Novel Hilbert Huang Transform Techniques for Bearing Fault Detection

By:

Shazali Osman

A thesis presented to the Lakehead University in fulfillment of the thesis requirement for the degree of Master of Science in Control Engineering

Lakehead University, Thunder Bay, Ontario, Canada

Abstract

Bearings are commonly used in rotary machinery; while up to half of machinery malfunctions could be related to bearing defects. A reliable bearing fault detection technique becomes vital to a wide array of industries to recognize an incipient bearing defect to prevent machinery performance degradation, malfunction, and unexpected breakdown. Many signal processing techniques have been suggested in literature to extract fault-related signatures for bearing fault detection, but most of them are not robust in real-world bearing health condition monitoring when signal properties vary with time. Vibration signals generated from bearings can be either stationary or nonstationary. If bearing defect-related signature is stationary, it is relatively easy to analyze using these classical data analysis techniques. However, bearing nonstationary signals are much more complex to analyze using these classical signal processing techniques, especially when slippage has occurred. Reliable fault detection still remains a challenging task, especially when bearing defect-related features are nonstationary. Two alternative approaches are proposed in this work for bearing fault detection: The first technique is based on analytical normality test, named Normalized Hilbert Haung Transform (NHHT). The second technique is based on information domain analysis, named enhanced Hilbert Haung Transform (eHHT). In the proposed NHHT technique, a novel strategy based on d'Agostino-Pearson normality analysis is suggested to demodulate feature functions and highlight feature characteristics for bearing fault detection. In the proposed eHHT, a novel strategy is proposed to enhance feature extraction based on the analysis of correlation and mutual information. The effectiveness of the proposed techniques is verified by a series of experimental tests corresponding to different bearing health conditions. Their robustness in bearing fault diagnostic is examined by the use of data sets from a different experimental setup.

Acknowledgements

First of all, I am very thankful to my Supervisor, Dr. Wilson Wang for his excellent guidance, support and patience to listen. His always-cheerful conversations, friendly behavior, and unique way to make his students realize their hidden research talents are extraordinary. I heartily acknowledge his constant encouragement and genuine efforts to explore possible funding routes for the continuation of my research studies. Great appreciation is given to my co-supervisor, Dr. Abdelhamid Tayebi, for his support, encouragement and valuable suggestions.

My sincere acknowledgements to Dr. Xiaoping Liu, and Dr. Krishnamoorthy Natarajan, for their useful suggestions and answering my queries. I would also like to thank Dr. Xiaoping Liu and Dr. Kefu Liu for their reviewing comments.

I will always remember my friends Rafeeq, Majed, and Nadeer with whom I have cherished some joyous moments and refreshing exchanges.

I wish to extend my utmost thanks to my relatives in Sudan, especially my parents and parents-in-law for their love and continuous support. Finally, my thesis would have never been in this shape without loving encouragement from my wife Mozn. Her invaluable companionship, warmth, strong belief in my capabilities, and her overall faith in me has always helped me to be assertive in difficult times. Her optimistic and enlightening boosts have made this extensive research task a pleasant journey.

Table of Contents

List of Figures

List of Tables

Chapter 1
Introduction

1.1 Overview

Rolling element bearings are used extensively in most rotating machines to support static and dynamic loads. Their performance is of the utmost importance in automotive industries, aerospace turbo machinery, chemical plants, power stations, and process industries that require precise and efficient performance. They have a great influence on the dynamic behavior of rotating machinery and act as a source of vibration and noise in these systems.

Any bearing in operation will unavoidably fail at some point. As a matter of fact, up to 50% of machinery defects are related to bearing faults [1]. Bearing defects can induce performance degradation, malfunction, and unexpected breakdown of the related machinery equipment which can also lead to economic loss and safety problems due to unexpected and sudden production stoppage. Thus, attempting to diagnose faults in complex rotary machines is often a difficult task for the operator as well as for plant maintenance. One way to increase operational reliability and thereby increase machine availability is to monitor faults in these bearings.

The recognition of incipient damage necessitates the identification of the state of a system, based on the variables monitored. The knowledge needed for such an absolute identification is often unavailable and continuous or regular measurements must be undertaken during the operation of a system. Fault diagnosis techniques are crucial for monitoring conditions in bearings. Current fault diagnosis techniques have a variety of limitations. Methods that are more effective need to be researched and developed for industrial machinery diagnostic activities.

The work presented in the adjoining sections of this thesis is a thorough investigation into application of selected Hilbert Huang based fault detection techniques to nonstationary signals collected from different bearing conditions.

This chapter provides the description of the literature review, objectives and strategies. The originality of this work and its contribution to the overall field of bearing fault diagnosis is also presented.

1.2 Literature Review

All rolling element bearing systems generate secondary effects during their operation. These secondary effects, in contrast to primary effects such as pressure and flow in pumps, are not directly utilized in the operation of the system. Examples of typical secondary effects are vibration, acoustic noise, heat generation, etc. In fact, often these secondary effects are major operational problems, and one would rather see a system operate without them. Fortunately, there is a positive side to these effects. Generated inside the rolling element bearings, they carry information about internal operation to the outside where they can be detected by sensors and analyzed to provide insight into the operation of the rolling element bearings.

Data sources for bearing condition monitoring include direct physical inspection (provided the machine can be shut down periodically for this purpose), non-destructive material inspection techniques, examination of lubricating oil and oil borne wear debris, and the analysis of dynamic values of various parameters generated while the machine is in operation [2]. Of the latter, the most commonly-used parameter is machine vibration, although noise, shaft torque and other parameters can be used when they provide significant data [3]. Analysis methods for dynamic signals, such as vibration, aim to derive the maximum condition information from the available data, and methods to enhance the usefulness of this information are constantly sought. Vibration analysis is the most popular diagnostic technique found in the case of rolling element bearings vibration analysis carried out in industries [4].

Vibration analysis is used as both a maintenance tool and as a production quality control tool for machinery systems. Vibration analysis as a maintenance tool, often called condition monitoring, enables establishment of a maintenance program based on an early warning [5]. The vibration-based analysis technique is non-intrusive and cost-effective, which makes it more attractive for defect detection [6]**.** Vibration monitoring of rolling element bearings is typically conducted using a case-mounted transducer, an accelerometer, a velocity pick-up, and occasionally a displacement sensor shown in Figure 1.1. Acceleration signals obtained from case-mounted sensors emphasize high frequency sources, while displacement signals emphasize lower frequency sources, with velocity signals falling between the extremes [7].

(a) (b) (c)

Figure 1.1. Images of case-mounted transducer: (a) accelerometer; (b) velocity pick-up; and (c) displacement sensor [8]

Currently there are many signal processing techniques in literature. Based on signal analysis strategies, these techniques can be classified into approaches in the time domain, frequency domain, and time-frequency domain analyses [9, 10].

1.2.1 Time domain techniques

When rolling elements of a bearing pass the defect location, wide band impulses are generated, and those impulses will then excite some of the vibrational modes of the bearing and its supporting structure. The excitation will result in the sensed vibration signals (waveforms) difference in either the overall vibration level or the vibration magnitude distribution, comparing them to those under fault-free conditions. Time-domain feature extraction can identify the signatures from the sensed time-domain waveforms (i.e., vibration signals) which are sensitive to bearing conditions. Depending on what underlying technology is used, time-domain feature extraction techniques can be further categorized into three groups: statistical-based, model-based, and signal processing-based approaches, all of which are detailed as follows:

(a) Statistical-based approach:

One of the most traditional time-domain feature extraction methods is to calculate descriptive statistics of vibration signals, including those measuring power content of vibration signals such as the root mean square (RMS) value; those measuring signal magnitude and pattern such as the peaks, the peak-to-peak intervals, the crest factor; and those measuring signal distribution such as the mean (1st

moment), the variance (2^{nd} moment), the skewness (3^{rd} moment), and the kurtosis (4^{th} moment) [11]. Definitions of those descriptive statistics can be found in many publications (e.g., [11, 12]) and thus will not be provided here. These descriptive statistics can be calculated directly on raw signals. However, for the descriptive statistics to be more effective in bearing condition monitoring, they are frequently calculated on filtered or processed signals and are sensitive to load and condition variations. These non-dimensional statistical parameters are very effective in identifying incipient fatigue spalling. But sometimes these parameters fail to indicate the defects due to the development of the failure and load/condition variation. For example, if the defect becomes severe, the crest factor and kurtosis value will reduce to normal values, and therefore will not be very reliable and will be unable to be used in isolation. Moreover, they cannot be used to directly indicate the location of the defect.

(b) Model-based approach:

Model-based feature extraction involves treating vibration signals as time series data and fitting them into a parametric time series model. The model parameters are then used as features. The most popular time series model used for bearing diagnosis is the autoregressive (AR) model. In [13], H. Endo *et al*. applied the AR model to vibration signals collected from an induction motor and used the AR model coefficients as extracted features. Other time series models such as the autoregressive moving average (ARMA), and other non-linear models such as neural networks and support vector machines, can also be used. Recent direction for model-based feature extraction seems to extend model-based approaches that work for both stationary signals and nonstationary signals [14].

(c) Signal processing based approach:

Classical digital signal processing includes filtering, averaging, correlation, and convolution. Another popular digital signal processing (DSP) technique is synchronous averaging [15]. More recently, several techniques rooted in chaos theory have been adapted to feature extraction, for example, correlation dimensions [16]. A useful technique in defect detection is the synchronous signal averaging technique (SSAT) [17]. The result of the SSAT is the signal average, which is the ensemble average of the angle domain signal, synchronously sampled with respect to the rotation of one particular shaft. In the resulting averaged signal (SA), the random noises as well as non-synchronous components are attenuated. The main advantage of the SSAT is the possibility to extract a simpler signal related to the gear of interest from a complex gearbox vibration signal. However this technique has a pivotal drawback related to the complexity of the measurement equipment, which requires an

additional sensor to measure the rotational shaft speed. In addition, the SA can be band-pass-filtered at the dominant meshing harmonic, and the application of the proper choice of transform function provides both amplitude and phase modulation functions. If the bandwidth of the bandpass filter is properly chosen, the demodulated signal will carry the information related to the gear fault [18].

1.2.2 Frequency domain techniques

Potential defects can be analyzed by the frequency domain spectrum of the vibration signal. In order to calculate the frequency spectrum of a sampled time signal, the Fast Fourier transform (FFT) algorithm can be used as a numerically efficient method [17]. It is important to note that all digital FFT methods assume stationary signals, periodic in the time window or single transients. Discrete Fourier transform (DFT) analysis of the time waveform has become the most popular method of deriving the frequency domain signal. The signature spectrum obtained can provide valuable information with respect to machine and bearing conditions [19]. Spectral analysis and spectrum comparison are commonly used frequency domain techniques. Envelope analysis or demodulating the time waveform prior to performing the FFT is also gaining popularity. Time-domain features are generally considered to be good for fault detection, but less effective for fault isolation, i.e., to determine the location of the defect, inner race, outer race, rolling elements, and cage [20]. For fault isolation, frequency-domain features are generally more effective. Frequency-domain feature extraction methods include spectral analysis, envelope analysis, cepstrum, and higher-order spectra, which are briefly discussed as follows:

(a) Spectral Analysis:

The most popularly-used method is the spectral analysis. A spectrum (more practically power spectrum) obtained from FFT of a vibration signal represents characteristic frequency of the signal. Either the entire spectrum or the frequency amplitudes at the bearing characteristic frequencies calculated from the power spectrum of vibration signals can be used as features. Power spectrum can be used to identify the location of the defect by relating the defect characteristic frequencies to the major frequency components in the spectrum [21]. For bearing fault detection and diagnosis, detailed knowledge of the bearing defect characteristic frequencies is required, which will be discussed next. Power spectral density (PSD) is considered to be one of the spectra smoothing techniques [22].

(b) Envelope Analysis:

Also known as amplitude demodulation or high frequency resonance technique (HFRT), envelope analysis is another widely-used frequency domain technique for bearing fault diagnosis [23]. Envelope analysis consists of two steps: band-pass filtering and enveloping. During bearing operation, wide band impulses are generated when rolling elements pass over the defect. Certain vibration modes of the bearing and its supporting structure will be excited by the periodic impulses. Band-pass filtering allows keeping only signal components around the resonance frequency. Enveloping then removes the structural resonance and preserves the defect impact frequency [17]. Thus, envelope analysis can be used for detecting incipient faults of bearings. The key to effective envelope analysis is strategically selecting the frequency band [24].

(c) Cepstral Analysis:

Cepstrum is defined as the power spectrum of the logarithm of the power spectrum of the signal, which can be used for detecting the periodicity of spectra. A defect in a bearing element (ball and races) generates impulses and the bearing and its structure respond to those impulses. Bearing vibration signals thus are the result of convolution between impulses and the system's response to these impulses, which lead to harmonic series in the spectra [10]. Cepstral analysis detects common spacing between the harmonics and has been used for bearing fault detection and diagnosis as demonstrated in [25].

(d) Higher Order Spectra:

This typically refers to bispectrum and trispectrum [26]. Higher order spectra are also called higher order statistics, since bispectrum and trispectrum are essentially the FT of the 3^{rd} and 4^{th} order statistics of signals. Higher order spectra (i.e., bispectrum or trispectrum) have proven to have more diagnostic information [10]. Advantages for using higher order spectra include additive Gaussian noise suppression, non-minimum phase system identification, non-linear systems detection and identification. Li *et al*. [27] presented bicoherence signal analysis for detection of faults in bearings. The rationale behind the bicoherence analysis is that interactive coupling between various frequencies and existing bearing fault frequencies can be amplified and detected by monitoring the statistical dependence or correlations between the energies in the corresponding frequency-combinations.

1.2.3 Time-frequency analysis techniques

Time-frequency analysis techniques analyze signals in both time and frequency simultaneously for identifying time-dependent variations of frequency components within the signal, which makes time-frequency analysis techniques a powerful tool for analyzing nonstationary signals. The most commonly used time-frequency analysis techniques for analysis of vibration signals of ball bearing for condition monitoring purposes are the short-time Fourier transform (STFT) [19], the Wigner-Ville distribution (WVD) [28], the wavelet transform (WT) [21, 29], and the Hilbert Huang transform [30, 31]. In this thesis, wavelets and HHT functions are categorized as separate groups due to their increasing popularity and various types of function used.

Other newly-developed time-frequency analysis techniques include spectral kurtosis and cyclostationary analysis. The first two approaches of STFT and WVD have many limitations that prevent their effectiveness. Such limitations represented in the resolutions between the two domains are due to a constant width of time and frequency windows when using STFT analysis (since the product of time and frequency resolution is a constant), as well as the possibility of negative energy levels and non-physical interference (cross) terms in the WVD, which can also result in signal deterioration. In the case of the STFT, the choice of the optimum window characteristics according to specific criteria can dramatically reduce the limitation imposed by the resolution. Such optimization has been shown in the case of the STFT-based Kurtogram [19, 24]. Due to the nature of wavelet, it also has its disadvantages such as the fact that the phase spectrum of the wavelet is not robust to noise once the signal is contaminated by noise, and also suffers from the shortcoming of border distortion and energy leakage. Moreover, overlapping will cause frequency aliasing and add interference term to scalogram, which usually occurs as a result of convolution.

The Wavelet transform is an advanced signal processing technique with a growing number of applications in machine fault diagnosis [29, 30]. Wavelet analysis has been established as one of the most suitable time-frequency analysis techniques due to its flexibility and efficient computational implementations, but in particular because of its inherent constant percentage bandwidth structure, which is appropriate to signals dominated by impulse responses at different frequencies. The use of wavelet analysis in detecting rolling element bearing fault has been investigated by different researchers such as J. Liu [21], who used continuous wavelet analysis to detect localized bearing defects based on vibration signals. The wavelet transform is the leading time-frequency signal analysis

method, and is commonly applied in the analysis of nonstationary signals; it, has been proven to be a very effective diagnostic tool. In 1996, W. J.Wang and P. D. McFadden [15] applied time synchronous averaging (TSA) and the continuous wavelet transform (CWT) to vibration data from a helicopter gearbox with a fatigue crack. The authors proposed a computation algorithm and demonstrated the first use incorporating a variable time-frequency resolution depending on the frequency for gear fault detection. Diagnosis was performed by analysis of the resulting wavelet scalogram contour maps. The improvement that the WT makes over the STFT is that it can achieve high frequency resolutions with sharper time resolutions.

The conventional wavelet transform is based on real valued wavelet function and scaling function, but the Morlet wavelet transform is a complex wavelet transform. Complex wavelet can solve some of the issues with which we are faced when using real wavelet, such as:

- Oscillation - since the mother wavelet is a band-pass filter, the wavelet coefficients will oscillate around singularities. It complicates the wavelet based processing;
- Shift variance - a small shift in the signal will greatly perturb the wavelet coefficient oscillating around singularities;
- Aliasing - the signal resulted after the wavelet transform will result in substantial aliasing. Only when the wavelet and the scaling coefficients are not changed, inverse DWT will cancel this aliasing. Upsetting the delicate balance between the forward and inverse transforms will lead to artifact errors in the reconstructed signal. Further discussions of related WT and details about implementation of the HHT method used in time-frequency domain are found in Chapter 2 and Chapter 3 respectively.

Vibration signals acquired from bearings can be either stationary or nonstationary. While stationary signals in the case of a defect on a fixed race can be characterized by time-invariant statistical properties such as the mean value, or statistical properties, signals such as defects on a rotating race or a rolling element defect are often considered nonstationary, especially within a short time frame for computational convenience. For nonstationary signals, since the statistical properties change over time, traditional spectral analysis becomes ineffective. Vibration signals from real-world bearings are almost always nonstationary since bearings are inherently dynamic (e.g., speed and load

condition change over time). Techniques used for tackling nonstationary signals include signal processing techniques [29, 30, 31].

1.3 Objectives and Strategies

The main goal of this research was to develop a novel signal processing method to enable proper diagnosis of rolling element bearing. The work program comprised of synthesis techniques from the fields of signal processing and pattern recognition with application to bearing condition monitoring.

To tackle the aforementioned challenges, the objectives of this work are to develop new techniques for more reliable fault detection in rolling element bearing. The focus will be initial bearing defects and nonstationary feature analysis, for which two approaches will be proposed. The first approach is a normalized Hilbert-Huang transform (NHHT) technique for nonstationary feature analysis and bearing fault detection. The proposed technique will be new as a novel intrinsic mode function IMF integration method based on analysis of normality information; its purpose is to enhance the distinctive IMFs for representative feature extraction and for bearing fault detection. The second approach is to develop an enhanced HHT (eHHT) technique for bearing fault detection. This proposed technique is new in the following aspects: (1) the measured vibration signal is denoised properly to improve signal-to-noise ratio; and (2) another novel approach based on the analysis of correlation and discrepancy of mutual information (MI) to choose the most distinctive functions for representative feature extraction.

Figure 1.2 illustrates the fault detection processes using the proposed techniques. Firstly, the measured vibration signal is collected by a data acquisition system, which is then denoised by the use of a minimum entropy deconvolution (MED) filter in the case of applying eHHT technique. Then, the signal is processed using the proposed techniques for bearing fault detection.

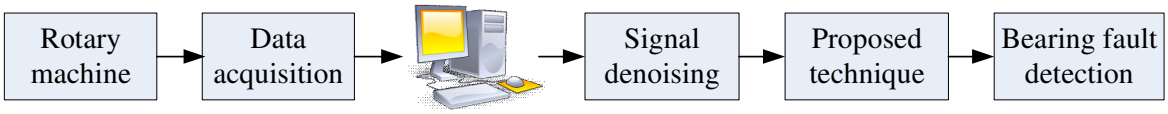


Figure 1.2. Flowchart of proposed technique for bearing condition monitoring.

1.4 Thesis Outline

This thesis report is organized as follows: theoretical background and fundamental methods for processing rolling element bearings used in industry, as well as bearing properties and basic geometry used to determine possible defects frequencies that can occur, all of which are provided in Chapter 2.

Chapter 3 presents the Normalized Hilbert Haung transform (NHHT) processing technique. This chapter also represents the improvements to the classical Hilbert Huang analysis to extract the representative features related to the bearing health conditions. The related result are demonstrated and compared to other popular signal processing techniques.

Chapter 4 discusses the minimum entropy deconvolution (MED) filter used to denoise collected signals that are further processed by the enhanced Hilbert Haung transform (eHHT) technique. The investigation results will demonstrate the enhanced Hilbert Haung transform (eHHT) as an effective signal processing technique for bearing fault detection, which is especially useful for nonstationary feature. The verification results are also compared to other signal processing techniques.

Chapter 5 demonstrates the robustness results from using data sets acquired from a different experimental setup examined for both NHHT and eHHT methods. In addition, multi-defect testing is conducted with results obtained, presented, and discussed.

At the end, Chapter 6 summarizes the importance of this work and conclusions are derived from the whole work. Possible future improvements and additions to the project such as decision-making, signal conditioning, and implementation issues are provided.

Chapter 2
Theoretical Background

2.1 Ball Bearing Geometry and Characteristic Frequencies

Based on rolling element structures, rolling element bearings can be grouped into two main types: (a) the ball bearing, which has point contact; and (b) the roller bearing, which provides line contact on both raceways. In general, rolling element bearings are designed to carry an axial and/or radial load while minimizing the rotational friction by placing rolling elements such as cylinders or balls between inner and outer races. There are different types of rolling element bearings, however ball bearings are the most economical since balls rather than cylinders are used in their construction. There are also different types of ball bearings such as thrust, axial, angular contact, and deep groove ball bearings. The measurement data sets used in this thesis work are from deep groove ball bearings.

A ball bearing is generally comprised of four principal parts: an inner ring or race, an outer ring or race, a ball complement, and a ball separator or cage. An example of a typical structure of deep groove ball bearing is provided in Figure 2.1. The inner race is fastened to the shaft and is grooved on its outer diameter to provide a circular ball raceway. The outer ring is mounted in the housing and contains a similar grooved circular ball raceway on its inner diameter. Normally the inner race carries the rotating element, but in some applications the inner race may be stationary and the outer race may carry the rotating element. Most of the defects on the fixed race such as cracks or pits occur at the locations related to the load zone since they are directly under the applied force. The balls serve to space the inner and outer raceways apart to provide for smooth relative motion between them. The cage serves to keep the balls' uniformity spaced inside the bearing, preventing them from rubbing against each other or bunching up on one side of the bearing. The rotating race and rolling element faults, on the other hand, can occur anywhere since the race is not stationary and rotating.

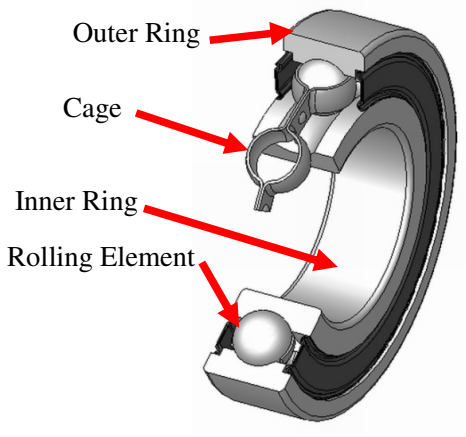


Figure 2.1. Ball bearing structure [32].

Bearing defects can be classified into distributed and localized faults. One of the basic mechanisms that is directly related to the operational conditions and initiates a localized defect is due to the Hertzian contact stresses between the rolling elements and the rings. This leads to bearing damage in due course, resulting in micro-pitting, smearing, indentation and plastic deformation, besides surface corrosion. Localized defects include cracks, pits and spalls in rolling surfaces. The presence of a defect causes a significant increase in vibration levels, which is used as the feature to be sought for diagnosing different machinery faults. Distributed defects include mainly wear, surface roughness, waviness and misaligned races. In application, many distributed defects originate as localized defect.

In general, the main concern for bearing health condition is to recognize a bearing defect at its earliest stage in order to prevent machinery performance degradation, where distributed faults originate from localized bearing defects (e.g., a pit or spall), therefore this research will focus on localized bearing faults.

As discussed in Section 1.2, currently the most commonly used technique is related to spectral analysis of bearing signals, in terms of characteristic frequency analysis. Consider a bearing as shown in Figure 2.2 with the following geometry; D: outer diameter; d_m: pitch diameter; d: bore diameter; d_b: ball diameter, w: raceway width; and α: is the contact angle.

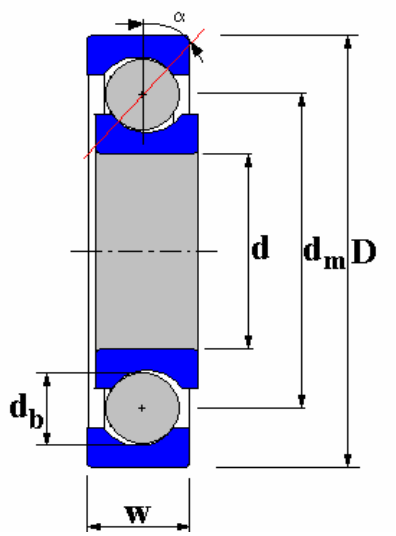


Figure 2.2. Geometry of ball bearing [33].

The cage diameter $\boldsymbol{d_c}$ can be approximated by

$$d_c = (d_i + d_o)/2 \tag{2.1}$$

where d_i and d_o are the inner and an outer race diameters respectively

Since the bearing may be loaded at an angle α from the radial plane, the effective inner and outer raceway diameters $d_{i;eff}$ and $d_{o;eff}$ can be expressed as

$$d_{i;eff} = d_c - d_b \cos\alpha \tag{2.2}$$

and

$$d_{i;eff} = d_c + d_b \cos\alpha \tag{2.3}$$

The frequency at which any rolling element passes a specific defect point on one of the races is called the ball pass frequency (BPF). For an outer race defect, this frequency is

$$BPFO = f_r (z/2)(1 - \frac{d_b}{d_m}\cos\alpha) \tag{2.4}$$

where z is number of balls; f_r is shaft frequency.

And for an inner race, the defect characteristic frequency is defined as

$$BPFI = f_r(z/2)(1+\frac{d_b}{d_m}\cos\alpha) \tag{2.5}$$

Also of interest for condition monitoring is the ball spin frequency of rolling element, which is expressed as

$$BPFR = f_r(z/2)\left((1+\frac{d_b}{d_m}\cos\alpha)^2\right) \tag{2.6}$$

As demonstrated by the aforementioned equations, the characteristic defect frequencies depend on kinematic principles such as the rotational speed and the location of the defect in a bearing. The presence of the defect frequencies in the direct or processed frequency spectrum is the sign of the fault; the signature of the defected bearing is spread across a wide frequency bandwidth and can easily be masked with low frequency machinery vibration and noise. The consecutive impact between the defect and rolling elements excites the resonances of the structure and the resonant frequencies dominate the frequency spectrum. Therefore, the characteristic defect frequencies cannot be easily detected because of their low amplitude with respect to resonant amplitudes.

Vibration has been used to determine the mechanical condition of machinery and their parts over the last fifty years. Many researchers have attempted different approaches and descriptors under various environmental conditions and tried to investigate the relationship between the tested bearing and changes in vibration response under operating condition [9]. Peak acceleration, RMS of overall acceleration, crest factor, spike energy, shock pulse, acoustics emission, cepstrum, and statistical descriptors such as mean, standard deviation, skewness and kurtosis have also been reported in evaluating the bearing condition [7, 20]. Details about adaptive systems and signal denoising methods will be provided next and in the following chapters.

2.2 Wavelet Analysis

Wavelet is defined as a waveform of the effectively limited duration that has an average value of zero. Unlike the FT analysis in which a signal is decomposed into its harmonic using global sinusoidal functions that go on forever, in wavelet analysis the signal is broken down into a series of

local basis functions called wavelets. Each wavelet is located in a different position on the time axis and is local in the sense that it decays to zero when sufficiently far from its center. At the finest scale, wavelets may be very short; at a coarse scale, they may be very long. The higher the resolution in time is required, the lower resolution in frequency has to be (Figure 2.3a). The larger the extension of the analysis windows is chosen, the larger is the value of Δt (t is used here to describe the time difference). This is demonstrated in Figure 2.3b.

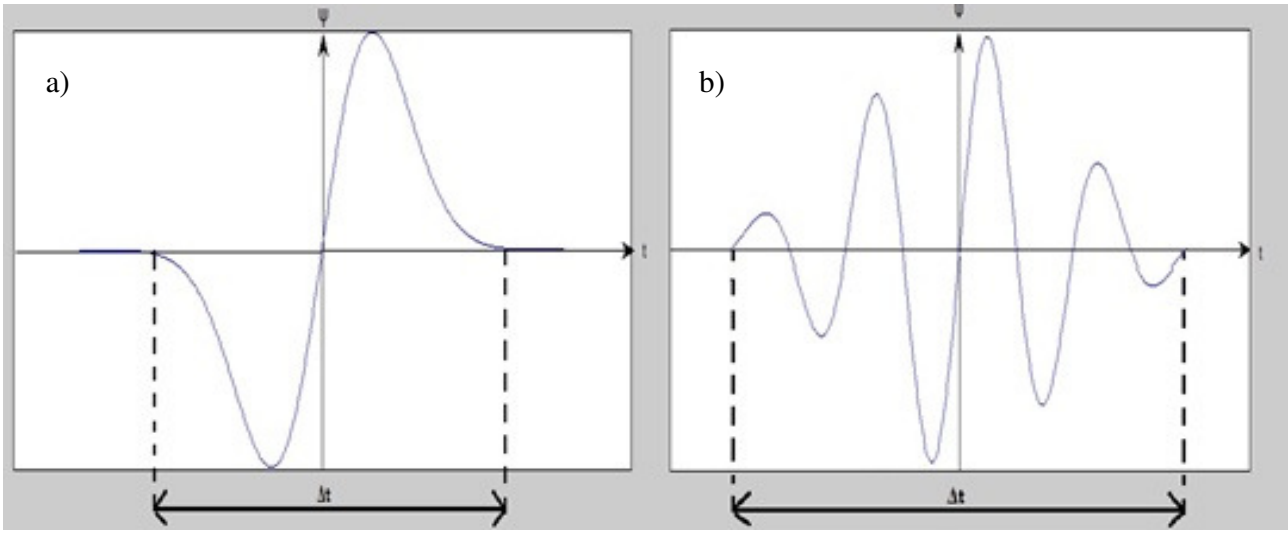


Figure 2.3. Wavelet transformation (WT): compared basis functions with compression factor [34].

A particular local feature of a signal can be identified from the scale and position of the wavelets into which it is decomposed. Wavelets are used as the reference in wavelet analysis and are defined as signals with two properties: admissibility and regularity. Admissibility refers to the referenced wavelet (or the mother wavelet), and must have a zero average in the time domain which implies that wavelets must be oscillatory [35]. Regularity specifies that wavelets have some smoothness and concentration in the time-frequency domains, which means that wavelets are oscillatory and compact signals. The wavelet transform is the leading time-frequency signal analysis method, and has been applied in the analysis of nonstationary signals [36]. Wavelet transform can be classified as: Continuous WT (CWT), Wavelet Packet Transform (WPT), and Discrete Wavelet Transform (DWT) [37], all of which will be briefly discussed as follows:

2.2.1 Continuous wavelet transform (CWT):

Given a continuous signal, the continuous wavelet transform possesses the ability to construct a time-frequency representation of that signal and offers very good time and frequency localization. The continuous wavelet transform is defined by

$$X_w(a,b) = \frac{1}{\sqrt{b}} \int_{-\infty}^{\infty} x(t)\psi^*\left(\frac{t-a}{b}\right)dt \begin{cases} a \in (-\infty, \infty) \\ b \in [0, \infty) \end{cases} \tag{2.7}$$

where $\psi^*(t)$ is the complex wavelet conjugate of the mother wavelet $\psi(t)$, a is the location parameter translation and is a real number, and b is the scaling (dilation) parameter and also a real number.

A mother wavelet must satisfy the following criteria

a) admissibility criterion is defined by

$$c_\psi = \int_0^{\infty} \frac{|\Psi(f)|^2}{|f|} df < \infty \tag{2.8}$$

where $\Psi(f)$ is the FT of the mother wavelet $\psi(t)$ so the inverse WT can exist. The inverse wavelet transform is given as follows

$$x(t) = \frac{1}{c_\psi} \int_0^{\infty} \int_{-\infty}^{\infty} \frac{1}{b^{5/2}} X_w \psi\left(\frac{t-a}{b}\right) da db \tag{2.9}$$

The Scalogram, i.e., absolute value and square of the output of the wavelet is defined by

$$Sc_x(a,b) = |X_w(a,b)|^2 = \frac{1}{|b|}\left|\int_{-\infty}^{\infty} x(t)\psi^*\left(\frac{t-a}{b}\right)dt\right|^2 \tag{2.10}$$

b) compact support and vanishing moments of the width of the support is not infinite, while support of the region of the mother wavelet is not equal to zero. Since all signals in nature can be represented by a polynomial, we can define the *k*th moment by

$$m_k = \int t^k \psi(t) dt \tag{2.11}$$

If $m_0 = m_1 = m_2 = \cdots = m_{p-1} = 0$, we can conclude that $\psi(t)$ has p vanish moment. At least the vanishing moment is not less than one, such that,

$$\int \psi(t) dt = 0 \tag{2.12}$$

Finally, the higher the order of vanishing moment, the higher the frequency function of the mother wavelet, and the more sensitive this wavelet to process with the high frequency. In this work, we are more interested in the definition and application of complex Morlet wavelet transform function.

A complex wavelet is defined as

$$\psi_c(t) = \psi_r(t) + j\psi_i(t) \tag{2.13}$$

where $\psi_r(t)$ is real and even, $\psi_i(t)$ imaginary and odd also $\psi_i(t)$ is the Hilbert transform of $\psi_r(t)$. The complex scaling is defined similarly. The choice of complex mother wavelet and complex scaling function is another important issue.

As an example, the Morlet wavelet is one of the most popular complex wavelets in application [38], which will also be used in this work, which is defined as

$$\psi(t) = \frac{1}{\sqrt[4]{\pi}} \left(e^{jw_0 t} - e^{\frac{w_0^2}{2}} \right) e^{\frac{t^2}{2}} \tag{2.14}$$

where w_0 is the central frequency of the mother wavelet. $e^{\frac{w_0^2}{2}}$ is used to correct the non-zero mean of the complex sinusoid, which can be negligible if $\psi_c = \psi_0 > 5$. The Morlet wavelet has similar form to Gabor transform, but with a major difference whereby the size of window in Gabor is fixed with respect to scaling parameters. The Morlet wavelet is found to be useful in bearing and machine condition monitoring [39], due to some of its specific characters such as:

- It can be used to extract impulse component because Morlet wavelet is more similar to an impulse;
- It can be adjusted to adapt to impulses with decaying rate;

- When combined with a proper denoising method, Morlet WT could be effective in analyzing nonstationary features related to mechanical faults.

2.2.2 Wavelet packet transform (WPT):

Different from the continuous wavelet transform, CWT applies the wavelet transform to the low pass result; the wavelet packet transform applies the wavelet transform step to both the low pass and the high pass result. WPT can be considered as a generalized form of the time-frequency analysis of the wavelet transform. It yields a family of orthonormal transform bases. By filtering the wavelet spaces, it can partition phase space in different ways. For example, Haar transform filters can be obtained by:

$$A_{j+1,i} = \left(even_{j,i}(t) + odd_{j,i}(t)\right)/2 \qquad \text{Haar Scaling (Low Pass) function} \qquad (2.15)$$

$$D_{j+1,i} = \left(even_{j,i}(t) - odd_{j,i}(t)\right)/2 \qquad \text{Haar Scaling (High Pass) function} \qquad (2.16)$$

where $A_{j+1,i}$ and $D_{j+1,i}$ are the low pass and high pass function respectively.

WPT is proven to be useful in bearing condition monitoring when it is used in combination with a suitable chosen cost function for the best basis algorithm [36]. The concern with wavelet packet is a potentially larger computational load, as well as lower sensitivity to maintain high probability of false detection [40, 41].

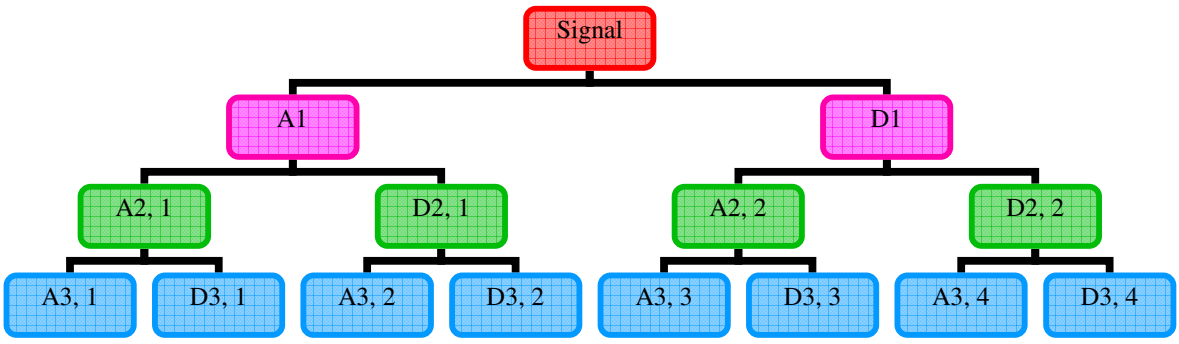


Figure 2.4. Envelope determination using WPT method.

2.2.3 Discrete wavelet transform (DWT):

The DWT is a subset of the WPT which generalizes the time-frequency analysis of the WT. DWT is the implementation of the wavelet using a discrete set of wavelet scales and translation, obeying the defined rules that were mentioned previously. This transform decomposes the signal into a mutually orthogonal set of wavelets, which is considered to be the main difference from the CWT. DWT has been applied in [42, 43] to detect incipient bearing faults.

2.3 Minimum Entropy Deconvolution (MED) Method

If a bearing is damaged (e.g., a fatigue pit on the fixed ring race), impulses are generated whenever rolling elements strike the damaged region. Due to the impedance effect of the transmission path, the measured signal using a vibration sensor, is a modulated signature of the defect-related impulses. To highlight defect-related impulses, a denoising process is applied using the MED filter.

The MED was originally proposed by Wiggins for deconvolving the impulsive sources from a mixture of signals [44]. The MED has shown its effectiveness in highlighting the impulse excitations from a mixture of responses [44], which has also been used for machinery system condition monitoring. For example, Endo *et al.* combined the MED, autoregressive models and wavelet analysis for fault detection in gear systems [13] and bearings [20]. The MED filter aims at highlighting impulses while minimizing the noise (i.e., entropy) associated with signal transmission path [45, 46]. Entropy minimization is achieved by maximizing signal kurtosis, which is sensitive to impulse-induced distortion in the tails of the distribution function.

Figure 2.5 illustrates the MED signal denoising operations. The signal x represents the original form of the defect impulses, and N_s represents the random noise interference. The structure filter g represents the impedance effect of the transmission path from the impulse to the measurement sensor, and $\otimes$ presents the convolution operation.

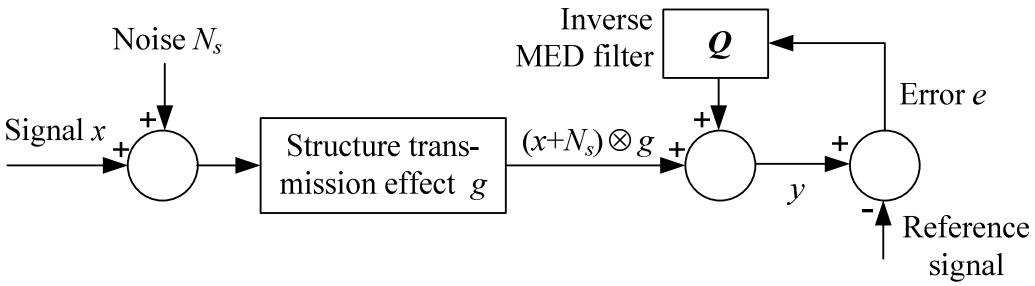


Figure 2.5. Inverse filtering process of MED.

The objective of the inverse MED filter $\boldsymbol{Q}$ is to find an optimal set of filter coefficients vector $\boldsymbol{q}$ to recover the original impulse signal by maximizing kurtosis or minimizing entropy. The kurtosis is determined as the 4th order statistic measurement of a signal (an objective function) such as

$$O_4(\boldsymbol{q}(l)) = \sum_{i=1}^{N} y^4(i) \Big/ \left[\sum_{i=1}^{N} y^2(i)\right]^2 \tag{2.17}$$

where y is the output signal using the inverse MED filter $\boldsymbol{Q}$ and N is the length of the signal.

The optimal filter coefficient vector $\boldsymbol{q}$ is achieved by optimizing the kurtosis of the objective function in equation (2.17), which is achieved by letting

$$\partial(O_4(\boldsymbol{q}(l))) / \partial(\boldsymbol{q}(l)) = 0 \tag{2.18}$$

The convolution of the inverse filter is generally given by

$$y(j) = \sum_{l=1}^{L_m} \boldsymbol{q}(l) z(j-l) \tag{2.19}$$

where z is the observed signal and L_m is the length of the MED filter. Delay l is used to make the inverse filter causal.

By using $\partial(y(j)) / \partial(q(l)) = z(j-l)$ and combing Equations 2.18 and 2.19 yield

$$\left[\sum_{i=1}^{N} y^2(i) \Big/ \sum_{i=1}^{N} y^4(i)\right] \sum_{i=1}^{N} y^3(i) z(i-l) = \sum_{p=1}^{L_m} \boldsymbol{q}(p) \sum_{i=1}^{N} z(i-l) z(i-p) \qquad (2.20)$$

Equation (2.20) can also be represented by $\boldsymbol{B} = \boldsymbol{A}\boldsymbol{q}$. where $\boldsymbol{B}$ on the left-hand side of equation (2.20); $\boldsymbol{A} = \sum_{i=1}^{N} z(i-l) z(i-p)$ is the Toeplitz autocorrelation matrix of observed signal z.

The MED is conducted by the use of the following algorithm [20, 44]:

1) Set the initial Toeplitz autocorrelation $\boldsymbol{A}$, then initialize the filter coefficient $\boldsymbol{q}^{(0)}$ as the delayed impulse. The autocorrelation matrix $\boldsymbol{A}$ is calculated once and is used repeatedly in the following iteration operations.

2) Determine the output signal $y^{(0)}$ using the input signal $z^{(0)}$ in Equation (2.19).

3) By using $\boldsymbol{B}^{(1)}$ obtained in equation (2.20) to determine the new optimal filter coefficients $\boldsymbol{q}^{(1)}$ by

$$\boldsymbol{q}^{(1)} = \boldsymbol{A}^{-1} \boldsymbol{B}^{(1)} \qquad (2.21)$$

4) Compute the error from the changes in filter coefficient values as

$$e = \left(\boldsymbol{q}^{(1)} - \mu \boldsymbol{q}^{(0)}\right) / \mu \boldsymbol{q}^{(0)} \qquad (2.22)$$

where $\mu = \left(E\left(\boldsymbol{q}^{(0)}\right)^2 / E\left(\boldsymbol{q}^{(1)}\right)^2\right)^{1/2}$ and $E(*)$is the expectation operator.

In operation, if $E(e) > T_E$ ($T_E = 0.01$ in this case), update the filter coefficients by repeating the aforementioned processes starting from step 1) to step 4). Otherwise if $E(e) \le T_E$, terminate the iteration. Additional condition to terminate the iteration process is set for the convergence of the algorithm over specified iterations if it is found to be diverging.

2.4 Shannon Entropy

Shannon entropy [47, 48] is a method of measuring the peakedness of a distribution, where the parameter is chosen according to the minimal Shannon entropy criterion using the following equations:

$$SE = -\sum_{k=1}^{N_{dm}} O_k \log(O_k) \tag{2.23}$$

where $O_k = \dfrac{|S_{ki}|}{\sum_{k=1}^{N_{dm}} S_k}$ (2.24)

S_k (k = 1, 2, …, N_{dm}) is the transform function (i.e., IMF in this application). Shannon entropy function has been proven to be an effective technique in [43] to synthesize the resulting wavelet coefficients in bearing fault detection.

2.4 Spares Shrinkage Code

Spares shrinkage code [49, 50] is a method of estimating the non-Gaussian data in noisy conditions. It can eliminate in-band noise and increase the signal-to-noise-ratio (SNR) of the measured signal. To present a sparse distribution, Hyvarinen [50] proposed the following function to model the super-Gaussian signal:

Given a signal x, its probability density can be obtained by

$$p(x) = \frac{1}{2\sigma} \frac{(\delta + 2)[\delta(\delta + 1)/2]^{\left(\frac{\delta}{2}+1\right)}}{\left[\sqrt{\delta(\delta + 1)/2} + |x/\sigma|\right]^{(\delta+3)}} \tag{2.25}$$

where δ is the parameter controlling the sparseness of the probability density function (PDF), and σ is the standard deviation. Based on the prior distribution in equation (3.19), $\hat{x}$ can be estimated using the following maximum likelihood estimate:

$$\hat{x} = sign(y) \max\left(0, \frac{|y| - a_s\sigma}{2} + \frac{1}{2}\sqrt{(|u + a_s\sigma|)^2 - 4\sigma^2(\delta + 3)}\right) \tag{2.26}$$

where $a_s = \sqrt{\delta(\delta+1)/2}$. This method is usually used to cancel the residual noise after implementing the WT analysis [50].

Chapter 3

Proposed Normalized Hilbert Huang Transform (NHHT)

3.1 Analysis of the Classical HHT

Hilbert-Huang Transform (HHT) is a data analysis tool developed in 1998 [51, 52], which can be used to extract the periodic components embedded within oscillatory data. It carries on the HHT after the signal is processed to obtain the signal energy integrity, and the precise time frequency distribution, simultaneously. It is also one auto-adaptive signal processing method, extremely suitable for the non-linearity and the non-steady process.

HHT includes the empirical mode decomposition (EMD) and Hilbert spectral analysis. In this method, arbitrary nonstationary signals were divided into a series of data sequences with different characteristic scales. After the first EMD treatment, each sequence is called an intrinsic mode function (IMF), and each IMF component subsequently gains correspondence to the Hilbert spectrum. In essence, this method has proven to be a stable process in processing nonstationary signals by decomposing the fluctuations or trends in the signal, and representing the frequency content of the original signal using the final instantaneous frequency and energy.

HHT is widely used in the nonstationary signal modulation in data analysis, in the seismic signal and structure analysis, in the bridge and building condition monitor research areas, and so on analysis [53]. At present, this method has been utilized in mechanical breakdown diagnosis [54]. This work, in view of the rolling bearing breakdown vibration signal non-steady characteristic, proposes an improved approach based on the HHT characteristic energy method. When the rolling bearing breaks down, it increases the natural frequency of the rolling bearing vibrating system. Consequently, it vibrates at resonant frequencies of the bearing system and the structures around it. This thesis utilizes HHT in the inherent function that distinguishes the rolling bearing active status and the diagnosis type.

As previously mentioned, HHT includes two processes: the EMD analysis and the HT. The HHT decomposes the signal into a series of IMF components. Each IMF can have variable amplitude and frequency along the time axis, but each IMF component must satisfy the following two conditions:

a) the number of extrema and the number of zero crossings must either be equal or differ at most by one in the whole data set; and

b) the mean value of the envelope defined by the local minima and the envelope defined the local maxima is zero.

Figure 3.1 illustrates the NHHT sifting process to extract condition-related IMFs. Specific procedures are used as follows:

1) Determine all the local extrema of the signal. The cubic spline line is used to connect all the local maxima to form the upper envelope x_U. Determine all the local minima, and then connect them using a cubic spline line to form the lower envelope x_L. These two envelopes should cover all of the signal data between them.

2) Calculate the mean of the two envelopes

$$\mu_{11}(t) = (x_U(t) + x_L(t))/2 \tag{3.1}$$

Accordingly, the difference between the signal $x(t)$ and the mean $\mu_{11}(t)$ will be

$$h_1(t) = x(t) - \mu_{11}(t) \tag{3.2}$$

3) Check if $h_1(t)$ is an IMF according to the above two conditions. If it is not, then $h_1(t)$ is taken as the primary data (original signal). Repeat steps 1) and 2), and

$$h_{11}(t) = h_1(t) - \mu_{11}(t) \tag{3.3}$$

where $\mu_{11}(t)$ is the mean of the upper and lower envelopes, $h_{11}(t)$ will be considered as a signal and will be further processed by repeating the above steps (sifting process) k times until $h_{1k}(t)$ satisfies the IMF conditions. The resulting $h_{1k}(t)$ will be the first IMF component, designated as c_1

$$c_1(t) = h_{1k}(t) \tag{3.4}$$

c_1 contains the finest scale or the shortest period component of the signal.

4) Compute the first residual $r_1(t)$

$$r_1(t) = r_0(t) - c_1(t) \tag{3.5}$$

in this case $r_0(t) = 0$, and $r_1(t)$ will be the primary or the new data signal.

5) Repeat steps 1) to 4) n times to obtain the first n^{th} IMF and residual $r_n(t)$

$$r_n(t) = r_{n-1}(t) - c_n(t), \quad n = 2, 3, 4 \ldots \tag{3.6}$$

6) Terminate the sifting process when the residual $r_n(t)$ becomes a monotone function and no further IMFs can be extracted. Then signal $\bar{x}(t)$ will be formulated by

$$\bar{x}(t) = \sum_{m=1}^{n} c_m + r_n(t) \tag{3.7}$$

where $c_m(t)$ represents the m^{th} IMF; $r_n(t)$ is the nth residual, which represents the signal average tendency.

Apply the HT to each IMF component, $\{c_1(t), c_2(t),\ldots, c_n(t)\}$, to determine the instantaneous frequency.

7) Formulate the signal $\hat{x}(t)$ as the real part of the following function

$$\hat{x}(t) = \Re eal\left(\sum_{m=1}^{n} a_m(t) e^{j\int \omega_m(t)dt}\right) \tag{3.8}$$

where $a_m(t)$ is the instantaneous frequency amplitude function determined by

$$a_m(t) = \sqrt{c_m(t)^2 + \left(HT\left[c_m(t)\right]\right)^2} \tag{3.9}$$

where $HT(*)$ specified is the Hilbert transform.

The instantaneous frequency phase function will be determined by using the phase angle θ determined using ($\theta = \tan^{-1}\left(\frac{c_m(t)}{HT[c_m(t)]}\right)$)

$$\omega_m(t) = \frac{d(\theta(t))}{dt} = \frac{d}{dt}\left(\tan^{-1}\left(\frac{HT\left[c_m(t)\right]}{c_m(t)}\right)\right) \tag{3.10}$$

The extracted instantaneous frequency information from equations (3.9) and (3.10) constitutes the Hilbert spectrum. The monotonic functions $r_n(t)$ are not considered in the analysis because they do

not influence the frequency content of the signal. These spectrums are used to determine the averaged Hilbert spectrum of signal in the frequency domain.

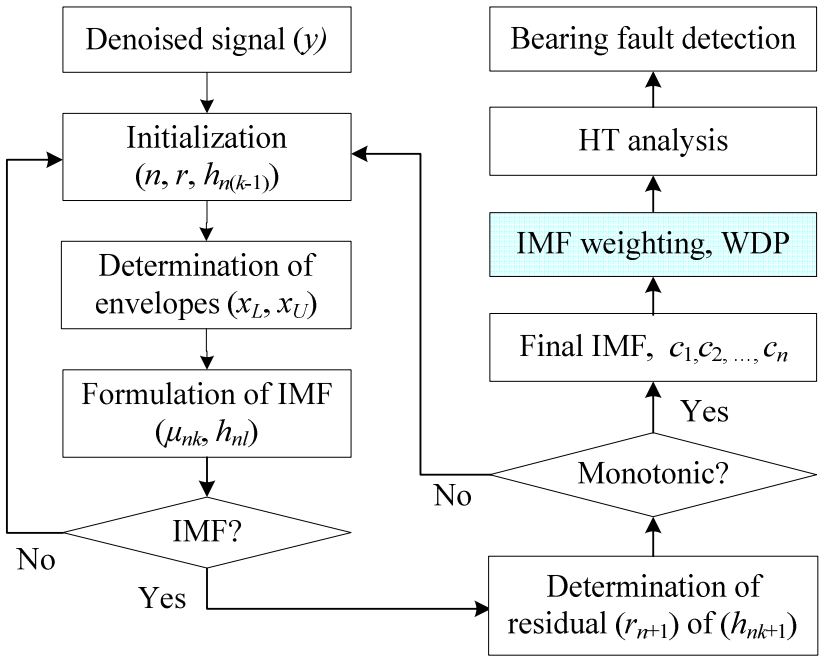


Figure 3.1. Flowchart of the proposed NHHT analysis process for bearing fault detection

As stated in Section 1.2, several HHT-based techniques have been proposed in literature for machinery defect detection [31, 56]. For example, Yan *et al*. [57] applied the HHT to process vibration signals for machinery health condition monitoring; however, their method could not effectively isolate distinctive condition-related IMFs for signal demodulation and features extraction. Yang *et al*. [58] applied the IMF envelope spectrum and support vector machine (SVM) for bearing fault detection, however the loss function used for support vector regression did not have clear statistical interpretation, whereas the SVM could move the problem of over-fitting from parameter optimization to model selection. Zvokelj *et al*. [59] combined the kernel principal component analysis with the ensemble EMD to model multiscale system dynamics; however it would be difficult to properly select the related function parameters.

To tackle the aforementioned challenges, the objective of this work is to develop a new normalized Hilbert-Huang (NHHT) technique for nonstationary feature analysis and bearing fault detection.

3.2 Proposed Normalized Hilbert Huang Transform (NHHT) and DP Analysis

This section deals with determining the weight of each IMF in terms of its contributions to normality and correlation to related condition. This integration, defined here as DP (integration of d'Agostino-Pearson (DP) normality test) is used to maximize combination output of the IMFs integration technique. Function non-normality becomes a problem when the sample size is small (less than 50), or when data is highly skewed, or leptokurtic. Normality becomes a serious concern when there is activity, especially clumps, in the tails of the data set.

In bearing fault detection, the classical skewness and kurtosis coefficient have some common disadvantages: they both have a zero breakdown value, and therefore they are very sensitive to outlying values. One single outlier can make the estimate become very large or small, making it difficult to interpret. Another disadvantage is that they are only defined on distributions having finite moments. DP test is tailored to detect departures from normality in the tails of the distribution as the properties also vary with respect to its bandwidth. DP incorporates both the skewness and the kurtosis to characterize the abnormality that exists in the distribution. The DP test has been proven to be an efficient overall test for normality, particularly for detecting normality due to asymmetry [60]. DP test for skewness and kurtosis is effective for detecting abnormality caused by asymmetry or non-normal tail heaviness, respectively. DP test for skewness and kurtosis has been suggested to be more powerful than the Shapiro-Wilk tests [61].

The d'Agostino-Pearson Omnibus test, on the other hand, analyzes the skewness and the kurtosis of your data with the Gaussian distribution, then compares these values with the values expected from a Gaussian distribution. The difference is then squared up and added to provide a sum of squares. This single sum of squares result is then used to calculate a p-value. A population, or its random variable X, is said to be normally distributed, its density function is given by

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\begin{cases}-\infty<x<\infty\\-\infty<\mu<\infty\\\sigma>0\end{cases} \tag{3.11}$$

where μ and σ are the mean and the standard deviation, respectively. The third and fourth standard moments are given by

$$\sqrt{b_1}=\frac{E(X-\mu)^3}{[E(X-\mu)^2]^{3/2}}=\frac{E(X-\mu)^3}{\sigma^3} \tag{3.12}$$

$$b_2=\frac{E(X-\mu)^4}{[E(X-\mu)^2]^2}=\frac{E(X-\mu)^4}{\sigma^4} \tag{3.13}$$

where E is the expected value operator and X is this case is c_i's obtained in Equation 3.4. These moments measure skewness and kurtosis, respectively. And for a normal distribution, they are 0 and 3, respectively.

D'Agostino and Pearson [60] proposed the test statistic K^2 that combines normalizing transformations of skewness and kurtosis, $Z\left(\sqrt{b_1}\right)$ and $\mathrm{Z}(b_2)$, respectively. The test statistic (DP)K^2 is given by

$$DP_i=Z_i^2\left(\sqrt{b_1}\right)+Z_i^2\left(b_2\right) \qquad ,i=1,2,3,\ldots,\ n_D. \tag{3.14}$$

where n_D is the number of selected function (i.e., ten in this work), in which each of the transformed sample skewness $Z\left(\sqrt{b_1}\right)$ test statistic is obtained by

$$Z\left(\sqrt{b_1}\right)=\frac{\ln\left(Y_1/\alpha_1+\sqrt{(Y_1/\alpha_1)^2+1}\right)}{\sqrt{\ln(W_1)}}, \tag{3.15}$$

where the corresponding Y_1, α_1 , and W_1 are obtained for each function (i.e. for each IMF) is given by

$$Y_1=\sqrt{b_1}\left(\frac{(n_D+1)(n_D+3)}{6(n_D-2)}\right)^{1/2}; \tag{3.16}$$

$$(W_1)^2=-1+\sqrt{2\beta_1(\sqrt{b_1})-1}\,; \tag{3.17}$$

$$\beta_1(\sqrt{b_1}) = \frac{3(n_D^2 + 27n_D - 70)(n_D + 1)(n_D + 3)}{(n_D - 2)(n_D + 5)(n_D + 7)(n_D + 9)} \tag{3.18}$$

$$\alpha_1 = \sqrt{\frac{2}{W_1^{\,2} - 1}} \tag{3.19}$$

The term $Z\left(\sqrt{b_1}\right)$ is test statistic and it's considered approximately normally distributed under the null hypothesis that the population follows a normal distribution.

And each of the transformed sample kurtosis $Z(b_2)$ test statistic is obtained by

$$Z(b_2) = \frac{1}{\sqrt{4.5A_2}}\left[\left(1 - \frac{2}{9A_2}\right) - \sqrt[3]{\frac{1 - 2/A_2}{1 + Y_2\sqrt{2/(A_2 - 4)}}}\right] \tag{3.20}$$

with A_2 and Y_2 obtained by

$$A_2 = 6 + \frac{8}{\sqrt{\beta_2(b_2)}}\left(\frac{2}{\sqrt{\beta_2(b_2)}} + \sqrt{1 + \frac{4}{\beta_2(b_2)}}\right); \tag{3.21}$$

$$\sqrt{\beta_2(b_2)} = \frac{6(n_D^2 - 5n_D + 2)}{(n_D + 7)(n_D + 9)}\sqrt{\frac{6(n_D + 3)(n_D + 5)}{n_D(n_D - 2)(n_D - 3)}} \tag{3.22}$$

$$Y_2 = \frac{b_2 - 3(n_D - 1)/(n_D + 1)}{\left(24n_D(n_D - 2)(n_D - 3)/(n_D + 1)^2(n_D + 3)(n_D + 5)\right)^{1/2}} \tag{3.23}$$

The normality hypothesis of the data is rejected for large values of the test statistic. Furthermore, according to [60], the test K^2 is approximately chi-squared distributed with two degrees of freedom.

The proposed integration weighted d'Agostino-Pearson (WDP) indicator is determined by [62]

$$WDP = \frac{\sum_{i=1}^{n_D} \frac{(DP_i)}{\max(DP)}\hat{x}_i}{mean\left((DP)\hat{x}_i\right)} \tag{3.24}$$

where n_D is the number of function (i.e. n_D = 10 IMFs in this case).

For the integration process of the WDP values for the resulting output signal of Hilbert-Huang analysis $\hat{x}$ obtained using Equation (3.8) (i.e. NHHT), a large DP number is expected for a faulty bearing since it indicates higher magnitudes of periodic features. This technique is designed to combine the more contributive parts of the signals that are condition-related for feature extraction. When d'Agostino Pearson indicators are realized for all IMF's, the resulting NHHT is determined as weighted IMFs in Equation (3.24). The most distinctive feature functions can be correlated to the highest indicator values among the IMFs. These weighted functions will then be used to recognize the most distinctive IMFs in the developed NHHT technique.

In implementation, the proposed integration process of NHHT technique uses an averaged spectrum over several segments of the signal to mitigate random noise. The effectiveness of the proposed NHHT technique including the realization of the IMF integration method will be verified experimentally in the following section.

WDP is integrated to estimate the normality measure of the IMFs. When DPs are realized for all functions, the WDP indicator is determined as the maximum output of the integration between WDP and HHT corresponding to each IMF. The most distinctive feature functions can be observed to be those with the highest DP and WDP indicators values among the IMFs. In implementation, the proposed integration process of NHHT technique uses the maximum spectrum over several segments of the signal to mitigate random noise.

3.3 Experimental Setup

The experimental setup employed for effectiveness tests in this work is shown in Figure 3.2. The system is driven by a 3-hp induction motor, with the speed range from 20 to 4200 RPM. Speed is controlled by a speed controller (VFD022B21A). A variable load is applied by a magnetic brake system through a bevel gearbox and a belt drive. The load is provided by a brake system through a belt drive. An optical sensor is used to provide a one-pulse-per-revolution signal for shaft speed measurement. A flexible coupling is utilized to dampen the high-frequency vibration generated by the motor. Two ball bearings (MB ER-10K) are fitted in the solid housings, which have the following parameters: rolling elements: 8; rolling element diameter: 7.938 mm; pitch diameter: 33.503 mm; and contact angle: 0 degree. The bearing on the left-hand side housing is used for testing. Accelerometers

(603C01) are mounted on the housing of the tested bearing to measure the vibration signals along two directions. Considering the structural properties, the signal measured vertically is utilized for analysis in this work, whereas the signal measured from the horizontal direction is used for verification. These vibration and reference signals are fed into a computer for further processing through a data acquisition board (NI PCI-4472) which has built-in anti-aliasing filters with the cut-off frequency set to half of the sampling rate. The sampling frequency varies with the shaft speed to collect 600-700 samples over each shaft rotation cycle.

In the tests, four bearing health conditions are considered: healthy bearings, bearings with outer race defect, bearings with inner race defect, and bearings with rolling element defect. Verification testing was also conducted with six different shaft speeds (600, 900, 1200, 1500, 1800, and 2100 RPM) and three loading levels (1, 2.5, and 5 N.m) are used to test each of the five bearing conditions (healthy, outer race defect, inner race defect, rolling element defect, and combination of all defects).

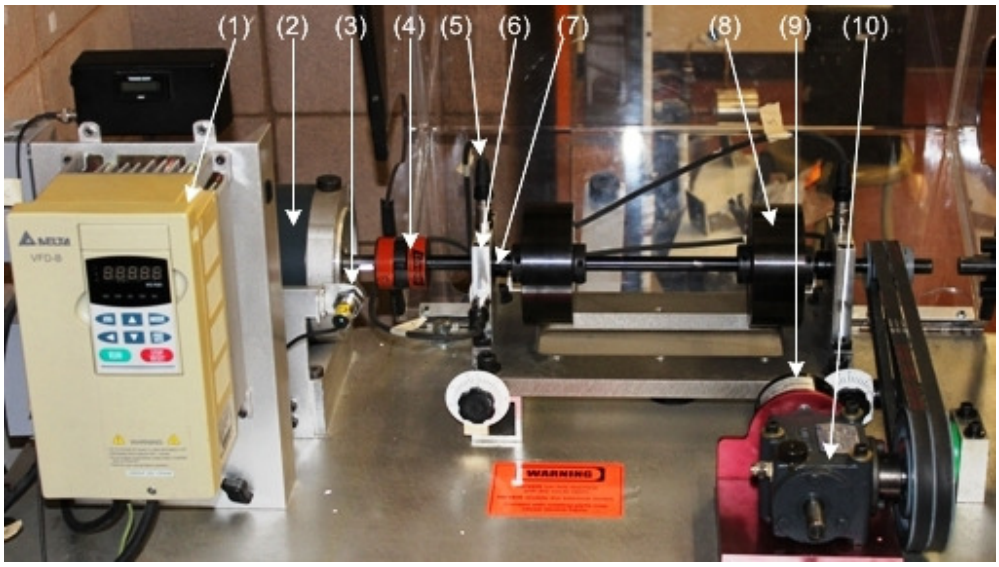


Figure 3.2. Verification tests experimental setup: (1) speed control; (2) motor; (3) optical sensor; (4) flexible coupling; (5) ICP accelerometer; (6) bearing housing; (7) test bearing; (8) load disc; (9) magnetic load system; (10) bevel gearbox.

All of the related techniques are implemented in MATLAB. The processing results of the performance comparison and the robustness examination are discussed in the following subsection.

3.4 NHHT Performance Evaluation Tests

The performance effectiveness of the proposed bearing fault detection technique will be compared to some other related techniques. Since the proposed technique involves both DP normality test and HHT processing, it will be specified as NHHT for simplicity. To examine the effectiveness of the NHHT, the classical Hilbert Huang technique without using DP, specified by HHT, will be used for comparison. To compare the performance of the NHHT, the commonly used HHT method combined using the Shannon entropy method discussed in Section 2.4 (Equation 2.23) and the wavelet transform, WT, will be applied, which are designated as SHHT and WT, respectively. The Morlet wavelet will be used in the WT analysis which will be performed over the resonance frequency band of [1000, 2000] Hz.

3.4.1 IMF integration using WDP

A few examples are used to demonstrate the implementation of the proposed IMF integration method in the NHHT technique. In general, the higher the values of these DP and WDP indicators, the more distinctive of the related information units [57].

1) Overview

Firstly, some examples are used to illustrate the properties and application of DP normality indicator. Figure 3.3 shows an example EMD and the corresponding IMFs representation in the time domain. Figure 3.4 shows some DP results corresponding to four bearing health conditions, which correspond to the shaft speed of 1800 RPM (or 30 Hz) and load torque of 2.5 Nm. To improve processing accuracy, the first ten IMFs will be used for the HHT analysis in this case, while only the first two IMFs are used for HHT analysis for general applications [52, 53]. For a healthy bearing (Figure 3.4a), the most distinguishable IMFs are the first and the eight IMF (instead of the first and second in the classical HHT analysis). For a bearing with an outer race defect (Figure 3.4b), the most distinguishable functions are the first and the third IMFs. For a bearing with an inner race defect (Figure 3.4c), the most distinguish IMF are determined to be the first and the second IMFs, which is consistent with the classical HTT analysis. For a bearing with a rolling element defect (Figure 3.4d), the most distinguishable functions become the first and the third IMFs. On the other hand, it is seen that higher order IMFs (higher than 2) may also contribute significantly to signal properties. As stated before, the

proposed NHHT aims to properly integrate these IMFs for advanced signal property analysis and more accurate bearing fault detection.

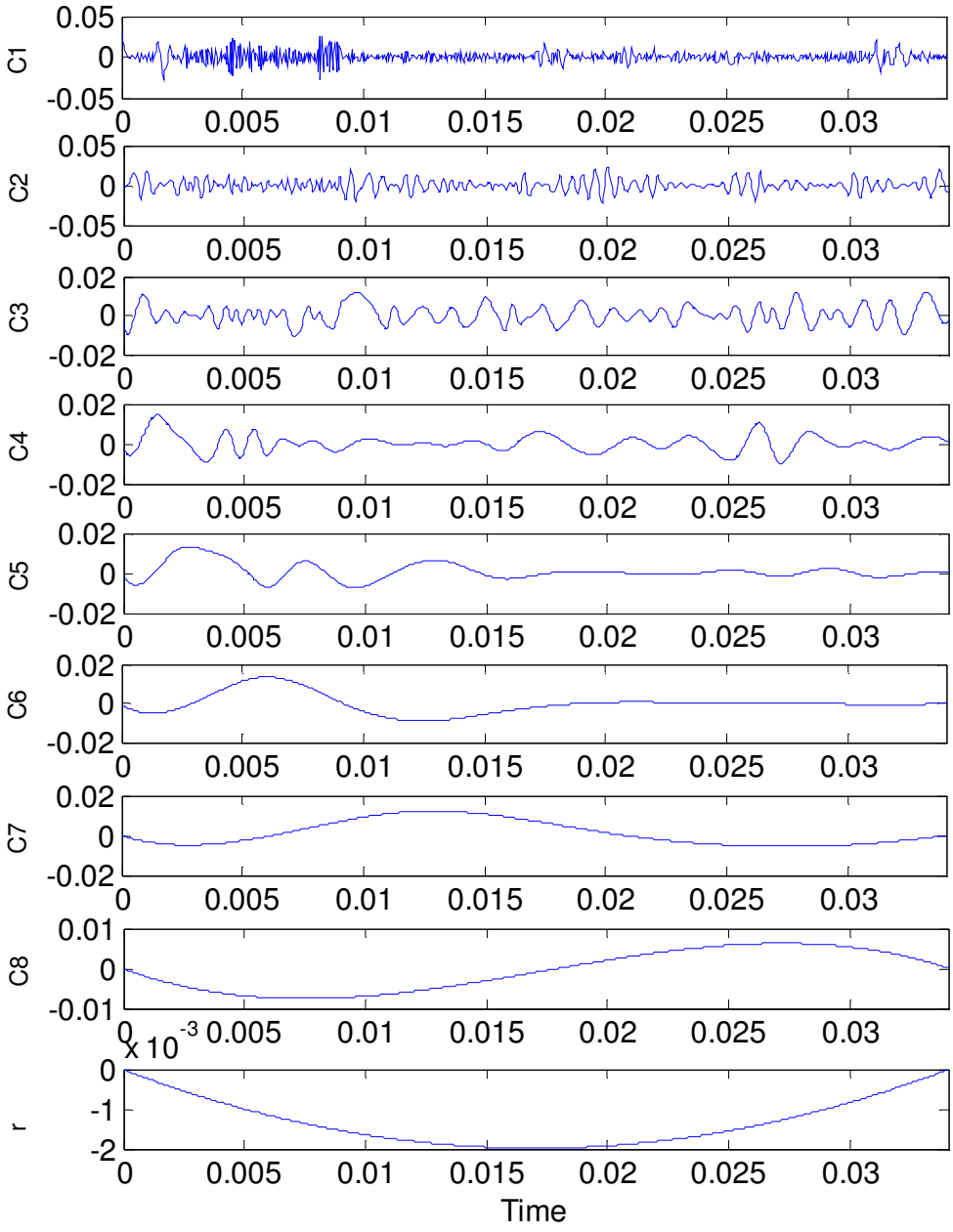


Figure 3.3. The EMD decomposed results of vibration signal of the healthy ball bearing.

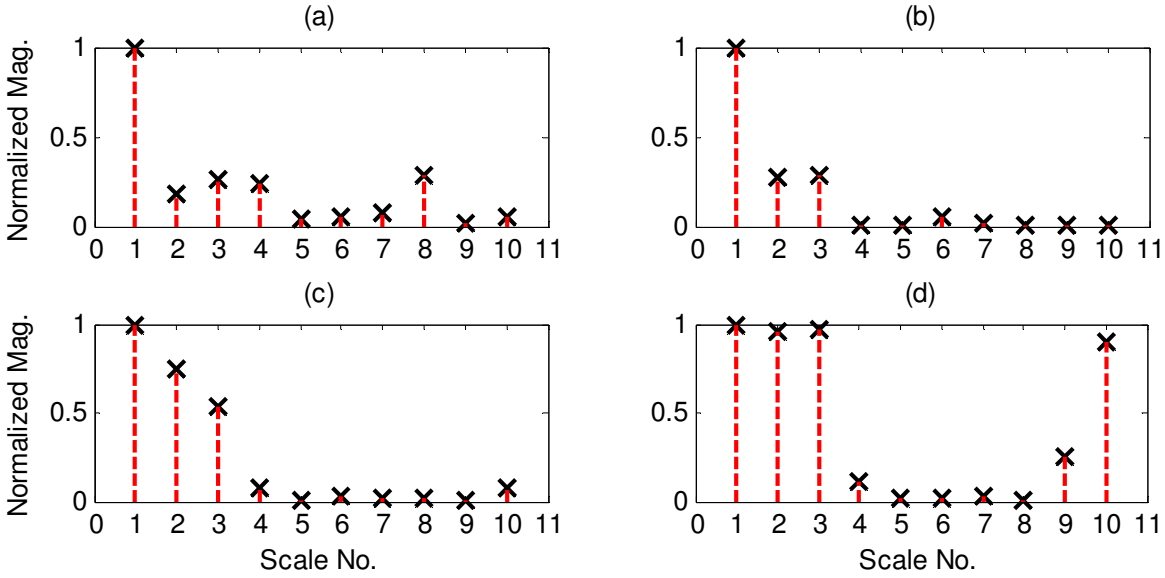


Figure 3.4. Demonstration of normalized DP indicator values versus IMF scale numbers corresponding to different bearing health conditions: (a) healthy bearing; (b) bearing with outer race defect; (c) bearing with inner race defect; (d) bearing with rolling element defect.

Figure 3.5 shows part of the collected vibration signals corresponding to different bearing conditions. Although some impulses could be recognized from these vibration signatures (e.g., 3.5a and 3.5d), it is difficult to diagnose bearing health conditions based solely on these original vibration patterns.

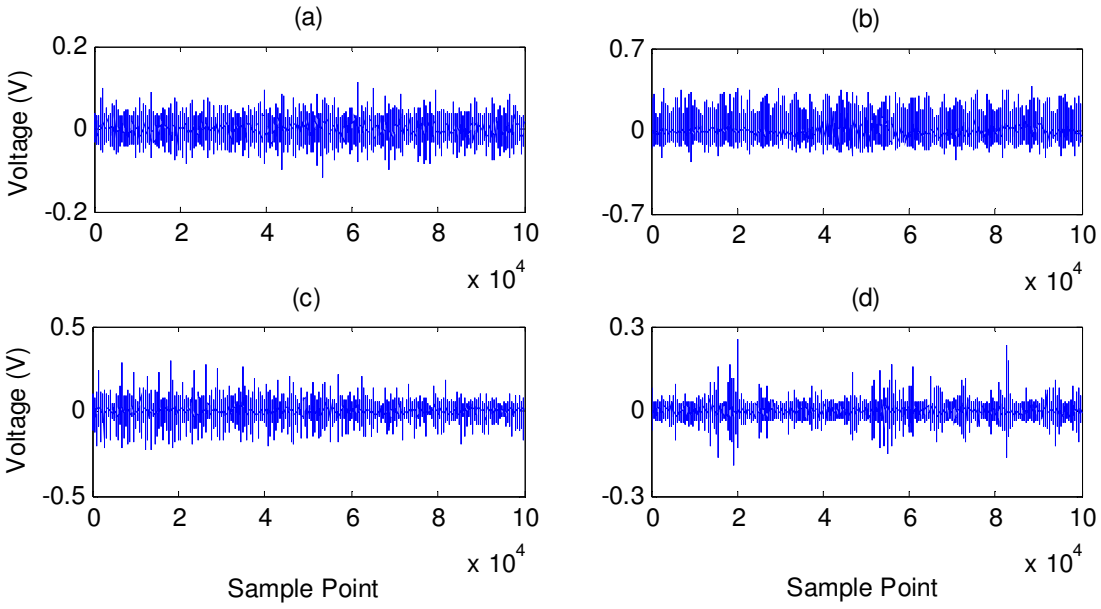


Figure 3.5. Part of collected vibration signals for bearings with different health condition: (a) a healthy bearing; (b) a bearing with outer race defect; (c) a bearing with inner race defect; (d) a bearing with rolling element defect.

3.4.2 Performance evaluation

2) Healthy bearing condition monitoring

Firstly, the bearings with healthy conditions are tested. Figure 3.6 shows the processing results using the related techniques and result obtained in Figure 3.4a. Figure 3.6a-3.6d demonstrates processing results based on power spectrum scales. Although power spectra have lower spectral magnitude than spectrum map, they provide better result representation and make spectral analysis easier, which will be used in this work.

In this case, the bearing characteristic frequency (f_r) is 30 Hz. Examining power spectral graphs in Figure 3.6a-3.6d, the bearing characteristic frequency can be identified by each technique. Comparing Figure 3.6d using the proposed NHHT and Figure 3.6c using the HHT, the weighted DP indicator can process the signal properly and highlight characteristic frequency component (in this case, the shaft speed 30 Hz). Comparing Figure 3.6d (NHHT), Figure 3.6a (SHHT), and Figure 3.6b (WT), the proposed NHHT technique outperforms all in enhancing the shaft speed information (with higher magnitude). Although the SHHT and WT have recognized the shaft speed (Figure 3.6a and

Figure 3.6b, respectively), its spectral map contains more frequency components due to leakage, which may reduce the reliability for bearing health condition monitoring.

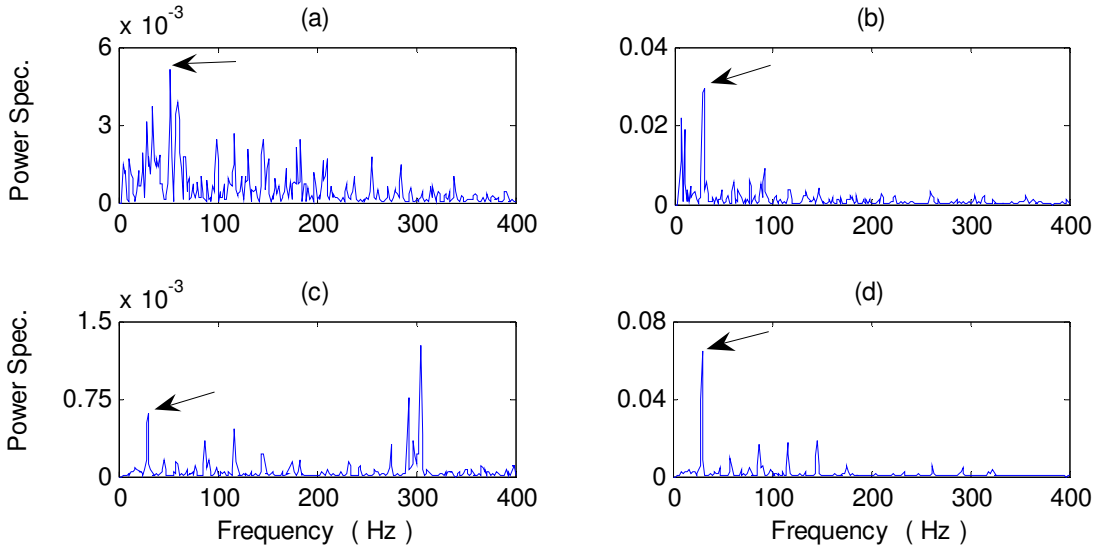


Figure 3.6. Comparison of processing results for a healthy bearing using the related techniques with a scale of the power spectrum: (a) SHHT; (b) WT; (c) HHT; and (d) NHHT. (Arrow indicates the characteristic frequency)

3) Outer race fault detection

When defect occurs on the fixed ring race of a bearing (the outer race in this case), its defect-induced resonance modes do not change over time. In this case, the characteristic frequency is BPFO or f_{OR} = 91 Hz. Figure 3.7 shows the processing results using the related methods and result obtained for Figure 3.4b. Comparing Figure 3.7d (NHHT) and Figure 3.7c (HHT), the DP can effectively process the signal and result in higher magnitude of the characteristic frequency component. It is seen that the proposed NHHT (Figure 3.7d) outperforms not only the SHHT (Figure 3.7a) with higher magnitude of characteristic frequency components, but also the WT (Figure 3.7b) with the defect frequency dominating the spectral map. The main reason is due to effective information processing using the proposed NHHT method.

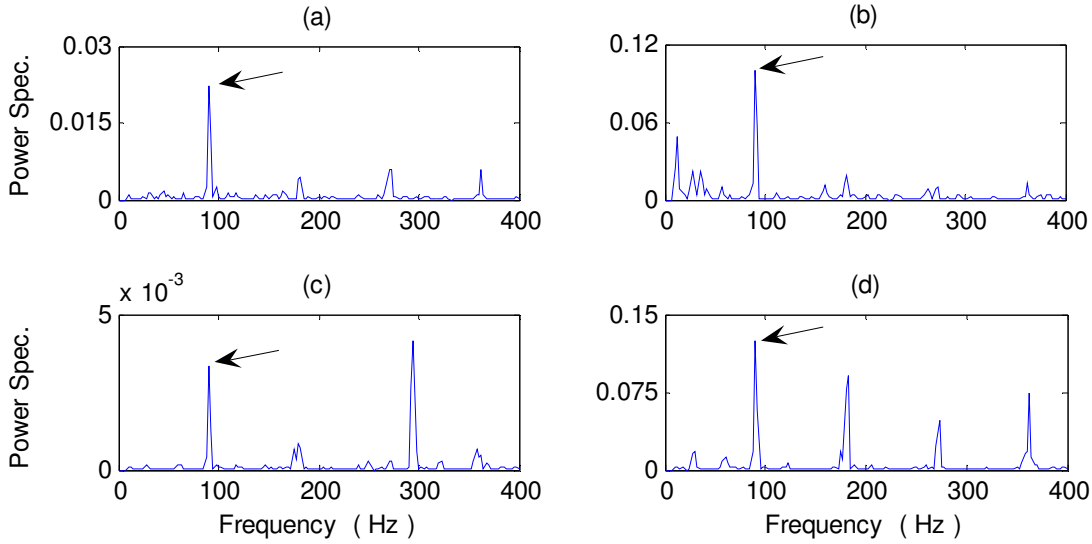


Figure 3.7. Comparison of processing results for a bearing with outer race fault detection with a scale of the power spectrum using the techniques of: (a) SHHT; (b) WT; (c) HHT; and (d) NHHT. (Arrow indicates the characteristic frequency)

4) Inner race fault detection

The detection of fault on rotating elements is considered more challenging than the detection of fault in the outer race; this is due to the fact that the inner race rotates and its resonance modes vary over time. In this case, the characteristic frequency BPFI or f_{IR} = 148 Hz. Processing results using the results obtained for Figure 3.4c and the related methods are shown in Figure 3.8. Comparing Figure 3.8d (NHHT) and Figure 3.8c (HHT), the weighted DP can effectively process the signal and improve signal-to-noise-ratio. Furthermore, the proposed NHHT (Figure 3.8d) outperforms both the SHHT (Figure 3.8a) and the WT (Figure 3.8b) due to its efficient information processing.

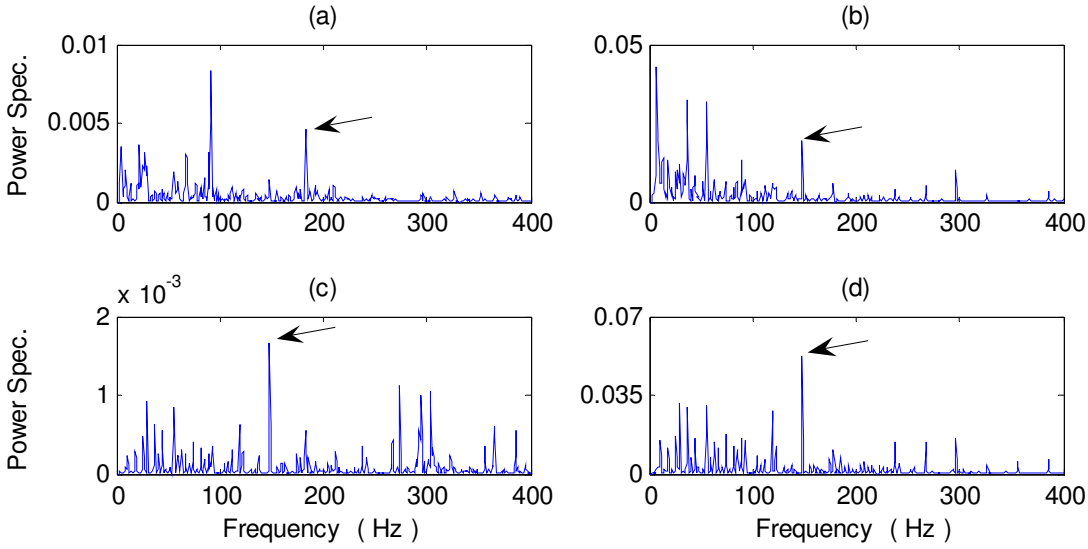


Figure 3.8. Comparison of processing results for a bearing with inner race fault detection with a scale of the power spectrum using the techniques of: (a) SHHT; (b) WT; (c) HHT; and (d) NHHT. (Arrow indicates the characteristic frequency)

5) Rolling element defect detection

The detection of fault on the rolling element (i.e., ball in this case) is considered the most challenging task in bearing fault detection. This is due to the fact that the rolling element rotates along different directions (it also slides), and its resonance modes change over time. In this case, the characteristic frequency is BPFR or f_{BD} = 119 Hz. Figure 3.9 shows the processing results using the related methods and result obtained for Figure 3.4d. It is clear that the proposed NHHT is the method that can detect the rolling element defect in this case (Figure 3.9d) much clearer than HHT alone (Figure 3.9c), this is due to its effective process and IMF information integration strategy. Neither SHHT (Figure 3.9a) nor WT (Figure 3.9b) can provide recognizable defect frequency signatures in this case with lower noise level and defect frequency dominating the spectral map.

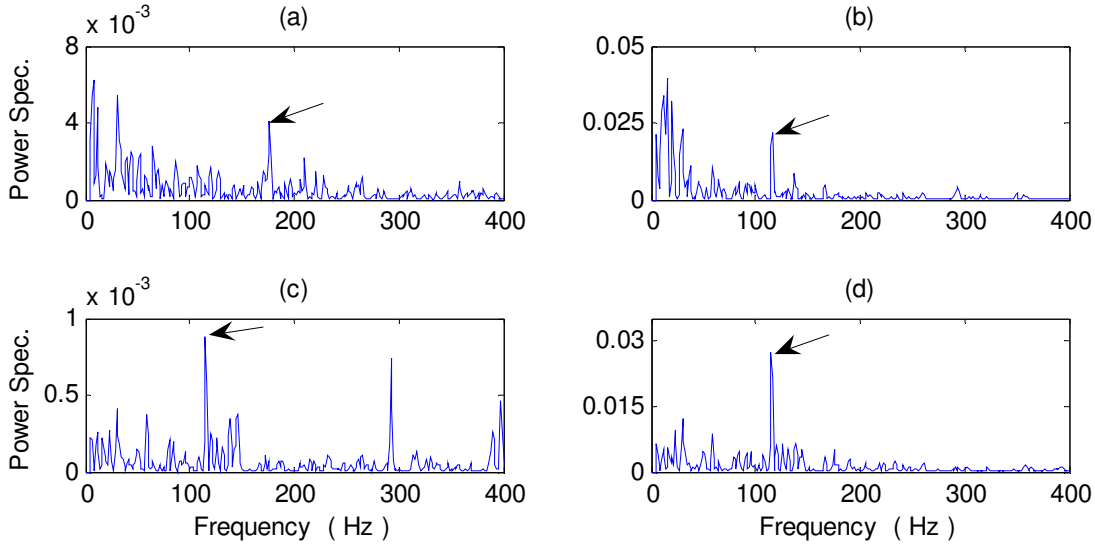


Figure 3.9. Comparison of processing results for a bearing with rolling element (ball) fault detection with a scale of the power spectrum using the techniques of: (a) SHHT; (b) WT; (c) HHT; and (d) NHHT. (Arrow indicates the characteristic frequency)

As demonstrated in the previous figures, the proposed NHHT has detected the incipient fault effectively for all four tested conditions of bearings. This is due to its effective informational process and enhancing defect-related features.

Chapter 4

The Enhanced Hilbert Huang Transform (eHHT) Technique

4.1 Minimum Entropy Deconvolution (MED) for Signal Denoising

As stated in Section 2.3, the minimum entropy deconvolution, MED, can be used to reduce the impedance effect of the transmission path of the signal from the bearing to the sensor. To highlight defect-related impulses, a denoising process is applied using the MED filter. MED was thought of as being useful in enhancing the definition of the impulses and widening the use of the Spectral Kurtosis (SK), which was discussed in Section 1.2.2d.

A MED filter is used as a denoising process after the signal is acquired and before applying the proposed eHHT techniques. This step is essential to reduce the noise level and to enhance some of defect-related impulses.

The MED filter searches for an optimum set of filter coefficients that recover the output signal (of an inverse filter) with the maximum kurtosis value. Many faults are known to manifest themselves as impulses in rolling element bearings. The MED filter was implemented by the objective function method (OFM) given in section 2.3 and [44]. The OFM is an optimization process designed to maximize the kurtosis of the MED filter output (minimizing entropy). OFM achieves this by changing the values of the MED parameters. The optimization process finishes when the values of the coefficients converge within the specified tolerance (Step 4) and equation 2.22).

In general, the MED filter parameters (e.g., filter length) are determined by trial and error. A series of simulations will be carried out corresponding to different number of impulses plus noise signals with different noise levels. The convergence of the MED filter can be measured by iteration number. Table 4.1 summarizes the simulation result of the MED filter length L_m corresponding to different impulse numbers and noise levels. In general, signal length does not influence the convergence of the MED filter; the signal length is set to 10,000 samples as an example. The criterion is to clearly recognize the impulses with respect to sufficiently large

kurtosis values. The MED filter convergence depends on the signal characteristics and the filter length. For example, as filter length varies from 10 to 400, the MED filter converges over 10 and 20 iterations with kurtosis values greater than 10. When the signal contains between 1 and 7 simulated impulses, the MED filter converges between 8 and 20 iterations with kurtosis greater than 13. However, if the signal contains too many impulses, the algorithm will treat the large number of impulses as noise. On the other hand, when noise is added to the tested signal with different magnitude levels (e.g., 20% to 60% of the signal's maximum magnitude), the MED filter converges over 7 to 16 iterations with kurtosis greater than 11.

Table 4.1. Summary of initial values of the MED filter

Parameters	Tested Range	Kurtosis	Iteration Number
Filter length	[10, 400]	≥ 10	[10, 20]
Impulse number	[1, 7]	≥ 13	[8, 20]
Noise level (%)	[20, 60]	≥ 11	[7, 16]

As an example, consider a simulated signal consisting of four impulses and some noise as illustrated in Figure 4.1a. The signal length is set to 10,000 samples. Figures 4.1b-4.1d illustrate the processing results of three MED filters with filter lengths L_m = 100, 300, and 500, respectively. The respective convergence in terms of kurtosis is illustrated in Figure 4.2. It is seen that these four distinctive impulses can be clearly highlighted if the filter length L_m = 300 (Figure 4.1b), in this case. Although the filter with length L_m = 500 has the highest kurtosis (Figure 4.2c) than filters with lengths of 100 (Figure 4.2a) and 300 (Figure 4.2b) respectively, it cannot recognize all three impulses due to phase distortion. On the other hand, as demonstrated in Figure 4.2, all three MED filters converge quickly over about 9-10 iterations only; that is, the convergence of the implemented MED filter is not sensitive to filter length. As a result, by trial and error, the filter length of 300 is chosen for this example.

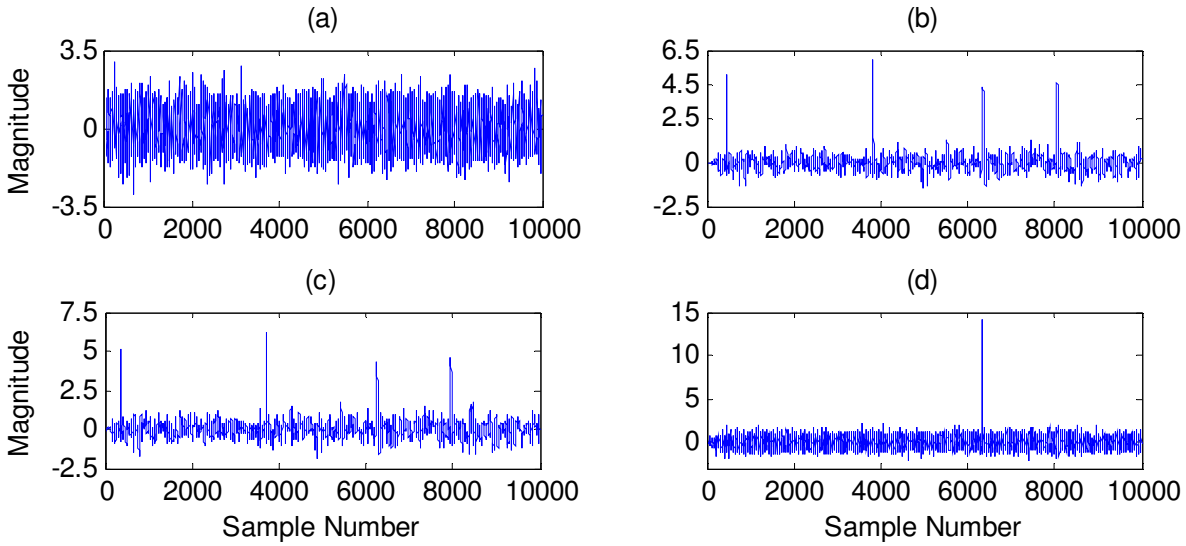


Figure 4.1. Response comparison of a test signal using MED filters with different lengths: (a) simulated input signal; (b) MED response with filter length of 100; (c) MED response with filter length of 300; (d) MED response with filter length of 500.

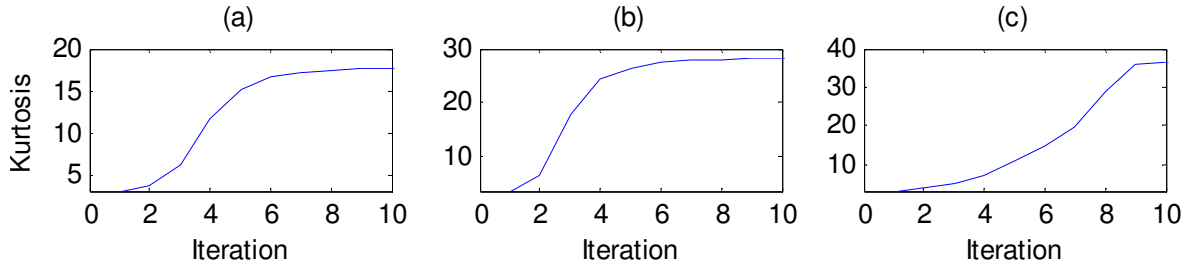


Figure 4.2. Convergence comparison of MED filters with different lengths: (a) filter length of 100; (b) filter length of 300; (c) filter length of 500.

After the vibration signal is denoised by using the adopted MED filter, it will be further processed by the use of the proposed eHHT technique as discussed in the following subsection.

4.2 Proposed Enhanced HHT (eHHT) Technique

Different from the proposed NHHT in Chapter 3, which applied a normality measure (the statistical domain) to integrate IMF information for HHT analysis [53, 54], a novel IMF selection method based on mutual information analysis (the information entropy domain) is proposed in this section to extract the most instinctive IMFs for nonstationary signal property analysis and bearing fault detection.

The Mutual Information (MI) can be used to measure the correlation between variables [58]. Different from the general autocorrelation function as well as the statistical normality analysis as applied the proposed NHHT in Chapter 3, the MI takes into account non-linear correlations. In addition, it is usually less sensitive to the size of the partition elements and is relatively easier for implementation.

As illustrated in Figure 4.3, consider two information spaces $H(F_1)$ and $H(F_2)$ of the random variables F_1 and F_2 with probability distributions of D_1 and D_2 respectively. Joint entropy, $H(F_1, F_2)$, represents the joint probability distribution of discrete random variables F_1 and F_2. In this work, F_1 and F_2 represent the signals (y and $\hat{x}$) that are obtained from using the MED filter and the Hilbert analysis (equation 2.19 and equation 3.8), respectively. The MI, $I(F_1, F_2)$, will be defined as

$$I(F_1, F_2) = H(F_1) + H(F_2) - H(F_1, F_2) \tag{4.1}$$

MI uses the joint probability and product probability to determine the correlation between two variables. The joint probability distribution, D_j, is a distribution defined in terms of both F_1 and F_2, while the product distribution D_p is a distribution defined by the product of two random variables having two known distributions (i.e. $D_p = D_1 * D_2$). $I(F_1, F_2)$ can be considered as the relative entropy between D_j and D_p , such that

$$I(F_1, F_2) = \sum_{f_1 \in F_1} \sum_{f_2 \in F_2} D_j(f_1, f_2) \log_{10}\left(D_p(f_1, f_2)\right) / [D_1(f_1) D_2(f_2)] \tag{4.2}$$

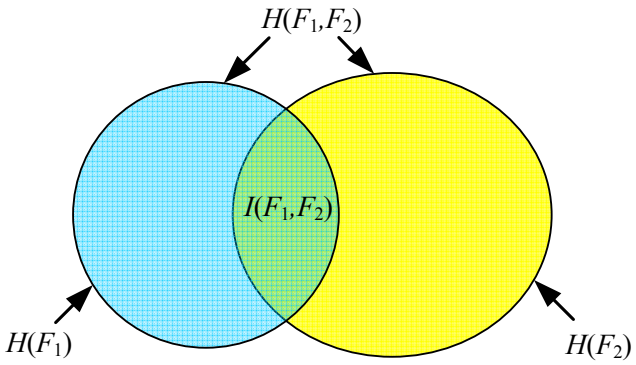


Figure 4.3. Illustration of the relationship between entropy $H(F_1)$ (left circle), $H(F_2)$ (right circle); mutual information $I(F_1, F_2)$ (intersection area); and joint entropy $H(F_1,F_2)$.

From equation (4.2), it is seen that the maximum MI can be achieved by either maximizing the dataset information (or entropies) $H(F_1)$ and $H(F_2)$, or by minimizing the joint entropy $H(F_1,F_2)$. The MI, $I(F_1, F_2)$, is usually sensitive to the changes that occur in the overlap statistics. Since MI takes into account joint data information, it will be superior to the general joint entropy in measuring information similarity. If the random variables F_1 and F_2 are independent, the joint probability distribution becomes the product distribution, or $D_j = D_P$. Thus, MI can measure the correlation between variables F_1 and F_2.

The proposed IMF selection technique consists of two indicators for linear and non-linear similarity measurement.

1) *Normalized Correlation Measure* (*NCM*). The proposed NCM indicator is used to measure linear similarity between the two distributions, which is defined as [59, 60]

$$NCM = \sum_{i=1}^{N} \| R(y, \hat{x}) \| / L_N \tag{4.3}$$

where $R(y, \hat{x}) = E[(y - E[y])(\hat{x} - E[\hat{x}])]/(\sigma_y \sigma_{\hat{x}})$ is the cross-correlation of the residual of the MED filter output y from Equation (3.13) and the demodulated signal $\hat{x}$ from Equation (4.8); $L_N = 2N\text{-}1$, where N is the length of original signal; $E(.)$ is the expectation operator; σ_y and $\sigma_{\hat{x}}$ are the standard deviations of y and $\hat{x}$, respectively.

2) *Deficiency of Mutual Information* (*DMI*). The proposed DMI indicator aims to provide a non-linear similarity measure to characterize the uniqueness of the extracted IMFs in processing representative features related to bearing health conditions, which is defined as [59, 60]

$$DMI = \| \hat{I} \| \Big/ \sqrt{\log_{10}(N)} \tag{4.4}$$

where $\hat{I} = \hat{I}(F_1, F_2) = \big(H(F_1) + H(F_2)\big) \big/ H(F_1, F_2)$ is a normalized MI; N is length of the signal being analyzed ($N > 1$); F_1 and F_2 represent the signals y and $\hat{x}$ obtained by using the MED filter in equation (2.19) and the Hilbert analysis in Equation (3.8), respectively.

DMI represents the relative information contained in each dataset F_1 and F_2 over the joint information $H(F_1,F_2)$. If the probability of overlap space from two distributions remains the same, the DMI is constant regardless of the orientation of distributions $H(F_1)$ and $H(F_2)$. That is, DMI is only dependent on the overlap information, but independent of the orientations of the constitute spaces $H(F_1)$ and $H(F_2)$. Therefore, it will be more robust than the joint entropy and the MI.

Both NCM and DMI are integrated (denoted as NCM/DMI) to estimate the similarities between signals. When NCM and DMI are realized for all the functions, the NCM/DMI indicator is determined as the integration of NCM and DMI corresponding to each IMF [60]

$$\text{NCM/DMI} = a_n (\text{NCM}) + b_n (\text{DMI}) \tag{4.5}$$

where coefficients a_n and b_n are determined based on applications and subscript n is the corresponding IMF number. In this case, they are selected as unity for simplicity. The most distinctive feature functions are chosen to be those with the highest NCM/DMI indicator values among the IMFs. These selected functions will contain maximum correlation information (or minimum entropy), which will then be used to recognize most distinctive IMFs in the developed eHHT technique.

In implementation, the proposed eHHT technique uses an averaged spectrum over several segments of the signal to mitigate random noise. The effectiveness of the proposed eHHT technique, including the realization of the IMF selection method, will be verified experimentally in the following section.

4.3 eHHT Performance Evaluation Tests

4.3.1 IMF selection using NCM/DMI

In this work, only the first ten IMFs are used for enhanced Hilbert spectral analysis because the signal energy is mainly contained in the first few IMFs. The Hilbert analysis is performed, but the monotonic functions, $r_n(t)$, are not considered in the analysis because they do not influence the frequency content of the signal. These spectrums are used to determine the averaged Hilbert spectrum of signal in the frequency domain. The enhanced Hilbert Huang transform (eHHT) differs from NHHT and the classical HHT by the utilization of the proposed correlation measure (NCM/DMI) to obtain the most distinguished functions that are condition-related, which are used in the analysis to obtain signature related features.

The performance effectiveness of the proposed bearing fault detection technique will be compared with other related techniques. Since the proposed technique involves both MED denoising filter and eHHT processing, it will be specified as MED-eHHT for simplicity. To examine the effectiveness of the MED denoising filter, the eHHT technique without using the MED filter, specified by eHHT, will be used for comparison. To compare the performance of the eHHT, the commonly-used HHT method and the wavelet transform, WT, will be applied, which are designated as MED-HHT and MED-WT, respectively. The Morlet wavelet will be used in the WT analysis which will be performed over the resonance frequency bands of [1500, 2500] Hz and [1000, 8000] Hz, respectively.

Firstly, a few examples are used to demonstrate the implementation of the proposed IMF selection method (i.e., NCM/DMI) in the eHHT technique. In general, the higher the values of these DMI and NCM indicators, the more distinctive the related information units.

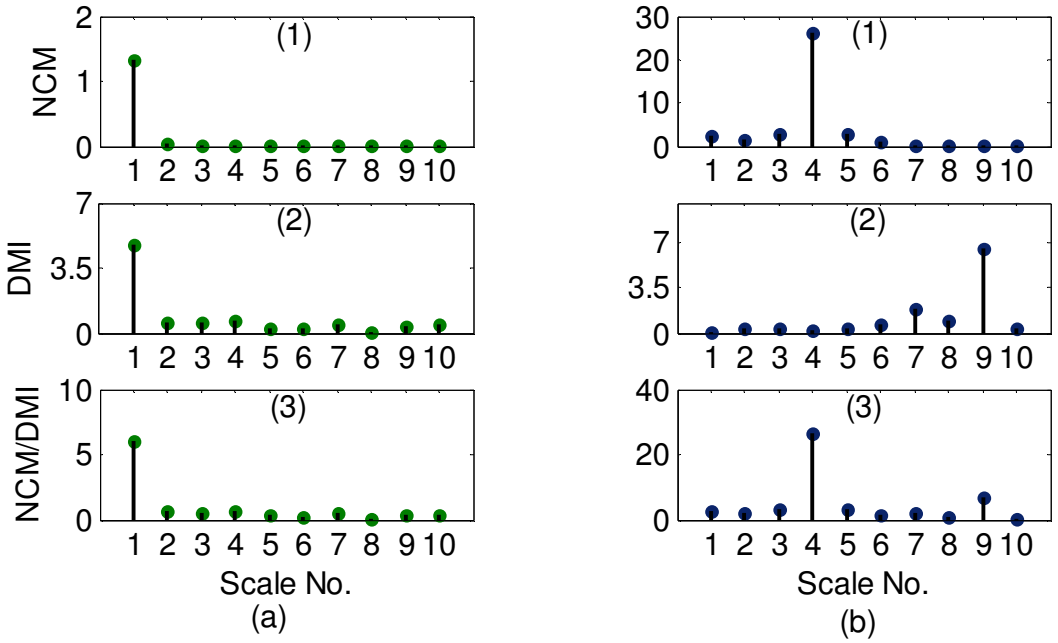
4.3.2 Performance Evaluation

1) Overview

A series of tests have been conducted corresponding to different load and speed operating conditions using the experimental setup as illustrated in Figure 3.2. The following examples correspond to the shaft speed of 1800 RPM (or 30 Hz) and load torque of 2.5 Nm.

Four bearing conditions are tested. Figure 4.4 shows some examples in using the information indicators NCM, DMI, and NCM/DMI for different bearing health conditions: a healthy bearing (Figures 4.4 a1-b3), a bearing with outer race defect (Figures 4.4 b1-b3), a bearing with inner race defect (Figures 4.4 c1-c3), and a bearing with rolling element (ball) defect (Figures 4.4 d1-d3), respectively. To improve processing efficiency, only two IMFs will be selected for eHHT and HHT analysis in this case. In general, the first two IMFs are chosen for HHT analysis [53, 54]. Different from the general HHT, the proposed eHHT will select the two most distinctive IMFs among the first 10 IMFs for analysis using the proposed NCM/DMI indicator.

For the healthy bearing (Figure 4.4 a3), the NCM/DMI indicator will select the first and fourth IMFs for eHHT analysis. For the bearing with an outer race defect (Figure 4.4 b3), the two most distinguishable functions are the fourth and ninth IMFs. For the bearing with an inner race defect (Figure 4.4 c3), the first and second IMFs are selected by the NCM/DMI indicator. For the bearing with a rolling element defect (Figure 4.4 d3), the NCM/DMI indicator selects the first and second IMFs for eHHT analysis.



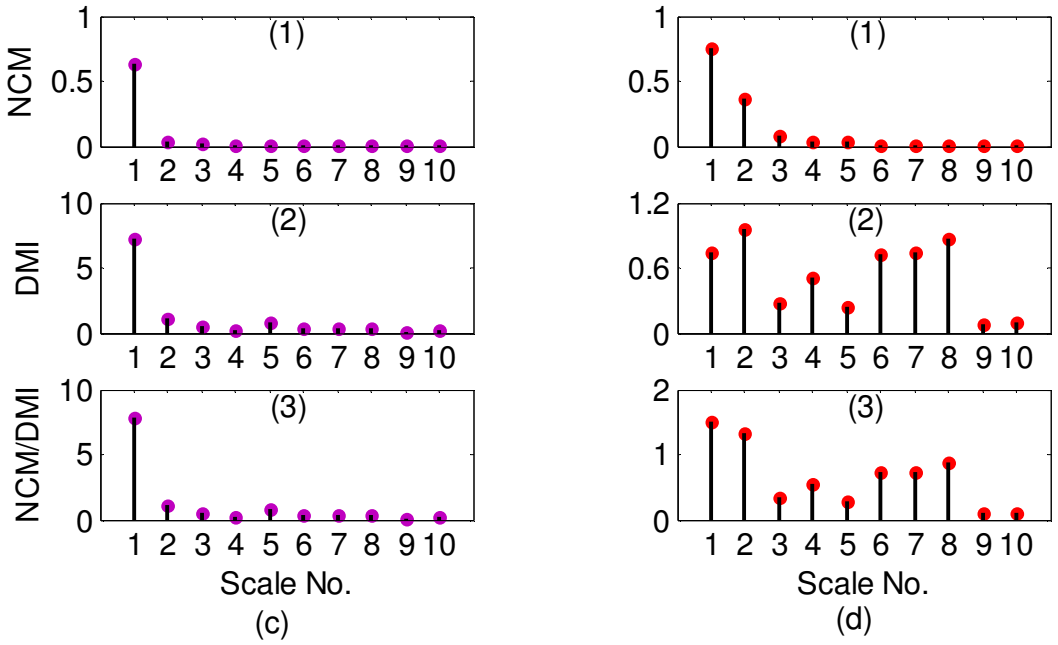


Figure 4.4. Demonstration of NCM/DMI indicator values versus IMF scale numbers corresponding to different bearing health conditions: (a1,a2,a3)-healthy bearing; (b1,b2,b3)-bearing with outer race defect; (c1,c2,c3)-bearing with inner race defect; (d1,d2,d3)-bearing with rolling element defect.

2) Healthy bearing condition monitoring

Firstly, the bearings with healthy conditions are tested. As shown in Figure 4.4 a3, the first and fourth IMFs are selected for eHHT analysis, while the HHT selects the first and second IMFs as in general. Figure 4.5 shows the processing results using the related techniques. Figures 4.5a-4.5d demonstrate processing results with normal spectrum scales, while Figures 4.5e-4.5h represent processing results based on power spectrum scales. It is seen that although power spectra have lower spectral magnitude, they provide better result representation and make spectral analysis easier, which will be used in this work.

In this case, the bearing characteristic frequency is 30 Hz. Examining power spectral graphs in Figures 4.5e-4.5h, the bearing characteristic frequency can be identified by each technique. Comparing Figure 4.5a using the proposed MED-eHHT and Figure 4.5b using the eHHT, the MED filter can denoise the signal properly and highlight characteristic frequency component (in this case, the shaft speed 30 Hz). Comparing Figure 4.5e (MED-eHHT) and

Figure 4.5g (MED-HHT), the proposed MED-eHHT technique outperforms MED-HHT in enhancing the shaft speed information (with higher magnitude). Although the MED-WT (Figure 4.5h) has relatively higher magnitude than the MED-eHHT (Figure 4.5e), its spectral map contains more frequency components due to leakage, which may reduce the reliability for bearing health condition monitoring.

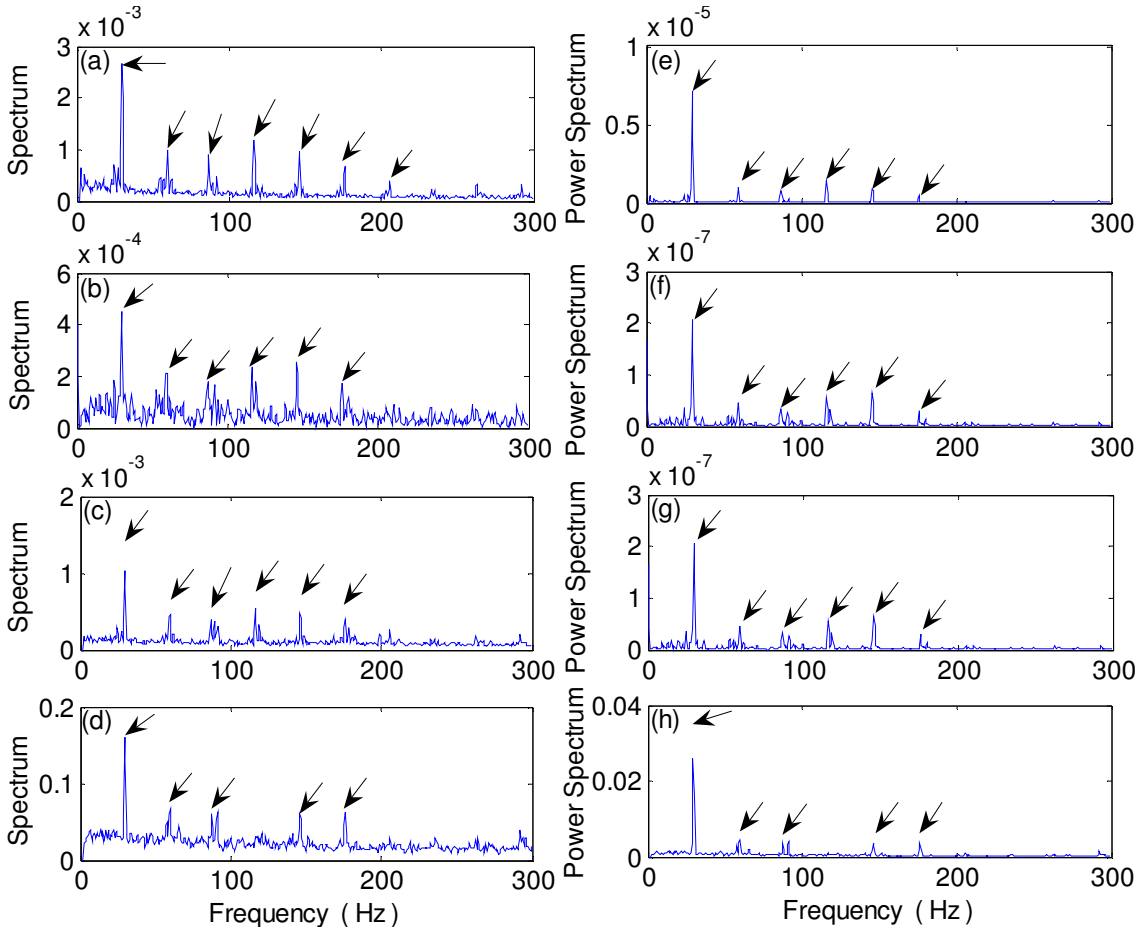


Figure 4. 5. Comparison of processing results for a healthy bearing using the related techniques: (a,e) MED-eHHT; (b,f) eHHT; (c,g) MED-HHT; (d,h) MED-WT. The results in the right column are scaled using spectral amplitude, and the results in the left column are with a scale of power spectrum. (Arrows point to the shaft speed and its harmonics)

3) Outer race fault detection

When defect occurs on the fixed ring race of a bearing (the outer race in this case), its defect-induced resonance modes do not change over time. In this case, the characteristic frequency is f_{OR} = 91 Hz. The fourth and ninth IMFs are selected for eHHT analysis, as shown in Figure 4.4 b3. Figure 4.6 shows the processing results using the related methods. Comparing Figure 4.6a (MED-eHHT) and Figure 4.6b (eHHT), the MED filter can effectively denoise the signal and result in higher magnitude of characteristic frequency components. It is seen that the proposed eHHT (Figure 4.6a) outperforms not only the classical HHT (Figure 4.6b) with higher magnitude of characteristic frequency components, but also the WT (Figure 4.6d) with lower noise levels. The main reason is due to effective information processing using the proposed NCM/DMI method.

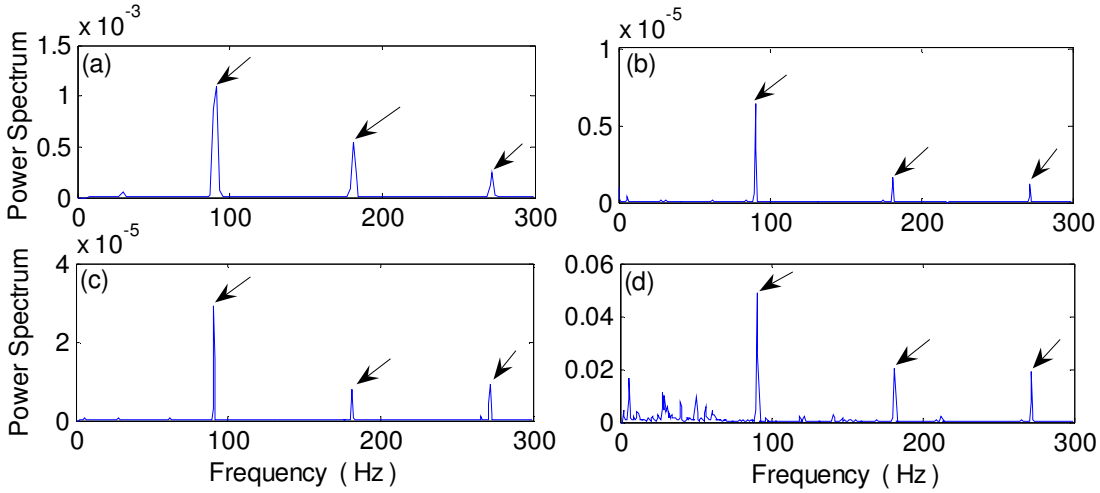


Figure 4.6. Comparison of processing results for a bearing with outer race fault detection using the techniques of: (a) MED-eHHT; (b) eHHT; (c) MED-HHT; (d) MED-WT. (Arrows point to the defect frequency and its harmonics)

4) Inner race fault detection

The detection of fault on rotating elements is considered more challenging than the detection of fault in the outer race; this is due to the fact that the inner race rotates and its resonance modes vary over time. In this case, the characteristic frequency f_{IR} = 148 Hz. As

shown in Figure 4.4 c3, the first and second IMFs are selected for eHHT analysis. Processing results using the related methods are shown in Figure 4.7. Comparing Figure 4.7a (MED-eHHT) and Figure 4.7b (eHHT), the MED filter can effectively denoise the signal and highlight characteristic frequency components. Furthermore, the proposed eHHT (Figure 4.7a) outperforms both the classical HHT (Figure 4.7b) and the WT (Figure 4.7d) due its efficient information processing.

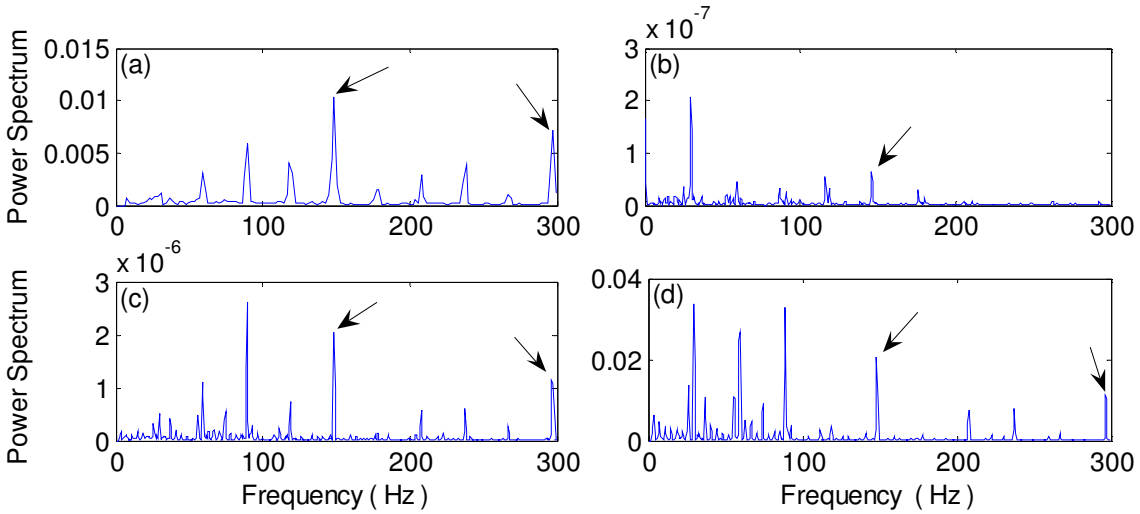


Figure 4.7. Comparison of processing results for a bearing with inner race fault detection using the techniques of: (a) MED-eHHT; (b) eHHT; (c) MED-HHT; (d) MED-WT. (Arrows point to the defect frequency and its harmonics)

5) Rolling element defect detection

The detection of fault on the rolling element (i.e., ball in this case) is considered the most challenging task in bearing fault detection. This is due to the fact that the rolling element rotates along different directions (it also slides), and its resonance modes change over time. In this case, the characteristic frequency is f_{BD} = 119 Hz. The first and second IMFs (in Figure 4.4 d3) are selected for eHHT analysis. Figure 4.8 shows the processing results using the related methods. It is clear that the proposed MED-eHHT is the only method that can detect the rolling element defect in this case (Figure 4.8a), due to its efficient MED denoising process and IMF information selection strategy. Neither eHHT (Figure 4.8b) nor MED-HHT (Figure 4.8c) can provide clear

defect frequency information in this case. Although the MED-WT can recognize the defect characteristic frequency 119 Hz (Figure 4.8d), it does not dominate the spectral map; other strong frequency components on the spectral map may result in false diagnostic result.

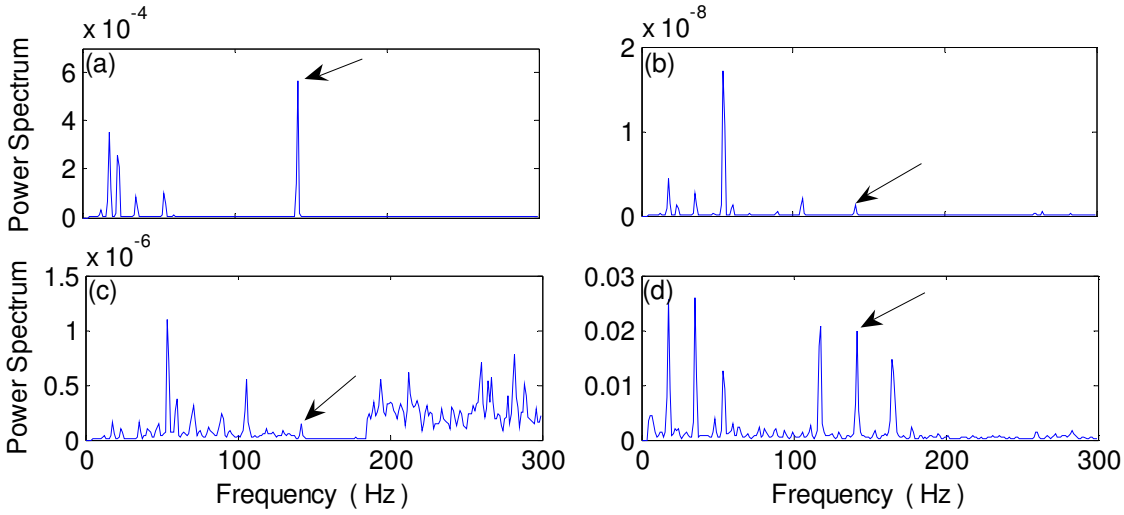


Figure 4.8. Comparison of processing results for a bearing with rolling element (ball) fault detection using the techniques of: (a) MED-eHHT; (b) eHHT; (c) MED-HHT; (d) WT. (Arrow points to the defect frequency)

As demonstrated in the previous figures, the proposed eHHT has detected the incipient fault effectively for all four tested conditions of bearings. This is due to its effective defect-related features extraction using the proposed NCM/DMI indicator integration.

Chapter 5

Robustness Verifications

5.1 Overview

To evaluate the robustness of the proposed NHHT and MED-eHHT techniques, different data sets from Case Western Reserve University (CWRU) [66] will be used for demonstration. Experiments were conducted using an experimental setup as shown in Figure 5.1. The system is driven by a 2-hp Reliance electric motor, with acceleration data measured at locations near to and remote from the motor bearings. Vibration data was collected using accelerometers, which were attached to the housing with magnetic bases. Accelerometers were placed at the 12-o'clock position at both the drive end and fan end of the motor housing. During some experiments, an accelerometer was attached to the motor supporting base plate as well. Vibration signals were collected using a 16-channel DAT recorder, and were post-processed in a Matlab environment. Digital data was collected at 12,000 samples per second with a sampling frequency. Speed and horsepower data were collected using the torque transducer/encoder and were recorded manually. Further details about the bearing properties, defect specifications, experimental setup, and test conditions can be found from [66]. The tested bearings have different health conditions (e.g., outer, inner, and rolling element faults). As an example, some data sets corresponding to shaft speed of 1,797 RPM (or $f_r \approx 30$ Hz) will be used for demonstration in this chapter. The corresponding characteristic frequencies are summarized in Table 5.1.

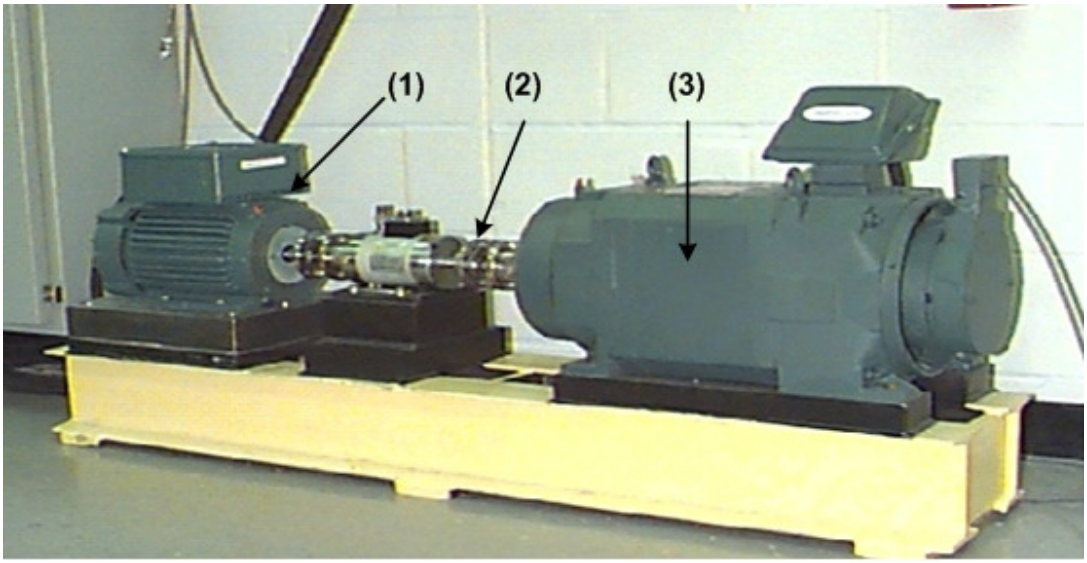


Figure 5.1. CWRU robustness test experimental setup [61]: (1) drive motor; (2) tested bearing; (3) load motor.

Table 5.1. Bearing characteristic frequencies at shaft speed of 30 Hz for bearings at CWRU [66]

Bearing Condition	Characteristic Freq (Hz)
Normal/healthy bearing	30
Outer race defect	107
Inner race defect	162
Rolling element defect	141

5.2 NHHT Robustness Tests

To evaluate the robustness of the proposed NHHT technique as proposed in Chapter 3, different data sets from the aforementioned CWRU public database on [66] will be used for demonstration. The processing results similar to related techniques used in Section 3.4 will also be shown for comparison.

1) Healthy bearing condition monitoring.

Figure 5.2 illustrates the comparison results for a healthy bearing with the characteristic frequency (i.e., the shaft speed) 30 Hz as listed in Table 5.1. Each technique can recognize the shaft speed information in this case. The proposed NHHT (Figure 5.2d), however, outperforms HHT (Figure 5.2c) and SHHT (Figure 5.2a) in terms of defect-related signatures due to its effective information processing using the information integration process, respectively. Although WT (Figure 5.2b) generates good processing results, it could not generate higher magnitude than NHHT (Figure 5.2d). NHHT can enhance spectral information of the shaft speed and its harmonics with reduced noise levels.

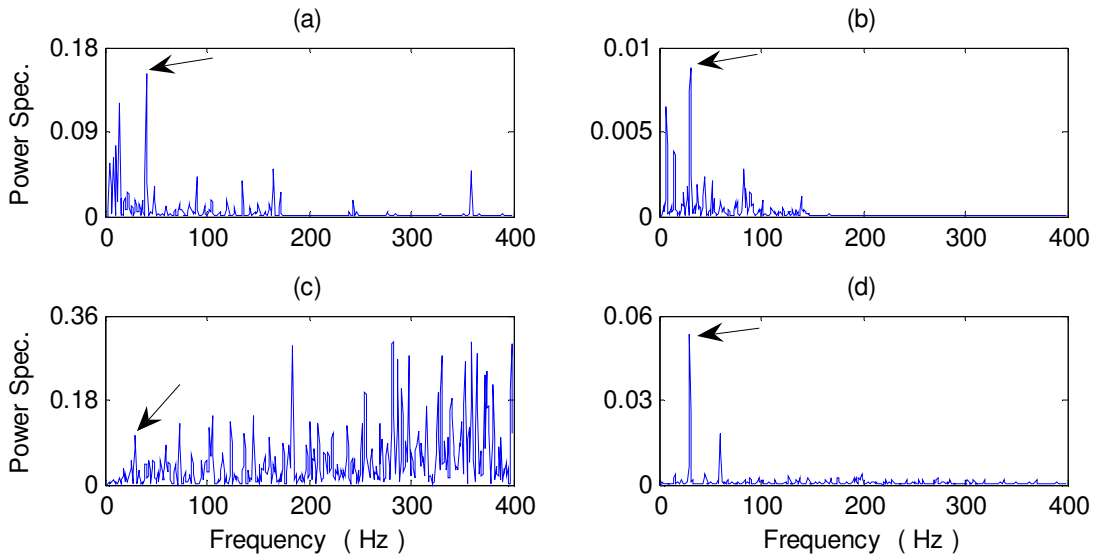


Figure 5.2. Comparison of processing results of CWRU dataset of monitoring healthy bearing with a scale of the power spectrum using the technique of: (a) SHHT; (b) WT; (c) HHT; (d) NHHT. (Arrow indicates the characteristic frequency)

2) Outer race defect detection.

Figure 5.3 shows the processing results for a bearing with an outer race defect with defect characteristic frequency of 107 Hz. It is seen that the proposed NHHT technique (Figure 5.3d)

outperforms HHT (Figure 5.3c), SHHT (Figure 5.3a), and WT (Figure 5.3b) in terms of recognizing defect and its harmonics related signatures.

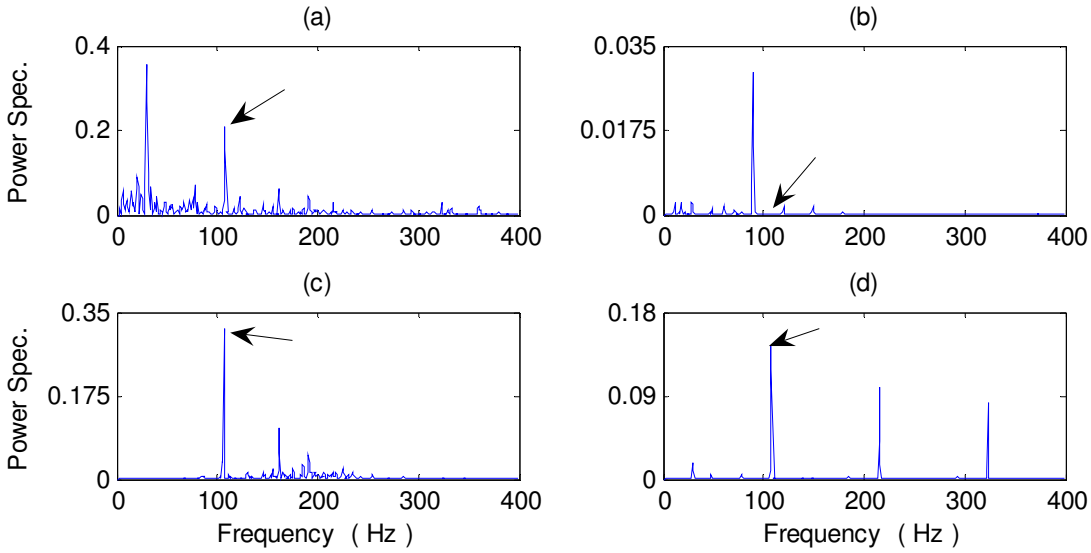


Figure 5.3. Comparison of processing results of CWRU dataset for bearing with outer race fault detection with a scale of the power spectrum using the techniques of: (a) SHHT; (b) WT; (c) HHT; (d) NHHT. (Arrow indicates the characteristic frequency)

3) Inner race fault detection.

Figure 5.4 shows the processing results of the related techniques for a bearing with inner race defect with f_{IR} = 162 Hz. In this case, the bearing fault can be clearly recognized by all the related techniques due to the relatively large size of the defect on the inner race. However, the proposed NHHT technique (Figure 5.4d) provides the best processing results in terms of spectral magnitude than SHHT (Figure 5.4a) and WT (Figure 5.4b) due to its efficient denoising process and information formulation approach. NHHT technique (Figure 5.4d) provides the related fault signatures with lesser noise level than HHT technique (Figure 5.4c).

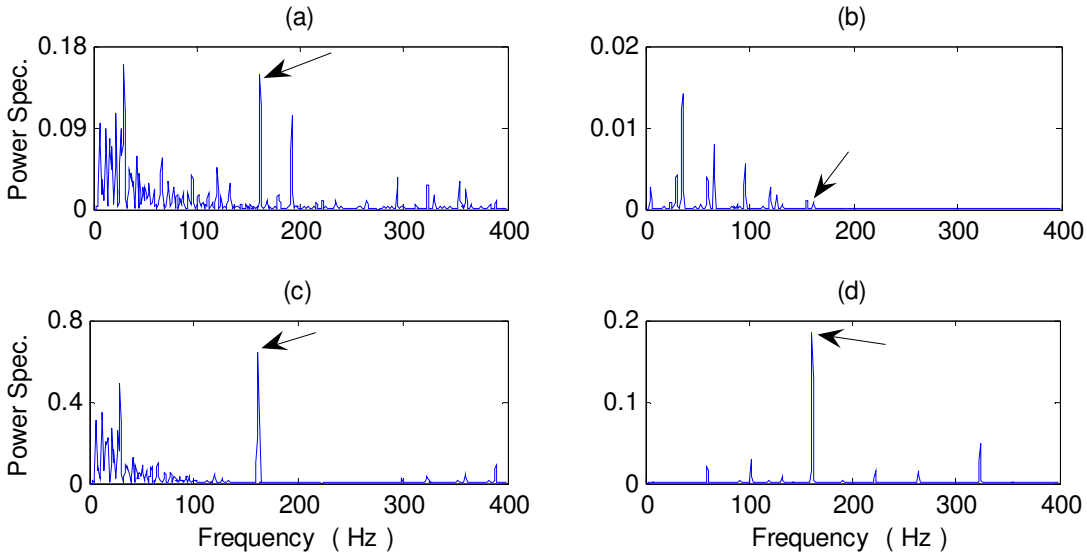


Figure 5.4. Comparison of processing results of CWRU dataset for bearing with inner race fault detection with a scale of the power spectrum using the techniques of: (a) SHHT; (b) WT; (c) HHT; (d) NHHT. (Arrow indicates the characteristic frequency)

4) Rolling element (ball) defect detection

Figure 5.5 illustrates processing results of the related techniques for a bearing with a rolling element defect f_{BD} = 141 Hz. Similarly, the proposed NHHT technique (Figure 5.5d) outperforms other related techniques; HHT (Figure 5.5c), SHHT (Figure 5.5a) and WT (Figure 5.5b). It can localize the disturbance more precisely and extract the related pulse information more effectively, which is important for nonstationary signal analysis.

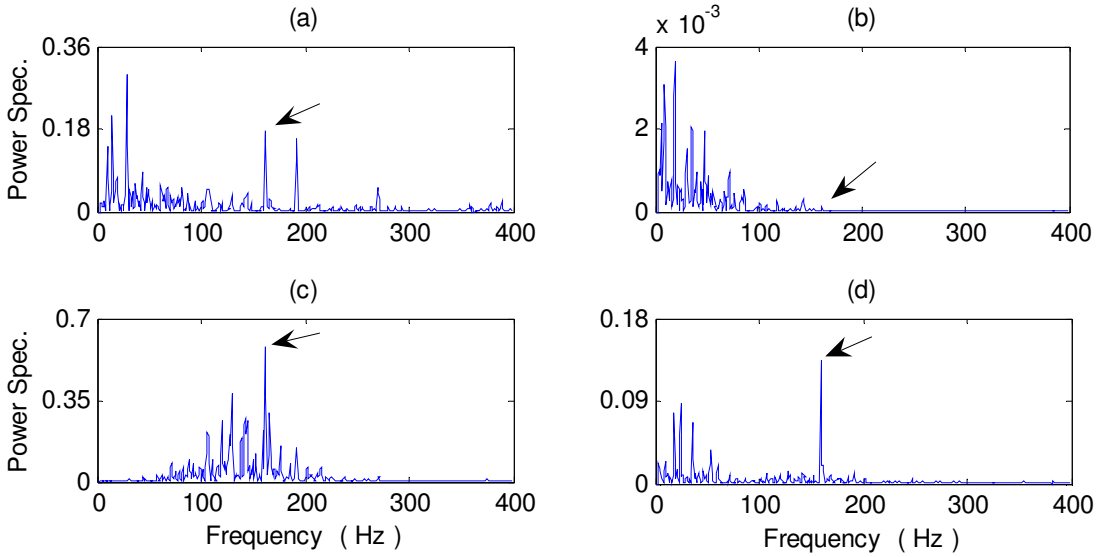


Figure 5.5. Comparison of processing results of CWRU dataset for bearing with rolling element (ball) fault detection with a scale of the power spectrum using the techniques of: (a) SHHT; (b) WT; (c) HHT; (d) NHHT. (Arrow indicates the characteristic frequency)

5.3. Robustness test for the proposed eHHT Technique

To evaluate the robustness of the proposed eHHT technique as proposed in Chapter 4, different data sets from the aforementioned CWRU public database on [66] will be used for demonstration. The processing results similar related techniques used in Section 4.3 will also be shown for comparison.

1) Healthy bearing condition monitoring

Figure 5.6 illustrates the comparison results for a healthy bearing with the characteristic frequency (i.e., the shaft speed) 30 Hz. Each technique can recognize the shaft speed information in this case. The proposed MED-eHHT (Figure 5.10a), however, outperforms eHHT (Figure 5.6b) and MED-HHT (Figure 5.6c) in spectral magnitude due to its effective denoising process using the MED filter. WT (Figure 5.16d) generates higher magnitude than MED-eHHT (Figure 6a); however, MED-eHHT can enhance spectral information of the shaft speed and its harmonics with less noise.

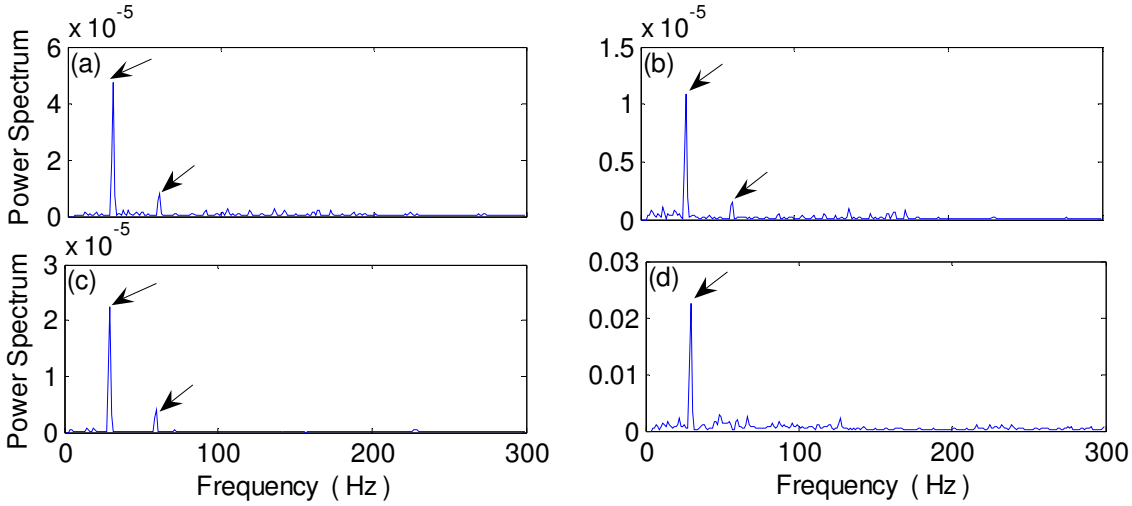


Figure 5.6. Comparison of processing results of CWRU dataset of monitoring healthy bearing using the technique of: (a) MED-eHHT; (b); eHHT; (c) MED-HHT; (d) MED-WT. (Arrows point to the shaft speed and its harmonics)

2) Outer race defect detection

Figure 5.7 shows the processing results for a bearing with an outer race defect with defect characteristic frequency of 107 Hz. It is seen that the proposed MED-eHHT technique (Figure 5.7a) outperforms eHHT (Figure 5.7b) and MED-HHT (Figure 5.7c) in terms of spectral magnitude. In this case, MED-WT cannot provide clear fault detection information (Figure 5.7d) as its characteristic frequency does not dominate the spectral map.

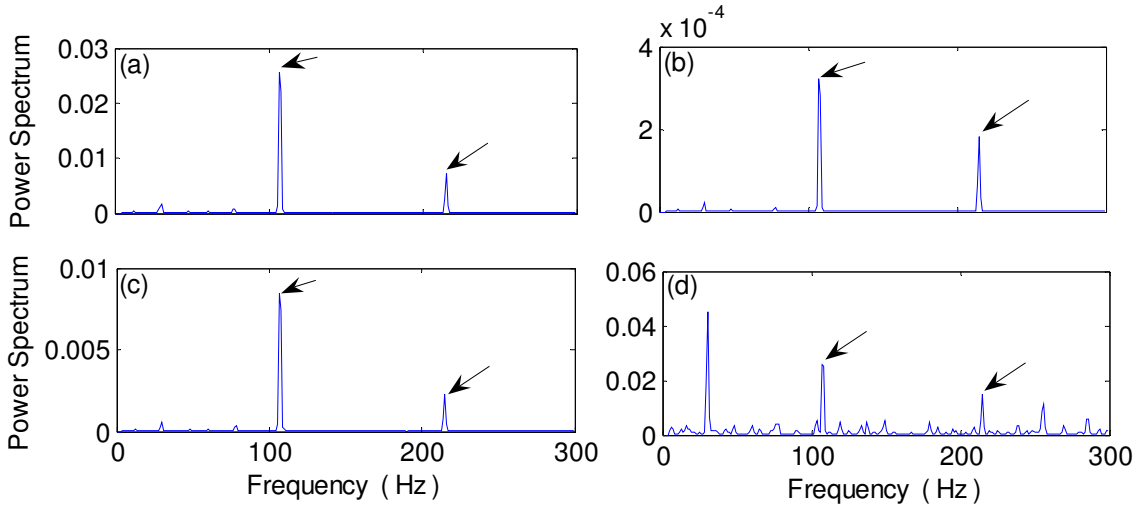


Figure 5.7. Comparison of processing results of CWRU dataset for bearing with outer race fault detection using the techniques of: (a) MED-eHHT; (b) eHHT; (c) MED-HHT; (d) MED-WT. (Arrows point to the defect frequency and its harmonics)

3) Inner race fault detection

Figure 5.8 shows the processing results of the related techniques for a bearing with inner race defect with f_{IR} = 162 Hz. In this case, the bearing fault can be clearly recognized by all of the related techniques due to the large size of the defect on the inner race. However, the proposed MED-eHHT technique (Figure 5.8a) provides the best processing results in terms of spectral magnitude than eHHT (Figure 5.8b) and MED-HHT (Figure 5.8c) due to its efficient denoising process and information formulation approach. In this case, although MED-WT provides clear fault information (Figure 5.8d), shaft speed and harmonics are still apparent in the spectrum.

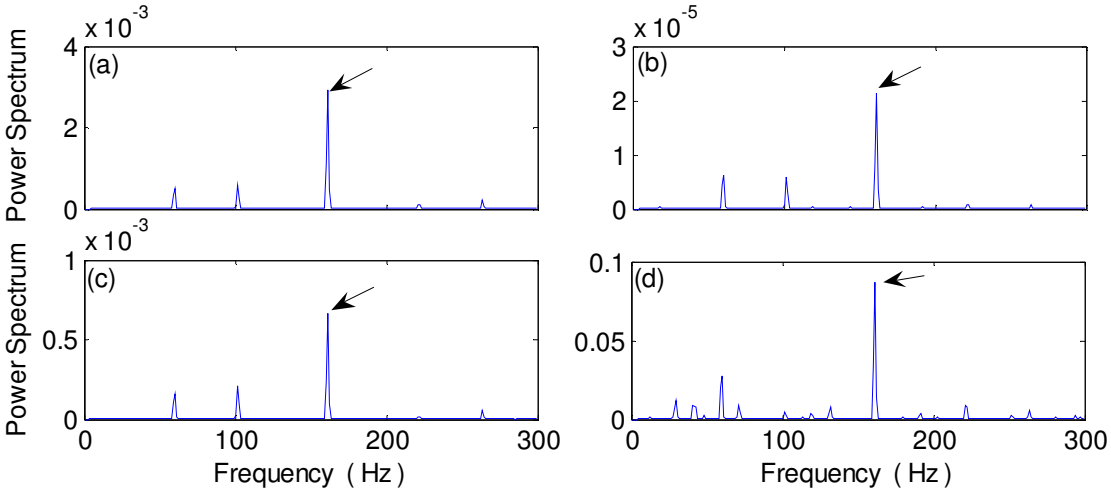


Figure 5.8. Comparison of processing results of CWRU dataset for bearing with inner race fault detection using the techniques of: (a) MED-eHHT; (b) eHHT; (c) MED-HHT; (d) MED-WT. (Arrows point to the defect frequency and its harmonics).

4) Rolling element (ball) defect detection

Figure 5.9 illustrates processing results of the related techniques for a bearing with a rolling element defect f_{BD} = 141 Hz. Similarly, the proposed MED-eHHT technique (Figure 5.9a) outperforms other related techniques such as the eHHT (Figure 5.9b), MED-HHT (Figure 5.9c) and MED-WT (Figure 5.9d). It can localize the disturbance more precisely and can extract the related pulse information more effectively, which is important for nonstationary signal analysis.

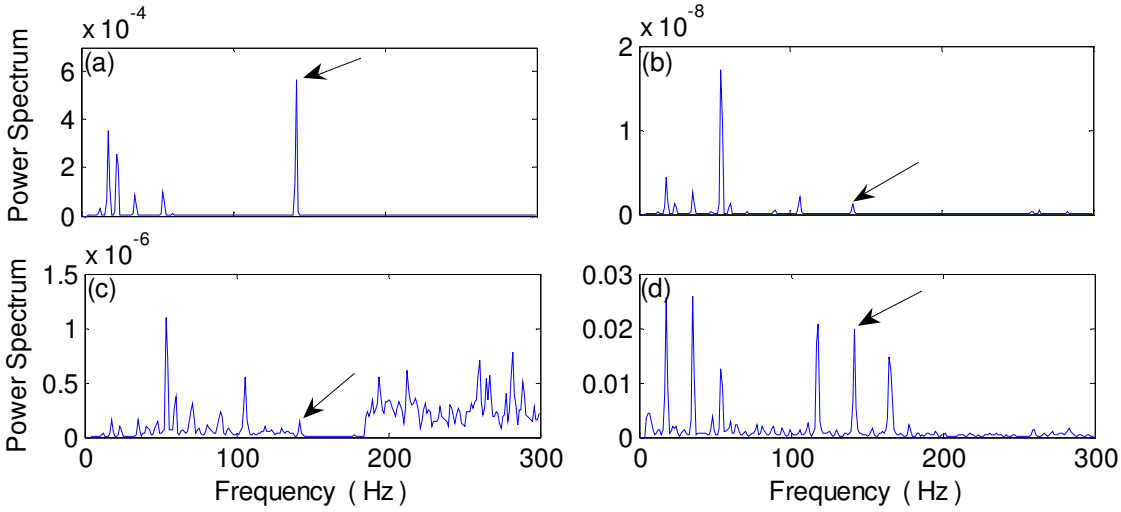


Figure 5.9. Comparison of processing results of CWRU dataset for bearing with rolling element (ball) fault detection using the techniques of: (a) MED-eHHT; (b) eHHT; (c) MED-HHT; (d) MED-WT. (Arrow points to the defect frequency)

5.4 Multi-Defects Detection Tests

It should be stated that this research focuses on the fault detection on initial bearing fault defects in order to prevent the related machinery performance degradation and malfunction. A bearing defect initiates from fatigue pitting (i.e., localized defect) on a bearing component, such as the fixed ring race, the rotating ring race, or a rolling element. If the damaged bearing remains in use, the defect will propagate to a wider area or to other bearing components. For example, if pitting occurs on the fixed ring race (usually the outer ring) first and the bearing fault cannot be detected properly (i.e., missed alarm), the pitted region cannot carry load anymore and consequently the unpitted region over the loading zone has to carry extra load. As a result, the contact stress increases, which leads to two effects: a) the unpitted region on the outer ring race is prone to fatigue damage due to higher contact stress, which leads to distributed fault condition; b) higher contact stress will damage other bearing components such as the inner ring race and rolling elements, which results in multiple defect condition.

In general, the currently available vibration-based fault detection techniques could be used to detect localized bearing defects; however, they cannot be used effectively to diagnose distributed defects, which is a limitation of vibration-based condition monitoring techniques [62]. In general, the distributed defect can be detected by the use of analysis of acoustic signals or lubricant properties [62].

As an example for multiple bearing defect detection, the tested bearing has simulated defects on its outer ring (fixed ring), inner ring, and a rolling element (a ball). The test is conducted using the same experimental setup as illustrated in Section 3.3. For comparison analysis of defect detection of multiple defects using the experimental setup from Section 3.1, a test result corresponding to shaft speed of 1800 RPM (i.e., 30 Hz) and load 2.5 is used as example to a typical fault detection analysis. The corresponding characteristic frequencies are f_r =30, f_{OR} = 91 Hz, f_{IR} = 148 Hz, and f_{BD} = 141 Hz. Figure 5.10 shows the processing results using NHHT technique. Similarly, the proposed NHHT technique (Figure 5. 10d) outperforms other related techniques such as HHT (Figure 5.10c), SHHT (Figure 5. 10a) and WT (Figure 5. 10b). It can localize the defect-related frequency and enhance the characteristic frequencies more clearly than other related techniques in this multiple bearing fault detection example.

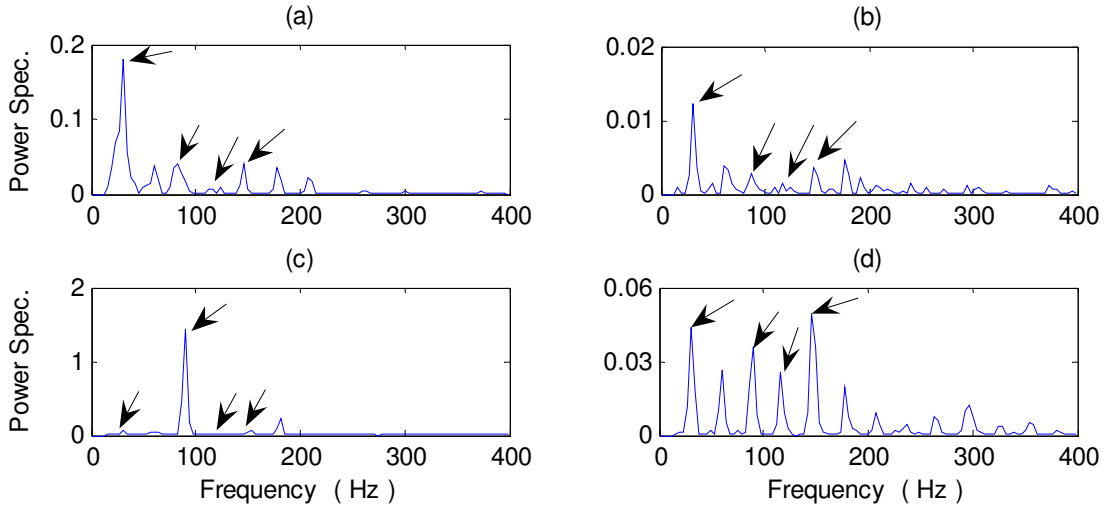


Figure 5.10. Comparison of processing results for bearing with combination of all faults detection with a scale of the power spectrum using the techniques of: (a) SHHT; (b) WT; (c) HHT; (d) NHHT. (Arrow indicates the characteristic frequency)

It should be stated that the currently available fault detection techniques could be used for localized and initial bearing fault detection only [62], that is, a single defect (not multiple defects). As a matter of fact, the result shown in Figure 5.10 is a specific example as the bearing housing is not coupled with any other mechanical equipment. If a bearing is used in a gearbox, the bearing signals are weak compared to gear signals, and will be modulated correspondingly. On a spectral map, bearing characteristic frequencies cannot be recognized as independent frequency components as shown in the graphs in the previous chapters. Instead, they become sidebands around these strong gear mesh frequencies and their harmonics. Correspondingly, reliable fault detection of even an initial bearing defect will become very challenging, not to mention multiple fault detection at our current technology levels.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

Rolling element bearings are widely used in rotary machinery. Reliable fault detection techniques are critically needed in wide array of industries to detect bearing defect at its earliest stage so as to prevent machinery performance degradation and malfunction. Many signal processing techniques have been proposed in the literature for bearing fault detection. However, these techniques are not robust in applications, especially in analyzing nonstationary features. The objective is this work is to develop new technologies for initial bearing defect detection.

This research focused on Hilbert-Huang transform, HHT, technique for feature extraction and analysis. Vibration signals are used for analysis in this work. The collected vibration signals are firstly processed by the use of MED filter to minimize the impedance effect of the signal transmission path and demodulate defect-related impulses. Then representative features are extracted and analyzed by the proposed Normalized Hilbert-Huang transform, NHHT. A novel IMF integration approach, based on the d'Agostino-Pearson (DP) normality measure, a weighted DP indicator, has been suggested for distinctive IMF integration and formulation. The effectiveness of the proposed NHHT technique has been verified by experimental tests corresponding to different bearing conditions. Test results have shown that the NHHT method can recognize more distinctive IMF functions than some popular approaches used in bearing fault detection analyses. The NHHT technique has been proven to be suitable for analyzing nonstationary signals, especially that of a bearing.

The second approach to improving HHT is integrating the IMF components from the standpoint of information entropy analysis. In the proposed enhanced HHT technique (or eHHT), a normalized correlation measure, NCM, and deficiency of mutual information, DMI, have been suggested for distinctive IMF selection and formulation. Firstly, test results have shown that the MED filter can effectively denoise the signal and highlight defect-related impulses. In the second part of the eHHT process, the most distinctive feature functions are chosen to be those with the highest NCM/DMI indicator values among the IMFs. These selected functions will contain

maximum correlation information (or minimum entropy), which will then be used to recognize most distinctive IMFs in the developed eHHT technique. The effectiveness of the proposed eHHT technique has been verified by experimental tests corresponding to different bearing conditions. The NCM/DMI method can recognize more distinctive IMF functions than the approach in classical HHT analysis. The eHHT technique has been proven to be suitable for analyzing nonstationary signals, especially that of a bearing.

To verify the robustness of the proposed techniques in this work, more tests have been conducted using different data sets from the CWRU public data base corresponding to different bearing conditions. Test results have shown that the proposed MED, NHHT and eHHT are efficient and robust signal processing techniques for nonstationary feature analysis and bearing fault detection. The NHHT and eHHT techniques outperform their corresponding classical fault detection techniques. They are efficient and robust techniques that have potential for bearing fault detection in real-world rotary machinery equipment.

Just like its name described, EMD is an empirical method. More than ten years after it was first introduced, EMD still lacked theoretical background. The research about its characters and limitations are highly based on different experiments. There are still many questions without answers. Some people compare the current situation of EMD with the wavelet. The principle of EMD is simple and intuitive. However, it seems that the intuitive character of the EMD alone makes it difficult to be described by mathematical theory. It is possible that there are more effective theories behind the EMD method, and EMD is just some light of those theories projected out of a window.

The difficulty in this project occurs when multiple defects exist within a bearing, the vibration signals will superimpose with each other. The characteristic waveform of a single defect signal cannot be found directly in the composed output data. In general, the currently available vibration-based fault detection techniques could be used to detect localized bearing defects; however, they cannot be used effectively to diagnose distributed defects.

6.2 Contributions from this work

Significant contributions have been made from the course of this work, which include two journal papers (one published) and one conference article:

[1] S. Osman and W. Wang, "A normalized Hilbert-Huang transform technique for bearing fault detection", Mechanical Systems and Signal Processing, (submitted July 2013).

[2] S. Osman and W. Wang, "An enhanced Hilbert–Huang transform technique for bearing condition monitoring", Measurement Science and Technology, 24, (13pp), 2013.

[3] S. Osman and W. Wang, "A new HHT technique for bearing health condition monitoring" Proceedings of the International Conference on Mechanical Engineering and Mechatronics Toronto, Ontario, Canada, August 2013.

6.3 Future Work

The following are some challenges from this work, which require advanced signal processing methods and further research analyses:

Mode mixing is a problem in EMD algorithm. It frequently appeared in EMD processing and is defined as a single IMF, consisting of either signals of widely disparate scales, or a signal of a similar scale residing in different IMF components. Usually, mode mixing is caused by signal intermittency, which is a type of disturbance signal without fixed form.

In this work, we focused on detecting and diagnosing bearing defects caused by fatigue spalling, which is the most predominant bearing failure mode. Further studies are required to include other bearing failure modes.

Although the presented methods have accurately recognized bearing conditions, the results were obtained using experimental data. In actuality, we deal with bearings mounted on real world application such as water treatment plant pumps, trains, etc., to validate robustness of proposed techniques. Vibration signals obtained from this environment are expected to have different characteristics than those obtained from a test rig in the lab.

References

[1] H. Hashemian, “State-of-the-art predictive maintenance techniques,” IEEE Transactions on Instrumentation and Measurements, 60, pp. 226–236, 2011.

[2] N. Tandon and B. Nakra, “The application of the sound intensity technique to defect detection in rolling element bearings,” Applied Acoustics, 29, pp. 207–217, 1990.

[3] D. Butler, “The shock-pulse method for the detection of damaged rolling bearings,” Non-destructive Testing, 6, pp. 92–95, 1973.

[4] J. Mathew and R. J. Alfredson, “The Condition Monitoring of Rolling Element Bearings Using Vibration Analysis”, Journal of Vibration and Acoustic, Stress and Reliability, 106, 447-453, 1984.

[5] N. Tandon, “A comparison of some vibration parameters for the condition monitoring of rolling element bearings,” Measurement, 12, pp. 285-289, 1994.

[6] A. Jardine, D. Lin, and D. Banjevic, “A review on machinery diagnostics and prognostics implementing condition-based maintenance”, Mechanical Systems and Signal Processing, 20, pp. 1483-1510, 2006.

[7] A. Choudhury and N. Tandon, “Vibration response of ball bearings in a rotor bearing system to a local defect under radial load,” Tribology, 128, n 2, pp. 252-261, 2006.

[8] N. Tandon and A. Choudhury, “An analytical model for the prediction of the vibration response of rolling element bearings due to a localized defect, Journal of Sound and Vibration, 205(3), pp. 275-292, 1997.

[9] S. Ericsson, N. Grip, E. Johansson, L. Persson, R. Sjoberg, and J. Stromberg, “Towards automatic detection of local bearing defects in rotating machines”, Mechanical Systems and Signal Processing, 19, pp.50-535, 2005.

[10] R.J. Alfredson and J. Mathew, “Time domain methods for monitoring the condition of rolling element bearings”, Institution of Engineers: Mechanical Engineering Transactions, ME10 (2), Australia, pp. 102-117, 1985.

[11] Images of accelerometer, velocity sensor, displacement sensor. www.intertechnology.com. (Retrieved July 2013).

[12] A. Choudhury and N. Tandon, “Vibration response of ball bearings in a rotor bearing system to a local defect under radial load,” Tribology, 128, pp. 252-261, 2006.

[13] Q. Miao, D. Wang, H. Huang, “Identification of characteristic component in frequency domain from signal singularities”, Review of Scientific Instruments, 81, 035113, pp. 1-7, 2010.

[14] J. Antoni, “The spectral kurtosis: a useful tool for characterizing nonstationary signals,” Mechanical Systems and Signal Processing, 20, pp. 282-307, 2006.

[15] C. Pachaud, R. Salvetat, and C. Fray, “Crest factor and kurtosis contributions to identify defects inducing periodical impulsive forces”, Mechanical Systems and Signal Processing, 11, pp. 903-916, 1997.

[16] H. Endo, R.B. Randall, “Enhancement of autoregressive model based gear tooth fault detection technique by the use of minimum entropy deconvolution filter”, Mechanical Systems and Signal Processing, 21, pp. 906-919, 2007.

[17] J.R. Stack, T.G. Habetler, and R.G. Harley, “Bearing Fault Detection via Autoregressive Stator Current Modeling”, IEEE Transactions on Industry Application, 40, pp. 740-747, 2004.

[18] P.D. McFadden and M.M Toozhy, “Application of synchronous averaging to vibration monitoring of rolling element bearings”, Mechanical Systems and Signal Processing, 14, pp. 891-906, 2000.

[19] W.J. Wang, P.D. McFadden, “Application of wavelets to gearbox vibration signals for fault detection”, Journal of Sound and Vibration, 192, pp. 927–939, 1996.

[20] R. B. Randall, J. Antoni, and S. Chobsaard, “The relationship between Spectral Correlation and Envelope Analysis in the Diagnostics of Bearing Faults and Other Cyclostationary Machine Signals”, Mechanical Systems and Signal Processing, 15 , pp.945-962, 2001.

[21] A.C. McCormick, and A.K. Nandi, “Cyclostationary in Rotating Machine Vibrations”, Mechanical Systems and Signal Processing, 12, pp. 225–242, 1998.

[22] T. Kaewkongka, Y. Au, R. Rakowski, and B. Jones, “A comparative study of short time Fourier transform and continuous wavelet transform for bearing condition monitoring,” International Journal of COMADEM, 6, pp. 41-48, 2003.

[23] N. Sawalhi, R.B. Randall, H. Endo, “The enhancement of fault detection and diagnosis in rolling element bearings using minimum entropy deconvolution combined with spectral kurtosis”, Mechanical Systems and Signal Processing, 21, pp. 2616-2633, 2007.

[24] J. Liu, W. Wang, F. Ma, “Bearing system health condition monitoring using a wavelet cross-spectrum analysis technique”, Journal of Vibration and Control, 11, pp. 1-11, 2011.

[25] K.F. Al-Raheem, A. Roy, K.P. Ramachandran, K.D. Harrison, S. Grainger, “Rolling Element Bearing Fault Diagnosis Using Laplace-Wavelet Envelope Power Spectrum”, EURASIP Journal on Advances in Signal Processing, 2007, pp.1-14, 2007.

[26] D. Ho and R.B. Randall, “Optimisation of Bearing Diagnostic Techniques using Simulated and Actual Bearing Fault Signals, Mechanical Systems and Signal Processing, 14, pp. 763 -788, 2000

[27] J. Antoni, “Fast computation of Kurtogram for the detection of transient faults”, Mechanical Systems and Signal Processing, 21, pp. 108-124, 2007.

[28] J. Liu, W. Wang, and F. Golnaraghi, “An Extended Wavelet Spectrum for Bearing Fault Diagnostics”, IEEE Transaction on Instrumentation and Measurements, 57, No. 12, pp. 2801-2812, 2008.

[29] P.D. McFadden and J.D. Smith, “ Vibration monitoring of rolling element bearings by the high-frequency resonance technique - a review”, Tribology International, 17, pp. 3–10, 1984

[30] J.C. Li, J. Ma, B. Hwang, “Bearing localized defect detection by bicoherence analysis of vibrations”, Transactions of the ASME Journal of Engineering for Industry, 117, pp. 625–629, 1995.

[31] W. J. Staszewski, K. Worden, and G. R. Tomlinson, “The frequency analysis in gearbox fault detection using the Wigner-Ville distribution and pattern recognition”, Mechanical Systems and Signal Processing, 11(5) 673–692, 1997.

[32] Rolling element bearing picture. http://en.wikipedia.org/wiki/Rollingelementbearing. (Retrieved November, 21 2012).

[33] Geometry of as typical ball bearing. http://ars.els-cdn.com/content/image/1-s2.0S0094114X06000516gr1.jpg. (Retrieved November,21 2012)

[34] Image of Wavelet transformation (WT): compared basis functions with compression factor https://commons.wikimedia.org/wiki/File:Basis_function_with_compression_factor.jpg (Retrieved July 2013).

[35] P. Tse, W.X. Yang, and H.Y. Tam, “Machine fault diagnosis through an effective exact wavelet analysis”, Journal of Sound and Vibration, 277, pp. 1005–1024, 2004.

[36] N. Tandon and N and A Choudhury, “A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings”, Tribology International, 32, pp. 469-480, 1999.

[37] Y. Yu and Y. Dejie, and Junsheng C., “A roller bearing fault diagnosis method based on EMD energy entropy and ANN,” Sound and Vibration, 294, pp. 269-277, 2006.

[38] Y. Lei, J. Lin, Z. He, and Y. Zi, “Application of an improved Kurtogram method for fault diagnosis of rolling element bearings”, Mechanical Systems and Signal Processing, 25, pp. 1738-1749, 2011.

[39] G.D. Manolakis, V.K. Ingle, and S.M. Kogon, “Statistical and adaptive signal processing”, McGraw-Hill, 2000.

[40] S. Mallat. “A Wavelet tour of signal processing”, Academic Press, 1999.

[41] Su, W., Wang, F., Zhu, H., Zhang, Z., and Guo, Z., “Rolling element bearing faults diagnosis based on optimal morlet wavelet filter and auto correlation enhancement”, Mechanical Systems and Signal Processing, 24, pp.1458-1472, 2010.

[42] H. Qiu, J. Lee, J. Lin, and G. Yu, “Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics”, Journal of Sound and Vibration, 289, pp. 1066-1090, 2006.

[43] A. M. Atto and Y. Berthoumieu, “Wavelet Packets of Nonstationary Random Process: Contributing Factors for Stationary and Decorrelation”, IEE Transaction on Information Theory, 58, pp. 317-330, 2012.

[44] C. J. Li and J. Ma, “Wavelet decomposition of vibration for detection of bearing-localized defect”, NDT&E international, 30, pp. 143-149, 1997.

[45] Z.K. Peng and F.L. Chu, “Application of the wavelet transform in machine condition monitoring and fault diagnosis: a review with bibliography”, Mechanical Systems and Signal Processing, 18, pp. 199-221, 2004.

[46] J. Liu, W. Wang, F. Golnaraghi, and K. Liu, “Wavelet spectrum analysis for bearing fault diagnosis”, Measurement Science and Technology, 19, 015105, pp. 1-9, 2008.

[47] R.A. Wiggins, “Minimum Entropy Deconvolution”, Geoexploration, 16, Elsevier Scientific Publishing, Amsterdam, pp. 21-35, 1978.

[48] C.A. Cabrelli, “Minimum entropy deconvolution and simplicity: A non-iterative algorithm”, Society of Geophysicists, 50, No. 3, pp. 394-413, 1984.

[49] J. Y. Lee and A.K Nandi, “Extraction of impacting signals using blind deconvolution”, Journal of Sound and Vibration, 232, pp. 945-962, 2000.

[50] S. Wensheng, F. Wang, H. Zhu, Z. Zhang, et al., “Rolling element bearing faults diagnosis based on optimal Morlet wavelet filter and autocorrelation enhancement”, Mech. Sys. Signal Proc. 24, 1458–1472, 2010.

[51] R.F. Nalewajski, “Applications of the information theory to problems of molecular electronic structure and chemical reactivity”, Int. J. Mol. Sci. 3, 237-59, 2002.

[52] W. He, Z. Jiang, and K. Feng, “Bearing fault detection based on optimal wavelet filter and sparse code shrinkage”, Measurement, 42, pp. 1092-1102, 2009.

[53] A. Hyvarinen, “Spare code shrinkage: Denoising of non-Gaussian data by maximum likehood estimation”, Neural Computation, 11, pp.1739-1768, 1999.

[54] N.E. Huang, Z. Shen, S.R. Long, M. Wu, H. Shih, N. Zheng, C. Yen, C.C. Tung, and H.H. Liu, “The empirical mode decomposition and the Hilbert spectrum for non-linear and nonstationary time series analysis”, Proceedings of the Royal Society of London Series A—Mathematical Physical and Engineering Sciences, 454, pp. 903–995, 1998.

[55] N.E. Huang, M. C. Wu, S.R. Long, S.R. Long, and W. Qu, P. Gloersen, and K.L. Fan, “A confidence for the empirical mode decomposition and the Hilbert spectral analysis”, Proceedings of the Royal Society of London Series A—Mathematical Physical and Engineering Sciences, 459, 2317-2345, 1998.

[56] N.E. Huang, Z. Shen, and S.R. Long, “A new view of non-linear water waves: the Hilbert spectrum”, Annual Review of Fluid Mechanics, 31, pp. 417–457, 1999.

[57] N.E. Huang, S. Riemenschneider, B. Liu, et al., Interdisciplinary Mathematical Sciences vol. 5: Hilbert-Huang Transform and its Applications”, Singapore World Scientific Publishing Co. Pte. Ltd, (2005), pp. 1-51.

[58] Y. Leia, J. Lina, Z. Hea, and M.J. Zuob, “A review on empirical mode decomposition in fault diagnosis of rotating machinery”, Mechanical Systems and Signal Processing, 35, pp.108–126, 2013.

[59] Q. Du and S. Yang, “Improvement of the EMD method and applications in defect diagnosis of ball bearings,” Measurement Science and Technology, 17, pp. 2355-2361, 2006.

[60] R. Yan and R. X. Gao, “Hilbert-Huang Transform-Based Vibration Signal Analysis for Machine Health Monitoring”, IEEE Transaction on Instrumentation and Measurement, 55, No. 6, pp. 2320-2329, 2006.

[61] Y. Yang, D. Yu, and J. Cheng, “A fault diagnosis approach for roller bearing based on IMF envelope spectrum and SVM”, Measurement, 40, pp. 943-350, 2007.

[62] M. Zvokelj, S. Zupan, and I. Prebil, “Non-linear multivariate and multiscale monitoring and signal denoising strategy using Kernel Principle Component Analysis Combined with Ensemble

Empirical Mode Decomposition method", Mechanical Systems and Signal Processing, 25, pp. 2631-2653, 2011.

[63] R.B. d'Agostino, A. Belanger, and R.B. d'Agostino Jr., "A suggestion for using powerful and informative test of normality", The American Statistician, 44 316-321, 1990.

[64] G. Poitras, "Economics Letters 90", pp.304–309, 2006.

[65] S. Osman and W. Wang, "A normalized Hilbert-Huang transform technique for bearing fault detection", Mechanical Systems and Signal Processing, (submitted July 2013).

[66] C. Studholme, D.L.G. Hill", D.J. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment, Pattern Recognition, 32", pp.71–86, 1999.

[67] S. Osman and W. Wang, "An enhanced Hilbert–Huang transform technique for bearing condition monitoring", Measurement Science and Technology, 24, (13pp), 2013.

[68] S. Osman and W. Wang, "A new HHT technique for bearing health condition monitoring" Proceedings of the International Conference on Mechanical Engineering and Mechatronics Toronto, Ontario, Canada, August 2013.

[69] CWRU, "Bearing Data Center, seeded fault test data". http://www.eecs.case.edu/ (Retrieved November, 28 2012).