Tool wear monitoring for face milling process with intelligent algorithms

Ensieh Sadat Hosseini Rooteh

A Thesis

in

The Department

of

Mechanical and Industrial Engineering

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Applied Science at

Concordia University

Montréal, Québec, Canada

September 2013

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared,

By : **Ensieh Sadat Hosseini Rooteh**

Entitled : **Tool wear monitoring for face milling process with intelligent algorithms**

and submitted in partial fulfilment of the requirements for the degree of

Master of Applied Sciences (Mechanical Engineering)

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the Final Examining Committee:

______________________________ Chair
Dr. R. Wuthrich

______________________________ Examiner
Dr. S. Engin

______________________________ Examiner
Dr. A. Aghdam External
Electrical and Computer Engineering

______________________________ Supervisor
Dr. Y. Zhang

Approved by:

____
Dr. S. Narayanswamy, MASc Program Director
Department of Mechanical and Industrial Engineering

Date: ____________________

____
Dean
Faculty of Engineering & Computer Science

ABSTRACT

Today, machining processes are among the most common and important industrial operations. Considering the wide applications of the machining processes, the importance of having an accurate and reliable monitoring system is clear. Therefore, several researches and studies have been performed to respond to this demand. Despite considerable efforts in this field, the achievement is not satisfactory and still the request for having an accurate, reliable, automatic, inexpensive and robust monitoring algorithm is remained without a good answer. This study tries to design monitoring algorithms with aforementioned specifications. The monitoring algorithms predict the amount of tool flank wear in the face milling process. They are designed based on the pattern recognition concept. The algorithms analyze signals in four steps: preprocessing, feature extraction, feature selection and classification. Descriptors, wavelet transform and S-Transform are applied for feature extraction. Principal component analysis (PCA) and independent component analysis (ICA) perform the feature selection step. Neural network (NN), which is an artificial intelligence method, classifies the data and makes the algorithms intelligent. By combining these methods, five intelligent algorithms are developed. The results show that the most accurate algorithm between these five algorithms is the combination of S-Transform, ICA and NN. Results also confirm the good performance of S-Transform for feature extraction comparing with the wavelet transform or descriptors. Applying the best designed algorithm, the effect of sensor fusion on the accuracy of algorithm and the ability of monitoring algorithm for working in different operating conditions are studied as well. It is also shown that the accuracy of the best designed algorithm for indicating the tool status, sharp or dull, is better than the accuracy of predicting the value of the tool wear. Applying S-Transform for machining monitoring and designing five intelligent, practical, inexpensive and accurate algorithms for tool wear prediction can be considered as the key outcomes of this thesis.

ACKNOWLEDGEMENTS

It would not have been possible to write this thesis without the help and support of the kind people around me, to only some of whom it is possible to give particular mention here.

Above all, I would like to thank my lovely husband and my dear parents for their endless support and great patience. My sisters and brothers, who have supported me throughout, as always, for which my mere expression of thanks likewise does not suffice.

I am grateful to my supervisor, Professor Youmin Zhang, for his support and guidance throughout my work. I highly appreciate him for being kind and considerate towards me.

Last, but not least, my special thanks to my friends, Nina, Iman, Javad, Abbas, Mahyar, Hadi, whom their help and encouragement motivated me during the course of this thesis work. I am also grateful to all those people who have given me support and I did not have the chance to thank them.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Introduction

1.1 Machining processes

Machining processes are cutting operations in which the shape of a stock changed with removing material and the desired part is produced. The objective of these processes is to produce parts with desired quality and as productively as possible [1]. Using machining processes has a long history and their role in manufacturing became more and more common and important during the time. Today, in different parts of industry many products are made with machining processes and many new demands are created in the market. Computer numerically controlled (CNC) machines do the machining processes automatically. The most common machining operations are turning, milling (face and end), and drilling:

- **Turning**: Turning, as shown in Fig. 1.1, is a material removal process, which is used to create rotational parts by cutting away unwanted material [3]. Turning operation is considered as the simplest machining operation. Unlike the other machining processes, turning cutters are single cutting edge and also they do not rotate.

- **Milling**: Face milling, as shown in Fig. 1.2, involves the use of multi-toothed rotating cutters to remove material from the surface of a workpiece [5]. End milling (Fig. 1.3) uses multi-tooth helical cutters to perform machining on the periphery of

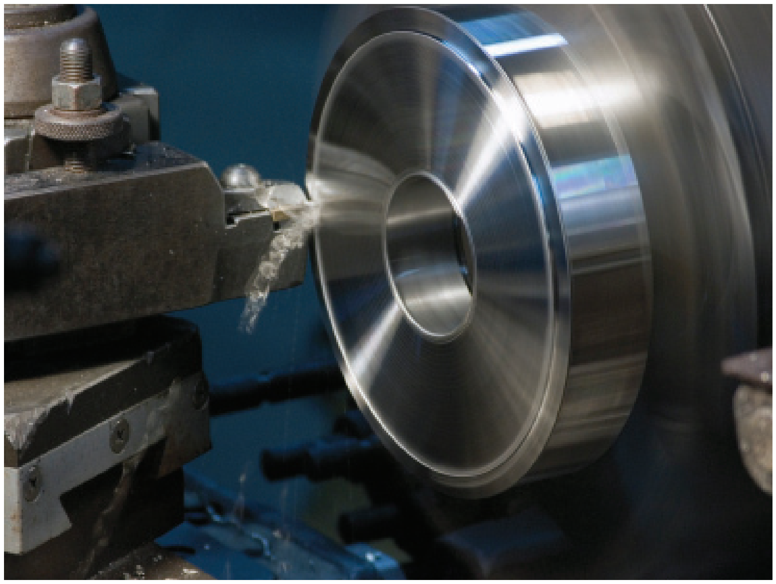

Figure 1.1: Turning operation [2].

workpieces [5]. End milling usually is in the final stage of the part production and commonly its duty is to finish the surface of the part. Accordingly, the accuracy of this operation is very important for production quality.

- **Drilling**: Drilling, as shown in Fig. 1.4, is a cutting process that uses a drill bit to cut or enlarge a hole in solid material [8]. This, in general, includes a complex three dimensional material removal operation [5]. Drilling process is one of the most commonly used machining processes in industry [9, 10]. For example, up to 50% of all machining operations in the U.S. involve drilling [11]. In addition, around 40% of metal removal operations in the aerospace industry involve the drilling process [12].

1.2 Machining process faults

Tool wear, tool breakage, and chatter are the most common faults happen during machining processes and affect the quality of productions.

- **Tool wear**: The changes in tool shape during the machining process which caused by contact between cutting tool and cutting surface or cutting tool and cutting chips,

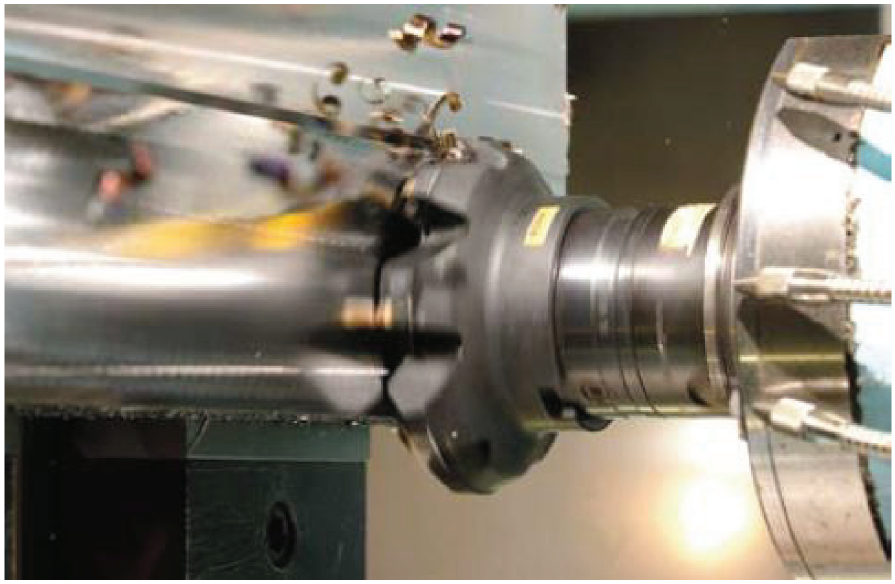

Figure 1.2: Face milling operation [4].

known as tool wear. This phenomenon changes the final workpiece dimension and quality. Also, it may lead to tool damage and long down time for the maintenance and increase the cost of production consequently [1]. Two most common tool wears are: flank wear and crater wear (Fig. 1.5). The moving chip which is in contact with the part creates crater wear and flank wear is occurred due to friction of the tool on the workpiece.

- **Tool breakage**: When a piece of cutting edge fractures abruptly during the machining process, a tool breakage happens. This phenomenon may happen due to thermal or mechanical overloading of cutting edge [14]. Tool failure makes 20% of the reported machine tools downtime and they increase the total production cost between 3% to 12% [15].

- **Chatter**: Chatter is the self-excited vibration of the machine tool that reflects the instability of the cutting process. Chatter is often a serious limitation for achieving higher rates of removal, as it adversely affects the surface finish (Fig. 1.6), reduces dimensional accuracy, and may damage the tool and machine. Therefore, machine tool chatter needs to be detected rapidly and corrected before it damages the workpiece,

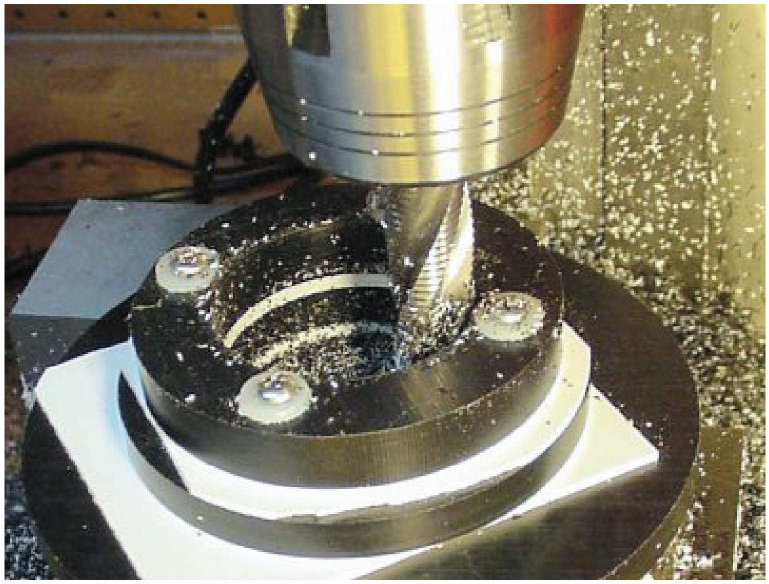

Figure 1.3: End milling operation [6].

tool, or the machine.

1.3 Machining process monitoring

In order to respond increasing demands of the market to have more productions in better quality, cheaper price and in the shorter time, manufactures are investigating innovative solution such as process monitoring. Process monitoring is indicating the state of the process with measuring parameters like force, vibration, sound, temperature etc. Process monitoring can decrease the cost of machining process by increasing the quality of products and reducing the tool breakage downtime and maintenance expenses.

Monitoring task has been doing with machine tool operators primarily. For example, they visually detect missing or broken tools and also chatter from the generated sound with the system. The recent monitoring methods are automated monitoring algorithms. These algorithms use filtered sensor measurements for determining the state of the process. Complex processes are monitored with the signal processing methods which analyze the recorded signals of the systems [1]. Artificial intelligence methods are among the

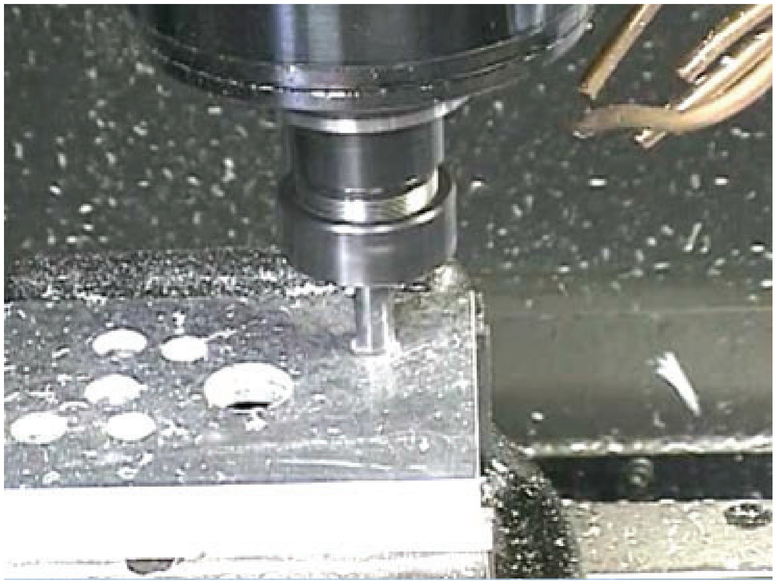

Figure 1.4: Drilling operation [7].

most common signal processing techniques which are implemented for process monitoring. They are mostly utilized for pattern recognition and for designing process monitoring algorithm [1].

In Pattern recognition concept, the monitoring task is performed with classifying data. The data is analyzed in several consecutive steps. These steps are preprocessing, feature extraction, feature selection, and data classification respectively. In preprocessing step, the recorded signals from machining operation is analyzed for removing and filtering noises. After preprocessing, to extract the information from raw data and reduced the dimension of them, each data will be transformed into a reduced representation called feature vector. This transformation process is called feature extraction. In the next step, if there are still redundant features in the feature vector which make the calculations time consuming and complex, the redundant features will be omitted from the feature vector with feature selection approach. Classification methods categorize feature vectors into the determined groups and accomplish the monitoring process. Artificial intelligence methods are often applied in classification step and make the monitoring algorithm intelligent. Figure 1.7 shows the monitoring algorithm which is designed based on the pattern recognition. Next

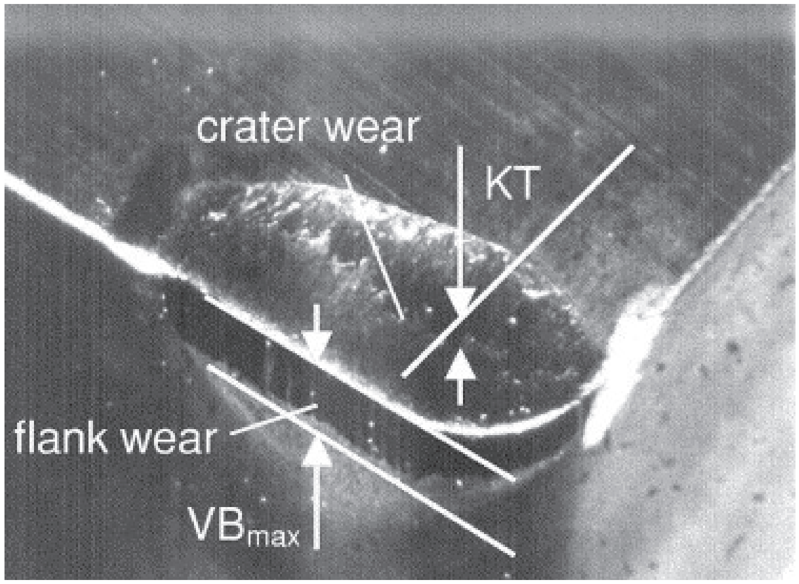

Figure 1.5: Flank and carater wear [13].

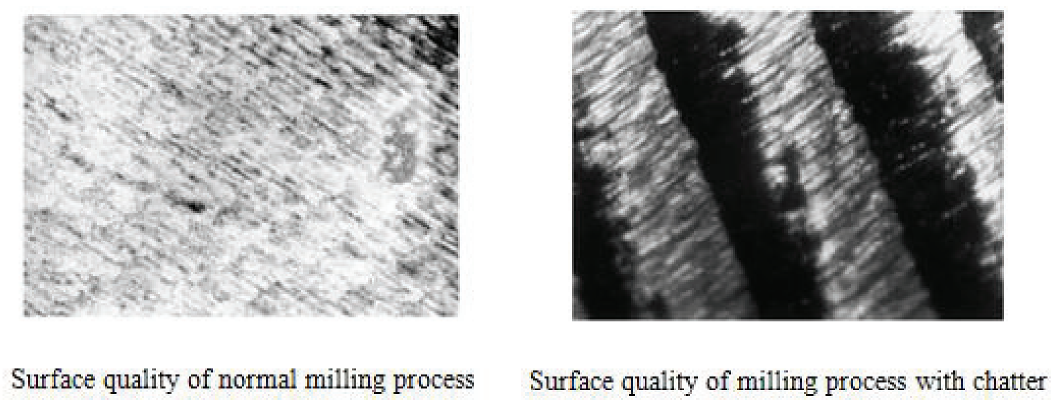

Figure 1.6: Chatter [14].

section reports a brief literature review on the process monitoring.

1.4 Literature review

In the field of process monitoring many studies have been conducted during the time. In many studies signals were analyzed with signal processing and artificial intelligence techniques and the results indicate the status of the systems. Cutting force, vibration, acoustic

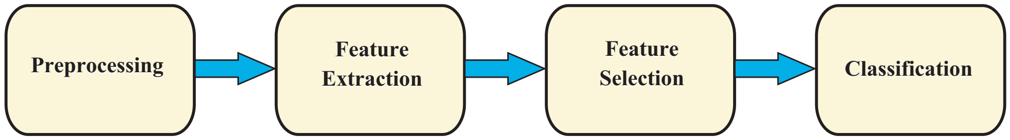

Figure 1.7: Monitoring algorithm.

Figure 1.8: Dynamometer [43].

emission (AE), current or power signals are the most common signals applied in these studies. Cutting force is considered to be the best variable to describe the cutting process [16]. The hidden information in force pattern is useful to evaluate the quality and geometric profile of the cutting surface [17]. Therefore, cutting-force has been used for tool condition monitoring [12, 15, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38] or surface roughness monitoring [39, 40, 41, 42]. Cutting forces increase when the tool gets dull. Therefore, tool wear can be easily detected by analyzing cutting force signals. The increase in cutting force due to the tool wear is dependent on the type of wear as well as the cutting conditions like cutting material, work material, etc. Dynamometer (Fig. 1.8) has been used to measure the cutting forces in literature. Nevertheless, it has some drawbacks which make its application limited in industry. First it is an expensive sensor and it can be easily destroyed in the rush environment of factories. Limited frequency response is another weakpoint which affects its application in industry [44, 45, 46]. Vibration monitoring is mainly applied for surface roughness prediction [47, 48, 49]. It has especially been used in turning process [47, 50, 51, 52] rather than milling process.Vibration monitoring also has wide application in cutting tool wear prediction. Vibration amplitude usually changes during the machining process due to change of tool condition from sharp to dull. Until a certain threshold for tool wear, it increases the frictional damping and decreases the vibrations. This threshold is dependent to the tool and workpiece material and the type of machining process. Beyond this threshold, tool wear increases the vibration. This increase is due to the stronger excitation caused by the larger cutting forces [53]. Vibration signals have also been used for tool breakage monitoring. Tool breakage changes cutting force and as a result it changes cutting vibrations. Therefore, vibration signals can be good variables to be analyzed for tool breakage monitoring. Accelerometer (Fig. 1.9) is the sensor for recording vibration signals. Although vibration signals have wide application in process monitoring and accelerometer has been applied to the monitoring systems successfully, there are a number of practical problems for using vibration signals for process monitoring or measuring the part accuracy [55]:

- The machining speed should be in a specific range.
- The amplitude of the signal is depending on the sensor distance from cutting edge. Also mounting sensor close to the cutting location increases the variability of the signal.
- The rush producing environment can strike the accelerometer and damage it or cause it to read the vibration inaccurately.

Vibration signals also are not accurate and reliable as much as other signals like cutting forces or AE signals [55]. The recorded signals with dynamometer and AE sensors are more robust for monitoring surface roughness and production quality. However, the advantage of simplicity and low cost for monitoring systems based on accelerometers have made them popular.

AE is the energy released in the form of mechanical vibration from a material when it is under stress. In machining operations this energy can be released from tool, workpiece or machine body. The stress may be generated by chip deformation, chip fracture, friction

Figure 1.9: Accelerometer [54].

existed between chip, workpiece and tool (two by two), tool breakage, flexible deformation of machine structures and thermal reactions of materials [56]. The frequency spectrum of AE usually spans the range $10kHz - 10MHz$ [56]. AE sensors (Fig. 1.10) are inexpensive and easy to install and they also have another positive point related to the frequency range of AE signals. The frequency range of AE signals is much higher than that of machine vibrations and environmental noises, which make it easy to distinguish AE signals [58]. In addition to these advantages, it has also some disadvantages which make its application limited. AE signals are sensitive to the sensor location and cutting parameters and also, it is hard to have physical understanding of these signals [59]. Sensory overload can considerably distort the AE signal, therefore these sensors need to be calibrated carefully for the range of cutting operation [60]. In the literature there is an argument about using AE sensors for tool wear diagnosis and replacing the dynamometer with this sensor. Although there are disagreement about tool wear, they all agree that AE sensors are good choice for cutting tool breakage detection [61, 62, 63]. During tool breakage and fracture large amount of AE is generated, therefore this signal can be used for cutting tool breakage detection.

The other sensors which have been used in literature are current sensors. They can measure machining forces indirectly. Current signals are proportional to the torque produced with DC motor which is proportional to the cutting forces. The limitations for using

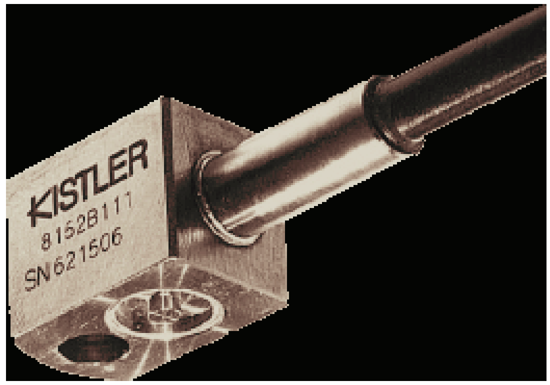

Figure 1.10: AE sensor [57].

current sensors are their nonlinear nature and limited sensing bandwidth. Nonlinearity imposes complicated calibration for using these sensors [56]. Due to limited sensing bandwidth, these sensors act as a low pass filter and they are suitable only for detecting slow events or when a fast response is not essential [45].

In literature different signal processing and artificial intelligence methods have been implemented for analyzing the recorded signals. For preprocessing analogue or digital filtering and signal segmentation have been used. Filters in sensors suppress high frequency noise or continuous biases and keep the recorded signal within the range of the frequency response of the sensor. Digital filtering keeps the sensor information which best correlates with the performance process variable of interest [64, 65]. Signal segmentation is for extracting the signal information when the process is in steady state. Segmentation is an optional step in preprocessing [16, 65]. Other signal processing methods in time or frequency or time-frequency domain have been used for feature extraction. Descriptors are the simplest features which have been used in some studies. Mean, root mean square (RMS), power spectral density (PSD) are some of these descriptors. Methods for time-frequency analysis are fast Fourier transform (FFT), short time Fourier transform (STFT), wavelet transform and S-Transform. Time-frequency methods can reveal the information or features of the signals which are hidden for other time–domain signal analysis. FFT is a common method of signal processing in process monitoring research [66, 67]. The averaging natures of the FFT calculations can distort its signal analysis performance when it is applied for analyzing signals from transient phenomenon like tool breakage [5]. Wavelet transform has been used in process monitoring widely [61, 63, 68, 69, 70, 71, 72, 73]. Continuous and discrete wavelet transform have been used for data compression filtering or for feature extraction. Although many studies have applied wavelet transform for feature extraction, few studies have implemented STFT and S-Transform [74]. These methods have not been investigated in the field of machining process monitoring while they receive much more attention in other field of study like medical diagnosis or electrical system monitoring.

Principal component analysis (PCA) is one of the most common methods for feature selection in process monitoring [75] and has many applications in different fields of study. Neural network, fuzzy logic, hidden Markov model (HMM) and support vector machine have been applied for classifying data and predicting the status of process in many studies. The main artificial intelligence techniques applied for monitoring machining processes are artificial neural networks (ANN) [21, 22, 23, 24, 25, 26, 28, 34, 35, 47, 67, 72, 76, 77], fuzzy logic systems [21, 24, 47, 76, 78] and neuro-fuzzy inference systems [79, 80]. Other artificial intelligence approaches such as Baysian network (BN) [18, 30, 81], hidden Markov models [12, 29, 31, 32, 33, 82, 83], evolutionary algorithms [84, 85] or support vector machines (SVM) [25, 71, 86, 87, 88] are gaining popularity in recent works, although they have not been widely used before.

As explained before process monitoring and fault diagnosis in the field of machining process has long history and many studies have been conducted in this field. However, there are still many unsolved challenges and problems that the researchers have struggled with them. Many studies are doing process monitoring indirectly. It means that they measure one parameter to calculate or predict the other parameter that they cannot measure directly. For example, they measure cutting force signals and analyze them to predict the value of tool wear. These indirect techniques have their own drawbacks. They are not reliable and robust due to the intricacies which the machining processes have and the uncertainty in the correlation between the process parameters and tool wear [89]. In monitoring approaches the sensor is key element for recording data. Sensors can rarely be placed at the desired point, and when located at other locations their performance in recording clear signals for reliable monitoring decreases [1]. Frequency response limitations of some sensors are the other difficulties in applying sensors. This limitation becomes more obvious when the sensor is implementing for chatter detection [1]. The other problem is related to the robustness against varying cutting conditions. The single sensor measurements do not seem to be robust in this case [1]. In some research, for analyzing the recorded data and making decision about the process status, the recorded data is comparing with an expected threshold. Analytical model are used to specify these thresholds. Analytical models have their own advantages and disadvantages. They can calculate the threshold with considering changes in the machine inputs such as feeds and speeds. However, they are often not accurate and need to be calibrated for the process. Calculating the expected value of measurements with empirical methods is simpler and more straightforward, although they would be only suitable for particular operations and can not be extrapolated to others [1]. The limitations of signal processing and artificial intelligence methods which have been used in literature also should be considered. Some of these methods need time consuming training or huge dataset for training. Also, some process monitoring algorithms proposed in literature are dependent on the operating conditions. It means that they just work in special cases or they have to be trained for any single case. Complexity of machining process is another barrier in machining operation monitoring. For example, milling process monitoring involved with many difficulties. Complex geometry of milling tool, the interrupted nature of milling process, and the large periodic components of force and torques make milling process monitoring complicated [61].

1.5 Thesis objectives

According to the above literature review, although many studies have been carried out in the field of process monitoring, there are still many unsolved problems in this field of study. Considering these shortages and at the same time by increasing the demand for having reliable, robust, accurate, practical and inexpensive monitoring methods from the industry, the need for doing more research in this field is obvious. This research is going to design a process monitoring algorithm which is more accurate, practical to be used in industry, intelligent, independent from operating conditions and inexpensive. Such an objective is to be achieved by investigating and developing effective monitoring algorithms with using signal processing and artificial intelligence methods, based on the pattern recognition. Among the signal processing methods, S-Transform received poor attention in the field of machining monitoring. In this thesis machining monitoring algorithm with assisting S-Transform for feature extraction will be designed and the accuracy of algorithm for predicting the value of tool wear will be calculated. The accuracy of this algorithm will be compared with the accuracies of other algorithms which use the conventional signal processing methods. In this way the effect of the S-Transform on the accuracy of algorithm will be investigated and the most accurate algorithm will be selected for doing some further investigations. The effect of sensor fusion on the accuracy of algorithms, the ability of the algorithms for working in different operating conditions and for detecting the sharp tool from the dull one will be investigated through the rest of the thesis.

1.6 Thesis organization

In this thesis, first a brief explanation about the experimental setup which provides the data for training the algorithms will come in chapter two. Then in chapter three, the signal processing and artificial intelligence methods implemented in this research will be explained. Chapter four will introduce five monitoring algorithms and with comparing the results, the most accurate monitoring algorithm will be selected. It also will discuss about the effect of S-Transform method on the accuracy of monitoring algorithm. In chapter five the effect of sensor fusion on the accuracy of algorithm and the ability of algorithm for working in variant operating conditions will be discussed. The accuracy of algorithm for detecting the sharp tool from dull one is another topic which will be investigated in chapter five. Conclusions and suggestions for the possible research works in future come in chapter 6.

Chapter 2

Experimental dataset

In this study the experimental data has been downloaded from the NASA website [90]. This dataset includes the recorded signals from runs on a milling machine in different operating conditions and also flank wear of tools during the recording signals have been measured. The data has been sampled with three sensors: current, acoustic emission and vibration sensors. In each cut, tool wear has been measured. The recorded data has been organized in a MATLAB struct array which is explained in the Table 2.1.

Table 2.2 shows variable parameters: depth of cut, feed and working material. These 16 cases represent 8 operating conditions: every two cases are assigned to one set of operating conditions. For example cases 1 and 9 have the same operating conditions.

2.1 Experimental setup

Machining center used for making this dataset is a Matsuura MC-510V. The spindle and table of this machining center make the basic setup of the experiment. Each table and spindle of this machining center has two sensors: acoustic emission sensor and vibration sensor. Also one current sensor records the spindle current signals.

The parameters of the experiment have been specified based on the industrial applicability and recommended manufacturers settings. Cutting speed was selected as 200 m/min

Table 2.1: Struct field names and description [90].

Field name	Description
Case	Case number (1-16)
run	Counter for experimental runs in each case
VB	Flank wear, measured after runs
time	Duration of experiment (restarts for each case)
DOC	Depth of cut (does not vary for each case)
feed	Feed (does not vary for each case)
material	Material (does not vary for each case)
smcAC	Spindle motor alternating current (AC)
smcDC	Spindle motor direct current (DC)
vib_table	Table vibration
vib_spindle	Spindle vibration
AE_table	Acoustic emission at table
AE_spindle	Acoustic emission at spindle

(or 826 rev/min). The selected values for depth of cut, feed and material are as given in Table 2.2. These choices make 8 different settings. All experiments have been done twice and the second time with the same parameters and with a second set of inserts. The size of the workpieces are 483 $mm \times 178$ $mm \times 51$ mm. Cutting tool was chosen as a 70 mm face mill with 6 inserts (Fig. 2.1). Based on the recommendation for roughing (Kennametal, 1985), the inserts KC710 was selected. They are coated with multiple layers of titanium carbide, titanium carbonitride, and titanium nitride (TiC/TiC-N/TiN) in sequence. These layers improve resistance to cratering and edge wear and they have the advantages of titanium carbide plus reduced face friction.

2.2 Tool wear

As a generally accepted parameter for evaluating tool wear, flank wear (VB) is used as a measure for tool wear. The VB is defined as the distance from the cutting edge to the end of the abrasive wear on the flank face of the tool. During the experiment the flank wear was measured with the help of microscope. For measuring the tool wear the insert was taken

Table 2.2: Experimental conditions [90].

Case	Depth of Cut (mm)	Feed (mm/rev)	Material
1	1.5	0.5	cast iron
2	0.75	0.5	case iron
3	0.75	0.25	cast iron
4	1.5	0.25	cast iron
5	1.5	0.5	steel
6	1.5	0.25	steel
7	0.75	0.25	steel
8	0.75	0.5	steel
9	1.5	0.5	cast iron
10	1.5	0.25	cast iron
11	0.75	0.25	cast iron
12	0.75	0.5	cast iron
13	0.75	0.25	steel
14	0.75	0.5	steel
15	1.5	0.25	steel
16	1.5	0.5	steel

out of the tool (Fig. 2.2).

2.3 Recorded signals

For each run of milling operation the signals have been recorded with six sensors located in six different positions. Fig. 2.3–2.6 show a sample of the recorded signals from these sensors.

As it can be seen in the figures the data recorded in entry and exit area of the signals are noisy. They suffer from the entry and exit non-stable conditions. Therefore, in this thesis the recorded data in the area of sample number 3000 to 6000 will be used for calculations.

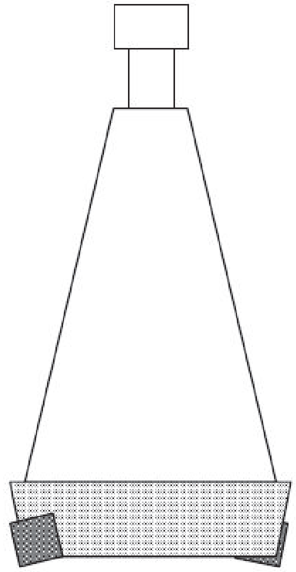

Figure 2.1: Tool and inserts of face mill [90].

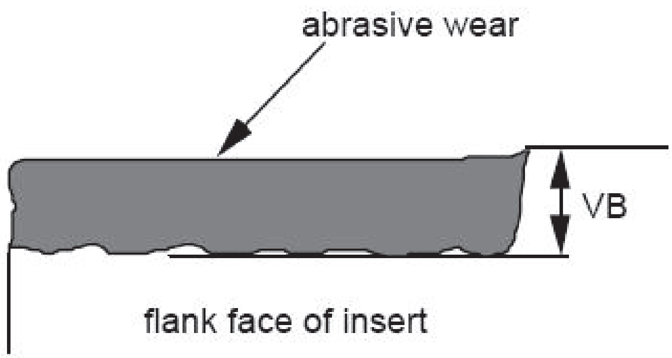

Figure 2.2: Tool flank wear (VB) as it is seen on the insert [90].

2.4 Chapter summary

In this chapter the experimental dataset applied in this study has been introduced. The characteristic of machining center, cutting tool, sensors and 8 operating conditions for recording signals were reported.

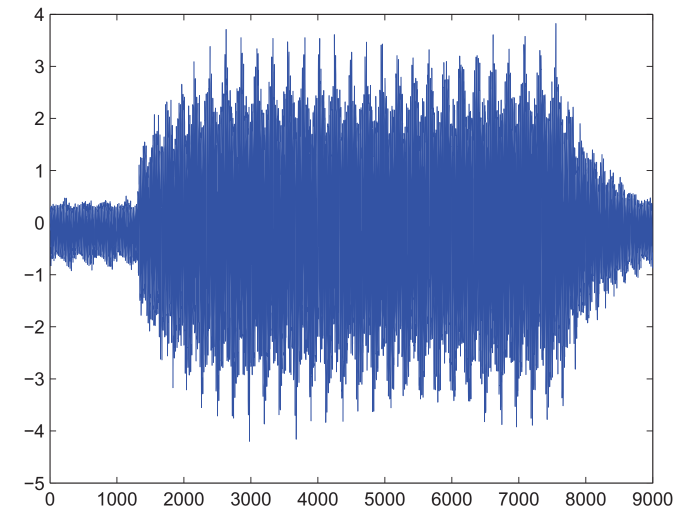

Figure 2.3: Spindle motor current signal (AC).

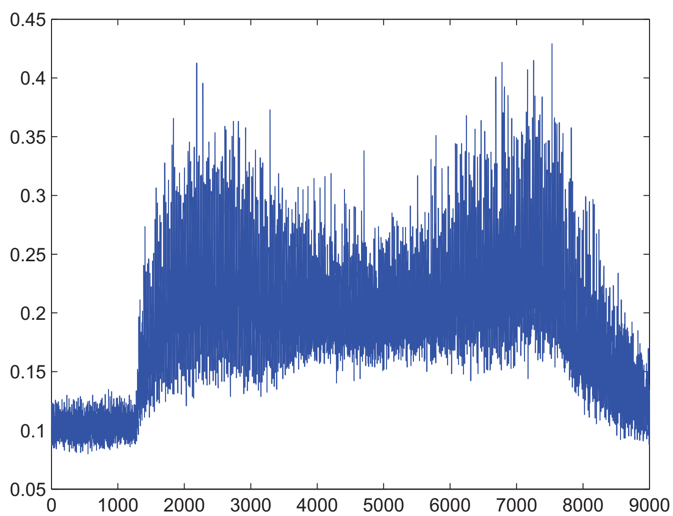

Figure 2.4: Acoustic emission signal (spindle).

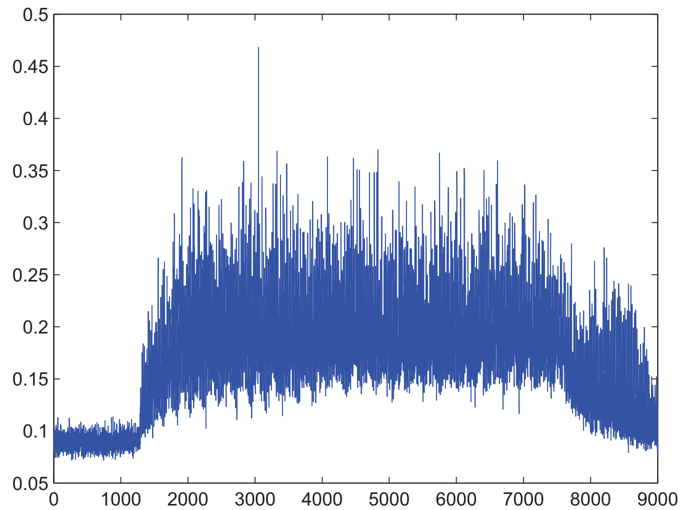

Figure 2.5: Acoustic emission signal (table).

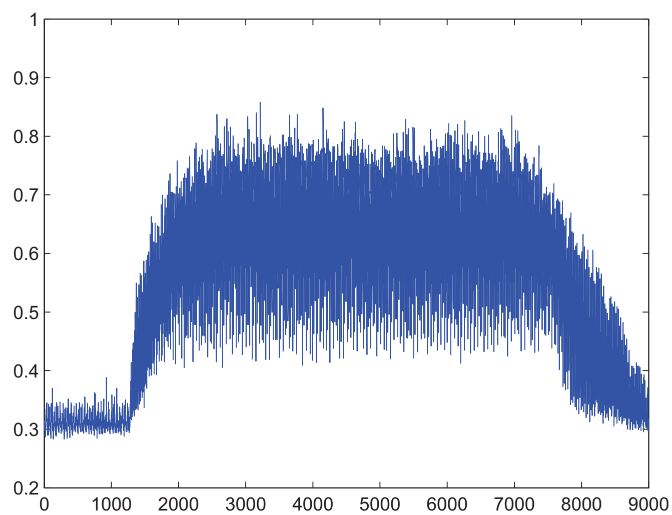

Figure 2.6: Vibration signal (spindle).

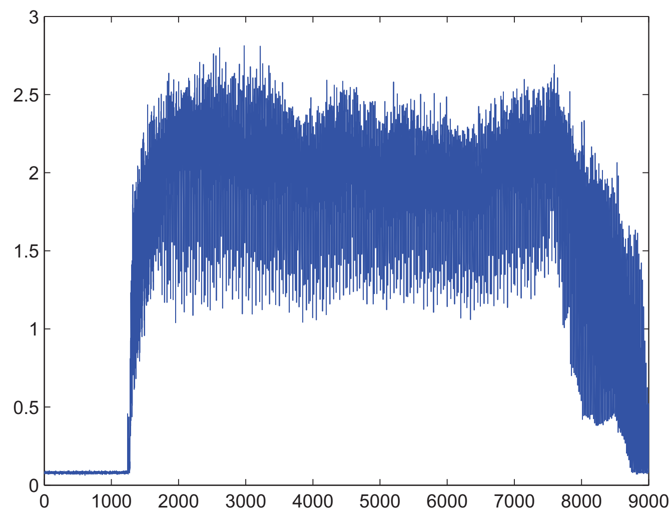

Figure 2.7: Vibration signal (table).

Chapter 3

Signal processing and artificial intelligence methods

The applied methods for data analysis in this study will be introduced in this chapter briefly. The methods are categorized based on their application for pattern recognition in monitoring algorithm. Preprocessing, feature extraction, feature selection and classification are the steps of monitoring algorithm. The dataset which has been utilized in this study, includes preprocessed data. Therefore, there is no need for doing the preprocessing again. In the following sections, other steps and the methods which will be implemented for accomplishing those steps will be explained respectively.

3.1 Feature extraction methods

Feature extraction methods transform the data into a reduced representation of them called feature vector. Working with feature vectors, instead of the original data, is less time consuming and more effective to have accurate results. Many methods have been innovated and applied for doing feature extraction. Statistical descriptors, continuous wavelet and S-Transform which will be implemented in this study for feature extraction, are explained in the following sections.

3.1.1 Descriptors

Descriptors summarize large, complex datasets (numerically and visually) to express their essence to the data analyst and to make further processing possible [91]. There are descriptors which measure the central tendency or dispersion of the data. A measure of central tendency is a typical value around which other figures congregate [92]. It is occasionally called an average or just the center of the distribution [93]. Geometric mean, harmonic mean, arithmetic average or mean, 50^{th} percentile or median, trimmed mean, most frequent value or mode are descriptors which measure the central tendency. These descriptors indicate the location of the center of the distribution, but they do not expose how the items are spread out on either side of the center. This characteristic, frequency distribution, is called dispersion. Small dispersion shows high uniformity of the data, while large dispersion indicates less uniformity [92]. In the following, some descriptors which measure the central tendency and also some descriptors for evaluating dispersion are introduced.

- **Measure of central tendency**:
  - **Geometric mean**: The geometric mean of a dataset which contains n observations is the n^{th} root of the product of the values. If $x_1, x_2, \cdots, x_n$ are observations then [92]:

$$GM = \sqrt[n]{x_1 x_2 \cdots x_n}. \tag{3.1}$$

  - **Harmonic mean**: Harmonic mean is defined as the reciprocal of the arithmetic average of the reciprocal of the data. If $x_1, x_2, \cdots, x_n$ are observations,

$$HM = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}, \tag{3.2}$$

for a frequency distribution [92],

$$HM = \frac{N}{\sum\limits_{i=1}^{n} f(\frac{1}{x_i})}. \quad (3.3)$$

- **Arithmetic average (mean)**: Arithmetic mean is defined as the sum of the data divided by the number of them. If the variable x assumes n values x_1, x_2,$\cdots$, x_n then the $\bar{x}$ is given by [92],

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i. \quad (3.4)$$

- **50$^{\text{th}}$ percentile (median)**: Median ($\tilde{X}$) is defined by

$$\tilde{X} = \begin{cases} X_{k+1} & \text{for odd } n = 2k+1, \\ \dfrac{X_k + X_{k+1}}{2} & \text{for even } n = 2k, \end{cases} \quad (3.5)$$

to be the middle or the average of the two middle data values after sorting. It is the 50$^{\text{th}}$ percentile using the percentile definition. Median is greater than or equal to at least 50% of data and also is less than or equal at least 50% of data. The positive point of median is that it is not affected by outlier values and is stable [94].

- **Trimmed mean**: The idea for designing r-trimmed mean is to protect the mean value against a few outliers and it is defined by

$$T_r = \frac{X_{r+1} + X_{r+2} + \ldots + X_{n-r-1} + X_{n-r}}{n-2r}. \quad (3.6)$$

Equation3.6 trims r outliers and averages the remaining values [94].

- **Most frequent value (mode)**: Mode is the value in a distribution, which happens most frequently. Mode has the most concentration of data around it [92]. It is more useful with discrete or coarsely rounded data [91].

Table 3.1: First three moment about the mean.

	Individual series
First moment about the mean, μ_1	$\frac{\sum(x-\bar{x})}{n} = 0$
Second moment about the mean, μ_2	$\frac{\sum(x-\bar{x})^2}{n} = \sigma^2$
Third moment about the mean, μ_3	$\frac{\sum(x-\bar{x})^3}{n}$

- **Measure of dispersion**:
  - **Interquartile range (IQR)**: This range is defined by the difference of the third and first quartiles,

$$IQR = Q_3 - Q_1. \tag{3.7}$$

    The larger value of IQR means that the data values are more dispersed [94].

  - **Mean absolute deviation (MAD)**: The mean absolute deviation about the sample median is defined by

$$MAD = \frac{1}{n}\sum_{i=1}^{n} |X_i - \tilde{X}|. \tag{3.8}$$

    This definition also can be written with sample mean $\bar{X}$ too. In this case the measure would be larger [94]. MAD is also sensitive to outliers but it does not suffer as much as standard deviation or variance [91].

  - **Central moment of all orders (moment)**: Central moments are the arithmetic mean of various powers of deviations taken from the mean of a distribution [92]. Table 3.1 shows the first three moments.

  - **Range (R)**: Range is the difference of the largest and smallest data. In terms of the order statistics [94]:

$$R = X_n - X_1. \tag{3.9}$$

  - **Variance**: Variance is calculated by the Eq. 3.10. This descriptor is the most commonly used value to measure the spread of data [94].

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}. \tag{3.10}$$

- **Standard deviation (std)**: The square root of the variance is called the sample standard deviation [94].

3.1.2 Wavelet transform

Wavelet was first introduced by Morlet [95]. Wavelets are mathematical functions that cut up data into different frequency components, and then study each component with a resolution matched to its scale. Continuous wavelet transform and discrete wavelet transform are two types of wavelet transform. Morlet formulized the continuous wavelet transform (CWT), as shown by Eq. 3.11:

$$W_x(a,b,\psi) = a^{-\frac{1}{2}} \int x(t)\psi^*(\frac{t-b}{a})dt, \tag{3.11}$$

where a is the scale parameter, b is the time parameter, $\psi(t)$ is an analyzing or mother wavelet, and $\psi^*(t)$ is the complex conjugate of $\psi(t)$. For $|a| < 1$, the wavelet is compressed version of the mother wavelet and is related mainly to higher frequencies, while for $|a| > 1$ the output of wavelet transform has a larger time width than the mother wavelet and correspond to lower frequencies. Therefore, wavelet has a time width which is adapted to their frequencies. This characteristic of wavelet is the main reason for the success of the Morlet wavelets in signal processing and time-frequency signal analysis. The resolution of the wavelet at different scales varies in the time and frequency domains. The resolution is coarse in the time domain and fine in frequency domain at large scale. When the scale a is decreasing, the resolution in the time domain becomes finer while the corresponding value in the frequency domain becomes coarser [95]. Figure 3.1 shows a signal with its

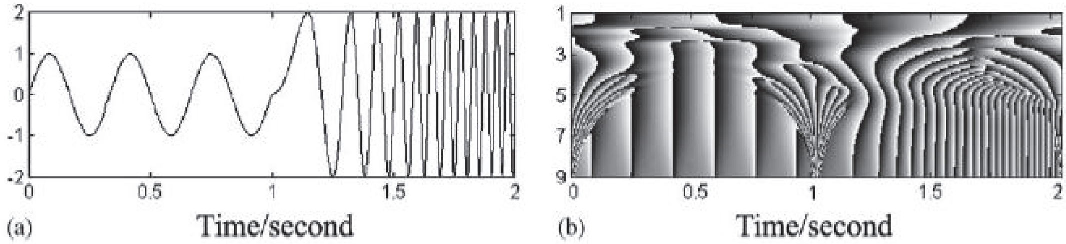

Figure 3.1: Continuous wavelet transform: A typical signal (a) and its wavelet phase map (b) (at the time of 1 s, the signals frequency changed and started to increase with time) [96].

continuous wavelet transform [96].

Later, Daubechies with Mallat [97, 98] developed wavelet from continuous to discrete signal analysis. In the discrete wavelet transform (DWT), the scale a and the time b are discretized as following:

$$a = a_0^m,\ b = na_0^m b_0, \tag{3.12}$$

where m and n are integers. The discrete $\psi_{a,b}(t)$ is:

$$\psi_{m,n}(t) = a_0^{-\frac{m}{2}} \psi(a_0^{-m} t - nb_0). \tag{3.13}$$

The new definition of discrete wavelet transform with the discretization of the scale parameter and time parameter is:

$$W_x(m,n,\psi) = a_0^{-\frac{m}{2}} \int x(t) \psi^*(a_0^{-m} t - nb_0) dt. \tag{3.14}$$

Discrete wavelet transform also can be considered as a high and low pass filter of discrete time domain signal (Fig. 3.2). This algorithm is known as Mallat algorithm or Mallat tree decomposition. $x[n]$ is discrete time domain signal and G_0 is low pass filter and H_0 is high pass filter. At each decomposition level, H_0 produces the detailed information, $d[n]$, and G_0 which is related to the scale, produces the coarse approximation, $a[n]$ [99]. Wavelet transform has different applications. It can be used for multi-scale analysis of

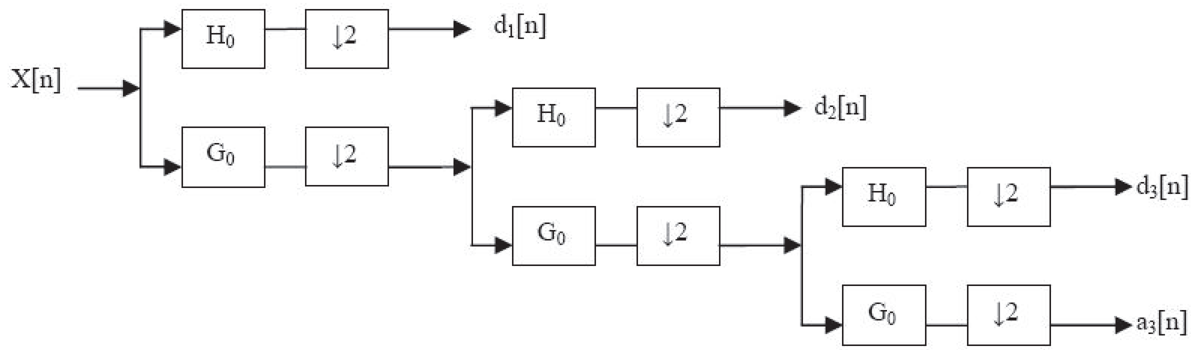

Figure 3.2: Three-level wavelet decomposition tree [99].

a signal through dilation and translation, so it can extract time-frequency features of a signal effectively [27]. Therefore, the wavelet transform is suitable for the analysis of non-stationary signals. Also due to the compact support of the basic functions used in the wavelet transforms, wavelets have good energy concentration properties and they can be implemented for fault feature extraction. Other applications of wavelet transform are signal denoising, system and parameter identification and signal compression.

Wavelet also has some shortcomings. It always suffers from the effects of the border distortion and energy leakage. Also, the phase spectrum of the wavelet is not robust to the noise, therefore once a signal is contaminated by noise, its phase spectrum will change greatly [67]. Additionally, since the definition of wavelet transform is based on the convolution, the occurrence of the overlapping is expected. The overlapping will cause undesirable frequency aliasing and bring the interference terms to the scalograms under certain conditions [67].

3.1.3 S-Transform

S-Transform combines the short-time Fourier transform and the wavelet transform using a Gaussian window [28] and it overcomes some of their disadvantages. Eq.3.15 shows the definition of S-Transform:

$$S(\tau,f)=\int_{-\infty}^{\infty}h(t)\frac{|f|}{\sqrt{2}\pi}e^{\frac{-(\tau-t)^2f^2}{2}}e^{-ift2\pi}dt. \tag{3.15}$$

The S-Transform can be viewed from two different perspectives: as a short-time Fourier transform with a variable window length, or as a special type of CWT with a Gaussian mother wavelet modified by adding a phase factor [28]. The Gaussian window is chosen because it is the most compact in time and frequency. In fact, S-Transform is a special case of the multiresolution Fourier transform [12]. S-Transform is a generalization of the Fourier transform to nonstationary time series. The S-Transform localizes the real and imaginary components of the spectrum independently, localizing the phase spectrum as well as the amplitude spectrum.

The discrete S-Transform is calculated by taking the advantage of the efficiency of the FFT and the convolution theorem [12]:

$$S[jT,\frac{n}{NT}]=\sum_{m=0}^{N-1}H[\frac{m+n}{NT}]e^{\frac{-2\pi^2m^2}{n^2}}e^{\frac{imj2\pi}{N}} \quad (n\neq 0). \tag{3.16}$$

For $n=0$, it is equal to the constant defined as:

$$S[jT,0]=\frac{1}{N}\sum_{m=0}^{N-1}h[mT], \tag{3.17}$$

where j, m and $n=0,1,...,N-1$. The sampling of the S-Transform is such that $S[jT,\frac{n}{NT}]$ has a point at each time sample and at each Fourier frequency sample. Before using S-Transform some preprocessing should be done on actual data: detrending (removing the very low harmonics), removing edge effects, using the analytic signal, estimating the noise level and resolution change [100]. S-Transform has the following unique properties [29]:

- A direct connection to the Fourier transform through the inverse S-Transform, and a direct connection to the time series through the direct signal extraction.

- Frequency invariant amplitude response.
- Absolutely referenced phase properties.
- Progressive resolution.

Absolutely referenced phase information and also frequency invariant amplitude response are the weak points for continuous wavelet transform. S-Transform has some drawbacks too. It requires higher complexity computation. The discrete S-Transform suffers from the problems related to sampling and finite length, giving rise to implicit periodicity in the time and frequency domains [74].

S-Transform can be implemented for signal analysis, power system disturbance recognition, system identification, fault diagnosis and condition monitoring.

3.2 Feature selection methods

Applying the most meaningful features in training monitoring algorithm, develops robust and reliable models for monitoring [101]. These features are selecting through feature selection techniques. Feature selection is the process of selecting a subset of relevant features for using in model construction. Feature selection eliminates redundant or irrelevant features. Irrelevant features provide no useful information in any context and redundant features do not have more information than the currently selected features. Feature selection is a part of feature extraction which is a more general concept compare with feature selection. Feature extraction creates new features from the original data, whereas feature selection returns a subset of the features [102]. In this study principal component analysis (PCA) and independent component analysis (ICA) are implemented for feature selection.

3.2.1 Principal component analysis (PCA)

Karl Pearson proposed principal component analysis (PCA) in 1901 [103] to analyze principal axes theorem in mechanics; later in 1930 Harold Hotelling developed and introduced PCA independently [104]. PCA is a mathematical procedure which applies an orthogonal transformation to convert the dataset. It changes the dataset which contains a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. This transformation also reduces the dimension of the dataset; the number of principal components is less than or equal to the number of original variables. This transformation organizes principal components based on the value of variance. Principal component with the highest variance is the first component and other components (which are orthogonal to each other or uncorrelated) will come after that based on their variance values. PCA can be calculated either by eigenvalue decomposition of a data covariance (or correlation) matrix or singular value decomposition (SVD) of a data matrix. Before doing PCA, usually mean centering (and normalizing) the data matrix is carried out [105]. Today, there are efficient ways for calculating SVD, therefore, it becomes the standard method for calculating PCA from a dataset now. In this formulation SVD of X (a n by p matrix which represent the dataset) expresses this matrix as follows (all the formulation comes from [106]):

$$X = U\Sigma W^T, \tag{3.18}$$

in which Σ is called the singular values of X. It is a p-by-p diagonal matrix of positive numbers $\sigma(k)$. U, left singular vectors of X, is an n-by-p matrix, the columns of which are orthogonal unit vectors of length n. W, right singular vectors of X, is a p-by-p matrix and its columns are orthogonal unit vectors of length p. In terms of this factorization, the $X^T X$, which is the covariance matrix of X, can be written as:

$$X^T X = W\Sigma U^T U\Sigma W^T = W\Sigma^2 W^T. \tag{3.19}$$

Comparing with the eigenvalue and eigenvector of $X^T X$, the right singular vectors W of X are equivalent to the eigenvectors of $X^T X$, and the singular values $\sigma(k)$ of X are equal to the square roots of the eigenvalues of $X^T X$.

Using the singular value decomposition the score matrix T, the full principal component decomposition of X, can be written as:

$$T = XW = U\Sigma W^T W = U\Sigma, \tag{3.20}$$

so each column of T is made by multiplying the left singular vectors of X and the corresponding singular value. If the first L largest singular values and their singular vectors were considered:

$$T_L = U_L \Sigma_L = X W_L. \tag{3.21}$$

The reduced matrix T_L would be the nearest possible matrix with rank L to the original matrix. This matrix between all the transformed data matrices with only L columns, maximizes the variance in the original data that has been preserved, while minimizing the total squared reconstruction error $||T - T_L||^2$.

Such dimensionality reduction can be a very useful step for processing high dimensional datasets, while still retaining as much of the variance in the dataset as possible. PCA has application in many fields, for example, clustering, regression analysis, denoising and pattern identifying. In clustering, PCA reduces the dimension of dataset and changes the dataset such as the cluster gets more visible and distinguishable. In regression analysis, it reduces the number of explanatory variable and by this way it reduces the chance of

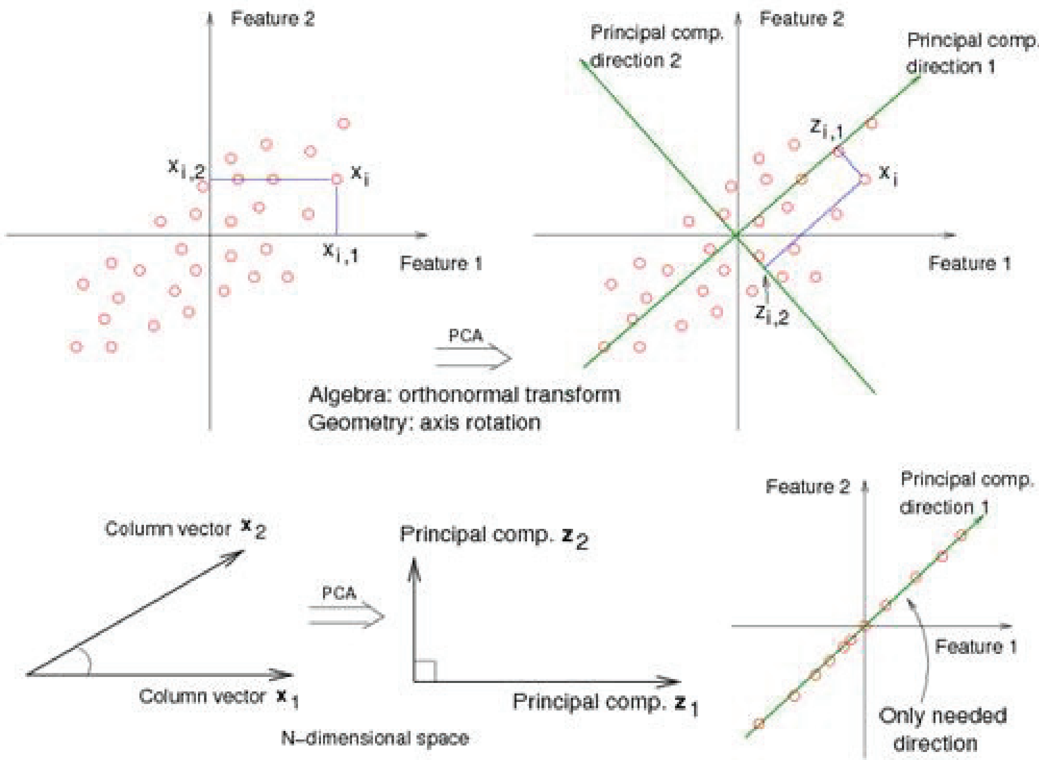

Figure 3.3: Illustration to the concept of principal component analysis [107].

overfitting the model and increases the generalization of conclusions. PCA has an important effect on regression analysis when the explanatory variables are correlated. PCA could have positive role for identifying patterns too. It can express the data with highlighting their similarities and differences. This is important when the dimension of dataset is high and finding patterns is hard. The other main advantage of PCA in identifying patterns is that once patterns in the data were specified, it can compress the data by reducing the number of dimensions without much loss of information (Fig. 3.3).

3.2.2 Independent component analysis (ICA)

Independent component analysis (ICA) is a mathematical method which reveals hidden factors of a dataset. ICA defines a generative model for the observed multivariate data. The dataset usually is a large database of samples. In this method, the data are considered to be linear mixtures of some unknown variables. The mixing system, the coefficient weights

Figure 3.4: Comparison between PCA and ICA [110].

multiplied to data, is also unknown. Variables called independent components of the data are assumed to be non-Gaussian and mutually independent. These independent components are found by ICA. This method is more powerful than other techniques like PCA, and it is able to find the hidden factors which other methods cannot [108].

The differences between PCA and ICA can be described by two factors. First, in PCA, components were organized based on the value of their variances but in ICA, there is no order of magnitude associated with each component. It means that there is no better or worst components. Second, the extracted components are invariant to the sign of the sources [109]. In PCA, principal components are orthogonal to each other, while in ICA, independent components are not necessarily orthogonal as shown in Fig. 3.4. Following explains the theoretical formulation of ICA briefly.

A random observed vector $X = [X_1, X_2, \cdots]^T$ is considered in which its m elements are mixtures of m independent elements of a random vector $S = [S_1, S_2, \cdots]^T$ given by:

$$X = AS, \tag{3.22}$$

where A is an $m \times m$ mixing matrix. The goal of ICA is to find the unmixing matrix W which is the inverse of A. That will give Y, the best possible approximation of S:

$$Y = WX = S, \tag{3.23}$$

For applying ICA it is necessary to make five assumptions:

- Statistical independence between each of the sources S_i from the sources vector S is assumed.
- The mixing matrix must be square and full rank. It means that the number of mixtures must be equal to the number of sources. Also the mixtures must be linearly independent from each other.
- In ICA, only source vector S could have stochasticity in the model. The model (the coefficient weight) should be noise free.
- The data should be centered. It means that the data should have zero mean.
- The source signals should not have a Gaussian probability density function (pdf) except for one single source that can be Gaussian [109].

ICA finds the independent components by maximizing the statistical independence of the estimated components. There are different definitions for statistical independency. Two common definitions of independence for ICA are:

- Minimization of mutual information
- Maximization of non-Gaussianity

The minimization of mutual information (MMI) family of ICA algorithms uses measures like maximum entropy. The non-Gaussianity family of ICA algorithms, motivated by the central limit theorem, uses negentropy (measure the distance from normality). It is remarkable to mention that the typical algorithms, in order to decrease the complexity of the problem in iterative algorithm, do centering and whitening as preprocessing steps.

ICA can analyze signals from many different sources like digital images, document databases, reducing noise in pictures, telecommunication, economic indicators and psychometric measurements. In many cases, the measurements are given as a set of parallel signals or time series. For these problems blind source separation is used to characterize them. Typical examples for ICA application are mixtures of simultaneous speech signals that have been picked up by several microphones, brain waves recorded by multiple sensors, interfering radio signals arriving at a mobile phone, or parallel time series obtained from some industrial processes.

3.3 Classification methods

Classification identifies that the new observation belongs to which set of specified categories. Classification works on the basis of a training set of data containing observations whose category membership is known [111]. Artificial intelligence methods are often applied in classification step and make the monitoring algorithm intelligent. There are many artificial methods for classification. In this study artificial neural network is used for data classification.

3.3.1 Artificial neural netwrok

Artificial neural networks (ANNs) may be defined as structures comprised of densely interconnected adaptive simple processing elements (neurons) that are capable of performing massively parallel computations for data processing and knowledge representation [112]. Neural networks have been used to solve a wide variety of tasks that are hard to solve using ordinary rule-based programming [113]. They are used in hardware and software to make machines, robots and programs intelligent. A neural network typically is defined with three parameters:

- The activation or transfer functions of layers.

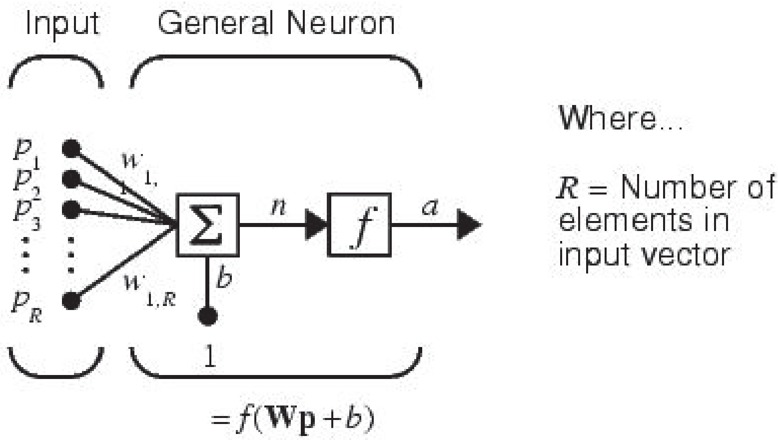

Figure 3.5: Neuron structure [114].

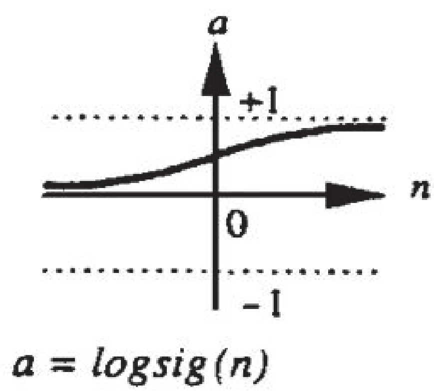

Figure 3.6: Transfer function of a neuron: logsig [114].

- Its topology which specifies the way its layers are connected.
- The learning or training algorithm of neural network.

Fig. 3.5 shows the model of a neuron with multiple inputs. In this figure P_1 to P_R are the elements of input vectors. $W_{1,1}$ to $W_{1,R}$ are weights and b is the bias related to this neuron and f is transfer function. A transfer function is the function that describes the output of a neuron given an input or set of inputs [115]. Some most common transfer functions are log-sigmoid (Fig. 3.6), tan-sigmoid (Fig. 3.7) and purelin (Fig. 3.8) [114]. Neurons make layers and they are organized based on the topology of network: feed-forward or feedback. Feed-forward networks (Fig. 3.9) have directed acyclic graph and feedback networks (Fig. 3.10) have cycles in their graphs.

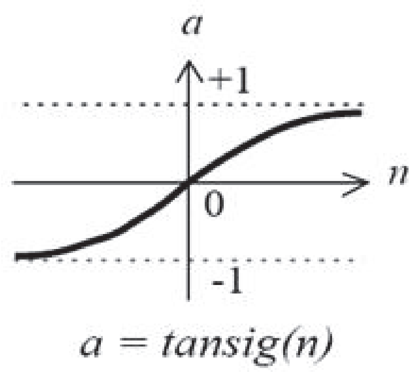

Figure 3.7: Transfer function of a neuron: tansig [114].

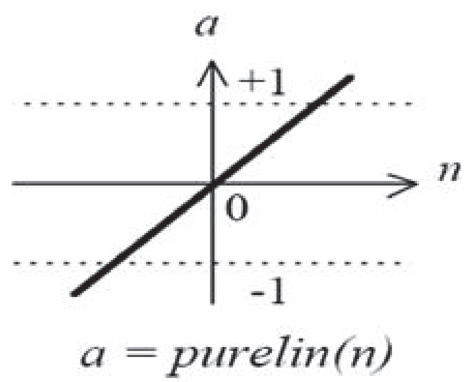

Figure 3.8: Transfer function of a neuron: purelin [114].

Learning neural network means finding the best function f for doing a specific desired goal based on the input data. To find this optimal f, a cost function will be defined and the learning methods try to minimize this cost function. There are three learning methods: supervised learning, unsupervised learning and reinforcement learning. In supervised learning, there are given pairs of input and output and the desired goal is to find a function f which connects them optimally. In this learning method a common cost function is mean squared error which minimizes the average squared error between network outputs and the given output values for all the given observations. Backpropagation algorithm for training neural network is a supervised learning method with mean square cost function. In unsupervised learning, some input data is given to the neural network and a cost function which could be any function of input data and output of neural network will be considered. The other method, reinforcement learning, also does not take the input data. In this learning

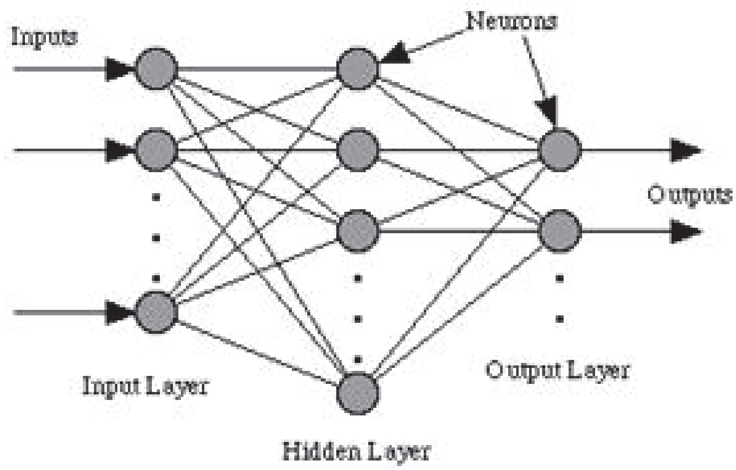

Figure 3.9: Feed-forward network [116].

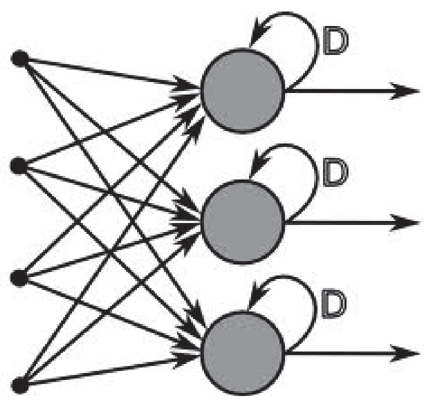

Figure 3.10: Feedback network [117].

method the input data will be generated and the goal is to minimize long-term cost.

Neural networks have wide variety of applications. They are applied for function approximation, classification and pattern recognition, data processing like filtering and clustering. They also could be used in control and robotics. Neural networks have special properties like learning ability, generalization, adaptivity and fault tolerance which make them applicable in different field of applications [118].

3.4 Chapter summary

In this chapter a brief introduction about the statistical and intelligence methods which will be used in this study, were explained. The methods were categorized based on their application for pattern recognition in monitoring algorithm. For each method a short explanation about the related technical issues with the application of the method and its properties was discussed.

Chapter 4

Tool wear monitoring algorithms

In this chapter, first, with applying two conventional signal processing methods for feature extraction in machining monitoring literature, descriptors and wavelet transform, three different monitoring algorithms are designed. In these algorithms neural network has been used for classification and PCA and ICA for feature selection. The algorithms are trained and examined with the introduced dataset in Chapter two. Then with using the S-Transform as feature extraction method, two new monitoring algorithms are developed with the same dataset, classification and feature selection methods. The accuracies of algorithms are compared to find the most effective monitoring algorithm and investigate the effect of S-Transform on the accuracy of algorithm. In this chapter, the dataset for developing the algorithms are the AC power, table and spindle vibration signals from case 1 and 9. These two cases have the same operating conditions.

4.1 Algorithm 1: Descriptors-NN

In this algorithm descriptors of each data will be considered as features for expressing it. For classifying the data, neural network method will be implemented. As it has been explained in the previous chapter, to express the central tendency and dispersion of the data the following descriptors are considered here:

- **For measure of central tendency**: geometric mean, harmonic mean, arithmetic average (mean), 50$^{\text{th}}$ percentile (median), most frequent value (mode), and trimmed mean.

- **For measure of dispersion**: interquartile range, mean absolute deviation (MAD), central moment of all orders (moment), range, standard deviation (std), and variance.

These descriptors can be simply calculated with MATLAB software. Therefore, here instead of each data, X, which is a 3001×1 matrix, there is a 12×1 feature vector. This feature vector has the benefit of reduced dimension and also keeps the distinctive information between the data.

$$[X]_{3001\times 1} \xrightarrow{\text{Calculate descriptors}} [Y]_{12\times 1}. \tag{4.1}$$

The arrays for matrix $[Y]$ are central tendency (from 1 to 6) and dispersion descriptors (from 7 to 12) respectively. Figures 4.1–4.3 show the value of descriptors for the AC power, table vibration and spindle vibration signals of case one. Fig. 4.1 shows that descriptor number five, most frequent value or mode, does not follow a distinctive trend in dataset and therefore, it cannot help for distinguishing the AC power signals and it is not a good feature to be in feature vector. Also Fig. 4.2 and 4.3 show that the descriptor number nine, central moment of all orders or moment, has the same value for all the data and then it cannot be a good feature for distinguishing the table vibration signals or the spindle vibration signals. Therefore, the feature vectors with eliminating these descriptors for each type of signals will be changed to a new feature vector:

$$[Y]_{12\times 1} \rightarrow [Z]_{11\times 1}. \tag{4.2}$$

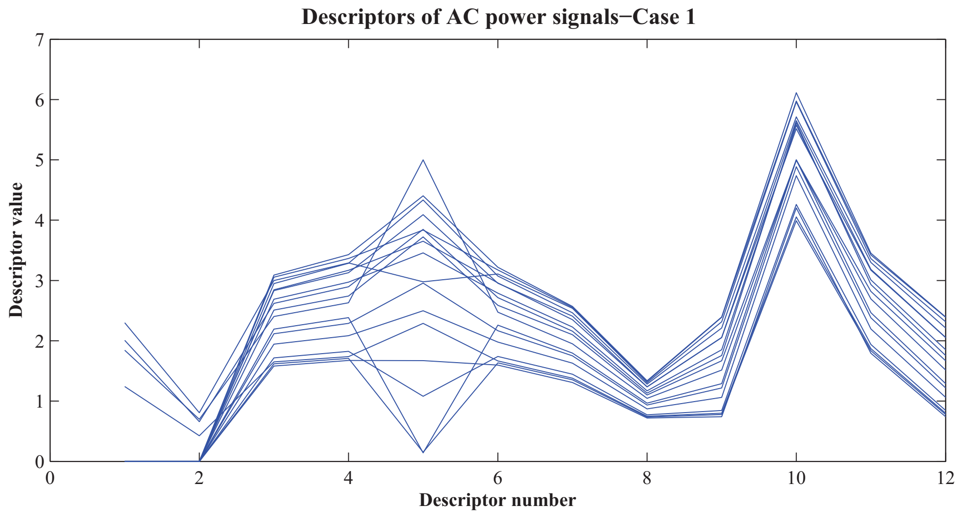

Figure 4.1: Descriptors of AC power signals.

Vectors A, B and C are samples of feature vectors of AC power, table vibration and spindle vibration signals:

$$[A] = \begin{bmatrix} 1.24 \\ 0.42 \\ 1.65 \\ 1.74 \\ 1.66 \\ 1.38 \\ 0.75 \\ 0.8 \\ 4.05 \\ 1.87 \\ 0.8 \end{bmatrix}, [B] = \begin{bmatrix} 1.82 \\ 1.8 \\ 1.83 \\ 1.85 \\ 1.78 \\ 1.85 \\ 0.26 \\ 0.18 \\ 0.06 \\ 1.55 \\ 0.06 \end{bmatrix}, [C] = \begin{bmatrix} 0.65 \\ 0.64 \\ 0.66 \\ 0.65 \\ 0.61 \\ 0.66 \\ 0.13 \\ 0.07 \\ 0.01 \\ 0.55 \\ 0.01 \end{bmatrix}. \tag{4.3}$$

After feature extraction step and making the dataset, $[A]_{11\times 26}$, the data should be classified. Since the dimension of feature vector is not too large, there is no need for doing dimension reduction and feature selection and the next step directly is classification. Neural

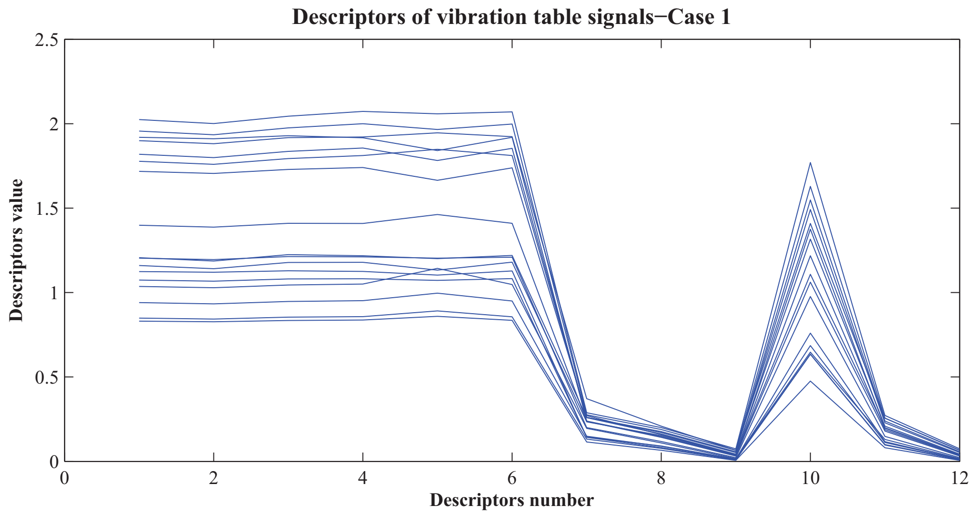

Figure 4.2: Descriptors of table vibration signals.

network is the classification method which will be used here. For categorizing each type of signal, AC power, table vibration and spindle vibration, one independent net will be trained. The goal of training neural network is predicting the value of tool wear for each new data. Therefore, for training this network the input is the signal and the output is the value of tool wear related to that signal.

For training the neural network, **nntool**, from MATLAB toolbar has been implemented. The dataset has been divided into two categorizes: training data and test data. With training data, neural network will be trained. With test data the effictiveness of the neural network will be evaluated. 20 signals of the dataset of case 1 and 9 are considered as training signals and 6 signals are the test data. Neural networks are considered as common feed-forward networks. The input layer of the neural network has 11 neurons which is equal to the number of input matrix arrays. The output layer gives the value of tool wear and it has one neuron. Number of hidden layers and their neurons will be specified during the training process. The training method is Levenberg-Marquardt backpropagation algorithm. Levenberg-Marquardt is a minimization method used for minimizing the errors in learning algorithm. Mean squared error of the validation data is an indicator for controlling the training process. The training will stop when mean squared error is increasing as

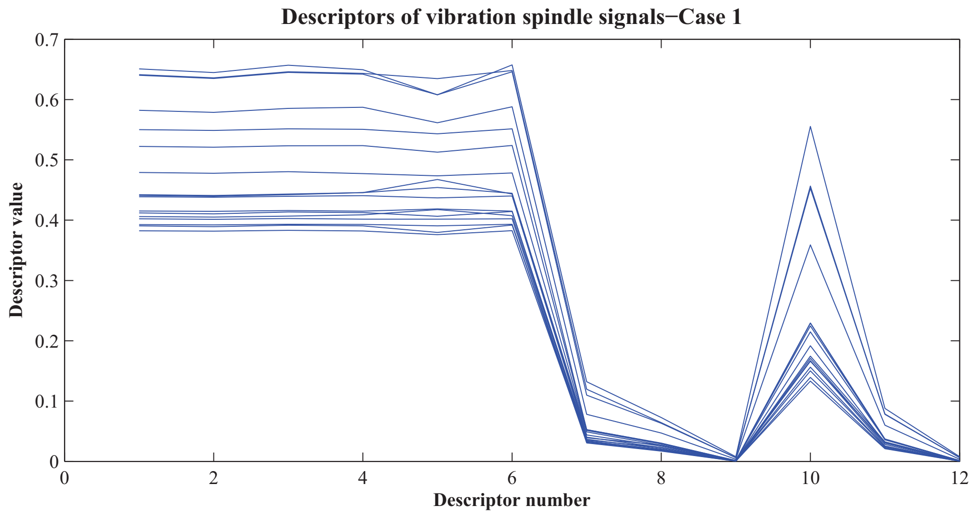

Figure 4.3: Descriptors of spindle vibration signals.

it shows that the generalization ability of the neural network is not improving any more. Since the goal of networks is to calculate the value of tool wear for signals and being as close as possible to the exact value of the tool wear, then the average of network error for calculating the tool wear value for test data can be a good parameter for comparing the trained networks and choosing the best one. This error is defined as following for each data:

$$\text{Error}(\%) = \frac{|\text{Desired output} - \text{Calculated output with network}|}{\text{Desired output}} \times 100. \tag{4.4}$$

The neural networks have been trained for all three different types of dataset: AC power, table vibration and spindle vibration. For all of them the best neural network which is the final choice, is a network with one hidden layer. For AC power and table vibration signals, hidden layer has 10 neurons and there are 5 neurons for vibration signal. In all of them the transfer functions for hidden layer is **tan-sigmoid** and for output layer is **purelin**. As an example, Fig. 4.4 shows a schematic structure of the best trained network for spindle vibration signals and Fig. 4.5 shows the trend of the training. The calculation of the average error of this network has come in the Table 4.1. Table 4.2 reports the number of neurons in hidden layer and the accuracy of the best trained networks for all three types of signals.

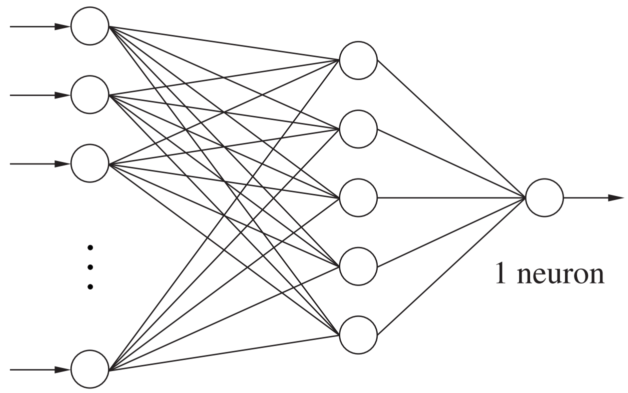

Figure 4.4: The schematic structure of the best network trained with Algorithm 1 for spindle vibration signals.

Table 4.1: Calculation of the average error of the trained networks with Algorithm 1 for spindle vibration signals.

Desired output	0.11	0.2	0.29	0.4	0.45	0.47
Calculated output	0.111	0.229	0.318	0.303	0.395	0.411
Error(%)	0.872	14.366	9.594	24.112	12.28	12.455
Average network error(%)	12.28					

4.2 Algorithm 2: Wavelet-PCA-NN

In this algorithm continuous wavelet method is doing the feature extraction. Principal component analysis reduces the dimension of feature vectors and the final feature vectors will be categorized with neural network.

Applying wavelet transform for analyzing data provides the opportunity of having time–frequency domain features at the same time. The simplest and oldest wavelet function, Haar, has been used here for wavelet analysis. This wavelet has the negative and positive values which is compatible with the available dataset which has negative and positive values. Also, since this function is not continuous, it is suitable for analyzing the signals that they may have sudden transition, such as monitoring machining tool breakage or damage [119]. Wavelet transform for each data is calculated for the scales between 1

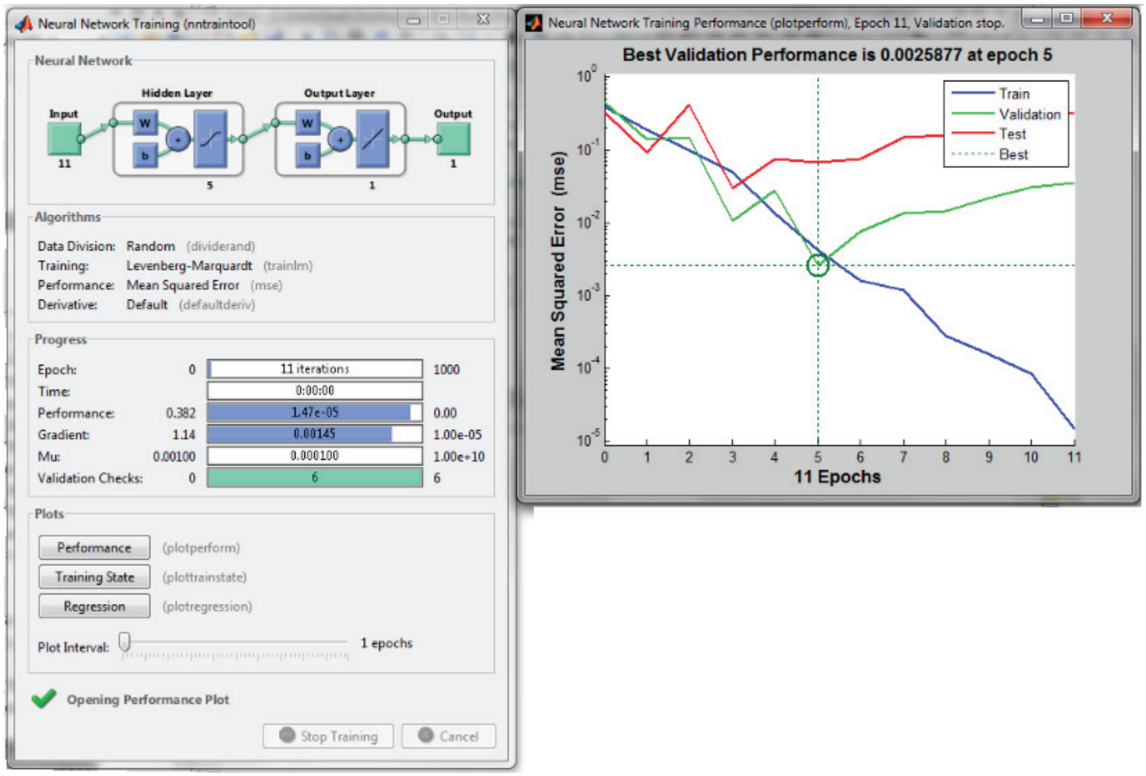

Figure 4.5: The trend of network training with Algorithm 1 for spindle vibration signals.

Table 4.2: Structure parameters and results of the trained networks with Algorithm 1.

Signal	Number of neurons in hidden layer	Network average error (%)	Network accuracy 100−error (%)
AC power	10	12.68	87.32
Table vibration	10	24.6	75.4
Spindle vibration	5	12.28	87.72

and 100. This range for scales is selected with trial and error. With command **cwt** in MATLAB, wavelet transform of each input data, $[X]_{3001\times 1}$ is calculated and the output is matrix $[Y]_{100\times 3000}$. Figures 4.6– 4.8 show the continuous wavelet transform for three signals of case 1. To have a vector instead of a matrix for each data, mean value of the matrix has considered as the final feature vector. Since the output of wavelet can be complex values, in order to do not have complex values in the dataset, absolute values of the feature vectors

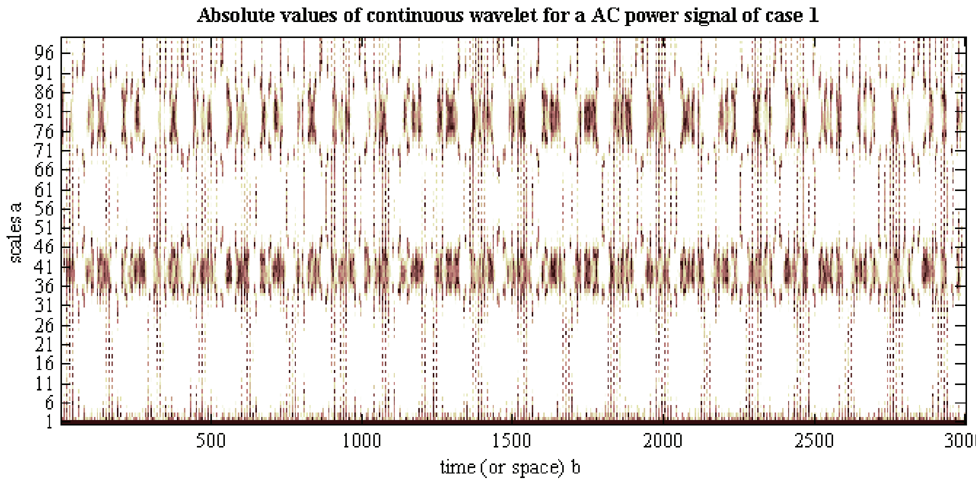

Figure 4.6: Continuous wavelet transform of a AC power signal.

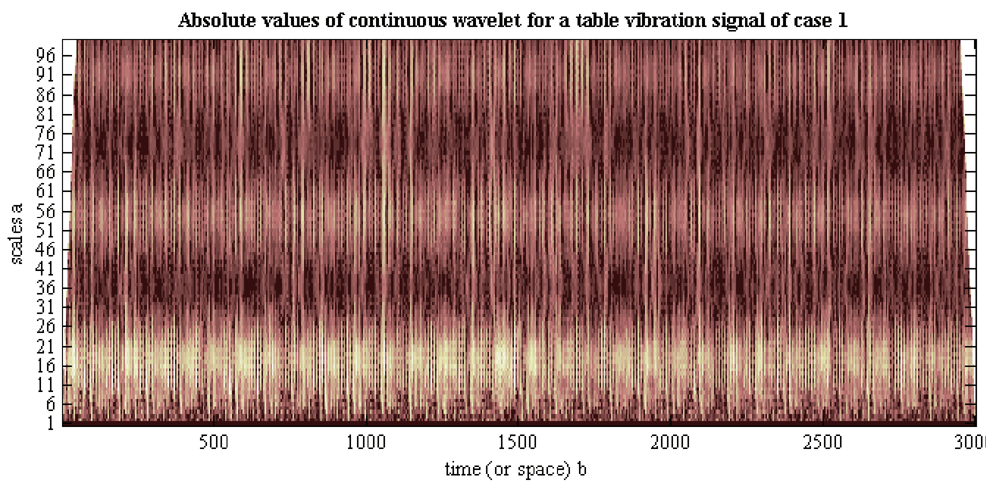

Figure 4.7: Continuous wavelet transform of a table vibration signal.

have been calculated.

$$[X]_{3001\times 1} \xrightarrow{\text{cwt}} [Y]_{100\times 3001} \xrightarrow{\text{mean}} [Z]_{1\times 3001}. \tag{4.5}$$

Principal component analysis method reduces the dimension of feature vectors to reduce the unreasonable complexity and time of calculations and also it reduces the required data for neural network training. Command **princomp** in MATLAB takes the data and returns the principal component scores. Component scores are the representation of

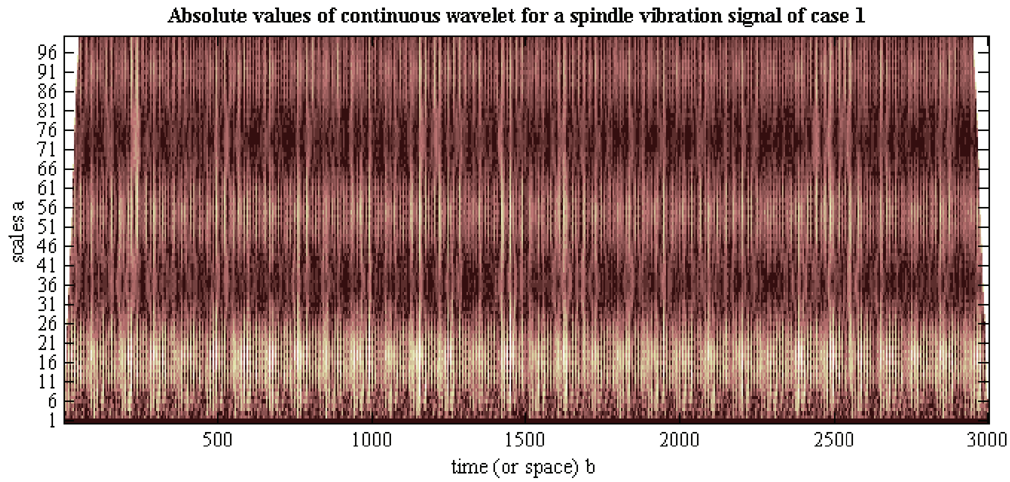

Figure 4.8: Continuous wavelet transform of a spindle vibration signal.

each data in the principal component space. For the present dataset, the matrix of the input signals is $[A]_{26\times3001}$ in which 26 is the number of data in this dataset. The output of **princomp** command on this input data is $[B]_{26\times3001}$. This matrix is the representation of matrix $[A]_{26\times3001}$ in the principal component space. Since for matrix $[A]_{26\times3001}$ number of observations (rows) is less than the number of variables (columns), $26 < 3001$, in matrix $[B]_{26\times3001}$ the value of columns 26 to 3001 are necessarily zero. Therefore, the database in new space can be defined with matrix $[C]_{26\times25}$.

$$[A]_{26\times3001} \xrightarrow{\text{PCA}} [B]_{26\times3001} \xrightarrow{\text{Delete zero columns}} [C]_{26\times25}. \tag{4.6}$$

As PCA arranges the principal components based on the largest possible variance, therefore the most important components for representing data in dataset are the first components in the output matrix of PCA. By taking the advantage of this characteristic, 5 first components are selected from the 25 components specified by PCA. By this way time consuming and complex calculations in the next step for training complex neural networks are more reduced. Therefore, the dimension of the dataset is as follows:

$$[C]_{26\times25} \rightarrow [D]_{26\times5}. \tag{4.7}$$

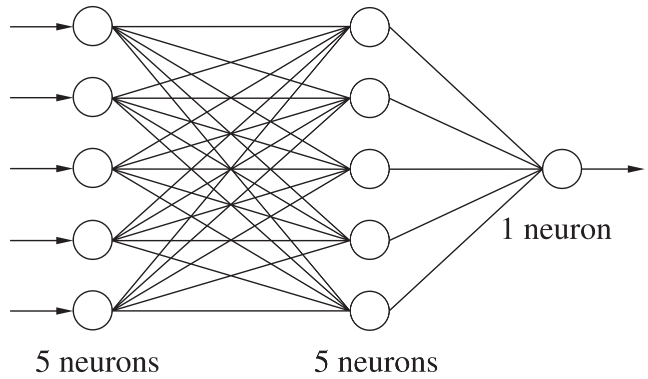

Figure 4.9: The schematic structure of the best network trained with Algorithm 2 for spindle vibration signals.

With this dataset three neural networks are trained for three categories of signals: AC power, table and spindle vibrations. As explained in the previous algorithm, neural networks are trained by use of **nntool** toolbar. The dataset has 26 signals, 20 signals for training and 6 signals for the test data. Considering the final dataset matrix $[D]_{26\times 5}$, the input layers of these networks have five neurons and the output layer have one neuron which gives the value of tool wear. The networks are feed-forward networks and they are trained with Levenberg-Marquardt backpropagation algorithm. As same as previous algorithm mean squared error is an indicator for controlling the training process. Also, the average of network error for calculating the tool wear value for test data is the scale for choosing the best trained network.

All the best trained networks for these three types of data, has one hidden layer. The network transfer function for hidden layer of all networks is **tan-sigmoid** and for output layer is **purelin**. As an example, Fig. 4.9 shows a schematic structure of the best trained network for spindle vibration signals and Fig. 4.10 shows the trend of training. Table 4.3 reports the number of neurons in hidden layer and the accuracy of the best trained networks for all three types of signals.

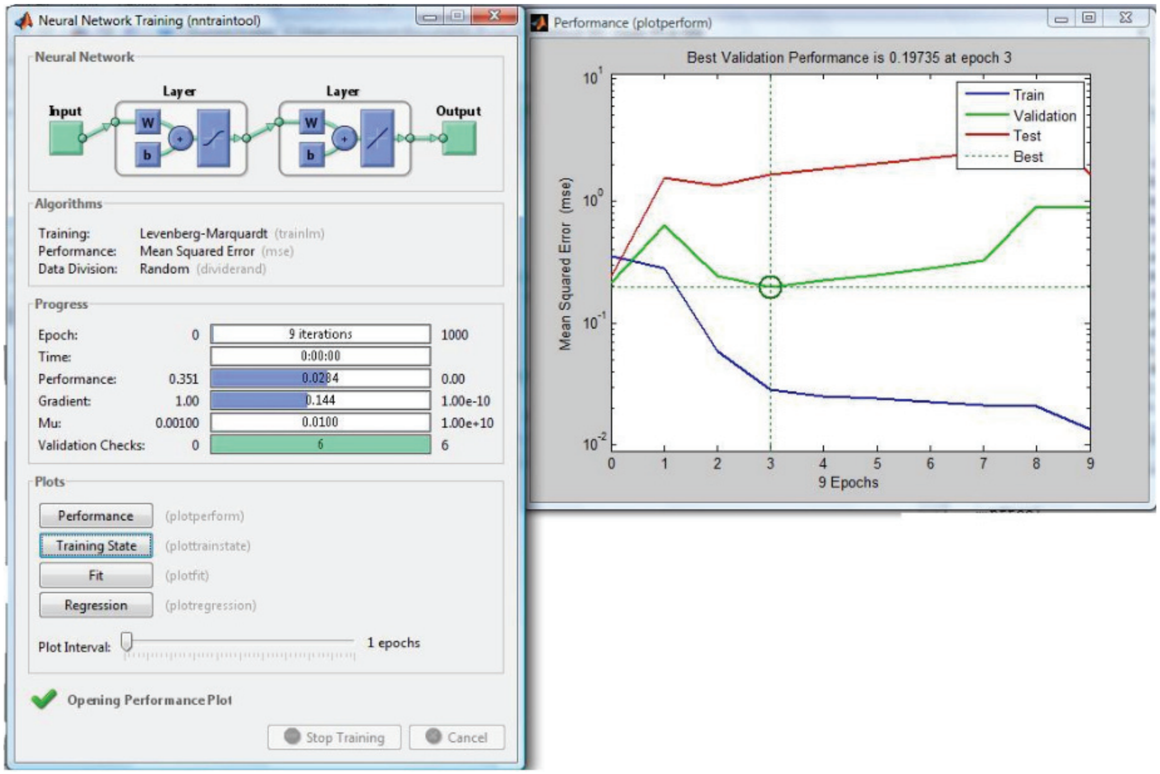

Figure 4.10: The trend of network training with Algorithm 2 for spindle vibration signals.

Table 4.3: Structure parameters and results of the trained networks with Algorithm 2.

Signal	Number of neurons in hidden layer	Network average error (%)	Network accuracy 100−error (%)
AC power	3	14.1	85.9
Table vibration	3	21.3	78.7
Spindle vibration	5	13.6	86.4

4.3 Algorithm 3: Wavelet-ICA-NN

This algorithm in feature extraction and classification is similar with the previous one, except the method used for dimension reduction is the independent component analysis (ICA). The output of wavelet transform is like before $[A]_{26\times3001}$. ICA method is applied to this dataset with using a MATLAB code [120] and the dimension of dataset is changed as follows:

$$[A]_{26\times3001} \xrightarrow{\text{ICA}} [B]_{26\times10}. \tag{4.8}$$

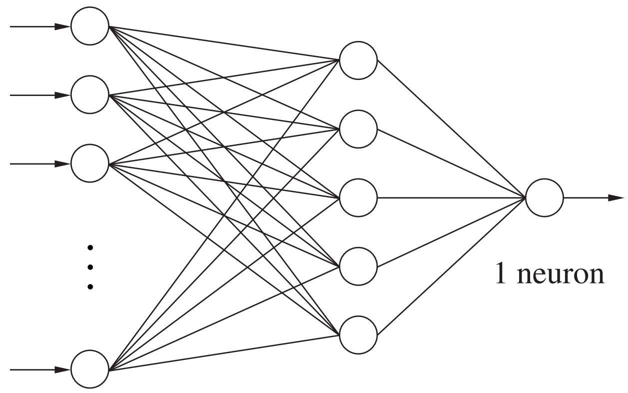

Figure 4.11: The schematic structure of the best network trained with Algorithm 3 for spindle vibration signals.

Table 4.4: Structure parameters and results of trained networks with Algorithm 3.

Signal	Number of neurons in hidden layer	Network average error (%)	Network accuracy 100−error (%)
AC power	3	19.9	80.1
Table vibration	5	17.9	82.1
Spindle vibration	5	6	94

This MATLAB code takes the number of independent components (the dataset will be expressed with these independent components) as an input data. Here this number is selected as 10. Therefore, in new representation of dataset, each data is presented with 10 features and the feature vector for each data is a 1×10 vector. Feed-forward neural networks with Levenberg-Marquet backpropagation algorithm are trained. All the networks (the best trained networks for AC power, table vibration and spindle vibration signals) have one hidden layer. The transfer function for hidden layers is **tan-sigmoid** and for output layer is **purelin**. Number of neurons of input layer is 10 (equal to the number of features for representing each data)and for output layer is one. As an example, Fig. 4.11 shows a schematic structure of the best trained network for spindle vibration signals and Fig. 4.12 shows the trend of training. Table 4.4 reports the number of neurons in hidden layer and the accuracy of the best trained networks for all three types of signals.

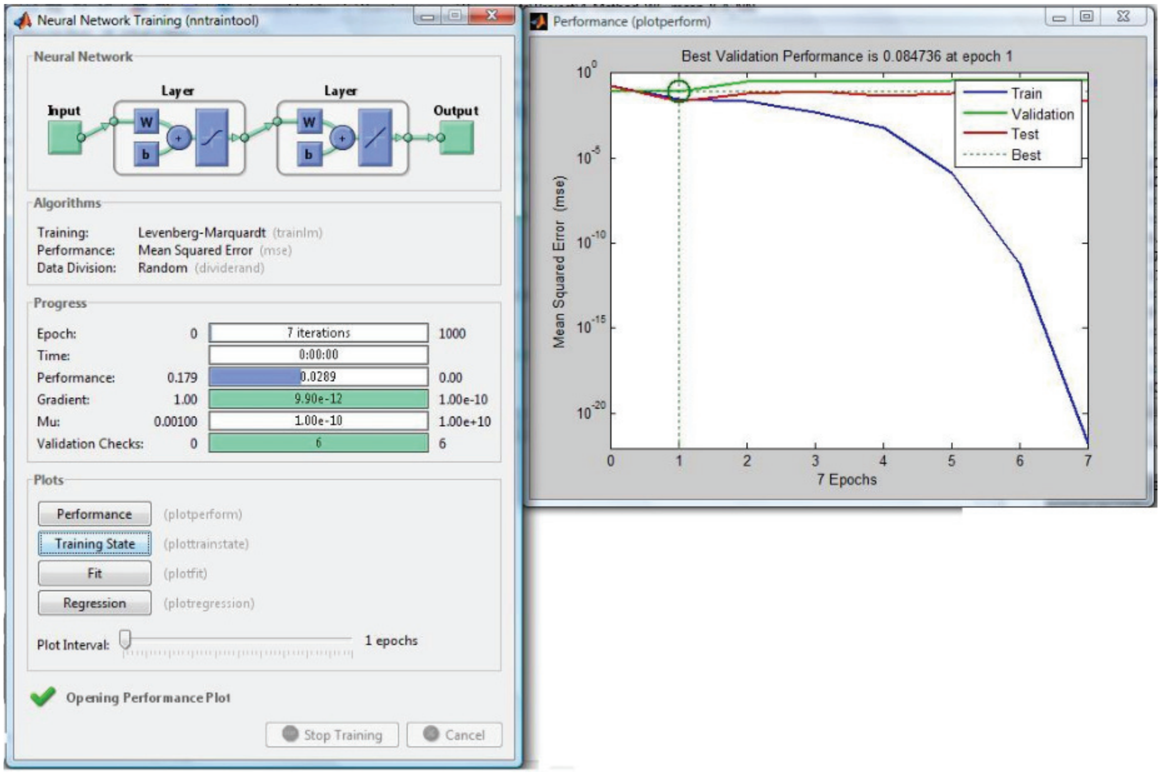

Figure 4.12: The trend of network training with Algorithm 3 for spindle vibration signals.

4.4 Algorithm 4: S-Transform-PCA-NN

After using common signal processing methods in the Algorithms 1-3, in Algorithm 4, 5, S-Transform is applied to investigate the effect of S-Transform on the accuracy of monitoring algorithm. In this algorithm S-Transform is applied for feature extraction and principal component analysis and neural network used for the feature selection and classification as the second algorithm.

As explained in Chapter 3, S-Transform has its own merits and demerits comparing with other feature extraction methods like wavelet transform [100]. For making us of the advantages of S-Transform with the hope of having more accurate results, it is investigated as Algorithm 4 here. As explained in Chapter 3, S-Transform of signals will be calculated in a range of frequencies. The available MATLAB code for calculating S-Transform takes this range of frequencies as an input in addition to the signals [120]. To make sure that the results of S-Transform return the best features which reflect the differences between signals

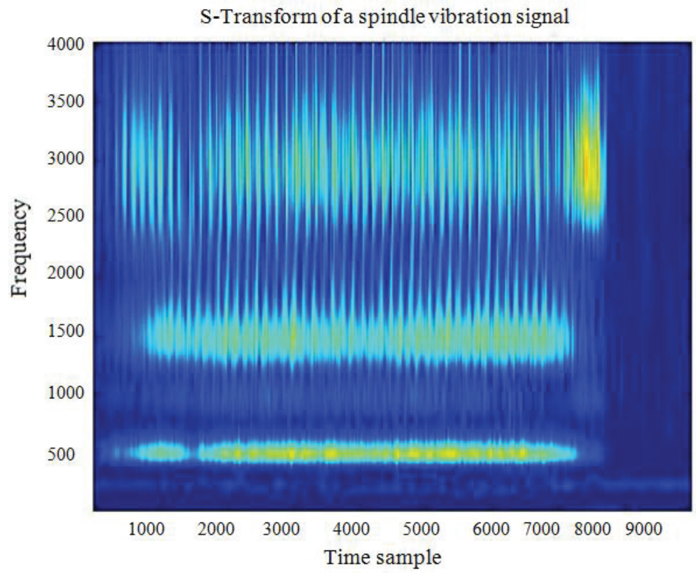

Figure 4.13: wide range frequency S-Transform analysis for a spindle vibration signal.

as much as possible, this range of frequencies should be investigated carefully. To do this, first few samples of each type of signals are analyzed in a wide range of frequencies. The results for spindle vibration is reported in Fig. 4.13. As it can be seen in this figure in the range of frequencies $300-700$ (Hz), the result of S-Transform is changing during the time with changing the situation of the tool status. Therefore, this range of frequencies has been selected for calculating the S-Transform of the signals. By applying S-Transform for each input signal there would be a feature matrix as following:

$$[X]_{3001\times 1} \xrightarrow{\text{ST}} [Y]_{401\times 3001}. \tag{4.9}$$

To avoid complex values, absolute value of arrays are used in following calculations. Also, to form a matrix based on the dataset, for each signal a feature vector is needed. By using mean value of the feature matrices, they are changed to feature vectors.

$$[Y]_{401\times 3001} \xrightarrow{\text{mean}} [Z]_{1\times 3001}. \tag{4.10}$$

By this way, $[A]_{26\times 3001}$ is feature vector matrix of the original dataset. The dimension of

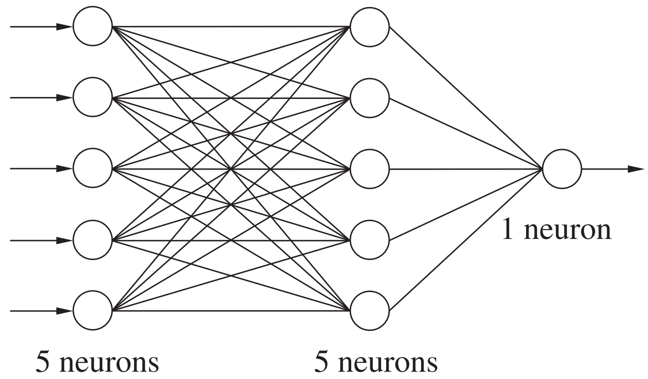

Figure 4.14: The schematic structure of the best network trained with Algorithm 4 for AC power signals.

this dataset is reduced with PCA method. With MATLAB command **princomp** which has been explained before, the new representation of dataset in principal component space is:

$$[A]_{26\times3001} \xrightarrow{\text{PCA}} [B]_{26\times3001}, \tag{4.11}$$

with deleting the zero columns, the reduced dataset becomes:

$$[B]_{26\times3001} \longrightarrow [C]_{26\times25}. \tag{4.12}$$

Considering 5 first features for describing the data, the dataset will be changed to new reduced matrix $[D]_{26\times5}$.

$$[C]_{26\times25} \longrightarrow [D]_{26\times5}. \tag{4.13}$$

Using this dataset three neural networks are trained for calculating the value of tool wear for AC power, table vibration, and spindle vibration signals. The networks are feed-forward type and they are trained with Levenberg-Marquet backpropagation algorithm. All the networks have one hidden layer and the transfer function for hidden layers is **tan-sigmoid** and for output layer is **purelin**. Number of neurons of input layer is 5 (equal to the number of features in feature vector of each data) and for output layer is one. As an example, Fig. 4.14 shows a schematic structure of the best trained network for AC power signals and

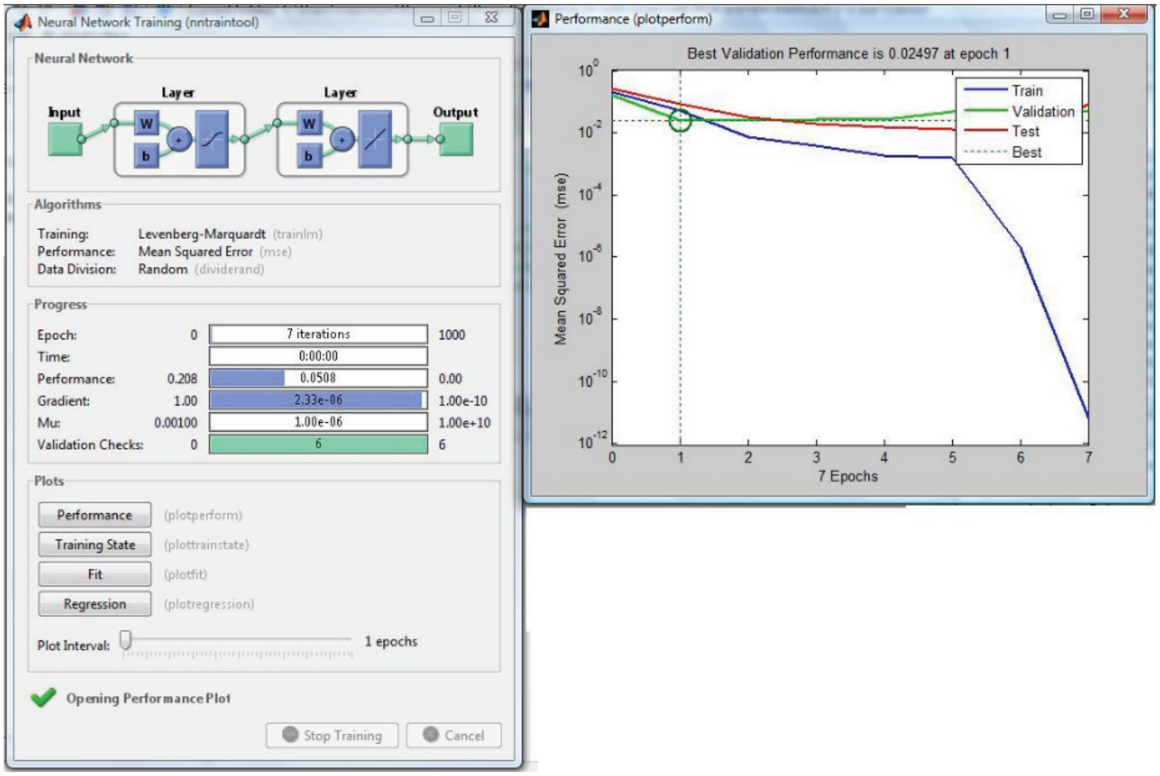

Figure 4.15: The trend of network training with Algorithm 4 for AC power signals.

Table 4.5: Structure parameters and results of trained networks with Algorithm 4.

Signal	Number of neurons in hidden layer	Network average error(%)	Network accuracy 100−error (%)
AC power	5	19.65	80.35
Table vibration	3	18.2	81.8
Spindle vibration	3	12.22	87.78

Fig. 4.15 shows the trend of training. Table 4.5 reports the number of neurons in hidden layer and the accuracy of the best trained networks for all three types of signals.

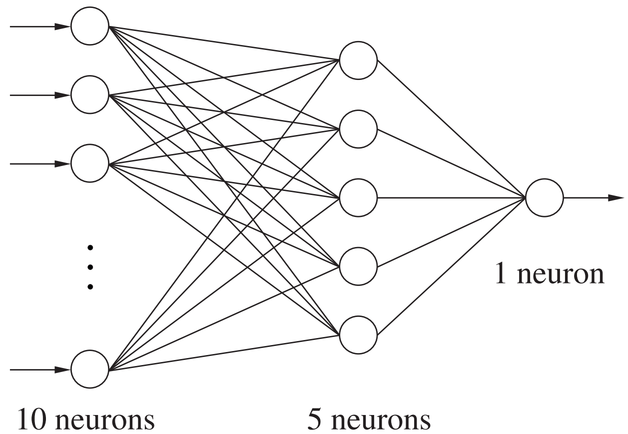

Figure 4.16: The schematic structure of the best network trained with Algorithm 5 for table vibration signals.

4.5 Algorithm 5: S-Transform-ICA-NN

In this algorithm, different from the previous algorithm which used PCA for feature selection, ICA method is applied.

$$[A]_{26\times 3001} \xrightarrow{\text{ICA}} [B]_{26\times 10}. \tag{4.14}$$

Like before, the networks are trained with the new reduced dataset. The input layer of networks have 10 neurons and the output layer has one neuron. All networks are feed-forward type and they are trained with Levenberg-Marquet backpropagation training method. The transfer functions are chosen as **tan-sigmoid** and **purelin** for hidden layer and the output layer respectively. As an example, Fig. 4.16 shows a schematic structure of the best trained network for table vibration signals and Fig. 4.17 shows the trend of training. Table 4.6 reports the number of neurons in hidden layer and the accuracy of the best trained networks for all the three types of signals.

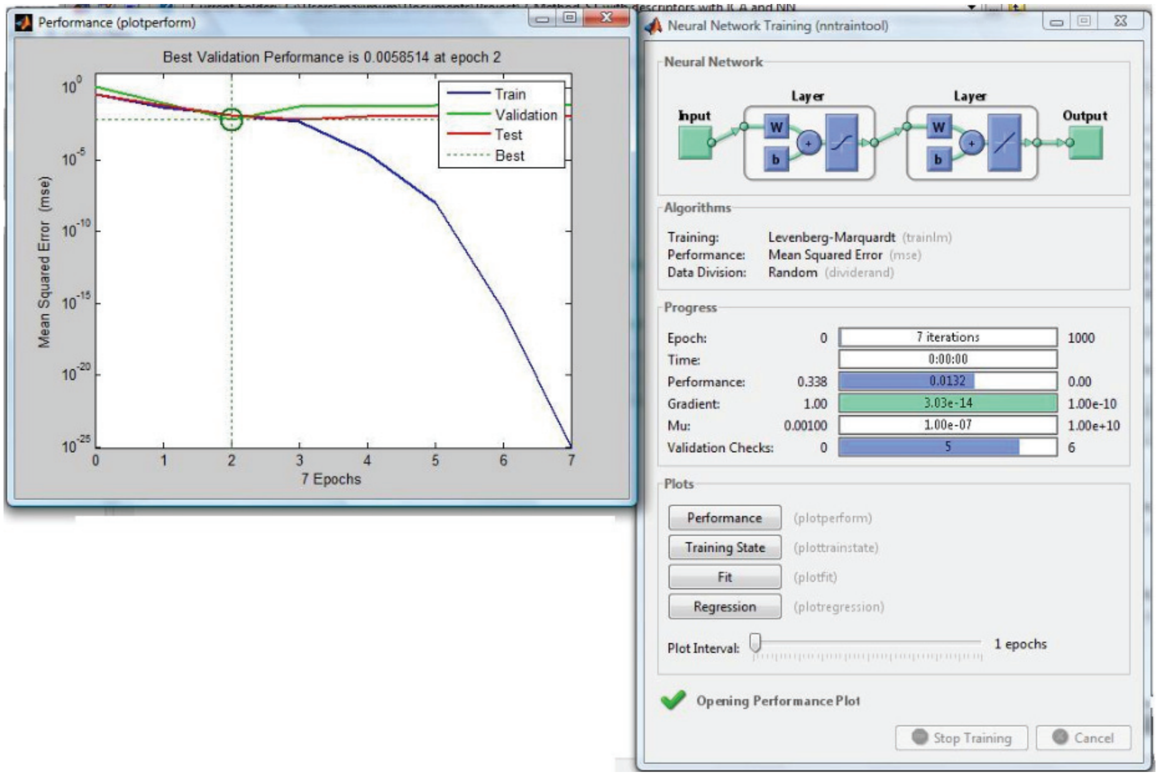

Figure 4.17: The trend of network training with Algorithm 5 for table vibration signals.

Table 4.6: Structure parameters and results of trained networks with Algorithm 5.

Signal	Number of neurons in hidden layer	Network average error (%)	Network accuracy 100−error (%)
AC power	5	13.54	86.46
Table vibration	5	15.93	84.07
Spindle vibration	5	11.6	88.4

4.6 Discussion

The above reported results in the tables show that all trained algorithms are highly accurate. Also based on the results, it can be concluded that the average accuracy of algorithms (the average of accuracies of three networks trained with one algorithm) are generally increasing from Algorithms 1 to 5. The accuracies of Algorithms 4 and 5, which they have used S-Transform as feature extraction method, are equal or more than the accuracies of other algorithms. It shows that S-Transform compares with descriptors and wavelet transform, has better performance in extracting the most informative features. Looking at each type of signals (AC power, spindle vibration, table vibration) also can find the most accurate algorithm. For example, the most accurate network for spindle vibration signals has been trained with the Algorithm 3 with 94% precision. The trained networks with AC power signals have the best average accuracy comparing with two other types of signals. The networks trained with table vibration signals have the lowest average accuracy.

For all monitoring algorithms, the training process was fast. In designing monitoring algorithms and training the neural networks, two different datasets have been used for training and test data. It has increased the generalization ability of the networks. However, due to some limitations like the small size of dataset or the difference between the experimental conditions which the signals are recorded with real rush industrial environment, the generalization ability decreases. Most of the monitoring algorithms proposed in the field of machining monitoring are suffering from these limitations. As it has been explained before, the training and test data are the recorded signals of case 1 and 9 which has the same operating conditions. Therefore, the neural networks have been trained for working in this set of operating conditions and for other set of operating conditions another trained networks are needed.

In next chapter, Algorithm 5 which has the best average accuracies for all three types of signals will be used for doing more investigations about its accuracy. First, the effect of sensor fusion on the accuracy of this algorithm will be evaluated. Then, the algorithm will be examined for working in different operating conditions and for detecting the sharp tool from the dull one.

Chapter 5

Further studies on the accuracy of the selected algorithm

In the previous chapter, monitoring algorithms have been trained with a single type of signals recorded in one specific operating conditions (case 1 and 9) and the accuracies of algorithms have been calculated for measuring the value of tool wear. In this chapter, the effect of sensor fusion on the results of the Algorithm 5 and its accuracy will be investigated. Also, since the operating conditions can change the pattern of the recorded signals, the effect of operating conditions like depth of cut, feed and workpiece material on the accuracy of algorithm will be studied as well. For this purpose different datasets which are the combination of signals from different operating conditions will be implemented for training and testing the monitoring algorithm. Finally, instead of just calculating the value of tool wear, the accuracies of algorithm for detecting the sharp tool from the dull one will be investigated.

5.1 Effect of sensor fusion on the accuracy of algorithm

Sensor fusion refers to the use of more than one sensor signal in a complementary manner to provide a more robust prediction of one or more machining attributes [101]. In previous chapter, the dataset for training and testing networks was included just one type of signals and for each type of signals one individual network was trained. In this chapter with sensor fusion, a dataset which combines the information of two different types of signals will be used for training and testing the algorithm. For example, instead of just using AC power signal, spindle vibration signal for the same data is also considered for representing the data. In this case, for one data, the AC power signal and the spindle vibration signal, both are analyzed with S-Transform and then the feature vector is made with extracted features of both of them.

$$\text{AC power signal} = [X]_{3001\times1} \xrightarrow{\text{ST-mean}} [x]_{401\times1},$$

$$\text{Spindle vibration signal} = [Y]_{3001\times1} \xrightarrow{\text{ST-mean}} [y]_{401\times1},$$

$$\Longrightarrow \text{Feature vector of data} = f = \begin{bmatrix} x \\ \\ y \end{bmatrix}_{802\times1}. \tag{5.1}$$

After making the feature vectors the new dataset is $[A]_{802\times26}$. With ICA method for feature selection, the input dataset to neural network is:

$$[A]_{802\times26} \xrightarrow{\text{ICA}} [B]_{10\times26}. \tag{5.2}$$

With the new dataset the neural network is trained. The neural network parameters are the same as explained in Chapter 4. The algorithm are trained for three pairs of signals:

1- AC power and spindle vibration signals,

2- AC power and table vibration signals,

3- Table vibration and spindle vibration signals.

Tables 5.1–5.3 show the accuracy of the trained algorithm for each pair of signals. In the tables, the results of individual networks have reported too. Results show that sensor fusion increase the accuracy of algorithm. Therefore, in case of having sensors available, using multiple sensors and making dataset with sensor fusion is completely recommended.

Table 5.1: Accuracy of Algorithm 5 for AC power and spindle vibration sensor fusion.

Dataset	AC power, Spindle vibration	AC power	Spindle vibration
Accuracy (%)	88.36	86.46	88.34

Table 5.2: Accuracy of Algorithm 5 for AC power and table vibration sensor fusion.

Dataset	AC power, Table vibration	AC power	Table vibration
Accuracy (%)	89.54	86.46	84.07

5.2 Effect of variable operating conditions on the accuracy of algorithm

Since a machining center works in different operating conditions, therefore, having a monitoring algorithm which can indicate the process status in different operating conditions would be very valuable to industrial applications. In this section the accuracy and efficiency of Algorithm 5 for monitoring in different operating conditions will be investigated. For this purpose the dataset for training the algorithm should include signals from all operating conditions that the algorithm is supposed to work under them. Here three different datasets are considered for the training algorithms. Each of these datasets includes signals from two set of operating conditions which are different in one condition. For example, the first dataset includes signals from cases 1, 9, 2 and 10 in which the operating conditions of cases 1 and 9 are different from 2 and 10. These operating conditions have same feed and material and different depth of cuts. By training and testing the algorithm using this dataset the accuracy of the algorithm for working in the operating conditions which have different depth of cuts will be examined. The results will show the accuracy of algorithm for working in the conditions which the depth of cut changes. Same as the first dataset, second and third dataset include signals from different operating conditions and are used for investigating the accuracies of algorithm when feed or workpiece material change (Table 5.4). First dataset is made with the signals from two operating conditions which are

Table 5.3: Accuracy of Algorithm 5 for table and spindle vibration sensor fusion.

Dataset	Table vibration, Spindle vibration	Table vibration	Spindle vibration
Accuracy (%)	89.6	84.07	88.34

Table 5.4: Three datasets.

Dataset	Cases	Difference
1	$(1,9),(2,10)$	Depth of cut
2	$(1,9),(4,12)$	Feed
3	$(1,9),(5,13)$	Work piece material

different in depth of cut as follow:

$$\text{Signal from cases 1 and 9} = [X]_{3001\times 1} \xrightarrow{\text{ST-mean}} [x]_{401\times 1},$$

$$\text{Dataset made by cases 1 and 9} = [a]_{401\times 26},$$

$$\text{Signal from cases 2 and 10} = [Y]_{3001\times 1} \xrightarrow{\text{ST-mean}} [x]_{401\times 1},$$

$$\text{Dataset made by cases 2 and 10} = [b]_{401\times 23},$$

$$\text{First dataset made by cases 1,9,2,10} = A = \begin{bmatrix} a & b \end{bmatrix}_{401\times 49}. \tag{5.3}$$

Second and third datasets are made like the first one:

$$\text{Second dataset made by cases 1,9,4,12} = B = \begin{bmatrix} a & c \end{bmatrix}_{401\times 48}, \tag{5.4}$$

$$\text{Third dataset made by cases 1,9,5,13} = C = \begin{bmatrix} a & d \end{bmatrix}_{401\times 47}. \tag{5.5}$$

Algorithm 5 is trained by these three datasets as it has been explained in Chapter 4. The best accuracies of the best trained networks for these datasets are as Table 5.6. With an overall look on the table, it can be concluded that the accuracy of a trained algorithm with one of these three datasets is worse than the algorithm trained to work in one operating condition

Table 5.5: Datasets prameters.

Cases	Depth of cut (mm)	Feed (mm/rev)	Material
$1,9$	1.5	0.5	Cast iron
$2,10$	0.75	0.5	Cast iron

Table 5.6: Accuracy of Algorithm 5 for different operating conditions.

Cases make dataset	AC power	Spindle vibration	Table vibration
cases $(1,9)$	86.46	88.34	84.07
cases $(1,9),(2,10)$	75.75	47.8	81.2
cases $(1,9),(4,12)$	61.72	76.75	73
cases $(1,9),(5,13)$	58.14	63.03	67

only. This is due to increase the diversity of patterns of the data in the dataset, which make the distinguishing data and predicting their tool wear harder and more complex. The results show that the accuracy of algorithm when it works in the condition in which the material is changing, is worse than other cases which other parameters like feed or depth of cut change. Also, the algorithm monitors the process more accurate when the depth of cut changes rather than feed.

5.3 Detecting the sharp and dull tool with the selected algorithm

In designing the previous monitoring algorithms, the goal of monitoring was to predict the amount of tool wear. Sometimes it is only required to know if a tool is sharp or dull rather than knowing the amount of tool wear. In definition, sharp tool is a tool which does not have tool wears more than a specific value, "a" If based on this definition, tool is not a sharp tool, it is called dull tool. In the simple way:

$$\begin{cases} \text{sharp tool} & \text{tool wear} < a\ (mm), \\ \text{dull tool} & \text{tool wear} \geq a\ (mm). \end{cases} \tag{5.6}$$

The amount of "a" depends on the expected production quality from machining with this tool. It can also be dependent on the type of machining process and tool material. Since in the algorithms presented in this study, the amount of flank tool wear is predicted, it is possible to indicate sharp tool from dull one too. By adding a simple comparison to the MATLAB code of algorithms, the final value predicted for flank tool wear is compared with "a". The output of algorithm indicates if the input signal has been produced with a sharp tool or with dull one. This code can get "a" as an input too. By this way, the algorithm would be compatible for doing monitoring in different desired qualities of production. The output of this algorithm for five supposed test signals is as following:

$$\begin{bmatrix} a & b & c & d & e \end{bmatrix} \xrightarrow{\text{Algorithm 5}} \begin{bmatrix} \text{sharp} & \text{sharp} & \text{dull} & \text{sharp} & \text{dull} \end{bmatrix}. \tag{5.7}$$

In this case the accuracy of the algorithm is its ability for distinguishing sharp tool from dull one. So, the accuracy is expressed with dividing the number of output predicted correctly to the total number of outputs.

$$\text{Accuracy (\%)} = \frac{\text{number of correct predictions}}{\text{number of total outputs}} \times 100. \tag{5.8}$$

Therefore, for the above supposed test signals, if the real status of their tools would be as below:

$$\begin{bmatrix} a & b & c & d & e \end{bmatrix} \longrightarrow \begin{bmatrix} \text{sharp} & \text{dull} & \text{dull} & \text{sharp} & \text{sharp} \end{bmatrix}. \tag{5.9}$$

Table 5.7: Tool status.

Data No.	1	2	3	4	5
Predicted flank wear (mm)	0.07	0.2	0.26	0.36	0.51
Predicted tool status	Sharp	Sharp	Sharp	Sharp	Dull
Real tool status	Sharp	Sharp	Sharp	Dull	Dull

Then the status of tool for signals "*b*" and "*e*" have been detected incorrectly and the accuracy of algorithm would be:

$$\text{Accuracy }(\%) = \frac{3}{5} \times 100 = 60\%. \tag{5.10}$$

To investigate the accuracy of the algorithm 5 for detecting sharp tool from dull one, as explained before, it is necessary to specify the value of "*a*" which is the boundary limit between sharp and dull first. The range for maximum flank wear for coated carbide tool which has been utilized here for roughing, is $0.3-0.5\ mm$ [121]. The value "*a*" is considered as $0.4\ mm$ which is in the middle of this range. Therefore, the definition of sharp and dull tool here is:

$$\begin{cases} \text{sharp tool} & \text{tool wear} < 0.4\ mm, \\ \text{dull tool} & \text{tool wear} \geq 0.4\ mm. \end{cases} \tag{5.11}$$

The results of the first algorithm for some random input data can be expressed in the new definition as the Table 5.7. The accuracies of Algorithm 5 which was trained with the datasets include signals from different operating conditions are calculated based on the new definition and reported in the Table 5.8. The table shows that the algorithm detects sharp and dull tool with 90% accuracy for the combination of 1, 9, 2 and 10 cases and with AC power signals. This is more than the accuracy of algorithm for calculating tool wear which was 75.75%. Comparing these two accuracies shown in the Table 5.8, it can be easily seen that the accuracies of the algorithms when they are used for detecting sharp and dull tool is more than when they have been used for calculating the amount of flank

Table 5.8: Comparing accuracy of detecting tool status with predicting tool wear.

Dataset	Signal	Accuracy of detecting sharp and dull tool (%)	Accuracy for predicting tool wear (%)
Cases $(1,9),(2,10)$	AC power	90	75.75
	Table vibration	70	81.2
	Spindle vibration	100	47.8
Cases $(1,9),(4,12)$	AC power	78	61.72
	Table vibration	67	73
	Spindle vibration	67	66.75
Cases $(1,9),(5,13)$	AC power	75	50.14
	Table vibration	87	67
	Spindle vibration	75	63.03

tool wear.

5.4 Chapter summary

In this chapter, the effect of sensor fusion on the accuracy of algorithm has been investigated. The results show that sensor fusion increases the accuracy of the algorithm. Therefore, with increasing the number of sensors, considering the financial limitations, it is possible to have more accurate algorithm for monitoring machining process and have productions with better quality. The ability of algorithm for working in different operating conditions was another topic which has been studied in this chapter. Investigations show that in this case the accuracy of algorithm decreases generally. This decrease is more for material change and less for feed and depth of cut change. Finally, a new definition for tool status has been proposed. In this definition tool may have one of these two status: sharp or dull. Tool is sharp if its flank tool wear is less than an indicated value of 0.4 *mm* as an example. Also, it is dull if the tool wear is more than 0.4 *mm*. The results of algorithm have been expressed in this new definition and the accuracies of algorithm have been reported. Comparing the accuracies of algorithm in this new definition with the accuracies for calculating tool wear, it can be concluded that the accuracies of proposed algorithm for indicating sharp tool from dull tool is more than its accuracies for calculating the amount of tool wear. Therefore, the selected algorithm is more reliable when it is used for distinguishing sharp and dull tools.

Chapter 6

Conclusion and future works

6.1 Conclusion

Machining processes are among the most common industrial operations utilized everyday in factories. Therefore, process monitoring and fault diagnosis of machining operations are very important and useful to the industry. Although considerable research has been conducted in this field, there is still lack of monitoring algorithm which is reliable, robust, accurate, inexpensive, automatic, practical and independent from the operating conditions. This research was planned to reach this goal and design such a monitoring algorithm with applying and combining conventional signal processing and artificial intelligence methods in the field of machining monitoring and also with the aid of the new signal processing method, S-Transform in this field. Five monitoring algorithms were trained by three different types of signals: AC power, spindle vibration and table vibration signals. Monitoring algorithms were designed based on the pattern recognition concept. It includes four steps: preprocessing, feature extraction, feature selection and classification. All algorithms use the same classification method which is the neural network. Neural network is one of the most accurate artificial intelligent methods for data classification. Various methods for feature extraction and selection were applied. For feature extraction, first, two conventional methods in the field of machining monitoring, descriptors and wavelet transform were implemented and then the new method, S-Transform, were used. Although descriptors are simple features that can be very easily and quickly calculated, the accuracy of the first algorithm is high. In the second and third algorithms, wavelet transform was applied as the feature extraction method to increase the accuracy of the algorithm with analyzing data in both time and frequency domains. The results of training show that these two algorithms are more accurate than the first algorithm. Also, based on the results, it is shown that the third algorithm is more accurate than the second one. This observation reveals that ICA has better efficiency than PCA for selecting the most informative features. To take the advantages of the S-Transform comparing with the wavelet, it has been applied in Algorithm 4 and 5. In the first step, S-Transform of the input signals was calculated in a wide range of frequencies. The best range of frequencies, which shows the variations of the process situation better, was selected to be applied in the next steps of the algorithm. The results show that Algorithm 5, which applies a combination of S-Transform and ICA, is the most accurate algorithm between the designed algorithms. Also from the result it can be concluded that S-Transform has the same or better performance in extracting the features comparing with two other feature extraction methods. In addition to the better performance which increases the accuracy, S-Transform has other positive points comparing with other feature extraction methods. For example, comparing with descriptors which analyze the data in time domain, S-Transform is a time-frequency method and can provide the benefit of seeing the changes happen in the frequency domain. Despite of descriptors, S-Transform also is robust to the noise and its performance would be less suffered if more than one fault would be existed. It also has the absolutely referenced phase information and frequency invariant amplitude response that the continuous wavelet transform does not have. The better performance for feature extraction and the other mentioned positive points make S-Transform's application to signal processing in machining monitoring research more feasible.

With Algorithm 5, the effect of sensor fusion on the accuracy of algorithm has been investigated in Chapter 5. The results confirm that the accuracy of algorithm increases with applying the sensor fusion. The results of Section 5.2 show that when the datasets include signals from two sets of operating conditions rather than one set of operating conditions, the accuracy of algorithm generally decreases. Also, the decrease in accuracy is maximum when the workpiece material changes. After changing in workpiece material, the change in feed decreases the accuracy considerably. Depth of cut decreases the accuracy less than other two parameters. This investigation shows the algorithm is less reliable to work in variable operating conditions. In the rest of Chapter 5, in order to detect sharp tool from the dull one, the definition of sharp and dull tool was added to the algorithm. In view of the results, it can be concluded that the accuracy of the algorithm for indicating the sharp and dull tool is more than the accuracy for predicting the amount of tool wear. The main contribution of this study is summarized as follows:

- Applying conventional and new signal processing and artificial intelligence methods based on the pattern recognition for designing machining process monitoring algorithm.

- Investigating the effect of S-Transform on the accuracy of monitoring algorithm and comparing its feature extraction performance with two other conventional signal processing methods in the field of machining monitoring.

- Using a real benchmark of milling process for training and testing the algorithms and having the good results in monitoring.

- Investigating the effects of sensor fusion and variable operating conditions on the accuracy of the best designed algorithm.

- Developing the designed algorithm for detecting the sharp and dull tool.

- Designing practical monitoring algorithm by applying the sensors which they are practical to use for the industry.

- Meet the goal of the thesis for designing an accurate, automatic, practical and inexpensive monitoring method.

6.2 Future work

Following suggestions can be considered for developing new monitoring algorithms or improving the accuracy of the existing algorithms in future studies:

- Applying new signal processing methods such as fast S-Transform which possesses the advantages of S-Transform while it has less complexity and computations for feature extraction.

- Applying other intelligent methods such as fuzzy logic for signal classification.

- Optimizing the number of features selected by PCA and ICA with optimization methods.

- Combining the features of two signal processing methods such as Wavelet and S-Transform for more accurate representation of the data. Using this approach, the monitoring algorithm can benefit from the both methods.

- Working on the robustness of the algorithm for handling different operating conditions more effectively.

- Application of developed monitoring algorithms to other types of CNC machining tools and processes.

Bibliography

[1] D. I. Nwokah, Y. Hurmuzlu, "The Mechanical System Design Handbook", CRC Press, 2001.

[2] http://www.rdb-engineering.co.uk/turning.html (Accessed on 2013-08-15).

[3] http://www.custompartnet.com/wu/turning (Accessed on 2013-08-1).

[4] http://www.myyellowcoat.com/smart-ideas/facemilling-ideas-performance/ (Accessed on 2013-08-15).

[5] A. G. Rehorn, J. Jiang, P. E. Orban, "State-of-the-art methods and results in tool condition monitoring: a review", International Journal Advanced Manufacturing Technology, vol. 26, pp. 693-710, 2005.

[6] http://www.thetoolanddieguy.com/archives/345 (Accessed on 2013-08-15).

[7] ecom.training.dupont.com (Accessed on 2013-08-15).

[8] http://en.wikipedia.org/wiki/Drilling (Accessed on 2013-08-1).

[9] L. Fu, S. Ling, "Neural network based online detection of drill breakage in micro drilling process", In Proceeding 9th International Conference Neural Information Process. (ICONIP02), pp. 2054-2058, 2002.

[10] H. M. Ertunc, K. A. Loparo, E. Ozdemir, and H. Ocak, "Real time monitoring of tool wear using multiple modeling method", In Proceeding IEEE International Conference Electric Machine Drives, pp. 687-691, 2001.

[11] R. J. Furness, T. Tsao, J. S. Rankin, M. J. Muth, and K. W. Manes, "Torque control for a form tool drilling operation", IEEE Transaction, Control System Technology, vol. 7, pp. 22-30, 1999.

[12] F. Camci, R. B. Chinnam, "Health-state estimation and prognostics in machining processes", IEEE Transaction on Automation Science and Engineering, vol. 7, pp. 581–597, 2010.

[13] http://www.sciencedirect.com/science/article/pii/S0925838802002323 (Accessed on 2013-08-1).

[14] S. Engin, "Metal machining and surface technology", Lecture note, course Mech 6421, Concordia university, winter 2013.

[15] P. Fua, W. Lib, L. Guo, "Fuzzy clustering and visualization analysis of tool wear status recognition", International Conference on Power Electronics and Engineering Application, 2011.

[16] J. F. Dong, K. V. R. Subrahmanyam, Y. S. Wong, G. S. Hong, A. R. Mohanty, "Bayesian-inference-based neural networks for tool wear estimation". International Journal of Advanced Manufacturing Technology, vol. 30, pp. 797-807, 2006.

[17] R. E. Haber, J. R. Alique, S. Ros, "Application of knowledge-based systems for supervision and control of machining processes", In handbook of software engineering and knowledge engineering, vol. 2, pp. 673-710, 2002.

[18] O. Geramifard, J. X. Xu, T. Sicong, J. H. Zhou, X. Li, “A multi-modal hidden Markov model based approach for continuous health assessment in machinery systems”, 37th Annual Conference on IEEE Industrial Electronics Society, 2011.

[19] T. Kalvodaa, Y. R. Hwang, “Analysis of signals for monitoring of nonlinear and non-stationary machining processes”, Sensors and Actuators A 161, pp. 39-45, 2010.

[20] Z. Uros, C. Franc, K. Edi, “Adaptive network based inference system for estimation of flank wear in end-milling”, Journal of Materials Processing Technology, vol. 209, pp. 1504-1511, 2009.

[21] X. Li, H. X. Li, X. P. Guan, R. Du, “Fuzzy estimation of feed-cutting force from current measurementa case study on intelligent tool wear condition monitoring”, IEEE Transactions on Systems, Man, and Cybernetics–Part C: Applications and Reviews, vol. 34, No. 4, 2004.

[22] I. Deiab, Kh. Assaleh, F. Hammad, “On modeling of tool wear using sensor fusion and polynomial classifiers”, Mechanical Systems and Signal Processing, vol. 23, pp. 1719-1729, 2009.

[23] E. Kandilli, M. Sonmez, H. M. Ertunc, B. Cakr, “Online Monitoring of tool wear in drilling and milling by multi-sensor neural network fusion”, Proceedings of the IEEE International Conference on Mechatronics and Automation, 2007.

[24] M. Malekiana, S. S. Parka, M. B. G. Jun, “Tool wear monitoring of micro-milling operations”, Journal of Materials Processing Technology, vol. 209, pp. 4903-4914, 2009.

[25] K. Chen, P. Fu, W. Cao, W. Li, “The application of multiple model fusion based on correctable weight in tool wear pattern recognizing”, Third International Workshop on Advanced Computational Intelligence, 2010.

[26] N. Ghosh, Y. B. Ravi, A. Patra, S. Mukhopadhyay, S. Paul, A. R. Mohanty, A. B. Chattopadhyay, “Estimation of tool wear during CNC milling”, Mechanical Systems and Signal Processing, vol. 21, pp. 466–479, 2007.

[27] H. M. Ertunc and C. Oysu, “Drill wear monitoring using cutting force signals”, Mechatronics, vol. 14, pp. 533–548, 2004.

[28] S. L. Chen, Y. W. Jen, “Data fusion neural network for tool condition monitoring in CNC milling machining”, International Journal of Machine Tools and Manufacture, vol. 40, pp. 381–400, 2000.

[29] P. Baruah, R. B. Chinnam, “HMMs for diagnostics and prognostics in machining processes”, International Journal of Production Research, vol. 43, pp. 1275–1293, 2005.

[30] D. A. Tobon-Mejia, K. Medjaher, N. Zerhouni, “CNC machine tool wear diagnostic and prognostic by using dynamic Bayesian networks”, Mechanical Systems and Signal Processing, vol. 28, pp. 167-182, 2012.

[31] O. Geramifard, J.X. Xu, T. Sicong, J.H. Zhou, X. Li, “Continuous health condition monitoring: a single hidden semi-Markov model approach”, Proceedings of the IEEE International Conference Prognosis and Health Management, 2011 .

[32] A. J. Vallejo, R. M. Menendez, J.R. Alique, “Intelligent monitoring and decision control system for peripheral milling process”, Proceeding of IEEE International Conference on Systems, Man and Cybernetics, 2008.

[33] A. Kumar, F.Tseng, R. B. Chinnam, “Role of hidden-Markov models for autonomous diagnostics of cutting tools”, Innovations in Intelligent Systems and Applications (INISTA), International Symposium on IEEE, 2011.

[34] P. T. Huang, J. C. Chen, “Neural network-based tool breakage monitoring system for end milling operations”, Journal of Industrial Technology, Vol. 16, 2000.

[35] T. J. Ko, D. W. Cho and M. Y. Jung, “On-line monitoring of tool breakage in face milling using a self-organized neural network”, Journal of Manufacturing Systems, Vol. 14, No. 2, pp. 80-90, 1995.

[36] S. Tangjitsitcharoen, “In-process monitoring and detection of chip formation and chatter for CNC turning”, Journal of Materials Processing Technology, vol. 209, pp. 4682-4688, 2009.

[37] N. Pongsathornwiwat, S. Tangjitsitcharoen, “Intelligent monitoring and detection of chatter in ball-end milling process on CNC machining center”. Computers and Industrial Engineering (CIE), 40th International Conference on IEEE, 2010.

[38] E. Kuljanic, M. Sortino, G. Totis, “Multisensor approaches for chatter detection in milling”, Journal of Sound and Vibration, vol. 312, pp. 672-693, 2008.

[39] J. Z. Zhang, J. C. Chen, “The development of an in-process surface roughness adaptive control system in end milling operations”, International Journal of Advanced Manufacturing Technology, vol. 31, pp. 877-887, 2007.

[40] P. G. Benardos, G. C. Vosniakos, “Prediction of surface roughness in CNC face milling using neural networks and Taguchis design of experiments”, Robotics and Computer-Integrated Manufacturing, vol. 18, pp. 343-354, 2002.

[41] R. Azouzi, M. Guillot, “On-line prediction of surface finish and dimensional deviation in turning using neural network based sensor fusion”, International Journal of Machine Tools Manufacturing, vol. 37, pp. 1201-1217, 1997.

[42] A. E. Ouafi, M. Guillot, A. Bedrouni, “Accuracy enhancement of multi-axis CNC machines through on-line neuro compensation”, Journal Intelligent Manufacturing, vol. 11, pp. 535-545, 2000.

[43] http://www.kistler.com/ca/en/product/force/9257B (Accessed on 2013-08-15).

[44] K. Jemielniak, J. Kosmol, “Tool and process monitoringstate of art and future prospects”, In Scientific Papers of the Institute of Mechanical Engineering and Automation of the Technical University of Wroclaw, vol. 61, pp. 90-112 , 1995.

[45] S. Dey, J. A. Stori, “A Bayesian network approach to root cause diagnosis of process variations”, International Journal of Machining Tools and Manufacturing, vol. 45, pp. 75-91, 2005.

[46] J. C. Chen, W. L. Chen, “A tool breakage detection system using an accelerometer sensor”, Journal of Intelligent Manufacturing, vol. 10, pp. 187-197, 1999.

[47] B. Samanta, “Surface roughness prediction in machining using soft computing”, International Journal of Computer Integrated Manufacturing, vol. 22, pp. 257–266, 2009.

[48] D. Y. Jang, Y. G. Choi, H. G. Kim, A. Hsiao, “Study of the correlation between surface roughness and cutting vibrations to develop an on-line roughness measuring technique in hard turning”, International Journal of Machining Tools and Manufacturing, vol. 36, pp. 453-464, 1996.

[49] J. C. Chen, L. H. Huang, A. X. Lan, S. Lee, “Analysis of an effective sensing location for an in-process surface recognition system in turning operations”, Journal of Industrial Technology, vol. 15, pp. 1-6, 1999.

[50] N. R. Abburi, U. S. Dixit, “A knowledge-based system for the prediction of surface roughness in turning process”, Journal of Robotics Computer Integrated Manufacturing, vol. 22, pp. 363-372, 2006.

[51] O. B. Abouelatta, J. Madl, "Surface roughness prediction based on cutting parameters and tool vibrations in turning operations". Journal of Material Processing Technology, vol. 118, pp. 269-277, 2001.

[52] S. S. Lee, J. C. Chen, "On-line surface roughness recognition system using artificial neural networks system in turning operations", International Journal of Advanced Manufacturing Technology, vol. 22, pp. 498-509, 2003.

[53] Y. M. Ertekin, Y. Kwon, T. L. Tseng, "Identification of common sensory features for the control of CNC milling operations under varying cutting conditions", International Journal of Machine Tools Manufacturing, vol. 43, pp. 897-904, 2003.

[54] http://www.kistler.com/ca/en/product/acceleration/8688A10 (Accessed on 2013-08-15).

[55] B. Bahr, S. Motavalli, T. Arfi, "Sensor fusion for monitoring machine tool conditions", International Journal Computer Integrated Manufacturing, vol. 10, pp. 314-323, 1997.

[56] Y. Zhou, P. Orban, S. Nikumb, "Sensors for intelligent machining-a research and application survey", In systems, Man and Cybernetics IEEE International Conference on Intelligent Systems for the 21st century, vol. 2, pp. 1005-1010, 1995.

[57] http://www.intertechnology.com/Kistler/Accelerometers_Model_8152B.htm (Accessed on 2013-08-15).

[58] X. L. Li, "A brief review: acoustic emission method for tool wear monitoring during turning", International Journal of Machine Tools Manufacturing, vol. 42, pp. 157-165, 2002.

[59] S. Y. Liang, R. L. Hecker, R. G. Landers, “Machining process monitoring and control: the state-of-the-art”, Journal Manufacturing Science E-T ASME, vol. 126, pp. 297-310, 2004.

[60] K. Jemielniak, “Some aspects of AE application in tool condition monitoring”, Ultrasonics, vol. 38, pp. 604-608, 2000.

[61] H. Cao, X. Chen, Y. Zi, F. Ding, H. Chen, J. Tan, Z. He, “End milling tool breakage detection using lifting scheme and Mahalanobis distance”, International Journal of Machine Tools and Manufacture, vol. 48, pp. 141-151, 2008.

[62] H. Wang, H. Shao, M. Chen, D. Hu, “On-line tool breakage monitoring in turning”, School of Mechanical Engineering, Shanghai Jiao Tong University, 1954.

[63] X. Li, D. Shen, Z. Yuan, “Discrete wavelet transform for tool breakage monitoring”, International Journal of Machine Tools and Manufacture, vol. 39, pp. 1935-1944, 1999.

[64] R. E. Haber, J. E. Jimenez, C. R. Peres, J. R. Alique, “An investigation of tool-wear monitoring in a high-speed machining process”, Sensors and Actuators A-Physical, vol. 116, pp. 539-545, 2004.

[65] A. R. Mohanty, A. B. Chattopadhyay, “Estimation of tool wear during CNC milling using neural network-based sensor fusion”, Mechanical System and Signal Processing, vol. 21, pp. 466-479, 2007.

[66] J. Kopac and S. Sali, “Tool wear monitoring during the turning process”, Journal of Materials Processing Technology, vol. 113, pp. 312–316, 2001.

[67] C. Scheffer, H. Kratz, P.S. Heyns, F. Klocke, “Development of a tool wear-monitoring system for hard turning”, International Journal of Machine Tools and Manufacture, vol. 43, 2003.

[68] R. T. Rene de Jesus, H. R. Gilberto, T. V. Ivan, J. C. J. Carlos, “Driver current analysis for sensorless tool breakage monitoring of CNC milling machines”, International Journal of Machine Tools and Manufacture, vol. 43, pp. 1529–1534, 2003.

[69] X. Li, G. Ouyang, Z. Liang, “Complexity measure of motor current signals for tool flute breakage detection in end milling”, International Journal of Machine Tools and Manufacture, vol. 48, pp. 371-379, 2008.

[70] X. Li, “On-line detection of the breakage of small diameter drills using current signature wavelet transform”, International Journal of Machine Tools and Manufacture, vol. 39, pp. 157-164, 1999.

[71] Z. Yao, D. Mei, Z. Chen, “On-line chatter detection and identification based on wavelet and support vector machine”, Journal of Materials Processing Technology, vol. 210, pp. 713-719, 2010.

[72] E. Kuljanic, G. Totis, M. Sortino, “Development of an intelligent multi sensor chatter detection system in milling”, Mechanical Systems and Signal Processing, vol. 23.5, pp. 1704-1718, 2009.

[73] L. Wang, M. Liang, “Chatter detection based on probability distribution of wavelet modulus maxima”, Robotics and Computer-Integrated Manufacturing, vol. 25, pp. 989-998, 2009.

[74] A. G. Rehorn, E. Sejdic, J. Jiang, “Fault diagnosis in machine tools using selective regional correlation”, Mechanical Systems and Signal Processing, vol. 20, pp. 1221–1238, 2006.

[75] M. Elangovan, S. Babu Devasenapati, N. R. Sakthivel, K. I. Ramachandran, “Evaluation of expert system for condition monitoring of a single point cutting tool using principle component analysis and decision tree algorithm”, Expert Systems with Applications, vol. 38, pp. 4450-4459, 2011.

[76] Z. Uros, C. Franc, K. Edi, “Adaptive network based inference system for estimation of flank wear in end-milling”, Journal of Materials Processing Technology, vol. 209, pp. 1504-1511, 2009.

[77] R. M. Menendez, S. Aguilar, C. A. Rodrguez, F. G. Elizalde, L. E. G. Castanon, “Sensor-fusion system for monitoring a CNC-milling center”, Lecture Notes in Computer Science, vol. 3789, pp. 1164–1174, 2005.

[78] P. Fua, We. Lib, L. Guo, “Fuzzy clustering and visualization analysis of tool wear status recognition”, International Conference on Power Electronics and Engineering Application, 2011.

[79] F. Dweiri, M. Al-Jarrah, H. Al-Wedyan, “Fuzzy surface roughness modeling of CNC down milling of alumic-79”, Journal of Material Processing Technology, vol. 133, pp. 266-275, 2003.

[80] S. P. Lo, “An adaptive-network based fuzzy inference system for prediction of workpiece surface roughness in end milling”, Journal of Material Processing Technology, vol. 142, pp. 665-675, 2003.

[81] M. Elangovan, K. I. Ramachandran, V. Sugumaran, “Studies on Bayes classifier for condition monitoring of single point carbide tipped tool based on statistical and histogram features”, Expert Systems with Applications, vol. 37, pp. 2059-2065, 2010.

[82] H. M. Ertunc, C. Oysu, “Drill wear monitoring using cutting force signals”, Mechatronics, vol. 14, pp. 533–548, 2004.

[83] K. Zhu, Y. S. Wong , G. S. Hong, “Multi-category micro-milling tool wear monitoring with continuous hidden Markov models”, Mechanical Systems and Signal Processing, vol. 23, pp. 547-560, 2009.

[84] O. Colak, C. Kurbanoglu, M. C. Kayacan, “Milling surface roughness prediction using evolutionary programming methods”, Material and Design, vol. 28, pp. 657-666, 2007.

[85] M. Brezocnik, M. Kovacic, M. Ficko, “Prediction of surface roughness with genetic programming”, Journal of Material Processing Technology, vol. 157-158, pp. 28-36, 2004.

[86] Y. Qian, J. Tian, L. Liu, Y. Zhang, Y. Chen, “A Tool wear predictive model based on SVM”, International conference Control and Decision conference IEEE, 2010.

[87] Y. W. Hsueh, C. Y. Yang, “Tool breakage diagnosis in face milling by support vector machine”, Journal of Materials Processing Technology, vol. 209, pp. 145-152, 2009.

[88] S. Cho, S. Asfour, A. Onar, N. Kaundinya, “Tool breakage detection using support vector machine learning in a milling process”, International Journal of Machine Tools Manufacturing, vol. 45, pp. 241-249, 2005.

[89] D. Dinakaran, S. Sampathkumar, N. Sivashanmugam, “An experimental investigation on monitoring of crater wear in turning using ultrasonic technique”, International Journal of Machine Tools and Manufacture, vol. 49, pp. 1234-1237, 2009.

[90] http://ti.arc.nasa.gov/tech/dash/pcoe/prognostic-data-repository/ (Accessed on 2013-08-1).

[91] MATLAB help/ Statistics Toolbox/User Guide/Descriptive Statistics

[92] V. Varalakshmi, “Statistics Higher Secondary”, College Road , Chennai- 600 006.

[93] http://en.wikipedia.org/wiki/Central_tendency (Accessed on 2013-08-1).

[94] J. H. Klotz, “A computational approach to statistics”, Department of Statistics University of Wisconsin at Madison, 2004.

[95] L. Denath, “Wavelet transforms and their applications”, Edward Brothers Incorporation, 2002.

[96] Z. K. Peng, F. L. Chu, “Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography”, Mechanical Systems and Signal Processing, vol. 18, pp. 199-221, 2004.

[97] I. Daubechies, “Ten Lectures on Wavelets”, Society for Industrial and Applied Mathematics, 1992, ISBN 0-89871-274-2.

[98] S. Mallat, “A Wavelet Tour of Signal Processing”, 2nd Edition, Academic Press, 1999, ISBN 0-12-466606-X.

[99] D. Sripath, “Efficient implementations of discrete wavelet transforms using FPGAs”, Master of Science, The Florida State University, USA, 2003.

[100] R.G. Stockwell, “Why use the S-Transform?”, Mathematics Subject Classification, 1991.

[101] J. V. Abellan-Nebot, F. R. Subirn, “A review of machining monitoring systems based on artificial intelligence process models”, International Journal of Advanced Manufacturing Technology, vol. 47, no. 1-4, pp. 237–257, 2010.

[102] http://en.wikipedia.org/wiki/Feature_selection (Accessed on 2013-08-1).

[103] K. Pearson, “On lines and planes of closest fit to systems of points in space”, Philosophical Magazine, vol. 2, pp. 559-572, 1901.

[104] H. Hotelling, “Analysis of a complex of statistical variables into principal components”, Journal of Educational Psychology, vol. 24, 417–441, 1933.

[105] H. Abdi., L. J. Williams, “Principal component analysis”, Wiley Interdisciplinary Reviews: Computational Statistics, vol. 2, pp. 433-459, 2010.

[106] http://en.wikipedia.org/wiki/Principal_component_analysis (Accessed on 2013-08-1).

[107] https://onlinecourses.science.psu.edu/stat857/book/export/html/11 (Accessed on 2013-08-1).

[108] A. Hyvrinen, E. Oja, "Independent component analysis: algorithms and application", Neural Networks, vol. 13, pp. 411–430, 2000.

[109] L. Dominic, S. Chartier, D. Gosselin, "An introduction to independent component analysis: InfoMax and Fast ICA algorithms", Tutorials in Quantitative Methods for Psychology, vol. 6.1, pp. 31–38, 2010.

[110] B. Pczos, "Introduction to independent component analysis", Lecture notes, university of Alberta, Nov. 2009.

[111] http://en.wikipedia.org/wiki/Classification_(machine_learning) (Accessed on 2013-08-1).

[112] E. Hosseini, Y. Zhang, Z. Tian "Comparison of parallel and single neural networks in heart arrhythmia detection by using ECG signal analysis", Annual Conference of the Prognostics and Health Management Society, 2011.

[113] http://en.wikipedia.org/wiki/Artificial_neural_network (Accessed on 2013-08-1).

[114] http://radio.feld.cvut.cz/matlab/toolbox/nnet/backpr52.html (Accessed on 2013-08-1).

[115] http://www.cse.unsw.edu.au/ billw/mldict.html (Accessed on 2013-08-1).

[116] www.geocomputation.org (Accessed on 2013-08-1).

[117] en.wikibooks.org (Accessed on 2013-08-1).

[118] A. Magaly de Paula Canuto, "Combining neural networks and fuzzy logic for applications in character recognition", PhD thesis, University of Kent, 2001.

[119] http://person.hst.aau.dk/enk/ST8/wavelet_tutotial.pdf. (Accessed on 2013-08-1).

[120] http://www.codeforge.com/article/33451 (Accessed on 2013-08-1).

[121] P. Davim, "Machining Fundamentals and Recent Advances", Springer, 2008.