A REVIEW ON SIGNAL PROCESSING TECHNIQUES FOR BEARING DIAGNOSTICS

Mohd Syahril Ramadhan Mohd Saufi, Zair Asrar Ahmad, Meng Hee Lim and Mohd Salman Leong
Institute of Noise and Vibration, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia

ABSTRACT

Bearing is the most widely used component in many applications such as home appliances, industrial applications and military applications. Bearing works continuously in harsh environment especially for industrial use where the environment factor may affect the bearing conditions. Conditions of bearing require a proper monitoring to prevent sudden failure which will cause financial loss and threaten human life. Thus, a proper maintenance is introduced called Condition-based Maintenance (CBM). CBM is a maintenance strategy that provide a guideline to monitor the asset condition based on the information collected. In CBM, there are several important steps for monitoring the asset condition which one of them is signal processing. Signal Processing is important due to data acquire from the sensor is heavily masked by the background noise (machine sound and environment sound). Therefore, a robust signal processing technique is required to eliminate the noise and provide good features for decision making. This paper tends to review on the signal processing utilised for bearing fault diagnosis from the previous researcher.

Key words: Signal processing, bearing fault diagnosis and condition based maintenance.

http://www.iaeme.com/IJMET/issues.asp?JType=IJMET&VType=8&IType=6

1. INTRODUCTION

The demand for condition monitoring for bearing condition have shown an increasing trend year by year. This is due to cost reduction require by the company to prevent unexpected failure which will cause major breakdown of the whole system where major financial loss and possible to threaten human life. Bearing components ease the translational and rotational movement of machine by reducing the friction of moving components. Industrial process requires the bearing components to work in extreme condition such as high temperature, high humidity and overstress that will lead the condition degradation’s time faster. Thus, Condition Based Maintenance (CBM) is introduced in the machinery field to avoid consequences failure
by providing decision for preventive maintenance. CBM has been used in bearing applications 
since the bearing components has high possibility to fail compare to other components. 
Around 1980’s, Condition Based Maintenance (CBM) is commercialised which the 
machine’s health condition can be examined using computer software by analysing the data 
from physical measurement such as vibration amplitude, rms, standard deviation etc. 
Vibration analysis has shown its effectiveness to diagnose any kind of machinery faults [1]. 
Based on the literature study, by far, vibration analysis is the highest usage for bearing 
diagnostics compare to other method such as acoustic measurement, oil analysis and motor 
current analysis. Figure 1 shows the CBM step for condition monitoring which one of them is 
signal processing. Signal processing is used to process raw signal by reducing the noise level 
without removing the important features in the signal. Instead of noise, there another factor 
that affect the signal acquired from the bearing condition which is complexity of bearing parts 
since they contain rolling elements, inner race, outer race and cage. During operating 
condition, these four components will interact each other cause the raw signal difficult 
interpret. Therefore, signal processing is needed to clean the bearing’s raw signal before 
feature extraction process. Signal processing method for bearing diagnostic has been paid a 
lot of attention from researchers. Many efforts have been done to improve signal processing 
method. Such efforts are improvement of non-adaptive signal processing to self-adaptive 
signal processing. Non-adaptive signal processing require user to select the parameter 
manually such as short-time Fourier transform and wavelet transform while the self-adaptive 
methods process the signal automatically without any selection of parameter from user.

Therefore, the paper reviewed on various signal processing tools utilised for bearing 
diagnosis. Section two will deal with the challenges of signal processing tools to analyse 
bearing signal. Also, three types of domain for signal analysis will be discussed. Signal 
processing techniques have been concluded in the section three.

2. SIGNAL PROCESSING AND ANALYSIS

Raw signal is always affected by noise and interference [2]. Thus, good signal processing 
tools is required to reduce noise and interference in the raw signal. Also, signal analysis able 
to turn the signal into more informative view or domain. Noise and interference affect the 
signal that will mislead the analysis of signal [3] and provide the wrong indication. Due to 
bearing’s structure and operating environment make it difficult to explain the signal result [4]. 
Based on Jayaswal et al. works, they discovered the factor that contributes to the interruption 
of bearing signal such as complexity of bearing component (inner race, outer race, ball 
bearing and cage), defect location, defect stage, and working condition (speed and load)[5].

2.1. Time Domain Signal Analysis

Time domain refers to analysis or display of signal axis with the function of amplitude and 
time. The advantage of time domain analysis is no data lost during the inspection but it often 
contains too much information. Time waveform analysis includes visual inspection of the 
vibration signals (time series data), time waveform parameters, probability density function, 
and probability density moments. Trending and comparison is used for a time waveform 
parameter. Time waveform parameter is a statistical value that can be calculated based on the 
raw vibration signal. Every signal contains several parameters that are very informative for
bearing diagnosis such as peak value, mean value, RMS value, kurtosis and skewness. Patil et al. reviewed on bearing signature analysis, where they stated that bearing is in good condition when kurtosis value close to 3 [6]. Besides, the probability density function represents the instantaneous amplitude value from a vibration signal within a certain amplitude range. Bearing in good condition produced the normal (Gaussian) distribution of probability density function. Probability density moments are based on the probability density function to provide first to fourth density moments for example: mean (first moment), standard deviation (second moment), skewness (third moments) and kurtosis (fourth moments) which are more informative. In addition, under different defect model, the statistical analysis will produce a different statistical feature of time domain data. Therefore, using time domain feature it is possible to diagnose the type of defect located on the bearing. Usually, the signal acquired after physical monitoring contains noise that affects the signal characteristics which may provide the wrong indication if the signal is directly used. A common method to reduce the noise is time synchronous averaging (TSA) which will reduce the noise level by the averaging process. Li et al. reviewed on TSA where they found that TSA technique capable to deal with vibration data by removing the non-coherent and non-synchronous components without the loss of information [7].

2.2. Frequency Domain Signal Analysis

Frequency domain refers to the function of the signal with respect to frequency instead of time. The main advantage of frequency domain over time domain is that the repetitive nature of the vibration signals is clearly displaced as peaks in the frequency spectrum at the frequency where the repetition takes place. Which means the defect may not cause a significant change to the whole signal and this signal can be affected by energy generated from non-defect related vibrations. But, the pulse produced at very short duration when the interaction occurs between the defect and rolling element bearings produce a significant change in a band of frequencies in which the non-fault related vibrations are relatively small. Frequency domain contains several analyses such as spectral analysis, power cepstrum analysis, power spectral density (PSD) analysis and envelope analysis. Fast Fourier Transform is one of the example methods by transforming time series to frequency spectrum (spectral analysis) that was proposed by J. Fourier. FFT based on periodic function which can be expressed as an infinite sum of periodic complex exponential function. Power cepstrum analysis is an inverse of Fourier Transform where the signal is converted to time domain that represents peaks corresponding to the period of the frequency in the spectrum. While PSD is derived after FFT analysis where the highest amplitude (power) at a given frequency can be obtained. PSD is useful for fault diagnosis and hidden periodicity finding. Envelope analysis is also known as high-frequency resonance techniques used to determine the resonance excites by the impacts. Several researchers utilised FFT during their research study. For example, FFT is utilised with AE analysis to diagnose the fault on air blower bearing. Gowid et al. proved the combination of FFT and AE is a powerful tool for diagnosing the fault in centrifugal equipment with a need of suitable classifier [8]. Hecke et al. utilised the inverse Fourier Transform (Power cepstrum) analysis in order to determine a new parameter that allows clear indication of all four fault in slow speed bearing [9]. However, an FFT-based method is not applicable for non-stationary signal analysis [10] due to various factor that affects the signal characteristics such as environment factor and failure in the machine [11]. Also, the inability to provide both time and frequency resolution restricted the effectiveness of frequency domain. Resolution problems of the Fourier Transform is due to Heisenberg’s Uncertainty Principle where a pair of physical properties cannot be known simultaneously.
2.3. Time-Frequency Domain Signal Analysis

2.3.1. Short Time Fourier Transform

In time-frequency analysis, the energy of power spectrum of waveform signal will be represented in both functions of time and frequency which better to reveal fault patterns for more accurate diagnostic. The ideal assumption that has been made for stationary time series is not useful for bearing diagnosis practice since the waveform produced will be attenuated by noise and other factors. Nowadays, researchers shifted to the time-frequency domain to analyse the energy distribution of the frequency components through time. Transient signal since the needs of time localisation of the spectral components. Thus, representations of time-frequency. Traditional techniques for time-frequency include the short time Fourier transform (STFT). STFT is one of the examples of time-frequency domain analysis. STFT use a sliding window to produce a spectrogram. In another word, they will divide the signal into small segments where these small segments of the signal will be assumed to be stationary. The width of segmented signal and window must be related to ensuring the stationarity of the signal. Next, Fourier transform is applied on each small segment. The STFT efficiency depends on the scale and type of window used for analysing the signal to obtain a good frequency resolution. STFT has some issue with the time and frequency resolution which makes the interpretation of signal is difficult. However, there is still a study using STFT is done by researchers. Fiorenzo et al. utilised STFT to obtain spectrograms and combine with deep learning to access bearing condition [12]. As mentioned, STFT has a problem with a resolution. Signals from the defect in the rotating machinery contain the informative parameter that reflects the condition of the components however, they will be masked by the background noise. So, the need of effective monitoring technique is required [13]. Over the past decades, a lot of signal processing techniques in time-frequency domains are developed which will discuss in the next section.

2.3.2. Wavelet Transform

Wavelet transform (WT) was developed as an alternative method to STFT. WT is a mathematical tool which adjusts consecutive data in the time domain to time-frequency domain using different translation and dilation function called ‘mother wavelet’. Wavelet transform decomposes signal into several scales at different levels of resolution and capable to analyse waveform data of bearing signal. Chen et al. reviewed on WT for rotating machinery fault diagnosis where a lot of WT development has been done such as Continuous WT (CWT), Discrete WT (DWT), Wavelet Packet Transform (WPT), Empirical WT (EWT), Dyadic WT, dual three complex WT (DTCWT) etc.[14]. However, the user need a good knowledge on choosing the best mother wavelet since different mother wavelet produces a different result. Many researchers put an effort to establish the proper way to select the optimal mother wavelet. For example, Hemmati et al. proposed ratio between kurtosis and Shannon entropy (KER) for mother wavelet selection. Authors select four orthogonal mother wavelet families which are Symlets DMeyer, Daubechies, and Coiflets where the mother wavelet is selected based on the highest KER values produced by the mother wavelet [15]. Also, Hemmati et al. summarised several techniques for mother wavelet selection like shape matching and maiming the cross-correlation function. Ngui et al. reviewed on mother wavelet selection on various application [16]. Generally, most of the time CWT, DWT and WPT are utilised for signal processing purpose. DWT decomposed signal into several levels comprised with the low pass approximation and high pass detailed coefficients where after first level only the detailed coefficient is decomposed further. Whereas, CWT able to work with every scale where the entire signal will be scaled and shifted over which sometimes lead to redundancy of information. Redundancy of information will consume a lot of time during
signal processing even though large dataset is useful for signal denoising and feature extraction. Also, DWT has several restrictions that limit its effectiveness for example shift invariances, aliasing, less directional selectivity, oscillation of wavelet coefficients and highly redundant representation that require higher computational cost. Unlike DWT and CWT, WPT decomposed both approximation and details to form a full binary tree. WPT enhanced signal decomposition capability in high-frequency regions which can be utilised to differentiate transient components in high-frequency characteristics.

Application of DWT, CWT and WPT has been performed in previous research. Khanam et al. implement DWT method by using symlet5 on vibration signal to predict the defect size on outer race ball bearing using discrete wavelet transform. Authors demonstrated the capability of wavelet analysis to split the peak corresponding to rolling element entry and exit since Symlet wavelet contains linear phase nature that maintains the sharpness of the signal. The defect size produced was 2.06 percent of standard deviation [17]. The technique of tracking the impulse occur during the roller element hit the entrance and exit of the defect are continue by another researcher with a different method. Hemmati et al. utilised WPT to determine bearing defect size by identifying the spike produced when roller element enters and exit the defect. Authors chose WPT instead of DWT due to as mentioned before DWT decomposed signal into approximation and detailed signals and only detailed components is decomposed further. Hence, the insufficient treatment occurs to the approximation level where the bearing fault impulses exist. Therefore, wavelet packet transform (WPT) has been used to analyse the time travel between two spikes in order to measure the defect size located on bearing components [15]. Wu et al. performed CWT to solve the problem of ICA. ICA unable to get independent source signals if each signal in N signals has not enough redundancy information [18]. The authors demonstrated that the final result indicated that the proposed scheme has the accuracy from 91.11 percent to 99.44 percent under varied speed [19]. Kumar et al. utilised CWT to process raw vibration signal and time marginal integration is calculated based on CWT coefficient. TMI is then de-noised by undecimated WT where then the feature is extracted from the de-noised TMI signal to predict the defect size. Author succeed to predict the outer and rolling element defect size with the maximum deviation 12.67 percent while the inner race defect size is unable to determine [20].

2.3.3. Empirical mode decomposition
Empirical mode decomposition (EMD) is a self-adaptive method for time-frequency signal processing where the non-stationary signal is decomposed into several intrinsic mode functions (IMFs). IMFs represent an oscillatory mode embedded in the signal and during the decomposition. In EMD, IMFs is decomposed from high to low frequency. Each IMF must fulfil several conditions which the number of zero crossings and number of extrema must be equal or slightly different at most by one in the whole data set. Next, at any point, the mean value of envelope defined by local maxima and local minima is zero [21]. EMD has shown a good performance on processing the nonlinear and non-stationary time domain signal. Lei et al. reviewed the EMD analysis on variety component on rotating machinery such as bearing, gear, and rotor [22]. From the summarization by these authors, EMD has several improvements such as Improved EMD, Multivariate EMD and Ensembled EMD. The application of EMD is quite extensive so far were a lot of researchers utilised EMD method since it able to deal with the non-linear and non-stationary signal. Mishra et al. observed that de-noising using EMD method is effective when dealing with unknown signal characteristics [23]. Due to the superior performance of EMD, a lot of researchers has adopted this method in fault diagnosis of rotating machinery. For example, Singh et al. utilised energy-based criterion to select the most significant IMFs. Also, authors combine pseudo-fault and EMD to filter the
most relevant fault information [24]. Lv et al. treated multi-sensor signal by combining multivariate EMD, multiscale reduction method and fault correlation factor analysis where the result obtained is useful for bearing problems [25]. Hiremath et al. reported EMD provided a better technique to detect the defect in the bearing and defect severity from AE measurement [26]. Nevertheless, EMD suffers a limitation of mode mixing. Thus, the development of EEMD is done to mitigate the limitation suffer by EMD. EEMD is an effective tool to decompose raw vibration data signal into several IMFs with different frequency band. The eemd process is a noise-assisted data analysis by adding Gaussian white noise into signals and develop multiple EMD to obtain multiple sets of IMFs [27]. The added white noise is to ensure the signal contain enough extrema to prevent mode mixing [28]. Zhao et al. compare EEMD with WT and high-frequency resonance technique (HFR). The result proved that EEMD provides a fine separation of the characteristic faults in the signal from the background noise and other interference [29]. Then, several authors implement Jensen Renyi Divergence with EEMD to assess degradation behaviour of rolling element bearings using vibration data. JRD is used to evaluate the sensitive IMFs by improving its sensitivity as a diagnostic parameter. JRD neglect interference of other unrelated components to the faults [30]. Zvokelj et al. proposed EEMD base multiple ICA to process AE and vibration signal for large dimension slewing bearing. The proposed scheme able to deal with high-dimensional multiscale data sets for fault diagnosis process [31]. In addition, Luo et al. extracted bearing fault parameter using EEMD as input parameters for their proposed hybrid system names as Gravitational search algorithm with an extreme learning machine for compound feature selection [32]. Liu et al. preprocessed the vibration signal from rolling element bearing using EEMD where the first three IMF containing useful information are selected [33].

### 2.3.4. Stochastic Resonance

Apart EMD, there is another method used in signal processing like Stochastic Resonance (SR). SR was introduced by Benzi [34]. This method is based on the effect of internal mechanism and external periodic forcing. Especially, SR is suitable for detecting weak signal that is submerged in background noise where Noise is treated as a resource to enhance the useful periodic signal in non-linear. SR is a useful method to enhanced weak components that mask by background noise and can be used for incipient fault diagnosis [35]. SR adjust nonlinear bi-stable system and the intensity of generated noise to be added to the raw signal. This adjustment is done to amplify the signal of interest. Sometimes, the signal acquired during monitoring is low. However, the signal can be boosted by adding white noise to signal that contain wide spectrum frequencies. The original signal will resonate and amplify with each other while not amplify the rest of white noise. Then, the signal-to-noise ratio will be increased. SR has been utilised in many fields where Loerincz et al. demonstrated that the output signal-to-noise ratio can be amplified if the optimal matching periodic signal, noise and dynamical system can be achieved [36]. Further, the undetectable signal can be filtered out effectively since the sensor able to detect the added white noise. It seems like EEMD method where white noise is utilised. However, the function of adding white noise for both signal processing has a different function where white noise in EEMD to avoid mode mixing while in SR is used to enhance the weak signal. In another word, SR method utilised on a weak signal where the useful signal is amplified instead of noise reduction [37]. But, SR still can be used as noise reduction where Zhong et al. applied noise reduction using SR in their research [38]. Li et al. enhanced vibration signal’s parameters for bearing fault diagnosis [37]. Authors emphasised that SR is effective in enhancing weak components from vibration signal. Also, Lai et al. utilised SR method during their study to diagnose fault from outer race bearing [35]. Omprusunggu et al. applied (SR) with an adaptation of multi noise tuning for multi-fault diagnosis on low rotational speed and low load rolling element bearings [39].
2.3.5. Variational mode decomposition

Since there is no established method for a variety of application, feature extraction techniques are still in development. Therefore, a lot of signal processing techniques is introduced to overcome the limitation of certain signal processing. EMD technique has high computational complexity and requires a large data series [40]. Also, EMD unable to avoid mode mixing problems. So, another method is introduced in bearing diagnostic area called Variational Mode Decomposition (VMD). VMD is developed by decomposing signal to form ensemble of band-limited intrinsic mode function. VMD adopted three basic concepts: Wiener Filtering, Hilbert Transform and Frequency Mixing which theoretically perform better than EMD [41]. The mode overestimation is reduced by lowering the errors between the extracted IMFs. In addition, the mode mixing problem can be reduced by reducing the influence of harmonic frequency and noise since each IMF is estimated according to its narrow-band property that make VMD method capable to find correct IMF more effective compared to EMD. VMD provide a proper way to extract the parameter from a noisy signal. The difference between EMD and VMD is that VMD decomposed input signal from low to high frequency while EMD does vice versa. Consequently, the result produced by VMD contains more high-frequency components compare to EMD due to decomposition techniques. Due to the nature of decomposition, VMD method is suitable to process high-frequency signal whereas EMD is suitable for low-frequency signal. Li et al. utilised VMD for bearing fault diagnosis. This author reported that VMD is effective to deal with weak and compound fault signal [42].

2.3.6. Spectral Kurtosis

Spectral kurtosis (SK) has been broadly used for fault diagnosis of rotating machines. Kurtosis is the fourth order of statistical parameter which value will close to zero if the signal is stationary and Gaussian noise. While large value will be produced when an impulsive signal that contains series of short transients like bearing fault signal [43]. Generally, SK is expressed by the ratio of fourth order cyclic cumulant and variance. (SK) was initially proposed by Dwyer where it was applied on the imaginary and real part of the STFT to overcome the deficiency of the power spectral density to spot and characterise the transient signal. Vrabie introduced unbiased estimator of SK by justifying the theoretical definition of SK [44]. Further, Antoni utilised Wold-Cramér decomposition for the SK estimation of non-stationary processes [45]. Also, Antoni was the first applied SK on bearing analysis. This method determines the presence of non-stationarities and specifies the event of frequency bands happen in the spectrum by calculating the kurtosis for each frequency line in frequency spectrum [46]. In addition, SK is used as a filter by reducing the background noise and select the most impulsive components from the signal. However, SK is easily hinder by noise and preprocessing procedures for the raw signal is necessary such as denoising and filter. Ruiz-Carcel et al. utilised SK on AE signal where they found that SK able to increase the signal to noise ratio of AE signal [47].

2.3.7. Others method

Apart from the above signal processing techniques, there is another method that suitable for bearing diagnostics. For example, Oscillatory Behavior-based Signal Decomposition (OBSD). This method decomposes a signal based on oscillatory behaviour instead of frequency or scale. By utilising the idea of Morphological Component Analysis (MCA), OBSD represent the signal based on high oscillatory wavelets and low oscillatory wavelets using Tunable Q-factor Wavelet Transform (TQWT). Since this method is quite new, the application of OBSD on bearing diagnosis field is limited. Huang et al. found that this method capable to analyse the raw signal well even though the interference frequency of the bearing’s signal is almost
identical to the fault characteristic frequency [48]. Based on this author’s work, this method is capable to deal with bearing signal. On the other hand, Largest Lyapunov method has shown it effectiveness to analyse signal of low-speed bearing. Caesarendra et al. used the largest Lyapunov exponent (LLE) algorithm to extract circular domain feature in low-speed bearing during condition monitoring. Authors find out this method can be as an alternative method if vibration-based FFT, time domain feature extraction and EMD are not applicable to deal with low energy non-linear vibration bearing signals [49]. Besides OBSD and LLE, local mean decomposition (LMD) is another that capable to process bearing signal. Liu et al. used local mean decomposition (LMD) for fault feature extraction on bearing [50]. LMD provide a series of modulated signal and envelope signal by decomposing the original signal onto a number of product functions (PF).

3. CONCLUSIONS
A brief review of signal processing used for bearing diagnosis has been presented in this paper. It is clear from various literature that Wavelet Transform and Empirical Mode Decomposition method are the most selected method to diagnose bearing signal. However, Variational Mode Decomposition has shown its effectiveness in dealing with noisy data. Since this method is new in bearing fault diagnosis, there are still a room for improvement where the selection number of mode for VMD still in open problems. Other techniques for signal processing tools have also been discussed.

REFERENCES


A Review on Signal Processing Techniques for Bearing Diagnostics


