Computational intelligence for predictive condition monitoring and approaches for online analysis

Torabi Jahromi, Amin

2013

Torabi Jahromi, A. (2013). Computational intelligence for predictive condition monitoring and approaches for online analysis. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/62075>

<https://doi.org/10.32657/10356/62075>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

**Computational Intelligence for Predictive
Condition Monitoring and Approaches for
Online Analysis**

AMIN TORABI JAHROMI

SUPERVISORS:

PROF. ER MENG JOO

DR. LI XIANG

**SCHOOL OF ELECTRICAL AND ELECTRONIC ENGINEERING
DIVISION OF CONTROL AND INSTRUMENTATION**

2013

**COMPUTATIONAL INTELLIGENCE
FOR PREDICTIVE CONDITION
MONITORING AND APPROACHES FOR
ONLINE ANALYSIS**

AMIN TORABI JAHROMI

Supervisors:

Prof. Er Meng Joo

Dr. Li Xiang

**School of Electrical and Electronic Engineering
Division of Control and Instrumentation**

**A thesis submitted to Nanyang Technological
University in partial fulfillment of the requirement for
the degree of Doctor of Philosophy**

2013

Acknowledgment

This thesis marks the result of my research at the NTU-EEE. The work is the results of ideas, support, and encouragement of many people. I sincerely hope that I do not forget anyone who has helped me along the way. Firstly, I would like to sincerely express my deepest gratitude and appreciation to my advisor, Prof. Er Meng Joo. He has trusted and helped me to be his advisee in NTU when I was almost going to NUS. Without his kind trust and wise guidance, this work could not been as good as it is now. So, “Thank you Prof. !”

I thank Dr Li Xiang in her role as my Co-Advisor and Dr. Lim, both from SIMTech, for their advices on the practical sides of the experiments in this study.

I would also thank the technicians of “Intelligent Robotics Lab”, “Machine Vision Lab”, “SIMTech Machining Lab” and “SIMTech Research Cluster staff and students” for their kind help and kindness as my research has benefitted a lot from these labs and centers.

This project is supported by A*STAR Science and Engineering Research Council (SERC) Singapore-Poland Fund. The authors thank Dr. Saeid Javidi, Mr. Abdollah Khazen, Mr Huang Sheng, and Mr. Kah-Chuan Shaw

for their helpful guidance and assistance in the experiments. Appreciation is also due to the Alignment Tool (Singapore) Pte Ltd for their technical support.

Let me not forget to thank the reviewers of this thesis. Nothing could improve the understandability and quality of this thesis and my future works better than their wise and thoughtful comments and guidance.

The continuous support and encouragement of my family and specially my lovely wife and my dear little boy kept my motivation high. I admire them for all these. I also thank all my NTU, SIMTech and Iranian and Singaporean friends, who in one way or another, have been of a great help to me during my study in NTU.

The last but not least word cannot be anything except **“All the praises and thanks belongs to God the Lord of the universes”**.

Summary

The process of high speed milling (HSM) is regarded as one of the most sophisticated and complicated manufacturing operations. In the past four decades, many investigations have been conducted on this process aiming to better understand its nature and improve the surface quality of the products as well as extending tool life. To achieve these goals, it is necessary to form a general descriptive reference model of the milling process using experimental data, thermo-mechanical analysis, statistical or artificial-intelligent (AI) models. Besides, increasing demands for more efficient milling processes, qualified surface finishing, and modeling techniques have propelled the development of more effective modeling methods and approaches.

In the first part this dissertation, an extensive literature survey of the state-of-the-art modeling techniques of milling processes is carried out, more specifically of recent advances and applications of AI-based modeling techniques. The comparative study of the available methods as well as the suitability of each method for corresponding types of experiments is also presented. In addition, the weaknesses of each method as well as open research challenges are presented. Therefore, a comprehensive comparison of recent developments in the field will be a guideline for choosing the most suitable

modeling technique for this process regarding its goals, conditions, and specifications.

After comprehensive study of the available methods for modeling HSM processes, to build up a proper condition monitoring system, sensor signals are to be utilized to form a reference model which non-intrusively reflects the performance of the system. Therefore, a desired reference model has to apply more efficient feature extraction and artificial intelligence (AI) techniques to be more repeatable and generalizable. Since milling signals are complex, a time-frequency analysis method, namely wavelet, is applied for feature extraction.

Considering the high dimension of the wavelet features, clustering methods are used for dimension reduction and also as an interpretation layer between the signal feature extraction subsystem and artificial intelligence blocks.

This research illustrates the performance of artificial intelligence based techniques for modeling of high speed end milling experimental data. Studied and developed methods are applied on wavelet features of force and vibration signals to illustrate the repeatability and accuracy of their results. It is shown that the proposed structure as well as the developed artificial intelligent method can present the status of the process and can be applied for fault diagnosis and TCM purposes. It is also discussed that how application of available data mining methods with a proper structure may improve the performance of existing reference models towards more efficient utilization of available experimental data and easily generalizable reference models.

Contents

List of Figures	ix
List of Tables	xvii
1 Introduction	1
1.1 Experimental Set-ups, Signals, Preprocessing, and Feature Ex- traction	4
1.1.1 Experimental Design and Signals	4
1.1.2 Feature Extraction, Feature Selection and Preprocess- ing Methods	10
1.1.3 Signal Analysis: Time–Frequency Approaches, Fourier Transform, Short Time Fourier Transform, and Wavelet Analysis	11
1.2 Classic Models for a High Speed Milling Process	13
1.2.1 Analytical Modeling and Numerical Difference Equa- tion Solution Methods	14
1.2.2 Statistical and Experimental Evaluation	18
1.2.3 Multiple Regression and Time Series Modeling	21

1.3	Artificial-Intelligence Based Analysis of High Speed Milling Processes	23
1.3.1	Bayesian Networks	23
1.3.2	Fuzzy Logic, Neural and Fuzzy-Neural Networks-Based Methods	26
1.3.3	Evolutionary Algorithms, Genetic Algorithms, Genetic Programming and Particle Swarm Optimization	34
1.3.4	Hidden Markov Models	40
1.3.5	Discrete-Event Based Intelligent Methods	52
1.3.6	Clustering and Classification Methods	54
1.4	Conclusions	57
2	Proposed Methods and Results	60
2.1	Experiment Setup	60
2.2	Continuous Wavelet Analysis, Flute-Cut separation, Energy Feature Extraction	66
2.3	Neural Network and Adaptive Neuro-Fuzzy Inference System based models and results	70
2.3.1	Neural Network Based Methods	70
2.3.2	Problem Description and Proposed Neural Network based Techniques	74
2.3.3	Discussions on Results of Neural Network Based Methods for milling processes	75

2.3.4	Application of Cascaded Feed Forward Neural Network (CFFNN) for Correlation of CWT Energy Features and Tool Wear	81
2.3.5	Conclusion on the application of Neural Network Based Methods for milling processes	87
2.4	Application of Clustering Methods on Wavelet Features	89
2.4.1	Problem Description on Clustering of Wavelet features	92
2.4.2	Proposed Clustering-Based Reference Model	94
2.4.3	Clustering Results and Discussions	102
3	Sequential Fuzzy C-Means Clustering Based DFNN	116
3.1	Sequential Fuzzy C Means Clustering	116
3.2	Introducing a new simplified Dynamic Fuzzy Neural Network based on sequential Fuzzy C-Means Clustering	120
3.2.1	Learning Algorithm	123
3.3	Application of Sequential FCM based DFNN for Modeling a HSM process	132
4	Conclusions	139
4.1	Conclusions and Contributions	139
5	Future Work Plans	165
6	Published Papers and Honors	169
6.1	Published papers	169
6.2	Honors and Awards	173

A Nomenclature	175
B Abbreviations and Acronyms	179
C Sensor Specifications	183
C.1 Dynamometer [Kis, a]	183
C.2 Accelerometer [Kis, b]	188
Bibliography	192

List of Figures

1.1	Tool condition monitoring and surface roughness prediction [Zhu et al., 2009].	2
1.2	Surface profile measurement details [Zain et al., 2010, Yang and Chen, 2001].	6
1.3	The factors which affect surface roughness and tool wear. . . .	7
1.4	Origins of the AE signal. Arrows indicate emitted acoustic waves [Diei and Dornfeld, 1987, Xiaoli and Zhejun, 1998, Pai and Rao, 2002, Mathew et al., 2008, Li, 2002, Cheng, 2009]. . .	9
1.5	Wavelet-based localization of the flute signals. Left is 3D presentation of the signal. Center is peak localization. Right is one of scales to show the effects of wavelet on denoising the signal. [Torabi Amin J., 2010b, Torabi Amin J., 2010a]	13
1.6	Analytical analysis of tool wear [Abrari et al., 1998], results of mathematical finite element and thermal softening of the milling processes, [Aurich and Bil, 2006, Lim et al., 1995] . . .	15

1.7	(a) An example of a sharp cutting edge, (b) An example of a worn cutting edge, a statistical approach is used for representation of the experimental results: (c) Cutting force increasing linearly with cutting distance, [Li et al., 2006a], (d) Development of flank wear under cutting speed 900 rpm [Li et al., 2006a].	18
1.8	TAN model for surface roughness, [Correa et al., 2009]. Surface roughness (R_a), Feed per tooth (fz), Depth of cut (D_c) tool diameter (diam), material hardness (HB), geometry radius ($Radius$) and geometry curvature ($geom$), spindle speed (rpm) and the resulting forces applied to all the cutting planes F_{XY} [Correa et al., 2009].	25
1.9	Multilayer feed-forward ANN structure used in [Zuperl and Cus, 2003] to model surface roughness (R_a).	28
1.10	Fuzzy-Nets system proposed by [Yang et al., 2006] to adaptively control the surface roughness according the predicted values for surface finishing.	31
1.11	Proposed GP-based structure to represent a nonlinear function.	35
1.12	Proposed GP-based structure to represent a nonlinear function.	37
1.13	The obvious change in the resulting signals from the beginning of the process to the end: I- Cutter No.1: (a) The 40th cut of the process with cutter No.2, (b) The 600th cut of the process with cutter No.2. II- Cutter No.2: (c) The 40th cut of the process with cutter No.1, (d) The 600th cut of the process with cutter No.1	50

1.14 A technique based on AFPNs for modeling and control of product and machining process quality [Kasirolvalad et al., 2006]. V: Vibration, F: Cutting Force , S: Spindle Speed, f: feed rate, t: MachiniR: Surface Roughness, Q: Machining Quality	53
2.1 Workpiece and milling tool, cutting direction, and cutting method	61
2.2 Workpiece and the arrangement of the sensors.	61
2.3 Surface quality measurements, Taylor-Hobson surface measurement instrument	63
2.4 Surface quality measurements, Motoyo surface measurement instrument	63
2.5 Surface profile measurement details [Zain et al., 2010, Yang and Chen, 2001].	64
2.6 Tool wear being measured using a 2D microscope	64
2.7 Alicona InfiniteFocus device for tool wear measurements.	65
2.8 Tool wear being measured with Alicona InfiniteFocus microscope	65
2.9 The Raw cutting Signal.	67
2.10 The obvious change between the sensor signal from the beginning of the cutting process to the end. (a) The 40th cut of the process with cutter 3T-12. (b) The 600th cut of the process with cutter 3T-12.	68

2.11	The obvious change between the sensor signal from the beginning of the cutting process to the end. (a) The 40th cut of the process with cutter 18-SC3. (b) The 600th cut of the process with cutter 18-SC3.	69
2.12	Multilayer feed-forward ANN structure use by [Zuperl and Cus, 2003] to model the surface roughness (R_a).	72
2.13	Bior Wavelet Analysis of the Cutting Force Signal in 2D and 3D, (a) Bior wavelet analysis of the cutting force signal in 2D. (b) Bior wavelet analysis of the cutting force signal in 3D. . .	76
2.14	Meyer Wavelet Analysis of the Cutting Force Signal in 2D and 3D, (a) Meyer wavelet analysis of the cutting force signal in 2D. (b) Meyer wavelet analysis of the cutting force signal in 3D.	76
2.15	Coiflets wavelet analysis of the cutting force signal in 2D and 3D (a)Coiflets wavelet analysis of the cutting force signal in 2D (b)Coiflets wavelet analysis of the cutting force signal in 3D	77
2.16	Daubechies wavelet analysis of the cutting force signal in 2D and 3D, (a) Daubechies wavelet analysis of the cutting force signal in 2D. (b) Daubechies wavelet analysis of the cutting force signal in 3D.	78
2.17	Morlet wavelet analysis of the cutting force signal in 2D and 3D, (a) Morlet wavelet analysis of the cutting force signal in 2D. (b) Morlet wavelet analysis of the cutting force signal in 3D.	78
2.18	Wavelet-Based Analysis of the Cutting Force Signal in 2D and 3D and Localization of the Flutes	79

2.19 Separated Revolution Signals of a 2-Flute Cutter.	80
2.20 Tool-Wear Prediction based on the Wavelet Features, Total System Schematic. (a) Tool-Wear Prediction based on the Wavelet Features. (b) Cascaded Feed Forward Neural Network.	83
2.21 The prediction of the Tool-Wear based on Wavelet Energy Features. Test is Conducted on the Same Cutter.	84
2.22 The prediction of the Tool-Wear based on Wavelet Energy Features. Test is Conducted on another Cutter.	84
2.23 The prediction of the Tool-Wear based on Wavelet Energy Features. Test is Conducted on the Same Cutter.	85
2.24 The prediction of the Tool-Wear based on Wavelet Energy Features. Test is Conducted on another Cutter.	85
2.25 Proposed Method Using Fuzzy C-Means Clustering Layer as Interpretation Layer.	96
2.26 Wavelet Analysis of the Entire Cutting Portion of the Signal. The top is the 2D wavelet analysis of the signal. Below is the original signal	97
2.27 Wavelet-based localization of the flute signals. Top-left is 3D presentation of the signal. Top-right is peak localization. Be- low is one of scales to show the effects of wavelet on denoising the signal.	98
2.28 Fuzzy C-means clustering is applied on extracted features of a destructive-test force signal with 3 classes.	104
2.29 Fuzzy C-means clustering is applied on extracted features of three destructive-test force signal with 3 classes.	105

2.30 Fuzzy C-means clustering is applied on extracted features of a destructive-test vibration signal with 3 classes.	106
2.31 Fuzzy C-means clustering is applied on extracted features of three destructive-test force signal with 3 classes.	107
2.32 K-means clustering is applied on extracted features of a destructive-tests signal features with 3 clusters. Crisp clustering results are shown. To present the density of the presence of cluster members, the distribution of presence of each individual cluster is also depicted above.	108
2.33 K-means clustering is applied on extracted features of four destructive-tests signal features with 3 clusters. As it is crisp clustering method, distribution of presence of each individual cluster is depicted above. X-axis shows the sequence of incoming samples that can represent time.	108
2.34 K-medoid clustering is applied on extracted features of a destructive-tests signal features with 3 clusters. Crisp clustering results are shown. To present the density of the presence of cluster members, the distribution of presence of each individual cluster is also depicted above.	110
2.35 K-medoids clustering is applied on extracted features of four destructive-tests signal features with 3 clusters. As it is crisp clustering method, distribution of presence of each individual cluster is depicted above. X-axis shows the sequence of incoming samples that can represent time.	110

2.36	Fuzzy C-means clustering is trained on extracted features of one destructive-test force and vibration signal with three clusters. It is verified on the above four destructive signals. X-axis shows the sequence of incoming samples that can represent time.	111
2.37	Gath-Geva clustering is trained on extracted features of one destructive-test force and vibration signal with three clusters. It is verified on the above four destructive signals. X-axis shows the sequence of incoming samples that can represent time.	112
2.38	An example of constructing IF-THEN rules from input/output pairs when the input data can be grouped into limited number of clusters [Wang, 1999], U represents input space, V represents output space and A_i^j and B^k s are the fuzzy sets defined on the input and output space, respectively.	114
3.1	Absorbion version of updating equation for cluster centers. .	118
3.2	The architecture of a dynamic fuzzy neural network [Wu et al., 2001]	122
3.3	Flowchart for sequential fuzzy c-means clustering based dynamic fuzzy neural network	131
3.4	Fuzzy C Means based DFNN Results of Wavelets Energy Features, Mean Relative Error (MRE): 91.8115%	133
3.5	Fuzzy C Means based DFNN Results of Wavelets Energy Features, Mean Relative Error (MRE): 97.4951%	134

3.6 Fuzzy C Means based DFNN Results of Wavelets Energy Features, Mean Relative Error (MRE): 95.7210%	134
---	-----

List of Tables

1.1	Mathematical Modeling Approaches to Machining Processes .	17
1.2	Experimental Modeling Approaches to Machining Processes .	20
1.3	Multiple Regression and Time Series Modeling Approaches to Machining Processes	22
1.4	Artificial Neural Networks and Support Vector Machine Mod- eling Approaches to Machining Processes	33
1.5	Fuzzy Neural Network Modeling Approaches to Machining Process	34
1.6	Genetic Algorithms, Fuzzy Petri Net, Particle Swarm Opti- mization Modeling Approaches to Machining Processes	39
1.7	Bayesian Networks and Hidden Markov Modeling Approaches to Milling Processes	44
1.8	Clustering Modeling Approaches to Machining Processes . . .	56
1.9	Advantages and Disadvantages of the AI Methods	58

Chapter 1

Introduction

¹Today, high speed machining (HSM) is widely applied to fulfill the overwhelming and increasing demands for producing vital pieces for various industrial sectors, specially in aerospace industries. The throughput of the machining process is a critical parameter for determining the quality of a production process. Large throughput as well as the surface quality of the product are directly related to the change in the total production rate and the overall gain. Early research in this area started in the late 70s and early 80s [Kline et al., 1982]. Afterwards, many approaches were proposed for the production process to improve performance and achieve the desired quality and final mass production. The common objective was to achieve better surface finishing and extended tool life.

A literature survey of the most popular information extraction and modeling techniques in this area is beneficial for clarifying the research issues and illustrating their weaknesses and achievements. The main goal of this

¹This chapter is consolidated into a journal paper and is now available online [Torabi Amin J., 2013b]

Figure 1.1: Tool condition monitoring and surface roughness prediction [Zhu et al., 2009].

chapter is to consolidate the available knowledge on modeling techniques of milling processes. This knowledge facilitates the extraction of the inherent relationship between all the effective cutting parameters, sensor signals, and process results of the milling process by choosing the most appropriate modeling technique [Liang et al., 2004, Torabi Amin J., 2009, Li, 2002, Prickett and Johns, 1999, Teti et al., 2010]. As a result, it will be easier to choose the proper Artificial Intelligence (AI) or classical approach for presenting the experimental results and providing the best reference model.

There are numerous modeling methods and approaches to provide a reference model for milling processes. Many of the state-of-the-art methodologies

will be covered in the present chapter.

One of the factors that distinguishes these methods is their applied feature extraction and data preprocessing approaches. After discussing different experimental set-ups, output signals, and other experimental issues in Section 1.1.1, some frequency-, time-, and time-frequency-based signal pre-processing and feature extraction methods will be covered in Sections 1.1.2 and 1.1.3.

Another important factor for grouping modeling methods is the algorithm which they use. Numerous modeling methods are applied to provide a non-intrusive monitoring means of tool life and surface performance in milling process. Analytical and physical modeling and their solution approaches such as partial differential equations and finite-element methods will be mentioned in Section 1.2.1. Their basic focus is the physical nature of the cutting phenomena and its mathematical aspects. In the following section, the statistical and experimental modeling and verification methods will be pointed out. These methods as covered in Section 1.1.1 are mostly used for preliminary illustration of the performance of the process as well as the study and verification of proposed approaches based on previous studies. Another common technique which is used for milling process modeling is multiple regression, which determines the optimal parameters for a predefined mathematical model. Basically it applies regression analysis to estimate unknown parameters of the model from the experimental data as described briefly in Section 1.2.3. Probabilistic modeling methods such as Bayesian networks (BNs) and hidden Markov models (HMM) will be summarized in Sections 1.3.1 and 1.3.4. They apply probability rules and relations to form a model for milling process monitoring and prediction. However, these meth-

ods are not as common as the methods based on neural networks, fuzzy logic, and their combinations, which will be covered in Section 1.3.2. Evolutionary approaches, such as genetic algorithms (GA) and particle swarm optimization (PSO), are also applied in this field. As Section 1.3.3 will present, they are mostly used in combination with other methods for optimization purposes. Rarely used are Discrete Event Systems (DEVS) and Petri Nets (PNs), which will also be covered in Section 1.3.5. The application of these methods to the field of milling processes depends on the interpretation of the events and transitions in the process. Different types of clustering methods and algorithms are also applied in milling process modeling. They are mostly applied to the signal features as the first layer for signal interpretation. Categorization and grouping of distinct signal features and associating them with different cutting phenomena are also the goals of the researches summarized in Section 1.3.6.

Finally, in Section 4.1, discussions of available techniques and research issues and some suggestions for future studies will be presented. Conclusions will be drawn in Section 1.4.

1.1 Experimental Set-ups, Signals, Preprocessing, and Feature Extraction

1.1.1 Experimental Design and Signals

Fig.1.1 gives an overview of the methodology in experimental knowledge extraction for developing a reference model for the process. The experimental

set-up is a crucial part of a machining technology investigation. Its design is highly related to the demands of industry and/or research interests. Therefore, a proper experimental design will facilitate the study of many aspects of the process. For example, the tool geometry, workpiece and tool material and coating, HSM centre specifications, cutting conditions, dry/semi-dry/wet machining, signals of interest, choice of sensors and sampling period, cutting parameters, and surface quality indicators are very important parameters to be determined before the experiment is initiated [Chakraborty et al., 2008, ASME, 1996, Zain et al., 2010, Yang and Chen, 2001, Ng et al., 2000].

As mentioned above, the surface quality which results is one of the crucial factors in this field. To quantify surface quality, many important tests are commonly applied, like surface roughness measurement R_a , X-Ray Diffraction (XRD), and Environmental Scanning Electron Microscopy (ESEM) [Chakraborty et al., 2008]. The last two are comparatively expensive and time-consuming and hence do not fulfill the cost per production requirements of the industry: only a limited number of researches covered these results [Chakraborty et al., 2008, ASME, 1996]. By contrast, surface roughness measurements such as R_a , R_v , and R_t , are not so costly and can be easily conducted in industrial environments [ASME, 1996, Zain et al., 2010]. They give a rough indication of the product quality and surface finishing [Zain et al., 2010, Yang and Chen, 2001].

Among the other roughness measurements, R_a is the most commonly used surface quality indicator. It is the arithmetic mean of the absolute ordinate values of the surface profile $Y(x)$ as presented in (1.1) and depicted in Fig.1.2. Its unit of measurement is the micrometer (μm) and the standards on surface

roughness are accessible in ISO:1302 4288:1996 [iso,].

$$R_a = \frac{1}{L} \int_0^L |Y(x)| dx \quad (1.1)$$

Figure 1.2: Surface profile measurement details [Zain et al., 2010, Yang and Chen, 2001].

It is experimentally established that when tool wear increases, it directly damages surface quality. Therefore, measuring the tool wear is another alternative for milling process quality measurement. However, surface quality and tool wear measurements are not easily carried out on-line on the workpiece and tool during the production process. Therefore, much research effort was directed to the relationship between them on the one hand and other process information on the other hand.

Practically, many factors affect the surface quality. The most famous ones are presented in Fig.1.3. A more general and comprehensive model or Design of Experiment (DoE) shall encompass more of these parameters.

In addition, the experimental set-up mentions the specifications of the machining centres, workpieces, and cutters. It also elucidates the milling

Figure 1.3: The factors which affect surface roughness and tool wear.

method and the properties of the workpiece. Furthermore, various tool geometries, coatings, cooling methods, cutting directions and angles, data acquisition methods, and equipments as well as the model of the sensors are usually included in the set-up descriptions. These sensors capture signals that reflect the internal behavior and hidden internal states of the milling process. Along with the cutting process information, they were usually applied to develop on-line tool status monitoring systems. These systems used pre-trained models to predict tool wear or surface roughness non-intrusively [Dey and Stori, 2005, Correa et al., 2008, Ho et al., 2009, Kasirolvalad et al., 2006, Kirby et al., 2006]. Another application of these models were to adjust the cutting conditions in a way that the desired surface roughness remained secure [Ho et al., 2009, Kirby et al., 2006].

There are many signals of interest in the milling process, such as acoustic emission, force, vibration, temperature [Zuperl et al., 2011, Rivero et al., 2008], and spindle and feed motor currents and voltage [Squillaro, 1993, McDowell, 1999, Shell, 1987, Bhattacharyya et al., 2008]. Among these, acoustic emission can be collected at low cost using easily installed high frequency acoustic sensors. Plastic deformation in the workpiece and chip, friction contact of flank face and workpiece, which lead to flank wear, friction of the tool rank face and the chip, which lead to crater wear, collisions between the tool and workpiece, tool edge chipping and chip breakage, were mentioned as possible origins of this signal, Fig.1.4, [Diei and Dornfeld, 1987, Xiaoli and Zhejun, 1998, Pai and Rao, 2002, Mathew et al., 2008, Li, 2002, Cheng, 2009].

Some of the milling signals such as force and vibration are the most popular signals for experimental research whereas other sensor signals such as acoustic emission and feed and motor currents are more common in industrial environments [Teti et al., 2010]. A lot of data acquisition methods and sensor devices were developed to deal with different process signals. Usually, all the information on the applied sensors appear in the experimental set-up part of the reports. Force sensors [Dey and Stori, 2005, Correa et al., 2008, Ho et al., 2009, Kasirolvalad et al., 2006, Kirby et al., 2006], vibration sensors [Dey and Stori, 2005, Correa et al., 2008, Ho et al., 2009, Kasirolvalad et al., 2006, Kirby et al., 2006], and acoustic emission sensors [Dey and Stori, 2005, Correa et al., 2008, Ho et al., 2009, Kasirolvalad et al., 2006, Kirby et al., 2006, Squillaro, 1993] were so far the most commonly applied, while electrical power and current [Squillaro, 1993, McDowell, 1999, Shell, 1987], workpiece and cutter temperature and thrust used to be more common in

older researches [Squillaro, 1993, McDowell, 1999, Shell, 1987, Rivero et al., 2008]. Having collected these signals, it is possible to correlate them with surface roughness or tool wear. After the proposed reference models are trained with signal features and process results, they are applied to facilitate non-intrusive prediction of the state of the produced surface and the tool during the milling process. The most famous modeling and feature extraction techniques will be discussed in the next sections for their advantages and disadvantages. It is clear that some of the available methods have some weaknesses due to the lack of technical developments (which might, however, be improved in current or future advances in AI modeling) to match with the characteristics of milling process. A discussion of intelligent machining research issues will be given in Section 4.1.

Figure 1.4: Origins of the AE signal. Arrows indicate emitted acoustic waves [Diei and Dornfeld, 1987, Xiaoli and Zhejun, 1998, Pai and Rao, 2002, Mathew et al., 2008, Li, 2002, Cheng, 2009].

1.1.2 Feature Extraction, Feature Selection and Pre-processing Methods

One useful approach to avoid a huge volume of signal data is to extract their most informative features to be used as the inputs of the reference model. As the raw data usually consists of continuous multi-dimensional complex signals, it is preferred to become noiseless well-structured dataset with the least possible dimensions.

Besides, time-based statistical features such as the mean, variance, kurtosis, skew, energy, etc. [Li et al., 2010], wavelet transform, Fourier transform, and noise filtering were the most common data processing methods used for denoising raw signals and extracting their features [Bassiuny and Li, 2007, Ng et al., 2000]. Then, feature selection methods such as Principal Component Analysis (PCA), mutual information/correlation, or evolutionary based optimization algorithms, were also applied to highlight the most descriptive features for application in other modeling layers. For example in [Zhou et al., 2006], the selection method was to apply a genetic algorithm (GA) for optimum feature subset selection with a modified wrapper-based multi-criteria approach. This method was also applied in [Li et al., 2010] to determine the features which were most correlated with the output parameters such as tool wear and surface roughness.

Some signal preprocessing actions were necessary before information extraction, e.g., the Hilbert transformation [Squillaro, 1993, Bassiuny and Li, 2007], wavelet analysis [Bassiuny and Li, 2007], or Fourier analysis [Yamaguchi et al., 2007, Haber et al., 2004].

To avoid redundant data and the resulting high computational complexity, the consistency of the available data must be verified. As an example, a Taguchi-based technique to check the dependencies between signal features was used in [Al-Habaibeh and Gindy, 2000]. The results showed that if the average inter-dependency of the feature data, which was the average of the elements of the ASM (Association Matrix) in the OAs (Orthogonal Arrays) method were comparatively high, the classification would be more successful. The more successful the classification, the more accurate the related modeling and prediction will be. Applying time-frequency analysis, it was also necessary to clean the noisy signal from the high peaks caused by chip breakages because of their interference with the classification methods [Xiaoli and Zhejun, 1998].

1.1.3 Signal Analysis: Time–Frequency Approaches, Fourier Transform, Short Time Fourier Transform, and Wavelet Analysis

The time series of signal samples from the milling signal are usually not easily interpreted without being processed by a signal processing method. Hence, time–frequency analysis of the signals are commonly used. An example of such methods is wavelet analysis, which were applied in recent years to many industrial process signals to provide meaningful information. It was experimentally established that wavelet analysis provides more informative data on the internal behavior of the process than the Fourier transform [Wu and Du, 1996, Chen and Li, 2007, Patra et al., 2007]. Yet, there were re-

searches that relied on pure frequency domain analysis (such as the Fourier Transform) [Li, 2002, Xiaoli and Zhejun, 1998, Cho et al., 2005, Merdol and Altintas, 2004, Morgan et al., 2007]. Details on the wavelet analysis method can be found in [Daubechies, 1990].

Extensive investigations was conducted on the advantages of employing wavelets. [Kasashima et al., 1995, Paul, 1994, Li, 2002].

For example, it was shown in that the energy content of each scale in wavelet analysis varies due to changes in tool wear. Therefore, power and vibration signals were decomposed into a finite number of scales, [Wang et al., 2002, Patra et al., 2007]. After detecting some signal features, the features were compared to a pre-trained feature bank called the codebook. Then, the state of the tool was decoded. High condition detection probability of the method that applied vector quantization (VQ) of wavelet features and the hidden Markov models (HMM) was presented. Also in [Wang et al., 1998, Li et al., 2000], the average energy of each scale in the discrete wavelet transform (DWT) of a vibration signal was shown to be a blueprint of the status of the tool.

The wavelet modulus maxima distribution was used in [Miao and Makis, 2007], where the method relied on the local maxima of the wavelet analysis, instead of directly using the features of the Continuous Wavelet Transform (CWT). Obvious changes in the time–frequency patterns were used as a sign for changes in tool condition [Li and Guan, 2004]. As depicted in Fig.1.5, a CWT analysis of the cutting signal was applied for peak localization and extraction of the energy features of several scales of the signal to be used as a pattern for tool changes.

Figure 1.5: Wavelet-based localization of the flute signals. Left is 3D presentation of the signal. Center is peak localization. Right is one of scales to show the effects of wavelet on denoising the signal. [Torabi Amin J., 2010b, Torabi Amin J., 2010a]

Wavelet packets were also applied, although rarely, for feature extraction in milling processes [Xiaoli and Zhejun, 1998]. In wavelet packets, a linear combination of basis wavelets are used as basis vectors and one aims at keeping the original properties of their parent wavelets. In addition to the above mentioned signal analysis techniques, Hilbert or Gabor transformations were used, although rarely, for condition detection [Squillaro, 1993, Bassiuny and Li, 2007, Serrano and Fabio, 1996]. However, as they are solely frequency domain analysis, the results of these techniques are not superior to those of wavelet analysis (which is time-frequency analysis).

1.2 Classic Models for a High Speed Milling Process

There are many classic methods for presenting the characteristics of milling processes. Their types and characteristics are determined regarding the infor-

mation that the model provides. The following subsections will cover many papers that only present the raw results of experiments or pre-processed signals for better understanding of the nature of the process as well as others that utilize information extraction techniques to determine the internal behavior of the process. The following subsections will briefly describe the classical methods, their methodology, applications, as well as their pros and cons.

1.2.1 Analytical Modeling and Numerical Difference Equation Solution Methods

Methodologies and Applications

A mathematical representation is a theoretical way, based on the physics of the milling process, to model it. In some papers, thermo-dynamical behavior of the cutting tool and workpiece were studied [marusich, 1995, Aurich and Bil, 2006, Yan and Strenkowski, 2006, Soo et al., 2004, Sutherland, 1988, Abrari et al., 1998, Zhao et al., 2010]. Thermo-mechanical equations, such as the mathematical dynamics and static equations regarded as two important aspects of the process, were represented and coupled together. Afterwards, two distinct analysis approaches usually were followed. The first approach was to solve the nonlinear partial differential equations according to specified boundary conditions [Lim et al., 1995, Sutherland, 1988, Abrari et al., 1998, Fontaine et al., 2007]. As a result, many interesting issues such as tool wear, surface profile, chip formation and geometry, and cutting force, could be determined mathematically. However, since most of these equations do not have analyti-

cal solutions, as a second approach, numerical methods were applied to solve these equations. Consequently, continuous partial differential equations were changed to finite-element discretized difference equations where temperature and surface membrane were state variables of the system [marusich, 1995, Aurich and Bil, 2006, Yan and Strenkowski, 2006]. Afterwards, changes in the shape of the surface, the formation of the chips, and the heat profile of the tool and workpiece were studied using computer simulations as depicted in Fig.1.6. In simpler studies, the resulting surface roughness or cutting force were looked for in 1D cutting simulations [Baek et al., 2001, Lim et al., 1995]. In these kinds of simulations, the tool tip was usually assumed to be a rigid material and only the tip-print on the face of the workpiece was tracked and its accuracy verified [Baek et al., 2001].

More recent simulations showed that if the analysis was extended to three dimensions, the results would be in better agreement with experiment. In [Soo et al., 2004, Liu and Loftus, 2006, Ko and Cho, 2005, Zhao et al., 2010], an accurate finite element solution to present Lagrangian equations in three dimensions was used and experimental verifications with a ball-nose end

Figure 1.6: Analytical analysis of tool wear [Abrari et al., 1998], results of mathematical finite element and thermal softening of the milling processes, [Aurich and Bil, 2006, Lim et al., 1995]

milling process were performed. Unlike the research in [Soo et al., 2004], [Liu and Loftus, 2006] used a more accurate model in which the rigidity of the cutting edge was not assumed, and the effects of the cutting tool geometry and material were considered in more detail. To ensure the validity of the model, both simulated and experimentally cut surfaces were compared according to their visible shapes, [Liu and Loftus, 2006]. The same studies were conducted on a dynamic model for a machining system, with a nonlinear force model and geometrical model of the trace of the cutter, and the resulting surface profile was depicted and analyzed for its changes with various cutting conditions [Paris et al., 2004].

Another research issue in this area is the formation of chips and deformations of the tool and workpiece as well as the resulting forces [Lim et al., 1995]. Here, the geometrical aspect of the milling process and the movement of the tool on the workpiece and milling period were more discussed, Fig.1.6. In [Soo et al., 2004], the ball-nose cutter performance was considered based on its geometrical and chemical characteristics. Based on geometry of the tool and cutting conditions, the amount of material which could be affected by the cutting edge was computed. On the other hand, further extension in this area led to application of two cutting-force models in parallel. One model was used to capture the rigidity of the media and the tool and another one was intended to capture the fluctuations in the results of the first model according to imperfections of the material.

Table 1.1: Mathematical Modeling Approaches to Machining Processes

Papers	Material	Conditions	Cutting Conditions	Centre	Signals	Analysis Target	Preprocessing Technique
[Yang and Chen, 2001]	6061 Aluminium	NM*	Vc Cn Fc Dc Td*	Fadal VMC-40	NA*	Ra*	NA
[Bassiuny and Li, 2007]	NM	Dry	Flutes	Mazak AJV 25/405	Feed-motor current	Tool breakage	Low pass filter, Hilbert transform
[marusich, 1995]	AISI 4340	NM	Cutting conditions, Geometry	NA	NA	Force, Chip formation, Temperature	Finite element techniques
[Aurich and Bil, 2006]	NM	Dry	Cutting conditions, Geometry	NA	NA	Force, Chip formation, Temperature	Finite element techniques
[Yan and Strenkowski, 2006]	NM	NM	Rake angle	NM	NM	Strain, Crack formation	Finite element techniques
[Soo et al., 2004]	Inconel 718, Titanium	Dry	Cutting conditions, Geometry	NA	NA	Stress, Chip formation, Strain	Finite element techniques
[Lim et al., 1995]	Steel	Dry	Boundaries, Cutting conditions, Geometry	Cincinnati Milacron	NA	Force, Chip formation	NA
[Liu and Loftus, 2006]	P20 hard tool steel	NM	Boundaries, Cutting conditions, Geometry	Hermle C600U	NA	Surface profile	NA
[Paris et al., 2004]	Aluminium	NM	Cutting conditions, Geometry	Huron KX8-FIVE	NA	Force, Chip formation	NA
[Yang et al., 2006]	Graphite ISO-680	Dry	Vc Fc Dc	NM	NA	Ra	Grey relational analysis
[Baek et al., 2001]	Steel (AISI1041)	NM	Cutting conditions, Geometry	NM	NA	Surface profile	NA

* See the Abbreviation Appendix

Pros and Cons

Aside from all its advantages, such as pre-estimation of the critical moments during the simulated milling process models and being less costly for seeing the estimated surface profile before any actual cutting, this approach is not practical. The reason is the available imperfections in the industrial processes and difficulties of perfect modeling and simulation of every individual process. Another weakness of this approach is that it cannot be directly extended to real-time monitoring. Therefore, these studies are not directly applicable to mass-production industrial environments.

(a)

(b)

(c)

(d)

Figure 1.7: (a) An example of a sharp cutting edge, (b) An example of a worn cutting edge, a statistical approach is used for representation of the experimental results: (c) Cutting force increasing linearly with cutting distance, [Li et al., 2006a], (d) Development of flank wear under cutting speed 900 rpm [Li et al., 2006a].

1.2.2 Statistical and Experimental Evaluation

Methodologies and Applications

As another non-AI approach, the statistical properties of surface measurement were often considered. Simple basic computations were performed on experimental results while pre-processing techniques were used to illustrate the differences of each individual experiment. This method was used in [Ng et al., 2000], where an end-milling process on Inconel 718 was investigated.

One of the challenging issues of milling hard-to-cut super-alloys is the heat generated in the cutting point which causes the tool to be welded onto the workpiece. Hence, the cutting direction and alignment of the workpiece against the tool as well as the coating material of the tool are studied as vital factors in the process. Horizontal downward cutting direction was compared to other cutting directions and it was determined to be the best milling direction for enhancing tool life [Ng et al., 2000]. The same method was used in [Li et al., 2006a] for representation of the experimental results of tool wear versus milling time. Figs.1.7(a) and 1.7(b) depict the progressive tool wear during the milling process. From the results which are depicted in Figs.1.7(c) and 1.7(d), it can be seen that an interesting similarity exists between the nonlinear nature of the tool wear and surface roughness, measured tool wear and sensor signal, namely force in [Li et al., 2006a] where peak value of force signal was applied for tool wear detection for its variations in each successive cut or equivalently with the cutting distance, Fig.1.7(c).

Similarly, there were many attempts to present these complex nonlinear relations for application in surface quality control [Rao et al., 2011, Nalbant et al., 2007]. For example, a non-intrusive monitoring system was proposed by [Nalbant et al., 2007] for observing force signal features and their relation to cutting speed. Finding an optimal range of cutting conditions which leads to the desired surface roughness and the least cutting force is another challenge in investigations of milling processes [c Yang, 2006, Baek et al., 2001, Zhang et al., 2007]. Normally, the investigations of the optimum cutting parameters are associated with proper DoE methods. Some of these methods are based on simple optimization techniques such as bisection in terms of the

resulting surface roughness [Baek et al., 2001].

Pros and cons

Although these methods are extensively used in the field for simulation, they are only useful when the goal is to present observable changes of the work-piece and similarities of simultaneous and preliminary experiment results. Although it is clear that they can present the performance and depict interesting aspects of the milling process, they can never be used for prediction and on-line applications because they are not represented in lumped mathematical functions.

Table 1.2: Experimental Modeling Approaches to Machining Processes

Papers	Material	Conditions	Cutting Conditions	Centre	Signals	Analysis Target	Preprocessing Technique
[Ng et al., 2000]	Inconel 718	Dry	Vc Cn Fc* Cutting direction Cn	MATSURA FX	Force, Vibration	Flank wear	FFT*
[Rivero et al., 2008]	Al 7075-T6	Dry	Cn	Fatronik	Force, Power, Temperature	Tool wear, Tool breakage Tool life	Time-Frequency analysis Power spectrum
[Yamaguchi et al., 2007]	NM*	NM	NM	JTEKT AHP	AE*, Force	Tool wear, Tool breakage Tool life	FFT
[Haber et al., 2004]	Steel	Dry	Cutting conditions	HS1000 Kondia	AE, Force	Tool wear, Tool breakage Tool life	FFT
[Li et al., 2006a]	Inconel 718	Dry	Cn	Mitsubishi BAE500	Force	Tool wear, Tool breakage Ra*	NM
[c Yang, 2006]	Graphite	Dry	Vc Fc Dc	NM	NA*	Ra	Grey relational analysis Taguchi design procedure Wavelet packets
[Zhang et al., 2007]	AISI H13	Wet	Vc Cn Fc Dc Td*	Fadal VMC-40	Temperature, Tool wear Power, Vibration	Tool wear	Discrete wavelet transform Wavelet analysis
[Wu and Du, 1996]	Steel	NM	Dc Vc Fc	NM	Force	Tool failure	Force dynamical model
[Kasashima et al., 1995]	Steel	NM	Dc Vc Fc	NM	Force	Tool failure	Force dynamical model
[Paul, 1994]	AISI 4340	NM	NA	NM	AE	Tool wear	Force dynamical model
[Sagherian and Elbestawi, 1990]	Aluminium	NM	Cutting conditions, Geometry	NM	Force	Surface profile	Force dynamical model
[Nalbant et al., 2007]	Inconel 718	Dry	Dc Vc Fc	T35 lathe CNC	Force	Force	NA

* See the Abbreviation Appendix

1.2.3 Multiple Regression and Time Series Modeling

Methodologies and Applications

For many years, multiple regression was the preferred statistical method to present milling process reference models because of its simplicity and the simple structure supposed for milling process models. In multiple regression method, the results of a milling process are basically supposed to be a linear combination of cutting parameters and signal features or their mathematical functions can be found in studies such as [Chakraborty et al., 2008, Mike S. Lou, 1998]. (However, in some papers this relation is not considered linear any more and cross-multiplication of the inputs and their nonlinear functions were also considered [Routara et al., 2009].) This makes the modeling simple and hence inaccurate compared to available AI approaches as the studies in [Muammer Nalbant and Toktas, 2007, Palanisamy et al., 2008, Asiltürk and Çunkas, 2011] presented.

Regression method is not only applicable to simple time features but also to more complex features of the signals or cutting conditions, such as the energy content of different scales in the wavelet decomposition of the signals. In [Choi et al., 2004], using wavelet packets, each scale of the analyzed signal were processed for the RMS value, and they were correlated to their respective cut numbers. The results illustrate that except for extreme cutting conditions, the output model can accurately predict the energy contents of wavelet scales as well as tool wear.

Time series modeling techniques such as autoregressive (AR) and autoregressive moving average (ARMA) models used to be frequently applied for

milling process modeling in early researches. Both AR and ARMA modeling approaches were utilized and compared for tool breakage detection in studies such as [Lombardo et al., 1997].

Pros and cons

This method is mostly used to present a simplified or preliminary model of milling process. It usually requires statistical parameter-determination of unknown variables of the models with predetermined mathematical structure for the milling process or cutter's performance on the workpiece. Although the method is simple and applicable, it is not as accurate as recent AI-based methods according to literature. Therefore some studies combine this method with AI architecture, [Li et al., 2010]. Another weakness of this method is that the fixed outline of the model that limits its accuracy and prevents new modes of the process from being captured and modeled.

Table 1.3: Multiple Regression and Time Series Modeling Approaches to Machining Processes

Papers	Material	Conditions	Cutting conditions	Centre	Signals	Analysis target	Preprocessing technique
[Chakraborty et al., 2008]	AISI 4340	Semi-dry, Dry	Vc Cn Fc Dc*	VMC Model ES-V3016	NA*	Tool wear	NA
[Mike S. Lou, 1998]	6061 Aluminium	NM*	Vc Cn Dc Fc	Fadal CNC	NA	Ra*	NA
[Choi et al., 2004]	AISI 1018 steel	NM	Vc Cn Dc Fc	Fadal CNC	Force	Tool wear	Wavelet transform
[Muammer Nalbant and Toktas, 2007]	AISI 1030 steel	Dry	Dc Vc Fc	CNC lathe	NA	Ra	NA
[Lela et al., 2009]	Steel St 52-3	NM	Dc Vc Fc	Spinner VC 560 NM	NA	Ra	NA
[Palanisamy et al., 2008]	AISI 1020 steel	NM	Dc Vc Fc	DYNA V4.5	NA	Tool wear	Comparison to ANN
[Routara et al., 2009]	Aluminium brass mild steel	Wet	Dc Vc Fc		NA	Ra	NA
[Bhattacharyya et al., 2008]	Mild steel	Dry	Dc Vc Fc	Plano-Miller	Force power	Tool wear	Frequency analysis

* See the Abbreviation Appendix

1.3 Artificial-Intelligence Based Analysis of High Speed Milling Processes

Generally, experiments in the high speed machining field aim to provide an acceptable infrastructure for representing a good descriptive model for a milling process. From section 1.2, it is clear that milling processes have a nonlinear multi-variable nature and the sensor signals are applied to represent roughly the state of the process. Modeling techniques for this type of process should have two general characteristics. Firstly, they must be able to describe the inherent properties of the experiment well. Secondly, they should also be repeatable and applicable to other experiments. The following subsections present the most commonly used AI techniques.

1.3.1 Bayesian Networks

Methodologies and applications

A Bayesian network (BN) is a probabilistic graphical model which represents a set of random variables and their probabilistic dependencies. It is one of the most famous decision making methods based on the statistical behavior of the process [Stephenson, 2000, Dey and Stori, 2005, Dong et al., 2006, Cetin et al., 2007, Kumar et al., 2008]. It applies an extended form of Bayes' rule for the conditional probability of a joint distribution of the variables and their parent variables as represented in Eq.(1.2). It facilitates making a belief network to provide the probability of different phenomena using the

probability of the presence of its parents.

$$P(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^n P(X_i | \text{parents}(X_i)), x_i \in X \quad (1.2)$$

To correlate the input data with the outputs, a classification algorithm is needed to define individual states of the system based on the existing classes. These states are then used in training the model for correlation of the input data with the states of the system. A famous and simple classifier is the *naïve classifier*, which applies Bayes' theorem (assuming independent variables). This assumption makes its training very fast and with comparatively less data. More advanced techniques have been developed in the literature, such as the Tree Augmented Naive Bayes (TAN). Since many process parameters are not really independent, TAN is preferable to the naïve classifier, [Correa et al., 2008, Correa et al., 2009].

A Bayesian network was used in [Correa et al., 2008] to present the surface finishing results of a milling process using 'Elvira' software [CICYT, 2005]. The graphical model represented the correlation between the cutting conditions, e.g., the spindle speed, feed per tooth, depth of cut, and the surface roughness produced. Naive and TAN classifiers were used as the learning paradigm. After validation and comparing the confusion matrices, it was shown that in many cases TAN-trained Bayesian networks were superior to Naive-trained BN [Correa et al., 2008].

Another similar report applies both Naive and TAN and compare their performance with ANN. Since the complex structure of ANN is not opaque comprehensive, in some papers a Bayesian network is preferred over ANN

Figure 1.8: TAN model for surface roughness, [Correa et al., 2009]. Surface roughness (R_a), Feed per tooth (f_z), Depth of cut (D_c) tool diameter (diam), material hardness (HB), geometry radius ($Radius$) and geometry curvature ($geom$), spindle speed (rpm) and the resulting forces applied to all the cutting planes F_{XY} [Correa et al., 2009].

[Correa et al., 2009]. It proposed a model for the surface roughness prediction where the correlations between the variables were clearly visualized, as shown in Fig.1.8.

Combined with support vector regression (SVR), BN was applied in [Dong et al., 2006] to detect tool wear and its performance was compared to another BN-multi layer perceptron (BN-MLP) combination. Force features are used as the inputs to the both networks and they were compared in terms of their prediction accuracy. It is shown that the former model is more accurate [Dong et al., 2006]. Also in [Dey and Stori, 2005], a Bayesian network was used for studying acoustic emission and spindle power metrics. Face milling and drilling processes were investigated and the applicability of BN to the prediction of their surface finishing results was compared. As a result, the root causes of many changes in the signal during the process were correlated to the available cutting conditions.

Pros and cons

Bayesian networks have been extensively used in the best cutting condition determination problem. Using proper signal features as inputs, there are many applications where BN is used for milling process modeling. Compared to other AI techniques, statistical models need more data for training to achieve the same level of accuracy. Since it is graphically representable, using a transition probability matrix makes all the significant and insignificant parameters in the process easily recognizable. Also, it is reported that BN is superior and can replace other existing statistical methods [Giles et al., 1999].

1.3.2 Fuzzy Logic, Neural and Fuzzy-Neural Networks-Based Methods

Methodologies and Applications

Artificial neural-network (ANN), fuzzy-logic, and their combinations such as fuzzy-nets (FN), are widely used in modeling HSM processes. They have also been shown to be capable of modeling not only end-milling but also other kinds of machining processes, providing an accurate approximation of the surface finishing [Ho et al., 2009, Kasirolvalad et al., 2006, Kirby et al., 2006, Hanna et al., 1996, Dweiri et al., 2003, Yang et al., 2006, Aliustaoglu et al., 2009]. In each report ANN is applied with a different algorithm and architecture. There are several options for network architecture for modeling such as cascade-feed forward back propagation (CFFBP), Elman BP (EBP), time-delay BP (TDBP), adaptive resonance theory 2 (ART2)-type

neural networks, perceptron, radial basis functions (RBF), self organizing maps (SOM), and support vector machines (SVM). Feed forward BP algorithm was extensively used in many modeling works. However, choosing the best structure is still an open problem. The details of the structure and connections between inputs/outputs, e.g., the number of hidden layers and their neurons, are also considered as an important issue. However, there is always a trade-off between the accuracy of the model and the computation time. In [Zain et al., 2010, Muammer Nalbant and Toktas, 2007, Zuperl and Cus, 2003] the optimum neural network structure was investigated. Yet, the countless number choices makes it extremely difficult to investigate all possible configurations.

A third issue in neural network is the choice of functions to be used in training, learning, and performance where the best choice is still an open problem. A multilayer feed forward neural network is depicted in Fig.1.9. There are several options to check the performance where MSE is reported to be the most applied [Zain et al., 2010].

Given the fact that the wavelet coefficients of the force signal in a normal and a broken tool are different, an ART2-type self learning neural network was designed to detect the patterns that reflect the failure of the tool from the force signal [Tansel et al., 1995].

In [El-Sonbaty et al., 2008] cutting conditions and vibration signal as well as the fractal geometry and self-similarity properties of the surface were used as a reference building block for all surface patterns and for determining fractal parameters .

Tool wear and surface roughness was correlated with cutting conditions

Figure 1.9: Multilayer feed-forward ANN structure used in [Zuperl and Cus, 2003] to model surface roughness (R_a).

and force features using a BPNN model in [Saglam and Unuvar, 2003]. However, since there were many choices for ANN modeling, choosing the best method and structure remained an open issue. For example, [Briceno et al., 2002, Lu, 2008, Venkatesh et al., 1997] compared the radial basis functions (RBF), back-propagation methods and dynamic models, to find the best structure. Using only one hidden layer and proper DoE, a model was presented that had the ability to capture the characteristics of the force signal given the cutting conditions. Then [Lu, 2008] obtained an approximation to the surface profile while [Briceno et al., 2002] showed that RBF is superior in the sense of a presented cost function, in the prediction of force features.

Some properties of ANN such as its topology, learning rate, and training methods, must be determined before they are applied. Over-fitting and slow-learning are the main challenges. The support vector machine (SVM) method has been developed to overcome these issues by minimizing the generalization error as well as by maximizing the separation margin rather than training

error. Its proposed solution is in the form of constrained linear quadratic programming methods.

To approximate a linear function such as $f(x) = w \cdot x + b$ in the measured feature space, [Vapnik, 2000] proposed an estimation qualification function, an ϵ -insensitive loss function:

$$L_\epsilon(y) = \begin{cases} 0 & \text{for } |f(x) - y| < \epsilon \\ |f(x) - y| - \epsilon & \text{otherwise} \end{cases}. \quad (1.3)$$

The objective function to be minimized by the SVM algorithm was:

$$\min \left[\|w\|^2 + C \sum_{i=1}^m (\zeta_i + \zeta_i^*) \right] \quad (1.4)$$

ζ_i and ζ_i^* are non-negative slack variables as indicators of the training data outside of the ϵ -insensitive neighborhood. There are very few parameters to be set in SVM methods compared to other methods as described in [Vapnik, 2000]. With all its benefits, SVM was used for force, power, and spindle displacement signals, to classify the signs of a broken tool [Hsueh and Yang, 2008, Hsueh and Yang, 2009].

Since there are many parameters to be determined in nonlinear equations for resulting surface roughness, like (1.6), methods like support vector regression (SVR) were also applied to determine the best values for the unknown parameters of these models [Prakasvudhisarn et al., 2009, Lela et al., 2009, Cho et al., 2005]. It has been shown to be superior to traditional regression modeling in [Lela et al., 2009].

Fuzzy Logic (FL) based tool wear monitoring was suggested in [Chen

and Susanto, 2003]. To predict flank wear, the authors utilized the maximum cutting force with other cutting conditions such as the feed-rate and the depth of the cut. Forming its rule base according to experimental and expert knowledge, it was able to estimate the existing flank wear. In [Kim and Jeon, 2011], a fuzzy-logic based controller was applied to a milling process to increase the metal removal rate (and lessen production time) while maintaining a constant cutting force. It applies the spindle and feed current instead of impractical, difficult-to-use dynamometers as sensor signals.

A combination of fuzzy logic and neural networks, the so-called fuzzy neural network (FNN) can be applied to many machining processes as a condition monitoring system, [Elbestawi and Dumitrescu, 2006]. For example, the Hybrid Taguchi Genetic-Learning Algorithm (HTGLA) was used in [Ho et al., 2009] to fit a nonlinear model to the R_a values of a BCCD experiment. The learning data is identical to that used in [Lo, 2003]. The aim was to compare the results of different choices for membership functions which were used in the Adaptive Neuro-Fuzzy Inference System (ANFIS).

As a complete example of a combined monitoring and control system, fuzzy-neuro adaptive surface roughness control (FN-ASRC) was applied in [Yang et al., 2006], where FN-ASRC was divided into two distinct parts. One was the fuzzy-neuro in-process surface roughness recognition (FN-IPSRR), which predicts the surface roughness, and the other sub-system is FN adaptive feed-rate control (FN-ASRC) suggesting appropriate modifications to the cutting conditions in order to achieve a determined surface roughness set point. Fig.1.10 illustrates the framework of the FN-ASRC method introduced in [Yang et al., 2006].

Figure 1.10: Fuzzy-Nets system proposed by [Yang et al., 2006] to adaptively control the surface roughness according the predicted values for surface finishing.

In order to develop the whole monitoring and control system, two distinct five-layer fuzzy-nets were used. The layers are the input, feature-extraction, relations, combination, and defuzzification layers. The fuzzy rules for identification and control were defined and conflicting rules were moved out of the rule base. Three-dimensional force sensors and a proximity sensor were utilized to study the milling process of Aluminium 6061. The process was stopped half-way for the FN-IPSRR system to predict the surface roughness for the rest of the path. Then, in order to achieve better surface roughness results, a feed rate modification was suggested by FN-ASRC based on the predicted results of the FN-IPSRR, [Yang et al., 2006]. Then the remaining half was cut according to the change of the feed rate for verification. The same process could be implemented for other cutting conditions and signals. To avoid severe damage to the tool and surface, fuzzy logic and neural net-

work techniques were applied on-line to control the feed rate in order to maintain a constant force signal [Zuperl et al., 2011, Kim and Jeon, 2011].

Fuzzy-Nets were also applied to model the milling process, [Kirby et al., 2006, Chen and Lou, 2000, Chen and Savage, 2001]. In this method, a number of membership functions were assigned to the input space and were fine-tuned in order to obtain the most accurate input/output model. Then, combinations of these membership functions were considered as possible associative rules in the model's rule base. Besides, conflicting or nonsensical rules were omitted. After all, only the rules with more occurrences and no conflicts remained. For performance verification, several designs were tested and the Fuzzy-Nets method performed acceptably in its surface roughness predictions. The above-discussed papers are summarized in Tables 1.4 and 1.5.

Pros and cons

This section has mentioned methods that have been used due to their ability to model nonlinear processes which are applicable to both experiments to determine the best cutting condition and to destructive tests. Since the literature has been established on these techniques, there have been plenty of different implementations of these methods concerning prediction and control. There are also reports that claim the repeatability of the models. However, none have claimed to be a universal reference model for milling processes and there is no consistency with regard to their requirements for inputs and outputs. Also, many types of models have yet to be developed and tested, such as the combination of neuro-fuzzy algorithms with other AI methods and dynamic fuzzy models [Wu and Er, 2000] for off-line and on-line moni-

Table 1.4: Artificial Neural Networks and Support Vector Machine Modeling Approaches to Machining Processes

Papers	Material	Conditions	Cutting conditions	Centre	Signals	Analysis target	Preprocessing technique
Artificial neural network							
[Zain et al., 2010]	Titanium	NM*	Vc Fc Dc*	NM	NA*	Ra*	NA
[Tang and Guo, 2008]	Stainless steel 304L	Dry	Vc Fc Dc	CNC lathe	NA	Surface profile	NA
[Natarajan et al., 2007]	Steel	NM	Dc Vc Fc	VDF lathe	Flank wear	Tool life	PSO*
[Correa et al., 2009]	Aluminium	NM	Cutting conditions, Geometry	Kondia HS1000	Force	Ra	NA
[Vallejo et al., 2008]	Aluminium	NM	Vc Cn Fc Dc Td	HS-1000 Kondia	AE*, Force, Vibration	Tool life	DFT*
[Zuperl and Cus, 2003]	Steel	NM	Vc Fc Dc Cn Fz Td	CNC lathe,	AE, Force, Vibration	Ra	NA
[Zain et al., 2010]	Titanium	NM	Vc Fc Rake angle	NM	NA	Ra	NA
[Muammer Nalbant and Toktas, 2007]	AISI 1030 steel	Dry	Dc Vc Fc	CNC lathe	NA	Ra	NA
[Tansel et al., 1995]	Kistler 9257A	NM	NA	NM	Force	Tool failure	Wavelet transformation
[El-Sonbaty et al., 2008]	NM	NM	Vc Cn Fc Dc Td Initial tool wear	NM	Vibration	Surface profile	Fractal geometry approach
[Briceno et al., 2002]	NM	NM	Dc Vc Fc	VMC-3016L	Force	Force features	Feature extraction
[Lu, 2008]	Stainless steel 304L	Dry	Dc Vc Fc	NM	AE, Vibration	Surface profile	FFT*
[Venkatesh et al., 1997]	NM	NM	Dc Vc Fc	NM	Force	Flank wear	Normalization
[Patra et al., 2007]	Mild steel	Dry	Dc Vc Fc	Tektronix A621	Power, Force	Flank wear	Wavelet
[Palanisamy et al., 2008]	AISI 1020 steel	NM	Dc Vc Fc	NM	NA	Tool wear	Comparison to MRM*
[Abu-Mahfouz, 2003]	1040 carbon steel	Dry	NA	Bridgeport	Vibration	Tool wear	FFT features
[Zuperl et al., 2011]	16MnCrSi5 XM steel	NM	Geometry Vc Fc Material	Heller	force	force	Actual Value
[Saglam and Unuvar, 2003]	AISI 1040	dry	Vc Fc Dc	Taksan	force	Tool Wear, Ra	ANOVA
Support vector machine							
[Prakasvudhisarn et al., 2009]	6061 Aluminium	NM	Dc Vc Fc	Fadal CNC	NA	Ra	PSO
[Lela et al., 2009]	Steel St 52-3	NM	Dc Vc Fc	Spinner VC 560	NA	Ra	NA
[Hsueh and Yang, 2008]	Cast-Iron Aluminium	NM	Vc Cn Fc Dc Td	Fadal VMC-40	Force	Tool break-age	Low pass filter
[Cho et al., 2005]	Aluminium steel AISI-1018	NM	Dc Vc Fc	Okuma ES-V3016)	Force, Power	Tool break-age	Support vector regression
[Hsueh and Yang, 2009]	7075 Aluminium	NM	NA	NM	Displacement, Power	Tool break-age	NA

* See the Abbreviation Appendix

toring systems. In addition, there are very few on-line prediction and control researches available in this field, which leaves space for more investigations on universal on-line models.

Table 1.5: Fuzzy Neural Network Modeling Approaches to Machining Process

Papers	Material	Conditions	Cutting conditions	Centre	Signals	Analysis target	Preprocessing technique
Fuzzy neural network							
[Ho et al., 2009]	NM*	NM	Vc Dc Fc*	NA*	NA	Ra*	NA
[Kirby et al., 2006]	6061-T6511 Aluminium	Dry	Vc Fc Dc	Storm CNC A50	Force, vibration	Ra	Fuzzification
[Al-Habaibeh and Gindy, 2000]	Aluminium	NM	Dc Fc Vc	Beaver CNC	AE* , Force, spindle acceleration	TCM*	Taguchi's signal/sensor selection
[Li et al., 2000]	Steel #45	Dry	Vc Cn Fc Dc Td	Makino-FNC74-A20	Feed current, Spindle current	TCM	Wavelet analysis
[Ho et al., 2009]	6061 Aluminium	NM	Vc Dc Fc		NA	Ra	Hybrid Taguchi-genetic learning algorithm
[Dweiri et al., 2003]	Alumic-79	NM	Vc Fc Dc flutes Td	NM	NA	Ra	Fuzzification
[Aliustaoglu et al., 2009]	NM	NM	NA	NM	Force, vibration, AE	Tool wear	Sensor fusion
[Lo, 2003]	6061 Aluminium	NM	Vc Dc Fc	NM	NA	Ra	NA
[Chen and Lou, 2000]	6061 Aluminium	NM	Vc Dc Fc	Fadal CNC	Vibration, proximity	Ra	Fuzzification
[Chen and Savage, 2001]	6061 Aluminium	NM	Vc Dc Fc	Fadal VMC-40	Vibration, Proximity	Ra	Fuzzification
[Li and Wu, 1988]	NM	NM	NA	NM	Force, torque	Tool wear	Fuzzification
[Chen and Su-santo, 2003]	Steel AISI-1018	NM	Dc Fc	Fadal VMC	Vibration, proximity	Tool wear	Fuzzification
[Torabi Amin J., 2010a]	Inconel 718	semi-dry	NA	Roder	Vibration, force, acoustic	Tool wear	Wavelet analysis
[Kim and Jeon, 2011]	Aluminium	NM	Geometry Vc Fc material	ACE-V30	Spindle and feed current	Force control	Actual value
[Chen and Su-santo, 2003]	Steel	NM	Dc Fc Material	Fadal	Force	Flank wear (Vb)	Actual value

* See the Abbreviation Appendix

1.3.3 Evolutionary Algorithms, Genetic Algorithms, Genetic Programming and Particle Swarm Optimization

Methodologies and Applications

The method of genetic algorithms (GA) is an optimization method based on evolutionary searching of the solution space. The idea was based on the works introduced in [Holland, 1992]. GA is used in the machining technology field for modeling issues wherever optimization is concerned. In [Shunmugam et al., 2000] and [Liu and Wang, 1999] for example, it was used for best cut-

ting condition determination (BCCD). With proper economic justifications for the cost of the process and modified limitations on the variables, many cost functions were defined for the process with some cutting conditions as optimization variables [Shunmugam et al., 2000, Liu and Wang, 1999]. The same problem was solved with a combination of simulated annealing (SA) and GA in [Wang et al., 2005]. However, since GA training requires random measurements on surface roughness and tool wear, it was not easy to be generalized in the available form or to be used for on-line analysis and prediction.

Figure 1.11: Proposed GP-based structure to represent a nonlinear function.

The genetic programming (GP) method was first introduced in the early 90s by Koza [J.R.Koza, 1992]. Basically, it is an evolutionary algorithm that makes the program perform better at a predefined way of modeling, Fig.1.11. The program is made to evolve and produce an optimal model that

matched the data. Theoretically, they are represented in the form of recursively evaluated and evolved tree structures. Every tree node has an operator function and every terminal node has an operand, making mathematical expressions easy to evolve. There are several implementations of this method for milling process modeling [Zain et al., 2008, Brezocnik et al., 2004, Achiche et al., 2002]. A general review of these methods was found in [Zain et al., 2008]. The method was used in [Brezocnik et al., 2004] and [Çolak et al., 2007] to represent the surface roughness in its dependence on the cutting conditions and the vibration signal. According to this method, an evolutionary algorithm investigated the best match for the experimental data by evolving the tree of operators and operands as modeling functions for the milling process using simple function genes and terminal genes. In [Razfar et al., 2010], a GA-optimized neural network (GONN) was applied to tool condition monitoring where GA was applied to fine-tune the neural network parameters. The performance of this model was also compared with that of particle swarm optimization (PSO)-based neural network (PSONN). In both the GA and PSO based approaches, these optimization methods were applied for determining the neural network parameters. Yet, there is no application of dynamic evolutionary algorithm to the field of milling process monitoring, diagnosis and prognosis.

The particle swarm optimization (PSO) method is a famous optimization procedure based on a direct search method which imitates social behavior in the presence of objectives. It was firstly introduced by Kennedy and Eberhart [Kennedy et al., 2001] and was used in several applications. It uses

Figure 1.12: Proposed GP-based structure to represent a nonlinear function.

an iterative formula for the swarms to approach global maxima,

$$\begin{aligned}
 v_{i,j} &= c_0 v_{i,j} + c_1 r_1 (\text{globalbest}_j - \dots \\
 x_{i,j}) &+ c_2 r_2 (\text{localbest}_{i,j} - x_{i,j}) + \dots \\
 c_3 r_3 (\text{neighbourhoodbest}_j - x_{i,j}) \\
 x_{i,j} &= x_{i,j} + v_{i,j}
 \end{aligned} \tag{1.5}$$

Due to its ability to search for the global optimum, it has been mostly applied in milling processes to determine the best cutting conditions. PSO was used for in the machining literature to find the best matching parameters for a proposed surface roughness model [El-Mounayri et al., 2003, Sagherian and Elbestawi, 1990].

$$R_a = \frac{10aR^b F^c}{S^d} \tag{1.6}$$

where R_a is the surface roughness, R is the radial depth of cut, F represents the feed factor, S is the spindle speed, and a, b, c, d are unknown parameters.

In [Prakasvudhisarn et al., 2009, El-Mounayri et al., 2003], particle swarm optimization was applied to the results of a support vector machine (SVM). The SVM determined the unknown parameters of the model in (1.6). Then, PSO was used to find the optimal cutting conditions [Prakasvudhisarn et al., 2009].

Because of its ability to find the optimal solution for most nonlinear objective functions, there is no specific limitation on using any predefined model for the process. For example, [Cus et al., 2007] used an artificial neural network (ANN) model for the force versus surface roughness and a PSO algorithm was applied to find the optimum cutting conditions. In [Natarajan et al., 2007], an ANN was applied to model the tool life dependent on the cutting conditions and flank wear. PSO is also utilized to optimize the ANN parameters.

Pros and cons

Table 1.6 presents the papers on these methods. Since genetic algorithms were not developed for dynamic training until recently, they were just used for off-line best cutting condition determination. Since the basic idea is to reach an optimum point for an objective function, it can be properly used for building a best fitting model on off-line raw data. However, there is no report that this method is capable of on-line adaptation. There are some studies that suggest merging dynamic learning with this method, [Milani, 2004], and so it might be applied in the future studies of milling process

Table 1.6: Genetic Algorithms, Fuzzy Petri Net, Particle Swarm Optimization Modeling Approaches to Machining Processes

Papers	Material	Conditions	Cutting conditions	Centre	Signals	Analysis target	Preprocessing technique
Genetic algorithm							
[Brezocnik et al., 2004]	6061 Aluminium	NM*	Vc Cn Dc Fc*	NM	Vibration	Ra*	Genetic programming
[Achiche et al., 2002]	AISI 1018 steel	NM	Fc	NM	Cutting force, Feed force	Tool wear	Fuzzification
[Çolak et al., 2007]	Aluminium 6061-T8	NM	Dc Vc Fc	VMC550	NA*	Ra	NA
[Wang et al., 2005]	NM	NM	Fc Vc Power	NM	NA	Fc Vc Power	Simulated annealing
[Shunmugam et al., 2000]	NM	NM	Dc Vc Fc	NA	NA	Dc Vc Fc	NA
[Liu and Wang, 1999]	T20 Grey cast	NM	Dc Vc	NM	NA	Fc	NA
[Razfar et al., 2010]	X20Cr13 Steel	NM	Dc Vc Fc Cutter engagement	Fil Fresatrici	Flank wear	Ra	ANN*, Compared with PSO*
Fuzzy Petri net and DEVS							
[Kasirolvalad et al., 2006]	NM	NM	Vc Fc T	NM	Force, Vibration, AE	Ra	Fuzzification
[Hanna et al., 1996]	Brass	Wet	Vc Fc Dc Td	MHPB500	Spindle speed, Feed rate	Ra	Fuzzification
Particle swarm optimization							
[El-Mounayri et al., 2003]	Aluminium	Dry	Dc Vc Fc	Fadal VMC	NA	Ra	NA
[Prakasvudhisarn et al., 2009]	6061 Aluminium	NM	Dc Vc Fc	Fadal CNC	NA	Ra	SVM*
[Natarajan et al., 2007]	Steel	NM	Dc Vc Fc	VDF lathe	Flank wear	Tool life	ANN
[Razfar et al., 2010]	X20Cr13 Steel	NM	Dc Vc Fc Cutter engagement	Fil Fresatrici	Flank wear	Ra	ANN, Compared with GA*

* See the Abbreviation Appendix

modeling. Another issue that exists with GP is the complex formulations and functions in the output. To make the modeling more meaningful, the output model have to reflect the mechanical nature of the process. This makes the model process more computationally intensive.

As an optimization method similar to genetic algorithms, PSO is also used to facilitate nonlinear model identification and parameter determination. Also, it can be used as a training method for other AI techniques to find the best fitting model for the milling process. Since it requires an existing nonlinear function, it might not be suitable for on-line data analysis and prediction. Perhaps with some modifications in the variable definitions, it might be able to work in real time as well as GA.

1.3.4 Hidden Markov Models

Methodologies and Applications

Hidden Markov models (HMMs) were first introduced in [Baum and Petrie, 1966] as ‘probabilistic functions of Markov chains’. Afterwards, several methods were introduced for modeling and their application was summarized in [Ephraim and Merhav, 2002]. To formulate an HMM model $\lambda = (\pi, A, B)$, usually N distinct (hidden) states q_i for the system are considered. The Markov chain is defined by the connecting transitions between q_i states. These connections are completely defined by the state transition matrix $A = [a_{ij}]$ where each element a_{ij} represents the probability of the corresponding transition.

$$a_{ij} = P(q_t = j | q_{t-1} = i), 1 \leq i, j \leq N \quad (1.7)$$

Since the a_{ij} s are probability values, the following axiomatic constraints are applied [Juang and Rabiner, 1991]:

$$a_{ij} \geq 0, \sum_{j=1}^N a_{ij} = 1, \forall i. \quad (1.8)$$

We may assume without loss of generality that the start time of the model is 0, at which the model has an initial condition. It is represented by the probability of each individual state at the initial time or the initial condition probability distribution, $\pi_i = P(q_0 = i)$ which is the i^{th} element of π . Then,

the probability of any chain of states will be

$$P(q|A, \pi) = \pi_{q_0} a_{q_0 q_1} \dots a_{q_{T-1} q_T}. \quad (1.9)$$

Since the states of the system are not always observable, the only thing that is available about the system is the observations O_t which is according to the changes in the system states. The relation between these observations and the states is declared by another probability matrix which is called the emission matrix,

$$B = \{b_i(O_t)\}_{i=1}^N, b_i(O_t) = P(O_t|q_t = i). \quad (1.10)$$

If all the state sequences up to the current state of the system are known, then the probability of the observed output of the system will be calculated by:

$$P(O|q, B) = b_{q_1}(O_1)b_{q_2}(O_2)\dots b_{q_t}(O_t) \quad (1.11)$$

Then, it is straight forward to have:

$$P(O, q|\pi, A, B) = \pi_{q_0} \prod_{t=1}^T a_{q_{t-1} q_t} b_{q_t}(O_t) \quad (1.12)$$

Then to find the probability of any observation it is easy to verify [[Juang and Rabiner, 1991](#)]:

$$\begin{aligned} P(O|\pi, A, B) &= \sum_q P(O, q|\pi, A, B) = \dots \\ &= \sum_q \pi_{q_0} \prod_{t=1}^T a_{q_{t-1} q_t} b_{q_t}(O_t), \lambda = (\pi, A, B) \end{aligned} \quad (1.13)$$

There are three major issues to be faced for developing an HMM structure to model a system. The first one is to compute the probability of an output event's happening in the available model λ (*Evaluation*). The second issue is to find the unknown parameters of the HMM model which best matches the observations O (*Estimation*). The last problem is to find out the most probable sequence of states q , regarding the observations O (*Decoding*). Further details of the available solutions to these three problems and many other applications of HMM were discussed in [Ephraim and Merhav, 2002, Bilmes, 2006, Fish et al., 2003]. HMM may also be considered for higher order processes. It can be shown that this high order HMM is easy to be converted to an equivalent first order HMM [Bilmes, 2006].

$$Q'_t \triangleq \{Q'_t, Q'_{t-1}, \dots, Q'_{t-n}\} \quad (1.14)$$

while Q_t is an n^{th} -order Markov chain.

$$\begin{aligned} &P(Q'_t = q'_t | Q'_{t-1} = q'_{t-1}, Q'_{t-2} = q'_{t-2}, \dots, Q'_1 = q'_1) \\ &= P(Q_{t-n:t} = q_{t-n:t} | Q_{1:t} = q_{1:t}) \\ &= P(Q_{t-n:t} = q_{t-n:t} | Q_{t-n-1:t} = q_{t-n-1:t}) \\ &= P(Q'_t = q'_t | Q'_{t-1} = q'_{t-1}) \end{aligned} \quad (1.15)$$

So, Q'_t will be a first order Markov chain. This will result that if there is enough training data, HMM can capture any high order relations [Bilmes, 2006, Fish et al., 2003]. In this way, HMM can be viewed as a generator and acceptor in the Language theory [Bilmes, 2006, Doostfateme and Kremer,

2005].

HMM has rarely been used in the literature to present a dynamical model of a process. Each paper has its own way to provide sequential data to HMM training algorithms such as Baum–Welch, known also as EM (the expectation maximization) method. It is a maximum-likelihood based method that finds the parameters of the state transition matrix and the output emission matrix from the internal states [Juang and Rabiner, 1991, Vallejo et al., 2008]. Originally, it seemed that HMM is not as accurate as other models for a nonlinear system. However, its aggregation with classification and nonlinear methods can lead to better results [Owsley et al., 1997b].

To provide data for training an HMM model for a milling process, [Wang et al., 2002] used the vector quantization (VQ) method based on the discrete wavelet decomposition of sensor signals which is briefly described in Section 1.1.2. Applying the codebook of the worn or sharp tool, its status was predicted by applying the current state and emission matrix. There are classification methods other than VQ that have been used to generate the input/output sequence for an HMM to model. For example, [Fish et al., 2003] suggested a modified classifier for HMM training the status of the tool in probabilistic terms rather than in binary output, for example Worn/Sharp states.

HMM was also used to correlate the observable changes in the energy content of different DWT scales of vibration signals with tool wear. Vibration signals were analyzed for some of their details and for each detail the changes in the energy are observed for a certain period of time [Wang et al., 1998]. After proper training, two distinct codebooks for sharp/worn tool were de-

veloped. Simulations showed that HMM models could successfully monitor and detect the internal status of a milling tool. Wavelet modulus maxima information was used in [Miao and Makis, 2007] to build a combined HMM model. It was shown that this feature had meaningful changes according to tool wear progress and it was applied to provide an accurate representation of machining condition. Therefore, three models for different states of the tool, i.e., normal, warning, and failure condition, were presented. The probability of each sequence of the data was estimated according to these models and the sensor signals and finally the highest probability is chosen as the real state of the system.

Table 1.7: Bayesian Networks and Hidden Markov Modeling Approaches to Milling Processes

Papers	Material	Condition	Cutting Conditions	Centre	Signals	Analysis Target	Preprocessing Technique
Bayesian networks							
[Dey and Stori, 2005]	AISI 4140	NM*	NM	NM	AE*, Spindle power	Tool wear, WH*	Feature extraction
[Correa et al., 2008]	F114 steel	NM	Vc Fc Dc Cn Fz Flutes Td*	Kondia HS1000	NA*	Ra*	K-means discretizer
[Correa et al., 2009]	Aluminium		Cutting conditions, Geometry	Kondia HS1000	Force	Ra	TAN* algorithm
[Lela et al., 2009]	Steel St 52-3	NM	Dc Vc Fc	Spinner VC 560	NA	Ra	NA
[Dong et al., 2006]	ASSAB718H	NM	NA	Makino CNC	Force	Tool wear	Support vector machine
hidden Markov models							
[Wang et al., 1998]	AISI 8620 steel	Dry	Fc Dc Vc	30hp lathe CNC	Vibration	Tool wear	Discrete wavelet transform
[Wang et al., 2002]	AISI 8620 steel	Dry	Fc Dc Vc	30hp lathe	Vector quantized vibration	Tool wear	Discrete wavelet transform
[Miao and Makis, 2007]	NM	NM	NM	NM	Vibration	Failure detection	Modulus maxima wavelet
[Vallejo et al., 2008]	Aluminium	NM	Vc Cn Fc Dc Td	HS-1000 Kondia	AE, Force, Vibration	Tool life	DFT*
[Owsley et al., 1997b]	NM	NM	NA	NM	Vibration	Tool monitoring	Spectral feature extraction
[Ertunc and Oysu, 2004]	Steel	Dry	Fc Dc Vc	MAHO 700S	Force, Torque, AE, Current	Tool wear	
[Fish et al., 2003]	AISI 4340 steel	Water-soluble	Dc Vc Fc	MAZAK H800	Vibration, AE	Tool wear	Classification

* See the Abbreviation Appendix

Pros and cons

HMM has only been applied in a few studies in the literature. Compared to BN, it has the benefit of being able to reflect the behavior of milling processes in the form of dynamic models rather than static models. It facilitates an estimation of the internal states of the system, needing only system outputs. As an essential issue, finding the probability distribution structure that fully describes the sequence of the signal features has been investigated in many researches. However, since it requires a large amount of data for training it seems less appropriate for modeling BCCD experiments. For destructive tests however, it might be used the same way that it is usually used in speech processing and recognition [Juang and Rabiner, 1991, Varga and Moore, 1990, Lee, 1990], because of the availability of AE sensors [Wang et al., 2002, Owsley et al., 1997b]. However, among the reports on the performance of HMM in milling processes, there are very few predictive accuracy comparisons with other AI techniques in the field. Hidden Markov models has been extensively used in speech processing. In order to apply it for resulting milling signals, we have to investigate the properties of signals that makes HMM suitable for their analysis. Speech is a non-stationary signal and the result of modulation on air pressure by articulatory apparatus [Juang and Rabiner, 1991]. Speech signal has frequency content up to several thousand hertz. However, the human's articulatory system does not perform more vibrations than 10 hertz per second. Speech processing and modeling has two aspects to be considered: (1) Short-time spectral analysis of the speech signal. (2) Long-term analysis of the changes in sound sequences due to changes

in articulatory system. HMMs are also applied in biological sciences where the application of hidden Markov models is the detection of distinct genes in DNA protein sequence. These genes are composed of some definite proteins in the form of a sequence [Regad et al., 2008, Thomas et al., 2007].

In speech signal processing, time-frequency properties of windowed discrete-sampled sound signal is analyzed. These windows are well defined in the literature to avoid the effects of discontinuity on the resulting signal spectrum. As a known solution to speech processing analysis, different output symbols are assigned to different spectral vectors of the speech signal which is called spectral labeling. This spectral vectors are the set of all the possible frequency distributions for speech. So, the entire speech can be supposed as a linear combination of the sequences of these spectral prototypes. In this sense, all the selected spectral vectors can be viewed as the different states of the articulatory system. In a complete speech, all the states interchange several times. Therefore, it makes HMM-based model a suitable choice to describe a probabilistic model for the speech. HMM analysis of these sequences can be done statistically based on the interchanges of the labels given to the input signal in long-term. Hypothetically, sequence of spectral observations in each time instance, O_t , starts from $t = 1$ to $t = T$. As the result, the spectral labeled sequence is $\{O_t\}_{t=1}^T = (O_1, O_2, \dots, O_T)$. Since, the composition of these observations makes the whole speech and each observations is repeated several times, it can easily be modeled by statistical models. To model these kinds of sequences, they can be described in the form of Markov chains. It statistically models the way that spectral prototypes interchange.

- Modeling Problems and Cutting Process Application Signal Issues

In the development of HMM structure to model a system, there are three major problems to be solved. First problem is to compute the probability of happening an output event regarding to available model, λ , (*Evaluation*). The second query is to find the unknown parameters of the HMM model, λ , which fits well to the observations, O_t , (*Estimation*). The last main problem is to find out the most probable sequence of states, q_i , regarding to the observations (*Decoding*). Further details of the available solutions to these problems in addition to many other applications on hidden Markov processes were discussed in [Ephraim and Merhav, 2002].

Besides the above-mentioned standard issues, there are some issues regarding HMM when it is to be applied for modeling cutting process signals. In others areas like speech processing where HMM is very well-performing, most of these issues are already solved due to the nature of speech signals while they are already open questions and issues for the milling signals. These issues are covered in the following subsections.

- Training Problem

As mentioned before, there are usually three major problems to be solved for a HMM to be defined and used. The first one is training or evaluation. As specified in many reports, this part of modeling process needs many amount of data [Regad et al., 2008, Juang and Rabiner, 1991, Ertunc and Oysu, 2004, Wang et al., 2002]. It might not be a significant issue in some of the fields like biology or text processing

or speech recognition which normally deal with the sequence of outputs, like proteins of DNA, sentences made up of countable alphabets and sequences of spectral patterns [Regad et al., 2008, Thomas et al., 2007, Juang and Rabiner, 1991]. But, even in speech processing it is not easily feasible [Juang and Rabiner, 1991]. This is an existing issue for milling processes as well and in some cases it is even worse than other processes. In best-cutting-condition-determination experiments, HMM requires many distinct experiments with different cutting conditions. It is not feasible for many industrial applications and academical researches due to its cost and time consumption. It is also reported that the reliability of the model decreases when there is not enough data for training. Therefore, some techniques for accuracy improvement have to be utilized [Juang and Rabiner, 1991]. For destructive tests, the situation is different. In one single destructive test, there are lots of data to be analyzed. However, the resulting model cannot be easily generalized to be used for other cutters due to the differences between cutters and cutting conditions. Besides, there are other issues of non-repeatability of patterns in destructive-test resulting signal which will be clarified in the following.

- Definition of States and Events of Hidden Markov Model

Second issue is the proper definition of the hidden states and events of the model which originates from a basic issue of finding a period of the signal for analysis. According to automata theory, each of the available spectral vectors of the resulting signal can be considered as alphabets

of the observations language² [Linz, 2006]. The time period of any of these samples should be properly defined to facilitate the spectral analysis and to avoid discontinuity problem which leads to many new and unnecessary patterns (spectral vectors). Concerning the milling process, these composing elements can be flute-cut signals or revolution signals that form the entire cut-signals. Analyzing these signals facilitates one to one correspondency of the signals and specific physical events in the cutting process, like shearing, plunging, scratching. Using time-frequency analysis, the original cutting signals are decomposed into many flute and revolution signals in order to form a chain of successive events and input pattern.

- Repeated Output Symbols vs. Unique and Unrepeated Symbols

Another issue for HMM modeling of the the cutting process is the continuous and non-repetitive change in the raw cutting data. After the proper separation of the cutting signal, it is clear that the flute and revolution signals in the beginning moments of the process does not have any similarity to the flute and revolution signals at the end of the cutting process. Fig.1.13 depicts the obvious differences between the signals at the beginning and at the end of destructive tests on two cutters. This is regardless of unrepeated changes of these signals in the amplitude and period. The amplitude of these signals are due to change for several reasons like increasing temperature and hence friction and/or the distance between the sensors and the cutting point.

²The terms like 'language' and 'alphabet' are commonly used technical terms in this theory.

Figure 1.13: The obvious change in the resulting signals from the beginning of the process to the end: I- Cutter No.1: (a) The 40th cut of the process with cutter No.2, (b) The 600th cut of the process with cutter No.2. II- Cutter No.2: (c) The 40th cut of the process with cutter No.1, (d) The 600th cut of the process with cutter No.1

Milling process in this last issue can be compared to its counterparts in other famous applications of HMM, like speech processing [Juang and Rabiner, 1991] and biology [Regad et al., 2008, Camproux et al., 2004, Bicego et al., 2009]. The elements of these sequences are repeated many times regardless of the time stamp of the signal. In contrast, the elements of the cutting signals are continuously changing and they do not have similarity to the past signals. Consequently, the resulting transition matrix of these HMM models for milling signals are expected to have semi-diagonal form. The reason is that there is no alteration between signal patterns of the worn cutter and the patterns of the flute signals in the earlier stages of destructive test.

- Model Accuracy

Another issue is the accuracy of the hidden Markov models. The root-causes of this issue for speech processing is discussed in [Juang and Rabiner, 1991]. For example, it is not necessary for speech signal to be non-stationary. The problem does exist in the case of cutting signal as

well.

Next issue is that in the training of HMM process, the input data must be classified to be labeled as an specific pattern. These labels are analyzed for probability distributions for the transitions between states. The same is for output of the model. The outputs have to be made discretized and cannot be modeled in its continuous form any more. Therefore, the accuracy of the output of predictions and detections, [Ertunc and Oysu, 2004, Wang et al., 2002, Vallejo et al., 2008, Atlas et al., 2000, Ertunc et al., 2001, Heck and McClellan, 1991, Owsley et al., 1997a, Jr. et al., 2005] might be unnecessarily affected.

- Repeatability

The repeatability of the HMM for machining process status detection has not been discussed vastly in the literature. Therefore, the final accuracy of the model may not increase this unless the number of hidden states and the input and output symbols are increased which results in another problem of computational complexity. Also, increasing the number of output states to compensate the discretization gap and error will increase the tendency of non-repeatability of the model.

Referring to literature, there are several challenges to modify the accuracy of the HMM model to be comparable to other AI techniques, some of which were addressed in this paper. It is also clear that the output of the milling process has not any similarity to other processes like speech, text and biological sequences from the self-repeatability point of view. Therefore, it is worth considering other AI techniques

instead of using HMM for milling process or addressing the challenging issue of HMM before its application.

Therefore, the issue of non-repeating patterns and discretization accuracy are also bottlenecks of the modeling problem. All these above mentioned issues makes the application of hidden Markov models a challenging idea in the field of milling process modeling.

1.3.5 Discrete-Event Based Intelligent Methods

Methodology and Applications

Although the cutting conditions and signals that are collected in high speed machining process are naturally continuous, the changes in the system can be interpreted as the results of successive incoming events of a discrete event system (DEVS).

Among the studies of DEVS modeling for milling processes, [Hanna et al., 1996] and [Kasirolvalad et al., 2006] presented an adaptive fuzzy petri net (AFPN) model for the high speed machining process. Petri net (PN) represents the states of the system and the input events in its graphical model. Like other discrete-event modeling systems, PN is generally an n -tuple DEVS machine represented with arrows as the incoming events or internal relationship and the states of the machine and circles or nodes representing the states, respectively [David and Alla, 1994].

Like its predecessor, fuzzy automata [Doostfatemeh and Kremer, 2005], fuzzy petri net (FPN) has membership values for states and transitions [Cardoso, 1999]. For example, proper membership functions and values are de-

Figure 1.14: A technique based on AFPNs for modeling and control of product and machining process quality [Kasirolvalad et al., 2006]. V: Vibration, F: Cutting Force, S: Spindle Speed, f: feed rate, t: MachiniR: Surface Roughness, Q: Machining Quality

defined for all the cutting conditions, collected signals and resulted surface roughness in [Kasirolvalad et al., 2006]. Using some fuzzy rules and the relations between the cutting conditions, [Kasirolvalad et al., 2006] suggested a PN and weighted transitions so that the dynamic behavior of the process in conjunction with its surface roughness could be represented. By defining two separate AFPNs for prediction and alarming algorithm, it tried to avoid unsatisfactory surface finishing by alerting an error in cutting parameters or tool's status, Fig.1.14.

Pros and cons

Very few DEVS models have been introduced for milling processes. As a positive point for this method, its visual representability by using meaningful states for the system can be mentioned. Representation of the events in the process as continuous or discrete can open the doors for different views

of the signals, their changes, and models. Besides, since there are not many researches in this field, the suitability of the DEVS models for milling processes is yet to be investigated. The appropriate input parameters also have not yet been widely investigated in comparison to neuro-fuzzy methods. An issue with this method is the way that the events and the states are defined and the transitions are considered in model structure.

1.3.6 Clustering and Classification Methods

Methodology and Applications

Clustering methods are meant to keep similar data together in clusters to facilitate a proper overview of the domain. There are two types of clustering which are commonly applied in milling process: hard/crisp clustering and fuzzy clustering. In the former, a datum can only belong to one cluster, but in fuzzy clustering methods, a datum can be a member of several clusters with a certain membership value. The process of assigning a datum to a cluster or some clusters depends on its distance, similarity, or connectivity to other data in that specific cluster [Farran et al., 2009, Berkhin, 2006].

Fuzzy C-means clustering is a famous clustering techniques [Yang, 1993]. It classifies the finite information into several classes based on some criteria. Given a finite set of data, the algorithm returns a list of cluster centers and a partition matrix. Each of its elements is a membership value of a datum that belongs to a specific cluster [Yang, 1993].

On the other hand, each datum is assigned to only one cluster in hard/crisp clustering, as in the k -means algorithm, where a datum is attributed to the

cluster with the nearest center. The center of each cluster is the arithmetic mean of all its members. Crisp clustering, such as k -means and k -medoids, are applied in [Torabi Amin J., 2011] to illustrate the applicability of such methods to modeling approaches.

The fuzzy C-means clustering method was used in [Xiaoli and Zhejun, 1998] on wavelet packet features of AE sensor signals and in [Torabi Amin J., 2010b, Torabi Amin J., 2011] on the energy contents of different scales of CWT of the force and vibration signals. Power consumption and vertical force are also clustered in [Wang et al., 1996]. Since the RMS value of each frequency band in an AE signal changes with different tool conditions, this signal feature is indicative of tool wear and surface roughness [Li et al., 2010]. As reported in [Xiaoli and Zhejun, 1998], four states for the tool wear with seven features each, compose the codebook of the clustering method. Fuzzy clustering on continuous and discrete wavelet analysis of AC servo-motor current signals of the spindle and feeder was used in [Li et al., 2000] for tool breakage detection and tool wear monitoring.

Classification methods have also been applied for milling condition detection purposes. From the experimental knowledge [Elbestawi et al., 1989] assumed five different classes for feature patterns of sensor signals, applying this knowledge with linear discrimination classification techniques.

Self organizing maps (SOM) can be considered as another clustering technique to reduce the dimensionality of the data. The new dimension depends on how the new sets of vectors are ordered. For example, for 2D-SOM, the code vectors are ordered in 2D and referred to by a code vector index such as $\bar{n}(i, j)$. To train the SOM, each training sample of the high dimensional

space, $\bar{t}_k$, is mapped to its nearest code vector member, and hence belongs to the corresponding class.

$$(\hat{i}, \hat{j}) = \arg \min_{i,j} \|\bar{n}(i, j) - \bar{t}_k\| \quad (1.16)$$

Then the code vector is updated by moving toward the training vector.

$$\bar{n}_{update}(i, j) \leftarrow \bar{n}(i, j) + f \left[r, \|(i, j) - (\hat{i}, \hat{j})\| \right] \cdot \frac{\bar{t}_k - \bar{n}(i, j)}{\|\bar{t}_k - \bar{n}(i, j)\|} \quad (1.17)$$

$f(\cdot)$ is a non-increasing function of the learning iteration number and the distance between the code vectors. So, in the learning procedure, all code vectors move towards the training vector depending on the iteration number and distance from the vector under which the last training vector was classified. SOM was used in [Owsley et al., 1997b] to reduce the dimension of the feature space of the time frequency blueprint of time windowed signals. It was also applied as a part of the rule generation procedure in [Li et al., 2010] in combination with a dynamic fuzzy regression modeling system.

Table 1.8: Clustering Modeling Approaches to Machining Processes

Papers	Material	Conditions	Cutting conditions	Centre	Signals	Analysis target	Preprocessing technique
Clustering and classification methods							
[Xiaoli and Zhejun, 1998]	Steel	NM*	Cutting conditions, geometry Fc Dc Vc*	Makino FNC 74-A20	AE*	Tool wear	Wavelet packet transform
[Wang et al., 1996]	EN1A	Dry		Cincinnati Sabre 500	Power, force	Flank wear	Signal feature extraction
[Fish et al., 2003]	AISI 4340 steel	Water-soluble	Dc Vc Fc	MAZAK H800	Vibration, AE	Tool wear	HMM* modeling
[Torabi Amin J., 2010b]	Inconel 718	semi-dry	NA*	Roder	Vibration, AE, force	Tool wear	Wavelet analysis
Self organizing maps							
[Owsley et al., 1997b]	NM	NM	NA	NM	Vibration	Tool monitoring	Spectral feature extraction
[Li et al., 2010]	Inconel 718	HRC32	dry	Roder	Vibration, force, AE	Tool wear	Fuzzy regression model

* See the Abbreviation Appendix

Pros and cons

The clustering of the available data of the process will lead to the generalization of the model as the clusters are easier to associate with tool status than were the pure signals. This issue mostly appears when there are different cutters involved. In addition, there are many uninvestigated and unclassified features to be studied, which leaves space for more research. Among them, time–frequency analysis features can be mentioned. These features can be applied in a more methodical way when clustering methods are involved. Besides, there are quite number of classification methods that have not been applied to the field of intelligent machining. Also, the combination of clustering methods with AI techniques remain to be investigated more extensively in the field so that the contribution of clustering methods be clarified. Clustering methods can also be used to investigate the similarities between different signal features. Finding these similarities, other AI techniques can be applied to map different classes to the different respective conditions of the tool and milling process. However, the number of classes and the structure of the classification method may determine its accuracy and they are open issues for further investigations.

1.4 Conclusions

This chapter investigated several commonly used methods for surface-finishing quality modeling of high speed milling processes. It covered many artificial intelligence methods as well as classic ones. The simpler methods are typically used for simple presentation of the behaviour of the process while AI-

Table 1.9: Advantages and Disadvantages of the AI Methods

Method	Advantages	Disadvantages	Target variable	Preprocessing methods	Signals
Mathematical modeling and numerical difference equation solution methods	A less costly method, detailed analysis, many simulation runs before actual running of the system	Can't deal with surface and tool imperfections, Incapable of real-time simulation and analysis	Ra*, tool breakage, force, chip formation, temperature, strain, crack formation, surface profile	Low pass filter, Hilbert transform, finite element techniques, grey relational analysis	Feed-motor current, force, vibration
Statistical and experimental evaluation	Outline of similarities and differences, easy to visualize	No mathematical model, never can be used on-line	Flank wear, tool wear, tool breakage, Ra, surface profile, force	FFT*, time frequency analysis, power spectrum, NM*, Grey relational analysis, Taguchi design procedure, wavelet packets, discrete wavelet transform, force dynamical model	AE*, force, temperature, tool wear, power, vibration
Multiple regression	Predetermined structure, not accurate enough	Simple to be used, simple and few computations, trained model can be used on-line	Tool wear, Ra	Comparison to ANN*, wavelet transform	Force, cutting conditions
Genetic algorithm and genetic programming	Optimization of the model, Suggesting new models	Time consuming, is not yet used for on-line machining, needs a pre-determined model to optimize	Tool wear, Ra	Genetic programming, fuzzification	Vibration, cutting force, feed force
Particle swarm optimization (PSO)	Parameter optimization of the model, can be combined with other methods, simple to implement	Needs a predetermined model, slow, cannot be used on-line	Tool life, Ra	Combined with SVM*, ANN	Flank wear
Bayesian networks and HMMs*	Presents hidden relations, visually representable, BNN* is superior to other statistical methods, good for generalization	Needs a lot of data, slow training, HMM is not reported accurate	Failure detection, workpiece hardness, Ra, tool wear	Feature extraction, <i>k</i> -means discretizer, TAN* algorithm, support vector machine, discrete wavelet transform, modulus maxima wavelet, DFT*, spectral feature extraction, classification	Spindle power, vector quantized vibration, AE, force, vibration, torque, current
Fuzzy-logic and neural network based methods	Good for non-linear models, well-established theory, high accuracy, fast in evaluation, easy to generalize	Overtraining problems, slow in training, ANN is not easily generalized to other cutters, fixed structure of model	Tool wear, flank wear, surface profile, force features, tool failure, Ra, tool condition monitoring	Fuzzification, Taguchi's Signal/sensor selection, wavelet analysis, sensor fusion, hybrid Taguchi-genetic learning algorithm, PSO, fractal geometry approach, FFT, normalization, compared to MRM*, FFT* features, DFT	Flank wear, AE, force, vibration, power, spindle acceleration, feed current, spindle current, vibration, proximity
Discrete-event based intelligent methods	Can be combined with other AI*, visually representable, new in machining field	Needs clear justification for event-based model, needs reasons for events, vague dealing with fixed parameters, definition of states of DEVS*	Ra	Fuzzification	Force, vibration, AE, spindle speed, feed rate
Clustering methods	Easy to generalize, lowers the data dimensions, can be combined with AI techniques, new in the machining field, easy to correlate the distinct classes with cutting phenomena	Low accuracy, some of the techniques have fixed and huge structure, slow training	Tool wear, Flank wear, Ra, Tool breakage	HMM modeling, Signal feature extraction, Wavelet packet transform, Support vector regression, PSO, Low pass filter	Force, power, displacement power, AE, vibration

* See the Abbreviation Appendix

based methods are applied for modeling, on-line monitoring, and predictive control. Based on these two categories, we investigated the state-of-the-art methods which are commonly used for both modeling and control. Since the nature of the process is multi-variable and nonlinear, most of these modeling approaches are found to be able to model such systems. Bayesian networks, fuzzy Petri-nets, hidden Markov models, and dynamic fuzzy-neural networks have proved to be the most suitable modeling techniques. On the other hand, there are many research gaps that need to be addressed in this field. Besides, there are very few research reports on the tool-production methods; tool attributes; and their effects and correlations with the sensor signals, surface-roughness and tool-degradation. Also, there is an obvious research gap as to presenting a single general model for milling processes where the available models are not expandable to other cutters (even those with similar attributes). Before obtaining a general descriptive model for milling processes, the field keeps being updated by new ideas based on fresh AI techniques and different features of the sensor signals.

In this chapter, many of the available modeling methods were discussed. Their benefits and disadvantages were presented and many research gaps in this field were identified. It will facilitate the selection of an appropriate modeling technique for different research purposes concerning milling processes. Also, some weaknesses from the research point of view in the field of machining technology were made clear.

Chapter 2

Proposed Methods and Results

2.1 Experiment Setup

¹ In all our case studies, the experiments are conducted on the 3-Axis Roders Tech RFM760 high speed milling center. For the workpiece material, Titanium and Inconel 718 were used. Several 2-flute and 3-flute micro-grain tungsten carbide ball-nose end mills were applied as the cutters to cut a surface with a slope of 60 degrees.

Most of our experiments were conducted on Inconel 718 because of its vast applications. Inconel 718 is a chemical composition of 53% Ni, 19% Cr, 18% Fe, 5% Nb, 3% Mo, 0.9% Ti, 0.5% Al and C balance [Ng et al., 2000]. It is considered as a hard-to-cut material in aerospace production field. Some other experiments were also conducted on Titanium workpiece.

The cutting process is conducted line by line from the top of the surface to the bottom as the cutter touches its flutes to the surface. Each surface

¹This chapter is consolidated into a journal paper and is now under consideration for publication at IEEE Systems Journal.

Figure 2.1: Workpiece and milling tool, cutting direction, and cutting method takes 300 to 400 cuts in our case studies depending on the depth of cut, cutting diameter, workpiece geometries, and cutting direction, Fig.2.1.

The destructive test experiments were conducted on the 3-Axis Roders Tech RFM760 high speed milling center. At certain preplanned moments of the process, the cutting process stopped for tool-wear and surface roughness measurements. Then a non-linear curve based on the tool-wear measurements were interpolated as the output of the process.

Figure 2.2: Workpiece and the arrangement of the sensors.

For experimental data acquisition, a vibration, force and acoustic sensor were attached to the workpiece and environment, Fig.2.2. For force sensor, Kistler 9257BA dynamometer with built-in 3-channel charge amplifier was applied on the workpiece, Section.C.1. For vibration, Kistler 8762A50 ceramic shear triaxial accelerometer was used while for capturing resulting acoustic signal, Kistler 8152B121 acoustic emission sensor were attached to the environment, Section.C.2. As the force sensors are not easy in their installation in industrial environments, vibration signal is another preferred choice for signal processing regarding the lower price of the sensor and its easy installation. Another type of sensor which is economically preferred is acoustic emission sensor. This sensor does not need to be mounted on the workpiece itself and is able to capture the changes in the cutting process.

For surface roughness measurements, we applied two different instruments. Taylor-Hobson instrument which was applied for surface measurements offline, Fig.2.3. Since this device is heavy and huge, it cannot be mounted in the HSM center and the measurements have to be conducted after the removal of the workpiece. The surface measurements were conducted in horizontal and vertical direction i.e. parallel and orthogonal to cutting direction. Another surface measurement device which was applied in our work was Motoyo surface measurement device, Fig.2.4. It has significantly smaller dimensions and was easily mountable in HSM center. It made the measurements easier without the need to remove the surface from the HSM for measurements.

Roughness measurement, R_a is the most commonly used surface quality indicator. It is the arithmetic mean of the absolute ordinate values of the

Figure 2.3: Surface quality measurements, Taylor-Hobson surface measurement instrument

Figure 2.4: Surface quality measurements, Motoyo surface measurement instrument

surface profile $Y(x)$ as presented in Eq.(2.1) and depicted in Fig.2.5. Its unit of measurement is the micrometer (μm) and the standards on surface roughness are accessible in ISO:1302 4288:1996 [iso,].

$$R_a = \frac{1}{L} \int_0^L |Y(x)| dx \quad (2.1)$$

For tool wear measurements, we have used two microscopes. The first microscope was able to capture only 2D pictures. Since in 2D photography, the ratios are not kept as are, we had to use the cutting edge of the tool as a scaling reference. Therefore, all the measurements were standardized ac-

Figure 2.5: Surface profile measurement details [Zain et al., 2010, Yang and Chen, 2001].

Figure 2.6: Tool wear being measured using a 2D microscope

cording to the measurements conducted on the known cutting edge thickness Fig.2.6.

The other microscope with more capabilities was Alicona InfiniteFocus which was able to capture the 3D pictures of the changes in the tool, Fig.2.7, Fig.2.8. The original tool edge picture was used as the reference and the rest of the pictures taken were compared to the reference for measurement of the tool edge wear.

For the microscope that did not have the 3D measurement capabilities, we have developed a method for more precise tool wear measurement. As

Figure 2.7: Alicona InfiniteFocus device for tool wear measurements.

Figure 2.8: Tool wear being measured with Alicona InfiniteFocus microscope

we were using an 2D photo to measure a 3D phenomena, the photo taking direction might vary each time. So we needed a reference for correction of the dimensions.

Since the primary angle of the applied tools was 100 micron, we measured the primary angle of the tool in each picture too. It gave us the correction factor which we could modify the results of the tool wear using this base.

2.2 Continuous Wavelet Analysis, Flute-Cut separation, Energy Feature Extraction

Wavelet processing of the signals has been applied in recent years and it leads to a considerable and meaningful analysis which facilitates a better understanding of the nature of the processes. In some experimental studies on machining processes it provides better information on the system and nature of the signals rather than Fourier transform [Wu and Du, 1996, Chen and Li, 2007, Patra et al., 2007]. The main reason behind choosing a data analysis method is the similarity between the signal under processing and the basis of the signal processing space. In milling process, the signals are mainly similar to mother wavelet functions rather than sinusoidal. Therefore, a wise choice to decompose the sensor signals are to apply wavelet analysis.

Just like other signal decomposition and analysis methods, wavelet is applicable on a energy limited time functions like $f(t)$. Fourier transform

Figure 2.9: The Raw cutting Signal.

look into the frequency content of the sensor signal,

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(w) e^{iwt} dt, F(w) = \int_{-\infty}^{+\infty} f(t) e^{-iwt} dt. \quad (2.2)$$

while the wavelet transform concerns about spectral content in the short period of time [Xiaoli and Zhejun, 1998, Daubechies, 1990]. Suppose $f(t)$ is a given signal and ψ is a mother wavelet function.

$$\psi_{a,b} = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right), a, b \in R, a \neq 0 \quad (2.3)$$

where a and b are dilation and translation factors, respectively.

Then the Continuous Wavelet Transform (CWT) is defined as following complex inner product:

$$w_f(a, b) = \int_0^{+\infty} f(t) \psi_{a,b}^*(t) dt \quad (2.4)$$

Figure 2.10: The obvious change between the sensor signal from the beginning of the cutting process to the end. (a) The 40th cut of the process with cutter 3T-12. (b) The 600th cut of the process with cutter 3T-12.

As the result the original function $f(t)$ can be rewritten in the following form:

$$f(t) = \frac{1}{c_\psi} \int_{-\infty}^{+\infty} \int_0^{+\infty} w_f(a, b) \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) da db \quad (2.5)$$

The application of this signal processing technique is not widely spread to the ball-nose surface finishing processes and there are very few papers in this area. In [Paul, 1994], a quick investigation on the benefits of wavelets' results is presented. Using wavelet packets, it is shown that the energy content of each scale in wavelet analysis is varying due to changes in the tool's status. To testify this idea, the energy content of some scales are shown to be changed from the moment that tool is new to the moment near tool-breakage. The difference between the shapes of the available signals are depicted in Fig.2.10,2.11.

In [Wang et al., 2002, Patra et al., 2007] some wavelet analysis of the vibra-

Figure 2.11: The obvious change between the sensor signal from the beginning of the cutting process to the end. (a) The 40th cut of the process with cutter 18-SC3. (b) The 600th cut of the process with cutter 18-SC3.

tion and power consumption signals were mentioned. Using wavelet packets, power and vibration signals are decomposed into several scales. Each scale shows different features regarding to the status of the tool. After detecting the current state, the features compared to a predefined feature bank is considered as a member of the codebook. Then, it will be distinguished that which state is the current state of the tool. Although there are questions on the suitability of the application of HMM for milling process by the author of the present study [Torabi Amin J., 2013a], the accuracy of the method which is a combination of Vector Quantization and the HMM is shown to be acceptable.

In the present work, first, using a supervised search among the scales of the CWT analysis of the force signal we find a leading signal for the flute-cut separation. Then, the energy content of several scales of each cut is considered as a blueprint of the inherent changes of the process. Then,

their features are used as inputs of the clustering-based tool-wear and surface roughness prediction system.

2.3 Neural Network and Adaptive Neuro-Fuzzy Inference System based models and results

2.3.1 Neural Network Based Methods

Artificial Neural-Network (ANN) and its combination with fuzzy logic like Fuzzy-Nets (FN) and ANFIS are widely used in modeling the HSM processes. They showed their capability not only in end-milling, but also in other kinds of machining processes, in providing an astonishing performance and approximation of the resulting surface-finish or tool-wear, [Ho et al., 2009, Kasirolvalad et al., 2006, Kirby et al., 2006, Hanna et al., 1996, Dweiri et al., 2003, Yang et al., 2006, Aliustaoglu et al., 2009].

There are several issues about the well developed ANN networks. The first issue is the network type that is applied. There are several choices for network architecture which is used for modeling such as Cascaded Feed-Forward Back Propagation (CFFBP), Elman BP (EBP), Time-Delay BP (TDBP), Adaptive Resonance Theory 2 (ART2)-type neural networks, Perceptron, Radial Basis Functions (RBF), SOM and SVM. In order to model the machining process, feed-forward BP algorithm has been used commonly in many researches. However, choosing the best structure is still an open

question and it is the main reason for many types of networks to be applied.

However, there is a trade-off between the accuracy of the model and the computation time as the second issue to be discussed. In [Zuperl and Cus, 2003] for example, feed forward back propagation is mentioned to be more precise and hence slow in training and testing while the radial basis functions are reported less accurate and comparatively faster.

The third issue is the structure of the network. The structure is considered as the way that input/output are connected. So, it is important that how many hidden layers are used and how many neurons each layer has. There are many papers which discuss about the optimum neural network structures [Zain et al., 2009, Muammer Nalbant and Toktas, 2007]. The forth issue is about the functions which are used in the training, learning and performance functions. A multilayer feed forward network is depicted in Fig.1.9. The input to neuron k in the hidden layer is as follows:

$$net - hidden = \sum_{j=1}^J C_{jk} i_j + \theta_k \quad (2.6)$$

where C_{ik} s are the weightings of the layer and i_j s are the input values to the network [Zain et al., 2009]. For the last layer, it is straight forward to drive:

$$net - output = \sum_{k=1}^K D_{kz} h_k + \phi_z \quad (2.7)$$

while D_{kz} are the weights and

$$h_k = f(net - hidden), o_z = f(net - output). \quad (2.8)$$

Figure 2.12: Multilayer feed-forward ANN structure use by [Zuperl and Cus, 2003] to model the surface roughness (R_a).

To choose a proper $f(\cdot)$ there are several options in the literature. Among them, 'logsig', 'purelin', 'fansig', and 'hardlin' are the most applied ones [Zain et al., 2009]. However, it is not easy to have a clear idea about the best choice for the network's function. As [Zain et al., 2009] declares, many papers use the 'logsig' transfer function. The transfer function is [Hassoun, 1995]:

$$f = \frac{1}{1 + e^{-net}}. \quad (2.9)$$

Then,

$$h_k = f(net - hidden) = \frac{1}{1 + e^{-\sum_{j=1}^J C_{jk} i_j + \theta_k}} \quad (2.10)$$

$$o_z = f(net - output) = \frac{1}{1 + e^{-\sum_{k=1}^K D_{kz} h_k + \phi_z}} \quad (2.11)$$

The resulting error will be defined according to how accurate the network follows the targeting value.

$$Error = \frac{1}{2} (o_t - o_z)^2 \quad (2.12)$$

The above mentioned function is called the performance function. There are several choices like Mean Square Error (MSE) and Mean Absolute Error (MAE), and Sum Square Error (SSE), Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). However, MSE is the most used one [Zain et al., 2009] and hence more appropriate. To train the network, the BP algorithm is applied. It uses the error to tune the weights of the network. The weight is usually updated using the 'delta rule' which uses the error and its changes to fine-tune the weights toward less errors in the sense of the performance function [Kosko and Burgess, 1998, Jain et al., 1996]. For the networks with hidden layers the weight-updating law can be written as follows:

$$\begin{aligned}\Delta C_{jk} &= -\eta \frac{\partial Error}{\partial C_{jk}} \\ \Delta \theta_k &= -\eta \frac{\partial Error}{\partial \theta_k} \\ \Delta D_{kz} &= -\eta \frac{\partial Error}{\partial D_{kz}} \\ \Delta \phi_z &= -\eta \frac{\partial Error}{\partial \phi_z}\end{aligned}\tag{2.13}$$

In [Tansel et al., 1995], ART2-type self learning neural network was used to detect the tool failure from the force signal. The method is based on the matching of the input and the feedback expectation. Besides the cutting conditions, the wavelet coefficients of force signal of the normal or new tool in several scales differ from the moment that the tool is broken or worn. Based on this fact, a neural network is designed to decide on the state of the tool.

ANN is not only applicable on prediction of surface roughness value, but

also on the surface profile. In [El-Sonbaty et al., 2008], cutting conditions and vibration is used to predict the surface profile. It uses the fractal's geometry and self-similarity properties of the surface as a reference building block for all the surface patterns and the determining fractal parameters.

Some properties of ANN must be determined before they were applied like topology, learning rate, training methods. They may also be encountered to over-fitting and slow-learning problems. Since there are many choices for ANN modeling, it is still an issue to choose the best structure. [Briceno et al., 2002, Lu, 2008, Venkatesh et al., 1997] are comparing the RBFs, back-propagation methods, and Dynamic models [Venkatesh et al., 1997]. Using only one hidden layer, based on White's theorem, and proper DOE the proposed model were presented to have the ability to capture the characteristics of the resulting force given the cutting conditions. Then, [Lu, 2008] renders the approximation of the surface profile while [Briceno et al., 2002] claimed that RBF has a better performance on resulting force prediction based on a reasonable cost function.

2.3.2 Problem Description and Proposed Neural Network based Techniques

The cutting process is a nonlinear multi-variable time varying process concerning its related parameters, like cutting conditions, signals and their features, tool preparation methods, tool geometry, tool wear and resulting surface roughness. So, to improve the products of the process, it necessitates to have a good model of the process. However, the framework of the feature ex-

traction and modeling was always an issue in the literature [Torabi Amin J., 2009]. Yet, there is no report of a global all-conditions considering method that can fully describe the cutting condition. So, the efforts to such model is still put in this area. In this part of research, the focus is on the repeated shapes of the signals resulted from the cutting process. Also, a noble method for flute-cut signal separation is presented. Then, a wavelet based feature extraction is presented. Wavelet analysis is chosen due to the similarity of the signals to the available standard wavelets. Also, for features and tool-wear correlation, neural network technique is applied. The repeatability of the results are also verified.

2.3.3 Discussions on Results of Neural Network Based Methods for milling processes

Many signals were collected during the cutting process. These signals visually carry the information about the process. It is observable that the shape of the cutting signals varies from an initial one to another state with different shapes, Fig.2.11 and Fig.2.10.

Since the shape and hence the frequency content of the signal is changing and the signal's shape is very similar to the original mother wavelet functions, specially DB and Bior, the spectral analysis of the signal has been done by these wavelet analysis. Besides, there are many benefits in wavelet-based processing and feature extraction of the signal that testifies the advantages of this method for signal pre-processing.

In addition to feature extraction, wavelet analysis is also applicable for de-

noising the signal from the high frequency industrial environment interfering signals. In our case study, it is used for peak localization of the flute-cutting signal.

Figure 2.13: Bior Wavelet Analysis of the Cutting Force Signal in 2D and 3D, (a) Bior wavelet analysis of the cutting force signal in 2D. (b) Bior wavelet analysis of the cutting force signal in 3D.

Figure 2.14: Meyer Wavelet Analysis of the Cutting Force Signal in 2D and 3D, (a) Meyer wavelet analysis of the cutting force signal in 2D. (b) Meyer wavelet analysis of the cutting force signal in 3D.

The flute-cut based cutting signal analysis is presented for the first time in the present work. Since the signal has an additive noise along, it is not an

easy task to localize the peaks of the signals. Therefore, the denoised version of the signal had to be used for peak-detection. Since we use the wavelet analysis for feature extraction, it is beneficial if it can also be applied for peak detection.

To find out the most proper mother-wavelet functions for the analysis of the cutting signal, many wavelet coefficients has been tested Fig.2.13, 2.14, 2.15, 2.16, 2.17. For the case of the present analysis, a mother wavelet which gives the smoothest approximation of the signal in some of its scales such that it could be thresholded for peak detection was needed. Since the energy distribution of many scales were used as features, it was desired for the signals have an obvious difference between lower and higher scales of CWT. Also, we have noticed that there is always a scale among these scales that the wavelet analysis provides a filtered noiseless smooth signal with the same base frequency of the original cutting signal.

Figure 2.15: Coiflets wavelet analysis of the cutting force signal in 2D and 3D (a)Coiflets wavelet analysis of the cutting force signal in 2D (b)Coiflets wavelet analysis of the cutting force signal in 3D

As the result of investigations for the best mother-wavelet analysis, since

Figure 2.16: Daubechies wavelet analysis of the cutting force signal in 2D and 3D, (a) Daubechies wavelet analysis of the cutting force signal in 2D. (b) Daubechies wavelet analysis of the cutting force signal in 3D.

Figure 2.17: Morlet wavelet analysis of the cutting force signal in 2D and 3D, (a) Morlet wavelet analysis of the cutting force signal in 2D. (b) Morlet wavelet analysis of the cutting force signal in 3D.

the output for some of their scales were properly noiseless, Bior and Daubechies were found suitable for the analysis. To facilitate the flute-separation process, a consistent search among the wavelet scales of the sensor-signals is conducted. Since there are 3D force, 3D vibration and 1D acoustic signals in our experiment design, wavelet analysis is performed on all 7 signals which leads to a complete decomposition of the signals in proper directions and var-

Figure 2.18: Wavelet-Based Analysis of the Cutting Force Signal in 2D and 3D and Localization of the Flutes

ious frequency bands. An interesting phenomena that was observed in the Roder high speed milling center during the cutting process, was the small drift of the speed during the cutting process. Therefore, the flute signal length which was supposed to be consistent by many reports in the field varies accordingly.

This problem and its proper solution is addressed for the first time in this work. For flute cut separation, first the sensor signals are pre-processed by wavelet-analysis. Then a threshold is set on the scales so that the peaks of the flute-signal can be detected properly. Because of the possible undesired peaks in some scales, there might be some inconsistencies in the length of the found flute cut signals in those scales of CWT. Therefore, there must be some verifications on the length of the flute-signals to be in some certain range. There is also another verification on the number of the flute-cutting

Figure 2.19: Separated Revolution Signals of a 2-Flute Cutter.

signals in a certain period of time. Since the cutting speed is pre-determined, the number of the revolution signals and the flute signals may not vary far from the expected one in a certain period of time.

Let v_c to be the cutting speed and *Sample/Second* the sampling ratio of the data-acquisition system. Also suppose that the cutter has *Flutes* number of flutes. Then the average samples per flute cut can be computed as the following:

$$\begin{aligned} v_c &= 10360rpm \\ sample/sec &= 50k \\ Flutes &= 3 \end{aligned} \tag{2.14}$$

$$E\{Flute - cutsamples\} = \frac{50,000 \times 60}{10,360 \times 3}$$

As mentioned above, because of undesired noise and peaks in the signals, the algorithm might find more flute cut signals surely with lengthier or shorter

width of signals. We have modified the algorithm so that it automatically search for the proper scales for flute cut separation. The results of the peak detection based on wavelet analysis can be seen in Fig.2.18.

Based on this localization, the flute-cut or revolution-cut signals are easily separable. Revolution-based separation for a 2-Flute cutter signals is depicted in the Fig.2.19, where the CWT is used to analyze the signal. This analysis includes a complete search on the several resulting signals of the process, the wavelet analysis of these signals and over 20 scales of each analysis.

Up to this level, the program can find the flute-signals properly. As a step forward, we refer to the corresponding wavelet scales and compute the energy content of each scale at that moment.

2.3.4 Application of Cascaded Feed Forward Neural Network (CFFNN) for Correlation of CWT Energy Features and Tool Wear

One of the modeling methods which is applied in our project is depicted in Fig.2.20. It is clear that each sensor's signal passes a CWT analysis block. This analysis is conducted on several scales. The CWT scale signals are the smoothed version of the sensor signals which are searched for their peaks regarding the expected flute-signal duration. They are also processed for the wavelet features which are used later as the input of other AI techniques such as CFFNN. As the result, any unreasonable flute cut signal is neglected and considered as a thresholding error. This type of analysis is tolerant to the slight drift of the cutting speed in HSM centers. This is the first time in the

literature that this problem is addressed and solved.

To check the repeatability and accuracy of the results, repeated random subsampling or monte carlo cross validation [Xu and Liang, 2001] is applied. Here we randomly select a subset of training data and then test the rest of the samples for testing. We also test the resulting trained model on another cutter to check the performance of the model. This method randomly splits the dataset into training and validation data. For each split, the model is fit to the training data, and predictive accuracy is calculated based on the evaluation of the model on validation data. Then the results are averaged over the splits. The benefit of applying this cross validation procedure which is that the proportion of the training/validation split is not dependent on the number of folds/iterations. The reported figures such as Fig.2.21, Fig.2.22, Fig.2.23 and Fig.2.24 are some among over 10 tests that are performed on the dataset.

An evolutionary dynamic fuzzy neural network is presented in [Pratama et al., 2011] where the parameters of DFNN is tried to be modified during the learning process using genetic algorithm. Its advantage is to evaluate the performance of DFNN based on its evolutionary developed parameters to find the optimal set of parameters. The reported accuracy on the process varies form 94 to 98%. However there is no testing another tool which makes it difficult to see the performance of the model in practical condition. Also, again it is pre-assumed that all data samples are available to the model with their desired value (measured tool wear or surface roughness) which is not the case in real conditions. Also the features are neither frequency domain features nor they are normalized. Setting all the novelties of the structure

Figure 2.20: Tool-Wear Prediction based on the Wavelet Features, Total System Schematic. (a) Tool-Wear Prediction based on the Wavelet Features. (b) Cascaded Feed Forward Neural Network.

aside, this model is not implementable in real industrial application due to above mentioned reasons.

The above method is applied on over several experiments, cutting conditions and several cutters with no report of unsuccessful flute cut or revolution cut signal detection. Having separated the flute signals, it is much easier to discuss the shape of the signals and address the cutting events according to different shapes. It is also facilitated to classify the signals and investigate the relation between the shapes of the signals and the sheering, rubbing and actual cutting phenomena during the milling process. The resulting classes can also be related to the different stages of the tool-wear process. For the present work, the energy features of the wavelet analysis of these flute signals is used as the input of the AI-based model. Any flute signal was already analyzed for its CWT scales. Energy content of each scale is computed and correlated related to the interpolated tool-wear value of the same moment of cutting. As mentioned in section 2.1, only in a limited number of times

Figure 2.21: The prediction of the Tool-Wear based on Wavelet Energy Features. Test is Conducted on the Same Cutter.

Figure 2.22: The prediction of the Tool-Wear based on Wavelet Energy Features. Test is Conducted on another Cutter.

during the tool life, tool-wear is measured.

To correlate the tool-wear and flute-based wavelet features, a cascaded neural network is used. The resulting configuration is tested on two different milling tool data. Half of data of the first tool is randomly selected to be used for training process. The rest half of the same tool and the whole data of another tool is used for testing and verification. Fig.2.21 presents the accuracy of the model on the same cutter's data. The accuracy according

Figure 2.23: The prediction of the Tool-Wear based on Wavelet Energy Features. Test is Conducted on the Same Cutter.

Figure 2.24: The prediction of the Tool-Wear based on Wavelet Energy Features. Test is Conducted on another Cutter.

to the Mean Relative Error (MRE) is calculated as 98.85%. However, for another cutter the accuracy is not the same and the calculated accuracy based on MRE is 85.97%. The resulting prediction and detection is depicted in Fig.2.22.

Comparing the above results with some other papers that use neural networks for modeling the milling processes we see an acceptable performance of our proposed neural network which works on wavelet energy features. For example [Li et al., 2006b] proposes a neural network for milling processes.

The authors use some time domain features such as maximum force level, total amplitude of force, average force, etc. The performance of neural networks on wavelet features leads to some improvements in terms of accuracy. In [Zuperl et al., 2006], all cutting parameters of the process are set as inputs of a neural network and the desired outputs of the model was to predict the cutting force, which can be considered as an inverse model of ours. The accuracy of back propagation and RBN networks presented there are 89.2%.

More recent works in this field, [Uros et al., 2009] applies ANFIS to correlate cutting conditions, time and force sensor signals as inputs and flank wear as output in a destructive test which can be considered as very similar to our study. As reported, the accuracy of ANFIS model to predict the flank wear is 93.6%. The major difference that this study has with our models is that they incorporate time as an input. It makes the modeling very dependent to conditions or in other words, the model will not be much useful if the process starts with a half worn tool or even if we have a stop in the cutting process for any reason. In contrast, our models in this study relies “only” on sensor signals which helps us to perform the diagnosis and prognosis from the start of the process whether a new tool is used or even a worn tool. Another superiority of our model is application of vibration and acoustic signals. It is a very important step forward as these sensors are much easier to be installed in industrial environments than force sensors. Another superiority of our results are normalization of sensor signals which we use in next chapters. This will make the model robust to the place of the sensor. It is obvious that the amplitude of sensor signals may drop if they be installed a little bit farther from the cutting point. It will affect those

models that apply the sensor signals “without” normalization.

Extreme learning machines (ELM) are also applied to the destructive tests on milling process, [Javed et al., 2012]. The deliverables of this method are the same as the deliverables of our methods. The features used in this study to predict the tool wear are again time domain features. The accuracy of the presented model based on ELM varies from 92% to 94% which is comparable to what we have earned in our studies. While the ELM method is interesting to be applied and may have a very bright future in diagnosis and prognosis of milling processes, usage of time domain features specially when they are not normalized has the above mentioned deficiency of lack of generalizability. This issue is mentioned in the conclusion of the study too. Based on the results of our study, we suggest the application of normalized frequency domain features to make the model more reliable.

2.3.5 Conclusion on the application of Neural Network Based Methods for milling processes

In many related researches on HSM processes, artificial neural networks have been utilized due to their capability of modeling the nonlinear processes. Both, best cutting condition determination experiments and destructive tests seem to be well described with these AI methods. There are plenty of different implementations of ANNs concerning the prediction and control matters. There are some reports that claim the repeatability of these models. Not all types of these models are investigated for a universal model for cutting process. Also, there is no consistency in the parameters which have been

used as the inputs. Therefore, there is still a visible gap on the completion of these studies with such models where there are no general meaningful signal features and model structures. In addition, there are very few on-line prediction and control researches available in the field which leaves space for more investigations on universal models. So, Cascaded Feed Forward Neural Network (CFFNN) was one approach among many on the investigations for a repeatable reference models for high speed ball-nose milling processes. A CFFNN based model is tried to fit the experimental data and is checked for its repeatability for some other cutters. The main goal here in this chapter was to show that an AI model can capture the main behavior of the milling process using its normalized wavelet features. Here we have to mention some achievements and novelties of this part of study:

- First, cutting signal into flute cuts. This was not in literature before. It enables the model to focus on each step of cutting process and is a considerable step forward for online tool condition monitoring.
- Second, application of normalized energy features of wavelet analysis. They are used as blueprint of tool states. Normalization makes the sensor signals independent of the placement of the sensor, a very considerable achievement in the literature of tool condition monitoring.
- Third, discarding the use of time factor as an input of the model. it enables the model to start prediction even on a worn or half worn tool while a model that incorporates time as an input, has to know the exact time passed the process. It makes it easier to apply and more industrial environment friendly.

After all, it is obvious that besides all successes and accuracies of such methods to model the milling process, the fixed structure is their main weakness. This is the root cause of some imprecisions and inflexiblenesses of such models in capturing the side changes in new tools.

Let us now step forward and check the possibility of a self constructive model. The models we applied so far had some deficiency. Their structure were fixed and once they are trained, they had to be used as were. However in industrial environment many parameters may drift or change and require the model to be slightly modified.

To get a more robust and repeatable model, we had to look for a signal organizing scheme that has the capability of repeatable results and modifiability for sequential organization during the process. Development of such models without loss of meaningfulness of the rules and model structure requires to check the applicability of clustering methods on our input data space. Next chapter will report our investigations on such issues.

2.4 Application of Clustering Methods on Wavelet Features

Continuous and Discrete Wavelet Transform were extensively used as a feature extraction for milling signals. It is reported that distribution of energy patterns in wavelet scales of cutting signals varies according to gradual change in tool condition. Then, the effect of the tool degradation on each frequency band can be detected by time-frequency analysis like wavelet analysis. As

reported in [Xiaoli and Zhejun, 1998], four states for the tool-wear with 7 features each, compose the codebook of the clustering method. Also in [Li et al., 2000] a fuzzy clustering approach were used for detecting the tool breakage of spindle on feeder current signals. To verify the suitability of our proposed method, the results of the clustering layer must fulfill two major characteristics. Firstly, the next AI layer must have the same activation level. This consideration can be achieved by normalization of the signals or the outputs of the interpretation layer. Another characteristic is the repeatability of the results among cutters of the same type. In our proposed method, these two characteristics were emphasized for the model. First, the signal features were converted into standardized outputs by the applied clustering method which is used by AI layer. Second goal is also attained by verifying the consistency and repeatability of clustering methods.

While early models for HSM process relied on tool-geometry and cutting conditions, state-of-the-art models consider time and/or other experimental factors to make models more accurate. Combining these features with other cutting conditions and specifications, the model will explain the changes in resulting surface quality and tool condition during the cutting process. It is clear that when the cutting tool changes, there must be always proper modifications in the model. Our idea is that, there must be an interpretation layer between resulting signals and the artificial intelligence model so that these changes never affect the entire model. It will be perfect improvement in the existing models if we can suggest a model with only one changing part according to tool replacement. So, in our proposed model, the only necessary modification is the amendment in this interpretation layer. This

change in the interpretation layer will protect the change in the weights of the artificial intelligence instead. As an example to verify our idea, many available clustering methods were applied on wavelet features of resulting signals. Clustering methods are the methods to keep similar data together in clusters so that another overview of the knowledge will be facilitated. In hard or crisp clustering methods, clustered data belongs to only one group while in fuzzy clustering methods incoming data can be a member of several clusters with a certain membership value. The process of assigning the incoming feature vector to one cluster or multiple clusters is according to the distance or connectivity of the data to other data in that specific cluster. Among the available clustering techniques, K-means, K-medoids, Gath-Geva and fuzzy C-means clustering are investigate in our research [Torabi Amin J., 2009, Torabi Amin J., 2013b].

Fuzzy c-means clustering method has been used in [Xiaoli and Zhejun, 1998] and [Wang et al., 1996] on AE sensor signals of the milling process while in [Li et al., 2000] it were used for tool breakage detection of spindle and feeder current signals. Wavelet packet transform and Discrete Wavelet Transform (DWT) was used for signal feature extraction respectively. It is observed that according to different tool conditions, RMS value of each frequency band in AE signal changes. Then, the effect on each frequency band detected by wavelet packets were investigated for different tool-wear conditions and the proper patterns to look for were solidified.

To verify suitability of the clustering methods, specially Fuzzy C-Means clustering, the results of this clustering layer must fulfill two major characteristics mentioned above. The two considerations will be ensured through

our method and verified in the following sections.

Section 2.4.3 presents outputs of the proposed clustering layer and discussions are made accordingly.

2.4.1 Problem Description on Clustering of Wavelet features

As described in Section 1, there are many methods to extract signal features to be used in tool performance estimation models. Some of them are time-domain while some others are frequency domain features. In our work, we focused on time-frequency, precisely wavelet, features of resulting cutting signals due to their similarity to some existing mother wavelets and ever increasing demands for time-frequency analysis. This requires time-based cutting phenomena identification and associating them with certain patterns in resulting signals. In our approach, feature extraction of signals are based on energy computation of different scales of wavelet analysis. The details will be provided in section 2.4.2.

There are many uninvestigated and unclassified features to be studied by clustering and classification methods which leaves space for more research. Also, there are quite number of classification methods that have not been applied in the field of intelligent machining. Besides, the combination of the classification methods and AI techniques has not been extensively investigated in the field. Classification methods can be used to investigate the similarities between different cutters' signal-features. Finding the similarities, other AI techniques can be applied to map different classes to different

respective conditions of the tool and cutting process. However, number of classes may determine the accuracy of the method and it is an open issue to determine this number. Similar to other AI techniques, the structure of the classification methods for machining process analysis need further investigations, in case there is not going to be a dynamic structure for modeling.

The model structure that is considered for the milling process have to lead to a generalizable descriptive model. This structure must be able to combine all available information on milling process, either from destructive tests or best-cutting-condition determination experiments [Torabi Amin J., 2009, Torabi Amin J., 2010b]. Our proposed structure is introduced in Fig.2.25. However, the applicability of each block of the model structure have to be verified separately before they are combined together. In [Torabi Amin J., 2010a, Torabi Amin J., 2010b], the applicability of some AI techniques were presented. In this part of the report, the contribution of clustering methods and the repeatability and consistency of their results is presented. Since the proposed structure is tried to be generalizable, it is based on the clustering of signal features. This characteristic makes the model more reluctant to changes due to tool replacement. Therefore, the results of clustering on the available experimental data of milling process will facilitate the generalization of the reference model. To verify the applicability of the clustering methods, some of them are applied on the extracted milling signal features.

As described in Section 1.1.3, there are many methods for signal feature extraction. Some are time-domain methods whereas the others are frequency domain which are used in tool condition monitoring. In our research, we focused on wavelet features of resulting cutting signals as they are shown to

carry consistent information on of the process [Wu and Du, 1996, Chen and Li, 2007, Patra et al., 2007]. Also the similarity between cutting signals and existing wavelets are another reason for this time-frequency analysis. Also, as it is time-frequency based analysis, it facilitates time-based identification of any undesirable phenomena in cutting process and associate them with certain patterns in resulting signals. The details are similar to method that is applied in our papers, [Torabi Amin J., 2010a, Torabi Amin J., 2010b] and were briefed in Section 2.2.

Here we try to push forward towards covering these gaps in literature:

- 1) There are many clustering methods that have not been applied in the field of intelligent machining.
- 2) The combination of the clustering methods and AI techniques has not been extensively investigated in the field.

Clustering methods can be applied to group the similarities between many cutters' signal-features. Finding the similarities, other AI techniques are to be applied to map different clusters to different respective conditions of the tool during process. The main challenge here is to determine the most suitable method among some well-known clustering methods while the criteria is smooth results while maintaining less complexity and generalizations for sequential online clustering.

2.4.2 Proposed Clustering-Based Reference Model

Experimental set-up section 2.1 highlights that there are three sensors to collect resulting signals of the cutting process. These signals are sampled

with relatively high sampling rates. To make them available for possible models of cutting process, some feature extractions were conducted which is wavelet analysis of vibration and force signals. This analysis is conducted in several scales to check variations of signal energy distribution due to changes in tool's status from new to old. Since there are some moments during milling process that the cutter is not cutting the surface in order to move to another line of cut, analysis starts with checking the presence of valid cutting signal. As in real world implementation, the cutting sound is applied for monitoring an ongoing cutting process, in our research we also applied Root Mean Square (RMS) value of acoustic emission signal as a sign of presence of the cutting signal. It is checked to be higher than a pre-specified level and in the moments of positive detection the signal is treated as a cutting signal.

After detecting the cutting portion of the signal, CWT of force and vibration signals were calculated. Taking some certain scales of wavelet analysis into consideration, the program looks for the peaks of the signal. The details of the peak detection method is described in Section 2.2. These peaks represent the moments that tool-flutes hit the surface of the material for cutting purpose. Later, every flute signal is localized and searched for its beginning moment. All flutes are kept for later analysis and beginning of the signals are referred later for the corresponding wavelet analysis and energy calculations of each individual flute cutting signal. According to the time of detected flute hits, all scales in wavelet analysis are cut. Energy content of each scale which is calculated like Eq.(2.15) creates unique patterns that can

be associated with a certain levels of tool-wear.

$$E(f) = \frac{1}{T} \sum_{k=1}^T f^2(k) \quad (2.15)$$

Figure 2.25: Proposed Method Using Fuzzy C-Means Clustering Layer as Interpretation Layer.

To show the consistency and repeatability of clustering methods, signal feature patters are unsupervisedly grouped by different clustering techniques such as K-means, K-medoids, fuzzy C-means and Gath-Geva algorithm. Some of these techniques such as K-means and K-medoids are considered as crisp clustering techniques where each data point belongs only to one specific cluster while the other clustering approaches like fuzzy C-means and Gath-Geva are fuzzy clustering where each data belongs to all cluster with a specifiable membership degree.

K-means clustering is a method of cluster analysis that partitions n num-

Figure 2.26: Wavelet Analysis of the Entire Cutting Portion of the Signal. The top is the 2D wavelet analysis of the signal. Below is the original signal

ber of data into k clusters where each data belongs to the cluster with closest mean. This method tries iteratively to optimize the following objective function:

$$\arg_S \min \sum_{i=1}^k \sum_{x_j \in S_i} \|x_j - \mu_i\|^2 \quad (2.16)$$

where k is the number of clusters and S is the set of all clusters, i.e. $S = \{S_1, S_2, S_3, \dots, S_k\}$ and μ_i s are the means of corresponding S_i .

The k-medoids clustering algorithm is quite similar to the k-means algorithm. Both of them are partitional clustering, i.e. they break the data space into clusters, and both attempt to minimize squared error, the distance between points labeled to be in a cluster and the center of that cluster. However, in k-medoids method the center of the cluster is chosen from the actual data points in the cluster.

Gath-Geva on the other hand, is considered as one type of fuzzy clustering approaches which has no preassumption on the size of clusters. Probabilisti-

cally, it assumes the clusters as summation of different Gaussian functions. The details on this method can be found in [Gath and Geva, 1989].

Fuzzy C-means technique was originally introduced by Jim Bezdek in 1981 as an improvement on earlier clustering methods [Bezdek, 1981]. It provides a method that shows how to group data points that populate a multidimensional space into a specific number of different clusters [MAT,].

Figure 2.27: Wavelet-based localization of the flute signals. Top-left is 3D presentation of the signal. Top-right is peak localization. Below is one of scales to show the effects of wavelet on denoising the signal.

Fuzzy C-means clustering algorithm is based on the iterative minimization of the objective function, called C-means functional. It is defined by Bezdek [Bezdek and Dunn, 1975, Bezdek, 1981, Pal et al., 1996] as:

$$J(X; U, V) = \sum_{i=1}^c \sum_{k=1}^N (\mu_{ik})^m \|x_k - v_i\|_A^2 \quad (2.17)$$

where $V = [v_1, v_2, \dots, v_c]$, $v_i \in R^n$ is the set of all n -dimensional centers, μ_s are the membership value of each data in a certain cluster and N is the number of data points. Also, $D_{ikA}^2 = \|x_k - v_i\|_A^2 = (x_k - v_i)^T A (x_k - v_i)$ which is weighted norm-2 of the distance.

The stationary points of Eq.(2.17) objective function is calculated by adjoining the following constraints to J by means of Lagrange multipliers:

$$\begin{aligned} \mu_{ij} &\in [0, 1], 1 \leq j \leq N, 1 \leq k \leq c \\ \sum_{k=1}^c \mu_{ij} &= 1, 1 \leq i \leq N \\ 0 < \sum_{i=1}^N \mu_{ij} &< N, 1 \leq k \leq c \end{aligned} \quad (2.18)$$

Then result will be:

$$J(X; U, V, \lambda) = \sum_{i=1}^c \sum_{k=1}^N (\mu_{ik})^m D_{ikA}^2 + \sum_{k=1}^N \lambda_k \left(\sum_{i=1}^c \mu_{ik} - 1 \right) \quad (2.19)$$

To minimize Eq.(2.19), the gradients of J with respect to U, V and λ have to be set to zero. If $D_{ikA}^2 > 0, \forall i, k, m > 1$, then (U, V) minimize Eq.(2.19) by iterating the following steps:

$$\mu_{ik} = \frac{1}{\sum_{j=1}^c (D_{ikA}/D_{jkA})^{2/(m-1)}}, 1 \leq i \leq c, 1 \leq k \leq N \quad (2.20)$$

and

$$v_i = \frac{\sum_{k=1}^N \mu_{ik}^m x_k}{\sum_{k=1}^N \mu_{i,k}^m} \quad (2.21)$$

Fuzzy c-means clustering is used in [Xiaoli and Zhejun, 1998] and [Wang et al., 1996] on AE sensor signals of the milling process. Also in [Li et al.,

2000] it were used for tool breakage of spindle and feeder current signals.

The Fuzzy Maximum Likelihood Estimates (FMLE) clustering algorithm employs a distance norm based on the fuzzy maximum likelihood estimates, proposed by [Gath and Geva, 1989, Bezdek and Dunn, 1975]:

$$D_{ik}(x_k, v_i) = \frac{\sqrt{\det(F_i)}}{\alpha_i} \dots \exp \left(\frac{1}{2} \left(x_k - v_i^{(l)} \right)^T F^{-1} \left(x_k - v_i^{(l)} \right) \right) \quad (2.22)$$

where

$$F_i = \frac{\sum_{k=1}^N (\mu_{ik})^m (x_k - v_i)(x_k - v_i)^T}{\sum_{k=1}^N (\mu_{ik})^m} \quad (2.23)$$

$$\alpha_i = \frac{\sum_{k=1}^n \mu_{ik}^m}{\sum_{i=1}^c \sum_{k=1}^n \mu_{ik}^m}$$

In the above, F is the positive definite matrix that controls the shape of the cluster, F_i is a fuzzy covariance matrix, v_i is the center of cluster i , α_i the a priori probability of selecting the cluster, and x_k is the k -th clustered data. The algorithm, fuzzy maximum likelihood estimates clustering algorithm, is reported to be able to detect clusters of varying shapes, sizes and densities [Gath and Geva, 1989]. The cluster covariance matrix is applied in conjunction with an “exponential” or “Gauss” distance, and the clusters are not constrained in volume.

To show the consistency and applicability of the clustering methods in milling process, the above mentioned clustering methods applied on milling process signal-features.

The training results show that membership values and cluster numbers assigned to patterns at different times in fuzzy clustering/crisp clustering methods highly reflect the status of the cutter. The repeatability of the

results of clustering method is also checked for different cutters, Section 2.4.3. It is shown that clustering methods can provide an appropriate platform to detect the status of the cutters of similar geometry. Applying these results facilitate a generalizable reference model that leads to robust non-intrusive prediction of cutter's performance.

As Fig.2.25 illustrates, clustering methods can be employed in proposed structure as an interpreting layer in general reference model for tool or resulting surface condition detection. As far as the applicability of this model in real-time condition monitoring is concerned, since these patterns are repeated in different experiments, it is applicable to real-time condition detection after training the model and adjusting the clustering subsystem. In the final structure, the output of the clustering method is fed to the AI-based block to modify and utilize the clustering results with other cutting conditions and tool specifications as well as results from different experiments on the tool such as Best Cutting Condition Determination Test (BCCDT), etc. We expect that this can lead to a more flexible and generalizable modeling structure.

Following the above motivation, clustering results show that changes in membership values of patterns at different times are highly due to changes in the status cutter. It is also checked for different cutters to confirm that the patterns are repetitive. The clustering results will be mentioned in the next section. It is shown that fuzzy C-means clustering can provide a very appropriate platform to determine the status of the cutters of similar geometry. This leads to a general reference model that facilitates prediction of performance of cutter non-intrusively. This clustering layer can be employed

as an interpreting layer in general reference model for tool condition or resulting surface quality detection. One issue is the applicability of this model in real-time analysis. Since these patterns are repeated in different cutters, it is applicable to real-time condition detection after training the model and adjusting the clustering subsystem.

Using this technique, an AI-based structure which applies clustering as an interpretation layer is presented Fig.2.25. In this structure, the output of the clustering method, as the preliminary judgment on the status of the tool, is fed to the AI model to adjust the clustering results with other cutting conditions and tool specifications.

2.4.3 Clustering Results and Discussions

As aforementioned in section 2.4.2, first analysis which is conducted on collected signals is wavelet analysis of whole cutting portion and the calculations on flute signal localization as depicted in Fig.2.26 and Fig.2.27 respectively. The former depicts the original signal and its 2D wavelet representation while in the latter, a portion of cutting signal which in this study these portions include whole the cutting portion, is analyzed for its CWT scales. Then, certain scales of the signal are looked for their peaks which also represent the peaks of original signal. The search is conducted on all force and vibration signals to ensure the validity of the peak locations.

Resulting wavelet based peak localization is depicted in Fig.2.27 where CWT analysis of force and vibration in various scales is searched for appropriate number of peaks according to cutting conditions, sampling rate and

time duration. Using the time label of the peaks, all resulting signals are decomposed into flute cutting signals. The Flute cutting signals and their energy distribution in different wavelet scales are then used as feature patterns to typify different moments of tool life. In Fig.2.27 it can be seen that all the peaks are successfully captured and location of the flute signals are determined considering the high level of environmental additive noise.

An important point to focus on here is the application of clustering methods for diagnosis and prognosis. Not always we want to predict the output in industry. There are some cases that industry prefers to have an alarm on their system to warns the moment it reaches to critical point. Here, according to the results and the simplicity of their implementation we may suggest clustering methods for fault diagnosis.

When the concern is prediction of the process, many proposed methods just use these features as the inputs to AI-based model. In contrast, based on discussions in [Wang, 1999], we suggest a clustering layer to take care of the interpretation of the input signal features and to determine the current state of the process. To verify the idea, clustering of the extracted features of the destructive test is examined first for its meaningful results and then for its repeatability and robustness. As the first try available fuzzy C-means clustering routine of MATLAB is applied to cluster these features and FCM clustering with 3 target classes is applied on extracted features of force signal. The result is depicted in Fig.2.28 where it is clear that the features are following a certain trend that distinguishes three different stages of the tool life. Since the classes are clearly separated without too much noise, the membership values of the features to these three classes can be considered as

a sign of tool-wear development.

Figure 2.28: Fuzzy C-means clustering is applied on extracted features of a destructive-test force signal with 3 classes.

It is important to note the input data is not labeled with the process status before clustering. Therefore, if one of the resulting clusters is going to be considered as a representative of “New Tool” signal patterns in Fig.2.28, it is related to “Red” membership values that have the higher membership values in the beginning of the experiment. The cluster with “Blue” membership value results can clearly be considered as the membership values according to “Normal Tool” cluster. And similarly the “Green” results are those patterns which are related to worn tool.

Since clustering results of a destructive test on one cutter may be totally different from others, the same procedure is applied on the extracted features of three slightly different cutters which only differ in their edge preparation and coating methods. Fig.2.29 shows that fuzzy clustering results are robust

regarding to cutters with similar-geometry.

Figure 2.29: Fuzzy C-means clustering is applied on extracted features of three destructive-test force signal with 3 classes.

However, aforementioned analysis are conducted on force signal which is comparatively less noisy than the vibration and acoustic signals. The issue with force sensor is that it might not be a suitable choice for industrial platforms due to its expenses and difficult and costly installation. Conversely, vibration and acoustic sensors are more eligible for these environments and comparing to force sensor, their resulting signals are more noisy. Thus, it is expected to have less quality results after clustering of vibration signal.

Having considered this fact, if clustering results of these signals follow a reasonable and distinguishable trend, it is worth replacing force signal in tool condition monitoring. Same analysis on vibration signal is conducted to check if the vibration features can deliver same information. Since vibration signal is more noisy than force, it is natural to have noisier response. Fig.2.30

shows the noisy clustering results of vibration features where they still follow a trend which can lead to a clear clue on the tool ageing. To ensure that the trend will be followed by other cutters, the test is repeated again for three cutters. Fig.2.31 illustrates that vibration wavelet features fall into the same clusters for three slightly different cutters.

Figure 2.30: Fuzzy C-means clustering is applied on extracted features of a destructive-test vibration signal with 3 classes.

Therefore, according to proposed structure in Fig.2.25, these clustering blocks can be applied to modify traditional AI-based models so that they can be more robust to the change of tools. Here, the outputs of block will be always membership values of extracted features to pre-specified classes and these values are always between 0 and 1 and hence the output level remains within certain bounds. Therefore, a clustering block can be replaced by another clustering block only by one successful destructive test. The rest elements and weights of the AI-based model might be trained through neces-

Figure 2.31: Fuzzy C-means clustering is applied on extracted features of three destructive-test force signal with 3 classes.

sary best cutting-condition determination experiments. These experiments are usually easy to conduct and for many typical cutters the results are already available in engineering handbooks and some other published resources. Our proposed method suggested a combined knowledge-based method to provide a simple reference model that gathers both benefits and information of destructive tests and best cutting condition determination tests.

To verify the applicability and consistency of the clustering results of wavelet features, first the clustering block was applied for training on one cutter experiment. Then each trained clustering method is verified for its clustering performance on 4 more cutters of the same geometry. Altogether, clustering methods are applied on five destructive test experiments. In these 5 destructive tests, ball-nose milling tool is applied from its “new” condition till the moment in which it is considered “worn”. Therefore, it is expected

Figure 2.32: K-means clustering is applied on extracted features of a destructive-tests signal features with 3 clusters. Crisp clustering results are shown. To present the density of the presence of cluster members, the distribution of presence of each individual cluster is also depicted above.

Figure 2.33: K-means clustering is applied on extracted features of four destructive-tests signal features with 3 clusters. As it is crisp clustering method, distribution of presence of each individual cluster is depicted above. X-axis shows the sequence of incoming samples that can represent time.

that different clusters represent different conditions of the tool. As the clustering method is an unsupervised learning scheme, the reasonableness of the results is verified when the membership values assigned to each cluster vary meaningfully with the gradual change in the process's status.

As the first try on clustering methods, K-means clustering with three target clusters is applied on extracted features of force signal. As aforementioned, this method is categorized under the crisp clustering methods. Therefore, its outputs are the sequence of cluster numbers clusters that the input data is associated with. After training with the data from one cutter, the clustering model is verified on other four cutters of the same geometry. To make the results clear and more understandable, the time distribution of the occurrence of each cluster is presented in Fig.2.33. It is clear that the features are following a similar trend that makes three different conditions of the tool life detectable. Since the cluster are clearly separated, the membership values of the features to these three clusters can be considered as a sign of tool-wear development.

Since the data is not labeled before clustering (unsupervised clustering), an explanation on the results can be drawn as follows. If one of the resulting clusters is going to be considered as a representative of "New Tool" signal feature patterns, it has to have the higher values in the beginning of the destructive test when the tool is new. Following the same analogy, the cluster that shows ever-increasing membership values during the destructive test have to be considered as the representative of the "Worn Tool" and the one that shows a bell-shaped trend must be considered as an indicator of the "Half-Worn" or "Normal" condition. Therefore, in Fig.2.33, the cluster whose

Figure 2.34: K-medoid clustering is applied on extracted features of a destructive-tests signal features with 3 clusters. Crisp clustering results are shown. To present the density of the presence of cluster members, the distribution of presence of each individual cluster is also depicted above.

Figure 2.35: K-medoids clustering is applied on extracted features of four destructive-tests signal features with 3 clusters. As it is crisp clustering method, distribution of presence of each individual cluster is depicted above. X-axis shows the sequence of incoming samples that can represent time.

membership values are depicted in “Red” is representing the “New” tool and “Green” membership values can represent the “Half Worn” tool while “Blue” graph represents the progress of the tool toward its worn condition.

Since the goal is to show the applicability of the clustering methods and to determine the most suitable type, other clustering methods were also tested for their performance. Fig.2.35 represents the results of K-medoids clustering method on milling signal features.

Figure 2.36: Fuzzy C-means clustering is trained on extracted features of one destructive-test force and vibration signal with three clusters. It is verified on the above four destructive signals. X-axis shows the sequence of incoming samples that can represent time.

Two other fuzzy clustering methods that has been experimented on milling process are Gath-Geva and Fuzzy C-Means (FCM) clustering. The preliminary results for FCM method has been reported in [Torabi Amin J., 2010b]. The results of method has been shown repeatable and consistent among cutters of the similar geometry. Its results on pure vibration signal are also shown to have the same characteristics and repeatability. Here, the same method applies both force and vibration wavelet features to form three clus-

ters. As it is depicted in Fig.2.36, pre-defined number of three clusters are able to capture the major changes in the tool condition. As compared to the Gath-Geva clustering results, Fig.2.37, which provides another fuzzy clustering method, Gath-Geva method has the advantage that the membership functions reach their maximum value of 1 in each individual cluster which makes the distinction between three stages of tool life easier. However, regarding the complexity of calculations, modifiability for sequential clustering as a way toward its application for online monitoring, FCM is the preferred method among all.

Figure 2.37: Gath-Geva clustering is trained on extracted features of one destructive-test force and vibration signal with three clusters. It is verified on the above four destructive signals. X-axis shows the sequence of incoming samples that can represent time.

All these methods are applied on the wavelet features of both force and vibration signals. As mentioned above, while it is shown that the results of pure vibration signal is more noisy than force signal analysis, the clustering methods can be applied and provide acceptable results [Torabi Amin J., 2010b]. Let us here do the analysis on force signal which provides less noisy

features comparing to vibration and acoustic emission signal. Same analysis on vibration signal is conducted to verify that the vibration features deliver same information.

For these models that apply artificial intelligence schemes directly on signal features, it is clear that clustering methods can contribute a lot for repeatability, generalizability, and online monitoring. Considering the above mentioned results, these clustering methods can modify traditional AI-based models in the way that they can be more robust to the change of tool conditions while incorporating the results of other tests on the tool such as BCCD experiments, etc. More details are explained in [Torabi Amin J., 2010a] and Section 2.4.2. However, there are some issues with the present algorithm. First issue is misclassification of some data patterns. As the results present, any online fault diagnosis or prognosis method have to be not too sensitive to misclassification of one input signal feature pattern in order to avoid any wrong decisions based on noisy data. Luckily, the process is slow enough for the condition detection system to get enough samples before stopping the machine or reporting any alarms. However, some of the clustering imperfections might have some origins in cutting process side phenomena.

Note that clustering analysis is not suitable to be placed as an interpretation layer in all models. An exception that can be mentioned here is the case where we have uniformly distributed input samples in input space. In such cases other modeling methods can be applied to correlate the input/output pairs, i.e to model such cases using fuzzy logic, rule base can be formed based on the input/output pairs of data. In contrast, cases that the input data can be clustered properly, the rules can easily be defined for a cluster of data

rather than for individual pairs of data, Fig.(2.38). The resulting rules can then be formulated and presented in the form of mathematical functions too using proper choice of fuzzifiers, defuzzifiers and inference engine. Please refer to [Wang, 1999] chapter 15 for more details on the benefits of application of clustering in fuzzy modeling. Therefore, it is important for us to show that the signal features in our study are able to be clustered using available clustering methods and application of this clustering algorithms will result into meaningful separation of the data samples in input space as well as reducing the dimension of the input space.

Figure 2.38: An example of constructing IF-THEN rules from input/output pairs when the input data can be grouped into limited number of clusters [Wang, 1999], U represents input space, V represents output space and A_i^j and B^k s are the fuzzy sets defined on the input and output space, respectively.

To verify such results, the clustering imperfections have to be correlated to the physical phenomena on the ball-nose face-milled surface or the tool. If there is find any correlation between these two happenings, new clusters come to the knowledge data base to make it more perfect for explaining the milling process. These investigations will make the future work of this study.

Comparing the results of clustering algorithms with some methods from the literature which use wavelet features of the signal, the superiority of clustering methods can be shown. In all such methods we do not expect the models to provide exact prediction of the outputs of the process, such as tool wear or surface roughness. As the output of clustering algorithms are membership value of the features to the clusters, we expect them to qualitatively compare these features.

For example in study [Zhai et al., 2011] which actually shares the same dataset as ours for one cutter, the quality of the process is shown to be reflected in standard deviation and magnitude of some levels of wavelet analysis of the cutting signal. However, the “pre-critical”, “critical” and “post-critical” stages of the cutting process were introduced based on the “average” magnitude of high frequency detail of the wavelet analysis of one cut signal. This can be in some cases unreliable as the signal might change drastically during one cut when the workpiece is hard-to-cut material. In contrast, based on the vectors of energy features of CWT of the signal, every “time window” of the signal, which is very small compared to the time for one cut, is analyzed by wavelet analysis, its features are extracted, clustered and the decision can be made on the moment. Another superiority in our study is the reliable use of vibration signals rather than force signals. This is a very important step forward considering the fact that vibration sensors are much easier to install.

Chapter 3

Development of a New Dynamic Fuzzy Neural Network (DFNN): Sequential Fuzzy C-Means Clustering Based DFNN

3.1 Sequential Fuzzy C Means Clustering

¹In previous sections we have shown that clustering methods, and fuzzy c-means specifically, are powerful tools for milling process status detection. The clustering scheme provides a platform for an artificial intelligent subsystem

¹This chapter is consolidated into a journal paper and is to be submitted to Fuzzy Sets and Systems journal.

to provide a prediction of the process quality, i.e either tool wear or surface roughness.

The definition of the fuzzy clustering were briefed in section 2.4.2. In the basic definition and solution by Bezdek in [Bezdek, 1981, Pal et al., 1996], iterating the equations Eq.(2.20) and Eq.(2.21) would result into the optimal solution for the functional proposed in Eq.(2.17).

Inspired by their alternating optimization (AO) solution between half-step defined at Eq.(2.20,2.21), knowing the possible weaknesses of the sequential clusterings mentioned in [Pal et al., 1996], we propose a sequential scheme for fuzzy clustering based on the idea of absorbtion of the samples to the cluster centers.

First, each new data sample is measured for its distance to the centers of the existing clusters. Then, just like the Bezdek's solution [Bezdek, 1981], the membership values of the new data object to each cluster is calculated as Eq.3.1 where u_{ki} is the membership value of the input data object i to cluster k :

$$u_{ki} = \left[\sum_{j=1}^c \left(\frac{\|x_i - p_k\|_A}{\|x_i - p_j\|_A} \right)^{\frac{2}{m-1}} \right]^{-1} \quad (3.1)$$

The norm A of x is defined as $\|x\|_A = \sqrt{x^T A x}$. If we want to express the Eq.(3.1) in terms of the distance between the data object and the cluster centers, there can be several choices for evaluating this membership value. Let us limit our choices to the following expressions:

$$u_{ki} = \frac{\left(\frac{1}{D_{ik}} \right)^{\frac{2}{m-1}}}{\sum_{j=1}^c \left(\frac{1}{D_{ij}} \right)^{\frac{2}{m-1}}} \quad (3.2)$$

Figure 3.1: Absorbion version of updating equation for cluster centers.

$$u_{ki} = \frac{e^{-w_{ki}D_{ki}}}{\sum_{j=1}^c (e^{-w_{ji}D_{ji}})} \quad (3.3)$$

where w_{**} s are arbitrary weights to the bell-shaped membership functions applied. Both of the above equations were applied in literature of clustering for different purposes [Wu and Er, 2000, Pal et al., 1996]. Then, let us modify the update equation of the cluster centers. In Eq.(2.21), the cluster centers are updated based on the membership values assigned to all available data objects. However, in online applications, it might be impractical to store all the input data. Therefore, instead of using all input data with their membership values, we suggest using the another form of update equation for cluster centers that does not need whole history of input data.

Our version of update equation for cluster centers applies the idea that each cluster can be represented by its center. However, in the update equation we suppose a weight for each cluster. This weight represents the number of the data in the input space that belongs to that specific cluster. Therefore, each cluster that represents the majority of the data samples carries more weight. It makes the cluster center more reluctant to sudden change to any

new input data specially when the data belongs to the cluster with quite small membership value. The idea is depicted in Fig.3.1.

The cluster center updating equation which is mentioned in Eq.(2.21) is based on the fact that all data in input space are available at the time of calculation. For each cluster the weight is defined as the summation on all membership values assigned to input data w.r.t. that specific cluster. This weight is similar to the denominator of Eq.(2.21) in Bezdek alternating optimization:

$$Uc_k(i) = \beta Uc_k(i - 1) + u_{ki} \quad (3.4)$$

In Eq.(3.4), $Uc_k(i)$ is the assigned weight to the cluster k with regards to the input sample i . β is a “forgetting factor” which can be a real number close to 1. Then, this weight is utilized for cluster center updating. First we calculate the cluster centers based on Bezdek’s FCM algorithm. Then, for updating the centers based on the new incoming samples, we suppose that any new input sample will update the cluster centers so that it attracts all cluster centers towards itself. However, any cluster center will be updated w.r.t its weight and the assigned membership values to the new input so that the more weight the cluster has, the less movement it will have towards the new input data and also the higher membership value the input has to one cluster, the more it will attracts the cluster center towards itself. This supposition is not far from the concept of cluster center update equation in Bezdek [Bezdek, 1981, Pal et al., 1996], Eq.(2.21) and depicted in Fig.3.1. In Bezdek’s equation, each cluster center can be seen as the Center of Gravity (CoG) of all data samples where each data sample is supposed to have a

weight equal to its assigned membership value. Let us suppose that by entering a new data object we have only two samples to calculate their CoG. One is definitely the input data sample and the other is the CoG of the rest of data which is represented by the calculated weight:

$$p_k(i) = p_k(i-1) + u_{ki} \frac{x_i}{U_c(i)} \quad (3.5)$$

Up to here, the number of clusters is supposed to be fixed and predetermined. There are some methods available in clustering literature that deals with the number of clusters and their optimal value. In next session another method for dynamic cluster generation and pruning the unnecessary clusters will be presented.

3.2 Introducing a new simplified Dynamic Fuzzy Neural Network based on sequential Fuzzy C-Means Clustering

Dynamic Fuzzy Neural Network (DFNN) was introduced in [Wu and Er, 2000]. The idea is to develop a fuzzy neural network that can develop the input/output neurons of the structure as well as evaluating their importance and in necessary cases, pruning the the neurons with least importance in the structure to reduce the complexity of the structure.

A brief explanation on DFNN structure can be found in papers like [Wu and Er, 2000, Wu et al., 2001, Liu and Er, 2012, Er and Deng, 2004] where

different versions and modifications on the basic idea were presented by the authors for different applications.

For the experiments we have applied the clustering methods and verified that the cluster analysis can be applied in tool condition monitoring. Then we correlated the clustering results with the target values. Hence, it looks beneficial to apply a model structure that can merge these two components together.

Dynamic Fuzzy Neural Networks is a modeling platform that applies the concept of data clustering and fuzzy neural networks to form a modeling structure for a given set of input output data. Since we have shown that the FCM clustering is the best clustering approach for our input data space, here we have to modify the DFNN version introduced in [Wu and Er, 2000]. There, Radial Basis Function (RBF) neural network were applied for clustering the input space. This layer in our proposed structure is replaced with the FCM sequential clustering that is introduced in Section.3.1.

The overall architecture of our proposed dynamic fuzzy neural network is presented in Fig.3.2. Suppose r is the dimension of the input variables which are, in our case, the wavelet energy features, x_i ($1, 2, \dots, r$). Each input feature or variable has u membership functions A_{ij} ($1, 2, \dots, u$). This fuzzification level is depicted in layer 2 of Fig.3.2. The membership functions here can be in the forms mentioned in Eq.3.2 or Eq.3.3.

In layer 3, each node stands for a IF-part of fuzzy rules in DFNN structure. So, if we apply the fuzzification definition in Eq.3.3, for j th rule R_j we

Figure 3.2: The architecture of a dynamic fuzzy neural network [Wu et al., 2001]

have:

$$\phi_j = \exp \left[-\frac{\sum_{i=1}^r (x_i - c_{ij})^2}{\sigma_j^2} \right] = \exp \left[-\frac{\|X - C_j\|^2}{\sigma_j^2} \right] \quad (3.6)$$

Here, X is a vector defined by $X = (x_1, \dots, x_r)$ in r dimensional space and C_j s are cluster centers.

In layer 4, normalization of the assigned membership values to an input variable is conducted. The actions in layers 3 and 4 were already included in our definition of sequential FCM clustering. In Eq.3.3 the concept of normalization is already implemented. Therefore, the output of N_j node is

$$\psi_j = \frac{\phi_j}{\sum_{k=1}^u \phi_k} = \frac{\exp \left[-\frac{\|X - C_j\|^2}{\sigma_j^2} \right]}{\sum_{k=1}^u \exp \left[-\frac{\|X - C_k\|^2}{\sigma_k^2} \right]}, j = 1, 2, \dots, u. \quad (3.7)$$

The above equation is equal to the Eq.3.3 where $D_{ij} = \|X - C_j\|^2$ and $w_{ki} = \sigma_j^{-2}$. Other membership functions like what applied in Eq.3.2 is also applicable.

Up to here, all the necessary calculations for DFNN were for clustering

the input space and were included in our proposed FCM clustering.

Layer 5 is the layer that the rules have to be correlated to the desired output variables. The idea here is to calculate the weighted summation of the fuzzified signals from normalization layer. Following what we have in [Wu and Er, 2000], let us suppose y is the output variable which is calculated as follows:

$$y(X) = \sum_{k=1}^u \alpha_k \psi_k = \frac{\sum_{k=1}^u \alpha_k \exp \left[-\frac{\|X - C_k\|^2}{\sigma_k^2} \right]}{\sum_{k=1}^u \exp \left[-\frac{\|X - C_k\|^2}{\sigma_k^2} \right]} \quad (3.8)$$

3.2.1 Learning Algorithm

The learning algorithm for the DFNN is quite the same as what mentioned in [Wu and Er, 2000]. However, we had to modify some parts due to the nature of our experiments.

In [Wu and Er, 2000], they suppose that they have the set of input data and their desired values. Therefore, they build up all the neurons from bottom up according to the resulting error comparing the desired and estimated values. Here in our research experiments, we have the same sets of data and the learning algorithm for such cases will not be so different. However, in real world applications, the input data is available while the output data is not available. Therefore we have to modify the learning scheme to fit such scenarios.

Criteria for Neuron Generation

Let us now suppose that we always have the input/output pairs. So, our concern is the training and application of sequential FCM based DFNN on normal cases. Since there is no limit on the number of clusters, we start with arbitrary number of clusters for initialization. Then, the sequential learning scheme will modify this number based on the following criteria.

1. Clustering Error:

Since our focus in this proposed clustering scheme is its success in clustering the input space, we define a criterion on new cluster to be considered. Since the algorithm is initiated with an arbitrary number of clusters, we define the following criteria for cluster generation.

In observation i , the set of new observation (X_i, t_i) arrives. We calculate the distance $D_i(j)$ between X_i and the existing cluster center C_j . A new cluster is considered if:

$$D_{\min} = \operatorname{argmin}(D_i(j)) > k_d \quad (3.9)$$

Where k_d is calculated differently from the one introduced in [Wu and Er, 2000]. Here, the initial value of k_d is calculated as the maximum distance of the initial cluster centers. However, this k_d will adaptively be modified during the process.

2. Estimation Error:

This error is based on the performance of the sequential FCM clustering

based DFNN. The structure is to provide a proper estimation of the output t_i according to the input observation vector X_i .

If the estimation error is defined as follows:

$$\|e_i\| = \|t_i - y_i\| \quad (3.10)$$

and if e_i exceeds than a pre-specified constant k_e , a new cluster, generalization to the location of clusters, or weights are considered. This k_e is not constant during the learning process and its modifications will be discussed later. however, it shows the desired accuracy of the DFNN structure.

The Concept of Hierarchical Learning

As presented in [Lee and Kil, 1991] and applied in [Wu and Er, 2000], the concept of hierarchical learning can be applied in DFNNs to accommodate the clusters. Initially, the main clusters are to be considered which cover the most important area of the input space. This is called rough learning or global learning in [Wu and Er, 2000]. Then, the “refinement” action starts and the error criterion and clustering quality criterion are gradually modified for the DFNN to capture the minor important clusters of the input space and the overall structure learns the most *troublesome* positions of the model.

Therefore, following the idea which is described in [Wu and Er, 2000], we modify the k_e and k_d during the learning process as following:

$$\begin{aligned} k_e &= \max(e_{\max} \times \varepsilon^i, e_{\min}) \\ k_d &= \max(d_{\max} \times \gamma^i, d_{\min}) \end{aligned} \quad (3.11)$$

Above,

- i i -th iteration;
- $e_{\max}$ predefined maximum error;
- $e_{\min}$ minimum error which defines the desired accuracy;
- ε convergence constant; The
- $d_{\max}$ largest distance between the cluster centers;
- $d_{\min}$ smallest distance acceptable between the cluster centers;
- γ decay constant.

idea here is to narrow down the acceptable distance of neighboring clusters to let the new clusters to join in. Also, it opens the space for new neurons to be generated for increasing the overall accuracy of DFNN structure.

Allocation and Relocation of Clustering Units

In [Wu and Er, 2000], Radial Basis Function (RBF) based clustering of the input space is applied. There, the width of the RBF units were important for generalization of the network. However, in sequential FCM based clustering, it is a choice to use either the Gaussian functions or the inverted distance function, both were discussed in Eq.3.2 and Eq.3.3 respectively.

Here, our solution is for allocation and relocation of cluster centers is

based on the concept of sequential FCM clustering. Therefore, in case a new cluster is going to be considered, the incoming input vector is considered as the center for new cluster. Then, the location of the cluster is modified due to the next incoming vectors so that the center be placed almost at the Center of Gravity of the other cluster mates. Two clusters also can be merged after their cluster centers become closer than the defined d_{min} .

Therefor, for four cases that might happen the decision of the learning algorithm will be as follows:

- (a) Case 1: $\|e_i\| > k_e, d_{min} > k_d$, A new cluster have to be considered. Here the initial cluster center will be the incoming input data X_i .
- (b) Case 2: $\|e_i\| < k_e, d_{min} < k_d$, which shows the acceptable performance of the DFNN for accommodating the observation (X_i, t_i) . There is no need for updating the clusters nor the weights.
- (c) Case 3: $\|e_i\| < k_e, d_{min} > k_d$, Here, the DFNN clusters fail to capture the incoming data well and yet, the DFNN is able to acceptably predict the input pair (X_i, t_i) . Therefore, the cluster centers will be updated according to their distance to the input sample X_i .
- (d) Case 4: $\|e_i\| > k_e, d_{min} < k_d$. In this case, the input data fell successfully into one existing cluster. However, the weights of the DFNN structure have to be updated towards the decreasing of the estimation error.

As a result, not all cases in the above will force the DFNN to

create a new rule. Later we will discuss the conditions for merging and/or pruning the existing clusters.

Weight Adjustment

In our research, we apply the same weight adjustment scheme for the neurons which is applied in [Wu and Er, 2000]. However, we have to note again that in our application which is modeling the HSM process, we do not always have the complete pairs of (X_i, t_i) incoming data. So, for the sake of completeness, let us introduce the weight adjustment scheme and then we will focus on the limitations of our application.

We suppose that all pairs of data samples are kept in memory. Then suppose u number of clusters are generated for n data samples. Let us combine all assigned fuzzy membership vales to all input data in a matrix called ψ as follows:

$$\psi = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{u1} & \dots & a_{un} \end{pmatrix} \quad (3.12)$$

If the output of the DFNN be a linear combination of the input values, the following equations will show the calculation of the output according to an input like X_j :

$$y_j = \sum_{i=1}^u w_{2i} \times \psi_{ij} \quad (3.13)$$

Or in its matrix form and for all input/output sets:

$$W_2\Psi = Y \quad (3.14)$$

Here without loss of generality, we skip the TSK model which is introduced in [Wu and Er, 2000] and focus on the S model as above. The estimation error is defined as:

$$\tilde{E} = \|T - Y\| \quad (3.15)$$

The optimal solution W_2^* to minimize the objective function $\tilde{E}^T\tilde{E}$ which minimizes the estimation error of the DFNN is given by the linear least squares (LLS) method:

$$\begin{aligned} W_2^*\Psi &= T \\ W_2^* &= T(\Psi^T\Psi)^{-1}\Psi^T \end{aligned} \quad (3.16)$$

Pruning the Neurons, Merging and Deleting the Clusters

Among the benefits of dynamic fuzzy neural networks is pruning the unnecessary or less important neurons and data clusters. It is undesirable to create and keep a rule which has low contribution to the input/output relation. One method to identify the importance of a rule in a fuzzy neural network is introduced in [Wu and Er, 2000, Wu et al., 2001, Liu and Er, 2012, Chen et al., 1991] and called Error Reduction Ratio (ERR) method. Let us follow this pruning

technique too and mention it here for completeness purpose.

1) *Error Reduction Ratio Method*: Calculating the transpose of the Eq.3.15 and Eq.3.16 , we have:

$$\begin{aligned}
 D &= H\theta + E \\
 D &= T^T \in \mathbb{R}^n \\
 H &= \Psi^T = (h_1 \cdots h_v) \in \mathbb{R}^{n \times v} \\
 \theta &= W_2^T \in \mathbb{R}^v
 \end{aligned} \tag{3.17}$$

Here, v is set according to the model we choose for fuzzy model. for example, for TSK model, $v = u \times (r + 1)$. Considering the dimensions of matrix H , it can be decomposed into its orthogonal basis.

$$H = KA \tag{3.18}$$

where, K is an $n \times v$ orthogonal matrix and A is a $v \times v$ upper triangular matrix. The above two equations result in:

$$D = KAW_2 + E = KG + E \tag{3.19}$$

Here, the solution G is given as follows:

$$\begin{aligned}
 G &= (K^T K)^{-1} K^T D \\
 g_i &= \frac{k_i^T D}{k_i^T k_i^T}, 1 \leq i \leq v
 \end{aligned} \tag{3.20}$$

ERR factor is introduced here as follows [Chen et al., 1991, Liu

Figure 3.3: Flowchart for sequential fuzzy c-means clustering based dynamic fuzzy neural network

and Er, 2012]:

$$err_i = \frac{g_i^2 k_i^T k^T}{D^T D} = \frac{k_i^T D^T}{k_i^T k^T D^T D}, 1 \leq i \leq v \quad (3.21)$$

Here a matrix called ERR matrix is introduced as $ERR = (\rho_1, \rho_2, \dots, \rho_u) \in \Re^{(r+1) \times u}$. All elements of ERR matrix are calculated using Eq.3.21.

The j th column of the ERR matrix is related to the j th rule.

The quantity η_i will represent the importance of j th rule. The higher η_i , the more significant the rule:

$$\eta_i = \sqrt{\frac{\rho_i^T \rho_i}{r+1}} \quad (3.22)$$

If η_i of one rule fails to be greater than a prespecified threshold k_{err} , the rule is not eligible enough to contribute in input/output relation of the DFNN structure. All the procedure is briefed in the Fig.3.3 for quick reference.

3.3 Application of Sequential FCM based DFNN for Modeling a HSM process

In previous session we tried to modify the well known dynamic fuzzy neural network for application in modeling the HSM process. In our sequential learning we tried to decrease the number of necessary data to keep in memory. If we refer to the initialization process for sequential FCM, we need at least one set of complete

Figure 3.4: Fuzzy C Means based DFNN Results of Wavelets Energy Features, Mean Relative Error (MRE): 91.8115%

input/output data before we start sequential learning. However, after sequential FCM starts, we keep only the same number of data for keeping the data set up and running as well as modifying the cluster centers and merging the necessary clusters and pruning the insignificant ones. The initial values for cluster centers will be determined by this set of data. This might seem limiting the capabilities of the clustering algorithm. However, it will help us manage the input data during the real world applications of cutting process.

We also have to note that in real world milling process, the input data do not have desired output along and therefore, have to be managed only by placing the cluster centers properly among them. Here we do not limit the number of necessary clusters and the sequential FCM based DFNN decides on their number. However,

Figure 3.5: Fuzzy C Means based DFNN Results of Wavelets Energy Features, Mean Relative Error (MRE): 97.4951%

Figure 3.6: Fuzzy C Means based DFNN Results of Wavelets Energy Features, Mean Relative Error (MRE): 95.7210%

since we apply traditional FCM for initialization of the cluster centers, we start with a priori defined number of clusters. Then, during the process, the algorithm decides on the number of necessary clusters. Since we apply the FCM clustering, the centers of the clusters are found and updated in the second layer by sequential FCM scheme and DFNN determination of significant rules.

Therefore, for online application of sequential FCM based DFNN for high speed milling process, where t_i s in the input pairs of data (X_i, t_i) are not accessible, only cluster centers can be updated. In necessary cases, only merging the clusters and creating the new ones is available.

However, for training purposes, all pairs of data are complete and the sequential FCM based DFNN can be trained and tested. Let us now compare the deliverables of this chapter with those algorithms with dynamic structures. One serious challenge in this study arises when it comes to industrial environment applicability of a method. Normally in industrial environments, a milling machine time is a very precious time. Therefore, stopping the machine to check the quality of surface only for training a dynamic model is not practical. Therefore, a model needs to either work on the saved data and then fix its structure or demand some stops during the process for new training samples.

The idea of sequential fuzzy c-means based DFNN is to incorporate input samples as well as output samples to organize the

model structure. Therefore, in addition to modifications based on the performance of the network on outputs, this learning paradigm focuses on the quality of the clusters formed in the input space. This is one of the key issues that can be mentioned among the positive points of this method.

For example, let us compare the structure and performance of some newly proposed dynamic models.

A Bayesian adaptive resonance theory-based fuzzy inference system is presented in [Oentaryo et al., 2011]. The structure and performance of the presented system is interesting and there are some plausible points that make this method considerable. As the authors claim:

- All training observations are sequentially (i.e., one-by-one) presented to the system
- At any time, only one training observation is seen/learned
- A training observation is discarded once the learning procedure for that particular observation is completed
- The learning system has no prior knowledge as to how many total training observations will be presented

These specifications are really necessary for an online monitoring system. However, to form an online prognosis and diagnosis scheme for milling process we have to note another characteristic of the process which is unavailability of the output in vast amount of milling time. As we previously pointed out, in milling

process we face an exceptional case where in only few moments of milling process the cutter and resulting surface can be examined. The proposed method in our work is more suitable in such cases because the modification of rules and weights and merging the clusters is done even when the output results are not available. In fact, the input and output space are considered independently.

The accuracy of the results in [Oentaryo et al., 2011] is mentioned in terms of non-dimensional error index. This index is 0.27 for a destructive test which is roughly acceptable. Comparing to our results we may see some fluctuations in the beginning of the predictions which is the result of training and fine tuning the rules while the performance of the model and the accuracy of the results are considerable when the process goes on, Fig.3.4, Fig.3.5 and Fig.3.6.

Fuzzy regression modeling were also presented in [Li et al., 2010] where a dynamic structure is proposed based on fuzzy rule-based generation and multiple regression modeling on the results. The idea is roughly similar to DFNN in terms of adjusting the center of clusters in rule layer of the model. The distinction however lies in the correlation of the results of rule layer to the output. In this method, the multiple regression is applied to adjust the parameters of a correlation model that relates the outputs of rule layer with the outputs and therefore it can be considered as an TSK model. In this paper, some fixed structure models are compared with the proposed fuzzy regression models and it shows that the

fuzzy regression model is superior over back propagation neural network or radial basis function network or even multiple regression models. Here they validate the model which is formed on one cutter on a portion of the same cutter's data which delivers a plausible accuracy result of 99%. However, the validation on another cutter is absent which makes us unable to compare their results with ours in more practical condition. Referring to the results we achieved with CFFNN in section 2.3.4, validation on the data of the same cutter usually delivers a more accurate model than validation on another cutter. For that case we reached the accuracy of 98% for test over the same cutter. However, the accuracy drops drastically when the validation cutter is not the same. We may stress on the performance of our proposed model which is over 91% in validation on a "different" cutter than the cutter used for training.

Chapter 4

Conclusions

4.1 Conclusions and Contributions

In this PhD study, benefits of application of clustering methods in analysis of milling-process resulting signals was investigated as a proper way for online monitoring of the process. To apply clustering method, first the features of force, vibration, and acoustic emission signal were extracted using wavelet analysis. Peak localization and energy computation of the flute cutting signals provide the proper features to be clustered by available clustering schemes. All sensor signals were checked to have consistent results. Since clustering results are invariant in all of the cutters of the same geometry, we concluded that they can be considered as an indication of tool's condition and surface degradation as these two factors are highly correlated. The validity of the idea was verified by its application on different cutters of similar geometry and

different coatings and tool lives.

As a noticeable achievement we have shown that vibration signal features are applicable to replace force signal features in milling-process condition detection. Then, a proposed method for artificial intelligence based modeling of ball-nose milling process is also presented. For this AI based modeling we had to look for a proper clustering scheme that provided the best results and was suitable for generalization for sequential clustering. Therefore, further studies on the clustering block of a proposed method for AI-based modeling of ball-nose milling process was presented after the preliminary studies in [Torabi Amin J., 2010b].

The benefits of utilization of clustering methods in milling tool condition detection system were also presented. In this part of our study, several crisp and fuzzy clustering methods were applied on sensor signal features. They are shown to provide similar analysis on the signal features which enabled us to develop a fault diagnosis scheme for milling process. Force and vibration signal features are checked to have consistent and repeatable results. Since different clustering results were almost invariant in detection of the status of all of the cutters of the same geometry, these methods can be applied in our proposed structure for a possible generalizable tool condition detection system.

Following the verification of benefits of clustering algorithms [Torabi Amin J., 2012], we have considered the most proper type of these

methods for the purpose of sequential learning. FCM was shown to be the most proper type among fuzzy clustering schemes as it is easily modifiable for sequential learning and delivering its interpretation on process status to next level of AI model.

This Report includes our investigations among several commonly used methods for modeling the surface finishing quality of a high speed machining process and discusses the advantages and disadvantages of the works which have been done in this area. Namely, these methods can be divided into two main categories, non-AI-based and AI-based approaches. The non-AI researches are typically used for presenting the behavior of the process while AI-based methods are mostly used for modeling, monitoring, fault diagnosis, and predictive control. Among on these two categories, we investigated the state-of-the-art methods which are commonly used for modeling, fault diagnosis, and control while considering on-line monitoring issues. Since the nature of the process is multi-variable and nonlinear, some of these modeling structures are found to be able to cope with such systems. Bayesian networks, fuzzy petri-nets, hidden Markov models and dynamic fuzzy-neural networks seem to be the most suitable methods which are already under investigation. On the other hand, yet there are many research gaps that need to be addressed. For example, there are very few research reports on the tool-production methods and tool attributes and their correlation to the resulting signals and surface-roughness and tool-degradation. Also, there is an obvious

research gap on a single general model for cutting process while the models are not expandable to other cutters with similar attributes. Until there is a general descriptive model available for cutting process, the field keeps being updated by application of new AI techniques and different features of the resulting signals. In the present report, many of the available modeling methods have been discussed. The pros and cons of them are shown and many research gaps in this field are identified and discussed. As a result, it facilitates the way to select an appropriate modeling technique for many milling processes and to control the surface quality based on the available models of the process. Also, it covers some present weaknesses from research point of view in the machining technology.

After brief survey on several AI and non-AI techniques which are used in this field, there are some points to be noted. Obviously, one of the matters is that the predictions resulted from any proposed approach must be accurate and repeatable enough. It is practically verified that AI-based methods are more accurate than the classic modeling approaches. It is also clear that each of the state-of-the-art modeling, inferencing and decision making methods in one way or another are able to predict the surface-roughness and tool-wear in a non-intrusive manner. As such, advances in these approaches result into more informative models in this field too. However, from the industry point of view, any suggested approach must be easy to be implemented and more accurate. The

learning speed, on-line applicability and flexibility of the model structure to deal with the changes of the system are the challenges to any new approach.

In order to provide an informative reference model, there are several kinds of concerned problems [Torabi Amin J., 2013b]. After brief survey on several AI and non-AI techniques which are used in this field, there was extensive studies on the milling process and many experiments was conducted. As the result, there was many models proposed and tested to develop a platform of comparison and to look for most suitable method(s). To continue the research, we have to go further to see in what ways we can modify or improve these models. Obviously, one of the matters is that the predictions resulted from these approaches must be accurate and repeatable enough. It is experimentally illustrated that AI-based methods are more accurate and more flexible than the non-AI ones. It is also clear that any of the state-of-the-art modeling, inferencing and decision making methods are able to predict the surface-roughness and tool-wear in a non-intrusive manner.

The AI-based methods that were presented in literature survey chapter are most commonly used and state-of-the-art methods in the field of machining. Some of their issues and weaknesses along with our solution to them are mentioned and discussed in this section. One of the benefiting characteristics that a complete research in this field would have to include is to carry an insightful comparison between available methods. For example, Aliustaoglu

and Ertunc showed and compared some methods which could be used in tool status monitoring [Aliustaoglu et al., 2009, Ertunc and Oysu, 2004]. There are not many such papers in the field that compare different methods together. Majority of the papers are concerning only one method and its capabilities to deal with the process. In the above mentioned paper for example, HMM, fuzzy inferencing and sensor fusion are well applied to detect the tool condition and resulting surface finishing. As such, fuzzy-nets are used to determine the state of the system and their applicability were discussed and compared in [Kirby et al., 2006]. Referring to respective sections of this study, it is obvious that although many AI techniques have been applied for tool wear detection or surface-roughness modeling, yet there are many other methods and approaches to be vastly investigated.

- Not all of these AI techniques are studied for their most proper structure.
- Many of the proposed methods have to be tuned for some parameters and they vary from one experiment design to another.
- There are not dynamic and intelligent methods which can be applied with the ability to form the model online, i.e to be able to model the process dynamically.

Some of the methods like Bayesian networks [Correa et al., 2008, Dey and Stori, 2005] and Petri nets [Kasirolvalad et al.,

2006] are applied for tool status and surface finishing prediction while the justification of the event-based models needs to be more studied and many advanced and intelligent event-based models like what is used in [Doostfatemeh and Kremer, 2005] have not been investigated in this field.

- Yet the real question remains unanswered about the benefits of the discrete event systems for modeling all the phenomena of the process.

All other methods are claimed to be successful to determine the internal state of the process of their concern. However, there is no report for any global model for a specific scenario of the cutting process.

An overview on the available techniques and their application in milling process modeling can be found in Table 1.9. It summarizes the works mentioned in the literature and presents their advantages and disadvantages. Advances in artificial intelligence, dynamic structure modeling techniques and clustering methods as well as data pre-processing schemes should be considered to affect the future of the investigations and provide better solutions for industry, such as better quality and more productivity.

To summarize all the works that have been done so far and to have a look to the future work, here we conclude by discussing the most important issues and our contributions:

High Speed Machining is a critical process which is conducted on

many aerospace workpieces. Since material and tool are expensive, there can not be any compromise on either side. Therefore, effective life of cutting tool and the status of the surface must be determined so that the tool can be replaced by a new one at the right time or the cutting process is modified accordingly. To predict the moment for termination of the process due to an occurred an abnormality (fault diagnosis), many experiments have been conducted on the performance of milling tools. The results are presented in many forms of modeling techniques. Some researches are visually presented these experimental results while others demonstrated them in closed form of mathematical models or artificial intelligence AI based structures [Torabi Amin J., 2009].

Most basic requirement of these models is its quality of being accurate and repeatable. Early proposed models were accurate only when the tool was new. In fact, these models were solely considering cutting conditions to be affective on surface roughness results. In contrast, there are some models that consider time and/or other factors to make models more accurate during destructive process. Instead of time as one of the inputs to the model, some of them used the side signals of cutting process such as force, vibration, and acoustic emission signals as input variables of reference model. Since these signals are continuous time, many researches apply feature extraction to reduce data dimensions and computation costs. These features are chosen to be informative

enough to be used in tool-wear model. There are many ways to extract the features from raw signals in literature. Some investigations apply time-domain features while others are using frequency features. Combining these features with other cutting conditions, the model will explain the changes in resulting surface quality and tool condition during the cutting process.

However, replacing the tools and some other cutting parameters of milling process affect the resulting signals which demands for new model that address the changes in extracted features and nature of the process. Therefore, there must be an interpretation layer between resulting signals and the artificial intelligence model so that these changes never affect the entire model. It is clear that when the tool is replaced, there must be always some modifications needed in the model. It will be perfect improvement in the existing proposed models if we can suggest one with only one subsystem changed due to tool replacement. So, in our proposed model, the only necessary modifications are the amendments in this interpretation layer. This change of the layer will protect the change in the weights of the artificial intelligence layer which is placed in next level. Besides, the model must overcome some other challenges such as frequency drift, flute-cut signal separability, etc which is briefed in Section 2.2. After all, there must be a model which can be used to modify the cutting conditions in case the tool is going to be used to extend its life.

In this PhD dissertation, an extensive literature survey has been

conducted on the state-of-the-art techniques. Such survey facilitates selection of proper method among the available methods for modeling the tool life or resulting surface roughness. It also illustrates some research directions which shall be followed by the further studies. Following that literature survey, many experiments have been conducted on different material and different tool geometries and coatings. All the cutting conditions and resulting signals along with tool and material specifications are available for model training and verification. As the result, many of the available methods were tried to check their suitability for the signal-input tool status-output model. There are also many other methods that have been checked as a part of the scope of this study [Torabi Amin J., 2009, Torabi Amin J., 2013b]. Among these methods neural networks, Adaptive Neuro-Fuzzy Inference Systems (ANFIS), Hidden Markov Models (HMM), clustering methods and Dynamic Fuzzy Neural Networks can be named.

We focused more on several types of clustering methods because of their positive benefits in the analysis of milling-process resulting signals is investigated. To make clustering method applicable, first the features of force and vibration are extracted using wavelet analysis. Peak localization and energy computation of the flute cutting signals provide the proper features to be clustered by fuzzy C-means clustering. Both signals are checked to have consistent results. Since clustering results are invariant in all of the cutters of the same geometry, it can be considered as an in-

dication of tool's condition and surface degradation. The validity of the idea is verified by its application on three different cutters. Also, vibration features show their applicability to replace force in milling-process condition detection [Torabi Amin J., 2011, Torabi Amin J., 2010b, Torabi Amin J., 2010a].

As the second subject of interest, Artificial Neural Networks (ANN) and their advanced versions which are capable of including fuzzy inferencing and structure varying will be concerned. In many related researches on HSM centers, artificial neural networks have been used due to their capability of modeling the nonlinear processes. For studies on Neural Network, we investigated many different structures for neural networks. Both, best cutting condition determination experiments and destructive test seem to be well described with these methods. There are plenty of different implementation of ANN concerning the prediction and control matters. There are some reports that claim the repeatability of these models. Not all types of these models are investigated for a universal model for cutting process. Also, there is no consistency in the parameters which have been used as the inputs. Therefore, there is still a visible gap on the completion of these studies with such models where there are no general meaningful features and model structures. In addition, there are very few on-line prediction and control researches available in the field which leaves space for more investigations on universal models. So, this part of the research is the preliminary work on the investigations for a repeatable refer-

ence models for high speed ball-nose milling processes. A cascaded Feed Forward model is tried to fit the experimental data and is checked for its repeatability for another cutter. Future research advances on a more general models and the models that can be used on-line are under investigation [Torabi Amin J., 2010a].

After all, there were also some obvious research gaps in this field which we aimed to address them in this present dissertation:

- i. One challenging area in industrial processes is to take better and more descriptive features out of the collected sensor signals using more suitable signal processing schemes and feature selection methods.
- ii. Lack of proper investigation of the data pre-processing methods is also obvious in the field of machining. For example, wavelet analysis and other state of the art pattern decomposition and extraction methods have only just recently been utilized for milling process signals, but they seem to be appropriate approaches for the extraction of the different properties of the signal.

In the present study, we addressed this gap. We have considered and applied different types of signal processing schemes, including time-domain features of the signal, frequency features of the signal, and finally its continuous wavelet analysis features. Among many possible signal processing methods, like time-domain analysis, frequency-domain analysis,

and time-frequency analysis, we have focused on time-frequency analysis as it focuses on the two main aspects of the signal. It is then made possible to investigate the variations of the signal in frequency domain with regards to time.

First we chose the type of mother wavelet function as it is always essential to apply the function which is in agreement with the type of signal that is under analysis. We have applied several mother wavelet functions and selected the one which fits our needs in signal processing. Section 2.2 explains the process of mother wavelet function selection in detail.

- iii. Unavoidable frequency drift of the signals and changes in their shape during their lifetime due to imperfections in mechanical parameters have not been addressed before the present study. These frequency drifts are naturally different from those can be found in the frequency analysis of the aging tool sensor signals.

Using wavelet analysis, in our project we have provided the solution for this problem for the first time in literature. The problem of frequency drift is due to imperfectness in electromechanical design of the HSM center and sometimes due to the effect of load on the milling machine. Despite the fact that we set the machining parameters and expect the timing for the hits of cutting edge to the cut surface to be regular, we observe that the flute hit peaks of the signal are slightly happen before or after the ideal expected time. Therefore, one

cannot go through the signal blindly and cut it into flute signals only based on milling parameters. This is another reason that we apply the time-frequency analysis.

Our solution to the peak localization after frequency drift in cutting signal is again based on the wavelet analysis of the signal. Since we apply continuous wavelet analysis features, this analysis is utilized for the purpose of detecting the flute signals in the condition of frequency drift too.

After applying continuous wavelet analysis, it becomes possible to go through different wavelet scales in various times. It is facilitated to look into different shapes of signal during the process. One of its applications that we have used in this study was to use a proper scale for flute cut separation. Using flute cut signals is the novel aspect of our work. This technique decomposes the cutting signal which is collected during the process into a set of signals which are resulted from one hit of the cutting edge of the milling tool to the workpiece. Each hit generates a peak in our cutting signal which is recorded by the attached sensors to the environment. These peaks together with some adding high frequency noise and non-cutting signal form the whole recorded signal.

To separate the useful parts of the signal, it is not possible to simply put a threshold on the signal and cut out the desired signal. The threshold is simply violated by peaks as well as some cutting signal abnormalities. Therefore, to extract the

flute-cut of the signal, we go through different scales of its wavelet analysis, we try an adaptive threshold on the signal in them scales and then try to find the expected flute-cut signals in an specified period of cutting time. This process is completely adaptive and there is no pre-assumption in the parameters applied.

As we are analyzing the signals for the purpose of tool condition monitoring, flute-signals are grouped based on their shape. Each shape can be correlated to its respective phenomena on the cut surface. In our work, we derived out the energy feature of each flute signal and then we showed that each of these features can successfully reflect the progress of the tool degradation and the condition of the surface roughness. However, using the energy feature can be a platform of using other features of the CWT of signal in the feature works.

- iv. Change in machine dynamics during long term running, which can lead to undesirable inaccuracy of reference model, is another issue to be focused on in monitoring systems.

In the solution based on cluster analysis, this problem is also addressed. Since no two machines are identical, no two tools are identical and no two surfaces are identical when we are considering their microscopic specifications, we have to think of a repeatable model. This model needs to be trained once for each working condition and then perform well on the rest

of the processes of the same condition. Therefore, grouping, classification, and clustering methods are needed to provide a repeatable result out of the collected cutting signal. Since generally there is no predetermined label for the signal features before the experiment, clustering algorithms are to be used rather than classification. Here we mean that a random collected cutting signal cannot be associated to a certain condition of the process at the time of collection. Therefore there is no output label for classification algorithm to group the signals based on that. Hence, we had to unsupervisedly cluster the signals first, and then decide which cluster goes to which level of tool condition. After first training, verifications were done to show the repeatability of the results of the cluster analysis. Therefore, as the first level of signal analysis and tool condition monitoring, we suggested cluster analysis.

- v. There are not many reports on the interpolation of the results from one type of cutter to another. Therefore, no matter how the cutters differ in their diameter or edge-preparation methods, for every new cutter the modeling must be repeated, which is expensive and time-consuming.

After number of destructive testings on milling tools, it seems it is unavoidable to do at least one destructive test for one type of cutter to train the reference model. However, it is crucial to know that decreasing the number of prerequisite experiments for modeling will decrease the total cost and production time.

So, based on available research in the literature and considering the available information through the published works, handbooks, and the collectable data from one destructive test, we suggested an structure which is able to incorporate other available results and interpolate the results of one cutter to another within an acceptable range of interpolation error.

This structure which is novel in literature is based on separation of the signal features layer from artificial intelligence layer using one interpreting layer. This interpretation layer applies cluster analysis to derive the information from the cutting signal features and providing it to the next layer. The advantage of such structure is modifiability of the final model due to the standardization of the signals and data which are handed to the next modeling level.

Another advantage of such model is the ability to integrate the results of destructive tests and the results of Best Cutting Condition Determination (BCCD) tests. The novelty of the proposed structure is that it separates the signal level from the AI level. Hence, the inputs to the AI level are the results of cluster analysis rather than the signal features and hence, it consists of normalized meaningful cluster membership values. Each signal feature when measured is examined for its membership to the available previously formed clusters. The output of such process is a preliminary evaluation of process status. Therefore, there is no need to change the

AI level from one tool to another provided that the cluster analysis is repeatable enough and can determine the status of the tool. In our study we have experimentally proved that for cutters of the same geometry the results of clustering level are completely repeatable [Torabi Amin J., 2011].

- vi. In the literature, there are very few papers that pay attention to the changes in the shape of the signal due to tool degradation, aging, or tool wear. These methods do not quantitatively investigate such changes. Mostly, they are limited to the use of frequency and time domain features and not the cross-correlation of the shape of the signals with corresponding tool edge phenomena.

Our solution which is based on flute-cut separation and clustering the flute-cut signals based on their shapes properly addressed the above mentioned issue. The wavelet analysis of the signal is mainly stress on the shape of the signal rather than its time features. Since wavelet analysis is used for peak localization, its features can be used for modeling too. These features are calculated per each flute-cut and can represent the shape of the signal.

Besides, we have provided the platform for process analysis, not limited to the milling process, based on any AI method that can deal with the shape of the sensor signals.

- vii. The effect of some production parameters of the cutters, such as edge-preparation methods, grinding quality, initial surface

roughness on the cutting edge, the geometrical cutting-edge design angles, and various coatings have not yet been investigated.

The scope for our study which was mostly concerning online monitoring and also limited budget and HSM machine time limited us not to consider all of the above parameters. For example, this study would not cover the design and study of the aspects related to cutter geometry, like geometrical cutting edge angles and their geometrical design. Therefore our cutters were mostly of the same brand and geometry. We also did not have access to information on grinding quality and edge-preparation of some other cutters which were purchased from the market. Therefore our model will have some deficiencies on this aspects and many more experiments and prior information are needed to make the model more perfect. Further studies in this field have to ensure access to some changes in these parameters. For example, the process can be studied aligned with some other studies on milling tool design, fabrication, and production.

Besides all above, we have noticed that the cutters of some certain coatings have longer tool life. Also we have noticed that geometry of the tool is an important factor in its performance.

Having considered the above-listed issues, the results of our experiments are addressing some of the important aspects of

the milling process. The results are discussed in Section 2.3.

- viii. There is apparent lack of investigation of the use of clustering, classification, and grouping methods in this field. One possible reason is the direct use of cutting conditions and signal features instead of clustering data in AI-based models.

Applying the signal features to the AI level of the model simplifies the modeling approach. However it will make the model limited to be used with one specific cutter model and limits the repeatability and generalizability.

Therefore our solutions mostly focus on grouping the signals and more specifically on clustering methods. We know that the cutting signals, the flute-cut shapes, and their features practically are not "labeled" inputs and therefore we have to group them based on their similarity rather than their label. If we can find a trend among the signal features that agrees the trends we physically have on tool degradation, then we can group the signals based on the same features in order to develop a model for monitoring the process.

Based on our studies on available grouping and clustering methods, we have found that Fuzzy C-Means clustering and its developed version for sequential learning are suitable choices for online monitoring. These are unsupervised learning methods and they adapt themselves with slight changes from one tool to another.

After training the clustering method, some evaluations were

also conducted to check the repeatability of the clustering approach. As it has been discussed in Section 2.4, many of the available clustering algorithms were tried for their performance on the extracted features of the flute-cut signals of the milling process. Many of them practically proved their repeatability. However the choice which is more computationally simple and yet is easily modifiable for sequential training is FCM algorithm. It was later developed in our work for better speed and sequential update and training.

- ix. There are many AI techniques that have not been used in the field of modeling of machining processes. As an example, a knowledge-based pattern recognition method like syntactic classification and modeling can be considered. To model a sophisticated pattern, it provides more simple patterns, called *primitives*, which are composed to make that complex one. Therefore a hierarchical model is presented for any similar pattern or simply any major pattern is decomposed to appropriate primitives as its building blocks [Jain et al., 2000]. It has been used in some articles to find the pre-specified shapes in the signals [Trahanias and Skordalakis, 1990]. This method was extensively used in speech processing [Lewis, 1992] and can be used for fault diagnostic and automatic failure sign distillation for tool condition detection.

In addition, the application of many variable structure AI techniques, namely like Dynamic Fuzzy Neural Network (DFNN)

[Wu and Er, 2000, Kim et al., 2005, Song et al., 2011], have not been studied in this field yet. They might replace the fixed structure of many of state-of-the-art structures and facilitate the generalization of those methods. The fixed structures of the available methods prevents them from being easily generalized and applied for on-line monitoring purposes.

We have used some artificial intelligence techniques, a novel proposed model structures, and some grouping methods which have not been applied in the field before, namely, FCM clustering, Gath-Geva, Fuzzy C-Means Sequential clustering. We also opened the platform for variable structure models. These models are proper choice for AI module with the inputs from sequential clustering level. As we know, in some possible cases the sequential clustering method might result into some new clusters that represent new states of the system.

A variable structure AI technique will serve properly and will find a better model for the process comparing to a fixed structure AI model [Wu and Er, 2000].

Another example could be the extreme learning method [Huang et al., 2006]. This method has proved to be applicable in on-line sequential learning. Similarly, other similar state of the art techniques that have not yet been used in the field can be applied.

- x. Some papers provide a solution for one specific experimental design in such a way that the results cannot easily be gen-

eralized to other design issues and conditions. As a result, many experiments are needed for modeling a new experimental design. The quality of the models to remain descriptive and useful in different scenarios is a critical issue.

In our approach, we did not only correlated the input features to the outputs of the process. This task has been done by others with various methods and the results were presented in literature. The superiority of a model is to provide a way for modeling a process that can be generalized easily.

In our proposed structure, we have applied one interpretation layer to the signal features which enables us to detach the AI layer of the model from the inputs. This interpretation model is shown to be updatable by one dummy experiment before the actual industrial cutting project starts. This simple test is to train a clustering layer which is simply able to determine the tool status. From that experiment onwards, the clustering layer keeps updated by each incoming feature and the slight changes in the tools will be compensated by this action of the reference model. Also, the AI section of the model is able to incorporate the outcome of the clustering layer and combines it with some other information such as cutting parameters, cutter and HSM machine specs, workpiece specs, etc.

Such model is able to work and be updated online and off-line, provide the information for Best Cutting Condition Determination experiments (BCCD) as well as destructive tests and

applications, and it is properly proposed for online supervisory monitoring, fault diagnosis and control.

Another benefit of such model is its modular structure. Each module or section can be determined independently and the experimental results can here be incorporated along with theoretical expected results, simulations results, and even engineering handbook results. For example, one module (clustering module) interprets the status of the tool considering the input signal features. The results of this module are standard primary evaluation of the tool status and is delivered to the AI module. The AI module gets the other parameter values of the experiment and evaluates the status of the surface or tool. In this part all of the above mentioned information can be incorporated.

- xi. The overall structures for generalizable monitoring and prediction system using the available modeling methods has not been considered in the field. This would seem to be a big gap needing to be covered in future studies.

To cover this gap in literature, the entire structure of the monitoring system has to be investigated for the best techniques to be applied in each part and their inter-connectivity and reasonable places in the structure.

In our proposed structure, we suggest the application of a clustering layer after signal collection and feature extraction part. The reasons for such suggestion is mentioned in Section

2.4 and summarized in previous items above. After application of a clustering layer, we develop an AI layer that have clustering results and experimental details as its input and surface roughness and tool wear as its output.

It is worth mentioning that the clustering layer by itself is sufficient for fault diagnosis purposes. It is completely able to capture the course behavior of the process and detect the status of the tool and surface. It is also shown that the clustering layer is able to be developed *online*. Depending on the properties of the process, the online monitoring and fault diagnosis system can be developed using the sequential experimental results. This process will effectively reduces the memory space required by the system and simplifies the model for fault diagnosis.

In addition to above, to model the development of tool wear, we need an AI layer to combine other information on the cutting process such as geometrical information of the tool, chemical properties of the tool and workpiece, and so on, with the results of clustering layer.

One approach is to apply neural networks to make the model. This basic approach, along with the proposed novel structure form a model that can be applied for tool wear development monitoring.

Another approach is based on the nature of the structure that we proposed and forms the future studies of this PhD study.

As we are applying an AI layer on the clustered results of extracted wavelet based features of the cutting signals, a generalization or development of such model is discussed in Section [2.4](#).

Then, an extension to the idea of Dynamic Fuzzy Neural Network (EDFNN) is presented that can combine both clustering and AI layer and present a unified model for the process. The benefits of such model is that it automatically adjusts the AI layer weights and properties, modify the fuzzy rules, prune the unnecessary ones, simultaneously with the proposed clustering layer organizing itself on the input data space. The details of this model is presented in Section [3](#).

Chapter 5

Future Work Plans

Future works will open the field to supplementary results based on the proposed structure on cutters of different geometry and various cutting conditions as well as investigating for correlation between surface results and resulting signal features.

The future of this research will mainly focus on the following factors:

- **Finding similar applications where the proposed fault diagnosis/prognosis and condition monitoring system is applicable.** There are many more applications that in general have the same characteristics as high speed milling process has. These processes might also need upgrade of their condition monitoring systems based on the theories and techniques developed in this dissertation.

Many of such processes are available in industry, namely oil industry deep sea drilling, power plants, gas turbines, grinding

process, micro-milling processes, etc. There are some observable signals in these processes which reflects their internal state. Applying the signal processing techniques similar to the ones applied in this research, the proper features for artificial intelligent bases monitoring system would be available. Then, based on the desired outputs of the monitoring system, these features would be correlated to those outputs. As a result, like the system developed in the present research, a monitoring structure would be developed for those processes as well.

- **Completing the proposed model for other milling tools using the results of future experiments.** The proposed method in this research will be more generalized using more experiments using several cutting conditions, machine parameters and even other milling centers. This requires many more experiments which requires more budget and time on its generalization.

Proposing other similar monitoring projects to the industry, will definitely open the doors for accessing to other experimental data that helps the generalization of the model.

The positive point is that the proposed model in this research is able to combine the results of two types of experiments, best cutting condition determination experiments and destructive tests.

- **Improvement of the proposed fuzzy c-means cluster-**

ing based dynamic fuzzy neural network. Like its predecessor, dynamic fuzzy neural network, that has been developed for many of its aspects like speed and accuracy, the fuzzy c-means clustering based dynamic fuzzy neural network is a newly developed AI technique and needs many more researches on its completeness, accuracy, simplification, application, etc.

The present research illustrated the applicability and benefit of this artificial intelligent modeling technique. It has to be applied in other modeling applications as well to check its deficiencies and benefits.

- **Developing more dynamic fuzzy neural networks for other purposes** Dynamic Fuzzy neural network is a dynamical modeling structure based on designing rule base and adaptation of clusters in the input space and pruning the unnecessary rules. This paradigm might be evolved using other search algorithms in the input space such as particle swarm optimization. The author is now working on this modeling technique and hopes that it seriously change the complex structure of the DFNN to a very simpler yet more natural one. It can be applied on modeling some behaviors of people in social networks and even in the society.
- **Looking for more simpler sensors for milling process monitoring** One of our main achievements in this study was introduced among the deliverables in clustering chapter. We

showed that application of vibration signal rather than force signal makes the fault diagnosis and prognosis process much more simple. The reason is the easy mounting of the vibration sensor on the workpiece. However, some sensors have still the potential to be investigated more as they are cheaper and easier to work with as they are not necessarily have to be in touch with the workpiece. Acoustic emission, current sensor of the spindle and the machine are among these sensors. The author believes that with proper mounting and denoising the signals of these sensors, they may also reflect the state of the process and hence are applicable for fault diagnosis and monitoring of many industrial processes and milling process in special.

Chapter 6

Published Papers and Honors

There are several journal and conference papers which are published from the present works. Also, the author with the group of his co-workers achieved a precious award. All of them will be mentioned in this chapter.

6.1 Published papers

Conference papers as main author:

- i. [Torabi Amin J., 2009] A. J. Torabi, E. M. Joo, L. Xiang, L. B. Siong, Z. Lianyin, P. S. Jie, Z. Junhong, S. Lin, H. Sheng, and J. T. T. Tijo, A survey on artificial intelligence technologies in modeling of High Speed end-milling processes, in IEEE/ASME International Conference on Advanced Intelligent Mechatronics,

2009. AIM 2009, 2009, pp. 320 325.
- ii. [Torabi Amin J., 2010b] A. J. Torabi, M. J. Er, X. Li, B. S. Lim, Z. L.Y, S. H, L. S, and G. O.P, Fuzzy Clustering of Wavelet Features for Tool Condition Monitoring in High Speed Milling Process, in Proceedings of the Annual Conference of the Prognostics and Health Management Society 2010, Portland, Oregon, USA, 2010.
 - iii. [Torabi Amin J., 2010c] Torabi A.J., et al, "Hidden Markov Model for Ball-Nose Tool Condition Monitoring" in proceedings of the Postgraduate Student Conference AOTULE 2010, Indonesia, Bandung, Nov 2010
 - iv. [Torabi Amin J., 2010a] A. J. Torabi, O. Massol, E. M. Joo, L. Xiang, L. B. Siong, G. O. Peen, H. Sheng, S. R. Gopikrishnan, Z. Lianyin, and S. Linn, Flute based analysis of ball-nose milling signals using continuous wavelet analysis features, in 2010 11th International Conference on Control Automation Robotics Vision (ICARCV), 2010, pp. 1359 1364.
 - v. [Torabi Amin J., 2012] A. J. Torabi, E. M. Joo, L. Xiang, L. B. Siong, Z. Lianyin, S. Linn, G. O. Peen, and C. C. Teck, Application of classical clustering methods for online tool condition monitoring in high speed milling processes, in 2012 7th IEEE Conference on Industrial Electronics and Applications (ICIEA), 2012, pp. 1249 1254.

Journal papers as main author:

- i. [Torabi Amin J., 2013b] Torabi A.J., et al A Survey on Artificial Intelligence Based Modeling Techniques for High Speed Milling

Processes, IEEE Systems Journal, 2013, [available online] [[Torabi Amin J., 2013b](#)].

- ii. Torabi A.J., et al, Application of Clustering Methods for Online Tool Condition Monitoring and Fault Diagnosis in High Speed Milling Processes, under revision for publication IEEE Systems Journal, 2014.

Co-authored in:

- i. L.Y. Zhai, M.J. Er, X. Li, B.S. Lim, S. Huang, J.H. Zhou, L. San and A.J. Torabi, Intelligent Monitoring of Surface Integrity and Cutter Degradation in High-speed Milling Processes, accepted at the Annual Conference of the Prognostics and Health Management Society 2010, October 10-16, 2010, Portland, Oregon, USA.
- ii. L. San, M. J. Er, X. Li, L. Y. Zhai, and A. J. Torabi, An enhanced online self-organizing fuzzy neural network, in 2010 11th International Conference on Control Automation Robotics Vision (ICARCV), 2010, pp. 2214 2220.
- iii. M. Pratama, M. J. Er, X. Li, L. San, J. O. Richard, L.-Y. Zhai, A. Torabi, and I. Arifin, Genetic Dynamic Fuzzy Neural Network (GDFNN) for Nonlinear System Identification, in Advances in Neural Networks ISNN 2011, D. Liu, H. Zhang, M. Polycarpou, C. Alippi, and H. He, Eds. Springer Berlin Heidelberg, 2011, pp. 525-534.
- iv. Pratama.M, M.J.Er, X.Li, Oentaryo.J.O, L.San, L-Y. Zhai, Amin J. Torabi, Arifin.I, Genetic Dynamic Fuzzy Neural Network for Nonlinear System Identification, International Symposium on Neural Network (ISNN),China ,2011.

- v. X. Li, M. J. Er, H. Ge, O. P. Gan, S. Huang, L. Y. Zhai, S. Linn, and A. J. Torabi, Adaptive Network Fuzzy Inference System and support vector machine learning for tool wear estimation in high speed milling processes, in IECON 2012 - 38th Annual Conference on IEEE Industrial Electronics Society, 2012, pp. 2821-2826.

6.2 Honors and Awards

IES Prestigious Engineering Achievement Award

Group Photo with University President

Appendix A

Nomenclature

Mathematical symbols in chapter 1

P	 probability distribution function
$L_\varepsilon(y)$	 ϵ -insensitive objective function
ξ_i and ζ_i^*	 non-negative slack variables
v	 velocity
x	 position
λ	 HMM model
q_i	 hidden states
A	 state transition matrix
B	 emission matrix
Q_t	 n th order Markov chain
Q'_t	 first order Markov chain
O_t	 spectral observations in each time instance
$\bar{n}(i, j)$	 code vector index
$\hat{t}_k$	 training sample of the high dimensional space

Mathematical symbols in chapter 2

$Y(x)$	 surface profile
--------	-----------------------

Mathematical symbols in section 2.2

$f(t)$	 signal
$F(w)$	 Fourier transform
ψ	 mother wavelet function
a	 dilation factor
b	 translation factor
$w(a, b)$	 continuous wavelet transform of a signal

Mathematical symbols in section 2.3

h_z	 f(net - hidden)
o_z	 f(net - output)
θ	 biases of the hidden layer neurons
C_{ij}	 weight between input neurons and hidden layer neurons
ϕ	 biases of the output layer neurons

D_{ij} weight between hidden neurons and output layer neurons

Mathematical symbols in section [2.4](#)

$E(f)$ energy of a signal f

S set of all clusters

μ membership value

V set of all n-dimensional centers

D_{ikA}^2 weighted norm-2 of the distance

F matrix that controls the shape of a cluster

F_i fuzzy covariance matrix

α_i a priori probability of selecting the cluster

v_i center of cluster i

Mathematical symbols in chapter [3](#)

u membership values

U weight assigned to a cluster

β forgetting factor

C_i center of i th RBF

σ RBF's width

R_j j th rule

γ decay constant

d_{min} smallest distance acceptable between the cluster centers

d_{max} largest distance between the cluster centers

ϵ convergence constant

e_{min} minimum error which defines the desired accuracy

W weights matrix

T target value matrix

D distance of the input from the center of RBF

err error reduction ratio

η importance of a rule

K_{err} threshold for err parameter

ρ_i significance of rule i

r number of existing rules

Appendix B

Abbreviations and Acronyms

AI	artificial intelligence
AISI	American Iron and Steel Institute
AFPN	adaptive fuzzy Petri-net
ANFIS	adaptive neuro-fuzzy inference system
ANN	artificial neural network
ASM	association matrix
BCCD	best cutting condition determination
BN	Bayesian network
BP	back propagation
CFFBP	cascaded feed forward back propagation
CNC	computed numerically controlled
Dc	depth of cut
DoE	design of experiment
DEVS	discrete event systems
DWT	discrete wavelet transform
EBP	Elman back propagation
ESEM	environment scanning electron microscopy
Fc	feed rate
Fz	feed per tooth
FFT	fast Fourier transform
FL	fuzzy logic
FN	fuzzy nets
FN-ASRC	fuzzy nets adaptive surface roughness control
FN-IPSR	fuzzy nets in process surface roughness recognition
FPN	fuzzy Petri net
EM	expectation maximization
GA	genetic algorithms
GONN	genetic algorithm optimized neural network
GP	genetic programming
HMM	hidden Markov model

HTGLA	hybrid Taguchi–genetic learning algorithm
ISO	International Organization for Standardization
LDA	linear discriminant analysis
MAE	mean absolute error
MAPE	mean absolute percentage error
MLP	multi-layer perceptron
MRE	mean relative error
MSE	mean square error
NA	not applicable
NM	not mentioned
NN	neural network
OA	orthogonal arrays
PSO	particle swarm optimization
PSOINN	particle swarm optimized neural network
PVD	physical vapour deposition
Ra	roughness profile, arithmetic average
Rq	roughness profile, root mean squared
Rv	roughness profile, maximum valley depth
Rp	roughness profile, maximum peak height
Rt	roughness profile, maximum height of the profile
RBF	radial basis function
RMSE	root mean square error
SA	simulated annealing
SOM	self organizing maps
SVM	support vector machine
SVR	support vector regression
TAN	tree augmented naive
TDBP	time-delay back-propagation
Vc	cutting speed
VMC	vertical machining centre

VQ		vector quantization
XRD		X-ray diffraction

Appendix C

Sensor Specifications

C.1 Dynamometer [Kis, a]

3-Component Dynamometer

with built-in Charge Amplifiers up to 10 kN

Type 9257BA, 9403, 5233A1

Quartz dynamometer for measuring the three orthogonal components of a force.

- Reliable force measurement system
- For measuring cutting forces measuring in milling, grinding and turning
- Built-in charge amplifiers – simple operation
- 4 switchable measuring ranges
- Conforms to CE per directive 89/336/EEC

Description

The dynamometer consists of four 3-component force sensors fitted under high preload between a base plate and a cover plate. The force components are measured practically without displacement.

The four sensors are fitted so that they are ground-isolated. This largely excludes ground loop problems.

The dynamometer is corrosion-resistant and protected against penetration by spray water and cutting fluid (IP 67).

The cover plate is fitted with a special thermal insulation layer which makes the dynamometer largely insensitive to temperature influences.

A 3-channel charge amplifier is built-in the dynamometer. Therefore, the output signal at the dynamometer is of low impedance. The integrated cable is connected to the control unit Type 5233A1. The control unit can select the four measuring ranges in two groups (F_x and F_y resp. F_z).

The dynamometer and the control unit meet CE requirements according to EU Directive 89/336/EEC and comply with EMC standards for industrial equipment (interference emission EN 50081-2 and interference immunity EN 50082-2). The control unit is easy to operate and contains power pack, keyboard with status displays together with a connector for signal input. The output voltages are proportional to the forces occurring.

Application

- General 3-component force measurement (dynamic and quasistatic)
- Cutting force measurements for optimization of the manufacturing process (temporary measurement)
- Cutting force measurements (turning, milling, grinding) for training purposes

Technical Data

Dynamometer Type 9257BA

Range 1	F _x , F _y	kN	-0,5 ... 0,5
	F _z	kN	-1 ... 1
Range 2	F _x , F _y	kN	-1 ... 1
	F _z	kN	-2 ... 2
Range 3	F _x , F _y	kN	-2 ... 2
	F _z	kN	-5 ... 5
Range 4	F _x , F _y	kN	-5 ... 5 ¹⁾
F _z for F _x and F _y ≤ 0,5 F _z	F _z	kN	-5 ... 10 ²⁾
Overload	F _x , F _y , F _z	kN	-7,5/7,5
F _z for F _x and F _y ≤ 0,5 F _z	F _z	kN	-7,5/15
Threshold		N	<0,01
Sensitivity	F _x , F _y	mV/N	10,0
Range 1	F _z	mV/N	5,00
Linearity, all ranges		% FSO	≤±1
Hysteresis, all ranges		% FSO	≤0,5
Cross talk		%	≤±3
Rigidity	C _x , C _y	kN/μm	>1
	C _z	kN/μm	>2
Natural frequency	f _n (x,y)	kHz	≈2,0
(mounted on flanges)	f _n (z)	kHz	≈3,5
Operating temperature range		°C	0 ... 60
Drift (charge amplifier)	F _x , F _y	N/s	≤±0,005
at 25 °C	F _z	N/s	≤±0,01
Ground isolation		MΩ	>100
Connecting cable (integral)	l	m	5
Degree of protection			IP 67
Weight		kg	7,4

¹⁾ Application of force inside and max. 25 mm above top plate area.

²⁾ Range for turning, application of force at point A.

Control Unit Type 5233A1

Number of channels		3
Number of ranges per channel		4
Output signal	V	±5
Cut-off frequency of the built-in filter	Hz	200
Output connector	3xBNC neg. 37 pin D-Sub	
Operating temperature range	°C	0 ... 60
Power supply (switchable)	V/AC %	230/115 +15/-22
Power consumption	VA	<23
Dimensions	mm	170x126x55
Weight	kg	1,52

Fig. 1: Dimensions of Dynamometer Type 9257BA

Fig. 2: Dimensions of Dynamometer Type 9257BA with tool holder Type 9403 mounted

Mounting

The dynamometer can be mounted with screws or clamps to any face-ground, clean mounting surface such as, for example, on a machine tool table. It must be noted that uneven contact surfaces may cause internal distortions, placing additional heavy stresses on the individual measuring elements and increasing the cross talk.

There are 14 M8x1,25 threaded holes are provided in the cover plate for clamping the force-inducing components such as turning tools and workpieces. The contact surfaces of the force-introducing parts must be surface ground to achieve good mechanical coupling to the cover plate.

The tool holder Type 9403 can be used for problem-free mounting of cutting tools up to a shaft cross-section of 26x26 mm.

The tool holder is not included in the parts supplied and must be ordered separately.

Accessories Included

• Cover	Type 3.322.275
• Mounting screw	6.150.058
• Connector	3.640.603
• Mounting screw	6.131.030

Optional Accessories

• Tool holder	Type 9403
(for cutting force measurements when turning)	
• Control Unit	5233A1
• Connecting cable	1500B5

Ordering Key

3-Component Dynamometer	Type 9257BA
-------------------------	-------------

Fig. 3: Measuring system

C.2 Accelerometer [Kis, b]

Annular Ceramic Shear Sensor

Type 8762A...

Light Weight, Voltage Mode, Triaxial Accelerometer

High sensitivity triaxial accelerometers that simultaneously measure vibration in three, mutually perpendicular axis (x, y and z). Designed primarily for modal analysis applications, the triaxial accelerometer features three tapped mounting surfaces that allow each axis to be hard mounted for calibration.

- Low impedance voltage mode
- Cube shaped, ceramic shear sensor
- Ultra low thermal transient response
- Durable hard anodized, ground isolated aluminium housing
- Conforming to CE

Description

Internal of the 8762A... accelerometer is a unique annular, shear sensor element that features extremely low thermal transient response, a high immunity to base strain and transverse acceleration. An advanced hybrid charge amplifier design provides outstanding phase response as well as a wide operating frequency range. The light weight aluminum housing is epoxy sealed and hard anodized coated to provide ground isolation.

Each of the three sensing elements is internally connected to a microelectronic circuit that converts the charge from the ceramic piezoelectric elements into a useable high level voltage signal at a low impedance output. The 8762A... accelerometer series will operate directly from the internal power source found in most FFT analyzers; from several Kistler Piezotron® power supply couplers or any industry standard IEPE (Integrated Electronic Piezo-Electric) compatible power source.

Application

The lightweight 8762A... triaxial accelerometer series, is highly desirable for measurement applications on light weight structures where mass loading must be kept to a minimum. The accelerometers are highly suited for multi-channel measurements; modal analysis measurements on automotive bodies and aircraft structures; general vibration measurements.

Accessing TEDS Data

Accelerometers with a "T" suffix are variants of the standard version incorporating the "Smart Sensor" design. Viewing an accelerometer's data sheet requires an Interface/Coupler such as Kistler's Type 5134B... or 5000M04 with TEDS Editor software. The Interface provides negative current excitation (reverse polarity) altering the operating mode of the PiezoSmart® sensor allowing the program editor software to read or add information contained in the memory chip.

Mounting

The 8762A... accelerometer series can be attached to the test surface by using a 10-32 stud inserted in any one of the three threaded mounting holes. Reliable and accurate measurements require that the mounting surface be clean and flat. The instruction manual for the 8762A... (8762A_002-233) provides detailed information regarding mounting surface preparation.

Technical Data

Specification	Unit	Type 8762A5	Type 8762A10	Type 8762A50
Acceleration range	g	±5	±10	±50
Acceleration limit	gpk	±8	±16	±80
Threshold, nom.	grms	0,0003	0,00035	0,0012
Sensitivity, ±5 %	mV/g	1 000	500	100
Resonant frequency, mounted, nom.	kHz	30	30	30
Frequency response, ±5 %	Hz	0,5 ... 6 000	0,5 ... 6 000	0,5 ... 6 000
Amplitude non-linearity	%FSO	±1	±1	±1
Time constant, nom.	s	1	1	1
Transverse sensitivity, nom.	%	<5	<5	<5

Environmental

Base strain sensitivity @ 250 µε	g/µε	0,004	0,004	0,004
Shock limit (0,2 ms pulse)	gpk	5 000	7 000	7 000
Temperature coefficient of sensitivity	%/°C	-0,06	-0,02	-0,02
Operating temperature range	°C	-55 ... 80	-55 ... 80	-55 ... 80
Type 8762A...T	°C	-40 ... 80	-40 ... 80	-40 ... 80

Output

Bias, nom.	VDC	11	11	11
Impedance	Ω	≤500	≤500	≤100
Voltage full scale	V	±5	±5	±5

Source

Voltage	VDC	20 ... 30	20 ... 30	20 ... 30
Constant current	mA	2 ... 18	2 ... 18	2 ... 18

Construction

Sensing element	Type	ceramic-shear	ceramic-shear	ceramic-shear
Case/base	material	aluminum hard anodized	aluminium hard anodized	aluminium hard anodized
Degree of protection case/connector (EN 60529)		IP66	IP66	IP66
Connector	Type	4-pin pos.	4-pin pos.	4-pin pos.
Ground isolated		yes	yes	yes
Mass	grams	23	23	23
Mounting (10-32 thd.x4 dp)	Type	stud	stud	stud

1 g = 9.80665 m/s², 1 inch = 25.4 mm, 1 gram = 0.03527 oz, 1 lbf-in = 0.1129 N·m

Bibliography

- [MAT,] Fuzzy clustering :: Tutorial (Fuzzy logic toolbox™). [online] <http://www.mathworks.com/access/helpdesk/help/toolbox/fuzzy/fp310.html>.
- [Lia,] Intelligent monitoring of surface integrity and cutter degradation in high-speed milling processes.
- [iso,] ISO 4288:1996 - geometrical product specifications (GPS) - surface texture: Profile method - rules and procedures for the assessment of surface texture. [online] <http://www.helmar.com.pl/helmar/biblioteka/pdf/9257ba.pdf>.
- [Kis, a] kisler 3-component dynamometer. [online] http://www.techstreet.com/standards/ISO/4288_1996?product_id=22877.
- [Kis, b] Kisler annular ceramic shear sensor. [online] http://www.intertechnology.com/Kistler/pdfs/Accelerometer_Model_8762A.pdf.
- [Çolak et al., 2007] Çolak, O., Kurbanoglu, C., and Kayacan, M. C. (2007). Milling surface roughness prediction using evolutionary programming methods. *Materials & Design*, 28(2):657–666.
- [Abrari et al., 1998] Abrari, F., Elbestawi, M., and Spence, A. (1998). On the dynamics of ball end milling: modeling of cutting forces and stability analysis. *International Journal of Machine Tools and Manufacture*, 38(3):215–237.
- [Abu-Mahfouz, 2003] Abu-Mahfouz, I. (2003). Drilling wear detection and classification using vibration signals and artificial neural network. *International Journal of Machine Tools and Manufacture*, 43(7):707–720.

- [Achiche et al., 2002] Achiche, S., Balazinski, M., Baron, L., and Jemielniak, K. (2002). Tool wear monitoring using genetically-generated fuzzy knowledge bases. *Engineering Applications of Artificial Intelligence*, 15(3-4):303–314.
- [Al-Habaibeh and Gindy, 2000] Al-Habaibeh, A. and Gindy, N. (2000). A new approach for systematic design of condition monitoring systems for milling processes. *Journal of Materials Processing Technology*, 107(1-3).
- [Aliustaoglu et al., 2009] Aliustaoglu, C., Ertunc, H. M., and Ocak, H. (2009). Tool wear condition monitoring using a sensor fusion model based on fuzzy inference system. *Mechanical Systems and Signal Processing*, 23(2):539–546.
- [Amador, 2005] Amador, J. J. (2005). Sequential clustering by statistical methodology. *Pattern Recognition Letters*, 26(14):2152–2163.
- [Asiltürk and Çunkas, 2011] Asiltürk, I. and Çunkas, M. (2011). Modeling and prediction of surface roughness in turning operations using artificial neural network and multiple regression method. *Expert Systems with Applications*, 38(5):5826–5832.
- [ASME, 1996] ASME (1996). Y14.36m - 1996 surface texture symbols. [online] http://catalog.asme.org/Codes/PrintBook/Y1436M_1996_Surface_Texture.cfm.
- [Atlas et al., 2000] Atlas, L., Ostendorf, M., and Bernard, G. (2000). Hidden markov models for monitoring machining tool-wear. In *Acoustics, Speech, and Signal Processing, 2000. ICASSP '00. Proceedings. 2000 IEEE International Conference on*, volume 6, pages 3887–3890 vol.6.
- [Aurich and Bil, 2006] Aurich, J. and Bil, H. (2006). 3D finite element modelling of segmented chip formation. *CIRP Annals - Manufacturing Technology*, 55(1):47–50.
- [Baek et al., 2001] Baek, D. K., Ko, T. J., and Kim, H. S. (2001). Optimization of feedrate in a face milling operation using a surface roughness model. *International Journal of Machine Tools and Manufacture*, 41(3):451–462.
- [Bassiuny and Li, 2007] Bassiuny, A. and Li, X. (2007). Flute breakage detection during end milling using Hilbert-Huang transform and smoothed nonlinear energy operator. *International Journal of Machine Tools and Manufacture*, 47(6):1011–1020.

- [Baum and Petrie, 1966] Baum, L. E. and Petrie, T. (1966). Statistical inference for probabilistic functions of finite state markov chains. *The Annals of Mathematical Statistics*, 37(6):1554–1563.
- [Berkhin, 2006] Berkhin, P. (2006). A survey of clustering data mining techniques. *Grouping Multidimensional Data*, pages 25–71.
- [Bezdek and Dunn, 1975] Bezdek, J. and Dunn, J. (1975). Optimal fuzzy partitions: A heuristic for estimating the parameters in a mixture of normal distributions. *Computers, IEEE Transactions on*, C-24(8):835–838.
- [Bezdek, 1981] Bezdek, J. C. (1981). *Pattern recognition with fuzzy objective function algorithms*. Kluwer Academic Publishers Norwell, MA, USA.
- [Bhattacharyya et al., 2008] Bhattacharyya, P., Sengupta, D., Mukhopadhyay, S., and Chattopadhyay, A. (2008). On-line tool condition monitoring in face milling using current and power signals. *International Journal of Production Research*, 46(4):1187–1201.
- [Bicego et al., 2009] Bicego, M., Pekalska, E., Tax, D. M., and Duin, R. P. (2009). Component-based discriminative classification for hidden markov models. *Pattern Recognition*, 42(11):2637–2648.
- [Bilmes, 2006] Bilmes, J. A. (2006). What HMMs can do. *IEICE - Trans. Inf. Syst.*, E89-D(3):869–891.
- [Bralla, 1999] Bralla, J. G. (1999). *Design for manufacturability handbook*. McGraw-Hill Professional.
- [Brezocnik et al., 2004] Brezocnik, M., Kovacic, M., and Ficko, M. (2004). Prediction of surface roughness with genetic programming. *Journal of Materials Processing Technology*, 157-158:28–36.
- [Briceno et al., 2002] Briceno, J. F., El-Mounayri, H., and Mukhopadhyay, S. (2002). Selecting an artificial neural network for efficient modeling and accurate simulation of the milling process. *International Journal of Machine Tools and Manufacture*, 42(6):663–674.

- [c Yang, 2006] c Yang, Y. K. (2006). Optimization of dry machining parameters for high-purity graphite in end-milling process. *Materials and Manufacturing Processes*, 21(8):832–837.
- [Camproux et al., 2004] Camproux, A., Gautier, R., and Tufféry, P. (2004). A hidden markov model derived structural alphabet for proteins. *Journal of Molecular Biology*, 339(3):591–605.
- [Cardoso, 1999] Cardoso, Janette; Camargo, H. E. (1999). *Fuzziness in Petri Nets*. Springer.
- [Cetin et al., 2007] Cetin, O., Ostendorf, M., and Bernard, G. (2007). Multirate coupled hidden markov models and their application to machining Tool-Wear classification. *Signal Processing, IEEE Transactions on*, 55(6):2885–2896.
- [Chakraborty et al., 2008] Chakraborty, P., Asfour, S., Cho, S., Onar, A., and Lynn, M. (2008). Modeling tool wear progression by using mixed effects modeling technique when end-milling AISI 4340 steel. *Journal of Materials Processing Technology*, 205(1-3):190–202.
- [Chen and Savage, 2001] Chen, J. and Savage, M. (2001). A Fuzzy-Net-Based multilevel In-Process surface roughness recognition system in milling operations. *The International Journal of Advanced Manufacturing Technology*, 17(9):670–676.
- [Chen and Susanto, 2003] Chen, J. and Susanto, V. (2003). Fuzzy logic based In-Process Tool-Wear monitoring system in face milling operations. *The International Journal of Advanced Manufacturing Technology*, 21(3):186–192.
- [Chen and Lou, 2000] Chen, J. C. and Lou, M. S. (2000). Fuzzy-nets based approach to using an accelerometer for an in-process surface roughness prediction system in milling operations. *International Journal of Computer Integrated Manufacturing*, 13(4):358.
- [Chen et al., 1991] Chen, S., Cowan, C. F. N., and Grant, P. (1991). Orthogonal least squares learning algorithm for radial basis function networks. *IEEE Transactions on Neural Networks*, 2(2):302–309.
- [Chen and Li, 2007] Chen, X. and Li, B. (2007). Acoustic emission method for tool condition monitoring based on wavelet analysis. *The International Journal of Advanced*

- Manufacturing Technology*, 33(9):968–976.
- [Cheng, 2009] Cheng, K. (2009). *Machining Dynamics Fundamentals, Applications and Practices*. Springer.
- [Cho et al., 2005] Cho, S., Asfour, S., Onar, A., and Kaundinya, N. (2005). Tool breakage detection using support vector machine learning in a milling process. *International Journal of Machine Tools and Manufacture*, 45(3):241–249.
- [Choi et al., 2004] Choi, Y., Narayanaswami, R., and Chandra, A. (2004). Tool wear monitoring in ramp cuts in end milling using the wavelet transform. *The International Journal of Advanced Manufacturing Technology*, 23(5):419–428.
- [CICYT, 2005] CICYT (2005). ELVIRA software. [online] <http://www.ia.uned.es/~elvira/index-en.html>.
- [Correa et al., 2009] Correa, M., Bielza, C., and Pamies-Teixeira, J. (2009). Comparison of bayesian networks and artificial neural networks for quality detection in a machining process. *Expert Systems with Applications*, 36(3, Part 2):7270–7279.
- [Correa et al., 2008] Correa, M., Bielza, C., Ramirez, M., and Alique, J. (2008). A bayesian network model for surface roughness prediction in the machining process. *International Journal of Systems Science*, 39(12):1181–1192.
- [Cus et al., 2007] Cus, F., Zuperl, U., and Gecevskas, V. (2007). High speed end-milling optimisation using particle swarm intelligence. *Journal of Achievements in Materials and Manufacturing Engineering*, 22(2):75–78.
- [Daubechies, 1990] Daubechies, I. (1990). The wavelet transform, time-frequency localization and signal analysis. *Information Theory, IEEE Transactions on*, 36(5):961–1005.
- [David and Alla, 1994] David, R. and Alla, H. (1994). Petri nets for modeling of dynamic systems : A survey. *Automatica*, 30(2):175–202.
- [Dey and Stori, 2005] Dey, S. and Stori, J. A. (2005). A bayesian network approach to root cause diagnosis of process variations. *International Journal of Machine Tools and Manufacture*, 45(1):75–91.
- [Diei and Dornfeld, 1987] Diei, E. and Dornfeld, D. (1987). Acoustic emission sensing of tool wear in face milling. *TRANS. ASME J. ENGN. IND.*, 109(3 , Aug. 1987):232–

247.

- [Dong et al., 2006] Dong, J., Subrahmanyam, K., Wong, Y., Hong, G., and Mohanty, A. (2006). Bayesian-inference-based neural networks for tool wear estimation. *The International Journal of Advanced Manufacturing Technology*, 30(9):797–807.
- [Doostfatemeh and Kremer, 2005] Doostfatemeh, M. and Kremer, S. C. (2005). New directions in fuzzy automata. *International Journal of Approximate Reasoning*, 38(2):175–214.
- [Duda et al., 2001] Duda, R. O., Hart, P. E., and Stork, D. G. (2001). *Pattern classification*. Citeseer.
- [Dweiri et al., 2003] Dweiri, F., Al-Jarrah, M., and Al-Wedyan, H. (2003). Fuzzy surface roughness modeling of CNC down milling of aluminic-79. *Journal of Materials Processing Technology*, 133(3):266–275.
- [El-Mounayri et al., 2003] El-Mounayri, H., Dugla, Z., and Deng, H. (2003). Prediction of surface roughness in end milling using swarm intelligence. In *Swarm Intelligence Symposium, 2003. SIS '03. Proceedings of the 2003 IEEE*, pages 220–227.
- [El-Sonbaty et al., 2008] El-Sonbaty, I., Khashaba, U., Selmy, A., and Ali, A. (2008). Prediction of surface roughness profiles for milled surfaces using an artificial neural network and fractal geometry approach. *Journal of Materials Processing Technology*, 200(1-3):271–278.
- [El-Sonbaty and Ismail, 1998] El-Sonbaty, Y. and Ismail, M. A. (1998). On-line hierarchical clustering. *Pattern Recognition Letters*, 19(14):1285–1291.
- [Elbestawi and Dumitrescu, 2006] Elbestawi, M. and Dumitrescu, M. (2006). Tool condition monitoring in machining - neural networks. In *Information Technology For Balanced Manufacturing Systems*, volume 220, pages 5–16. Springer.
- [Elbestawi et al., 1989] Elbestawi, M., Marks, J., and Papazafiriou, T. (1989). Process monitoring in milling by pattern recognition. *Mechanical Systems and Signal Processing*, 3(3):305–315.
- [Ephraim and Merhav, 2002] Ephraim, Y. and Merhav, N. (2002). Hidden markov processes. *Information Theory, IEEE Transactions on*, 48(6):1518–1569.

- [Er and Deng, 2004] Er, M. J. and Deng, C. (2004). Online tuning of fuzzy inference systems using dynamic fuzzy q-learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 34(3):1478–1489.
- [Ertunc et al., 2001] Ertunc, H. M., Loparo, K. A., and Ocak, H. (2001). Tool wear condition monitoring in drilling operations using hidden markov models (HMMs). *International Journal of Machine Tools and Manufacture*, 41(9):1363–1384.
- [Ertunc and Oysu, 2004] Ertunc, H. M. and Oysu, C. (2004). Drill wear monitoring using cutting force signals. *Mechatronics*, 14(5):533–548.
- [Farran et al., 2009] Farran, B., Ramanan, A., and Niranjana, M. (2009). Sequential hierarchical pattern clustering. *Pattern Recognition in Bioinformatics*, pages 79–88.
- [Fish et al., 2003] Fish, R., Ostendorf, M., Bernard, G., and Castanon, D. (2003). Multilevel classification of milling tool wear with confidence estimation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(1):75–85.
- [Fontaine et al., 2007] Fontaine, M., Moufki, A., Devillez, A., and Dudzinski, D. (2007). Modelling of cutting forces in ball-end milling with tool-surface inclination: Part i: Predictive force model and experimental validation. *Journal of Materials Processing Technology*, 189(1-3):73–84.
- [Fred and Jain, 2003] Fred, A. L. N. and Jain, A. K. (2003). Robust data clustering.
- [Gath and Geva, 1989] Gath, I. and Geva, A. B. (1989). Unsupervised optimal fuzzy clustering. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 11(7):773–780.
- [Giles et al., 1999] Giles, C. L., Omlin, C. W., and Thornber, K. K. (1999). Equivalence in knowledge representation: automata, recurrent neural networks, and dynamical fuzzy systems. *Proceedings of the IEEE*, 87(9):1623–1640.
- [Haber et al., 2004] Haber, R. E., Jimić, J. E., Peres, C. R., and Alique, J. R. (2004). An investigation of tool-wear monitoring in a high-speed machining process. *Sensors and Actuators A: Physical*, 116(3):539–545.
- [Halkidi et al., 2002] Halkidi, M., Batistakis, Y., and Vazirgiannis, M. (2002). Cluster validity methods: part i. *SIGMOD Rec.*, 31(2):40–45.

- [Hanna et al., 1996] Hanna, M., Buck, A., and Smith, R. (1996). Fuzzy petri nets with neural networks to model products quality from a CNC-milling machining centre. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans.*, 26(5):638–645.
- [Hansen and Jaumard, 1997] Hansen, P. and Jaumard, B. (1997). Cluster analysis and mathematical programming. *Mathematical Programming*, 79(1):191–215.
- [Hassoun, 1995] Hassoun, M. H. (1995). *Fundamentals of artificial neural networks*. The MIT Press.
- [Heck and McClellan, 1991] Heck, L. and McClellan, J. (1991). Mechanical system monitoring using hidden markov models. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on*, volume 0, pages 1697–1700, Los Alamitos, CA, USA. IEEE Computer Society.
- [Ho et al., 2009] Ho, W., Tsai, J., Lin, B., and Chou, J. (2009). Adaptive network-based fuzzy inference system for prediction of surface roughness in end milling process using hybrid taguchi-genetic learning algorithm. *Expert Systems with Applications*, 36(2, Part 2):3216–3222.
- [Holland, 1992] Holland, J. H. (1992). *Adaptation in natural and artificial systems*. MIT press Cambridge, MA.
- [Hoppner, 2002] Hoppner (2002). Speeding up fuzzy c-means: using a hierarchical data organisation to control the precision of membership calculation. *Fuzzy Sets and Systems*, 128(3):365–376.
- [Hsueh and Yang, 2008] Hsueh, Y. and Yang, C. (2008). Prediction of tool breakage in face milling using support vector machine. *The International Journal of Advanced Manufacturing Technology*, 37(9):872–880.
- [Hsueh and Yang, 2009] Hsueh, Y. and Yang, C. (2009). Tool breakage diagnosis in face milling by support vector machine. *Journal of Materials Processing Technology*, 209(1):145–152.
- [Huang et al., 2006] Huang, G., Zhu, Q., and Siew, C. (2006). Extreme learning machine: Theory and applications. *Neurocomputing*, 70(1-3):489–501.

- [Jain et al., 2000] Jain, A., Duin, R., and Mao, J. (2000). Statistical pattern recognition: a review. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(1):4–37.
- [Jain et al., 1996] Jain, A., Mao, J., and Mohiuddin, K. (1996). Artificial neural networks: a tutorial. *Computer*, 29(3):31–44.
- [Jain et al., 1999] Jain, A. K., Murty, M. N., and Flynn, P. J. (1999). Data clustering: a review. *ACM computing surveys (CSUR)*, 31(3):264–323.
- [Javed et al., 2012] Javed, K., Gouriveau, R., Zerhouni, N., Zemouri, R., and Li, X. (2012). Robust, reliable and applicable tool wear monitoring and prognostic: Approach based on an improved-extreme learning machine. In *2012 IEEE Conference on Prognostics and Health Management (PHM)*, pages 1–9.
- [Jr. et al., 2005] Jr., A. V., Nolazco-Flores, J., Morales-Menéndez, R., Sucar, L., and Rodríguez, C. (2005). Tool-Wear monitoring based on continuous hidden markov models. In *Progress in Pattern Recognition, Image Analysis and Applications*, pages 880–890.
- [J.R.Koza, 1992] J.R.Koza (1992). *Genetic Programming*. The MIT Press, Cambridge, MA.
- [Juang and Rabiner, 1991] Juang, B. H. and Rabiner, L. R. (1991). Hidden markov models for speech recognition. *Technometrics*, 33(3):251–272.
- [Kasashima et al., 1995] Kasashima, N., Mori, K., Ruiz, G. H., and Taniguchi, N. (1995). Online failure detection in face milling using discrete wavelet transform. *CIRP Annals - Manufacturing Technology*, 44(1):483–487.
- [Kasirolvalad et al., 2006] Kasirolvalad, Z., Motlagh, M., and Shadmani, M. (2006). An intelligent modeling system to improve the machining process quality in CNC machine tools using adaptive fuzzy petri nets. *International Journal of Advanced Manufacturing Technology*, 29(9-10):1050–1061.
- [Kennedy et al., 2001] Kennedy, J., Eberhart, R. C., and Shi, Y. (2001). *Swarm intelligence*. Springer.
- [Kim and Jeon, 2011] Kim, D. and Jeon, D. (2011). Fuzzy-logic control of cutting forces in CNC milling processes using motor currents as indirect force sensors. *Precision*

Engineering, 35(1):143–152.

- [Kim et al., 2005] Kim, K., Whang, E. J., Park, C.-W., Kim, E., and Park, M. (2005). A TSK fuzzy inference algorithm for online identification. In Wang, L. and Jin, Y., editors, *Fuzzy Systems and Knowledge Discovery*, number 3613 in Lecture Notes in Computer Science, pages 179–188. Springer Berlin Heidelberg.
- [Kirby et al., 2006] Kirby, E. D., Chen, J. C., and Zhang, J. Z. (2006). Development of a fuzzy-nets-based in-process surface roughness adaptive control system in turning operations. *Expert Systems with Applications*, 30(4):592–604.
- [Kline et al., 1982] Kline, W., DeVor, R., and Lindberg, J. (1982). The prediction of cutting forces in end milling with application to cornering cuts. *International Journal of Machine Tool Design and Research*, 22(1):7–22.
- [Ko and Cho, 2005] Ko, J. H. and Cho, D. (2005). 3D Ball-End milling force model using instantaneous cutting force coefficients. *Journal of Manufacturing Science and Engineering*, 127(1):1–12.
- [Kolatch, 2001] Kolatch, E. (2001). Clustering algorithms for spatial databases: A survey. *PDF is available on the Web*.
- [Kosko and Burgess, 1998] Kosko, B. and Burgess, J. C. (1998). Neural networks and fuzzy systems. *The Journal of the Acoustical Society of America*, 103:3131.
- [Kumar et al., 2008] Kumar, A., Tseng, F., Guo, Y., and Chinnam, R. (2008). Hidden-Markov model based sequential clustering for autonomous diagnostics. In *Neural Networks, 2008. IJCNN 2008. (IEEE World Congress on Computational Intelligence). IEEE International Joint Conference on*, pages 3345–3351.
- [Kumar et al., 2007] Kumar, P., Krishna, P. R., Bapi, R. S., and De, S. K. (2007). Rough clustering of sequential data. *Data & Knowledge Engineering*, 63(2):183–199.
- [Lee, 1990] Lee, K. F. (1990). Context-dependent phonetic hidden markov models for speaker-independent continuous speech recognition. *Readings in speech recognition*, pages 347–65.
- [Lee and Kil, 1991] Lee, S. and Kil, R. M. (1991). A gaussian potential function network with hierarchically self-organizing learning. *Neural Networks*, 4(2):207–224.

- [Lela et al., 2009] Lela, B., Baji, D., and Jozi, S. (2009). Regression analysis, support vector machines, and bayesian neural network approaches to modeling surface roughness in face milling. *The International Journal of Advanced Manufacturing Technology*, 42(11):1082–1088.
- [Lewis, 1992] Lewis, D. D. (1992). Evaluation of phrasal and clustered representations on a text categorization task. In *Proceedings of the Fifteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 37–50.
- [Li et al., 2006a] Li, H., Zeng, H., and Chen, X. (2006a). An experimental study of tool wear and cutting force variation in the end milling of inconel 718 with coated carbide inserts. *Journal of Materials Processing Technology*, 180(1-3):296–304.
- [Li and Wu, 1988] Li, P. G. and Wu, S. M. (1988). Monitoring drilling wear states by a fuzzy pattern recognition technique. *J. Eng. Ind.(Trans. ASME)*, 110(3):297–300.
- [Li, 2002] Li, X. (2002). A brief review: acoustic emission method for tool wear monitoring during turning. *International Journal of Machine Tools and Manufacture*, 42(2):157–165.
- [Li et al., 2010] Li, X., Er, M., Lim, B., Zhou, J., Gan, O., and Rutkowski, L. (2010). Fuzzy regression modeling for tool performance prediction and degradation detection. *International Journal of Neural Systems*, 20(5):405–419.
- [Li and Guan, 2004] Li, X. and Guan, X. P. (2004). Time-frequency-analysis-based minor cutting edge fracture detection during end milling. *Mechanical Systems and Signal Processing*, 18(6):1485–1496.
- [Li et al., 2000] Li, X., Tso, S. K., and Wang, J. (2000). Real-time tool condition monitoring using wavelet transforms and fuzzy techniques. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 30(3):352–357.
- [Li et al., 2006b] Li, X., Zhou, J., Zeng, H., Wong, Y. S., and Hong, G. S. (2006b). An intelligent predictive engine for milling machine prognostic monitoring. In *2006 IEEE International Conference on Industrial Informatics*. IEEE.
- [Liang et al., 2004] Liang, S. Y., Hecker, R. L., and Landers, R. G. (2004). Machining process monitoring and control: The State-of-the-Art. *Journal of Manufacturing Science*

- and Engineering*, 126(2):297–310.
- [Liao, 2005] Liao, T. W. (2005). Clustering of time series data—a survey. *Pattern Recognition*, 38(11):1857–1874.
- [Lim et al., 1995] Lim, E. M., Feng, H., Menq, C., and hang Lin, Z. (1995). The prediction of dimensional error for sculptured surface productions using the ball-end milling process. part 1: Chip geometry analysis and cutting force prediction. *International Journal of Machine Tools and Manufacture*, 35(8):1149–1169.
- [Linz, 2006] Linz, P. (2006). *An Introduction to Formal Language and Automata*. Jones and Bartlett Publishers.
- [Liu and Er, 2012] Liu, F. and Er, M. J. (2012). A novel efficient learning algorithm for self-generating fuzzy neural network with applications. *International Journal of Neural Systems*, 22(01):21–35.
- [Liu and Loftus, 2006] Liu, N. and Loftus, M. (2006). Prediction of surface quality from Ball-Nose milling in High-Speed machining applications. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, 220(4):571–578.
- [Liu and Wang, 1999] Liu, Y. and Wang, C. (1999). A modified genetic algorithm based optimisation of milling parameters. *The International Journal of Advanced Manufacturing Technology*, 15(11):796–799.
- [Lo, 2003] Lo, S. (2003). An adaptive-network based fuzzy inference system for prediction of workpiece surface roughness in end milling. *Journal of Materials Processing Technology*, 142(3):665–675.
- [Lombardo et al., 1997] Lombardo, A., Masnata, A., and Settineri, L. (1997). In-process tool-failure detection by means of AR models. *The International Journal of Advanced Manufacturing Technology*, 13(2):86–94.
- [Lu, 2008] Lu, C. (2008). Study on prediction of surface quality in machining process. *Journal of Materials Processing Tech.*, 205(1-3):439–450.
- [marusich, 1995] marusich (1995). Modelling and simulation of High-Speed Machining. *International journal for numerical methods in engineering*, 38:3675–3694.

- [Mathew et al., 2008] Mathew, M., Pai, P., and Rocha, L. (2008). An effective sensor for tool wear monitoring in face milling: Acoustic emission. *Sadhana*, 33(3):227–233.
- [Mcdowell, 1999] Mcdowell, R. L. (1999). Us patent number: Usp005904457, detecting tool wear by thermal monitoring of workpiece. [online] <http://www.freepatentsonline.com>.
- [Merdol and Altintas, 2004] Merdol, S. D. and Altintas, Y. (2004). Multi frequency solution of chatter stability for low immersion milling. *Journal of Manufacturing Science and Engineering*, 126(3):459–466.
- [Miao and Makis, 2007] Miao, Q. and Makis, V. (2007). Condition monitoring and classification of rotating machinery using wavelets and hidden markov models. *Mechanical Systems and Signal Processing*, 21(2):840–855.
- [Mike S. Lou, 1998] Mike S. Lou, Joseph C. Chen, C. M. L. (1998). Surface roughness prediction technique for cnc end-milling. *Journal of Industrial technology*, 15(1):1–6.
- [Milani, 2004] Milani, A. (2004). Online genetic algorithms. *International Journal of Information Theories and Applications*, 11:20–28.
- [Morgan et al., 2007] Morgan, G., Cheng, R. Q., Altintas, Y., and Ridgway, K. (2007). An expert troubleshooting system for the milling process. *International Journal of Machine Tools and Manufacture*, 47(9):1417–1425.
- [Muammer Nalbant and Toktas, 2007] Muammer Nalbant, H. G. and Toktas, I. (2007). Comparison of regression and artificial neural network models for surface roughness prediction with the cutting parameters in cnc turning. *Modelling and Simulation in Engineering*, 15(1).
- [Murtagh, 1983] Murtagh, F. (1983). A survey of recent advances in hierarchical clustering algorithms. *The Computer Journal*, 26(4):354–359.
- [Nalbant et al., 2007] Nalbant, M., AltIn, A., and Gokkaya, H. (2007). The effect of cutting speed and cutting tool geometry on machinability properties of nickel-base inconel 718 super alloys. *Materials & Design*, 28(4):1334–1338.
- [Natarajan et al., 2007] Natarajan, U., Periasamy, V., and Saravanan, R. (2007). Application of particle swarm optimisation in artificial neural network for the prediction of tool

- life. *The International Journal of Advanced Manufacturing Technology*, 31(9):871–876.
- [Ng et al., 2000] Ng, E., Lee, D., Dewes, R., and Aspinwall, D. (2000). Experimental evaluation of cutter orientation when ball nose end milling inconel 718(TM). *Journal of Manufacturing Processes*, 2(2):108–115.
- [Oberg et al., 2004] Oberg, E., Jones, F. D., Horton, H. L., and Ryffel, H. H. (2004). *Machinery’s handbook*. Industrial Press New York.
- [Oentaryo et al., 2011] Oentaryo, R., Er, M. J., San, L., Zhai, L., and Li, X. (2011). Bayesian ART-based fuzzy inference system: A new approach to prognosis of machining processes. pages 1–10.
- [Osei-Bryson and Inniss, 2007] Osei-Bryson, K. and Inniss, T. R. (2007). A hybrid clustering algorithm. *Computers & Operations Research*, 34(11):3255–3269.
- [Owsley et al., 1997a] Owsley, L., Atlas, L., and Bernard, G. (1997a). Automatic clustering of vector time-series for manufacturing machine monitoring. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on*, volume 4, page 3393, Los Alamitos, CA, USA. IEEE Computer Society.
- [Owsley et al., 1997b] Owsley, L., Atlas, L., and Bernard, G. (1997b). Self-organizing feature maps and hidden markov models for machine-tool monitoring. *Signal Processing, IEEE Transactions on*, 45(11):2787–2798.
- [Pai and Rao, 2002] Pai, P. S. and Rao, P. K. R. (2002). Acoustic emission analysis for tool wear monitoring in face milling. *International Journal of Production Research*, 40(5):1081.
- [Pal et al., 1996] Pal, N. R., Bezdek, J. C., and Hathaway, R. J. (1996). Sequential competitive learning and the fuzzy c-means clustering algorithms. *Neural Networks*, 9(5):787–796.
- [Palanisamy et al., 2008] Palanisamy, P., Rajendran, I., and Shanmugasundaram, S. (2008). Prediction of tool wear using regression and ANN models in end-milling operation. *The International Journal of Advanced Manufacturing Technology*, 37(1):29–41.
- [Paris et al., 2004] Paris, H., Peigne, G., and Mayer, R. (2004). Surface shape prediction in high speed milling. *International Journal of Machine Tools and Manufacture*,

44(15):1567–1576.

- [Patra et al., 2007] Patra, K., Pal, S. K., and Bhattacharyya, K. (2007). Application of wavelet packet analysis in drill wear monitoring. *Machining Science and Technology: An International Journal*, 11(3):413.
- [Paul, 1994] Paul, D. (1994). Detection of change in processes using wavelets. In *Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, pages 174–177.
- [Peng et al., 2011] Peng, Z. B., Li, J. L., Han, X. F., and Yan, J. (2011). The Real-Time monitoring tool states based on wavelet analysis. *Advanced Materials Research*, 189-193:881–886.
- [Prakasvudhisarn et al., 2009] Prakasvudhisarn, C., Kunnapapdeelert, S., and Yenradee, P. (2009). Optimal cutting condition determination for desired surface roughness in end milling. *The International Journal of Advanced Manufacturing Technology*, 41(5):440–451.
- [Pratama et al., 2011] Pratama, M., Er, M.-J., Li, X., Gan, O. P., Oentaryo, R., Linn, S., Zhai, L., and Arifin, I. (2011). Tool wear prediction using evolutionary dynamic fuzzy neural (EDFNN) network. In *IECON 2011 - 37th Annual Conference on IEEE Industrial Electronics Society*, pages 4739–4744.
- [Prickett and Johns, 1999] Prickett, P. W. and Johns, C. (1999). An overview of approaches to end milling tool monitoring. *International Journal of Machine Tools and Manufacture*, 39(1):105–122.
- [Rao et al., 2011] Rao, B., Dandekar, C. R., and Shin, Y. C. (2011). An experimental and numerical study on the face milling of Ti-6Al-4V alloy: Tool performance and surface integrity. *Journal of Materials Processing Technology*, 211(2):294–304.
- [Razfar et al., 2010] Razfar, M. R., Asadnia, M., Haghshenas, M., and Farahnakian, M. (2010). Optimum surface roughness prediction in face milling X20Cr13 using particle swarm optimization algorithm. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, 224(11):1645 –1653.

- [Regad et al., 2008] Regad, L., Guyon, F., Maupetit, J., Tufféry, P., and Camproux, A. (2008). A hidden markov model applied to the protein 3D structure analysis. *Computational Statistics & Data Analysis*, 52(6):3198–3207.
- [Rivero et al., 2008] Rivero, A., de Lacalle, L. L., and Penalva, M. L. (2008). Tool wear detection in dry high-speed milling based upon the analysis of machine internal signals. *Mechatronics*, 18(10):627–633.
- [Routara et al., 2009] Routara, B., Bandyopadhyay, A., and Sahoo, P. (2009). Roughness modeling and optimization in CNC end milling using response surface method: effect of workpiece material variation. *The International Journal of Advanced Manufacturing Technology*, 40(11):1166–1180.
- [Sagherian and Elbestawi, 1990] Sagherian, R. and Elbestawi, M. (1990). A simulation system for improving machining accuracy in milling. *Computers in Industry*, 14(4):293–305.
- [Saglam and Unuvar, 2003] Saglam, H. and Unuvar, A. (2003). Tool condition monitoring in milling based on cutting forces by a neural network. *International Journal of Production Research*, 41(7):1519 – 1532.
- [Serrano and Fabio, 1996] Serrano, E. and Fabio, M. (1996). Application of the wavelet transform to acoustic emission signals processing. *Signal Processing, IEEE Transactions on*, 44(5):1270–1275.
- [Shell, 1987] Shell, M. (1987). Us patent number: Usp004694686, cutting tool wear monitor. [online] <http://www.freepatentsonline.com>.
- [Shunmugam et al., 2000] Shunmugam, M. S., Reddy, S. V. B., and Narendran, T. T. (2000). Selection of optimal conditions in multi-pass face-milling using a genetic algorithm. *International Journal of Machine Tools and Manufacture*, 40(3):401–414.
- [Smid, 2003] Smid, P. (2003). *CNC programming handbook: a comprehensive guide to practical CNC programming*. Industrial Press Inc.
- [Song et al., 2011] Song, H., Miao, C., Wuyts, R., Shen, Z., D’Hondt, M., and Catthoor, F. (2011). An extension to fuzzy cognitive maps for classification and prediction. *IEEE Transactions on Fuzzy Systems*, 19(1):116 –135.

- [Soo et al., 2004] Soo, S. L., Aspinwall, D. K., and Dewes, R. C. (2004). 3D FE modelling of the cutting of inconel 718. *Journal of Materials Processing Technology*, 150(1-2):116–123.
- [Squillaro, 1993] Squillaro, Jerome C., M. J. C. L. (1993). Us patent number: Usp005266929, metal cutting tool wear indicator method and system. [online] <http://www.freepatentsonline.com>.
- [Stephenson, 2000] Stephenson, T. A. (2000). An introduction to bayesian networks theory and usage. [online] <ftp://ftp.idiap.ch/pub/reports/2000/rr00-03.pdf>.
- [Sutherland, 1988] Sutherland, J. W. (1988). A dynamic model of the cutting force system in the end milling process. *Diss. Abstr. Int.*, 49(1):293.
- [Tang and Guo, 2008] Tang, Y. and Guo, H. (2008). HSM strategy study for hardened die and mold steels manufacturing based on the mechanical and thermal load reduction strategy. *Journal of University of Science and Technology Beijing, Mineral, Metallurgy, Material*, 15(6):723–728.
- [Tansel et al., 1995] Tansel, I. N., Mekdeci, C., and Mclaughlin, C. (1995). Detection of tool failure in end milling with wavelet transformations and neural networks (WT-NN). *International Journal of Machine Tools and Manufacture*, 35(8):1137–1147.
- [Teti et al., 2010] Teti, R., Jemielniak, K., O'Donnell, G., and Dornfeld, D. (2010). Advanced monitoring of machining operations. *CIRP Annals - Manufacturing Technology*, 59(2):717–739.
- [Theodoridis and Koutroumbas, 2003] Theodoridis, S. and Koutroumbas, K. (2003). *Pattern recognition*. Academic Press.
- [Thomas et al., 2007] Thomas, S., Martinez, L. I., Westenberger, S., and Sturm, N. (2007). A population study of the minicircles in trypanosoma cruzi: predicting guide RNAs in the absence of empirical RNA editing. *BMC Genomics*, 8(1):133.
- [Torabi Amin J., 2009] Torabi Amin J., e. a. (2009). A survey on artificial intelligence technologies in modeling of high speed end-milling processes. In *Advanced Intelligent Mechatronics, 2009. AIM 2009. IEEE/ASME International Conference on*, pages 320–325.

- [Torabi Amin J., 2010a] Torabi Amin J., e. a. (2010a). Flute based analysis of ball-nose milling signals using continuous wavelet analysis features. In *Proceedings of 11th International Conference on Control, Automation, Robotics and Vision, ICARCV 2010*.
- [Torabi Amin J., 2010b] Torabi Amin J., e. a. (2010b). Fuzzy clustering of wavelet features for tool condition monitoring in high speed milling process. In *Proceedings of the Annual Conference of the Prognostics and Health Management Society 2010, Portland, Oregon, USA*.
- [Torabi Amin J., 2010c] Torabi Amin J., e. a. (2010c). Hidden markov model for ball-nose tool condition monitoring. In *Proceedings of the Postgraduate Student Conference AOTULE 2010, Indonesia, Bandung*.
- [Torabi Amin J., 2011] Torabi Amin J., e. a. (2011). Clustering methods for tool condition monitoring, repeatability and application. In *Submitted to IECON 2011, Melbourne Australia*.
- [Torabi Amin J., 2012] Torabi Amin J., e. a. (2012). Application of classical clustering methods for online tool condition monitoring in high speed milling processes. In *2012 7th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, pages 1249–1254.
- [Torabi Amin J., 2013a] Torabi Amin J., e. a. (2013a). Challenges of hidden markov models for modeling of ball-nose end-milling processes. *submitted to the Journal of Engineering Letters*.
- [Torabi Amin J., 2013b] Torabi Amin J., e. a. (2013b). A survey of modeling techniques for milling processes. *IEEE Systems Journal*.
- [Trahanias and Skordalakis, 1989] Trahanias, P. and Skordalakis, E. (1989). An efficient sequential clustering method. *Pattern Recognition*, 22(4):449–453.
- [Trahanias and Skordalakis, 1990] Trahanias, P. and Skordalakis, E. (1990). Syntactic pattern recognition of the ECG. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 12(7):648–657.
- [Uros et al., 2009] Uros, Z., Franc, C., and Edi, K. (2009). Adaptive network based inference system for estimation of flank wear in end-milling. *Journal of Materials Processing*

- Technology*, 209(3):1504–1511.
- [Vallejo et al., 2008] Vallejo, A. J., Morales-Menéndez, R., and Alique, J. R. (2008). On-line cutting tool condition monitoring in machining processes using artificial intelligence.
- [Vallejo and Morales-Menendez, 2009] Vallejo, A. J. and Morales-Menendez, R. (2009). Cost-effective supervisory control system in peripheral milling using HSM. *Annual Reviews in Control*, In Press, Corrected Proof.
- [Vapnik, 2000] Vapnik, V. N. (2000). *The nature of statistical learning theory*. Springer Verlag.
- [Varga and Moore, 1990] Varga, A. P. and Moore, R. K. (1990). Hidden markov model decomposition of speech and noise. In *proc. ICASSP*, volume 90, pages 845–848.
- [Venkatesh et al., 1997] Venkatesh, K., Zhou, M., and Caudill, R. J. (1997). Design of artificial neural networks for tool wear monitoring. *Journal of Intelligent Manufacturing*, 8(3):215–226.
- [Wang et al., 1998] Wang, L., Mehrabi, M. G., and Jr, E. K. (1998). Tool wear monitoring in reconfigurable machining systems through wavelet analysis. *Journal of Engineering for Industry*.
- [Wang et al., 2002] Wang, L., Mehrabi, M. G., and Kannatey-Asibu, J. (2002). Hidden markov model-based tool wear monitoring in turning. *Journal of Manufacturing Science and Engineering*, 124(3):651–658.
- [Wang, 1999] Wang, L.-X. (1999). *A Course in Fuzzy Systems and Control*. Prentice-Hall press, USA.
- [Wang et al., 1996] Wang, Z., Lawrenz, W., Rao, R. B. K. N., and Hope, T. (1996). Feature-filtered fuzzy clustering for condition monitoring of tool wear. *Journal of Intelligent Manufacturing*, 7(1):13–22.
- [Wang et al., 2005] Wang, Z., Rahman, M., Wong, Y., and Sun, J. (2005). Optimization of multi-pass milling using parallel genetic algorithm and parallel genetic simulated annealing. *International Journal of Machine Tools and Manufacture*, 45(15):1726–1734.

- [Wolbrecht et al., 2000] Wolbrecht, E., D’ambrosio, B., Paasch, R., and Kirby, D. (2000). Monitoring and diagnosis of a multistage manufacturing process using bayesian networks. *AI EDAM*, 14(01):53–67.
- [Wu and Er, 2000] Wu, S. and Er, M. J. (2000). Dynamic fuzzy neural networks-a novel approach to function approximation. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 30(2):358–364.
- [Wu et al., 2001] Wu, S., Er, M. J., and Gao, Y. (2001). A fast approach for automatic generation of fuzzy rules by generalized dynamic fuzzy neural networks. *IEEE Transactions on Fuzzy Systems*, 9(4):578–594.
- [Wu and Du, 1996] Wu, Y. and Du, R. (1996). feature extraction and assessment using wavelet packets. *Mechanical Systems and Signal Processing*, 10(1):29–53.
- [Xiaoli and Zhejun, 1998] Xiaoli, L. and Zhejun, Y. (1998). Tool wear monitoring with wavelet packet transform-fuzzy clustering method. *Wear*, 219(2):145–154.
- [Xu and Liang, 2001] Xu, Q.-S. and Liang, Y.-Z. (2001). Monte carlo cross validation. *Chemometrics and Intelligent Laboratory Systems*, 56(1):1–11.
- [Yamaguchi et al., 2007] Yamaguchi, T., Higuchi, M., Shimada, S., and Kaneeda, T. (2007). Tool life monitoring during the diamond turning of electroless Ni-P. *Precision Engineering*, 31(3):196–201.
- [Yan and Strenkowski, 2006] Yan, J. and Strenkowski, J. S. (2006). A finite element analysis of orthogonal rubber cutting. *Journal of Materials Processing Technology*, 174(1-3):102–108.
- [Yang and Chen, 2001] Yang, J. L. and Chen, J. C. (2001). A systematic approach for identifying optimum surface roughness performance in end-milling operations. *Journal of Industrial Technology*, 17(2).
- [Yang et al., 2006] Yang, L., Chen, J., Chow, H., and Lin, C. (2006). Fuzzy-nets-based in-process surface roughness adaptive control system in end-milling operations. *International Journal of Advanced Manufacturing Technology*, 28(3-4):236–248.
- [Yang, 1993] Yang, M. S. (1993). A survey of fuzzy clustering. *Mathematical and Computer Modelling*, 18(11):1–16.

- [Yu, 2010] Yu, S. (2010). Hidden semi-Markov models. *Artificial Intelligence*, 174(2):215–243.
- [Zain et al., 2008] Zain, A. M., Haron, H., and Sharif, S. (2008). An overview of GA technique for surface roughness optimization in milling process. In *Information Technology, 2008. ITSIm 2008. International Symposium on*, volume 4, pages 1–6.
- [Zain et al., 2009] Zain, A. M., Haron, H., and Sharif, S. (2009). Prediction of surface roughness in the end milling machining using artificial neural network. *Expert Systems with Applications*, In Press, Corrected Proof.
- [Zain et al., 2010] Zain, A. M., Haron, H., and Sharif, S. (2010). Prediction of surface roughness in the end milling machining using artificial neural network. *Expert Systems with Applications*, 37(2):1755–1768.
- [Zhai et al., 2011] Zhai, L. Y., Er, M. J., Li, X., Gan, O. P., and San, L. (2011). Wavelet-based analysis of force signals for monitoring and prediction of cutting performance in high-speed milling processes. 19(3).
- [Zhang et al., 2007] Zhang, J. Z., Chen, J. C., and Kirby, E. D. (2007). Surface roughness optimization in an end-milling operation using the taguchi design method. *Journal of Materials Processing Technology*, 184(1-3):233–239.
- [Zhao et al., 2010] Zhao, J., Zhang, X. F., and Luo, H. B. (2010). Simulation of Three-Dimensional dynamic cutting forces and chatter stability in Ball-End milling. *Key Engineering Materials*, 443:302–307.
- [Zhou et al., 2006] Zhou, J., Li, X., Gan, O. P., Han, S., and Ng, W. K. (2006). Genetic algorithms for feature subset selection in equipment fault diagnosis. In *Engineering Asset Management*, pages 1104–1113. Springer London, London.
- [Zhu et al., 2009] Zhu, K., Wong, Y., and Hong, G. (2009). Wavelet analysis of sensor signals for tool condition monitoring: A review and some new results. *International Journal of Machine Tools and Manufacture*, 49(7-8):537–553.
- [Zuperl and Cus, 2003] Zuperl, U. and Cus, F. (2003). Optimization of cutting conditions during cutting by using neural networks. *Robotics and Computer-Integrated Manufacturing*, 19(1-2):189–199.

- [Zuperl et al., 2006] Zuperl, U., Cus, F., Mursec, B., and Ploj, T. (2006). A generalized neural network model of ball-end milling force system. *Journal of Materials Processing Technology*, 175(1-3):98–108.
- [Zuperl et al., 2011] Zuperl, U., Cus, F., and Reibenschuh, M. (2011). Neural control strategy of constant cutting force system in end milling. *Robotics and Computer-Integrated Manufacturing*, 27(3):485–493.