operatorname{Re}\left(W_w^{s,\sigma}(t)\right)\right) = \operatorname{var}(w(t))\int_{-\infty}^{\infty}[\frac{c}{\sqrt{s}}e^{-(\frac{\sigma}{s})^2(t'-t)^2}\cos\frac{2\pi}{s}(t'-t)]^2 dt' \\
&\approx \operatorname{var}\left(\operatorname{Im}\left(W_w^{s,\sigma}(t)\right)\right) = \operatorname{var}(w(t))\int_{-\infty}^{\infty}[\frac{c}{\sqrt{s}}e^{-(\frac{\sigma}{s})^2(t'-t)^2}\sin\frac{2\pi}{s}(t'-t)]^2 dt' \approx \operatorname{var}(w(t))
\end{aligned}
\qquad \text{(3-17b)}
$$

According to Eqs. (3-15) and (3-17b), $W_V^{s,\sigma}(t)$ is a complex Gaussian random variable, and the real and imaginary parts of this random variable have means $\operatorname{Re}(W_v^{s,\sigma}(t))$ and $\operatorname{Im}(W_v^{s,\sigma}(t))$ respectively and variance equal to the background noise variance.

In order to form the envelope of the wavelet transformed signal, we denote $X(t)=\mathrm{Re}(W_V^{s,\sigma}(t))$ and $Y(t)=\mathrm{Im}(W_V^{s,\sigma}(t))$ and obtain:

$$Z(t)=\left|W_V^{s,\sigma}(t)\right|=\sqrt{X(t)^2+Y(t)^2} \tag{3-18}$$

Denoting the joint probability density function of $X(t)$ and $Y(t)$ by $f_{X,Y}(x,y)$, the cumulative probability distribution of $Z(t)$, $F_Z(z)$, can be written as (Papoulis and Pillai 2002):

$$F_{Z(t)}(z)=P(Z\leq z)=\int_{y=-z}^{z}\int_{x=-\sqrt{z^2-y^2}}^{\sqrt{z^2-y^2}} f_{X,Y}(x,y)dxdy$$

The Leibnitz differentiation rule is used to find the pdf of the wavelet coefficient modulus as follows:

$$\frac{\partial}{\partial z}\int_{a(z)}^{b(z)} f(x,z)dx=\int_{a(z)}^{b(z)}\frac{\partial f}{\partial z}dx+f(b(z),z)\frac{\partial b}{\partial z}-f(a(z),z)\frac{\partial a}{\partial z}$$

Letting $g(y,z)=\int_{x=-\sqrt{z^2-y^2}}^{\sqrt{z^2-y^2}} f_{X,Y}(x,y)dx$, one has:

$$f_Z(z)=\frac{\partial}{\partial z}F_Z(z)=\frac{\partial}{\partial z}\int_{y=-z}^{z} g(y,z)dy=\int_{-z}^{z}\frac{\partial}{\partial z}g(y,z)dy+g(z,z)-g(-z,z)=\int_{-z}^{z}\frac{\partial}{\partial z}g(y,z)dy$$

and

$$\frac{\partial}{\partial z}g(y,z)=\frac{\partial}{\partial z}\int_{x=-\sqrt{z^2-y^2}}^{\sqrt{z^2-y^2}} f_{X,Y}(x,y)dx=\frac{z}{\sqrt{z^2-y^2}}\left\{f_{X,Y}(\sqrt{z^2-y^2},y)+f_{X,Y}(-\sqrt{z^2-y^2},y)\right\}$$

so that

$$f_Z(z)=\int_{-z}^{z}\frac{z}{\sqrt{z^2-y^2}}\left\{f_{X,Y}(\sqrt{z^2-y^2},y)+f_{X,Y}(-\sqrt{z^2-y^2},y)\right\}dy \tag{3-19}$$

According to Eqs. (3-15) and (3-16), $X(t)$ and $Y(t)$ are affine transformations of independent Gaussian random variables. As a result, linear combinations of the two or $aX(t)+bY(t)$ are Gaussian random variables for all a and b. Consequently one can consider the joint normality for $X(t)$ and $Y(t)$ (Papoulis and Pillai 2002). For this purpose, $C_{X,Y}$ is defined as the covariance of $X(t)$ and $Y(t)$, i.e.,

$$C_{XY} = \frac{c^2}{s}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E\{w(t')w(t'')\} e^{-(\frac{\sigma}{s})^2(t'-t)^2} e^{-(\frac{\sigma}{s})^2(t''-t)^2} \cos\frac{2\pi}{s}(t'-t)\sin\frac{2\pi}{s}(t''-t)dt'dt'' = 0 \tag{3-20}$$

Therefore

$$f_{X,Y}(x,y) = \frac{1}{2\pi v^2} e^{\frac{-\left[\left(x-\mathrm{Re}\left[W_v^{s,\sigma}(t)\right]\right)^2+\left(y-\mathrm{Im}\left[W_v^{s,\sigma}(t)\right]\right)^2\right]}{2v^2}} \tag{3-21}$$

Now letting (Papoulis and Pillai 2002)

$$\mu(t) = \sqrt{\mathrm{Re}(W_v^{s,\sigma}(t))^2 + \mathrm{Im}(W_v^{s,\sigma}(t))^2} \tag{3-22a}$$

$$y(t) = z(t)\sin\theta \tag{3-22b}$$

$$\mathrm{Re}(W_v^{s,\sigma}(t)) = \mu(t)\cos\phi \tag{3-22c}$$

$$\mathrm{Im}(W_v^{s,\sigma}(t)) = \mu(t)\sin\phi \tag{3-22d}$$

and substituting Eqs. (3-21) and (3-22) into Eq. (3-19), one obtains

$$\begin{aligned}
f_Z(z) &= \frac{1}{2\pi v^2}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{z}{z\cos\theta}\left\{ e^{\frac{-[(z\cos\theta-\mu(t)\cos\phi)^2+(z\sin\theta-\mu(t)\sin\phi]^2}{2v^2}} + e^{\frac{-[(-z\cos\theta-\mu(t)\cos\phi)^2+(z\sin\theta-\mu(t)\sin\phi]^2}{2v^2}} \right\} z\cos\theta d\theta \\
&= \frac{ze^{-\frac{z^2+\mu(t)^2}{2v^2}}}{2\pi v^2}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (e^{z\mu(t)\cos(\theta-\phi)/v^2} + e^{-z\mu(t)\cos(\theta+\phi)/v^2})d\theta \\
&= \frac{ze^{-\frac{z^2+\mu(t)^2}{2v^2}}}{2\pi v^2}(\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{z\mu(t)\cos(\theta-\phi)/v^2}d\theta + \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} e^{z\mu(t)\cos(\theta-\phi)/v^2}d\theta) \\
&= \frac{ze^{-\frac{z^2+\mu(t)^2}{2v^2}}}{2\pi v^2}\int_0^{2\pi} e^{z\mu(t)\cos(\theta-\phi)/v^2}d\theta = \frac{ze^{-\frac{z^2+\mu(t)^2}{2v^2}}}{\pi v^2}\int_0^{\pi} e^{z\mu(t)\cos(\theta)/v^2}d\theta
\end{aligned}$$

i.e.,

$$f_Z(z) = \frac{ze^{-(z^2+\mu(t)^2)/2v^2}}{v^2} I_0(\frac{z\mu(t)}{v^2}) \tag{3-23}$$

where I_0 is the modified Bessel function of the first kind and zeroth order (Abramowitz and Stegun 1972) defined as:

$$I_0(\kappa) = \frac{1}{\pi}\int_0^{\pi} e^{\kappa\cos(\theta)}d\theta$$

This is the expression for a Rician distributed random variable. Figure 3.3 shows this distribution for different values of v^2 and $\mu(t)$. Letting $H(t)=\frac{Z(t)}{v}$ and $a(t)=\frac{\mu(t)}{v}$ and denoting the pdf of the random variable $H(t)$ by $f_H(h)$, Eq. (3-23) becomes (Rice 1945):

$$f_H(h)=he^{-(\frac{h^2+a(t)^2}{2})}I_0(a(t)h) \tag{3-24}$$

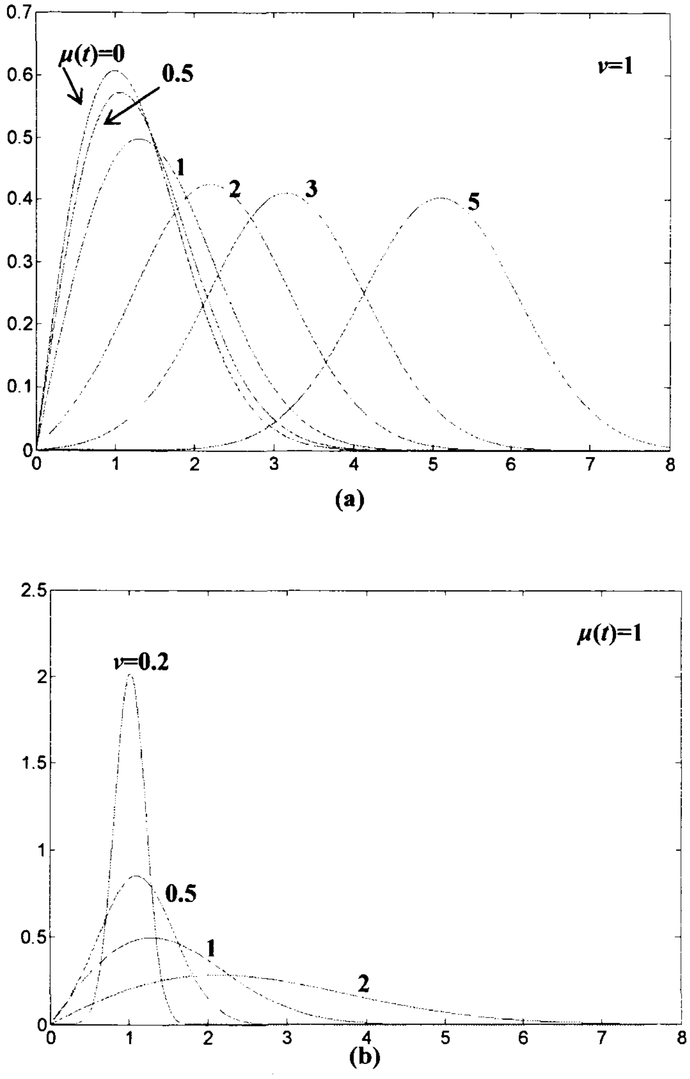


Figure 3.3 Rician distribution for (a) $v=1$ and different $\mu(t)$, (b) $\mu(t)=1$ and different v

3.5 Behavior of SI for different values of local SNR (*a*(*t*))

As shown in Eq. (3-14), $W_v^{s,\sigma}(t)$ is a periodic function of time with period T_p. Hence, for a large M, the SI can be written as:

$$\eta(s,\sigma)=\frac{\exp\left[\int_{-MTp}^{(M+1)Tp}\frac{1}{(2M+1)T_p}\ln\left(\left|W_v^{s,\sigma}(t)\right|\right)dt\right]}{\frac{1}{(2M+1)T_p}\int_{-MT_p}^{(M+1)T_p}\left|W_v^{s,\sigma}(t)\right|dt}\approx\frac{\exp\left[\int_0^{T_p}\frac{1}{T_p}E\{\ln(Z(t))\}dt\right]}{\frac{1}{T_p}\int_0^{T_p}E\{Z(t)\}dt}\qquad(3\text{-}25)$$

Replacing $Z(t)$ with $H(t)v$ in Eq. (3-25) leads to

$$\eta(s,\sigma)\approx\frac{\exp\left[\ln v+\int_0^{T_p}\frac{1}{T_p}E\{\ln(H(t))\}dt\right]}{\frac{v}{T_p}\int_0^{T_p}E\{H(t)\}dt}=\frac{\exp\left[\int_0^{T_p}\frac{1}{T_p}E\{\ln(H(t))\}dt\right]}{\frac{1}{T_p}\int_0^{T_p}E\{H(t)\}dt}\qquad(3\text{-}26)$$

The above equation provides a general expression of SI. As mentioned earlier, the SI approaches zero for signals of impulsive nature. However, the SI behavior should be analyzed when the signal is corrupted with background noise. More detailed analysis is given in the following.

3.5.1 SI behavior under low local SNR

For very low SNR, i.e., $a(t)$<<1, the following holds (Abramowitz and Stegun 1972):

$$\lim_{a(t)h\to 0} I_0(a(t)h)=\frac{1}{\Gamma(1)}=1\qquad(3\text{-}27a)$$

From the above result, when $a(t)$ approaches zero, Eq. (3-24) can be written as

$$\lim_{a(t)\to 0} f_H(h)=he^{-h^2/2}\qquad(3\text{-}27b)$$

Equation (3-27b) is the expression of Rayleigh distribution. Considering Eq. (3-27b), Eq. (3-26) can be written as:

$$\eta(s,\sigma)=\frac{\exp[E\{\ln(H)\}]}{E\{H\}}\qquad(3\text{-}28)$$

where H is a Rayleigh distributed random variable. As $E\{\ln(H)\} = \int_0^\infty \ln(h) h e^{-\frac{h^2}{2}} dh$, the following is obtained with a change of variable $r = h^2$:

$$\begin{aligned} E\{\ln(H)\} &= \int_0^\infty \ln(\sqrt{r})\sqrt{r}\, e^{-\frac{r}{2}} \frac{1}{2\sqrt{r}} dr = \frac{1}{4}\int_0^\infty \ln(r) e^{\frac{-r}{2}} dr = \frac{1}{4}\int_0^\infty (\ln\frac{r}{2} + \ln 2) e^{-\frac{r}{2}} dr \\ &= \frac{1}{4}\left[\int_0^\infty \ln\frac{r}{2}\, e^{-\frac{r}{2}} dr + \ln 2 \int_0^\infty e^{-\frac{r}{2}} dr\right] = \frac{1}{4}\left[2\int_0^\infty \ln\frac{r}{2}\, e^{-\frac{r}{2}} d\frac{r}{2} + 2\ln 2 \int_0^\infty e^{-\frac{r}{2}} d\frac{r}{2}\right] \\ &= \frac{1}{4}\left[-2\gamma + 2\ln 2\right] = -\frac{1}{2}\gamma + \frac{1}{2}\ln 2 = -\frac{1}{2}\left(\gamma + \ln\frac{1}{2}\right) \end{aligned} \tag{3-29a}$$

where $\gamma = \lim_{s\to\infty}\left(\sum_{m=1}^{s}\frac{1}{m} - \ln(s)\right) = 0.5772156649\ldots$ is the Euler's constant.

In addition, the denominator of Eq. (3-28) can be written as (Gradshteyn and Ryzhik 2000):

$$E\{H\} = \int_0^\infty h^2 e^{-\frac{h^2}{2}} dh = \sqrt{\frac{\pi}{2}} \tag{3-29b}$$

Substituting Eqs. (3-29a) and (3-29b) into Eq. (3-28) yields:

$$\lim_{a(t)\to 0} \eta(s,\sigma) = \frac{\exp\left[-\frac{1}{2}\left(\gamma + \ln\frac{1}{2}\right)\right]}{\sqrt{\frac{\pi}{2}}} = 0.845501287\ldots \tag{3-30}$$

According to Eq.(3-30), for very low SNR or in the absence of any faults where the vibrations behave randomly, the SI approaches a constant number.

3.5.2 SI behavior under higher local SNR

In this section we analyze the expected value and the expected logarithm of a Rician distributed random variable for a larger $a(t)$. The mean of such a random variable with distribution shown in Eq. (3-23) is (Rice 1945):

$$E\{Z(t)\} = \nu\sqrt{2}\Gamma(\frac{3}{2})\,{}_1F_1(-\frac{1}{2},1,-\frac{\mu(t)^2}{2\nu^2}) \tag{3-31}$$

where ${}_1F_1$ is a confluent hypergeometric function (Abramowitz and Stegun 1972). According to (Abramowitz and Stegun 1972, Eq. 13. 1. 5), ${}_1F_1$ can be expressed as:

$${}_1F_1(-\frac{1}{2},1,-\frac{\mu(t)^2}{2\nu^2}) = \frac{\Gamma(1)}{\Gamma(3/2)}\left(\frac{\mu(t)^2}{2\nu^2}\right)^{\frac{1}{2}}\left[1+O\left(\left(\frac{\mu(t)^2}{2\nu^2}\right)^{-1}\right)\right]$$

Substituting the above equation into Eq. (3-31), we obtain:

$$E\{Z(t)\} = \nu\sqrt{2}\Gamma\left(\frac{3}{2}\right)\frac{\Gamma(1)}{\Gamma(3/2)}\left(\frac{\mu(t)^2}{2\nu^2}\right)^{\frac{1}{2}}\left[1+O\left(\left(\frac{\mu(t)^2}{2\nu^2}\right)^{-1}\right)\right] = \mu(t)\left(1+O\left(\left(\frac{\mu(t)^2}{2\nu^2}\right)^{-1}\right)\right) \tag{3-32}$$

where $O(x)$ is the order of the argument x. In the above equation, the relation $\Gamma(1)=1$ is used.

According to Eq. (3-32), as $a(t)=\frac{\mu(t)}{\nu}$ increases the mean of $Z(t)$ approaches $\mu(t)$ or the modulus of the wavelet coefficients of the noise free vibration signal.

Moreover, considering $a'(t)=\frac{a(t)}{\sqrt{2}}$ and $H'(t)=\frac{H(t)}{\sqrt{2}}$, Eq. (3-24) can be written as:

$$f_{H'}(h') = 2h'e^{-(h'^2+a'(t)^2)}I_0(2a'(t)h') \tag{3-33}$$

From the above expression we find the expected logarithm of $H'(t)$ as follows:

$$E\{\ln(H'(t))\} = \int_0^\infty 2h'\ln h' e^{-(h'^2+a'(t)^2)}I_0(2a'(t)h')dh' = 2e^{-a'(t)^2}\int_0^\infty h'\ln h' e^{-h'^2}I_0(2a'(t)h')dh' \tag{3-34}$$

On the other hand $I_0(2a'(t)h')$ can be expressed as (Abramowitz and Stegun 1972):

$$I_0(2a'(t)h') = \sum_{k=0}^{\infty}\frac{(a'(t)h')^{2k}}{k!\Gamma(k+1)} \tag{3-35}$$

Substituting Eq. (3-35) into Eq. (3-34) leads to:

$$\begin{aligned} E\{\ln(H'(t))\} &= 2e^{-a'(t)^2}\int_0^\infty h'\ln h' e^{-h'^2}\sum_{k=0}^{\infty}\frac{(a'(t)h')^{2k}}{k!\Gamma(k+1)}dh' \\ &= 2e^{-a'(t)^2}\sum_{k=0}^{\infty}\frac{a'(t)^{2k}}{k!\Gamma(k+1)}\int_0^\infty \ln h' e^{-h'^2}h'^{2k+1}dh' \end{aligned} \tag{3-36}$$

Letting $r = h'^2$ yields:

$$\int_0^\infty \ln h' e^{-h'^2} h'^{2k+1}\, dh' = \int_0^\infty \frac{1}{4} \ln r e^{-r} r^k \sqrt{r} \frac{1}{\sqrt{r}}\, dr = \frac{1}{4} \int_0^\infty \ln r e^{-r} r^k\, dr$$

The following can be obtained from Eq. (4.352.1) in (Gradshteyn and Ryzhik 2000):

$$\frac{1}{4} \int_0^\infty \ln r e^{-r} r^k\, dr = \frac{1}{4} \Gamma(k+1) \Psi(k+1) \qquad (3\text{-}37)$$

where $\Psi(n)$ is the Euler's psi function defined as $\Psi(n) = -\gamma + \sum_{k=1}^{n-1} k^{-1}$ and γ as mentioned before is the Euler's constant. Replacing the integral of Eq. (3-36) with Eq. (3-37) leads to:

$$E\{\ln(H'(t))\} = 2e^{-a'(t)^2} \sum_{k=0}^{\infty} \frac{a'(t)^{2k}}{k!\Gamma(k+1)} \frac{1}{4} \Gamma(k+1)\Psi(k+1) = \frac{1}{2} e^{-a'(t)^2} \sum_{k=0}^{\infty} \frac{a'(t)^{2k} \Psi(k+1)}{k!} \qquad (3\text{-}38)$$

It is proven in (Lapidoth and Moser 2003) that:

$$e^{-a'(t)^2} \sum_{k=0}^{\infty} \frac{a'(t)^{2k} \Psi(k+1)}{k!} = \ln(a'(t)^2) - Ei(-a'(t)^2) \qquad (3\text{-}39)$$

where $Ei(x)$ is the exponential integral, defined as:

$$Ei(x) = -\int_{-x}^{\infty} \frac{e^{-t}}{t}\, dt$$

Substituting Eq. (3-39) in Eq. (3-38) results in:

$$E\{\ln(H'(t))\} = \frac{1}{2} \ln(a'(t)^2) - \frac{1}{2} Ei(-a'(t)^2) \qquad (3\text{-}40)$$

As $a'(t)^2$ increases, the exponential integral in Eq. (3-40) approaches zero. Considering the relations $a'(t) = \frac{\mu(t)}{\nu\sqrt{2}}$ and $H'(t) = \frac{Z(t)}{\nu\sqrt{2}}$, one obtains:

$$E\left\{\ln(\frac{Z(t)}{\nu\sqrt{2}})\right\} = \ln(\frac{\mu(t)}{\nu\sqrt{2}}) - \frac{1}{2} Ei\left(-\left(\frac{\mu(t)}{\nu\sqrt{2}}\right)^2\right) \Rightarrow E\{\ln(Z(t))\} = \ln(\mu(t)) - \frac{1}{2} Ei\left(-\left(\frac{\mu(t)}{\nu\sqrt{2}}\right)^2\right)$$

(3-41)

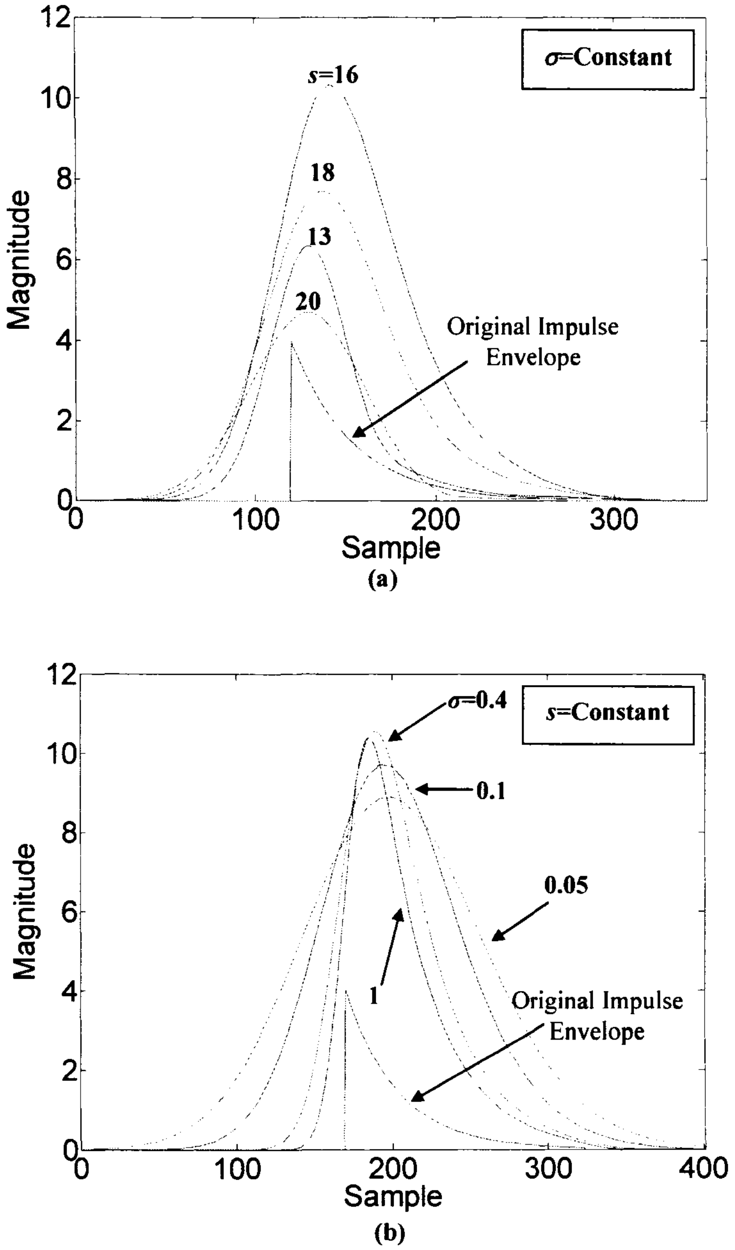
According to the above equation, as the local SNR or $a(t)$ becomes larger, the expected logarithm of $Z(t)$ approaches $\ln(\mu(t))$. Substituting Eqs. (3-32) and (3-41) in Eq. (3-25) yields:

$$\eta(s,\sigma)=\frac{\exp\left[\frac{1}{T_p}\int_0^{T_p}\ln\left(\mu(t)\right)-\frac{1}{2}Ei\left(-\left(\frac{\mu(t)}{\nu\sqrt{2}}\right)^2\right)dt\right]}{\frac{1}{T_p}\int_0^{T_p}\mu(t)\left(1+O\left(\left(\frac{\mu(t)}{\nu\sqrt{2}}\right)^{-2}\right)\right)dt} \qquad (3\text{-}42)$$

Though not very accurately but Equation (3-42) would provide us with some sense of the degree of sensitivity of the smoothness index to the background noise intensity. According to this equation, as $a(t)=\frac{\mu(t)}{\nu}$ increases, SI approaches its noise free environment (pure impulsive vibration) level which is closer to zero due to the impulsive nature of the faulty bearing vibration. This is also clear intuitively.

3.6 Effect of wavelet parameters on smoothness index

Figure 3.4a shows the wavelet coefficient moduli associated with different scales for a constant shape factor. These moduli are calculated for a single noise free simulated impulse. The scale corresponding to resonance frequency is 16. As one can see, the magnitude of the wavelet coefficient modulus at scale 16 is the highest of all the scales for all samples. Since noise variance ν^2 remains constant as shown in Eq. (3-17b), the local SNR, $a(t)=\frac{\mu(t)}{\nu}$ increases and according to Eq. (3-42) the SI approaches its noise-free environment level. This was also expected intuitively -- it is well known that the resonance frequency band corresponds to the high SNR frequency region.



(a)

(b)

Figure 3.4 (a) Wavelet coefficient modulus at different scales with fixed shape factor σ, (b) Wavelet coefficient modulus for different shape factors at the scale associated with the resonance frequency

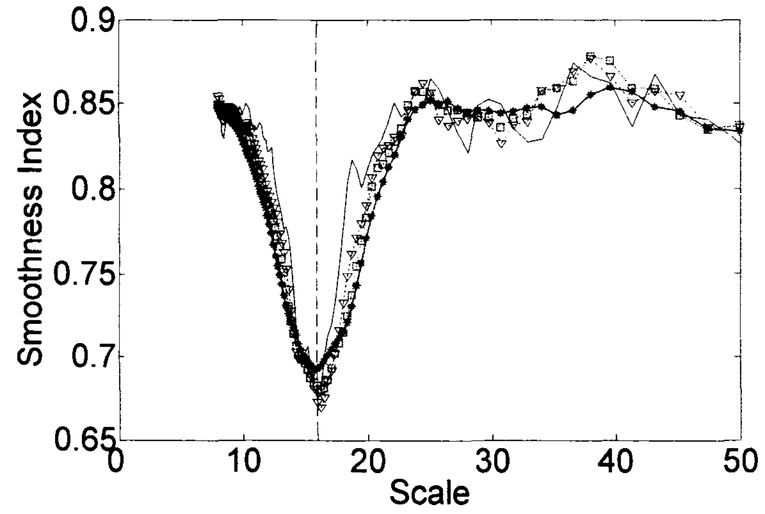
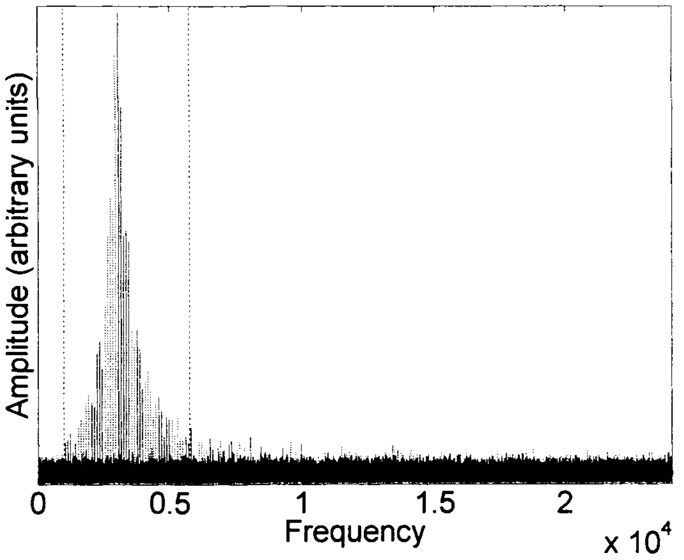


Figure 3.5 SI of the wavelet coefficient modulus versus scale for different shape factors calculated based on the simulated signal (Figure 3.8a) with noise added (scale corresponding to the resonance frequency is shown by the vertical dashed line) (— σ^2 =0.07, ∇ σ^2 =0.19, □ $\sigma^2 = 0.31$, * $\sigma^2 = 0.61$)

As shown in Figure 3.5, the lowest SI appears at a scale very close to the scale associated with resonance frequency which is in agreement with the above explanations. The effect of shape factor σ can also be explained in a similar manner. Figure 3.4b shows the modulus of the wavelet coefficients at the scale associated with resonance frequency for different shape factor values. As one can see, by decreasing the shape factor, the magnitude of the wavelet coefficient modulus increases in most of sampling points whereas again according to Eq. (3-17b) the noise variance remains constant. As explained earlier, this in fact results in a higher local SNR. To illustrate, we consider Figure 3.6. Figure 3.6a and Figure 3.6b show the power spectral density (PSD) of the simulated signal of Figure 3.8a with different noise intensity levels. Figure 3.6 shows that increasing noise leads to a narrower high SNR band in the frequency domain. This means that a smaller shape factor should be chosen to make the bandwidth of the filter narrower. However, wavelet transform is the weighted averaging and lowering the shape factor σ increases this averaging effect which results in a flatter signal with a higher SNR. This phenomenon can also be seen in Figure 3.4b. Besides, this smoothing effect causes the neighboring impulses to alias. It should also be noted that the decrease in σ has decreasing or advert influence on the wavelet coefficients through the constant *c* given in Eq. (3-2a). This effect is stronger for smaller shape factors.

Considering an extreme case where the bandwidth of the filter is extremely narrow (very small shape factor) and only the center frequency could pass, the corresponding envelope signal would be a flat function. On the other hand, for very large bandwidth (large shape factor) all the frequency components can pass through the filter and no noise reduction could take place. Hence the proper parameter is obtained when a desirable balance between the increase of SNR and flatter time signal is achieved so that the fault generated impulses can be better identified from the result.



(a)

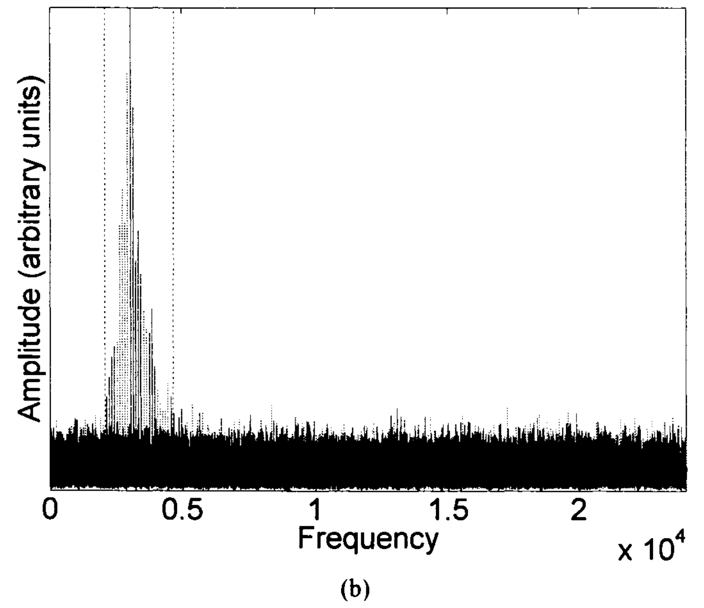


Figure 3.6 PSD of the simulated signal from Figure 3.8a: (a) SNR of -7 dB, (b) SNR of -13 dB. The frequencies between the vertical dashed lines indicate the approximate high SNR band

3.7 Comparison of proposed SI minimization method with Kurtosis maximization criterion

The proposed Smoothness Index quantifies the impulsiveness of the bandpass filtered signal. A similar criterion is the Kurtosis value defined below:

$$Kurt(S) = \frac{\frac{1}{N}\sum_{n=1}^{N}(S(n)-\mu)^4}{\left(\frac{1}{N}\sum_{n=1}^{N}(S(n)-\mu)^2\right)^2} - 3$$

where $S(n)$ is a time series and μ is the mean value. This index is a measure of non-Gaussianity of a time series and has higher values for signals with higher level of non-Gaussian characteristics. As the time signals containing impulsive features are believed to form a non-Gaussian data set, this index has been widely used in fault diagnosis research also for wavelet parameter selection. However, as explained before this criterion is very sensitive to the outliers in the data. This sensitivity may lead the parameter selection algorithm to wrong frequency bands of the measured vibration and consequently results in poor de-noising performance. Such a drawback of the Kurtosis maximization criterion is more clearly shown in an example. Figure 3.7 shows an interval of simulated white Gaussian noise. This data set also contains two large outliers, deliberately added to the data. As proved above, the ratio of the geometric mean to the arithmetic mean calculated for the envelope of this simulated white noise should be very close to the constant 0.8455. On the other hand the kurtosis value calculated for the same envelope should be close to zero. According to the results, as expected the $r_{G/A}$ calculated for the envelope signal was found as 0.8453 while the kurtosis value was 9.194. The same indices were recalculated for the envelope of the original simulated noise with no outliers in the data, $r_{G/A}$ was found as 0.8455 and Kurtosis value reduced to 0.252. This result shows the considerable sensitivity of Kurtosis to the outliers in the data which may be the result of noise or measuring flaws often encountered in practice. From this observation, it is reasonable to believe that the proposed Smoothness Index (SI) is a more robust criterion for fault detection and parameter selection.

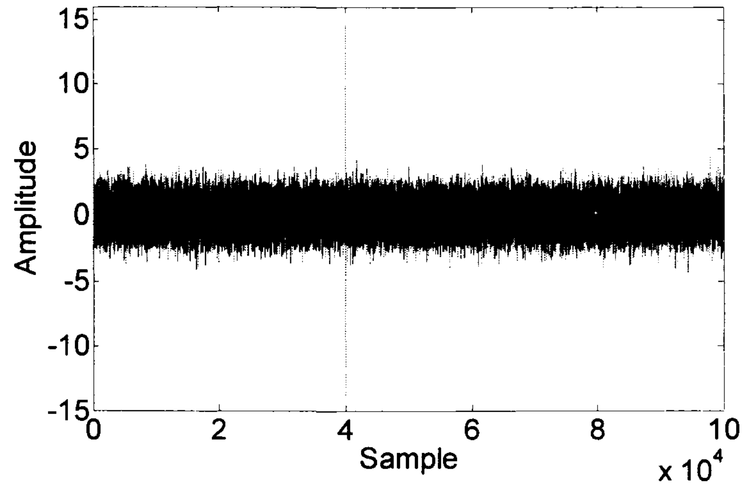
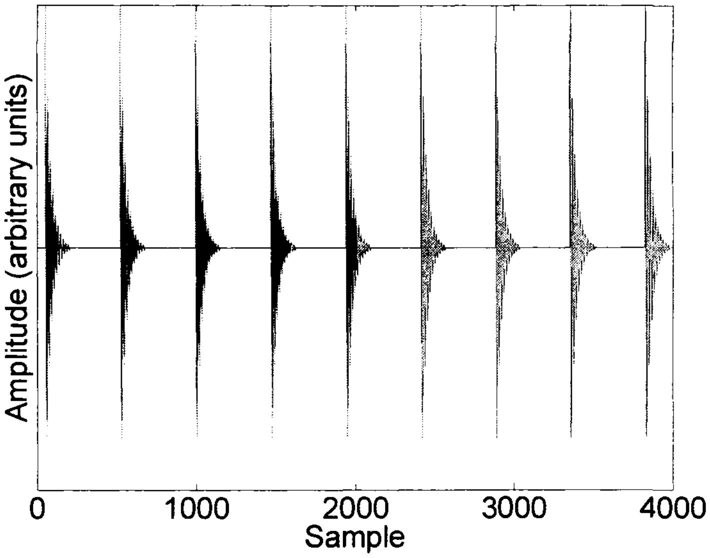


Figure 3.7 Simulated white Gaussian noise with two large outliers included in the signal

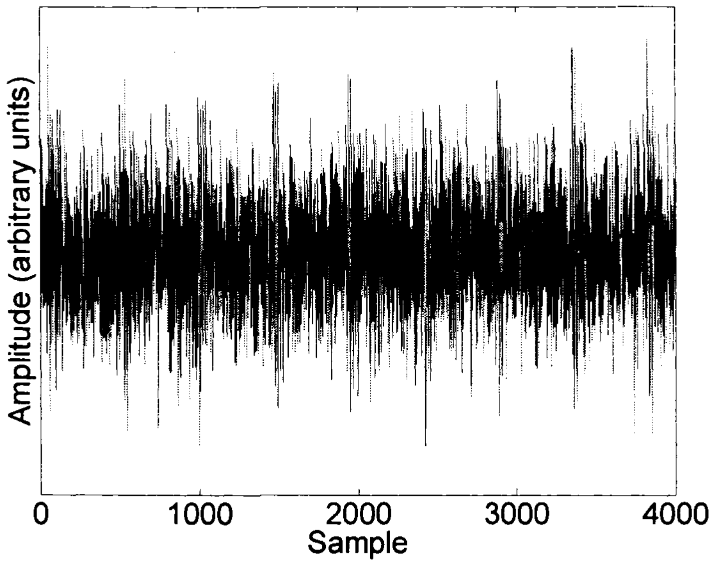
3.8 Performance assessment of the proposed index using the simulated data

3.8.1 Simulated signal and noise

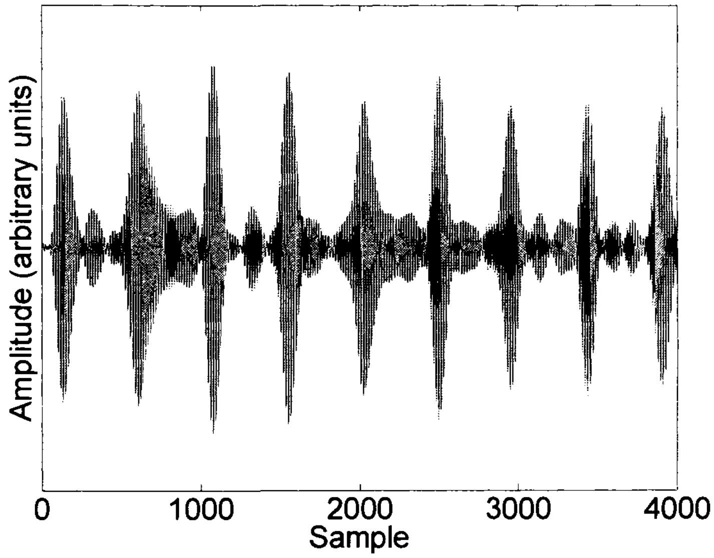
A considerable amount of noise is added to the simulated signal shown in Figure 3.8a. The resulting signal is shown in Figure 3.8b. This signal is then wavelet transformed using Gabor wavelet at the scales corresponding to the range $0.04F_{Nyq} \leq f \leq 0.25F_{Nyq}$. F_{Nyq} is one half of the sampling frequency (48000 Hz). This range is discretized with the step of 50 Hz. The squared shape factor is also considered over the range $0.01 \leq \sigma^2 \leq 1$ with the step of 0.01. The real part of the wavelet coefficients corresponding to the minimum SI achieved for scale s =15.9468 and σ^2= 0.1 is presented as the de-noised version of the simulated signal in Figure 3.8c. As expected, the best scale corresponds to the frequency of 3010 Hz which is very close to the resonance frequency of the simulated signal (3000 Hz). Figure 3.8d illustrates the SI found for different s-σ^2 combinations.



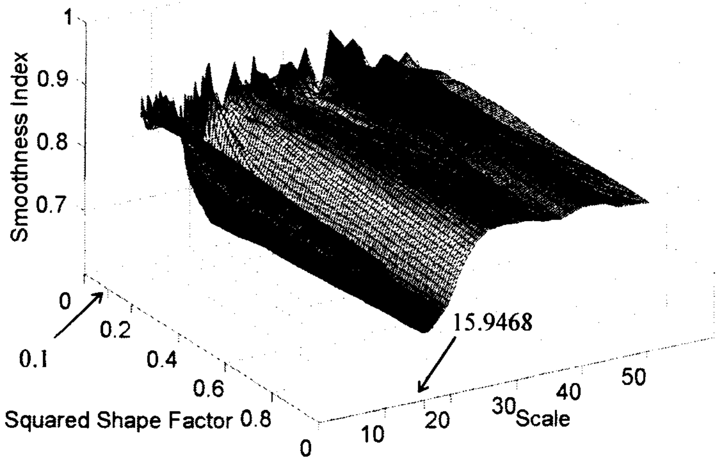
(a)


Amplitude (arbitrary units)
0
1000
2000
3000
4000
Sample

(b)


Amplitude (arbitrary units)
0
1000
2000
3000
4000
Sample

(c)

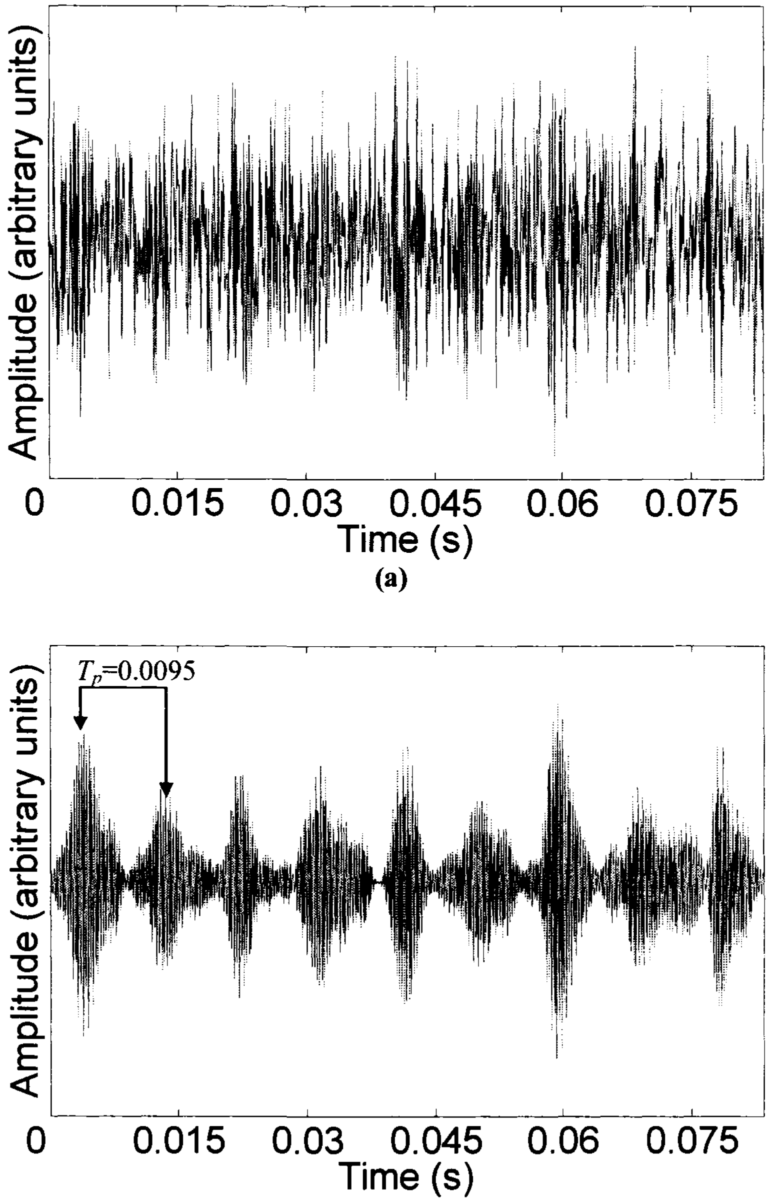


(d)

Figure 3.8 (a) Simulated faulty bearing signal at 48000 Hz sampling rate, 1739.4 rpm rotational speed, fault characteristic frequency = 3.5×rotational speed (Hz) and resonance frequency of 3000 Hz corresponding to scale 16, (b) Simulated signal with noise added, the resulting noisy signal has an SNR of -7, (c) De-noised signal using the scale and the shape factor corresponding to the minimum SI, (d) SI for different s-σ^2 combinations

3.8.2 Real bearing signal with additional simulated noise

To assess the performance of the proposed method, we took real bearing signals sampled at 12000Hz (with faults on the outer and inner races) from (Case Western Reserve University bearing data center, 2006) and increased the noise levels. The signal-noise mixtures are shown in Figure 3.9a and Figure 3.10a. Their de-noised results are plotted in Figure 3.9b and Figure 3.10b, respectively. The fault generated impulses can be clearly identified in the de-noised signal and the time interval between two consecutive impulses matches with the characteristic fault frequencies of 107.30 Hz and 162.18 Hz for the outer and inner race faults respectively. Figure 3.11a and Figure 3.11b show the SI values obtained for all combinations of scale-shape factor.



(a)

(b)

Figure 3.9 (a) Measured vibration from a bearing with outer race fault (Case Western Reserve University bearing data center, 2006) with added noise, (b) De-noised version of the same signal using the parameters corresponding to minimum SI (Sampling rate: 12000 samples/s)

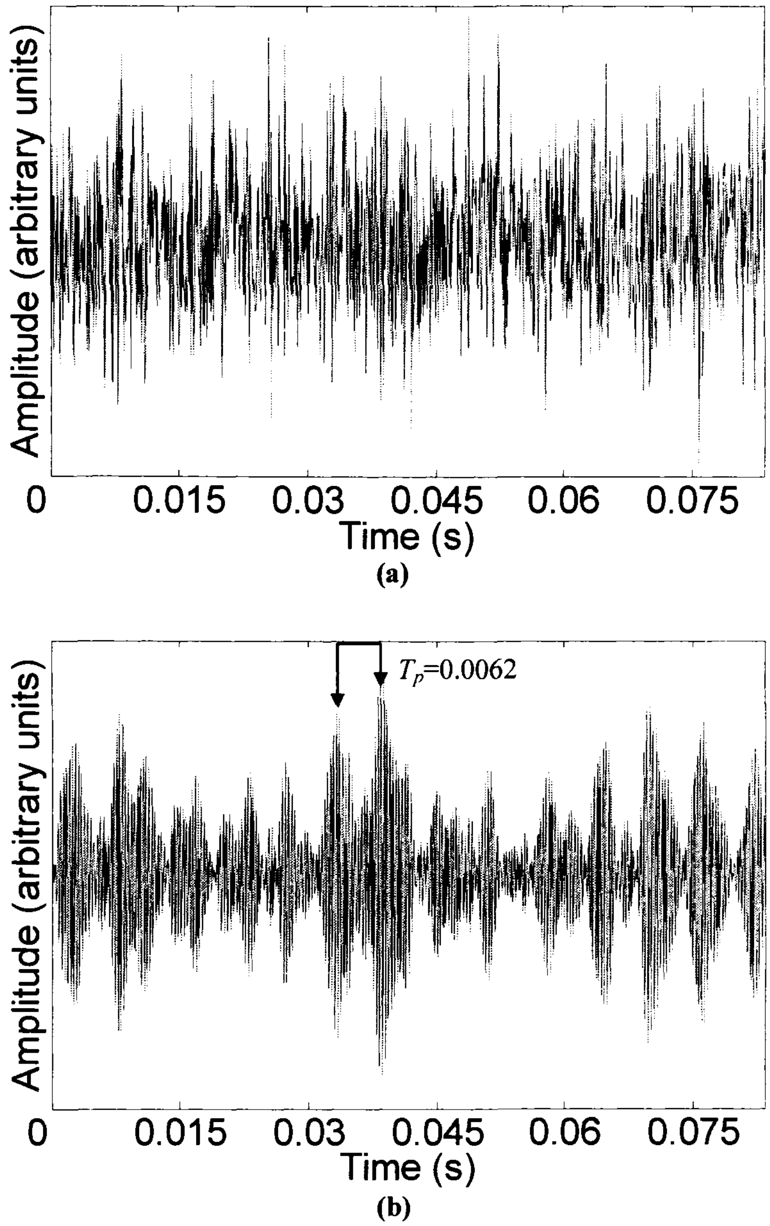
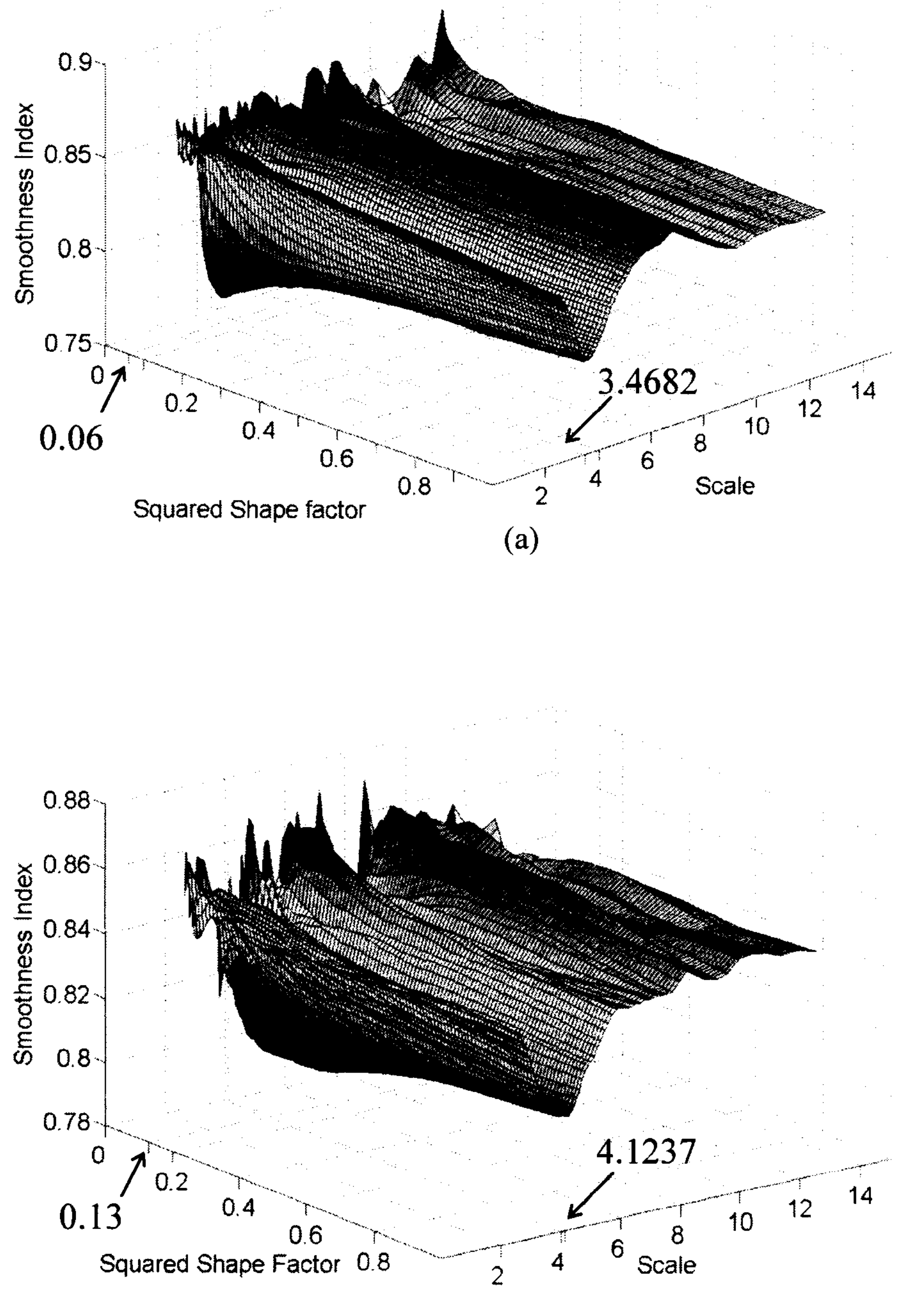


Figure 3.10 (a) Measured vibration from a bearing with inner race fault (Case Western Reserve University bearing data center, 2006) with added noise, (b) De-noised version of the same signal using the parameters corresponding to minimum SI (Sampling rate: 12000 samples/s)



(a)

(b)

Figure 3.11 SI of the vibration signals shown in (a) Figure 3.9a and (b) Figure 3.10a for different s-σ^2 combinations

3.9 Experimental Evaluation

The proposed method was further evaluated using the vibration data measured in our lab. The experiment was carried out using a SpectraQuest Machinery Fault Simulator (MFK-PK5M) as shown in Figure 3.12. Two bearings (type ER10K) were used to support a two-mass rotor test kit. The two mass rotors were well balanced and are 2" thick, 4" in diameter and 11.1 lbs each. The rotors were fitted into a 5/8" steel shaft to provide the radial load. The simulator was driven by a 3-hp AC motor with a Hitachi controller (SJ200-022NFU). The shaft speed was set at 1326 RPM (22.1Hz). The left bearing has a pre-seeded single fault on the outer race (generated by SpectraQuest) with a characteristic frequency of 67 Hz (= $3.052f_r$). A Montronix model VS100-100 accelerometer with 100 mV/g sensitivity and 1-12 kHz sensitivity range was used to measure the vibration signal. The signal was fed to an NI AT-MIO-16DE-10 DAQ card and then collected through LabVIEW. The signal processing was done using MATLAB on a Pentium® 4/2.52GHz PC.

Considering the fact that it may not be realistic to install an accelerometer in the near vicinity of a bearing in an industrial setting, we mounted the accelerometer at a spot on the simulator base that is away from the faulty bearing (Figure 3.12). In addition to the rotor load, we also connected a gearbox to the driving shaft using a belt shown on the figure to generate more interference. The vibration data were acquired at 20000 samples/sec. A portion of the measured data is plotted in Figure 3.13a and is de-noised using the proposed method. To find the best (s-σ^2) combination that yields the minimum SI, we searched the scale range corresponding to the frequency interval [200, 6000] Hz and the σ^2 range from 0.01 to 1. The search step lengths were 50 Hz for frequency and 0.01 for σ^2. The best (s-σ^2) combination was (23.53, 0.24) corresponding to the minimum SI of 0.7467. The de-noising result associated with this (s-σ^2) combination is displayed in Figure 3.13b. As shown in the figure, the time interval between two consecutive impulses is about 0.015 seconds which precisely reflects the fault characteristic frequency 67 Hz as mentioned above.

For comparison, the scale and shape factor values were also selected by maximizing the Kurtosis value. The (s-σ^2) combination associated with the maximum Kurtosis (8.8151) was (6.55, 0.5). The de-noising result is shown in Figure 3.13c. While some impulsive features can be seen from the result, no periodic component corresponding to the fault characteristic frequency of the bearing can be detected.

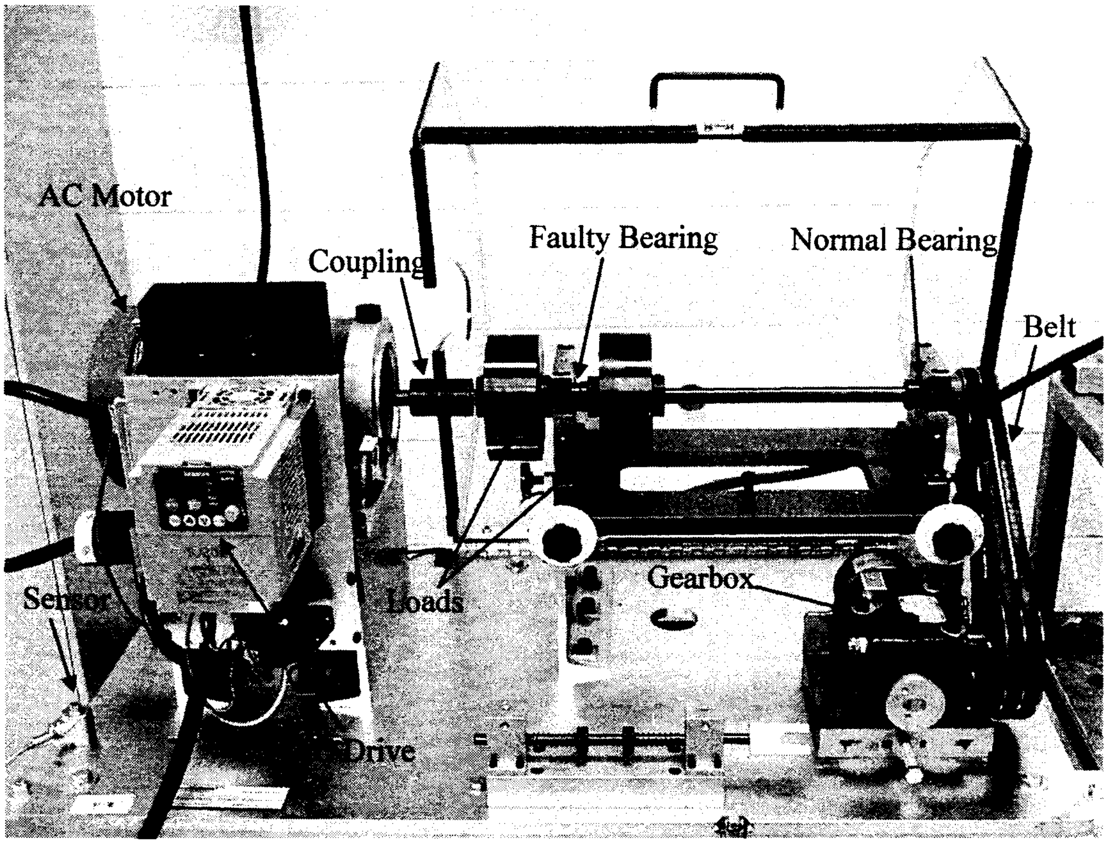
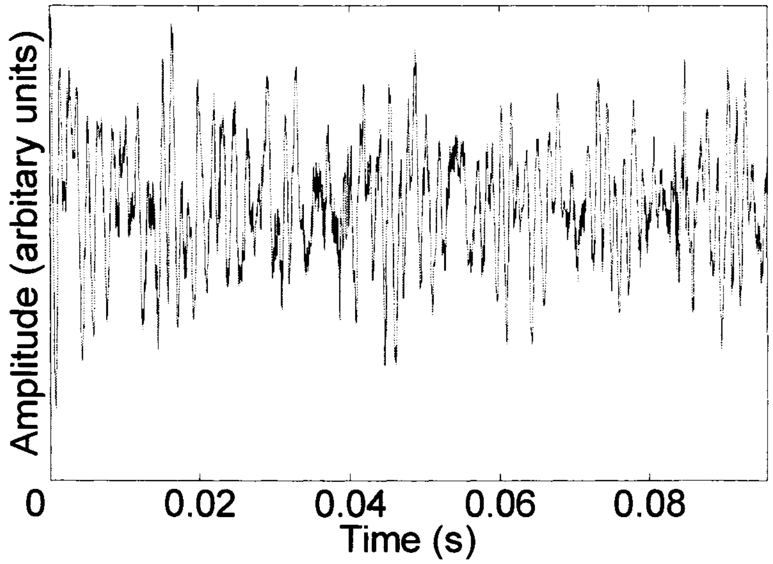
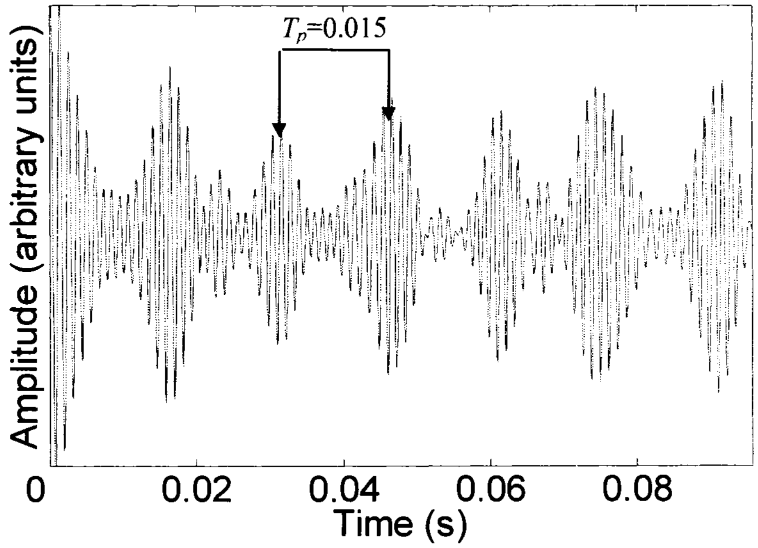


Figure 3.12 Experimental Setup


Amplitude (arbitrary units)
0
0.02
0.04
0.06
0.08
Time (s)

(a)


T_p=0.015
Amplitude (arbitrary units)
0
0.02
0.04
0.06
0.08
Time (s)

(b)

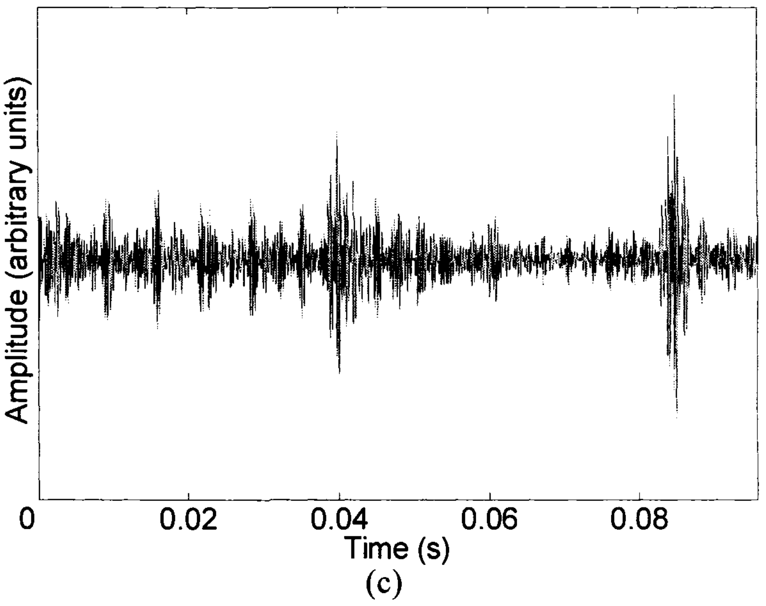


Figure 3.13 (a) Measured vibration signal, (b) de-noising result using SI minimization, (c) de-noising result based on Kurtosis maximization (Sampling rate: 20000 samples/s)

3.9 Conclusion

A SI-guided search approach has been proposed to find the best combination of wavelet scale and shape factor and hence the best Gabor daughter wavelet for de-noising impulsive signals. In this study, it is proven that the SI value approaches 0.8455 for data without impulse components or with extremely low SNR. It is also shown that the SI value decreases if a proper combination of wavelet scale and shape factor is selected. The proposed method has been tested using both simulated and experimental data and performed very well. However, the proposed de-noising method has certain drawbacks. These drawbacks are associated with the proposed wavelet parameter selection criteria in particular and the wavelet filter based de-noising approach in general. Related issues are addressed in the next chapter.

Chapter 4: Scale Selection and Spectral Subtraction

In the previous chapter a novel wavelet parameter selection method was proposed. This method performed very well as shown by the tests using both simulated and experimental data. However, there are some limitations that may adversely affect the de-noising results. These limitations are either due to the wavelet parameter selection criteria or inherited from the wavelet filter based scheme as a de-noising approach. These are discussed in the following sections and proper solutions are proposed. The final de-noising algorithm is presented thereafter in this chapter.

4.1 Drawbacks of the proposed de-noising algorithm

In chapter 3, a smoothness index (SI) was used to quantify the smoothness level for the envelope of the bandpass filtered signal. The SI value was treated as a measure of performance of the de-noising algorithm. Through the SI minimization process, proper center frequency and bandwidth of a bandpass filter were found.

In this approach the term 'proper' refers to the parameters leading to a time signal with the highest impulsive nature which presumably corresponds to the fault generated impulses. This viewpoint may not be valid while there might exist other signal components with the same featuring characteristic but not related to the fault generated impulses. On the other hand, this process requires SI calculation for all combinations of scale and shape factor on the chosen discretized ranges which is obviously not computationally efficient.

Furthermore, wavelet filter based de-noising cannot eliminate the in-band noise or the noise with the frequency content that is the same as the frequency range covered by the daughter Gabor wavelet. This limitation affects the performance of wavelet filter based method. In addition, as proven in chapter 3 when the noise intensity increases, the SI approaches a constant, 0.8455. As a result, in the presence of coloured additive noise, SI minimization process may direct the scale or the center frequency of the bandpass filter towards a wrong frequency band.

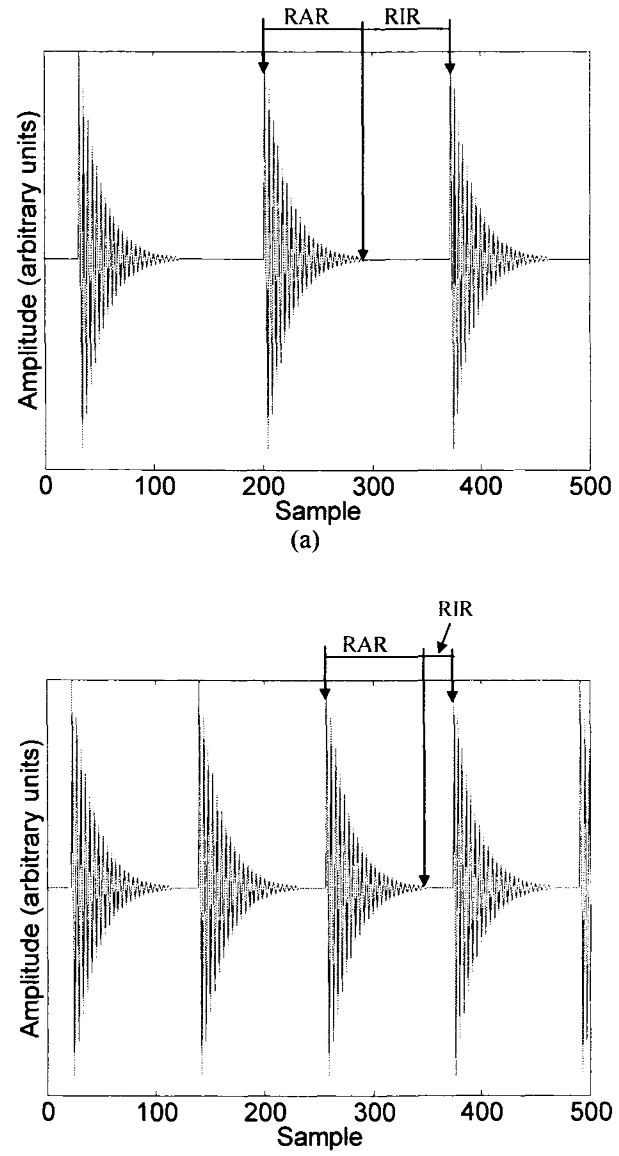
To tackle the above drawbacks, a new scale selection method is proposed to make the de-noising method less susceptible to other signal or noise components of impulsive nature. Furthermore, to improve the performance of the filter based de-noising method, the problem of in-band noise is investigated and a spectral subtraction technique is proposed.

4.2 Scale Selection

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. The proper shape factor value is then found in a separate step by minimizing SI at the selected scale. This method involves much less computational effort.

4.2.1 Scale selection based on instantaneous frequency

As shown in chapter 3, the best scale corresponds to the resonance frequency excited by the fault impacts. Accordingly, a resonance frequency estimation method would provide us with the proper scale value. The vibration signal from a faulty bearing is periodic. Each period contains two ranges: a resonance active range (RAR) and a resonance inactive range (RIR) as shown in Figure 4.1a and b. When the shaft speed goes up, the RAR length remains unchanged or 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 (Figure 4.1a and b).



(a)

(b)

Figure 4.1 Simulated faulty bearing signal at 12000Hz sampling rate and rotational frequency of (a) 20 Hz and (b) 29 Hz

On the other hand, considering:

$$f(t) = a(t)\cos\varphi(t),$$

the instantaneous frequency of $f(t)$ is given by $\varphi'(t)$. The fault generated impulse can be modeled as:

$$S(t) = A.e^{-\beta t}\cos(\omega_0 t)u(t)$$

where u(t) is the units step function. The instantaneous frequency found for the RAR would be the excited resonance frequency ω_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 data sets measured from the same bearing at two different rotational frequencies. It is reasonable to expect the probability of the resonance frequency to increase more, compared to the other frequencies with an increase in the rotational speed. To show this idea more clearly, define $\varphi'^{s}_{m}(n)$ as the instantaneous frequency at sampling time nT_s of the vibration signal measured at sampling frequency $\frac{1}{T_s}$ and rotational speed ω_m, and denote the empirical probability distribution of $\varphi'^{s}_{m}(n)\ (n = 1 \ldots N_a)$ by $\Pr^{s}_{m}(f)$ given as:

$$\Pr^{s}_{m}(f) = \frac{N_f}{N_a}$$

N_f = number of samples with instantaneous frequency of $\varphi'^{s}_{m}(n) = f$

N_a = total number of samples

Following a rise in the shaft rotational speed, the probability density change for each frequency will be:

$$\Delta \Pr^{s}_{i,j}(f) = \Pr^{s}_{j}(f) - \Pr^{s}_{i}(f) \quad \text{for} \quad \omega_j > \omega_i \tag{4-1}$$

Then the resonance frequency can be identified as:

$$f_{res} = \arg\max_{f} \left\lfloor \Delta \Pr^{s}_{i,j}(f) \right\rfloor$$

where f_{res} is the estimated resonance frequency. The same concept can be applied for a decrease in the rotational speed.

Though the Hilbert transform is a simple approach for calculating instantaneous frequency (Cohen 1995), it results in a considerable error in discrete time domain and with the presence of noise. For this reason, the instantaneous frequencies are calculated as follows. In doing so, the mother wavelet is written as:

$$\psi(t) = g(t)e^{i\theta t} \tag{4-2}$$

It has been shown (Mallat 1998) that the wavelet transform of a signal defined as $f(t) = a(t).\cos\varphi(t)$ can be expressed as:

$$W_f(s,u) = \frac{\sqrt{s}}{2}a(u).e^{i\varphi(u)}\left(\hat{g}(s[\xi - \varphi'(u)]) + \varepsilon(u,\xi)\right) \tag{4-3}$$

where $\hat{g}(f)$ is the Fourier transform of $g(t)$, $\xi = \frac{\theta}{s}$, and $\varphi'(t)$ is the derivative of $\varphi(t)$. The corrective term $\varepsilon(u,\xi)$ is negligible if $a(t)$ and $\varphi'(t)$ have small variations over the support of $\psi_{s,u}(t)$ defined in Eq. (2-2) 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. (3-1) and assuming a proper shape factor, both of the above conditions would satisfy. Neglecting the corrective term in Eq. (4-3) leads to:

$$\frac{|W_f(s,u)|^2}{s} = \frac{1}{4}a^2(u)\left|\left(\hat{g}(s[\xi - \varphi'(u)])\right)\right|^2 \tag{4-4}$$

From Eq. (3-5), one has

$$\hat{g}(f) = c\sqrt{\frac{\pi}{\sigma^2}}e^{-(\frac{\pi^2}{\sigma^2}).f^2} \tag{4-5}$$

According to Eq.(4-5), Eq. (4-4) 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. (4-4). Figure 4.2 shows a flowchart of the proposed de-noising algorithm.

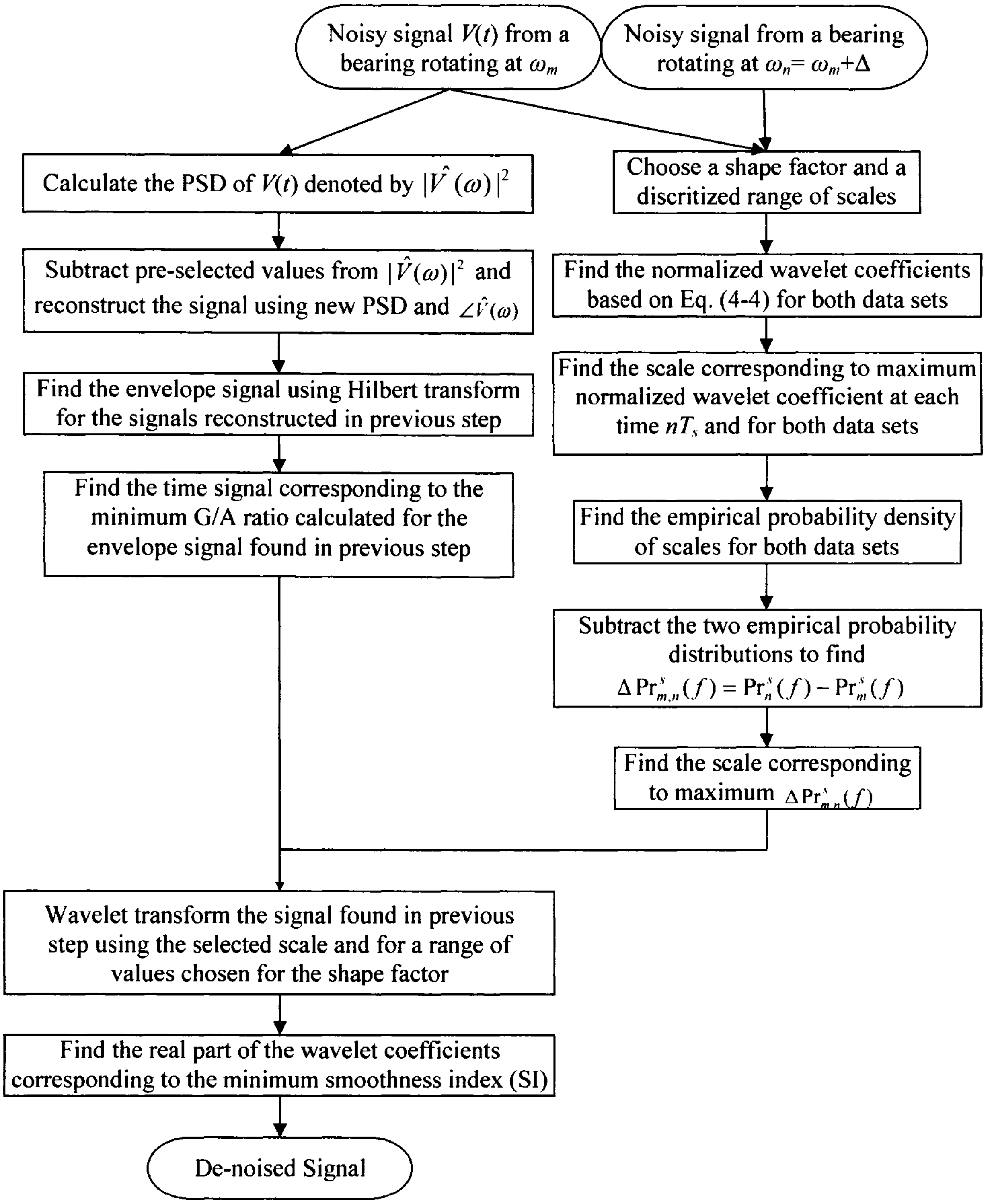


Figure 4.2 Flowchart of the proposed de-noising method

4.3 Spectral Subtraction

As explained in previous sections, bandpass filtering the signal around the resonance frequency band results in a signal which better reflects the impulsive nature of the vibrations of a faulty bearing. However, a drawback of the above approach is that the in-band noise, i.e. the noise with frequency components in the range covered by the daughter Gabor wavelet is not eliminated through the de-noising process. Due to this flaw, the bandpass filtering will not result in considerable enhancement for signals with very low SNR. In other words, the fault impulses may not be clearly identified because of the strong in-band noise. To ameliorate the wavelet filter based de-noising method, spectral subtraction can be used prior to bandpass filtering. Spectral subtraction is a very effective approach for noise elimination. It has been widely used in speech signal processing (Lim and Oppenheim 1979) but is often overlooked in machinery fault diagnosis research. As a result, there are very few studies (Dron, Bolaers and Rasolofondraibe 2004) applying this method in the above mentioned area.

In this method the PSD (Power Spectral Density) of the noise is estimated and subtracted from the PSD of the raw data in order to find an estimate of the PSD of the useful part of the signal. This estimate will later be used to reconstruct the purified signal.

Consider:

$$V(t) = v(t) + N(t) \qquad (4\text{-}6)$$

where $V(t)$ is the measured vibration, $v(t)$ is the faulty bearing vibration and $N(t)$ is the corrupting noise. Then:

$$|\hat{V}(\omega)|^2 = |\hat{v}(\omega)|^2 + |\hat{N}(\omega)|^2 + \hat{v}(\omega).\hat{N}^*(\omega) + \hat{v}^*(\omega).\hat{N}(\omega) \qquad (4\text{-}7)$$

In order to estimate $|\hat{v}(\omega)|^2$, one needs to find an approximation of the last three components in Eq. (4-7). Subtraction of this estimate from the PSD of the measured signal yields an approximation of the PSD of the faulty bearing vibration. In speech signal processing, noise is uncorrelated with the speech. This fact makes it possible to ignore the last two terms in Eq. (4-7). On the other hand, a good estimate of $|\hat{N}(\omega)|^2$ can be found by analyzing the intervals where speech is not present (Lim and Oppenheim

1979). However, none of these is applicable to the vibration signal measured from faulty bearings.

In this case the corrupting noise also contains the vibration from other sources (e.g. shaft imbalance, gear meshing, etc) which is not necessarily uncorrelated with bearing vibrations due to mechanical engagements. However, the fault generated impact usually excites a resonance in the system at much higher frequency than the vibration generated by other machine elements (McFadden and Smith 1984b). As a result, it can be assumed that the major part of the in-band noise consists of the Gaussian background noise present in any measurement device. To elaborate, consider:

$$V(t) = G(t) + N_2(t) = v(t) + N_1(t) + N_2(t) \tag{4-8}$$

where $V(t)$ is the measured vibration, $v(t)$ is faulty bearing vibration, $N_1(t)$ is the noise correlated with $v(t)$ and caused by other vibration (e.g., shaft imbalance, gear meshing, etc.) and $N_2(t)$ is the Gaussian background noise uncorrelated with $v(t)$ and $N_1(t)$. From Eq. (4-8), one has:

$$|\hat{V}(\omega)|^2 = |\hat{v}(\omega) + \hat{N}_1(\omega)|^2 + |\hat{N}_2(\omega)|^2 \tag{4-9}$$

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 approximated by:

$$|\tilde{\hat{G}}(\omega)|^2 = |\hat{V}(\omega)|^2 - |\tilde{\hat{N}}_2(\omega)|^2 \tag{4-10}$$

where $\tilde{\hat{N}}_2(\omega)$ and $\tilde{\hat{G}}(\omega)$ are estimates of $\hat{N}_2(\omega)$ and $\hat{G}(\omega)$ respectively.

Given $|\tilde{\hat{G}}(\omega)|$, an estimate of $|\hat{v}(\omega) + \hat{N}_1(\omega)|$, to estimate $v(t) + N_1(t)$ the phase information is also required. Here it is assumed that the phase information is relatively unimportant so that $\angle\left(\hat{v}(\omega) + \hat{N}_1(\omega)\right)$ can be approximated by $\angle\left(\hat{V}(\omega)\right)$.

Bearing in mind that the main purpose of the spectral subtraction in this context is to reduce the in-band noise which would usually correspond to a narrow frequency band, it is reasonable to assume that the Gaussian background noise $N_2(t)$ is 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 proven in chapter 3, the presence of white noise in the measured vibration signal leads to the larger values of the SI closer to the constant 0.8455. It was also shown that the presence of noise causes larger SI and the elimination of this noise will lead to smaller SI. As a result it is expected to see this value decrease by reducing the intensity of the corrupting white Gaussian background noise. This can be done by subtracting a constant value from the PSD of the vibration signal for all the frequencies. Although this approach may also reduce frequency components other than that of the white Gaussian 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 $\left[|N_{\min}|^2, |N_{\max}|^2\right]$ are considered for $|\tilde{N}_2(\omega)|^2$. The value on this range which minimizes the G/A ratio calculated for the envelope of the spectral subtracted signal is chosen as the best estimate for $|\hat{N}_2(\omega)|^2$. As Eq. (4-10) also yields negative values, it is modified as follows to ensure a non-negative result:

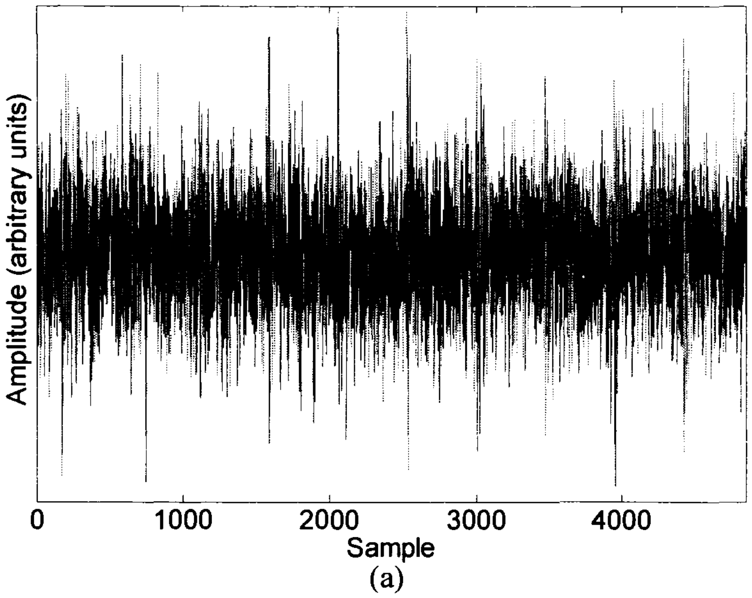
$$|\tilde{\hat{G}}(\omega)|^2 = \begin{cases} |\hat{V}(\omega)|^2 - |\tilde{\hat{N}}_2(\omega)|^2 \; for |\hat{V}(\omega)|^2 > |\tilde{\hat{N}}_2(\omega)|^2 \\ 0 \quad Otherwise \end{cases} \tag{4-11}$$

Finally $v(t) + N_1(t)$ is estimated using:

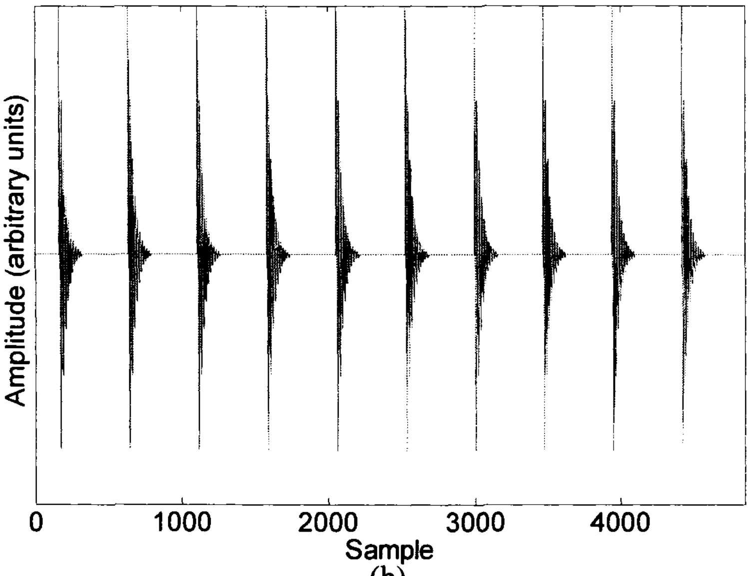
$$\tilde{\hat{G}}(\omega) = |\tilde{\hat{G}}(\omega)| . \exp\left[j.\angle\hat{V}(\omega)\right] \quad \text{and}$$

$$\tilde{G}(t) = F^{-1}\left[\tilde{\hat{G}}(\omega)\right] \tag{4-12}$$

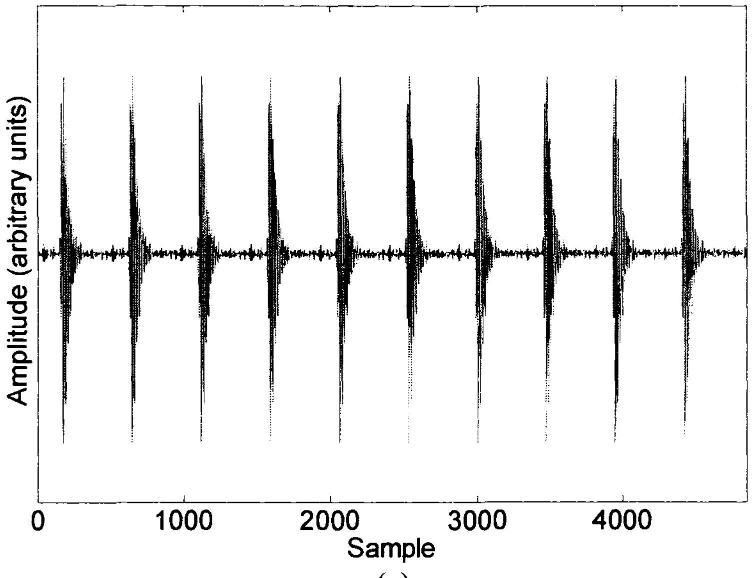
To examine the effect of approximating $\angle\left(\hat{v}(\omega) + \hat{N}_1(\omega)\right)$ by $\angle\left(\hat{V}(\omega)\right)$, simulated faulty bearing vibration signal corrupted by additive white Gaussian noise shown in the Figure 3.8b was spectral subtracted. The signal is then reconstructed by applying the above phase approximation approach and using the phase information of the original simulated signal without noise. The reconstruction results are shown in Figure 4.3c and d respectively. It should be noted that the results presented in Figure 4.3c and d are obtained by the spectral subtraction without using wavelet filtering. 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. The noisy signal and the original simulated signal are also shown for the same intervals on Figure 4.3a and b respectively. The comparison of Figure 3.8c and Figure 4.3c shows the superior performance of the spectral subtraction in compare 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. As a result, 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. (4-12) is bandpass filtered using the wavelet filter based de-noising method explained in previous sections to eliminate $N_1(t)$.



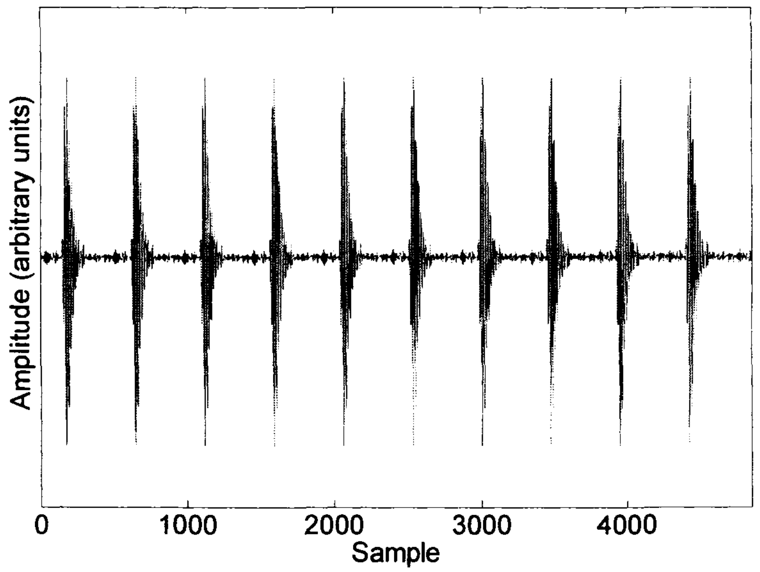
(a)


Amplitude (arbitrary units)
0
1000
2000
3000
4000
Sample

(b)


Amplitude (arbitrary units)
0
1000
2000
3000
4000
Sample

(c)



(d)

Figure 4.3 (a) Simulated signal plus noise, (b) simulated signal, (c) spectral subtracted signal using approximation of the phase information, (d) spectral subtracted signal using the phase information of the original simulated signal shown in part (b)

4.4 Experimental Evaluation

The proposed method is further evaluated using the vibration data measured in our lab. Experimental setup is the same as described in Chapter 3. The experiment is carried out using a SpectraQuest Machinery Fault Simulator (MFK-PK5M) as shown in Figure 4.4. The shaft speed was set at 630 RPM (10.5Hz). The left bearing has a pre-seeded single fault on the outer race with a characteristic frequency of 32 Hz (= $3.052f_r$). The signal was fed to an NI AT-MIO-16DE-10 DAQ card and then collected through LabVIEW. The signal was processed using MATLAB on a Pentium® 4/2.52GHz PC.

The accelerometer was mounted on the simulator base at a spot that is away from the faulty bearing and closer to the belt and the driven gearbox (Figure 4.4). The vibration data were acquired at 20000 samples/sec. A portion of the measured data is plotted in Figure 4.5a and 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 was acquired at 780 RPM (13Hz) rotational speed. Following the spectral subtraction, SI is minimized using the estimated scale and for a range of values chosen for σ^2.

The proper (s-σ^2) combination is found as (50, 0.32) and the minimum SI is 0.7398. The de-noising result associated with this (s-σ^2) combination is displayed in Figure 4.5b. As shown in the figure, the time interval between two consecutive impulses is about 0.0313 seconds which precisely reflects the fault characteristic frequency 32Hz. Figure 4.5c shows the de-noising result with spectral subtraction step eliminated from the algorithm. Though some impulsive features can still be detected from the result, comparison of Figure 4.5b and Figure 4.5c indicates that a better de-noising performance can be achieved by adding the spectral subtraction component to the bandpass filtering algorithm.

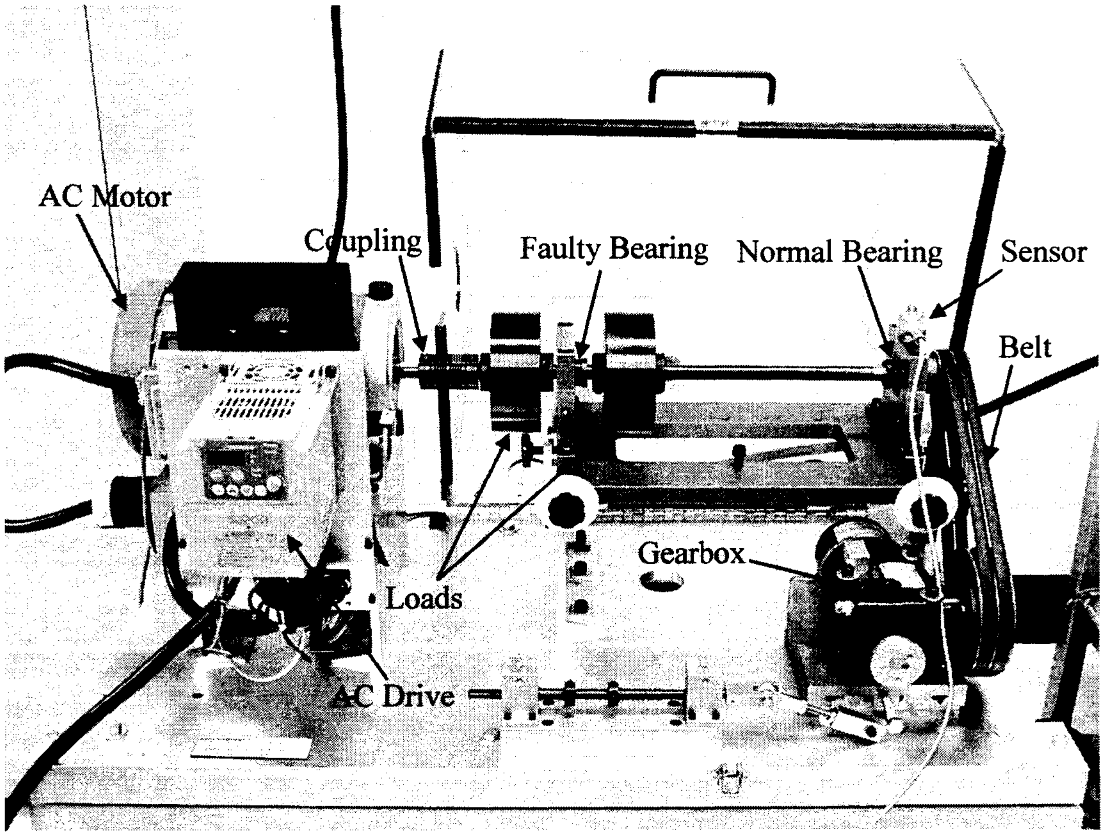
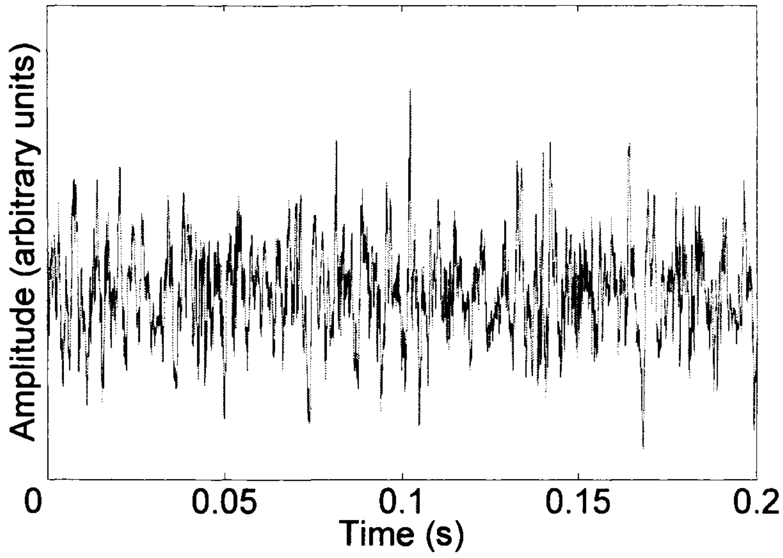
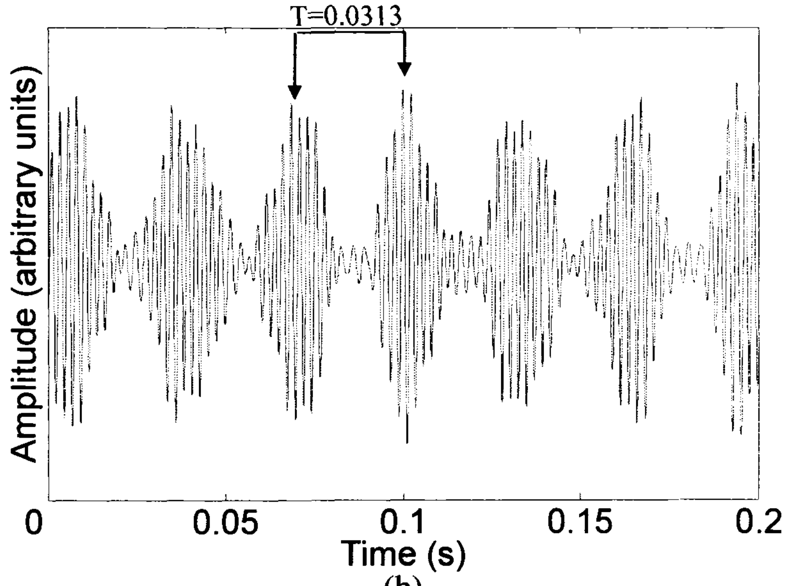


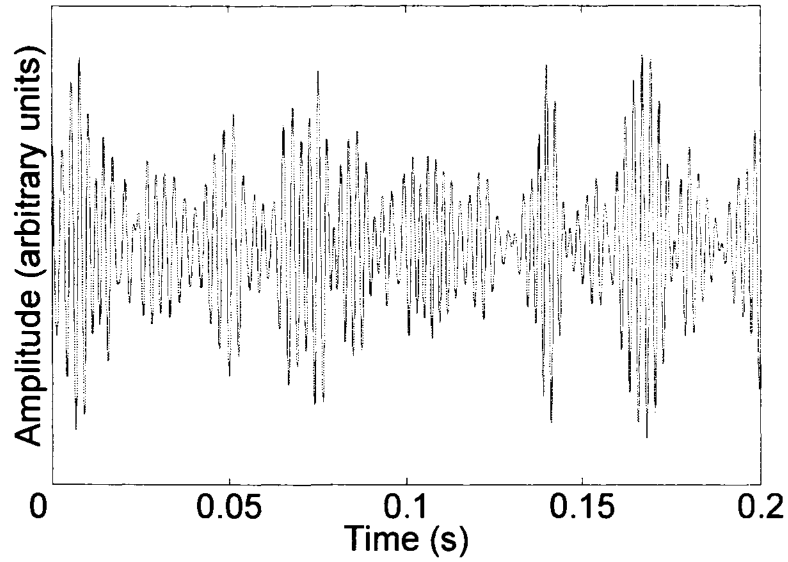
Figure 4.4 Test setup #1


Amplitude (arbitrary units)
0
0.05
0.1
0.15
0.2
Time (s)

(a)


T=0.0313
Amplitude (arbitrary units)
0
0.05
0.1
0.15
0.2
Time (s)

(b)



(c)

Figure 4.5 (a) Original measured vibration signal, (b) De-noised through the spectral subtraction followed by wavelet filter de-noising, (c) De-noising result obtained using wavelet filtering alone (Sampling rate: 20000 samples/s)

In another experiment, the two rotors are removed from the setup and faulty bearing is installed on the belt-end of the shaft. The sensor (accelerometer) is installed on the structure at a spot away from the faulty bearing as shown in Figure 4.6. Vibration signal was acquired at two rotational frequencies: 12Hz and 13.2Hz. The two data sets were fed to the resonance frequency estimation algorithm to find the proper scale value. A portion of the measured raw data from both cases are shown in Figures 4.7a and 4.8a and de-noised using the proposed joint wavelet filtering and spectral subtraction. The de-noising results are shown in figures 4.7b and 4.8b. As one can see the proposed scale and shape-factor selection algorithms perform very well. In the de-noised results, the fault generated impulses can be clearly identified in both cases. The time intervals between two consecutive impulses are 0.0248 (sec) for Figure 4.7b and 0.0273 for Figure 4.8b which precisely match the fault characteristic frequencies of 40.3Hz and 36.6Hz respectively. Figures 4.7c and 4.8c show the de-noising result with spectral subtraction step eliminated from the algorithm. Obviously, the elimination of this pre-processing step has once again led to a weaker de-noising performance.

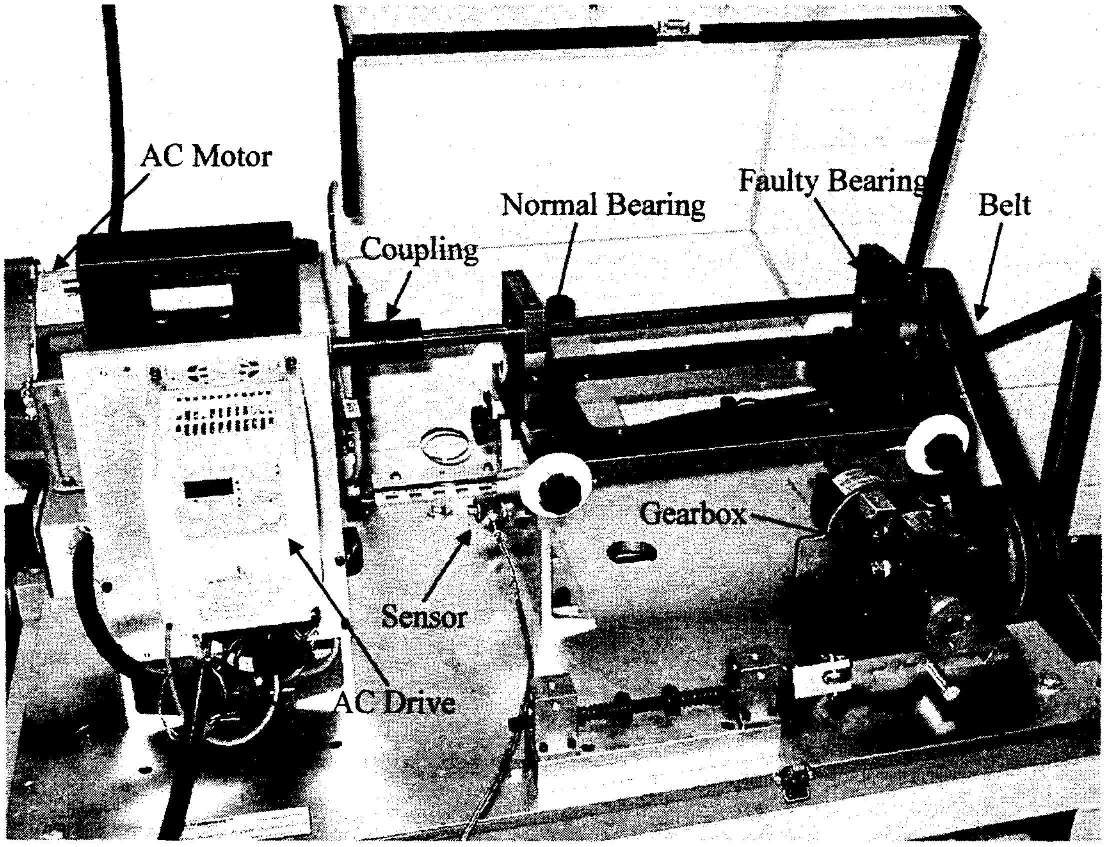
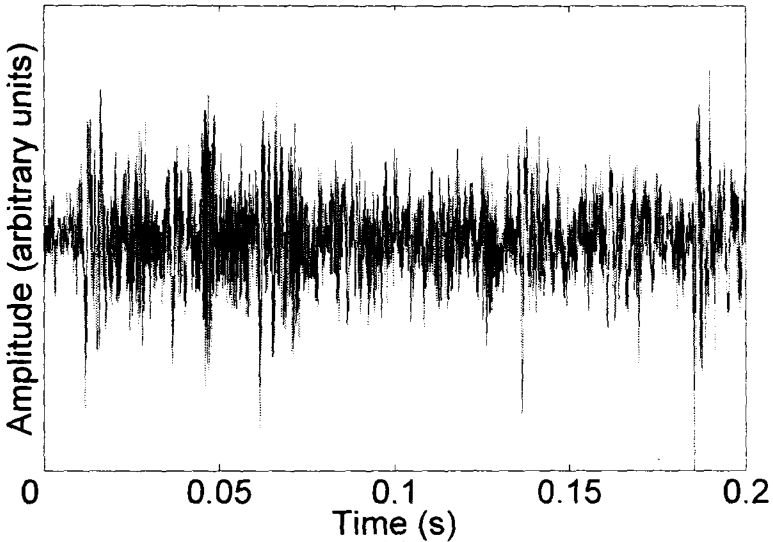
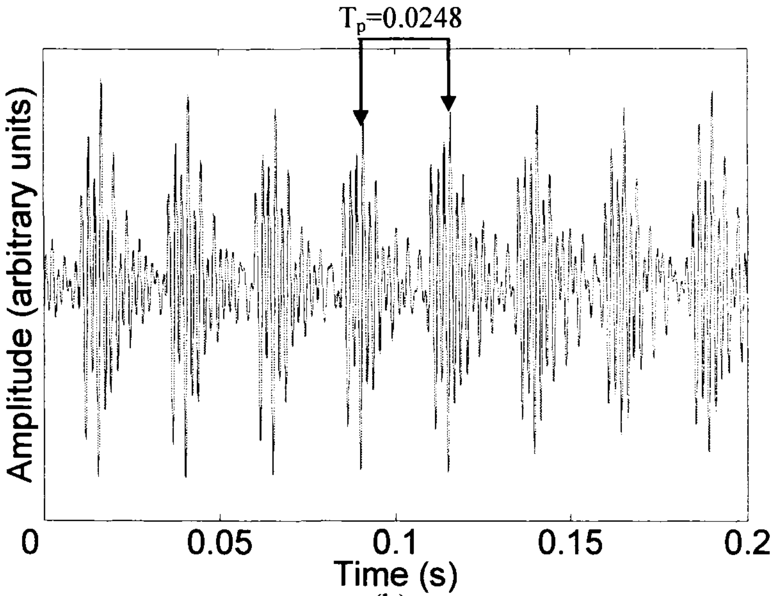


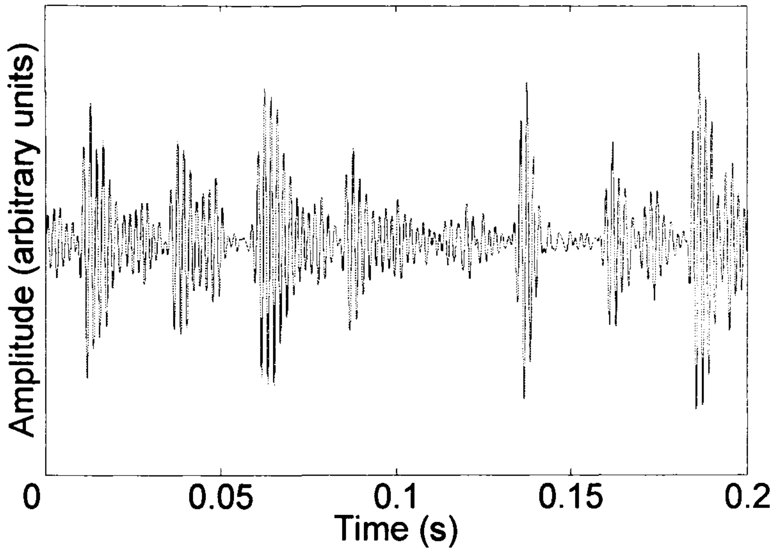
Figure 4.6. Test setup #2



(a)

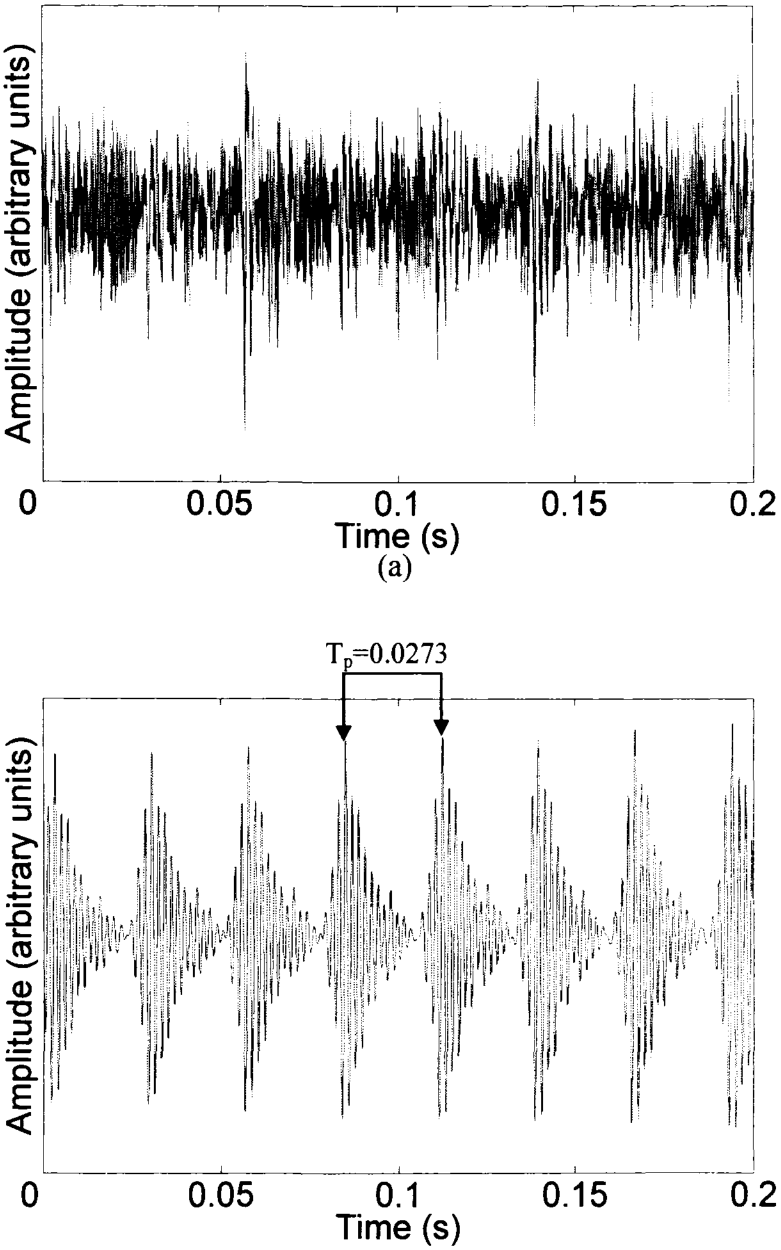


(b)



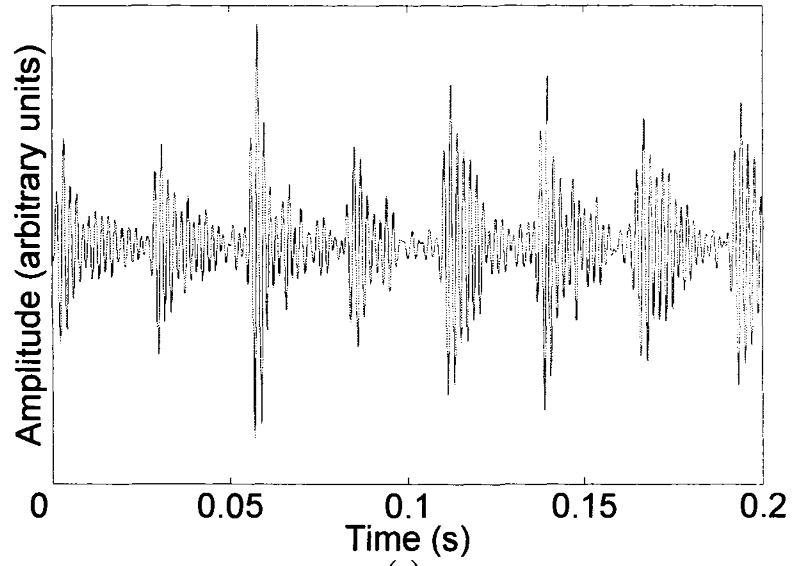
(c)

Figure 4.7. (a) Vibration signal measured at 13.2Hz rotational frequency, (b) De-noising result using spectral subtraction followed by wavelet filter de-noising, (c) De-noising result obtained using wavelet filtering alone (Sampling rate: 20000 samples/s)


Amplitude (arbitrary units)
0
0.05
0.1
0.15
0.2
Time (s)
Tp=0.0273
Amplitude (arbitrary units)
0
0.05
0.1
0.15
0.2
Time (s)

(a)

(b)



(c)

Figure 4.8. (a) Vibration signal measured at 12Hz rotational frequency, (b) De-noising result using spectral subtraction followed by wavelet filter de-noising, (c) De-noising result obtained using wavelet filtering alone (Sampling rate: 20000 samples/s)

4.5 Conclusion

Though the wavelet filtering method presented in Chapter 3 performed well, its performance deteriorates substantially when the signal contains in-band noise with frequency content identical to the frequency band covered by the daughter Gabor wavelet. In this chapter, a spectral subtraction approach is proposed to remove the in-band noise prior to wavelet filtering. The spectral subtraction is based on the assumption of white uncorrelated in-band noise. The power spectral density of such corrupting noise is estimated by minimization of the ratio of the geometric mean to the arithmetic mean calculated for 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 like shaft imbalance, gear meshing, etc. To find the proper scale value, the resonance phenomenon due to the fault generated impacts is utilized. Proper shape factor is then found in a separate step using the selected scale and by minimizing the smoothness index. The experimental result showed good performance of the proposed scale and shape factor selection algorithms. It also illustrated that the performance of the wavelet filter based de-noising method can be improved by pre-processing the signal using spectral subtraction. The proposed de-noising algorithm aims to enhance the vibration signals measured from faulty bearings. However, it should be noted that the same principles might be applied to vibrations of other machinery components as well. In the next chapter, enhancement of the signals measured from an oil debris sensor is investigated.

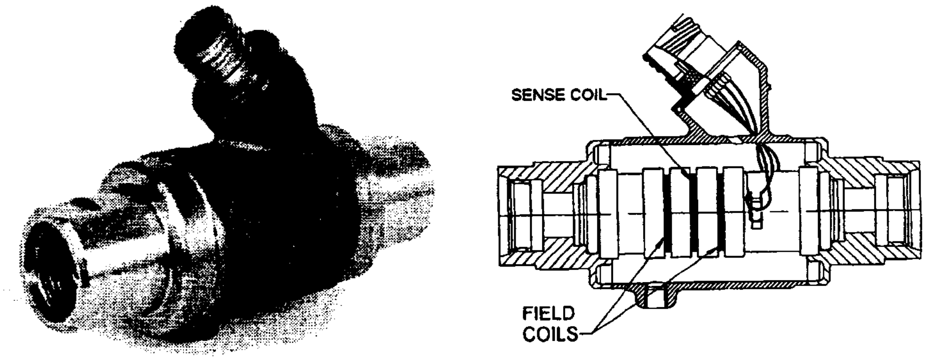
Chapter 5: Interference Removal from Signals Collected by an Oil Debris Sensor

Besides the vibration analysis approach, oil debris monitoring is another widely used method for machine health monitoring. In this method wear severity of machinery components can be assessed by analysing the lubricating oil and measuring the level of presence of metallic particles. Such data can be collected through special sensors designed to detect passage of metallic particles as well as their size and nature. This information provides us with a direct perception of the degree of wear in the mechanical system. Maintenance decisions can be made by the comparison between such data and tabulated data measured from the same kind of machinery at different health states. However, similar to the case of vibrations measured from faulty bearings, these signals are also contaminated by noise and interferences. This inevitably affects the performance of the wear estimation system. In the following sections the problem of such interferences is addressed.

5.1 Oil Debris Monitor

ODM (Oil Debris Monitor) is an on-line oil condition monitoring system. This system is installed on the oil return lines and before the oil filter to provide a full flow passage way for the lubricating oil. It can detect the metallic particles that pass through it (Industry Canada, 2007). The ODM was first developed for the monitoring of the F22 Advanced Tactical Fighter engine. The operation of the ODM is based on sensing the electromagnetic disturbances caused by passing metallic particles. Figure 5.1a and b show the oil debris sensor. The output signature resulting from the passage of a metallic particle is shown in Figure 5.2.

This signature is affected by the nature and size of the passing metallic particle. The phase of the signature depends on the nature (ferromagnetic or non-ferromagnetic) of the passing particle and the amplitude depends on the mass of the particle for ferromagnetic metals and to the surface area of the particle for the non-ferromagnetic metals (Miller and Kitaljevich 2000). Following the measurements using the sensor, the built-in software counts the number of such signatures contained in the signal and estimates the size and nature of each particle through the phase and amplitude of each one. As a result, an estimate of the wear level could be achieved. If necessary an alarm regarding the health state of the machinery would provide time for scheduled maintenance and consequently reduce the occurrences of unplanned production delays or in-flight shut-downs in the aircrafts (Industry Canada 2007, Miller and Kitaljevich 2000). The smallest particle that can be detected by the sensor depends on the sensor bore size. For ½" sensors used in F119 engines this minimum size is 125 microns (Miller and Kitaljevich 2000).



(a) (b)

Figure 5.1 (a) Oil debris sensor, (b) Sensor cross section (Miller and Kitaljevich 2000)

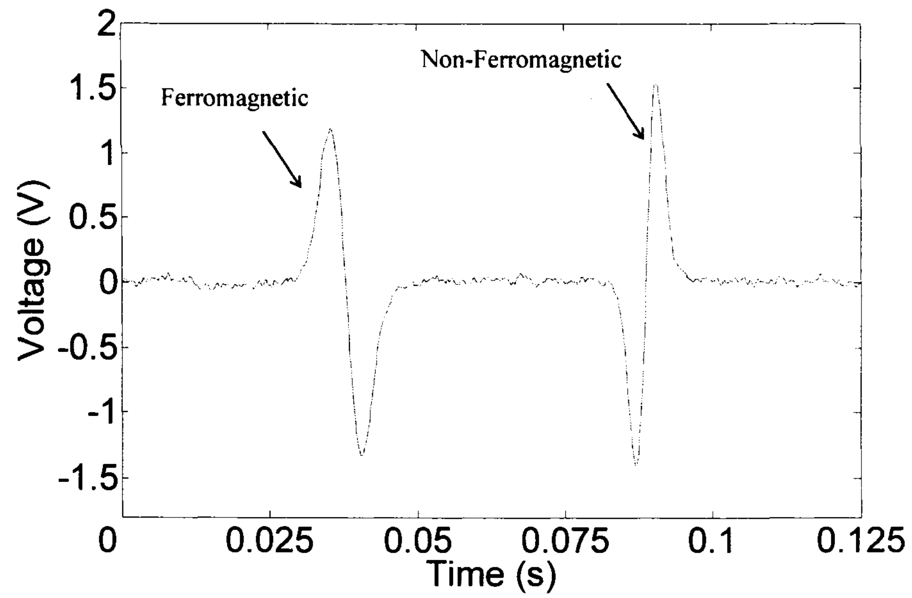


Figure 5.2 Sensor output in response to the passage of a metallic particle (Sampling rate: 8000 samples/s)

5.2 Interferences

Like any other measuring device, the performance of such system is affected by the noise or interfering signals. One such interference results from the vibrations of the oil debris sensor. The effect of these vibrations on the output signal can be simulated as the addition of a combination of modulated sinusoids to the signal composed of the signatures of passing particles. Figure 5.3a shows the output of an oil debris sensor with consecutive passage of two particles and in the absence of vibration interferences, sampled at 8000 Hz. Figure 5.3b displays the expected output of the system in the presence of structural vibrations. This was obtained by the addition of the simulated vibration interferences to the original signal of Figure 5.3a. As one can see it is not possible to detect the passage of metallic particles using the noisy signal. An Adaptive Noise Cancellation (ANC) technique is proposed in this work to mitigate the problem.

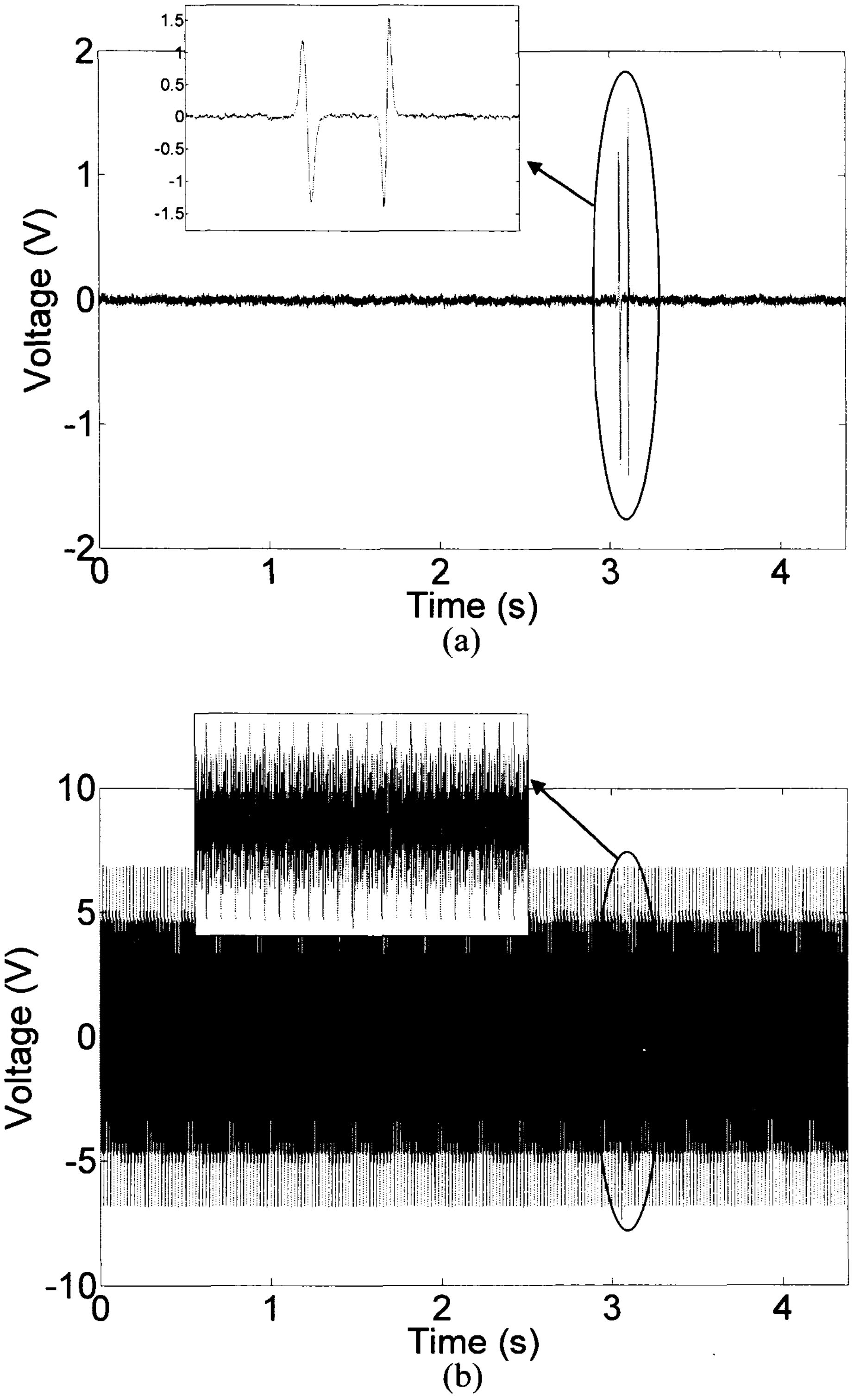


Figure 5.3 (a) The output of the sensor due to two consecutive metallic particle passages (b) Output $+ 2\sin(1400\pi t)(1+\cos(900\pi t)) + 2\sin(2000\pi t)(1+\cos(700\pi t))$ (Sampling rate: 8000 samples/s)

5.3 Adaptive Noise Cancellation (ANC)

A filtering process designed to eliminate the noise and pass the signal components could provide a good approximation of the original signal. A similar approach was taken in previous chapters to increase the SNR of the vibration signals measured from faulty bearings. These filters can be fixed or adaptive. A fixed filter can be designed when adequate prior knowledge about both signal and noise is available. However an adaptive system which adjusts certain parameters adaptively can compensate for the lack of such information.

The frequency content of the particle signals are mainly dependent on the speed in which the metallic particles pass through the sensor. This is defined by the lubricating oil flow speed. The vibration interferences on the other hand depend on the vibrations of the structure where the sensor is mounted. Accordingly, it is obvious that both signal and interference attributes are totally dependent on the working conditions of the machinery and can not be previously known for a proper filter design. Consequently, an adaptive system capable of adjusting filter parameters according to the working conditions of the machinery is required.

Adaptive noise cancellation makes use of a reference input correlated with the corrupting noise. This input is filtered to find an estimate of the corrupting noise that minimizes the MSE. Subtraction of this filtered signal from the primary input results in a signal with higher SNR. Figure 5.4 shows the block diagram of the adaptive noise cancellation approach.

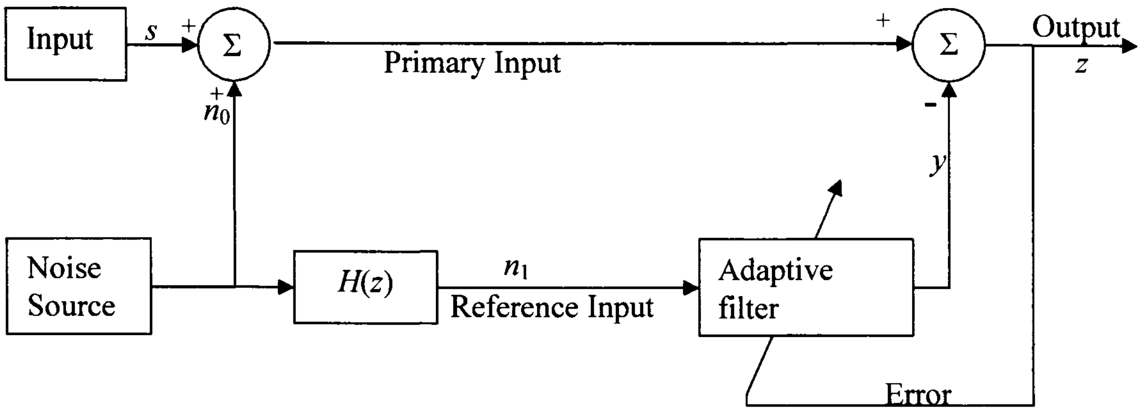


Figure 5.4 Adaptive Noise Cancellation technique

Assume that s, n_0, n_1 and y in Figure 5.4 are statistically stationary and have zero means. Assume also that s is uncorrelated with n_0 and n_1 but n_0 is correlated with n_1. The output z is

$$z = s + n_0 - y$$

Squaring both sides of the equation and taking the expectation yields (Widrow *et al* 1975):

$$\begin{aligned} E\left\{z^2\right\} &= E\left\{s^2\right\} + E\left\{(n_0 - y)^2\right\} + 2E\left\{s(n_0 - y)\right\} \\ &= E\left\{s^2\right\} + E\left\{(n_0 - y)^2\right\} \end{aligned}$$

As the above equation shows, by minimizing $E\left\{z^2\right\}$ the signal power $E\left\{s^2\right\}$ remains unaffected and $E\left\{(n_0 - y)^2\right\}$ approaches zero by proper adjustment of the filter coefficients. As a result, filter output y represents an estimate of noise n_0 in the minimum MSE sense. Let us consider the linear filtering process shown in the Figure 5.5.

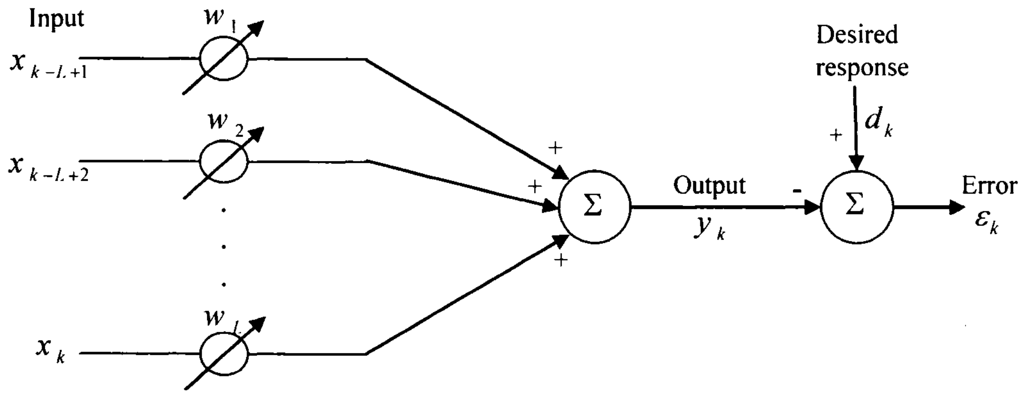


Figure 5.5 Adaptive linear filter

In Figure 5.5, x and w are input and filter weight respectively, y is the output of the filter and d is the desired output of the filter. The purpose of this filtering process is to find an estimate of a desired signal d in the minimum MSE sense through linear filtering the signal x. Comparing this filtering process with the adaptive noise cancellation explained above, it can be seen that in the ANC (Adaptive Noise Cancellation) the primary input plays the role of the desired signal. Assuming linear filtering with weight vector W of length L, one has (Widrow and Stearns 1985)

$$y_k = \sum_{l=1}^{L} w_l x_{k-L+l}$$

The above equation can be expressed in matrix form as follows

$$y_k = W X_k^T$$

where $X_k^T = [x_{k-L+1} \quad x_{k-L+2} \quad \dots \quad x_k]$ and $W^T = [w_1 \quad w_2 \quad \dots \quad w_L]$

Accordingly

$$\varepsilon_k = d_k - y_k = d_k - W^T X_k$$

$$\varepsilon_k^2 = d_k^2 + W^T X_k X_k^T W - 2 d_k X_k^T W \tag{5-1}$$

Taking the expectation, one obtains

$$E\{\varepsilon_k^2\} = E\{d_k^2\} + W^T E\{X_k X_k^T\} W - 2E\{d_k X_k^T\} W \tag{5-2}$$

or

$$E\{\varepsilon_k^2\} = E\{d_k^2\} + W^T R W - 2P^T W$$

where

$R = E\{X_k X_k^T\}$ is the input correlation matrix and $P = E\{d_k X_k^T\}$ is the cross correlation between the desired response and the input components. The performance surface given in Eq. (5-2) is a quadratic function of the weight vector W (Haykins 2002). Taking the partial derivative of the MSE with respect to the weight vector W yields

$$\Delta = \frac{\partial \varepsilon}{\partial W} = 2RW - 2P$$

When the weight matrix W is set at its optimum value, the above gradient is equal to zero and hence

$$W_0 = R^{-1} P$$

The above equation is an expression of the well-known Wiener-Hopf equation. However, in practice matrices R and P are not available. For this reason, a stochastic gradient approach known as Least Mean Square method is applied here. In this approach the noisy gradient is calculated from a single realization of matrix X and ε by taking the derivative of equation (5-1) with respect to W. In other words at each iteration in the adaptive process, there is a gradient estimation of the form

$$\hat{\nabla}_k = -2\varepsilon_k X_k$$

The corresponding steepest descent algorithm for filter weight matrix W update is defined as

$$W_{k+1} = W_k - \mu \hat{\nabla}_k$$

where μ is a gain constant or learning rate that regulates the speed and stability of adaptation.

To apply ANC for interference removal from the output signal of the oil debris sensors, we need a reference signal correlated with the structural vibrations. Such reference signal can be collected from a vibration sensor installed on the structure where the oil debris sensor is mounted. Consequently, the signal measured from the vibration sensor would be well correlated with the corrupting interferences and can lead to successful enhancement of the oil condition data. However, this method requires the addition of a vibration sensor and related hardware to the system which should be avoided as it would considerably complicate the system.

5.4 Adaptive line enhancement (ALE)

Adaptive line enhancement was first applied to the classical detection problem of finding a sine wave in noise (Widrow and Stearns 1985). The block diagram of the method is illustrated in Figure 5.6. In this method it is assumed that there is no correlation between the noise samples. This approach is appealing because it uses a delayed version of the primary input as the reference signal. It is therefore particularly suitable for the oil debris sensor interference problem because no extra sensors and associated hardware are needed.

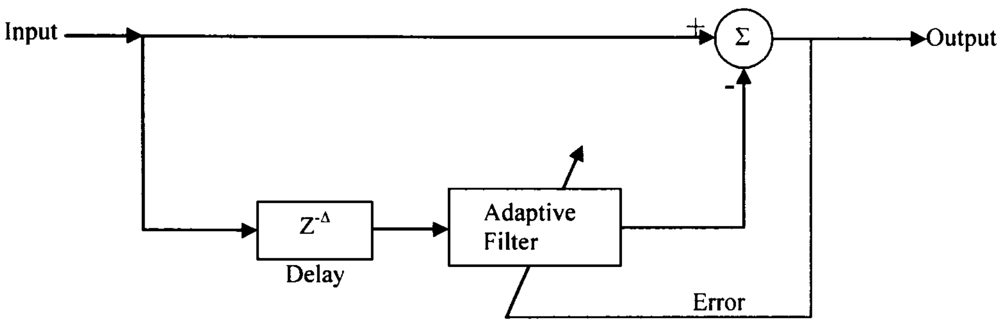


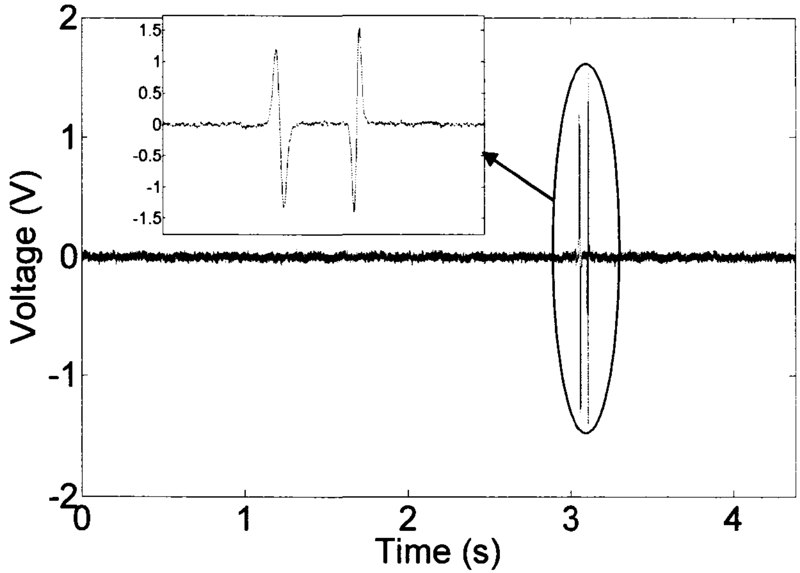
Figure 5.6 Adaptive Line Enhancement technique

In the case of oil debris sensor, it is expected that the passage of metallic particles follow a random nature and there should be no correlation between the corresponding signatures. On the other hand, it is reasonable to assume that the rate of change of the vibration nature of any mechanical system is slow compared to the adaptation rate of the filter. In other words a delayed version of the measured signal will be correlated with the original signal due to such vibration interferences. Following the same idea of using the ALE technique for sine wave detection, it would be possible to detect and eliminate such interferences.

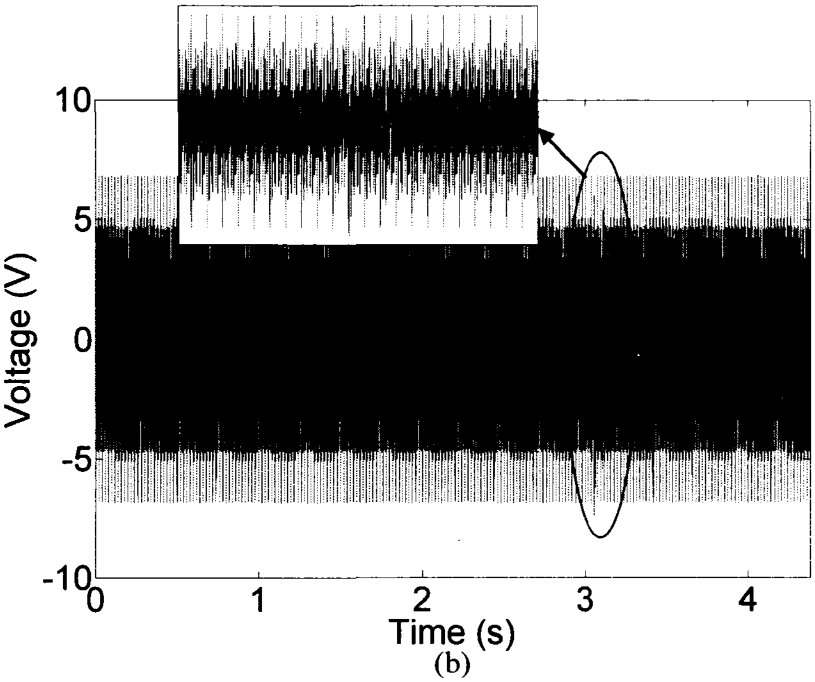
On major difference between the two approaches is that in oil debris monitoring we are not interested in the interferences while sinusoidal signal was the feature to be extracted in ALE. Furthermore, in contrast to other adaptive systems, in this case very high adaptation rate is unfavourable as it may also cause the useful signatures to be eliminated or distorted. As the characteristics of these signatures are later used to assess the wear level of the system, it is very important that the adaptive system leaves the particle signals intact. To prevent such a phenomenon, filter weight vector is updated following a number of iterations. The update value is then calculated by averaging the weight vector changes found in these iterations. It should be noted that this method is unable to remove the white Gaussian noise included in the signal while noise samples are also expected to be uncorrelated.

5.5 Performance assessment of the ALE technique using experimental signal plus simulated interference

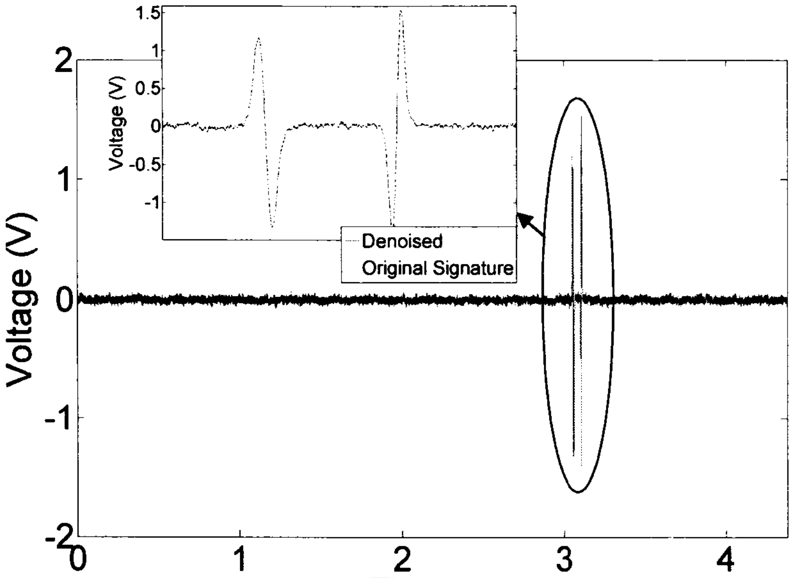
Figure 5.7a shows the oil debris sensor signal output due to the consecutive passage of two metallic particles in the absence of any vibration interferences with a sampling frequency of 8000 Hz. Figure 5.8a shows the signal of Figure 5.7a with additional simulated white Gaussian noise. Figure 5.7b and Figure 5.8b show the mixture of the oil particle signal (shown in Figure 5.7a and Figure 5.8a respectively) and different simulated interferences. Figure 5.7c and Figure 5.8c show the de-noised version of the same signals using the ALE approach. The de-noising results may be better appreciated by close-up views of the particle passing portions. In Figure 5.7c the true particle signal is plotted (dotted line) and is superimposed on top of the de-noised result (solid line). As one can see the difference between the original and the recovered particle signal is barely noticeable.


2
1
0
-1
-2
0
1
2
3
4
Voltage (V)
Time (s)
1.5
1
0.5
0
-0.5
-1
-1.5

(a)

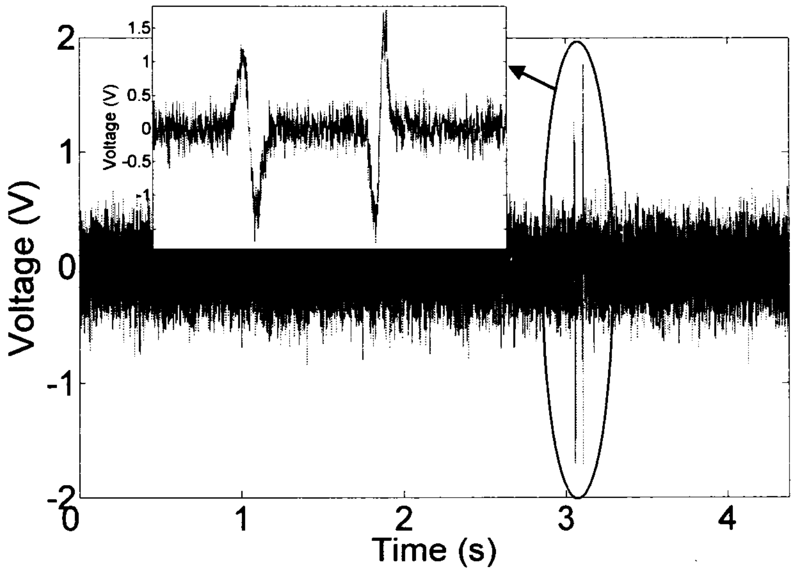

10
5
0
-5
-10
0
1
2
3
4
Voltage (V)
Time (s)

(b)

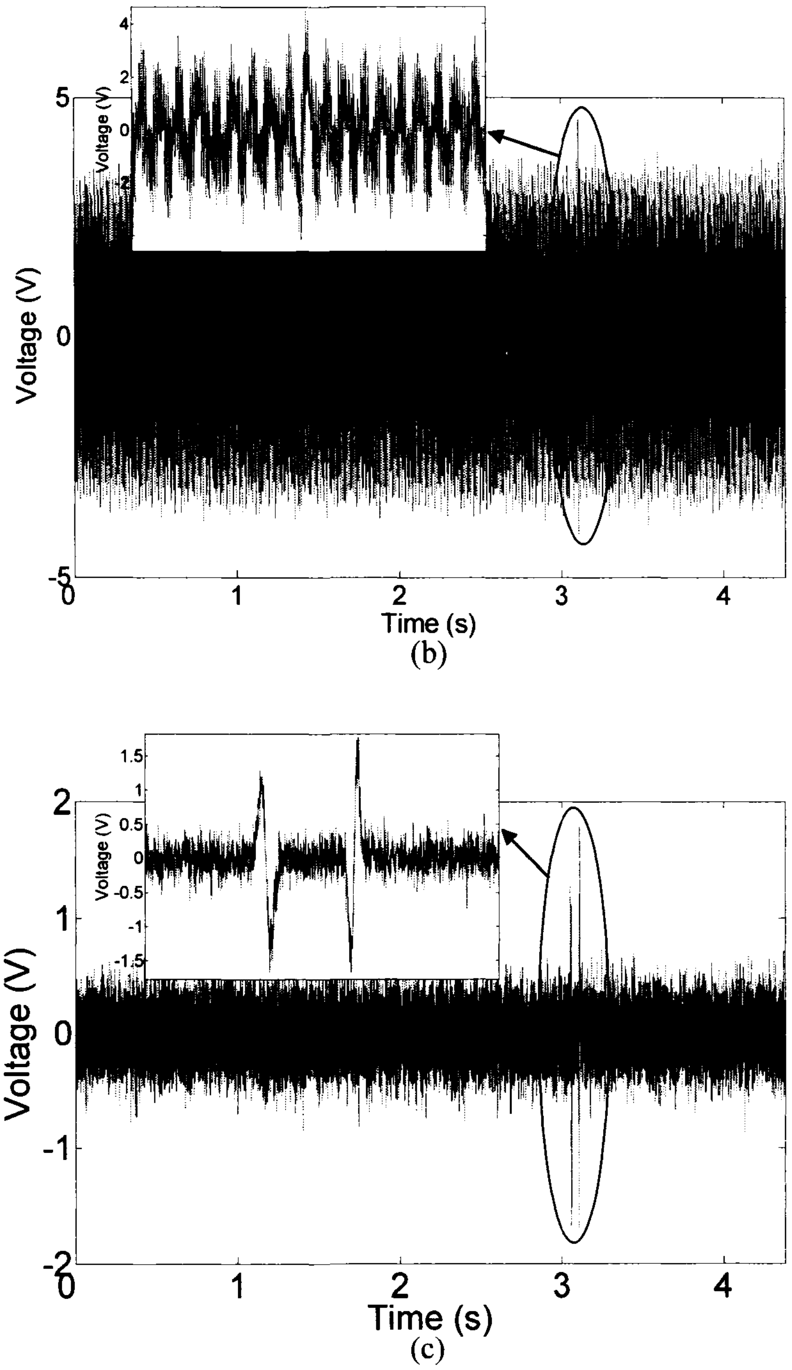


(c)

Figure 5.7 (a) Oil debris sensor output signal,
(b) Output $+2\sin(1400\pi t)(1+\cos(900\pi t))+2\sin(2000\pi t)(1+\cos(700\pi t))$, (c) restored signal using ALE, original output is superimposed on the magnified portion (dotted line) , the difference is unnoticeable (Sampling rate: 8000 samples/s)



(a)



(b)

(c)

Figure 5.8 (a) Oil debris sensor output shown in Figure 5.7a + additional simulated noise (b) Signal shown in part (a)+ $2\sin(1000\pi t)\cos(800\pi t)+2\sin(3000\pi t)\cos(4500\pi t)$, (c) restored signal using ALE (Sampling rate: 8000 samples/s)

5.6 Wavelet threshold de-noising

Though not as harmful as the vibration related interferences, the background noise due to the wiring and measurement system flaws also affects the performance of wear assessment system. This is true specifically for the very small particles where the corresponding signatures can be easily masked by the corrupting noise. As explained in previous section and also according to the simulation results, ALE technique is unable to remove the white Gaussian background noise. The reason is that, just like between particle signatures, there is no correlation between such noise samples. To further enhance the signal processed by the Adaptive Line Enhancement step, a wavelet threshold de-noising scheme is proposed.

As explained in Chapter 2, for the Gaussian corrupting noise, by applying hard thresholding rule and using the following threshold (Mallat 1998):

$$T_m = \sigma_m \sqrt{2\log_e N}$$

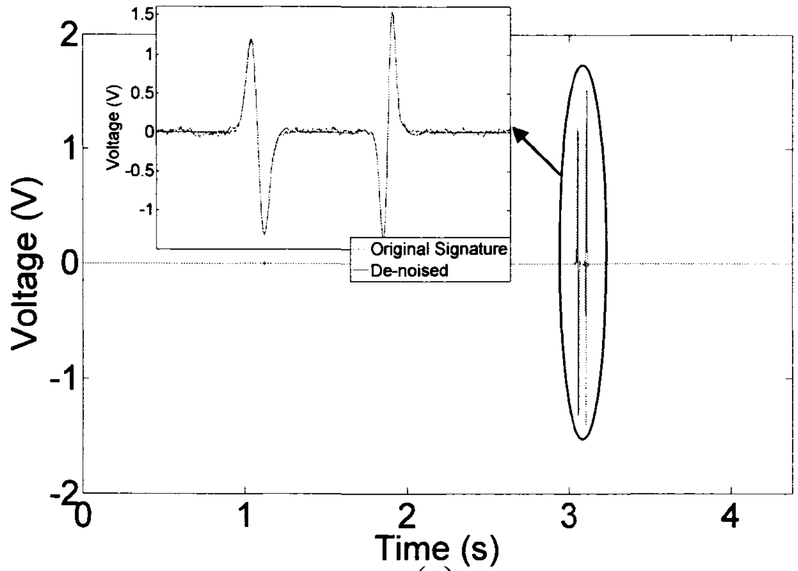
where N is the number of data points, it is possible to have an estimation error that is within a factor of $2log_2N$ of an ideal selection error. In this method noise variance at different scale levels should be known. One possible approach for noise variance estimation would be to measure the system noise in the absence of any passing particles or structural vibrations. However, it should be noted that the system background noise characteristics may vary in accordance with the working conditions.

On the other hand, a variance estimator insensitive to large outliers in the data set was introduced as (Mallat 1998):

$$\tilde{\sigma}_m = \frac{1}{0.6745} Med\left(\left|\langle W, \psi_{m,n}\rangle\right|\right)_{0\le n<\frac{N}{2^m}}$$

Following the same concept, for the signal containing particle signatures and white Gaussian noise, it is possible to consider the particle signatures as the outliers in the data set and use the above variance estimation method to calculate the threshold value. In this case no prior knowledge about the background noise is required.

This approach was tested on the simulated signals that were previously enhanced in the ALE step. Symlet 4 orthogonal wavelet was chosen as the mother wavelet. Figure 5.9 shows the de-noising results. The original particle signatures (dotted line) are also superimposed on the de-noised signal (solid line) in the close up view. As one can see the signatures of the passing particles are practically free from any remaining noise or interferences. The improvement achieved by the addition of the threshold de-noising step to the ALE algorithm can be more clearly seen by comparing Figure 5.8c and Figure 5.9b.



(a)

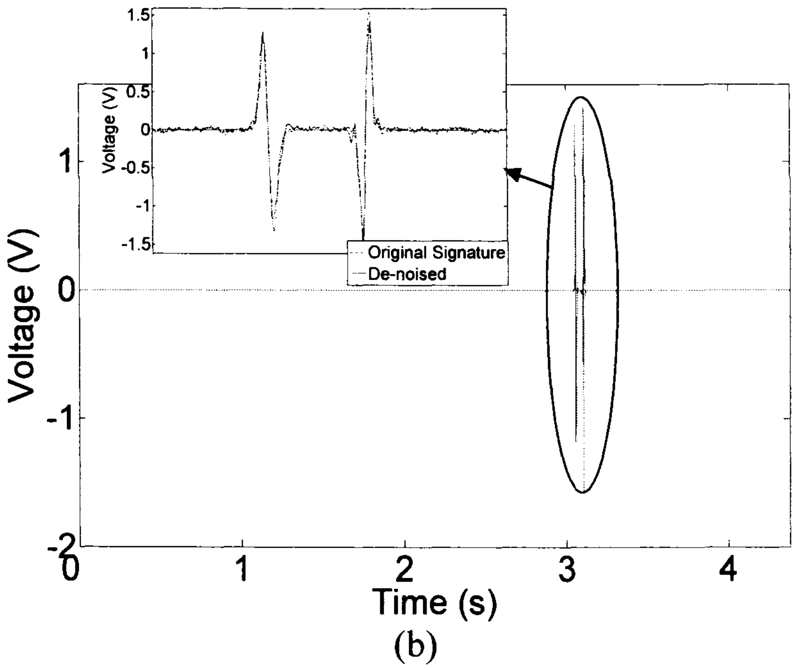


Figure 5.9 (a) Wavelet de-noised ALE result shown in Figure 5.7c, (b) wavelet de-noised ALE result shown in Figure 5.8c, original measured output is superimposed on the magnified portion (dotted line) (Sampling rate: 8000 samples/s).

5.7 Experimental evaluation

In this section, the ALE method is further evaluated using the signal acquired from oil debris sensor mounted on a shaker. Metal particles passed through the sensor while the vibration was introduced by the shaker. Sampling frequency was set at 8000 Hz. The measured signals are shown in Figure 5.10a and Figure 5.11a. ALE was applied to better reveal the signals. The restored signals are shown in Figure 5.10b and Figure 5.11b. These results were further enhanced using the proposed threshold de-noising scheme. Final results are illustrated in Figure 5.10c and Figure 5.11c. As one can see the original signature of the passing metallic particles can be clearly identified from the result.

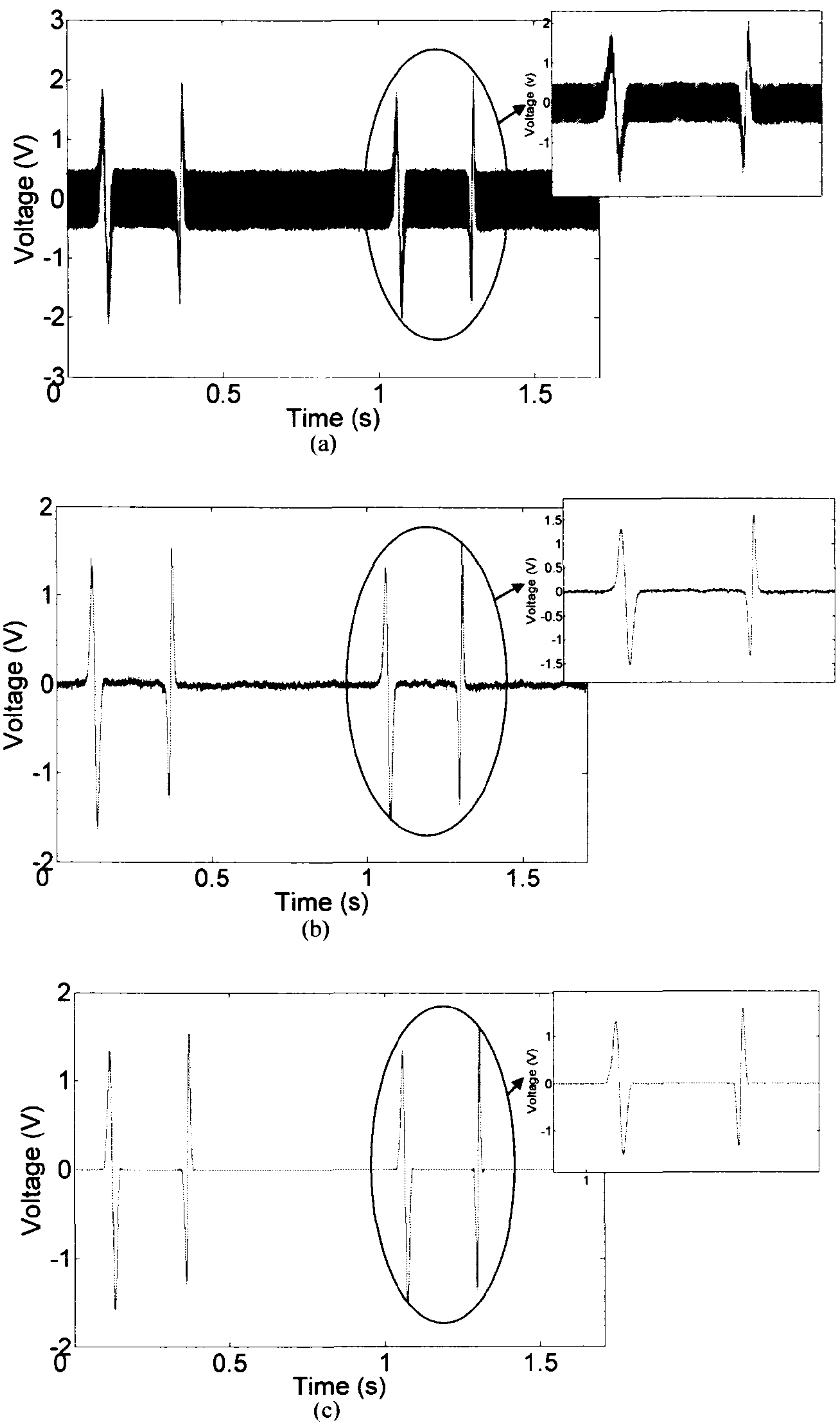


Figure 5.10 (a) Measured signal, (b) Result of the interference removal using ALE, (c) threshold de-noised version of ALE result shown in part (b) (Sampling rate: 8000 samples/s)

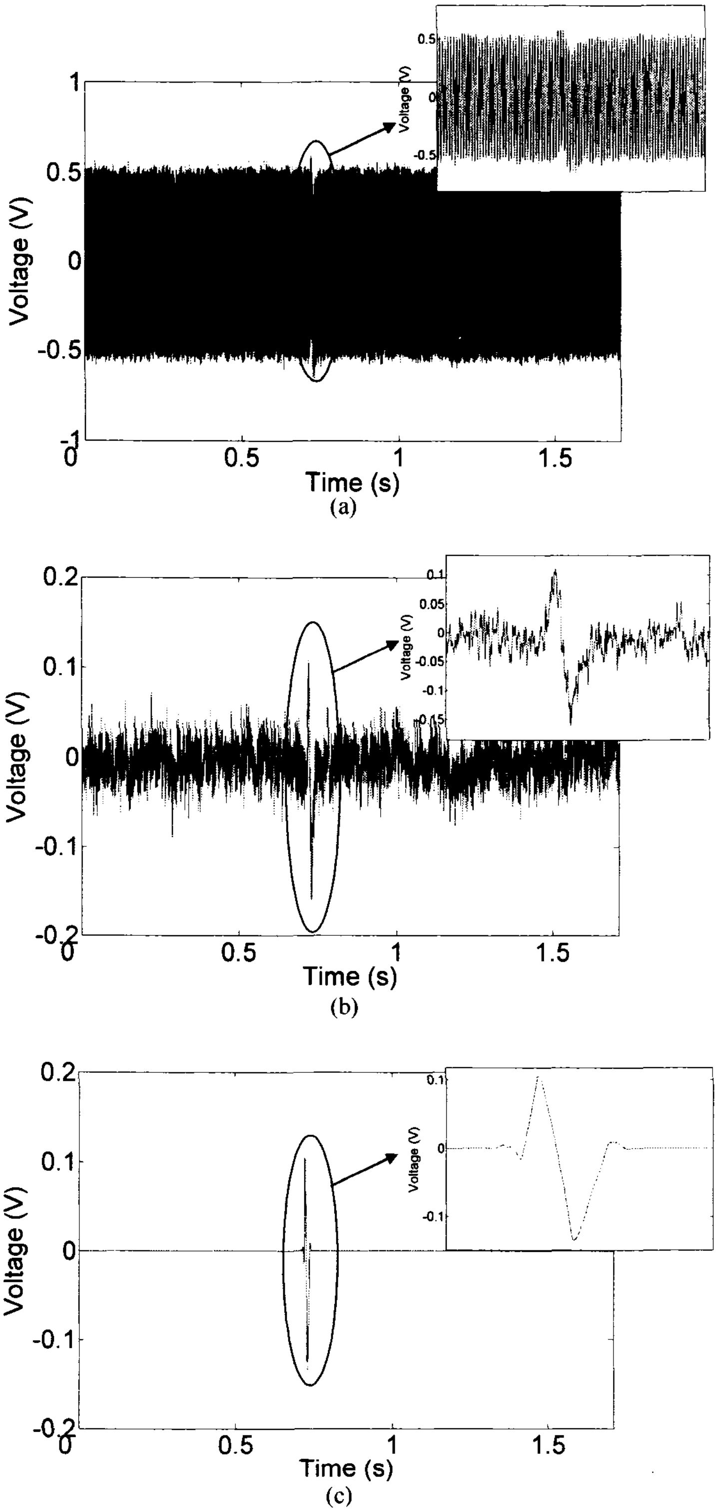


Figure 5.11 (a) Measured signal, (b) Result of the interference removal using ALE, (c) threshold de-noised version of ALE result shown in part (b) (Sampling rate: 8000 samples/s)

5.8 Conclusion

In this chapter, the ALE technique was proposed to tackle the problem of interferences that affects the performance of the oil debris sensors. The application of this technique is based on the fact that the passage of metallic particles holds a random nature and there should be no correlation between the signatures corresponding to the passing particles. On the other hand, it is assumed that the rate of change of the vibration nature of the mechanical system is slow compared to the adaptation rate of the filter. However, ALE is unable to remove the white Gaussian background noise. The signal processed by the ALE algorithm was further enhanced through a threshold de-noising scheme. The proposed two step de-noising approach has been evaluated using both simulated and experimental data and performed very well.

Chapter 6: Conclusion and Future Research

6.1 Conclusions

The main goal of this thesis is to enhance two major sources of information for machinery condition monitoring, namely vibration signal measured from faulty bearings and the signal measured from an oil debris sensor. This goal has been achieved by the development of the following.

(1) *A wavelet filter based de-noising algorithm*

This algorithm is mainly used for purification of the vibration signals. For bearings, the Gaussian background noise and the interferences from other machinery components like shaft imbalance and gear meshing form the fundamental sources that corrupt the signature of a faulty bearing. A wavelet filter based de-noising method has been proposed to remove such interferences from the measured vibration. To do so, parameters of daughter wavelet namely scale and shape factor have to be selected.

A SI-guided search approach has been proposed to find the best combination of wavelet scale and shape factor and hence the best Gabor daughter wavelet for de-noising impulsive signals measured from faulty bearings. In this study, it is proven that the SI value approaches 0.8455 for data without impulse components or with extremely low SNR. It has also been shown that the SI value decreases if a proper combination of wavelet scale and shape factor is selected. The proposed method has been tested using both simulated and experimental data and has performed very well.

Nevertheless, this approach requires a global search for minimum SI using all combinations of scale and shape-factor which is not computationally efficient. In addition, the de-noising algorithm is based on the assumption that the impulsive features of the measured vibration are formed by the vibrations of the faulty bearing while other signal components like electrical interferences from AC drives may have very similar characteristics. Furthermore, as the de-noising is performed by bandpass filtering the vibration signal, the in-band noise cannot be eliminated.

(2) *A resonance frequency estimation algorithm to select the scale and a spectral subtraction method for removing the in-band noise*

To mitigate the above drawbacks in the proposed wavelet filter based de-noising process, the algorithm is modified by incorporating a novel scale selection scheme and a spectral subtraction method. In the scale selection algorithm, prior knowledge of an important characteristic of the fault generated impulses, i.e., excitation of the resonance frequency of the bearing, structure or sensor is exploited. Shape factor is later found using the selected scale and through SI minimization.

In the modified algorithm, the spectral subtraction is applied prior to wavelet transform. The main purpose of the spectral subtraction is to eliminate the in-band noise with frequency content identical to the frequency band covered by the daughter Gabor wavelet. The power spectral density of such corrupting noise is estimated by minimizing the ratio of the geometric mean to the arithmetic mean calculated for the envelope of the reconstructed time signal. The resulting time signal is then wavelet transformed to eliminate the interfering vibration signal from other sources like shaft imbalance, gear meshing, etc.

The proposed combined spectral subtraction and wavelet de-noising algorithm was successfully tested using experimental data. The results showed good performance of the scale and shape factor selection algorithms and also illustrated that the wavelet filter based de-noising method can be improved considerably by pre-processing the signal using spectral subtraction.

(3) *An ALE algorithm for oil debris data de-noising*

For the signal measured from an oil debris sensor, vibrations of the structure where the sensor is mounted constitute the main source of interference. To remove such interferences, ALE method has been proposed. This approach is suitable because the passage of the metallic particles through the sensor holds a random nature and generally there is no correlation between the corresponding signatures. This approach has been applied to both simulated and experimental data and performed reasonably very well.

6.2 Future Research

6.2.1 De-noising faulty bearing vibrations

A close inspection of Figure 3.5, Figure 3.8d and Figure 3.11 shows a local minimum near the scale corresponding to the resonance frequency. This means that while the selected scale may not be the exact local minimum, an optimization algorithm with an arbitrary shape factor and the chosen scale as the starting point may lead us to this minimum. Preliminary results using SQP (Sequential Quadratic Programming) showed the feasibility of the method. However, problems were encountered in the convergence process. Further analyses are required in order to find a more suitable optimization algorithm and adjust the different parameters involved in the process.

Furthermore, analysis of the vibrations measured from faulty bearings with faults of different sizes showed dependence of the minimum SI on fault severity. Figure 6.1 shows minimum SI for measured vibration from faulty bearings of different fault severities versus fault size. Zero fault size represents a normal bearing. As one can see, the start of the fault is manifested by a sudden decrease in minimum SI. However, this decreasing phenomenon ends at a certain point. Further deterioration of the fault is accompanied by a reverse trend of the SI value. This may be explained by the fact that the progression of the fault is followed by flattening of the fault geometry. The flatter fault geometry leads to less severe impact and hence higher SI. A thorough research on the matter may lead to a severity estimation index composed of different indices including minimum SI.

In addition, as explained, several different criteria have been used so far to select the proper combination of scale and shape factor in wavelet filter based de-noising. A detailed comparison between these methods in different working conditions may be considered as a topic for further research.

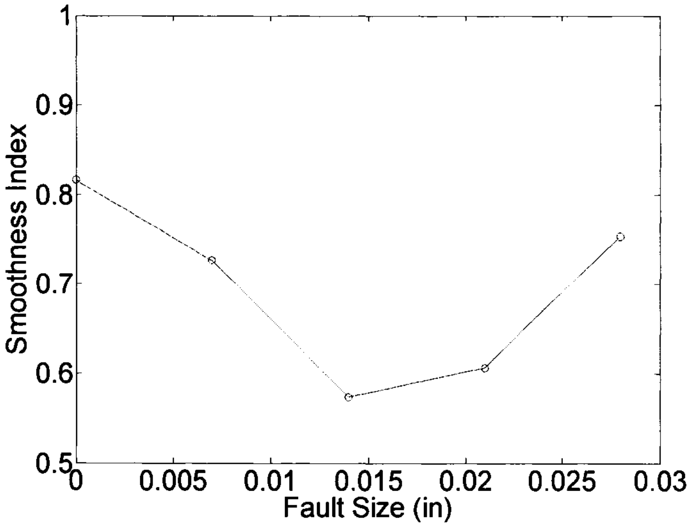


Figure 6.1 Minimum SI versus fault size

6.2.2 Enhancement of the oil debris signal

Performance of the proposed method is dependent on different parameters such as filter length, delay length and the number of averages. Further research on the matter may provide additional insight about the effect of such parameters on the performance of the algorithm and leads to adjustment guidelines. Automatic or adaptive adjustment techniques may also be the outcome of such research, which should improve the performance of the proposed method and make it more convenient for practical use.

References:

Abramowitz, M. and Stegun, A. I., 1972, *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, (New York: Dover Publications).

Antoni J. and Randal, R. B., 2006, The spectral kurtosis: application to the vibratory surveillance and diagnosis of rotating machines. *Mechanical Systems and Signal Processing*, 20(2), 308-331.

Braun, S., 1986, *Mechanical signature analysis*, (London: Academic Press).

Case Western Reserve University bearing data center, 2006, http://www.eecs.cwru.edu/ laboratory/ bearing.

Chang, S. G., Bin, Y. and Vetterli, M., 2000, Adaptive wavelet thresholding for image de-noising and compression, *IEEE Transactions on Image Processing*, 9(9), 1532-1546.

Cohen, L., 1995, *Time-Frequency Analysis*, (Englewood Cliffs, N.J.:Prentice Hall).

Coifman, R. R. and Wickerhauser, M. V., 1992, Entropy-based algorithms for best basis selection. *IEEE Transactions on Information Theory*, 38(2), 713-718.

Dempsey, J. P., 2000, A comparison of vibration and oil debris gear damage detection methods applied to pitting damage. NASA report number: TM-2000-210371, Prepared for the *13th International Congress on Condition Monitoring and Diagnostic Engineering Management*, Houston, Texas

Donoho, D. L., 1995, De-noising by soft-thresholding. *IEEE Transactions on Information Information Theory*, 41(3), 613-627.

Donoho, D. L. and Jonestone, I. M., 1994, Ideal spatial adaptation via wavelet shrinkage. *Biometrika*, 81, 425-455.

Dron, J. P., Bolaers, F. and Rasolofondraibe , L., 2004, Improvement of the sensitivity of the scalar indicators (crest factor, kurtosis) using a de-noising method by spectral subtraction: application to the detection of defects in ball bearings. *Journal of Sound and Vibration*, 270, 61-73.

Ericsson, S., Grip, N., Johansson, E., Persson, L., Sjoberg, R. and Stromberg, J., 2005, Towards automatic detection of local bearing defects in rotating machines. *Mechanical Systems and Signal Processing*, 19(3), 509-535.

Gradshteyn, I. S. and Ryzhik, I. M., 2000, *Table of Integrals, Series and Products*. 6th ed., (San Diego: Academic Press).

Grossmann, A. and Morlet, J., 1984, Decomposition of hardy functions into square integrable wavelets of constant shape. *SIAM Journal of Mathematical Analysis*, 15(4), 723-736.

Hansen, M. and Bin, Y., 2000, Wavelet thresholding via MDL for natural images, *IEEE Transactions on Information Theory*, 46(5), 1778-1788.

Harris, T. A., 1984, *Rolling bearing analysis*, 2nd ed., (New York: JOHN WILEY & SONS).

Haykins, S., 2002, *Adaptive filter theory*, (Upper Saddle River, N.J.: Prentice Hall).

Hyvarinen, A., 1999, Sparse code shrinkage: de-noising of Non-Gaussian data by maximum likelihood estimation. *Neural Computation*, 11(7), 1739-1768.

Hyvarinen, A., Karhunen, J. and Oja, E., 2001, *Independent component analysis*, (New York: John Wiley & Sons).

Industry Canada, 2007, http://strategis.ic.gc.ca/app/ccc/search/navigate.do?language=eng&portal=1&subPortal=&estblmntNo=123456056194&profile=completeProfile

Kay, S. M., 1979, The effect of noise on the autoregressive spectral estimator. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 27(5), 478-485.

Kotzalas, M. N. and Harris, T. A., 2001, Fatigue failure progression in ball bearings. *Transactions of ASME Journal of Tribology*, 123(2), 238-242.

Lapidoth, A. and Moser, S. M., 2003, Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels. *IEEE Transactions on Information Theory*, 49(10), 2426-2467.

Lim, J. S. and Oppenheim, A. V., 1979, Enhancement and bandwidth compression of noisy speech, *Proceedings of the IEEE*, 67, 1586-1604.

Lin, J. and Qu, L., 2000, Feature extraction based on Morlet wavelet and its application in mechanical fault diagnosis. *Journal of Sound and Vibration*, 234(1), 135-148.

Lin, J. and Zue, M. J., 2003, Gearbox fault diagnosis using adaptive wavelet filter, *Mechanical Systems and Signal Processing*, 17(6), 1259-1269.

Lin, J., Zue, M. J. and Fyfe, K. R., 2004, Mechanical fault detection based on the wavelet de-noising technique. *Journal of Vibration and Acoustics*, 126(1), 9-16.

Mallat , S., 1998, *A Wavelet Tour of Signal Processing*, (San Diego: Academic Press).

Marshal, A. W. and Olkin, I., 1979, *Inequalities: Theory of majorization and its applications,* (Boston: Academic press).

Markel, J. D. and Gray, A. H., 1976, *Linear prediction of speech,* (New York: Springer-Verlag).

McFadden, P. D. and Smith, J. D., 1984a, Model for the vibration produced by a single point defect in a rolling element bearing. *Journal of Sound and Vibration*, 96(1), 69-82.

McFadden, P. D. and Smith, J. D., 1984b, Vibration monitoring of rolling element bearing by the high-frequency resonance technique- a review. *Tribology International*, 17, 3-10.

McFadden, P. D. and Smith, J. D., 1985, The vibration produced by multiple point defects in a rolling element bearing. *Journal of Sound and Vibration*, 98(2), 263-273.

Miller, J. and Kitaljevich, D., 2000, In-line oil debris monitor for aircraft engine condition assessment, *IEEE Aerospace Conference Proceedings*, Big Sky, MT, USA, 6, 49-56.

Nikolaou, N. G. and Antoniadis, I. A., 2002, Demodulation of vibration signals generated by defects in rolling element bearings using complex shifted Morlet wavelets. *Mechanical Systems and Signal Processing*, 16(4), 677-694.

Papoulis, A. and Pillai, S. U., 2002, *Probability, Random Variables and Stochastic Processes*, 4th ed., (New York: McGraw-Hill).

Qiu, H., Lee, J., Lin, J. and Yu, G., 2003, Robust performance degradation assessment methods for enhanced rolling element bearing prognostics. *Advanced Engineering Informatics*, 17(3-4), 127-140.

Qiu, H., Lee, J., Lin, J. and Yu, G., 2006, Wavelet filter-based weak signature detection method and its application on rolling element bearing prognosis. *Journal of Sound and Vibration*, 289(4-5), 1066-1090.

Rao, R. M. and Bopardikar, A. S., 1998, *Wavelet Transforms: Introduction to Theory and Application*, (Reading, MA: Addison Wesley).

Rice, S. O., 1945, Mathematical analysis of random noise. *Bell System Technical Journal*, 24, 46-156.

Widrow, B., Glover, J. R., McCool, J. M., Kaunitz, J., Williams, C. S., Hearn, R. H., Zeidler, J. R., Dong, E. and Goodlin, R. C., 1975, Adaptive noise cancelling: principles and applications. *Proceeding of the IEEE*, 63(12), 1692-1716.

Widrow, B. and Stearns, S. D., 1985, *Adaptive signal processing*, (Englewood Cliffs, N.J.: Prentice-Hall).