Ph. D. Thesis

Study on Intelligent Condition Diagnosis Based on Vibration Information and Support Vector Machine for Plant Machinery

Graduate School of Bioresources Mie University

XUE HONGTAO

March, 2014
Contents

Abstract .................................................................................................................. i-iii

Chapter 1 Introduction

1.1 Background ........................................................................................................ 1
1.1.1 Significance of machine maintenance ......................................................... 1
   A. Improving efficiency ..................................................................................... 1
   B. Saving energy ............................................................................................... 2
   C. Safety and pro-environment ....................................................................... 3
1.1.2 Development of machine maintenance techniques .................................... 3
1.1.3 Condition-based maintenance .................................................................... 5
1.1.4 Data acquisition techniques ....................................................................... 7
1.1.5 Signal processing techniques ...................................................................... 9
   A. Time domain analysis .................................................................................. 10
   B. Frequency domain analysis ....................................................................... 10
   C. Time-frequency domain analysis ................................................................ 11
1.1.6 Artificial intelligence techniques ............................................................... 13
1.1.7 Symptom parameters for intelligent condition diagnosis .......................... 14
   A. Dimensional symptom parameter (DSP) ..................................................... 14
   B. Non-dimensional symptom parameter (NDSP) ......................................... 15
1.2 Objective of this thesis .................................................................................... 15
1.2.1 Research on effective extraction of fault signals ........................................ 16
1.2.2 Research on symptom parameters for diagnosis ......................................... 16
1.2.3 Research on an intelligent diagnosis method ............................................. 16
1.3 Outline of the thesis ....................................................................................... 18
References ............................................................................................................... 20
Chapter 2  Condition Diagnosis Method based on Symptom Parameters and Support Vector Machine

2.1 Introduction .................................................................................................................................................. 27
2.2 Symptom parameters and Sensitivity Evaluation ......................................................................................... 28
   2.2.1 Symptom parameters in frequency domain ....................................................................................... 28
   2.2.2 Discrimination index (DI) .................................................................................................................... 29
2.3 Condition diagnosis method based on SVMs .............................................................................................. 30
   2.3.1 Support vector machines (SVMs) ......................................................................................................... 30
   2.3.2 Condition Diagnosis Method based on SPs and SVMs ................................................................. 32
2.4 Verification experiment ............................................................................................................................... 33
   2.4.1 Experimental system .......................................................................................................................... 33
   2.4.2 Acquisition of diagnostic system by SVMs ....................................................................................... 36
   2.4.3 Diagnosis and verification .................................................................................................................. 39
   2.4.4 Comparison with diagnosis result by Neural Network (NN) ............................................................. 40
2.5 Conclusions .................................................................................................................................................. 42
References ......................................................................................................................................................... 42

Chapter 3  Sequential Diagnosis Method using Support Vector Machine and Relative Ratio Symptom Parameters for Precisely Distinguishing Structural Faults of Rotating Machinery

3.1 Introduction .................................................................................................................................................. 45
3.2 Relative ratio symptom parameters (RRSPs) ............................................................................................. 46
   3.2.1 Distinctive characteristics of structural faults .................................................................................. 46
   3.2.2 Relative ratio symptom parameters (RRSPs) .................................................................................. 47
3.3 Intelligent diagnosis using SVMs and RRSPs ............................................................................................. 49
   3.3.1 Support vector machines (SVMs) ....................................................................................................... 49
3.3.2 Intelligent diagnosis using SVM and RRSPs ........................................53
3.3.3 System of sequential diagnosis .......................................................54
3.4 Application to an experiment .............................................................55
  3.4.1 Experimental system and signal conditions ..................................55
  3.4.2 Sensitivity of RRSPs ...................................................................56
  3.4.3 Fault diagnosis ...........................................................................59
3.5 Conclusions .....................................................................................63
3.5 References .......................................................................................64

Chapter 4 Condition Diagnosis Method using Distinctive Frequency Components and Support Vector Machines under Varied Operating Conditions

4.1 Introduction .....................................................................................68
4.2 Distinctive frequency components (DFCs) of structural faults ..........70
  4.2.1 Rotating frequency and its high-order harmonic components .......70
  4.2.2 Distinctive frequency components (DFCs) of structural faults ....72
  4.2.3 Normalization and extraction of distinctive frequency components (DFCs) .... 74
4.3 Fault diagnosis based on SVMs .........................................................76
  4.3.1 Support vector machines (SVMs) ...............................................76
  4.3.2 Fault diagnosis based on DFCs and SVMs ..................................80
4.4 Application to an experiment ............................................................83
  4.4.1 Experimental system and signal conditions .................................83
  4.4.2 Data acquisition and DFC extraction .........................................85
  4.4.3 Fault diagnosis ........................................................................86
    4.4.3.1 Training data acquisition and building diagnostic system ....86
    4.4.3.2 Diagnosis classification ......................................................90
4.5 Conclusions ...................................................................................96
Chapter 5  Diagnosis Method of Multi-faults State using Optimal Composition of Symptom Parameters and Fuzzy Support Vector Machine

5.1 Introduction ................................................................................................................. 100
5.2 Vibration data .................................................................................................................. 102
5.3 Non-dimensional symptom parameters (NSPs) ................................................................. 106
5.4 Organizing optimal composition of symptom parameters .................................................. 107
  5.4.1 Condition diagnosis approach ...................................................................................... 107
  5.4.2 Principal component analysis for selecting symptom parameters ................................. 108
  5.4.3 Organizing optimal composition of symptom parameters ............................................ 111
5.5 Generation method of synthetic symptom parameter by SVM ............................................ 112
  5.5.1 Acquisition of optimal hyper-plane by SVM ............................................................... 112
  5.5.2 Synthetic symptom parameter .................................................................................... 115
5.6 Condition diagnosis using possibility theory ...................................................................... 116
  5.6.1 Possibility theory for condition diagnosis ................................................................. 116
  5.6.2 Diagnosis and verification ........................................................................................... 119
  5.6.3 Comparing the performances of OCSPs ..................................................................... 122
5.7 Conclusions .................................................................................................................... 122
References .......................................................................................................................... 123

Chapter 6  Precision Diagnosis Method for a Centrifugal Pump using Statistic Filter, Support Vector Machine and Possibility Theory

6.1 Introduction ...................................................................................................................... 127
6.2 Experimental centrifugal pump system for condition diagnosis ................................................................. 129
6.3 Feature extraction using statistic filter ...................................................................................................... 131
  6.3.1 Statistic filter ....................................................................................................................................... 131
  6.3.2 Analysis of experimental signals using statistic filter ........................................................................... 133
6.4 Non-dimensional symptom parameter for intelligent diagnosis ............................................................ 134
6.5 Application of SVM for synthetic symptom parameter ............................................................................ 137
6.6 Fuzzy diagnosis using possibility theory .................................................................................................. 138
  6.6.1 Fuzzy inference using possibility theory .............................................................................................. 139
  6.6.2 Building diagnosis system .................................................................................................................. 140
  6.6.3 Diagnosis classification ....................................................................................................................... 142
6.7 Conclusions .............................................................................................................................................. 145
References ......................................................................................................................................................... 146

Chapter 7 Conclusions and Future Research

7.1 Conclusions ................................................................................................................................................ 150
7.2 Future research ......................................................................................................................................... 153
Abstract

Condition diagnosis plays a significant role in modern equipment management. Establishment of an intelligent system of condition diagnosis not only improves the productivity, but also reduces the cost of maintenance and the risk of unexpected failures. In the field of condition diagnosis of the plant machinery, particularly rotating machinery, vibration information is widely used to detect a fault and identify the fault types. Certainly, condition diagnosis based on vibration information depends largely on feature extraction. Only when the features of vibration information are sensitively extracted in any condition of a machine can condition diagnosis be effective. However, plant machines are operating under unsteady conditions, even if the machines are in the normal state, rotating speed and operating load can vary. These can influence the spectrum feature of the vibration information measured. Moreover, vibration information often contain strong noise, especially at an early stage of a fault. Therefore, it is difficult to extract the features of vibration information. In addition, when building an intelligent system for diagnosing the condition of plant machinery, symptom parameters (SPs) and artificial intelligence (AI) are required. A high sensitive SP can express features information of machine conditions. However, in most cases of condition diagnosis for plant machinery, the sensitivity of some SPs is not high. The main reasons can be explained as follow: (1) the pure features are ambiguous in vibration information, because a fault is at an early stage, or the measure point is farther from failure part; (2) previous work namely feature extraction is unsatisfied; (3) the selected SPs cannot sensitively reflect the conditions of monitoring machine. In the case of AI, neural network (NN), genetic algorithm (GA), support vector machine (SVM), etc. have some special advantages as well as some disadvantages. For example, NN and GA will never converge when the first-layer data has the same values in different states, and SVM is only a two-class classifier. Moreover, these methods
cannot deal with ambiguous classification problems.

In order to extract effectively the features of vibration information, improve the ambiguous relationship between SPs and machine conditions, strengthen the sensitivity of SPs, and increase the efficiency of condition diagnosis at an early stage, this thesis has focused on studying condition diagnosis based on vibration information and support vector machine for plant machinery.

This thesis proposes statistic filter is performed to extract fault features from vibration information measured. Statistic filter is a method of signal processing by statistical tests of spectrums between normal information and measured information. In the field of machinery diagnosis, statistic filter is used to smooth away noise, what’s more, leave fault features unchanged.

This thesis defines many dimensional symptom parameters (DSPs) and non-dimensional symptom parameters (NDSPs) in time domain and in frequency domain for automatic diagnosis, which reflect the features of vibration information measured in plant machinery. Moreover, optimal composition of symptom parameters (OCSPs) that is a SPs’ group from multiple directions is proposed to improve the ambiguous relationship between SPs and machine conditions. In addition, distinctive frequency components (DFCs) and relative ratio symptom parameters (RRSPs) are proposed to exclusively detect and identify structural faults of rotating machinery. DFCs are the values of the spectrum features of structural faults, and RRSPs are a new type of SPs defined by the features of structural faults. Finally, discrimination index (DI) and principal component analysis (PCA) are employed to evaluate the sensitivity of DSPs, NDSPs or RRSPs, respectively. Each type of SPs has its good points. In the case of the diagnosis of structural faults of rotating machinery, the sensitivity of RRSPs is much higher, as the capability of DFCs is more effective.

This thesis proposes a sequential diagnosis method for monitoring and diagnosing the operating condition of rotating machinery, and the diagnosis approach is constructed on the basis of vibration information and support vector machines (SVMs).
DSPs, NDSPs, RRSPs and DFCs are regarded as the objects of a follow-on process, respectively. SVMs are used to establish the sequential diagnosis system for faults detection and the identification of fault types.

This thesis presents a fuzzy diagnosis method based on SVMs and possibility theory (PT). Soft margin SVMs are used to merge NDSPs or OCSPs into synthetic symptom parameters (SSPs) aiming to increase the diagnosis’ efficiency. PT is used to convert the probability distribution function of a SSP into a possibility function, and then the possibility function is regarded as the membership function for fuzzy inference.

Many practical examples of fault detection and the identification of fault types in rotating machinery are provided to verify that all of the proposed methods are effective.
Chapter 1

Introduction

1.1 Background

1.1.1 Significance of machine maintenance

Machinery, the most important requirement for any manufacturing unit except human resource, is used to get maximum productivity and meet the supply demand. With the rapid development of scientific technology, machines are growing large, complex, digital and intelligent, and their functions become perfect. No matter how well all machines design and make, parts can loosen or wear or deteriorate over time since they are operating under certain stress or load in the real environment, a fault would naturally occur. Not only does the fault lead to adverse effects on equipment performance and product quality, but the fault may also intensify secondary faults that could be created. In severe cases, machinery is damaged, environment is polluted and human is harmed. Therefore, it is very important to take care of their functions and ensure safety and reliability throughout a life of modern equipment, this is machine maintenance. An importance of an effective maintenance can be denoted by the keywords called efficiency, energy, safety and pro-environment [1-7].

A. Improving efficiency

In the increasingly fierce competition, corporations focus their attentions on how to reduce production cost and improve production efficiency. However, sudden failure causes important equipment to stop running. Nor is it only heavy losses for production that is a lot of cost and time for repairing and to impact directly their profit, reliability
and development.

The main purpose of regular maintenance is to ensure that all equipment required for production is operating at 100% efficiency at all times. Through short daily inspections, cleaning, lubricating, making minor adjustments, and monitoring main parts or positions, minor problems can be detected and corrected before they become major problems that can shut down a production line. Therefore, an effective maintenance can improve production efficiency to enhance competitive capacity [1, 3, 4].

B. Saving energy

Once a fault occurs and has not been timely repair, useless electric power increases and the service life of the parts is greatly affected. For example, when the misalignment of rotating shaft occurs, useless energy increases ever more steeply as the misalignment’s degree, shown as Fig. 1.1. Moreover the life of bearing and mechanical seal is shortened sharply, shown as Fig. 1.2. Therefore, a regular maintenance is to keep equipment running at optimal efficiency, to save energy and prolong operational life [8].

Fig. 1.1 The relationship of misalignment of rotating shaft and useless electric power
Chap. 1 Introduction

As safety and pro-environment are also an important issue to look at all machines need to be in a perfect running condition. A defected machine can lead to a major accident. In recent years, some accidents, fire and explosion are caused by the defected machine, have occurred frequently in the manufacturing. Then many people were hurt or killed, and environment such as water, soils and atmosphere has also been polluted in varying degrees [1]. Therefore, a regular maintenance is to help determine the condition of in-service equipment to prevent beforehand a major accident.

Obviously, machine maintenance is greatly useful tool to ensure equipment running safely, increase efficiency and reliability, save energy, avoid heavy economic loss and a major accident, and promote modern equipment management. Therefore, maintenance techniques develop rapidly.

1.1.2 Development of machine maintenance techniques

Before the Second World War, almost all equipment was simple and over-designed.
Then the consequences of failure did not have a strong influence and the effect was neglected, the plant was running until failure occurred, and when it did it was either replaced or repaired. There were no actions and techniques that were taken to detect or prevent failure. However, the typical maintenance practices were basic and routine maintenance, reactive breakdown service (fix it when it broke) and corrective maintenance. Therefore, the period is the first generation of machine maintenance techniques.

Between the Second World War and the late 1970s, the mechanization increased and the manufacturing became more complex, and cost, longevity and availability were regarded as important factors to achieve the business objectives. Therefore, maintenance was considered as a technology, periodic, planned and preventive maintenance programs, such as planned preventive maintenance, time based maintenance and system for planning and controlling work, were allowed to develop and implement. Since time was spent to disrupt normal operations and maintenance cost increased, the programs were criticized for imposing quite often unnecessary treatments. The period is the second generation of machine maintenance techniques.

Third generation of machine maintenance techniques belongs to the time period within 1980 and 2000. In the period, industries became even more automated and complex. Reliability, availability and maintainability, as well as quality, safety and environment were considered very important. Then condition based maintenance, reliability centered maintenance, computer aided maintenance management and information system began to be used in the industry [9]. However, the techniques were unripe.

It was not until recently that maintenance has gained recognition as potential profit generator and becomes more and more part of the integrated business concept. Moreover, there is a growing trend towards proactive strategies such as predictive and preventive maintenance, also a shift from “fire-fighting” maintenance to use-based maintenance and increasingly towards condition-based maintenance [10].
1.1.3 Condition-based maintenance

Condition-based maintenance (CBM) is a maintenance strategy that uses real-time operating performances to monitor and assess the deviations in both the quality of the product and the machine condition, to increase productivity, to lower operating costs, and to take machine health to the next level for the lifetime of the equipment. Basically, CBM not only differs from preventive maintenance by basing maintenance need on the actual condition of the machine rather than on some preset schedule, but also extends the concepts of predictive maintenance by using data from both on-line operating performances and off-line maintenance tests [1, 7, 11-13].

A CBM program commonly consists of four key steps namely data acquisition, feature extraction, diagnostics and prognostics, as shown in Fig. 1.3.

Data acquisition, an essential step of a CBM program, is a process of collecting and storing right data (information) from targeted physical assets for identifying machinery fault or not. It may involve various types of information, such as, vibration, temperature, pressure, speed, sound, voltage, etc. The techniques of data acquisition will be introduced particularly in the section 1.1.4 of the chapter.

Feature extraction is also called signal processing. Usually, the program of feature
extraction involves data processing, analysis, interpretation and extracting useful information from data collected in upper program. Moreover, signal collected in normal condition is a standard of detection, diagnostics and prognostics, then is called normal signal. If one or more machinery faults occur, there is feature signal that differs from normal signal and is useful to further diagnostics and prognostics in measured signal. It is difficult and important to extract feature signal. The commonly methods of feature extraction be introduced particularly in the section 1.1.5 of the chapter.

Diagnostics is a main category of techniques for maintenance decision support in a CBM program. Diagnostics is frequently described to deal with fault detection, fault identification. The first, fault detection, is the process of detecting fault when fault occurs. The second process, fault identification, is to determine the nature of the fault, the degree of the fault, the location of the fault and the time of detection. Fault detection is primary objective of a CBM program, even fault identification is triggered on the basis of the detection of a fault, so fault detection is commonly termed basic diagnosis. Together, Fault identification is termed precise diagnosis. Over the last decade or so, efforts have been undertaken to bring artificial intelligence techniques for diagnostics. In section 1.1.6, artificial intelligence techniques in support of a CBM program will be introduced.

The last process of a CBM program, prognostics, focuses on predicting whether a fault is impending and estimating how soon and how likely a fault will occur. Prognostics predicts the future performance or condition of equipment by assessing the extent or rate of deviation or degradation of equipment from its expected normal operating conditions [14]. The science of prognostics is based on the analysis of failure modes, detection of early signs of wear and aging, and fault conditions. The information is commonly provided from diagnostics. Technical approaches to building models in prognostics can be categorized broadly into data-driven approaches, model-based approaches, and hybrid approaches. Prognostics are used to make operation and maintenance decisions. Use of prognostics enables transition from maintenance based
on current conditions of equipment to predictive maintenance. Therefore, prognostic is much more efficient than diagnostics to achieve zero-downtime performance. Nevertheless, prognostics cannot completely replace diagnostics since in practice there are always some faults and failures which are not predictable. Besides, prognostics, like any other prediction techniques, cannot be 100% sure to predict faults and failures. In the case of unsuccessful prediction, diagnostics can be a complementary tool for providing maintenance decision support. In addition, diagnostics is also helpful to improving prognostics in the way that diagnostic information can be useful for preparing more accurate event data and hence building better CBM model for prognostics. Furthermore, diagnostic information can be used as useful feedback information for system redesign. A CBM program can be used to do diagnostics and prognostics.

With the development of artificial intelligence techniques, the concept of machinery diagnostics and prognostics transform from supervised learning methods (demanding human intervention) to unsupervised learning methods (reducing or severing human intervention). The term “unsupervised” implies ability to learn by itself without human supervision. Therefore, the supervisory reasoning algorithms in performing effective diagnostics and prognostics will be challenged.

Certainly, the core technology of CBM is condition monitoring which, in principle, involves data acquisition and extracting useful information. The information helps to identify whether the asset health has deviated from the normal. If so, fault diagnosis and prognosis often follow. Finally, a decision, regarding when and what maintenance tasks are to be performed, is taken [15]. A review on the research and development in machinery diagnostics and prognostics in CBM can also be found in [16].

1.1.4 Data acquisition techniques

Data acquisition is the process of collecting information that measure real-world physical conditions and converting the resulting samples into digital numeric values that
can be manipulated by a computer [17, 18]. Data acquisition system typically gets involved with three techniques.

The first, sensors, convert physical parameters to electrical signals. Since the physical parameters of interest can as diverse as temperature, pressure, light, force, position or vibration, there are many different types of sensor. However, each sensor converts usually one parameter to electrical quantities such as voltage [19]. For example, accelerometer converts proper acceleration of the vibration in rotating machinery that is exposed to dynamic loads to voltage signal. Now, single-axis or multi-axis models of accelerometer are available to detect magnitude and direction of the proper acceleration, as a vector quantity, and can be used to sense orientation, coordinate acceleration, vibration, shock, and falling in a resistive medium. At the same time, each sensor has certain capabilities and limitations that are described by the main characteristics such as range, resolution, sensing frequency, accuracy, size, opt environment, reliability, drift and cost. Thus, the selection of a sensor is based on matching the operating characteristics of sensors to the requirements of an application [20].

The second, analog signal conditioning, converts sensor signals into a form that can be converted to digital values. Nearly all sensor signals must be conditioned by analog circuitry before they can be digitized and used by a computer [18]. This conditioning often includes amplification and filtering. Amplification is necessary for the signal’s amplitude to fit within a reasonable portion of ADC’s (analog-to-digital converter’s) dynamic range. If ADC’s full dynamic range will be used, a minimum of information is lost. Of course, if the sensor signal is amplified too much, some of the signal will be clipped and severely distorted. However, noise or unwanted signal artifacts exist on analog signals. Then filter must usually be performed to remove noise or unwanted signal artifacts. More common filters are low-pass filter, high-pass filter and band-pass filter. The proper filter is selected to exhibit wanted signal.

The last, data conversion, is to convert conditioned sensor signals to digital values, and is at the heart of data acquisition system. As previously noted, most transducers
have analog outputs, usually voltage or current, which represent the physical parameters being measured, such as temperature, pressure, position or speed. Besides, digital parameters have discrete levels that vary by steps instead of continuously, and most digital electronic equipment uses binary values, which have two possible states, called true (on or 1) and false (off or 0) [18]. Then analog signals must be converted to binary representations through an analog-to-digital converter (ADC). Now, a multitude of techniques are used to improve resolution and sampling rate which are the most important ADC parameters. An ADC’s resolution is the smallest change it can detect in a measurement, and its value is actually a percentage of full-scale reading. Then high resolution is usually desirable in an ADC. As for sampling rate, it is the ADC specification most often examined, because an ADC’s sampling rate should be much higher than twice the maximum signal frequency in practice [1].

In recent years, there has been tremendous growth in data acquisition techniques, data acquisition software is produced. Data acquisition software not only collects wanted data, but also controls data analysis and eventual display in the first time. Of course, a detailed review about data acquisition techniques using personal computers (PCs) can also be found in [18].

### 1.1.5 Signal processing techniques

There are a large number of signal processing techniques that can be used in order to extract interesting information from a measured signal. In the case of machine maintenance, particularly condition-based maintenance, machines with moving parts give rise to vibrations, and the status of each machine yield to a peculiar vibration feature [21-25]. Therefore, vibration signals are interested to detect faults and identify fault type. Up to now vibration method is the most useful approach for condition diagnosis of machinery, especially rotating machinery, and the rate of vibration method in all diagnostic methods is 66% [26]. The key is to analyze the time of occurrence, or the frequency ranges of the feature, or both, from vibration signals, and extract useful
information to reflect the feature. Normally, the program is called signal processing or feature extraction. Here, the common techniques of signal analysis are introduced respectively.

A. Time domain analysis

Traditional time-domain analysis directly depends on simple statistical parameters to evaluate the measured time-domain signal, then to give characteristic feature about potential defects. For example, mean and standard deviation are respectively used to reflect central tendency and variability, peak and root mean square values are referred to the overall vibration level, skewness and kurtosis are shape parameters that estimate the deviation of a distribution from the normal distribution. These statistical parameters are simple to implement, however they are rather insensitive tools for defect detection. Now, a popular time-domain analysis approach is the synchronous signal averaging technique (SSAT) [27], the result of the SSAT is the signal average, which is the ensemble average of the angle domain signal, synchronously sampled with respect to the rotation of one particular shaft. In the resulting averaged signal (SA), the random noise as well as non-synchronous components is attenuated. The main advantage of the SSAT is the possibility to extract a simple and interesting signal from complex vibration signals. In addition, more advanced approaches of time-domain analysis apply time series models to waveform data. The main idea of time series modeling is to fit the waveform data to a parametric time series model and extract features based on this parametric model. The popular models used in the literature are the narrow-band demodulation model, the auto-regressive model and the auto-regressive moving average model [28, 29].

B. Frequency domain analysis

Frequency domain is a function of frequency that is used to represent a signal by its amplitude and phase at each component frequency. It not only simplifies the understanding of the waveform, but also isolates certain frequency components of interest. In order to calculate the frequency spectrum of a sampled time signal, the Fast
Fourier Transform is widely used as a conventional and efficient method [30]. In early studies, Fourier transform techniques such as Fourier series expansion (FSE), Fourier integral transform (FIT) and discrete Fourier transform (DFT) have been widely used in science and engineering. With the development of large-scale integration (LSI) and the associated microprocessor technology, fast Fourier transform (FFT) analysis became cost effective for general application. Fourier analysis has been the dominating signal analysis tool for fault detection. Unfortunately, there are some crucial restrictions of the Fourier transform. For example, the signal that will be analyzed must be periodic or stationary, and the resulting Fourier spectrum will make little physical sense [31-32]. Then some useful auxiliary tools, such as graphical presentation of spectrum, frequency filters, envelope analysis, are applied to analyze the spectrum of non-stationary signals. Hilbert transform is a useful and classical tool in envelope analysis, a detailed review that Hilbert transform has been used to diagnose machine faults can be found in [33]. Descriptions of the above-mentioned techniques for FFT-based spectrum can be found in [34].

C. Time-frequency domain analysis

Although previously mentioned spectral methods such as Fourier transform is the standard spectral analysis technique, it performs poorly at analyzing non-stationary signals. However, localized defects generally introduce non-stationary signal components [31], which cannot be properly described by ordinary spectral methods. This drawback can be overcome by time-frequency analysis techniques. Therefore, in later studies, time-frequency analysis methods are widely used to detect faults since they can determine not only the time of occurrence but also the frequency ranges of the location.

Time-frequency analysis can simultaneously generate a signal in both the time and frequency domains, using various time-frequency representations. The practical motivation for time-frequency analysis is that classical Fourier analysis assumes that
Chap. 1 Introduction

signals are infinite in time or periodic, while many signals in practice are of short duration, and change substantially over their duration. One of the most basic forms of time-frequency analysis is the short-time Fourier transform (STFT). The idea of STFT is to divide the whole waveform signal into segments with short-time window and then apply Fourier transform to each segment. However, narrow time windows mean poor frequency resolution. This trade-off between time and frequency resolution is the main disadvantage of the STFT, which can be solved by the use of other time-frequency techniques such as Wigner-Ville distribution (WVD), Choi-Williams distribution (CWD) and wavelet transform (WT). The WVD method is a bilinear transform, it is not based on signal segmentation and thus provides better time-frequency resolutions compared to the STFT, but produces the interference terms. These interference terms make interpretation of the estimated distribution difficult, but are overcome by using CWD. CWD, a transform that represents the spectral content of non-stationary signal as a two-dimensional time-frequency map, can filter out the cross-terms result from the components differ in both time and frequency center [35-37]. Other transform for time-frequency analysis is wavelet transform (WT).

WT is different from conventional Fourier Transform and modern time-frequency analysis. For stationary or non-stationary signals, WT not only gives a better representation to produce a high frequency resolution at low frequencies and a high time resolution at high frequencies with long duration low frequencies and short duration high frequencies [38-40], but also, like FFT, is a fast, linear operation that operates on a data vector whose length is an integer power of two, transforming it into a numerically different vector of the same length and an invertible transform [41]. Therefore, WT have become one of the most important and powerful tool of signal representation. WT is classified into discrete wavelet transform (DWT) and continuous wavelet transform (CWT). Since DWT and CWT are continuous-time (analog) transforms, they can be used to represent continuous-time (analog) signals. However each transform is suitable for different application. For example, CWT operates over every possible scale and
translation whereas DWT uses a specific subset of scale and translation values or representation grid [42-43].

1.1.6 Artificial intelligence techniques

In [44], artificial intelligence (AI) was firstly proposed as a concept. Since the 1980s several applications of AI had been developed in almost every field of engineering and management. Now, AI has pervasive applications in virtually all fields: engineering, science, education, medicine, business, accounting finance, marketing, economics, stock market and law [45, 46].

AI is basically a computer system performing as a substitute for intelligent functions of human beings. It mimics methods of learning, reasoning and solving problems in human beings through knowledge gathering. There are various subfields of AI such as distributed artificial intelligence, computational intelligence and robotics. In operations management including maintenance, we will focus on distributed artificial intelligence and computational intelligence.

Distributed artificial intelligence (DAI) is a subfield of artificial intelligence which has for more than two decades now, been investigating knowledge models, as well as communication and reasoning techniques that computational agents might need to participate in societies composed of computers. More, generally, DAI is concerned with situations in which several systems interact in order to solve a common problem. There are two main areas of research in DAI, distributed problem solving (DPS) and multi-agent system. DPS considers how solving a task of a particular problem can be divided among a number of modules that cooperate in dividing and sharing knowledge about the problem and about its evolving solution. A multi-agent system is concerned with the behavior of a collection of autonomous agents aiming at solving a given problem.

Computational intelligence is the study of adaptive mechanisms to enable or facilitate intelligent behavior in complex and changing environment. Computational intelligence
techniques include expert system (knowledge-based system), artificial neural networks (ANN), fuzzy classifier models (FCM), fuzzy logic systems (FLS) and genetic algorithm (GA) techniques, GA assisted ANN and support vector machines (SVM). They possess learning, adaptation, and classification capabilities, and have been successfully used for failure diagnosis and condition monitoring [47-50].

In practice, however, it is not easy to apply AI techniques due to the lack of efficient procedures to obtain training data and specific knowledge, which are required to train the models. So far, most of the applications in the literature just used experimental data for model training. Although many studies have been carried out to investigate the use of neural networks for automatic diagnosis of machinery condition, most of these methods have been proposed to deal with discrimination of fault types collectively [51-52]. However, the conventional neural networks cannot adequately reflect the possibility of ambiguous diagnosis problem, and will never converge when the first-layer symptom parameters have the same values in different states [53].

1.1.7 Symptom parameters for intelligent condition diagnosis

When a computer is used for condition diagnosis of plant machinery, symptom parameters (SPs) are required to express the information of machine condition [54-56]. If we can find one or several SPs by which most of faults can be identified, the equipment condition should be easily identified [57]. So many SPs have been defined in the field of condition diagnosis of machinery [58]. Commonly, SPs are classified into two types, namely, dimensional symptom parameter (DSP) and non-dimensional symptom parameter (NDSP).

A. Dimensional symptom parameter (DSP)

DSP can express the magnitude of a signal, such as mean, peak, root mean square, etc. Generally, the value of a DSP increases in proportion to the grade of a fault, then DSP is often used to detect whether machine condition is normal. Now, DSPs are almost used...
in the marketing equipment which is employed to diagnose rotating machine. However, the value of a DSP will change with the change of equipment, revolution and load, so there is no unified standard for diagnosing machine condition [1].

**B. Non-dimensional symptom parameter (NDSP)**

NDSP, such as skewness and kurtosis, crest factor and shape of wave, can reflect the shape of a signal. Because the value of a NDSP cannot be easily influenced by the change of equipment, revolution and load, NDSP is used to identify different faults. But the variation of a NDSP is not monotonous with increasing the grade of a fault, so it is difficult to set up a danger standard [1].

However, in most practical cases of plant, the measured data cannot be always satisfied because of the measuring techniques and manner of the inspectors [59]. In addition, since the mechanisms and the features of machine faults cannot be perfectly clarified by theory, it is difficult to set up accurate mathematical models. Moreover, plant machine conditions are more complex, and then fault types to be identified are too many. Thus, it is very hard to find one or several SPs that can simultaneously identify all of those faults. Particularly, it is difficult to judge the relationship between fault states and the SPs by a theoretical approach. Therefore, it should be directed against specific machine to find SPs for intelligent condition diagnosis.

**1.2 Objective of this thesis**

Rotating machinery, which accounts for approximately 40% of all machinery, is one of the most commonly used types of equipment and plays an important role in daily production [1]. Almost all manufacturing processes involve correlated rotating machines. Certainly, the failures of rotating machinery occur most frequently. Therefore, the objective of this thesis is to develop a quick and correct fault diagnostic system that should collect and process all the available information to detect any typical changes of
Chap. 1 Introduction

rotating machinery at an early stage. Several main researches are considered and briefly introduced as follows.

1.2.1 Research on effective extraction of fault signals

The diagnosis method using vibration signals has been widely used in practice, and the used proportion is about 66% [1]. Therefore, the thesis devotes much of its attention to the application of this method at an early stage of fault condition. However, at an early stage of fault condition, fault signals are essentially weak, even appear intermittently. Particularly, when some factors such as dangerous operation environment, poisonous material, are considered, the places that sensors are installed away from the failure part, fault signal are weaker. Moreover, there is very little possibility that the signals in normal condition are provided in plant. Though signal processing techniques are used to find fault signals, it is difficult to extract effectively fault signals form measured signals. Therefore, the first objective of this thesis is to research an effective extraction method for fault signals from vibration signals.

1.2.2 Research on symptom parameters for diagnosis

Symptom parameters (SPs), as a bridge between a computer and machine condition, are widely used for intelligent condition diagnosis. However, fault types are too much, and the mechanism of some fault is difficultly elucidated, and then it is difficult to find one or several SPs that the sensitivities are high. Therefore, the second objective of this thesis is to define exclusive symptom parameters for diagnosis that will been used to diagnose effectively some type faults or some rotating machine.

1.2.3 Research on an intelligent diagnosis method

Many intelligent methods such as neural networks (NN) and genetic algorithm (GA) have potential applications in pattern recognition and fault diagnosis. Many studies have been carried out to investigate the use of neural networks for automatic diagnosis of
machinery, and most of these methods have been proposed to deal with discrimination of fault types collectively [60-63]. However, these traditional intelligent methods, which are normally based on statistical learning, require a large number of training samples. Nevertheless, in most cases of practical machinery diagnosis, measured signals are not sufficient for conditioning the learning data. Moreover, these methods cannot reflect the possibility of ambiguous diagnosis problems, and will never converge when the first layer symptom parameters have the same values in different states.

Support vector machine (SVM) is a relatively new computational learning method based on the statistical learning theory. SVM uses a structural risk minimization principle to minimize an upper bound based on an expected risk, which leads to a better generalization performance. SVM is applied to diagnose automatically whether machine condition is normal or not, and discriminate fault types [64-65]. The objective of this thesis is to expand the application of SVM for fault diagnosis with common structural faults and main part or equipment such as bearing and centrifugal pump.

Structural faults easily occur in rotating machinery. Not only does any one of structural faults lead to adverse effects on equipment performance and product quality, but the fault may also intensify secondary faults such as bearing faults that could be created. Bearing, an important part of rotating machinery, is one of the most parts which are damaged easily. In [58] the failure rate of bearings in rotating machinery is about 33%. The failure of a bearing may lead to cease production, even the breakdown of a rotating machine. Therefore, in order to increase production efficiency and guarantee plant safety, it is the most important to detect any one of structural faults or bearing faults at an early stage. In addition, centrifugal pumps are used in a variety of different applications, such as for lifting thin liquids as well as highly dense liquids, such as muddy and sewage water, paper pulp, chemicals, etc. Some of these installations are crucial for a large system to work. Faults of pumps can cause the breakdown of a whole system, and lead to substantial economic losses. Therefore, condition diagnosis of a pump system at an early stage is very important [66-68].
1.3 Outline of the thesis

As mentioned previously, in automated manufacturing, process plants or major equipment operating, condition-based maintenance (CBM) is preferred to try to maintain the correct equipment at the right time. Despite its usefulness, there are several challenges encountered in completing the full circle of CBM. Firstly, condition diagnosis based on vibration signals depends largely on feature extraction. Only when the feature of the signal is sensitively extracted in any condition of a machine can condition diagnosis be effective. But plant machines are operating under unsteady conditions, and the measured vibration signals often contain strong noise, so it is difficult to extract the features of machine conditions from vibration signals. Secondly, symptom parameters (SPs) and artificial intelligence (AI) are required in an intelligent system of CBM. However, it is very hard to find one or several SPs that can perfectly identify all of the faults, and the high sensitive SPs are different for different machines. Moreover, the conventional AI techniques cannot deal with ambiguous classification problems. But these techniques have numerous advantages over conventional fault diagnostic approaches. Besides giving improved performance these techniques are easy to extend and modify. These can be made adaptive be the incorporation of new data or information. Then this thesis aims at these problems, and devotes to studying intelligent condition diagnosis for rotating machinery. Outline of this thesis is described as follows:

In chapter 2, a sequential diagnosis system is established using symptom parameters (SPs) and support vector machines (SVMs), to detect faults and identify fault types. SVMs can not only obtain good convergence, but also have a strong learning ability to deal with the fault diagnosis for small samples of learning data, and perform linear classification in a feature space for non-linear classification by kernel function. Moreover, sequential diagnosis system is proposed to overcome SVM’s disadvantage that SVM is only a two-class classifier. In addition, condition diagnosis using neural network (NN) is also built. Two different diagnosis methods are applied to diagnose
structural faults of rotating machinery. The diagnosis results are shown to verify that the performance of SVM is better than the performance of NN.

In chapter 3, relative ratio symptom parameters (RRSPs) are defined as a new type of exclusive symptom parameters (ESPs) to detect and identify structural faults of rotating machinery. Discrimination index (DI) is employed to evaluate the sensitivities of RRSPs and general symptom parameters (GSPs). The compared results are shown to verify that the sensitivity of RRSPs is very higher, even if at an early stage. Finally, RRSPs and support vector machines (SVMs) are used to build a sequential diagnosis system. The efficiency of sequential diagnosis method proposed in the chapter has been verified by applying it to a practical structural faults diagnosis of rotating machinery, and its performance is high.

In chapter 4, the spectrum features of structural faults under varied operating conditions, including no-load, fixed load, varied loads and different defect sizes, are analyzed to detect the common distinctive frequency components (DFCs). The normalized DFCs not only decrease feature differences from a state with varying operational conditions, but also replace effective symptom parameters (SPs). Then DFCs, as the object of a follow-on process, are used to build an automatic diagnosis method for structural faults of rotating machinery. These proposed methods have been applied to diagnose structural faults in rotating machinery, and verification results have shown that the methods are accurate and practical.

In chapter 5, a novel condition diagnosis method using optimal composition of symptom parameters (OCSPs) and fuzzy support vector machine (FSVM) is proposed to detect multi-faults state at an early stage. Principal component analysis (PCA) is employed to select higher sensitive common non-dimensional symptom parameters (NSPs) from multi-direction signals, and then to obtain an optimal composition namely OCSPs. Secondly, OCSPs and soft margin SVM are used to obtain an optimal hyper-plane, to define synthetic symptom parameter (SSP). Finally, a fuzzy diagnosis method with possibility theory is proposed for practical diagnosis of simple faults and
multi-faults of a bearing. In addition, the different composition SPs from simple direction and two directions signals are planned to compare the performance of OCSPs. The efficiency of the method proposed in this chapter is verified by applying them to bearing diagnosis.

In chapter 6, an intelligent diagnosis for a pump system using statistic filter, support vector machine (SVM) and possibility theory (PT) on the basis of the vibration signals, is proposed to detect faults and identify fault types at an early stage. The diagnosis system can add new diagnosing information into its knowledge base. At the same time, statistic filter, a method of signal processing that failure signal is extracted by statistical tests of spectrums between normal signal and fault signal, is adopted to extract the feature signals. Practical diagnosis examples for incipient faults of the centrifugal pump are given to show that those faults can be precisely identified by this method.

In chapter 7, the conclusions that can be drawn from the work accomplished in this thesis are summarized and the future research is also provided.

References


[28] P. D. McFadden: Detecting Fatigue Cracks in Gears by Amplitude and Phase


[51] A. Saxena and A. Saad: Evolving an Artificial Neural Network Classifier for


Chapter 2

Condition Diagnosis Method based on Symptom Parameters and Support Vector Machine

2.1 Introduction

In the field of condition monitoring and machinery diagnosis, symptom parameter (SP) is extremely important and effective tool. If we can find good symptom parameters which sensitively reflect the feature of the machine states, automatic diagnosis for mechanical failure is possible [1-4]. However, in many cases of the condition diagnosis for rotating machinery in a real plant, there are many problems as follows: (1) Effective SPs cannot be easily found because of too many fault categories [3, 5-6]; (2) The sensitivity of SP is not high under the influence of noise and operating conditions to be diagnosed machine [7-10]; (3) Traditional Condition Diagnosis Methods such as Neural Networks based on statistical learning require a large number of training samples. However, in most cases of practical machinery, measured signals are not so sufficient for the condition of learning data [11-13]. The intelligent methods, namely neural networks, genetic algorithms, etc., often cannot converge when learning, so the work of condition diagnosis is in passive situation. In order to resolve the above problems, this paper proposes a condition diagnosis method based on support vector machines (SVMs) and symptom parameters (SPs). SPs can be calculated by using the measured vibration signals, and then evaluated by the Discrimination Index (DI) with SVMs to classify states of rotating machinery, many verification results show that the methods presented in this paper are accurate and practical.
2.2 Symptom parameters and Sensitivity Evaluation

2.2.1 Symptom parameters in frequency domain

Many symptom parameters (SPs) have been defined for condition diagnosis of machinery [10]. Here, eight dimensionless SPs, which are usually used for the failure diagnosis of plant machinery in frequency domain, are shown as follows:

\[
P_1 = \frac{\sum_{i=1}^{N} x_i^2}{\sigma} \quad (2.1)
\]

\[
P_2 = \frac{\sigma}{x} \quad (2.2)
\]

\[
P_3 = \frac{\sum_{i=1}^{N} (x_i - \bar{x})^3}{N \cdot \sigma^3} \quad (2.3)
\]

\[
P_4 = \frac{\sum_{i=1}^{N} (x_i - \bar{x})^4}{N \cdot \sigma^4} \quad (2.4)
\]

Here, \(x_i\) is digital data of vibration signal, and \(N\) is the number of \(x_i\). \(\bar{x}\) and \(\sigma\) are the mean value and the standard deviation of \(x_i\), namely \(\bar{x} = \frac{\sum_{i=1}^{N} x_i}{N}\), \(\sigma = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \bar{x})^2}{N - 1}}\).

\[
P_5 = \frac{\sum_{i=1}^{I} (f_i - \bar{f})^3 \cdot F(f_i)}{\sigma_f^3 \cdot I} \quad (2.5)
\]
\[ P_6 = \frac{\sum_{i=1}^{I} (f_i - \bar{f})^4 \cdot F(f_i)}{\sigma_f^4 \cdot I} \] (2.6)

\[ P_7 = \sqrt{\frac{\sum_{i=1}^{I} f_i^4 \cdot F(f_i)}{\sum_{i=1}^{I} f_i^2 \cdot F(f_i)}} \] (2.7)

\[ P_8 = \frac{\sum_{i=1}^{I} f_i^2 \cdot F(f_i)}{\sqrt{\sum_{i=1}^{I} F(f_i) \cdot \sum_{i=1}^{I} f_i^4 \cdot F(f_i)}} \] (2.8)

Here, \( I \) is the number of spectrum line, \( f_i \) is the frequency, \( F(f_i) \) is the power spectrum. \( \sigma_f = \sqrt{\frac{\sum_{i=1}^{I} (f_i - \bar{f})^2 \cdot F(f_i)}{I}} \) and \( \bar{f} = \frac{\sum_{i=1}^{I} f_i \cdot F(f_i)}{\sum_{i=1}^{I} F(f_i)} \).

### 2.2.2 Discrimination index (DI)

Supposing that \( x_1 \) and \( x_2 \) are values of a symptom parameter (SP) calculated from the signals measured in state 1 and state 2, respectively, and conforming respectively to the normal distributions \( N(\mu_1, \sigma_1) \) and \( N(\mu_2, \sigma_2) \). Here, \( \mu_1, \mu_2 \) and \( \sigma_1, \sigma_2 \) are the average and the standard deviation of the SP. The larger the value of \( |x_1-x_2| \) is, the higher the sensitivity of distinguishing the two states by the SP. Because \( z = x_2 - x_1 \) also conforms to the normal distribution \( N(\mu_2 - \mu_1, \sqrt{\sigma_1^2 + \sigma_2^2}) \), there is the following density function about \( z \):

\[ f(z) = \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} \exp\left\{ -\frac{(z - (\mu_2 - \mu_1))^2}{2(\sigma_1^2 + \sigma_2^2)} \right\} \] (2.9)

where, \( \mu_2 \geq \mu_1 \) (the same conclusion can be drawn when \( \mu_1 \geq \mu_2 \)). The probability can be
calculated with the following formula:

$$P_0 = \int_{-\infty}^{0} f(z) dz$$  \hspace{1cm} (2.10)$$

where, $1-P_0$ is called the “discrimination rate (DR)”. With the substitution:

$$\mu = \frac{z - (\mu_2 - \mu_1)}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$  \hspace{1cm} (2.11)$$

into Formulas (2.9) and (2.10), the $P_0$ can be obtained by:

$$P_0 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-DI} \exp\{-\frac{\mu^2}{2}\} d\mu$$  \hspace{1cm} (2.12)$$

where, the distinction index (DI) is calculated by:

$$DI = \frac{|\mu_2 - \mu_1|}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$  \hspace{1cm} (2.13)$$

It is obvious that the larger value of DI, the larger value of Discrimination Rate (DR) will be, and therefore, the better the SP will be [3, 14]. When the value of DI is 1.65, distinction rate (DR) is 95%.

2.3 Condition diagnosis method based on SVMs

2.3.1 Support vector machines (SVMs)

SVMs are a relatively new computational learning method based on Statistical Learning Theory (SLT), and the basic idea is to create the Optimal Hyper-plane [15-17]. Given samples $S (x_i, y_i) (i=1,2, ,N)$ , $N$ is the number of samples, samples are assumed
to have two classes, \( x_i \in x = \mathbb{R}^d \), \( y_i \in y = \{+1, -1\} \), each of classes associates with labels be \( y_i = 1 \) for class A and \( y_i = -1 \) for class B respectively. SVMs try to place a linear boundary between two different classes A and B. the boundary can be expressed as Eq. (2.14), where \( \omega \) and \( b \) are used to define the position of separating hyper-plane. And by transforming to Quadratic Problem as Eq. (2.15), the hyper-plane is maximized. Finally, the decision function as Eq. (2.16) is made using \( \text{sign}(f(x)) \) to create separating hyper-plane.

\[
f(x) = \omega \cdot x + b = \sum_{i=1}^{N} \omega_i \cdot x_i + b = 0 \tag{2.14}
\]

\[
Q(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N} \alpha_i \alpha_j y_i y_j (x_i x_j) \tag{2.15}
\]

\[
y = \text{sign}(f(x)) = \text{sign}(\omega \cdot x + b) \tag{2.16}
\]

SVMs’ decision function is similar with neural network, as the output \( y \) is a linear combination of intermediate nodes, while each intermediate node is corresponding with a support vector, as shown in Fig. 2.1. SVMs can perform a non-linear data into a feature space by kernel function. SVMs also exhibit good generalization characteristic when fault samples are few. These have important action in the field of mechanical fault diagnosis.

![SVMs' classification diagram](image-url)
2.3.2 Condition Diagnosis Method based on SPs and SVMs

This paper presents a condition diagnostic method based on SPs and SVMs for intelligent machine condition monitoring and fault diagnosis. The flowchart of the method is shown in Fig. 2.2. The data of SPs, as the object of follow-on process, are extracted from the vibration signals in plant for condition diagnosis. Put them into SVMs as the training data then the optimal hyper-plane obtained. If the training data can be linearly separated by SVMs, optimal hyper-plane will be obtained by linear classification. If the training data cannot be linearly separated, appropriate kernel function will be chosen to search the optimal hyper-plane using SVMs. But using kernel function, the training data are mapped to higher-order feature space. It follows that the generalizing capability decreases and the computational complexity increases with the dimension becomes more. So the linear classification is used as much as possible [8]. After the optimal hyper-plane obtained, the testing data are analyzed using condition diagnosis method.

![Flowchart of Condition Diagnosis Method based on SPs and SVMs](image)

**Fig. 2.2 Condition diagnosis based on SPs and SVMs**

SVMs is a 2-class classifier, however many conditions such as structural faults will
be often diagnosed, and it is difficult to find the excellent SPs that can be used together. For efficient diagnosis, the sequential diagnosis is proposed, shown as Fig. 2.3. It consists of a number of SVMs. The first SVM is used to detect faults while the others are used to identify fault types.

![Sequential diagnosis based on SPs and SVMs](image)

**Fig. 2.3** Sequential diagnosis based on SPs and SVMs

### 2.4 Verification experiment

#### 2.4.1 Experimental system

Fig. 2.4 shows the experimental equipment used in low rotation speed for diagnosis test, and the place of accelerometers. The most commonly occurring structural faults in rotating machinery are shaft misalignment, unbalance and looseness states. These structural faults are shown as Fig 2.5 and were artificially made to simulate the failure at an early stage. The accelerometer (PCB MA352A60) with a bandwidth from 5 Hz to 60 kHz and 10 mV/g output was used to measure the vibration signals of the horizontal direction in the normal (N), unbalance (U), looseness (L) and misalignment (M) states,
respectively. The vibration signals of the normal, unbalance, looseness and misalignment states are shown in Fig. 2.6. These states are structural faults in low frequency domain, so the feature of 0~200Hz is showed in the spectrum through low pass filter. In addition, the sampling frequency of signal measurement is 5000Hz, and the sampling time is 20s, the rotational speed is 1500rpm. For calculating the SPs shown in formula (2.1) ~ (2.8), the vibration signals in each states are divided into 20 parts.

![Experimental equipment for fault diagnosis](image)

**Fig. 2.4 Experimental equipment for fault diagnosis**

(A) Unbalance state  (B) Looseness state  (C) Misalignment state

![Most commonly occurring structural faults](image)

**Fig. 2.5 The most commonly occurring structural faults**

According to the sequential diagnosis proposed in the section 2.3.2, the diagnostic system is designed three steps. In the first step, normal (N) state will be distinguished
from four states namely normal state and three commonly occurring structural faults. Then three structural faults are considered as another state namely abnormal (ULM) state. In the second step, unbalance (U) state will be identified from three structural faults. Similarly, structural faults except unbalance state are regarded as another state (LM). In the third step, looseness (L) and misalignment (M) states will be separated. However, since the experiment was to simulate structural faults at the early stage, the DI and DR values of SPs in each step are not large, shown in Table 2.1. Therefore, two SPs are selected by the values of DI in each step, shown in Table 2.1, to detect fault and identify fault type.

![Vibration signals and feature spectra in each state](image)

**Table 2.1 The DI and DR value of SPs**

<table>
<thead>
<tr>
<th>States</th>
<th>Value</th>
<th>P_1</th>
<th>P_2</th>
<th>P_3</th>
<th>P_4</th>
<th>P_5</th>
<th>P_6</th>
<th>P_7</th>
<th>P_8</th>
<th>SPs selected</th>
</tr>
</thead>
<tbody>
<tr>
<td>N vs. ULM</td>
<td>DI</td>
<td>0.44</td>
<td>1.21</td>
<td>0.49</td>
<td>0.69</td>
<td>0.96</td>
<td>0.82</td>
<td>1.18</td>
<td>0.68</td>
<td>P_2, P_7</td>
</tr>
<tr>
<td></td>
<td>DR</td>
<td>67%</td>
<td>89%</td>
<td>69%</td>
<td>75%</td>
<td>83%</td>
<td>79%</td>
<td>88%</td>
<td>75%</td>
<td></td>
</tr>
<tr>
<td>U vs. LM</td>
<td>DI</td>
<td>0.65</td>
<td>0.48</td>
<td>0.69</td>
<td>0.67</td>
<td>0.71</td>
<td>0.69</td>
<td>1.19</td>
<td>0.55</td>
<td>P_5, P_7</td>
</tr>
<tr>
<td></td>
<td>DR</td>
<td>74%</td>
<td>69%</td>
<td>75%</td>
<td>75%</td>
<td>76%</td>
<td>75%</td>
<td>88%</td>
<td>71%</td>
<td></td>
</tr>
<tr>
<td>L vs. M</td>
<td>DI</td>
<td>0.91</td>
<td>0.42</td>
<td>0.58</td>
<td>0.90</td>
<td>1.32</td>
<td>0.75</td>
<td>1.09</td>
<td>0.84</td>
<td>P_5, P_7</td>
</tr>
<tr>
<td></td>
<td>DR</td>
<td>82%</td>
<td>66%</td>
<td>72%</td>
<td>82%</td>
<td>91%</td>
<td>77%</td>
<td>86%</td>
<td>80%</td>
<td></td>
</tr>
</tbody>
</table>
2.4.2 Acquisition of diagnostic system by SVMs

In the first step, $P_2$ and $P_7$ are selected as the object of follow-on process. Then we firstly build $x_i = (P_2, P_7)$ $(i = 1, 2, \cdots, 8)$ from the SPs of normal state, and $x_j = (P_2, P_7)$ $(j = 1, 2, \cdots, 8)$ from the SPs of other states, to combine to the training data $X_n = [x_i, x_j]$, and define +1 for N and -1 for ULM in Table 2.2. Secondly by solving Eq. (2.15), Lagrange coefficient $\alpha$ as Table 2.3 is obtained under the conditions as follows:

Subject to:

$$\alpha_i \geq 0, i = 1, N \quad (2.17)$$

$$\sum_{i=1}^{N} \alpha_i y_i = 0 \quad (2.18)$$

<table>
<thead>
<tr>
<th>$P_2$</th>
<th>$P_7$</th>
<th>$Y_n$</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.001</td>
<td>0.748</td>
<td>1</td>
<td>N</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>0.017</td>
<td>0.746</td>
<td>1</td>
<td>N</td>
</tr>
<tr>
<td>0.235</td>
<td>0.776</td>
<td>-1</td>
<td>ULM</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>0.891</td>
<td>0.543</td>
<td>-1</td>
<td>ULM</td>
</tr>
</tbody>
</table>

Table 2.2 Training data of normal (N) state and abnormal (ULM) state

<table>
<thead>
<tr>
<th>N</th>
<th>ULM</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.45E-9</td>
<td>9.92E-9</td>
</tr>
<tr>
<td>1.11E-8</td>
<td>176.868</td>
</tr>
<tr>
<td>1.53E-8</td>
<td>3.75E-9</td>
</tr>
<tr>
<td>1.4E-8</td>
<td>4.67E-9</td>
</tr>
<tr>
<td>9.91E-9</td>
<td>3.05E-9</td>
</tr>
<tr>
<td>176.868</td>
<td>1.51E-9</td>
</tr>
<tr>
<td>1.56E-8</td>
<td>4.28E-9</td>
</tr>
<tr>
<td>1.25E-8</td>
<td>9.1E-10</td>
</tr>
</tbody>
</table>

Table 2.3 Lagrange coefficient $\alpha$

Only if $\alpha_i > 0$, $x_i$ is called support vector (SV). When $\alpha$ is very small, the influence is very small for the optimal hyper-plane, so we take it as 0. Obviously, the fewer
support vectors are, the less the constraints, the stronger the generalization ability of the SVMs.

Thus this training obtained 2 SVs accounted for 12.5% of training samples. Finally, \( \alpha^*(\alpha_i > 0) \) express the \( \omega \) and \( b \) by Eq. (2.19), (2.20), then optimal hyper-plane can be confirmed as Fig. 2.7, and the training time is 0.01 seconds, the margin is 0.106. The optimal hyper-plane is considered as the classifier of the diagnostic system in the first step.

\[
\omega = \sum \alpha^* y_i x_i (x_i \in SV) \quad (2.19)
\]
\[
b = y_i - \omega \cdot x_i (x_i \in SV) \quad (2.20)
\]

![Fig. 2.7 Optimal Hyper-plane of normal (N) state and abnormal (ULM) state](image)

Similarly, in the second step, unbalance (U) and other abnormal states (LM) will be distinguished. \( P_5 \) and \( P_7 \) are selected for training input data. But we used linear classification for the training data, the corresponding hyper-plane is as shown in Fig. 2.8 (A). Obviously there are some mistakes, so we adopt non-linear classification, and select polynomial function and Gaussian RBF function to train. The hyper-planes are shown in Fig. 2.8 (B), (C), and the training results are shown in Table 2.4.

For the polynomial function and the Gaussian RBF function, execution time is the same, but using the Gaussian RBF function, the margin is larger, and support vectors are fewer, so the generalization ability is stronger. Therefore, the hyper-plane using the
Gaussian RBF function is the optimal hyper-planes of the diagnostic system in the second step.

Table 2.4 Training results corresponding with three different kernel functions

<table>
<thead>
<tr>
<th>Kernel</th>
<th>Execution time</th>
<th>Margin</th>
<th>Support Vectors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>0.01sec</td>
<td>0.000009</td>
<td>16 (100.0%)</td>
</tr>
<tr>
<td>Polynomial</td>
<td>0.01 sec</td>
<td>0.000185</td>
<td>16 (100.0%)</td>
</tr>
<tr>
<td>Gaussian RBF</td>
<td>0.01 sec</td>
<td>0.045363</td>
<td>11 (68.8%)</td>
</tr>
</tbody>
</table>

Fig. 2.8 Optimal hyper-planes corresponding with three different kernel functions

- Unbalance state (U) □ other abnormal state (LM) ◇ support vector

Fig. 2.9 Optimal hyper-plane from looseness (L) state and Misalignment (M) state

In the final step, loose state (L) and misalignment state (M) will be distinguished. P₅ and P₇ are also selected for training input data. The hyper-plane of linear classification is shown as Fig. 2.9. The training time is 0.01 seconds. The margin is 0.0586, and the
number of SVs is 3. It is clear that linear classifications can lead the good results, so there is no need to use non-linear classification. While the classifier of the diagnostic system in the third step, shown as Fig. 2.9, is obtained, the diagnostic system is acquired.

### 2.4.3 Diagnosis and verification

In order to verify the diagnostic capability of the method proposed in this study, the test data, measured in each known state and not used to train the diagnostic system, are used. When inputting the test data into the diagnostic system, they can correctly and quickly diagnose those faults with the possibility steps of the corresponding states. The test data and the diagnosis results are shown in Table 2.5, Table 2.6 and Table 2.7.

<table>
<thead>
<tr>
<th>Test data and diagnosis result in the first step</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>State</strong></td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>ULM</td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>ULM</td>
</tr>
</tbody>
</table>

All the testing results shown above verified that the condition diagnosis method based on SVMs and the SPs proposed in this paper are available in machinery fault diagnosis.
Table 2.6 Test data and diagnosis result in the second step

<table>
<thead>
<tr>
<th>State</th>
<th>Test data</th>
<th>Diagnosis result</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P5 0.205</td>
<td>Y1 1</td>
</tr>
<tr>
<td>U</td>
<td>P7 0.774</td>
<td>Judge U</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td></td>
<td>0.082</td>
<td>Y1 1</td>
</tr>
<tr>
<td>LM</td>
<td>P5 0.317</td>
<td>Y1 -1</td>
</tr>
<tr>
<td></td>
<td>P7 0.803</td>
<td>Judge LM</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td></td>
<td>0.810</td>
<td>Y1 -1</td>
</tr>
</tbody>
</table>

Table 2.7 Test data and diagnosis result in the third step

<table>
<thead>
<tr>
<th>State</th>
<th>Test data</th>
<th>Diagnosis result</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P5 0.317</td>
<td>Y1 1</td>
</tr>
<tr>
<td>L</td>
<td>P7 0.803</td>
<td>Judge L</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td></td>
<td>0.166</td>
<td>Y1 1</td>
</tr>
<tr>
<td>M</td>
<td>P5 0.592</td>
<td>Y1 -1</td>
</tr>
<tr>
<td></td>
<td>P7 0.602</td>
<td>Judge M</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td></td>
<td>0.909</td>
<td>Y1 -1</td>
</tr>
</tbody>
</table>

2.4.4 Comparison with diagnosis result by Neural Network (NN)

In order to compare the performances of SVM and neural network (NN), the construction of NN for diagnosis is shown in Fig. 2.10. The number of neurons in the hidden layer is sixty, the training goal is 0.001, and the number of training epochs is $10^5$. In addition, according to the sequential diagnosis shown in Fig. 2.3, the diagnosis system using NN is also designed three steps, and the neurons of the first layer are the SPs which are selected by the values of DI in each step. The trained result is that NNs
have converged in the first and second steps, but the performance of NN is only 0.1037 in the third step.

“state x” means normal (N), unbalance (U), looseness (L), misalignment (M) states

Fig. 2.10 Neural network for the diagnosis

The test data shown in Table 2.5, Table 2.6 and Table 2.7 are inputted into the trained NN in each step. The diagnostic results are shown in Table 2.8. It is obvious that the diagnosis result of the second step is the best and the accuracy of the third step is lower. Moreover, when the diagnosis system is obtained, the training time using SVMs is about 0.1 second in each step, but NNs are used to take 200 seconds or more.

Table 8 Diagnosis results using NNs in each step

<table>
<thead>
<tr>
<th>Each step</th>
<th>State</th>
<th>Diagnosis accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>First step</td>
<td>normal</td>
<td>83.3%</td>
</tr>
<tr>
<td></td>
<td>abnormal</td>
<td>66.7%</td>
</tr>
<tr>
<td>Second step</td>
<td>unbalance</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>other abnormal</td>
<td>91.7%</td>
</tr>
<tr>
<td>Third step</td>
<td>looseness</td>
<td>66.7%</td>
</tr>
<tr>
<td></td>
<td>misalignment</td>
<td>50%</td>
</tr>
</tbody>
</table>
In addition, although diagnosis accuracy may be raised by changing the construction of NN, such as the number of middle layer, and the number of neurons in the hidden layer, or increasing the number of the training data, the performance of NN is worse than the performance of SVM.

### 2.5 Conclusions

When building an intelligent diagnosis system for rotating machinery, in order to improve the efficiency and the accuracy of the condition diagnosis for plant rotating machinery and distinguish fault types at an early stage, a new condition diagnosis method based on SVMs and SPs was proposed. The methods proposed here can not only replace neural networks and genetic algorithms in the fault diagnosis, but also have a strong learning ability to deal with the fault diagnosis for small samples of learning data, and perform linear classification in a feature space for non-linear classification by kernel function. The practical examples of fault diagnosis for a rotating machine were shown to verify the efficiency of the method proposed in this paper. We will widely use this method to the field of machinery diagnosis.

### References


Chapter 3

Sequential Diagnosis Method using Support Vector Machine and Relative Ratio Symptom Parameters for Precisely Distinguishing Structural Faults of Rotating Machinery

3.1 Introduction

In a real plant, unbalanced, looseness, misalignment, etc. states are the common faults and easily occur. These faults are normally seen and show respective characteristics in the low frequency band, so these faults in the low frequency band are called “structural faults” [1-2]. Structural faults not only impact on equipment performance or product quality, but also add an excessive stress to circumference parts, such as bearing and gear, thereby causing secondary faults. So it is a very important issue that structural faults are diagnosed in early stage [3]. And then the fault caused in early stage is called “early-stage fault”. The degree of early-stage fault is slight, the abnormal characteristics are unclear.

In the field of condition monitoring and machinery diagnosis, symptom parameter (SP) is extremely important as an effective tool. If good SPs that sensitively reflect the feature of the machine states can be found, automatic diagnosis for mechanical failure is possible [1, 4-6]. However, it is not easy to find effective SPs for early-stage faults. And traditional intelligent methods such as neural networks (NN), which are based on statistical learning, require a large number of training samples. In most cases of practical machinery diagnosis, measured signals are not sufficient for conditioning the learning data [7~10]. Moreover, these methods often do not converge during the
learning phase of these networks [11-13].

To solve these problems, after analyzing the features from structural faults of rotating machinery, relative ratio symptom parameters (RRSPs) are defined to identify structural faults of rotating machinery. At the same time, SVMs are adopted as a novel intelligent classifier. SVMs use a structural risk minimization principle to minimize an upper bound based on an expected risk, which leads to a better generalization performance. This study proposes a new intelligent diagnosis method based on RRSPs and SVMs in early stage using sequential diagnosis. These proposed methods have been applied to diagnose structural faults of rotating machinery in early stage, and verification results will show that the methods are accurate and practical.

3.2 Relative ratio symptom parameters (RRSPs)

3.2.1 Distinctive characteristics of structural faults

Structural faults have many categories. However the rotational frequency and these high-order harmonic components of structural faults are not only more prominent than in normal states, but these frequencies and components also show respective characteristics, even though early-stage faults [1].

Generally, in the unbalanced state, the spectrum captures a rotating frequency \( f \), that is the main and dominating 2nd-order harmonic component; this harmonic has sometimes appeared clearly; however, high-order harmonic components of a rotational frequency \( i \cdot f, (i \geq 3) \) have hardly appeared. Moreover, the level of the vibration is related to the channel measuring the signal, and the amplitude is proportional to the second power of the rotational speed. In contrast, the vibration level in the horizontal direction is the strongest, and the vibration level in the axial direction is a little weaker.

When the looseness state is occurring, the influence from the rotational speed is large, and the noise level is also comparatively large, and there is a possibility that impact will
occur. Therefore, the high-order harmonic components of rotational frequency appear with significant amplitude.

While in the misalignment state in the frequency components, $2f_r$, $3f_r$ or $4f_r$, … appear, but then $2f_r$, $3f_r$ or $4f_r$ is significant and dominating, and the amplitude of $f_r$ is minor. Compared with the unbalanced state, the influence of the amplitude from the rotational speed is much lower. However the vibrating level of the axial direction is larger.

On top of these faults, bent shaft, shaft crack also show those characteristics in the low frequency band. The rotational frequencies and their high-order harmonic components can demonstrate the features of the vibration signals and reflect structural faults.

### 3.2.2 Relative ratio symptom parameters (RRSPs)

According to the distinctive characteristics of structural faults, relative ratio symptom parameters (RRSPs) are defined as in Eq. (3.1) ~ (3.10) for diagnosing structural faults of rotating machinery, so RRSPs are also called “exclusive symptom parameters: ESPs” for structural faults.

\[
P_1 = \frac{P_d(f_r)}{P_n(f_r)} \frac{\sum_{i=2}^{20} P_d(i \cdot f_r)}{\sum_{i=2}^{20} P_n(i \cdot f_r)} \tag{3.1}
\]

\[
P_2 = \frac{P_d(2f_r)}{P_n(2f_r)} \frac{P_d(f_r)}{P_n(f_r)} \tag{3.2}
\]

\[
P_3 = \frac{P_d(3f_r)}{P_n(3f_r)} \frac{P_d(f_r)}{P_n(f_r)} \tag{3.3}
\]
Here, \( f_r \) is the rotating frequency. \( P_5(f_r) \), \( P_6(f_r) \) are the spectrum value at \( f_r \) in normal state and abnormal state respectively, \( P_5(i \cdot f_r) \) and \( P_6(i \cdot f_r) \) are the high-order harmonic value at \( i \cdot f_r \) in normal and abnormal state respectively.

\[
P_5 = \frac{P_5(f_r)}{P_5(f_r)}
\]

Here, \( N_n \) and \( N_d \) are the rotation speed of the machine in normal and abnormal state respectively.

\[
P_6 = \frac{\sum_{<1.5f_r} P_6(f)}{\sum_{>1.5f_r} P_6(f)}/\frac{\sum_{<1.5f_r} P_6(f)}{\sum_{>1.5f_r} P_6(f)}
\]

\[
P_7 = \frac{\sum_{i=1}^{n} P_7(f_i)}{\sum_{i=1}^{n} P_7(i \cdot f_r)}/\frac{\sum_{i=1}^{n} P_7(f_i)}{\sum_{i=1}^{n} P_7(i \cdot f_r)}
\]

\[
P_8 = \frac{A_{ad}/A_{rd}}{A_{an}/A_{rn}}
\]

Here, \( A_{ad} \) and \( A_{an} \) are the vibration level of shaft direction and radial direction in normal state respectively, \( A_{ad} \) and \( A_{ad} \) are the vibration level of shaft direction and radial direction in abnormal state respectively. (Vibration level is Standard deviation or root mean square value etc.)
Here, $A_v$ and $A_h$ are the vibration level of vertical direction and horizontal direction in normal state respectively, $A_{vd}$ and $A_{hd}$ are the vibration level of vertical direction and horizontal direction in abnormal state respectively.

\[
P_9 = \frac{A_{vd}}{A_{hd}} \frac{A_{vn}}{A_{hn}}
\]

Here, $\beta_n$ and $\beta_d$ are the skewness in normal state and abnormal state respectively.

### 3.3 Intelligent diagnosis using SVMs and RRSPs

#### 3.3.1 Support vector machines (SVMs)

SVMs are a relatively new computational learning method based on the statistical learning theory, and the basic idea is to create an optimal hyper-plane, as shown in Fig. 3.1, in a 2-dimensional situation [14–18]. Fig. 3.1 shows a series of points for two different classes: the black points for class A and the white circles for class B. The SVMs try to place a linear boundary between the two different classes and orient it in such way that the margin represented by the dotted line is maximized. Furthermore, SVMs attempt to orient the boundary to ensure that the distance between the boundary...
and the nearest data point in each class is maximized. Then, a boundary, such as H, is placed in the middle of this margin and between the two points. The nearest data points used to define the margin are called the support vectors and are represented by the circles. When the support vectors have been selected, the rest of the data feature set is not required as the support vectors can contain all the basic information needed to define the classifier.

When given samples \( S (x_i, y_i) (i = 1, 2, \ldots, N) \), where \( N \) is the number of samples, the samples are assumed to have two classes, \( x_i \in X = \mathbb{R}^N \), \( y_i \in Y = \{-1, +1\} \), and each of classes associates with labels: \( y_i = 1 \) for class A and \( y_i = -1 \) for class B. The goal of SVMs is to define an optimal hyper-plane that divides \( S \) so that all samples with the same label are on the same side of the hyper-plane, and the distance between two classes and the hyper-plane is maximized. In the case of linear data, the hyper-plane can be expressed as follows:

\[
f(x) = \omega \cdot x + b = \sum_{i=1}^{N} \omega_i \cdot x_i + b = 0, (\omega \in \mathbb{R}^N, b \in \mathbb{R})
\]

(3.11)

Where \( \omega \) is \( N \)-dimensional vector, \( x \) is the input data, \( b \) is a classification threshold, and \( \omega, b \) are used to define the position of separating the hyper-plane. The decision function is made using \( \text{sign}(f(x)) \) to create separating hyper-planes that classifies the input data as either belonging to class A or class B. A separating hyper-plane should satisfy the constraints:

\[
f(x_i) = 1, \text{if } y_i = 1;
\]

\[
f(x_i) = -1, \text{if } y_i = -1.
\]

(3.12-1)

or it can be presented in complete equation:

\[
y_i (\omega \cdot x_i + b) \geq 1, i = 1, \ldots, N
\]

(3.12-2)
The SVM method tries to find a unique separating hyper-plane by minimizing $\|\omega\|$ under the constraining conditions. Here $\|\omega\|$ is the Euclidean norm, and the distance between the hyper-plane and the nearest data points of each class is $2/\|\omega\|$. By introducing the Lagrange multiplier $\alpha_i$, the optimal hyper-plane problem transformed into a quadratic problem (QP) as follows:

Maximize

$$Q(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N} \alpha_i \alpha_j y_i y_j x_i x_j$$  \hfill (3.13)

Subject to

$$\sum_{i=1}^{N} \alpha_i y_i = 0$$  \hfill (3.14-1)

$$\alpha_i \geq 0, \ i = 1, \ldots, N$$  \hfill (3.14-2)

Thus, by solving the QP, the coefficient $\alpha_i$ is obtained, which is required to express $\omega$ and $b$ using Eq. (3.15) and Eq. (3.16). When $\alpha$ is very small, the influence is negligible for the optimal hyper-plane, so we take it as zero. If $\alpha_i > 0$, $x_i$ is called a support vector (SV). Obviously, the fewer number of support vectors, the fewer number of constraints, and the SVMs’ generalization ability will be stronger.

$$\omega^* = \sum_{i \in SV} \alpha_i y_i x_i$$  \hfill (3.15)

$$b^* = y_i - \omega^* x_i, i \in SV$$  \hfill (3.16)

When the SVMs are trained, the decision function can be written as

$$y = \text{sign}(\omega^* \cdot x + b^*) = \text{sign}(\sum_{i \in SV} \alpha_i y_i x_i, x + b^*)$$  \hfill (3.17)
For a linear non-separable case, SVMs perform a non-linear mapping of the input vector \( x \) from the input space into a higher dimensional feature space and create the optimal hyper-plane in the feature space; the principle is shown in Fig. 3.2.

This mapping is determined by the kernel function \( K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j) \). In fact, when applying a kernel function, the learning in the feature space does not require an explicit evaluation of \( \phi \). Any function that satisfies Mercer’s theorem can be used as a kernel function to compute a dot product in the feature space. There are different kernel functions used in SVMs. In this work, linear, polynomial and Gaussian RBF functions were evaluated and formulated in Table 3.1.

![Fig. 3.2 Transformation to feature space from input space](image)

**Table 3.1 Formulation for kernel functions**

<table>
<thead>
<tr>
<th>Kernel function</th>
<th>( K(x_i, x_j) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>( x_i \cdot x_j )</td>
</tr>
<tr>
<td>Polynomial</td>
<td>( (x_i \cdot x_j + 1)^q )</td>
</tr>
<tr>
<td>Gaussian RBF</td>
<td>( \exp\left{-\frac{</td>
</tr>
</tbody>
</table>

Upon selecting the appropriate kernel function, classification accuracy is improved, so the selection of the appropriate kernel function is very important, and influences the goal function in Eq. (3.18) and the decision function in Eq. (3.19).

\[
Q(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N} \alpha_i \alpha_j y_i y_j K(x_i, x_j) 
\]  

(3.18)
\[
y = \text{sign}(\omega^* \cdot x + b^*) = \text{sign}(\sum \alpha_i^* y_i K(x_i, x) + b^*)
\] (3.19)

SVMs’ decision function is similar to that of neural networks as the output \( y \) is a linear combination of intermediate nodes, while each intermediate node corresponds to a support vector, as shown in Fig. 3.3. SVMs can transform non-linear data into a feature space using a kernel function. SVMs can construct an optimal classifier, which has a strong generalization capacity when only a small amount of training samples are available.

These features have an important role in the field of mechanical fault diagnosis.

![Fig. 3.3 SVMs’ classification diagram](image)

**3.3.2 Intelligent diagnosis using SVM and RRSPs**

![Fig. 3.4 Intelligent diagnosis based on SVMs and RRSPs](image)
This paper presents a method using SVM and RRSPs for intelligent machine condition monitoring and fault diagnosis. The flowchart of the method is shown in Fig. 3.4. The data of RRSPs, as the object of follow-on process, are extracted from the vibration signals in plant for intelligent diagnosis. Put them into SVM as the training data then the optimal hyper-plane obtained. If the training data can be linearly separated by SVM, optimal hyper-plane will be obtained by linear classification. If the training data cannot be linearly separated, appropriate kernel function will be chosen to search the optimal hyper-plane using SVM. But using kernel function, the training data are mapped to higher-order feature space. It follows that the generalizing capability decreases and the computational complexity increases with the dimension becomes more. So the linear classification is used as much as possible [16]. After the optimal hyper-plane obtained, the testing data are analyzed using intelligent diagnosis method.

3.3.3 System of sequential diagnosis

![Flowchart of the method](Image)

Fig. 3.5 System of sequential diagnosis using multi-SVMs

SVM is a 2-class classifier, however many conditions such as structural faults will be often diagnosed, and it is difficult to find the excellent SPs that can be used together. For
efficient diagnosis, then using multi-SVMs, the system of sequential diagnosis is found to detect many faults and identify fault types from rotating machinery in early stage, shown as Fig. 3.5 [19]. The signal from the rotating machine is captured in four conditions: normal, unbalanced, looseness and misalignment. The main aim of fault diagnosis is to recognize the state of the machine. The first step is finding whether the device is normal state (N) or abnormal state (UN). If abnormal, the next step is to distinguish the state between unbalanced state (U) and other abnormal states (LM). The last step is to segregate the fault into looseness state (L) or misalignment state (M).

3.4 Application to an experiment

3.4.1 Experimental system and signal conditions

Fig. 3.6 Experimental equipment for intelligent diagnosis

Fig. 3.6 shows the experimental rotating machine with middle rotation speed for diagnosis test, and the measuring points of accelerometers. The most commonly occurring structural faults of rotating machinery in early stage are shaft misalignment, unbalanced and looseness states. The accelerometer is used to measure vibration signals. The sampling frequency of signal measurement is 5000Hz, and the sampling time is 20s.
the rotational speed is 1500rpm. The time-domain waveforms of vibration signals in each state are shown in Fig. 3.7. For calculating the RRSPs shown as Eq. (3.1) ~ (3.10), the vibration signals in each state are divided into 20 parts, and the spectra with 0~300Hz is obtained by a low-pass filter.

![Vibration signals in each state](image)

(A) Normal state (B) Unbalanced state (C) Looseness state (D) Misalignment state

Fig. 3.7 The time-domain waveforms of vibration signals in each state

### 3.4.2 Sensitivity of RRSPs

In order to evaluate the sensitivity of a SP for distinguishing two states, discrimination index (DI) is defined as follows [1].

\[
DI = \frac{|\mu_2 - \mu_1|}{\sqrt{\sigma_1^2 + \sigma_2^2}} \tag{3.20}
\]

Here, \( \mu_1 \) and \( \mu_2 \) are the mean values of SP calculated by the data in state 1 and state 2 respectively. \( \sigma_1 \) and \( \sigma_2 \) are their standard deviations. When the value of DI is 1.65, discrimination rate (DR) is 95%, and when DI is 0.8, DR is 78.8%. Here, DR is calculated as follows.

\[
DR = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-DI} \exp(-\frac{\mu_1^2}{2})d\mu \tag{3.21}
\]

It is obvious that the larger value of DI would get, the larger value of DR will be, and
consequently, the better the SP will be. Using SVMs for diagnosis classification, when DI is more than 1.65, linear classification is usually utilized to create optimal hyper-plane, when DI is from 0.8 to 1.65, appropriate kernel function is selected to classify in feature space, but DI is less than 0.8, the sensitivity of this SP is very lower, this SP cannot be used [3].

The values of DI between the RRSPs of normal (N) and abnormal (UN), unbalanced (U) and other abnormal (LM), looseness (L) and misalignment (M) in the experiment are shown in Fig. 3.8. The percentage that the values DI of the RRSPs index more than 1.65, from 0.8 to 1.65, from less than 0.8 are 46.7%, 30%, 23.3% respectively, the utilization rate of the RRSPs is 76.7%.

However, a large set of symptom parameters have been defined in the pattern recognition field [20-23]. Here, 10 SPs, commonly used for the fault diagnosis of plant machinery, are shown as follows, so these parameters are called “general symptom parameters: GSPs” [1].

\[
P_{11} = \sqrt{\frac{\sum_{i=1}^{N} x_i^2}{N}} \quad (3.22)
\]

\[
P_{12} = \sigma / \sqrt{\bar{x}} \quad (3.23)
\]

\[
P_{13} = \frac{\sum_{i=1}^{N} (x_i - \bar{x})^3}{N \cdot \sigma^3} \quad (3.24)
\]
Chap. 3 Sequential Diagnosis Method using SVMs and RRSPs

\[ P_{14} = \frac{\sum_{i=1}^{N} (x_i - \bar{x})^4}{N \cdot \sigma^4} \]  
(3.25)

\[ P_{15} = P_{12} / \sigma \]  
(3.26)

Here, \( x_i (i=1,2,\ldots,N) \) is digital data of vibration signal, \( N \) is the number of data. \( \bar{x} \) and \( \sigma \) are the mean value and the standard deviation.

\[ P_{16} = \frac{\sum_{i=1}^{I} (f_i - \bar{f})^3 \cdot F(f_i)}{I \cdot \sigma^3} \]  
(3.27)

\[ P_{17} = \frac{\sum_{i=1}^{I} (f_i - \bar{f})^4 \cdot F(f_i)}{I \cdot \sigma^4} \]  
(3.28)

\[ P_{18} = \sqrt{\frac{\sum_{i=1}^{I} f_i^4 \cdot F(f_i)}{\sum_{i=1}^{I} f_i^2 \cdot F(f_i)}} \]  
(3.29)

\[ P_{19} = \frac{\sum_{i=1}^{I} f_i^3 \cdot F(f_i)}{\sqrt{\sum_{i=1}^{I} F(f_i) \sum_{i=1}^{I} f_i^4 \cdot F(f_i)}} \]  
(3.30)

\[ P_{20} = \sum_{i=1}^{I} F(f_i) \]  
(3.31)

Here, \( I \) is the number of spectrum lines, \( f_i \) is the frequency from 0 Hz to the maximum analysis frequency, \( F(f_i) \) is the power spectrum values at frequency \( f_i \). \( \bar{f} \) and \( \sigma \) are the mean value and the standard deviation of the analysis frequency, shown as

\[ \bar{f} = \frac{\sum_{i=1}^{I} f_i \cdot F(f_i)}{\sum_{i=1}^{I} F(f_i)} \quad \text{and} \quad \sigma = \sqrt{\frac{\sum_{i=1}^{I} (f_i - \bar{f})^2 \cdot F(f_i)}{I}}. \]
The DI values of GSPs in three steps in the experiment are shown in Fig. 3.9. The percentage that the DI values of GSPs index more than 1.65, from 0.8 to 1.65, from less than 0.8 are 13.3%, 43.3%, 43.4% respectively, the utilization rate of the GSPs is 56.6%.

It is obvious that the sensitivity and utilization rate of RRSPs are high, and it has been verified by the different structural faults experiments. Even if there will be a slight structural fault at an early stage, the sensitivity of RRSPs can be high. So it is very efficient to diagnose structural faults of rotating machinery using RRSPs.

![Discrimination index of GSPs](image)

**Fig. 3.9 Discrimination index of GSPs**

### 3.4.3 Fault diagnosis

According to sequential diagnosis as shown in Fig. 3.5, in the first step, normal state will be distinguished from three abnormal states. Firstly $P_7$ and $P_8$, the larger DI value of RRSPs (>1.65), are selected from normal state and abnormal states, to build the training data, and $Y_n$ is defined +1 for N and -1 for UN to represent two states in Table 3.2. Secondly using SVM, the training data are trained to achieve Lagrange coefficient $\alpha$ as Table 3.2. there are 3 SVs for expressing the $\omega$ and $b$ by Eq. (3.15), (3.16), then the hyper-plane is obtained as shown in Fig. 3.10, and the execution time is 0.01 seconds, the margin is 0.1345. Obviously, the training data are divided correctly, so the hyper-plane is optimal for the standard of intelligent diagnosis in the first step.
Chap. 3 Sequential Diagnosis Method using SVMs and RRSPs

![Optimal hyper-plane of normal state and abnormal state in the first step](image)

Table 3.2 Training data in the first step and Lagrange coefficient trained

<table>
<thead>
<tr>
<th>Signal state</th>
<th>Training data</th>
<th>Lagrange coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Xₙ</td>
<td>Yₙ</td>
</tr>
<tr>
<td>No.</td>
<td>P₇</td>
<td>P₈</td>
</tr>
<tr>
<td>Normal state</td>
<td>(N)</td>
<td></td>
</tr>
<tr>
<td>x₁</td>
<td>1.06</td>
<td>0.98</td>
</tr>
<tr>
<td>x₂</td>
<td>0.99</td>
<td>1.01</td>
</tr>
<tr>
<td>x₃</td>
<td>1.06</td>
<td>0.99</td>
</tr>
<tr>
<td>x₄</td>
<td>0.99</td>
<td>1.02</td>
</tr>
<tr>
<td>x₅</td>
<td>0.96</td>
<td>0.99</td>
</tr>
<tr>
<td>x₆</td>
<td>1.01</td>
<td>1.02</td>
</tr>
<tr>
<td>Abnormal state</td>
<td>(UN)</td>
<td></td>
</tr>
<tr>
<td>x₇</td>
<td>1.39</td>
<td>1.31</td>
</tr>
<tr>
<td>x₈</td>
<td>1.62</td>
<td>1.28</td>
</tr>
<tr>
<td>x₉</td>
<td>1.68</td>
<td>1.26</td>
</tr>
<tr>
<td>x₁₀</td>
<td>0.81</td>
<td>1.21</td>
</tr>
<tr>
<td>x₁₁</td>
<td>0.74</td>
<td>1.17</td>
</tr>
<tr>
<td>x₁₂</td>
<td>0.9</td>
<td>1.22</td>
</tr>
</tbody>
</table>

Similarly, in the second step, unbalanced (U) and other abnormal states (LM) will be distinguished. P₄ and P₉ (0.8<DI<1.65) are selected for training data. But linear classification is used for the training data, the corresponding hyper-plane is shown in
Fig. 3.11 (A). Obviously two states cannot be well diagnosed, so non-linear classification will be adopted and selected polynomial function and Gaussian RBF function for training. The hyper-planes are shown in Fig. 3.11 (B), (C), and the training results are shown in Table 3.3.

![Hyper-planes](image)

(A) Linear  (B) Polynomial  (C) Gaussian RBF

- Unbalanced state (U)
- Other abnormal state (LM)
- SVs

Fig. 3.11 Optimal hyper-plane corresponding with kernel functions in the second step

<table>
<thead>
<tr>
<th>Kernel function</th>
<th>Margin [unit]</th>
<th>Execution Time [sec]</th>
<th>support vectors (SVs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>0.000007</td>
<td>0.1</td>
<td>14 (93.3%)</td>
</tr>
<tr>
<td>Polynomial</td>
<td>0.022481</td>
<td>0.01</td>
<td>5 (33.3%)</td>
</tr>
<tr>
<td>Gaussian RBF</td>
<td>0.006093</td>
<td>0.01</td>
<td>7 (46.7%)</td>
</tr>
</tbody>
</table>

Table 3.3 Training results corresponding with kernel functions in the second step

For the polynomial function and the Gaussian RBF function, execution time is same, but using the polynomial function, the margin becoming larger, and SVs being fewer, so the generalization ability goes stronger. The three functions are also applied to classify the same test data, the classification precision is 80%, 100%, 93.3% respectively. It is verified that the precision of classification is the highest using the polynomial function by the experiment results. Thus the hyper-plane shown in Fig. 3.11(B) is optimal for the standard of intelligent diagnosis in the second step.
In the third step, $P_2$ and $P_9$ (0.8<DI<1.65) are selected to obtain the training data for identifying looseness state (L) and misalignment state (M). Firstly the training data are trained using linear classification. However the corresponding hyper-plane is not satisfying as shown in Fig. 3.12 (A). Then polynomial function is selected to create the hyper-plane as shown in Fig. 3.11 (B). Obviously the hyper-plane is optimal for the standard of intelligent diagnosis in the third step.

Then the intelligent system of sequential diagnosis has been built to detect many faults and identify fault types from structural faults of rotating machinery. In order to verify the effectiveness of the sequential diagnosis method proposed in this study, the vibration signals, which states (N, U, L, M) are known, are measured from early-stage structural faults of rotating machinery, and are calculated RRSPs for the test data. Then the test data are input to the intelligent system of sequential diagnosis, and are judged intelligently in each step. Table 3.4 is the results of sequential diagnosis in each step.

All the testing results shown above verified that the sequential diagnosis method based on RRSPs and SVMs proposed in this paper is available in early-stage fault structural faults diagnosis.
3.5 Conclusions

When building an intelligent system of sequential diagnosis for rotating machinery, in order to improve the efficiency and the accuracy of the intelligent diagnosis for plant rotating machinery and distinguish fault types at an early stage, a new intelligent diagnosis method using RRSPs and SVMs was proposed, and the effectiveness was proved according to the experiment. The superiority of the method proposed in this paper can be explained by the following points:

Table 3.4 Diagnostic results from early-stage structural faults of rotating machinery

<table>
<thead>
<tr>
<th>Test signal</th>
<th>Test data $X_t$</th>
<th>Output data $Y_t$</th>
<th>Diagnosis result</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$1^{st}$ step</td>
<td>$2^{nd}$ step</td>
<td>$3^{rd}$ step</td>
</tr>
<tr>
<td>Normal state (N)</td>
<td>0.938 1.004</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.915 1.018</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Unbalanced state (U)</td>
<td>0.387 1.354</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.242 3.145</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>Looseness state (L)</td>
<td>0.033 0.775</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.250 1.482</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>Misalignment state (M)</td>
<td>0.207 2.267</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.280 0.937</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>Remarks</td>
<td>$1^{st}$ step: 1-normal state (N), -1-abnormal state (UN)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$2^{nd}$ step: 1-unbalanced state (U), -1-other abnormal state (LM)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$3^{rd}$ step: 1-looseness state (L), -1-misalignment state (M)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
(1) RRSPs are a new type of exclusive symptom parameters (ESPs) for detecting and identifying structural faults of rotating machinery. The sensitivity of RRSPs is very higher, even if there will be a slight fault at an early stage, it is visible that the values of RRSPs can change.

(2) The diagnosis method based on SVMs not only has a strong adaptable ability to deal with fault diagnosis with small number of training samples, but also solves the non-convergent problem that frequently occurs when traditional diagnostic methods are used.

(3) Diagnosis method using SVMs can perform a linear classification in a feature space for non-linear classification using a kernel function; therefore, when the feature of fault signal is not obvious at an early stage, the resulting identification of the equipment condition is also excellent.

(4) The efficiency of sequential diagnosis proposed here has been verified by applying it to a practical structural faults diagnosis of rotating machinery, and its performance is high. In fact, the high performance of sequential diagnosis using RRSPs and SVMs is attributed primarily to RRSPs’ high sensitivity and SVMs’ generalization capability.

This method will be used widely in the of condition monitoring and machinery diagnosis.

3.5 References


[3] Hongtao Xue and Peng Chen: Condition Diagnosis Method Based on Symptom


(2007)


Chapter 4

Condition Diagnosis Method using Distinctive Frequency Components and Support Vector Machines under Varied Operating Conditions

4.1 Introduction

Rotating machinery, which accounts for approximately 40% of all machinery, is one of the most commonly used types of equipment and plays an important role in daily production [1]. Almost all manufacturing processes involve correlated rotating machines. When any one of these machines deviates beyond their specified limits or when any component of the structures of these machines loses its designed structural integrity, a fault may occur. Not only does the fault lead to adverse effects on equipment performance and product quality, but the fault may also intensify secondary faults that could be created. For example, misaligned machines result in high vibrations and premature wear of the bearings, leaking shaft seals or hot couplings; this misalignment can then reduce the machines’ performance. This type of fault is referred to as a misalignment state. In addition to misalignment, there are often unbalanced states and looseness states, which are called “structural faults” [1]. Structural faults easily occur in rotating machinery; therefore, the detection of these faults is extremely important. A quick and correct fault diagnostic system should collect and process all the available information to detect any typical changes that are indicative of the faults; the system should also help avoid product quality problems and facilitate preventive maintenance [2-3].

Generally, when a computer is used for on-line monitoring and condition diagnosis of
plant machinery, symptom parameters (SPs) are required to express the information indicated by vibration signals [3-5] and then, to detect the incipient fault of the machine components and reduce the possibility of catastrophic damage and down time [6-10]. These SPs include time domain features, such as mean, root mean square, variance, skewness and kurtosis [11-12] as well as frequency domain features, such as the feature frequency and the amplitudes of the fast Fourier transform (FFT) spectrum [13-15]. However, some plants face various problems:

1. It is not easy to find effective SPs for some faults because the fault categories are large and complex [16].
2. The sensitivity of SPs is not high when under the influence of noise and varied operating conditions [17]. Indeed, SPs from different operating conditions are very different and the same training data cannot be used to diagnose the faults from different operating conditions.
3. Traditional intelligent methods such as neural networks (NN), which are based on statistical learning, require a large number of training samples. However, in most cases of practical machinery diagnosis, measured signals are not sufficient for conditioning the learning data. Moreover, these methods often do not converge during the learning phase of these networks.

To solve these problems, after analyzing the features of structural faults under varied operating conditions, including no-load, fixed load, varied loads and different defect sizes, the common distinctive frequency components (DFCs) are detected to distinguish the structural faults. At the same time, SVMs are adopted as a novel intelligent classifier. SVMs use a structural risk minimization principle to minimize an upper bound based on an expected risk, which leads to a better generalization performance. This study proposes a structural fault diagnosis based on DFCs and SVMs under varied operating conditions using sequential diagnosis. Fig. 4.1 shows the flow diagram for the proposed procedure. These proposed methods have been applied to diagnose structural faults in rotating machinery, and verification results will show...
that the methods are accurate and practical.

![Flow chart of fault diagnosis system](image)

**Fig. 4.1 Flow chart of fault diagnosis system**

### 4.2 Distinctive frequency components (DFCs) of structural faults

#### 4.2.1 Rotating frequency and its high-order harmonic components

The rotational frequency of a rotating machine is the number of revolutions per unit time, mostly represented as per second, and the rotational speed \((N)\) is represented as revolutions per minute (rpm). Thus, the relation between the rotational frequency \((f_r)\) and rotational speed \((N)\) is shown in Eq. (4.1).

\[
f_r = \frac{N}{60}
\]  

(4.1)

High-order harmonic components of the rotational frequency \((i \cdot f_r, \ i = 2, 3, \cdots)\) are the corresponding multiples of the rotating frequency \((f_r)\). Fig. 4.2 shows the frequency components that are of special concern extracted from a structural fault signal using vibration spectrum analysis. Here \(f_{so}\) is the natural frequency of the rotor system, \(f_{uo}\)
Chap. 4 Condition Diagnosis Method using DFCs and SVMs

is the natural frequency of other parts, and \( f_{\text{max}} \) is the maximum frequency for analysis [1].

![Figure 4.2 Frequency components for analysis](image)

![Figure 4.3 Part of spectrum of an abnormal state](image)

In theory, when a machine is operating properly at a constant speed, the rotational frequency and its high-order harmonic components are easily calculated. However, in a real plant, for various reasons such as voltage, load, the summation of driving forces and working resistances cannot constantly be zero; therefore, the rotational speed fluctuates. Therefore the experimental rotational frequency and its high-order harmonic components can also fluctuate at values close to those corresponding theoretical values, and this is especially obvious under abnormal conditions. The fluctuating value of the rotational frequency varies directly with the fluctuation of the speed, so the fluctuating range is related to the stability of the experimental system. Fig. 4.3 is an example where the rotational speed is 1200 rpm in a misalignment state, and the maximum value of the fluctuating speed is 120 rpm. The theoretical values of \( 3f_r, 4f_r, 5f_r \) and \( 6f_r \) correspond...
to 60 Hz, 80 Hz, 100 Hz and 120 Hz, respectively. However the actual values are 61.9 Hz, 81.2 Hz, 98.6 Hz and 121 Hz.

4.2.2 Distinctive frequency components (DFCs) of structural faults

Structural faults are normally seen in the low frequency band. By contrasting and analyzing the spectrums from many experiments, the rotational frequency and its high-order harmonic components of structural faults are not only more prominent than in normal states, but these frequencies and components also show respective characteristics.

Generally, in the unbalance state, the spectrum captures a rotating frequency \( f_r \) that is the main and dominating 2nd-order harmonic component; this harmonic has sometimes appeared clearly; however, high-order harmonic components of a rotational frequency \( i \cdot f_r \) (\( i \geq 3 \)) have hardly appeared. Moreover, the level of the vibration is related to the channel measuring the signal, and the amplitude is proportional to the second power of the rotational speed. In contrast, the vibration level in the horizontal direction is the strongest, and the vibration level in the axial direction is a little weaker. Fig. 4.4 (A) and (B) are shown as an example of the spectrums between the normal state and an unbalanced fault in the horizontal direction.

When the looseness state is occurring, the influence from the rotational speed is large, and the noise level is also comparatively large, and there is a possibility that impact will occur. Therefore, the high-order harmonic components of rotational frequency appear with significant amplitude. The features of the looseness state in the vertical direction are shown as Fig. 4.4 (C).

While in the misalignment state in the frequency components, \( 2f_r \) or \( 3f_r, 4f_r, \ldots \) appear, but then \( 2f_r, 3f_r \) or \( 4f_r \) is significant and dominating, and the amplitude of \( f_r \) is minor. Compared with the unbalanced state, the influence of the amplitude from the rotational speed is much lower. However the vibrating level of the axial direction is
larger. Fig. 4.4 (D) is shown as an example of the misalignment fault in the axial direction.

Fig. 4.4 Comparison of spectrums between the normal state and structural faults at a speed of 1200 rpm, the rectangles represent the fluctuating ranges of $i \cdot f_r$: (A) the normal states in the horizontal direction; (B) an unbalanced state in the horizontal direction; (C) a looseness state in the vertical direction; (D) a misalignment state in the axial direction.
The rotational frequencies and their high-order harmonic components can demonstrate features of the vibration signals and reflect structural faults. In fact, the features of structural faults have manifested mainly in the frequency components from \( f_r \sim 5f_r \), and sometimes can extend the predictability of the frequency components. Therefore, the frequency components are called “distinctive frequency components (DFCs)” to diagnose structural faults.

### 4.2.3 Normalization and extraction of distinctive frequency components (DFCs)

When a machine is working at a set speed, the objective conditions, such as load or defect size of an abnormal condition, may sometimes change; at the same state, the features of the FFT spectrums will not change, though their amplitudes may change. To compare spectrum values from two signals under varied operating conditions, the frequency spectrums must first be normalized under the same standard according to the formula shown in Eq. (4.3).

\[
\mu_A = \frac{1}{M} \sum_{i=1}^{M} A(f(i)) \\
H(f(i)) = \frac{A(f(i))}{\mu_A} (i = 1, 2, \ldots, M)
\]

Here \( M \) is the number of the most salient frequency band. \( f(i) \) is the frequency at the point, \( A(f(i)) \) is the amplitude value corresponding \( f(i) \). \( \mu_A \) is the mean value of the frequency in the most salient frequency band. \( H(f(i)) \) is the normalized amplitude value. If \( A(f_r) \) and \( A(2f_r) \) are the amplitudes of the spectrum at the points \( f_r, 2f_r \) from one signal under a no load operating condition, \( A'(f_r) \) and \( A'(2f_r) \) are from another signal under the 120 kg load, \( \mu_A \) and \( \mu'_A \) are the mean-values. Generally

\[
A'(f_r) > A(f_r) > 0
\]
\[
A'(2f_r) > A(2f_r) > 0
\]
\[
\mu'_A > \mu_A > 0
\]  
(4.5)  
(4.6)

\(H(f_r), H(2f_r)\) and \(H'(f_r), H'(2f_r)\) are the corresponding amplitudes normalized from the signal differences.

\[
\frac{H(f_r)}{H(2f_r)} = \frac{A(f_r)}{\mu_A} \frac{A(2f_r)}{\mu_A} = \frac{A(f_r)}{A(2f_r)}
\]  
(4.7)

\[
\frac{H'(f_r)}{H'(2f_r)} = \frac{A'(f_r)}{\mu_A} \frac{A'(2f_r)}{\mu_A} = \frac{A'(f_r)}{A'(2f_r)}
\]  
(4.8)

\(\frac{H'(f_r)-H(f_r)}{H(f_r)}\) is the normalized rate difference for the same state under varied operating conditions, \(\frac{A'(f_r)-A(f_r)}{A(f_r)}\) is the original rate difference.

The relation is

\[
\frac{H'(f_r)-H(f_r)}{H(f_r)} = \frac{A'(f_r)-A(f_r)}{A(f_r)} = \frac{H'(f_r)}{H(f_r)} - 1 \frac{A'(f_r)}{A(f_r)} - 1
\]

\[
= \frac{A'(f_r)}{A(f_r)} \cdot \frac{\mu_A - \mu'_A}{\mu'_A} < 0
\]  
(4.9)

After the spectrum from a signal is normalized using Eq. (4.3), it is shown that the feature cannot change using Eq. (4.7) and (4.8), and the normalized rate difference under the varied operating conditions can be reduced. Thus, it is beneficial to detect the common DFCs under varied operating conditions.

The DFCs’ values are the amplitude of the spectrum at the corresponding points \(f_r, 2f_r, \ldots, Nf_r\). Here \(N\) is the number of the most rotating frequency components that can reflect the fault feature under varied operating conditions. Because the rotating frequency and its high-order harmonic components can fluctuate near those corresponding theoretical values in a real plant, the DFCs’ values for the structural
faults such as \( H_i = H(i \cdot f_r) \) \( (i = 1, 2, \cdots, N) \) can be extracted using a multi-band pass filter in the range \( (i \cdot f_r - \Delta f, i \cdot f_r + \Delta f) \). Here \( \Delta f \) is the fluctuating range from the theoretical value. For example in Fig. 4.4, the rectangles represent the fluctuating ranges. The data from the DFCs, as the object of follow-on process, are extracted from the vibration signals collected in plant for condition diagnosis.

### 4.3 Fault diagnosis based on SVMs

#### 4.3.1 Support vector machines (SVMs)

SVMs are a relatively new computational learning method based on the statistical learning theory, and the basic idea is to create an optimal hyper-plane, as shown in Fig. 4.5, in a 2-dimensional situation [18–20]. Fig. 4.5 shows a series of points for two different classes: the black points for class A and the white circles for class B. The SVMs try to place a linear boundary between the two different classes and orient it in such way that the margin represented by the dotted line is maximized. Furthermore, SVMs attempt to orient the boundary to ensure that the distance between the boundary and the nearest data point in each class is maximized. Then, a boundary, such as \( H \), is placed in the middle of this margin and between the two points. The nearest data points used to define the margin are called the support vectors and are represented by the circles. When the support vectors have been selected, the rest of the data feature set is not required as the support vectors can contain all the basic information needed to define the classifier.
When given samples \( S(x_i, y_i) \) \( (i=1, 2, \ldots, N) \), where \( N \) is the number of samples, the samples are assumed to have two classes, \( x_i \in X = \mathbb{R}^N \), \( y_i \in Y = \{+1, -1\} \), and each of classes associates with labels: \( y_i = 1 \) for class A and \( y_i = -1 \) for class B. The goal of SVMs is to define an optimal hyper-plane that divides \( S \) so that all samples with the same label are on the same side of the hyper-plane, and the distance between two classes and the hyper-plane is maximized. In the case of linear data, the hyper-plane can be expressed as follows:

\[
f(x) = \omega \cdot x + b = \sum_{i=1}^{N} \omega_i \cdot x_i + b = 0, (\omega \in \mathbb{R}^N, b \in \mathbb{R}) \quad (4.10)
\]

Where \( \omega \) is \( N \)-dimensional vector, \( x \) is the input data, \( b \) is a classification threshold, and \( \omega \) and \( b \) are used to define the position of separating the hyper-plane. The decision function is made using \( \text{sign}(f(x)) \) to create separating hyper-planes that classifies the input data as either belonging to class A or class B. A separating hyper-plane should satisfy the constraints:

\[
f(x_i) = 1, \text{ if } y_i = 1;
\]

\[
f(x_i) = -1, \text{ if } y_i = -1; \quad (4.11-1)
\]

or it can be presented in complete equation:

\[
y_i(\omega \cdot x_i + b) \geq 1, i = 1, \ldots, N \quad (4.11-2)
\]

The SVM method tries to find a unique separating hyper-plane by minimizing \( \|\omega\| \) under the constraining conditions. Here \( \|\omega\| \) is the Euclidean norm of, and the distance between the hyper-plane and the nearest data points of each class is \( 2/\|\omega\| \). By introducing the Lagrange multiplier \( \alpha_i \), the optimal hyper-plane problem transformed into a quadratic problem (QP) as follows:

Maximize
Chap. 4 Condition Diagnosis Method using DFCs and SVMs

\[ Q(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1, i \neq j}^{N} \alpha_i \alpha_j x_i x_j y_i y_j \]  

(4.12)

Subject to

\[ \sum_{i=1}^{N} \alpha_i y_i = 0 \]  

(4.13-1)

\[ \alpha_i \geq 0, i = 1, \ldots, N \]  

(4.13-2)

Thus, by solving the QP, the coefficient \( \alpha_i \) is obtained, which is required to express \( \omega \) and \( b \) using Eq. (4.14) and (4.15). If \( \alpha_i > 0 \), \( x_i \) is called a support vector (SV). When \( \alpha \) is very small, the influence is negligible for the optimal hyper-plane, so we take it as zero. Obviously, the fewer number of support vectors, the fewer number of constraints, and the SVMs’ generalization ability will be stronger.

\[ \omega^* = \sum_{i \in SV} \alpha_i y_i x_i \]  

(4.14)

\[ b^* = y_i - \omega^* x_i, i \in SV \]  

(4.15)

When the SVMs are trained, the decision function can be written as

\[ y = sign(\omega^* \cdot x + b^*) = sign(\sum \alpha_i^* y_i x, x + b^*) \]  

(4.16)

For a linear non-separable case, SVMs perform a non-linear mapping of the input vector \( x \) from the input space into a higher dimensional feature space and create the optimal hyper-plane in the feature space; the principle is shown in Fig. 4.6.
This mapping is determined by the kernel function $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$. In fact, when applying a kernel function, the learning in the feature space does not require an explicit evaluation of $\phi$. Any function that satisfies Mercer’s theorem can be used as a kernel function to compute a dot product in the feature space. However, taking into account the noise with the error penalty parameter $C$, the optimal hyper-plane separating the data can be obtained. There are different kernel functions used in SVMs. In this work, linear, polynomial and Gaussian RBF functions were evaluated and formulated in Table 4.1.

Table 4.1 Formulation for kernel functions

<table>
<thead>
<tr>
<th>Kernel</th>
<th>$K(x, x_i)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>$x \cdot x_i$</td>
</tr>
<tr>
<td>Polynomial</td>
<td>$(x \cdot x_i + 1)^q$</td>
</tr>
<tr>
<td>Gaussian RBF</td>
<td>$\exp(-\left|x - x_i\right|^2/2\sigma^2)$</td>
</tr>
</tbody>
</table>

Upon selecting the appropriate kernel function and finding optimal parameter values, classification accuracy is improved, so the selection of the appropriate kernel function and the model parameters is very important, and influences the goal function in Eq. (4.17) and the decision function in Eq. (4.18).
4.3.2 Fault diagnosis based on DFCs and SVMs

This work presents a method based on DFCs and SVMs for intelligent machine condition monitoring and fault diagnosis. The flowchart of the method is shown in Fig. 4.8. The data from the DFCs, as the object of a follow-on process, are extracted from the vibration signals collected in a plant and are used for condition diagnosis. These data are then put into the SVMs as the training data and the optimal hyper-planes are
obtained. Generally if the training data can be linearly separated using SVMs, an optimal hyper-plane will be obtained by linear classification. If the training data cannot be linearly separated, an appropriate kernel function and its optimal parameter value will be chosen to search for the optimal hyper-plane using the SVMs. When using a kernel function, the training data are mapped to a higher-order feature space. It follows that the generalizing capability decreases and the computational complexity and dimensions increase. As far as possible, namely, if the diagnosis accuracy of the linear classifier is higher than 70%, the linear classifier is selected to perform to construct the diagnosis system; else the non-linear classification will be selected. Thus the training classification accuracy is always higher than 70% when the diagnostic system is set up. After the optimal hyper-plane is obtained, the testing data are analyzed using an intelligent diagnosis method. Now, we are also studying on the fuzzy-linear classification and shall report the result in the near future.

Fig. 4.8 Fault diagnosis based on DFCs and SVMs

It is well known that SVMs are a two-class classifier, and structural faults are large and complex. In order to achieve an intelligent and efficient diagnosis, sequential diagnosis was proposed [1, 16]. Then the diagnostic system was designed to consist of a number of SVMs and was able to diagnose the faults one by one as shown in Fig. 4.9.
The first SVMs are used to detect faults while the others are used to identify the fault type. In order to improve performance, it is recommended that conditions that represent more frequent faults should be placed in the diagnostic system.

![Flow chart of diagnostic system using DFCs and SVMs](image)

Fig. 4.9 Flow chart of diagnostic system using DFCs and SVMs

The first step is to set up the diagnostic system. In other words, the DFCs of two classes are selected to recombine the training data in each step, and the SVMs are trained to obtain the optimal hyper-plane. In the first step, the diagnosis goal is to detect the faults. So these two classes are the normal and abnormal states. In a normal state, the DFCs $x_i$ ($i = 1, 2, \cdots, I$) are extracted to associate with labels and to recombine the training data $(x_i, y_i)$ ($i = 1, 2, \cdots, I$) for a normal state. However, the abnormal state includes all of the faults, such as an unbalanced, looseness and misalignment state. So the DFCs for these faults are extracted to recombine the DFCs of the class; then, the training data $(x_j, y_j)$ ($j = 1, 2, \cdots, J$) are obtained to associate with labels $y_j = -1$, ($j = 1, 2, \cdots, J$). These data are input into the first SVMs to train using a linear function first. If the classification accuracy is not high, a non-linear function is selected to train until the optimal hyper-plane is detected. In the second step, with fault A as the diagnosis goal state, the labels are used to index fault A, and the labels index other faults. Similarly, in the following step, the labels are used to index the diagnostic goal
Chap. 4 Condition Diagnosis Method using DFCs and SVMs

state, and the training data are recombined to obtain the optimal hyper-plane. Then the
diagnostic system has been set up.

The second step is to diagnose whether a fault has occurred or not and to judge the
fault type for a new signal measured at the same speed. Using the same method, the
DFCs are extracted to establish the test data with labels. When the test data are input
into the diagnosis system, the system first classifies the data using Eq. (4.16) or Eq.
(4.18) with the output being 1 or -1. If the output is 1 or more, it indexes the signal as
normal. Otherwise, the signal is abnormal, and the DSCs switch to judge whether the
signal is a fault. If not, the process moves to the next step, and this is repeated until the
fault type is determined. To quantitatively analyze this method of classification, the
classification accuracy $P$ is proposed as follows.

$$P = \frac{N}{N_0}$$  \hspace{1cm} (4.19)

Here $N_0$ is the total number of test states. $N$ is the number of the states judged correctly.
Many experiments have shown that condition judgment is correct when the
classification accuracy is more than 60%.

4.4 Application to an experiment

4.4.1 Experimental system and signal conditions

Fig. 4.10 shows that the testing set-up is equipped with a rotating machine, loading
equipment and an accelerometer. The rotating machine is driven by a variable speed
3AC motor (0.5 hp) with speeds up to 2000 rpm. At most, 2T power can be transported
by the loading equipment (Model: RCS2-RA13R), and the transported power may be
either a fixed load or varied load. If a varied load is selected, the setting range of the
power is divided into $2^n$ parts, such as 120, 130, 140, 150 kg, as a half-periodic sine
wave, and the power is represented in 0.5 s time. Three accelerometers are mounted on the three directions (horizontal, vertical and axial direction) of the bearing housing to acquire the vibration signals with a sampling frequency of 5000 Hz.

![Fig. 4.10 The rotating machine used for the tests](image)

Table 4.2 Fault types and defect sizes for the tests

<table>
<thead>
<tr>
<th>Fault type</th>
<th>Defect size</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>U: unbalance</strong></td>
<td>S: small</td>
<td>Two flanges are loaded with masses of 10 g respectively.</td>
</tr>
<tr>
<td></td>
<td>M: medium</td>
<td>Two flanges are loaded with masses of 20 g respectively.</td>
</tr>
<tr>
<td></td>
<td>B: large</td>
<td>Two flanges are loaded with masses of 30 g respectively.</td>
</tr>
<tr>
<td><strong>L: looseness</strong></td>
<td>S: small</td>
<td>The screws on the lip of the right bearing are loosed 1 turn.</td>
</tr>
<tr>
<td></td>
<td>B: large</td>
<td>The screws on the lip of the right bearing are loosed 2 turns.</td>
</tr>
<tr>
<td><strong>M: misalignment</strong></td>
<td>S: small</td>
<td>The shaft deviates at an angle of 1° from the center line.</td>
</tr>
<tr>
<td></td>
<td>M: medium</td>
<td>The shaft deviates at an angle of 2° from the center line.</td>
</tr>
<tr>
<td></td>
<td>B: large</td>
<td>The shaft deviates at an angle of 3° from the center line.</td>
</tr>
</tbody>
</table>

The experiment has three set operating conditions which include: no load, fixed load at 120 kg and varied load in the range of 120 to 150 kg. In each condition, three typical
types of structural faults in the rotating machinery, namely a shaft misalignment, unbalanced and looseness states, were performed. To examine whether faults were able to be detected at an early stage, each negligible abnormality was set as shown in Table 4.2. Table 4.2 describes how the three fault conditions were tested and each defect size. Moreover, the rotational speed set values were 1000 rpm and 1500 rpm, which correspond to the theoretical values for the rotational frequencies of 16.7 Hz and 25 Hz.

4.4.2 Data acquisition and DFC extraction

The vibration signals, including normal and faults caused by different defect sizes, were obtained from the experimental system under different operating conditions. The sampling time was 20 seconds, and each data set included the signals from three channels. Then, 162 different data sets were acquired, and each data set corresponds to a different defect, at a different speed, collected with different channels and under a different operating conditions.

In the present work, in order to decrease the influence of experimental fluctuations, each raw signal \( X_n \) was normalized using the following formula

\[
X'_n = \frac{X_n - \mu}{\sigma}
\]  (4.20)

Here \( \mu \) is the mean-value, \( \sigma \) is the standard deviation and \( X'_n \) is the normalized signal.

Then, the normalized signal \( X'_n \) was divided into 24 parts and each part of the signal had 4096 samples, and each of these signals was processed to extract the DFCs. To further extract the valid DFCs, each of these raw signals was transformed into the frequency domain and the spectrum was normalized using Eq. (4.3). Since the rotating machine shown as Fig. 4.10 was influenced by inputted voltage and load during the experiments, the maximum value of the fluctuating speed was about 15% (180 rpm) of
the rotational speed, the DFCs were extracted from the specified ranges, with the most fluctuating value being 3 Hz, and the DFCs included frequency components at the corresponding points.

To validate the proposed method, each data set was solved. Comparison of the mean value of the 24 parts was done between the varying operating conditions. Here, taking the signals collected with a speed of 1000 rpm in the horizontal direction as an illustration, and these signals included signals that were normal and contained faults produced from different defect sizes. The DFCs extracted from three operating conditions are presented in Fig. 4.11.

Fig. 4.11 (A) shows the mean values of the DFCs obtained from a normal state, and even if the operating conditions changed, the DFCs’ values were very close to each other. Fig. 4.11 (B) shows the DFCs obtained from an unbalance state resulting from 9 different experimental conditions. As the operating condition or defect size changed, the DFCs’ values also changed but the rates of change were very small. Fig. 4.11 (C) and (D) also show the same results obtained from looseness and misalignment state. The most important thing indicated was that the feature from each state remained the same, which has important implications for DFCs fault diagnosis.

4.4.3 Fault diagnosis

4.4.3.1 Training data acquisition and building diagnostic system

To diagnose the faults under varied operating conditions, the diagnostic system was found using the method described in Fig. 4.9, and the diagnosis goal and the components of the training data in each step were determined as shown in Table 4.3.

In the first step, the goal was to diagnose whether a fault was occurring or not. The two classes were the normal state and the abnormal state. The diagnostic goal was the normal state, which was class A with labels $y_i = 1$, $(i = 1, 2, \cdots, I)$. The training data $y_i = 1$, $(i = 1, 2, \cdots, I)$ for class A included three DFC data sets with normal states
The normal state

An unbalance state (small defect)

An unbalance state (medium defect)

An unbalance state (large defect)

A looseness state (small defect)

A looseness state (large defect)

A misalignment state (small defect)

A misalignment state (medium defect)

A misalignment state (large defect)

Fig. 4.11 The DFCs from three operating conditions: ■ no-load, ■ fixed load at 120kg, ■ varied load 120~150 kg (rotational speed: 1000 rpm; channel: horizontal direction)
from three operating conditions: no-load, fixed load and varied load. Class B was the abnormal state which included unbalanced, looseness, misalignment and other states, and each state included many data sets corresponding to different defect sizes and operating condition. The DFCs from each state were selected to recombine the training data $x_j \ (i = 1, 2, \cdots, J)$ for class B and were labeled $y_j = -1, \ (j = 1,2,\cdots,J)$.

Table 4.3 Diagnosis goal and training data in each step

<table>
<thead>
<tr>
<th>Step</th>
<th>Diagnosis goal</th>
<th>Training data $X_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Class A</td>
</tr>
<tr>
<td>1</td>
<td>Normal</td>
<td>Normal state</td>
</tr>
<tr>
<td>2</td>
<td>Unbalanced</td>
<td>Unbalanced state</td>
</tr>
<tr>
<td>3</td>
<td>Looseness</td>
<td>Looseness state</td>
</tr>
<tr>
<td>4</td>
<td>Misalignment</td>
<td>Misalignment state</td>
</tr>
</tbody>
</table>

In the first step, the goal was to diagnose whether a fault was occurring or not. The two classes were the normal state and the abnormal state. The diagnostic goal was the normal state, which was class A with labels $y_i = 1, \ (i = 1, 2,\cdots,I)$. The training data $y_i = 1, \ (i = 1,2,\cdots,I)$ for class A included three DFC data sets with normal states from three operating conditions: no-load, fixed load and varied load. Class B was the abnormal state which included unbalanced, looseness, misalignment and other states, and each state included many data sets corresponding to different defect sizes and operating condition. The DFCs from each state were selected to recombine the training data $x_j \ (i = 1, 2,\cdots,J)$ for class B and were labeled $y_j = -1, \ (j = 1,2,\cdots,J)$.
Then the training data were input into the SVM for training. Table 4.4 presents the results of the SVM according as the three kernel functions with different parameter values. In Table 4.4, \(d\) is the degree of the polynomial, and \(\sigma\) is the width of the Gaussian RBF kernel parameter. It is obvious that the accuracy of linear classification is lower and Gaussian RBF with the width parameter \(\sigma=0.4\) is the best with high training and testing accuracy. Therefore, the optimal hyper-plane using the Gaussian RBF kernel was applied to diagnose whether a fault was occurring or not in the first step.

Table 4.4 Classification results according as kernel functions with the error penalty parameter \(C=0.1\)

<table>
<thead>
<tr>
<th>Kernel function</th>
<th>Parameter value</th>
<th>Classification accuracy (%)</th>
<th>Number of SVs</th>
<th>Training time(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Training</td>
<td>Testing</td>
<td></td>
</tr>
<tr>
<td>Linear</td>
<td></td>
<td>65.5</td>
<td>53.6</td>
<td>44</td>
</tr>
<tr>
<td>Polynomial</td>
<td>(d=1)</td>
<td>84.1</td>
<td>71.8</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>(d=2)</td>
<td>88.7</td>
<td>74.9</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>(d=3)</td>
<td>79.6</td>
<td>69.4</td>
<td>37</td>
</tr>
<tr>
<td>Gaussian RBF</td>
<td>(\sigma=0.2)</td>
<td>86.8</td>
<td>73.2</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>(\sigma=0.4)</td>
<td>96.5</td>
<td>91.7</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>(\sigma=1.6)</td>
<td>90.2</td>
<td>84.1</td>
<td>45</td>
</tr>
</tbody>
</table>

Similarly, the training data from each step were established in accordance with the components of two classes given in Table 4.3 and used to find the optimal hyper-plane of each step. If the accuracy of linear classification cannot satisfy the requirement of the system, Gaussian RBF with the width parameter \(\sigma=0.4\) is used to build the diagnostic system.
4.4.3.2 Diagnosis classification

To verify the proposed diagnostic method, the test data sets were acquired in accordance with the fault types and defect sizes shown in Table 4.2. However, the operating conditions included the previously mentioned conditions and two additional conditions: a fixed load of 150 kg and a varied load of 0 to 120 kg. The rotational speed was also set to 1000 rpm and 1500 rpm.

Each data set using these conditions was normalized using Eq. (4.20) and divided into 24 parts, and the DFCs from each part were extracted to recombine the test data using the same method. At the same time, using labels $y_i = 1 (i = 1, 2, \ldots, 24)$, the state of the test data was assumed in accordance with the state of diagnostic goal in each step. Then the test data were input into the system to begin to diagnose using the optimal hyper-plane from the first step. The labels indicated if the fault was recognized. If the hypothesis that the state was the normal state was correct, the normal state was indicated; otherwise, the state was indicated as the abnormal state. The classification accuracy $P$ was calculated. If the accuracy was more than 60%, the signal was judged as normal and the system was stopped automatically. If $P$ was less than 60% and the signal had many features of an abnormal state, the test data were input to the SVMs for the second step of judging until the fault type was determined. Here, taking the signals with a 1000 rpm speed in the horizontal direction as an illustration, the diagnosis results are shown in Table 4.5.

There are 45 different data sets presented in Table 4.5. Here the normal state had 5 different data sets from five operating conditions. Each data set was diagnosed separately. The diagnostic results were in accord with the original state. The unbalanced state had 15 data sets that corresponded to different sized defects and different operating conditions. These signals diagnosed that the fault had occurred in the first step, and in the second step, 14 signals were judged in the unbalanced state. Only one signal could not be judged, and it was input to the third step and diagnosed as
Table 4.5 The diagnostic results from various states with a 1000 rpm speed in the horizontal direction under five different operating conditions

<table>
<thead>
<tr>
<th>Test signal</th>
<th>Defect size</th>
<th>Operating condition</th>
<th>Accuracy (%) of each step</th>
<th>Diagnosis result</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Normal</td>
<td></td>
<td>No load</td>
<td>91.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 120 kg</td>
<td>83.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 150 kg</td>
<td>70.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 0~120 kg</td>
<td>62.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 120~150 kg</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>Unbalanced</td>
<td>small</td>
<td>No load</td>
<td>20.8</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 120 kg</td>
<td>37.5</td>
<td>79.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 150 kg</td>
<td>33.3</td>
<td>66.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 0~120 kg</td>
<td>29.2</td>
<td>83.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 120~150 kg</td>
<td>8.3</td>
<td>70.8</td>
</tr>
<tr>
<td>Unbalanced</td>
<td>medium</td>
<td>No load</td>
<td>8.3</td>
<td>79.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 120 kg</td>
<td>16.7</td>
<td>70.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 150 kg</td>
<td>29.2</td>
<td>66.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 0~120 kg</td>
<td>16.7</td>
<td>66.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 120~150 kg</td>
<td>4.2</td>
<td>83.3</td>
</tr>
<tr>
<td>Unbalanced</td>
<td>large</td>
<td>No load</td>
<td>4.2</td>
<td>41.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 120 kg</td>
<td>8.3</td>
<td>79.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 150 kg</td>
<td>4.2</td>
<td>87.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 0~120 kg</td>
<td>12.5</td>
<td>70.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 120~150 kg</td>
<td>4.2</td>
<td>83.3</td>
</tr>
</tbody>
</table>
### Table 4.5 (continued)

<table>
<thead>
<tr>
<th>Test signal</th>
<th>Defect size</th>
<th>Operating condition</th>
<th>Accuracy (%) of each step</th>
<th>Diagnosis result</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Looseness</td>
<td>small</td>
<td>No load</td>
<td>0</td>
<td>54.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 120 kg</td>
<td>33.3</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 150 kg</td>
<td>91.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 0–120 kg</td>
<td>4.2</td>
<td>62.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 120–150 kg</td>
<td>54.2</td>
<td>8.3</td>
</tr>
<tr>
<td></td>
<td>large</td>
<td>No load</td>
<td>0</td>
<td>33.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 120 kg</td>
<td>8.3</td>
<td>45.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 150 kg</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 0–120 kg</td>
<td>50</td>
<td>45.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 120–150 kg</td>
<td>0</td>
<td>4.2</td>
</tr>
<tr>
<td>Misalignment</td>
<td>small</td>
<td>No load</td>
<td>70.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 120 kg</td>
<td>50</td>
<td>12.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 150 kg</td>
<td>8.3</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 0–120 kg</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 120–150 kg</td>
<td>41.7</td>
<td>16.7</td>
</tr>
<tr>
<td></td>
<td>medium</td>
<td>No load</td>
<td>8.3</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 120 kg</td>
<td>8.3</td>
<td>8.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 150 kg</td>
<td>58.3</td>
<td>62.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 0–120 kg</td>
<td>37.5</td>
<td>37.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 120–150 kg</td>
<td>41.7</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>large</td>
<td>No load</td>
<td>8.3</td>
<td>16.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 120 kg</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 150 kg</td>
<td>8.3</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 0–120 kg</td>
<td>37.5</td>
<td>8.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 120–150 kg</td>
<td>54.2</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 4.6 The Diagnostic results from various states with a 1500 rpm speed in the vertical direction under five different operating conditions

<table>
<thead>
<tr>
<th>Test signal</th>
<th>Defect size</th>
<th>Operating condition</th>
<th>Accuracy (%) of each step</th>
<th>Diagnosis result</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Normal</td>
<td></td>
<td>No load</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 120 kg</td>
<td>79.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 150 kg</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 0–120 kg</td>
<td>70.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 120–150 kg</td>
<td>95.8</td>
<td></td>
</tr>
<tr>
<td>Unbalanced</td>
<td>small</td>
<td>No load</td>
<td>8.3</td>
<td>66.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 120 kg</td>
<td>20.8</td>
<td>70.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 150 kg</td>
<td>16.7</td>
<td>45.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 0–120 kg</td>
<td>8.3</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 120–150 kg</td>
<td>41.7</td>
<td>83.3</td>
</tr>
<tr>
<td></td>
<td>medium</td>
<td>No load</td>
<td>4.2</td>
<td>54.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 120 kg</td>
<td>41.7</td>
<td>95.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 150 kg</td>
<td>8.3</td>
<td>79.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 0–120 kg</td>
<td>16.7</td>
<td>66.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 120–150 kg</td>
<td>8.3</td>
<td>70.8</td>
</tr>
<tr>
<td></td>
<td>large</td>
<td>No load</td>
<td>0</td>
<td>70.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 120 kg</td>
<td>0</td>
<td>70.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 150 kg</td>
<td>0</td>
<td>62.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 0–120 kg</td>
<td>4.2</td>
<td>79.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 120–150 kg</td>
<td>0</td>
<td>79.2</td>
</tr>
</tbody>
</table>
Chap. 4 Condition Diagnosis Method using DFCs and SVMs

Table 4.6 (continued)

<table>
<thead>
<tr>
<th>Test signal</th>
<th>Defect size</th>
<th>Operating condition</th>
<th>Accuracy (%) of each step</th>
<th>Diagnosis result</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Looseness</td>
<td>small</td>
<td>No load</td>
<td>4.2</td>
<td>29.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 120 kg</td>
<td>4.2</td>
<td>4.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 150 kg</td>
<td>8.3</td>
<td>20.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 0–120 kg</td>
<td>8.3</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 120–150 kg</td>
<td>0</td>
<td>45.8</td>
</tr>
<tr>
<td></td>
<td>large</td>
<td>No load</td>
<td>0</td>
<td>33.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 120 kg</td>
<td>0</td>
<td>8.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 150 kg</td>
<td>0</td>
<td>33.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 0–120 kg</td>
<td>4.2</td>
<td>8.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 120–150 kg</td>
<td>0</td>
<td>58.3</td>
</tr>
<tr>
<td>Misalignment</td>
<td>small</td>
<td>No load</td>
<td>50</td>
<td>4.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 120 kg</td>
<td>91.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 150 kg</td>
<td>95.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 0–120 kg</td>
<td>29.2</td>
<td>16.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 120–150 kg</td>
<td>29.2</td>
<td>20.8</td>
</tr>
<tr>
<td></td>
<td>medium</td>
<td>No load</td>
<td>58.3</td>
<td>12.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 120 kg</td>
<td>20.8</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 150 kg</td>
<td>0</td>
<td>12.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 0–120 kg</td>
<td>41.7</td>
<td>20.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 120–150 kg</td>
<td>33.3</td>
<td>8.3</td>
</tr>
<tr>
<td></td>
<td>large</td>
<td>No load</td>
<td>54.2</td>
<td>12.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 120 kg</td>
<td>12.5</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed load at 150 kg</td>
<td>41.7</td>
<td>8.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 0–120 kg</td>
<td>45.8</td>
<td>33.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varied load 120–150 kg</td>
<td>0</td>
<td>8.3</td>
</tr>
</tbody>
</table>
a looseness state. Similarly, 2 looseness-state signals were diagnosed incorrectly as normal and unbalanced states, and 3 misalignment signals were diagnosed incorrectly as normal and looseness states. The diagnostic results from the other signals were in accord with their original states.

Similarly, for the signals with a 1500 rpm speed in the vertical direction, the training data for each step were established using the signals with the same speed and channel to establish the diagnostic system. Then test signals were acquired and diagnosed, and the diagnosis results are shown in Table 4.6.

The results in Table 4.5 and Table 4.6 show that 10 cases of three operating conditions that have been learnt by the system, and 3 cases of two operating conditions that have not been learnt, were judged incorrectly, the corresponding Diagnostic accuracies are 81.48% and 91.67% respectively. Moreover the latter is better than the former.

All the test data sets were diagnosed using the diagnostic system corresponding to the rotational speed and channel. Table 4.7 shows the diagnostic accuracies.

<table>
<thead>
<tr>
<th>Rotational speed (rpm)</th>
<th>Signal channel</th>
<th>Number of test signals</th>
<th>Number of signals diagnosed correctly</th>
<th>Diagnostic accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>Horizontal</td>
<td>45</td>
<td>39</td>
<td>86.7</td>
</tr>
<tr>
<td></td>
<td>Vertical</td>
<td>45</td>
<td>35</td>
<td>77.8</td>
</tr>
<tr>
<td></td>
<td>Axial</td>
<td>45</td>
<td>29</td>
<td>64.4</td>
</tr>
<tr>
<td>1500</td>
<td>Horizontal</td>
<td>45</td>
<td>37</td>
<td>82.2</td>
</tr>
<tr>
<td></td>
<td>Vertical</td>
<td>45</td>
<td>38</td>
<td>84.4</td>
</tr>
<tr>
<td></td>
<td>Axial</td>
<td>45</td>
<td>31</td>
<td>68.9</td>
</tr>
</tbody>
</table>
Viewing the overall diagnostic results, on the one hand, the test data from the two operating conditions added were diagnosed and the accuracies were very high. This ameliorated that fact that varying the operating condition restricts the machinery fault diagnosis in a real plant. On the other hand, the diagnostic accuracies were high if using the test signals from the horizontal and vertical directions; however, the diagnosis accuracies in the axial direction were not so high.

4.5 Conclusions

To intelligently diagnose structural faults in rotating machinery at an early stage, a new intelligent diagnosis method based on DFCs and SVMs was proposed, and the effectiveness was demonstrated experimentally. The superiority of the method proposed in this paper can be explained by the following points:

1. DFCs are features extracted from the vibration signal spectrum. Time and effort must be expended to find effective SPs when traditional methods are used.

2. The normalized DFCs decrease feature differences from a state with varying operational conditions. The method proposed will not be restricted by the varying operation conditions of rotating machinery.

3. The method proposed here not only has a strong adaptability to deal with fault diagnosis with small number of training samples. The method can also solve the non-convergent problem that frequently occurs when traditional diagnostic methods are used.

4. Diagnostic methods using the SVMs can perform a linear classification in a feature space for a non-linear classification using a kernel function; therefore, when the features of a fault signal is not obvious at an early stage, the resulting identification of the equipment condition is also excellent.

This method will be used widely in the field of machinery diagnosis.
References

[17] Peng Chen, Toshio Toyota and Zhengjia He: Automated Function Generation of Symptom Parameters and Application to Fault Diagnosis of Machinery under Variable Operating conditions; IEEE TRANSACTIONS ON SYSTEMS, No. 6, pp. 775-781 (2001)

Chapter 5

Diagnosis Method of Multi-faults State using Optimal Composition of Symptom Parameters and Fuzzy Support Vector Machine

5.1 Introduction

When diagnosing rotating machinery, the cases that two or more faults had been detected simultaneously often arise [1]. Here, the state that two or more faults have occurred simultaneously is called "multi-faults state". The diagnosis of a multi-faults state is an increasingly active research domain [2]. However, multiple faults may compensate or cover their features, so the diagnostic information reflecting the feature of multi-faults state is ambiguous, especially at an early stage. Moreover, in most cases of condition diagnosis for rotation machinery, the values of symptom parameter calculated from vibration signals for condition monitoring and fault diagnosis are ambiguous. The main reasons can be explained as follows: (1) When rotation speed and load of rotation machinery vary while vibration signals is being measured and a fault is in an early stage, the signal contains strong noise, stronger noise than the actual failure signal may lead to misrecognition of useful information for diagnosis. (2) The statistical objectivity of the measured signal cannot always be satisfied because of the measuring techniques and manner of the inspectors [3]. Therefore, it is more difficult to solve the ambiguous problem of fault diagnosis.

Roller bearings are an important part and widely used in rotating machinery. The failure of a bearing may cause the breakdown of a rotating machine, and furthermore, serious consequences may arise due to the failure. Therefore, fault diagnosis of roller
bearings is most important for guaranteeing production efficiency and plant safety. Although fault diagnosis of bearings is often artificially carried out using time or frequency analysis of vibration signals, there is a need for a reliable, automated diagnosis method thereof. Neural networks (NN) have potential applications in automated detection and diagnosis of machine failures [4-9]. However, a conventional NN cannot adequately reflect the possibility of ambiguous diagnosis problems, and will never converge, when the symptom parameters, input to the first layer of the NN, have the same values in different states [3].

To solve these problems, this paper proposes a novel condition diagnosis method using optimal composition of symptom parameters (OCSPs) and fuzzy support vector machine (FSVM), to detect multi-faults state at an early stage. The flowchart for the condition diagnostic procedure proposed in this study is shown in Fig. 5.1. The non-dimensional symptom parameters (NSPs) are defined for reflecting the feature of vibration signals measured in each state. Since vibration signals from different direction sensors are analyzed to show different features, the higher sensitive NSPs from multiple directions are selected by principal component analysis to form an optimal composition. Then the OCSPs of two states, as the object of follow-on process, are inputted into a SVM, to obtain an optimal hyper-plane. The optimal hyper-plane is used to define a synthetic symptom parameter (SSP). It is verified that the sensitivity of the SSP is very higher. Finally, a fuzzy diagnosis method with possibility theory is also proposed, in which the conditions of the machinery can be well identified sequentially. Practical examples of diagnosing a roller bearing which is used in a centrifugal fan are provided to verify the effectiveness of the proposed method. The verification results show that the faults that often occur in roller bearing, such as outer race defect, inner race defect, rolling element defect, outer race-rolling element defect, and inner race-rolling element defect, are effectively identified by the proposed method.
5.2 Vibration data

Fig. 5.2 shows the experimental system which is a centrifugal fan (TERAL CLF3) used for roller bearing fault diagnosis. A 2.2 kW induction motor (SB-JR Super Line...
Motor) with 3 phases and a maximum revolution of 1420 rpm is employed to drive the fan through 2 V-belts, and its rotating speed can be varied by a speed controller. Moreover, the bearings (Type: NU204) were utilized, and the specification of the bearings is listed in Table 5.1.

Table 5.1 Specification of the bearings

<table>
<thead>
<tr>
<th>Contents</th>
<th>Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outer diameter</td>
<td>47 mm</td>
</tr>
<tr>
<td>Inner diameter</td>
<td>20 mm</td>
</tr>
<tr>
<td>Width</td>
<td>14 mm</td>
</tr>
<tr>
<td>Roller diameter</td>
<td>7 mm</td>
</tr>
<tr>
<td>Number of rollers</td>
<td>11</td>
</tr>
<tr>
<td>Contact angle</td>
<td>0 rad</td>
</tr>
</tbody>
</table>

(A) outer race defect (O)  (B) inner race defect (I)  (C) roller element defect (E)

(D) outer race-rolling element defect (OE)  (E) inner race-rolling element defect (IE)

Fig. 5.3 Defect types for roller bearing fault diagnosis test
Is case of the practical diagnosis, three single-fault states, namely, the outer race defect (O), the inner race defect (I), the roller element defect (E), and two multi-faults states, such as the outer race-rolling element defect (OE), and the inner race-rolling element defect (IE), have often occurred in a roller bearing. In order to research whether these faults of roller bearing are detectable at an early stage, the faults, shown in Fig. 5.3, were artificially made with the use of a wire-cutting machine, and the sizes of the faults are shown in Table 5.2.

Table 5.2 Sizes of the faults in roller bearing

<table>
<thead>
<tr>
<th>Fault type of roller bearing</th>
<th>Size of defect (width * depth)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outer race defect (O)</td>
<td>0.3 * 0.05 mm</td>
</tr>
<tr>
<td>Inner race defect (I)</td>
<td>0.3 * 0.05 mm</td>
</tr>
<tr>
<td>Rolling element defect (E)</td>
<td>0.3* 0.05 mm</td>
</tr>
<tr>
<td>Outer race-rolling element defect (OE)</td>
<td>(O) 0.3 * 0.05 mm, (E) 0.3 * 0.05 mm</td>
</tr>
<tr>
<td>Inner race-rolling element defect (IE)</td>
<td>(I) 0.3 * 0.05 mm, (E) 0.3 * 0.05 mm</td>
</tr>
</tbody>
</table>

In this work three accelerometers (Type: PCB MA352A60; Sensitivity: 10mV/g; Frequency: 5-60000 Hz) were mounted on the three directions (horizontal, vertical and axial direction) of the bearing housing to acquire the vibration signals with a sampling frequency of 50kHz, and the sampling time is 20s. As the fan usually works at a constant speed, the vibration signals are measured at a rotation speed of 1000 rpm. Then the time-domain waveforms of vibration signals in each state are shown in Fig. 5.4.

Since those faults appear in the high frequency domain, a high-pass filter with 10 kHz cut-off frequency is used to cancel the low frequency noise in the vibration signals. The power spectrums of the signals can be obtained by FFT, and be used for calculating the non-dimensional symptom parameters (NSPs).
Fig. 5.4 The time-domain waveforms of vibration signals in each state and direction; six conditions: (A) normal state, (B) outer race defect, (C) inner race defect, (D) Rolling element defect, (E) outer race-rolling element defect, (F) inner race-rolling element defect; three directions: (1) horizontal direction, (2) vertical direction, (3) axial direction.
5.3 Non-dimensional symptom parameters (NSPs)

When a computer is used for the condition diagnosis of plant machinery, symptom parameters (SPs) are required to express the information indicated by a signal measured for diagnosing machinery faults. A good symptom parameter can correctly reflect states and the condition trend of plant machinery [10-12]. Many symptom parameters have been defined in the pattern recognition field. In this study, six non-dimensional symptom parameters (NSPs) in the time domain for the condition diagnosis of roller bearing are defined.

\[
p_1 = \frac{\sum_{i=1}^{N} (x_i - \bar{x})^3}{N \cdot \sigma^3}
\]  
\[
p_2 = \frac{\sum_{i=1}^{N} (x_i - \bar{x})^4}{N \cdot \sigma^4}
\]

Here, \(x_i(i=1,2,\cdots,N)\) is digital data of vibration signal, \(N\) is the number of data. \(\bar{x}\) and \(\sigma\) are the mean value and the standard deviation.

\[
p_3 = \frac{\sum_{i=1}^{N_p} (x_{pi} - \bar{x}_p)^3}{N_p \cdot \sigma_{p}^3}
\]
\[
p_4 = \frac{\sum_{i=1}^{N_p} (x_{pi} - \bar{x}_p)^4}{N_p \cdot \sigma_{p}^4}
\]

Here, \(x_{pi}(i=1,2,\cdots,N_p)\) is the peak data of vibration signal, \(N_p\) is the number of data. \(\bar{x}_p, \sigma_{p}\) are the mean value and the standard deviation.
Here, \( x_i \) (\( i = 1, 2, \ldots, N_v \)) is the valley data of vibration signal, \( N_v \) is the number of data. \( \bar{x}_v \) and \( \sigma_v \) are the mean value and the standard deviation.

Since vibration signals of the roller bearing from three directions have been measured, \( P, P_i, P_j \) (\( i = 1, 2, 3 \)) are used to reflect NSPs of a condition from horizontal, vertical and axial direction, respectively.

### 5.4 Organizing optimal composition of symptom parameters

In order to effectively reflect the different features of vibration signals from multiple directions, and automatically extract higher sensitive symptom parameters for detecting and distinguishing faults, principal component analysis (PCA) are proposed to select good SPs in each direction. Moreover, optimal composition of symptom parameters is proposed as the object of follow-on process.

#### 5.4.1 Condition diagnosis approach

In many cases of condition diagnosis, symptom parameters are defined to reflect the features of vibration signals measured in each state. Though it is difficult to find one or more symptom parameters that can identify all of the faults simultaneously, the symptom parameters for identification of two states are easy to identify. Therefore, in order to solve these problems, a precision diagnosis method is proposed, as shown in Fig. 5.5. In the first step, it is distinguished whether the diagnosis condition is normal state (N) or not. When normal, a diagnostic system can stop automatically. Otherwise,
the faults of roller bearing have occurred, and then diagnostic system will come into the step of precision diagnosis. Precision diagnosis is performed to identify the fault types. Since there are five types of faults in the experiment, five patterns are used to perform fuzzy diagnosis, as shown in Fig. 5.5. In each pattern, one fault is selected as diagnosis target, to build a fuzzy diagnosis with other faults, respectively. Finally, possibility theory is used to integrate the results of five patterns, and then diagnosis condition is judged.

![Flowchart for condition diagnosis](image)

**Fig. 5.5 Flowchart for condition diagnosis**

### 5.4.2 Principal component analysis for selecting symptom parameters

Principal component analysis (PCA) is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of uncorrelated variables called principal components.
PCA was invented in 1901 by Karl Pearson. Now it is mostly used as a tool in exploratory data analysis and for making predictive models. PCA can be done by eigenvalue decomposition of a data covariance matrix or the singular value decomposition of a data matrix. These decompositions are usually performed after mean centering the data for each attribute. The results of a PCA are usually discussed in terms of component scores and loadings [13]. In the last few years, PCA has been applied to process fault diagnosis (identification) [14-16].

Define a data matrix with size $m \times n$, where $m$ is the number of identifying states and $n$ is the number of symptom parameters, whose covariance matrix has eigenvalue $\lambda_i$ and eigenvector $a_i$ and $i = 1 - n$ with $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_n$. Principal components $Z_i$ and the cumulative contribution rate of the principal components $\eta_i$ can be calculated as follows:

\[
\begin{bmatrix}
Z_1 \\
Z_2 \\
M \\
Z_n
\end{bmatrix} = \begin{bmatrix}
a_{i1} & a_{i2} & \cdots & a_{in} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
M & M & M & M \\
a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix} \begin{bmatrix}
P_1 \\
P_2 \\
M \\
P_n
\end{bmatrix} = A \cdot P
\]

(5.7)

\[
\eta_i = \frac{\sum_{j=1}^{i} \lambda_j}{\sum_{k=1}^{n} \lambda_k}
\]

(5.8)

Where $\lambda_i$ is the standard deviation of the principal component, $a_{ij}$ is the weight coefficient of the principal component, $P_i$ is a symptom parameter, $i = 1 - n$ and $j = 1 - m$.

Using PCA, original data can be converted into many principal components. The first few principal components contain most of the information and the discriminatory features. The weight coefficient of the principal components can express the importance of the symptom parameters for each principal component. Therefore, the symptom parameters that have high sensitivity for detecting faults can be selected by the weight coefficient for the first few principal components.
Table 5.3 Symptom parameters selected by PCA in horizontal direction

<table>
<thead>
<tr>
<th>Two states</th>
<th>weight coefficient</th>
<th>Cumulative contribution rate $\eta_1$</th>
<th>NSPs selected by PCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>N, O</td>
<td>P_1: -1.0, P_2: -0.6, P_3: 1.0, P_4: 0.73, P_5: 0.97, P_6: -0.9</td>
<td>0.91</td>
<td>P_3, P_5</td>
</tr>
<tr>
<td>N, I</td>
<td>P_1: -1.0, P_2: 0.5, P_3: 0.99, P_4: -0.4, P_5: 1.0</td>
<td>0.88</td>
<td>P_4, P_6</td>
</tr>
<tr>
<td>N, E</td>
<td>P_1: -0.6, P_2: 0.89, P_3: 0.71, P_4: 1.0, P_5: -1.0, P_6: -0.9</td>
<td>0.93</td>
<td>P_2, P_4</td>
</tr>
<tr>
<td>N, OE</td>
<td>P_1: -1.0, P_2: -0.7, P_3: 0.98, P_4: 0.6, P_5: 1.0, P_6: -0.8</td>
<td>0.87</td>
<td>P_3, P_5</td>
</tr>
<tr>
<td>N, IE</td>
<td>P_1: -1.0, P_2: -0.9, P_3: 0.79, P_4: 1.0, P_5: -0.8, P_6: 0.89</td>
<td>0.90</td>
<td>P_4, P_6</td>
</tr>
<tr>
<td>O, I</td>
<td>P_1: -0.7, P_2: 0.6, P_3: 0.83, P_4: -0.7, P_5: 1.0</td>
<td>0.85</td>
<td>P_3, P_6</td>
</tr>
<tr>
<td>O, E</td>
<td>P_1: 0.89, P_2: -1.0, P_3: -0.7, P_4: 1.0, P_5: 0.7, P_6: -0.9</td>
<td>0.91</td>
<td>P_1, P_4</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>OE, IE</td>
<td>P_1: -1.0, P_2: 1.0, P_3: 0.6, P_4: 0.94, P_5: -0.7, P_6: -0.8</td>
<td>0.88</td>
<td>P_2, P_4</td>
</tr>
</tbody>
</table>

Table 5.4 Selection results of NSPs from three directions by PCA

<table>
<thead>
<tr>
<th>Two states</th>
<th>Two NSPs selected by PCA in each direction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Horizontal</td>
</tr>
<tr>
<td>N, O</td>
<td>P_3, P_5</td>
</tr>
<tr>
<td>N, I</td>
<td>P_4, P_6</td>
</tr>
<tr>
<td>N, E</td>
<td>P_2, P_4</td>
</tr>
<tr>
<td>N, OE</td>
<td>P_3, P_5</td>
</tr>
<tr>
<td>N, IE</td>
<td>P_4, P_6</td>
</tr>
<tr>
<td>O, I</td>
<td>P_3, P_6</td>
</tr>
<tr>
<td>O, E</td>
<td>P_1, P_4</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>OE, IE</td>
<td>P_2, P_4</td>
</tr>
</tbody>
</table>

In this study, the two best NSPs that contain the most information and have high
sensitivity are selected by PCA, as shown in Eq. (5.7), in each direction. As an example, parts of the selection results from six NSPs in horizontal direction are shown in Table 5.3. Similarly, higher sensitive NSPs are also selected in vertical and axial directions, as shown in Table 5.4.

### 5.4.3 Organizing optimal composition of symptom parameters

<table>
<thead>
<tr>
<th>Two states</th>
<th>Symptom parameters (SPs) as the object of follow-on process</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plan 1</td>
<td>P₃, P₅, P₂⁻, P₆⁻, P₁⁻, P₂⁻</td>
</tr>
<tr>
<td>Plan 2</td>
<td>N, O</td>
</tr>
<tr>
<td>Plan 3</td>
<td>N, I</td>
</tr>
<tr>
<td>Plan 4</td>
<td>N, E</td>
</tr>
<tr>
<td>Plan 5</td>
<td>N, OE</td>
</tr>
<tr>
<td>Plan 6</td>
<td>N, IE</td>
</tr>
<tr>
<td>Plan 7</td>
<td>O, I</td>
</tr>
<tr>
<td>Plan 8</td>
<td>O, E</td>
</tr>
<tr>
<td>Plan 9</td>
<td>OE, IE</td>
</tr>
</tbody>
</table>

Table 5.5 Different objects of follow-on process in diagnostic system
In order to effectively diagnose multi-faults state at an early stage, the paper proposes that higher sensitive SPs from multiple directions are used to organize a composition as the object of follow-on process. In order to compare the performances of OCSPs and some higher sensitive SPs from simple direction, the different objects of follow-on process are planned, as shown in Table 5.5. The SPs of plan 1, plan 2 and plan 3 are from simple direction, the SPs of plan 4, plan 5 and plan 6 are from two directions, and the SPs of plan 7 are from three directions. Here, the SPs of plan 7 are used to introduce the performance of OCSPs, as an example.

5.5 Generation method of synthetic symptom parameter by SVM

5.5.1 Acquisition of optimal hyper-plane by SVM

As shown in chapter 2.3.1, a SVM is a relatively new computational learning method [17-19]. The basic idea of applying SVM to pattern classification can be stated as follows: first, map the input vectors into one features space, possible in higher space, either linearly or nonlinearly, which is relevant with the kernel function. Then, within the feature space from the first step, seek an optimized linear division, that is, construct a hyper-plane which separates two classes. It can be extended to multi-class. SVMs training always seek a global optimized solution and avoid non-convergent problem, so it has ability to deal with a large number of feature. The goal of a SVM is to define a hyper-plane which divides S, such that all the points with the same label are on the same side of the hyper-plane while maximizing the distance between two classes A, B and the hyper-plane. The boundary can be expressed as Eq. (5.9), where \(\omega\) and \(b\) are used to define the position of optimal hyper-plane. And by transforming to quadratic problem as Eq. (5.10), the hyper-plane is maximized. Finally, the decision function as Eq. (5.11) is made using \(\text{sign}(f(x))\) to create optimal hyper-plane.
It is rare that linear separation is possible on the real problem of pattern recognition. However, we might prefer a solution that better separates the bulk of the data while ignoring a few weird noise documents. That is soft margin SVM [20-21]. When the two classes are not linearly separable, the condition for the optimal hyper-plane can be relaxed by including an extra term:

\[ y_i (x_i^T w + b) \geq 1 - \xi_i , \quad (i = 1, \Lambda , m) \]  

For minimum error, \( \xi_i \geq 0 \) should be minimized as well as \( \|w\| \), and the objective function becomes [22]:

Minimize:

\[ w^T w + C \sum_{i=1}^{m} \xi_i^k \]  

Subject to: \( y_i (x_i^T w + b) \geq 1 - \xi_i \), and \( \xi_i \geq 0 \); \( i = 1, \Lambda , m \)  

Here \( C \) is a regularization parameter that controls the trade-off between maximizing the margin and minimizing the training error. Small \( C \) tends to emphasize the margin while ignoring the outliers in the training data, while large \( C \) may tend to over fit the training data.

Then the optimization problem for soft margin classification becomes that Lagrange multipliers \( \alpha_i (i = 1, \Lambda , N) \) are found such that Eq. (5.10) is maximized and

\[ \sum \alpha_i y_i = 0 \]  
\[ 0 \leq \alpha_i \leq C , \quad (i = 1, \Lambda , N) \]
Neither the slack variables $\xi_i$ nor Lagrange multipliers for them appear in the dual problem. All we are left with is the constant $C$ bounding the possible size of the Lagrange multipliers for the support vector data points. As before, the $x_i$ with non-zero $\alpha_i$ will be the support vectors. The solution of the quadratic problem is of the form:

$$\omega^* = \sum_{i \in SV} \alpha_i y_i x_i$$

$$b^* = y_i (1 - \xi_i) - \omega^* x_i, i \in SV$$

When the SVMs are trained, the optimal hyper-plane and its equation are also obtained. For example, soft margin SVM and two SPs are used to obtain the optimal hyper-plane, as shown in Fig. 5.6.

![Figure 5.6 Example of optimal hyper-plane constructed by soft margin SVM](image)

Of course, for a linear non-separable case, SVM can perform a non-linear mapping of the input vector $x$ from the input space into a higher dimensional space, where the mapping is determined by kernel function. But the vibration signals from different conditions have only changing tendency, but not obvious boundary.

Then in the field of mechanical fault diagnosis, soft margin SVM has a strong generalization capacity when only a small amount of training samples are available. These features have an important role in the field of mechanical fault diagnosis.
5.5.2 Synthetic symptom parameter

Soft margin SVM is used to construct the optimal classification hyper-plane, and the function corresponding to the optimal classification hyper-plane is called “the optimal classification function” as follows.

\[ f(x) = \omega \cdot x + b \]  

(5.19)

In this study, in order to increase the diagnostic sensitivity of the SPs, the optimal classification function obtained using soft margin SVM and primitive SPs is defined as synthetic symptom parameter (SSP).

\[ SSP = \omega \cdot \left[ P_1 \quad P_2 \quad \cdots \quad P_n \right] + b \]  

(5.20)

![Diagram]

Fig. 5.7 The efficiency of SSP

To explain the efficiency of SSP defined using soft margin SVM and primitive SPs, two SPs are used to obtain the optimal classification hyper-line (L) when class A and class B are distinguished. Then a line perpendicular to the line (L) is SSP-axis, as shown in Fig. 5.7. \( f_a(P_i) \) and \( f_b(P_i) \) are the \( P_i \) distributions of class A and class B, \( f_a(P_i) \)
and \( f_a(P) \) are the \( P \) distributions of class A and class B, \( f_A(SSP) \) and \( f_B(SSP) \) are the \( SSP \) distributions of class A and class B. It is obvious that the overlap of \( f_A(SSP) \) and \( f_B(SSP) \) is smaller than those of others. Then the sensitivity of \( SSP \) is higher than \( P \) and \( P \). Then the OCSPs of plan 7 and soft margin SVM are used to generate \( SSPs \), as shown in Table 5.6.

Table 5.6 SSPs generated by OCSPs of plan 7 and soft margin SVM

<table>
<thead>
<tr>
<th>Two states</th>
<th>synthetic symptom parameter (SSP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N, O</td>
<td>( y=2.12P_3-0.342P_5-1.756P_2+6.09P_6+0.002P_1-1.892P_2-2.051 )</td>
</tr>
<tr>
<td>N, I</td>
<td>( y=-0.26P_4-1.58P_5+0.386P_1+2.118P_5'-2.012P_7'+5.007P_7'-11.967 )</td>
</tr>
<tr>
<td>N, E</td>
<td>( y=4.423P_2-3.001P_4+0.002P_3'+1.905P_6'-1.529P_7'+2.201P_7'+0.998 )</td>
</tr>
<tr>
<td>N, OE</td>
<td>( y=-0.413P_3-2.549P_5-0.001P_5+7.703P_6'-2.567P_3'-1.115P_3'-4.001 )</td>
</tr>
<tr>
<td>N, IE</td>
<td>( y=6.251P_4+1.482P_6-2.739P_3-0.203P_4-1.194P_1'+0.845P_7'+1.083 )</td>
</tr>
<tr>
<td>O, I</td>
<td>( y=0.927P_3+3.331P_6-2.056P_3'-0.3423P_3'-3.415P_2'+2.270P_5'-0.992 )</td>
</tr>
<tr>
<td>O, E</td>
<td>( y=3.001P_1+2.152P_5-0.206P_2'-1.929P_3'+0.803P_2'-3.418P_6'+2.709 )</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>OE, IE</td>
<td>( y=-0.903P_2-2.402P_4+2.196P_5+0.0805P_5'-5.018P_4'+0.092P_6'-3.087 )</td>
</tr>
</tbody>
</table>

5.6 Condition diagnosis using possibility theory

In this section, possibility theory is introduced for condition diagnosis, and the condition diagnosis of simple-fault state and multi-faults state in roller bearing is shown to verify that the method proposed in this paper is effective.

5.6.1 Possibility theory for condition diagnosis

Possibility theory is a mathematical theory for dealing with certain types of uncertainty and is an alternative to probability theory. Professor L.A. Zadeh first
introduced possibility theory in 1978 as an extension of his theory of fuzzy sets and fuzzy logic [23]. D. Dubois and H. Prade further contributed to its development [24]. Recently, possibility theory has been used for fault diagnosis [25-26]. In [25-26], possibility theory was applied to condition diagnosis in rotating machinery under variable rotating speeds to process the uncertain relationship between the symptoms and fault types.

For fuzzy inference, the membership function of a SP is necessary [25-26]. This can be obtained from probability density functions of the SP using possibility theory. When a SP conforms to the normal distribution, it can be changed to possibility distribution function \( p(x_i) \) by the following formulae.

\[
p(x_i) = \sum_{k=1}^{N} \min \{ \lambda_i, \lambda_k \}, (i, k = 1, 2, \cdots, N) \quad (5.21)
\]

\( \lambda_i \) and \( \lambda_k \) can be calculated as follows:

\[
\lambda_i = \int_{x_i-1}^{x_i} \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{(x - \bar{x})^2}{2\sigma^2} \right\} dx \quad (5.22)
\]

\[
\lambda_k = \int_{x_{k-1}}^{x_k} \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{(x - \bar{x})^2}{2\sigma^2} \right\} dx \quad (5.23)
\]

![Fig. 5.8 Probability density function \( f(x) \) and possibility function \( p(x) \)](image)

Where, \( N \) is the division number of the domain of the SP namely \([\bar{x} - 3\sigma, \bar{x} + 3\sigma]\), \( \bar{x} \) and \( \sigma \) are the mean and the standard deviation of the SP, respectively. \( x_i = \bar{x} + \)
\[
\frac{6i-3N}{N} \cdot \sigma, \text{ and } x_k = \bar{x} + \frac{6k-3N}{N} \cdot \sigma. \text{ Fig. 5.8 shows an illustration of the possibility function and the probability density function.}
\]

\[p(h)\]

**Fig. 5.9 Matching examples of possibility function**

When intelligently identifying a state, the possibility function of the SP is used to constitute the membership function for diagnosis using possibility theory. Fig. 5.9 shows the matching examples of possibility function. The matching degrees between the possibility functions \(p_i(h), p_j(h)\) of model state \(i, j\) and the possibility functions \(p_t(h)\) of diagnostic state, is calculated by the following formulas.

\[
w_i = p_i(h) \wedge p_i(h)\big|_{h=\bar{h}} \quad (5.24)
\]

\[
w_j = p_j(h) \wedge p_t(h)\big|_{h=\bar{h}} \quad (5.25)
\]

Where \(\bar{h}\) is the mean value of the SP in diagnostic state. Moreover, the unknown state except state \(i, j\) is the model state \(x\), its possibility functions \(p_x(h)\) is calculated by the following formula (5.26). And the matching degree between \(p_x(h)\) and \(p_t(h)\) is calculated by the following formula (5.27).

\[
p_x(h) = \max\{1 - [p_i(h) + p_j(h)], 0\} \quad (5.26)
\]

\[
w_x = p_x(h) \wedge p_t(h)\big|_{h=\bar{h}} \quad (5.27)
\]
Therefore, the probability, a diagnostic state should be diagnosed with model state \( i, j \) and \( un \), respectively, can be calculated by the following formulas. Then it is judged that the diagnostic state is the model state with the most probability.

\[
\omega_i = \frac{w_i}{w_i + w_j + w_x} \quad (5.28)
\]
\[
\omega_j = \frac{w_j}{w_i + w_j + w_x} \quad (5.29)
\]
\[
\omega_x = \frac{w_x}{w_i + w_j + w_x} \quad (5.30)
\]

### 5.6.2 Diagnosis and verification

In order to achieve an intelligent and efficient condition diagnosis, the training data are inputted into soft margin SVM, to define a SSP for diagnosing two states. Then possibility functions of all SSPs are used to build diagnostic system. In order to verify the diagnostic capability of the method proposed in this study, the test data, measured in each known state and not used for the pre-calculated possibility function, are used. The practical diagnostic examples in the first pattern of precision diagnosis step are shown in Fig. 5.10.

In the first pattern of precision diagnosis step, outer race defect (O) is selected as diagnosis target. Then outer race defect (O) is regarded as model states with inner race defect (I), roller element defect (E), outer race-rolling element defect (OE), inner race-rolling element defect (IE), respectively. Here, \( p_o(h) \), \( p_i(h) \), \( p_e(h) \), \( p_{oe}(h) \) and \( p_{ie}(h) \) express the possibility functions of the SSPs in outer race defect (O), inner race defect (I), roller element defect (E), outer race-rolling element defect (OE) and inner race-rolling element defect (IE), respectively. \( p_i(h) \) expresses the possibility functions of the SSPs for diagnosing state. Moreover, \( W_o \), \( W_i \), \( W_e \), \( W_{oe} \), \( W_{ie} \) and \( W_x \) are the matching degrees between the diagnosing state and the model states. \( \omega_o \), \( \omega_i \), \( \omega_e \), \( \omega_{oe} \), \( \omega_{ie} \) and \( \omega_x \) are calculated by Eq. (5.28), Eq. (5.29) and Eq. (5.30), as shown in Table 5.7. Therefore, the probability that the diagnostic state is outer race defect is 0.698.
Fig. 5.10 The practical examples in the first pattern of precision diagnosis step: (A) diagnosis models of outer race defect (O) and inner race defect (I); (B) diagnosis models of outer race defect (O) and roller element defect (E); (C) diagnosis models of outer race defect (O) and outer race-rolling element defect (OE); (D) diagnosis models of outer race defect (O) and inner race-rolling element defect (IE).
Table 5.7 Diagnosis results in the first pattern of precision diagnosis step

<table>
<thead>
<tr>
<th>No.</th>
<th>Model states</th>
<th>Probability of a diagnostic state</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>State 1</td>
<td>State 1</td>
</tr>
<tr>
<td>A</td>
<td>O</td>
<td>I</td>
</tr>
<tr>
<td>B</td>
<td>O</td>
<td>E</td>
</tr>
<tr>
<td>C</td>
<td>O</td>
<td>OE</td>
</tr>
<tr>
<td>D</td>
<td>O</td>
<td>IE</td>
</tr>
</tbody>
</table>

Table 5.8 Diagnosis results using the OCSPs of plan 7

<table>
<thead>
<tr>
<th>Test state</th>
<th>Diagnosis probability of diagnostic target in each step</th>
<th>Judge state</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>In the first step</td>
<td>In the second step</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>UN</td>
</tr>
<tr>
<td>N</td>
<td>0.911</td>
<td>0.089</td>
</tr>
<tr>
<td>O</td>
<td>0.244</td>
<td>0.756</td>
</tr>
<tr>
<td>I</td>
<td>0.257</td>
<td>0.743</td>
</tr>
<tr>
<td>E</td>
<td>0.189</td>
<td>0.811</td>
</tr>
<tr>
<td>OE</td>
<td>0.173</td>
<td>0.827</td>
</tr>
<tr>
<td>IE</td>
<td>0.195</td>
<td>0.805</td>
</tr>
</tbody>
</table>

Similarly, in the second, third, fourth, fifth pattern of precision diagnosis step, inner race defect (I), roller element defect (E), outer race-rolling element defect (OE), inner race-rolling element defect (IE) are selected as diagnosis target, respectively. The same diagnostic state is used to diagnose. The diagnosis results, that the diagnostic state is I, E, OE, IE respectively, are 0.347, 0.298, 0.475 and 0.264. All results of five patterns in precision diagnosis step considered, it is obvious that the diagnostic state is outer race defect. Moreover, the test data of five different states are inputted into the diagnosis system, respectively. The diagnosis results, as shown in Table 5.8, are in accord with the test states.
5.6.3 Comparing the performances of OCSPs

As shown in chapter 5.4.3 and Table 5.5, the different OCSPs and higher sensitive SPs from simple direction are planned. Here, the same vibration signals are used to organize the SPs of plan 1~plan 6, and then the diagnosis system is trained, respectively. Finally, the test data, measured in each known state and not used for the pre-calculated possibility function, are used to verify the diagnostic capability. The diagnosis results are shown in Table 5.9.

Table 5.9 Diagnosis results with different objects of follow-on process in each plan

<table>
<thead>
<tr>
<th>Test state</th>
<th>Diagnosis probability in each plan</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Plan 1</td>
</tr>
<tr>
<td>N</td>
<td>0.714</td>
</tr>
<tr>
<td>O</td>
<td>0.508</td>
</tr>
<tr>
<td>I</td>
<td>0.485</td>
</tr>
<tr>
<td>E</td>
<td>0.534</td>
</tr>
<tr>
<td>OE</td>
<td>0.536</td>
</tr>
<tr>
<td>IE</td>
<td>0.547</td>
</tr>
</tbody>
</table>

In plan 1~plan 3, there are the cases that the diagnosis probabilities are less than 0.5. Thus, erroneous judge will necessarily happen. In plan 4~plan 7, the diagnosis probabilities are more than 0.5, all test states can be automatically and correctly identified. However, the diagnosis probabilities of plan 7 are higher to enhance the reliability of diagnosis.

5.7 Conclusions

To intelligently diagnose structural faults in rotating machinery at an early stage, a
new intelligent diagnosis method based on OCSPs and SVMs was proposed, and the effectiveness was demonstrated experimentally. The superiority of the method proposed in this paper can be explained by the following points:

(1) OCSPs are the synthetical diagnosis information using multi-directions SPs. The application of OCSPs is very strong, even if there will be a slight fault at an early stage, the diagnostic information of OCSPs can be strong.

(2) The diagnosis method based on SVMs not only has a strong adaptable ability to deal with fault diagnosis with small number of training samples, but also solves the non-convergent problem that frequently occurs when traditional diagnostic methods are used.

(3) Diagnosis method using OCSPs and SVMs has more capacity of reliability and robustness. Even if linear classification is only used, the resulting identification of the equipment condition is also excellent when the feature of fault signal is not obvious at an early stage.

(4) The efficiency of intelligent diagnosis proposed here has been verified by applying it to a practical structural faults diagnosis of rotating machinery, and its performance is high. In fact, the high performance of intelligent diagnosis using OCSPs and SVMs is attributed primarily to OCSPs’ high sensitivity and SVMs’ generalization capability.

This method will be used widely in the of condition monitoring and machinery diagnosis.

References

Chap. 5 Diagnosis Method of Multi-faults State using OCSPs and FSVM


(1986)


Chapter 6

Precision Diagnosis Method for a Centrifugal Pump using Statistic Filter, Support Vector Machine and Possibility Theory

6.1 Introduction

Centrifugal pumps are one of the most important elements in almost all industries. They plan an important role in industries and it requires continuous monitoring to increase the availability of the pump. The pumps are widely used in food industry, waste water treatment plants, agriculture, oil and gas industry, paper and pulp industry, etc. [1-4]. However, faults of pump can cause a high rate of energy loss with associated performance degradation, even the breakdown of a whole system, and then lead to substantial economic losses. Therefore, monitoring of pumps is necessary to prevent a decrease in their efficiency. Condition monitoring of rotating machinery is very important for guaranteed production efficiency and safety of a machine in terms of system maintenance and process automation [5-7]. There are the different approaches which have been used for fault detection of centrifugal pumps. In [8-10] those studies have used indirect parameters or interrelated parameters to identify the faults, but not considered the features of a vibration signature.

Vibration signature is the most revealing information on the condition of rotating machinery [10-11]. When a rotating machine is operating properly, the vibration is small and constant, however, when some fault develops and some of the dynamic process changes, there will be changes in vibration spectrum observed [10]. Then accurate information is a crucial aspect of a maintenance regimen [12]. In the field of machinery
diagnosis, vibration signals are often used for fault detection and condition discrimination [13-15]. In [16] fault diagnosis based on vibration signals has been widely used in practice, and the used proportion is about 66%. And in [17] the study has shown that the vibration level changes with different faults of pump. Therefore, diagnosis based on vibration analysis seems to be the most reliable way to monitor pump state. However, vibration signals measured at any point of a centrifugal pump often contains strong noise, and then the feature extraction for fault diagnosis is very difficult. Particularly, in an early stage of a fault, effects of noise are so strong that the features of a fault are not evident. Moreover, the features of pump faults which easily occur in centrifugal pump, such as cavitation, impeller unbalance and shaft misalignment, are extracted in different frequency domains, and then it is also difficult to identify the fault type. Therefore, the effective method for carrying out automatic diagnosis in diagnosis system of a pump has not been established yet.

Furthermore, in the case of intelligent diagnosis of the plant machinery, many studies have been carried out to investigate the use of neural networks (NNs) for automatic diagnosis, most of these methods have been proposed to detect faults and identify fault types [10, 18]. However, NNs cannot reflect the ambiguous diagnosis problems, and will never converge during the learning phase of these networks [19].

To solve these problems, we propose an intelligent diagnosis for a pump system using statistic filter, support vector machine (SVM) and possibility theory (PT) on the basis of the vibration signals, to detect faults and identify fault types at an early stage, and its flowchart is shown in Fig. 6.1. The statistic filter is used to detect and extract the feature signals of pump faults across optimum frequency regions, and the feature signals extracted are reflected by non-dimensional symptom parameters (NDSPs). In order to increase the diagnosis’ sensitivity, synthetic symptom parameter (SSP) is defined with the function of the optimal classification hyper-plane obtained using SVM. Then the possibility distributions of the SSP are used to detect faults and identify fault types by possibility theory. Practical examples of the detection of incipient faults for the
centrifugal pump are provided to verify the effectiveness of the proposed method.

Fig. 6.1 Flowchart of intelligent diagnosis

6.2 Experimental centrifugal pump system for condition diagnosis

Fig. 6.2 Photograph of the experiment centrifugal pump system

Fig. 6.2 shows that the testing set-up consists of a pump, a motor and a close-loop water piping system. The SF-JRO series motor is employed to drive the pump through a
coupling, and the rotating speed can be varied through control panel. The experimental pump is a centrifugal pump described as fellows: HONDA Pump, Type: HAS, Head: 40 m, Output: 3.7 kW and Capacity: 7.5m³/h. The capacity of the tank is based on the maximum flow rate. But then the flow rate of pump can be also adjusted by the valve control system. Five accelerometers are used to acquire the vibration signals with a sampling frequency of 50 kHz. The sensor locations are shown in Fig. 6.3. One is mounted on the pump housing, two sensors in the piping direction and its vertical direction of the pump inlet respectively, other sensors in the piping direction and its horizontal direction of the pump outlet respectively.

In order to examine whether incipient faults of a pump were able to be detected, normal state (N) and two typical faults, namely cavitation (C) and shaft misalignment (M), were performed while the pump speed was at 3500 rpm and the valves in the suction and discharge lines were fully open. Moreover, each fault state was divided into three steps (mild, medium, severe) according to the defect degree to measure the vibration signals. Here, the numbers of 1, 2 and 3 are used to express each fault in mild, medium or severe step. For example, C₁ expresses cavitation state in mild step, and M₃ expresses shaft misalignment state in severe step. Then the time-domain waveforms of vibration signals on the pump housing in each state are shown in Fig. 6.4.
Fig. 6.4 Time-domain waveforms of vibration signals on the pump housing in each state. (a) Normal state. (b) Cavitation state in mild step (C1). (c) Cavitation state in medium step (C2). (d) Cavitation state in severe step (C3). (e) Shaft misalignment in mild step (M1). (f) Shaft misalignment in medium step (M2). (g) Shaft misalignment in severe step (M3).

6.3 Feature extraction using statistic filter

6.3.1 Statistic filter

Statistic filter is a very powerful filter which accepts tickets that have selected statistical properties, which has been widely used in many fields [16, 20-21]. In the field
of machinery diagnosis, a signal measured in normal condition is a standard of diagnosis, but a noise when pure features of abnormal signal are extracted. Therefore, it is necessary to smooth away noise, but leave abnormal signal unchanged [22]. Here, statistic filter based on statistical tests of spectrum is proposed, as shown in Fig. 6.5.

Fig. 6.5 Principle of statistic filter based on statistical tests of spectrum

When normal signal and diagnosis signal have been measured, we can divide the signals into $N$ signal parts. Spectrum analysis is performed to reflect each part of vibration signal, and the spectrum content $F_{ij}(f_k) (j = 1, 2, ..., N)$ of frequency $f_k (f_k = k \cdot f_{\text{max}}/M, k = 1, 2, ..., M)$ is obtained, where $i$ is the style of signal, such as normal signal (n) and diagnosis signal (d), $f_{\text{max}}$ and $M$ are the maximal frequency and the number of spectrum analysis. Thus the spectrum components $F_n(f_k)$ and $F_d(f_k)$ of frequency $f_k$ of normal signal and diagnosis signal are shown as follows.

$$F_n(f_k) = \{F_{n1}(f_k), F_{n2}(f_k), ..., F_{nM}(f_k)\} \quad (6.1)$$

$$F_d(f_k) = \{F_{d1}(f_k), F_{d2}(f_k), ..., F_{dM}(f_k)\} \quad (6.2)$$

In order to determine whether there is significant difference between the spectrum components $F_n(f_k)$ and $F_d(f_k)$ of frequency $f_k$ of normal signal and diagnosis signal, statistical test is used to provide a mechanism for making quantitative decision.
Here, the full hypothesis $H_0$ is “there is significant difference between the spectrum components $F_n(f_k)$ and $F_d(f_k)$ of frequency $f_k$ of normal signal and diagnosis signal”. When $H_0$ is rejected, in other words, the difference between the two is smaller than others, therefore, the spectrum components $F_d(f_k)$ of frequency $f_k$ of diagnosis signal will be removed as noise by a filter. Otherwise, $F_d(f_k)$ are left with no change. A simple example is used to explain a statistic filter based on statistical tests of spectrum, as shown in Fig. 6.6. Thus, the components extracted from diagnosis signal on the basis of normal signal can help to exhibit the pure feature of abnormal condition, and can be reconstructed in time domain by inverse fast Fourier transform (IFFT).

Fig. 6.6 Explanation of statistic filter based on statistical tests of spectrum. ■ is the value of the spectrum of normal signal; ● is the value of the spectrum of diagnosis signal. In these points, components corresponding to $f_2, f_4, f_6, f_8, ..., f_M$ between normal signal and diagnosis signal are not significant difference, but components indicated by arrow (▲) such as $f_3, f_5, f_7, ..., f_M$ of diagnosis signal are so higher than normal signal. Then components indicated by arrow (▲) are only extracted to detect pure feature of abnormal signal.

**6.3.2 Analysis of experimental signals using statistic filter**

To make the recomposed signals comparable regardless of differences in magnitude, the recomposed signals of each state are normalized by the following formula.
\[ x_i = \frac{x'_i - \mu'}{\sigma'} \]  

(6.3)

where \( x'_i \) \((i = 1, 2, \cdots, N)\) are the original signal series, \( \mu' \) and \( \sigma' \) are the mean and standard deviation of \( x'_i \), \( x_i \) \((i = 1, 2, \cdots, N)\) are the recomposed signal series after normalization.

To extract useful features from the measured various signals, the statistic filter based on statistical tests of spectrum is used. In present work, the normalized signal of each state has been divided into 16 signal parts and the spectrums of 16 signal parts have been gathered statistics in accordance with the unit of frequency, respectively. Then t-test is chosen to test whether the full hypothesis \( H_0 \) is rejected, and the statistic filter based on statistical tests of spectrum is performed. Fig. 6.7 is the spectrums of vibration signals from beginning to end in each state and by statistic filter.

### 6.4 Non-dimensional symptom parameter for intelligent diagnosis

For intelligent diagnosis, symptom parameters are needed that can sensitively detect the occurrence of any fault and distinguish the fault category. A large set of symptom parameters has been defined in the pattern recognition field. Here, the non-dimensional symptom parameters (NDSPs) in time domain, commonly used for the fault diagnosis of plant machinery, are considered.

When given the digital data \( x_i \) \((i = 1, 2, \cdots, N)\) of the vibration signal, where \( N \) is the number of the signal, the 4 NDSPs in the time domain are shown as follows.

\[
P_1 = \frac{\sum_{i=1}^{N} |x_i|^p}{N_p} \tag{6.4}
\]

\[
P_2 = \frac{\sum_{i=1}^{N} (x_i - \overline{x})^3}{N \cdot \sigma^3} \tag{6.5}
\]
Fig. 6.7 Spectrums of vibration signals from beginning to end in each state and by statistic filter. (b) cavitation state in mild step (C₁). (c) cavitation state in medium step (C₂). (d) cavitation state in severe step (C₃). (e) shaft misalignment in mild step (M₁). (f) shaft misalignment in medium step (M₂). (g) shaft misalignment in severe step (M₃).
\[ P_3 = \frac{\sum_{i=1}^{N} (x_i - \bar{x})^4}{N \cdot \sigma^4} \]  

(6.6) 

\[ P_4 = \frac{P_i}{\sigma} \]  

(6.7) 

where, \( \bar{x} \) and \( \sigma \) are the mean and standard deviation of the vibration signal \( x_i \). \( |x_i| \) is the absolute value of the vibration signal \( x_i \), \( |\bar{x}| \) is the mean of \( |x_i| \), and \( |x_i|_p \) is the peak value of \( |x_i| \). 

Popularly, the sensitivity of a symptom parameter (SP) is judged by discrimination index (DI) [23-24] as following Eq. (6.8). Here, \( \mu_1, \sigma_1 \) and \( \mu_2, \sigma_2 \) are the mean and standard deviation of the SP in state 1 and state 2. And then it is proved that the larger the value of the DI, the higher the sensitivity of the SP will be, and therefore, the better the SP will be. Moreover, when the DI value of two states is more than 1.65, the discernment rate of 95% is obtained or more. 

\[ DI = \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \]  

(6.8) 

Table 6.1 DI values of the normal state and each fault state in different step 

<table>
<thead>
<tr>
<th>Two states</th>
<th>( P_1 )</th>
<th>( P_2 )</th>
<th>( P_3 )</th>
<th>( P_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>N, M₁</td>
<td>0.47</td>
<td>0.84</td>
<td>0.85</td>
<td>0.63</td>
</tr>
<tr>
<td>N, M₂</td>
<td>0.49</td>
<td>0.82</td>
<td>0.93</td>
<td>0.57</td>
</tr>
<tr>
<td>N, M₃</td>
<td>0.55</td>
<td>0.91</td>
<td>0.90</td>
<td>0.62</td>
</tr>
<tr>
<td>N, C₁</td>
<td>0.35</td>
<td>0.76</td>
<td>0.83</td>
<td>0.41</td>
</tr>
<tr>
<td>N, C₂</td>
<td>0.30</td>
<td>0.79</td>
<td>0.85</td>
<td>0.39</td>
</tr>
<tr>
<td>N, C₃</td>
<td>0.38</td>
<td>0.68</td>
<td>0.79</td>
<td>0.44</td>
</tr>
</tbody>
</table>

Then the DI values of the normal state and each fault state in different step are calculated as shown in Table 6.1. However, the DI values of these SPs are almost lower
at an incipient stage of a fault. Therefore, in this study, time and effort must not be expended to find which SPs are effective, but the new method, those primitive SPs are integrated an effective SP for diagnosis by introducing weight coefficient, is proposed. Here, this effective SP is called “synthetic symptom parameter (SSP)”.

6.5 Application of SVM for synthetic symptom parameter

Fig. 6.8 Optimal classification hyper-plane for identifying state1 and state 2 using soft margin SVM; The white circles (○) represent state 1, the black points (●) represent state 2, \( x_i \) \((i = 1, 2, 3)\) represent 3-dimensional sample vector. The colorized plane is the optimal classification hyper-plane.

As shown in chapter 5.5, soft margin SVM is used to obtain an optimal classification hyper-plane, as shown in Fig. 6.8, and then define a new synthetic symptom parameter (SSP) [25-30]. When original SPs are inputted into SVM and the ranges of some discernment error are prepared, the optimal classification hyper-plane as shown in Eq. (6.9) is constructed. Here, \( \omega \) is the optimal vector of the weight coefficients. Then the optimal classification hyper-plane is defined as synthetic symptom parameter (SSP). Therefore, the weight coefficients \( \omega \) decide the diagnostic sensitivity of the SSP.
\[ f(x) = \omega \cdot x + b \]  \hspace{1cm} (6.9)

In the study, the vibration signals of each fault have been measured in mild, medium and severe step. In order to enhance the diagnostic ability of the method proposed in the paper, the SPs of each fault in three steps and normal state are used to define new SSPs, shown in Table 6.2.

Table 6.2 New SSPs of each fault and normal state

<table>
<thead>
<tr>
<th>Two states</th>
<th>Synthetic symptom parameter (SSP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N, M</td>
<td>( y = -1.4P_1 + 0.7P_2 - 3.3P_3 - 2.5P_4 - 0.7 )</td>
</tr>
<tr>
<td>N, C</td>
<td>( y = 0.9P_1 + 1.0P_2 - 1.7P_3 - 3.1P_4 + 1.1 )</td>
</tr>
</tbody>
</table>

Table 6.3 New SSPs of each fault and normal state

<table>
<thead>
<tr>
<th>Two states</th>
<th>DI value of SSP</th>
<th>Two states</th>
<th>DI value of SSP</th>
</tr>
</thead>
<tbody>
<tr>
<td>N, M_1</td>
<td>1.29</td>
<td>N, C_1</td>
<td>1.12</td>
</tr>
<tr>
<td>N, M_2</td>
<td>1.41</td>
<td>N, C_2</td>
<td>0.97</td>
</tr>
<tr>
<td>N, M_3</td>
<td>1.55</td>
<td>N, C_3</td>
<td>1.28</td>
</tr>
</tbody>
</table>

Then the SSPs of each fault in different step and normal state are calculated, and the DI values of the SSPs are also calculated as shown in Table 6.3. It is obvious that the DI value of SSP is larger than the DI values of an individual SP. Therefore, SSPs are used to raise the discernment sensitivity.

6.6 Fuzzy diagnosis using possibility theory

In this section, possibility theory is used to build a fuzzy diagnosis system, and some usual faults of a centrifugal pump have been diagnosed to verify that the method
proposed in this paper is effective.

6.6.1 Fuzzy inference using possibility theory

For fuzzy inference, the membership function of a SP is necessary [31-32]. In this paper, possibility theory is also used to change to possibility distribution function from probability density function of a SP, as shown in chapter 5.6. However, the matching method of possibility function is different. The paper proposes that the matching method of possibility function is the degree of the common area between two possibility functions. Fig. 6.9 shows the matching examples of possibility function. The common area, between the possibility functions \( p_i(h) \), \( p_j(h) \) of model state \( i, j \) and the possibility functions \( p(x) \) of diagnostic state, is calculated by the following formulas.

\[
S_i = \int_{\bar{x}_d - 3\sigma_d}^{\bar{x}_d + 3\sigma_d} \psi_i(x) dx, \quad \psi_i(x) = \min\{ p_i(x), p(x) \} \tag{6.10}
\]

\[
S_j = \int_{\bar{x}_d - 3\sigma_d}^{\bar{x}_d + 3\sigma_d} \psi_j(x) dx, \quad \psi_j(x) = \min\{ p_j(x), p(x) \} \tag{6.11}
\]

Where \( \bar{x}_d \) and \( \sigma_d \) are the mean and the standard deviation of a SP in diagnostic
Chap. 6 Precision Diagnosis Method for a Centrifugal Pump using ST, SVM, PT

state. Moreover, the unknown state except state $i, j$ is the model state \( u \), its possibility functions $p_u(x)$ is calculated by the following formula (6.12). And the common area between $p_u(x)$ and $p(x)$ is calculated by the following formula (6.13).

$$p_x(h) = \max\{1 - [p_i(h) + p_j(h)], \quad 0\} \quad (6.12)$$

$$S_{un} = \int_{x_{ri} - 3\sigma}^{x_{ri} + 3\sigma} \psi_{un}(x)dx, \quad \psi_{un}(x) = \min\{p_u(x), p(x)\} \quad (6.13)$$

Therefore, the probability, a diagnostic state should be diagnosed with model state $i, j$ and $u$, respectively, can be calculated by the following formulas. Then it is judged that the diagnostic state is the model state with the most probability.

$$\omega_i = \frac{S_i}{S_i + S_j + S_{un}} \quad (6.14)$$

$$\omega_j = \frac{S_j}{S_i + S_j + S_{un}} \quad (6.15)$$

$$\omega_{un} = \frac{S_{un}}{S_i + S_j + S_{un}} \quad (6.16)$$

### 6.6.2 Building diagnosis system

SSP, as the object of a follow-on process, is obtained using soft margin SVM and primitive 4 NDSPs, and then the possibility functions of SSPs, as the membership function, are used to build a fuzzy diagnosis system, as shown in Fig. 6.10. In order to achieve an intelligent and efficient condition diagnosis, sequential diagnosis is proposed. In the first step, normal state and misalignment state are regarded as diagnosis target, and then the unknown state except two states is also regarded as diagnosis target. Therefore, possibility functions of SSPs of normal state, misalignment state and unknown state are used to diagnose whether the diagnosis signal is normal state, misalignment state or unknown state.

When normal state or misalignment state, the diagnosis system can stop automatically, and the data of the diagnosis signal will be added to the model data. Otherwise, the condition cannot be identified, then diagnosis system will come into the second step.
In the second step, normal state and cavitation state are regarded as diagnosis target, and then the unknown state except two states is also regarded as diagnosis target. Thus, the second step is designed to consist of possibility functions of SSPs of normal state, cavitation state and unknown state. Similarly, when normal state or cavitation state, the diagnosis system can stop automatically, and the data of the diagnosis signal will be added to the model data. If unknown state, the vibration signal is analyzed until the fault type is confirmed. And then the diagnosis system will add a step, the data of the signal will become the model data of the fault. And so on, the diagnosis system can diagnose more faults. In this study, simulation experiment is that normal state and two typical faults of centrifugal pump were performed. Therefore, the experimental data are
processed into training data, and then the training data are input SVMs to obtain the SSPs for diagnosis. Then the possibility functions of SSPs in each step are used to build the diagnosis system.

6.6.3 Diagnosis classification

To verify the proposed diagnostic method, the test data, measured in each known state and not used for the pre-calculated possibility functions, are used. The practical diagnostic examples in the first and second steps are shown in Fig. 6.11 and Fig. 6.12.

Fig. 6.11 shows the practical diagnosis example in the first step. In Fig. 6.11, normal state (N), misalignment state (M) and unknown state except two states (UN) are diagnosis model states. \( p_N(x) \), \( p_M(x) \) and \( p_{UN}(x) \) express the possibility functions of the SSPs in normal state (N), misalignment state (M) and unknown state except two states (UN), respectively. Here, four known states (N, M1, M2, M3) are used as the test data, and \( p(x) \) is used to express the possibility functions of the SSPs for the test state. Moreover, \( S_N \), \( S_M \) and \( S_{UN} \) are the matching degrees between the test state and the model states. \( \omega_N \), \( \omega_M \) and \( \omega_{UN} \) are calculated by Eq. (6.14), Eq. (6.15) and Eq. (6.16), as shown in Table 6.4. Each state has been correctly diagnosed so that it might understand.

<table>
<thead>
<tr>
<th>Test state</th>
<th>N</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnosis probability of each model state</td>
<td>N $p_N(x)$</td>
<td>0.647</td>
<td>0.197</td>
<td>0.098</td>
</tr>
<tr>
<td></td>
<td>M $p_M(x)$</td>
<td>0.332</td>
<td>0.768</td>
<td>0.812</td>
</tr>
<tr>
<td></td>
<td>UN $p_{UN}(x)$</td>
<td>0.031</td>
<td>0.035</td>
<td>0.090</td>
</tr>
<tr>
<td>Judge state</td>
<td>N</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
</tbody>
</table>

Table 6.4 Diagnosis results in the first step
Fig. 6.11 Practical diagnosis examples of four vibration signals in the first step. (A) Normal state (N). (B) Shaft misalignment in mild step (M₁). (C) Shaft misalignment in medium step (M₂). (D) Shaft misalignment in severe step (M₃).
Fig. 6.12 Practical diagnosis examples of four vibration signals in the second step. (A) Normal state (N). (B) Cavitation state in mild step (C1). (C) Cavitation state in medium step (C2). (D) Cavitation state in severe step (C3).
Fig. 6.12 shows the practical diagnosis example in the second step. In Fig. 6.12, normal state (N), cavitation state (C) and unknown state except two states (UN) are diagnosis model states. $p_N(x)$, $p_C(x)$ and $p_{UN}(x)$ express the possibility functions of the SSPs in normal state (N), cavitation state (C) and unknown state except two states (UN), respectively. Here, four known states (N, C, C, C) are used as the test data, and $p(x)$ is used to express the possibility functions of the SSPs for the test state. Moreover, $S_N$, $S_C$ and $S_{UN}$ are the matching degrees between the test state and the model states. $\omega_N$, $\omega_C$ and $\omega_{UN}$ are calculated by Eq. (6.14), Eq. (6.15) and Eq. (6.16), as shown in Table 6.5. Each state has been correctly diagnosed so that it might understand.

<table>
<thead>
<tr>
<th>Test state</th>
<th>Diagnosis probability of each model state</th>
<th>N</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>0.577</td>
<td>0.332</td>
<td>0.282</td>
<td>0.110</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.388</td>
<td>0.616</td>
<td>0.702</td>
<td>0.802</td>
</tr>
<tr>
<td></td>
<td>UN</td>
<td>0.035</td>
<td>0.052</td>
<td>0.016</td>
<td>0.088</td>
</tr>
<tr>
<td>Judge state</td>
<td>N</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
</tbody>
</table>

### 6.7 Conclusions

To intelligently diagnose some typical faults of a centrifugal pump at an early stage, a new intelligent fuzzy diagnosis method based on vibration signals was proposed, and the effectiveness was demonstrated experimentally. The superiority of the method proposed in this paper can be explained by the following points:

1. Statistic filter is a method of signal processing that failure signal is extracted by statistical tests of spectrums between normal signal and fault signal. The application of statistic filter is very effective, even if there will be a slight fault at an early stage, statistic filter can extract effectively pure feature signal.
(2) Soft margin SVM is used to define synthetic symptom parameter (SSP), and then to raise the discernment sensitivity of original SPs.

(3) The efficiency of intelligent fuzzy diagnosis proposed here has been verified by applying it to a practical diagnosis for incipient faults of the centrifugal pump, and its performance is high. In fact, the high performance of intelligent diagnosis using possibility functions of the SSPs is attributed primarily to effective feature extraction by statistic filter, SVMs’ generalization capability and SSPs’ high sensitivity.

In the near future, the method proposed in this paper will be applied to condition diagnosis in various types of rotating machinery in a real plant.

References

Chap. 6 Precision Diagnosis Method for a Centrifugal Pump using ST, SVM, PT


148
Chap. 6 Precision Diagnosis Method for a Centrifugal Pump using ST, SVM, PT


Chapter 7

Conclusions and Future Research

7.1 Conclusions

In the field of condition diagnosis of the plant machinery, particularly rotating machinery, vibration information is widely used to detect a fault and identify the fault types. Furthermore, the utilization of vibration information is effective and the used proportion is about 66%. Certainly, condition diagnosis based on vibration information depends largely on feature extraction. Only when the features of vibration information are sensitively extracted in any condition of a machine can condition diagnosis be effective. However, plant machines are operating under unsteady conditions, even if the machines are in the normal state, rotating speed and operating load can vary. These can influence the spectrum feature of the vibration information measured. Moreover, vibration information often contain strong noise, especially at an early stage of a fault. Therefore, it is difficult to extract the features of vibration information. Furthermore, when building an intelligent system for diagnosing the condition of plant machinery, symptom parameters (SPs) and artificial intelligence (AI) are required. A high sensitive SP can express useful information of machine conditions. However, in most cases of condition diagnosis for plant machinery, the sensitivity of some SPs is not high. The main reasons can be explained as follow: (1) the pure feature signals are weak in vibration signal, because a fault is at an early stage, or the measure point is farther from failure part; (2) previous work namely feature extraction is unsatisfied; (3) the selected SPs cannot sensitively reflect the conditions of monitoring machine. In the case of AI,
neural network (NN), genetic algorithm (GA), support vector machine (SVM), etc. have some special advantages as well as some disadvantages. For example, NN and GA will never converge when the first-layer data have the same values in different states, and SVM is only a two-class classifier. Moreover, these methods cannot deal with ambiguous classification problems.

Therefore, in the introduction of this thesis, I have considered three main research objectives that this thesis aimed to overcome these questions. In order to extract effectively the features of vibration signals, improve the ambiguous relationship between SPs and machine conditions, strengthen the sensitivity of SPs, and increase the efficiency of condition diagnosis at an early stage, this thesis has focused on intelligent condition diagnosis based on support vector machine approach for rotating machinery.

The main conclusions are summarized as follows:

(1) The thesis proposed an intelligent condition diagnosis for rotating machinery, and the main diagnosis approach was support vector machine (SVM). SVMs not only have obtained good convergence, but also have accurately, quickly and automatically detected whether machine condition was normal or not, and identified fault types.

(2) The thesis proposed a sequential diagnosis system based on support vector machine (SVM). Sequential diagnosis system not only avoided that time and effort must be expended to find effective symptom parameters (SPs) for multiple faults, but also overcame SVM’s disadvantage that SVM is only a two-class classifier.

(3) The thesis defined many types of dimensional symptom parameters (DSPs) and non-dimensional symptom parameters (NDSPs) in the time domain and in the frequency domain for reflecting the features of vibration signals measured in machine condition. Moreover, relative ratio symptom parameters (RRSPs) were defined to exclusively detect and identify structural faults of rotating machinery.

At the same time, discrimination index (DI) and principal component analysis
(PCA) were employed to evaluate the sensitivity of DSPs, NDSPs or RRSPs, respectively. Each one type of SPs has its good points. In the case of structural faults of rotating machinery, the sensitivity of RRSPs was very higher.  

4) The thesis proposed common distinctive frequency components (DFCs) which are the spectrum features of structural faults extracted from vibration signals in rotating machinery. The normalized DFCs not only decreased feature differences from a state with varying operational conditions, but also replaced effective symptom parameters (SPs). Moreover, DFCs, as the object of a follow-on process, were applied to diagnose structural faults in rotating machinery, and its efficiency was verified.  

5) The thesis proposed soft margin SVM is applied in fault diagnosis. Since the vibration signals from different conditions have only changing tendency, but not obvious boundary, soft margin SVM is used to obtain an optimal hyper-plane of linear classification in two conditions. The practical examples of condition diagnosis based on soft margin SVM verified that soft margin SVM has a strong generalization capacity when only a small amount of training samples are available.  

6) The thesis proposed that soft margin SVM is used to merge several symptom parameters (SPs) into a synthetic symptom parameter (SSP) for increasing the diagnosis’ efficiency. Practical examples of condition diagnosis for bearing faults and pump faults verified that the method was effective.  

7) The thesis proposed optimal composition of symptom parameters (OCSPs) are applied in the field of condition diagnosis. OCSPs are the synthetical diagnosis information using symptom parameters from multiple directions or sensors. The practical example of diagnosing simple faults and multi-faults of a roller bearing verified the application of OCSPs was very strong, even if at an early stage of a fault.  

8) The thesis proposed statistic filter is applied for fault diagnosis. Statistic filter is a
new feature extraction method that failure signal is extracted from vibration signal measured in machine condition by statistical tests of spectrums between normal signal and measured signal. The practical example of incipient faults of the centrifugal pump verified statistic filter can extract effectively pure feature signal, even if at an early stage of a fault.

(9) The thesis proposed a fuzzy diagnosis using possibility theory to detect faults and identify fault types. The possibility function of a symptom parameter (SP) was established by possibility theory to achieve fuzzy inference, and then to process the ambiguity problem of condition diagnosis and represent complex relationships between symptoms and fault types. Practical examples of condition diagnosis for multi-faults of a roller bearing and incipient faults of the centrifugal pump verified that the fuzzy diagnosis method was effective.

(10) The intelligent diagnosis methods proposed in this thesis had been used to successfully monitor and diagnose practical rotating machinery, such as a centrifugal fan, a centrifugal pump. These diagnosis techniques will be applied to condition diagnosis in various types of rotating machinery in a real plant.

### 7.2 Future research

This thesis has mainly focused on the research of intelligent condition diagnosis based on support vector machine approach for rotating machinery. Certainly, many studies on condition diagnosis for plant machinery have been carried out, and achieved excellent results. However, many technical problems on condition diagnosis are not yet resolved for real plant, which offer perspectives for future research. In the following, the key objective of future research will be outlined.

First, I shall focus on further development and integration of diagnosis approaches for rotating machinery. Although a large number of diagnosis methods have been proposed, each method is applicable on a given machine, a special part or some typical faults.
Integration of methods into a diagnosis tool being capable of addressing the entire diagnosis problem is required.

Second, I shall promote the development of prognostic approaches and life prediction in plant machinery. So long as the useful information such as symptom parameters (SPs) of every diagnosis has been collected, the degree of the fault will be evaluated. Moreover, if only the number of efficient diagnosis will reach up to the required quantity, the variation trend of the machine will also be predicted, and then the remaining life of the machine will be predicted. Thus, not only will the effectiveness of condition-based maintenance (CBM) be enhanced, but also the optimal time of machine or part replacement will be judged effectively.