



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

Doctor of Philosophy

**Fault Diagnosis and Size Estimation of Rolling Element
Bearing under Time-varying Speed Conditions**

The Graduate School
of the University of Ulsan
Department of Mechanical and Automotive Engineering

Guang-Quan Hou

**Fault Diagnosis and Size Estimation of Rolling
Element Bearing under Time-varying Speed
Conditions**

Supervisor: Professor Chang-Myung Lee

Author: Guang-Quan Hou


**Department of Mechanical and Automotive Engineering
University of Ulsan**

**A dissertation submitted to the faculty of the University of Ulsan in
partial fulfillment the requirement for the degree of Doctor of
Philosophy in the Department of Mechanical and Automotive
Engineering.**

Ulsan, Korea

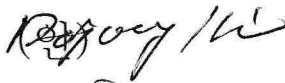




December 5th, 2018

Approved by



Professor Chang-Myung Lee

GUANG-QUAN HOU 의 공학박사학위 논문을 인준함

심사위원장	김도중	(인) 
심사위원	박성태	(인) 
심사위원	이병룡	(인) 
심사위원	전두환	(인) 
심사위원	이장명	(인) 

울산대학교 대학원

2018년 12월

ABSTRACT

Fault Diagnosis and Size Estimation of Rolling Element Bearing under Time-varying Conditions

Guang-Quan Hou

Department of Mechanical and Automotive Engineering

The Graduate School

University of Ulsan

Rolling element bearings are one of the most significant elements and frequently-used components in mechanical systems. Rolling bearing fault diagnosis and failure prognostics are helpful for preventing equipment failure and predicting the remaining useful life (RUL) to avoid catastrophic failure. Therefore, reliable fault detection is necessary to ensure productive and safe operations.

Spall size is an important fault feature for the RUL prediction, and most of research work has focused on estimating the fault size under constant speed conditions. However, estimation of the defect width size under time-varying speed conditions is still a challenge work. In this paper, a novel method is proposed to solve this problem.

The influence of speed variation on fault size estimation was investigated in the follow-up study. A resampling method was used to eliminate the effect of speed variation. The defect size can be calculated with the angle duration, which is measured from the identified entry and exit points.

To obtain better understanding of defect size estimation, a dynamic vibration model of a defective rolling bearing is established. The changes of contact deformation and force of the defective bearing pattern are obtained. The entry and exit events can be

identified by these illustrations. Then, two defect size estimation models are introduced from the small size model to the large model.

Based on the edited cepstrum and LMD (EC-LMD) algorithm, entry and exit events were enhanced to achieve a better diagnosis result than the classical methods. An improved LMD method is proposed to eliminate the end effect by the DTW technique. In order to verify the effectiveness of EC-LMD method, the experiment was performed with a fault bearing. The diagnosis results from the experimental data illustrated that the EC-LMD method could improve the diagnosis performance.

With the previous research work, a novel signal processing method combining EC-LMD, resampling and continuous wavelet transform was proposed for estimating the fault size of rolling element bearing under time-varying speed conditions. The combination method could not only diagnose the bearing fault but also estimate the fault size under time-varying speed condition. In order to prove the effectiveness and stability of this combination method, the real experiments were carried out. The estimation results show that the proposed method can effectively estimate the defect size on the outer race under time-varying speed conditions.

An intelligent rolling bearing fault diagnosis method was proposed. EC-LMD was used to pre-process the signal for extracting good features. The feature extraction was done by the MFE. Laplacian score was used to select the fault feature by reorder the scale factors. SVM is used to evaluate the classification performance. The experimental results showed that the different categories of rolling bearings are effectively identified by the proposed method.

ACKNOWLEDGMENTS

First of all, I would like to express my sincere gratitude to Prof. Chang-Myung Lee, my supervisor, for his invaluable guidance, encouragement, and living supporting during my Ph.D. program. Sincere gratitude is also expressed to the members of the advisory committee in the graduate school of the University of Ulsan, who are Prof. Byung-Ryong Lee, Prof. Do-Joong Kim, Prof. Jung-Soo Ryue, and Prof. Du-Hwan Chun and the other professors who had taught me and helped me in the fields of mechanical, automotive Engineering, sound and vibration.

Specifically, I wish to thank Prof. Hui He and associate Prof. Hao Liu, the supervisors of my M.D. program in the Liaoning University of Technology, for their encouragement, guidance, and help during the period of the graduate study. I wish to express my thanks to senior brothers Zhen-Hua Xu, Zhi Qiu, and Cong-hao Liu, for their recommendations, encouragements and help during my study in Korea.

Special thanks are extended to the members in the Sound and Vibration laboratory of the University of Ulsan, especially to Huan-Yu Dong, Peng Wang, Qi Wu, Min Chen, Hang Su, and Bably Das for their advice on this manuscript and the help of life in Korea. I would like to thank the members of the SCIEN company, who are Woo-Jin Kim, Chae-Rok Lim, and Young-Wook Bae, for their guidance and help for experiments implementation during my Ph.D. program. Same thanks are also regarded to my friends, Ya-Li Yang, Jing Liu, Jue Wang, Jin-Cheng Wang, Ji Liu, Tian-Jun Zhou, Ying-Xiao Yu, etc., for their help in my daily life.

Finally, I would like to express my sincere gratitude to my parents for their understanding and encouragement.

CONTENTS

ABSTRACT	i
ACKNOWLEDGMENTS.....	iii
CONTENTS	iv
LIST OF FIGURES.....	vi
LIST OF TABLES.....	ix
ABBREVIATIONS	x
NOMENCLATURES	xii
Chapter 1 Introduction	1
1.1 Research background	1
1.2 Review of bearing fault diagnosis and enhancement techniques.....	3
1.2.1 Pre-processing method	3
1.2.1.1 Time synchronous averaging (TSA)	4
1.2.1.2 Linear prediction	5
1.2.1.3 Cepstrum pre-whitening.....	5
1.2.2 Enhancement techniques	6
1.2.2.1 Spectral kurtosis (SK)	6
1.2.2.1 Discrete wavelet transform	8
1.2.2.2 Local mean decomposition	11
1.3 Time-frequency analysis for fault size estimation	12
1.3.1.1 Short time Fourier transform.....	12
1.3.1.2 Wigner-Ville distribution.....	14
1.3.1.3 Continuous wavelet transform	15
1.4 Research purposes and methods.....	16
1.5 Main contents and organization	16
Chapter 2 Effects of speed variation on fault diagnosis and size estimation.....	18
2.1 Introduction	18
2.2 Bearing degradation with contact fatigue.....	18
2.2.1 Response of the fault bearing	22
2.3 Bearing fault diagnosis.....	23
2.3.1 Fault detection	23
2.3.2 Bearing fault severity analysis	25
2.4 Estimation of the defect size	27
2.5 Influences of the rotating speed on vibration signal	30
2.5.1 Spectral analysis under time-varying conditions	31
2.5.2 Parameters of the vibration signal under time-varying conditions ..	32
2.5.3 Transmission path under time-varying conditions	33
2.6 Order tracking	34
2.7 Conclusions	37
Chapter 3 Dynamic vibration model of a defective bearing.....	38
3.1 Introduction	38

3.2	Bearing dynamic model	38
3.2.1	Kinematics of the roller.....	39
3.2.2	Contact deformation during roller passing the defect area	41
3.2.3	Hertzian contact model	41
3.2.4	Damping force.....	42
3.2.5	Vibration equations of motion.....	43
3.2.6	Analysis result	44
3.2.6.1	Bearing defective response	44
3.2.6.2	Different defect sizes with large depth size.....	45
3.2.6.3	Different defect sizes with small depth	48
3.2.6.4	Load effect.....	50
3.3	Estimation model of the defect size	50
3.3.1	Estimation model of small defect size	51
3.3.2	Estimation model of the large defect size	54
3.4	Conclusions	55
Chapter 4	Bearing fault component extraction	57
4.1	Introduction	57
4.2	Cepstrum	58
4.3	Fault component extraction by LMD	59
4.3.1	Introduction of LMD.....	59
4.3.2	DTW-LMD method.....	62
4.4	Experiment setup.....	68
4.4.1	Fault diagnosis	69
4.4.2	Fault size estimation.....	73
4.5	Conclusions	77
Chapter 5	Time-frequency analysis for fault size estimation	78
5.1	Introduction	78
5.2	Morlet wavelet	78
5.3	Parameter optimization	80
5.4	Experiment setup.....	81
5.5	Results and discussion.....	83
5.6	Conclusions	90
Chapter 6	Bearing Fault Identification	91
6.1	Introduction	91
6.2	Feature extraction.....	91
6.2.1	Fuzzy entropy.....	92
6.2.2	Multi-scale fuzzy entropy	93
6.3	Laplacian Score for feature selection	94
6.4	Pattern classification with Support Vector Machines (SVM).....	96
6.5	Experimental validation	99
6.6	Conclusions	107
Chapter 7	Conclusions	108
	REFERENCES.....	111

LIST OF FIGURES

<i>Figure 1-1. Paving of the frequency/frequency resolution plane[13].</i>	7
<i>Figure 1-2. Combinations of center frequency and bandwidth for the 1/3-binary tree kurtogram estimator [13].</i>	8
<i>Figure 1-3. Mother wavelets of wavelet family.</i>	10
<i>Figure 1-4. ‘Symmlet 8’ wavelet at various scales and locations.</i>	10
<i>Figure 1-5. A multi-level analysis and synthesis process of DWT [34].</i>	11
<i>Figure 1-6 Procedure of the STFT [34].</i>	14
<i>Figure 2-1. progression of the damage across the raceway[60]</i>	20
<i>Figure 2-2. Typical signals and envelope signals of local faults in a rolling element bearing [61].</i>	21
<i>Figure 2-3. Three Evolution of the averaged amplitude of the defective frequency[69].</i>	26
<i>Figure 2-4 Vibration signal of a defect bearing. (a) Diagram of a rolling element traveling into a line spall defect located on the outer raceway. (b) Typically measured vibration response of a line spall defect [73].</i>	29
<i>Figure 2-5 A illustration of defect size estimation. (a) Envelope signal of defective fault. (b) Illustration of the entry and exit events [77].</i>	30
<i>Figure 2-6. Schematic representation of ideal vibration signals in constant (top) and variable speed (bottom) operation with bearing outer raceway fault</i>	32
<i>Figure 2-7. Sampling signal under constant speed condition. (a) Speed file, (b) Sampling point in the time domain, (c) Sampling point at the angular domain, (d) Sampled signal.</i>	34
<i>Figure 2-8. Sampling signal under time-varying speed condition with a constant time interval. (a) Speed file, (b) Sampling point in time domain, (c) Sampling point at the angular domain, (d) Sampling signal.</i>	35
<i>Figure 2-9. Sampling signal under time-varying speed condition with constant angle interval. (a) Speed file, (b) Sampling point in time domain, (c) Sampling point at the angular domain, (d) Sampling signal.</i>	36
<i>Figure 3-1. Diagram of bearing dynamic model.</i>	39
<i>Figure 3-2. Relative position of the components for the jth roller [89].</i>	40
<i>Figure 3-3. Experimental and simulated vibration response.</i>	45
<i>Figure 3-4. Load direction and distribution on the bearing.</i>	46
<i>Figure 3-5. Defect information of different sizes with large depth.</i>	47
<i>Figure 3-6. Simulation result of different defect sizes with large depth value. (a) Contact deformation (b) Contact force.</i>	48
<i>Figure 3-7. Defect information of different width sizes with the small depth value.</i>	49
<i>Figure 3-8. Simulation result of different defect sizes with the small depth. (a) Contact deformation, (b) Contact force.</i>	49
<i>Figure 3-9. Simulation result of the dynamic model with different loads. (a) Contact deformation of the roller, (b) Contact force of the roller.</i>	50
<i>Figure 3-10. Events of a roller passes the defect area.</i>	51
<i>Figure 3-11. Estimation model</i>	51
<i>Figure 3-12. Model of roller passing over a small defect.</i>	52

<i>Figure 3-13. Model of roller passing over a small defect.</i>	54
<i>Figure 4-1. A flowchart of the cepstrum editing procedure.</i>	59
<i>Figure 4-2. A cost matrix with the minimum-distance warp path.</i>	64
<i>Figure 4-3. Result of DTW processing. (a) the original signals of the two sequences. (b) Warped signal of two sequences.</i>	65
<i>Figure 4-4. A simulation signal.</i>	66
<i>Figure 4-5. Decomposition result of the original signal by LMD.</i>	67
<i>Figure 4-6. Extension signal with DTW.</i>	67
<i>Figure 4-7. Decomposition result of improved LMD method.</i>	67
<i>Figure 4-8. Test rig of the rolling bearing with outer race fault.</i>	68
<i>Figure 4-9. A line Defect in the outer race.</i>	69
<i>Figure 4-10. Speed signal. (a) Key-phasor signal. (b) Speed signal.</i>	70
<i>Figure 4-11. Original signal and its order spectrum. (a) Original signal. (b) Order spectrum of the resampled signal.</i>	70
<i>Figure 4-12. Envelope analysis by low-pass filter.</i>	71
<i>Figure 4-13. Envelope analysis by fast kurtogram. (a) Spectrum of Kurtosis. (b) Envelope spectrum of the filtered signal by SK.</i>	72
<i>Figure 4-14. Flowchart of the fault diagnosis approach.</i>	72
<i>Figure 4-15. Pre-whitening signal by edited cepstrum.</i>	73
<i>Figure 4-16. Decomposition result of the pre-whitening signal by improved LMD.</i>	74
<i>Figure 4-17. Envelope spectrum of PFI.</i>	74
<i>Figure 4-18. Vibration signal after pre-whitening.</i>	75
<i>Figure 4-19 Vibration signal by the proposed method.</i>	76
<i>Figure 5-1. Morlet wavelet with different center frequency values (bandwidth is 3).</i>	79
<i>Figure 5-2. Effect of bandwidth parameter on Morlet wavelet in the time domain.</i>	80
<i>Figure 5-3. Effect of bandwidth parameter on Morlet wavelet in the frequency domain.</i>	80
<i>Figure 5-4. Rolling bearing test rig.</i>	82
<i>Figure 5-5. Different fault sizes on the outer races.</i>	82
<i>Figure 5-6. Vibration and speed signal of a bearing with different fault sizes on the outer race: (a, c, e) original vibration signal. (b, d, f) original speed signal.</i>	84
<i>Figure 5-7. Pre-whitening vibration signal by edited cepstrum: (a) 0.75-mm. (b) 1.5-mm. (c) 2.5-mm.</i>	85
<i>Figure 5-8. LMD decomposition result of the resampled signal with different fault sizes on outer race: (a) 0.75 mm, (b) 1.5 mm, (c) 2.5 mm</i>	86
<i>Figure 5-9. Relation between wavelet entropy and the center frequency with a bandwidth of 400: (a) 0.75 mm, (b) 1.5 mm, (c) 2.5 mm.</i>	87
<i>Figure 5-10. Relation between wavelet entropy and bandwidth parameters with the center frequency setting: (a) 0.75 mm, (b) 1.5 mm, (c) 2.5 mm.</i>	87
<i>Figure 5-11. Relation between wavelet entropy and center frequency with optimal bandwidth: (a) 0.75 mm, (b) 1.5 mm, (c) 2.5 mm.</i>	88
<i>Figure 5-12. CWT analysis of the outer race fault: (a) CWT scalogram of 0.75-mm defect, (b) CWT scalogram of 1.5-mm defect, (c) CWT scalogram of 2.5-mm defect. (d-f) Detailed analysis of corresponding defect size.</i>	89
<i>Figure 6-1. Find the maximum margin of the two classes.</i>	97

<i>Figure 6-2. Rolling bearing experiment system.</i>	99
<i>Figure 6-3. Flow chart of the proposed fault diagnosis algorithm.</i>	101
<i>Figure 6-4. vibration acceleration signal of each rolling bearing condition.</i>	102
<i>Figure 6-5. MFE over 20 scales of the pre-processing signal with the average of fifty trails.</i>	102
<i>Figure 6-6. LS result of MPE over 20 scales.</i>	103
<i>Figure 6-7. Distribution of RMS and kurtosis features (Legends of 1-8 are corresponding class labels of Table 6-1).</i>	104
<i>Figure 6-8. Distribution of MFE features over 20 (Legends of 1-8 are correspond class labels of Table 6-1).</i>	105
<i>Figure 6-9. Distribution of selected MFE features by LS (Legends of 1-8 are corresponding class labels of Table 6-1).</i>	105
<i>Figure 6-10. Classification results of SVM using RMS and kurtosis feature.</i>	106
<i>Figure 6-11. Classification results of SVM using selected MFE.</i>	107

LIST OF TABLES

<i>Table 3-1. Geometrical parameters of the test bearing</i>	44
<i>Table 4-1. Fault characteristic order</i>	69
<i>Table 4-2. Estimation results of the defect size on the outer race</i>	76
<i>Table 5-1. Geometrical parameters of the test bearing</i>	83
<i>Table 5-2. Estimation results of the defect size on the outer race</i>	90
<i>Table 6-1. Details of experimental data sets</i>	100

ABBREVIATIONS

RUL	Remaining useful life
IAS	Instantaneous angular speed
SSA	Synchronous signal averaging
FCF	Fault characteristic frequency
TSA	Time synchronous averaging
AR	Autoregressive
CPW	Cepstrum pre-whitening
WT	Wavelet transform
EMD	Empirical mode decomposition
LMD	Local mean decomposition
SK	Spectral kurtosis
STFT	Short-time Fourier transform
DWT	Discrete wavelet transforms
SVD	Singular value decomposition
TFA	Time-frequency analysis
WVD	Wigner-Ville distribution
CWT	Continuous Wavelet Transform
BPFO	Ball pass frequency of outer ring
BPFI	Ball pass frequency of inner ring
BSF	Ball spin frequency
FTF	Fundamental train frequency
RMS	Root-Mean-Square

OT	Order tracking
PF	Product function
DTW	Dynamic time warping
ED	Euclidean distance
EC-LMD	Edited cepstrum and local mean decomposition
LPF	Low pass filter
AEn	Approximate entropy
SEn	Sample entropy
FE	Fuzzy entropy
MFE	Multiscale fuzzy entropy
LS	Laplacian Score
SVM	Support vector machines

NOMENCLATURES

$F(n, f)$	Complex envelope of frequency signal
$K(f)$	Spectral kurtosis value
ψ	Mother wavelet
ψ^*	Complex conjugate of mother wavelet
A	Approximations of discrete wavelet decomposition
D	Details of discrete wavelet decomposition
$\omega(t)$	Window function
$X(\tau, f)$	Fourier transform of segment signal
τ	Time index
W_c	Continuous wavelet transform of a time signal
d	Diameter of the rolling element
D_p	Pitch diameter of bearing
n	Number of rolling elements
f_r	Shaft speed in frequency
α	Contact angle
$d(t)$	Impulse component of a defective bearing
$q(t)$	Variations of the load distribution experienced
$a(t)$	Variations of the transfer function between the measurement location and the point at fault location
$m(t)$	Decay of a unit impulse
O_{BPO}	Ball pass order of outer ring
O_{BPI}	Ball pass order of inner ring
O_{BP}	Ball pass order

O_{FT}	Fundamental train order
$H_{fault}(t)$	Fourier transform of x_{fault}
$H(f)$	Frequency response function (FRF) of the entire transmission path
ω_c	Angle speed of the bearing cage
M_o	Mass of the outer raceway attached to support housing
M_i	Mass of the inner raceways with the shaft
m	Mass of the rolling elements
F_l	static load
O_{in}	Centers of the inner race
O_{out}	Centers of the outer race
P_j	Center of the j th roller
r	Radii of the roller
$\alpha_{out,j}$	Angular position of the j th roller
$\delta_{in,j}$	Contact deformation between j th roller and inner race
$\delta_{out,j}$	Contact deformation between j th roller and outer race
$\gamma(\alpha)$	Shape function of the defect
Q	Contact force
K	Contact stiffness
$Q_{in,x}, Q_{in,y}, Q_{out,x}, Q_{out,y}$	Total contact forces between roller and races in the x-axis and y-axis directions
$F_{d,in,j}, F_{d,out,j}$	Radial contact damping forces between j th roller and

	both raceways
c_b	Damping coefficient of roller
c_s	Damping coefficients of the housing
K_{lin}	Linear stiffness of the bearing
K_s	Stiffness of support shaft
$F_{d,in,x}, F_{d,out,x}, F_{d,in,y}, F_{d,out,y}$	Total contact damping forces between the rollers and both raceways in the x-axis and y-axis directions
P1, P2	Center of the roller
P3, P4	Location of entry and exit points
L	Defect size of the rolling element bearing
θ_{cage}	Angle distance of the cage element between the entry and exit points
f_{cage}, f_{shaft}	Speed of cage and shaft
q	quefreny
z_i	extrema values of the original signal
m_i, a_i	Local mean value and the amplitude envelope
W	Warp path
f_b	Bandwidth parameter of the Morlet wavelet
f_c	Center frequency parameter of the Morlet wavelet
$\mu(d_{ij}^m, n, r)$	Fuzzy function
$FuEn(m, n, r)$	Fuzzy entropy
f_{ri}	i th sample of the r th feature
L_r	Laplacian Score of the r th feature
S_{ij}	Weight matrix
ω	Normal direction of hyperplane

α_i

Lagrange multipliers

Chapter 1 Introduction

1.1 Research background

Rolling element bearings are one of the critical components in rotating machines, where faults such as the spalling and the pitting are commonly observed in the course of normal operation [1]. Therefore, the diagnosis of rotating machinery faults and failure prognostics are necessary to prevent equipment failure or predict the remaining useful life (RUL) to avoid catastrophic failure.

Failure prognostic predicts the RUL of a component with conducting an analysis of the obtained monitoring data. This enables failure prediction in machines, resulting in benefits such as shorter downtimes, higher operation reliability, reduced operations and maintenance cost, and more effective maintenance and logistics planning. An increasing number of bearing prognostics studies have been conducted. However, most existing studies mainly focus on machines operating at a constant speed, which may not be practical considering the machines and associated systems in real applications becoming increasingly complex. As such, a diagnostic and prognostic study on bearings operating at variable speeds has been conducted in this research.

Fatigue failure is a common fault type of rolling element bearing. A defect will be generated by fatigue failure. Spalling size is very small in the early fault stage and is deteriorated gradually until the bearing totally broken, so that spall size can be used as a good fault feature for the RUL prediction. In a number of studies [2, 3], the signature of the vibration signal originating from the passage of a rolling element over the spalled area has been reported as being composed of two main parts. The first originates from the entry of the rolling element into the fault, while the second results from the exit of the rolling element as it strikes the trailing edge of the fault. As the defect size increases,

the separation between the two points, i.e. the time to impact, increases and if the entry and exit events can be successfully extracted from the vibration signal, the size of the fault can be estimated. Sawalha and Randall [4] illustrated more details explanation of the two events by observing the vibration signatures of seeded faults. To improve the estimation performance, Sawalhi et al. [5, 6] proposed other methods to process the vibration signal, such as the autoregressive inverse filtration, synchronous averaging, energy envelopes, and numerical differentiation.

Other signal processing methods have also been proposed to measure the defect size of a bearing under different conditions. Jena et al. [7] measured the different defect sizes in inner race and outer race of the bearing under constant speed. The ridge spectrum method which derived from the Continuous Wavelet Transform was proposed to obtain an obvious indication of the time duration between the entry point and the exit point. Moustafa et al. [8] estimated the different seeded fault widths under low speed with an instantaneous angular speed (IAS) technique. The IAS could effectively reveal shaft speed variation of a bearing with a fault in the outer race when the rolling element passed through the defect area. Khanam et al. [9] detected the different fault sizes in the outer race of ball bearing using the discrete wavelet transform analysis. The entry and exit events were pointed out clearly in the decomposed signal, and a good estimation of the defect size was obtained. Wang et al. [10] proposed a vibration signal processing methodology for extracting the fault size of naturally generated and observing the propagating of bearing fault under high-speed conditions. Entry and exit events from the vibration signal were enhanced by the tacho-less synchronous signal averaging (SSA) and the wavelet transform.

Previous research has focused on estimating the defect size of the bearing under constant speed condition. The signal enhancement processing methods have been based on the assumption of constant speed condition [11]. However, rotating machinery sometimes works under time-varying speed condition. For such working condition, the

amplitude and fault characteristic frequency (FCF) of rolling element bearing vibration signal will be influenced by the time-varying speed [12]. Hence, the envelop analysis and other enhancement techniques which based on the constant speed condition cannot be applied directly.

Thus, it can be a challenging task for rolling element bearings fault size estimation due to the harsh and variable working conditions.

1.2 Review of bearing fault diagnosis and enhancement techniques

In practice, bearings often operate under speed variation conditions. Some representative cases often exist in wind turbines, mining equipment, and rotating machinery during speed-up and ramp-down processes. In such operation conditions, the repetition frequency of bearing transient impulses also varies with time and hence the corresponding signals are non-stationary in nature. As a result, the traditional signal processing approaches (e.g. envelope analysis-based methods [13-16]) developed for signal analysis at a constant rotating speed will lead to spectral smearing and false diagnosis. Thus, some signal processing methods are used to extract the fault component from the original signal. This section will introduce some common signal processing methods on bearing fault diagnosis.

1.2.1 Pre-processing method

A bearing fault signal based on time-varying speed condition consists of impulse fault component, determined component and the random noise component. It is usually advantageous therefore to remove such discrete frequency noise before proceeding with bearing diagnostic analysis.

1.2.1.1 Time synchronous averaging (TSA)

TSA is easily applicable as a byproduct of the order tracking of the signal, always performed in order to manage the speed fluctuation [17]. It is used to separate the impact component produced by a fault bearing from the original vibration signal with noise and resonance components. The conventional TSA divides the sensory signal into a number of segments based on the rotating speed and takes ensemble average for the divided segments. In practice, the averaged signal can be obtained by averaging all signal segments each corresponding to one period of a synchronizing signal. It can be done by follows:

$$y(t) = \frac{1}{M} \sum_{m=0}^{M-1} x(t+mT) \quad (1.1)$$

where y is the averaged signal, $x(t)$ is the original signal, M is the number of segments. McFadden et al. [18] utilized TSA to isolate the vibration of planet gears and estimate the distribution of damage on the inner race of the beating [19]. To improve the performance of TSA, some development methods have proposed. Ha et al. [15] propose autocorrelation-based TSA to diagnose the gears fault in planetary gearboxes of a wind turbine. Siegel et al. [20] provide an enhanced method which combined the TSA with empirical mode decomposition that can determine the health of rolling element bearings.

However, as demonstrated in [21], the application of TSA is not able to delete completely the effect of deterministic components, such as gear coupling and misalignments, when speed is not constant, even if the frequency of the vibration source is known. Thus, this method can not be used on the fault diagnosis of bearing under time-varying speed condition.

1.2.1.2 Linear prediction

Linear prediction is basically a way of obtaining a model of the deterministic (i.e. “predictable”) part of a signal, based on a certain number of samples in the immediate past, and then using this model to predict the next value in the series. The residual (unpredictable) part of the signal is then obtained by subtraction from the actual signal value. Autoregressive (AR) model is a time sequence analysis method whose parameters comprise important information of the system condition, and an accurate AR model can reflect the characteristics of a dynamic system. Sawalhi et al. [2] used the AR model to enhance the fault detection by removing the constant components from vibration signal. However, the AR model is based on the stationary condition. Thus, this method cannot be used on the fault diagnosis of bearing under time-varying speed condition.

1.2.1.3 Cepstrum pre-whitening

Considering mechanical systems such as gears and motors, the idea of cepstrum pre-whitening (CPW) is based on the fact that deterministic excitations related to shaft harmonics and gear meshing are periodic, but not sinusoidal, and therefore produce a spectrum with multiple harmonics of the first excitation frequency. This results in periodicity of the spectrum of the signal, with peaks equally spaced in the frequency domain. In the cepstral domain [22], the periodicity of the spectrum results in a peak at a quefrequency equal to the period of the base frequency of the multi-harmonic vibration. The real cepstrum can be used to edit the log amplitude spectrum of stationary signals and combined with the original phase to achieve edited time signals [23]. This finding has given rise to the development of cepstrum editing methods for the separation of deterministic signal content from stochastic content.

Initially, most of the research focused on developing a cepstrum editing procedure to selectively set certain cepstral peaks belonging to masking discrete frequencies to zero [24, 25]. The idea here is mainly to filter out the deterministic frequencies while preserving the rest of the signal's content. Lately, there has been an increasing usage of a so-called cepstrum pre-whitening method [5, 17]. Instead of setting only a selection of peaks to zero, this method sets the whole real cepstrum to zero, except for the zero frequency. This technique is very easy to implement and has a very low computational cost.

1.2.2 Enhancement techniques

It is difficult to estimate the defect size of a bearing on the outer race. This is because of the impulse component of the signal smeared into the noise signal. For extracting the impact component of the signal, some signal processing techniques can be used to extract the fault information, such as spectral kurtosis, wavelet transform (WT), the empirical mode decomposition (EMD) and local mean decomposition (LMD) for the non-stationary signal.

1.2.2.1 Spectral kurtosis (SK)

Spectral kurtosis is a signal processing method which can extract transient impact component from vibration signal. It provides a method to select the frequency band which contains a component of maximum impulsivity [26]. The spectral kurtosis of a signal can be computed from the short time Fourier transform (STFT). The non-stationary random time series $y(n)$ can be obtained as given in [13, 14]:

$$y(n) = \int_{-1/2}^{1/2} F(n, f) e^{j2\pi fn} dW_x(f) \quad (1.2)$$

where $dW_x(f)$ is the increment of orthonormal spectral, $F(n,f)$ is a complex envelope of the signal for each frequency position.

Thus, the spectral kurtosis can be defined in term of the fourth-order normalized cumulant [14]:

$$K(f) = \frac{E(|F(n, f)|^4)}{\left(E(|F(n, f)|^2)\right)^2} - 2 \quad (1.3)$$

where E is the expectation operator. $K(f)$ is the spectral kurtosis value at the frequency f . The SK map formed a function of frequency and window length and its maximum values gives the optimal central frequency and bandwidth of the band-pass filter. Figure 1.1 shows the general paving of the (frequency/frequency resolution) plane

A further simplification is proposed and is called fast kurtogram. It involves the computation of the kurtosis of the coefficients obtained at the output of filter-banks of quasi-analytic filters with central frequency and bandwidth value. Figure 1-2 shows the general paving of the (frequency/frequency resolution) plane by 1/3-binary tree kurtogram estimator.

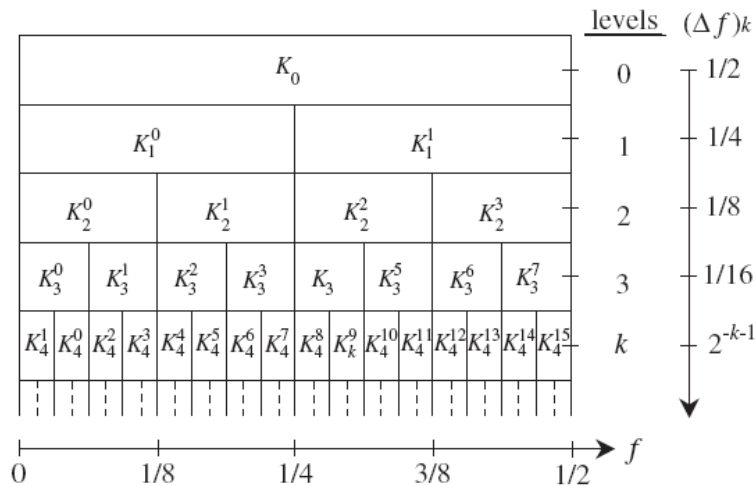


Figure 1-1. Paving of the frequency/frequency resolution plane [13].

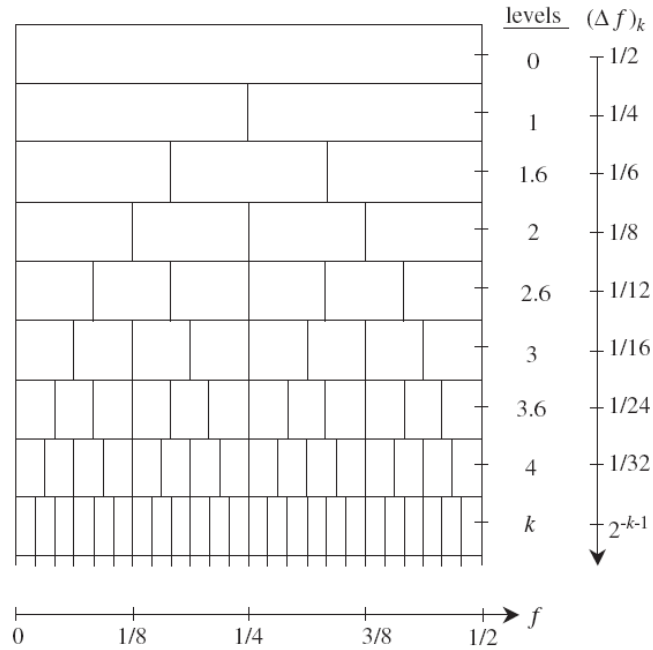


Figure 1-2. Combinations of center frequency and bandwidth for the 1/3-binary tree kurtogram estimator [13].

The fast kurtogram gives a strict way to find the best frequency band automatically to filter the signal. Moreover, it detects in which frequency band transients take place and it returns the complex envelope for each selected band. Kurtogram is calculated based on the STFT or filters. But, limits the performance of the kurtogram in extracting transient impact component under low signal-to-noise ration and non-Gaussian noise. To overcome the above shortcomings of kurtogram and to further enhance its accuracy in discovering characteristics and detecting faults, more precise filters need to be developed and incorporated into the kurtogram algorithm [27].

1.2.2.1 Discrete wavelet transform

The Wavelet transform is classified as continuous and discrete wavelet transforms (DWT). The continuous wavelet transform is calculated by the convolution of the signal and a wavelet function. A wavelet function is a small oscillatory wave, which contains

both the analysis and the window function. However, the discrete wavelet transform uses filter banks to analyze and synthesize a signal. This technique, in fact, provides powerful multi-resolution analysis in both time and frequency domain and thereby becomes a rather useful tool to extract the transitory features of non-stationary vibration signals produced by the faulty bearing. In order to extract the fault feature from the signals more effectively, an appropriate wavelet base function should be selected. The discrete wavelet transform is derived from the discretization of continuous wavelet transform by adopting the dyadic scale and translation to reduce the computational time and can be expressed after [28] by the following equation:

$$DW(j, k) = \frac{1}{\sqrt{2^j}} \int_{-\infty}^{+\infty} s(t) \psi^* \left(\frac{t - 2^j k}{2^j} \right) dt \quad (1.4)$$

where j, k are integers, 2^j and $2^j k$ are the scale and translation parameter, ψ is the mother wavelet, ψ^* is the complex conjugate of the mother wavelet.

Four kinds of wavelets from the wavelet family are shown in Figure 1-3. As the excellent adaptability of the wavelet, wavelet analysis provides a high time resolution for high frequency, at the same time, a high-frequency resolution for low frequency. To show the adaptability of the wavelet, the ‘Symmlet 8’ wavelet at various scales and locations is shown in Figure 1-4.

In particular, the DWT decomposes the signal in its high-scale, low-frequency components, named approximations (A) and its low-scale, high-frequency components, which is called details (D). The decomposition process, with the approximation and detail components, is represented in Figure 1-5. This decomposition process can be iterated, and the successive approximations are decomposed in turn: one signal is broken down into many lower-resolution components to obtain the so-called wavelet

decomposition tree. As it could be imagined, the analysis process is iterative and so, in theory, it can be continued indefinitely. It has been applied widely in rotating machinery fault diagnostics due to the good capabilities for non-stationary signals [29-33]. However, in reality, the decomposition can proceed only until a certain suitable number of levels based on the nature of the signal, or on a suitable criterion such as entropy.

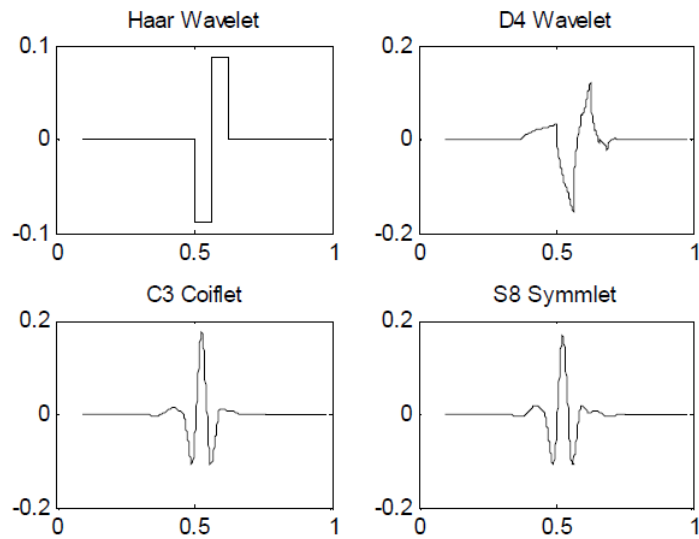


Figure 1-3. Mother wavelets of wavelet family.

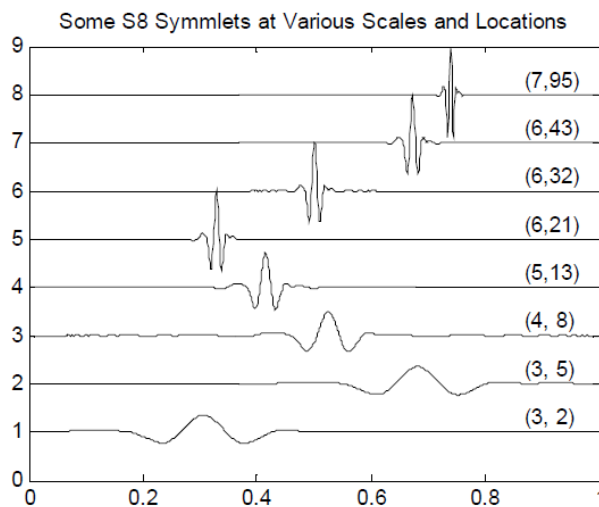


Figure 1-4. 'Symmlet 8' wavelet at various scales and locations.

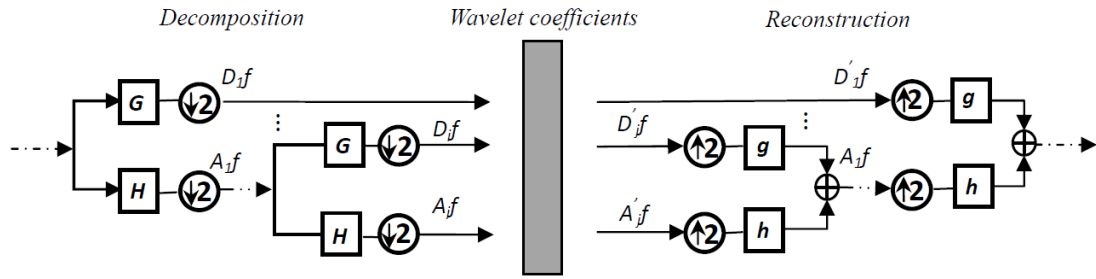


Figure 1-5. A multi-level analysis and synthesis process of DWT [34].

1.2.2.2 Local mean decomposition

LMD is a novel signal processing method which was proposed by the Smith [35]. It is an adaptive signal processing method and has been successfully used on the bearing fault diagnosis [36-39]. It has similar properties with EMD. In contrast to EMD, the LMD method uses the moving average to get the amplitude envelope instead of cubic spline interpolation. Therefore, the overshooting and undershooting effect caused by cubic interpolation can be eliminated [40].

A series of product functions (PFs) can be obtained by the LMD processing method, each of PFs is a component of the original signal which contains multi-component[38]. The amplitude envelope and instantaneous frequency can be obtained via to the PFs, these features are useful to the fault diagnosis of the rolling element bearing. Cheng et al. [36] detected the work condition of rotating machinery with the fault features extracted by the LMD. Sun et al. [41] used the LMD to obtain the envelope spectrum entropy which can effectively identify the leakage in the gas pipeline. Tian et al. [39] combined the LMD and the singular value decomposition (SVD) to recognize the rolling bearing fault under variable conditions and obtained a good result. Li, Y et al. [42] proposed an algorithm to identify the fault patterns of the rolling bearing based on the LMD. Han et al. [37] Successfully classified the fault pattern of the rolling

bearing by the sample entropy and energy ratio features that were extracted by the LMD method.

However, the LMD also suffers from problems as a new signal processing technique and need to be improved in the practical application. The main shortage of the LMD is the end effect [43]. During the LMD procedure, the local mean function and magnitude envelope function by smoothing the local extreme values of the signal. The starting and ending points cannot be identified whether they are extreme points or not. The improved LMD method will be introduced in Chapter 4.

1.3 Time-frequency analysis for fault size estimation

Unfortunately, frequency domain analysis is not able to process non-stationary vibration signals, especially for the time-varying speed condition. Time-frequency analysis could be a good solution for solving this problem because it could investigate waveform signals in both time and frequency domain. There are many time-frequency analysis (TFA) methods have been widely proposed and applied in fault bearing diagnosis, such as short time Fourier transform (STFT), Wigner-Ville distribution (WVD), wavelet analysis. What's important, the digital signal processing techniques need to be carefully selected according to the characteristics of the signals of interest in practical engineering.

1.3.1.1 Short time Fourier transform

STFT is seen as a method that partitions the non-stationary signal into a series of the small segment with a short window, which can be assumed to be locally stationary. Then, the frequency spectrum of each segment can be obtained, and the procedure of STFT is shown in Figure 1-6. Given a time signal $x(t)$, STFT is obtained by

multiplying the signal itself by a window function $\omega(t)$ centered in t :

$$STFT\{x(t)\} = X(\tau, f) = \int_{-\infty}^{+\infty} x(\tau)\omega(\tau-t)e^{-j2\pi f\tau} d\tau \quad (1.5)$$

$$x(t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} X(\tau, f)e^{+j2\pi f\tau} d\tau df \quad (1.6)$$

where $\omega(t)$ is a window function (Hanning or Gaussian window), $x(t)$ is the original signal to be processed, $X(\tau, f)$ is Fourier transform of segment signal with a window function, τ is the time index and f is frequency position. The major disadvantage of the STFT is the resolution tradeoff between time and frequency. Once the window function and its length are determined, the time-frequency resolution is fixed. This resolution is determined by the width of the window function and the relationship between time-frequency resolution and window length can be expressed as,

$$\Delta t = 1/(\Delta f \times N) \quad (1.7)$$

where N is window function length, Δf and Δt are the frequency and time resolutions respectively. So, a good resolution in the frequency domain but the poor resolution in the time domain can be provided by a large window length. On the contrary, a good resolution in the time domain and poor resolution in the frequency domain is achieved by a small window length. This particular limitation is due to the use of a single window for all frequencies that cause the same resolution of analysis at all locations in the time-frequency plane. Thus, the STFT can only be used to process the non-stationary signals with slow change.

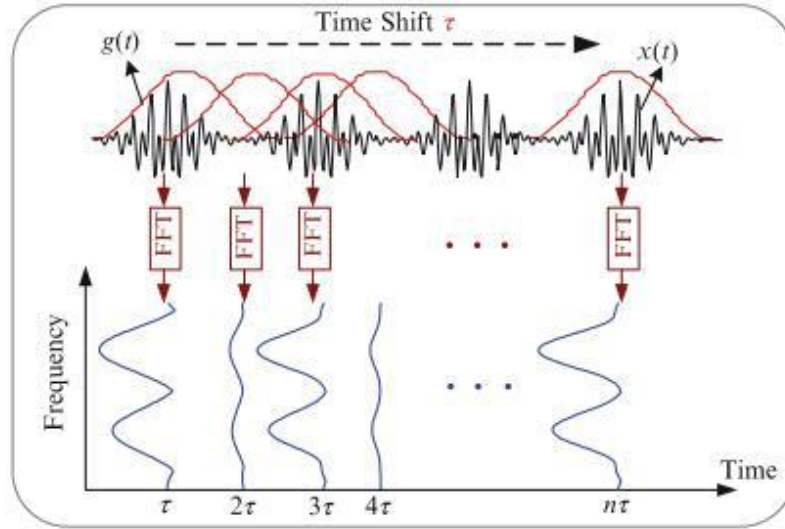


Figure 1-6 Procedure of the STFT [34].

1.3.1.2 Wigner-Ville distribution

A better resolution is given by the Wigner-Ville distribution (WVD) which is able to overcome this spectrogram limitation since it is not based on signal segmentation. Wigner-Ville distribution defined as:

$$WVD(t, f) = \int_{-\infty}^{+\infty} x\left(t + \frac{\tau}{2}\right)x\left(t - \frac{\tau}{2}\right)e^{-j2\pi f\tau} d\tau \quad (1.8)$$

This distribution satisfies many desirable mathematical properties. In particular, the WVD is always real-valued, it preserves time and frequency shifts and satisfies the marginal properties.

As the WVD is a bilinear function of the signal, the quadratic superposition principle applies:

$$W_{x+y}(t, f) = W_x(t, f) + W_y(t, f) + 2\Re\{W_{x,y}(t, f)\} \quad (1.9)$$

where $W_{x,y}(t, f)$ is the cross-WVD of x and y.

$$W_{x,y}(t, f) = \int_{-\infty}^{+\infty} x\left(t + \frac{\tau}{2}\right)y^*\left(t - \frac{\tau}{2}\right)e^{-j2\pi f\tau} d\tau \quad (1.10)$$

Thus, the WVD of two signals is not just the sum of the signal's WVDs, but also of their cross-WVD. Cross-terms can make the WVD difficult to interpret, especially if the components are numerous or close to each other. The more so in the presence of noise, cross-terms between signal components and noise exaggerate the effects of noise and cause rapid degradation of performance as the SNR decrease. Due to these reasons, cross-terms are often regarded as the fundamental limitation on the applicability of quadratic time-frequency methods.

1.3.1.3 Continuous wavelet transform

Continuous Wavelet Transform (CWT) uses basis functions with translation and dilation parameters instead of simple complex exponentials unlike Fourier transform [32]. It adds a dilation parameter along with a time translation parameter to the analysis. Hence, it is best suited for time-frequency analysis, and effective in analyzing non-stationary signals like run-up vibration data of the rotor. CWT of a time signal $x(t)$ is obtained by solving the following convolution integral:

$$W_c(a, b) = \int_{-\infty}^{+\infty} x(t) \bar{\psi}_{a,b}(t) dt \quad (1.11)$$

where $x(t)$ is the signal, a is the scale, b is the translation, $\psi_{a,b}(t)$ is the mother wavelet:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (1.12)$$

The CWT is certainly one of the best transforms for singularity detection. Impact faults could be detected by finding the singularity in the signal. So, its identification is based on the use of the local maxima lines through finding the abscissa, where the wavelet modulus maxima converge at fine scales. The ‘‘continuity’’ of the method stands in the scales at which it operates. It could be processed for each scale possible in the shifting because the analyzing wavelet is shifted smoothly over the full domain of the

analyzed function. Thus, the CWT is used in this research work.

1.4 Research purposes and methods

This research is aimed to develop a new bearing fault diagnosis system under time-varying speed conditions. Bearing fault diagnosis contains three parts: fault detection, fault identification and fault prediction. This research will put the focus on solving them.

We proposed a new method named EC-LMD to extract the fault component and improved the detection result. The defect size is used as a feature to predict the RUL. However, based on the description of the existing research methods, we can conclude that the most suitable methods to estimate the defect size of rolling element bearing is under constant speed condition. This research proposed a new method to measure the defect size of a defective bearing under time-vary speed condition. The signal is transformed from time domain to angle domain by resampling method. The angle duration when the roller passes over the defect area can be measured based on the energy distribution. proposes an intelligent rolling bearing fault diagnosis method. EC-LMD was used to pre-process the signal for extracting good features. The feature extraction was done by the MFE. MFE could effectively characterize the complexity of the pre-processed signal. Laplacian score was used to select the fault feature by reordering the scale factors. SVM is used to evaluate the classification performance.

To validate the proposed method, some experiments were performed. The experimental results illustrate that the proposed methods obtained good performance.

1.5 Main contents and organization

The remainder of this research is organized as follows.

Chapter 2 studies some critical parameters of bearing fault diagnosis and fault size

estimation. Fault diagnosis method is introduced detailly. To measure the defect size, the entry and exit events are illustrated. Furthermore, the effects of speed variation on fault diagnosis and defect size estimation of a defective bearing are elaborated.

Chapter 3 a dynamic vibration model of a defective bearing is established. The changes of contact force and deformation of the roller when it passes a defect area are described by this dynamic model. It can help us to easily understand the vibration response when the roller travels the defect area. An estimation model is illustrated for measuring the defect size of a defect bearing under time-varying speed conditions.

Chapter 4 proposes a new fault diagnosis method of a rolling element bearing under time-varying speed conditions. The proposed method can remove the determining component of the vibration signal and extract the impact component effectively. An experiment is used to verify the effectiveness of the proposed method.

Chapter 5 proposes a time-frequency analysis method to estimate the defect size of a defect bearing. The entry and exit points can be identified exactly based on the energy distribution plot. The proposed method was validated experimentally.

Chapter 6 proposes an intelligent rolling bearing fault diagnosis method. EC-LMD was used to pre-process the signal for extracting good features. The feature extraction was done by the MFE. MFE could effectively characterize the complexity of the pre-processed signal. Laplacian score was used to select the fault feature by reordering the scale factors. SVM is used to evaluate the classification performance. The experimental results showed that the different categories of rolling bearings can be effectively identified by the proposed method.

Chapter 7 concludes.

Chapter 2 Effects of speed variation on fault diagnosis and size estimation

2.1 Introduction

Fault diagnosis and failure prognostic predict the RUL of a component with conducting an analysis of the obtained monitoring data. Theories of fatigue life prediction have been developed to estimate the remaining useful life of bearings [44-46]. Among these theories, the geometry of the bearing and other factors, such as material properties, lubricant, temperature, load, and speed, are considered for the estimation of the fatigue life probability in bearings. Some of them were developed based on discrete finite element methods [27, 46, 47] and some others account for the debris of a bearing [45]. The purpose of these bearing failure principles is to predict the bearing life before the first indication of fatigue or illustrating the fatigue with a severity level. The different working condition, bearing life criterion may be different. Therefore, there is no single failure standard is applicable for all conditions (the bearing is operating under different conditions). However, defect size of bearing can be a good fault feature for determining the severity of defect in bearings at the early stages to predict the RUL. In this chapter, how to measure the fault size is illustrated.

2.2 Bearing degradation with contact fatigue

Generalized roughness, brinelling, and single-point faults (spalls) are generally two kinds of bearing faults [48]. Contact fatigue, which usually leads to spall faults, accounts for a large percentage of bearing failure [49-52]. Ferreira et al. [49] conducted a reliability analysis for rail bearings. About 47,000 bearings under real work conditions were investigated and found that around 80% of the bearings failed caused by contact

fatigue: 58% is outer-race fatigue, 13% is inner race fatigue, 3% is roller fatigue and 4% is micro fatigue. Generally, fatigue failure is characterized when the alternating shear stress reaches a maximum value [50]. The flaking phenomenon will occur in the raceways or the rolling elements caused by the merger of the microcracks initiated under the repeated contact stress condition [50].

Bearing degradation caused by contact fatigue typically progresses in three stages: crack nucleation and propagation, spalling and fast deterioration.

Phase 1. In the crack nucleation and propagation stage, subsurface cracks grow in the races without any visible modification of the contact surfaces. Once the cracked is loaded repeatedly, the subsurface crack will propagate along the bearing races. By theoretical analysis and experimental observations, Chen et al. [53] concluded that subsurface cracks initiate which caused by the combined action of the inclusion local stress field and the maximum reversed shear stress. The shortest visible crack length is 10 - 15 μm from the observations, and the nucleation period for a crack is only about 10% of the bearing's useful life. The illustrations showed that subsurface cracks, initiated at the same time from both tips of the inclusions in the steel and grow in opposite directions. Deng et al. [54] illustrated that the direction of subsurface cracks are approximately parallel to the spalling surface along the axial direction.

Phase 2. The second stage is spalling, which happens when the material removal from the races as a result of repeated stresses. During this stage, the spall propagating speed is faster. As shown in many studies [55-59], the bearing fault can be detected at this degradation stage. It is the best choice to take action to protect the machine at this stage.

Phase 3. After the spall is developed in the bearing, the component deteriorates rapidly to complete failure. A sudden increase in system vibration can be observed, and damage to other components in the system may result if the operation is not terminated.

Figure 2-1 shows the growth of the defect.

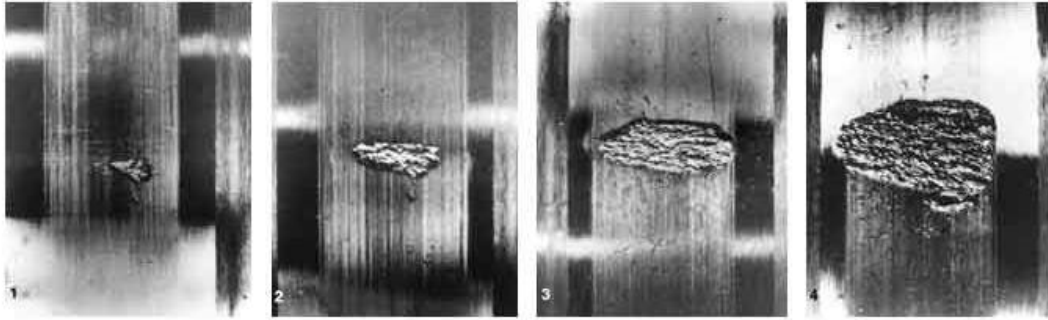


Figure 2-1. progression of the damage across the raceway [60]

In practice, the bearing degradation stage as discussed in the previous section (crack nucleation and propagation), when the crack grows under surface, it is difficult to diagnose the bearing fault due to the mechanism of vibration measurement. At Phase 2, the degradation stage (spalling), the impacts of each rolling element with the spall generate periodically vibration bursts. Each impulse signal is produced at a defect fault frequency which is calculated by the geometry of the bearing, the rotating speed and fault position, i.e. the characteristic fault frequencies of bearing (BPFO, BPFI) [48] as the Figure 2-2 shown. The four characteristic fault frequencies of a ball bearing can be expressed by the following equations:

$$BPFO = \frac{n}{2} f_r \left(1 - \frac{d}{D} \cos \alpha \right) \quad (2.1)$$

$$BPFI = \frac{n}{2} f_r \left(1 + \frac{d}{D} \cos \alpha \right) \quad (2.2)$$

$$BSF = \frac{D}{2d} f_r \left[1 - \left(\frac{d}{D} \cos \alpha \right)^2 \right] \quad (2.3)$$

$$FTF = \frac{1}{2} f_r \left(1 - \frac{d}{D} \cos \alpha \right) \quad (2.4)$$

where

BPFO- ball pass frequency of outer ring (Hz),

$BPFI$ - ball pass frequency of inner ring (Hz),

BSF - ball spin frequency (Hz),

FTF - fundamental train frequency (Hz),

d - diameter of the rolling element (mm),

D - pitch diameter of bearing (mm),

n - number of rolling elements,

f_r - shaft speed in frequency (Hz),

α - contact angle.

For the time-varying speed condition, the characteristic fault frequencies of the rolling element bearing are also time-varying based on the equation 2.1-2.4. To solve this problem, the order tracking method is necessary to be used. The order tracking method will be introduced in the later section.

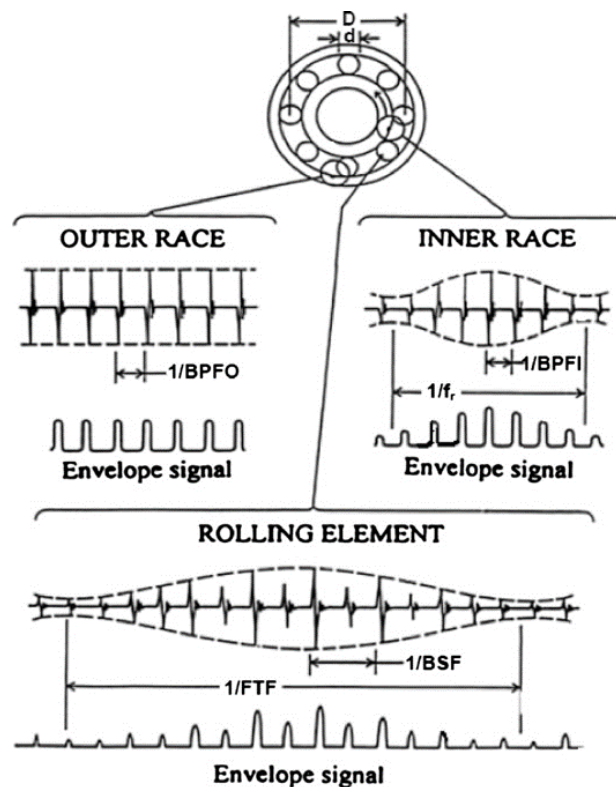


Figure 2-2. Typical signals and envelope signals of local faults in a rolling element bearing [61].

2.2.1 Response of the fault bearing

A periodic impulse signal was produced when the bearings existed a spalling fault. This periodic impulse signal is not strictly periodic (i.e. almost-periodic) due to randomly varying slip in the contact between the rollers and races. The varying slip is caused inevitably by the variation of the load angle, which determines the effective rolling radius, and the position of each rolling element in the bearing [61]. Common bearing models are established by taking natural vibration bursts as a (pseudo-) random carrier with a power spectral density characterized by the resonance frequencies of the machine.

Many developed theoretical models of fault bearing were proposed with considering different parameters or conditions that could affect the vibration response of the faults. McFadden et al. [62, 63] described the vibration signals produced by a single point fault on the inner race of a rolling element bearing. The response of the bearing and machine to the impulses produced by the defect is described as:

$$x(t) = [d(t)q(t)a(t)] * m(t) \quad (2.5)$$

where $d(t)$ is the impulse component produced by rolling elements hitting the defect area, $q(t)$ is the variations of load distribution experienced caused by the spall around the bearing, $a(t)$ is the variations of the transfer function between the measurement location and the point at spall, and $m(t)$ is a real function expressing the decay of a unit impulse. Therefore, the key point of the bearing fault diagnosis is to extract the fault component from the vibration signal. In this research, we propose a new method to extract the fault component effectively.

2.3 Bearing fault diagnosis

2.3.1 Fault detection

For bearings running under a known constant speed and geometry, the fault characteristic frequencies can be calculated by the equations as described in Section 2.2. The fault component may not produce an obvious change in the full vibration signal, and this may be masked by non-fault related vibrations signal which generated higher energy. However, it also can generate a change in a frequency band in which the non-fault related vibration is very small.

Envelope analysis has been widely utilized in bearing fault diagnostics. It can effectively detect the impulse signal from the vibration signals of the defective bearing. Compared to the intervals between each impulse signal, the time duration of one impulse signal is very short and its energy is distributed across a very wide frequency range. Consequently, several resonances of the fault bearing and the surrounding structure are excited by these impacts. This impulsive excitation of a bearing is normally repetitive. Hence, the energy is concentrated in a narrow band, which makes it much easier to detect than that of the dispersed energy. To obtain good results with envelope analysis method, a frequency band in which the impulsive components excited by the defect need to be selected. Then, the envelope of a filtered signal $x(t)$ is obtained by calculating the amplitude of the complex signal which constituted by the real part of $x(t)$ and its Hilbert transform as the imaginary part. The Hilbert transform of a time domain signal can be expressed by following:

$$\text{hilbert}[x(t)] = \tilde{x}(t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x(\tau)}{t - \tau} d\tau \quad (2.6)$$

Then, the envelope of the filtered signal is obtained:

$$\text{envelop}[x(t)] = \sqrt{x^2(t) + \tilde{x}^2(t)} \quad (2.7)$$

Finally, the envelope spectrum can be obtained with a spectrum analysis of the squared envelope signal. Thus, the characteristic fault frequencies may be presented in envelope spectrum.

To better understand the envelope analysis, an analytical investigation of the envelope analysis is established. The model can be described by taking into account the vibration signal of a defective bearing, after passing through the high-pass filter. Then, the vibration at some ringing frequency has the form of a high-frequency carrier (the system natural frequency) and a low-frequency modulation (the defect frequency, each time a defect is encountered):

$$x = A_1(1 + A_2 \cos(\omega_1 t)) \cos(\omega_2 t) \quad (2.8)$$

Where A_1 and A_2 are constants, ω_1 is the fault frequency of the system and ω_2 is the natural frequency of the system. When the signal is enveloped, the squared signal would be:

$$x^2 = A_1^2(1 + A_2^2 \cos^2(\omega_1 t) + 2A_2 \cos(\omega_1 t)) \cos^2(\omega_2 t) \quad (2.9)$$

Then, expanding the equation (2.9):

$$\begin{aligned} \frac{x^2}{A_1^2} &= A_2 \cos(\omega_1 t) - \frac{A_2^2}{4} \cos(2\omega_1 t) + \frac{2A_2 + 1}{4} \cos(\omega_2 t) \\ &- \frac{A_2}{2} \cos((2\omega_2 - \omega_1)t) - \frac{A_2}{2} \cos((2\omega_2 + \omega_1)t) - \frac{A_2^2}{8} \cos((2\omega_1 - 2\omega_2)t) \\ &- \frac{A_2^2}{8} \cos((2\omega_1 + 2\omega_2)t) + \frac{2A_2 + 1}{4} \end{aligned} \quad (2.10)$$

From the equation (2.10), the fault frequency ω_1 can be found as a separate frequency component with A_1 . The spectrum of the squared envelope signal detects the obvious fault frequency ω_2 of the modulation. However, the envelope analysis technique is suitable for constant speed conditions. It can not be used on the time-varying speed conditions directly.

2.3.2 Bearing fault severity analysis

When a spall propagated at the beginning of bearing failure, fundamental fault characteristic frequency and its harmonics (multiple integers of the fault characteristic frequency) will occur at the spectrum. An effective condition monitoring system should be established to detect and evaluate the severity or measure the spall size in bearings at the early stages. By this way, remedial action can be taken in time or a schedule of the replacement can be made at a convenient time. Normally, the severity of the fault size in rolling element bearings is associated to the level of the transient vibration, when each rolling element hits the spall at different locations within the load zone. Thus, the remaining useful life can be determined based on the defect size in bearing. On the hand, the existence harmonics of the fault characteristic frequency can also indicate the condition and propagating of a spall in a bearing [64].

More studies have been done on the relationship between fault severity and defect size. Igarashi and Hamada exhibited that: with the size of the defect changing, the width and shape of pulses in time and frequency domain of the vibration signals will change [65, 66]. With the fault developed, the energy will be redistributed in the frequency domain, and the magnitude of the characteristic frequency components also will increase [67]. To show the relationship between the magnitude of the vibration signal and defect progress, some studies of run-to-failure on bearings have been done [68-70]. In these studies, the trend of vibration and some parameters, such as Root-Mean-Square (RMS) and kurtosis values are detected. The results of these studies show that the RMS of vibration signals exhibits an overall increasing trend to the end of life of the bearing. But the RMS value fluctuates at some stages. The fluctuation of the averaged RMS values of the amplitude of the defect bearing in the frequency spectrum with the growth of defect size in the surface of raceway [69] is shown in Figure 2-2. The fluctuation of RMS value is caused by a phenomenon known as 'healing'. When the defect occurred

in the bearing, there are sharp edges of the defect area. After repeatedly hitting, the sharp edges are smoothed. This phenomenon suppresses the increase of the signal level because of the fault growth. However, as the fault grows deeply, the general vibration level increases again. Therefore, the RMS value cannot be used as a fault feature to detect the trend of the defect development reliably. This is because of the bearing might be at the stage of the defect size has grown while the RMS trend exhibits a low level.

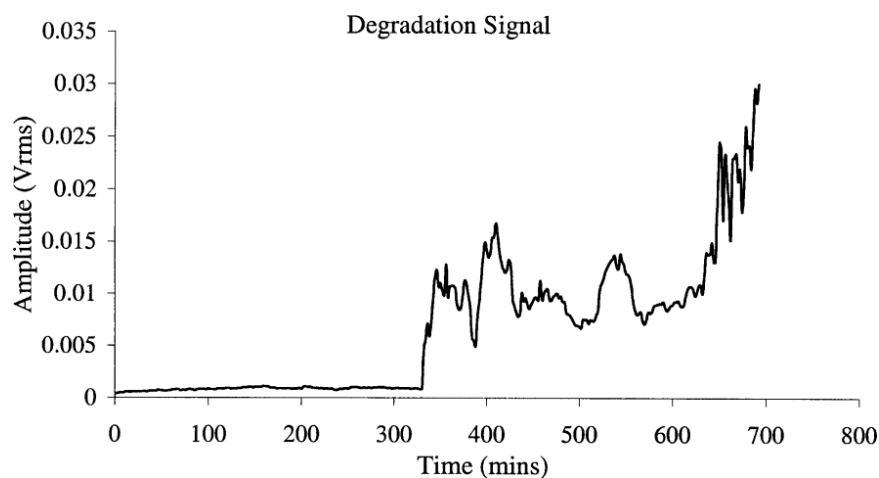


Figure 2-3. Three Evolution of the averaged amplitude of the defective frequency [69].

Williams et al. [71] exhibited run-to-failure tests of bearings by monitoring the trend of the kurtosis value both of vibration and acoustic signal. When the rolling elements pass over the line spall fault, it generates impulse components, and this results in high kurtosis values. The trend kurtosis value oscillates with time as the defect size grows. However, when the defect size raises beyond the spacing of the rollers, the distribution of the signal will like the same distribution to the normal condition with the kurtosis value decreases. The phenomenon of the kurtosis value reduction could also be explained due to the gradually worn sharp edges of the fault.

Thus, although the kurtosis value can be a good fault feature for detecting impact component of the fault bearing, it cannot be utilized to detect non-impact type bearing faults as the kurtosis value fluctuates with the size of the defect. As another possibility

to study severity measures, to indirectly measure the defect size, the actual geometric arc length of a defect in the bearing can be determined from the vibration signal and used for predicting the remaining useful lifetime. This thesis aims to consider the characteristics of the vibration signal in defective bearings and propose a novel defect size estimation method to diagnose the bearing fault. The following section contains an overview of defect size estimation methods.

2.4 Estimation of the defect size

Studies of defect size estimation of defective bearing with a line spall defect illustrate that when roller passes over the defect area, two main vibration characteristics events: entry and exit events [4] are used to measure the defect size. Most fault size measurement methods proposed previously are based on measuring the time duration between these two events. The characteristic of the entry and exit events is introduced in this section.

Sawalhi and Randall [72] illustrated the characteristics of the measured vibration signal in defective bearings by the experimental testing on rolling element bearings with line spall defects. The vibration response of a defective bearing when the roller enters the defect area is shown in Figure 2-4. When the roller enters the defect leaf end, the “destressing” process starts and then the amplitude decreases to the local minimum point C. After passing through point C, the amplitude begins to increase and reaches the first peak at point A. Then, the signal oscillates at a frequency until the beginning of a high-frequency response at point B. Sawalhi and Randall [5] supposed that the entry point was identified at point A which is the local maxima of the low-frequency zone between point C and B, while the exit point was identified at point B which is the largest high-frequency responses in magnitude in the multiple impact zone shown Figure 2-4 (b). The entry event is expressed with a low-frequency response vibration

signal when a rolling element rolls into the line spall defect zone [3, 73]. When the rolling element departs the defect end, the exit event is expressed as a wide range of frequencies, including the bearing resonance harmonic frequencies. The frequencies of bearing resonance are excited when the rolling element hits the exit point of a defect, in addition to the bearing stiffness changes rapidly, which caused by the rolling element re-stresses between the raceways when the rolling element totally leaves the defect zone [74]. Some studies [3, 75] illustrated that the exit event detected in experimental results often occurs with multiple impacts instead of one impact. Singh et al. [76] obtained the same conclusions by the simulation results of the defective bearings. Simulation results exhibit that multiple impulse components generate when the roller strikes the raceway successively as it is re-loaded after the exit point. If the defect size is small and the roller doesn't travel on the defect surface, the impulse component of the vibration response caused when the center of a rolling element is the middle position of the defect size [4]. Zhao et al. [77] used the harmonic product spectrum to recover the fault impulse component. The defect size is estimated based on the double impact phenomenon as Figure 2-5 shown. However, the maximum error of the estimation result is 19.5%. This is caused as the first impact is not the entry point.

In summary, the purpose of these studies is to measure the defect size of a defective bearing by identifying point A (local maxima of the low frequency zone) and point B (maximal amplitude of the high frequency event) and doubling the distance between points A and B. Moazen et al. [75] proposed a nonlinear dynamic model of a defective bearing to identify the entry and exit events detailly. The path of rolling elements in the defect zone is illustrated by the analysis results. The results show that the point C (shown in Figure 2-4 (b)) is the actual entry point and the exit point is among the high-frequency impact zone. However, the previous research works estimate the defect size under constant speed conditions. This research proposes a new method to estimate the defect size under time-varying speed conditions.

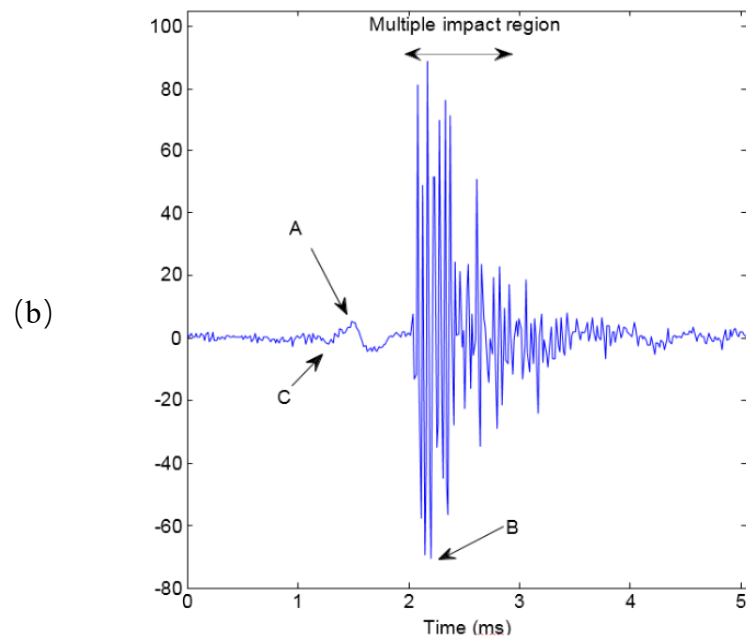
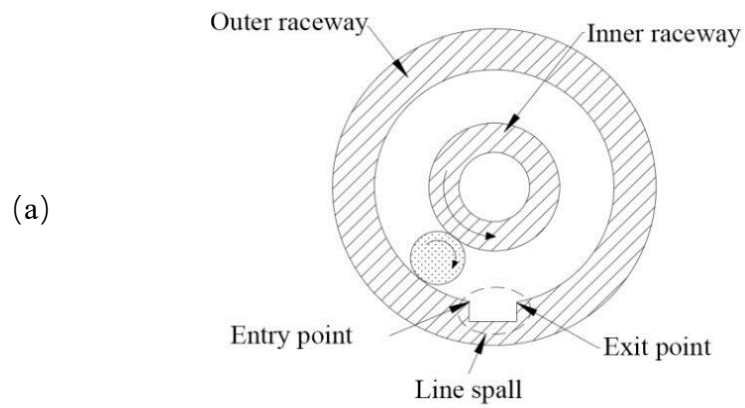


Figure 2-4 *Vibration signal of a defect bearing. (a) Diagram of a rolling element traveling into a line spall defect located on the outer raceway. (b) Typically measured vibration response of a line spall defect [73].*

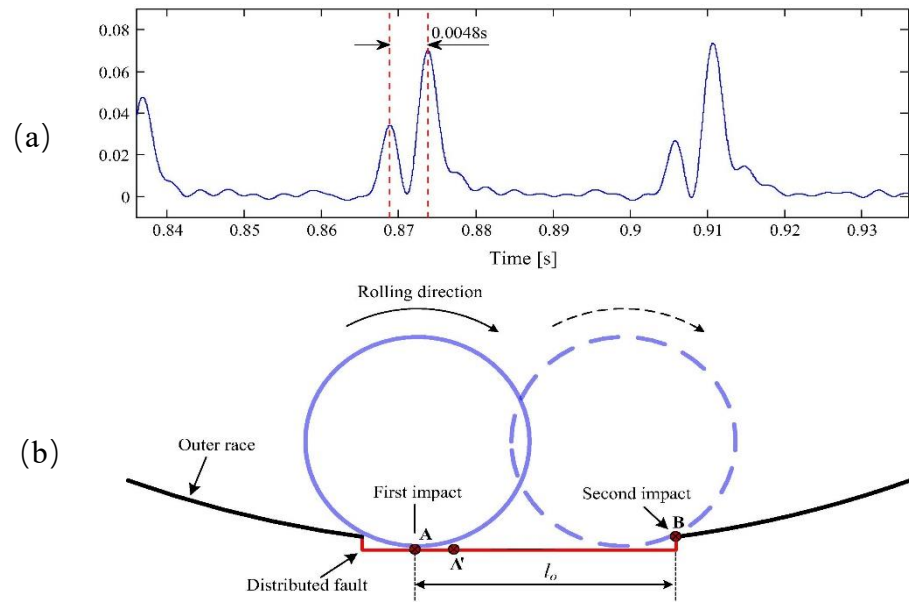


Figure 2-5 *A illustration of defect size estimation. (a) Envelope signal of defective fault. (b) Illustration of the entry and exit events [77].*

2.5 Influences of the rotating speed on the vibration signal

Vibration analysis techniques which frequency based are effective only under the constant speed conditions. However, most of the machinery work at different kinds of condition. Such as start-up or coast-down, load changing. The characteristics of vibration signal under time-varying conditions are usually with time-varying frequency and amplitude. As a result, the fault characteristic frequencies of the defect bearing are generally time-varying. For such conditions, frequency smearing would happen unavoidably when the frequency spectrum analysis is used directly to vibration signal, and it will bring undesirable effect to detect the fault of bearings. It is inappropriate to apply most traditional fault diagnosis techniques because these techniques are based on an assumption that the signal is stationary or quasi-stationary, for which the speed is constant [78].

2.5.1 Spectral analysis under time-varying conditions

The spectral analysis is a most widely used method to detect a fault of the rolling element bearing. The fault characteristic frequencies of a fault bearing can be detected from the spectrum [79]. When the machine is running at a constant speed, the fault characteristic frequencies of a fault bearing are constant. But, when the machine is running under time-varying speed conditions, the fault characteristic frequencies of a fault bearing also change with the rotating speed. As a result, the smearing phenomenon occurred in the spectrum. Furthermore, the fault characteristic frequencies can hardly be detected. Figure 2-6 illustrates the influence of rotating speed on the vibration spectrum. It can be seen obviously that the characteristic fault frequency varied with the rotating speed. At faster speeds, vibration amplitudes are higher, and time intervals of impulses are shorter. In contrast, low amplitude and longtime interval occur at slow speeds. The fault type or fault severity cannot be identified. For time-varying speed conditions, the corresponding characteristic fault orders of the bearing can be obtained by rewriting the equations (2.1-2.4) in the angular domain as follows:

$$O_{BPI} = \frac{n}{2} \left(1 + \frac{d}{D} \cos \alpha \right) \quad (2.11)$$

$$O_{BPO} = \frac{n}{2} \left(1 - \frac{d}{D} \cos \alpha \right) \quad (2.12)$$

$$O_{BS} = \frac{D}{d} \left(1 - \left(\frac{d}{D} \cos \alpha \right)^2 \right) \quad (2.13)$$

$$O_{FT} = \frac{1}{2} \left(1 - \frac{d}{D} \cos \alpha \right) \quad (2.14)$$

where

O_{BPO} - ball pass order of outer ring,

O_{BPI} - ball pass order of inner ring,

O_{BS} - ball pass order,

O_{FT} - fundamental train order.

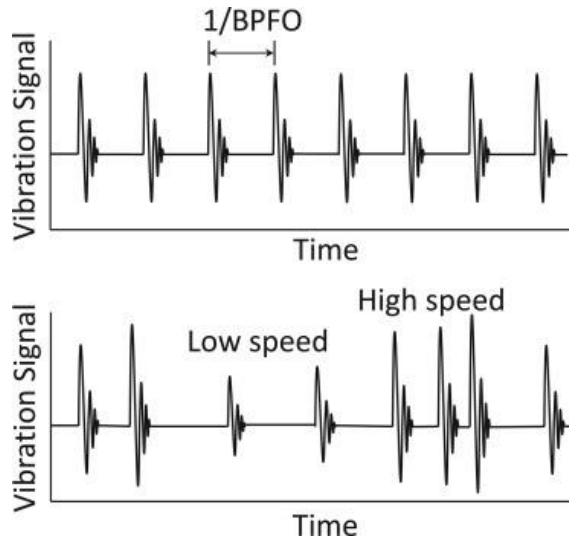


Figure 2-6. Schematic representation of ideal vibration signals in constant (top) and variable speed (bottom) operation with bearing outer raceway fault

2.5.2 Parameters of the vibration signal under time-varying conditions

For the bearing fault detection, it is a simple way to diagnose the bearing by some fault parameters from the vibration signal. Some statistical parameters, such as the peak-to-peak value, root-mean-square (RMS), skewness, kurtosis, and some other mean value-based indices have been used to detect the bearing fault and predict the remaining useful lifetime of a bearing. However, these statistical parameters cannot be used as a fault feature to detect the trend of the defect development reliably under time-varying conditions. This is because of the vibration levels of a bearing will change with the rotating speed (the higher speed, vibration levels are higher. In contrast, lower levels occur at lower speeds). This indicates such statistical parameters can not be stable for a health condition as their values change with the rotating speed of machinery. Therefore, those parameters cannot be used as reliable fault features to detect the bearing fault and predict RUL of rotating machinery under time-varying speed conditions.

2.5.3 *Transmission path under time-varying conditions*

The vibration signals are obtained by a vibration transducer, which is mounted on the machinery cover. The vibration signals are affected by the transfer function of the components from the position vibration source to the vibration sensor. If the transfer function is a linear system, the frequency response of the measured vibration signals can be express by the fault excitation $x_{fault}(t)$ and its vibration response $x(t)$ as follows:

$$X(f) = H_{fault}(f)H(f) \quad (2.15)$$

where $H_{fault}(f)$ and $X(f)$ are the Fourier transforms of $x_{fault}(t)$ and $x(t)$, respectively. $H(f)$ represents the frequency response function (FRF) of the entire transmission path including the effects from shafts, bearings and casing. From the equation (2.15), it can be found that the measured vibration signal is amplitude and phase modulated through the transmission path. Furthermore, the frequency response function of the transmission path is affected by the frequency. The effect of the modulation is constant when the bearing under constant speed and it can be ignored in the diagnostic procedure. However, for the time-vary speed conditions, the distortion in both of the amplitude and phase of vibration signals will occur as the effect of modulation. Thus, the effect of the transmission path should be taken into consideration. In summary, it can be known that the vibration signals of the bearings under time-varying speed conditions are complicated and non-stationary. The vibration signals are not only frequency-modulated by the varying speed, but also amplitude-modulated and phase-modulated due to the transmission effect. Therefore, feature extraction and fault diagnosis are challenging work under time-varying conditions.

2.6 Order tracking

Techniques have been proposed to process non-stationary vibration signals [80-84]. Order tracking (OT) is widely used among those techniques, which can eliminate the effect of speed fluctuation [85]. The main idea of OT is to transform the time-varying speed vibration signal in the time domain to the angular domain with constant angle interval sampling. Thus, the effect caused by the time-varying speed is removed by resampling the signal into the angular domain. By this way, the resampled signal is not dependent on the rotating speed. As a result, the order spectrum is defined by the spectrum with respect to the angular domain signal.

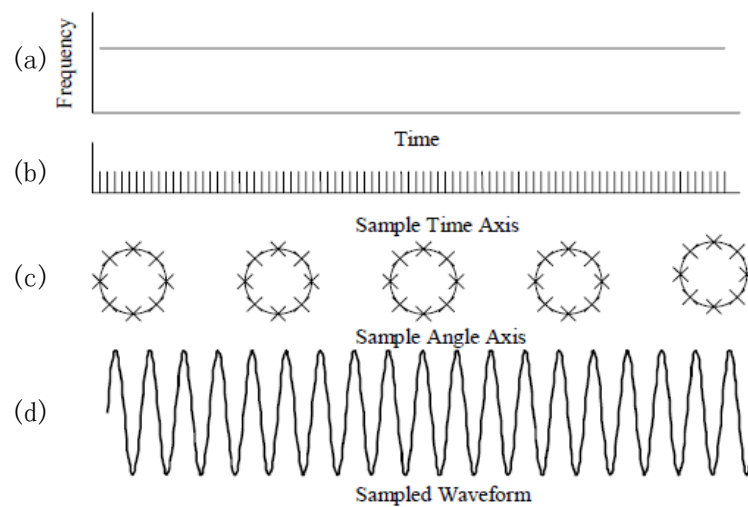


Figure 2-7. Sampling signal under constant speed condition. (a) Speed file, (b) Sampling point in the time domain, (c) Sampling point at the angular domain, (d) Sampled signal.

Figure 2-7 shows a signal sampled under constant speed condition. Figure 2-7 (a) is the speed information. The signal is obtained by the sensors with a constant sampling rate value. Thus, the time interval between each sample is equal. The Figure 2-7 (c) show the marks where on the shaft's axis. It can be seen the sample spacing relates to the shaft's angle. The sampled waveform is shown in Figure 2-7 (d). For the constant

speed condition, the constant time interval in the time domain leads to a constant angular interval in the angle domain. It's no different of the sampling signal with the time domain or angle domain. The spectrum can be obtained by doing the Fourier transform to the sampled signal.

Figure 2-8 shows a sampled signal under time-varying condition. As the sampling rate is constant, the sampling time interval is also constant. However, the angular position of each sample changes with the shaft frequency. The sampled waveform is shown in Figure 2-8 (d). As the speed increases, the spectrum of the signal will not hold in one spectral line in the frequency domain. As a result, the smearing problem is generated. Further, it will lead to smeared fault characteristic frequency and make the fault harder to detect.

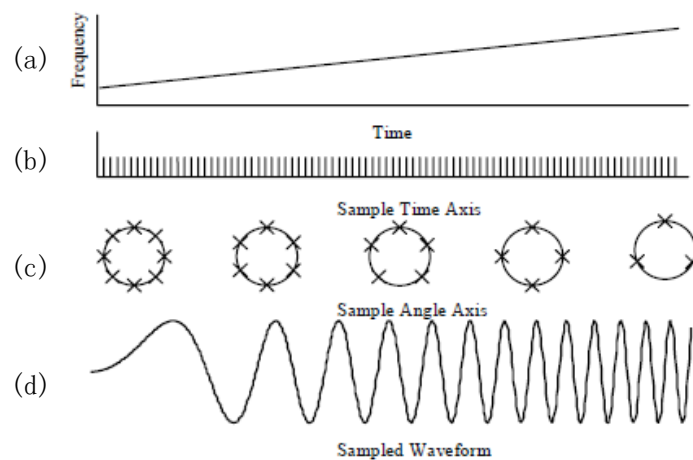


Figure 2-8. Sampling signal under time-varying speed condition with a constant time interval. (a) Speed file, (b) Sampling point in time domain, (c) Sampling point at the angular domain, (d) Sampling signal.

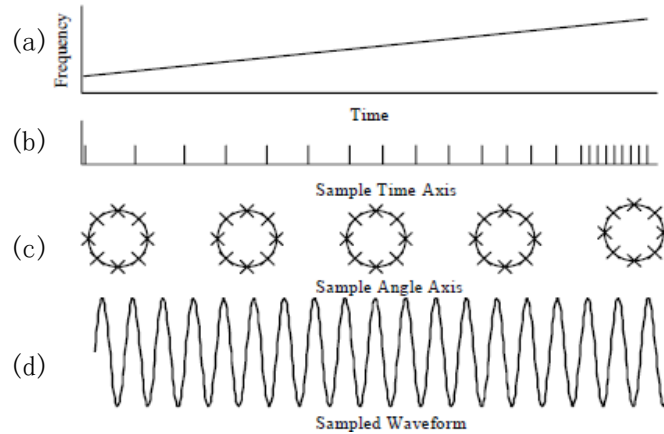


Figure 2-9. Sampling signal under time-varying speed condition with constant angle interval. (a) Speed file, (b) Sampling point in time domain, (c) Sampling point at the angular domain, (d) Sampling signal.

Figure 2-9 shows the resampled signal in angle domain with a constant angular interval. For the resampling procedure, the speed information is obtained by a tachometer. Then the signal can be resampled from the time domain to the angle domain. Once the signal was resampled to the angle domain, the order spectrum is obtained by doing the FFT to the resampled signal.

The main steps of the resampling process include [86-88]:

(1) Synchronous acquisition of the vibration and key phasor signal (constant-time increment).

(2) Obtaining the speed of the shaft from the keyphasor signal, and the total phase can be calculated.

(3) Setting the resampling rate based on the maximum value of shaft speed, and then obtaining the even-angle increment and corresponding sampling time.

(4) Interpolating the vibration signal according to the even-angle increment.

Provided that the shaft rotating angle of the roller passing over the defect area is obtained, the defect size can be estimated. Therefore, the defect width can be calculated

from the signal in the angle domain.

2.7 Conclusions

This chapter introduces the background of bearing diagnostics including different bearing failure modes, modeling of bearing vibration due to a localized spall defect, signal processing methods and damage severity estimation. The effects of speed variation on bearing fault diagnosis are illustrated. Thus, the scope of this research is to localize bearing defect analysis under time-varying speed conditions. To get a better understanding of the defective signature, a dynamic model is established in the next chapter.

Chapter 3 Dynamic vibration model of a defective bearing

3.1 Introduction

To estimate the defect size of a line spall defect, it is important to identify the entry and exit events accurately from the vibration signal. This chapter presents a dynamic vibration model of defective bearing to illustrate the relationship between the contact force variation and the entry and exit events. We can conduct the vibration response features of the entry and exit events by the simulation dynamic mode. The relationship between these features and the defect size is introduced with the bearing system dynamic model. An improved defect size estimation model of a defective bearing under the time-varying speed conditions is proposed.

3.2 Bearing dynamic model

As chapter 2 introduced, previous studies have suggested that once a roller passes over defect area on the raceways, the entry and exit events were identified based on the vibration response. The time duration between them is taken as a parameter to measure defect size based on the constant speed condition. In order to provide adequate evidence to identify the entry and exit points, a dynamic model of a defective bearing is established. By this way, it's easier to understand the vibration response of a defective bearing. Firstly, a kinematic analysis is used as the foundation of the bearing dynamic model. Then, the Hertzian contact theory is applied to derive the nonlinear system equations, which are solved numerically. Finally, the simulation result is discussed, and the defective pattern is summarized.

A nonlinear dynamic model of a defective bearing is used (shown in Figure 3-1) to illustrate the vibration response of a defective rolling element bearing. The model includes the masses of the outer raceway attached to the support housing (M_o), inner raceways with the shaft (M_i), and rolling elements (m). The static load F_l applies on the shaft in the y-direction.

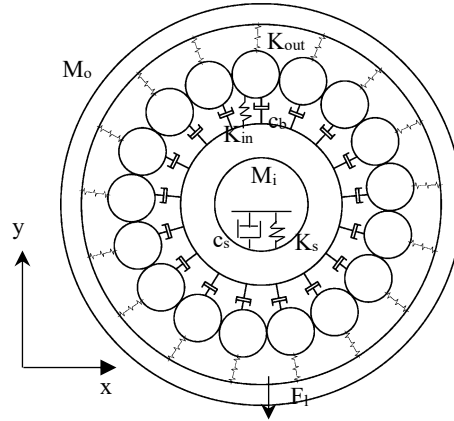


Figure 3-1. Diagram of bearing dynamic model.

3.2.1 Kinematics of the roller

Assuming that there is no slippage between the shaft and the two inner raceways, the rotating speed of the cage is obtained by:

$$\omega_c = 2\pi f_s \left(1 - \frac{d \cos \alpha}{D} \right) \quad (3.1)$$

where ω_c is the cage speed, D is pitch diameter of the bearing, d is the ball diameter and α is the contact angle.

The relationships between the position of each component of a bearing are shown in Figure 3-2. $O_{in}(x_i, y_i)$ and $O_{out}(x_o, y_o)$ are the centers of the inner and outer raceways, $P_j(p_{x,j}, p_{y,j})$ is the center of the j th roller, R_i and R_o are the radii of the inner and outer raceway and r is the radii of the roller. The angular position of the j th roller

on outer raceway is given:

$$\alpha_{out,j} = \omega_c t + \frac{2\pi}{n}(j-1) + \alpha_0 \quad (3.2)$$

where α_0 is the initial value of the cage position and n is the number of rollers.

The roller-raceway contact deformations are obtained as follows:

$$\delta_{in,j} = r + R_i - OP_{in,j} \quad (3.3)$$

$$\delta_{out,j} = r - R_o + OP_{out,j} \quad (3.4)$$

where

$$OP_{in,j} = \sqrt{(p_{x,j} - x_i)^2 + (p_{y,j} - y_i)^2} \quad (3.5)$$

$$OP_{out,j} = \sqrt{(p_{x,j} - x_o)^2 + (p_{y,j} - y_o)^2} \quad (3.6)$$

The relationship between $\alpha_{in,j}$ and $\alpha_{out,j}$ is shown as follows:

$$\cos \alpha_{in,j} = \frac{(x_o + OP_{out,j} \cos \alpha_{out,j} - x_i)}{OP_{in,j}} \quad (3.7)$$

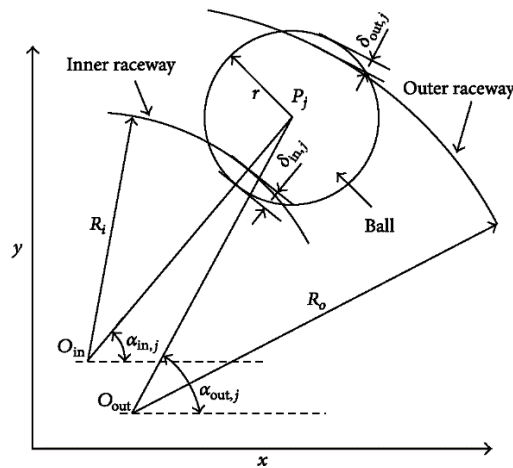


Figure 3-2. Relative position of the components for the j th roller [89]

3.2.2 Contact deformation during roller passing the defect area

The contact deformation changes sharply when the roller passes the defect area. For modeling defect area on the raceways, a shape function $\gamma(\alpha)$ is built. This function can be used to contain spall faults of any geometry [72, 74]. A rectangular shaped, sharp-edged bearing defect on the outer raceway can be modeled as:

$$\gamma(\alpha_{out,j}) = \begin{cases} \lambda, & \alpha_{entry} < \alpha_{out,j} < \alpha_{exit} \\ 0, & otherwise \end{cases} \quad (3.8)$$

where α_{entry} and α_{exit} are the angular position of the entry and exit points, and λ is the depth of the defect.

Thus, the contact deformation when the roller passes the defect area can be obtained by follows:

$$\delta_{in,j} = r + R_i - OP_{in,j} - \gamma(\alpha_{out,j}) \quad (3.9)$$

$$\delta_{out,j} = r - R_i + OP_{in,j} - \gamma(\alpha_{out,j}) \quad (3.10)$$

3.2.3 Hertzian contact model

When roller rolls on the raceways, Hertzian contact theory can be used to calculate the contact force between the roller and the raceways over a long period, as shown in the following:

$$Q = K \times \delta^z \quad (3.11)$$

Where Q is the contact force, K is the contact stiffness, δ is the contact deformation, z is equal to 10/9 for the cylindrical roller bearings and 1.5 for ball bearings [90].

Because the Hertzian contact force only occurs when a rolling element and a raceway contact. The contact force is equal to zero when the contact deformation is equal or less than zero.

The contact forces between the j th roller and both raceways are given:

$$Q_{in,j} = \begin{cases} K_{in} \delta_{in,j}^{10/9}, & \delta_{in,j} > 0 \\ 0, & \delta_{in,j} \leq 0 \end{cases} \quad (3.12)$$

$$Q_{out,j} = \begin{cases} K_{out} \delta_{out,j}^{10/9}, & \delta_{out,j} > 0 \\ 0, & \delta_{out,j} \leq 0 \end{cases} \quad (3.13)$$

where K_{in} is the stiffness of the inner race, and K_{out} is the stiffness of the outer race.

The total contact forces between the rollers and both raceways in the x-axis and y-axis directions are calculated by follows:

$$Q_{in,x} = \sum_{j=1}^n Q_{in,j} \times \cos \alpha_{in,j} \quad (3.14)$$

$$Q_{in,y} = \sum_{j=1}^n Q_{in,j} \times \sin \alpha_{in,j} \quad (3.15)$$

$$Q_{out,x} = \sum_{j=1}^n Q_{out,j} \times \cos \alpha_{out,j} \quad (3.16)$$

$$Q_{out,y} = \sum_{j=1}^n Q_{out,j} \times \sin \alpha_{out,j} \quad (3.17)$$

3.2.4 Damping force

The radial contact damping forces between j th roller and both raceways are given:

$$F_{d,in,j} = \begin{cases} c_b \times \delta'_{in,j}, & \delta_{in,j} > 0 \\ 0, & \delta_{in,j} \leq 0 \end{cases} \quad (3.18)$$

$$F_{d,out,j} = \begin{cases} c_b \times \delta'_{out,j}, & \delta_{out,j} > 0 \\ 0, & \delta_{out,j} \leq 0 \end{cases} \quad (3.19)$$

where c_b is the damping coefficient of the rollers is given as follows [91]:

$$0.25 \times 10^{-5} \times K_{lin} \leq c_b \leq 2.5 \times 10^{-5} \times K_{lin} \quad (3.20)$$

where K_{lin} is the linear stiffness of the bearing, here K_{lin} is equal to 3.34×10^4 N/mm.

The damping coefficients of the housing and shaft are given as

$$c_s = \frac{fa_l \times K_s}{\omega} \quad (3.21)$$

where fa_l is the loss factor based on the material, ω is the resonance frequency, and K_s is the support shaft stiffness. For the simulation analysis, the parameters are given as $fa_l=0.01$, $\omega=30$, $K_s = 3.7 \times 10^4$ N/mm.

The total contact damping forces between the rollers and both raceways in the x-axis and y-axis directions are calculated by follows:

$$F_{d,in,x} = \sum_{j=1}^n F_{d,in,x} \times \cos \alpha_{in,j} \quad (3.22)$$

$$F_{d,out,x} = \sum_{j=1}^n F_{d,out,x} \times \cos \alpha_{out,j} \quad (3.23)$$

$$F_{d,in,y} = \sum_{j=1}^n F_{d,in,y} \times \sin \alpha_{in,j} \quad (3.24)$$

$$F_{d,out,y} = \sum_{j=1}^n F_{d,out,y} \times \sin \alpha_{out,j} \quad (3.25)$$

3.2.5 Vibration equations of motion

So far, the vibration equations for the inner and outer raceways in the x-axis and y-axis directions are obtained based on previous work:

$$M_i x_{in}'' + c_s x_{in}' + K_s x_{in} + Q_{in,x} + F_{d,in,x} = 0 \quad (3.26)$$

$$M_i y_{in}'' + c_s y_{in}' + K_s y_{in} + Q_{in,y} + F_{d,in,y} = -M_i g \quad (3.27)$$

$$M_o x_{out}'' + F_{d,out,x} - Q_{out,x} = 0 \quad (3.28)$$

$$M_o y_{out}'' + F_{d,out,y} - Q_{out,y} = -M_o g - F_l \quad (3.29)$$

The parameters of each component of the bearing are given in Table 3-1. The

vibration equations of motion are solved by the fourth-order Runge-Kutta algorithm and commercial software Matlab.

Table 3-1. Geometrical parameters of the test bearing

Parameter	Value
Number of rolling elements N_b	17
Contact angle α	14°02'10
Ball diameter d	8.2 mm
Pitch diameter D_p	45.6 mm
Outer race inner radius R	27.05 mm
Radius of the inner and outer raceway R_i	15 mm
Radius of the outer raceway R_o	31 mm
c_b	0.8 Ns/mm
K_{in}, K_{out}	8.98×10^5 N/mm
M_o, M_i	0.513, 10.25 Kg

3.2.6 Analysis result

3.2.6.1 Bearing defective response

The purpose of establishing the dynamic model is to illustrate the contact procedure starting from when the roller enters the defect at the entry point to when it restresses back to its normal load carrying capability at the exit point. The experimental and simulated vibration responses are shown in Figure 3-3. The rotating speed is 600 rpm, and the defect size is 2.5 mm. In the experiment, the accelerometer was mounted on the outer race.

By comparing the two plots in Figure 3-3, the dynamic model is able to describe some key features in the experimental results. As the simulated result is agreeable with the experimental data, the effectiveness of the dynamic model is validated. Then these key features can be used to measure the defect size in the outer race.

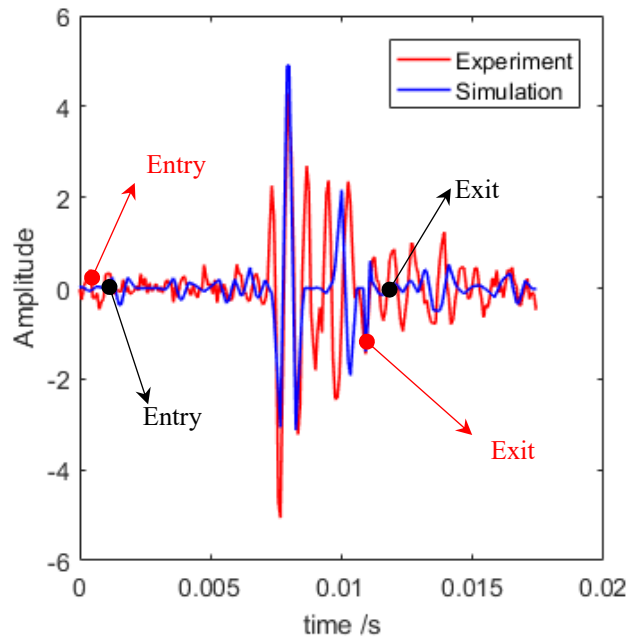


Figure 3-3. Experimental and simulated vibration response.

3.2.6.2 Different defect sizes with large depth size

For this simulation analysis, two different defect sizes (2 and 12 mm) are used to analyze the contact force variation. In order to illustrate easily, the defect size is expressed by the angular domain as the Figure 3-5. The shape of the defect is square with depth 100 μm . A static load force of 100 N is applied on the bearing housing in the vertical radial direction as Figure 3-4 shown. The load distribution zone is from 90° to 270° . The center of the defect size is chosen at the maximum load force of the load zone.

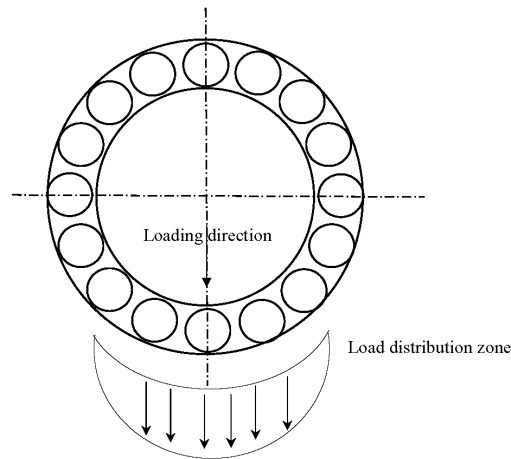


Figure 3-4. Load direction and distribution on the bearing.

Figure 3-6 shows the contact force and deformation of the roller with different size under the static load of 100 N. For the bearing without defect, the contact force and deformation of the roller changes with the acting load force value. The maximum value of the contact force and deformation occurs at the maximum value of the load zone in the vertical radial direction. The values of the contact force and contact deformation are zeros when the roller enters the defect area. This is because that in the large depth case, the deformations of the roller and raceways are zeros as the roller loses its load carrying capability, so there is no contact force between the roller and raceways. When the roller loses its load carrying on the defect area, the load is redistributed by increased loading on the rollers outside the defect area. When the roller leaves the defect area, high amplitude of the contact force and deformation is generated as the roller restresses back to its normal load carrying capability.

Thus, the entry point can be identified by observing the relationship between the contact force and deformation. It reveals that when roller enters the defect area, the contact force and deformation amplitude decrease as the de-stressing of the roller in the defect area. At this stage, the center of the roller passed through the entry point of the defect. The contact deformation and the contact forces applying on the roller become zero. As the result, the energy will lose.

The exit point also can be identified by the same observation of the contact force and deformation variation. When a roller traveled to the end edge of the defect, the contact force increases as a result of the roller may alternately hit inner and outer raceways and then it decreased. Finally, the roller re-stresses back to its normal load carrying capacity [92]. As the Figure 2-4 shown, multiple impacts were generated when the roller hit the exit of the defect in previous studies and this position was called exit point [75, 76]. However, the center of the roller is not at the end edge of the defect. Based on the observation of the bearing dynamic model, the exit event includes two components: the first is the high-frequency impulse response component which is produced when the roller hit the end edge of the defect. The second is the low-frequency response component which is generated when the roller restresses back. Therefore, the exit point can not be identified by detecting the maximum local amplitude of the high-frequency response.

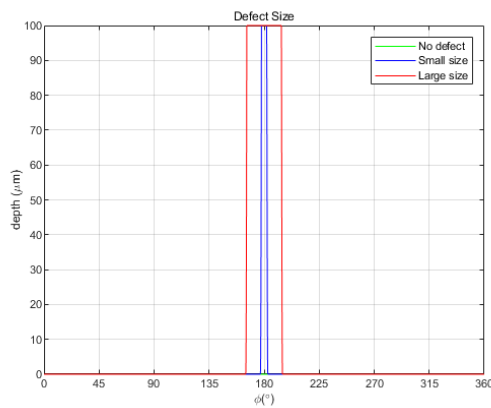


Figure 3-5. Defect information of different sizes with large depth.

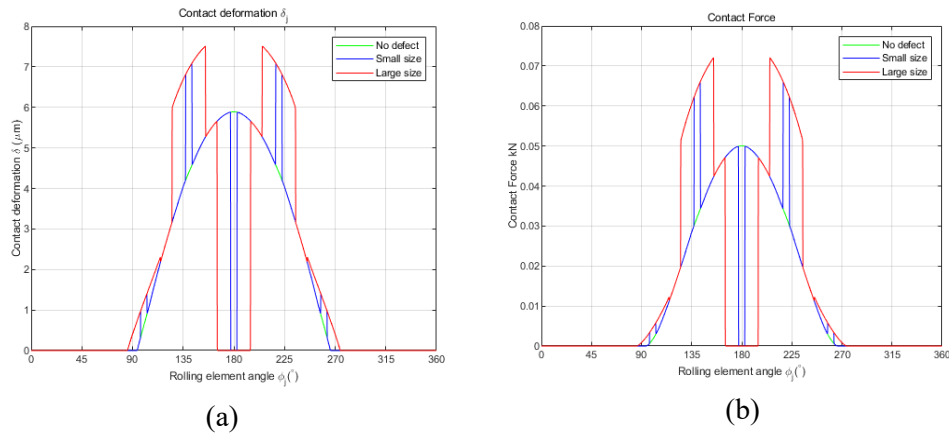


Figure 3-6. Simulation result of different defect sizes with large depth value. (a) Contact deformation (b) Contact force.

3.2.6.3 Different defect sizes with small depth

For this simulation analysis case, two different defect sizes (3 and 12 mm) are used. In order to illustrate easily, the defect size is transformed to the angular domain as Figure 3-7 shown. The shape of the defect is square with depth 9 μm . The center of the defect size is chosen at the maximum load force of the load zone. The load force is 100 N in the vertical radial direction.

Figure 3-8 shows the contact force and deformation of the roller with different sizes under the static load of 100 N. The values of the contact force and contact deformation are not zeros when the roller enters the defect area. This is because of in the small depth case, the roller will travel on the defect surface. The contact force and deformation of the roller will decrease firstly as the de-stressing effect. After that, the contact force and deformation will increase with the load increase. At the same time, the roller still holds part of the load, and the load is also redistributed by increased loading on the rollers outside the defect area. The peak value of the contact force and deformation occurs at the vertical radial direction. The contact force and deformation will decrease with the load decreasing. When the roller leaves the defect area, high amplitude of the contact

force and deformation is also generated as the roller is reloaded.

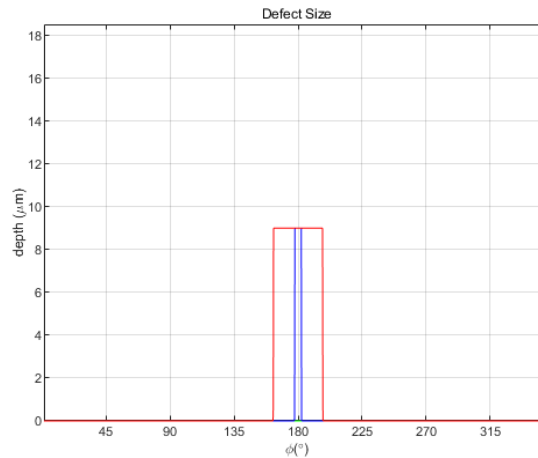


Figure 3-7. Defect information of different width sizes with the small depth value.

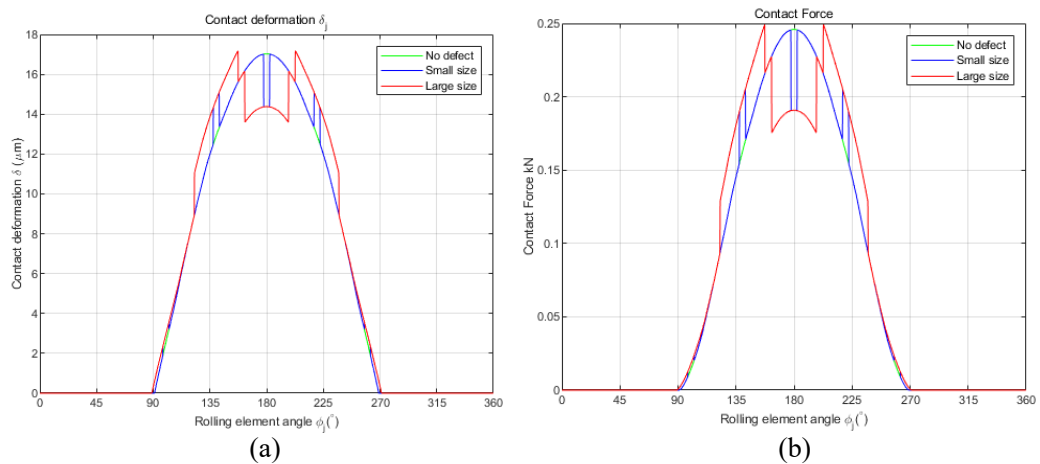


Figure 3-8. Simulation result of different defect sizes with the small depth. (a) Contact deformation, (b) Contact force.

In summary, energy is lost when roller enters the starting edge of a defect, and high energy is generated when the roller hits the ending edge of the defect. Once the roller leaves the end edge of the defect, the low energy will be generated as the roller restresses back to normal load carrying capacity. Thus, the entry and exit points can be identified by the energy distribution of the vibration response signal. This will be introduced in Chapter 5.

3.2.6.4 Load effect

Three different load forces (100, 600, 1200 N) are applied to the bearing housing in the vertical radial direction as Figure 3-4 shown. The defect size is 2 mm, and the depth is 100 μm .

Figure 3-9 (a) and (b) show the contact deformation and force variation of the estimated number roller under different applied loads. As a result, with the applied load increase, the corresponding displacement of the roller relative also increases.

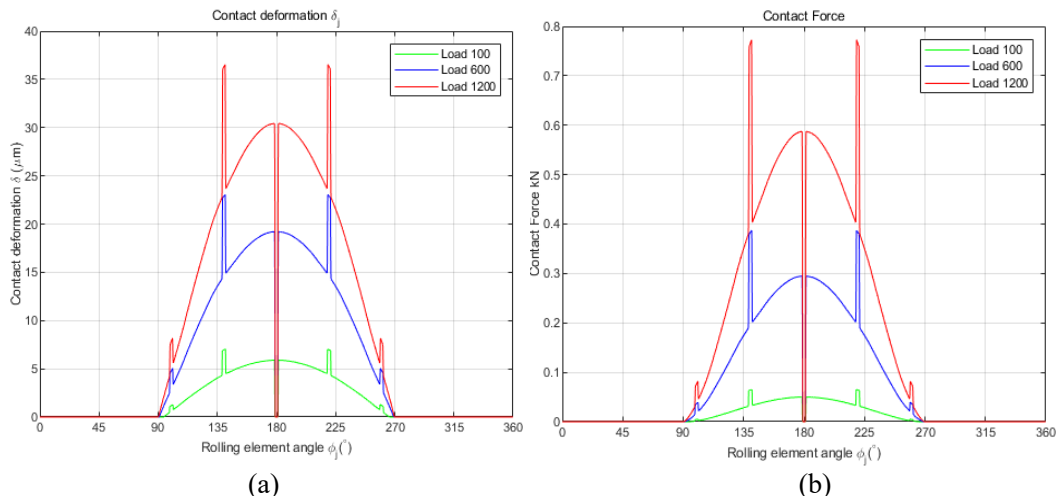


Figure 3-9. Simulation result of the dynamic model with different loads. (a) Contact deformation of the roller, (b) Contact force of the roller.

3.3 Estimation model of the defect size

Once the entry and exit points are identified, the time duration between the entry and exit points can be measured and the defect size can be estimated. The time duration of the roller passing over the defect is defined as Figure 3-10 shown. The next assignment is to estimate the defect size with the time duration obtained from the vibration signal. The details of defect size estimation will be introduced in this section.

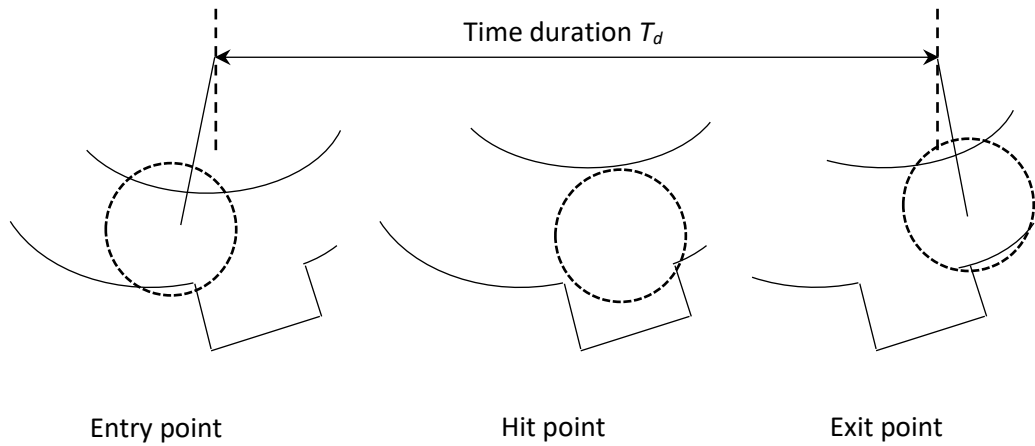


Figure 3-10. Events of a roller passes the defect area.

3.3.1 Estimation model of small defect size

For the small defect size condition, the roller doesn't travel the surface of the defect. The defect size can be estimated by the model as shown in Figure 3-11. The center of the raceways is O, P1 is the center of the roller where at the entry point P3 and P2 is the center of the roller where at the exit point P4.

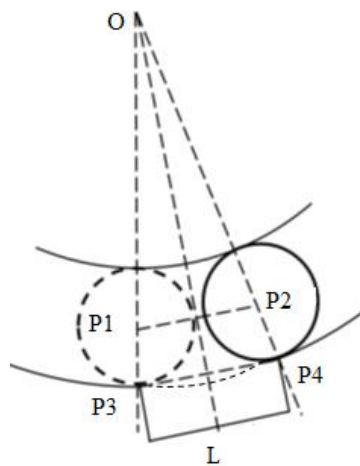


Figure 3-11. Estimation model

Assuming the shaft speed is constant, and no slipping between the roller and cage.

The velocity of the roller cage is given:

$$V_c = \omega_c (r + R_i) \quad (3.30)$$

where r is the radius of the roller, R_i is the radius of the inner race, ω_c is the angular velocity of the cage and it can be calculated by equation 2.4 in Chapter 2. So, the equation can be updated:

$$V_c = \pi \times f_r \times \left(1 - \frac{d}{D} \cos \alpha\right) \times (r + R_i) \quad (3.31)$$

where f_r is the shaft speed in Hertz, d is the diameter of the roller, D is the pitch diameter, and α is the contact angle. The time duration from P1 to P2 is T_d . Then, the defect size L can be estimated by following:

$$L = V_c \times T_d \times \frac{R_i + d}{R_i + d/2} \quad (3.32)$$

From the above equation, it can be found that the estimation of the defect size is affected by the shaft speed. For the time-vary speed condition, it is not suitable. Therefore, the estimation model needs to be developed for estimating the defect size under time-varying conditions.

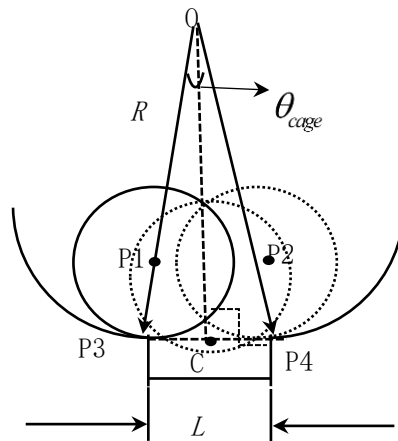


Figure 3-12. Model of roller passing over a small defect.

As Figure 3-12 shown, the center of the raceways is O, P1 is the center of the roller where at the entry point P3 and P2 is the center of the roller where at the exit point P4. C is the center of the defect width. It is obviously seen that the OC and P3P4 are perpendicular to each other. Thus, the defect size can be estimated as follows:

$$L = 2 \times R \times \sin(\theta_{cage}/2) \quad (3.33)$$

where R is the outer race's inner radius, L is the fault width, θ_{cage} is the angular distance of the cage element between the entry and exit points. The θ_{cage} can be obtained by the following equation:

$$\theta_{cage} = \theta_{shaft} \times \frac{f_{cage}}{f_{shaft}} \quad (3.34)$$

where θ_{shaft} is the angular distances of the shaft, when the roller passes over the defect area. f_{cage} and f_{shaft} are the cage and shaft speeds, respectively. The relation between f_{cage} and f_{shaft} is the following:

$$f_{cage} = \frac{f_{shaft}}{2} \times \left(1 - \frac{d}{D_p} \times \cos(\alpha) \right) \quad (3.35)$$

where d is the roller diameter, α is the contact angle and D_p is the pitch diameter. The equation 3.35 can be rewritten as:

$$L = 2 \times R \times \sin \left(\theta_{shaft} \left(1 - \frac{d}{D_p} \times \cos(\alpha) \right) / 4 \right) \quad (3.36)$$

Once the angular distance is obtained, the fault size can be estimated. Therefore, the signal needs to be resampled from the time domain to the angle domain.

3.3.2 Estimation model of the large defect size

For the large defect size condition, the roller travels on the surface of the defect. The defect size can be estimated by the model as shown in Figure 3-13. P'1 is the center of the roller where at the entry point, P'2 is the center of a roller at the location of the roller starting to travel over the defect area. P'3 is the center of the roller where at the hitting point. P'4 is the center of the roller where at the exit point. The defect size is divided into three parts: $L1$, $L2$, and $L3$. The θ_1 , θ_t and θ_2 are the corresponding angular travel.

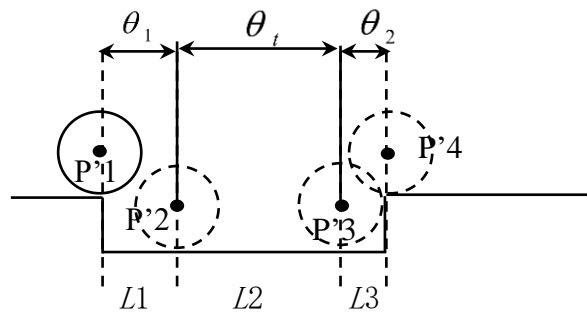


Figure 3-13. Model of roller passing over a small defect.

L1 phase: At this phase, the angular duration between the P'1 and P'2 only is decided by the geometry parameters of a bearing and the maximum local contact deformation between the roller where at the entry point and inner and outer raceways. When the roller enters the defect at P'2, the contact deformation of the bearing is related to the distance of the roller become empty load [74, 75, 93]. The angular travel θ_1 can be obtained by the following equation [94]:

$$\theta_1 = \sqrt{\frac{2r\delta_{\max} - \delta_{\max}^2}{(r+R)^2}} \quad (3.37)$$

where R is the radius of the outer race, r is the radius of the roller and δ_{\max} is the maximum contact deformation of the roller. The δ_{\max} can be estimated as the following

equation based on the equation 3.11-3.12.

$$\delta_{\max} = \left(\frac{Q_{\max}}{K_n} \right)^{\frac{9}{10}} \quad (3.38)$$

where Q_{\max} is the maximum radial load force acting on the roller, and K_n is the the stiffness of inner or outer race.

Thus, the distance $L1$ can be measured as follows:

$$L_1 = \sin(\theta_1) \times R \quad (3.39)$$

L2 phase: At this phase, the roller travels on the surface of the defect. The angular duration θ_t of the center of a roller is measured from the point P'2 to the first time hitting the exit point (point P'3). Thus, the distance $L2$ can be calculated with the angular duration:

$$L_2 = 2 \times R \times \sin(\theta_t/2) \quad (3.40)$$

L3 phase: At this phase, the angular duration of the center of a roller also only is decided by the geometry parameters of a bearing and the maximum local contact deformation between the roller where at the entry point and inner and outer raceways. Thus, the distance $L3$ is equal to $L1$. The total defect size is: $L = L1 + L2 + L3$.

3.4 Conclusions

This chapter establishes a nonlinear dynamic model of a defective bearing, a localized defect on the outer raceway, to describe the vibration response of the system. This model can clearly illustrate the changes of contact deformation and force of the defective bearing pattern. The entry and exit events can be identified by these key features. Then, two defect size estimation models from the small to the large size are proposed to measure the defect size of a rolling element bearing under time-varying

speed conditions. The small model is validated in Chapter 5 with the experimental results.

Chapter 4 Bearing fault component extraction

4.1 Introduction

The fault signal contains complicate information due to the characteristics of the rotating machinery [95]. How to extract the fault feature is the key point to identify the failure type of rolling bearing. To extract the good features, the signal processing techniques are widely used [96]. By this way, the rolling bearing fault can be diagnosed effectively. Kinds of signal processing techniques can be used to extract the fault information, such as the fast Fourier transform (FFT) for the periodic and stationary signal, wavelet transform (WT), the empirical mode decomposition (EMD) and local mean decomposition (LMD) for the nonlinear and non-stationary signal. The FFT is not suitable for the non-stationary signal. The WT needs to select the mother wavelet and decomposition levels. If these parameters are selected inappropriate, the wrong result will be obtained. The EMD and LMD have similar characteristics, and both of them are adaptive signal processing methods. In contrast to EMD, the LMD method uses the moving average to get the amplitude envelope instead of cubic spline interpolation. The overshooting and undershooting effect caused by cubic interpolation can be eliminated. The signal of a fault bearing based on time-varying speed conditions consists of an impulse fault component, determined component and random noise component. To extract the fault component, a new fault diagnosis method of rolling element bearing is proposed in this chapter.

4.2 Cepstrum

The cepstrum is the inverse Fourier transform of the log spectrum [24, 25]. The main purpose of using the cepstrum pre-whitening (CPW) technique is to separate the determined component from the vibration signal [17, 23]. Once the constant component signal has been removed, the entry and exit events are enhanced. For a given vibration signal $x(t)$, the cepstrum is defined as follows:

$$C(q) = F^{-1} \{ \ln(A(f)) + j\varphi(f) \} \quad (4.1)$$

where q is the quefrency, and $X(f)$ is the Fourier transform of $x(t)$:

$$X(f) = F \{ x(t) \} = A(f)e^{j\varphi(f)} \quad (4.2)$$

The real cepstrum can be obtained by taking the real part in equation (4.1).

Cepstrum analysis can concentrate the harmonic components into a series of peaks. These peaks in the quefrency domain indicate the periodic harmonic components in the spectrum. It is a simple way to separate to deterministic components from the vibration signal by editing the amplitude of the real cepstrum. The cepstrum editing procedure is shown in Figure 4-1. There are two ways to edit the real cepstrum signal. The first one is setting a zero value for the whole real cepstrum (except possibly at zero quefrency), so that the discrete harmonics and resonances are eliminated in the frequency domain. The pre-whitened signal can be obtained by recombining the edited cepstrum signal with the phase information of the original signal and inverse transforming to the time domain. The other one is eliminating the deterministic excitations by removing the peaks with filtering operations, which is a cepstrum editing procedure. The edited cepstrum signal is then transformed to the frequency domain. The edited real cepstrum signal can be obtained by recombining the edited cepstrum signal with the phase information of the original signal and inverse transforming to the time domain.

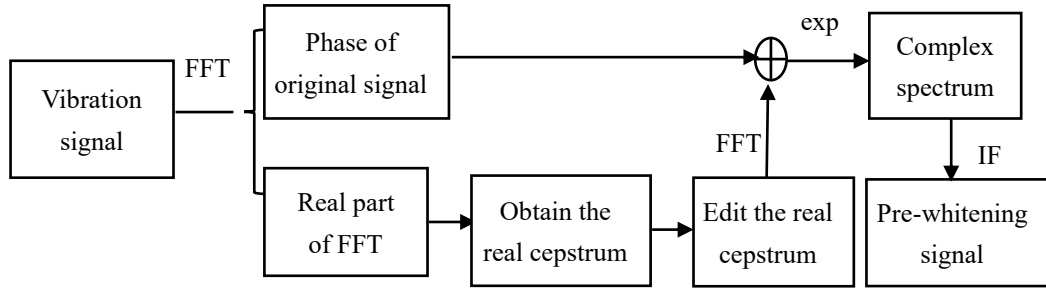


Figure 4-1. A flowchart of the cepstrum editing procedure.

By using the CPW technique, the entry and exit events of rolling elements passing the defect area were enhanced. However, it is still hard to determine the obvious entry and exit points in the time domain due to the effect of random noise and speed variation. In the next section, the LMD method is used to solve this problem.

4.3 Fault component extraction by LMD

4.3.1 Introduction of LMD.

A series of product functions (PFs) can be obtained by the LMD processing method, each of PFs represents a mono-component of the original signal which contains multi-components. The LMD method has been widely used to extract fault features for diagnosing rolling element bearing faults [37, 39, 97]. It can also be used to extract the fault component by selecting an appropriate PF which contains the fault component signal.

For a given signal $u(t)$, the LMD decomposition procedure is shown as following steps [36]:

Step.1. Extract the extrema values $z_i (i=1, \dots, M)$ of the original signal $u(t)$.

Step.2. Obtaining the local mean value m_i and the amplitude envelope estimate a_i via the two successive extrema.

$$m_i = \frac{z_i + z_{i+1}}{2}, i = 1, 2, \dots, M-1 \quad (4.3)$$

$$a_i = \frac{|z_i + z_{i+1}|}{2}, i = 1, 2, \dots, M-1 \quad (4.4)$$

These local means are linked by straight lines with extending between successive extrema. The continuous local mean function $m_{11}(t)$ is formed through moving averaging. The local magnitude envelope function $a_{11}(t)$ can be obtained by the same smoothing procedure as the local means.

Step.3. The continuous local mean function is subtracted from the original signal $u(t)$:

$$h_{11}(t) = u(t) - m_{11}(t) \quad (4.5)$$

The demodulated amplitude $s_{11}(t)$ is obtained with the envelope function $a_{11}(t)$:

$$s_{11}(t) = \frac{h_{11}(t)}{a_{11}(t)} \quad (4.6)$$

Ideally, if $s_{11}(t)$ was a pure frequency modulated signal, its envelope function $a_{12}(t)$ should satisfy the condition: $a_{12}(t) = 1$. If $a_{12}(t) \neq 1$, $s_{11}(t)$ is taken as the new signal, and the procedures are repeated steps 1-3. The process will stop until $s_{1n}(t)$ is the pure frequency modulated signal. The iteration can be shown as follows:

$$\begin{cases} h_{11}(t) = x(t) - m_{11}(t) \\ h_{12}(t) = s_{11}(t) - m_{12}(t) \\ \vdots \\ h_{1n}(t) = s_{1(n-1)}(t) - m_{1n}(t) \end{cases} \quad (4.7)$$

where the $s_{1n}(t)$ is:

$$\begin{cases} s_{11}(t) = h_{11}(t) / a_{11}(t) \\ s_{12}(t) = h_{12}(t) / a_{12}(t) \\ \vdots \\ s_{1n}(t) = h_{1n}(t) / a_{1n}(t) \end{cases} \quad (4.8)$$

Step.4. The envelop signal can be obtained as:

$$a_1(t) = a_{11}(t)a_{12}(t)\cdots a_{1n}(t) = \prod_{p=1}^n a_{1p}(t), \lim_{n \rightarrow \infty} a_{1n}(t) = 1 \quad (4.9)$$

Thus, the envelope function $a_1(t)$ is expressed as the instantaneous amplitude. The instantaneous phase is:

$$\theta_1(t) = \arccos(s_{1n}(t)) \quad (4.10)$$

The instantaneous frequency can be defined as:

$$f_1(t) = \frac{1}{2\pi} \frac{d\theta_1(t)}{dt} \quad (4.11)$$

Step.5. The first $PF_1(t)$ can be obtained from the product envelope function $a_1(t)$ and frequency modulated signal $s_{1n}(t)$:

$$PF_1(t) = a_1(t)s_{1n}(t) \quad (4.12)$$

Step.6. New data $u(t)$ can be obtained by subtracting $PF_1(t)$ from the original data. Then, steps 1-5 are repeated k times until $u_k(t)$ is a constant or doesn't contain oscillations. Finally, the original signal can be reconstructed:

$$x(t) = \sum_{p=1}^k PF_p(t) + u_k(t) \quad (4.13)$$

As previously introduced, the end effect problem existed on the LMD. Researchers have proposed some improvement methods to eliminate it.

Huang et al. [98] proposed the mirror periodic method to extend the two ends data for handling the end effect. The data was only extended by using the extreme values which close to the ends, the characteristics of data were not considered. Considering the tendency of the data, Gai et al. [99] proposed a wave matched processing method. However, a choppy signal was generated near the ends. By using the neural network, Yong-jun et al. [100] extended the data by the adaptive prediction method. But

parameter selection and long computation time should be considered on the applications. An adaptive extending method based on the spectral coherence was proposed to eliminate the end effect on the LMD [97]. However, it just separated the signal into a number of segments, some information of the signal missed when the segment was very large. Wang et al. [101] used the mean local extrema on the ends to process the boundary effect of LMD. But the matched signal cannot reveal the similarity of the original signal.

Although these improvement methods could eliminate the end effect, the limitations also cannot be overcome due to the non-stationary and nonlinear characteristic of the rolling element bearing. Aiming at improving the performance of the LMD, a novel waveform extension method based on the dynamic time warping (DTW) is proposed in this section.

DTW is a well-known pattern recognition method and has been successfully applied in many fields as its simplicity and effectivity [102]. For example, the spoken words recognition, fingerprint and signature verification and fault diagnosis [103]. The DTW method can find an optimal warp path to calculate the distance between two-time series with different length, the smaller value of distance, more similar time series are. By stretching or shrinking the optimal warp path, time series will be transformed in one-to-one correspondence dimension [104]. DTW can be applied to on the time series with a different phase. As the vibration signal is nonlinear and non-stationary, the DTW method can be used to extend the data dynamically by searching the waves that matched with the ends data from the original signal. By this way, the end effect of the LMD can be eliminated. The improved LMD based on the DTW method is proposed.

4.3.2 DTW-LMD method

To obtain a good signal waveform extension, it is necessary to find the desired signal waveform that has the same feature with one end of the signal. The DTW can find an

optimal warping path to measure the distance between two waveforms is similar but with a different phase. By this way, the similarity between the two waveforms can be determined.

Given two data on the time domain, a sequence X of length S , and the other sequence Y of length N , the procedure of the DTW can be shown as follows:

$$X = x_1, x_2, \dots, x_i, \dots, x_S \quad (4.14)$$

$$Y = y_1, y_2, \dots, y_j, \dots, y_N \quad (4.15)$$

A matrix D with $S \times N$ dimensions is constructed and every element of D is the corresponding distances of the alignment between each point. Then, the warp path W is constructed. It can be expressed as:

$$W = w_1, w_2, \dots, w_k, \dots, w_P \quad (4.16)$$

where P is the length of the warp path W , and $w_k = (i, j)$, $i=1,2,3,\dots,S$, $j=1,2,3,\dots,N$, $k=1,2,3,\dots,P$. $\text{Max}(S, N) \leq P \leq S + N$.

For the warp path matrix, it must satisfy a few of constraints when it is constructed [105]:

1) The beginning of the warp path should be the first point of each two sequences, that is $w_1 = (1, 1)$.

2) The end of the warp path should be the last point of each two sequences, that is $w_p = (S, N)$.

3) The adjacent point $w_k = (i, j)$ and $w_{k+1} = (i', j')$ should be constrained by $i' \in (i, i+1)$, $j' \in (j, j+1)$.

The best warping path can be obtained by the dynamic programming as follows:

$$D(i, j) = d(i, j) + \min\{D(i-1, j-1), D(i-1, j), D(i, j-1)\} \quad (4.17)$$

where the $d(i, j)$ is the distance of the current cell. $D(i, j)$ is the cumulative distance of $d(i, j)$ and the minimum cumulative distances from the three adjacent cells. The warp path is shown as Figure 4-2.

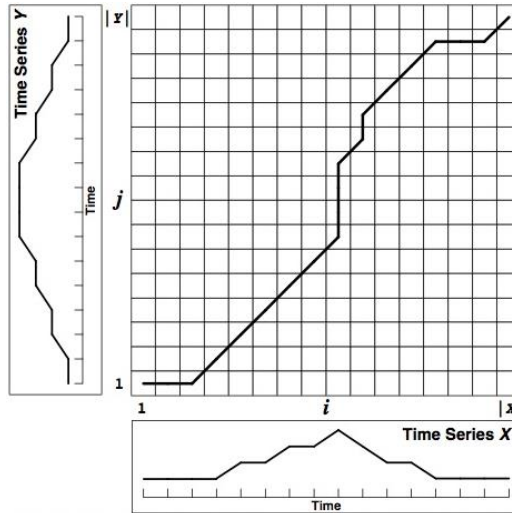


Figure 4-2. A cost matrix with the minimum-distance warp path.

Figure 4-3 shows a simulation result of the DTW processing. Two sequences with different phase are used in Figure 4-3 (a). It is obviously found that the distance value of the DTW is smaller than the Euclidean distance (ED) value. Unlike the ED, the DTW method considers the phase condition and finds the least distance between two series. Therefore, the DTW is a suitable method for identifying the similarity between two sequences.

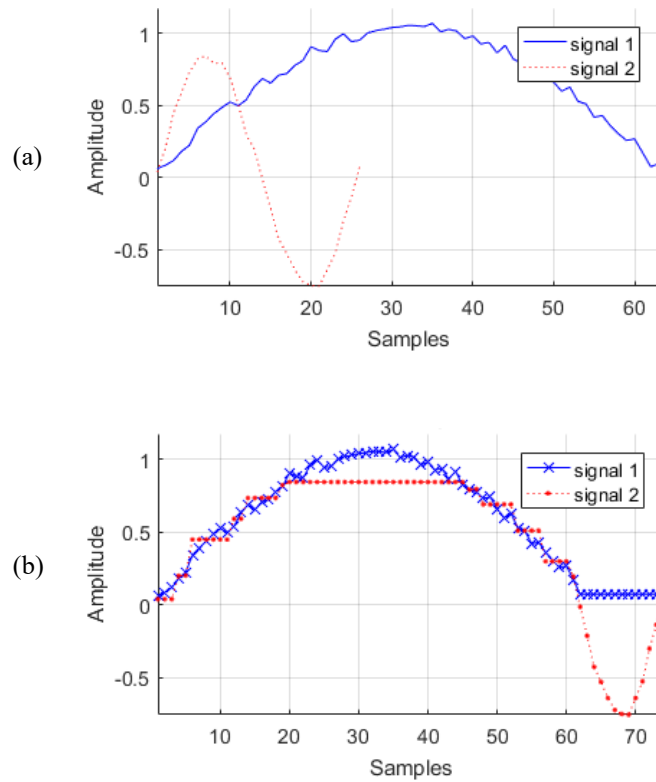


Figure 4-3. Result of DTW processing. (a) the original signals of the two sequences. (b) Warped signal of two sequences.

Thus, a novel method is proposed, called DTW-LMD which eliminates the end effect of the LMD by a new extending method with the DTW. Given a signal x , N is the length of x . The method can be processed as the following steps:

1. Find the extremal points of the data.
2. Construct a data based on the left end-point, first maximum and first minimum. The data length is M . Then, separate the signal into K segments based on the extremal points, except the first and last of them. Each segment contains one maximum and one minimum value.
3. Find the minimum distance between the constructed data and the segment signal by the DTW. Then the matched signal and optimal warp path can be obtained. Use the previous segment of the matched signal for extending left side signal. The extension data is extended on the left side with the optimal warp path.

4. Extend the right-side data. Find the matched signal by the same way of the left side. Use the next segment of the matched signal for extending the right-side signal.

5. Process the extended data by LMD.

6. Obtain the original data length from the DTW-LMD decomposition result.

Simulation analysis is used to validate the DTW- LMD. The simulation signal is shown in Figure 4-4. It contains two components.

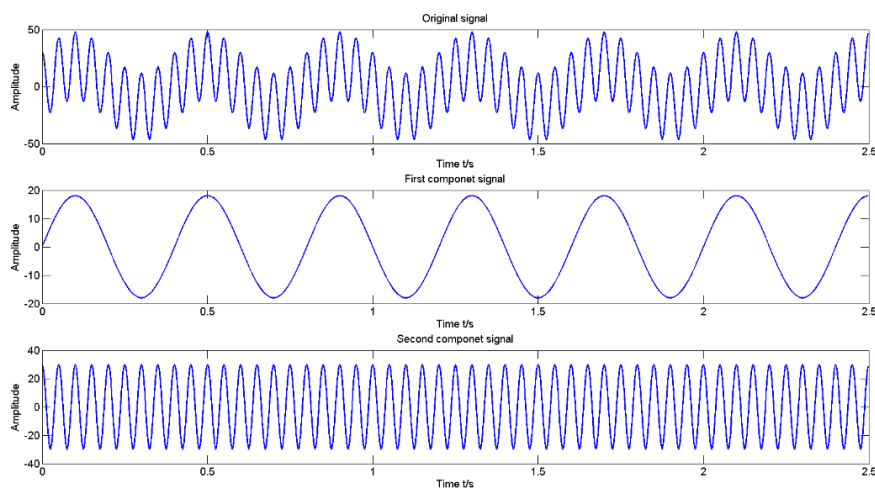


Figure 4-4. A simulation signal.

Figure 4-5 shows the decomposition result of the original LMD method. It can be seen obviously the end effect occurs at the end of the decomposition signal. To solve this problem, the data is extended by the DTW. The extended signal is shown in Figure 4-6. The proposed extension method can keep the inherent characteristic and trend of the original signal.

The decomposition result of the improved LMD method is shown in Figure 4-7. It can be seen that end effects are eliminated at both ends of the decomposition components.

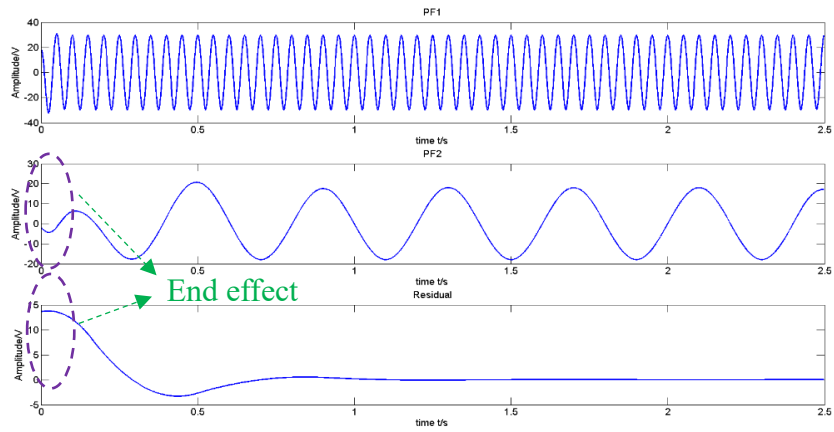


Figure 4-5. Decomposition result of the original signal by LMD.

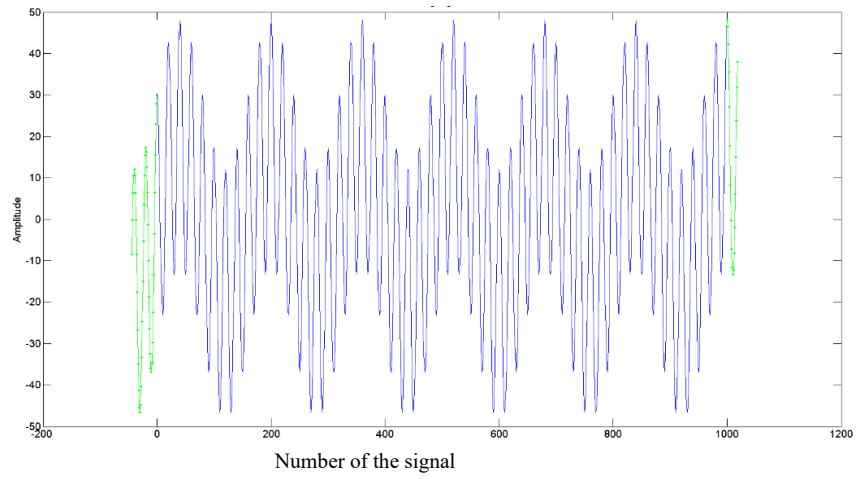


Figure 4-6. Extension signal with DTW.

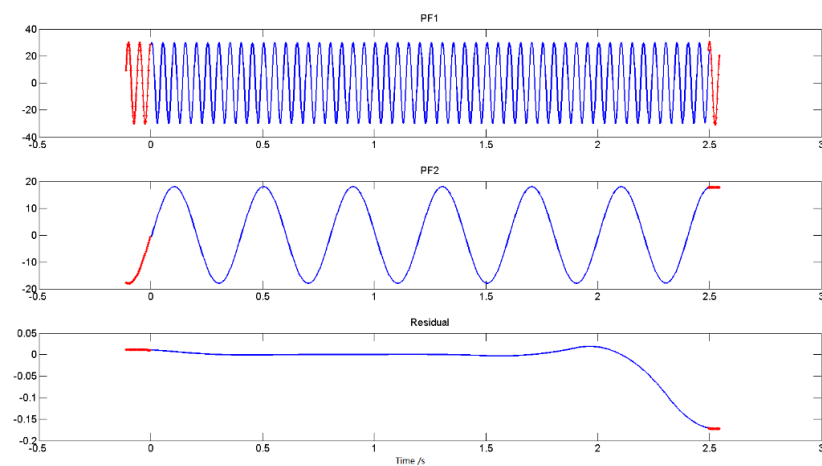


Figure 4-7. Decomposition result of improved LMD method.

Combining the edited cepstrum and improved LMD, a new fault diagnosis method namely EC-LMD is proposed. An experiment is used to illustrate the validity of the proposed method in the next section.

4.4 Experiment setup

A laboratory experiment was conducted with a defect on the outer race to validate the proposed method. The experimental setup is shown in Figure 4-8. The test system consists of a 3-phase AC motor, a shaft supported by two bearings (one is normal, and another is faulty), and an AC controller. The vibration signal is collected by an acceleration sensor that is mounted on the housing of the faulty bearing. The speed signal is collected by a tachometer mounted on the end of the shaft. All the data are obtained by an acquisition card (National Instrument) with a sampling frequency of 12000 Hz. The defect size in outer race of the bearing is 1.5 mm (as Figure 4-9 shown). It is made by electric discharge machining. The fault characteristic orders of the rolling element bearing are listed in Table 4-1.

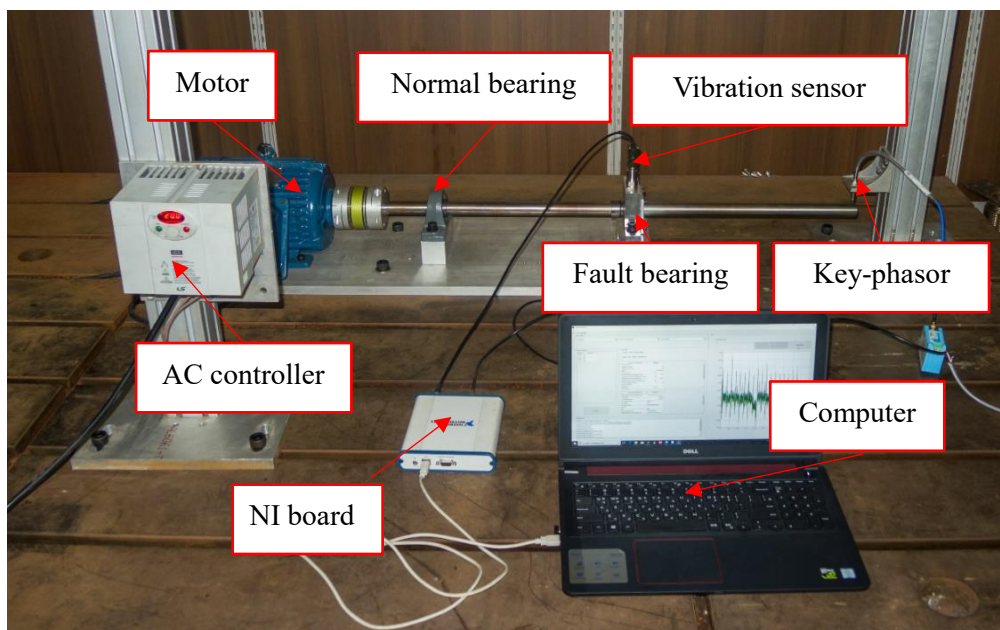


Figure 4-8. Test rig of the rolling bearing with outer race fault.

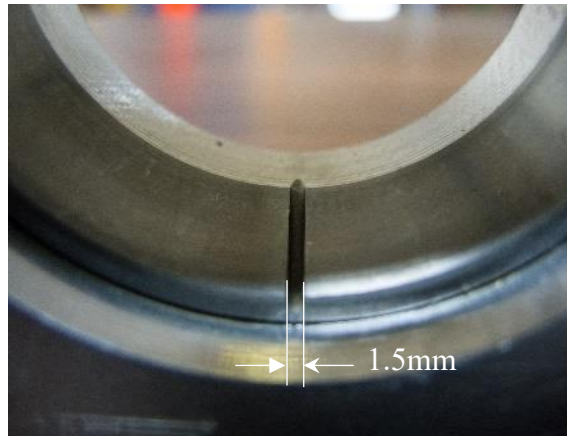


Figure 4-9. A line Defect in the outer race.

Table 4-1. Fault characteristic order.

Fault type	Order
Outer race fault (OF)	7.0155
Inner race fault (IF)	10.0373
Rolling element fault (RF)	5.1893
Cage fault (CF)	0.4096

The key-phasor signal is shown in Figure 4-10 (a). The speed signal is obtained based on the impulses of the key-phasor signal. Figure 4-11 shows the original vibration signal and its order spectrum. The fault characteristic order is not found in the order spectrum. This is caused by the impact signal is smeared in the determined component or resonance component.

4.4.1 Fault diagnosis

a) Envelop spectrum by Low pass filter (LPF)

The fault characteristic order can be detected effectively by the envelope analysis. First, the original signal is resampled with the speed signal. Then the resampled signal is filtered by a low-pass filter. The envelope spectrum is shown in Figure 4-12. The fault

characteristic order and its harmonics are found from the envelope spectrum. However, other peaks also exist. This indicates that the order spectrum contains some other components (deterministic components).

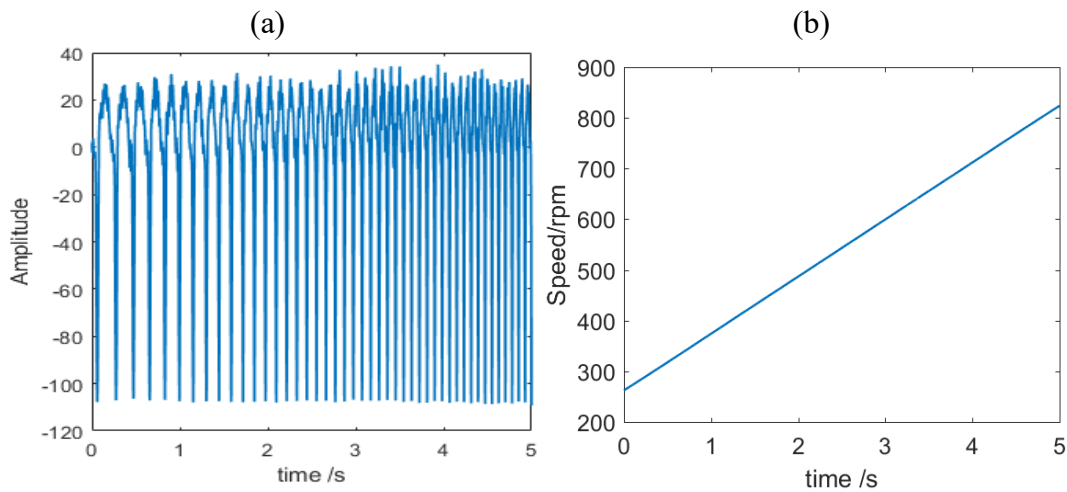


Figure 4-10. Speed signal. (a) Key-phasor signal. (b) Speed signal.

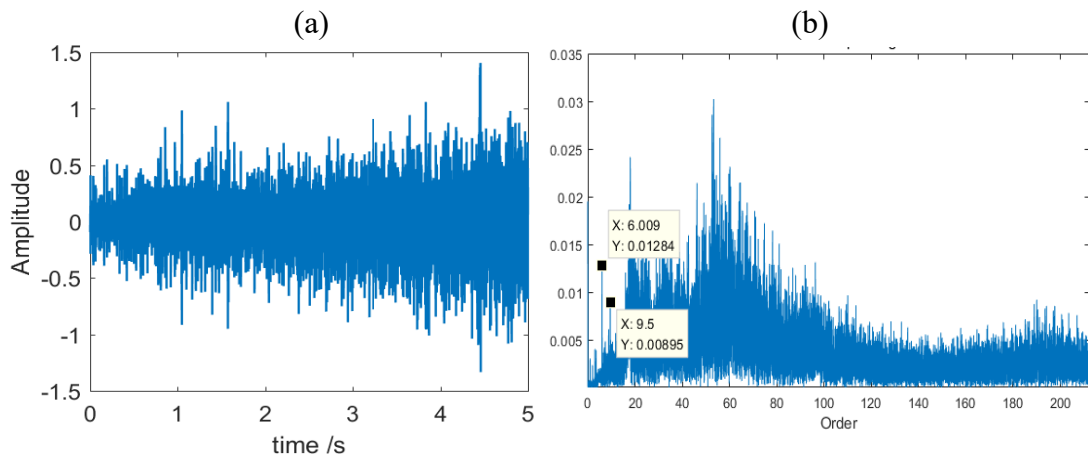


Figure 4-11. Original signal and its order spectrum. (a) Original signal. (b) Order spectrum of the resampled signal.

b) Envelope spectrum by fast kurtogram

The kurtogram is also a popular technique on fault detection of rolling element bearing, which can select the optimal parameters of the band-pass filter. The center frequency and band width are selected based on the kurtosis spectrum. Figure 4-13 shows the result of envelope analysis by kurtogram. The kurtosis spectrum is shown in Figure 4-

13 (a). The optimal bandwidth and center frequency are selected based on the maximum kurtosis value. The envelope spectrum is obtained by the optimal band-pass filter as Figure 4-13 (b) shown. The characteristic order is found in the order spectrum. At the same time, the determining component is also found, and the fault diagnosis result may be affected by these peak values. The envelope spectrum result is better than the LPF method.

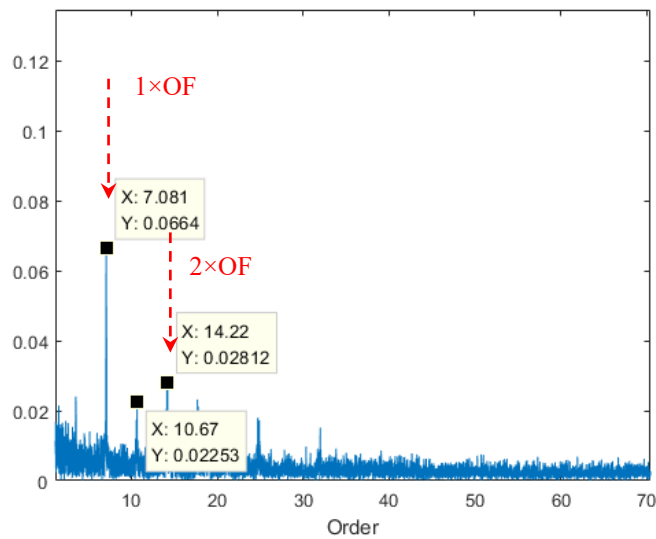


Figure 4-12. Envelope analysis by low-pass filter.

c) Proposed method

The whole procedure of the proposed method is shown in Figure 4-14. Firstly, the edited cepstrum technique is used to remove the determined component of the signal, and the pre-processed signal is resampled with a constant angle interval by a tachometer signal to eliminate the effect of speed variation. Then, the impact component of the fault bearing is extracted from the vibration signal by the LMD method. Lastly, the order spectrum is obtained by the envelope analysis.

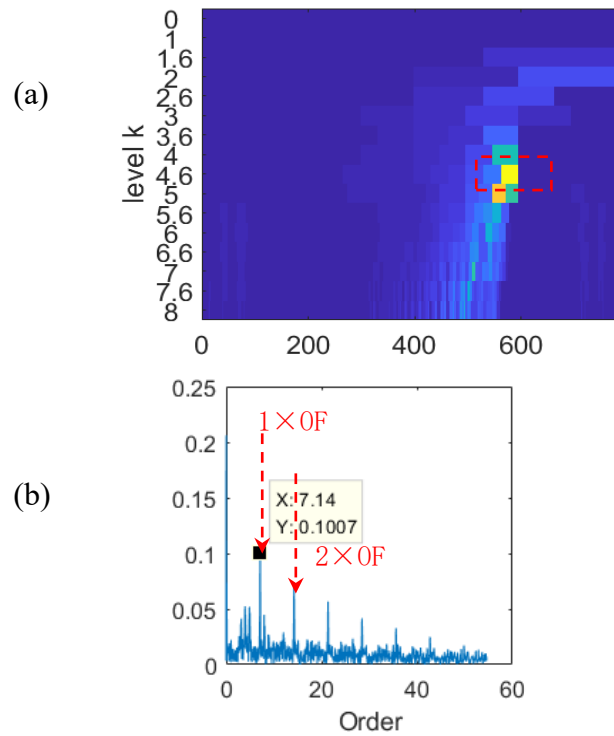


Figure 4-13. Envelope analysis by fast kurtogram. (a) Spectrum of Kurtosis. (b) Envelope spectrum of the filtered signal by SK.

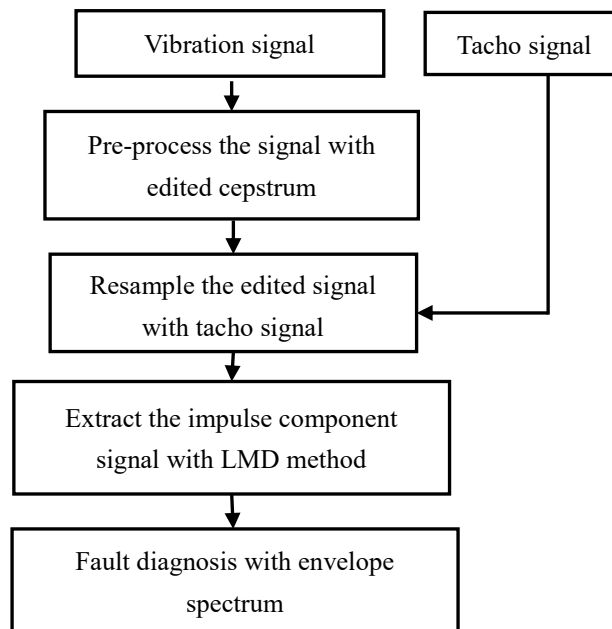


Figure 4-14. Flowchart of the fault diagnosis approach.

The resampled signal after the edited cepstrum processing is shown in Figure 4-15. The amplitude of the vibration signal decreased by removing the determining

component. During the resampling procedure, the signal is firstly upsampled by a factor 10 to avoid aliasing. The vibration signal is resampled by recombining with the speed signal, which was obtained from key-phasor data.

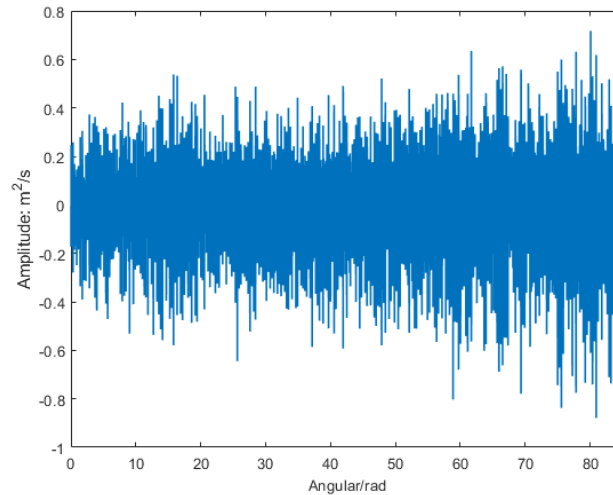


Figure 4-15. Pre-whitening signal by edited cepstrum.

The LMD decomposition result of the resampled signal is shown in Figure 4-16. Several PFs are obtained with the improved LMD process. The first PF is selected for further analysis because it has the biggest correlation coefficient value and keeps the most information of the original signal. The envelope spectrum of PF1 is shown in Figure 4-17. The fault characteristic order and its harmonics are obviously found. There are no other components in the order spectrum. The outer race fault is detected easily and correctly.

4.4.2 Fault size estimation

To enhance the entry and exit events, the edited cepstrum technique is used to remove the determined component of the signal. The pre-processed signal is resampled by a constant angle interval sampling with a tachometer signal to eliminate the effect of speed variation. The impact component of the fault bearing is obtained with the LMD

method. By utilizing the LMD method, the response of entry event and the impulse response of exit event were enhanced when rolling elements passed over the defect area on the outer race.

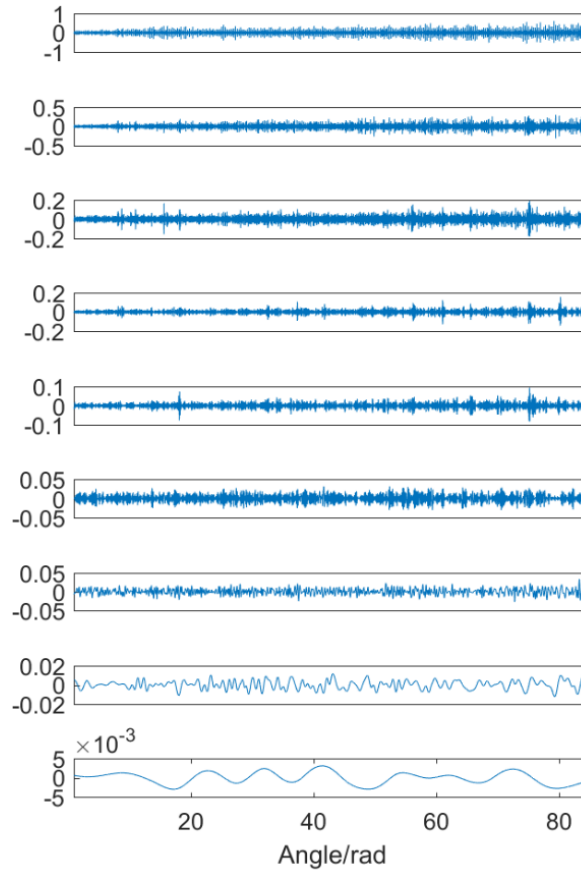


Figure 4-16. Decomposition result of the pre-whitening signal by improved LMD.

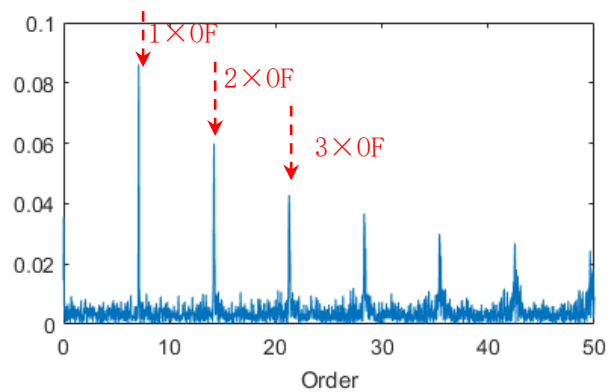


Figure 4-17. Envelope spectrum of PF1.

The pre-whitening signal is shown in Figure 4-18. The entry and exit events can not be found. To enhance the entry and exit events, LMD method is used. The result is shown in Figure 4-19. As previous illustrated, when the roller entered the defect area, the deformation and contact force decreased as the distress effect. The entry point was identified by the starting point of the local maximum value zone. When the roller departed from the defect area at the exit point, it restressed back to its normal load carrying capacity. As a result, the exit point was identified at the end of the vibration level decreasing sharply.

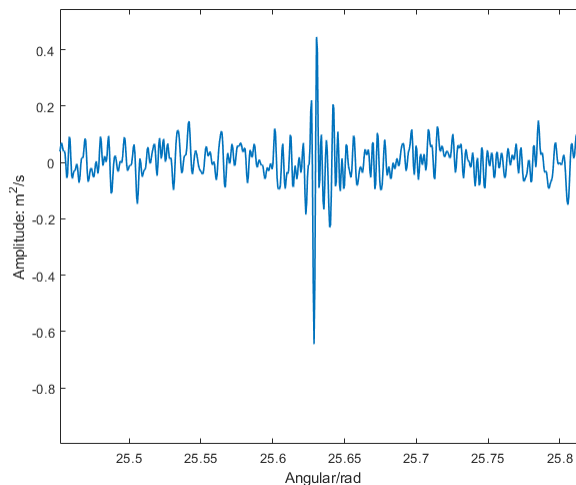


Figure 4-18. Vibration signal after pre-whitening.

The angle duration when the roller passes the defect area is measured with 0.1357 rad in Figure 4-19. The estimated value of the defect size is 1.5035 mm. To check the error and deviation values of the estimation result, 10 portions of the extracted signal which contain an impact component are selected. The estimation result is shown in Table 4-2.

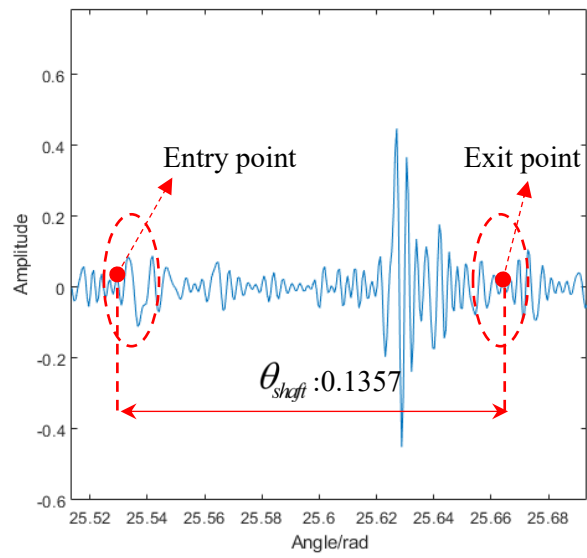


Figure 4-19 Vibration signal by the proposed method.

Table 4-2. Estimation results of the defect size on the outer race.

Number	Values (mm)
1	1.5035
2	1.4404
3	1.4071
4	1.3939
5	1.2864
6	1.3972
7	1.775
8	1.39
9	1.4437
10	1.4548

The mean and deviation values are obtained based on Table 4-2, and their values are 1.4492 and 0.1277, respectively. This measurement is manual to identify the entry and exit points. However, it is still hard to identify the entry and exit points on the angular domain. In addition, this estimation model did not consider the speed slipped between the roller and raceways. These factors may lead to estimation errors.

4.5 Conclusions

A new method, named EC-LMD, was proposed to extract the fault component from the vibration signal in this chapter. The edited cepstrum is used to removing the determining component. To solve the end effect, an improved LMD is developed by the DTW method. The DTW found an optimal wrapped path to extend the signal. Simulation result shows that the improved method can eliminate the end effect effectively. Then the impact component was extracted with the improved LMD method. An experiment, an outer race fault in a bearing, is used to validate the proposed method. Compared with the fast kurtogram, the proposed method obtained a better result of detecting the bearing fault.

At the same time, the entry and exit events were enhanced by the proposed method. Once the entry and exit point were identified, the angle duration was measured on the angular domain signal. Thus, the defect size was estimated with the angle duration based on the estimation model. The result showed that the estimation defect size by using the proposed method is 1.4492 mm. The estimation accuracy is still low as it is hard to identify the entry and exit points. To identify the entry and exit points more exactly, a time-frequency analysis will be introduced in the next chapter.

Chapter 5 Time-frequency analysis for fault size estimation

5.1 Introduction

As previous section introduced, the key point of the defect size estimation is to identify the entry and exit events correctly. However, it is a little difficult to identify the entry and exit points exactly on the time domain signal.

Energy will lose when a roller enters the leading edge of a defect, and high energy is generated when a roller hits the ending edge of the defect. Therefore, the position of the entry and exit points can be identified by the time-frequency analysis, which can represent the energy distribution of the vibration signal. As the characteristic of high time-frequency resolution, the continuous wavelet transform is used to obtain the exact position of the entry and exit points [106]. The Morlet wavelet has been widely used to diagnose the rolling bearing fault as its shape is similar to the impulse component of the rolling element bearing signal [107]. To obtain the optimal wavelet, it is very important to select appropriate parameters of Morlet wavelet (bandwidth and center frequency). Researchers have proposed many methods to optimize the bandwidth and center frequency [108-110]. Once the optimized bandwidth and center frequency obtained, the denoised signal will be processed by continuous wavelet transform based on Morlet wavelet.

5.2 Morlet wavelet

It is very important to select a proper mother wavelet function for wavelet analysis. The Morlet wavelet is selected as the mother wavelet to analyze the signal because of the

shape of the wavelet is similar to the impulse component. The function of the Morlet wavelet is defined as:

$$\psi(x) = \frac{1}{\sqrt{f_b \pi}} e^{-\frac{x^2}{f_b}} e^{j2\pi f_c x} \quad (5.1)$$

Its Fourier transform is:

$$\Phi(f) = e^{-\pi^2 f_b (f - f_c)^2} \quad (5.2)$$

where f_b is the bandwidth parameter, and f_c is the center frequency.

Eq. (5.1) and (5.2) show that the time-frequency resolution of the wavelet depends on the bandwidth f_b and center frequency f_c [108]. The larger the central frequency f_c is, the faster the Morlet wavelet will oscillate, which results in a lower frequency resolution as Figure 5-1 shown.

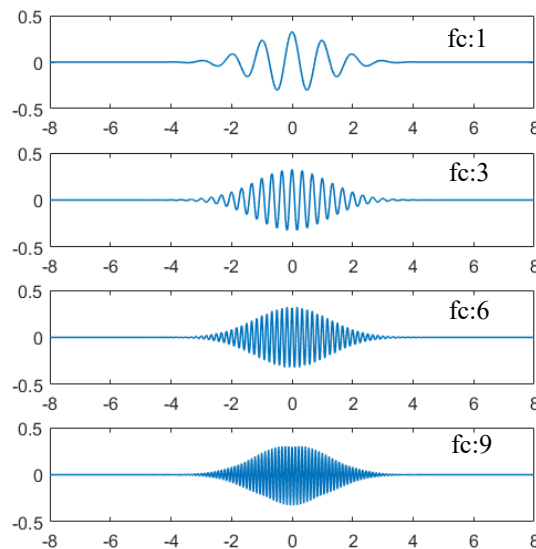


Figure 5-1. Morlet wavelet with different center frequency values (bandwidth is 3).

The bandwidth parameter controls the oscillation attenuation of the Morlet wavelet. A larger f_b value results in a better frequency resolution, at the cost of a lower time resolution as Figure 5-2 and 5-3 shown. Thus, to obtain the best time-frequency

resolution, the parameters need to be optimized.

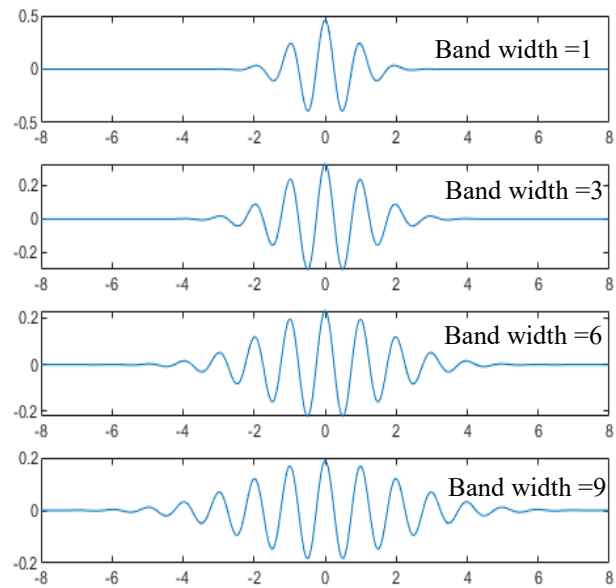


Figure 5-2. Effect of bandwidth parameter on Morlet wavelet in the time domain.

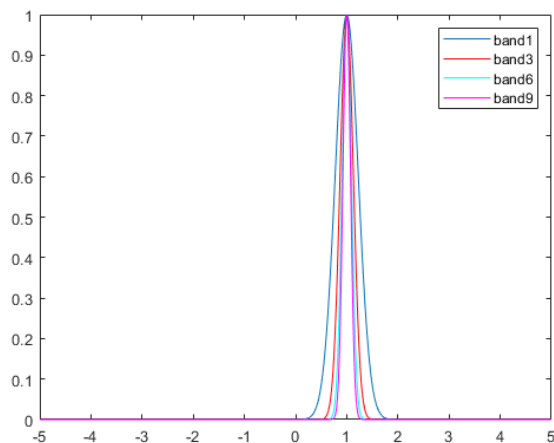


Figure 5-3. Effect of bandwidth parameter on Morlet wavelet in the frequency domain.

5.3 Parameter optimization

To select the optimal parameters of the wavelet, the “sparsity” is usually utilized as a rule to optimize the parameters. As we know, the wavelet entropy can measure the sparsity of a signal effectively. The wavelet entropy is defined as the energy-entropy of

the wavelet coefficients.

Jiang et al. [108] proposed an optimal method to select the parameters of the Morlet wavelet based on the modified Shannon wavelet entropy. The modified wavelet entropy is defined as:

$$En(f_b) = -\sum_{i=1}^M p_i^k \log p_i^k, \sum_{i=1}^M p_i^k = 1, f_b \in [N, M], f_c = k \in [A, B] \quad (5.3)$$

where p_i^k is calculated by:

$$p_i^k(f_b) = |W_c^k(a_i, b)| / \sum_{j=1}^M |W_c^k(a_j, b)| \quad (5.4)$$

Thus, the parameters of the Morlet wavelet can be optimized by the following steps. Firstly, an initial bandwidth $f_b \in [N, M]$ and center frequency $f_c \in [A, B]$ are given. To set the initial center frequency, a large initial bandwidth value is chosen. The Shannon entropy is calculated by increasing the center frequency from A to B . The initial center frequency is set based on the minimum wavelet entropy. Then, the bandwidth is increased from N to M , and the wavelet entropy $E_n(f_b)$ is calculated based on the initial center frequency. The optimized f_{ob} is selected based on the corresponding minimum wavelet entropy. By calculating the wavelet entropy by increasing the center frequency f_c from A to B with f_{ob} , the optimal f_c value is obtained based on the corresponding minimum wavelet entropy.

5.4 Experiment setup

Laboratory experiments were conducted with different fault defect sizes on the outer race to validate the proposed method. The experimental setup is shown in Figure 5-4. The test system consists of a 3-phase AC motor, a shaft supported by two rolling bearings, and an AC controller. The vibration signal is collected by an acceleration sensor that is mounted on the housing of the faulty bearing. The speed signal is collected

by a tachometer mounted on the end of the shaft. All the data are obtained by an acquisition card (National Instrument) with a sampling frequency of 12000 Hz.

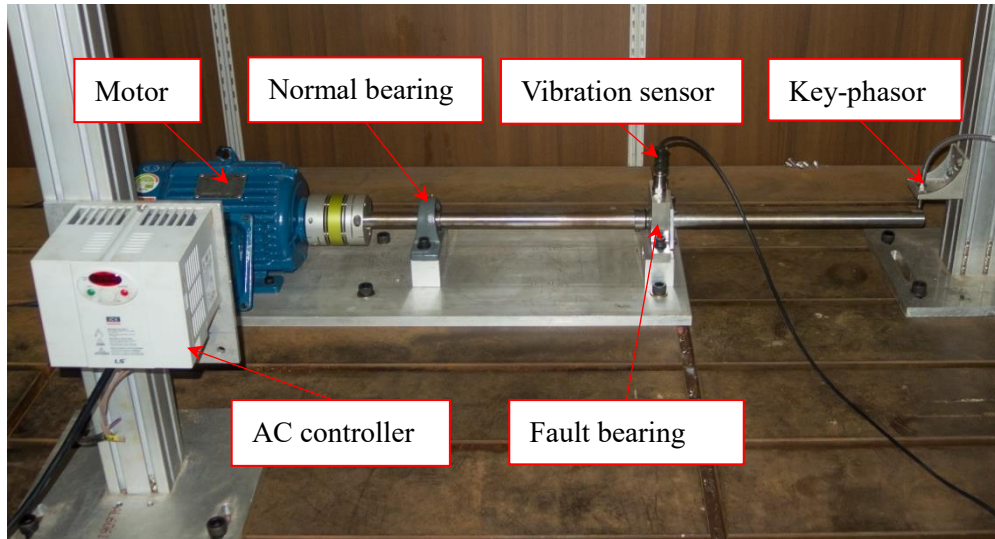


Figure 5-4. Rolling bearing test rig.

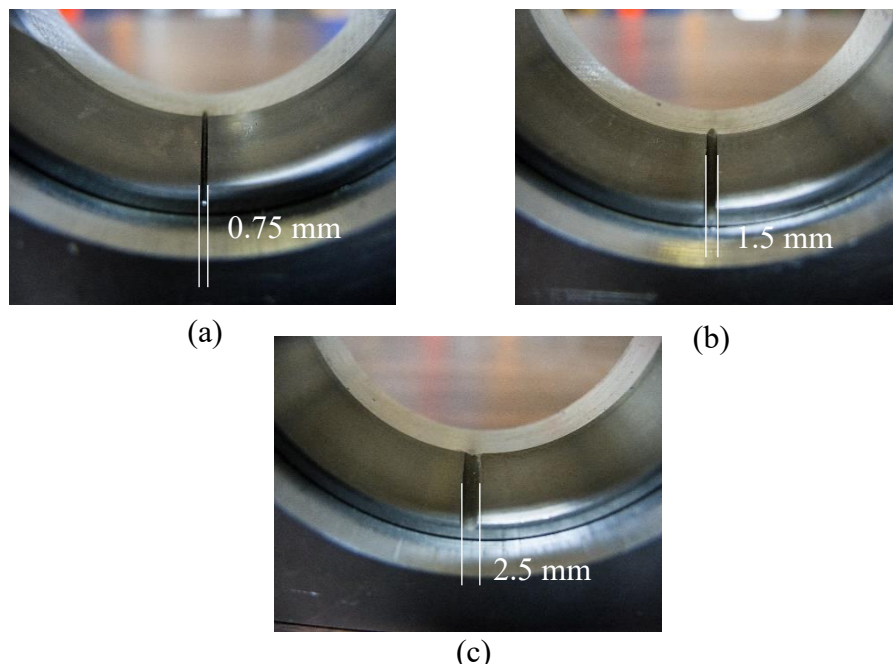


Figure 5-5. Different fault sizes on the outer races.

Three sets of the rolling element bearings (NTN 30206) were used to estimate the defect size on the outer races. Three different defect sizes on the outer race (0.75, 1.5, and 2.5 mm) were made by electric discharge machining (as Figure 5-5 shown). The parameters of the rolling element bearing are listed in Table 5-1.

Table 5-1. Geometrical parameters of the test bearing.

Parameter	Value
Number of rolling elements (N_b)	17
Contact angle (α)	14°02'10
Ball diameter (d_{ball})	8.2 mm
Pitch diameter (D_p)	45.6 mm
Outer race inner radius (R)	27.05 mm

5.5 Results and discussion

Figure 5-6 shows the original vibration signal and corresponding speed signal. The amplitude of the vibration signal increased with the shaft speed increased. The impact component is more obvious with the defect size increases. To remove the determined component, the edited cepstrum is used to process the vibration signals and the results are shown in Figure 5-7. The amplitude of the pre-whitened signal decreased after removing the determined component. However, the entry and exit points still could not be identified. It is because of the entry and exit events of the vibration signal are masked by the random noise.

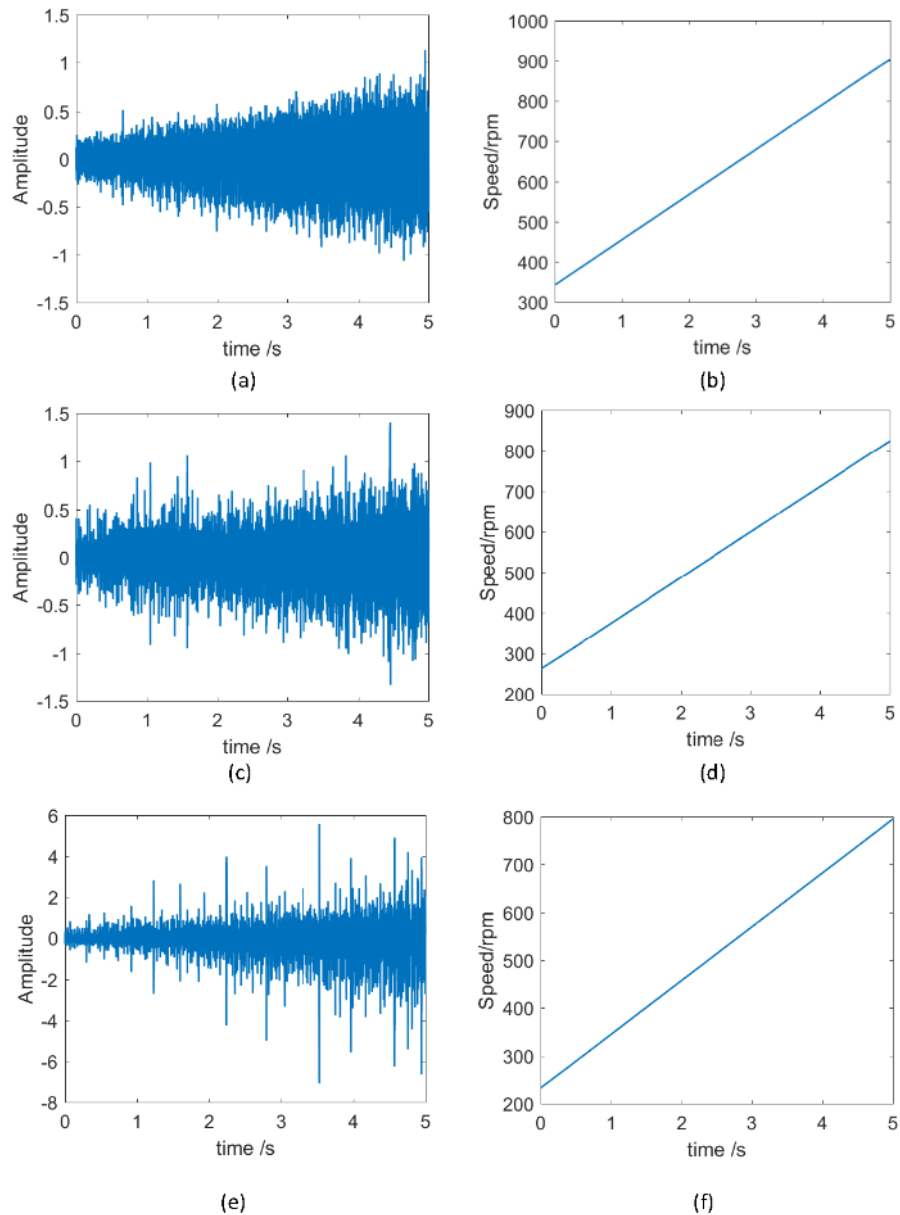


Figure 5-6. Vibration and speed signal of a bearing with different fault sizes on the outer race: (a, c, e) original vibration signal. (b, d, f) original speed signal.

To estimate the defect size of a bearing under time-varying speed conditions, the signal must be resampled from the time domain to the angle domain. Thus, the processed signal by the edited cepstrum method needs to be resampled to the angle domain before using the LMD method to extract the impact fault component.

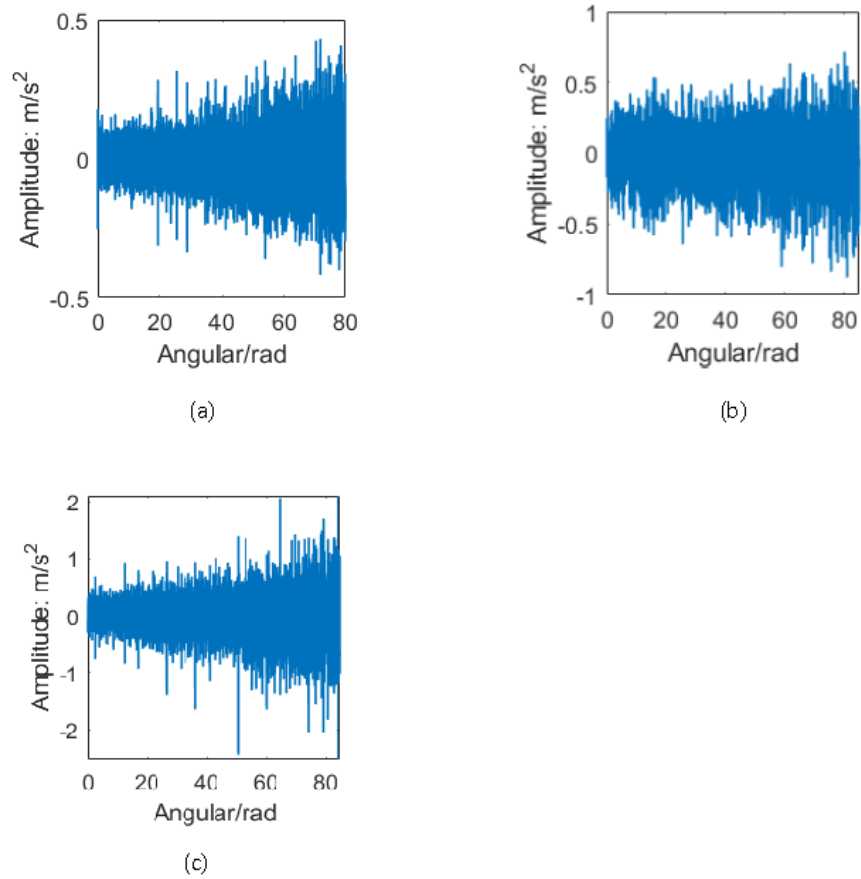


Figure 5-7. Pre-whitening vibration signal by edited cepstrum: (a) 0.75-mm. (b) 1.5-mm. (c) 2.5-mm.

During the resampling procedure, the signal is first upsampled by a factor of 10 to avoid aliasing. The speed information is obtained by a tachometer. Then the signal is resampled from the time domain to the angle domain. The LMD decomposition results of the resampled signal with different fault sizes are shown in Figure 5-8. After decomposing the vibration signals by LMD method, several PFs are obtained. The first PF is selected for further analysis because it has the biggest correlation coefficient value and keeps most of the information from the original signal. By utilizing the LMD signal denoising method, the response of entry event and exit event when rolling elements passed over the defect area on the outer race were enhanced.

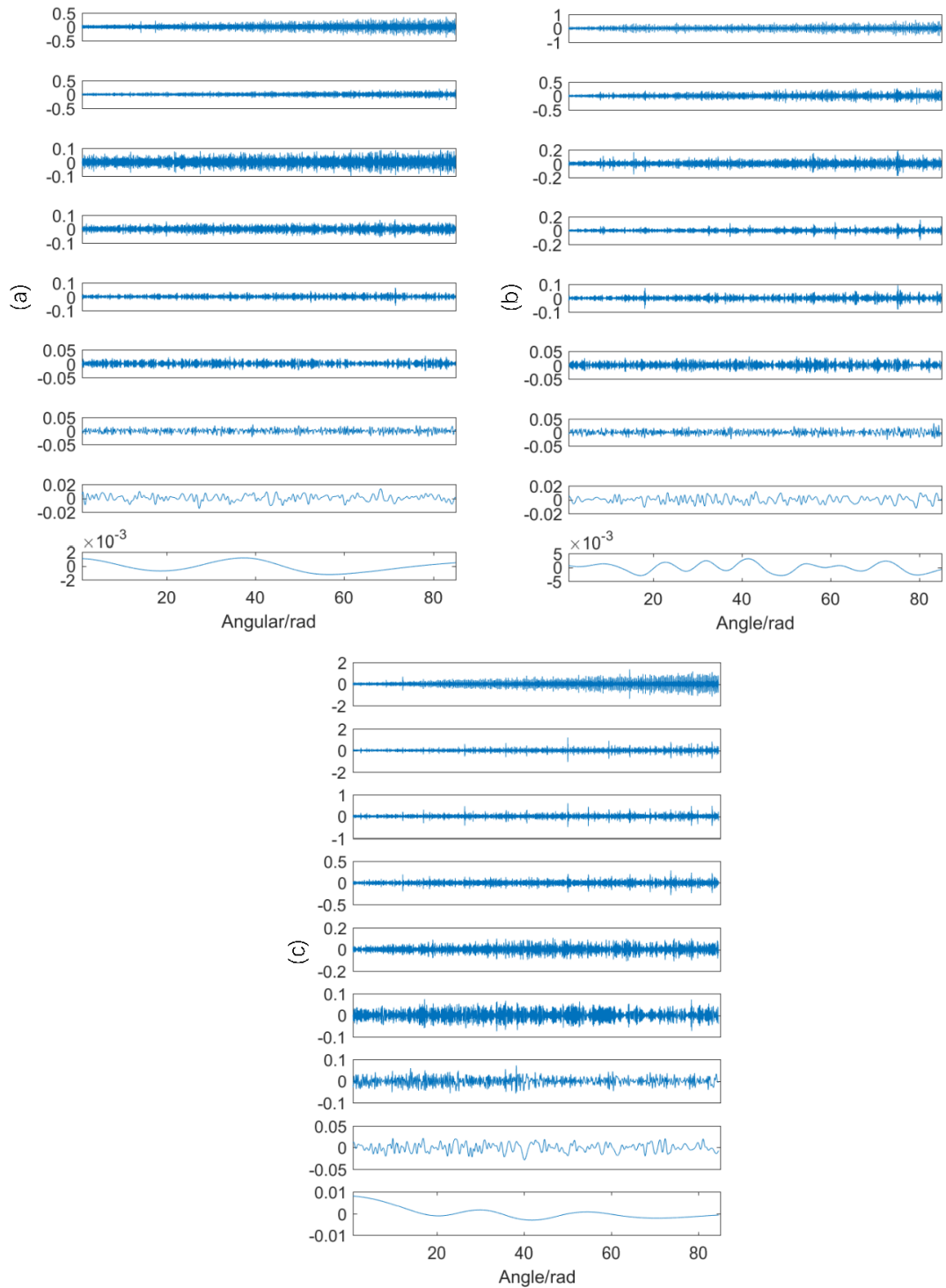


Figure 5-8. LMD decomposition result of the resampled signal with different fault sizes on outer race: (a) 0.75 mm, (b) 1.5 mm, (c) 2.5 mm

When using the continuous wavelet transform to process the denoised signal, the bandwidth and center frequency first need to be selected. The optimal parameters are selected based on the minimum Shannon wavelet entropy. The maximum order of this study is 400. The bandwidth is initialized as 400 to set the initial center frequency value. The initial bandwidth range is 1 to 30 with interval increment of 0.1, and the center frequency range is 0.1 to 5 with interval increment of 0.05. Figure 5-9 shows the initial center frequency. Figure 5-10 and 5-11 show the optimal parameters of Morlet wavelet for each defect size. Then, the CWT method is used to analyze the signal. At low speed (below 300 rpm), it is difficult to estimate the defect size due to the low signal energy and small response of the sensor in the low-frequency range [8]. Thus, the defect size of a defective bearing is measured above 300 rpm in this research.

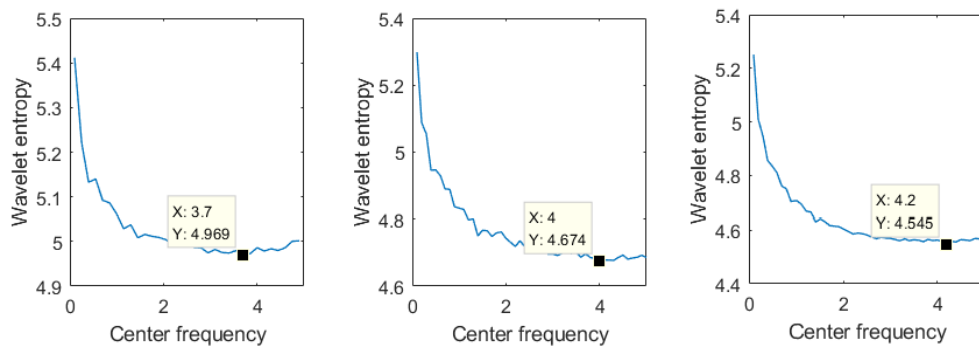


Figure 5-9. Relation between wavelet entropy and the center frequency with a bandwidth of 400: (a) 0.75 mm, (b) 1.5 mm, (c) 2.5 mm.

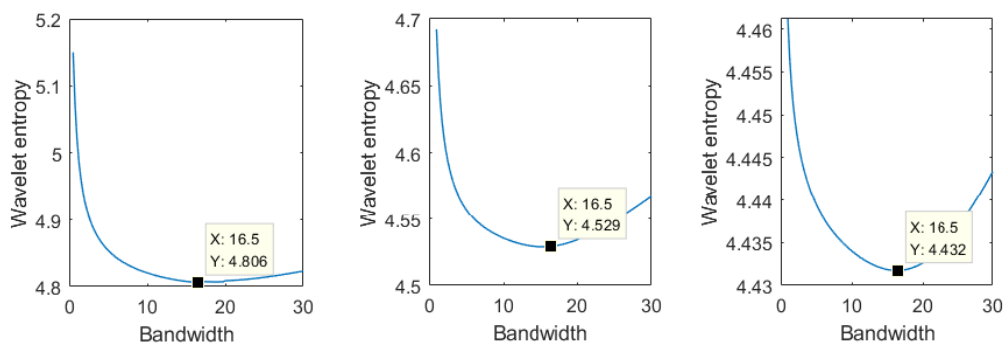


Figure 5-10. Relation between wavelet entropy and bandwidth parameters with the center frequency setting: (a) 0.75 mm, (b) 1.5 mm, (c) 2.5 mm.

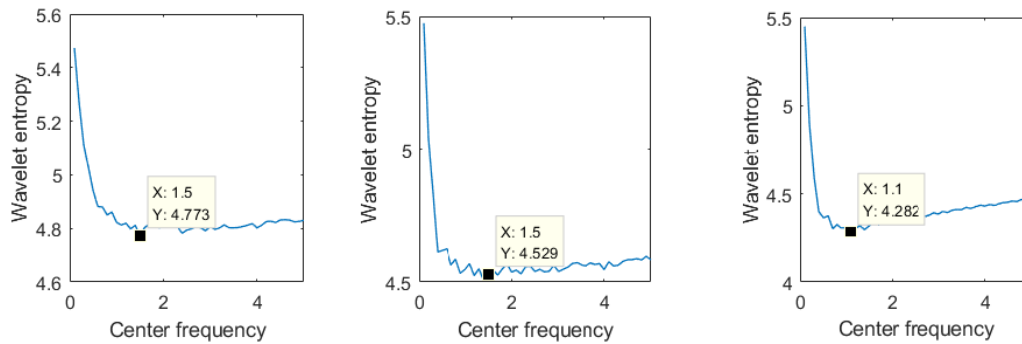


Figure 5-11. Relation between wavelet entropy and center frequency with optimal bandwidth: (a) 0.75 mm, (b) 1.5 mm, (c) 2.5 mm.

To better demonstrate the result by the proposed method, a portion of the signal was selected to show the estimation result. Figure 5-12 (a)-(c) show the scalogram analysis results of the CWT with different defect widths on the outer race and the details of the time-frequency analysis are shown in Figure 5-12 (d)-(f). The impact response of the fault rolling bearing is like a sharp spike. Jena et al. [7] illustrated how to measure the defect width by using the CWT scalogram. The high energy zone is generated when the roller hits the end edge of the defect. The portion data is selected around this high energy zone. When roller enters into the start edge of the defect area, the energy decreased as the de-stressing effect. When roller leaves the end edge of the defect, low energy zone will be generated as the roller re-stresses back to its normal load condition. Thus, the entry point can be identified at the starting point of the low energy zone (pre-side of high energy zone). The exit point can be identified at the end of a low energy zone (post side of high energy zone). Once the entry and exit points identified, the angle duration between the two points is obtained. The defect width is measured by the angle duration and the proposed estimation model.

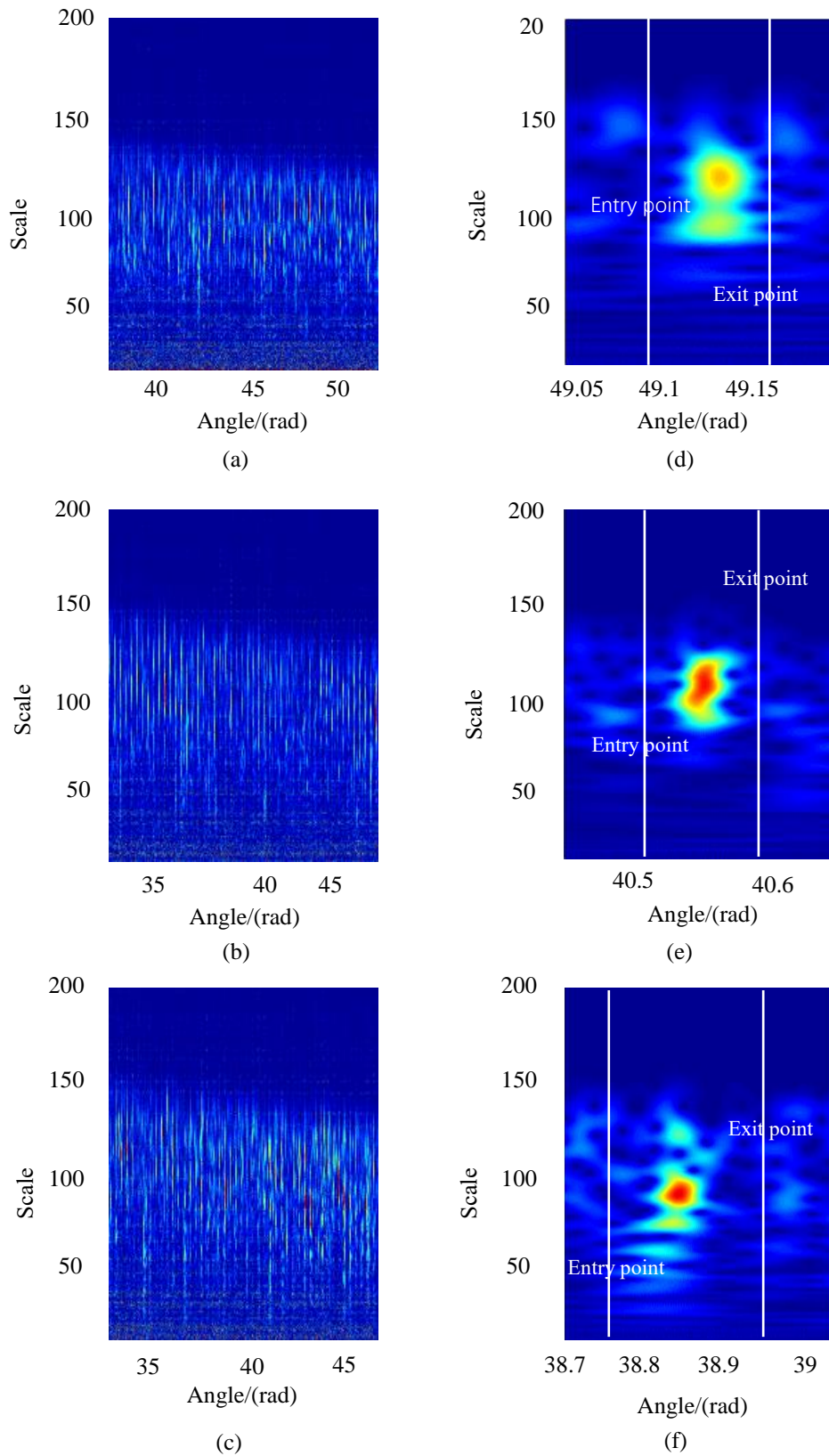


Figure 5-12. CWT analysis of the outer race fault: (a) CWT scalogram of 0.75-mm defect, (b) CWT scalogram of 1.5-mm defect, (c) CWT scalogram of 2.5-mm defect. (d-f) Detailed analysis of corresponding defect size.

10 portions of the CWT scalogram are used to measure the defect size of the defective bearing. Each of the portions contains one impact component. The estimation result is shown in Table 5-2. The mean estimation value and deviation are obtained from the scalogram. The deviation values for the different defect sizes are 0.078, 0.066, and 0.05. The results show that the defect size on the outer race can be measured with the proposed method.

Table 5-2. Estimation results of the defect size on the outer race.

Defects (mm)	Mean	Std. deviation
0.75	0.787	0.078
1.5	1.549	0.066
2.5	2.522	0.05

5.6 Conclusions

A novel method was proposed to estimate the defect size on the outer race of a rolling bearing under time-varying speed conditions. To enhance the entry and exit events, the edited cepstrum was used to remove the determined components. To eliminate the speed variation effect and estimate the defect size, the edited signal was resampled with a constant angle interval. The LMD method was used to extract the transient impulse component from the resampled signal. The CWT is used to provide full information about the energy distribution.

The entry and exit events of roller passing over the defect area were identified clearly by the CWT spectrum. The average deviation of the estimated defect size was 6.5%. The estimation results showed that the proposed method can effectively estimate the defect size of a defective bearing under time-varying speed conditions.

Chapter 6 Bearing Fault Identification

6.1 Introduction

In recently, artificial intelligence technique has been extensively applied in rolling element bearing fault diagnosis. The fault identification of a rolling element bearing is a pattern recognition task. Thus, the artificial intelligent technique is a good choice to solve this problem. Generally, an intelligent diagnosis system consists of four steps: data collection, feature extraction, data dimension reduction, and pattern classification [111]. This chapter will propose a new fault feature classification method to identify the fault type of a rolling element bearing under time-varying speed conditions.

6.2 Feature extraction

The fault patterns of a rolling element bearing on the time domain are diverse and complex. As an introduced previous chapter, RMS and kurtosis are commonly used to identify the bearing fault. However, these features are not appropriate for the time-varying speed conditions. Benefit from the development of nonlinear dynamic theory, many nonlinear feature extraction methods have been used to extract the fault features hidden in collected vibration signals that may not be effectively identified by using other existing methods. Such as approximate entropy (AEn), sample entropy (SEn) and fuzzy entropy (FE). Nevertheless, the shortage of the AEn is lower uniformly estimated value for the short data length. The sample entropy can get better performance than the AEn. But an unacceptable result appeared on the real application. The fuzzy entropy used the fuzzy membership function instead of Heaviside function. A continuous value is obtained compared with the AEn and sample entropy. The fuzzy entropy has been

successfully used as a feature to identify the rolling bearing fault types and improved by many researchers. The multiscale fuzzy entropy (MFE) was developed to check the self-similarity of original data. Thus, the MFE method is utilized to extract the fault features from the pre-processing signal.

6.2.1 Fuzzy entropy

The fuzzy entropy value can be obtained by the following steps:

- 1) Construct an m dimension vector for a given time series $\{u(i):1 \leq i \leq N\}$ at time i which is normalized, and its length is N .

$$X_i^m = \{u(i), u(i+1), \dots, u(i+m-1)\} - u_0(i), i = 1, 2, \dots, N - m + 1. \quad (6.1)$$

Where $u_0(i)$ is mean value of the m consecutive $u(i)$ values.

$$u_0(i) = \frac{1}{m} \sum_{k=0}^{m-1} u(i+k) \quad (6.2)$$

- 2) Find the maximum distance between X_i^m and components X_j^m ,

$$\begin{aligned} d_{ij}^m &= d[X_i^m, X_j^m] \\ &= \max_{0 \leq k \leq m-1} \{|[u(i+k) - u_0(i)] - [u(j+k) - u_0(j)]|\}, i, j = 1, 2, \dots, N - m, i \neq j. \end{aligned} \quad (6.3)$$

where d_{ij}^m is the maximum distance vector.

- 3) Obtain the similarity vector D_{ij}^m by the fuzzy function $\mu(d_{ij}^m, n, r)$, it can be calculated as the following formula:

$$D_{ij}^m = \mu(d_{ij}^m, n, r) = e^{-\ln 2 (d_{ij}^m / r)^n} \quad (6.4)$$

where, n is the gradient and r is the parameter of the boundary.

- 4) Then the ϕ^m function can be obtained by the following:

$$\varphi^m(n, r) = \frac{1}{N-m} \sum_{i=1}^{N-m} \left(\frac{1}{N-m-1} \sum_{\substack{j=1 \\ j \neq i}}^{N-m} D_{ij}^m \right) \quad (6.5)$$

$m+1$ dimensional vector can be obtained by repeating the steps 1-4:

$$\varphi^{m+1}(n, r) = \frac{1}{N-m} \sum_{i=1}^{N-m} \left(\frac{1}{N-m-1} \sum_{\substack{j=1 \\ j \neq i}}^{N-m} D_{ij}^{m+1} \right) \quad (6.6)$$

5) Calculate the fuzzy entropy of the time series with φ^m and φ^{m+1} :

$$FuEn(m, n, r) = \lim_{N \rightarrow \infty} \left[\ln \varphi^m(n, r) - \ln \varphi^{m+1}(n, r) \right] \quad (6.7)$$

If the length of the sequence is finite, $FuEn(m, n, r)$ can be estimated as:

$$FuEn(m, n, r) = \ln \varphi^m(n, r) - \ln \varphi^{m+1}(n, r) \quad (6.8)$$

6.2.2 Multi-scale fuzzy entropy

As a single scale entropy may cause the ambiguous results in the real application, multi-scale entropy is proposed to measure the complexity of time sequences [112]. It can measure the complexity of signals by different scales. As this advantage, the multi-scale fuzzy entropy is developed to extract the features. The MFE is calculated based on the fuzzy entropy and the details are described as the following steps [113]:

1) Construct the coarse-grained vector $\{y_i^\tau\}$ for a given time sequence

$\{u(i) : 1 \leq i \leq N\}$. The sequences are separated by a window of length τ .

$\{y_i^\tau\}$ can be obtained according to the following equation:

$$y_j^\tau = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} u(i), 1 \leq j \leq N/\tau \quad (6.9)$$

where τ is the scale factor.

- 2) Calculate the *FuEn* of each coarse-grained time sequences with different scale factor τ based on equation 6.1-6.8, then the MFE values can be obtained:

$$MFE(u, \tau, m, r) = FuEn(y_j^\tau, m, r) \quad (6.10)$$

Note that the r is calculated by the $r = \lambda * SD$ and SD is the standard deviation of $u(t)$. To calculate the MFE values, it needs to select appropriate parameters. The embedding dimension m is usual 2. It is because of a large m needs a large N and will lead to losing the information of the signal. If the gradient of the boundary r was too narrow, the noise effect is enhanced. On the contrast, if the r was too broad, it could keep most of the information. It is suggested to set the r by 0.1-0.25*SD of the original signal. The boundary gradient of the exponential function n is convenient to fix to 2. A large τ will lead to heavy computation. However, a small τ will lead to losing the information of the signal.

In this research, the parameters of the MFE are selected as the followings:

$$\tau=20, \text{ embedding dimension: } m=2, r=0.15 * SD, n=2$$

6.3 Laplacian Score for feature selection

Theoretically, the fault patterns recognition can be accomplished by the extracted MFE features in 20 scales. However, a feature vector in 20 scales is a high dimension. It will enhance the time-consuming, and even may lead to the information inefficient for fault identification of the rolling bearing. Thus, how to select the most interrelated fault information from the 20 features is very important. By this way, it avoids the dimension disaster and improves the performance of classification and the efficiency of rolling bearing automatically fault diagnosis. In this research, the Laplacian Score (LS) algorithm is utilized to reduce the feature vectors dimension via their importance and distinguishability. LS selects the feature vectors based on Laplacian Eigenmaps and

Locality Preserving Projection. Its basic idea is to estimate the features according to their locality preserving power.

A given data which the number of groups is m and each group data has n features. Assume that L_r represents the Laplacian Score of the r th feature, $r = 1, \dots, n$. Let f_{ri} represents the i th sample of the r th feature, $i = 1, \dots, m$. The main calculation procedures of the LS algorithm can be written as follows [113]:

1. Construct the nearest neighbor graph G with m nodes, where the i th node corresponds to x_i , and the edge is put between nodes i and j , if x_i and x_j are “close”. When the label information is available, one can put an edge between two nodes sharing the same label.

2. Define the weight matrix as the following equation:

$$S_{ij} = \begin{cases} e^{-\frac{|x_i - x_j|}{t}} & \text{if nodes } i \text{ and } j \text{ are connected} \\ 0 & \text{otherwise} \end{cases} \quad (6.11)$$

3. Calculate graph Laplacian L .

$$L = D - S, D = \text{diag}(SI) \quad (6.12)$$

4. The Laplacian Score of the r th feature can be written as follows:

$$L_r = \frac{\sum_{ij} (f_{ri} - f_{rj})^2 S_{ij}}{\text{Var}(f_r)} = \frac{\tilde{f}_r^T L \tilde{f}_r}{\tilde{f}_r^T D \tilde{f}_r} \quad (6.13)$$

where $f_r = [f_{r1}, f_{r2}, \dots, f_{rm}]$, $\text{Var}(f_r)$ is the estimated variance of the r th feature, \tilde{f}_r is the matrix that can be obtained by removing the mean from the samples.

$$\tilde{f}_r = f_r - \frac{f_r^T D I}{I^T D I} I \quad (6.14)$$

Based on the equation 6-13, it can be seen that the bigger S_{ij} value, smaller the Laplacian Score value, which represents the characteristic of the good feature.

Therefore, the dimension of the feature can be reduced by using the LS method. It not only enhances the identification efficiency but also provides feature vector information.

6.4 Pattern classification with Support Vector Machines (SVM)

Many intelligence fault diagnosis methods of the rolling element bearing have been proposed. Such as neural network, fuzzy logic, immune genetic [55, 114-116]. However, some limitations exist among these methods, for example, slow convergence velocity and over-fitting. Particularly, the fault data is not easy to obtain. The size of the samples is small. Thus, it may lead to a bad result. It is a bottleneck when using the neural network to identify the fault types. Comparing with these methods, support vector machine achieves good performance as it is not influenced by these limiting factors.

The support vector machine is a machine learning algorithm based on the structured risk minimization principle. The advantages of SVM are great flexibility, dimension insensitivity, shorter training time and good generalization performance. The goal of SVM is to find a hyperplane to separate different classes. The hyperplane is a boundary which makes the maximum distance between the boundary and the closest data points of each class. Once the optimal hyperplane obtained, the maximum margin is defined. Furthermore, the samples in all classes closest to this margin are called support vectors. Then, the classifier is established by the information of the support vectors.

Given a sample data (x_i, y_i) , with $i = 1, \dots, M$, x_i is the input sample, M is the number of samples, and y_i is the label of x_i . For a linear case, the hyperplane can be expressed as

$$f(x) = \omega \cdot x + b = 0 \quad (6.15)$$

where ω is the normal direction of the hyperplane, b is a parameter of the

hyperplane position.

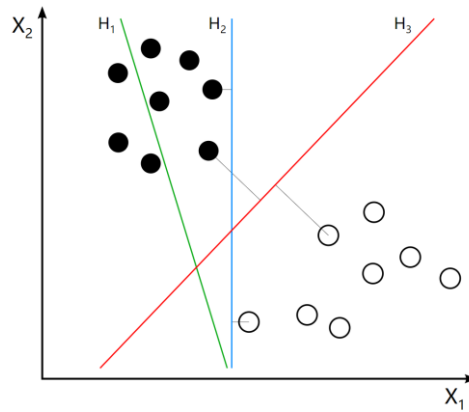


Figure 6-1. Find the maximum margin of the two classes.

The input data should be met with the following constraints

$$f(x_i) = \begin{cases} \geq 1, & \text{for } y_i = 1 \\ \leq -1, & \text{for } y_i = -1 \end{cases} \quad (6.16)$$

as a complete equation

$$y_i(\omega \cdot x_i + b) \geq 1 \quad i = 1, \dots, M \quad (6.17)$$

The nearest samples lie instead on the planes $f(x) = \pm 1$ and so the margin distance between them is $2/\|\omega\|$. Thus, the pair of hyperplanes can be obtained by finding the maximum margin by minimizing $\|\omega\|^2$, subject to constraints 6.17. The optimization problem is:

$$\begin{aligned} \min_{\omega, b, \gamma} \quad & \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^M \gamma_i \\ \text{subject to} \quad & \begin{cases} y_i(\omega \cdot x_i + b) \geq 1 - \gamma_i \\ \gamma_i \geq 0 \end{cases} \end{aligned} \quad (6.18)$$

where γ_i is positive slack variable, which is necessary to allow misclassification. The objective function will produce the classification error, so a generalization parameter C is presented. The greater generalization parameter, the higher misclassification error will be, as well as the heavier of the punishment.

To simplify this problem with the Kuhn-Tucker condition, the positive Lagrange multipliers α_i is introduced. For equality constraints, the Lagrange multipliers are unconstrained [117]. This gives Lagrangian:

$$L_p(\omega, b, \alpha) = \frac{1}{2} \|\omega\|^2 - \sum_{i=1}^M \alpha_i (y_i (\omega \cdot x_i + b) - 1) \quad (6.19)$$

To obtain a stationary point of $L_p(\omega, b, \alpha)$ over ω and b , the partial derivatives are applied to ω and b , respectively

$$\frac{\partial L}{\partial \omega} = 0, \frac{\partial L}{\partial b} = 0 \quad (6.20)$$

Combine equation (6.20) with (6.19),

$$\omega = \sum_{i=1}^M \alpha_i y_i x_i, 0 = \sum_{i=1}^M \alpha_i y_i \quad (6.21)$$

Substituting equation (6.21) into equation (6.19), the dual formulation is obtained:

$$\text{Max } L_D(\alpha) = \sum_{i=1}^M \alpha_i - \frac{1}{2} \sum_{i=0}^M \sum_{j=0}^M \alpha_i \alpha_j y_i y_j x_i x_j \quad (6.22)$$

Takin the gradient of $L_D(\alpha)$ with respect to α_i , the optimal α^* is obtained.

Then ω and b are calculated by α^* as follows:

$$\begin{aligned} \omega^* &= \sum_{i=1}^M \alpha_i^* y_i x_i \\ b^* &= y_i - \sum_{i=1}^M \alpha_i^* y_i x_i x_i \end{aligned} \quad (6.23)$$

With the parameters ω^* , α^* and b^* , the decision function is expressed as:

$$f(x) = \text{sign}(\omega^* x + b^*) \quad (6.24)$$

For the non-linear classification case, the input data firstly are mapped onto a high-dimensional feature space. Thus, the different classes can be classified by using the linear classification.

There is a class of functions $K(x, x_j)$ with the following property. There is a linear space S and a function φ mapping x to S such that:

- **Linear:**

$$K(x, x_j) = \langle \varphi(x), \varphi(x_j) \rangle \quad (6.25)$$

- **Polynomials:** For some positive integer p ,

$$K(x, x_j) = (1 + x'x_j)^p \quad (6.26)$$

- **Gaussian Radial basis function (RBF):**

$$K(x, x_j) = \exp(-\|x - x_j\|^2) \quad (6.27)$$

6.5 Experimental validation

The purpose of this section is to examine the utility of the proposed algorithm for analyzing the real rolling bearing vibration data. The vibration signals are obtained from the Bearing Data Center of Case Western Reserve University [118].

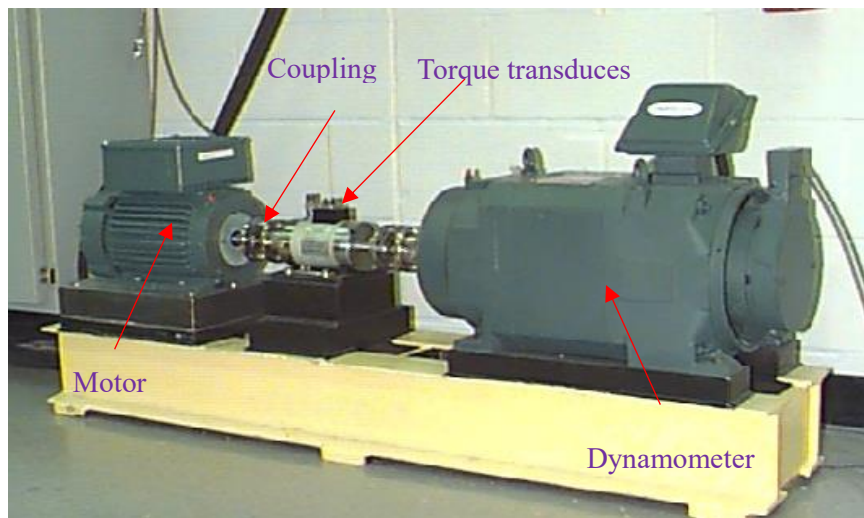


Figure 6-2. Rolling bearing experiment system.

The experiment system was shown by Figure 6-2. The type of the tested rolling bearings is 6205-2RS JEM SKF deep groove ball bearing. Vibration data was collected

by an accelerometer, which was attached to the housing with magnetic bases. The sampling frequency is 12000 Hz and the shaft rotating speed of the motor changes from 1730 to 1797 RPM. The dataset of vibration signal consisted of the normal condition and fault working condition. The fault condition contains the inner race fault, outer race fault, ball fault. For each fault type, single point faults were introduced to the test bearings using electro-discharge machining with different fault sizes of 7, 14, 21, 28 mils.

In this thesis, it is an eight-class recognition problem according to the different fault categories. For each group, there are 50 samples and a total number of the samples is 400. Out of which 160 samples are randomly selected to train the SVM classifier and the residual 240 samples are used for testing. The detailed numbers of samples description for each condition are shown in Table 6-1. The whole steps of the fault identification are presented by Figure 6-3.

Table 6-1. Details of experimental data sets.

Fault class	Fault size (inch)	Fault severity	Number of training data	Number of test data	Class label
IRF	0.007'	Slight	20	30	1
(Inner race fault)	0.014'	Medium	20	30	2
	0.021'	Severe	20	30	3
BF	0.007'	Slight	20	30	4
(Ball fault)	0.028'	Very severe	20	30	5
ORF	0.007"	Slight	20	30	6
(Outer race fault)	0.014"	Medium	20	30	7
Normal	0	Normal	20	30	8

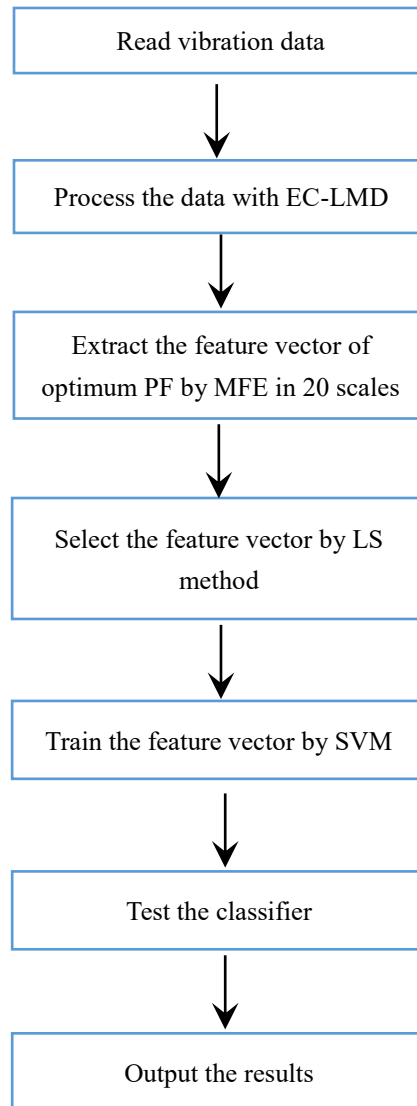


Figure 6-3. Flow chart of the proposed fault diagnosis algorithm.

As the collected vibration data under different fault patterns often have different characteristics, it is hard to identify the fault categories from each other only by the time domain waveforms as shown in Figure 6-4. To obtain the good feature vectors, vibration signals are processed by the proposed EC-LMD method.

MFE was utilized to extract the fault features from the selected PF component under 20 scales for each bearing condition. It means the dimension of the obtained feature space was 20 in the beginning analysis. The MFE values by 20 scales of the data

were shown by Figure 6-5. It was not easy to identify the fault types clearly by all the scales of MFE. This is because feature values are mixed at some scales.

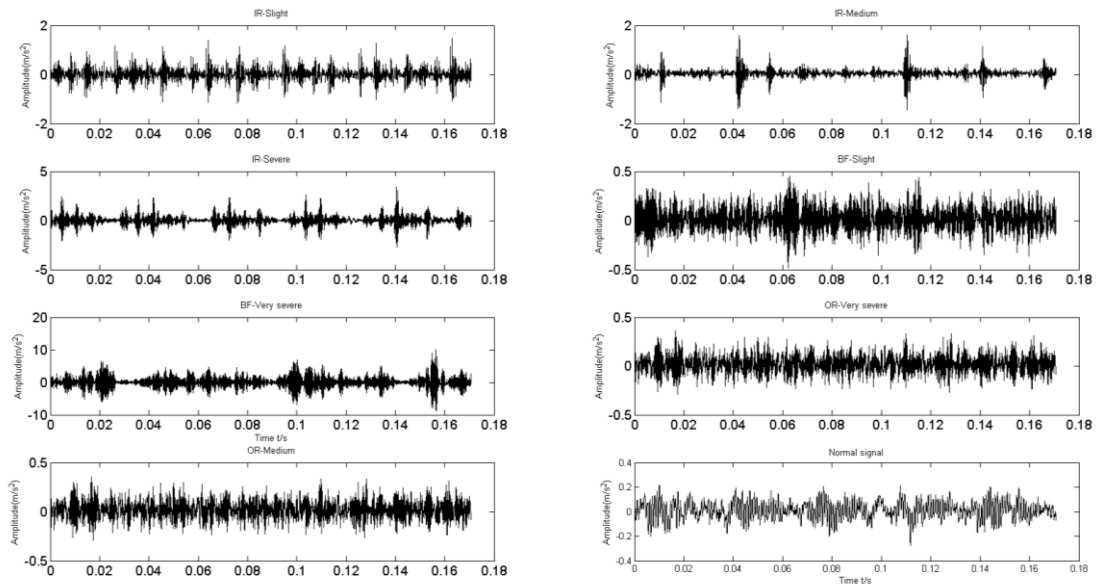


Figure 6-4. vibration acceleration signal of each rolling bearing condition.

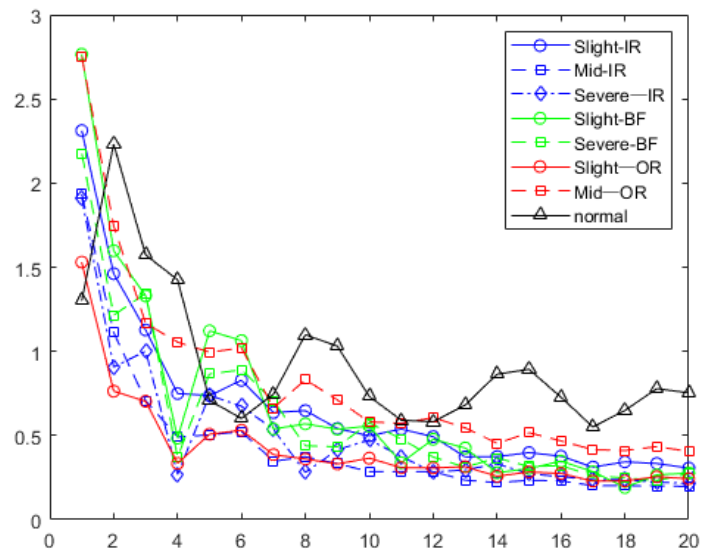


Figure 6-5. MFE over 20 scales of the pre-processing signal with the average of fifty trails.

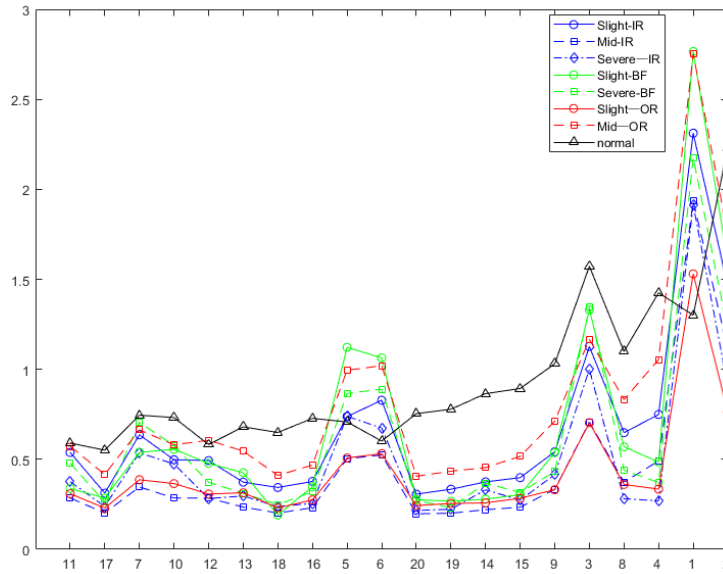


Figure 6-6. LS result of MPE over 20 scales.

If MFE over 20 scales was all taken as the feature vector, it would be time-consuming and reduce the classification accuracy rate. The new order of MFE by LS selection was listed as follows:

$$LS_{11} < LS_{17} < LS_7 < LS_{10} < LS_{12} < LS_{13} < LS_{18} < LS_{16} < LS_5 < LS_6 < LS_{20} < LS_{19} < LS_{14} < LS_{15} < LS_9 < LS_3 < LS_8 \\ < LS_4 < LS_1 < LS_2.$$

Note that the subscript of LS is the scale factors and the MFE values were ranked and replotted in Figure 6-6. Then the first front four features ($\tau=11,17,7,10$) with most important information were chosen to construct a new feature vector. Naturally, the new feature vector was fed into the classifier SVM to obtain the model of the pattern recognition. When using the SVM to train the data, the training samples need to be mapped from the nonlinear space to linear by the kernel function. The Gaussian kernel function gives a special solution by transforming the nonlinear space to the linear conditions. Thus, it was selected as the kernel function in this research. At the same time, 160 samples were selected randomly from the whole dataset as the training data and the residual 240 samples as the testing data.

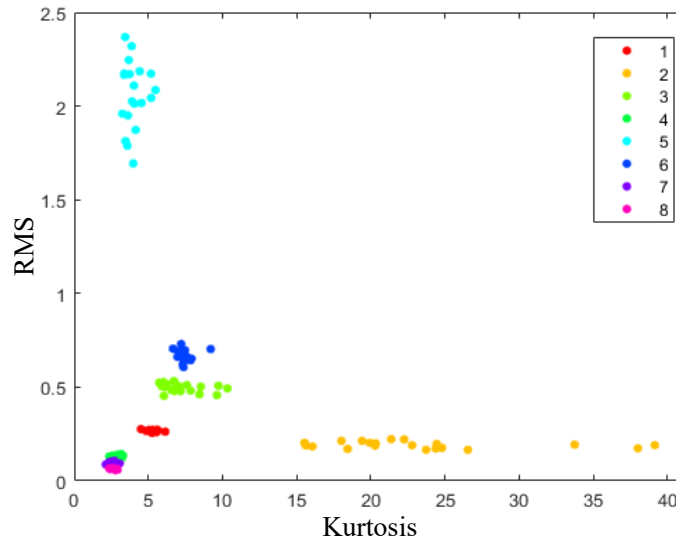


Figure 6-7. Distribution of RMS and kurtosis features (Legends of 1-8 are corresponding class labels of Table 6-1).

In order to illustrate the advantage of the proposed method, the RMS and kurtosis were used to construct the feature vector. The signal was firstly pre-processed by the EC-LMD to extract the faulty component. Figure 6-7 showed the distribution of the RMS and kurtosis features. There was a mixed area at type 4, type 7 and type 8. As a result, it's hard to identify these three fault types.

The sample distribution of MFE over 20 scales was shown by Figure 6-8. The original MFE over 20 scales was too high, the recognition rate may decrease with high dimension. To enhance the identification efficiency, the important scales were selected by the LS method. The feature selection result was shown in Figure 6-9. Compared with the original MFE, the selected features made it easy to identify the fault types.

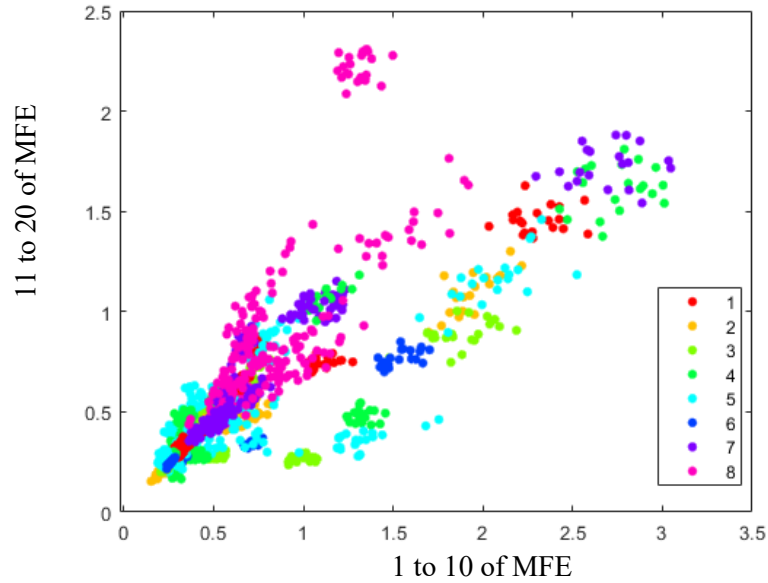


Figure 6-8. Distribution of MFE features over 20 (Legends of 1-8 are correspond class labels of Table 6-1).

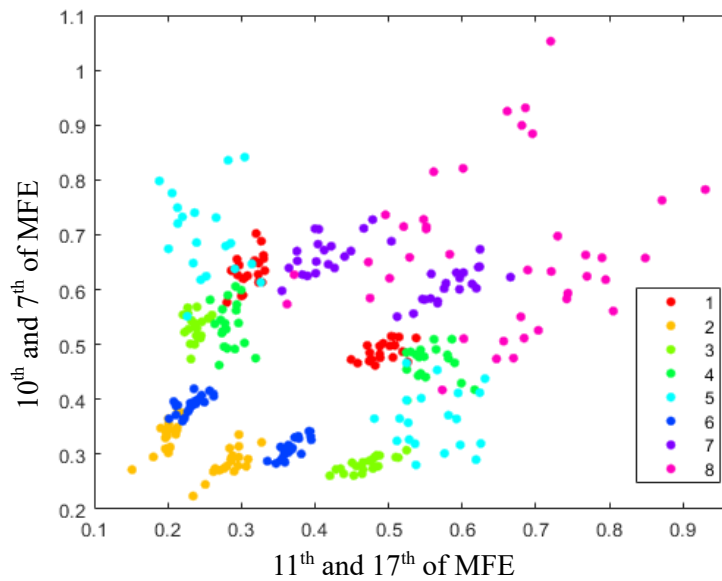


Figure 6-9. Distribution of selected MFE features by LS (Legends of 1-8 are corresponding class labels of Table 6-1).

The SVM was used to train the data for obtaining a model to identify the fault types based on the extracted features. After obtaining the model, the test data was used to validate its effectiveness. First, the feature vector of the RMS and kurtosis were fed into SVM to accomplish the pattern recognition. The SVM outputs of the testing data

were shown in Figure 6-10. There were 3 test samples misclassified. The classification accuracy is 98.75%. By the result of the classification analysis, the RMS and kurtosis couldn't provide enough fault information to identify the fault types. It was because the RMS and kurtosis were affected by the speed. They are insufficient to illustrate the dynamic variation of the system and result in lower identification accuracy. Therefore, it was essential to analyze the vibration signal by the MFE method.

The classification results of the proposed method were shown by Figure 6-11. There was only one testing sample misclassified. The average recognition accuracy achieves to 99.58%. The comparison results show that the proposed method achieves a good classification result, which is appropriate and effective in fault diagnosis of the rolling element bearing.

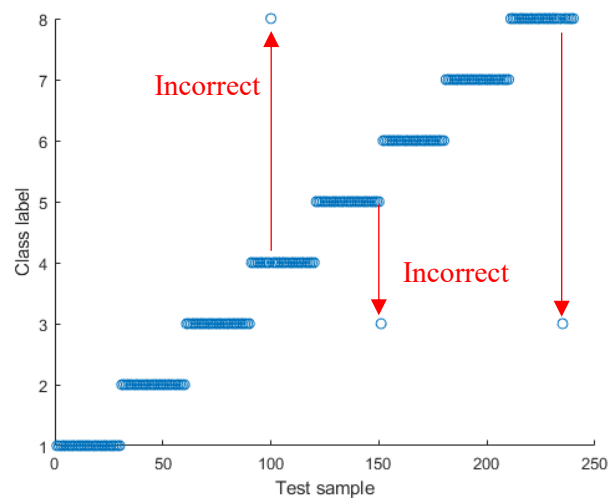


Figure 6-10. Classification results of SVM using RMS and kurtosis feature.

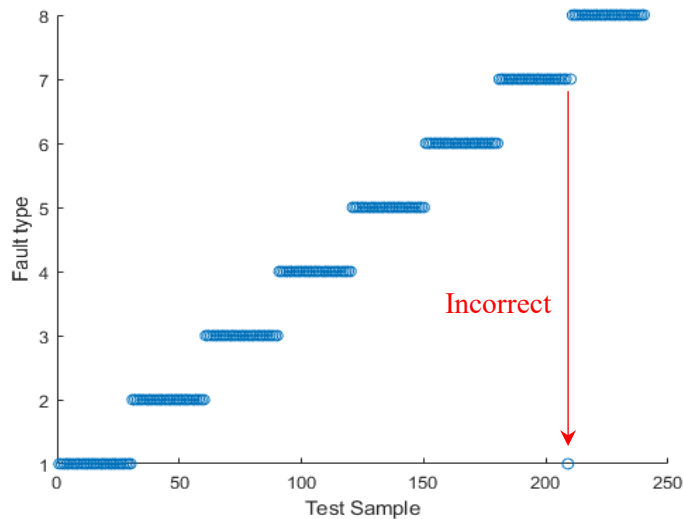


Figure 6-11. Classification results of SVM using selected MFE.

6.6 Conclusions

An intelligent rolling bearing fault diagnosis method was proposed to identify the fault types. The EC-LMD method was used to extract the fault component from the vibration signal. The feature vector was obtained by the MFE. MFE could effectively characterize the complexity of the pre-processed signal. Laplacian score was used to select the fault feature by reordering the scale factors. SVM is used to evaluate the classification performance. As the results shown, the fault types were recognized effectively by the proposed method.

The comparison results show that the proposed method achieves a good classification result, which is appropriate and effective in fault diagnosis of the rolling element bearing.

Chapter 7 Conclusions

This research work mainly focuses on the fault diagnosis and defect size estimation of rolling element bearing under time-varying speed conditions. The fault diagnosis and defect size estimation of the rolling bearing are affected by the speed variation. The common defect size estimation methods are based on the time domain and the constant speed condition. However, for the time-varying speed conditions, these methods are not suitable as the shaft speed varies. If the angle duration between the entry and exit points of the defect known, the defect size could be estimated by the geometric parameters of the rolling bearing. Therefore, to measure the defect size of a rolling element bearing under time-varying conditions, the signal needs to be resampled from the time domain to the angle domain.

To better understand the vibration response of a defective bearing, a dynamic model is established. The variations of contact deformation and force when the roller passed over the defect area were illustrated by the analysis result. The entry and exit events are identified by these illustrations. Then, two defect size estimation models are developed for the time-varying speed conditions.

A vibration signal of a defective bearing based on time-varying speed conditions consists of a fault component, determined component and random noise component. The cepstrum pre-whitening (CPW) technique is used to separate the determined component from the vibration signal. The determined component is eliminated by removing the peaks on the real cepstrum.

To extract the impact component, LMD method is used. To solve the end effect, an improved LMD is proposed by the DTW method. The DTW found an optimal wrapped path to extend the signal. Simulation result shows that the improved method can eliminate the end effect effectively. Combining the edited cepstrum with improved

LMD, the entry and exit events are enhanced. An experiment, an outer race fault in a bearing, is used to validate the proposed method. Compared with the fast kurtogram, the proposed method obtained a better result for detecting the bearing fault.

To identify the entry and exit events exactly, the energy distribution of the CWT coefficients is used. The optimal parameters of Morlet wavelet are selected by the wavelet entropy. The high energy zone is generated when the roller hit the end edge of the defect. When the roller enters into the start edge of the defect area, the energy decreased as the de-stressing effect. When roller leaves the end edge of the defect, low energy zone will be generated as the roller re-stresses back to its normal load condition. Thus, the entry point was identified at the starting point of a low energy zone (pre-side of high energy zone). The exit point was identified at the end of a low energy zone (post side of high energy zone). The experimental results showed that the entry and exit events of roller passed over the defect on the outer race were identified clearly from the CWT spectrum. The average deviation of the estimated defect size was 6.5%. The estimation results showed that the proposed method could effectively estimate the defect size on the outer race under time-varying speed conditions.

An intelligent rolling bearing fault diagnosis method was proposed. The EC-LMD method was used to pre-process the signal for extracting good features. The MFE was taken as the feature vectors. Laplacian score was used to select the fault feature by reordering the scale factors. SVM is used to evaluate the classification performance. As the results shown, the fault types were recognized effectively by the proposed method.

Future work will focus on follows:

1. obtaining defect size automatically to predict the remaining useful lifetime of rolling element bearing. This may be realized by the image processing techniques.
2. The load variation effect may be taken consideration to estimate the defect size.

As the result shows that the deformation is affected by the applying load. It should be considered when estimating the defect size.

3. Conduct the minimum defect size that the proposed method can estimate. It decides whether the proposed method can be used on the real application or not.
4. Analyze the effect of distributed defects. Since generalized roughness defects have distributed small defect with fluctuating sizes, it needs to conduct the model that can estimate defect size both small and large line spall.

REFERENCES

1. Ming, A.B., et al. *Spalling Size Evaluation of Rolling Element Bearing Using Acoustic Emission*. 2013. Trans Tech Publ.
2. Sawalhi, N., R.B. Randall, and 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*, 2007. **21**(6): p. 2616-2633.
3. Epps, I.K., *An investigation into vibrations excited by discrete faults in rolling element bearings*. 1991.
4. Sawalhi, N. and R.B. Randall, *Vibration response of spalled rolling element bearings: Observations, simulations and signal processing techniques to track the spall size*. *Mechanical Systems and Signal Processing*, 2011. **25**(3): p. 846-870.
5. Sawalhi, N., W. Wang, and A. Becker, *Vibration signal processing for spall size estimation in rolling element bearings using autoregressive inverse filtration combined with bearing signal synchronous averaging*. *Advances in Mechanical Engineering*, 2017. **9**(5): p. 1687814017703007.
6. Ismail, M. and N. Sawalhi. *Bearing spall size quantification based on geometric interpretation of vibration envelope energy*. in *13th international conference on condition monitoring and machinery failure prevention technologies, Paris*. 2016.
7. Jena, D.P. and S.N. Panigrahi, *Precise measurement of defect width in tapered roller bearing using vibration signal*. *Measurement*, 2014. **55**: p. 39-50
8. Moustafa, W., et al., *Low speed bearings fault detection and size estimation using instantaneous angular speed*. *Journal of Vibration and Control*, 2014. **22**(15): p. 3413-3425.
9. Khanam, S., N. Tandon, and J.K. Dutt, *Fault Size Estimation in the Outer Race of Ball Bearing Using Discrete Wavelet Transform of the Vibration Signal*. *Procedia Technology*, 2014. **14**(Supplement C): p. 12-19.
10. Wang, W., N. Sawalhi, and A. Becker, *Size estimation for naturally occurring bearing faults using synchronous averaging of vibration signals*. *Journal of Vibration and Acoustics*, 2016. **138**(5): p. 051015.
11. Ming, A.B., et al., *Fault feature extraction and enhancement of rolling element bearing in varying speed condition*. *Mechanical Systems and Signal Processing*, 2016. **76–77**: p. 367-379.
12. Zhao, D., J. Li, and W. Cheng, *Feature Extraction of Faulty Rolling Element Bearing under Variable Rotational Speed and Gear Interferences Conditions*. *Shock and Vibration*, 2015. **2015**: p. 9.
13. Antoni, J., *Fast computation of the kurtogram for the detection of transient*

- faults*. Mechanical Systems and Signal Processing, 2007. **21**(1): p. 108-124.
14. Antoni, J. and R.B. Randall, *The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines*. Mechanical Systems and Signal Processing, 2006. **20**(2): p. 308-331.
 15. Ha, J.M., et al., *Autocorrelation-based time synchronous averaging for condition monitoring of planetary gearboxes in wind turbines*. Mechanical Systems and Signal Processing, 2016. **70-71**: p. 161-175.
 16. Morhain, A. and D. Mba, *Bearing defect diagnosis and acoustic emission*. Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, 2003. **217**(4): p. 257-272 1350-6501.
 17. Borghesani, P., et al., *Application of cepstrum pre-whitening for the diagnosis of bearing faults under variable speed conditions*. Mechanical Systems and Signal Processing, 2013. **36**(2): p. 370-384.
 18. McFadden, P.D., et al., *The Detection of Seeded Faults in an Epicyclic Gearbox by Signal Averaging of the Vibration*. 1990: Department of Defence, Defence Science and Technology Organisation, Aeronautical Research Laboratory.
 19. McFadden, P.D. and M.M. Toozhy, *Application of synchronous averaging to vibration monitoring of rolling element bearings*. Mechanical Systems and Signal Processing, 2000. **14**(6): p. 891-906.
 20. Siegel, D., et al., *Novel method for rolling element bearing health assessment—A tachometer-less synchronously averaged envelope feature extraction technique*. Mechanical Systems and Signal Processing, 2012. **29**: p. 362-376.
 21. Borghesani, P., et al., *Order tracking for discrete-random separation in variable speed conditions*. Mechanical Systems and Signal Processing, 2012. **30**: p. 1-22.
 22. Childers, D.G., D.P. Skinner, and R.C. Kemerait, *The cepstrum: A guide to processing*. Proceedings of the IEEE, 1977. **65**(10): p. 1428-1443.
 23. Peeters, C., P. Guillaume, and J. Helsen, *A comparison of cepstral editing methods as signal pre-processing techniques for vibration-based bearing fault detection*. Mechanical Systems and Signal Processing, 2017. **91**: p. 354-381.
 24. Barbini, L., et al., *Phase editing as a signal pre-processing step for automated bearing fault detection*. Mechanical Systems and Signal Processing, 2017. **91**: p. 407-421.
 25. Sawalhi, N. and R.B. Randall. *Cepstrum editing (liftering) to remove discrete frequency signals, leaving a signal dominated by structural response effects and enhance fault detection in rolling element bearings. The eight international conference on condition monitoring and machinery failure prevention technologies, CM MFDT Cardiff Bay on*. 2011.
 26. Wang, Y., et al., *Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: A review with applications*. Mechanical Systems and Signal Processing, 2016. **66-67**(Supplement C): p. 679-698.
 27. Lei, Y., et al., *Application of an improved kurtogram method for fault diagnosis of rolling element bearings*. Mechanical Systems and Signal Processing, 2011. **25**(5): p. 1738-1749.
 28. Prabhakar, S., A.R. Mohanty, and A.S. Sekhar, *Application of discrete wavelet*

- transform for detection of ball bearing race faults*. Tribology International, 2002. **35**(12): p. 793-800.
29. Chen, J., et al., *Wavelet transform based on inner product in fault diagnosis of rotating machinery: A review*. Mechanical Systems and Signal Processing, 2016. **70**: p. 1-35.
 30. Lin, J. and L.S. Qu, *Feature extraction based on Morlet wavelet and its application for mechanical fault diagnosis*. Journal of Sound and Vibration, 2000. **234**(1): p. 135-148.
 31. Liu, H., A.N. Cartwright, and C. Basaran, *Moiré interferogram phase extraction: a ridge detection algorithm for continuous wavelet transforms*. Applied Optics, 2004. **43**(4): p. 850-857.
 32. Liu, Z., et al., *A hybrid fault diagnosis method based on second generation wavelet de-noising and local mean decomposition for rotating machinery*. ISA Transactions, 2016. **61**: p. 211-220.
 33. Yang, D.-M. *Induction motor bearing fault detection with non-stationary signal analysis*. in *2007 IEEE International Conference on Mechatronics*. 2007. IEEE.
 34. Zhi, Q., *Sound Quality Improvement for Vehicle Noise Based on A Delayless Narrowband Active Sound Quality Equalizer Algorithm*, in *Department of Mechanical and Automotive Engineering*. 2014, University of Ulsan.
 35. Smith, J.S., *The local mean decomposition and its application to EEG perception data*. Journal of the Royal Society Interface, 2005. **2**(5): p. 443-454.
 36. Cheng, J., Y. Yang, and Y. Yang, *A rotating machinery fault diagnosis method based on local mean decomposition*. Digital Signal Processing, 2012. **22**(2): p. 356-366.
 37. Han, M. and J. Pan, *A fault diagnosis method combined with LMD, sample entropy and energy ratio for roller bearings*. Measurement, 2015. **76**: p. 7-19.
 38. Liu, H. and M. Han, *A fault diagnosis method based on local mean decomposition and multi-scale entropy for roller bearings*. Mechanism and Machine Theory, 2014. **75**: p. 67-78.
 39. Tian, Y., et al., *Rolling bearing fault diagnosis under variable conditions using LMD-SVD and extreme learning machine*. Mechanism and Machine Theory, 2015. **90**: p. 175-186.
 40. Wang, L., et al., *Time–frequency analysis based on ensemble local mean decomposition and fast kurtogram for rotating machinery fault diagnosis*. Mechanical Systems and Signal Processing, 2018. **103**: p. 60-75.
 41. Sun, J., et al., *Natural gas pipeline small leakage feature extraction and recognition based on LMD envelope spectrum entropy and SVM*. Measurement, 2014. **55**: p. 434-443.
 42. Li, Y., et al., *A fault diagnosis scheme for rolling bearing based on local mean decomposition and improved multiscale fuzzy entropy*. Journal of Sound and Vibration, 2016. **360**: p. 277-299.
 43. Fengwei, W., et al., *Study on Methods for Improving LMD End Effect by Gaussian Process Based on the Particle Swarm Optimization Algorithm*. 2016.
 44. Lundberg, G. and A. Palmgren, *Dynamic capacity of rolling bearings*. 1947:

Generalstabens Litografiska Anstalts Förl.

45. Takemura, H., Y. Matsumoto, and Y. Murakami, *Development of a new life equation for ball and roller bearings*. SAE transactions, 2000: p. 261-267.
46. ZARETSKY, E.Y., *Fatigue criterion to system design, life, and reliability*. Journal of Propulsion and Power, 1987. **3**(1): p. 76-83.
47. Ioannides, E. and T.A. Harris, *A new fatigue life model for rolling bearings*. Journal of Tribology, 1985. **107**(3): p. 367-377.
48. Natu, M., *Bearing Fault Analysis Using Frequency Analysis and Wavelet Analysis*. International Journal of Innovation, Management and Technology, 2013. **4**(1): p. 90.
49. Ferreira, J., J. Balthazar, and A. Araujo, *An investigation of rail bearing reliability under real conditions of use*. Engineering Failure Analysis, 2003. **10**(6): p. 745-758.
50. Gurumoorthy, K., B.D. Faye, and A. Ghosh, *Handling abuse causes premature bearing failures*. Case Studies in Engineering Failure Analysis, 2013. **1**(4): p. 235-242.
51. Gurumoorthy, K. and A. Ghosh, *Failure investigation of a taper roller bearing: A case study*. Case Studies in Engineering Failure Analysis, 2013. **2**(1): p. 110-114.
52. Li, Y., et al., *Adaptive prognostics for rolling element bearing condition*. Mechanical systems and signal processing, 1999. **13**(1): p. 103-113.
53. Longching, C., C. Qing, and S. Eryu, *Study on initiation and propagation angles of subsurface cracks in GCr15 bearing steel under rolling contact*. Wear, 1989. **133**(2): p. 205-218.
54. Deng, S., et al., *Subsurface crack propagation under rolling contact fatigue in bearing ring*. Science China Technological Sciences, 2013. **56**(10): p. 2422-2432.
55. Li, B., et al., *Neural-network-based motor rolling bearing fault diagnosis*. IEEE transactions on industrial electronics, 2000. **47**(5): p. 1060-1069.
56. Li, C.J. and S. Li, *Acoustic emission analysis for bearing condition monitoring*. Wear, 1995. **185**(1-2): p. 67-74.
57. Shiroishi, J., et al., *Bearing condition diagnostics via vibration and acoustic emission measurements*. Mechanical systems and signal processing, 1997. **11**(5): p. 693-705.
58. Taha, Z. and K. Widiyati, *Artificial neural network for bearing defect detection based on acoustic emission*. The International Journal of Advanced Manufacturing Technology, 2010. **50**(1-4): p. 289-296.
59. Wang, Z., *Fault Diagnosis Method based on Fuzzy Support Vector Machines and Self-organizing Map Neural Network*. Vol. 4. 2012. 139-148.
60. Morales-Espejel, G.E. and A. Gabelli, *The Progression of Surface Rolling Contact Fatigue Damage of Rolling Bearings with Artificial Dents*. Tribology Transactions, 2015. **58**(3): p. 418-431.
61. Randall, R.B. and J. Antoni, *Rolling element bearing diagnostics—A tutorial*. Mechanical systems and signal processing, 2011. **25**(2): p. 485-520.

62. McFadden, P. and J. Smith, *Model for the vibration produced by a single point defect in a rolling element bearing*. Journal of sound and vibration, 1984. **96**(1): p. 69-82.
63. McFadden, P. and M. Toozy, *Application of synchronous averaging to vibration monitoring of rolling element bearings*. Mechanical Systems and Signal Processing, 2000. **14**(6): p. 891-906.
64. Harris, T.A., *Rolling bearing analysis*. 2001: John Wiley and sons.
65. Igarashi, T. and H. Hamada, *Studies on the vibration and sound of defective rolling bearings: First report: Vibration of ball bearings with one defect*. Bulletin of JSME, 1982. **25**(204): p. 994-1001.
66. Yan, R. and R.X. Gao, *Approximate entropy as a diagnostic tool for machine health monitoring*. Mechanical Systems and Signal Processing, 2007. **21**(2): p. 824-839.
67. McFadden, P. and J. Smith, *The vibration produced by multiple point defects in a rolling element bearing*. Journal of sound and vibration, 1985. **98**(2): p. 263-273.
68. Gebraeel, N., A. Elwany, and J. Pan, *Residual life predictions in the absence of prior degradation knowledge*. IEEE Transactions on Reliability, 2009. **58**(1): p. 106-117.
69. Gebraeel, N., et al., *Residual life predictions from vibration-based degradation signals: a neural network approach*. IEEE Transactions on industrial electronics, 2004. **51**(3): p. 694-700.
70. Swansson, N.S. and S. Falavaro, *Applications of vibration analysis to the condition monitoring of rolling element bearings*. 1984, AERONAUTICAL RESEARCH LABS MELBOURNE (AUSTRALIA).
71. Williams, T., et al., *Rolling element bearing diagnostics in run-to-failure lifetime testing*. Mechanical systems and signal processing, 2001. **15**(5): p. 979-993.
72. Sawalhi, N. and R. Randall, *Simulating gear and bearing interactions in the presence of faults: Part II: Simulation of the vibrations produced by extended bearing faults*. Mechanical Systems and Signal Processing, 2008. **22**(8): p. 1952-1966.
73. Sawalhi, N. and R. Randall, *Vibration response of spalled rolling element bearings: Observations, simulations and signal processing techniques to track the spall size*. Mechanical Systems and Signal Processing, 2011. **25**(3): p. 846-870.
74. Petersen, D., et al., *Analysis of bearing stiffness variations, contact forces and vibrations in radially loaded double row rolling element bearings with raceway defects*. Mechanical Systems and Signal Processing, 2015. **50**: p. 139-160.
75. Moazenahmadi, A., D. Petersen, and C. Howard, *A nonlinear dynamic model of the vibration response of defective rolling element bearings*. Proc of Australian Acoustics, 2013.
76. Singh, S., et al., *Analyses of contact forces and vibration response for a defective rolling element bearing using an explicit dynamics finite element*

- model*. Journal of Sound and Vibration, 2014. **333**(21): p. 5356-5377.
77. Zhao, M., et al., *Detection and recovery of fault impulses via improved harmonic product spectrum and its application in defect size estimation of train bearings*. Measurement, 2016. **91**: p. 421-439.
 78. Lin, J. and M. Zhao. *A review and strategy for the diagnosis of speed-varying machinery*. in *2014 International Conference on Prognostics and Health Management*. 2014.
 79. Ma, J. and C.J. Li, *Gear defect detection through model-based wideband demodulation of vibrations*. Mechanical Systems and Signal Processing, 1996. **10**(5): p. 653-665.
 80. Borghesani, P., et al., *The velocity synchronous discrete Fourier transform for order tracking in the field of rotating machinery*. Mechanical Systems and Signal Processing, 2014. **44**(1): p. 118-133.
 81. Fyfe, K. and E. Munck, *Analysis of computed order tracking*. Mechanical Systems and Signal Processing, 1997. **11**(2): p. 187-205.
 82. Zhao, M., et al., *Tachless Envelope Order Analysis and Its Application to Fault Detection of Rolling Element Bearings with Varying Speeds*. Sensors, 2013. **13**(8).
 83. Pan, M.C. and Y.F. Lin, *Further exploration of Vold–Kalman-filtering order tracking with shaft-speed information—I: Theoretical part, numerical implementation and parameter investigations*. Mechanical Systems and Signal Processing, 2006. **20**(5): p. 1134-1154.
 84. Pan, M.C. and Y.F. Lin, *Further exploration of Vold–Kalman-filtering order tracking with shaft-speed information—II: Engineering applications*. Mechanical Systems and Signal Processing, 2006. **20**(6): p. 1410-1428.
 85. Wang, Y., et al., *Rotating speed isolation and its application to rolling element bearing fault diagnosis under large speed variation conditions*. Journal of Sound and Vibration, 2015. **348**: p. 381-396.
 86. Fyfe, K.R. and E.D.S. Munck, *Analysis of computed order tracking*. Mechanical Systems and Signal Processing, 1997. **11**(2): p. 187-205.
 87. Guo, Y., et al., *Envelope order tracking for fault detection in rolling element bearings*. Journal of Sound and Vibration, 2012. **331**(25): p. 5644-5654.
 88. Wang, K.S., D. Guo, and P.S. Heyns, *The application of order tracking for vibration analysis of a varying speed rotor with a propagating transverse crack*. Engineering Failure Analysis, 2012. **21**: p. 91-101.
 89. Kong, F., et al., *A Vibration Model of Ball Bearings with a Localized Defect Based on the Hertzian Contact Stress Distribution*. Shock and Vibration, 2018. **2018**: p. 14.
 90. Sunnersjö, C.S., *Rolling bearing vibrations—The effects of geometrical imperfections and wear*. Journal of Sound and Vibration, 1985. **98**(4): p. 455-474.
 91. Khanam, S., J. Dutt, and N. Tandon, *Impact force based model for bearing local fault identification*. Journal of Vibration and Acoustics, 2015. **137**(5): p. 051002.
 92. Tandon, N. and A. Choudhury, *A Theoretical Model to Predict the Vibration*

- Response of Rolling Bearings in a Rotor Bearing System to Distributed Defects Under Radial Load*. Journal of Tribology, 1999. **122**(3): p. 609-615.
93. Moazenahmadi, A., et al. *Defect size estimation and analysis of the path of rolling elements in defective bearings with respect to the operational speed*. 2014. Institute of Noise Control Engineering.
 94. Moazen Ahmadi, A., C.Q. Howard, and D. Petersen, *The path of rolling elements in defective bearings: Observations, analysis and methods to estimate spall size*. Journal of Sound and Vibration, 2016. **366**: p. 277-292.
 95. Zhang, X., Y. Liang, and J. Zhou, *A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM*. Measurement, 2015. **69**: p. 164-179.
 96. Cempel, C. and M. Tabaszewski, *Multidimensional condition monitoring of machines in non-stationary operation*. Mechanical Systems and Signal Processing, 2007. **21**(3): p. 1233-1241.
 97. Guo, W., et al., *Elimination of end effects in local mean decomposition using spectral coherence and applications for rotating machinery*. Digital Signal Processing, 2016. **55**: p. 52-63.
 98. Huang, D., J. Zhao, and J. Su, *Practical implementation of the Hilbert-Huang transform algorithm*. Acta Oceanologica Sinica, 2002. **25**(1): p. 1-11.
 99. Gai, Q., X.-J. Ma, and H.-Y. Zhang, *New method for processing end effect in local wave method*. JOURNAL-DALIAN UNIVERSITY OF TECHNOLOGY, 2002. **42**(1): p. 115-117.
 100. Yong-jun, D., et al., *Comments and modifications on EMD method*. Chinese Science Bulletin, 2001. **46**(3): p. 257-263.
 101. Wang, Y., Z. He, and Y. Zi, *A demodulation method based on improved local mean decomposition and its application in rub-impact fault diagnosis*. Measurement Science and Technology, 2009. **20**(2): p. 025704.
 102. Tian, Y., Z. Wang, and C. Lu, *Self-adaptive bearing fault diagnosis based on permutation entropy and manifold-based dynamic time warping*. Mechanical Systems and Signal Processing.
 103. Sakoe, H., et al., *Dynamic programming algorithm optimization for spoken word recognition*. Readings in speech recognition, 1990. **159**: p. 224.
 104. Han, T., X. Liu, and A.C.C. Tan, *Fault diagnosis of rolling element bearings based on Multiscale Dynamic Time Warping*. Measurement, 2017. **95**: p. 355-366.
 105. Salvador, S. and P. Chan, *Toward accurate dynamic time warping in linear time and space*. Intelligent Data Analysis, 2007. **11**(5): p. 561-580.
 106. Liu, H., et al., *Adaptive spectral kurtosis filtering based on Morlet wavelet and its application for signal transients detection*. Signal Processing, 2014. **96**: p. 118-124.
 107. Su, W., et al., *Rolling element bearing faults diagnosis based on optimal Morlet wavelet filter and autocorrelation enhancement*. Mechanical Systems and Signal Processing, 2010. **24**(5): p. 1458-1472.
 108. Jiang, Y., et al., *Feature extraction method of wind turbine based on adaptive*

- Morlet wavelet and SVD*. Renewable Energy, 2011. **36**(8): p. 2146-2153.
109. Cao, Y., et al., *A method for extracting weak impact signal in NPP based on adaptive Morlet wavelet transform and kurtosis*. Progress in Nuclear Energy, 2018. **105**: p. 211-220.
 110. He, P., P. Li, and H. Sun, *Feature Extraction of Acoustic Signals Based on Complex Morlet Wavelet*. Procedia Engineering, 2011. **15**: p. 464-468.
 111. Hoang, D.-T. and H.-J. Kang, *A survey on Deep Learning based bearing fault diagnosis*. Neurocomputing, 2018.
 112. Costa, M., A.L. Goldberger, and C.-K. Peng, *Multiscale entropy analysis of complex physiologic time series*. Physical review letters, 2002. **89**(6): p. 068102.
 113. Zheng, J., et al., *A rolling bearing fault diagnosis method based on multi-scale fuzzy entropy and variable predictive model-based class discrimination*. Mechanism and Machine Theory, 2014. **78**: p. 187-200.
 114. Bhalla, D., R.K. Bansal, and H.O. Gupta, *Function analysis based rule extraction from artificial neural networks for transformer incipient fault diagnosis*. International Journal of Electrical Power & Energy Systems, 2012. **43**(1): p. 1196-1203.
 115. Li, X., et al., *Rolling element bearing fault detection using support vector machine with improved ant colony optimization*. Measurement, 2013. **46**(8): p. 2726-2734.
 116. Saravanan, N. and K.I. Ramachandran, *Incipient gear box fault diagnosis using discrete wavelet transform (DWT) for feature extraction and classification using artificial neural network (ANN)*. Expert Systems with Applications, 2010. **37**(6): p. 4168-4181.
 117. Nasrabadi, N.M., *Pattern recognition and machine learning*. Journal of electronic imaging, 2007. **16**(4): p. 049901.
 118. Bearing Data Center, C.W.R.U.; Available from: <http://csegroups.case.edu/bearingdatacenter/pages/download-data-file>.