New Jersey Institute of Technology [Digital Commons @ NJIT](https://digitalcommons.njit.edu/)

[Theses](https://digitalcommons.njit.edu/theses) [Electronic Theses and Dissertations](https://digitalcommons.njit.edu/etd)

1-31-1991

Development of techniques of artificial intelligence in identifying the different machined surfaces

Yi-Ching Chou New Jersey Institute of Technology

Follow this and additional works at: [https://digitalcommons.njit.edu/theses](https://digitalcommons.njit.edu/theses?utm_source=digitalcommons.njit.edu%2Ftheses%2F2466&utm_medium=PDF&utm_campaign=PDFCoverPages)

Part of the Manufacturing Commons

Recommended Citation

Chou, Yi-Ching, "Development of techniques of artificial intelligence in identifying the different machined surfaces" (1991). Theses. 2466. [https://digitalcommons.njit.edu/theses/2466](https://digitalcommons.njit.edu/theses/2466?utm_source=digitalcommons.njit.edu%2Ftheses%2F2466&utm_medium=PDF&utm_campaign=PDFCoverPages)

This Thesis is brought to you for free and open access by the Electronic Theses and Dissertations at Digital Commons @ NJIT. It has been accepted for inclusion in Theses by an authorized administrator of Digital Commons @ NJIT. For more information, please contact digitalcommons@njit.edu.

Copyright Warning & Restrictions

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a, user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use" that user may be liable for copyright infringement,

This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve violation of copyright law.

Please Note: The author retains the copyright while the New Jersey Institute of Technology reserves the right to distribute this thesis or dissertation

Printing note: If you do not wish to print this page, then select "Pages from: first page $#$ to: last page $#$ " on the print dialog screen

The Van Houten library has removed some of the personal information and all signatures from the approval page and biographical sketches of theses and dissertations in order to protect the identity of NJIT graduates and faculty.

ABSTRACT

Title of thesis: Development of techniques of artificial intelligence in indentifying the different machined surfaces

Yi-Ching Chou, Master of Science in Manufacturing Engineering, Jan. 1991

Thesis directed by: Dr. Nouri Levy Associate Professor Department of Mechanical and Industrial Engineering

This thesis deals with the identification of different machined surfaces with artificial inteligence techniques. There are five kinds of machining operations: grinding, turning, milling, EDM and waterjet which determine the characteristics of resulting surfaces. Topology of various specimens is measured using the surfanylizer. Twenty one parameters are extracted from the measurements to describe the characteristics. In this thesis, we develop neural and expert systems that recoginize and classify surfaces according to the machining methods.

$\frac{1}{2}$ DEVELOPMENT OF TECHNIQUES OF ARTIFICIAL INTELLIGENCE IN INDENTIFYING THE DIFFERENT MACHINED SURFACES

by YI-CHING CHOU

Thesis submitted to the Faculty of the Graduate School of the New Jersey Institute of Technology in partial fulfillment of the requirements for the degree of Master of Science in Manufacturing Engineering

 1991

APPROVAL

Title of Thesis: Development of Techniques of Artificial Intelligence in identifying the different machined surfaces.

Name of Candidate: Yi-ching Chou Master of Science in Manufacturing Engineering, 1991

Thesis and Abstract Approved:

Dr. Nouri Levy, Advisor Date Associate Professor Department of Mechanical and Industrial Engineering

Signatures of other members:

Signatures of other members: $\frac{1}{\pi}$ $\frac{10}{19}$ /991
of the thesis committee. Dr. Keith O'Brien Date Professor Department of Mechanical and Industrial Engineering

Dr. R. Y. Chen Date $\frac{\sqrt{a_{\mu} \sqrt{c_{\mu}}} \sqrt{f}}{2}$ Professor Department of Mechanical and Industrial Engineering

1991

VITA

Name: Yi-Ching Chou

Permanent address:

Degree and date to be conferred: M.S.M.E., Jan. 1991

Date of birth:

Place of birth:

Major: Manufacturing Engineering

ACKNOWLEDGEMENT

First, I would like to express my sincere appreciation to my advisor, Dr. Nouri Levy, Associate Professor, New Jersey Institute of Technology, for his guidance and advice. Without his valuable source and equipment, this work could not have been done so quickly, and successfully. Also, I would like to thank Mr. Zhou G. Y. for his help in designing the measurement apparatus and choosing surface quality specifications.

I am also very thankful to Dr. Keith O'brien and Dr.R. Chen for devoting their precious time to reviewing my work and providing constructive suggestion.

It is also a pleasure to express my thanks to the people who have been of assistance to me in the preparation of this thesis. Finally, thanks to my wife. parents, in—laws, sister and brothers that I would like to dedicate this thesis to them for their support.

Contents

1 INTRODUCTION 1

 $\ddot{}$

List of Tables

 $\overline{}$ m,

List of Figures

Chapter 1 INTRODUCTION

In industry, there are various machining methods resulting in different finished surfaces which can be controlled or designed to meet specific requirements. Frequently. finished surfaces which are produced by different machining methods cannot be distinguished easily and need very experienced experts for proper identification of the method of machining used.

The American National Standard ANSI B46.1-1978 has provided a set of surface quantitative parameters derived from specific measurements to descirbe surface texture[1]. By means of these parameters, machined surfaces can be described. But there is no standard procedure that can determine how surface texture is related to the method of machining. The difficulty in identifying the different machined surfaces is that the value of the parpmeters describing the surfaces often overlap. Therefore, it is necessary to devise methods to integrate all parameters and extricate the patterns that characterize each machining method.

In artificial intelligence field, there are two particular techniques, *neural network* and *expert system.* Neural networks embody gestalt functions such as the understanding of images, pictures, and graphic information. Expert systems are typified by logical functions such as rules, concepts, and calculations[2]. Knowledge incoporated in an expert system is made explicit in the form of rules which are based on human expertise. Unlike traditional expert systems, neural networks generate their own rules by learning from being shown examples.

In this thesis, we use profilometer probe to measure 25 profiles on each machined specimen. The obtained signal is amplified and then sent to a microprocessor where A/D conversion, digital filtering and some signal treatment are done. The data is then transmitted to an IBM—PC for data analysis and storage. After calculation, each reading will be represented by 21 parameters. Then, we try to use *neural network* and *expert system* techniques to build identification systems which will identify the method of machining of an examined surface. The neural network learning is achieved through backpropargation method. The expert system will adopt rules which are based on statistical properties of the parameters. Finally, we use some readings in a test set to evaluate their performance.

Chapter 2 NEURAL NETWORK AND EXPERT SYSTEM

In this chapter, we describe the background of two artificial intelligence techniques, *neural network* and *expert system.*

2.1 What is a Neural Network?

Neural Network is composed of a number of very simple processing elements that communicate through a rich set of interconnections with variable weights or strengths. These simple processing elements, are called *neurodes.*

Real neurodes, such as those in the human brain, are extremely complicated devices, with a myriad of parts, subsystems, and control mechanisms. They use a variety of electrochemical pathways to communicate each other. Each neurode receives a large number of individual input signals that together constitute an input pattern. This input pattern causes the neurode to reach some level of activity. If the activity is strong enough, the neurode, like its biological analog, generates a signal. Output signal transmits over the neurode's interconnects to other neurodes or to the outside world. For all intents and purposes, the neurodes computes a mapping, or a function that associates a given level of input stimulus to a particular level of output activity. This overall mapping is usually called the transfer function for neurodes.

Figure 2.1 shows a basic neurode. The inputs are combined into a signed signal value in the box labeled signed Weighted Sum of Inputs. This result is the Total Input. The Total Input is run through another function known as the Activation Function. The network generally takes the activation function to he the identity mapping, so that the neurode's total input and activation values are identical.

Next, the activation value is passed through a function called the Neuron Transfer Function, which produces the neuron's output. The interesting behaviors of neural network are substantially dependent on the character of the neuron transfer function. By means of transfer function, the critical factor which determines the nature of the activity patterns is called *Synapse Weight[3].*

Learning in a neural network consists of changing the value of these weights. The learning laws, which dictate by how much and in which directions the weights will be changed are the most important in a designer's tool box. In application, we implement *backpropagatzon network* to the identification system. Backward error propagation is the most successful learning law in many fields. In next subsection, it will be described in detail.

2.1.1 Features of Backpropagation Systems

Backpropagation is a supervised learning scheme by which a layered feed forward network, and the error generated by network's output will produce another information to feed back to the input layer immediately. For this important feature, hierarchical structure and nonlinear transfer function are essential. At the minimum, the network must be composed of three layers: input layer, middle layer, and output layer. The input layer accepts pattern from the outside world, the output layer presents the network's response back to the outside world, and the middle layer inter-connects with them.

Figure 2.1: The internal representation of neurodes.

 $\bar{\bar{t}}$

The middle layer is also called hidden layer which creates a map relating each input pattern to a unique output response. The higher the interlayer connectivity , the better the ability of network to build a representation or a model of the input data[4]. Figure 2.2. shows a simple neural network architecture.

The input, activation, and output functions of a neurode can usually combined into one function, the neurode transfer function. The learning rule corresponds to gradient minimization of the average of error E(this average is computed over all input vectors in a specific training set).

$$
E = \sum_{i=1}^{n} (D_i - R_i)^2
$$
 (21)

 $n =$ number of training facts

 $D =$ Desire output

 $R = Real$ output from network

The difficulty is to get the minimum error. First, let us see how to calculate the error. The weighted output error is propagated from each neurode in the output layer back to the neurodes of the middle layer. The weights used to compute these errors are those on the interconnects between the various neurodes. Thus, each middle-layer neurode can calculate exactly how big a contributation it made to the error of each neurode in the output layer. This means that an output neurode with a large weight from a particular neurode in the middle layer can probably "blame" much of its output error on input signals from that particular middle-layer neurodes. Thus a large weight and large error in the output layer will generate a large contributation to that middle-layer neurode's own error. If a middle-layer neurode has a large weight that produce a large output error on interconnection with output-layer neurode, the next step is to generate a large weight change for that middle-layer neurode even if the middle--layer neurode had generated a very

 $\overline{}$

Figure 2.2: Example of a neural network architecture.

small or even zero output.

To use derivative of neurode activation function with the address on weight change will cause the maximum change. The derivative of a function can be thought of as the slope of the function at each point along its curve. A negative slope means that the function decreases in value as the value of its variable increases. It is why we choose the negative derivative of neurode active function.

In computer simulation, using numerical analysis to calculate the derative of function is time consuming. The best way is to find the exact form of derative function. Sigmoid function own this properties [5].

$$
NET = X_1 W_1 + X_2 W_2 + \dots + X_n W_n
$$
 (2.2)

$$
OUT = \frac{1}{1 + e^{-NET}}\tag{2.3}
$$

$$
OUT' = OUT^2 \times e^{-NET} \tag{2.4}
$$

 $X = input vector$

 $W =$ weight of neuron

 $OUT = output of neuron from sigmoid transfer function$

 $OUT' = the derivative of OUT$

Moreover, the derivative , or slope, of sigmoid curve is shaped like a bell, with the largest values in the middle range of summed inputs and very small values at both extremes, as shown in Figure 2.3.

Multiplying the error correction by this bell—shaped derivative function ensures us that the weights on the neurode we are considering will change on slightly as the summed input approaches either very low or very high values. Because of the two advantages, we know that choosing sigmod not only prevent unfairly blamming mid—layers neurodes for errors in the output layer, but also ensure the stability of the network as a whole. For each mid—layer neurode, we must propagate the error

 $\frac{1}{2}$

Figure 2.3: Sigmoid function, transferfunction of back propergation network.

from the output layer back to it, compute the sum of the weighted errors, multiply it by the derivative of the sigmoid function, and then use this computed error value in the delta rule to modify the weights. When these computations have been made for all of the mid—layer neurodes, this particular iteration of the network is completed.

2.1.2 Building a Neural Network for Surface Identification

The network will learn to recognize surfaces by learning to associate readings from samples with the machining method they represent. Using surface assessment method based on ANSI standard, we collect readings which are tranformed into 21 parameters to adequately characterize the machined surfaces. Then we choose a set of readings as the training set. Another set of readings is used for testing the performance of the network after training.

2.2 What is an Expert System?

An *expert system,* also called Knowledge—Based System [KBS], is a computer software that uses knowledge and inference procedure to solve specific classes of problems.

Expert Systems have been implemented to solve problems, ranging from medical diagnosis to trouble shooting generator, locomotives, and telephone networks. Rule—based systems use simple if—then rules to express the domain knowledge, and rule interpreter that links together such rules into a complex and long chain of inference. Much of the knowledge in expert system is heuristic in nature, and consists of rules of thumb[6].

All expertise is built up of small, modular chunks of knowledge which include facts and rules, and the overall intelligent behavior emerges from the combination of many chunks.

Each rule has the form:

IF \langle list of premises $>$ THEN \langle conclusion $>$.

For example:

If it is raining, then the ground is wet.

The performance level of an expert system primarily depends on a function of the rule size and quality of the knowledge base that it processes.

Figure 2.4 shows the most important modules that makes up a general rule based expert system.

2.2.1 Use of Expert System

An expert system is a knowledge—based program that provides "expert quality" solutions to problems in a specific domain. Generally, its knowledge is extracted from human experts in the domain and it attempts to emulate their methodology and performance.

Knowledge from countless fields can be used in an expert system. The main criterian for determining whether or not a subject is the answer to this question: Can knowledge on this subject be translated? If answer is yes, this subject is suitable to be used in an expert system[7].

In the nomenclature common to the artificial intelligence literature, these three components of problem—solving systems are known as:

- 1. The database.
- 2. The operators or production rules.
- 3. The control strategies[8].

Expert systems do not simulate human mental architecture in detail, like neural network. They are practical programs that use *heuristic* strategies developed by humans to solve specific classes of problems.

Figure 2.4: Architecture of a typical expert system.

2.2.2 Expert System Technology

Heuristics

Heuristics are criteria, methods, or principles for deciding which among several alternatives courses of action promises to be the most effective in order to achieve some goal. They represent compromises between two requirements: the need to make such criteria simple and , at the same time , the desire to see them discriminate correctly between good and bad choices[8].

It is the nature of good heuristics both that they provides a simple means of indication which among several courses of a action is to he prefered, and that they are not necessarily guaranteed to identify the most effective course of action, but do so sufficiently often[9].

The time required to find out an exact solution always takes a long time for some problem. Heuristics play an efective role in such problems by indicating a way to reduce the number of evaluation and to obtain solutions within reasonable time constraints.

Probability—Based Heuristics

Probability is a language for quantifying the unpredictability of our environment. Probabilistic modles specify a set of possible instances of the environment together with their associated degree of likelihood[8].

Because the ultimate test for the success of heuristic methods is that they work well "most of the time", and because probability theory is our principal formalism for quantifying concepts such as "most of the time", it is not only natural that probabilistic models should provide a formal ground for quantitatively evaluating the performance of heuristic methods. More over, it is equally natural that probabilistic model, whenever available, be consulted in the process of devising heuristic methods, or in selecting the parameters that govern these methods, so as to guarantee that only a small fraction of problem will escape an adequate treatment.

Using Uncertainty and Certainty Measures

First, consider the rule as follow:

IF

Student is sick.

THEN

Student would not be present in class.

Clearly, this rule is heuristic in nature. For the converse of the above rule is an implication:

IF

Student would not be present in class.

THEN

Student is sick.

This is an example of *abductive* reasoning. Formally, abduction states that: from P \Rightarrow Q, it is possible to infer P by Q[10].

Abduction is an unsound rule of inference, meaning that the conclusion is not necessarily true for every interpretation in which the premises are true. Like this example, it is possible that student have traffic problems, sleep too late, etc.

Although abduction is unsound, it is often essentional to solving problems. Such as this problem, Modus pones cannot be applied and rule must be used in an abductive fashion. Measuring uncertainty is adapted in this situation.

Uncertainty results from the use of abductive inference as well as from attempts to reason with missing or unreliable data. To get around this problems, we can attatch some measure of confidence to the conclusion.

Confidence factor

Confidence factor is the degree of confidence one has in a fact or relationship. As used in artificial intelligence, it contrasts with probabilities, which is the

14

likelihood than an event will occur[11].

Certain theory makes some simple assumption for creating confidence measures and has some equally simple rules for combining these confidence as the program moves toward its conclusion.

When experts put together the rule base, they must agree on a CF to go with each rule. This CF reflects their confidence in the rule's relibility. Certainty measures may be adjusted to tune the system performance, although slight variations in the confidence measure tend to have little effect on the overall running of the system.

2.2.3 Building an Expert System for Surface Identification

First, we analyse all numeric value of the parameters which are collected by measuring the surface topography. Second, we incoperate the observations in two rules. Confidence factors are calculated and are associated with the conclusions of each rule. By the combination of confidence factors ,the expert system can successfully identify different machined surfaces.

In chapter 6, we will describe how to build the identification expert system in detail.

Chapter 3

MEASUREMENT APPARATUS AND MEASUREMENT PROCEDURE

3.1 Apparatus

The measurement system is composed of three major units:

- 1. Daedal moving table control unit.
- 2. Surfanaylizer-4000 unit.
- 3. Dell computer unit.

As shown in Figure 3.1. When the probe of a profilmeter traces through the specimen surface, the signal is received and amplified. The signal is then processed through a microprocessor, A/D converter and digital filter. The processed data are then transmitted to PC and stored on a data file.

3.2 Measurements Errors

Measurements errors should be minimized by taking the following precautions[12]:

• Surface to be measured should be free of any material that would interfere with the measurement.

Figure 3.1: Measuring system

- Before measuring, any flaw should be avoided.
- In each reading of per profile, due to the signal procession delay, the front multi—hundred points data should be truncated.
- Because the signal analysis is subjected to the limitations of low and high precision voltmeter, the precision should be timely changed according to a different range of roughness.

3.3 Measurement Procedure

Based on the measuring system, surfaces can be analysed systematically by three methods: single profile method, multi-parallel profile method, and random process method. A general typology of solid surfaces and the analysis method are described in figue 3.2.

From G.Y. Zhou's work[13], we know EDM, waterjet and grinding machined surfaces belong to isotropic surfaces. Shaping, milling, and turning machined surfaces are classified as determinstic ones. So we adopts multi—parallel profile method, is indicated in Figure 3.2.

As shown in Figure 3.1, the specimen is mounted on table, using stepping motor controller to move specimen from one profile to another profile. The process requires two steps: First, the probe should go back to the original point, as shown in Figure 3.3. Second, each profile will have definite span with near ones. The position control is accomplished by two position sensors, daedal positioning table and control unit.

For each specimen, all test data will be recorded into one file. The more details of testing steps are as follows:

- Clear specimen surface.
- Mount it on table.

١

a f

Figure 3.2: The classification of solid surfaces

Ì.

 $\frac{1}{2}$

Figure 3.3: The track of stylus

- Turn on surfanaylizer (amplifier, microprocessor and recorder).
- set up environment.
- Turn on stepping motor controller and stylus controller (daedal moving table controller).
- Turn on profilemeter (drive unit and return control).
- Turn on PC (dell computer center command system, monitor and printer).
- Execute test program.
- Operate 25 readings for each specimen.
- Exchange specimen and repeat execuation of measurement program.

3.4 Measurement Environment Setup

To measure a solid surface, a common method is to move a sharp—pointed stylus over the surface and to translate its motion perpendicular to the surface into a meter reading proportional to the roughness average. Some procedures should be carefully followed in order to obtain accurate results:

- 1. To ensure that stylus follows exactly the contour of the surface being measured, a force is needed to push it against the surface. Values of stylus force have been specified which will enable roughness measurement to be made accurately on materials. For soft material, it is a special case. But here, all test materials belong to hard ones.
- 2. During the testing, a sufficient length of surface must be traversed to ensure that the full reading characteristic of the surface is obtained. Besides, the front 400 sample points are deleted to ensure the data, to be accurate.
- 3. Because the roughness reading may vary depending on the location of the sampling profile on the surface, 25 measurements per specimen are processed in order to get adequate measurements.
- 4. In general, surface will contain irregularities with a large range of widths. Instruments are designed to respond only to irregularities with a large range of width. When irregularities of small width tend to he important, more significant values will be obtained when a small cutoff value is used [12].

According to those notes, the environment setups are as follow:

- Cutoff $= 0.8$ mm.
- Drive speed $= 0.25$ mm/s.
- Sampling length = 5, Cutoff = 1, Profile = 1.
- Travel length $= 4$ mm.
- Sampling points of whole length $= 3000$ for one profile.
- Measure 25 profile. Then, *points* \times *profile* = 3000 \times 25 = 75,000.
- Sampling points interval = 1.25 μ m(0.25mm/s), 12.5 μ m(2.5m/s)
- Profile span = 40μ m (motor moves 80 steps)
- Measuring area $=4 \times 1 mm^2$
- Motor steps $= 10,000$ /rev
- Linear motion table resolution: 5μ m

Figure 2.4 shows the sampling grid for measuring specimen.

Figure 3.4: Sampling grid for multi-profile method.

Chapter 4

GENERATION AND PRE-PROCESSING OF DATA

In this chapter, detail specifications of surface characteristics are described. For more detail, please see ASME standard committee B46—Classification and Designation of Surface Quantities. The descriptions of the specimens used from different machining methods, are shown in Appendix A.

Twenty one parameters, shown below, are calculated from each measurement and are used to describe the surface.

P0: Roughness Average

P1: Root—Mean—Square(rms) Roughness

P2: Ten—Point Height . The average distance between the five highest peaks and the five deepest valleys within the sampling measured from a line parallel to the mean line and not crossing the profile.

P3: Maximum Peak—to—Valley Roughness Height .

This is the distance between two line parallel to the mean line that contacts the extreme upper and lower points on the profile within the roughness sampling length.

P4: Maximum Peak—to—Meanline Depth Roughness.

P5: Maximum Valley—to—Meanline Height Roughness.

P6: Skewness. A measure of symmetry of the profile about a mean line , offer a convenient way to illustrate load carrying capacity, porosity, and characteristics of nonconventional machining process.

P7: Kurtosis. A measure of the ADF sharpness. In addition, it quantitatively describes the randomness of a profile's shape relative to that of perfectly random surface which has a kurtosis of 3.

P8: Mean Wavelength.

P9: Average of total sum of each two nearby points' slope.

P10, P11, P12: A reference line is drawn parallel to the mean line and at a preselected or predetermined distance from it to intersect the profile in one or more subtended lengths. The sequences are 25%, 50%, and 75% of the distance between mean line.

P13: The peak value number which over least-square line.

P14: Density of peaks on profile.

P15: Mean curvature of peak points.

P16: Wave band width.

P17: Second power spectral density(PSD) moment.

P18: Fourth PSD moment.

P19: Density of summit on profile.

P20: Number of peaks on the whole measuring length.

4.1 Pre-Processing of Data

4.1.1 Measurements

There are twenty-one parameters in each measurement. Twenty-five measurements are made on each specimen. The parameters can be viewed as the components of a 21-dimension vector.

As shown in table 4.1, specifications of readings are follows:

4.1.2 Filtering of Extreme Measurements and Treatment of Raw Data

Processed measurements yields the values of 21 parameters. For further manipulation, the 21 parameters can be conveniently considered as components of a vector representing a surface measurement.

We can, therefore, evaluate the closeness of two measurements by dot multiplication of the vectors representing them. After dividing the dot product by the lengths of the two vectors, the cosine of the angle between the two vectors is calculated. For parallel vectors, i.e. simillar measurements, the vectors would be parallel and the cosine of the angle between them is equal to unity.

The measurements viewed as vectors are further processed as follows:

- Calculate the average vector from the 25 vectors which correspond to to the number of measurements per sample..
- Obtain the dot product of each of the vectors by the average vector.
- Divide each dot product by the length of the two vectors.
- Filter the extreme vectors, the average of the readings which are within confidence interval is the mean vector.

According to statistics, when sampling number is small, and data histogram shows as bell shape, this population can be classified as t—distribution. In addtion, because average value sometimes is different from the mean value for whole distributation, the confidence $(1-\alpha\%)$ interval is decided by the following equation.

$$
\overline{X} - t_{\alpha/2} \frac{s}{\sqrt{n}} < \mu < \overline{X} + t_{\alpha/2} \frac{s}{\sqrt{n}} \tag{4.1}
$$

Where

 \overline{X} : average of samples.

s : standard deviation of samples.

n: sample number, $n < 30$.

n-1: degree of freedom.

 α : confidence.

Figure 4.2, for example, show the angles between each measurement (from 1 to 25) and the confidence interval. The vectors within confidence interval insures that the mean reading will be located in the interval with 99% confidence. We also call these mean vectors "representative vectors" or "representative readings". Later, these vectors will be treated as our neural network training sets and expert system database.

4.2 Data Preparation

4.2.1 Testing Set

As described in the last section, readings among confidence interval are used to calculate the mean vector with 99% confidence. In our identification procedure, the confidence readings will be treated as testing set. The number of confidence readings in the testing set shown in table 4.2.

4.2.2 Constructing Learning Set of Network and Database of Expert system

For network:

The set of the vectors of specimens is treated as training set $#1$. Three more training sets are used with some modifications to set $#1$, as explained in chapter 5.

All data will be scaled before being input to network, because backpropargation network use sigmoid transfer function. For more detailed description, please see chapter 5.

For expert system:

÷,

For the expert system, the data is composed of parameters' ranges, standard deviation of each parameter, and the average of vectors. Please see chapter 6.

 $\hat{\mathbf{r}}$

Table 4.1: Specifications of measurement.

Machining	NO. of Samples	NO. OF CONFIDENCE READING
Grinding	6	65
Turning	5	65
Milling	5	53
EDM	5	44
Waterjet	9	79
TOTAL	30	306

Table 4.2: Number of confidence readings for specimens

Chapter 5

APPLICATION OF BACKPROPAGATION NETWORK

In this chapter, we use $BRAINMARKER^R[14]$ software to simulate the backpropergation network. The following details are necessary to build up the identification network.

5.1 Input Layer

The input layer of the network receives input from the outside world and send output off to the hidden layer. Two items are required to input data into the network, one is an input layer of neurons, and the other is an encoding scheme. The function of the encoding algorithm is to take input data and convert it into a form which is suitable for presenting to the network. The input encoding turns raw data into a sequence of numeric values that the network can understand. The input encodings for a neural network can be either local or distributed in nature[14]. We choose the former. Each input neuron represents one component of the representative vector. This is equivalent to one measurement parameter. As shown in Table 5.2, the input feature is for grinding sample $#1$.

5.2 Output Layer

The output layer of the network sends the output to the outside world. By looking at the output of this layer we can tell what solution the network has arrived at, for any particular input. Like inputs, the outputs of a network can be either local or distributed. We still use the local, i.e. each neuron corresponds to one machining methods. An example of such an local output set is our network, in which the output of the network consists of 14 numbers. An input vector is represented by an output vector. Turning sample $#1$ input test vector might result in the following output:

0.02 grinding $+$ 0.00 milling $+$ 0.88 turning $+$ 0.1 EDM $+$ 0.00 Waterjet $+$ 0.99#1. The numbers indicate the degree of "confidence " that the input vector represents the corresponding machining. In this case, the network is indicating than the machining method is turning with a confidence of 0.88.

5.3 Hidden Layer

A hidden layer is composed of neurons which may be connected to neurons in other layers, or perhaps to each other, but which do not interact directly with the environment. The hidden layer recodes the input into a form that captures the correlations in the training set. This allows the network to generalize in a natural way from previously learned facts to new, or noisy inputs. The correlational information is stored in the connection matrix between the input and hidden layers. The output of the hidden layer is decoded into a meaningful form by the matrix between the hidden and output layers.

How many hidden neurons should we implement to the network? Histograms of weight matrix are useful for checking on the mental health of network[14]. If the histograms resemble bell shaped curves, the network own a lot of spare capacity.

If the histograms are split into disconnected segments, the reason is using a very small number of neurons and the histogram are hard to interpret. There is no hard and fast rule. We have to watch the network and find out how the histograms looks for our specific configuration. There is one general rule: as more and more connections piles up at -8, the network is getting more and more hopeless. As shown in figure 4.1, we list some histograms of different hidden neuron numbers. From top to bottem, they represent distribution of weight matrix which neuron number are 40, 60 and 90. The left ones present histogram of connection matrix between input layer and hidden layer. The right ones are about histogram between hidden layer and output layer. Table 5.1. lists the performance and convergence iterations of network using training set $#1$ with different hidden neuron number. It is obvious that network using less hidden neuron number may has lower rate of accuracy, and the hidden neuron number does not affect the iteration number of convergence.

hidden neuron no.	performance of network	steps
40	78 %	12720
50	85 %	11250
60	84 %	11730
70	88 %	11920
80	90 %	13960
90	90 %	11640
100	91 %	12270

Table 5.1: Performance and iterations of network using different hidden neuron number

5.4 Data type, Precision, Renormalization

We choose 21 numerical parameters to represent machined surface characteristics. The parameters' values are input in double precision. The software BRAIN-MAKER accepts numerical values in the range of $10 \times e^+308$ and $-10 \times e^-308$. The data is normalized so that all parameters values are scaled to fall in the range of 0

Figure 5.1: Histogram of weight matrix with different neuron number. From top to bottem, the number are 40, 60 and 90.

to 1. The scaling is done as follows:

Input to neuron =
$$
(C - Min.)/(Max.-Min.)
$$

Where

 $C = input$ value from real world

Min. and Max. $=$ From data collection, the minimum and maximum values of C

5.5 Learning parameters

• Learning rate:

We already mentioned that the optimum goal is to move the weight vector (W) in a direction so that the value of function (F) will be smaller at the new value Obviously, the best approach would be to move in the W direction in which F is decreasing most rapidly. Thus, a reasonable learning law is

$$
\vec{W_{new}} = \vec{W_{old}} - \alpha \times \nabla(F(W)) \tag{5.1}
$$

Where

 \vec{W} : weight vector

 $F(W)$: mean square error function

 α : learning rate

V: gradient

In generally, if $\alpha > 0$, we often choose 1 as learning rate. With high learning rates, W_{new} sometimes "zooms" past the solution. When using learning rate smaller than 1, it slows down training, and does not help convergence.

• Training tolerance

If any individual neuron in the network is incorrect, the network is considered to have gotten this fact wrong, and training will continue. For example, 0.7 is the correct answer for a certain fact, a training tolerance of 0.1

means that any output from 0.6 to 0.8 is consider correct. We use default value of 0.1.

5.6 Ordering of the Input Vectors

Neural network learns most effectively when the training facts are in random order^[14]. Convergence could be slowed if input vectors representing different machining methods are not randomly presented to the network during training.

5.7 Training and Testing Sets

To check the network performance, there are two methods. One way is to be sure that training sets are large enough and to show that further increasing the size of training set does not affect performance. Another way is to test the performance of the network on both the training and testing sets and show that the network predicts correctly the machining methods most of the time.

The first thing is to check the amount of available data to see if it is enough or not. Too little data causes failure of neurocomputing to generalize. We use four training sets. Two set consist only 30 representative vectors. The other two include both representative vectors and extra 93 readings which are within the confidence interval.

We know that neural network can generalize from the training set examples to the entire problem environment. Since a real-world input vector is assumed to lie close to the training set examples, the output of the welltrained network should be reasonably related to the outputs it would give for the similar examples in the training set. If the input is far away from any

training examples, then the output of the network cannot be expected to be meaningful.

Too many training facts would cause the training time to be unnecessary long. The other disatvantage is the phenomenon of "overtraining" when simillar or same training vectors appear repeatedly which leads to poor performance of the network when presented with new input.

5.8 Description of Identification Network

This network will evaluate various machined surfaces which are generated by various machining methods. The networks give an output expressed in output neuron activations which can be interpreted in terms of machining methods and the number of sample. Examples: "measurement belongs to turning sample $#1$ ", or "measurement belongs to waterjet sample $#9$ ", and so on. High single neuron activation will mean high confidence, and vise versa.

Constructing training set. The first thing we need to do is to list the training set. Here is a single input vector, followed by the corresponding desired output.

facts

I PO P1 P2 ... P20

D 1 0 0 0 0 1 0 0 0 0 0 0 0

This fact is about the description of grinding sample #2. PO to P20 input and D represents desired output. Note that we do not label I or D in the real training facts files. In order for the neural network software to recognize the input data, we must enter each fact serially as shown in Table 5.2.

Part of a training set is shown in Figure 5.2.

G:grinding T:turning M:milling E:EDM W:waterjet
I: Input set
D: Desired output

 \bullet

 $\mathcal{L}_{\mathcal{A}}$

Figure 5.2: Part of training set #1 of identification network.

 $\ddot{}$

Constructing testing set The network's performance is measured using two different data set: *training set* and *testing set.* The training set is a seperate body of data from the testing set. While the training set is used directly for training of the network, the testing set is used only for the evaluation of its performance.

We mentioned earlier the phenomenon of "overtraining". In fact, as the training set becomes larger, the error estimated by using the test set typically decreases for a while, then begins to increase again. If we were somehow able to increase the size of the training set to the point where further increase in size would not change the error values, then this could be taken as evidence that the error reaches a minimum.

Part of the test set $#1$ are listed in Figure 5.3.

Specifications for the network

This specification of the neural network consists in specifying the number of input, hidden and output neurons.

The networks have input layers of 21 or 22 inputs. one hidden layers of 100 or 120 processing elements, and 14 processing elements in the output layer.

Detail description are given in Figure 5.4 and Table 5.2.

1.44 1.81 11.3 12.1 -6.9 5.2 -0.37 3.12 45.5 0.25 01 0.69 0.97 141 0.08 0.14 21 2 0.04 0.01 0.009 325 1.4 1.8 10.3 10.5 -4.1 6.4 0.23 2.69 47.1 0.24 002 0.27 0.77 142 0.08 0.14 21 2 0.04 0.01 0.009 327 1.5 1.9 12.0 12.5 -5.8 66 001 307 47.1 0.25 003 029 091 146 0.09 0.15 21.5 0.05 0.01 0.009 339 1.51 1.92 12.0 12.5 -5.8 66 001 307 47.1 0.25 003 039 091 146 009 0.15 21.5 0.05 0.01 0.009 339 1.48 1.87 12.4 13.0 -5.6 7.3 002 3.23 49.0 0.23 0.01 033 0.89 148 0.09 0.13 209 0.05 0.01 0.009 339 1.41 1.83 14.6 15.3 -8.8 6.5 -0.07 4.1 46.3 0.24 0.06 0.76 0.99 142 0.08 0.15 21.9 0.04 0.01 0.009 333 1.3 1.7 14.0 14.4 -6.7 7.6 0.18 4.13 49.8 0.21 001 039 097 149 0.09 0.13 21 1 0.04 0.015 0.010 343

................

Figure 5.3. Part of the test set

SPECIFICATION:

TYPE NETWORK: BACKPROPAGATION NETWORK SIZE: I, H, O **INPUT LAYER:** NUMBER OF NEURONS: I TRANSFER FUNCTION: LINEAR LINEARING RULE: NONE **HIDDEN LAYER:** NUMBER OF NEURONS: H TRANSFER FUNCTION: SIGMOD LEARNING RULE: CUMULATIVE - DELTA NUMBER: J **OUTPUT LAYER:** NUMBER OF NEURONS: O TRANSFER FUNCTION: LINEAR LEARNING RULE: NONE **LEARNING RATE: 1** TRAINING TOLERANCE: 0.1 NUMBER OF FACTS IN TRAINING SETS: M

Figure 5.4: Specification of network.

Desire Output	Example (grinding sample $#2$)
grinding	1
turning	Ω
milling	
EDM	
waterjet	O
sample $#1$	0
sample $#2$	
sample $#3$	O
sample $#4$	
sample $#5$	
sample $#6$	
sample $#7$	O
sample $#8$	
sample $#9$	

Table 5.2: Series of desired output for grinding sample #2.

				Training set $\# I H O $ NO. of Training Facts
	21		14	-30
	22	100	14 ₁	-30
2		$120 +$	14	-123
	22			123

Table 5.3: Specification table of network (I: no. of input neuron, H: no. of hidden neurons, 0: no. of output neurons.)

Chapter 6

DESCRIPTION OF THE EXPERT SYSTEM FOR IDENTIFICATION

6.1 Data Analysis

TThe first step to build an expert system is to analyes the existing knowledge. Figure 6.1 shows the ranges for the Roughness Average for for the various machining methods. Appendix B, shows the ranges for each of the other 20 parameters for the various machining methods. From those figures, it is clear that most parameters' ranges overlap. In these figures, the X—axis indicates the magnitude of the parameters. The Y—axis shows the machining methods.

6.2 Parallel and Sequential Procedure

The problem of overlapping parameters requires an expert system which can evaluate all the possibilities. When a specimen is presented, how much information an expert system should receive in order to make a decision will be an important factor to make us decide which of the above of procedures will be adopted.

Roughness Average

 \hat{b}

Figure 6.1: Ranges of the parameter 0

There are two major decision procedures: parallel and sequential, as shown in Figure6.12.

Below are their definitions:

- Parallel decision procedure The system will get in all the information on a subject and then make a decision.
- Sequential decision procedure The system is always guided by the last piece of information received.

In general, a parallel procedure will make decisions which are as good and usually better than a sequential procedure[15]. No doubt that identification expert system needs all information to make the best judgement. Our system will use parallel decision procedure.

6.3 Adding Rules

The expert system will process the specimen data and will output the likelihood that the specimen is generated from each possible machining method. The highest likelihood will be used to identify the machining method. Let RESULT [j] represents the value of the combined confidence factor for machining method j. (j=1 represents grinding, j=2 represents turning, etc.) The selected machining method m will maximize the value of RESULT [j], so that

$RESULT[m] > RESULT[i],$

for all $i \neq m$.

The system consists of two rules:

Sequental procedures

The system will make decision if the data was enough, else it request more input.

Figure 6.2: Parallel and sequential procedures

```
Rule 1 
IF 
LOCALMIN [j][i] < P[i] < LOCALMAX [j][i] 
THEN
CF1 [j][i] = 1ELSE 
CF1 [j][i] = -1
```
Rule 2

IF

 $MEAN[j][i] - 3 \times STDEV[j][i] < P[i] < MEAN[j][i] + 3 \times STDEV[j][i]$ **THEN** $CF2$ [j][i] = relative probability of P [i] with MEAN [j][i] ELSE $CF2 [j][i] = -1$

Combine Rule 1 and Rule 2:

FOR $j=1$ to 5 FOR $i=1$ to 21 RESULT $[j]$ = RESULT $[j]$ + CF1 $[j][i]$ + CF2 $[j][i]$; END END

The First Rule

To calculate RESULT [j], for a given input vector, we consider the minimum and maximum of each parameter for different machining methods. We use LOCALMIN [j][i] and LOCALMAX [j][i] (i=1 to 21, for parameters. j=1 to 5, for five kinds of machining methods) to represent them. By observing

all parameters' ranges, the first rule will contribute to RESULT [j].

First rule is based on the theory of *uniform distributation* (A random variable having distribution in this family arises by choosing a number from some bounded interval in such a way that all numbers are equally likely to be chosen)[16]. Rule 1 tests if a given parameters of the input measurement falls within the range of the corresponding parameter of each of the machining methods. Thus, if an input parameters falls within the range for a machining method m, the confidence factor RESULT $[m]$ is increased by unity. If not, the confidence factor is reduced by the same amount.

The Second Rule

This rule is a typical abduction heuristic, also based on probability theorem. In this rule, we made two assumptions:

- 1. There exists a seperate and unique mean value MEAN [j][i] of each parameter for every machinning methods. For example: grinding $(j = 1)$, the P [i] $(i=1 \text{ to } 21)$ would be near MEAN[1][i] and far away from MEAN [2][i], MEAN [3][i], etc.
- 2. Random variable *P[i],* belong to family of *normal distribution.*

Rule 2 calculates the distance between an input parameter from the mean value of the corresponding parameters which are related to various methods of machining. The distance is converted into an estimate of the probability that the input parameter belongs to one of the various methods of machining.

A normal distributation has the famous "bell shape" [17]. To calculate the confidence factor of $P[i]$ in normal distributation, there are two methods. The first way is numerical analysis, integrates the area under "bell shape"

curve excluded the area between $P[i]$ and MEAN [j][i]. The other is to use relative probability. For probability of MEAN [j][i] equal one, the relative probability of P[i] is

$$
e^{\frac{-X^2}{Y}}
$$

Where

 $X = P[i] - MEAN[j][i]$ $Y = 2 \times STDEV[j][i]^2$

The system picks the highest value of RESULT [j] and deduces that j is the machining method that generated the specimen.

After testing, we found that the identification rate is up to 94%.

Chapter 7

RESULTS AND DISCUSSION ON USING BACKPROPAGATION NETWORK

7.1 Identification Results of Network

Description of the Training Sets

We constructed four learning sets :

- 1. Set $#1:30$ representative vectors which are average of all vectors within the confidence interval.
- 2. Set $#2: 30$ representative vectors, with vector length added as the twenty—second parameter for each fact.
- 3. Set #3 : 30 representative vectors and 93 vectors which are randomly selected from 306 confidence readings.
- 4. Set $#4$: Same as set $#3$ but adding length of vector to each training fact.

Testing Sets

- 1. Set $#1$ and set $#2$: Include all 306 vectors to test the network performance which were trained by training set $#1$ and $#2$.
- 2. Set $#3$ and set $#4$: These are to test the networks which are trained by the training set $#3$ and $#4$.

part of the results of testing the network performance using the test set are shown in Table 7.1 to Table 7.5. The rest of the tables are shown is Appendix C. In the tables, the first column of the table lists the reading number. The second column lists the degree between the testing reading and representative reading. From the third to sixth columns, the results from the various network which use different training sets are shown.

In these tables, T represents the reading which is included in the training set. The results are prsented in a simple form. For example: : $99e + 94#1$. It means : the network has 99% confidence that the specimen is from grinding and 94% confidence that the specimen is simillar to sample#1.

- 1. Table 7.1 presents the results of network identification of grinding sample #1. We
- 2. Table 7.2 presents the results about turning sample #2.

When turning specimens were being measured, because of their round shape, hand—adjustment of span was used instead of using positioning table. The precision become lower and the confidence interval become wider. Fewer readings will be filtered. Twenty—two readings are included into the testing set for turning sample $#2$. We also can see some missclassification occured.

3. Table 7.3 presents the results about milling sample #1. See the testing result about reading $#1$. Using training set $#3$ and $#4$ can help decrease miss—classification.

4. Table 7.4 and 7.5 present the identification results about EDM sample $#2$ and waterjet sample $#1$.

7.2 Discussion

To judge what kind of training set is the fittest one, we have to evaluate network performance. As shown in Figure 31 to Figure 33, M represents the number of error miss—classification of machining methods, S represents the number of miss—identification about sample number(for example: sample $\#1$, sample $\#5$, etc.).

Table 34 shows the conclusion. We can see that using training set $#1$ and $#2$, the accuracy are about 91%. For network useing set $#3$, the rate is up to 99%.

Note that network in identifying different operations works well. It always higher than 94%. If it uses training set 3, it is up to 99%.

The results shows that using backpropagation network can successfully identify machined surfaces which are produced by different methods: grinding, turning, milling, EDM or waterjet. However, for telling the difference of the specific machining condition of each specimen, only 85.9% accurcy could be reached when network uses learning sets #3.

Ŷ,

 \mathbb{R}^2

Table 7.1: Identification results about grinding sample $#1$.

 $\frac{1}{\lambda}$

- 30 - 7

Table 7,2: Identification results about turning sample #2.

 ~ 2 \bar{z}

 $\epsilon = 1.35$

 $54\,$

 $\frac{1}{2}$

 $\frac{1}{2}$, $\frac{1}{2}$

 \mathcal{A}^{\perp}

Table 7.3: Identification results about milling sample #1.

i.

ju.

 $\bar{\mathbf{r}}$

 $\dot{\gamma}$

 $\frac{1}{2}$

- 30

Table 7.4: Identification results about EDM sample #1.

 $\overline{}$

۳

 $\frac{1}{\beta}$

 $\tilde{\mathcal{S}}_t$

بتعر

 $\sum_{i=1}^N$

 $\chi^{(0)}$

Table 7.5: Identification results about waterjet sample #1.

Turning Sample #			
Total No. of		14	
Readings			
Miss-classification			
(Using training			
set $#1)$			
Miss-classification			
(Using training			
set $#2)$			
Miss-classification			
(Using training			
set $#3)$			
Miss-classification			
(Using training			
set $#4$			

Table 7.6: Miss-classification of grinding and turning specimens.

 $\frac{1}{2}$

 \mathcal{M}

EDM Sample #							
Total No. of Readings			8	13			
Miss-classification (Using training set $#1)$							
Miss-classification (Using training set #2)							
Miss-classification (Using training set #3)							
Miss-classification (Using training set #4							

Table 7.7: Miss-classification of milling and EDM specimens.

 $\ddot{}$

 $\frac{1}{2}$

 $\mathcal{L}_{\mathcal{A}}$

 $59\,$

Ġ,

 $\ddot{}$

 $\ddot{}$

Table 7.8: Miss-classification of waterjet specimens.

 $\hat{\mathbf{z}}$

 $\sum_{i=1}^{n}$

 $\hat{\mathcal{A}}$ ϵ

بالمنزر

Table 7.9: Conclusion table of accuracy

 $\bar{\nu}$.

 $\frac{1}{3}$

Chapter 8

CHOOSING PROPER PARAMETERS IN USING IDENTIFICATION EXPERT SYSTEM

As metioned in chapter 6, the parallel decision expert system is built to identify different machined surfaces. The accuracy is approximately 93.5% (20 error estimation in 306 testing facts). But it still can be inscreased by selecting useful parameters and filtering out variables that have weak classificatory power.

First, we have to inspect the population of each parameter. In chapter 6, we know that the expert system needs to consider the local minimum and local maximum, as well as mean value and standard deviation for different machining methods. It is reasonable to assume that parameters whose distributations are steep, i.e. with small standard deviations, would be more useful as classificatory variables, than parameters with large standard deviation. In statistics, the property of the steepness is known as *kurtosis.* Figure 8.1 lists all MEAN $[j][i]$ and Figure 8.2 lists the kurtosis. The parameters are 6, 7, 8, 15, 17, 18, 20. Using these parameters only to calculate the total confidence level, the rate of success is raised from 93% to 98%.

 \mathcal{A}

Figure 8.1: Mean[i][j]

 $\hat{\boldsymbol{z}}$

.
St

 $\omega_{\rm f}$

Ņ

 $\label{eq:1} \begin{array}{l} \mathbf{v} = -\frac{1}{2} \mathbf{v} \frac{\partial \mathbf{v}}{\partial \mathbf{v}} \\ \mathbf{v} \frac{\partial \mathbf{v}}{\partial \mathbf{v}} \frac{\partial \mathbf{v}}{\partial \mathbf{v}} \end{array}$

Figure 8.2: Kurtosis[i][j]

À,

Chapter 9

CONCLUSIONS AND FUTURE WORK

9.1 CONCLUSION

- 1. In identification network, adding vector length as the 22th can increase rate of convergence during training by 20%. Although the network converged more fast, we did not gain any improvement in accuracy when using training set #4. Vector of lenghth is not an important parameters for network in identifying different machined surfaces.
- 2. Even where all parameters overlap, the network as well as expert systems are able to learn and recoginize new input and be able to classify correctly.
- 3. An advantage of Neural Network versus Expert Systems, is that while in expert system, we have to assume a certain populatuion distriburion (normal is our case), the neural network outp erformed the expert system without any statistical assumption.
- 4. In expert system, choosing only parameters 6, 7, 8,15, 17, 18 and 20 to construct the testing set and knowledge base, the accuracy of performance increased by 4% comparing with using total 21 parameters.

9.2 FUTURE WORK

The backpropergation networks are useful for identifying different machined surfaces future work can continue towards improving the robustness of the system to noise, and to improve it performance.

- Introducing noise into the data and try to see how the number of hidden neurons increases to keep the same accuracy.
- Discover method or methods using Neural Network technology to single out the input variables that are most relevant to the classificatory performance of the network.
- Exploring the impact of choosing more than one layer of hidden neurons on the performance of network for classificatory tasks.

Appendix A

SPECIFICATIONS AND MEASURING RESULTS OF SPECIMENS

All of the representative vectors and machining for all samples are listed on following tables:

 $\bigg\{$

Table A.1: Representative vectors of grinding specimens.

 $\overline{}$

¥

y.

 $\overline{\mathcal{N}}$

 $\frac{1}{2}$

Table A,2: Machining conidition **of grinding specimens.**

Table A.3: Representative vectors of turning specimens.

Ų,

 ϵ

 \bullet

 \mathfrak{a}

 $\tilde{\mathbf{r}}$

 $\ddot{\cdot}$

ķ.

 $\sigma_{\mathcal{D}}^{(1,0)}$

 $\ddot{\cdot}$

 $\label{eq:2} \frac{1}{2} \sum_{i=1}^n \frac{1}{2} \sum_{j=1}^n \frac{1}{$

 $\sum_{i=1}^n \frac{1}{2}$

ju.

Table A.4: Machining condition of **turning specimens.**

Table A.5: Representative vectors of milling specimens.

 $\tilde{2}$ $\tilde{\mathcal{A}}_{\mathbf{a}}$.

 $\ddot{}$ $\overline{}$

 \bar{L}_A

ż

×

Â,

 $\frac{1}{2}$

÷

 \bullet .

٠

À

÷

 $\frac{1}{4}$

 $\ddot{}$

٠

 $\ddot{}$

 $\alpha_{\rm s}$ ÷

 $\frac{1}{\sqrt{2}}$

 $\ddot{\cdot}$

 Ω

 $\overline{\mathcal{L}}$

٠

 $\mathcal{L}^{\text{max}}_{\text{max}}$

Ļ,

Table A.6: Machining condition of **turning specimens.**

Table A.7: Representative vectors of EDM specimens.

 $\frac{1}{\sqrt{2}}$

×

j.

 \bullet

 \mathcal{I}_4

 \mathcal{V}_A

 $\bar{1}$

 \mathcal{L}_{max}

Table A.8: Machining condition **of EDM specimens.**

Table A.9: Representative vectors of waterjet specimens.

÷

×

 λ

À

٠

 $\overline{9}$

 \backslash

 $\chi\lesssim 1$

Table A.10: Machining condition of waterjet specimens.

ţ.

À

 $77\,$

Appendix B RANGES OF PARAMETERS

Figure 1 to figure 10 show the ranges of parameters for the various machining methods. In figures, the Y-axis shows the machining methods : 1. grinding, 2. turning, 3. milling, 4. EDM, 5. waterjet.

Figure B.I: Ranges of the parameter 1 and 2

 $\bar{\zeta}$

Maximum Peak-to-Meanline Depth Roughness

Figure B.2: Ranges of the **parameter 3 and 4**

 $\ddot{}$.

Maximum Valley-to-Meanline Height Roughness

 $\ddot{\S}$

Figure B.3: Ranges of the parameter 5 and 6

 \bar{z}

 $\frac{1}{\Lambda}$

Ÿ.

Figure B.4: Ranges of the parameter 7 and 8

Average of Total Sum of Each Two Nearby Points' Slope

Predetermined Distance-25%

Figure B.5: Ranges of the parameter 9 and 10

Predetermined Distance-75%

Figure B.6: Ranges of the parameter 11 and 12

 $\frac{1}{3}$

The Peak Value Number Which Over Least-Square Line

 \mathbf{I}_p

Figure B.7: Ranges of the parameter 13 and 14

Ň

Mean Curvature of Peak Points

Ä

Figure B.8: Ranges of the parameter 15 and 16

 $\bar{\ell}_i$

Å

Second Power Spectral Density (PSD) Moment

Fourth PSD Moment

Figure B.9: Ranges of the parameter 17 and 18

 $\frac{1}{3}$

Number of Peaks on the Whole Measuring Length

Figure B.10: Ranges of the parameter 19 and 20

 $\bar{\bar{\lambda}}$

Appendix C

RESULTS ON USING BACKPROPAGATION NETWORK

Identification results about testing sets are list following. Note that all results are displayed in simplified way. It have been described in chapter 7.

Table C.1: Identification results about grinding sample #2.

Table C.2: Identification results about grinding sample #3.

 \mathcal{L}

 $\frac{1}{3}$

Table C.3: Identification results about grinding sample #4.

 \sim

 $\frac{1}{2}$

¥

Table C.4: Identification results about grinding sample #5.

Table C.5: Identification results about grinding sample #6.

 χ

Table C.6: Identification results of turning sample #1.

Table C.7: Identification results of turning sample #3.

 $\langle \cdot \rangle$

Table C.8: Identification results of turning sample #4.

Table C.9: Identification results of turning sample $#5$.

 $\hat{\boldsymbol{\alpha}}$

Table C.10: Identification results of milling sample #2.

 $\bar{\beta}$

Table C.11: Identification results of milling sample #3.

Table C.12: Identification results of milling sample #4.

Table C.13: Identification results of milling sample #5.

 $\langle \cdot, \cdot \rangle$

 $\bar{\mathbf{r}}$

Table C.14: Identification results of EDM sample #3.

 $\bar{\mathbf{z}}$

 \bar{r}

Table C.15: Identification results of EDM sample $#4$.

Table C.16: Identification results of EDM sample #5.

 $\hat{\lambda}$

Table C.17: Identification results of EDM sample $\#6$.

 \bar{z}

 \sim ω

 ~ 5 .

Table C.18: Identification results of waterjet sample #2.

Table C.19: Identification results of waterjet sample #3.

 $\hat{\mathcal{A}}$

Table C.20: Identification results of waterjet sample #4.

 \sim \sim

 $\mathcal{L}^{\text{max}}_{\text{max}}$

Table C.21: Identification results of waterjet sample #5.

Table C.22: Identification results of waterjet sample $#6$.

 Λ

 \rightarrow

Table C.23: Identification results of waterjet sample $#7$.

 \bar{z}

 $\frac{1}{2}$

Table C.24: Identification results of waterjet sample $#8$.

 $\frac{1}{\sqrt{2}}$

 $\frac{3}{4}$

 $\ddot{}$

Table C.25: Identification results of waterjet sample $#9$.

 $\langle \sigma_{\rm{eff}} \rangle$

Bibliography

[1] Lawrence E. Doyle (1985), *Manufacturing Processes and Materials for Engineers,* 3rd edition, pp. 444.

 $\mathbf{1}$

- [2] Paul Harmon (1987), *Expert system strategies Newsletter,* vol.3, no. 12, pp.6-11.
- [3] California Scientific Software (Jan., 1989), *BrainMaker—Neural Network Simulation Software,* 3rd edition, pp. 50-51.
- [4] Maureen Caudill (1990), Charles Butler, *Naturally Intelligent Systems,* pp. 184.
- [5] Philip D. Wasserman (1980), *Neural Computing Theory and Practice,* pp. 18.
- [6] Anne Moose, Terry Schussler, Dan Shafer (1989), *VP-Expert, Rule-Based Expert System Development Tool,* pp. 1-2.
- [7] Hart, Anna (1988), *Knowledge Acquisition for expert systems,* pp. 13.
- [8] Judea Pearl (1984), *Heuristics—Intelligence Search Strategies for Computer Problem Solving,* 2nd edition, pp. 3, pp.6., pp. 124.
- [9] Elaine Rich (1983), *Artificial Intelligence,* pp.36.
- [10] George F. Luger, William A. Stubblefield (1989), *Artificial Intelligence and the Design of Expert Systems,* pp. 309.
- [11] Siegel, Paul (1986),Expert *System: A non—programmer's guide to development and applications,* pp. 38.
- *[12] ANSI/ASME standard committee B46.1-1985, Classification and Designation of Surface Quantities.*
- [13] G. Y. Zhou, M. C. Leu, S. X. Dong *(1989)Measurement and Assessment of Topography of Machined Surfaces,* pp4.,pp. 13.
- [14] "CSS" (Jan., 1990), *BrainMaker—Neural Network Simulation Software,* 4th edition, ch. 7, pp. 17-18, ch.5, pp11-22.
- [15] Chris Naylor (1987), *Build Your Own Expert System,* 2nd Edition, pp. 51-54.
- [16] Keinosuke Fukunaga (1972), *Introduction to Statistical Pattern Recognition.*
- [17] Norman S. Matloff (1988), *Probability Modeling and Computer Simulation,* pp. 86.
- [18] Philip D. Wasserman (1989), *Neural Computing Theory and Practice,* pp. 10.
- [19] Casimir C. "Casey" Klimasauskas, NeuralWare, Incorported (1988), *NeuralWorks—An Introduction to Neural Computing,* revision 2.00, pp.7.