Drilling Process Monitoring

---- drill wear prediction and drilling conditions recognition with newly generated features

(ドリル加工のモニタリング —工具摩耗予測・加工状態認識の新方法)

学位取得年月 2014 年 9 月

徐 杰

Drilling Process Monitoring

---- drill wear prediction and drilling conditions recognition with newly generated features

A Doctoral Dissertation

by

Xu Jie

Contents

List of illustrations I

List of tables IV

Abstract ..... A

Chapter 1 Introduction 1

1.1 Research Background 1

1.1.1 Manufacturing, Machining and Drilling 1

1.1.2 Intelligent Machining 2

1.1.3 Machining Monitoring 4

1.1.4 Drilling Process Monitoring 7

1.2 New Ideas and the Purpose 7

1.2.1 New Ideas of Feature Generation and Extraction, Drill Condition Recognition Diagnostic 7

1.2.2 Purpose and Configuration of the Research 9

1.3 Design of the Research 10

1.3.1 Applied Technologies 10

1.3.2 Experimental Design 10

1.4 Organization of the Dissertation 10

Chapter 2 Literature Review 12

2.1 Monitoring Scopes 12

2.2 Sensors/Sensor systems 14

2.3 Signal Analysis 16

2.3.1 Signal Pre-processing 16

2.3.2 Time Domain Analysis 17

2.3.3 Frequency Domain analysis 20

2.3.4 Time-Frequency Domain analysis 21

2.3.5 Without Signal Analysis 24

2.4 Decision-making Support Systems 25

2.4.1 Threshold 26

2.4.2 Statistical 27

2.4.3 Fuzzy logic 27

2.4.4 Neural network 27

2.4.5 Hidden Markov model 29

Chapter 3 Methodologies 30
3.1 Sensor system and General Setup 30
3.2 Signal Segmentation 31
3.3 Feature Generation 31
3.3.1 The Relations between the Thrust Force and Torque Signals 31
3.3.2 The Novel Feature Generation Method 34
3.3.3 The Features Grouping 37
3.4 Feature Extraction 37
3.4.1 Fourier Transform and Wavelet Transform 37
3.4.2 Wavelet Packet Decomposition and Reconstruction (WPD&R) 38
3.5 Feature Selection 40
3.5.1 The Necessity for Applying PCA 40
3.5.2 The PCA Algorithm 42
3.6 Decision-making support AI System 43
3.6.1 The Architecture and the Training Procedure of the BPNN 44
3.6.2 The Training Algorithm of the BPNN 45
3.6.3 The Pattern Recognition by BPNN 47
Chapter 4 Findings of the Relations between Features and Drilling Process 48
4.1 Workpiece Material 48
4.1.1 Static features 48
4.1.2 Dynamic Features 50
4.1.3 Analysis 53
4.2 Drill Diameter 54
4.2.1 Static Features 55
4.2.2 Dynamic Features 56
4.3 Spindle Speed 58
4.3.1 Static Features 58
4.3.2 Dynamic Features 59
4.4 Feed Rate 61
4.4.1 Static Features 62
4.4.2 Dynamic Features 63
4.4.3 Investigation of the Effect of Feed Rate 65
4.5 Drill Wear 67
4.5.1 Static Features 68

4.5.2 Dynamic Features 69
4.6 Summary 71
Chapter 5 Drill Wear Prediction and Drilling Conditions Recognition 73
5.1 Data Preparation 73
5.2 Drill Wear prediction with Feature Selected by PCA 74
5.2.1 Feature selection by PCA 74
5.2.2 Training with features selected by PCA 77
5.3 Drill Wear Prediction Directly with the BPNN 78
5.3.1 Finding the major cluster 78
5.3.2 Finding the major frequency band 79
5.3.3 Comprehensive evaluation 80
5.4 Drilling Conditions Recognition 82
5.4.1 Workpiece material 83
5.4.2 Drill diameter 85
5.4.3 Spindle speed 86
5.4.4 Feed rate 87
5.5 Summary 89
Chapter 6 Conclusion 91
References 92
Acknowledgement 98

List of illustrations

Figure 1-1 A typical intelligent machining system architecture 3
Figure 1-2 Machining monitoring technologies in sequences of common practical situations 6
Figure 1-3 Share of different scopes among previous research 6
Figure 1-4 Machine coordinate and tool coordinate in turning [18] 7
Figure 1-5 Evaluation method of fluctuation in the dynamic components [18] 8
Figure 1-6 Phase plane and reference rectangle [19] 8
Figure 2-1 Locations and patterns of tool wear on a twist drill [23] 13
Figure 2-2 AE-RMS over the life of a 1.5 mm diameter twist drill [63] 15
Figure 2-3 Typical analog preprocessing for AE signals [16] 17
Figure 2-4 Segmentation of the spindle power signal into ‘‘*n*’’ equidistant sections [63] 17
Figure 2-5 The influence of drill wear to kurtosis and skewness of vibration signal data points [44] 18
Figure 2-6 The impedance conceptual model of a drilling system [59] 19
Figure 2-7 Motor impedance signals of drilling processes using normal and breakage drills [59] 19
Figure 2-8 Torque signals for a sharp drill and a corner worn drill [45] 20
Figure 2-9 The 2 flat regions of WT coefficients [25] 22
Figure 2-10 Waveforms of WT coefficients of spindle current signals for a normal and a damaged drill [54] 23
Figure 2-11 A 10 state HMM superimposed onto the joint scatter plot [49] 25
Figure 3-1 The signal system and general experimental set up 30
Figure 3-2 Three stage of the drill process and signal Segmentation 31
Figure 3-3 The cross-correlation of original and DC removed sampled thrust force and torque 32
Figure 3-4 The coherence of thrust force and torque 34
Figure 3-5 Twist drill geometry [70] 34
Figure 3-6 The conversion of thrust force and torque 35
Figure 3-7 Definition of static components 36
Figure 3-8 Definition of dynamic components 36
Figure 3-9 Orbit diagram of cutting force signals in dual directions measured by integrated force [76] 41
Figure 3-10 Architecture of a 3 layered feed forward neural network 44
Figure 3-11 Back propagation algorithm procedure 45
Figure 3-12 Mathematical structural illustration of the BPNN 46
Figure 3-13 The application of BPNN for pattern recognition [82] 47
Figure 4-1 The relations between the workpiece material and the average features 49
Figure 4-2 The relations between the workpiece material and the RMS features 49
Figure 4-3 The relations between the workpiece material and the delta features 50
Figure 4-4 The relations between the workpiece material and the standard deviation features 50

Figure 4-5 The relations between the workpiece material and the velocity features 51
Figure 4-6 The sampled *Fp* values for S45C and Ti-alloy 51
Figure 4-7 The frequency domain of the *Fp* for S45C and Ti-alloy in frequency range (0-250Hz) 52
Figure 4-8 The frequency domain of the *Fp* for S45C and Ti-alloy in frequency range (250-1000Hz) 52
Figure 4-9 The relation between the workpiece material and the geometry feature 53
Figure 4-10 Radar chart of the material properties 54
Figure 4-11 The relations between drill diameter and the average features 55
Figure 4-12 The relations between drill diameter and the RMS features 55
Figure 4-13 The relations between drill diameter and the delta features 56
Figure 4-14 The relations between drill diameter and the standard deviation features 56
Figure 4-15 The relations between drill diameter and the velocity features 57
Figure 4-16 The relation between drill diameter and the *Fr_ConvHullArea* 57
Figure 4-17 The relations between spindle speed and the average features 58
Figure 4-18 The relations between spindle speed and the RMS features 59
Figure 4-19 The relations between spindle speed and the delta features 59
Figure 4-20 The relations between spindle speed and the standard deviation features 60
Figure 4-21 The relations between spindle speed and the velocity features 60
Figure 4-22 The relations between spindle speed and the *Fr_ConvHullArea* 61
Figure 4-23 The rectangular coordinate plot of *Ft* to *Fp* with different feed rate 62
Figure 4-24 The relations between feed rate and the average features 62
Figure 4-25 The relations between feed rate and the RMS features 63
Figure 4-26 The relations between feed rate and the delta features 63
Figure 4-27 The relations between feed rate and the standard deviation features 64
Figure 4-28 The relations between feed rate and the velocity features 64
Figure 4-29 The relations between feed rate and the *Fr_ConvHullArea* 65
Figure 4-30 Average and variation of the chip thickness with different feed rate 66
Figure 4-31 Hole side wall roughness for different feed rate 66
Figure 4-32 The definition of drill corner wear 67
Figure 4-33 The relations between drill corner wear and average features of forces 68
Figure 4-34 The relations between drill corner wear and static, delta and RMS features of θ 68
Figure 4-35 The relations between drill corner wear and RMS features of forces 69
Figure 4-36 The relations between drill corner wear and delta features of forces 69
Figure 4-37 The relations between drill corner wear and standard deviation features of forces 70
Figure 4-38 The relations between drill corner wear and velocity features of forces 70
Figure 4-39 The relations between drill corner wear and *Fr_ConvHullArea* 71
Figure 5-1 The accumulative contribution rate for the 19 principal components 75
Figure 5-2 The first 3 eigenvectors from features to PCs 76
Figure 5-3 The prediction result with features selected by PCA 78
Figure 5-4 Prediction performance of cluster 3 features in different frequency bands 79

Figure 5-5 The values in the first eigenvector from dynamic features to PC1 80
Figure 5-6 The prediction result by directly using BPNN 81
Figure 5-7 Relative importance of different features and condition parameters 82
Figure 5-8 The recognition result for workpiece materials with features selected by PCA 83
Figure 5-9 The recognition result for workpiece materials with features of cluster 3 84
Figure 5-10 Recognition result of drill diameter with PCA selected features 85
Figure 5-11 Recognition result of drill diameter with cluster 3 features 86
Figure 5-12 Recognition result of spindle speed with PCA selected features 86
Figure 5-13 Recognition result of spindle speed with cluster 3 features 87
Figure 5-14 Recognition result of feed rate with PCA selected features 88
Figure 5-15 Recognition result of feed rate with cluster 3 features 88
Figure 5-16 The relative importance of different features in the BPNN for feed rate recognition ... 89

List of tables

Table 1-1 Subjects of the fundamental functional actions for different machining paradigms 3
Table 1-2 Major machining monitoring scopes 5
Table 2-1 Classification of different monitoring scopes in previous literatures 12
Table 2-2 Categories of sensors deployed in previous literatures 15
Table 2-3 Different time domain signal features in previous literatures 18
Table 2-4 Summary of the application of WT and WPT in previous literatures 22
Table 2-5 Different decision-making support systems applied in previous literatures 26
Table 3-1 Features of *Ft, Fp and Fr* 35
Table 3-2 The grouping of different features 37
Table 3-3 Wavelet packet transform decompose 40
Table 3-4 Frequency bands of the decomposed sub-signals 40
Table 4-1 Conditions for different workpiece materials 48
Table 4-2 Relevant material properties 54
Table 4-3 Experimental Conditions for different drill diameters 54
Table 4-4 Conditions for experiments with different spindle speed 58
Table 4-5 Conditions for experiments with different feed rate 61
Table 4-6 Experimental conditions for evaluation of drill corner wear to features 67
Table 5-1 Condition parameter values for 256 data patterns 73
Table 5-2 The eigenvalues and ACR for 256×19 data matrix 75
Table 5-3 Number of the PCs and ACR for all of the 16 frequency bands 76
Table 5-4 Clustering of the 19 features for evaluation 79
Table 5-5 The prediction performances for 3 clusters 79

Abstract

The automation of machining processes has made great progress in the last several decades, and the future direction of machining technology development is the intellectualization. In this research, the term “machining” is constrained in the scope of conventional “metal cutting processes”, which specifically are turning, milling, drilling, sawing, planing, grinding, etc. There are 3 basic fundamental functional actions for machining processes which are monitoring, decision making and control.

This dissertation concerns the monitoring of drilling process. The prediction of drill corner wear and the recognition of drilling conditions are both investigated. New features generation, extraction and selection methods, how are these features are affected by the drilling condition, and the result of the prediction of drill corner wear and recognition of drilling conditions are mainly discussed in this dissertation.

In chapter 1, at first, related back ground is introduced, new ideas and methods are proposed, and the purpose of this work is announced, then the research design is explained, at last, the organization of the dissertation is summarized.

In chapter 2, a state-of-the-art literature review on tool condition monitoring is offered. Different methodologies and technologies applied are categorized and compared.

In chapter 3, the methodologies applied in this research are detailed. First the sensor system and the general experimental set up are introduced, and then the signal processing method and detailed feature generation procedure are discussed. After that the feature extraction method using wavelet packet transform is presented and then the feature selection methods are explicated. At last section the artificial neural network (ANN) model based regression and recognition approaches are specified.

In chapter 4, how the generated and extracted features are affected by drilling condition parameters such as workpiece material, drill diameter, spindle speed, feed rate and the drill corner wear are revealed.

In chapter 5, the prediction of drill wear and the recognition of drilling condition parameters are carried out. Two different features selection methods are applied, and one is using principal component analysis (PCA) for sub signals in each frequency band, the other is directly using ANN to select the major feature cluster and major frequency band.

In chapter 6, conclusions are made.

This dissertation examines the integrated procedure and methods for drilling process monitoring on the scopes of drill wear prediction and drilling conditions recognition. It is found that the dynamic features are more sensitive to the drilling status than static features, and they also contribute to the principal components more than static ones, but the change of drilling conditions leads to a more clear change of static features than dynamic ones. Good drill wear prediction and drilling condition recognition results are obtained.

Chapter 1 Introduction

In this chapter, at first, related back ground is introduced, new ideas and methods are proposed, and the purpose of this work is announced, then the research design is explained, at last, the organization of the dissertation is summarized.

In section 1, brief back ground of this research is provided. First, definitions and relations of manufacturing, machining, and drilling are carried out. Then the development of machining technologies is briefed and the concept, requirement and architecture of intelligent machining are presented. After that, review and summary of machining monitoring are specified.

In section 2, new ideas of feature generation and extraction, and drill condition recognition diagnostic are presented, and the purpose of this research is pointed out.

In section 3, the methodologies applied and the design of this research is explicated.

In section 4, the organization of the dissertation is provided.

1.1 Research Background

1.1.1 Manufacturing, Machining and Drilling

Manufacturing is a very heterogeneous collection of production activities which refer to a range from handicraft to high tech, but most commonly applied to industrial production, in which raw materials are transformed into finished goods on a large scale. The major processes of manufacturing according to the similarity of functions can be generally categorized as casting, molding, forming, machining, joining and assembling.

Machining is the broad term used to describe the removal of material from a workpiece and is one of the most important manufacturing processes [1]. In manufacturing realm, machining is the removal of the unwanted material from the workpiece so as to obtain a finished product or part of the desired size, shape and surface quality.

Machining is a part of the manufacture of many metal products, but it can also be used on materials such as metals, alloys, woods, plastics, ceramics, and composites. Moreover, recently, advanced machining techniques such as electrical discharge machining (EDM), electro-chemical erosion, laser cutting, and water jet cutting, have been developed to shape metal workpieces. However, in this work, the term machining is constrained in the scope of conventional metal cutting processes, which specifically are turning, milling, drilling, grinding, etc.

Drilling is a cutting process that uses a drill bit to cut or enlarge a hole of circular cross-section in solid materials. Drilling is one of the most important metal cutting operations, which comprises a major part of all metal cutting operations [2].

Drilling is often carried out as one of the last steps in the manufacturing production of a part, thus high demands in terms of process reliability are therefore placed on the operation. And furthermore, comparing to other machining processes (such as turning and milling), drilling can be seen as more complicated in terms of the kinematics and dynamics, the process control, the chip removal and thereby

the dissipation of the cutting heat.

Therefore, developments of technologies for drilling process are highly demanded.

1.1.2 Intelligent Machining

Machining has traditionally been one of the major operations within most manufacturing systems. However, machining processes are inherently complex, nonlinear, multivariate and often subjected to various unknown external disturbances. Therefore, traditionally, a machining process is usually performed by a skilled operator, who uses his decision-making capabilities based on the intuition and rules of thumb gained from experience.

Since many new theories and methods have been contributed to machining science, many new technologies have been brought to machining industries. And given the importance of machining to most industries, machine tools have often led the way in the development of automation technology.

Machine tool automation began in 1950s with the introduction of numerical control (NC) technology which opened doors in the early 1970s to computer numerical control (CNC) and direct numerical control (DNC) machining centers that enhanced product accuracy and uniformity. Latterly, machine tool dynamometers were used with machine tools to measure, monitor and control forces generated during machining processes which are closely related to product accuracy and surface integrity. The recent trend in machining is to add intelligence to the machining process to automate the machining processes[3].

Automation and intellectualization at the process level for machining operations and machine tools has been a focus of research attention in both academia and industry for several decades. Research in this area has carried strong expectations in the context of increased productivity, improved part quality, reduced costs, and relaxed part design constraints.

An intelligent machining system can be compared with a human operator. A number of sensors provide feedback to the system in a way the sensory organs provide feedback to a human operator. And then the feedback is processed by an automated system based on a computer in real time like by the human brain, and also by which the appropriate decisions are made to react optimally to different conditions. For a human operator, the machining technique comes from training and practice; while for an intelligent machining system, the automation is based on the physics of machining processes. The physics of machining processes are usually converted in mathematical forms by models based on the data obtained from sensors. Modeling of machining processes for an intelligent machining system is usually accomplished by soft computing with artificial neural networks (ANN) or fuzzy sets which can generate approximate solutions of machining process physics.

From human machining to intelligent machining, the machining paradigms can be categorized as: human machining, NC machining, monitored NC machining and intelligent machining in a sequence that the requirement of manual intervention decreases.

There are 3 basic fundamental functional actions for machining processes which are monitoring, decision making and control. Table 1-1 shows the subjects of the actions for different machining paradigms, in which “H” means human and “M” means machine.

Table 1-1 Subjects of the fundamental functional actions for different machining paradigms

Paradigms / Actions	Human Machining	NC Machining	Monitored NC Machining	Intelligent Machining
Monitoring	H	H	M	M
Decision Making	H	H	H	M
Control	H	M	M	M

The definitions and basic requirements for intelligent machining are generally the same but with some slight differences. The following is one representative statement. Intelligent machining means a blend between machining processes (turning, milling, drilling, grinding) and sensors, intelligent monitored by hardware and software applications dedicated to this task[4]. The most important requirements for the intelligent techniques to be applied in machining systems are the abilities for integration of multiple sensor information, for real-time functioning, for effective knowledge representation and for learning or adaptiveness [5].

So an intelligent machining system should also have the ability to learn from the machining processes as a human operator getting more experienced with practice, besides to achieve the 3 basic fundamental functional actions intelligently. Usually this function is accomplished by an expertise system.

Figure 1-1 shows example architecture of an intelligent machining system.

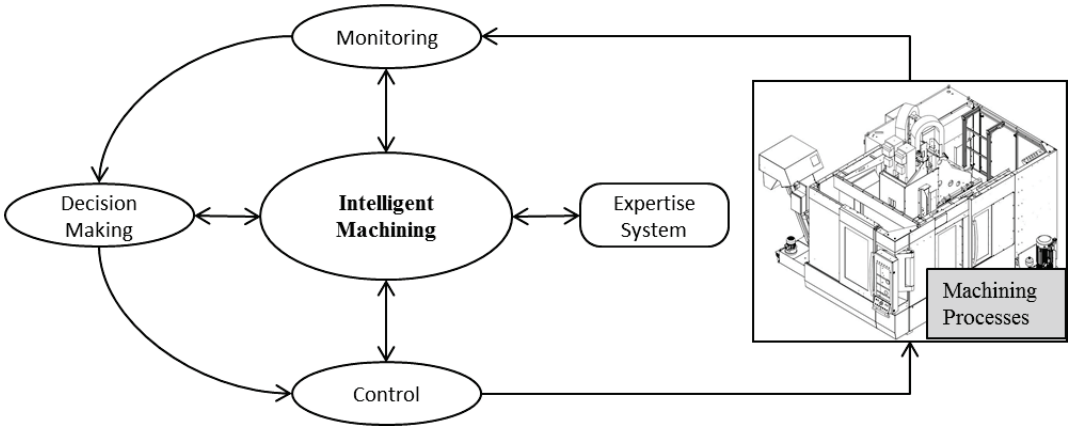

Figure 1-1 A typical intelligent machining system architecture

In this intelligent machining system, sensors that can measure the forces, vibrations, temperature, and sound etc. during machining are installed on the machine tools to monitor the machining performances. Mathematical models that correlate the relationship between the measured sensor signals and the state of machining are formed. The mathematical models are coded into real-time algorithms that monitor the machining process and send commands to CNC for corrective actions, which is the decision making procedure. Then the control unit will take proper actions to adjust the machine tool status in an optimal way. As the machining instances getting more abundant, the expertise system can grow more experienced and thus more intelligent. A knowledge-based optimization for intelligent machining and a machining scenario and its applications to intelligent machining operations based on similar processes were discussed in [6] and [7].

1.1.3 Machining Monitoring

Generally, an intelligent machining system requires the capabilities for automatic monitoring, control and diagnostics of the machining processes. And the success of manufacturing process automation hinges primarily on the effectiveness of process monitoring and control systems [8]. Furthermore, the key issue for an unattended and automated machining system is the development of reliable and robust monitoring systems [9]. In many cases, the decision making is regarded as part of the machining monitoring for its strong relation to and heavy dependency on the machining monitoring.

There have been many excellent reviews reported detailing the range of problems to be addressed in machining monitoring. First, a brief review of these reviews is presented, then based on which the research on machining monitoring is summarized in two dimensions, latitudinal and longitudinal.

Tönshoff et al [10], in 1988, reviewed researches describing conventional and enhanced methods for the monitoring and control of machining processes with a limitation to cutting and grinding machine tools. The monitoring subjects related in this paper were machine, tool, process, cutting condition and workpiece. The components required by monitoring and control systems were sensor, signal conditioning, model and strategy as mentioned in this paper. And the requirements of future monitoring and control systems are worked out as “multi sensor” and “multi model”.

Shiraishi published his great trilogy of critical reviews about in-process measurement, monitoring and control techniques in machining processes in 1988 and 1989, segmented by the different monitoring technologies for tools, workpiece and the cutting process and machine. In the first part [11], measurement technologies for tool wear and tool failure were described and categorized into direct and indirect methods. In the second part [12], technologies for the quality assurance of machined products by in-process measurement was surveyed. The workpiece quality was represented by dimension, profile and surface roughness and the measuring methods were categorized as ones by mechanical, optical, pneumatic, ultrasonic electric and temperature detection ways. In the third part [13], Shiraishi reviewed the in-process techniques for cutting processes and machine tools in which the monitoring of cutting processes covered cutting force, chatter, tool chip, tool/work collision and identification of a cutting process and, the monitoring of machine tools involved driving systems, bearing and rotating systems and temperature control.

Dornfeld [14], in 1994, reviewed research focused on the monitoring of the cutting process for the purpose of determining the state of the operation, in which the states included the condition of cutting tool, the formation and behavior of the chip, and factors affecting the workpiece such as chatter and part dimensions, and the subjects included sensors, signal processing and sensor fusion. And technologies such as acoustic emission, sensor fusion, autoregressive models, and neural networks were specially emphasized.

Dong-Woo, Cho et al [15], in 1999, offered a review of the state of machining process monitoring research in South Korea, which concerned the monitoring of chatter, tool failure and wear, built-up edge, chip, and etc., with technologies such as sensor, signal processing and decision making algorithms.

Liang, Steven Y. et al [8], in 2004, discussed the evolution of machining process monitoring and control technologies and conducted an in-depth review of the state-of-the-art of these technologies over

the then past decade. For machining process monitoring, important attributes in machining such as surface texture, dimensional accuracy, tool condition and chatter detection were included, and sensor fusion was specially discussed.

Abellan-Nebot and Subirón [9], in 2009, reviewed the six key issues involved in the development of intelligent machining systems: (1) sensor systems applied to monitor machining processes, (2) signal processing techniques, (3) sensory features applied in modelling machining processes, (4) the sensory feature selection and extraction methods, (5) the design of experiments required to model a machining operation and (6) artificial intelligence techniques.

Teti et al [16], in 2010, reviewed the past contributions and provided an marvelous, up-to-date, and yet the most comprehensive survey of sensor technologies, signal processing, and decision making strategies for process monitoring. At first, the history of sensorial perception and knowledge acquisition was presented followed by the survey of technologies about sensors and sensor systems for machining. And then, advanced signal processing methods and their application were detailed. And furthermore, the "monitoring scopes" was brought up to present the different goals of machining monitoring. After that, schemes, techniques and paradigms used to develop decision making support systems were reviewed and their relevant applications were presented. The industrial initiatives, experiences and applications of machining monitoring were also exemplified. At last, an outlook on future challenges and trends was made.

Kovač and Mankova [17], in 2011, reviewed the past contributions in machining monitoring areas and provided an up-to-date comprehensive survey of methodology overview, sensor technologies, signal processing, decision making strategies for process monitoring and integrated workpiece quality evaluation, as well as tool wear measuring technique using vision system. Application examples including sensor systems were reported and future challenges and trends in sensor based machining operation monitoring were also presented.

On the basis of the massive amounts of research and conspicuous number of papers, we could summarize the work on machining monitoring in a latitudinal dimension ---- scopes (what is monitored), and in a longitudinal dimension ---- technologies (how to monitor).

The major monitoring scopes can be summarized as follows: machine tool state, tool conditions, workpiece quality, and process conditions, as shown in Table 1-2.

Table 1-2 Major machining monitoring scopes

Machining Monitoring Scopes														
Machine			Tool			Workpiece			Chip			Process		
Chatter	Drive systems	Gears and bearings	Conditon	Failure	Wear	Dimensinal accuracy	Surface texture	Surface integrity	Form	Disposal	Breakage	State	Fault	Variation

The sequence of adopted technologies (addressed as methodologies in some cases) in machining

monitoring can be surmised as: sensor/sensor system, signal processing, feature generation, feature extraction, feature selection and decision making as shown in Figure 1-2.

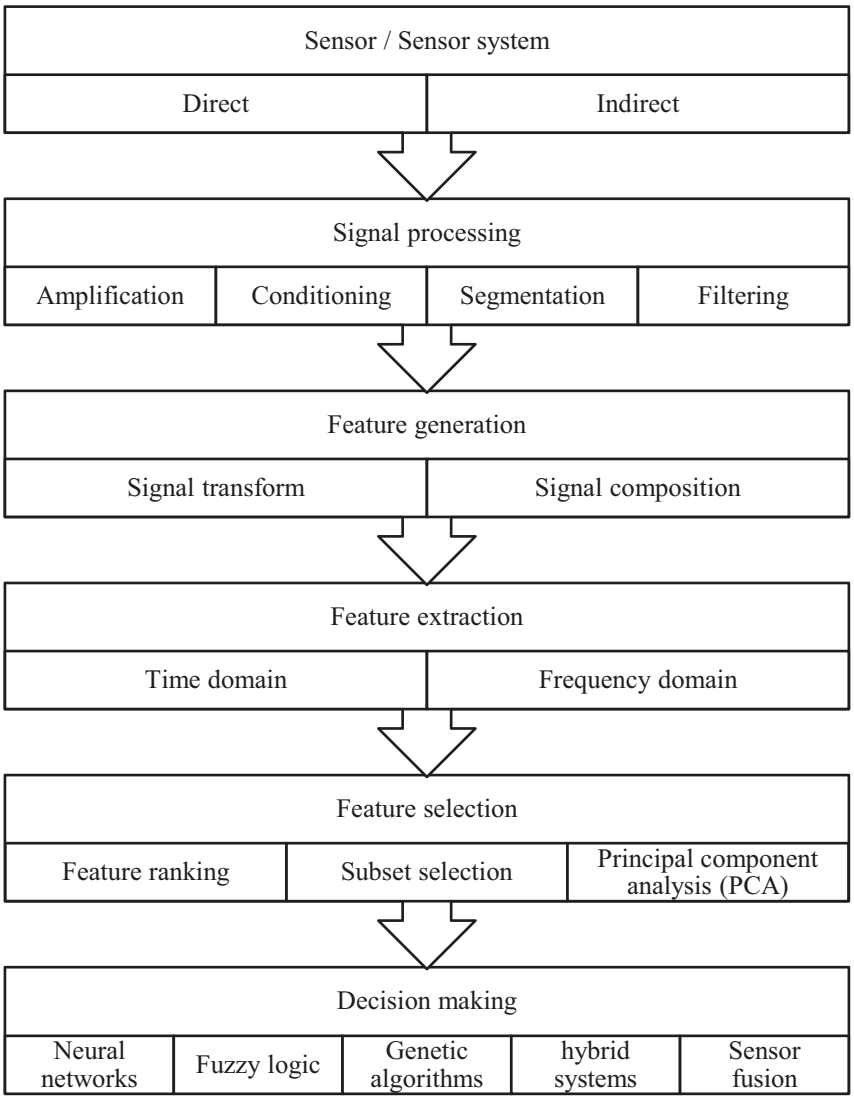

Figure 1-2 Machining monitoring technologies in sequences of common practical situations

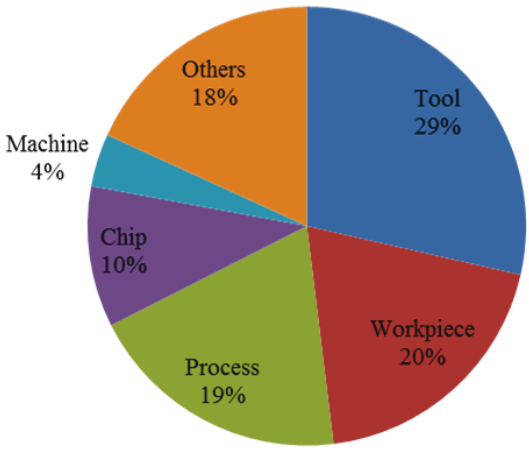

Figure 1-3 Share of different scopes among previous research

Among the monitoring scopes, tool condition monitoring holds a relatively major share, as shown in Figure 1-3, about 29% of the previous research concerned about tool monitoring. This data is calculated from the reviews discussed above with an average measure from 77 papers.

1.1.4 Drilling Process Monitoring

Drilling process monitoring is one of the most important parts among machining monitoring and it has raised quite a lot of interest among researchers and has consequently been studied in a number of research projects by a number of research organizations.

Especially for tool condition monitoring in drilling, a great many methods and technologies has been invented and employed and many papers have been published. In Chapter 2, a state-of-the-art comprehensive literature review on tool condition monitoring in drilling will be presented, which covers all the methodologies and technologies mentioned above.

1.2 New Ideas and the Purpose

1.2.1 New Ideas of Feature Generation and Extraction, Drill Condition Recognition Diagnostic

Ryo et al [18], used dynamic components of the resultant force of principal force and feed force to indicate the adhesion of tool-chip interface and predict the surface finish in turning successfully. Firstly, they transformed the measured 3 dimensional cutting forces from the machine coordinate (defined by feed direction, axis direction, radius direction and tool tip) into tool coordinate (defined by rake face, major flank, minor flank and tool tip) as shown in Figure 1-4.

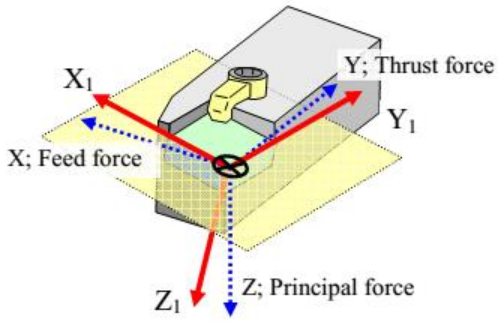

Figure 1-4 Machine coordinate and tool coordinate in turning [18]

Then, on the tool rake face, the tool end edge direction (X_1) and side edge direction (Y_1) form a rectangular coordinate, with the converted feed force (F'_{X1}) and the converted thrust force (F'_{Y1}) match X_1 and Y_1 respectively, as shown in Figure 1-5.

Therefore, the locus of the resultant force of F'_{X1} and F'_{Y1} in the rectangular coordinate will like that shown in Figure 1-5. Then they defined the direction from the origin to the center point of the locus G(x_1,y_1) as the chip flow direction, and the perpendicular direction as the chip width direction. At last the assumption that the dynamic components (such as Δx, Δy, θ) can indicate the chip condition and consequently the workpiece surface quality even tool wear.

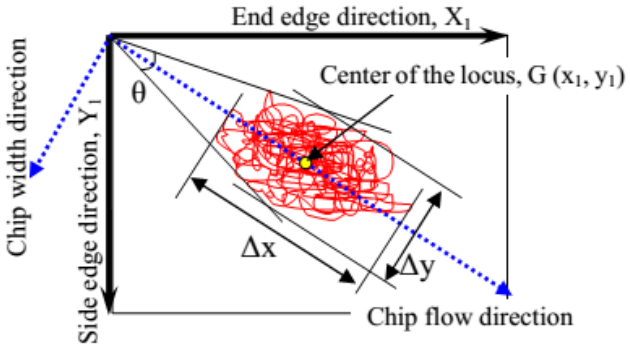

Figure 1-5 Evaluation method of fluctuation in the dynamic components [18]

Ertunc and Loparo [19], exploited percentages of the number of data points inside the reference rectangle to track tool wear with a "phase plane method". The phase plane method is based on plotting the data signals on the Cartesian plane (torque or spindle power on the *x*-axis and thrust or servo power on the *y*-axis) and determining the boundary values of a reference rectangle for the data obtained from drilling a hole with a sharp tool. As illustrated in Figure 1-6, the rectangle is intended to contain the steady state values of the data points, i.e. without the entry and exit parts of the force and power signals.

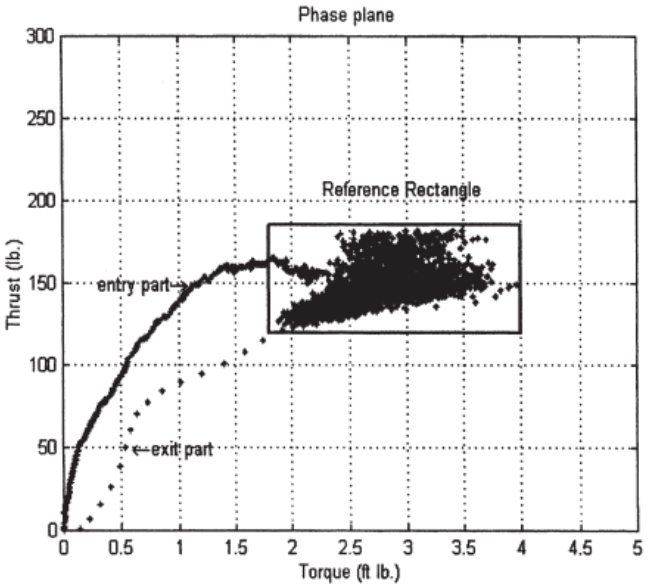

Figure 1-6 Phase plane and reference rectangle [19]

Inspired by these two researches, a conversion method which transforms the thrust force and torque into equivalent thrust force and principal force is proposed. Then a rectangular coordinate is employed with the equivalent thrust force as horizontal axis and the equivalent principal force as vertical axis. And then new features are generated based on the locus of the resultant force. Detailed methodology is illustrated in Chapter 3, 3.3 Feature Generation.

Wu, Ya and Du, R. introduced a new method of feature extraction and feature assessment using a wavelet packet transform for monitoring of machining processes [20]. The procedures of this method proceed in the following sequences.

Step 1, calculate the wavelet packet transform of the training samples;

Step 2, select the principal component feature wavelet packets;

Step 3, reconstruct a time domain signal using the selected feature packets;

Step 4, calculate features from the reconstructed signal (mean values and variance values).

They improved this method and make it functional in [21]. Two more steps were added to make the method able to indicate the tool condition:

Step 5, calculate the alarm thresholds;

Step 6, calculate the average number of threshold crossing points.

The above two works were scarce examples of the few works that used wavelet packets reconstruction to get leaner time domain signals to extract features.

Enlightened by these works, a new method to extract features using wavelet packet reconstruction is originated. However, unlike choosing the principal component packets by the energy of them in the frequency domain, all of the packets are reconstructed severally and from each band the features are extracted. The detailed feature extraction method is provided in Chapter 3, 3.4 Feature Extraction.

Pattern recognition technologies have been widely used to recognize different tool wear states with cutting conditions and process features in machining monitoring, while in reverse, it was hardly used in the recognition of cutting conditions with tool wear states and process features.

A neural network pattern recognition method is applied to classify and recognize the cutting conditions with process features and tool wear states in this work, to provide more information for the diagnostic in an intelligent machining system.

This part of work is detailed in Chapter 5, 5.4 Drilling Conditions Recognition.

1.2.2 Purpose and Configuration of the Research

The use of multiple sensors to increase the capability of any intelligent system has received considerable attention in recent years. One of the several advantages of using multiple sensors is that it can improve ‘observability’. Given the goal of efficient and effective diagnostics and prognostics, sensor fusion is often the way to go. However, one should carefully test the hypothesis that additional sensors might lead to an improvement in the overall performance. In the absence of such a proof, sensor fusion should be avoided for it puts an extra burden on the model developer and could negatively affect the performance (in terms of both computational complexity and effectiveness).

The guiding ideology of this research is to lean the system, to improve the reliability, by using fewer sensors but gaining more information. Therefore, the only sensor used in this research is a dynamometer which can provide the thrust force and torque generated during drilling. Nevertheless, many efforts have been made to obtain information as much as possible, such as new feature generation methods and the application of different feature selection methods.

The purpose of this research is to clarify the new ideas and assumptions, therefore make contributions to the drilling monitoring, drilling automation and then intelligent drilling.

The scope of this drilling monitoring research is restricted; first, the materials are several kinds of metals, then the tools used are common twist drills made of high speed steel (HSS), and the highest cutting speed is around 30 m/s, at last, the drilling depth is less than 3 times of the drill diameter.

1.3 Design of the Research

1.3.1 Applied Technologies

The new methods applied to fulfil sensor signal providing, feature generation, and feature extraction in this research have been briefly introduced above. Other technologies required for machining monitoring applied in this research are as follows.

First, signal segmentation. The whole drilling process of a single workpiece can be divided into 3 stages, according to the cutting edges position. The aimed segmentation is intercepted from the medial stage between entry stage and penetration stage, as illustrated in Chapter 3, 3.2 Signal Segmentation.

For feature selection, both feature ranking and principal component analysis (PCA) are used and compared. They are introduced in details in Chapter 3, 3.5 Feature Selection.

For decision making support system, an artificial neural network (ANN) is applied, and trained with back propagation algorithm, which is introduced in Chapter 3, 3.6 Decision-making support AI System.

1.3.2 Experimental Design

Before using the newly defined features to fulfill a specific monitoring task, the relations between the features and machining conditions should be investigated first. So, the primary task is to verify the existence of, then find out what are the relationships.

The parameters used to represent the drilling conditions are workpiece material, drill diameter, spindle speed, feed rate, and drill wear. A series of experiments with different instances for one condition parameter but same constant value for the others are conducted. Then the relations are obtained and compared to those between conventional features and conditions.

Concretely, workpiece materials used are cast iron, S45C, stainless steel and Titanium alloy; drill bit diameters are 7, 8, 9, 10 mm; spindle speed and feed rate are different for different workpiece materials and drill bit diameters but within a commonly used range.

For collection data to train the neural network, totally 256 different instances are investigated, which covers each 4 different instances for workpiece materials, drill diameter, spindle speed and feed rate. For each instance, drill wear is measured.

1.4 Organization of the Dissertation

The dissertation is organized as follows.

Chapter 1 gives an introduction of this research, which contains background information, new ideas and purpose, what is the research about, how the research is designed, and what methods are applied.

Chapter 2 offers a state-of-the-art literature review on tool condition monitoring in drilling. Different methodologies and technologies applied are categorized and compared.

Chapter 3 discusses all of the methods related in this research. First in the sensor section, modeling of drilling process and forces and the applied piezoelectric dynamometer used are introduced. Then in the signal segmentation section, 3 stages of drilling process are defined and the method of obtaining the aimed segmentation is explained. And in the feature generation section, conversion method of thrust and torque is detailed, based on which the feature generation method is explicated. In the feature

extraction section, wavelet packet transform composition and reconstruction algorithms are interpreted. After that, in feature selection section, theories of feature importance ranking and principal component analysis (PCA) are explained. At last, the back propagation training algorithm for feed forward neural network learning is presented.

Chapter 4 reveals the findings of the relations between the features and drilling conditions. The four drilling condition parameters are workpiece material, drill diameter, spindle speed and feed rate. The effect of drill corner wear to different features is also investigated.

Chapter 5 presents the drill wear prediction and drilling condition recognition by using the methodologies. The feature selection is conducted by the PCA and direct use of BPNN.

Finally, in chapter 6 conclusions are made.

Chapter 2 Literature Review

In this chapter, a summary of the monitoring scopes, methods, signal analysis and diagnostic techniques for drilling process that have been tested and reported in the previous 45 literatures ([19]–[63], sorted by time) are presented. Only indirect monitoring methods such as force, vibration, sound, temperature and current measurements are covered, i.e. direct monitoring methods based on vision, dimensional measurement etc. are not included.

In the first section, monitoring scopes of drilling process are categorized into abnormal detection, wear states monitoring, and life prediction. And beneath them several sub categories are further created. Within these categories, it is known that drill breakage detection, drill wear degree classification and drill wear value estimation are the most attracted research scopes.

In the second section, different signal analysis methods are classified as preprocessing, time domain analysis, frequency analysis and time-frequency analysis. Different feature acquiring methods are also included for different signal analysis methods. Special cases that without any signal analysis are particularly illustrated.

In the last section, a detailed introduction of the usage of different decision-making support systems is presented, and the decision-making support systems are classified into 5 categories: threshold, statistical, fuzzy logic, neural network and Hidden Markov models.

2.1 Monitoring Scopes

Among all of the 45 papers, only one work ([41]) concerned about detecting the state of chip disposal in drilling, other 44 literatures are all about the monitoring of drill bit states, which are categorized into three major types here: abnormal detection, wear states monitoring, and life prediction. Detailed classification is shown in Table 2-1.

Table 2-1 Classification of different monitoring scopes in previous literatures

Categories	Sub categories	Literatures
Abnormal detection	Breakage/fracture	[22], [23], [26], [31], [34], [36], [37], [39], [42], [59]
	Damage/worn	[25], [54]
Wear monitoring	Wear type	[23], [31], [44], [62]
	Wear degree	[19]–[21], [24], [27], [29], [32], [33], [35], [38], [40], [45], [49], [52], [55], [61]
	Wear value	[28], [30], [39], [46]–[48], [50], [51], [53], [57], [58], [60]
Life prediction	Holes to failure	[43]
	Time to failure	[56]
	Tool utilization	[63]

As shown in Table 2-1, 12 literatures are concerned about the abnormal detection in drilling, and they are divided in to 2 sub categories: breakage and damage. In [22], a new eddy current sensor was

used to measure torque at drill shank and analyzed the correlation between signal dynamics in frequency 400-500 Hz to the drill fracture. In [23], different feature of the forces were proved to have relations to different kinds of wears, which were chisel wear, flank wear and corner wear. The AE signal burst was used as a triggering signal to examine force change in [26], and it was described that if the force drops below the preset threshold then it was considered to be a tool failure. Both tool wear and failure were discussed in [31]. Li et al, applied different methods in real-time detection of drill breakage presented in [34], [36], [37], [39]. Normal and abnormal drill state signals were compared in [42], and neural networks were employed to classify them. The input impedance of the spindle motor was used as monitoring signature for detecting drill breakage in [59]. In [25] and [54], sensor signals from normal tools and damaged ones were compared.

For wear monitoring, sub-scopes of wear type, wear degree and wear value are created according to the different focused wear properties.

Locations and patterns of tool wear on a twist drill are shown in Figure 2-1. In [23], chisel wear, flank wear, crater wear and land wear were indicated by 4 levels: very little, small, medium and large or serious, and it was shown that the diversity in drill wear form lead to changes in both the cutting force and the torque. In [31], it was illustrated that chisel wear, corner wear, margin wear and flank wear have different power spectra in the vibration signals of transverse direction. In [44], effects of chisel wear, crater wear, flank wear, edge wear and corner wear to the vibration signals were investigated. And in [62], drills with irregular edge, worn cutting edge, worn outer corner and crater wear on edges were used to conducted experiments and the differences of AE signals were studied.

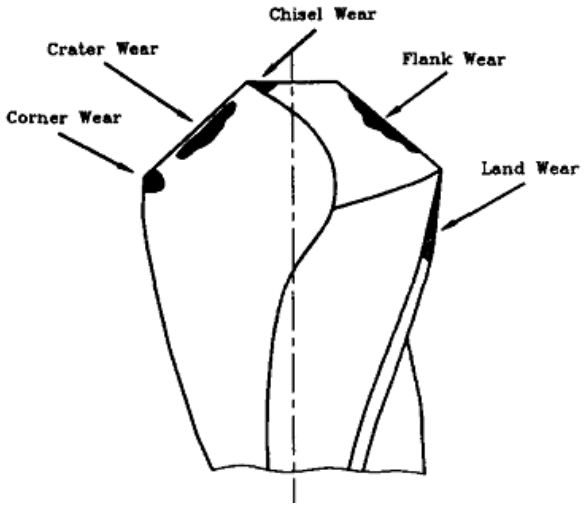

Figure 2-1 Locations and patterns of tool wear on a twist drill [23]

Literatures in wear degree sub category divided the drill wear into several degrees or different states by the wear width, area or drilled holes. Noori-Khajavi and Komanduri in [24], [27], [29], tried to classify the drill wear into 3 states defined by the corner wear area automatically. Wu et al defined the tool conditions in 4 sets, new, initial wear, normal wear and worn, by the average flank wear in [20], [21]. Li et al classified the tool conditions as initial wear, normal wear, acceptable wear, severe wear and failure in [32], [33] and [38], and as initial wear, acceptable wear and severe wear in [35]. Ertunc et al used run number of drill hole to divide the drill status into sharp, workable and dull. In [49], drill

health states were represented by good, medium and worn-out. Progressing tool wear in the total drill life was discussed together with the changes in the different process signals in [52]. Tool status were classified into sharp, workable and dull in [55], same as in [61].

In literatures categorized in wear value, specific drill wear values were presented and the works were dedicated to estimate or predict the wear values. Lin and Ting conducted tests using drills with 0.1, 0.5 and 0.9 mm flank wear to obtain data to train a neural network in [28], [30]. A fuzzy and defuzzy method was applied to estimate the tool wear in [39]. Dependence between tool flank wear and main forces and spindle power was found in [46]. Experiments using drills with various flank wears measured when drilled a certain number of holes were conducted to collect data in [47]. Panda et al obtained 49, 52, 52 and 64 data instances from experiments with different drill wear respectively in [48], [51], [53] and [60]. The effect of drill wear on the electrical power consumption was studied in [50]. The effect of drill flank wear to spindle current was researched in [57] and it was applied to predict drill wear with neural network in [58].

It should be noticed that researches in sub categories wear degree and wear value are different in the major objective. The former ones are aiming at recognized the drill wear status, which always classified into several levels according to the wear degree, usually by using fuzzy logic methods. The later ones are concentrated on estimating or predicting the exact current wear value, constantly by employing neural networks.

Literature [43], [56] and [63] devote to predict the drill life with 3 different approaches, which are by means of achieving the relations between sensor signals and holes to failure, time to failure and utilization percentage respectively.

2.2 Sensors/Sensor systems

In the midst of 45 literatures, only one paper [43] did not use any sensor at all, with the cutting time that the tool has been used (sec) as the only information beside the cutting conditions. Others used one or more sensors without exception. The sensors applied in these researches can be categorized as power/current, force/torque, acoustic emission (AE), vibration and sound. Temperature is also an important condition parameter which can reflect drill wear state. However, it is much more difficult to deploy temperature sensor into a drilling system. So it is rarely used in previous researches and not involved here. The detailed sensor categories and their sub categories are listed in Table 2-2, with corresponding literatures referred.

An eddy current sensor integrated in the drill shank was developed in [22], based on the principal that the strains will change the micro-magnetic properties. Hall current transducers are the most frequently used sensors for monitoring current signals in industrial environment. Electricity carried through a conductor will produce a magnetic field that varies with current, and a Hall sensor can be used to measure the current without interrupting the circuit. As which can be found in Table 2-2, Hall sensors were applied in many works. In advanced CNC machining systems, the current signal data of spindle drivers or feed servo motors can be acquired from the CNC control unit directly, and this method was adopted in [41], [46], [54]. There were some other kinds of sensors that can generate power consumption curves of motors were used in [19], [40], [55], and the output of these sensors was

in terms of horsepower (HP) or Watts. Sensors used in [42] and [50] were not specified. However, it could be speculated that a hall transducer and a power cell were used in the former and later system respectively.

Table 2-2 Categories of sensors deployed in previous literatures

Categories	Sub categories	Literatures
Power/current	Eddy current	[22]
	Hall current transducer	[21], [34]–[37], [39], [57]–[59], [63]
	Machine controller	[41], [46], [54]
	Power cell	[19], [40], [55]
	Unspecified	[42], [50]
Force/torque	Thrust and torque	[19], [23], [24], [27]–[30], [40], [45], [47]–[53], [60]
	Thrust	[25], [26], [55], [61]
	Torque	[59], [62]
	Others	[24], [27], [29]
Acoustic emission	Breakage detection	[26], [34], [37]
	Wear monitoring	[32], [33], [52], [62], [63]
Vibration	On spindle	[20], [23], [56]
	On workpiece	[31], [38], [44], [60], [61]
Sound	Microphone	[61]

The force/torque sensors used in the previous works were all dynamometers based on the piezoelectric effect. Most of them monitored thrust force and torque simultaneously, but some of them monitored only thrust or only torque with other sensors, which were listed as follows. Thrust and vibration were monitored in [25]. Thrust and AE were monitored in [26]. Thrust and motor load were monitored in [55]. Thrust, sound and vibration were monitored in [61]. Torque and spindle current were monitored in [59]. Torque and AE were monitored in [62]. Strains in two orthogonal directions to the drill axis were monitored together with thrust and torque in [24], [27] and [29].

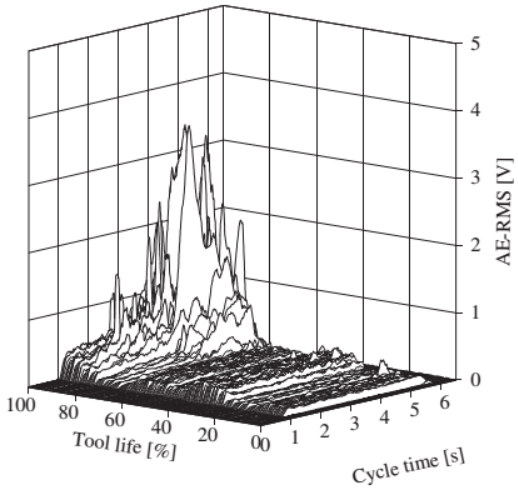

Figure 2-2 AE-RMS over the life of a 1.5 mm diameter twist drill [63]

Acoustic emission (AE) is generated during a variety of metal cutting processes. It is generated anytime during normal machining. However, there is a distinct difference among the amplitudes of AE signals obtained during fracture, chipping and normal machining. This fact made it possible to monitor tool fracture with an AE sensor, like that had been reached in [26], [34] and [37]. Moreover, the major advantage of using AE to monitor the tool condition is the frequency range of AE signal is much higher than the frequency of machine vibrations and environmental noises. The AE-RMS signal showed a very significant increase in magnitude during the final 20–40% of drill's tool life, as shown in Figure 2-2 from [63]. Thus the AE signals were also used to monitor the drill wear.

All vibration sensors used in the literatures were piezoelectric sensors but one in [25], a laser vibrometer was applied to adjust the drill centering and vertical position. Since the laser vibrometer was not used to offer signal, it was not included in the vibration category. The installation position of a vibration sensor could be on the spindle box or on the workpiece or fixture, as listed in Table 2-2. However, the vibration direction also could be axial (feed direction) or transverse (radial direction). Unfortunately, only in [31] and [60], both two directions vibrations were explicitly accounted been monitored at the same time. It could be conjectured from the experimental set up illustrations or text descriptions that axial vibrations were monitored in [23], [38] and transverse one were monitored in [20], [44], [56]. Neither illustration or text description mentioned about the vibration direction in [61].

Besides the force and vibration measurements, machine sound data were also collected through a microphone placed in the direct vicinity of the workpiece in [61].

Constitutionally, the sound and AE are both vibrations. However, they are utilized to differentiate the vibrations by the frequency and amplitude properties. The duration of the AE bursts due to tool fracture is approximately 1 msec. So, at least 1 kHz sampling frequency is needed [26]. For example, a 50 Hz high pass and a 1 MHz filter were applied to pre-process the AE signal in [34]. The sound frequency is between 20 Hz to 20 KHz usually and vibration frequency is commonly up to several hundred KHz.

2.3 Signal Analysis

On account of no sensor application in the work, [43] is not discussed in this section.

In the other 44 literatures, signal processing and feature acquiring procedures were not exactly in the order or contained all that listed in Figure 1-2. And there are always very strong causal associations between the signal processing methods and the feature acquiring approaches. So, these two major methodologies are synthetically talked over in this section.

First, the analog and digital preprocessing methods for sensor signals are discussed.

Then, since the feature selection process can be rather involved and might utilize time-domain analysis methods, frequency domain analysis methods (such as FFT and STFT), as well as joint time-frequency-domain analysis methods (such as wavelets), the signal analysis are particularly presented in time domain, frequency domain and time-frequency domain.

2.3.1 Signal Pre-processing

The analog signal from the sensor usually cannot be connected directly to the A/D converter but

needs pre-processing by a conditioner specific to the sensor (piezotron coupler, charge amplifier, etc.). For example, a typical procedure of analog AE signal pre-processing follows the pattern schematically shown in Figure 2-3 [16].

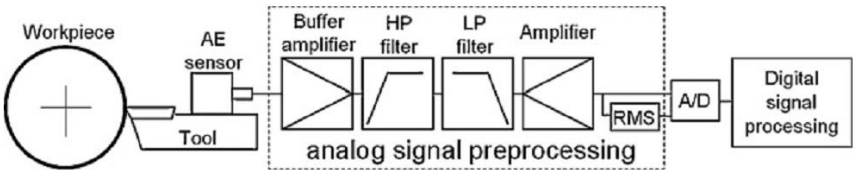

Figure 2-3 Typical analog preprocessing for AE signals [16]

In some other cases, the filtering is completed after the signal has been digitalized.

It is also known that generally a full drill process can be divided into 3 stages according to the drill tip position relative to the workpiece surface. And furthermore, the process parameters (such as thrust and torque) are also affected by the drill depth. Hence, signal segmentation is frequently conducted for signal analysis in drilling.

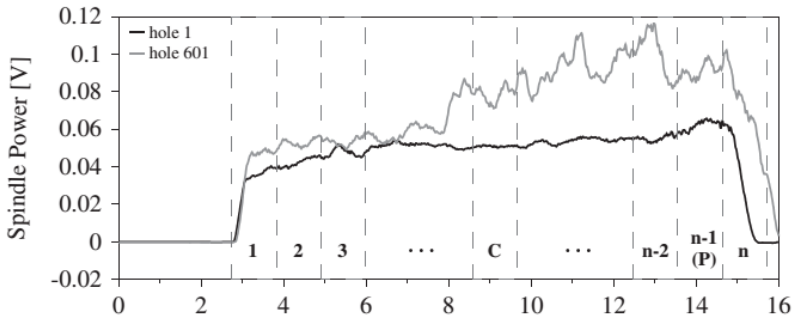

Figure 2-4 Segmentation of the spindle power signal into ‘‘*n*’’ equidistant sections [63]

An instance is shown as Figure 2-4 [63]. It depicts the spindle power recorded for two drilling cycles of a 1.5mm drill. The black curve was recorded from the 1st hole, the grey curve was from the 601-th hole. It can be observed from the illustration that during the second half of the drilling cycle, the rise of the spindle power becomes much more significant.

In recent common cases, analog preprocessing is in charge of the signal amplification, while digital preprocessing accomplishes other works like filtering, segmentation and conditioning etc., on the strength of much more powerful computers.

2.3.2 Time Domain Analysis

Time domain analysis, also termed as statistical analysis, can generate signal features such as average value, max value, RMS, variance etc. There are also some particularly defined features or signatures been used in the literatures. Different signal features obtained by time domain analysis are listed in Table 2-3.

Since average, max, RMS, standard deviation and variance are frequently used; only kurtosis, skewness and energy are introduced here. And for the particularly defined features, the impedance and transient time are chosen to be explained here.

Table 2-3 Different time domain signal features in previous literatures

General features	Average/mean	[23], [26]–[28], [30], [35], [40], [44], [47], [48], [50], [51], [53], [55], [59]–[61], [63]
	Max	[26], [59], [61]
	RMS	[57], [58], [61]
	Standard deviation/Variance	[27], [44], [59], [61], [63]
	Kurtosis	[31], [44]
	Skewness	[44]
	Energy	[23]
Particularly defined features	Impedance	[59]
	MAMP, MVMP	[62]
	Mechanistic approach features	[45]
	Phase plane method features	[19], [45]
	Transient time	[45]

Kurtosis and skewness of the vibration signal data monitored by an accelerometer were used to indicate different drill wear types in literature [44], as depicted in Figure 2-5.

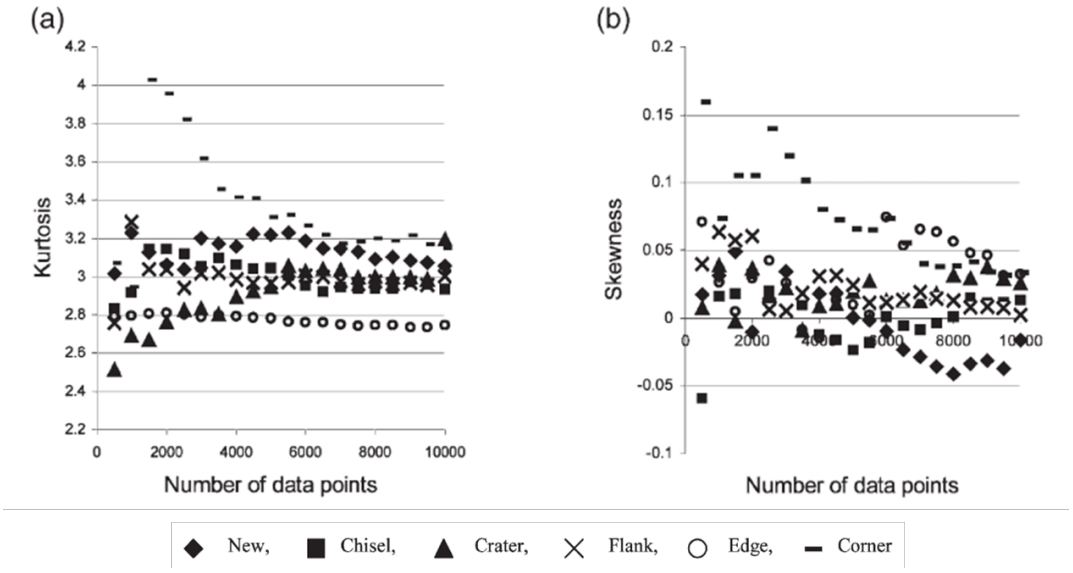

Figure 2-5 The influence of drill wear to kurtosis and skewness of vibration signal data points [44]

The information involved in the total energy of the signal has very limited connection with time. To avoid this weakness, a parameter called signal "short- time energy" which means the energy of a signal involved in a particular time period was defined in [23].

The investigation results on the effectiveness of using the input “impedance” of the spindle motor as monitoring signature for detecting drill breakage in micro-drilling was reported in [59]. Figure 2-6 illustrates the method conceptually.

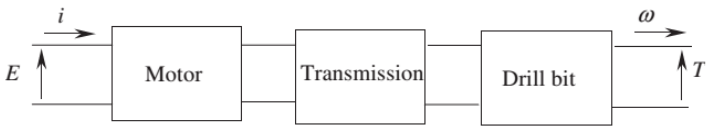

Figure 2-6 The impedance conceptual model of a drilling system [59]

As accounted in the paper, this input impedance is an inherent system property of these subsystems together and can be evaluated via literally dividing the effect variable (voltage or force) by the flow variable (current or velocity) at the point. Since this change of drilling condition is time varying, the electrical impedance also varies accordingly in time domain. To calculate the impedance, the voltage and current signals were acquired by Hall Effect probes, and dynamometer was installed to measure the drilling torque.

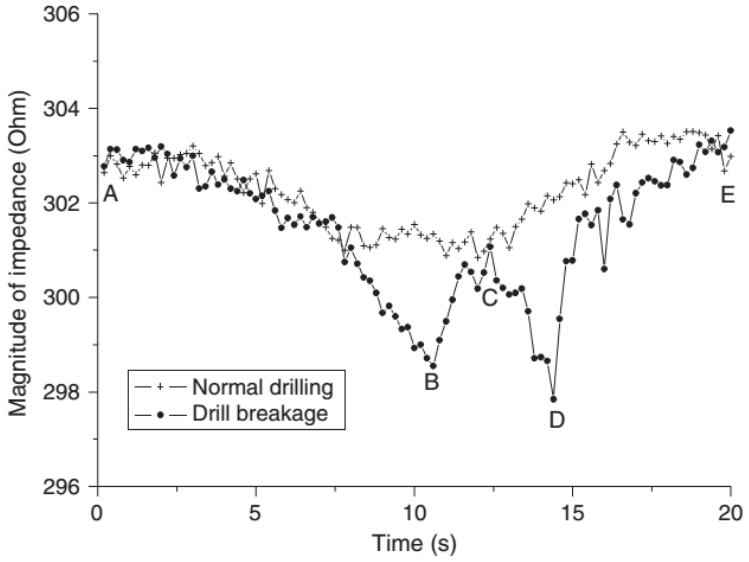

Figure 2-7 Motor impedance signals of drilling processes using normal and breakage drills [59]

The motor impedance signals for normal drilling and drilling with a drill breakage are shown in Figure 2-7. Differences can be found in this two scatter lines.

The transient time was used as a parameter to monitor drill corner wear in [45]. As corner wear progresses, the point angle of drill reduces and the cutting lips lengthen because the tool loses its sharp edge. Consequently, the transient time, that is the time for entry of the tool into the workpiece, increases.

As it is illustrated in Figure 2-8, the entry stage periods (defined as transient time) of the torque signals for a sharp drill and a corner worn drill is different. Thus, a threshold value of the transient time is set to stop the drilling process.

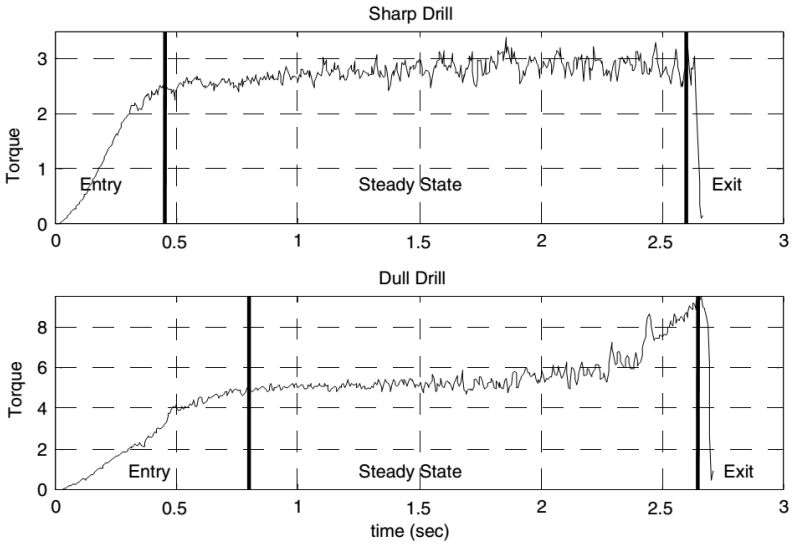

Figure 2-8 Torque signals for a sharp drill and a corner worn drill [45]

2.3.3 Frequency Domain analysis

Frequency domain analysis is used to inspect the dynamic components contained in the signals which cannot be detected by time domain analysis. Fast Fourier transform is the most frequently used method to get the spectra form of a signal.

Drill fracture and the dynamic components in torque within a certain frequency band was found related in [22].

Power spectral density (PSD) was used as the frequency domain indicator in [24], [27] and [29].

In physical world, the signal is usually a wave, such as an electromagnetic wave, random vibration, or an acoustic wave. The power spectral density (PSD) of the signal, when multiplied by the appropriate factor, describes the power contributed to the wave, by a frequency, per unit frequency. Power spectral density is commonly expressed in watts per hertz (W/Hz). Detailed continuous and discrete PSD calculation algorithms can be found in [27].

It was found in these papers that the change in PSD with respect to drill wear is very clear. 9 groups of data distributed into 3 sets, according to 3 different corner wear, were obtained by experiments. There was found much difference between the PSD curves of thrust, torque, axial vibration and transverse vibration signals for drilling processes with drills of different corner wears. And for those have the same corner wear, all the PSD plots are coincident.

Moreover, it was found that the change of area under the PSD plots was considered instead of the power at one frequency, because integration decreases the error. The change in the area under the PSD plots of all four sensor signals was plotted against the total corner wear, and it was observed that the change of area under the PSD plots of three sensor signals, namely, thrust, torque, and strain in the *X*-direction (transverse direction), showed good correlation with drill wear.

Cepstrum analysis is used to identify a series of harmonics or side bands in the power spectrum and to estimate their relative strength in [31]. The cepstrum is defined as the inverse Fourier transform of the logarithm of the power spectrum [64].

The cepstra calculated varied considerably from one hole to another. However, the cepstra

calculated for the signal measured during drill breakage showed consistently a peak at a quefrency related to the time of one spindle revolution. This phenomenon was observed in all drilling tests performed for detecting breakage in this investigation. Thus, monitoring the existence of a peak in the cepstra corresponding to the time of one spindle revolution can be used as an index for detecting drill breakage.

Moreover, cepstra were calculated for each of the vibration signals measured in both *Y* (transverse) and *Z* (axial) directions, namely *Py* and *Pz*, respectively. Then the ratio between the cepstra in the *Y* and *Z* directions was then defined as cepstra ratio (*Py/Pz*). And it was found that the cepstra ratio can indicate the drill breakage effectively.

Similar but different from [27], the area under the vibration spectrum was considered as a monitoring feature for drill wear, to avoid the effects of random variations in peak amplitudes (rather than using the variation of the peak at any one frequency). And a linear relation was found between the number of drilled holes to the area under the vibration spectrum.

2.3.4 Time-Frequency Domain analysis

The spectral analysis method is the most commonly used signal processing technique in tool breakage detection. But spectral analysis has a good resolution only in the frequency domain but a very bad resolution in the time domain. That is, some signal information in time domain is lost in the spectral analysis process. Even though the sensor signals detected during machining are essentially unstable, the FFT averages the frequency composition over the duration of the signal with fixed resolution of the entire frequency spectrum. Therefore, a time-frequency method which can give both good resolutions in time domain and frequency domain is highly demanded.

Wavelet transform (WT) can extract more information in the time domain at different frequency bands, and it has been widely used in signal analysis of machining monitoring. The WT decomposes a signal through the wavelet scale function and scaled and shifted versions of the mother wavelet. Practically, it can be reduced to filtering the signal by high-pass and low-pass filters derived from the wavelet and the scaling function. The discrete wavelet transform (DWT) decomposes the signal into the scaling coefficients (approximations A) and the wavelet coefficients (details D) by convolution of the signal and impulse response of the low-pass and high-pass filters.

Another type of WT is the wavelet packet transform (WPT) where both approximations and details are decomposed, generating many more frequency bands. This provides more opportunities to find useful signal features. On the other hand, for n levels of decomposition, the DWT produces 2^n sets of coefficients.

Both WT and WPT were frequently used in the previous works. Table 2-4 gives a summary of the application of WP and WPT signal analysis. Meanwhile, the feature acquiring methods for them both are separately listed with literatures applied.

Table 2-4 Summary of the application of WT and WPT in previous literatures

Categories	Sub categories	Literatures
Wavelet Transform (WT)	RMS	[32], [38]
	Envelope detection	[34], [37]
	Encoding	[25]
	Waveform	[36], [39], [54]
Wavelet Packet Transform (WPT)	RMS	[33], [58]
	Energy	[20], [21]
	Statistical	[20]
	Fisher	[56]
	Reconstruction	[20], [21]

There is another time-frequency signal analysis method used in [42], called as the harmonic wavelet analysis, is also been discussed.

2.3.4.1 Wavelet Transform (WT)

The RMS (root mean square) values of the WT decomposed results for the AE and vibration signals were used as indices to estimate the tool state in [32] and [38].

The envelope detection method was used to determine if a maximum value of WP decomposed results is from tool breakage or noise in [34] and [37].

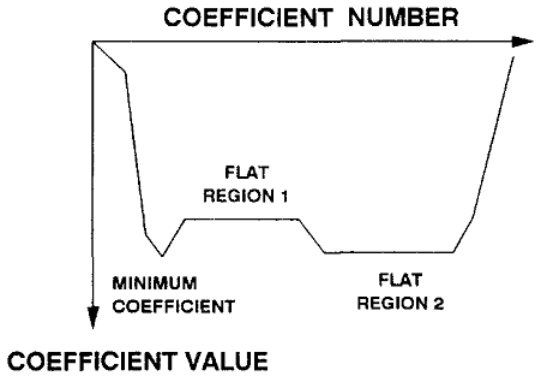

Figure 2-9 The 2 flat regions of WT coefficients [25]

6 encoded parameters of the wavelet coefficients (E_i; i=1…6) were used as inputs to the neural network in [25]. These 6 parameters were specially defined according to the particular 2 flat regions found in the WT coefficient plot as shown in Figure 2-9.

For flat region 1, E_1, E_2 were defined as Equation 2-1 and Equation 2-2.

$$E_1 = C_{mean}/C_{min} \quad \text{Equation 2-1}$$

$$E_2 = C_{var} \quad \text{Equation 2-2}$$

C_{mean}, C_{min} and C_{var} represent the average, minimum and variance of the coefficient values with in flat region 1.

And another parameter E_3 was defined as the number of coefficient values which were out of the

band defined as$(C_{mean}, 2C_{mean} - C_{mean})$.

For flat region 2, E_4, E_5, E_6 were defined in a similar way.

In some papers, no features were acquired from the WT coefficients but only the waveform of the coefficients was investigated to offer a correlation to the drilling process. For instance, in [36] and [39], it was found that an obvious peak arise in the WP coefficients when drill breakage happens.

In [54], the waveform of the WT coefficients was also investigated but with responding asymmetry values been defined. Figure 2-10 shows comparison of two consecutive pulses from WT coefficients of spindle power current signals with a normal drill in subfigure (a) and a damaged drill in subfigure (b).

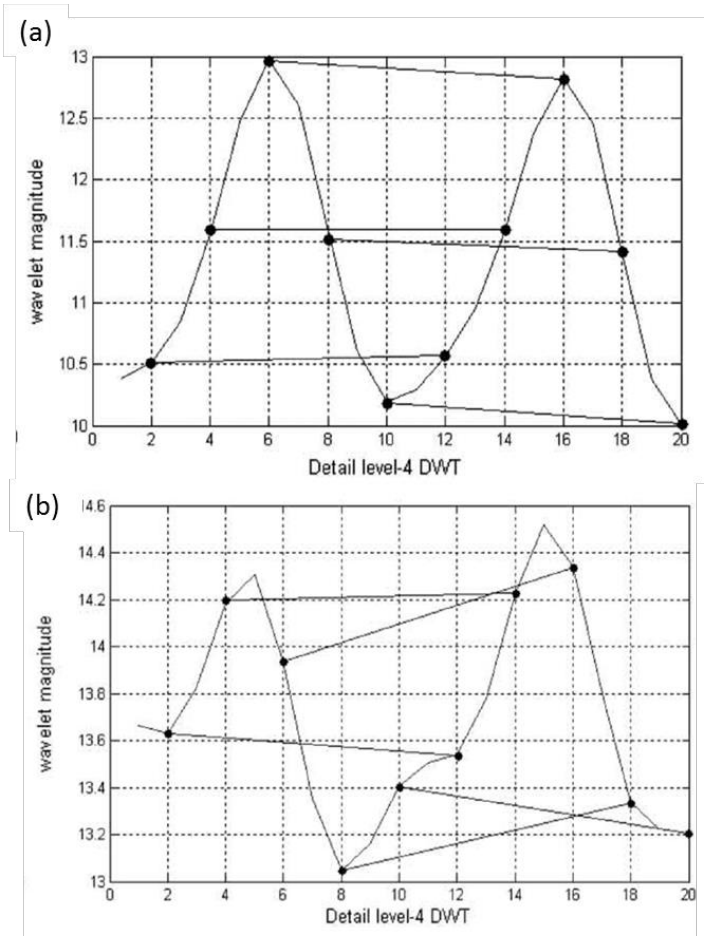

Figure 2-10 Waveforms of WT coefficients of spindle current signals for a normal and a damaged drill [54]

The asymmetry value A is defined as Equation 2-3.

$$A = \sum_{i=1}^{M}(C_{i+M} - C_i)^2 \qquad \text{Equation 2-3}$$

M means the number of values comprising a single pulse, and C_i means the WT coefficient values.

2.3.4.2 Wavelet Packet Transform (WPT)

The RMS values of WPT coefficients were calculated and applied as the features in [33] and [58]. And the feature values of different frequency bands were compared by the correlations to the drill wear. Then the features were selected according to the comparison results.

The energy of the wavelet packets were used to rearrange the sequences of the different decomposed packets, on the basis of the higher the energy the more the importance.

Since the WPT coefficient values can also be treated statistically, the mean, variance, kurtosis, etc. values may also be used as features. The peak-to-valley value was used as a certain index to describe the characteristic of the data in [20].

It was pointed out in [56] that direct manipulation of a whole set of node energies is ineffective since the space normally has very high dimensionality, and the existence of undesired components makes the classification unnecessarily difficult. Thus, the Fisher's criterion was applied to select the feature subsets. The criterion function is defined as Equation 2-4.

$$J_{fk}(i,j) = \frac{|u_{i,fk} - u_{j,fk}|^2}{\sigma_{i,fk}^2 + \sigma_{j,fk}^2} \qquad \text{Equation 2-4}$$

where $u_{i,fk}$, $u_{j,fk}$ are the mean values of the *k-th* feature *fk*, for class *i* and *j*, and $\sigma_{i,fk}^2$, $\sigma_{j,fk}^2$ are the variance of the *k-th* feature *fk* for class *i* and *j*.

The WPT decomposition and reconstruction were used in [20] and [21]. The feature wavelet packets were used to reconstruct the principal component of the signal, which is done by choosing the feature packets and setting the other packets to zero. The details of the WPT decomposition and reconstruction are presented in Chapter 3.

The feature packets were chosen first as the first several packets by ranking with the packet energy. Then the effectiveness of the feature packets was evaluated based on the reconstructed signal. The assessment criterion was based on four indices:

(1) the cross-correlation between the reconstructed signal and the original signal,

(2) the cross-coherence between the reconstructed signal and the original signal,

(3) the correlation of the difference between the reconstructed signal and the original signal,

(4) the power spectrum of the difference between the reconstructed signal and the original signal.

2.3.4.3 Harmonic Wavelet Analysis

The harmonic wavelet analysis method applied in [44] and the features generated method were calculated as follows:

1. The input *f(t)* time history of the vibration signal was represented in segments of 4096 data points.

2. The DFFT (discrete FFT) algorithm was used to give the 2048 Fourier Coefficients *F(w)*.

3. The Fourier Coefficients were multiplied with a frequency form wavelet base function $\emptyset_i^*$ and then C_i was obtained: $C_i = F_i \times \emptyset_i^*$, *i=0-2047.*

4. The IFFT (inverse FFT) of the generated series was computed to obtain $c_r(t)$, *r* = 0-2047.

5. 16 averaged wavelet coefficients were extracted from the result of step 4 by grouping and averaging each adjacent 128 coefficients $c_r(t)$.

The 16 averaged harmonic wavelet coefficients (HWCs) form a feature vector which served as an input pattern to the neural network.

2.3.5 Without Signal Analysis

The performance of any monitoring system can greatly depend on the quality of the features extracted during the feature extraction process. While 'sophisticated' features could make the task of the inference model easy, this often comes at the expense of time (for feature acquiring), computation-

complexity, and cost. Thus, in some papers, the raw signal data (not more than simple preprocessing) was used directly, without a single feature been extracted.

In [42], an unsupervised learning neural network was used to extract the "principal components" from the raw spindle power signal without prior definition of what these principal components are.

In [49], hidden Markov models (HMMs) were employed for performing diagnostics and prognostics in drilling which almost eliminated the need for any feature extraction. The HMM training data sets were sampled from thrust force and torque from the time instant the drill bit penetrated the workpiece to the time instant the drill tip protruded out the other side of the workpiece. The data was collected until the drill bits reached a state of total physical failure.

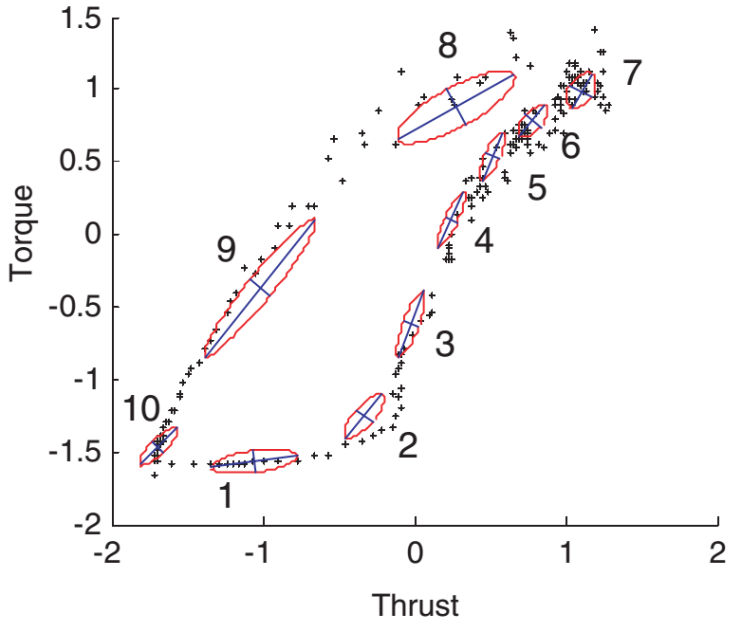

Figure 2-11 A 10 state HMM superimposed onto the joint scatter plot [49]

It is illustrated in Figure 2-11 that a particular HMM had 10 hidden states, hence the 10 ellipsoids. The location of the ellipse corresponded to the mean vector of the observable bivariate Gaussian density and the major and minor axes represented the Eigen vectors of the covariance matrix.

2.4 Decision-making Support Systems

There were many kinds of decision-making support systems applied in the previous works. However, the clustering, mapping and decision surface were seldom used and found not so efficient in [24] and [29], thus they were not of discussion here.

The rest of them can be summarized as those listed in Table 2-5.

Table 2-5 Different decision-making support systems applied in previous literatures

Categories	Sub categories		Literatures
Threshold	N/A		[19]–[21], [23], [26], [31], [34], [54], [55]
Statistical	Mechanical models		[28], [35], [47]
	Regression		[56], [57]
Fuzzy logic	N/A		[33], [39], [61]
Neural network	Supervised learning	BPNN	[30], [41], [44], [47], [48], [51], [57], [58], [60]
		RBNN	[60]
		FNN	[32], [38]
		FBPNN	[53]
		OLL	[43]
	Unsupervised learning		[25], [42]
Hidden Markov model	N/A		[19], [40], [45], [49], [55]

2.4.1 Threshold

The threshold method was proved to be very effective on basis of the well-selected features and well-grounded thresholds, thus it was widely used by researchers.

A series of discriminants were applied to judge the drill wear states in [23]. The sampled thrust force and torque were checked through these criteria, and if a certain discriminant was satisfied, the corresponding drill state, such as worn chisel, worn flank, broken edge, etc., was detected.

The AE signal burst was used as a triggering signal to examine force change in [26], and if the force drops below the preset threshold, it was considered to be a tool failure.

The instantaneous Ratio of the Absolute Mean Value (RAMV) and the kurtosis (K) of the spindle current and vibration signals were calculated and tested by threshold values to trigger the examinations of further steps in [31]. The further examinations, such as the cepstra ratio and the area under the power spectrum were also based on threshold method.

A judgment procedure was given in [20] to detected the states of drill wear. The features applied were the particularly defined peak-to-valley index and crest factor index of the reconstructed signals of vibration signals by WPT.

The envelope detection method applied in [34] can also be regarded as a threshold method, for the envelope can also be taken as a threshold band.

A decision fusion center algorithm (DFCA) was proposed in [19], which combines the outputs of other individual methods (the HMM methods, the phase plane method, the transient time method and torque model method) to make a global decision about the wear status of the drill. The produced global decision variable (U) was then finally used to make the decision of drill status.

An alarm threshold band was established with the processing of spindle current signals in [21], then the feature values for the monitored moment was calculated and the decision making is done in two steps: (1) check the threshold crossing, (2) calculate the number of threshold crossing to determine whether an alarm shall be given.

The threshold method for the defined asymmetry value in [54] was accounted to be a possible way to make the drill change decision.

Similar to [19], a DFCA and its threshold were used to determine the wear status of the tool during the drilling operations.

2.4.2 Statistical

There were two kinds of statistical methods for decision-making as shown in Table 2-5. One is mechanical model approach and the other is regression. They are both based on the regression theory but with the difference that the former one was built up by mechanical mechanisms and with machining parameters while the later one was a statistical fitting.

Two models were employed in [28], which expressed thrust force and torque as functions of the feed rate, drill diameter and the flank wear. The spindle speed was ignored because the effect of cutting speed on the cutting force signals is relatively insignificant in the cases studied.

Three similar yet different in the coefficients equations were established for both the spindle motor current and the feed motor current in [35]. The difference between them was the drill wear, which was 0.2mm, 0.5mm and 0.8mm in the three equations respectively. Thus, these equations can be used to classify the drill wear state.

The tool wear was directly expressed by an equation with thrust force, torque, spindle speed and torque in [47].

A regression model of the spindle motor current was presented in [57], which related the spindle speed, feed-rate, drill diameter and average drill flank wear.

In [56], prediction of degraded performance is accomplished using Auto-regressive Moving Average (ARMA) model by the degradation trend from the logistic regression module.

2.4.3 Fuzzy logic

A fuzzy clustering method (FCM) was applied in [33], which successfully classified the tool state by the RMS values of the selected wavelet packets of decomposed AE signals.

Fuzzy classification and “defuzzification” were used in [39], to estimate the drill wear. The defuzzification is a mapping from a space of fuzzy values into that of the crispy universe, that is, the drill wear value can be estimated with the fuzzy coefficients.

In [61], a tool wear condition monitoring technique based on a two-stage fuzzy logic scheme was presented. In the first stage, statistical parameters derived from thrust force, machine sound and vibration signals were used as inputs to fuzzy process; and the crisp output values of this process were then taken as the input parameters of the second stage. In the second stage, a Takagi–Sugeno fuzzy model was applied to generate an output, which was conclusively taken into a threshold function to assess the condition of the tool.

2.4.4 Neural network

There are three major learning paradigms for a neural network (NN), each corresponding to a particular abstract learning task. These are supervised learning, unsupervised learning and reinforcement learning. The former two were used in the previous literatures and discussed here.

Generally, a supervised learning NN is used for tasks like pattern recognition (also known as classification) and regression (also known as function approximation), while an unsupervised learning NN is used for tasks such as clustering, the estimation of statistical distributions, compression etc. Specifically in tool condition monitoring, the supervised learning NN is usually used for the estimations or prediction of the drill wear value while the unsupervised learning NN is commonly used for the classification of the different tool states. However, exceptions did exist in the previous works.

2.4.4.1 Supervised learning neural network

The back propagation algorithm was the most popular one in training the neural networks among supervised learning paradigms.

The drill wear value was prediction in [30], [47], [48], [51], [57], [58], [60], with back propagation neural networks (BPNN).

In [41], the output of the back propagation neural networks was used to indicate the chip disposal state, where a value below 0.5 mean a good state while a value over 0.5 indicate a bad state.

In [44], the outputs of the back propagation neural networks were a vector of values which represented the possibility of being a certain kind of drill wear state. There were 6 values in the vector, and they represented the following states: sharp (new), chisel wear, crater wear, flank wear, edge fracture and corner wear.

Hybrid model of the neural network and fuzzy logic called a fuzzy neural network (FNN) was applied in [32] and [38], which provided an output vector capable to indicate the tool states.

There was another kind of combination of neural network and fuzzy logic which was applied in [53] called fuzzy back-propagation neural network (FBPNN). The fuzzy neuron is the basic element of a fuzzy back-propagation neural network. In the fuzzy neuron, both the input vector and the weight vector are represented by a triangular LR-type fuzzy number.

The radial basis function network (RBFN) and back propagation neural network (BPNN) were compared in [60] and it was found that BPNN predicted the wear more accurately compared to RBFN but RBFN learnt the pattern much faster.

A tailor-made feed forward network was proposed to predict tool life in terms of the number of holes to failure in [43], which was carried out by adopting the optimization layer by layer (OLL) method.

2.4.4.2 Unsupervised learning neural network

Adaptive resonance theory (ART2)-type neural networks (ART2NN) were applied in [25], to detect severe tool damage just before complete tip breakage occurs. Unsupervised ART2-type neural networks can monitor a signal based on previous experience and can update itself automatically while it is monitoring the signals. It works like this: when an ART2NN receives an input pattern, this pattern is compared to already known patterns, if the input pattern is matched with a known pattern in memory, the weights of the model are changed to update the category, and if the new pattern cannot be classified in a known category, it is coded and classified as a new category.

A feature extractor based on unsupervised learning neural network was used in [42]. The details were not provided in the paper but it was described that the “feature extractor” can extract the

"principal components" from the raw signal without prior definition of what those principal components were.

2.4.5 Hidden Markov model

Hidden Markov models (HMMs) can be considered as a method of modeling pattern classes that consist of time-series data and then comparing patterns to recognize or classify new data.

The force and power signals in the form of time series were used to build the HMM and to monitor the machining process in [40]. Two methods were proposed using HMMs for tool wear condition monitoring in drilling operations: the bargraph method and the multiple modeling method. These two methods were combined with a decision fusion center algorithm (DFCA) in [19]. Similar work was done in [45] and [55].

In [49], a method with employing hidden Markov models (HMMs) for carrying out both diagnostic and prognostic activities for metal cutting tools was presented. The method employed HMMs for modelling sensor signals emanating from the machining process and in turn, it was able to identify the health state of the cutting tool and even estimate the remaining useful life (RUL) of the cutting tool.

Chapter 3 Methodologies

In this chapter, the methodologies applied in this research are illustrated in detail. First the sensor system and the general experimental set up is introduced in section 1. Then the signal segmentation method is demonstrated in section 2. And the section 3 offers a detailed feature generation procedure. The modeling of drilling process is referred to provide a theoretical explanation. In section 4, the feature extraction method using wavelet packet transform is presented and then the feature selection methods are explicated in section 5. In the last section 6, the artificial neural network (ANN) model based regression and recognition approaches are specified.

3.1 Sensor system and General Setup

The sensor system comprises a Kistler 9365B piezoelectric dynamometer, a Kistler 5073 charge amplifier (Frequency response 20 kHz) and a Yokogawa DL750 scope recorder.

The piezoelectric dynamometer generates continuous thrust and torque signals in a form of electric charge. The charge amplifier then amplifies and converts the charge signals into voltage signals. After that the voltage signals are sampled and recorded by the scope recorder as digital data.

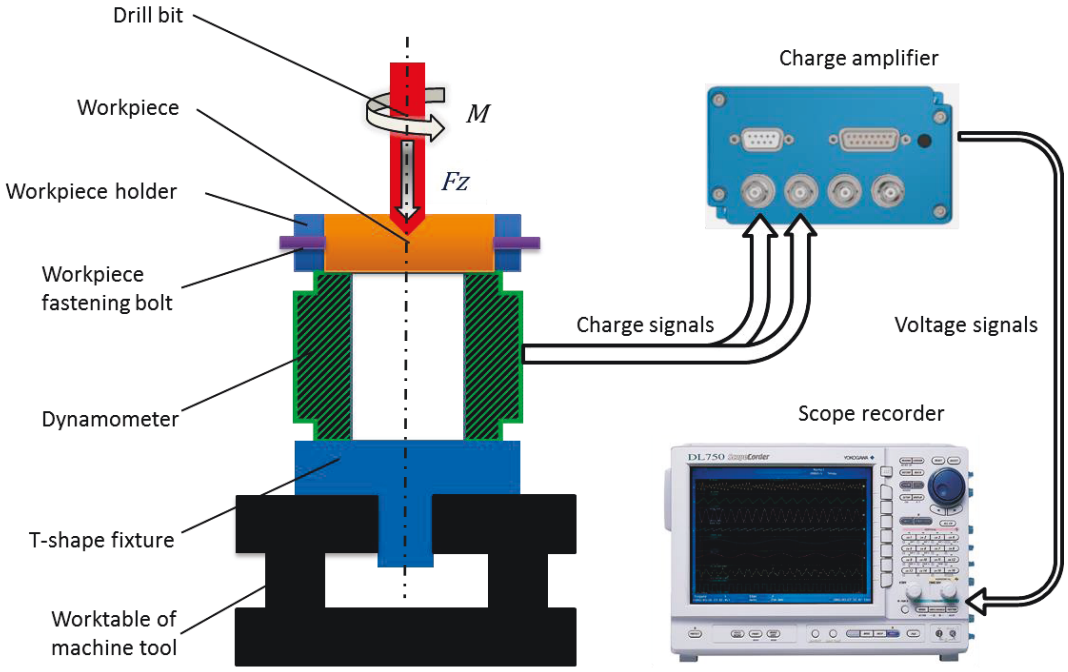

Figure 3-1 The signal system and general experimental set up

Figure 3-1 demonstrates the signal system and the experimental set up. The experiments were conducted on a Mitsubishi CNC machining center, which has 3 independently controlled axes. The worktable has 2 horizontal degrees of freedom and several T-slots in the left-right direction. A T-shape fixture is used to mount the dynamometer fastening on the worktable. Upon the dynamometer, a workpiece holder is fixed with 4 bolts. The workpiece holder has a cylindrical hole where a section of workpiece bar is inserted. Two screws are used to tighten the connection from side direction.

3.2 Signal Segmentation

Forces increase when the cutting edges are not entirely inside the work piece, where the drill depth is less than 3 mm (taking an 8 mm diameter drill bit for example, its point height is about 2.4mm) from the workpiece surface which is called entry stage. The forces decrease in another case where the drill depth is around the workpiece thickness named penetration stage. The forces are relatively stable in the medial stage between the entry stage and the penetration stage as long as the drill depth is within certain limits.

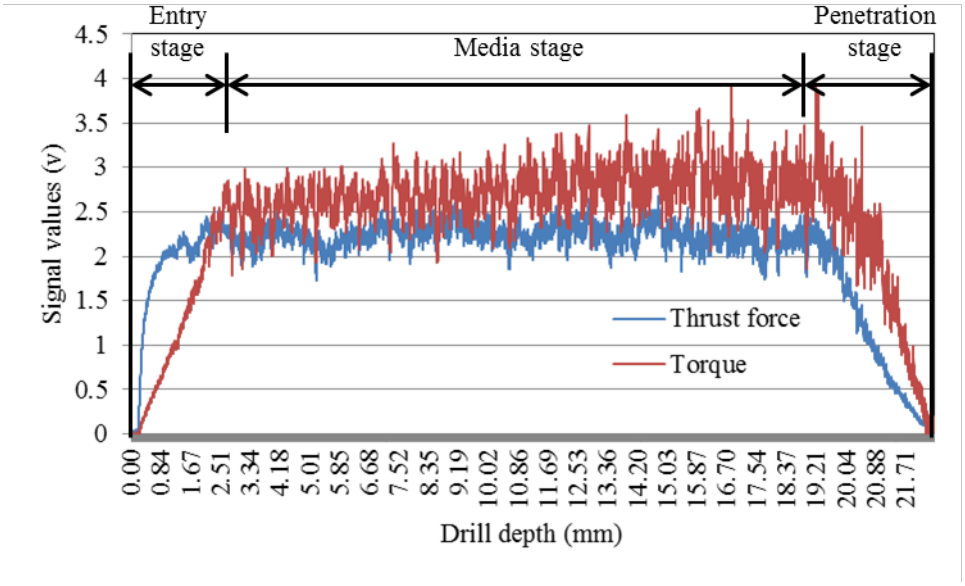

Figure 3-2 Three stage of the drill process and signal Segmentation

The 3 stages of a drilling process with an 8 mm drill bit and a workpiece with 20 mm thickness are illustrated in Figure 3-2 (the upper signal is torque signal and the lower is thrust force signal). To obtain static & dynamic features under a relatively stable condition, data intercepted during the stage between the former two stages named medial stage is used for further analysis. Thus, the aimed segmentation is intercepted from the signals in the medial stage.

3.3 Feature Generation

As aforementioned in Chapter 1, the relations between the thrust force and torque contain useful information for drilling monitoring. However, the correlation has not been investigated and quantized. So, first in this section, the correlation between the thrust force and torque signals is investigated. Moreover, the features employed in these previous works were not sophisticated enough to obtain the essence hidden in the correlation. Thus, a novel method which generates new features by converting the thrust force and torque into a rectangular coordinate is presented in the following parts.

3.3.1 The Relations between the Thrust Force and Torque Signals

There are several frequently used and effective ways to evaluate the correlations between two signals applied here, which are the cross-correlation, correlation coefficient and coherence. The former two inspect the correlation in the time domain and the later one in the spectra.

The instance data segment is intercepted from the medial stage shown in Figure 3-2. The sampling frequency is 20 KHz and the data length is 4096 for each signal.

3.3.1.1 Cross-correlation of thrust force and torque signals

In signal processing, cross-correlation is a measure of similarity of two waveforms as a function of a time-lag applied to one of them. This is also known as a sliding dot product or sliding inner-product. The cross-correlation function provides a measure of similarity between one signal and a time delayed version of the other signal.

Here we define one signal as $x(t)$ and the other as $y(t)$, then the cross-correlation function $CrCorr(\tau)$ can be defined as[65] :

$$CrCorr(\tau) = \int_{-\infty}^{+\infty} x(t)y(t+\tau)dt \qquad \text{Equation 3-1}$$

τ means the shift time of $y(t)$.

For two discrete time signals, for instances, two sequences $x(n)$ and $y(n)$, the cross-correlation $CrCorr(k)$ can be defined as[66]:

$$CrCorr(k) = \sum_{n=-\infty}^{\infty} x(n)y(n+k) \quad k = 0, \pm 1, \pm 2, \cdots \qquad \text{Equation 3-2}$$

Here index k takes only integer values. In Equation 3-2, $x(n)$ is fixed and for positive values of k, the sequence $y(n)$ is shifted left by k units while for negative values of k, the sequence $y(n)$ is shifted right by k units.

In our case here, the thrust force signal $Th(n)$ and torque signal $T(n)$ both have a length 4096, thus the shift index k can be set in the range of $[-4095,4095]$.

The cross-correlation result of thrust force and torque sequences is shown in Figure 3-3.

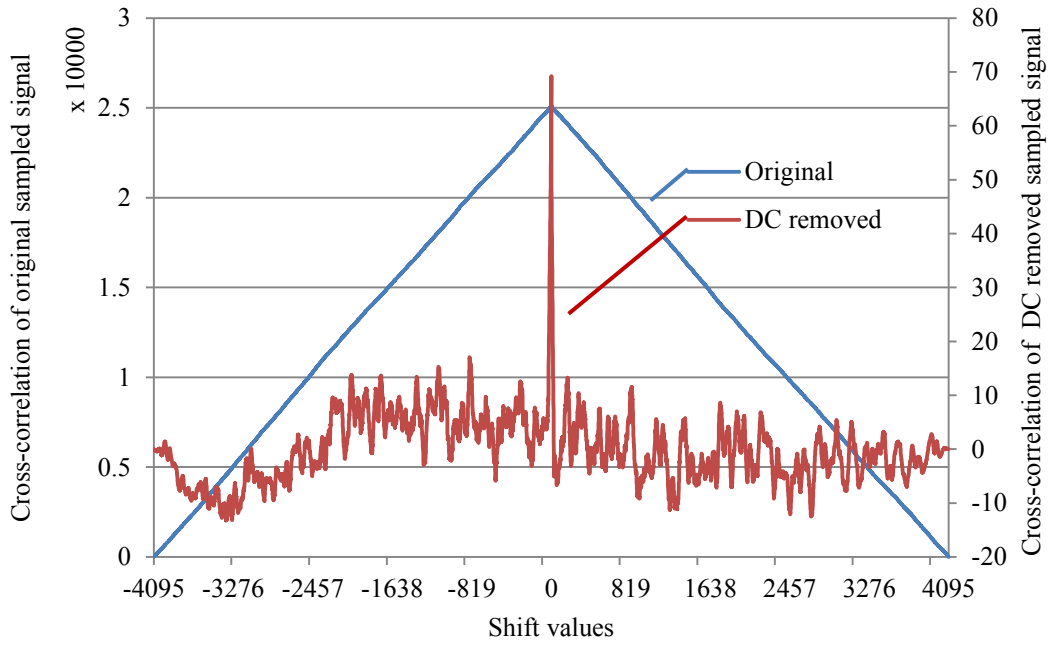

Figure 3-3 The cross-correlation of original and DC removed sampled thrust force and torque

The cross correlation values for the original sampled signals and that removed of average are both

calculated. It's known from Figure 3-3 that when the shift value is zero, which means no shift for torque sequence, the cross correlation reaches its maximum value. It implies that high coherence of these two signals can be obtained without delay of any of them, i.e., a simultaneous change exists in these two signals. Thus, the calculation of correlation coefficient and coherence is carried out without time shift.

3.3.1.2 The Cross-correlation Coefficient of Thrust Force and Torque

The cross-correlation coefficient is also known as the product-moment coefficient of correlation or Pearson's correlation. The cross-correlation coefficient is a widely used measure of the correlation between two sections of two different but random variables.

The cross-correlation coefficient of two sequences of random variables, $x(n)$ and $y(n)$, is denoted by $CrCorrCoef$. The precise mathematical definition of cross-correlation coefficient of two random variable sequences is given by [67]:

$$CrCorrCoef = \frac{Cov(x,y)}{\sqrt{Cov(x,x) \times Cov(yy)}} \quad \text{Equation 3-3}$$

where $Cov(x,y)$ means the covariance of $x(n)$ and $y(n)$, and $Cov(x,x)$ and $Cov(y,y)$ mean the autocovariance of $x(n)$ and $y(n)$.

The covariance between $x(n)$ and $y(n)$ is defined as[68]:

$$Cov(x,y) = E[(x - E(x))(y - E(y))] \quad \text{Equation 3-4}$$

where $E(x)$ is the expected value of x, also known as the mean of x.

The cross-correlation coefficient of $Th(n)$ and $T(n)$ is 0.5772, which means a high correlation between thrust force and torque.

3.3.1.3 The Coherence of Thrust Force and Torque

The spectral coherence is a statistic that can be used to examine the relation between two signals or data sets. The coherence (sometimes called magnitude-squared coherence) between two signals $x(t)$ and $y(t)$, is a real-valued function that is defined as[69]:

$$Coh_{xy}(f) = \frac{|CPSD_{xy}(f)|^2}{PSD_{xx}(f)PSD_{yy}(f)} \quad \text{Equation 3-5}$$

where $CPSD_{xy}(f)$ means the cross power spectral density of $x(t)$ and $y(t)$ at the frequency f, and $PSD_{xx}(f)$ and $PSD_{yy}(f)$ mean the power spectral density of $x(t)$ and $y(t)$ at the frequency f.

Figure 3-4 shows the coherence values of thrust force and torque in different frequencies. It can be seen that the coherence values are around 0.6 along the frequencies which means a high relation between the thrust force and torque.

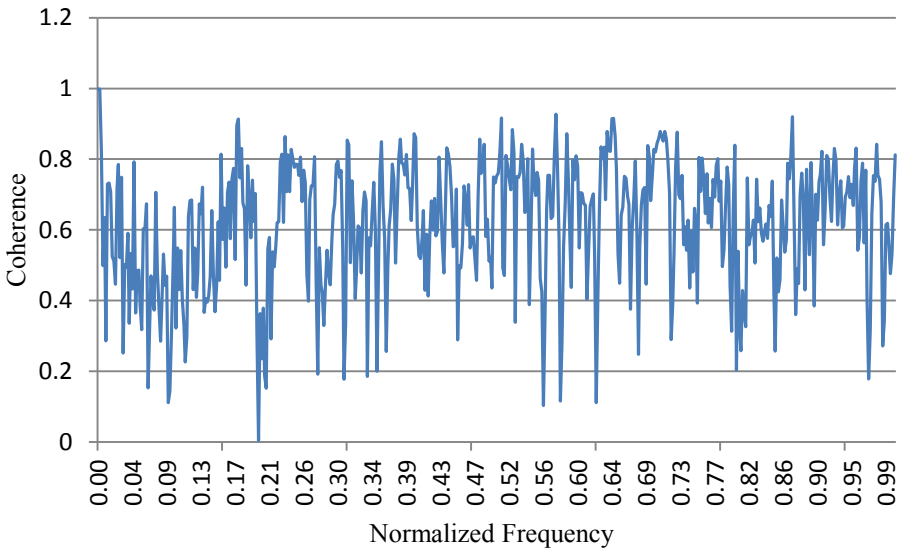

Figure 3-4 The coherence of thrust force and torque

3.3.2 The Novel Feature Generation Method

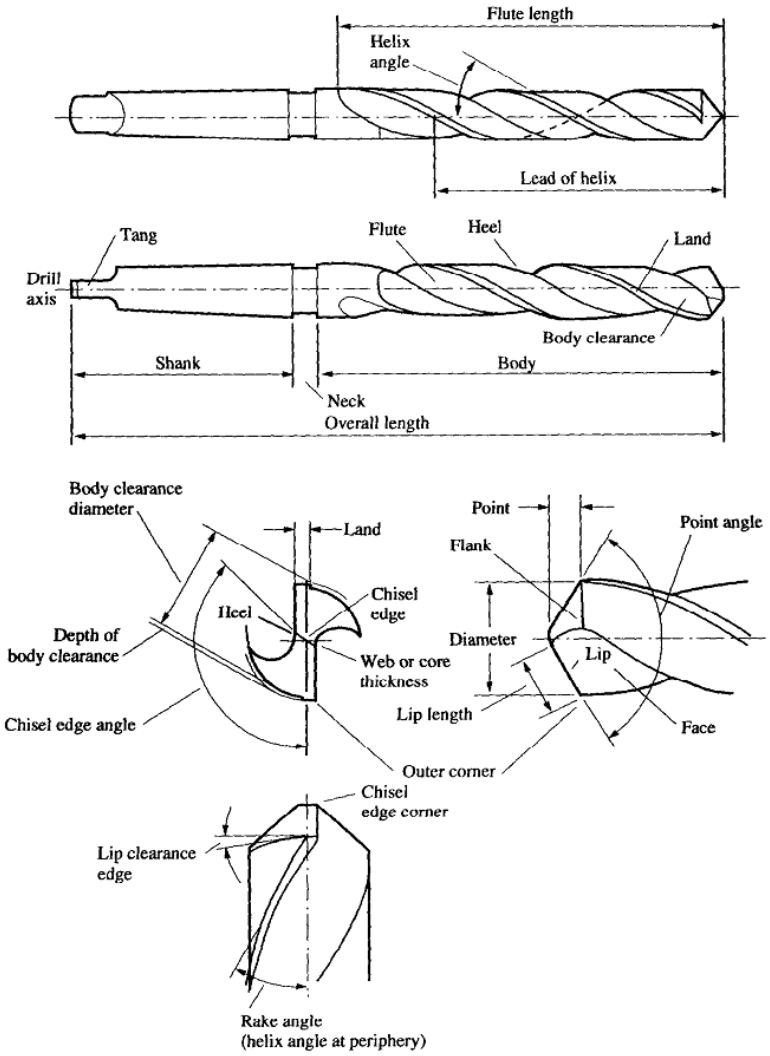

Figure 3-5 Twist drill geometry [70]

Twist drill geometry and force modeling of drill process are very complicated as discussed in [70]–[72]. The general used definition of the geometry parameters is shown in Figure 3-5.

Owing to the complexity of the geometry of drill bit, usually, cutting in the chisel region was treated as orthogonal cutting with different cutting speeds while the cutting forces along the cutting lips were represented as a series of oblique sections [73]. However, the force modeling is too complicated to be applied to signal analysis. Therefore, a simplified force model is required.

The thrust (*Fz*) and torque (*Mz*) monitored by the dynamometer during the drilling process are conversed to equivalent thrust force (*Ft*) and principal force (*Fp*) on two symmetrical points at each cutting lip as shown in Figure 3-6.

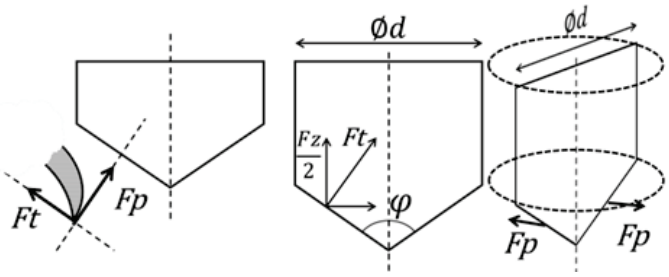

Figure 3-6 The conversion of thrust force and torque

The equivalent thrust force (*Ft*) and equivalent principal force (*Fp*) can be drawn as the following equations:

$$Ft = Fz/(2 * sin(\varphi/2)) \quad \text{Equation 3-6}$$

$$Fp = 2Mz/d \quad \text{Equation 3-7}$$

in which φ means the point angle and d means the diameter of the drill bit.

Then a rectangular coordinate is applied with the horizontal axis as principal force and vertical axis as thrust force. Therefore, any point in this rectangular coordinate means the resultant force of the equivalent thrust force and principal force. The data sampled and processed from the experiment become a series of the resultant force (*Fr*) vectors in time sequences and to be the trajectory of the resultant force (*Fr*).

In previous works, most of them used features generated from thrust force and torque independently, that is, equivalent to generate features from *Ft* and *Fp* independently, because *Ft* and *Fp* can be regarded as *Fz* and *Mz* multiplied with certain constants. However, none of them converted *Fz* and *Mz* into forces which can create a third force: the resultant force (*Fr*). Consequently, new features can be generated with the resultant force (*Fr*). Features generated from *Ft, Fp* and *Fr* are listed in Table 3-1.

Table 3-1 Features of *Ft, Fp and Fr*

Force	Features
Ft	$\overline{Ft}$, *ΔFt* ,*Ft_RMS*, *Ft_STDEV*, *Ft_Vel*
Fp	$\overline{Fp}$, ΔFp, *Fp_RMS*, *Fp_STDEV*, *Fp_Vel*
Fr	$\bar{\rho}$, $\bar{\theta}$, *Δρ*, *Δθ*, *ρ_RMS*, *θ_RMS*, *Fr_STDEV*, *Fr_Vel*, *Fr_ConvHullArea*

$\overline{Ft}$ and $\overline{Fp}$ are the mean values of the equivalent thrust force and principal force, as shown in Figure 3-7, and the average features of the resultant force are $\bar{\theta}$ and $\bar{\rho}$ which can be described as:

$$\bar{\theta} = atan(\overline{Ft}/\overline{Fp}) \qquad \text{Equation 3-8}$$

$$\bar{\rho} = \sqrt{\overline{Ft}^2 + \overline{Fp}^2} \qquad \text{Equation 3-9}$$

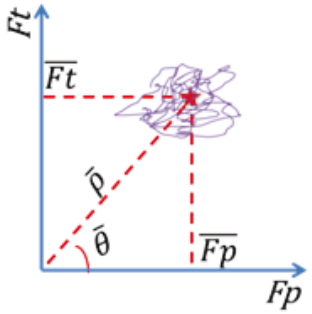

Figure 3-7 Definition of static components

ΔFt and ΔFp are the fluctuation ranges of *Ft* and *Fp* and $\Delta\theta$ and $\Delta\rho$ indicate the variation range of the resultant force in polar form as shown in Figure 3-8.

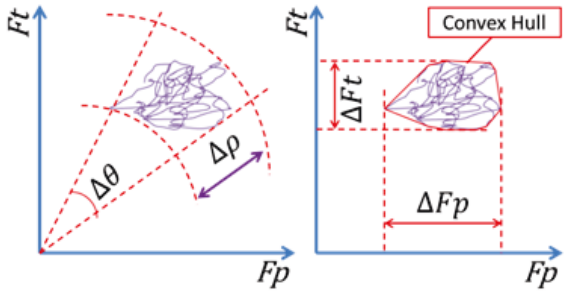

Figure 3-8 Definition of dynamic components

Ft_RMS, Ft_STDEV represent the root mean square (RMS) and the standard deviation of *Ft.* And, *Ft_Vel* means the variation velocity for Ft, which is defined by Equation 3-10. Similar rules adopted by those for *Fp*, ρ and θ.

$$Ft\_Vel = \frac{1}{t_0}\sum_{i=2}^{k}|Ft_i - Ft_{i-1}| \qquad \text{Equation 3-10}$$

In Equation 3-10 k indicates the number of data points in the period of t_0.

The definition of the standard deviation and velocity features of *Fr* are slightly different to one feature vectors. "*Fr_STDEV*" is defined as the average distance from each trajectory point (Ft_i, Fp_i) to the center point $(\overline{Ft}, \overline{Fp})$ by Equation 3-11 and, "*Fr_Vel*" is defined as the path length of the resultant force trajectory in unit time, i.e. the velocity of the resultant force track in the previously declared coordinate illustrating by Equation 3-12.

$$Fr\_STDEV = \frac{1}{n}\sum_{i=1}^{n}\sqrt{(Ft_i - \overline{Ft})^2 + (Fp_i - \overline{Fp})^2} \quad \text{Equation 3-11}$$

$$Fr\_Vel = \frac{1}{t_0}\sum_{j=2}^{m}\sqrt{(Ft_j - Ft_{j-1})^2 + (Fp_j - Fp_{j-1})^2} \quad \text{Equation 3-12}$$

In Equation 3-11, *n* means the number of data points during the whole sampling period; while in Equation 3-12 *m* indicates the number of data points in the period of t_0.

Intuitively, the concentration of the distribution of data points directly indicates the stability of the resultant force. Therefore, "*Fr_ConvHullArea*" is defined as the area of the convex hull of the discretely distributing data points as shown in Figure 3-8.

3.3.3 The Features Grouping

And according to the statistical methods, features listed in Table 3-1 are grouped into static features, and dynamic features. The static features comprise average and RMS features of the forces, and the dynamic features contain delta features, standard deviation features, velocity features and the geometry feature.

Table 3-2 The grouping of different features

Group	Sub groups	Features
Static features	Average	$\overline{Ft}$, $\overline{Fp}$, $\bar{\rho}$, $\bar{\theta}$
	RMS	*Ft_RMS*, *Fp_RMS*, *ρ_RMS*, *θ_RMS*
Dynamic features	Delta	*ΔFt*, *ΔFp*, *Δρ*, *Δθ*
	Standard deviation	*Ft_STDEV*, *Fp_STDEV*, *Fr_STDEV*
	Velocity	*Ft_Vel*, *Fp_Vel*, *Fr_Vel*
	Geometry	*Fr_ConvHullArea*

The detailed grouping is shown in Table 3-2.

3.4 Feature Extraction

All of the features generated in Table 3-1 are time domain features. To examine their performances in frequency domain, reconstruct of the time domain signals is necessary. Wavelet packet transform decomposition and reconstruction are used to decompose the signal into series of sub-signals stand for different frequency bands. Then, with the reconstructed time domain sub-signals, the features can be extracted using the feature generation method aforementioned in 3.3.2.

3.4.1 Fourier Transform and Wavelet Transform

It is well known that an energy limited signal (i.e. a square integrable signal), $f(t)$, can be decomposed by its Fourier transform $F(\omega)$ as:

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} F(\omega)e^{i\omega t}d\omega \quad \text{Equation 3-13}$$

Where

$$F(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-i\omega t} dt \qquad \text{Equation 3-14}$$

Equation 3-14 is called the Fourier transform and Equation 3-13 is called the inverse Fourier transform. The Fourier transform implies that the signal $f(t)$ can be decomposed into a family of harmonics $e^{i\omega t}$ and the weight coefficients $F(\omega)$, which represent the amplitude of the harmonics in $f(t)$.

The wavelet transform is defined in a similar manner. However, instead of using the harmonics $e^{i\omega t}$, the wavelet transform uses wavelet bases:

$$\psi_{s\tau}(t) = \frac{1}{s}\psi(\frac{t-\tau}{s}) \qquad \text{Equation 3-15}$$

where *s* represents the frequency, *t* represents the time shift or 'location', and ψ(·) is called a mother wavelet function. Accordingly, a signal $f(t)$ can be decomposed into[74]:

$$f(t) = \frac{1}{c_\psi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} W_s[f(\tau)] \frac{1}{s}\psi(\frac{t-\tau}{s})\, ds d\tau \qquad \text{Equation 3-16}$$

where c_ψ is a constant depending on the base function, and $W_s[f(t)]$ is the wavelet transform defined as:

$$W_s[f(\tau)] = \int_{-\infty}^{+\infty} f(t) \frac{1}{s}\psi(\frac{t-\tau}{s}) dt \qquad \text{Equation 3-17}$$

Equation 3-16 is called the wavelet transform and Equation 3-17 is called the inverse wavelet transform (reconstruction).

Wavelet transforms decompose a signal into various components at different time windows and frequency bands, and all of which form a surface in the time–frequency plane. The size of the time window is controlled by the translation while the length of the frequency band is controlled by the dilation. Hence, one can examine the signal at different time windows and frequency bands by controlling translation and dilation.

3.4.2 Wavelet Packet Decomposition and Reconstruction (WPD&R)

Wavelet packets are particular linear combinations of wavelets. They form bases that retain the orthogonality, smoothness and locational properties of their parent wavelets. The wavelet packet transform is the most generalized signal decomposition method. In wavelet analysis, a signal is decomposed into a low-frequency component known as approximate, and a high-frequency component known as detail. The approximate is then decomposed into a second level of approximation and detail, and this process is repeated. However, in wavelet packet analysis, the approximate as well as the detail parts can be decomposed.

One of the most commonly used discrete wavelet transforms is the binary orthogonal wavelet transform. At *j*-th resolution, let $A_j[f(t)]$ be the operator of the approximation of the signal *f(t)* and $D_j[f(t)]$ represent the detail of the signal *f(t)* and it has been verified[75]:

$$A_j[f(t)] = f(t) * \emptyset_j(t) \quad \text{Equation 3-18}$$

$$D_j[f(t)] = f(t) * \psi_j(t) \quad \text{Equation 3-19}$$

where $\phi_j(t)$ are smooth scaling orthogonal bases, $\psi_j(t)$ are orthogonal wavelet bases and "*"denotes convolution. Furthermore, $\phi_j(t)$ and $\psi_j(t)$ are correlated through a pair of quadrature mirror filters $h(t)$ and $g(t)$ defined below:

$$\phi_j(t) = h(t) * \phi_{j-1}(t) \quad \text{Equation 3-20}$$

$$\psi_j(t) = g(t) * \psi_{j-1}(t) \quad \text{Equation 3-21}$$

Therefore, the discrete binary wavelet transform is then obtained:

$$A_j[f(t)] = \sum_k h(k-2t) * A_{j-1}[f(t)] \quad \text{Equation 3-22}$$

$$D_j[f(t)] = \sum_k g(k-2t) * A_{j-1}[f(t)] \quad \text{Equation 3-23}$$

where t=1,2,…, N; j=1,2,…,J and J=$\log_2 N$.

Let operators H and G be the convolution sum defined below:

$$H\{\cdot\} = \sum_k h(k-2t) \quad \text{Equation 3-24}$$

$$G\{\cdot\} = \sum_k g(k-2t) \quad \text{Equation 3-25}$$

Then Equation 3-22 and Equation 3-23 could be:

$$A_j[f(t)] = H\{A_{j-1}[f(t)]\} \quad \text{Equation 3-26}$$

$$D_j[f(t)] = G\{A_{j-1}[f(t)]\} \quad \text{Equation 3-27}$$

Equation 3-26 and Equation 3-27 are the approximation and detail of the j-1th approximation, for the j-1th detail, similar process can be drawn:

$$A_j^{Dj-1}[f(t)] = H\{D_{j-1}[f(t)]\} \quad \text{Equation 3-28}$$

$$D_j^{Dj-1}[f(t)] = G\{D_{j-1}[f(t)]\} \quad \text{Equation 3-29}$$

Finally, let $P_j^i(t)$ be the ith packet on jth resolution, then, the wavelet packet transform can be computed by the recursive algorithm below:

$$P_j^{2i-1}(t) = H\{P_{j-1}^i(t)\} \quad \text{Equation 3-30}$$

$$P_j^{2i}(t) = G\{P_{j-1}^i(t)\} \quad \text{Equation 3-31}$$

where, t=1, 2, . . . ,2^{J-j}; i=1, 2, . . . ,2^j; j=1, 2, . . . , J, and J=$\log_2 N$.

Table 3-3 Wavelet packet transform decompose

Resolution	Signal/Packet							
0	f(t)							
1	P_1^1				P_1^2			
2	P_2^1		P_2^2		P_2^3		P_2^4	
3	P_3^1	P_3^2	P_3^3	P_3^4	P_3^5	P_3^6	P_3^7	P_3^8
…	…							
j	P_j^i							

Table 3-3 shows the decomposed packets $P_j^i(t)$ in a binary wavelet packet decomposition tree.

The signal reconstruction formula of the wavelet packet transform is as follows:

$$P_j^i(t) = 2[\bar{H}P_{j+1}^{2i-1}(t) + \bar{G}P_{j+1}^{2i}(t)] \qquad \text{Equation 3-32}$$

The operators $\bar{H}$and $\bar{G}$ are the conjugate of *H* and *G*.

The reconstructed sub-signals are extracted from different frequency bands. Table 3-4 shows the frequency bands of different sub-signals. The range values listed are the ratios which should multiply the sampling frequency to get the real bands.

Table 3-4 Frequency bands of the decomposed sub-signals

Resolution	Frequency							
0	0-1							
1	0-1/2				1/2-1			
2	0-1/4		1/4-1/2		1/2-3/4		3/4-1	
3	0-1/8	1/8-2/8	2/8-3/8	3/8-1/2	1/2-5/8	5/8-3/4	3/4-7/8	7/8-1
…	…							
j	For P_j^i, (i-1)*2^{-j} - i*2^{-j}							

In this work, a 4 level resolution decomposition and reconstruction of the 20 KHz sampled signals are conducted, thus, the observation frequency band resolution is 625 Hz (20 KHz/2/16).

3.5 Feature Selection

3.5.1 The Necessity for Applying PCA

For each two sampled channel time domain signals, 19 features are generated as discussed in 3.3.2. And 16 sub-signals are reconstructed with a 4 level WPD&R. Consequently, 304 features are obtained for each drilling instance. In general, different features extracted from multiple machining process instances can be regarded as a high-dimensional multivariate random matrix composed of several vectors formed by different extracted features. Here suppose that *N* example drilling processes are conducted and the signals are samples, then an *N×304* matrix which contains 304 feature vectors is obtained after the feature extraction. It is not feasible to input the whole matrix to the prediction model

without any feature selection or dimensionality reduction procedure. Because some of them are abundant of useful information while a lot of them contain plentiful noises, and there are high relations between these feature vectors. Therefore, the feature selection process is quite essential.

The objective of feature selection is to preserve as much of the relevant information as possible by removing redundant or irrelevant information in acquired sensory signals. The main feature selection technique is principal component analysis (PCA), which has been widely used in system identification and dimensionality reduction in dynamic systems[9]. By implementing PCA, the complexity of modelling processes can be reduced and new feature vectors can be reconstructed. PCA transforms a number of correlated sensory features into new uncorrelated features (or principal components), thus reducing the complexity of modelling processes.

Shi and Gindy applied PCA to calculate the covariance matrix estimated by two directional cutting forces (the saw width and length direction, specifically) signals in broaching[76]. The eigenvectors and corresponding eigenvalue can be obtained. Coincidentally, the estimated eigenvectors and eigenvalues have obvious geometrical meaning and are appropriate for feature extraction. For an instance, the first and second eigenvalue corresponds to the length of the major axes and minor axes of scatter ellipse respectively. Additionally, the first eigenvector indicates the inclination angle of the major axes in the scatter ellipse. As a result, PCA is an efficient tool to extract the features from two perpendicular sensory signals in the form of scatter ellipse, i.e. the length (a/b) of the major/minor axes and inclination angle (β). The ellipse obtained by PCA is illustrated in Figure 3-9.

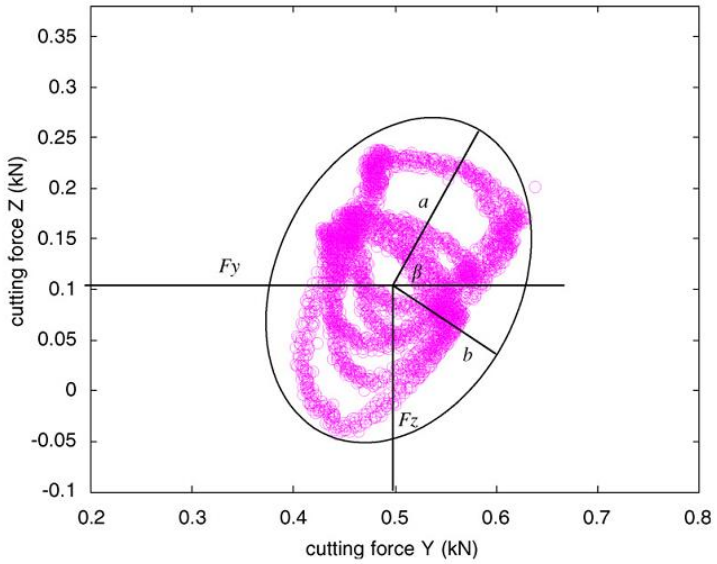

Figure 3-9 Orbit diagram of cutting force signals in dual directions measured by integrated force [76]

Unfortunately, the plot of *Ft* and *Fp* in the rectangular coordinate in our work doesn't show a distribution area similar to an ellipse. Thus, the PCA is not used to calculate the length of the major axes and minor axes from *Ft* and *Fp* signals directly. However, the PCA principal did inspire the innovation of the feature generation method. Moreover, the PCA can effectively select the principal features from the feature matrix, though the expression would be different from that in two dimensional spaces.

3.5.2 The PCA Algorithm

Principal component analysis (PCA) is a statistical procedure that uses orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. The number of principal components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has the largest possible variance (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to (i.e., uncorrelated with) the preceding components. Principal components are guaranteed to be independent if the data set is jointly normally distributed. PCA is sensitive to the relative scaling of the original variables.

PCA is mathematically defined as an orthogonal linear transformation that transforms the data to a new coordinate system such that the greatest variance by some projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on[77].

The PCA algorithm is detailed in [78] as follows.

Given the *d*-dimensional (here d = 304) feature vectors of M training samples, a data matrix $X = \{x_1, x_2, \dots, x_M\}$, where $x_i^T = \{F_{ij}, j = 1,2, \dots, d\}$, can be constructed. The corresponding covariance matrix R_X is as follows:

$$R_X = \sum_{i=1}^{M}(x_i - \bar{x})(x_i - \bar{x})^T \qquad \text{Equation 3-33}$$

where

$$\bar{x} = \frac{1}{M}\sum_{i=1}^{M} x_i \qquad \text{Equation 3-34}$$

is the mean feature vector of all samples of X and T denotes the transportation. Accordingly, the corresponding eigenvalues of the matrix X can be obtained by solving the following equation:

$$\lambda v = R_X v \qquad \text{Equation 3-35}$$

where λ and ν are the eigenvalue and the eigenvector of R_X respectively.

From Equation 3-35, most d eigenvalues can be obtained. The eigenvalues are usually arranged in the descending order: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$. And the corresponding eigenvector of $\lambda_j (j = 1,2, \dots, d)$ is denoted as $v_j (j = 1,2, \dots, d)$. The j-th principal component feature of a sample pattern x_i, denoted as f_{ij}, can be obtained by projecting x_i onto the direction of the eigenvector v_j:

$$f_{ij} = v_j^T (x_i - \bar{x}) \qquad \text{Equation 3-36}$$

The eigenvectors span a d-dimensional orthogonal space and the principal component (PC) representations can be obtained by projecting X onto the space.

The capability of each PC representation can be evaluated based on its contribution to the machine condition classification. To effectively represent and classify machine conditions whereas reducing the computation load, only the most representative PCs are needed. One of the traditional selection methods is the accumulative contribution rate (ACR), which selects the first m PCs based on the eigenvalues. Following the notation above, ACR is defined as follows:

$$R_m = \frac{\sum_{i=1}^{m} \lambda_i}{\sum_{i=1}^{d} \lambda_i} \quad \text{Equation 3-37}$$

From a mathematical point of view, R_m indicates the percentage of the total variance in the original feature set vectors (FSVs) explained by the first m PCs. Using the ACR method with a specific threshold, the first m PCs are then chosen as the new feature vector.

Denote $PC_i (i = 1,2, \dots, d)$ as the i-th PC representation of the signal samples, it is a vector with the same length as the j-th original FSV $F_j (j = 1,2, \dots, d)$. Based on the definition of correlation coefficient, the correlation between PC_i and F_j can be computed by

$$\rho_{ij} = \left| \frac{E\{[PC_i - E(PC_i)][F_j - E(F_j)]\}}{\sqrt{D(PC_i)} \times \sqrt{D(F_j)}} \right| \quad \text{Equation 3-38}$$

where $E(\cdot)$ represents the mean and $D(\cdot)$ represents the variance. The correlation vector $\rho_i = [\rho_{i1}, \rho_{i2}, \dots, \rho_{id}]$ reflects the strength of the relationship between the PC_i and the other statistical features. The bigger the correlation coefficient ρ_{ij} is, the more similar the i-th PC is to the j-th statistical feature.

In summary, the m principal components are selected to represent the d features generated from drilling process by a matrix transform defined in Equation 3-36, with a big shrink of the feature number but still possession of the information.

3.6 Decision-making support AI System

Artificial neural networks (ANNs) are computational models inspired by animals' central nervous systems (in particular the brain) that are capable of machine learning and pattern recognition. They are usually presented as systems of interconnected "neurons" that can compute values from inputs by feeding information through the network.

A feed forward neural network (FFNN) is an artificial neural network where connections between the units do not form a directed cycle. The FFNN was the first and simplest type of artificial neural network devised. In this network, the information moves in only one direction, forward, from the input nodes, through the hidden nodes (if any) and to the output nodes. There are no cycles or loops in the network.

Back propagation, an abbreviation for "backward propagation of errors", is a common method of training artificial neural networks.

Here, an artificial feed forward backward propagation neural network (FFBPNN) is applied to support the decision-making in drilling monitoring.

3.6.1 The Architecture and the Training Procedure of the BPNN

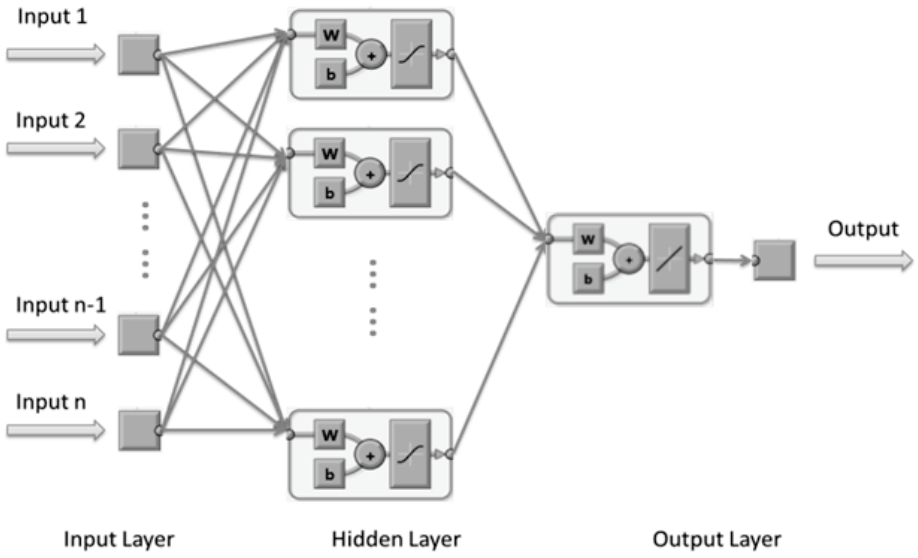

Figure 3-10 Architecture of a 3 layered feed forward neural network

A multilayer feed-forward neural network consists of at least three layers. The three layers are input layer, hidden layer and output layer [79]. Figure 3-10 shows the architecture of a back propagation neural network (BPNN) model. The input layer receives information from the external sources and passes this information to the network for processing. The hidden layer, composed by several sigmoid neurons, receives information from the input layer, and does all the information processing. And the output layer, consist of a linear neuron receives processed information from the network, and sends the results out to an external receptor.

There are 3 main learning paradigms for neural networks: supervised, unsupervised and hybrid [80]. In supervised learning, the network is provided with a correct answer (or answer set, targets), according to which it is trained to produce answer (output) as close as possible to the correct answer. In contrast, unsupervised learning does not require a correct answer and it explores the underlying structure in the data to organize or category them. The hybrid learning, just as its name implies, combines supervised and unsupervised learning. Naturally the prediction of drill wear requires supervised learning with actual drill corner wear as the correct answer.

The back propagation learning algorithm based on error-correction rule within supervised learning is one of the most popular learning methods for multi-layer feed forward neural networks. Functioning of back propagation proceeds in three stages, namely learning or training, testing or inferences and validation [51]. The data is commonly divided into 3 sets randomly: training set, validation set and testing set. The training set is used to train the network; the validation set is for avoiding overfitting and the testing set is used to obtain the performance characteristics such as accuracy, sensitivity, specificity and so on of the neural network.

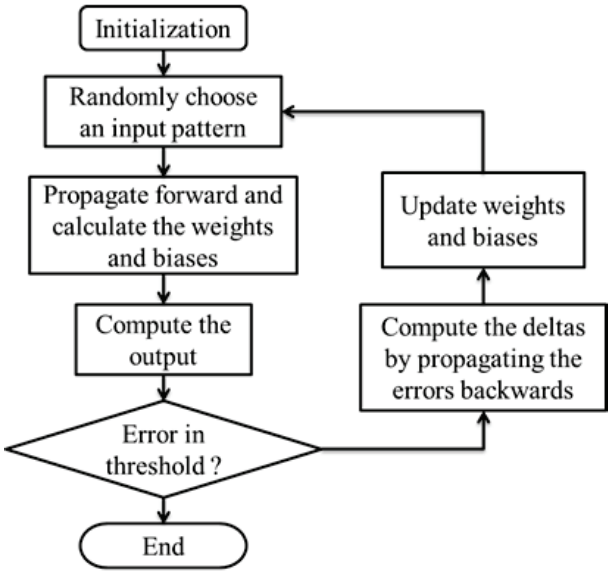

Figure 3-11 Back propagation algorithm procedure

Figure 3-11 shows the procedure of back propagation algorithm. And the main processes of back propagation algorithm proceeds as follows: first, inputs are presented to the network and errors are calculated; second, sensitivities are propagated from the output layer to the first layer; then, weights and biases are updated [81].

3.6.2 The Training Algorithm of the BPNN

The training of the BPNN first starts with the normalization of the input data, usually to values between 0.1 and 0.9, by the following equation:

$$I_i = \frac{F_i - F_{min}}{F_{max} - F_{min}} \times 0.8 + 0.1 \qquad \text{Equation 3-39}$$

where F_{max} and F_{min} mean the maximum and minimum input data, F_i means the i-th input and I_i is the normalized input.

A l-m-n (l input neurons, m hidden neurons, and n output neurons) architecture of a back propagation neural network model is illustrated in Figure 3-12. The interconnection weight between the input layer and the hidden layer is v_{ji} which represents the interconnection of i-th node of the input layer to j-th node of the hidden layer. And those between the hidden layer and output layer are denoted as w_{kj}. And to simplify the expression, the output of the hidden neuron is marked as P_j. The k-th output of the neural network is represented as O_k.

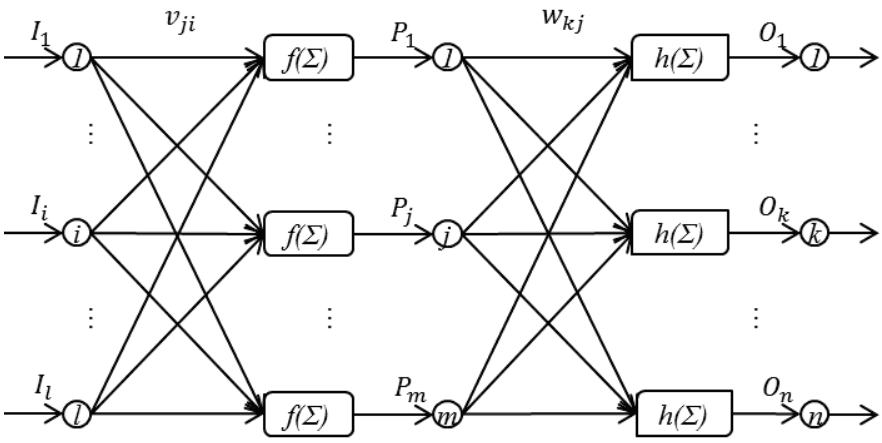

Figure 3-12 Mathematical structural illustration of the BPNN

First the output of the hidden neuron P_j is calculated by:

$$P_i = f(\sum_{i=1}^{l} v_{ji} \times I_i) \qquad \text{Equation 3-40}$$

And then the output O_k is calculated by:

$$O_k = h(\sum_{j=1}^{m} w_{kj} \times P_j) \qquad \text{Equation 3-41}$$

In Equation 3-40 and Equation 3-41, $f()$ and $h()$ are the transfer function for the hidden neurons and the output neurons, which usually are the sigmoid function given by:

$$f(x) = \frac{1}{1+e^{-x}} \qquad \text{Equation 3-42}$$

During training, the predicted output is compared with the desired output, and the mean square error is calculated. If the mean square error is more than a prescribed limiting value, it is back propagated from output to input, and weights are further modified till the error or number of iterations is within a prescribed limit.

Mean square error, E is defined as:

$$E = \sum_{k=1}^{n} \frac{1}{2}(D_k - O_k)^2 \qquad \text{Equation 3-43}$$

where D_k is the k-th target ouput.

Then, for the weight update, each weight-synapse follows the following steps: (1) multiply its output delta and input activation to get the gradient of the weight; (2) subtract a ratio (percentage) of the gradient from the weight. This ratio (percentage) influences the speed and quality of learning; it is called the learning rate. The greater the ratio, the faster the neuron trains; the lower the ratio, the more accurate the training is.

The modification value Δu_i for the weight u_i is calculated using gradient descent by:

$$\Delta u_i = -\eta \frac{\partial E}{\partial u_i} \qquad \text{Equation 3-44}$$

where η is the learning rate.

3.6.3 The Pattern Recognition by BPNN

While using BPNN for pattern recognition, the output number of the network is set the same as the pattern number, with each output corresponding to each right pattern. The value of the right pattern is set to 1 usually in a training target data set while values of others are set as zero. After the training process, test data is put into the network and output vector is obtained. The output value will be close to 1 for the corresponding pattern and close to 0 for others. Then the recognition accuracy is further calculated.

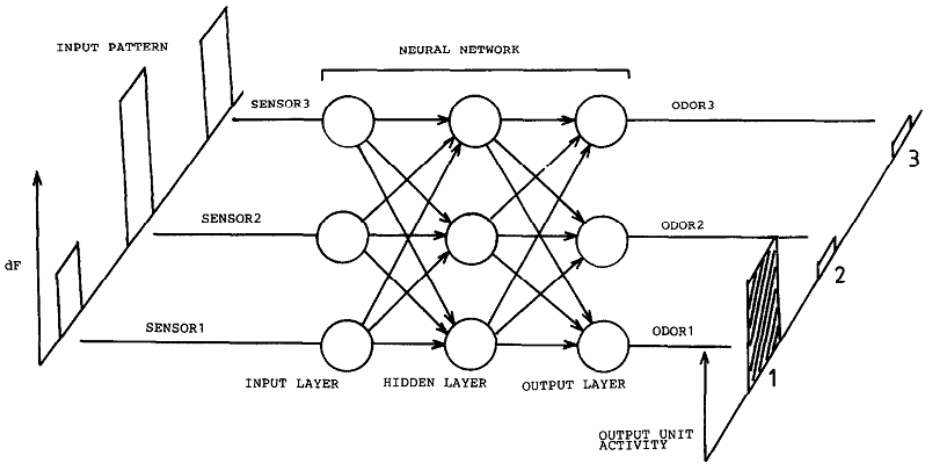

Figure 3-13 The application of BPNN for pattern recognition [82]

BPNN were used to recognize different odors in [83] and [82]. Training data for different odor patterns acquired from different sensor signal features were used to train the BPNN. After training, an odor recognition test was performed. If output unit 1 reacted strongly, as shown in Figure 3-13, the odor was identified as odor 1.

In this research, BPNN is used to recognize different drilling conditions. Training data obtained from experiments with different values for one drilling condition parameter while same value for others is used to train a BPNN, and from the recognition accuracy, the effect of the drilling condition parameter can be evaluated.

Chapter 4 Findings of the Relations between Features and Drilling Process

In this chapter, the relations between the drilling process conditions and the features generated and extracted above are revealed. The drilling process conditions are expressed by several parameters, which are workpiece material, drill diameter, spindle speed, feed rate and the drill corner wear.

To maintain the consistency between these different situations, all of the features are extracted from the first frequency band (equivalent sampling frequency 1.25 kHz) signals in this chapter if there is no specification.

It also has been known that as the hole depth increases, the chips tend to cluster together and clog the flutes, causing increased forces, poor hole quality, elevated drill temperatures, and drill breakage[84]. Therefore the workpiece thickness is set at 15 mm, and the segmentations positions of the intercepted data are the same for all the conditions.

Expect for the investigation of drill wear influence, new drills are used to conduct experiments to eliminate the effect of drill wear to different features.

4.1 Workpiece Material

In order to establish a comprehensive system which covers most frequent situations in common production activities, 4 different workpiece materials (cast iron, S45C, SUS304, Ti-alloy) were investigated in this paper.

Table 4-1 Conditions for different workpiece materials

Spindle Speed	400 rpm
Feed Rate	0.08 mm/rev.
Drill Bit Diameter	8 mm
Workpiece Material	cast iron, S45C, SUS304, Ti-alloy
Workpiece Thickness	15 mm
Lubrication	Dry

The drilling condition parameters of experiments for different workpiece materials are listed in Table 4-1.

4.1.1 Static features

4.1.1.1 Average features

Figure 4-1 shows the relations between the workpiece materials and the Average features. As shown in the legend, Ft_m, Fp_m, ρ_m and θ_m represent $\overline{Ft}$, $\overline{Fp}$, $\bar{\rho}$ and $\bar{\theta}$ respectively.

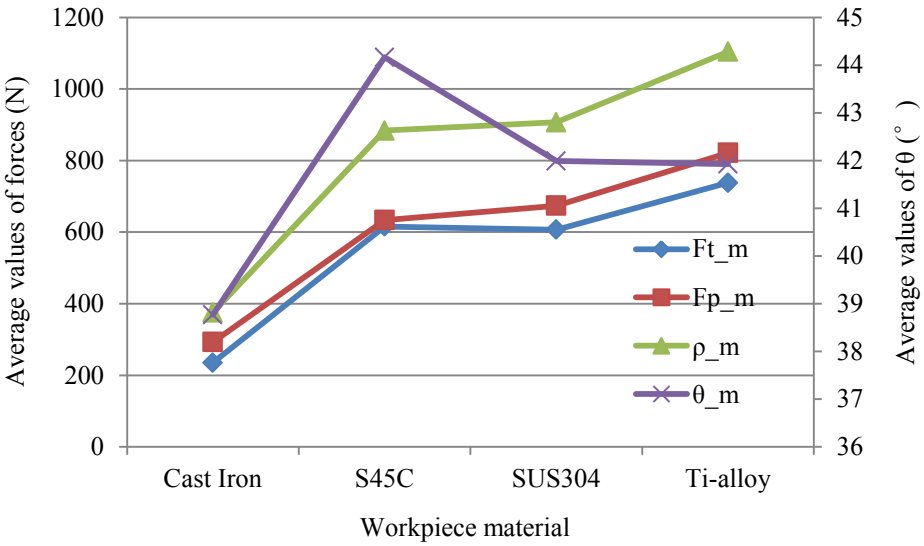

Figure 4-1 The relations between the workpiece material and the average features

The $\overline{Ft}$, $\overline{Fp}$, $\bar{\rho}$ increase for the different workpiece materials as shown in Figure 4-1, due to the increase of cutting difficulty. $\bar{\theta}$ increases comparing with cast iron and S45C, while decreases from S45C to Ti-alloy, which means that the cutting states for S45C, SUS304 and Ti-alloy are similar but differ to that of cast iron.

4.1.1.2 RMS features

Figure 4-2 illustrates the relations between the workpiece materials and the RMS features, where Ft_RMS, Fp_RMS, ρ_RMS and θ_RMS represent the RMS values of the sampled *Ft*, *Fp*, ρ and θ sequences.

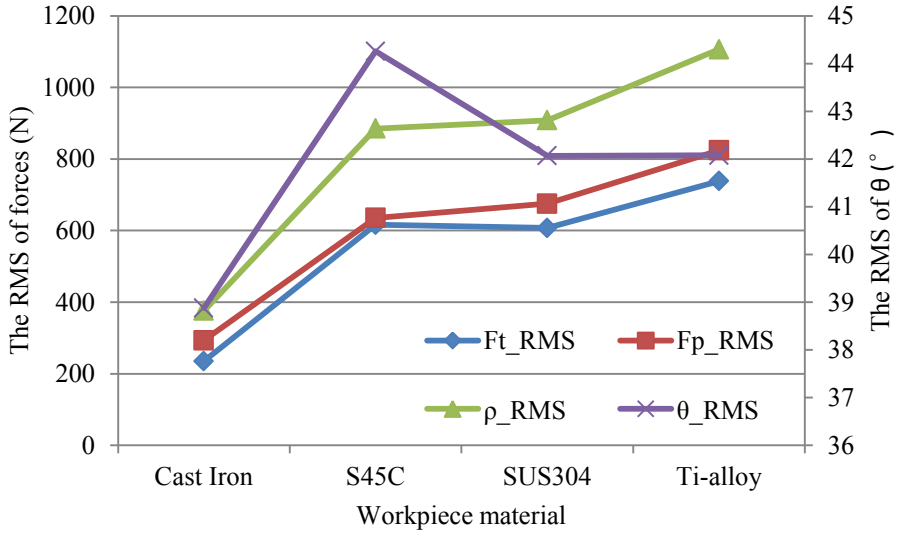

Figure 4-2 The relations between the workpiece material and the RMS features

It is shown that the RMS features of the forces increase as the workpiece material changes, and still there is an obvious difference between the cast iron and others. The trends of the RMS features are very close to those of the average features, while different to that of the delta features, due to their different

statistical algorithm. In Figure 4-2, no drop is found as that in Figure 4-3.

4.1.2 Dynamic Features

4.1.2.1 Delta features

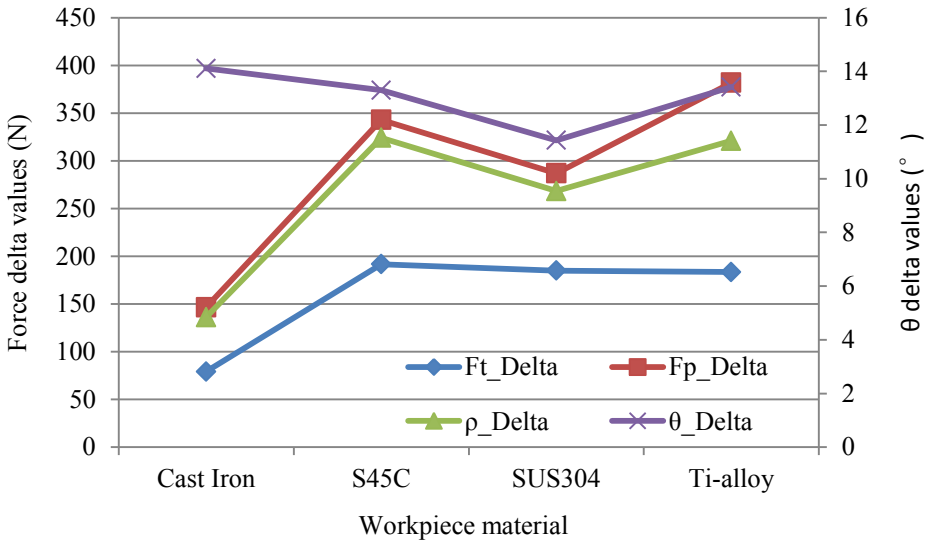

Figure 4-3 The relations between the workpiece material and the delta features

In Figure 4-3, Ft_Delta, Fp_Delta, ρ_Delta and θ_Delta mean ΔFt, ΔFp, $\Delta \rho$. $\Delta \theta$ respectively. The feature values for S45C, SUS304 and Ti-alloy are close and much larger than those of the cast iron. And there is a drop of the delta features for SUS304 which can be noticed.

4.1.2.2 Standard deviation features

Figure 4-4 illustrates the relations between the workpiece materials and the standard deviation values of the equivalent thrust force, principal force and their resultant force.

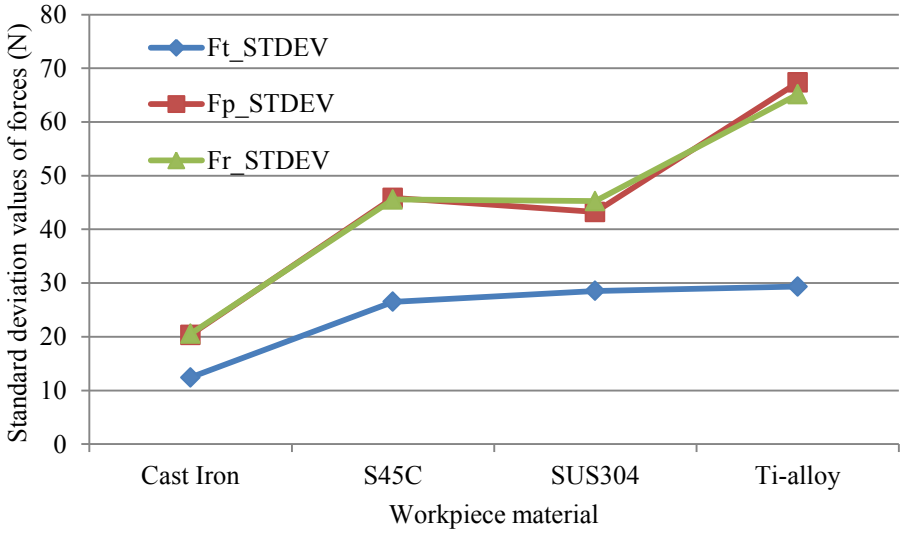

Figure 4-4 The relations between the workpiece material and the standard deviation features

It can be seen that the standard deviation values of the principal force and the resultant force is

larger than those of the thrust force, which implies the principal force affords the major deviation in the resultant force.

4.1.2.3 Velocity features

The relations between the workpiece materials and the velocity features are shown in Figure 4-5.

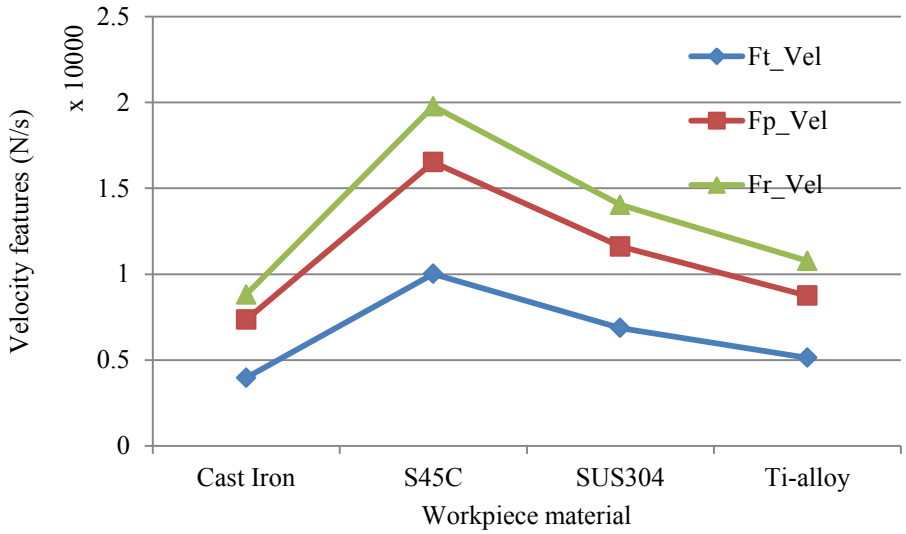

Figure 4-5 The relations between the workpiece material and the velocity features

The velocity features show a quite difference to other features, as it can be seen, the velocity features for S45C to Ti-alloy decrease instead of increase, which means that a big variation is not the determinant of a big velocity feature. The data for comparison is shown in Figure 4-6. Here the sampling frequency is set as 2000 Hz to downsize the data volume.

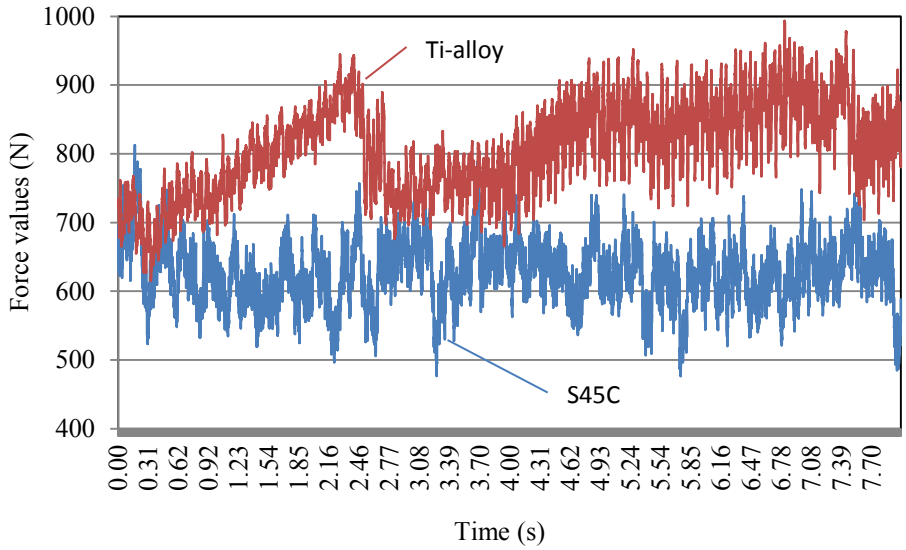

Figure 4-6 The sampled *Fp* values for S45C and Ti-alloy

The upper value line is the sampled data of the *Fp* with Ti-alloy while the lower one is that of S45C in Figure 4-6, and the features are obtained from this data in this section. It can be observed that the upper line has a larger variation range than the lower one, which implies larger delta features, and also

as discussed before, the RMS and standard deviation of the Ti-alloy data are larger than S45C. According to the definition of the velocity feature of *Fp*, it is calculated by the quotient of a certain value and the sampling time. The sampling time here is 8 seconds and the certain value is defined as the sum of the abstract values of the difference between the current sampled value and the next sampled value. If the difference between the adjacent two data points is very large for the whole sampled data sequences, it means that the fluctuation frequency is near the sampling frequency. And furthermore, if the data has a larger fluctuation magnitude for the lower frequencies, then its velocity feature will be smaller, and vice versa.

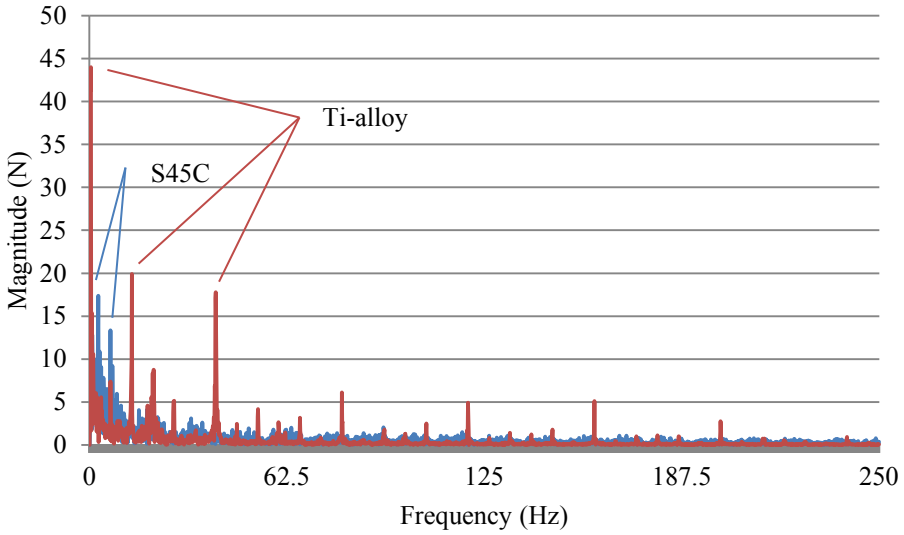

Figure 4-7 The frequency domain of the *Fp* for S45C and Ti-alloy in frequency range (0-250Hz)

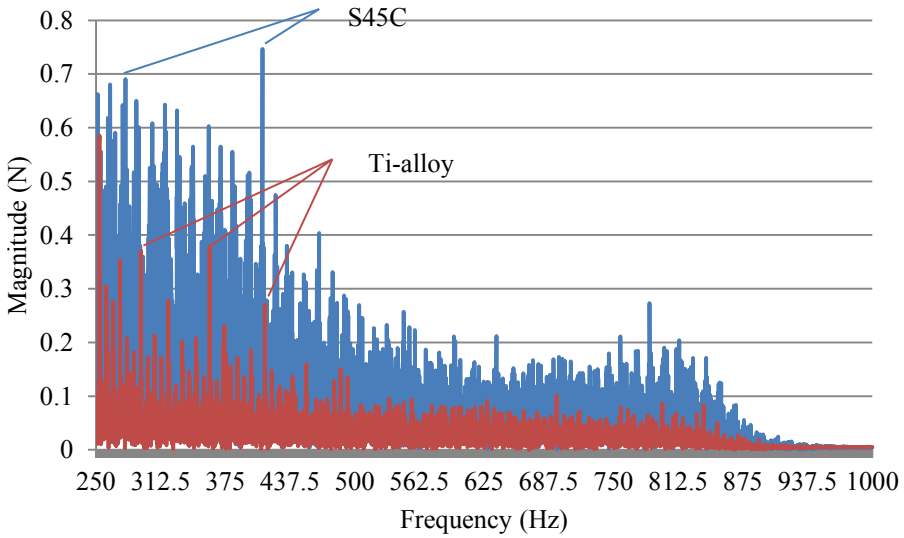

Figure 4-8 The frequency domain of the *Fp* for S45C and Ti-alloy in frequency range (250-1000Hz)

The frequency domain plots of *Fp* for S45C and Ti-alloy are shown in Figure 4-7 and Figure 4-8, where the former one contains the data in frequency range of 0-250 Hz while the later one illustrates that in frequency range of 250-1000 Hz. It can be observed that in the lower frequency range, the data has higher magnitudes for Ti-alloy, and in the higher frequency range the data has higher magnitudes

for S45C. Thus it is known that the velocity features involve the frequency information of the data.

4.1.2.4 The Fr_ConvHullArea

Figure 4-9 demonstrates the relation between the workpiece material and the geometry feature, i.e., *Fr_ConvHullArea.*

Figure 4-9 The relation between the workpiece material and the geometry feature

Similar to the delta features, there is a decrease of the value for SUS304. Like the velocity features, the convex hull area of the scatter points is also not decided by the statistical properties of the data. Data with high variation can have a low convex hull area for different distribution geometry.

4.1.3 Analysis

A radar chart method was proposed in [85] to calculate the Difficult-to-Cut Rating (DTCR). The mechanical properties, tensile strength, hardness and elongation, the thermal properties, thermal conductivity and specific heat capacity, and the physical property, density, were taken into account in plotting the radar chart. The radar chart is composed of four axes, which indicate the hardness (*H*), the tensile strength (*T.S.*), the elongation (*El*) and the thermal parameter (*T.P.*). The thermal parameter is defined in Equation 4-1.

$$T.P. = (K\rho c)^{-0.5} \qquad \text{Equation 4-1}$$

In Equation 4-1, K means the thermal conductivity, ρ indicates the density and c represents the specific heat capacity.

The cast iron used in this research is FC-250, and the Ti-alloy used is Ti-6Al-4V. Relevant properties of different materials are shown in Table 4-2. And by the same method, the DCTR value 0.33 for cast iron can be calculated. Therefore, the sequence for these four materials in the increase of difficulty is cast iron, S45C, SUS304 and Ti-alloy.

Table 4-2 Relevant material properties

Materials / Properties	Cast iron	S45C	SUS304	Ti-alloy
H	180	200	150	300
T.S. (MPa)	250	600	630	1000
El (%)	1	20	40	12
T.P.	0.00206	0.00215	0.00419	0.00789
DTCR	0.33	1.00	2.06	2.80

The feature values generally increase as workpiece material changes from cast iron to Ti-alloy. However, for some features, the value change patterns are different, which are determined by all of the properties in a comprehensive way. The radar chart of the properties for the four materials is shown in Figure 4-10, which shows the differences of properties more clearly.

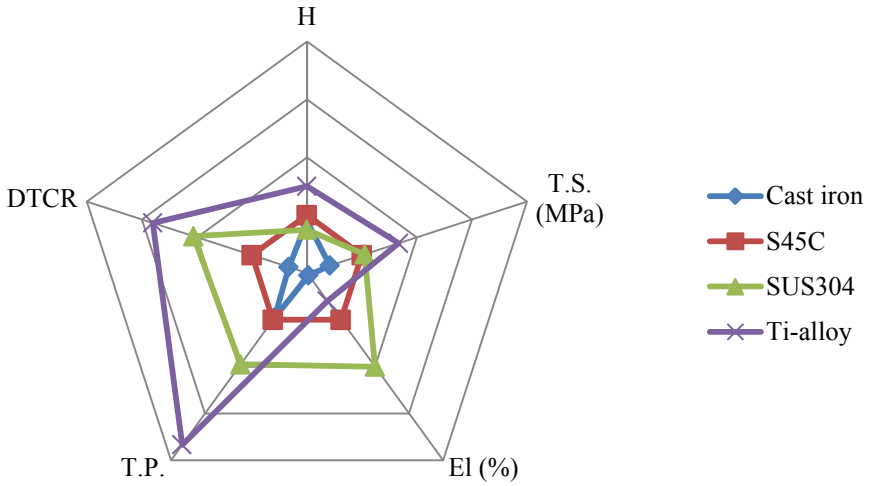

Figure 4-10 Radar chart of the material properties

4.2 Drill Diameter

Drills with 4 different diameters were used to obtain data under conditions shown in Table 4-3.

Table 4-3 Experimental Conditions for different drill diameters

Spindle Speed	800 rpm
Feed Rate	0.08 mm/rev.
Drill Bit Diameter	7, 8, 9, 10 mm
Workpiece Material	S45C
Workpiece Thickness	15 mm
Lubrication	Dry

4.2.1 Static Features

4.2.1.1 Average features

As shown in Figure 4-11, $\overline{Fp}$ and $\bar{\rho}$ increase as the drill diameter increases. $\overline{Ft}$ for diameter 7,8,9 are similar and for 10 is much larger. The average features of the resultant force $\bar{\theta}$ decreases for diameter 7,8,9 but increases for 10.

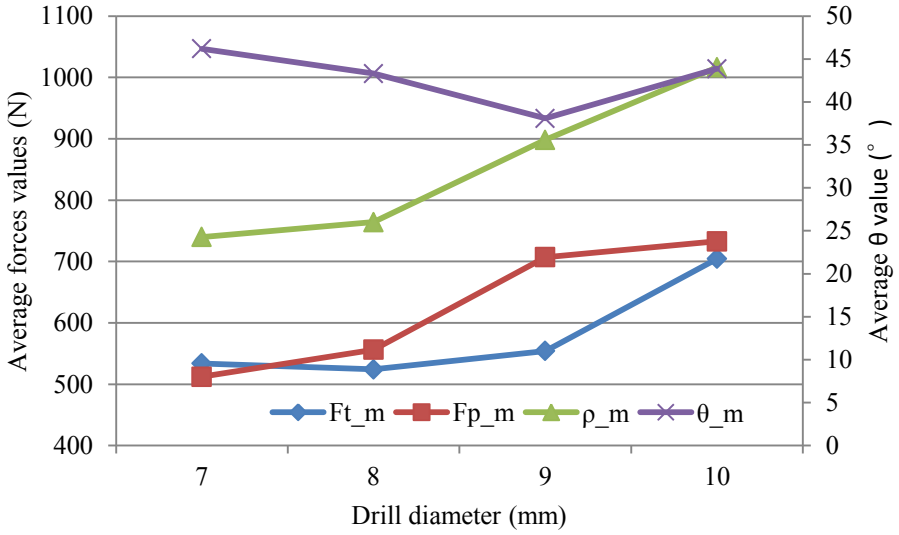

Figure 4-11 The relations between drill diameter and the average features

The increase of the average values of forces, $\overline{Ft}$, $\overline{Fp}$ and $\bar{\rho}$, is due to the increase of drill diameter, which needs more power to make. The $\bar{\theta}$ has a decrease trend as the drill diameter increases from 7-9 mm, but increases from 9-10 mm, which is caused by the significant increase of $\overline{Ft}$.

4.2.1.2 RMS features

Similar changes can be found as average features for the RMS features shown in Figure 4-12.

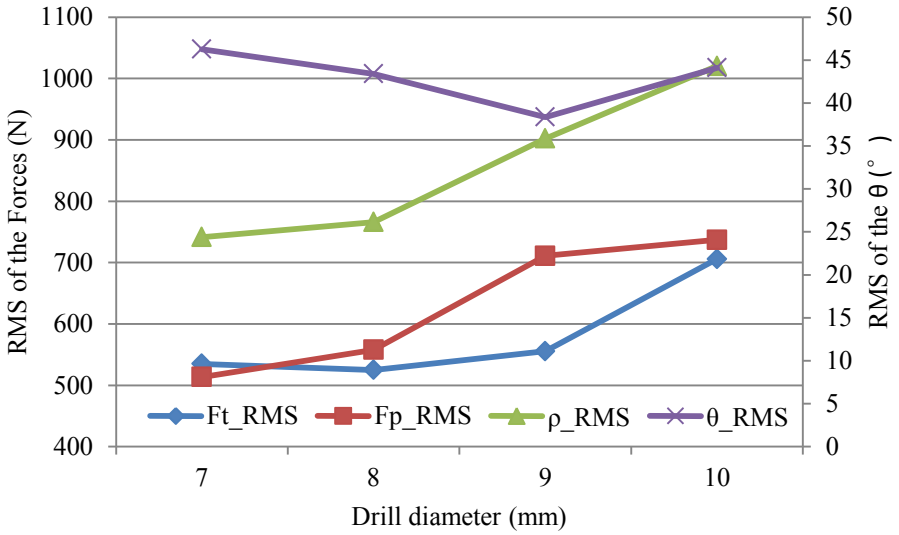

Figure 4-12 The relations between drill diameter and the RMS features

It shows that RMS features of forces increase as drill diameter increases. The increase of *Fp_RMS* is bigger than that of *Ft_RMS* during the drill diameter increases from 7-9 mm, and it is smaller from 9-10 mm, which causes the change of the changing trend for *θ_RMS*.

4.2.2 Dynamic Features

4.2.2.1 Delta Features

Figure 4-13 illustrates the relations between drill diameter and the delta features. It shows that *ΔFp, Δρ. Δθ* increase from 7-9 and decrease a little at 10 mm. *ΔFt* increases slightly from about 200N to around 300N.

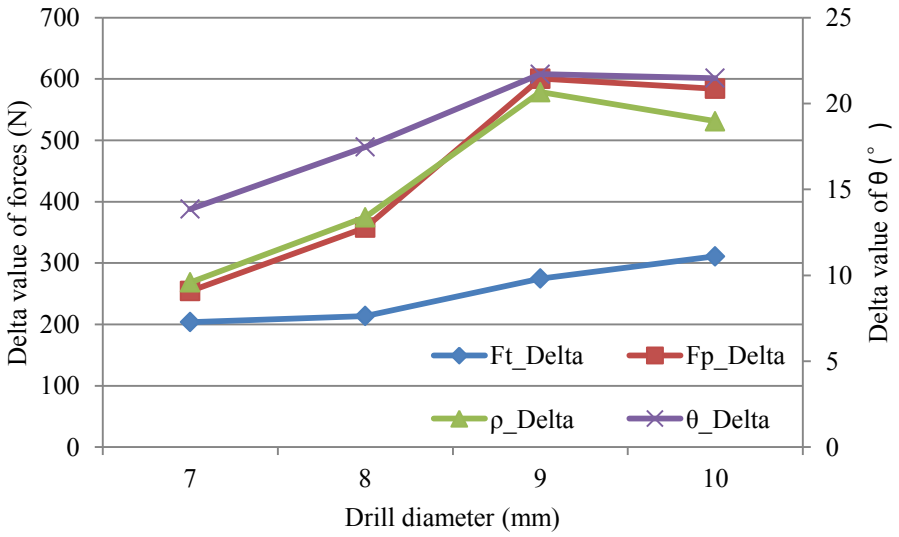

Figure 4-13 The relations between drill diameter and the delta features

4.2.2.2 Standard Deviation Features

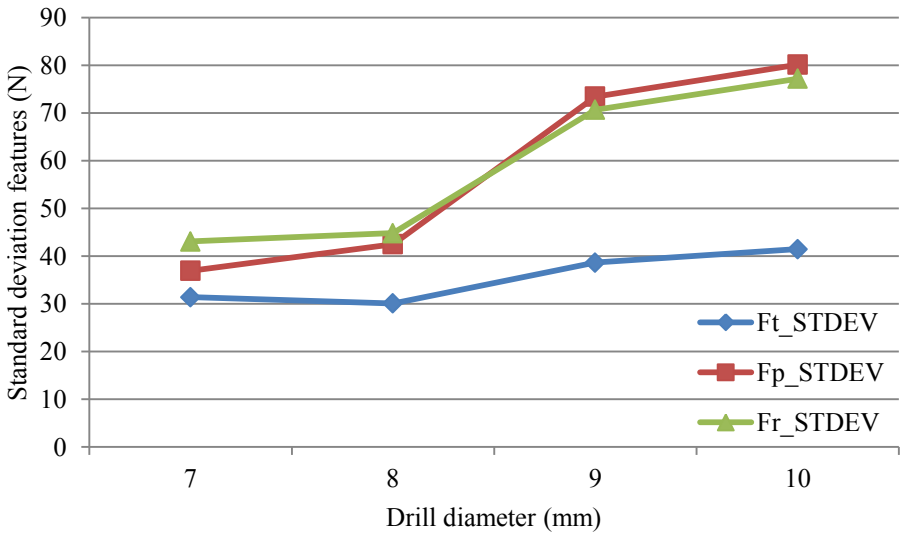

Figure 4-14 The relations between drill diameter and the standard deviation features

It can be found in Figure 4-14 that the standard deviation of *Fp* and *Fr* variation range is about 4

times of that of *Ft*, and it means that drill diameter effects *Fp* more than *Ft*.

4.2.2.3 Velocity Features

Figure 4-15 shows that the 3 velocity features all increase as the drill diameter increases. And the increase between 9 and 10 is much more than that between 7, 8 and 9.

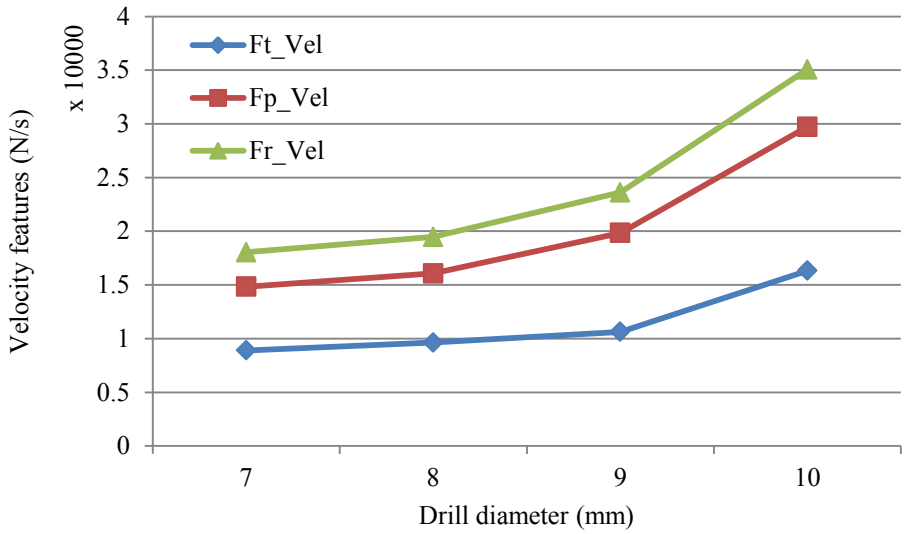

Figure 4-15 The relations between drill diameter and the velocity features

It can be found that the velocity features of the forces are affected by the drill diameter in a same manner, and thus they can be used to indicate the drill diameter with high accuracy.

4.2.2.4 The Fr_ConvHullArea

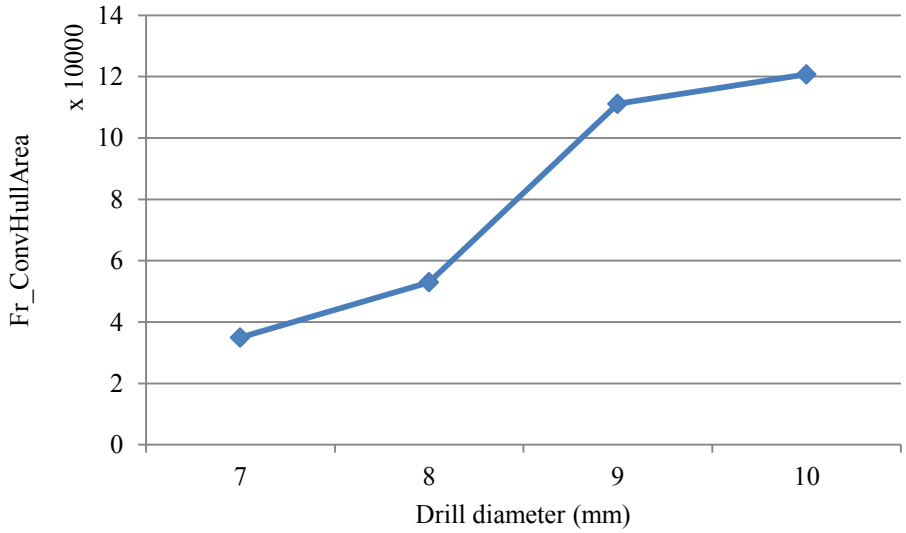

Figure 4-16 The relation between drill diameter and the *Fr_ConvHullArea*

Drill diameter has a strong effect to the *Fr_ConvHullArea* as shown in Figure 4-16. The *Fr_ConvHullArea* increases 4 times from near 30000 at 7 mm to around 120000 at 10 mm. And differing from the relations between other features, the increase between 8 mm and 9 mm is larger

which implying a good supplement for the indicating of drill diameter.

4.3 Spindle Speed

Table 4-4 shows the conditions for experiments with different spindle speed. Because HSS drill is employed in this study, the range of spindle speed is not very wide.

Table 4-4 Conditions for experiments with different spindle speed

Spindle Speed	400, 600, 800, 1000 rpm
Feed Rate	0.2 mm/rev.
Drill Bit Diameter	8 mm
Workpiece Material	S45C
Workpiece Thickness	15 mm
Lubrication	Dry

4.3.1 Static Features

4.3.1.1 Average features

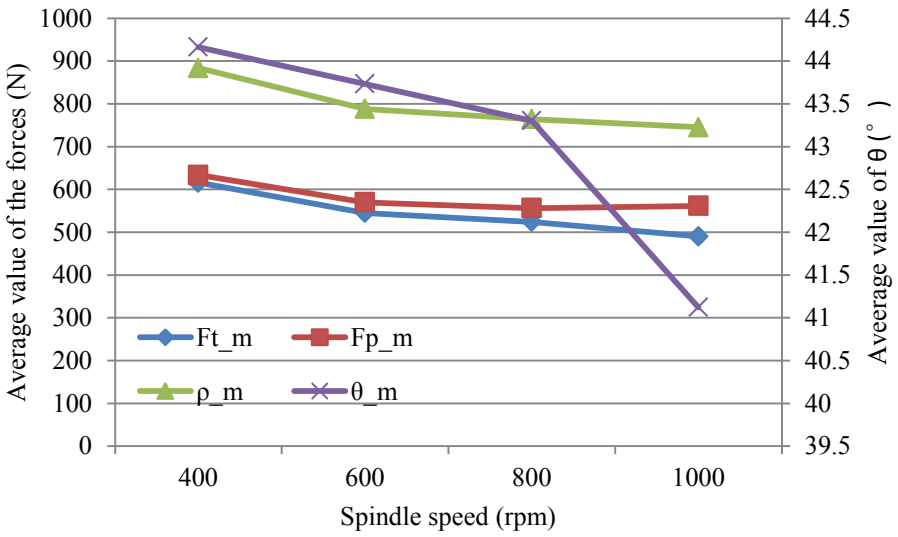

Figure 4-17 The relations between spindle speed and the average features

As shown in Figure 4-17, the average features decrease as spindle speed increases, this is due to more heat is generate at higher spindle speed which soften the workpiece.

4.3.1.2 RMS Features

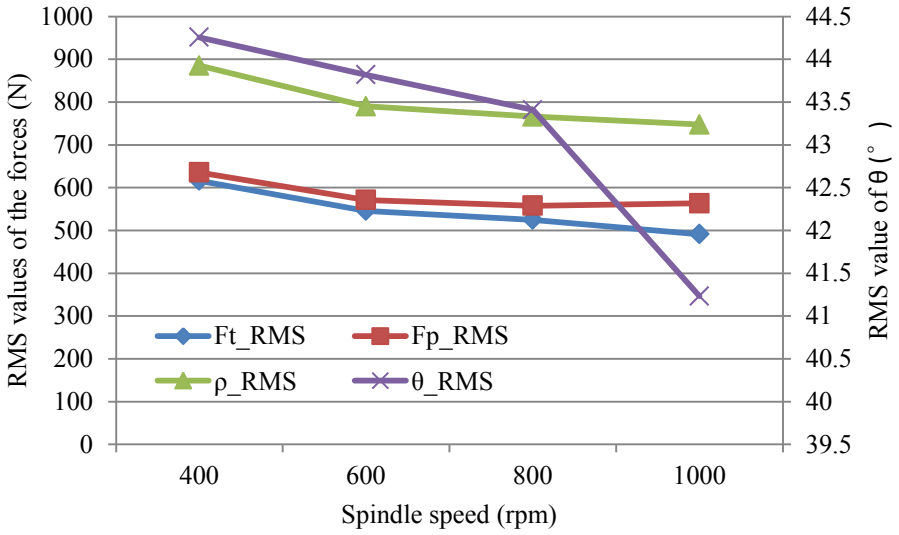

Figure 4-18 The relations between spindle speed and the RMS features

The spindle speed also has little influence to the RMS features as demonstrated in Figure 4-18. The RMS values of forces change in a range of 200N and the RMS value of θ changes about 3°.

4.3.2 Dynamic Features

4.3.2.1 Delta Features

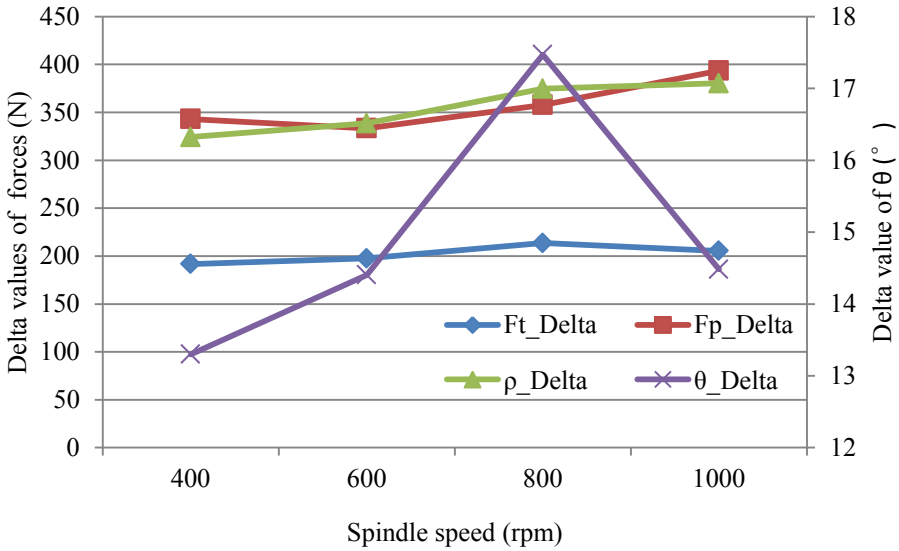

Figure 4-19 The relations between spindle speed and the delta features

The spindle speed has little effect to the delta features as shown in Figure 4-19. The *Ft_Delta* almost remains the same and, *Fp_Delta* and *ρ_Delta* vary within the range of 100N. *θ_Delta* changes slightly from about 13° to 18°.

4.3.2.2 Standard Deviation Features

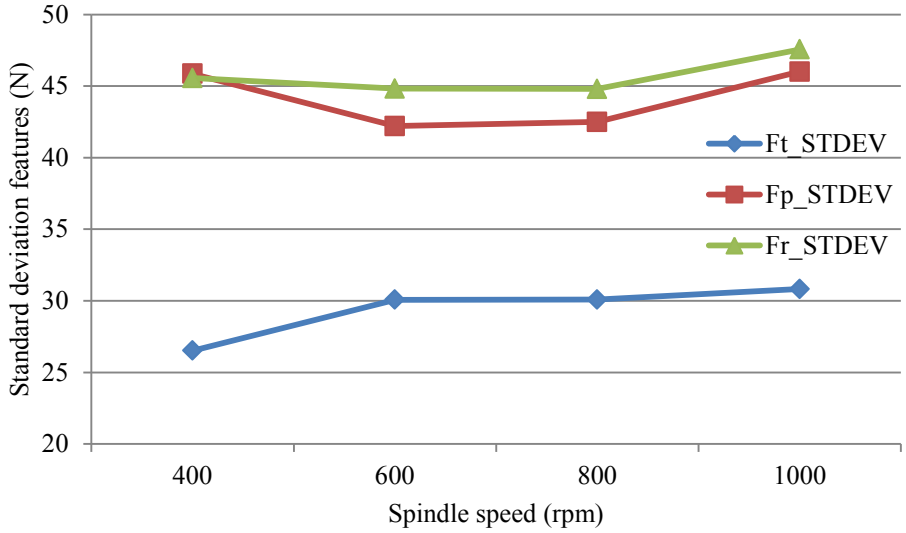

Figure 4-20 The relations between spindle speed and the standard deviation features

Little changes of the standard deviation features have also been found in Figure 4-20, where *Ft_STDEV*, *Fp_STDEV* and *Fr_STDEV* change about 5N.

4.3.2.3 Velocity Features

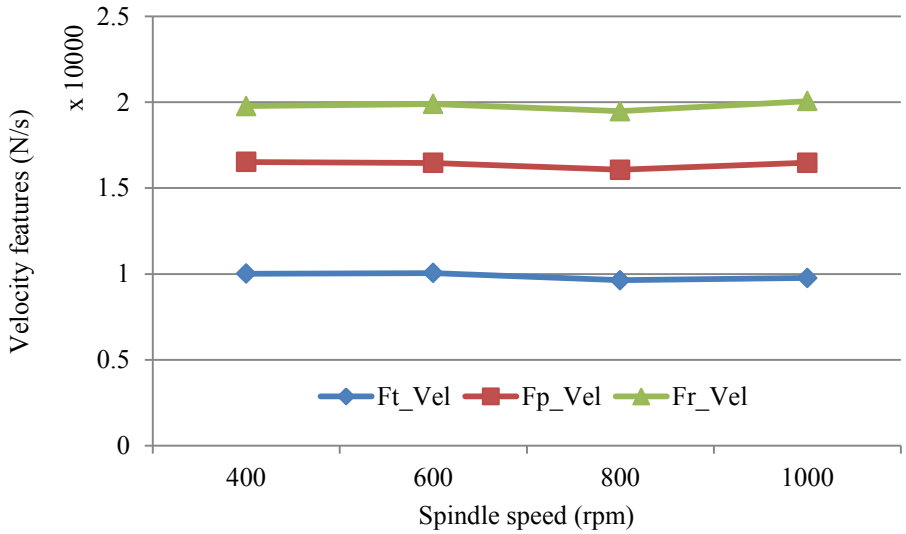

Figure 4-21 The relations between spindle speed and the velocity features

The velocity features are almost the same in spite of the increase of spindle speed as illustrated in Figure 4-21.

4.3.2.4 The Fr_ConvHullArea

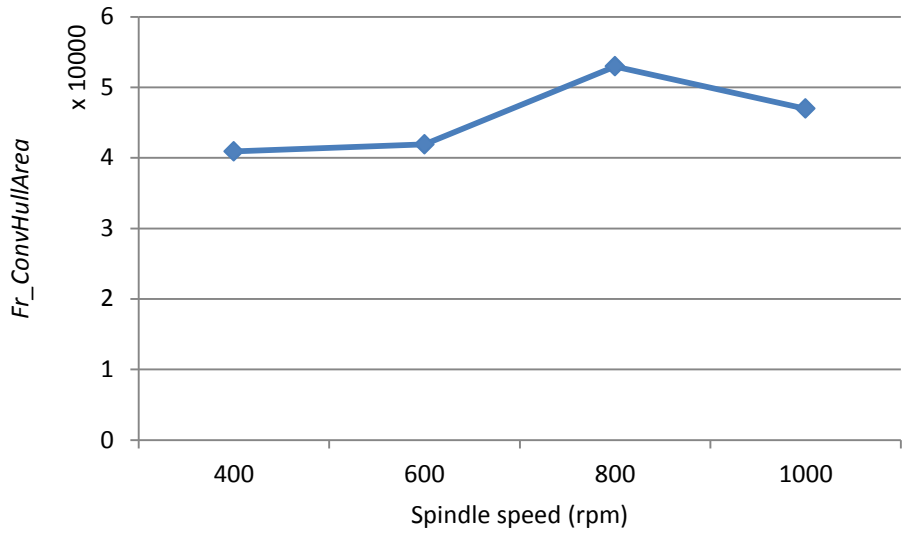

Figure 4-22 The relations between spindle speed and the *Fr_ConvHullArea*

Finally, for the *Fr_ConvHullArea*, again no much difference is found for different spindle speed as shown in Figure 4-22.

4.4 Feed Rate

Table 4-5 shows the conditions for experiments with different feed rate.

Figure 4-23 illustrates the rectangular coordinate plot of *Ft* to *Fp* with different feed rate.

Table 4-5 Conditions for experiments with different feed rate

Spindle Speed	745 rpm
Feed Rate	0.05, 0.1, 0.2, 0.3, 0.4 mm/rev.
Drill Bit Diameter	8 mm
Workpiece Material	S45C
Workpiece Thickness	15 mm
Lubrication	Dry

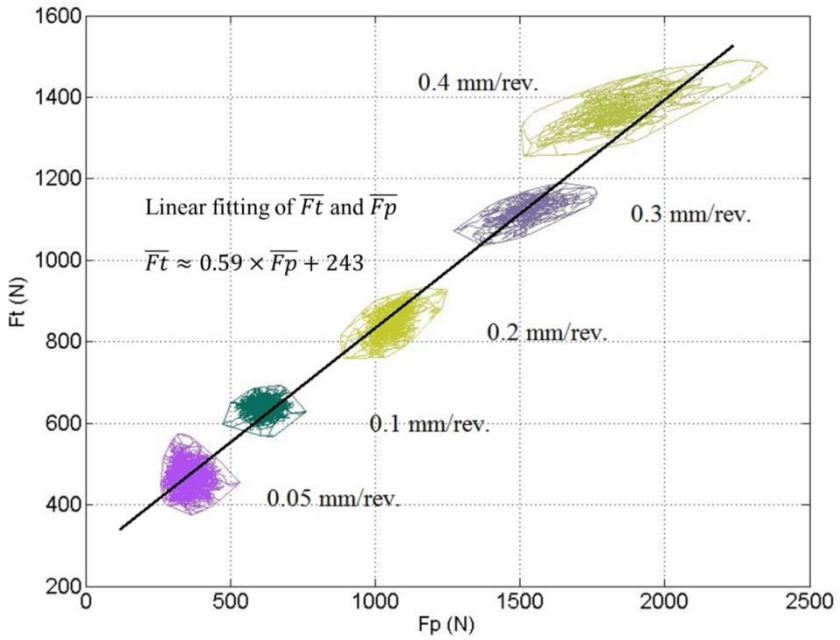

Figure 4-23 The rectangular coordinate plot of *Ft* to *Fp* with different feed rate

It is a very clear distribution of the resultant force trajectories as plotted in Figure 4-23. It shows that along with the increase of feed rate, both the position and distribution of the trajectories change. And it's also been noticed that the center points of trajectories of the resultant force keep in line, indicating a certain ratio between $\overline{Ft}$ and $\overline{Fp}$. The ratio value 0.59 can be calculated by linear fitting under conditions listed in Table 4-5.

4.4.1 Static Features

4.4.1.1 Average features

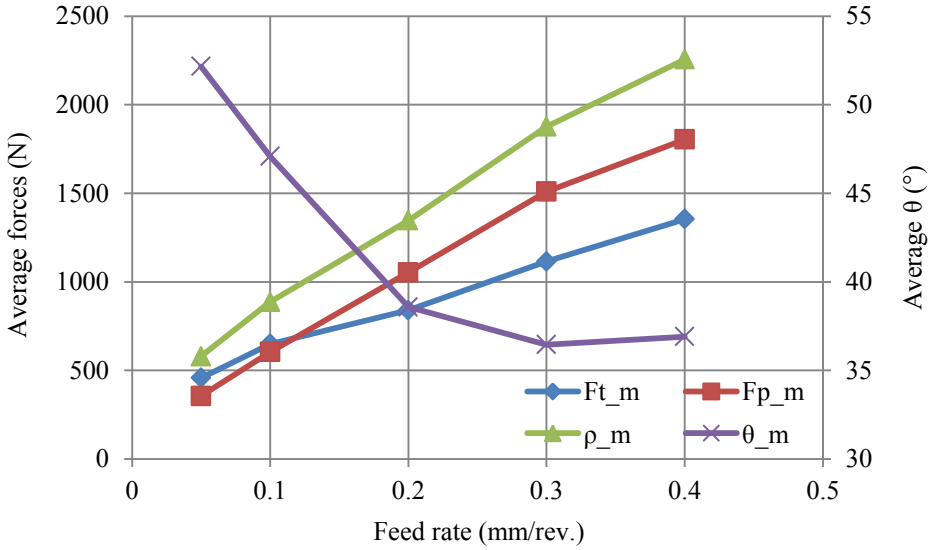

Figure 4-24 The relations between feed rate and the average features

It shows that the average values of forces increase linearly as feed rate increases, and the average θ decreases which means the increase of $\overline{Fp}$ is more than that of $\overline{Ft}$.

4.4.1.2 RMS Features

Figure 4-25 shows the relation between the RMS features and feed rate.

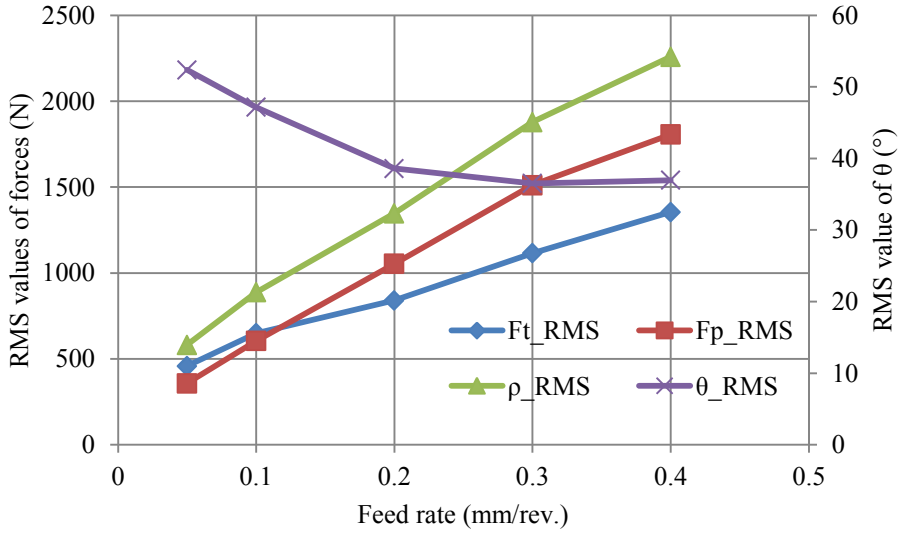

Figure 4-25 The relations between feed rate and the RMS features

Similar to the average features, the RMS values of the forces increase linearly and the RMS of θ decrease.

4.4.2 Dynamic Features

4.4.2.1 Delta Features

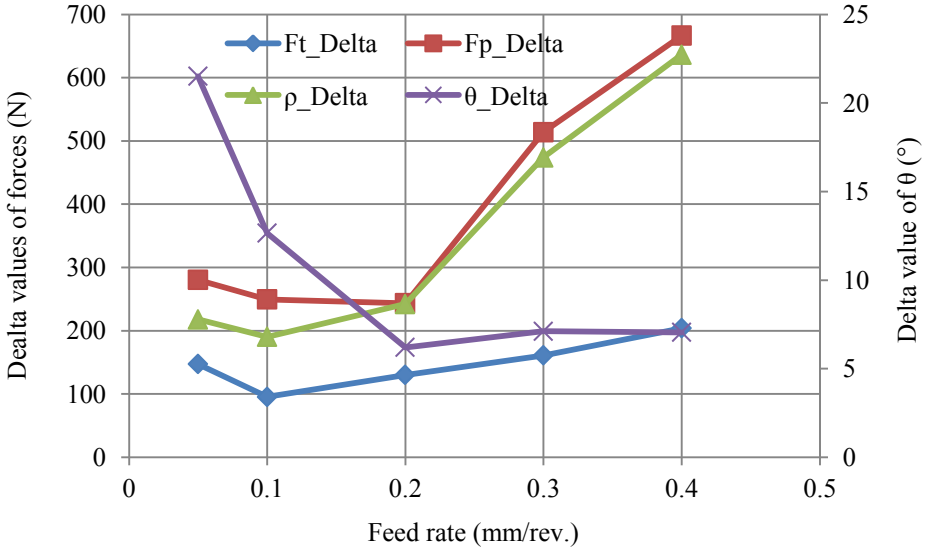

Figure 4-26 The relations between feed rate and the delta features

From Figure 4-26 it can be observed that ΔFt varies about 100N for different feed rate, while ΔFp is much more sensitive, which varies from about 200N to about 700N. And consequently, the $\Delta\rho$ and $\Delta\theta$ also change violently.

It also can be noticed that the ΔFp and $\Delta\rho$ values are very close when feed rate is 0.05, 0.1 and 0.2

mm/rev., which means there is a limitation of delta features to indicate the feed rate while it is very low.

4.4.2.2 Standard Deviation Features

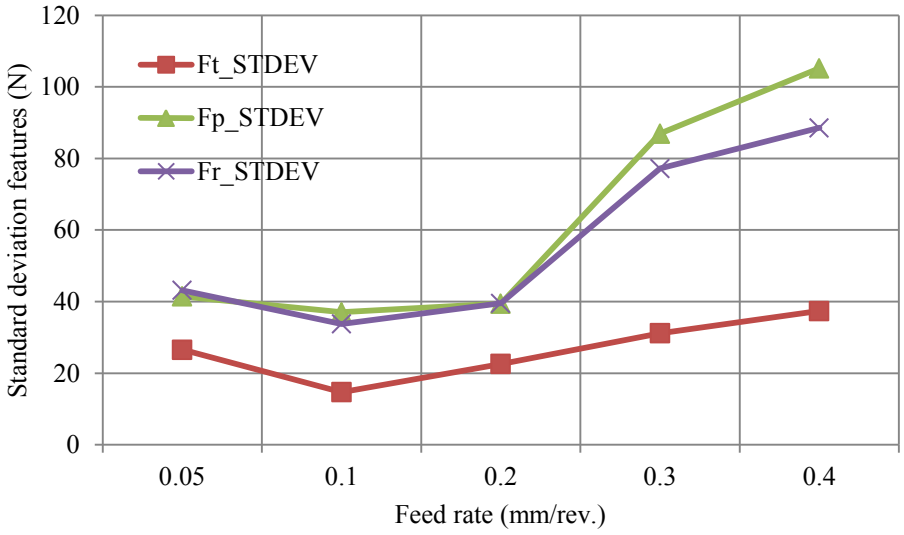

Figure 4-27 The relations between feed rate and the standard deviation features

Similar to delta features, the standard deviation of *Ft* changes a little while those of *Ft* and *Fr* change a lot from 0.2 mm/rev. to 0.4 mm/rev.

4.4.2.3 Velocity Features

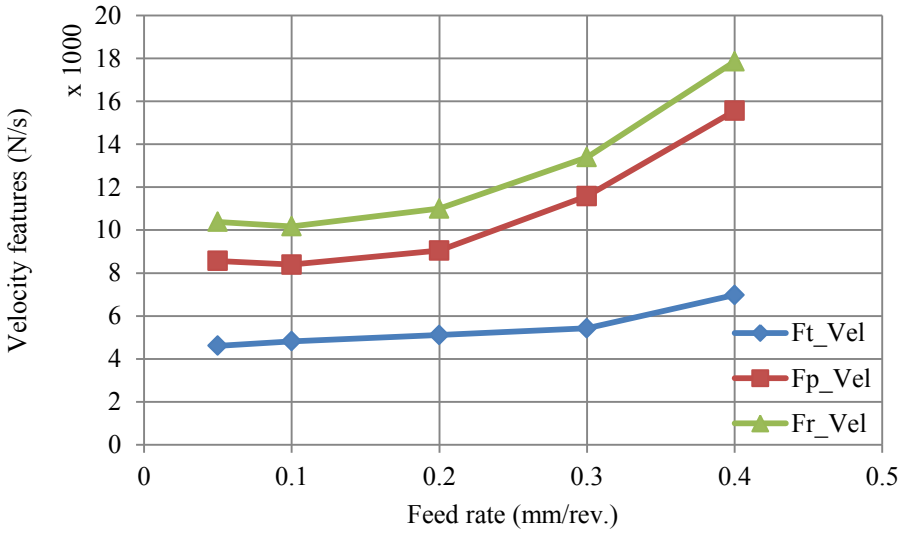

Figure 4-28 The relations between feed rate and the velocity features

The lower feed rate cannot affect the velocity features very much as shown in Figure 4-28. *Fp_Vel* and *Fr_Vel* change a little with feed rate from 0.05 to 0.2 mm/rev., however they increase over 60% when feed rate increases from 0.2 to 0.4 mm/rev.

4.4.2.4 The Fr_ConvHullArea

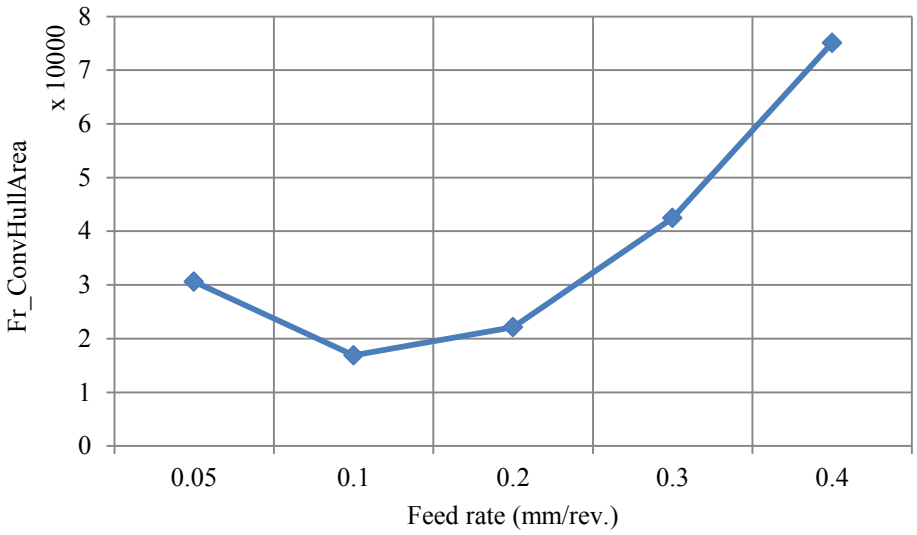

Figure 4-29 The relations between feed rate and the *Fr_ConvHullArea*

Similar result of the effect from feed rate to *Fr_ConvHullArea* can be found in Figure 4-29. The difference between *ConvHullArea* with feed rate from 0.05 to 0.2 mm/rev. is low but with feed rate from 0.2 to 0.4 mm/rev. is very high, where the value when feed rate is 0.4 mm/rev is almost four times. of the value when feed rate is 0.2 mm/rev.

4.4.3 Investigation of the Effect of Feed Rate

In the above results, the average features of the forces show a good correlation to the feed rate in the whole range, while the dynamic features, such as delta features, standard deviation features, do not have a good correlation to lower feed rate (0.05-0.2 mm/rev. here). The reason is probably the stability of the drilling processes is limited with lower feed rates.

To investigate this phenomenon, new parameters concerned with the cutting status are necessary.

During a drilling operation, the chips are formed along the cutting lip and moved up following the drill helix angle. The drill geometry has a complicated effect on the cutting forces [86].

Tarng and Li [87], [88] developed a simple method to predict chatter limit in drilling. In this model, the uncut chip thickness was decomposed into two components: a mean value and a variable chip thickness. The chip thickness was the input to the cutting process (described by a transfer function) yielding the thrust force. The thrust force excited the machine tool (also described by a transfer function), which causes vibration affecting the chip thickness.

A 3D chip formation model [89] was presented to describe the interaction between the cutter and the workpiece. The model predicted the dynamic forces and chatter limit well.

Therefore, chip thickness is measured to investigate the drilling status.

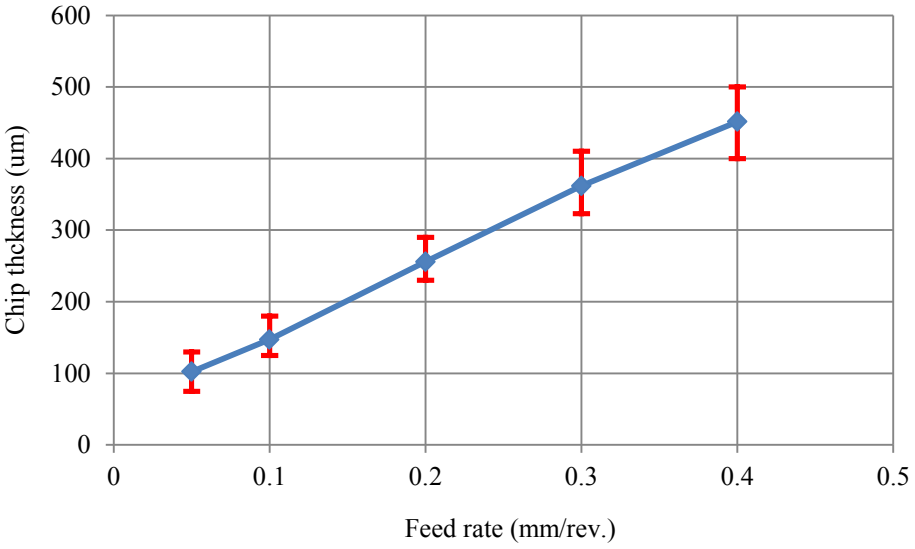

Figure 4-30 Average and variation of the chip thickness with different feed rate

Figure 4-30 shows the average and variation of the chip thickness with different feed rate. The rhombus markers mean the average values of the chip thickness and the error bars represent their variation ranges. It can be observed that there is a linear relation exiting between the feed rate and chip thickness; however, the variation ranges of the chip thickness do not increase linearly with the feed rate.

The variations range of the chip thickness for feed rate from 0.05 to 0.4 mm/rev. are 55, 55, 60, 87, and 100 um respectively. When the feed rate is within 0.05-0.2 mm/rev, the chip thickness variation ranges are approximate, which means the drilling status with feed rate at 0.05 and 0.1 is not very good and therefore result in similar dynamic features.

An unstable drilling process would also lead to bad hole quality, due to the unexpected scratch to the hole sidewall by the anomalous chips. Thus the hole sidewall roughness is measured to reflect the drilling status.

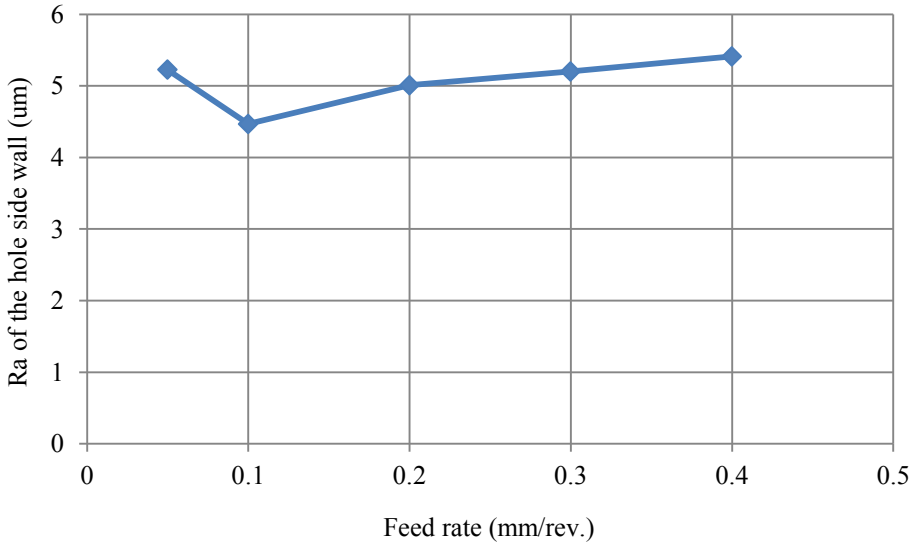

Figure 4-31 Hole side wall roughness for different feed rate

As shown in Figure 4-31, the hole sidewall roughness for the one with 0.05 mm/rev. feed rate is

higher than the contiguous ones, which can effectively explain the high dynamic features for lower feed rates.

4.5 Drill Wear

Wear is another major factor of the instability of a cutting process. Worn tools generally cause an increase in the static as well as dynamic components of the cutting force components, because of increased friction[90].

To predict drill corner wear, the extracted features should keep pace with the corner wear during the drill life. A life test has been conducted under conditions listed in Table 4-6.

Table 4-6 Experimental conditions for evaluation of drill corner wear to features

Spindle Speed	745 rpm
Feed Rate	0.15 mm/rev.
Drill Bit Diameter	8 mm
Workpiece Material	S45C
Workpiece Thickness	20 mm
Lubrication	Dry

Since the value of corner wear is small and the feature values are very high and unstable, thus plots with the features as vertical axis and corner wear as horizontal axis are very un-clear at revealing the relations between them. Therefore, the accumulated total drill depth is applied as a scale to measure both the corner wear and features. In the following figures, the horizontal axis means the accumulated drill depth, the left vertical axis stands for the normalized features and the right vertical axis represents the measured corner wear.

The experiments were carried out in the following way: during the drill bit's life, drilling of holes under the above listed conditions is repeated, and for drilling of one single hole, the drill depth is 20 mm, once the accumulated total drill depth reached a certain value, the real corner was measured and the thrust and torque were sampled for feature extraction for this very drilling process.

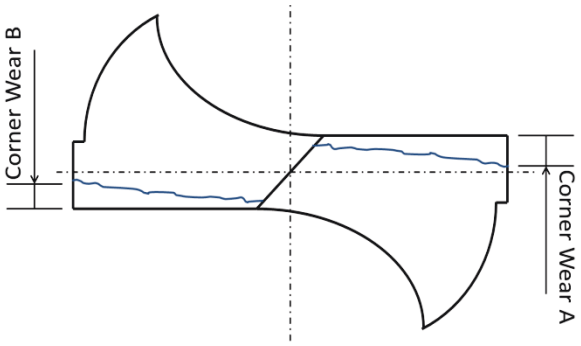

Figure 4-32 The definition of drill corner wear

As shown in Figure 4-32, the corner wear is defined as the average value of corner wear A and corner wear B which means the wear on each outer margin of the drill edge.

4.5.1 Static Features

4.5.1.1 Average features

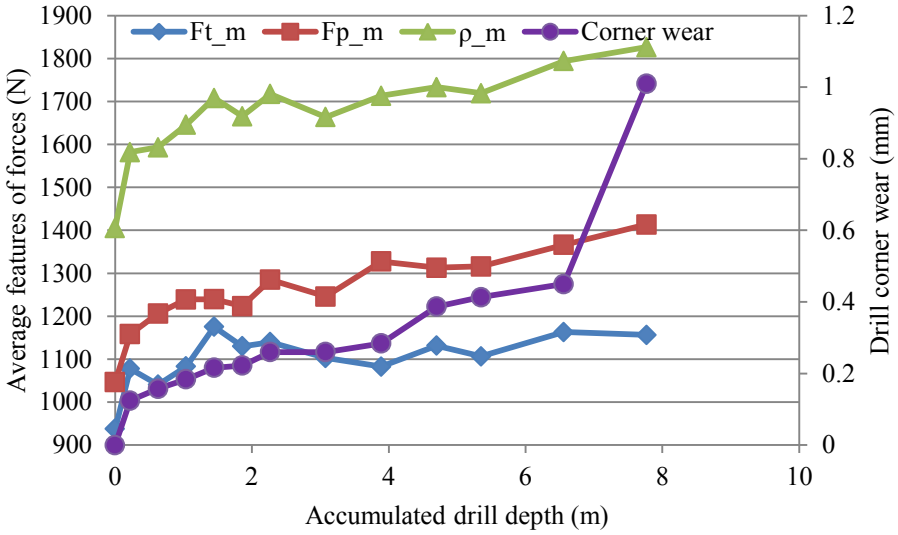

Figure 4-33 The relations between drill corner wear and average features of forces

As shown in Figure 4-33, as drill corner wear increase, the average features of the forces have increase trends, and as shown in Figure 4-34, the average value of θ decreases. However, there are big irregular values especially for *Ft*. The average *Ft* reaches the maximum value when the accumulated drill depth is 1.45m, where it is far from the maximum drill corner wear. Moreover, along with the increase of drill corner wear, the average features have increase about 25%, 35%, 30% for forces and decreases about 10% for the θ. Therefore, the corner wear does not have a strong effect to average features.

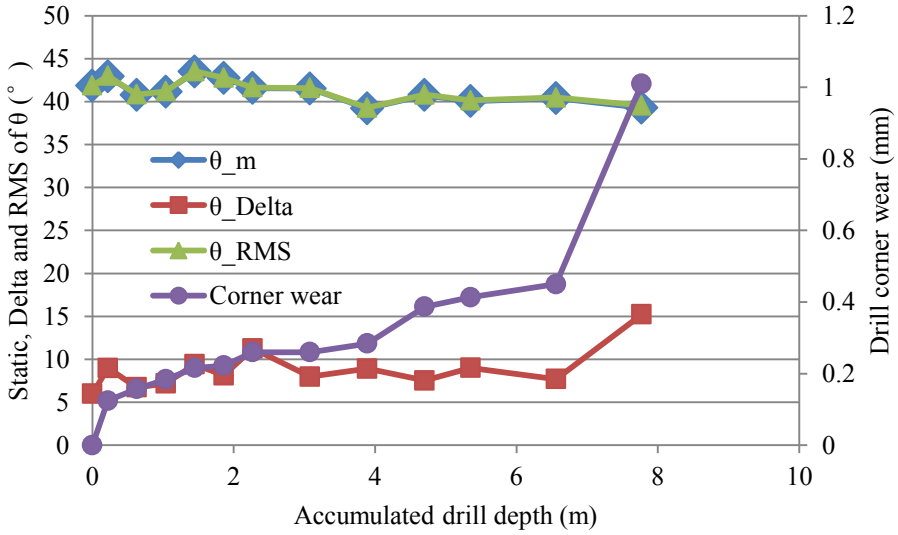

Figure 4-34 The relations between drill corner wear and static, delta and RMS features of θ

4.5.1.2 RMS Features

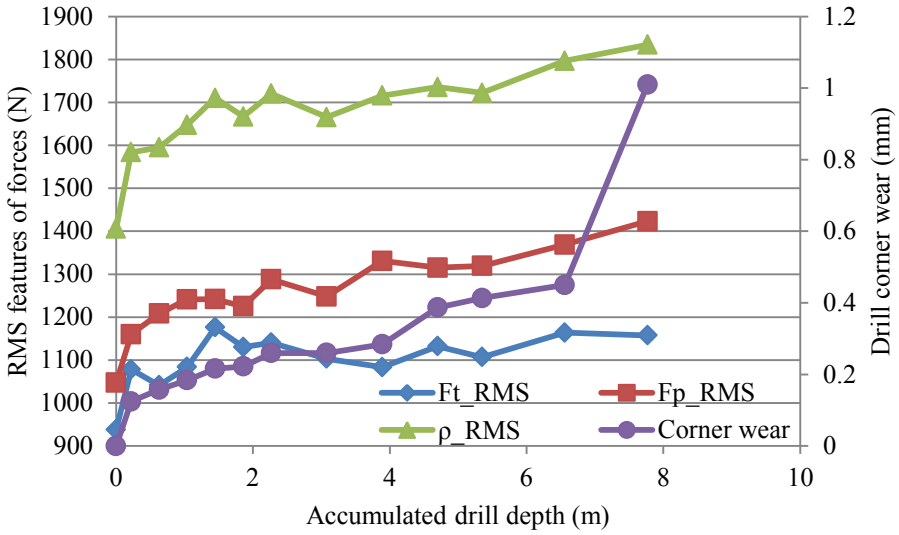

Figure 4-35 The relations between drill corner wear and RMS features of forces

The changing patterns of the RMS features are similar to that of average features as shown in Figure 4-35. The increase rates are low and there is no clear 3 stages in coincide with the drill corner wear, which are the initial state, the steady increase stage and the severe increase stage.

4.5.2 Dynamic Features

4.5.2.1 Delta Features

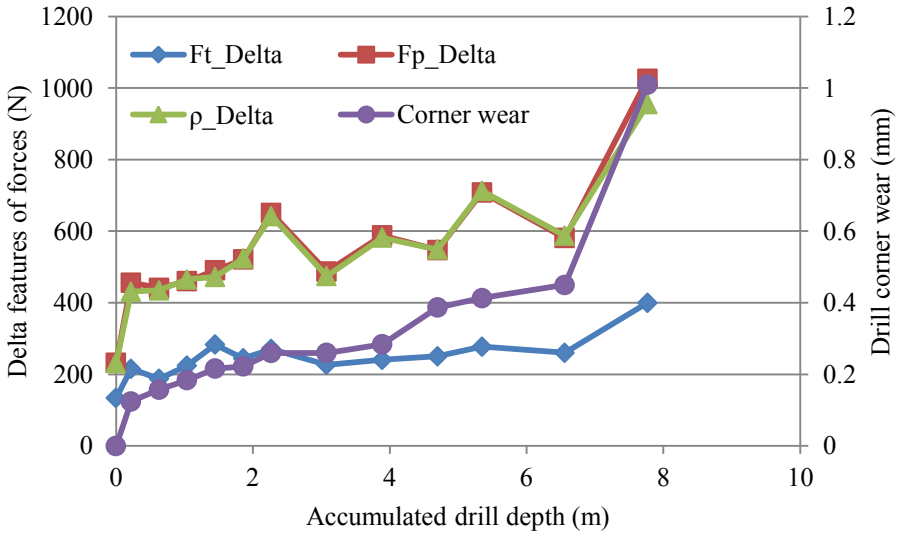

Figure 4-36 The relations between drill corner wear and delta features of forces

The delta features of forces increase as drill corner wear increases as shown in Figure 4-36. Especially for *Fp* and *Fr*, they increase 500% to the original value when the drill corner wear reaches 1.01 mm. and they show a typical 3 stages increase according to the 3 stages of wear state.

The delta feature of θ, that is *θ_Delta* shown in Figure 4-34, increases a little bit and the largest

change takes place at the last two instances where drill wear almost doubles.

4.5.2.2 Standard Deviation Features

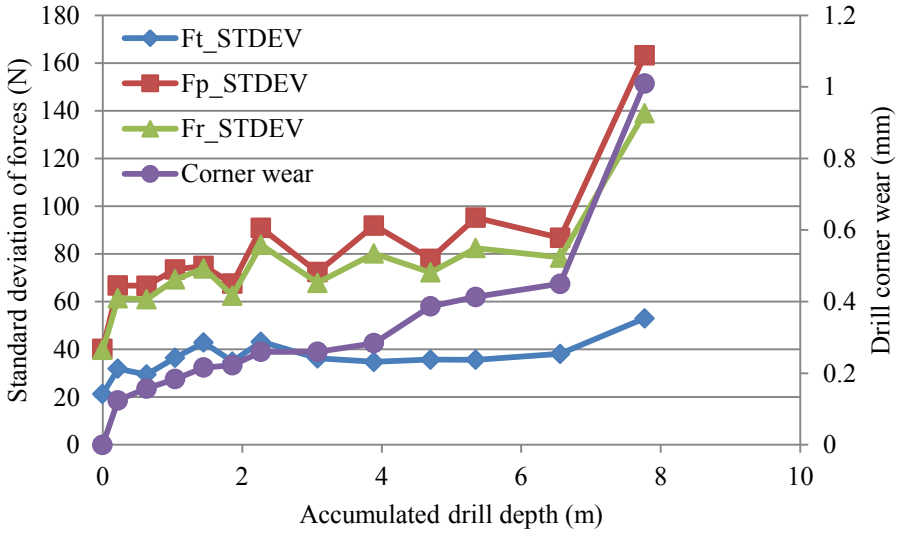

Figure 4-37 The relations between drill corner wear and standard deviation features of forces

The standard deviation features of *Ft* does not show a good relation to drill corner wear, however those of *Fp* and *Fr* increase in pace with the corner wear with a big increase rate and clear 3 stages just like the drill corner wear does.

4.5.2.3 Velocity Features

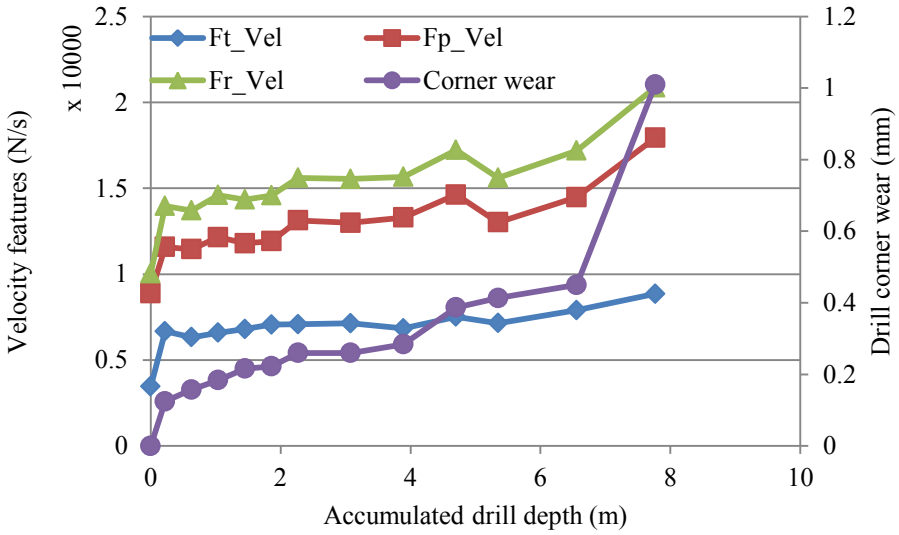

Figure 4-38 The relations between drill corner wear and velocity features of forces

The drill corner wear affects velocity features, but without great increases in the third stage as shown in Figure 4-38.

4.5.2.4 The Fr_ConvHullArea

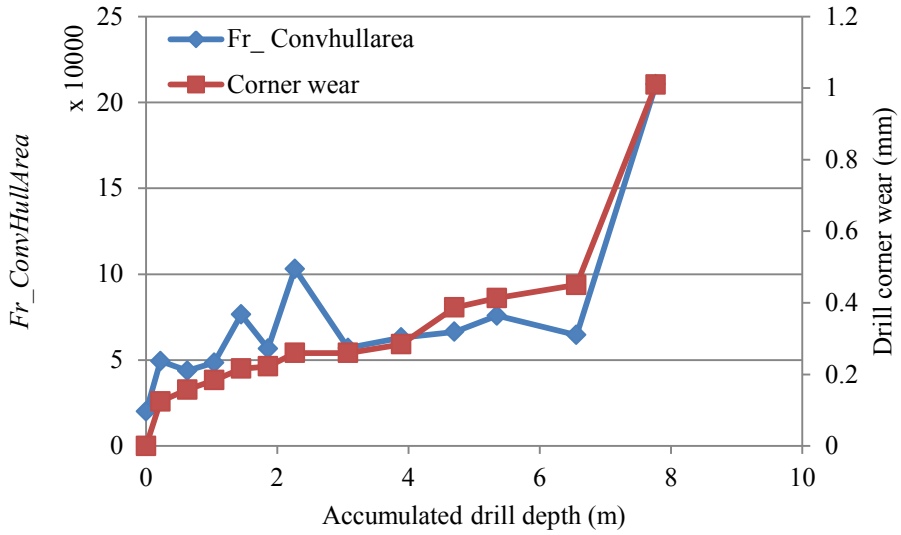

Figure 4-39 The relations between drill corner wear and *Fr_ConvHullArea*

The changing pattern of *Fr_ConvHullArea* stays close to that of drill corner wear as shown in Figure 4-39. However there exist some irregular values when the drill corner wear is small. Therefore, *Fr_ConvHullArea* and velocity features would be good complement for each other.

4.6 Summary

In this chapter, the relations between drilling conditions and the features are discussed, it can be found that different conditions have different effects to different features, and further the following summaries can be made.

As the workpiece getting more difficult to cut, the values of the features generally increase except for the velocity features. The reason is that the velocity features have a certain concern to the magnitude difference at various frequencies.

The bigger the drill diameter, the higher the feature values has been found, and commonly the features of *Fp* and *Fr* are more sensitive than *Ft*.

Spindle speed does not have much influence to the features comparing to other condition parameters, and a slight decrease of the cutting forces can be detected due to the increased heat generation speed which softens the workpiece.

The feed rate has significant effects to the features, and it not only changes the location but also the distribution patterns of the resultant force trajectories. It has been discovered that the average and RMS features increase linearly to the feed rate. And other features have clear distinguishing values when the feed rate is high but similar values when it is low. Further investigation has evinced that the cutting status is not so good which causes fluid chip thickness and low sidewall roughness.

Drill wear is another major cause of the variety of the features. Forces are getting higher when drill wear proceeds, and the dynamic features are more sensitive to the drill wear.

In the most circumstances, the change of drilling conditions leads to a more clear change of static features than dynamic features. Whilst the dynamic features are more sensitive to the status or the

stability of the drilling process, such as that discussed in 4.4.3 and the drill wear status presented in 4.5.2.

And for the static features, the average ones and the RMS ones are very similar in the values and the variation forms, thus, some of them can be eliminated for remove of redundancy.

Chapter 5 Drill Wear Prediction and Drilling Conditions Recognition

In this chapter, based on the methodologies discussed in Chapter 3 and the relations found in Chapter 4, the prediction of drill wear and the recognition of drilling condition parameters are carried out. The purpose of the work in this chapter is to predict the drill wear in a wide range of different drilling conditions by training the neural network with condition parameters and carefully selected features which are obtained from specially designed experiments.

5.1 Data Preparation

As discussed in Chapter 4, workpiece material, drill diameter, spindle speed and feed rate are all capable of making changes to the features, some effective at the static and some the dynamic. And these 4 condition parameters are the most various ones in the producing activities. Totally 256 instances of experimental patterns, under 4 different values for each of the 4 condition parameters, are conducted and investigated, to obtain the training data.

Table 5-1 Condition parameter values for 256 data patterns

Drill diameter (mm)	**Cast iron & S45C**		**SUS304 & Ti-alloy**	
	Feed rate (mm/rev.)	Spindle speed (rpm)	Feed rate (mm/rev.)	Spindle speed (rpm)
7	0.05	400,600,800,1000	0.06	200,300,400,500
	0.08	400,600,800,1000	0.08	200,300,400,500
	0.12	400,600,800,1000	0.10	200,300,400,500
	0.15	400,600,800,1000	0.12	200,300,400,500
8	0.05	400,600,800,1000	0.06	200,300,400,500
	0.08	400,600,800,1000	0.08	200,300,400,500
	0.12	400,600,800,1000	0.10	200,300,400,500
	0.15	400,600,800,1000	0.12	200,300,400,500
9	0.05	400,600,800,1000	0.06	200,300,400,500
	0.08	400,600,800,1000	0.08	200,300,400,500
	0.12	400,600,800,1000	0.10	200,300,400,500
	0.15	400,600,800,1000	0.12	200,300,400,500
10	0.05	400,600,800,1000	0.06	200,300,400,500
	0.08	400,600,800,1000	0.08	200,300,400,500
	0.12	400,600,800,1000	0.10	200,300,400,500
	0.15	400,600,800,1000	0.12	200,300,400,500

The detailed condition parameter values are shown in Table 5-1. Uncoated high speed steel (HSS)

drill bits with diameters of 7, 8, 9, 10 mm are used for all of the 4 workpiece materials. Experimental condition parameters for cast iron and S45C are the same and those for SUS304 and Ti-alloy are the same, to simplify the experiments within the drill bit endurance.

Drill corner wear is measured in the same way illustrated in 4.5, before and after the drilling test, and the average value is adopted as the feature of drill corner wear in the data base.

Workpieces are manufactured into suitable shape with the thickness of 20 mm.

As discussed above, 19 features are generated from the forces, which are converted from the sampled thrust and torque, and after the feature extraction by a 4-level-16-bands WPT decomposition and reconstruction, 304 monitored features are obtained. Together with the 4 condition parameters and the drill corner wear, 309 features for 256 instances constitute the knowledge data base.

5.2 Drill Wear prediction with Feature Selected by PCA

In this section, the feature selection procedure for all of the frequency bands is presented, and the evaluation of the influence of drilling conditions is also carried out.

5.2.1 Feature selection by PCA

It is not proper to use all of the 304 features to conduct PCA, because there is a limit of the observations-to-variables ratio (or subject-to-item ratio in some other circumstances) of PCA to get a better result.

It was found that[91], in general, minimum sample sizes appear to be smaller for higher levels of communality; minimum sample sizes appear to be smaller for higher ratios of the number of variables to the number of factors; and when the variables-to-factors ratio exceeds 6, the minimum sample size begins to stabilize regardless of the number of factors or the level of communality. And there was another study[92] which examined whether sample size or subject-to-item ratio is more important in predicting important outcomes in PCA, and the results indicated an interaction between the two, where the best outcomes occur in analyses where large sample sizes and high ratios are present.

So the feature selection is conducted for all of the frequency bands one by one. The first frequency band for instance, the principal components of the 19 features are obtained by PCA with the 256 samples of 19 variables matrix which has a subject-to-item ratio 13.47; and then, similar algorithm is applied for the rest 15 frequency bands.

5.2.1.1 Investigation of the first frequency band

The data matrix X comprises of $d=19$ dimensions and sample size $M=256$ is used to conduct the PCA. That is, $X = \{x_1, x_2, \dots, x_{256}\}^T$, and $x_i = \{F_{ij}, j = 1,2, \dots ,19\}$.

The scale difference of value of the raw features is very big, and some of them even do not have the same physical unit. Thus the standardizing of the features values is necessary before the PCA. The standardization method can turn the columns in the matrix to vectors with zero mean and standard deviation of 1, as used in [93] and [94], which is:

$$SF_{ij} = \frac{F_{ij} - \overline{x_j}}{SD_j} \quad (i = 1,2,\ldots,256, j = 1,2,\ldots,19) \qquad \text{Equation 5-1}$$

where,

$$\overline{x_j} = \frac{1}{256}\sum_{i=1}^{256} F_{ij} \qquad \text{Equation 5-2}$$

$$SD_j = \left[\frac{1}{256-1}\sum_{i=1}^{256}(F_{ij} - \overline{x_j})^2\right]^{1/2} \qquad \text{Equation 5-3}$$

After that, the eigenvalue vector λ, defined in Equation 3-35, and accumulative contribution rate (ACR), defined in Equation 3-37, can be calculated.

Table 5-2 The eigenvalues and ACR for 256×19 data matrix

N.O. (i)	λ_i	ACR	N.O. (i)	λ_i	ACR
1	12.34297416	64.96%	11	0.010261049	99.98%
2	3.190998499	81.76%	12	0.003245495	100.00%
3	1.698378168	90.70%	13	0.00051437	100.00%
4	1.046581874	96.20%	14	0.000215103	100.00%
5	0.372173257	98.16%	15	7.14041E-05	100.00%
6	0.163824956	99.03%	16	5.20157E-05	100.00%
7	0.085789731	99.48%	17	3.24073E-06	100.00%
8	0.044020871	99.71%	18	1.17217E-07	100.00%
9	0.023501537	99.83%	19	2.61778E-08	100.00%
10	0.017394129	99.92%			

The eigenvalues vector and ACR for the 256×19 data matrix of the feature selection is listed in Table 5-2. The accumulative contribution rates for the 19 features are shown in Figure 5-1. It shows that the first 3 principal components get an ACR of over 90%, thus, they are chosen as the 3 new PCs.

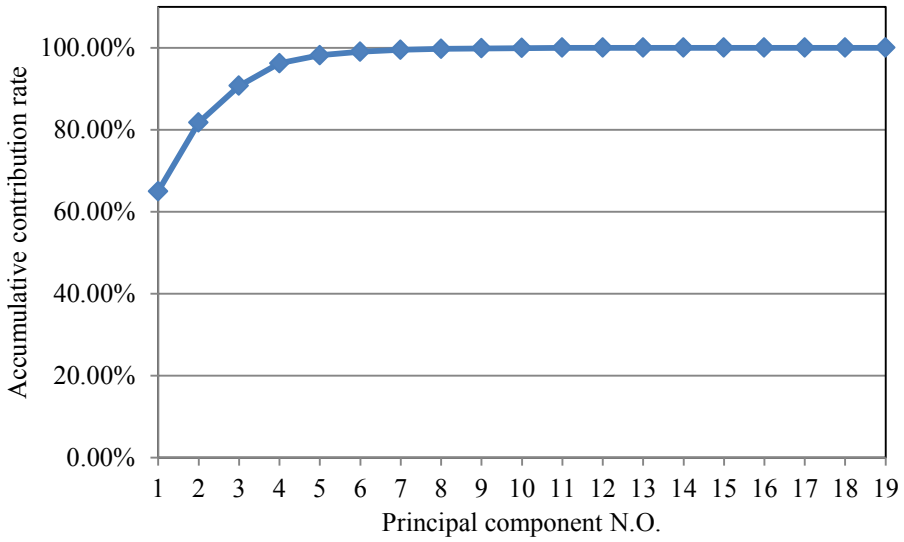

Figure 5-1 The accumulative contribution rate for the 19 principal components

And furthermore, the first 3 eigenvectors, as defined in Equation 3-31, can be used to examine the relations between the different features and the new PCs, i.e., the contributions of different features to the PCs.

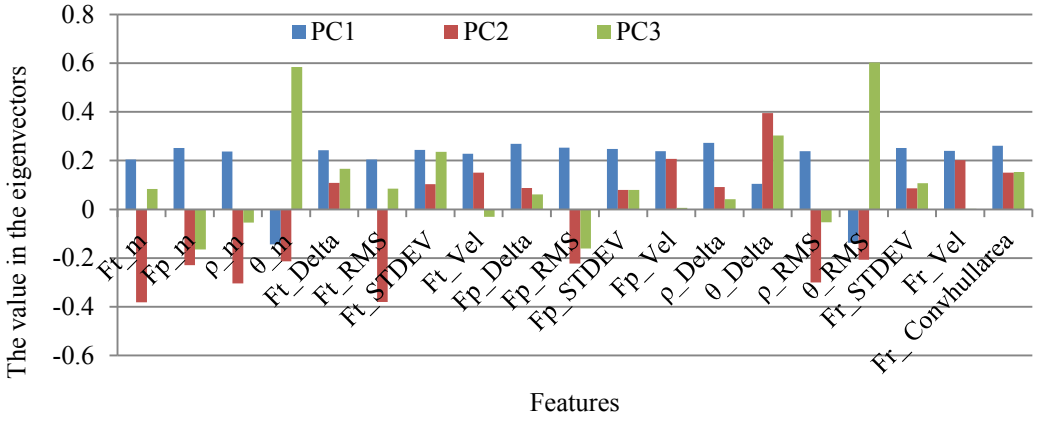

Figure 5-2 The first 3 eigenvectors from features to PCs

Figure 5-2 shows the values of the first 3 eigenvectors, plotted from features to the PCs. It can be observed that for the first PC, i.e., PC1, the values are close for most of the features, which means the first PC has fair relations to most of the features. And for PC2, values of the static features are negative ones, while of the dynamic ones are positive, and this means that PC2 contains more information of the dynamic components of the data. For PC3, high values locate at features related to θ, and that means their relation is stronger than that between others as compensation.

5.2.1.2 Investigation of feature selection in all frequency bands

The number of the PCs for all of the 16 frequency bands and the corresponding ACR are listed in Table 5-3.

Table 5-3 Number of the PCs and ACR for all of the 16 frequency bands

Frequency band	Number of PCs	ACR	Frequency band	Number of PCs	ACR
1	3	90.70%	9	3	93.43%
2	3	92.37%	10	2	85.27%
3	3	87.35%	11	3	92.53%
4	3	89.99%	12	3	92.21%
5	3	89.45%	13	3	88.23%
6	3	90.59%	14	3	88.51%
7	3	93.33%	15	3	89.75%
8	3	91.59%	16	3	85.53%

The threshold of the ACR is set as 85%. And it can be obtained that except for the 10^{th} frequency band, 3 PCs are selected to represent the original features. Therefore, the feature number is reduced to 47 from 304, and the reduction rate is 84.5%.

With the eigenvectors for each of the PC of all the frequency bands, the feature data matrix then can be transferred into a 256×47 matrix, which contains 256 samples for 47 PCs of the original features. Together with the 4 condition parameters, 51 inputs are generated for the training of the neural network.

5.2.2 Training with features selected by PCA

5.2.2.1 The initiation of the BPNN

The performance of a BPNN model depends on the number of hidden layers, the number of neurons in the respective hidden layers, the learning rate, and momentum coefficient[90]. Therefore, the initiation of the BPNN should be explicated.

The BPNN employed here is a 3 layered feed forward neural network, with only one hidden layer and 20 neurons in the hidden layer. Hyperbolic tangent sigmoid transfer function is adopted as the transfer function in the hidden layer while linear function is applied in the output layer.

The BPNN training method applied is the Levenberg-Marquardt algorithm (LMA) as presented in [95]. The LMA was originally designed to serve as an intermediate optimization algorithm between the Gauss–Newton (GN) method and gradient descent algorithm, and address the limitations of each of those techniques.

The increment of weights Δw can be obtained by Equation 5-4 as presented in [96].

$$\Delta w = [J^T J + uI]^{-1} J^T e \qquad \text{Equation 5-4}$$

Where J is the Jacobian matrix, μ is the learning rate which is to be updated using the β depending on the outcome. In particular, μ is multiplied by decay rate β ($0<\beta<1$).

The input data sets are normalized into the range of [-1, 1] by Equation 5-5.

$$f(x) = 2 * \frac{x - x_{min}}{x_{max} - x_{min}} - 1 \qquad \text{Equation 5-5}$$

Where x is the actual value vector, x_{max} is the maximum value of x, x_{min} is the minimum value of x, and $f(x)$ is the normalized value corresponding to x.

Among all of the 256 samples, 70% of them are randomly selected to train the BPNN, 15% are used for validation, and 15% are for testing.

The training objective is the minimization of mean square error (MSE) with the LMA, the MSE is defined as[97]:

$$MSE = \frac{1}{N} \sum_i^N (T_i - O_i^m)^2 \qquad \text{Equation 5-6}$$

where N is the total number of training datasets, T_i is the target output of the i-th dataset i.e. experimental output of the i-th dataset, and O_i^m is the output from the ANN model on the m-th iteration when the i-th dataset is considered as the network input.

5.2.2.2 The prediction result

The scatter plot of the predicted corner wear to measured corner wear is illustrated in Figure 5-3. The MSE of the prediction result is 2.22×10^{-4}.

Linear regressions of the training data set, validation data set and testing data set are also calculated

for evaluation respectively. As it can be seen that the R^2 values of the regressions are 0.9424, 0.9222 and 0.909. And the R^2 value for the whole data set is 0.9156.

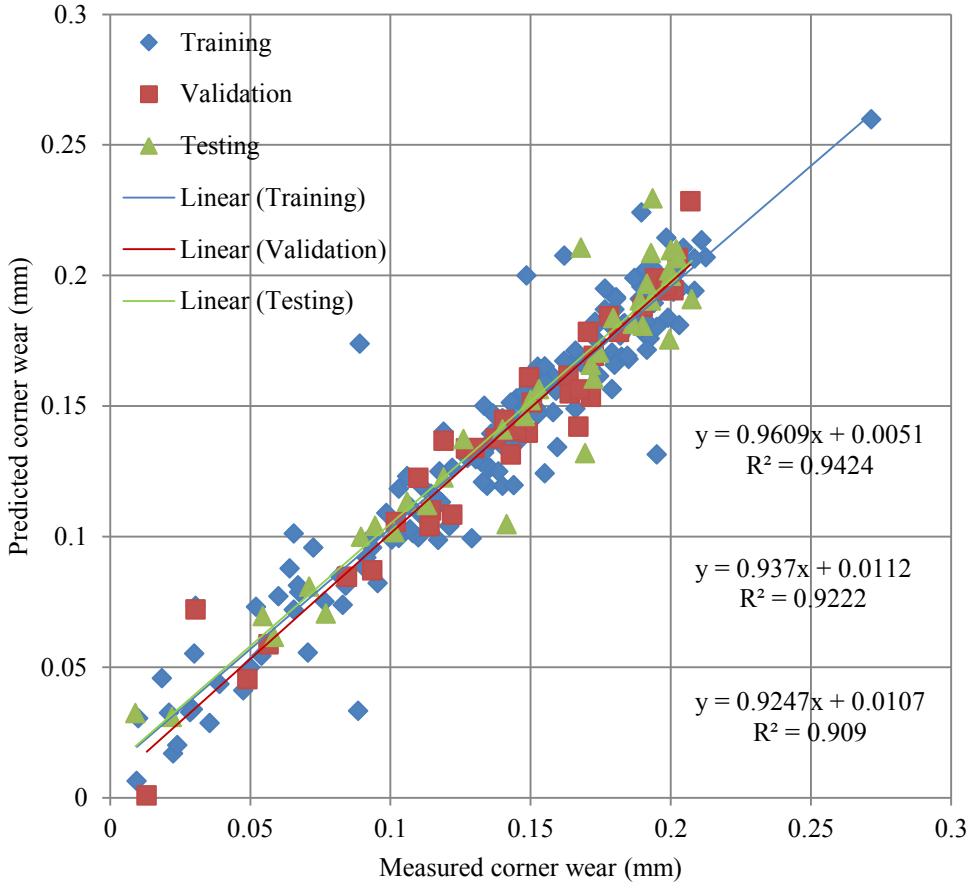

Figure 5-3 The prediction result with features selected by PCA

5.3 Drill Wear Prediction Directly with the BPNN

Another way to evaluate the influences of different features is directly using the BPNN. This method proceeds in three steps:

First, group the 19 features into different clusters, and use the features in each cluster and acquired from the full frequency band to training the BPNN, and then obtain the major cluster;

Second, use the data of major cluster features of different frequency bands to train the BPNN, and then get the major frequency band;

At last, use the data of the major cluster features in the major frequency band to train the BPNN, and the final evaluation is done by the importance ranking.

5.3.1 Finding the major cluster

The 19 features are divided into 3 clusters differing in the dynamic features as shown in Table 5-4.

As discussed in Chapter 4, the static features can indicate the drilling conditions well, thus they are added to the 3 clusters simultaneously. Meanwhile, the average features and RMS ones are very close

to each other, thus only the average one are used here. And for the dynamic features, in cluster 1, are delta features, and in cluster 2 are standard deviation features, and in cluster 3, they are totally newly defined in this work.

Table 5-4 Clustering of the 19 features for evaluation

	Static features	Dynamic features
Cluster 1	$\overline{Ft}$, $\overline{Fp}$, $\bar{\rho}$, $\bar{\theta}$	ΔFt, ΔFp, $\Delta\rho$, $\Delta\theta$
Cluster 2		*Ft_STDEV*, *Fp_STDEV*, *Fr_STDEV*
Cluster 3		*Ft_Vel*, *Fp_Vel*, *Fr_Vel*, *Fr_ConvHullArea*

A training data set, for cluster 1 with 4 static and 4 dynamic making together an 8 features and 4 condition parameters for 256 samples, is generated, and similarly, 7 and 8 features and 4 condition parameters for cluster 2 and 3.

The BPNN initiation for these 3 clusters are the same, and same to the one used in 5.2.2, with the only exception that the preliminary values of the weights are all set to zero to maintain consistency.

Table 5-5 The prediction performances for 3 clusters

	Cluster1	**Cluster2**	**Cluster3**
MSE	1.00E-03	0.0011	8.00E-04
R^2	0.6237	0.5776	0.6834

The BPNN prediction performances for 3 clusters are list in Table 5-5. It is known that the cluster 3 features can predict corner wear more accurate than other features under the current conditions.

5.3.2 Finding the major frequency band

Since cluster 3 features perform better, they are selected to extract features in different frequency bands. For 16 frequency bands, the cluster 3 features are calculated, and for all of the 256 samples. For each frequency bands, 8 features and 4 condition parameters of 256 instances are applied as inputs, and the measured corner wear is adopted as the targets as before.

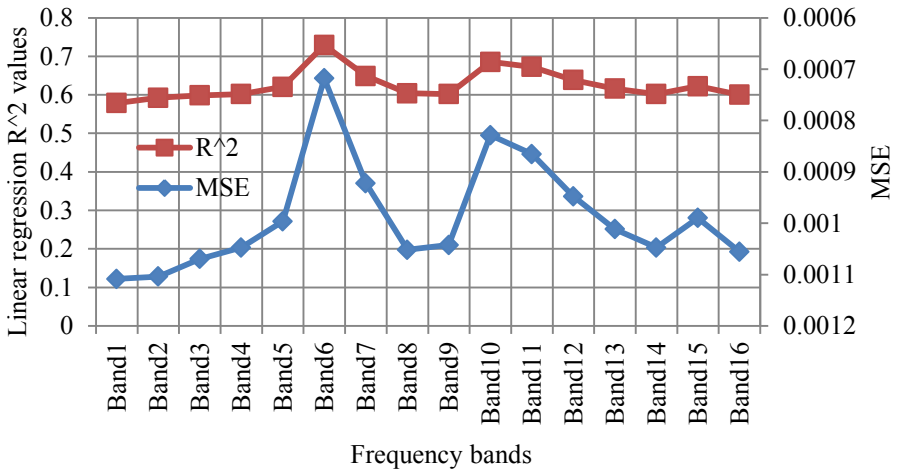

Figure 5-4 Prediction performance of cluster 3 features in different frequency bands

Figure 5-4 shows the prediction performance of the cluster 3 features in different frequency bands. Since the smaller the MSE value the better the performance, the MSE axis is reversed for easy comprehension. With the cluster 3 features in frequency band 6, the highest R^2 and lowest MSE can be obtained.

To investigate the reason for the good performance of frequency band 6, the PCA calculations of the dynamic features in cluster 3 in different frequency bands are conducted. The 4 dynamic features are *Ft_Vel, Fp_Vel, Fr_Vel* and *Fr_ConvHullArea*. For each of them, 16 features are extracted from the 16 frequency bands, and repeated for 256 examples. A 16 dimensional 256 samples matrix is then obtain for each of the 4 dynamic features.

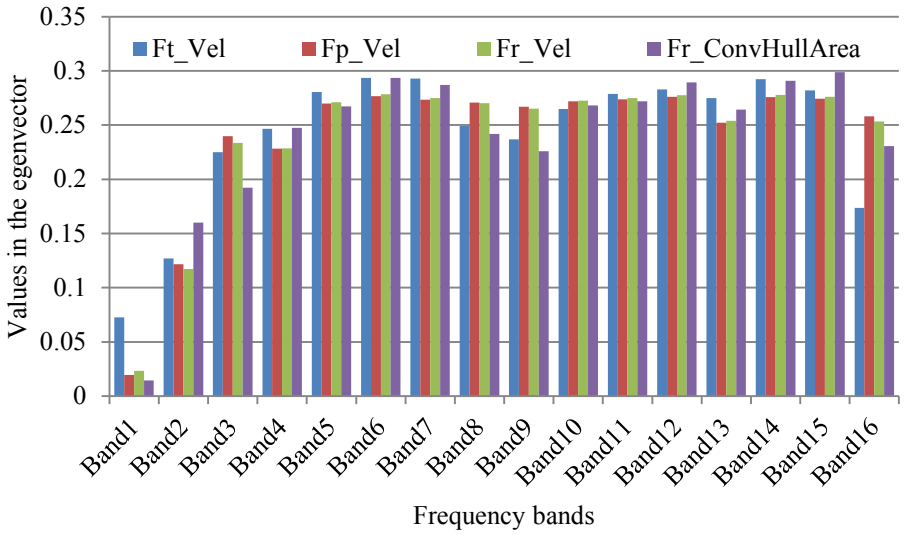

Figure 5-5 The values in the first eigenvector from dynamic features to PC1

As shown in Figure 5-5, the values in the first eigenvectors of the 4 PCA calculation results for 4 dynamic features, which mean the contributions of features in different frequency bands to PC1 (the first PC), vary between different features. And it can be observed that the values from frequency band 6 features to PC1 are higher than other frequency bands for all of the 4 dynamic features.

5.3.3 Comprehensive evaluation

Features in cluster 3 of frequency band 3 are selected to train the BPNN, since they make a comprehensive better performance. The initial values of weights are random values instead of zeros in this occasion and the hidden layer neuron number is still 20. The best values of 20 patterns of training are selected for illustration.

The prediction result is shown in Figure 5-6. The MSE of the result is 3.6×10^{-4} and the linear regression R^2 is 0.8648 for all of the 256 instances. And the evaluation values for training, validation and testing data sets are shown in Figure 5-6.

Garson [98] proposed a method, later modified by Goh [99] for partitioning the neural network connection weights in order to determine the relative importance of each input variable in the network, which is generally called the Garson's algorithm. The procedure of the Garson's algorithm can be demonstrated by Equation 5-7 to Equation 5-10, considering the neural network with *I-H-1* structure

consisting of I inputs, H hidden neurons and 1 single output.

$$P_{(h,i)} = Weights\_IH_{(h,i)} * Weights\_HO_{(h)} \quad \text{Equation 5-7}$$

$$Q_{(h,i)} = \frac{P_{(h,i)}}{\sum_{i=1}^{I} P_{(h,i)}} \quad \text{Equation 5-8}$$

$$S_i = \sum_{h=1}^{H} Q_{(h,i)} \quad \text{Equation 5-9}$$

$$R_i = \frac{S_i}{\sum_{i=1}^{I} S_i} \quad \text{Equation 5-10}$$

In Equation 5-7, $Weights\_IH_{(h,i)}$ means the weight of input i to hidden neuron h and $Weights\_HO_{(h)}$ indicates the weight of hidden neuron h to the output. The P, Q and S are process variables and R_i is the relative importance of the input i.

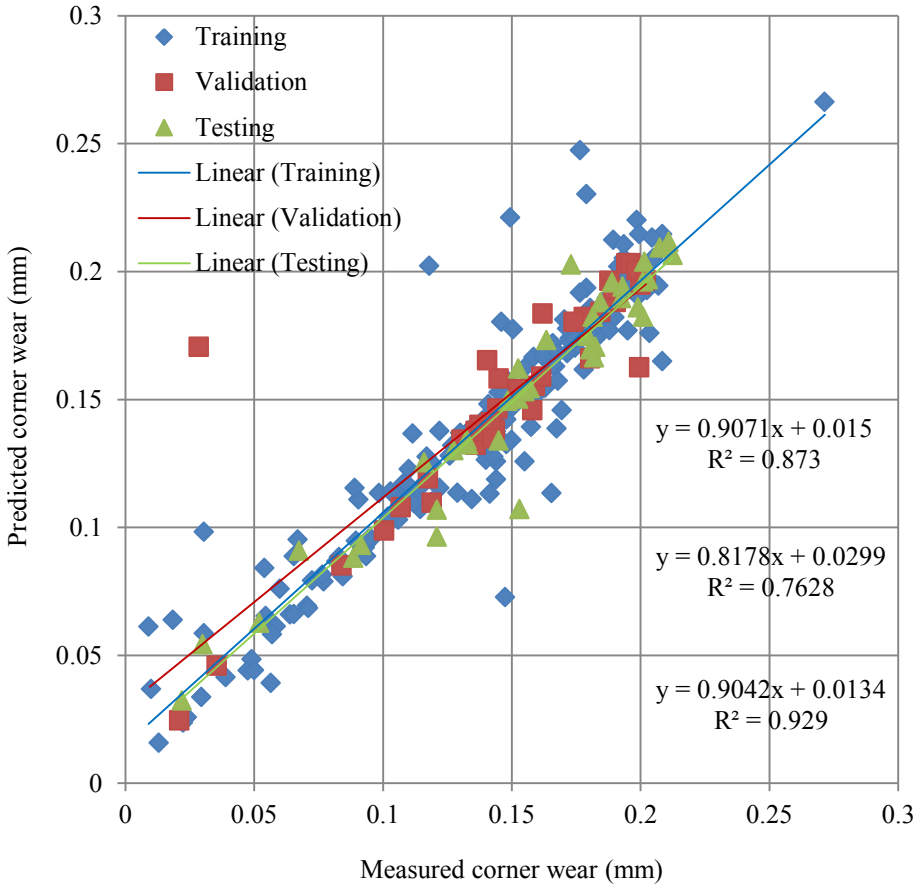

Figure 5-6 The prediction result by directly using BPNN

Here a 12-20-1 BPNN is employed and the relative importance value of each input R_i ($i = 1,2,\ldots,12$) is shown in Figure 5-7.

It can be obtained from the figure that comparing to the drilling condition parameters and dynamic features, static features have a relatively lower importance to the result of corner wear prediction in this very situation. The *spindle speed* has the lowest relative importance comparing with other parameters of drilling conditions in accordance with its lower influence to the static and dynamic features

discussed above. It is also been noticed that the relative importance values of features generated from *Ft* are lower than that of *Fp*, which can be explained by the less sensitivity of them to the drill corner wear.

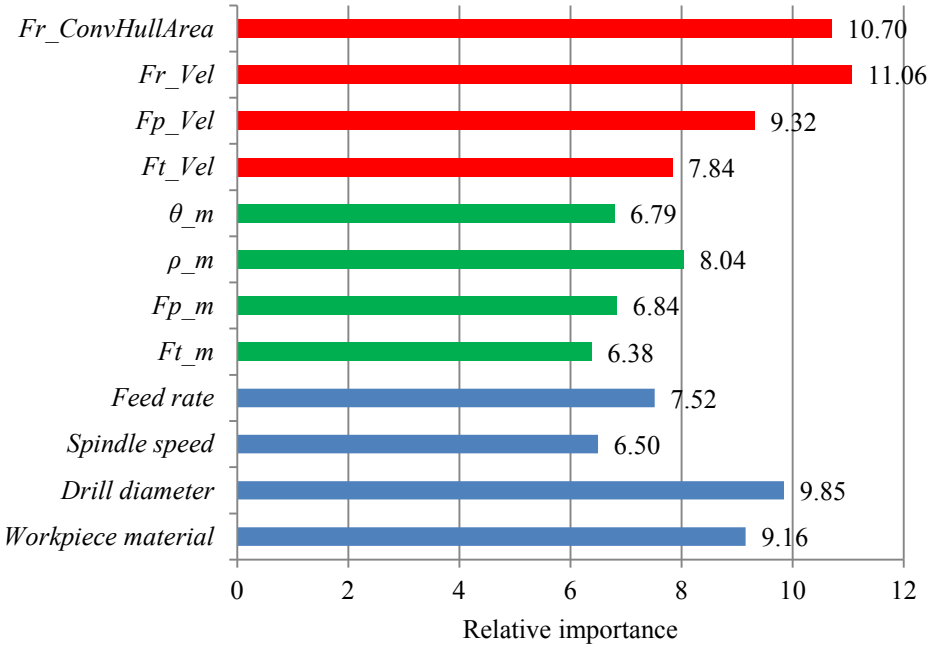

Figure 5-7 Relative importance of different features and condition parameters

5.4 Drilling Conditions Recognition

The BPNN can also be used to recognize the drilling conditions with the features extracted from the monitored signals. Comparing the recognition result with the actual condition parameters is useful for diagnostics of the drilling status. For instance, if the recognized pattern by the well trained BPNN is far from the set up pattern, then some unexpected anomaly might be taking place.

In this section, examples of training BPNN for drilling condition pattern recognition are presented. The data used is the same as that for drilling wear prediction. The training processes are conducted with features selected by PCA as well as by the direct use of BPNN similar to the prediction processes.

Four condition parameters are involved here, which are workpiece material, drill diameter, spindle speed and feed rate. However, the pattern numbers of them are not the same, due to the experimental set up differences shown in Table 5-1.

The BPNN employed in this section is also a 3 layered feed forward network with 20 neurons in the hidden layer. The input numbers and output number are different and decided by the feature numbers and pattern numbers. The initial values of the connection weights are all set as zero for consistency. The different recognition results are compared and discussed.

It should be noted that, the highest drill corner wear of the drill bits used for experiments is about 0.3 mm, which is regarded as medium wear in most practical situations. Therefore, drill corner wear is not concerned in the recognition operations.

5.4.1 Workpiece material

As listed in Table 5-1, four kinds of workpieces are used, thus four patterns or classes are pre-generated as the targets. The target data is then actually a 256×4 matrix, whose 256 rows mean the 256 pattern samples and 4 columns represent 4 workpiece materials respectively. If the workpiece material used is cast iron, then the first column would be valued as 1 and the other 3 as zeros, and the rest can be done in the same manner. The outputs of the BPNN accordingly compose a 256×4 matrix, with each row as the recognition result corresponding to each input data.

5.4.1.1 With features selected by PCA

As discussed above, 47 features are selected from the 304 features extracted from 16 frequency bands. These 47 features and the other 3 drilling condition parameters for 256 samples compose a 256×50 matrix, i.e. the training knowledge data base.

Figure 5-8 illustrates the recognition results for workpiece material patterns with features by PCA.

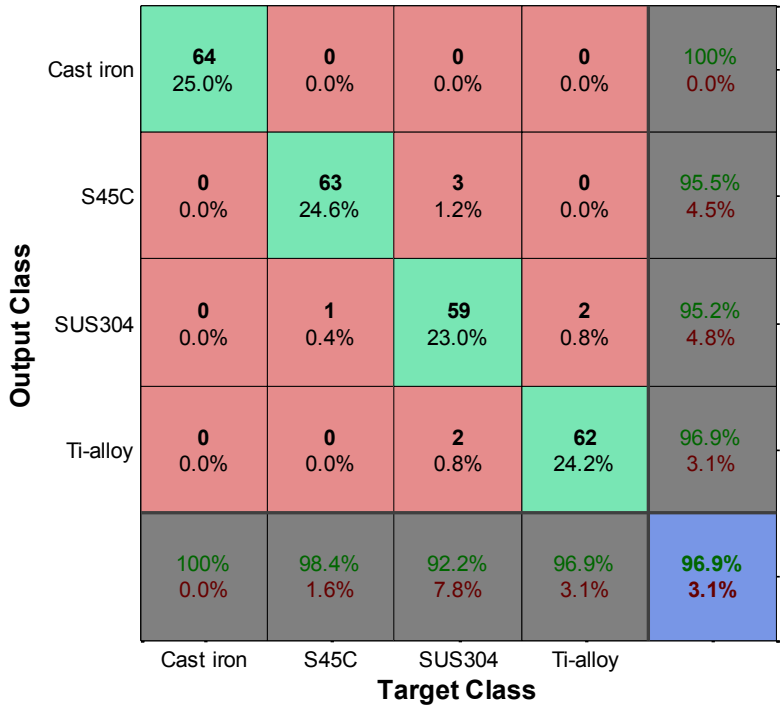

Figure 5-8 The recognition result for workpiece materials with features selected by PCA

In Figure 5-8, the rows are 4 possible output patterns and the 4 columns represent 4 target classes. For each, the output of the BPNN could be any of the 4 classes resulting in the plus one value in the right column and row.

The first left column shows that all of the 64 instances with cast iron as workpiece material have entirely been recognized, and the top row shows that no other training sample data leads to a mistaken output of cast iron. This is reasonable as discussed in 4.1, the feature values of data sampled from processes with cast iron as workpiece material are much different from others.

The second left column has a 98.4% successful rate, which calculated by 1 of the 64 error rate. And in the second top row, some column has nonzero values meaning some other target pattern is

improperly recognized as S45C.

The medial column shows the result for instances with SUS304 as workpiece material. It is demonstrated that 2 of them are mistakenly regarded as Ti-alloy and 3 as S45C, and resulting in a 7.8% error rate.

The fourth column provides the result for Ti-alloy. It is shown that 2 of the 64 instances are wrongly deemed as SUS304.

Values in the diagonal grids represent the numbers of correct results, apparently the bigger the better. Overall a recognition rate, as high as 96.9%, is obtained.

5.4.1.2 With cluster 3 features in frequency band 6

Since the cluster 3 features in frequency band 6 are capable to predict drill corner wear well, they are selected as inputs to train the recognition BPNN. The training data comprises of 256 samples of the other 3 condition parameters and the 8 features in cluster 3, in a form of 256×11 matrix. The target data is the same as that applied in 5.4.1.1.

The recognition result for workpiece materials with features of cluster 3 in frequency band 6 is illustrated in Figure 5-9.

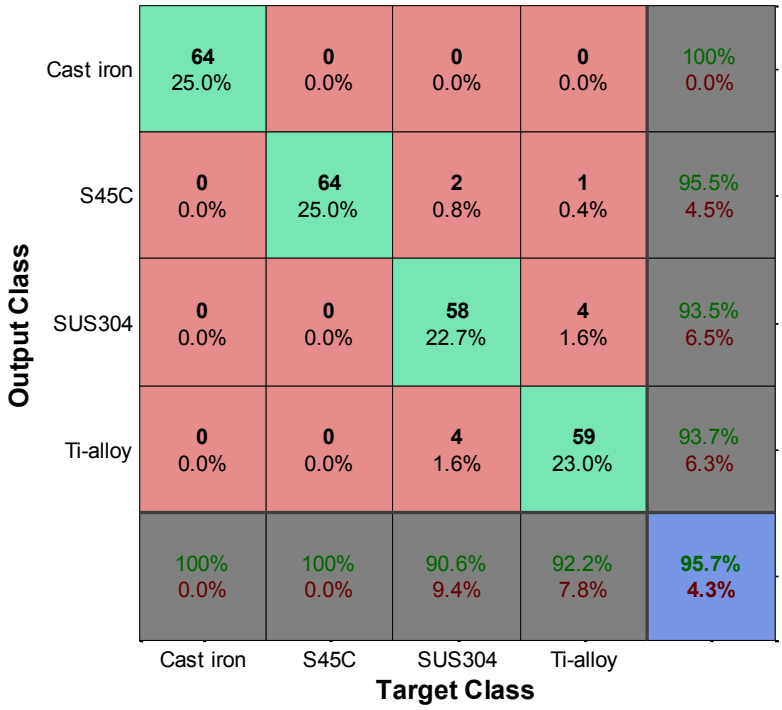

Figure 5-9 The recognition result for workpiece materials with features of cluster 3

Just like the result shown in the first left column and top row of Figure 5-8, the recognition of cast iron is completely successful. The second left column has a 100% successful rate, which means that all of the 64 instances with S45C as workpiece material are precisely recognized. The medial column shows the result for instances with SUS304 as workpiece material. It is demonstrated that 4 of them are mistakenly regarded as Ti-alloy and 2 as S45C. The fourth column provides result of Ti-alloy, as it is shown that 4 are wrongly deemed as SUS304 and 1 as S45C.

And the overall recognition rate is 95.7%, close to the PCA one.

5.4.2 Drill diameter

As listed in Table 5-1, drill bits with four different drill diameters are used in the experiments, thus, like the workpiece material, four patterns or classes of the targets are produced.

The diameter 7 mm is represented as (1, 0, 0, 0) in the data matrix, and similar rules for 8, 9 and 10 mm ones.

5.4.2.1 Training the BPNN with PCA selected features

The training data is a matrix of 256 samples for 47 PC features and 3 condition parameters. The result is shown in Figure 5-10.

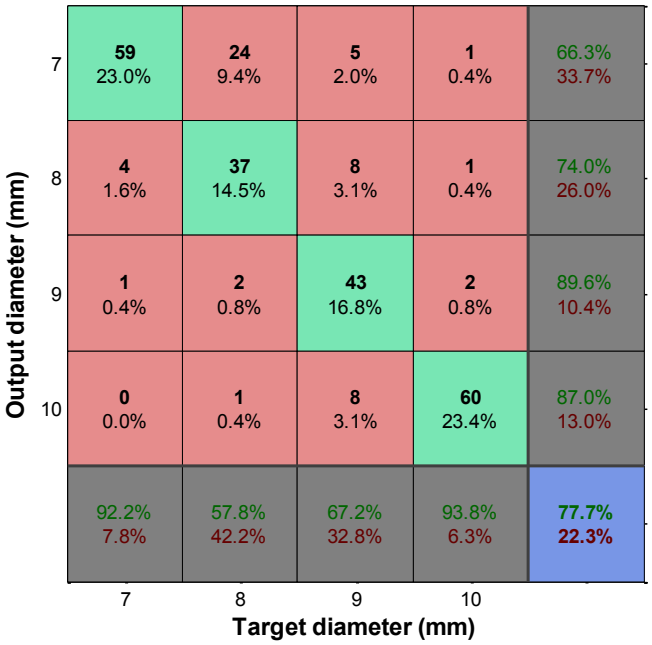

Figure 5-10 Recognition result of drill diameter with PCA selected features

As demonstrated in Figure 5-10, the recognition accuracy rate for 7 mm and 10 mm are as high as 92.2% and 93.8%, and for 9 mm it is 67.2%, while the lowest is for 8 mm, only 57.8%. The overall correct rate is 77.7%.

For the 9 mm pattern, 5 instances are mistaken as 7 mm pattern, 8 instances as the 8 mm pattern and 8 as the 10 mm one. It shows a more possibility to be a wrong pattern next to itself, same as that for 7 mm and 8 mm target patterns, which means the adjacent diameters are more easily confused.

5.4.2.2 With cluster 3 features

The 8 cluster 3 features and 3 drilling condition parameters of 256 samples compose another training data for recognizing the drill diameter. The result is shown in Figure 5-11.

The overall accuracy rate rises up to 84.4%, and the recognition rate for 8 mm and 9 mm instances has been improved a lot. Especially for the 8 mm pattern, 54 instances are correctly recognized.

The difference between the recognition results with PCA features and cluster 3 features is mainly caused by their different characters. The PCA selected features are more partial to the dynamic components contained in the signals while the cluster 3 features cover both the static and dynamic

features.

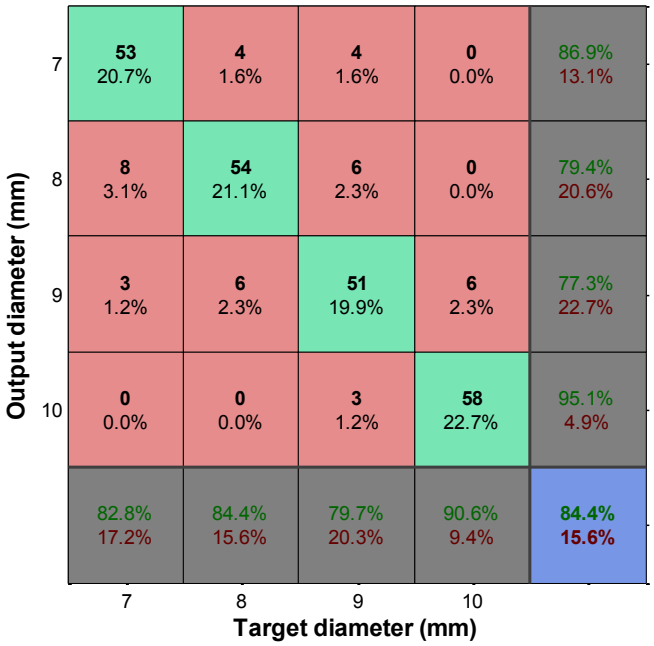

Figure 5-11 Recognition result of drill diameter with cluster 3 features

5.4.3 Spindle speed

Seven different spindle speeds are involved in the experiments which makes the target data a 256×7 matrix. Similarly the BPNN are trained with features both selected by PCA and in cluster 3.

5.4.3.1 With PCA selected features

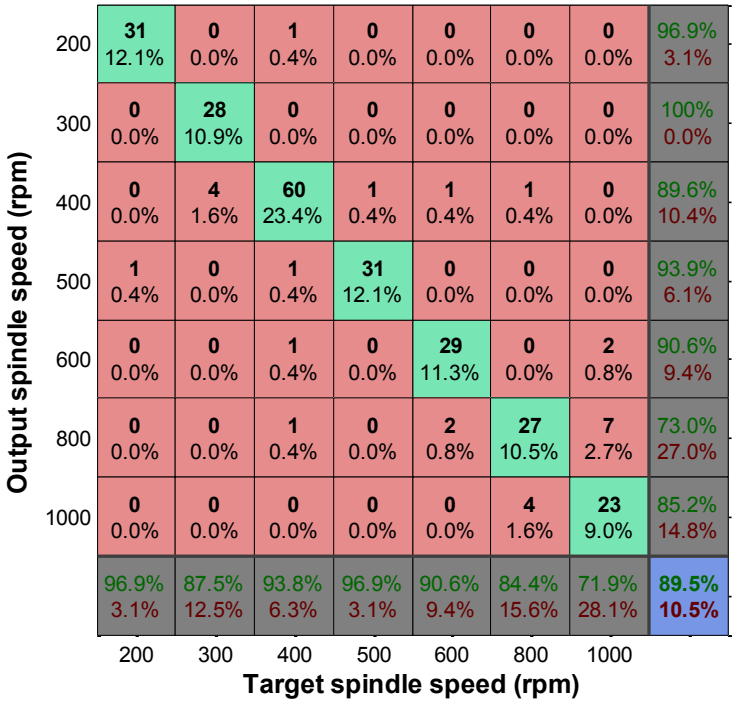

Figure 5-12 Recognition result of spindle speed with PCA selected features

As demonstrated in Figure 5-12, 89.5% of the 256 instances are correctly recognized. The largest error occurs in the last column, where 28.1% of the instances of 1000 rpm spindle speed are mistaken. 7 of them are regarded as 800 rpm pattern and 2 as 600 rpm pattern, which is mainly due to the tiny difference between the force features as discussed in 4.3. The difference between the features gets smaller from 800 rpm to 1000 rpm than that from 400 rpm to 800 rpm.

5.4.3.2 With cluster 3 features

The result trained with the cluster 3 features in demonstrated in Figure 5-13.

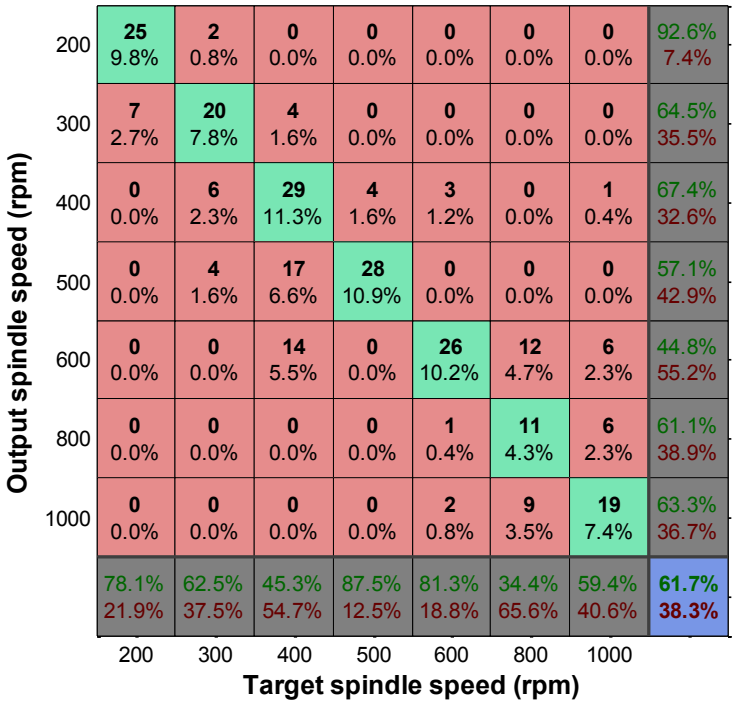

Figure 5-13 Recognition result of spindle speed with cluster 3 features

It is shown that the overall accuracy is 61.7%, much lower than that with PCA features.

The high errors take place when spindle speed is 400 rpm and 800 rpm, while the errors of 500 rpm and 600 rpm are much lower. The reason is that the workpiece material diversity for the later ones is bigger than the former ones. Workpiece materials for 500 rpm are SUS304 and Ti-alloy while for 600 rpm workpiece materials are cast iron and S45C.

5.4.4 Feed rate

As listed in Table 5-1, six different feed rates are used to conduct the drilling experiments, which are 0.05, 0.06, 0.08, 0.1, 0.12 and 0.15 mm/rev. Therefore the target data set is a 256×6 matrix.

5.4.4.1 With PCA selected features

The recognition result with PCA selected features is shown in Figure 5-14, an overall 80.5% accuracy rate is obtained.

It can be observed that the correct rates for 0.05 and 0.15 mm/rev. patterns are both 100%. The reason for 0.15 mm/rev. pattern is easily get because the feature values are much higher than others.

The good distinguish result between 0.05 and 0.06 mm/rev. patterns is that the workpiece material can be a strong indicator since they were separately.

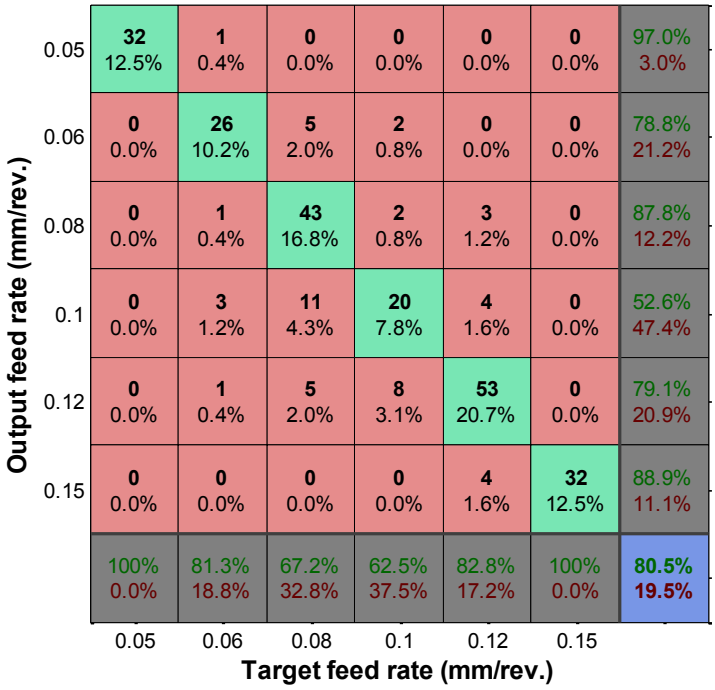

Figure 5-14 Recognition result of feed rate with PCA selected features

The situation for 0.08 and 0.1 mm/rev. patterns are much more severe, with 67.2% and 62.5% recognition rate obtained respectively.

5.4.4.2 With cluster 3 features

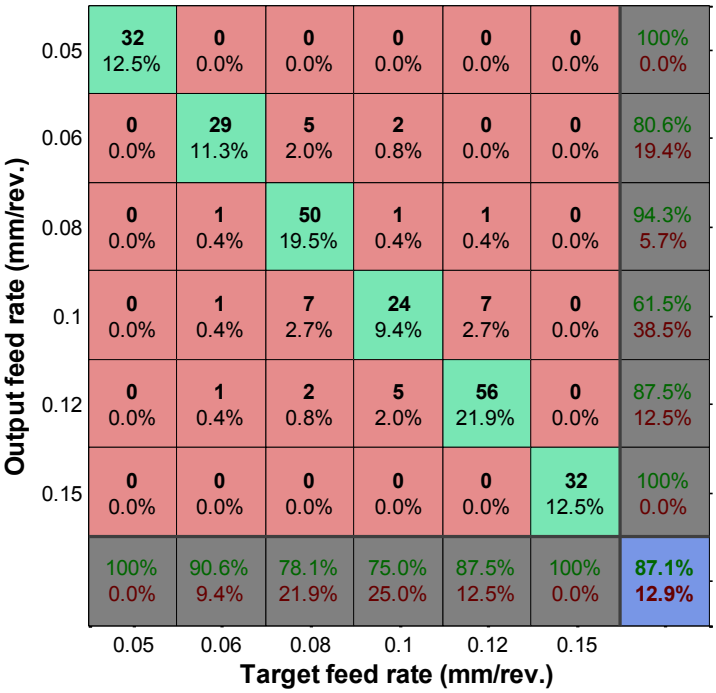

Figure 5-15 Recognition result of feed rate with cluster 3 features

The result is improved by using the cluster 3 features as demonstrated in Figure 5-15. It shows that the total precision rate is 87.1%, with big promotes for both 0.08 and 0.1 mm/rev. patterns. And the correct rates remain 100% for 0.05 and 0.15 mm/rev. patterns.

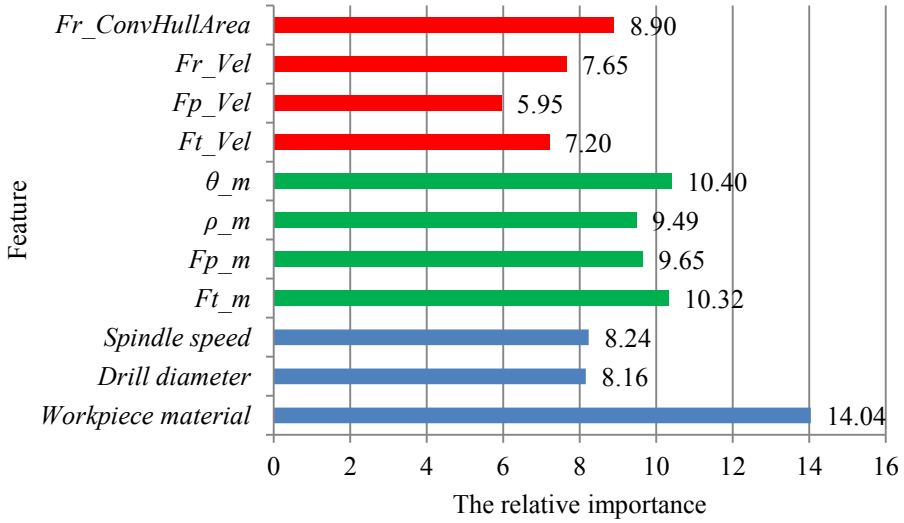

Figure 5-16 The relative importance of different features in the BPNN for feed rate recognition

The melioration of the recognition rate is on account of the contribution of static features. As demonstrated in Figure 5-16, the relative importance of the average features is higher than those of the dynamic ones.

It is also necessary to notice that the relative importance of workpiece materials is very high, which means the distribution of the feed rate patterns is not equal to different workpiece materials, and it should be avoided in further works.

5.5 Summary

In this chapter, the prediction of the drill corner wear and the recognition of the drilling condition parameters are carried out based on the data acquired from experiments under a wide range. The principal components analysis (PCA) is applied to select the principal components (PCs) from the features generated and extracted from the signals sampled during drilling, and then the PCs are taken as new features and are used for drill wear prediction and condition recognition. Another approach which employed BPNN directly is used to find the most effective features, which first find the best feature cluster and then excavate the best frequency band.

The first PC has similar correlations to both the static and dynamic features, while the second and third PCs are more relevant to the dynamic features. The static features are important because they indicate the drilling conditions well and for the dynamic features, the drilling status.

For each frequency band, 2 or 3 PCs are selected to represent the 19 features and for all of the frequency bands, 47 PCs are selected from the 304 features.

The 19 features are chosen and grouped into 3 clusters, and it is found that the cluster 3 features, which comprised of the average, velocity and geometry features, are more effective. And furthermore, the cluster 3 features in frequency band 6 have higher variance leading to better prediction result.

The relative importance, calculated from the weights in the BPNN, of each selected feature show that the dynamic features play a more important role than the static features for cluster 3.

The cluster 3 features are found more effective in recognizing the drilling conditions which strongly related to static features, than the PCA selected ones. Such as for the recognition of feed rate, it is discovered that static features have higher relative importance in the neural network. But for spindle speed, with the PCA selected PCs, the result is better for spindle speed affecting the frequency performance of the force signals.

Chapter 6 Conclusion

This dissertation has described an integrated procedure for drilling process monitoring on the scopes of drill wear prediction and drilling conditions recognition. A novel conversion approach based on the sampled thrust force and torque signals has been presented, producing new forces and capturing information of their correlation. The equivalent thrust force (*Ft*), equivalent principal force (*Fp*) and their resultant force (*Fr*) are obtained from a rectangular coordinate. Various features of these forces are generated with different calculation algorithms, among which the velocity of the force value change and geometry measure of the distributed resultant force trajectory is specially defined. The features are grouped and the effects of drilling condition parameters are evaluated to all of the features in each group. Experiments are conducted under a wide range of conditions to build a big enough data base for the training of neural networks devoting to the drill wear prediction and drilling condition monitoring. The PCA method and the direct using BPNN method are applied to select principal or important features. The features are evaluated by their effectiveness and importance in the artificial neural network models.

It is found that, the change of drilling conditions leads to a more clear change of static features than dynamic features, whilst the dynamic features are more sensitive to the status or the stability of the drilling process. And it's also been noticed that the velocity features have a certain concern to the frequency information of the force time series signals and the geometry feature keeps pace with the increase of drill corner wear more conformably.

The PCA of the data shows that the first principal component (PC) has similar correlations to both the static and dynamic features, while the second and third PCs are more relevant to the dynamic features. The high variation of the static features is caused by the diverse drilling conditions and for the dynamic features is different drilling status.

The direct use of BPNN indicates that with the feature cluster consist of average, velocity and geometry features higher prediction accuracy can be obtained. And with the help of PCA, it is explained that in the frequency band which has the most correlation to the PC features are more effective.

The relative importance method reveals that, in the prediction of drill corner wear, dynamic features, especially the newly defined features, have more effect, while for the drilling condition recognition, the static features are more contributory.

However, there are some disadvantages in this research. First the highest drill corner wear of the drill bits used in the experiments is not adequate, which limits the accuracy and the practicability of the method. Second, the drilling condition parameters overlap in some occasions, which affect the fairness of the evaluation.

References

[1] J. Davim, *Machining*. London: Springer London, 2008.

[2] H. K. Tönshoff, W. Spintig, W. König, and A. Neises, "Machining of Holes Developments in Drilling Technology," *CIRP Ann. - Manuf. Technol.*, vol. 43, no. 2, pp. 551–561, Jan. 1994.

[3] A. Y. Helmi and E.-H. Hassan, *Machining Technology, Machine Tools and Operations*. Raton Boca, USA: CRC Press, 2008, p. 633.

[4] A. Albu, L. Stanciu, and G. Belgiu, "An approach of artificial intelligent machining process," in *2011 6th IEEE International Symposium on Applied Computational Intelligence and Informatics (SACI)*, 2011, pp. 123–126.

[5] D. Barschdorff and L. Monostori, "Neural networks—Their applications and perspectives in intelligent machining," *Comput. Ind.*, vol. 17, no. 2–3, pp. 101–119, Nov. 1991.

[6] S. B. Billatos and P.-C. Tseng, "Knowledge-based optimization for intelligent machining," *J. Manuf. Syst.*, vol. 10, no. 6, pp. 464–475, Jan. 1991.

[7] S. Takata, "Generation of a Machining Scenario and Its Applications to Intelligent Machining Operations," *CIRP Ann. - Manuf. Technol.*, vol. 42, no. 1, pp. 531–534, Jan. 1993.

[8] S. Y. Liang, R. L. Hecker, and R. G. Landers, "Machining Process Monitoring and Control: The State-of-the-Art," *J. Manuf. Sci. Eng.*, vol. 126, no. 2, p. 297, 2004.

[9] J. V. Abellan-Nebot and F. Romero Subirón, "A review of machining monitoring systems based on artificial intelligence process models," *Int. J. Adv. Manuf. Technol.*, vol. 47, no. 1–4, pp. 237–257, Jul. 2009.

[10] H. K. Tönshoff, J. P. Wulfsberg, H. J. J. Kals, W. König, and C. A. van Luttervelt, "Developments and Trends in Monitoring and Control of Machining Processes," *CIRP Ann. - Manuf. Technol.*, vol. 37, no. 2, pp. 611–622, Jan. 1988.

[11] M. Shiraishi, "Scope of in-process measurement, monitoring and control techniques in machining processes — Part 1: In-process techniques for tools," *Precis. Eng.*, vol. 10, no. 4, pp. 179–189, Oct. 1988.

[12] M. Shiraishi, "Scope of in-process measurement, monitoring and control techniques in machining processes—Part 2: In-process techniques for workpieces," *Precis. Eng.*, vol. 11, no. 1, pp. 27–37, Jan. 1989.

[13] M. Shiraishi, "Scope of in-process measurement, monitoring and control techniques in machining processes—Part 3: In-process techniques for cutting processes and machine tools," *Precis. Eng.*, vol. 11, no. 1, pp. 39–47, Jan. 1989.

[14] D. A. Dornfeld, "In process recognition of cutting states," *JSME Int. journal. Ser. C, Dyn. Control. ...*, vol. 37, no. 4, pp. 638–650, 1994.

[15] D.-W. Cho, S. J. Lee, and C. N. Chu, "The state of machining process monitoring research in Korea," *Int. J. Mach. Tools Manuf.*, vol. 39, no. 11, pp. 1697–1715, Nov. 1999.

[16] R. Teti, K. Jemielniak, G. O'Donnell, and D. Dornfeld, "Advanced monitoring of machining

operations," *CIRP Ann. - Manuf. Technol.*, vol. 59, no. 2, pp. 717–739, Jan. 2010.

[17] P. Kovač and I. Mankova, "A review of machining monitoring systems," *J. Prod. Eng.*, vol. 14, no. 1, pp. 1–6, 2011.

[18] R. Tezuka and K. Sekiya, "The Relationship between Dynamic Components of Cutting Force and Adhesion of Tool-Chip Interface," *Key Eng. Mater.*, vol. 408, pp. 469–472, 2009.

[19] H. . Ertunc and K. . Loparo, "A decision fusion algorithm for tool wear condition monitoring in drilling," *Int. J. Mach. Tools Manuf.*, vol. 41, no. 9, pp. 1347–1362, Jul. 2001.

[20] Y. Wu and R. Du, "FEATURE EXTRACTION AND ASSESSMENT USING WAVELET PACKETS FOR MONITORING OF MACHINING PROCESSES," *Mech. Syst. Signal Process.*, vol. 10, no. 1, pp. 29–53, Jan. 1996.

[21] Y. Wu, P. Escande, and R. Du, "A New Method for Real-Time Tool Condition Monitoring in Transfer Machining Stations," *J. Manuf. Sci. Eng.*, vol. 123, no. 2, pp. 339–347, May 2001.

[22] E. Brinksmeier, "Prediction of tool fracture in drilling," *CIRP Ann. Technol.*, vol. X, no. 2, pp. 97–100, 1990.

[23] G. S. Li, W. S. Lau, and Y. Z. Zhang, "In-process drill wear and breakage monitoring for a machining centre based on cutting force parameters," *Int. J. Mach. Tools Manuf.*, vol. 32, no. 6, pp. 855–867, Dec. 1992.

[24] A. Noori-Khajavi and R. Komanduri, "On Multisensor Approach to Drill Wear Monitoring," *CIRP Ann. - Manuf. Technol.*, vol. 42, no. 1, pp. 71–74, Jan. 1993.

[25] I. N. Tansel, C. Mekdeci, O. Rodriguez, and B. Uragun, "Monitoring drill conditions with wavelet based encoding and neural networks," *Int. J. Mach. Tools Manuf.*, vol. 33, no. 4, pp. 559–575, Aug. 1993.

[26] J. M. Lee, D. K. Choi, and C. N. Chu, "Real-Time Tool Breakage Monitoring for NC Turning and Drilling," *CIRP Ann. - Manuf. Technol.*, vol. 43, no. 1, pp. 81–84, Jan. 1994.

[27] A. Noori-Khajavi and R. Komanduri, "Frequency and time domain analyses of sensor signals in drilling—I. Correlation with drill wear," *Int. J. Mach. Tools Manuf.*, vol. 35, no. 6, pp. 775–793, Jun. 1995.

[28] S. C. Lin and C. J. Ting, "Tool wear monitoring in drilling using force signals," *Wear*, vol. 180, no. 1–2, pp. 53–60, Jan. 1995.

[29] A. Noori-Khajavi and R. Komanduri, "Frequency and time domain analyses of sensor signals in drilling—II. Investigation on some problems associated with sensor integration," *Int. J. Mach. Tools Manuf.*, vol. 35, no. 6, pp. 795–815, Jun. 1995.

[30] S. C. Lin and C. J. Ting, "Drill wear monitoring using neural networks," *Int. J. Mach. Tools Manuf.*, vol. 36, no. 4, pp. 465–475, Apr. 1996.

[31] T. El-Wardany, D. Gao, and M. Elbestawi, "Tool condition monitoring in drilling using vibration signature analysis," *Int. J. Mach. Tools Manuf.*, vol. 36, no. 6, pp. 687–711, Jun. 1996.

[32] L. XIAOLI, Y. YINGXUE, and Y. ZHEJUN, "On-line tool condition monitoring system with wavelet fuzzy neural network," *J. Intell. Manuf.*, 1997.

[33] L. Xiaoli and Y. Zhejun, "Tool wear monitoring with wavelet packet transform—fuzzy

clustering method," *Wear*, vol. 219, no. 2, pp. 145–154, Sep. 1998.

[34] X. Li, "Real-time detection of the breakage of small diameter drills with wavelet transform," *Int. J. Adv. Manuf. Technol.*, vol. 14, no. 8, pp. 539–543, Aug. 1998.

[35] X. Li and S. K. Tso, "Drill wear monitoring based on current signals," *Wear*, vol. 231, no. 2, pp. 172–178, Jul. 1999.

[36] L. Xiaoli, "On-line detection of the breakage of small diameter drills using current signature wavelet transform," *Int. J. Mach. Tools Manuf.*, vol. 39, no. 1, pp. 157–164, Jan. 1999.

[37] X. Li, S. Dong, and Z. Yuan, "Discrete wavelet transform for tool breakage monitoring," *Int. J. Mach. Tools Manuf.*, vol. 39, no. 12, pp. 1935–1944, Dec. 1999.

[38] X. Li, S. Dong, and P. K. Venuvinod, "Hybrid Learning for Tool Wear Monitoring," *Int. J. Adv. Manuf. Technol.*, vol. 16, no. 5, pp. 303–307, Apr. 2000.

[39] X. Li, S. K. Tso, S. Member, and J. Wang, "Real-time tool condition monitoring using wavelet transforms and fuzzy techniques," *IEEE Trans. Syst. Man Cybern. Part C (Applications Rev.*, vol. 30, no. 3, pp. 352–357, 2000.

[40] H. M. Ertunc, K. a. Loparo, and H. Ocak, "Tool wear condition monitoring in drilling operations using hidden Markov models (HMMs)," *Int. J. Mach. Tools Manuf.*, vol. 41, no. 9, pp. 1363–1384, Jul. 2001.

[41] H.-Y. Kim and J.-H. Ahn, "Chip disposal state monitoring in drilling using neural network based spindle motor power sensing," *Int. J. Mach. Tools Manuf.*, vol. 42, no. 10, pp. 1113–1119, Aug. 2002.

[42] B. Brophy, K. Kelly, and G. Byrne, "AI-based condition monitoring of the drilling process," *J. Mater. Process. Technol.*, vol. 124, no. 3, pp. 305–310, Jun. 2002.

[43] K. Vishy, Karri;Tossapol, "Tool Condition Monitoring in Drilling Using Artificial Neural Networks," in *computer science*, 2003, vol. 2903, pp. 293–301.

[44] I. Abu-Mahfouz, "Drilling wear detection and classification using vibration signals and artificial neural network," *Int. J. Mach. Tools Manuf.*, vol. 43, no. 7, pp. 707–720, May 2003.

[45] H. M. Ertunc and C. Oysu, "Drill wear monitoring using cutting force signals," *Mechatronics*, vol. 14, no. 5, pp. 533–548, Jun. 2004.

[46] D. Axinte and N. Gindy, "Assessment of the effectiveness of a spindle power signal for tool condition monitoring in machining processes," *Int. J. Prod. Res.*, vol. 42, no. 13, pp. 2679–2691, Jul. 2004.

[47] C. Sanjay, M. L. Neema, and C. W. Chin, "Modeling of tool wear in drilling by statistical analysis and artificial neural network," *J. Mater. Process. Technol.*, vol. 170, no. 3, pp. 494–500, Dec. 2005.

[48] A. K. Singh, S. S. Panda, D. Chakraborty, and S. K. Pal, "Predicting drill wear using an artificial neural network," *Int. J. Adv. Manuf. Technol.*, vol. 28, no. 5–6, pp. 456–462, May 2005.

[49] P. Baruah and R. B. Chinnam, "HMMs for diagnostics and prognostics in machining processes," *Int. J. Prod. Res.*, vol. 43, no. 6, pp. 1275–1293, Mar. 2005.

[50] F. Alsulaiman, M. Baseer, and a Sheikh, “Use of electrical power for online monitoring of tool condition,” *J. Mater. Process. Technol.*, vol. 166, no. 3, pp. 364–371, Aug. 2005.

[51] S. PANDA, A. SINGH, D. CHAKRABORTY, and S. PAL, “Drill wear monitoring using back propagation neural network,” *J. Mater. Process. Technol.*, vol. 172, no. 2, pp. 283–290, Feb. 2006.

[52] R. Heinemann, S. Hinduja, and G. Barrow, “Use of process signals for tool wear progression sensing in drilling small deep holes,” *Int. J. Adv. Manuf. Technol.*, vol. 33, no. 3–4, pp. 243–250, Apr. 2006.

[53] S. S. Panda, D. Chakraborty, and S. K. Pal, “Monitoring of drill flank wear using fuzzy back-propagation neural network,” *Int. J. Adv. Manuf. Technol.*, vol. 34, no. 3–4, pp. 227–235, Aug. 2006.

[54] L. A. Franco-Gasca, G. Herrera-Ruiz, R. Peniche-Vera, R. D. J. Romero-Troncoso, and W. Leal-Tafolla, “Sensorless tool failure monitoring system for drilling machines,” *Int. J. Mach. Tools Manuf.*, vol. 46, no. 3–4, pp. 381–386, Mar. 2006.

[55] V. M. Natarajan, U; Arun, P; Periasamy, “A decision fusion algorithm for tool condition monitoring in drilling using Hidden Markov Model (HMM),” *INDIAN J. Eng. Mater. Sci.*, vol. 13, no. April, pp. 103–109, 2006.

[56] J. Yan and J. Lee, “A Hybrid Method for On-line Performance Assessment and Life Prediction in Drilling Operations,” in *2007 IEEE International Conference on Automation and Logistics*, 2007, pp. 2500–2505.

[57] K. Patra, S. K. Pal, and K. Bhattacharyya, “Artificial neural network based prediction of drill flank wear from motor current signals,” *Appl. Soft Comput.*, vol. 7, no. 3, pp. 929–935, Jun. 2007.

[58] K. Patra, S. Pal, and K. Bhattacharyya, “Application of wavelet packet analysis in drill wear monitoring,” *Mach. Sci. Technol.*, vol. 11, no. 3, pp. 413–432, 2007.

[59] L. Fu, S.-F. Ling, and C.-H. Tseng, “On-line breakage monitoring of small drills with input impedance of driving motor,” *Mech. Syst. Signal Process.*, vol. 21, no. 1, pp. 457–465, Jan. 2007.

[60] S. S. Panda, D. Chakraborty, and S. K. Pal, “Flank wear prediction in drilling using back propagation neural network and radial basis function network,” *Appl. Soft Comput.*, vol. 8, no. 2, pp. 858–871, Mar. 2008.

[61] C. Aliustaoglu, H. M. Ertunc, and H. Ocak, “Tool wear condition monitoring using a sensor fusion model based on fuzzy inference system,” *Mech. Syst. Signal Process.*, vol. 23, no. 2, pp. 539–546, Feb. 2009.

[62] M. P. Gómez, A. M. Hey, J. E. Ruzzante, and C. E. D’Attellis, “Tool wear evaluation in drilling by acoustic emission,” *Phys. Procedia*, vol. 3, no. 1, pp. 819–825, Jan. 2010.

[63] R. Heinemann and S. Hinduja, “A new strategy for tool condition monitoring of small diameter twist drills in deep-hole drilling,” *Int. J. Mach. Tools Manuf.*, vol. 52, no. 1, pp. 69–76, Jan. 2012.

[64] A. V. Oppenheim and R. W. Schafer, “Dsp history - From frequency to quefrency: a history of the cepstrum,” *IEEE Signal Process. Mag.*, vol. 21, no. 5, pp. 95–106, Sep. 2004.

[65] R. K. R. Yarlagadda, *Analog and Digital Signals and Systems*. Boston, MA: Springer US, 2010,

pp. 39–67.

[66] Udayashankara, *Real Time Digital Signal Processing*. PHI Learning Pvt. Ltd., 2010, pp. 67–82.

[67] L. Li and G. E. Caldwell, "Coefficient of cross correlation and the time domain correspondence," *J. Electromyogr. Kinesiol.*, vol. 9, no. 6, pp. 385–389, Dec. 1999.

[68] G. Upton and I. Cook, *A Dictionary of Statistics*. Oxford University Press, 2002, p. 104.

[69] J. S. Bendat and A. G. Piersol, *Random Data*. Wiley-Interscience, 1986.

[70] S. V. M. Selvam and C. Sujatha, "Twist drill deformation and optimum drill geometry," *Comput. Struct.*, vol. 57, no. 5, pp. 903–914, Dec. 1995.

[71] E. Abele and M. Fujara, "Simulation-based twist drill design and geometry optimization," *CIRP Ann. - Manuf. Technol.*, vol. 59, no. 1, pp. 145–150, Jan. 2010.

[72] K. Sambhav, P. Tandon, and S. G. Dhande, "Geometric modeling and validation of twist drills with a generic point profile," *Appl. Math. Model.*, vol. 36, no. 6, pp. 2384–2403, Jun. 2012.

[73] J. . Strenkowski, C. . Hsieh, and A. . Shih, "An analytical finite element technique for predicting thrust force and torque in drilling," *Int. J. Mach. Tools Manuf.*, vol. 44, no. 12–13, pp. 1413–1421, Oct. 2004.

[74] C. R., Ronald, Y. Meyer, S. Quake, and M. V. Wickerhauser, *Wavelets and Their Applications*. 1994, pp. 363–379.

[75] Y. Meyer, "Wavelets and operators," *London Math. Soc. Lect. Notes Ser.*, vol. 37, pp. 256–365, 1989.

[76] D. Shi and N. N. Gindy, "Tool wear predictive model based on least squares support vector machines," *Mech. Syst. Signal Process.*, vol. 21, no. 4, pp. 1799–1814, May 2007.

[77] Jolliffe I.T., *Principal Component Analysis*, 2nd ed. Springer, NY, 2002, p. 28.

[78] Q. He, R. Yan, F. Kong, and R. Du, "Machine condition monitoring using principal component representations," *Mech. Syst. Signal Process.*, vol. 23, no. 2, pp. 446–466, Feb. 2009.

[79] I. a Basheer and M. Hajmeer, "Artificial neural networks: fundamentals, computing, design, and application.," *J. Microbiol. Methods*, vol. 43, no. 1, pp. 3–31, Dec. 2000.

[80] A. Jain, J. Mao, and K. Mohiuddin, "Artificial neural networks: A tutorial," *IEEE Comput.*, 1996.

[81] P. V. Biron, "Backpropagation: Theory, architectures, and applications," *J. Am. Soc. Inf. Sci.*, vol. 48, no. 1, pp. 88–89, Jan. 1997.

[82] T. Nakamoto, K. Fukunishi, and T. Moriizumi, "Identification capability of odor sensor using quartz-resonator array and neural-network pattern recognition," *Sensors Actuators B Chem.*, vol. 1, no. 1–6, pp. 473–476, Jan. 1990.

[83] K. Ema, M. Yokoyama, T. Nakamoto, and T. Moriizumi, "Odour-sensing system using a quartz-resonator sensor array and neural-network pattern recognition," *Sensors and Actuators*, vol. 18, pp. 291–296, 1989.

[84] J. C. Mellinger, O. Burak Ozdoganlar, R. E. DeVor, and S. G. Kapoor, "Modeling Chip-

Evacuation Forces and Prediction of Chip-Clogging in Drilling," *J. Manuf. Sci. Eng.*, vol. 124, no. 3, p. 605, 2002.

[85] Y. YAMANE and K. SEKIYA, "An evaluation of difficulty in machining difficult-to-cut materials by using difficult-to-cut rating," *J. Japan Soc. Precis. Eng.*, vol. 70, no. 3, pp. 407–411, 2004.

[86] M. Pirtini and I. Lazoglu, "Forces and hole quality in drilling," *Int. J. Mach. Tools Manuf.*, vol. 45, no. 11, pp. 1271–1281, Sep. 2005.

[87] Y. S. Tarng and T. C. Li, "Detection and Suppression of Drilling Chatter," *J. Dyn. Syst. Meas. Control*, vol. 116, no. 4, p. 729, 1994.

[88] Y. S. Tarng and T. C. Li, "Adaptive pattern recognition of drilling chatter," *J. Mater. Process. Technol.*, vol. 48, no. 1–4, pp. 247–253, Jan. 1995.

[89] J. A. Yang, V. Jaganathan, and R. Du, "A new dynamic model for drilling and reaming processes," *Int. J. Mach. Tools Manuf.*, vol. 42, no. 2, pp. 299–311, Jan. 2002.

[90] S. Pal, P. S. Heyns, B. H. Freyer, N. J. Theron, and S. K. Pal, "Tool wear monitoring and selection of optimum cutting conditions with progressive tool wear effect and input uncertainties," *J. Intell. Manuf.*, vol. 22, no. 4, pp. 491–504, Sep. 2009.

[91] D. J. Mundfrom, D. G. Shaw, and T. L. Ke, "Minimum Sample Size Recommendations for Conducting Factor Analyses," *Int. J. Test.*, vol. 5, no. 2, pp. 159–168, Jun. 2005.

[92] J. W. Osborne and A. B. Costello, "Sample size and subject to item ratio in principal components analysis," *Pract. Assessment, Res. Eval. A peer-reviewed Electron. J.*, vol. 9, no. 11, p. 8, 2004.

[93] B. Liu, G. J. Keeler, J. T. Dvonch, J. a. Barres, M. M. Lynam, F. J. Marsik, and J. T. Morgan, "Temporal variability of mercury speciation in urban air," *Atmos. Environ.*, vol. 41, no. 9, pp. 1911–1923, Mar. 2007.

[94] H. Yu, J. Wang, H. Xiao, and M. Liu, "Quality grade identification of green tea using the eigenvalues of PCA based on the E-nose signals," *Sensors Actuators B Chem.*, vol. 140, no. 2, pp. 378–382, Jul. 2009.

[95] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the Marquardt algorithm.," *IEEE Trans. Neural Netw.*, vol. 5, no. 6, pp. 989–93, Jan. 1994.

[96] A. Suratgar, M. Tavakoli, and A. Hoseinabadi, "Modified Levenberg–Marquardt method for neural networks training," *World Acad Sci Eng Technol*, pp. 46–48, 2005.

[97] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Prentice Hall PTR Upper Saddle River, 1994.

[98] G. D. Garson, "Interpreting neural-network connection weights," *AI Expert*, vol. 6, no. 4, pp. 46–51, Apr. 1991.

[99] a. T. C. Goh, "Back-propagation neural networks for modeling complex systems," *Artif. Intell. Eng.*, vol. 9, no. 3, pp. 143–151, Jan. 1995.

Acknowledgement

The writing of this dissertation has been one of the most significant academic challenges I have ever had to face. I would never have been able to finish my dissertation without the guidance of my committee members, help from lab mates, scholarship by C.S.C. and support from my friends and family.

I would like to express my deepest gratitude to my advisor and the dissertation defense committee chair, Dr. YAMADA, for his excellent guidance, caring, patience, and providing me with an excellent atmosphere for doing research. His wisdom, knowledge and commitment to the highest standards inspired and motivated me.

I would like to thank my dissertation defense committee members, Dr. SHINOZAKI, Dr. SASAKI, Dr. HINO and Dr. TANAKA, for your brilliant comments and suggestions.

I would like to thank Mr. SEKIYA, for his management of the machines tool, devices and instruments in the lab and his help in my experiments.

I would like to thank Mr. KOJIMA and Mr. ISHII for their kind help with both the academic experiments and my living in Japan.

I would also like to express my gratitude to China Scholarship Committee for the financial support.

Many friends have encouraged, entertained, helped, and supported me during these years. I greatly value the friendships with them and I would like to thank Mr. XU Ming, Ms. ZHOU Yuchen and all my friends with my honor and respect.

My parents have encouraged me, corrected me and believed in me from the day I started learning, and I would never stop endeavoring to make them proud.

HAN Yi, my soul mate and my dearest love, who gives me energy and power to face the hardship, is the woman I would love till the last day of my life.