

4

Tool Condition Monitoring in Manufacturing Systems Using Neural Networks

- 4.1 [Introduction](#)
- 4.2 [Machining Tool Conditions](#)
Tool Wear Mechanism • Forms of Tool Wear • Flank Wear • Crater Wear • Groove Wear
- 4.3 [Sensors and Signal Processing](#)
Dynamic Force • Acoustic Emission (AE) • Wavelet Packet Analysis of AE and Force Signals • Vibration (Acceleration) • Coherence Function of Cross Vibration Signals
- 4.4 [Feature Extraction](#)
- 4.5 [Neural Network Architectures](#)
Multi-layer Perceptron (MLP) • Kohonen Networks • ART2 Networks
- 4.6 [Tool Condition Identification Using Neural Networks](#)
MLP for Force Sensor with Simple Pre-Processing[32] • Feature requirements • An Integrated Fault Diagnosis Scheme • Experiment • Discussion of Results • Single-ART2 Neural Network with Acoustic Emission Sensing • Transient Tool Condition Identification • Tool Wear Monitoring • Single-ART2 Neural Network with Acoustic Emission and Force Sensing • Multi-ART2 Neural Network with Force and Vibration Sensing • Feature Information
- 4.7 [Conclusions](#)
- [References](#)

G.S. Hong

National University of Singapore

M. Rahman

National University of Singapore

Y.S. Wong

National University of Singapore

4.1 Introduction

Tool condition monitoring is crucial to the efficient operation of any machining processes where the cutting tool is in constant or intermittent contact with the workpiece material and is subject to continuous wear. It presently acquires greater importance than ever as manufacturing systems are increasingly

required to provide greater automation and flexibility while maintaining a high level of productivity [10]. The more recent computer numerical controllers can be programmed to monitor the time spent by a cutting tool in machining and automatically changes the tool when the total machining time spent by the tool reaches its tool life. This tool life is determined experimentally by conducting a controlled set of machining tests. The useful life of the tools tends to be conservatively taken and may be wasted, resulting in frequent tool changes and longer machine downtime, thereby decreasing the system productivity. On the other hand, there may be tools that fail prematurely compared with the average tool. Tools can also fail earlier when used in conditions not similar to those employed in the experimental determination of the tool life. A premature tool failure can result in damage to the workpiece and disrupt the automated machining operation. Hence, a suitably developed tool wear monitoring technique is needed to utilize the tool more efficiently while preventing premature tool failure. Several tool wear monitoring techniques have been developed and reported [7, 26, 6, 39, 8, 35, 43, 42, 44]. A few commercial monitoring systems have also been developed for use with CNC machines [31]. The monitoring approaches are typically based on acoustic emission, motor current, vibration, and force sensing, or their combination, and are primarily developed for application in roughing operations. Further research aims to improve the reliability of the monitoring system to minimize false alarms and allow the system to be used over a wider range of machining conditions. More recent attempts adopt sensor fusion techniques [33, 42] so as to rely on more than one type of sensor inputs for more robust deduction of the state of the cutting tool. A general scheme of such an approach is shown in Fig. 4.1. It consists of two stages: a sensing and preprocessing stage and a sensor fusion stage. In the sensing and preprocessing stage, the signal from a sensor is conditioned and processed to derive information pertinent to the machining process. This preprocessing is commonly referred to as *feature extraction*. Important features obtained from a sensor signal or a set of sensor signals are then integrated or “fused” together using an appropriate sensor fusion technique to determine the tool condition. The objective is to associate the different extracted feature pattern with a corresponding wear condition. One of the techniques used is to utilize the associative capability of neural network to arrive at a more reliable conclusion on the tool condition.

This chapter discusses some neural network applications to tool condition monitoring for turning process. Different types of tool conditions are first introduced in Section 4.2. Section 4.3 then presents various common sensor used for tool condition monitoring and sensing parameters that are sensitive to tool wear conditions. However, the information embedded in the sensors is convoluted such that the association of the tool condition with the measured sensor signal is not apparent. Some preprocessing techniques to extract the information from the relevant sensor signals are described. Thus, Section 4.4 discusses the importance of feature extraction process to enhance the sensitivity and robustness of tool condition predictions. Due to the variance in the cutting condition and the chaotic nature of the cutting process, a single feature is normally not sufficient for reliable tool condition prediction. However, an increased number of feature components inherently imply more complex heuristic rules to be employed to associate each feature pattern with a corresponding tool condition. A popular approach to such a problem is to utilize the learning capability of neural network [16, 8, 32, 33] to combine these feature components to produce some indices. Such reduced indices are usually more manageable such that simple heuristic law can be applied to associate these indices with a corresponding tool conditions. Section 4.5 introduces some common neural network architectures used in tool condition monitoring processes. In Section 4.6, various case studies are discussed to illustrate the application of these neural network examples to tool condition monitoring.

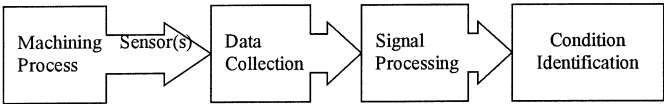


FIGURE 4.1 A general tool condition monitoring scheme.

4.2 Machining Tool Conditions

Machining can be performed efficiently if the tool geometries are very precise. Unfortunately, as the cutting process continues, wear changes the shape of the cutting edge and ultimately the tool life gets terminated when such changes are detected. In the following section, the wear mechanisms and the forms of tool wear are discussed.

Tool Wear Mechanism

The main mechanisms by which the cutting tools wear are attrition wear, abrasion wear, diffusion wear, edge chipping, and plastic deformation of cutting edge [25].

- **Attrition wear:** attrition wear is caused by the plucking out of microscopic fragments from the tool surface. In this wear mechanism, the work material seized to the tool is subsequently carried away by the moving chip and this imposes local tensile stresses on the cutting edge and many tiny particles may be torn out from the tool [41].
- **Abrasive wear:** abrasive wear is caused by the penetration and ploughing out of the hard particles from a softer surface. The inclusions in materials, such as carbides, oxides, and nitrides, that are harder than the tool material cause abrasive wear of the tool [36]. Since the built up edge (BUE) possesses high hardness, their fragmented torn out parts may also contribute to the abrasive wear of tools.
- **Diffusion wear:** diffusion wear takes place when the soluble atoms of the tool material diffuse into the work material across tool-work interface and are swept away with the chip [36]. This may also be caused by the diffusion of the work material into the tool and thus weakening the structure. Machining temperature, solubility, and cutting speed are the main controlling factors for diffusion rate.
- **Edge chipping:** at the start of a cut, when cutting with an uneven depth of cut or during interrupted cutting, sudden loads may be imposed on the cutting edge. Under such conditions, brittle carbide and ceramic tools may crack or fracture [40]. Cracks may also be formed due to thermal and mechanical fatigue arising from interrupted cutting [37]. If these cracks propagate it may cause small fragments of the tool to break away.
- **Plastic deformation of cutting edge:** during machining, the cutting edge of a tool is subjected to very high normal loads. At high cutting speeds, very high temperature is generated at the cutting edge of the tool and that may cause the tool material to soften and deform under high compressive stresses. This causes the tool tip to become rounded and blunt as a result of plastic deformation. The blunt cutting edge may become an additional source for further heat generation as it rubs against the freshly-machined work surface, causing further softening of the tool material and resulting in plastic collapse [40].

Different forms of tool wear and deformation processes, together with the regions on a tool in which these are likely to take place are shown in [Fig.4.2](#).

Forms of Tool Wear

Tool wear is a common cause of, and inevitable precursor, to tool failure in machining processes. The extent of the tool wear has a strong influence on the surface finish and dimensional integrity of the workpiece. Main forms of tool wear are flank wear, crater wear, nose wear and groove (or notch) wear. [Figure 4.3](#) shows the various wear and fracture surfaces that may be present on a worn cutting tool [38]. The measurement criteria for different types of wear are shown in the [Fig. 4.4](#) [38].

Flank Wear

The most commonly observed phenomenon is the flank wear. The width of the wear land, VB ([Fig. 4.4](#)), is often used as a quantitative parameter of tool wear and it can easily be measured quite accurately. With

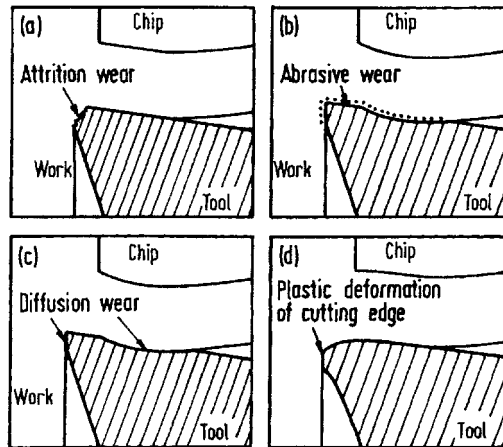


FIGURE 4.2 Wear mechanisms on cutting tools and their locations: (a) attrition wear, (b) abrasion wear, (c) diffusion wear, and (d) plastic deformation of cutting.

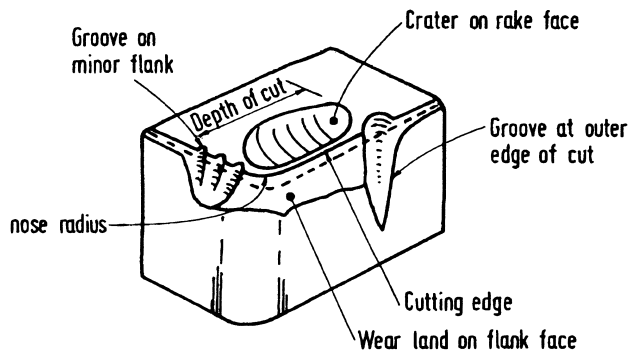


FIGURE 4.3 Various wear and fracture surfaces that may be present on a worn cutting tool.

the increase of flank wear, friction between the tool flank and the newly machined work surface increases and that leads to higher cutting forces and temperatures. Excessive flank wear may cause severe vibration leading to inefficient machining.

Crater Wear

Another form of common tool wear that appears is a crater on the rake face where the chips move over the tool surface. The formation of the crater usually starts at a distance from the cutting edge. With the elapse of time, it gradually becomes deeper and may lead to the weakening and breakage of the cutting edge. The crater wear is most common when cutting is done at relatively cutting speed on high-melting-point metals. The depth of crater, KT (Fig. 4.4), is often used as a quantitative parameter of tool wear.

Groove Wear

Groove wear may be observed at the flanks (both major and minor) at the position where the chip crosses the edge of the tool. If the groove wear becomes very deep, it may cause the tool to fracture. In practice, the tool life, which is a measure of the useful time span of the tool, is determined either by the flank wear or crater wear. Tool wear is attributed to several factors, such as the properties of the tool and workpiece materials, cutting conditions, tool geometry, and chip formation. When a tool wears, it affects

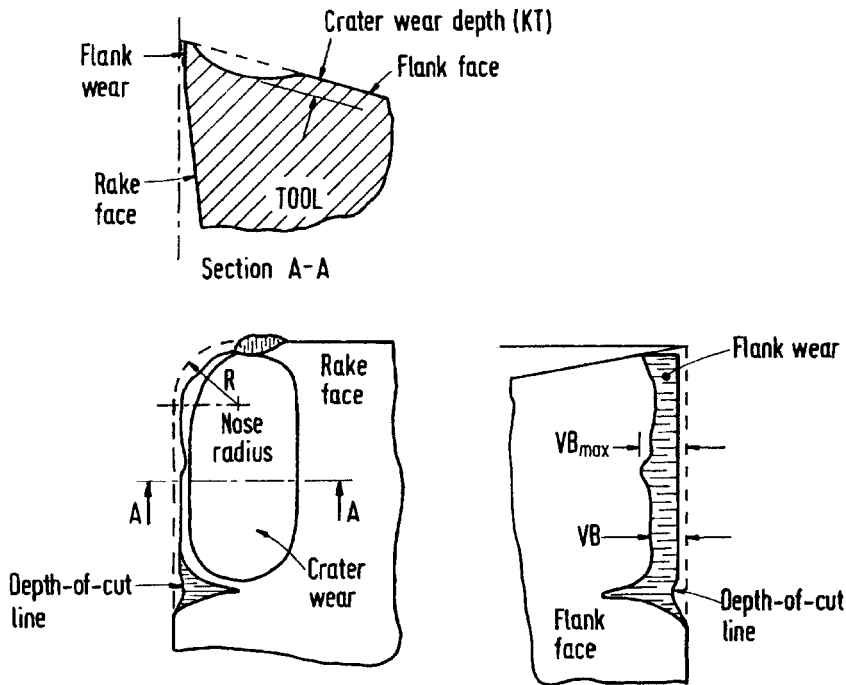


FIGURE 4.4 The measurement criteria for different types of wear.

the cutting force and the stress waves generated at the tool-workpiece interface due to friction and deformation. At times, there is increased vibration in the machine.

4.3 Sensors and Signal Processing

Tool conditions that have adverse effects on the machining process and need to be monitored include tool wear, various forms of tool damage, chatter, and chip breakage [30]. When the condition of a tool deteriorates, it affects the cutting force and the stress waves generated at the tool-workpiece interface due to friction and deformation. Sometimes, there is increased vibration in the machine. Thus, sensing methods measuring force, directly or indirectly (e.g., torque, current, or power of the spindle motor), acoustic emission, and vibration can be used to determine tool wear. The original force, (AE) and vibration signals require signal conditioning and processing to extract useful information. Commonly used tool wear sensing methods are based on cutting force and acoustic emission sensing. The following sub sections present examples of preprocesses force, AE and vibration signals and their correlation with tool conditions.

Dynamic Force

In the frequency domain, the dynamic tangential force has been found to exhibit a characteristic trend that can be used to indicate the extent of wear of the cutting tool [21]. Figure 4.5 shows the magnitude of the dynamic tangential force at the natural frequency of the tool overhang for a turning tool [27]. It shows a monotonic increase with the flank wear. As the tool approaches failure, it displays a relatively sharp decline. When the cutting tool is new, the sharp edge of the tool minimizes contact between the tool and the workpiece. Hence, the dynamic force is initially small. As the tool wears, the contact surface area between the tool and the workpiece increases, resulting in rubbing and increases in the dynamic tangential force. For an uncoated tool without grooves, as the tool approaches failure, crater wear becomes

ASSAB 60 Uncoated Tungsten Carbide P30 Without Grooves

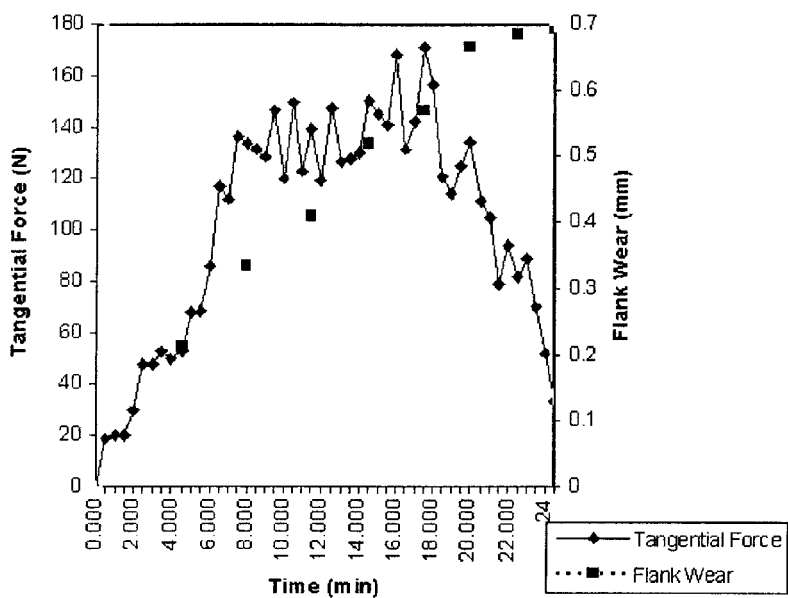


FIGURE 4.5 Dynamic tangential force at tool overhang frequency (turning).

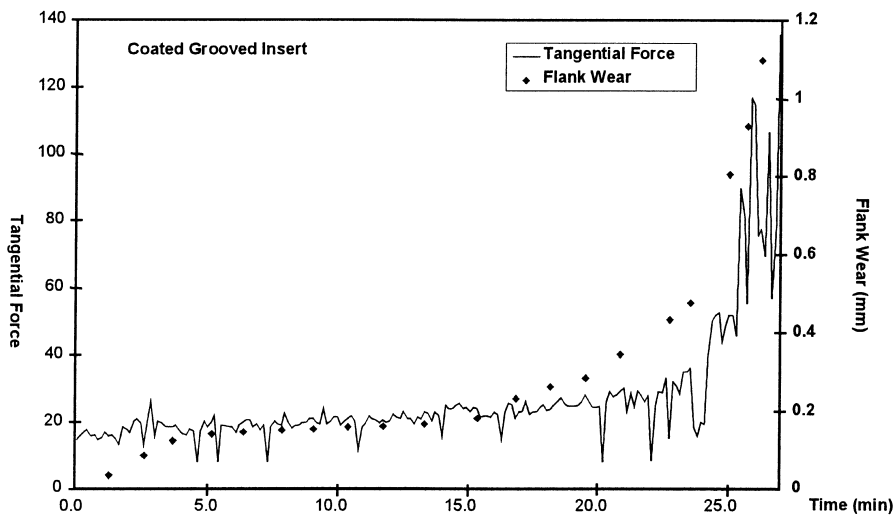


FIGURE 4.6 Increase in dynamic force component near tool failure.

significant. As the depth and width of the crater increases, it has the effect of increasing the effective normal rake angle (i.e., making the cutting edge relatively sharper) and correspondingly reduces the dynamic tangential force.

The decrease in the dynamic force component at the natural frequency of the tool overhang does not always occur. It depends on the manner at which the crater wear progresses and the geometry of the crater so formed. In the case of coated and grooved inserts, there is a sharp increase in the dynamic force component near tool failure. This continues until tool failure as can be seen in Fig. 4.6. During the cutting operation, the tool insert shows little crater wear because the chip-breaker groove does not allow chips to flow continuously across the rake face of the insert. Hence, crater wear is not the cause of tool failure

for this type of inserts. In the case of a coated insert, flank wear increases at a slow rate at the beginning of the tool life because of the titanium nitride (TiN) coating. When the coating is worn off, the flank wear proceeds at a faster rate until tool failure.

Generally, the use of the amplitude change in the dynamic force component is dependent on the insert characteristics and the machining conditions. This is one of the problems of relying on only a single feature information, such as that obtained in either a time or frequency domain, which in the aforementioned case, is the peak frequency amplitude in the frequency domain.

Acoustic Emission (AE)

AE refers to the emission of elastic stress waves due to rapid changes in strain energy as a result of structural change in the material, such as during plastic deformation, fracture, or phase change. The frequency range of the AE signal is much higher than that of the vibrations in the machine tool and environmental noise. Therefore, a relatively uncontaminated signal can be obtained by using a high-pass filter. Figure 4.7 shows the flank wear of a turning insert, the corresponding resonant force at the natural frequency of the tool overhang and the root mean square (RMS) and band power in the frequency band of 300 to 600KHz of the AE signal. Both the resonant force and AE-RMS show a surge in amplitude as the tool approaches failure. The band power in the frequency band of 300 to 600 kHz indicates increased bursts of AE in the higher frequency range, due most probably to higher incidents of cracking and chipping of the insert.

Wavelet Packet Analysis of AE and Force Signals

Signal processing approaches of the AE and force signals for the purpose of feature extraction typically utilize the time-domain or frequency-domain analytical method. These approaches include the use of FFT, statistical analysis, and stochastic modeling (such as AR, ARMA). They provide only the time or frequency domain information that is generally more suitable for the analysis of stationary process. The machining process, on the other hand, may not necessarily be stationary. A global signal processing method which incorporates both the time and the frequency domains is more suitable for the analysis of a non-stationary process. The short-time Fourier transform (STFT) is currently the standard method for the analysis of a non-stationary process. The short-time Fourier transform (STFT) is currently the standard method for the time-frequency analysis of signals [4]. Another commonly used method is the Wigner distribution, which originates from the classical works of Wigner in 1932. Du et al. [9] have used the exponential time-frequency distribution of acoustic emission for tool wear study in turning. In the last few years, a new method called the wavelet analysis has been developed with much progress both in the theoretical and applied areas. Like the Wigner distribution and short-time Fourier transform (STFT), the wavelet transform provides time-frequency analysis of signals but the wavelet transform adapts the window. A short window is needed to achieve the required fine resolution in the high-frequency range while a long window is needed to encompass the low-frequency range. The Wigner distribution and windowed Fourier transform use a uniform window size which can result in an uneven, and low frequency resolutions. Meanwhile, the Wigner distribution contains interference terms which are undesirable in practical applications. The wavelet analysis is a more advanced signal processing method. It offers the significant advantage of multi-resolution analysis of signals. Hence, both long and short windows can be used to capture the desired signal features. As a generalization of the wavelet transform, the wavelet packet (WP) decomposition together with the best basis algorithm have been developed by Coifman et al. [5], which can represent the signal in a most compact way. The wavelet packet analysis allows for the representation of the signal in both the time and frequency domains. Another appealing characteristic is that the computation time of this algorithm is as fast as the FFT and on-line implementation is possible. The wavelet packets contain modulated waveforms which have good time-frequency localization properties. Each waveform is called a time-frequency atom and can be assigned three parameters: frequency (f), scale(s), and position (p). The definition of the wavelet packets $\{W_{f,s,p}\}$ and wavelet packet decomposition are given in [11].

Materials	Workpiece	ASSAB760
	Tool Insert	Coated-Grooved Tungsten Carbide
Conditions	Cutting Speed (m/min)	230
	Feed Rate (mm/rev)	0.4
	Depth of Cut (mm)	2.0

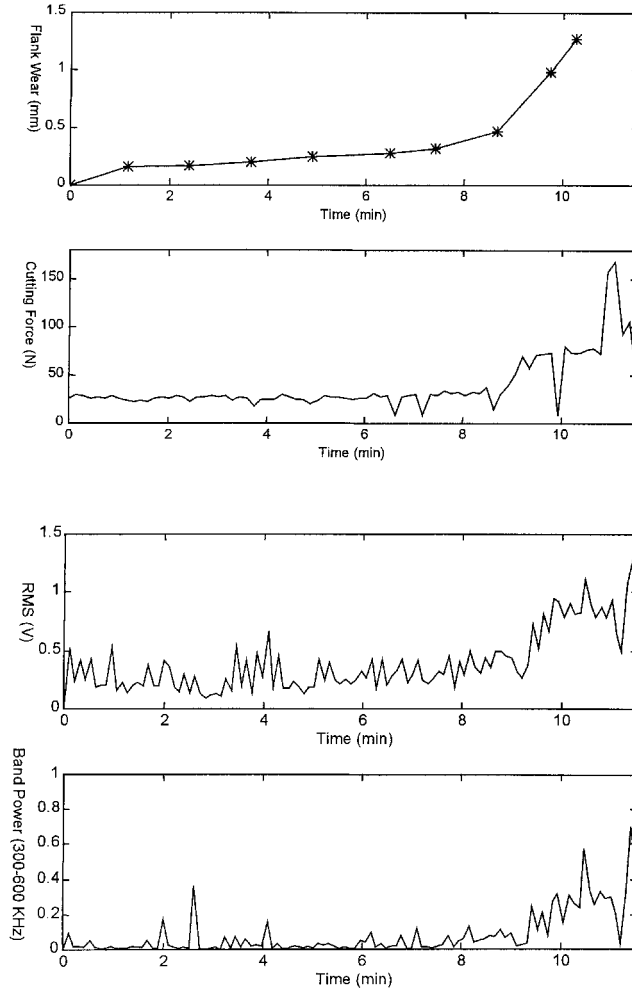


FIGURE 4.7 Flank wear, cutting force at tool overhang frequency and RMS and band power of acoustic emission (turning).

Given a signal x , the wavelet packet transform can be represented as a linear mapping:

$$WP: \begin{cases} X^n \Rightarrow W^n \\ \mathbf{x} \rightarrow \mathbf{w}(f, s, p) = WP(\mathbf{x}) = \langle \mathbf{x}, W_{f,s,p} \rangle \end{cases} \quad (4.1)$$

The above mapping represents the orthonormal wavelet transform. Hence, there exists a unique inverse transform to reconstruct the signal x :

$$WP^{-1}: \begin{cases} W^n \Rightarrow X^n \\ \mathbf{w}(f, s, p) \rightarrow \mathbf{x} = WP^{-1}(\mathbf{x}) = \langle \mathbf{w}(f, s, p), W_{f,s,p} \rangle \end{cases} \quad (4.2)$$

In the above wavelet packet transform, any basis sets from the orthonormal bases can be used to represent the analyzed signal. This provides more freedom in deciding which basis is to be used to represent the given signal. There must be a *best-basis* among them which can represent the signal in the most compact way (i.e., using the least number of coefficients). In order to find this best-basis, Coifman and Wickerhauser [5] introduced the concept of the information cost function which is defined to be real-valued. Then the cost function is used as a measure to search for its minimum over all bases in the wavelet packet library to obtain such a *best basis* for the signal. It should be larger when the coefficients are roughly the same size and smaller when most coefficients are negligible, but only a few coefficients need to be retained. One of the cost functions proposed is the entropy cost function defined as follows. The Shannon-Weaver entropy of a sequence $x = \{x_j\}$ is:

$$H(x) = -\sum_j p_j \log p_j \quad (4.3)$$

where $p_j = \frac{|x_j|^2}{\|x\|^2}$. For this entropy, $\exp(H(x))$ is related to the number of coefficients needed to represent the signal to a fixed mean square error. The search for the best-basis requires $O(N \log N)$ operations.

Through the above process, a compact set of wavelet packet coefficients is obtained. A phase-plane plot is employed to graphically represent the time-frequency properties of the analyzed signal. In the phase-plane plot, the wavelet packet coefficients are displayed on a 2-D time-frequency plane. In this representation, each wavelet packet coefficient is associated with a time t and frequency f , with its time and frequency uncertainty amount Δt and Δf , respectively. The result is interpreted as a rectangular patch of dimensions Δt by Δf , located around (t, f) on the phase plane. The smallest area (Δt by Δf) of the rectangular patch is limited by the *Heisenberg uncertainty principle*. The patch is assigned a color or gray scale in proportion to the amplitude of the corresponding coefficient. Figure 4.8 gives a phase plane of a wavelet packet. It is obvious that the phase plane representation of the signal can give us the global view of its time-frequency feature patterns.

The best-basis wavelet packet decomposition provides an efficient and flexible scheme for time-frequency analysis of non-stationary signals. The major advantages of this method lie in the following:

- It globally optimizes the signal representation to provide a compact and sparsity representation of the signal pattern, i.e., the one with the fewest significant coefficients; and
- It possesses the computational speed of FFT.

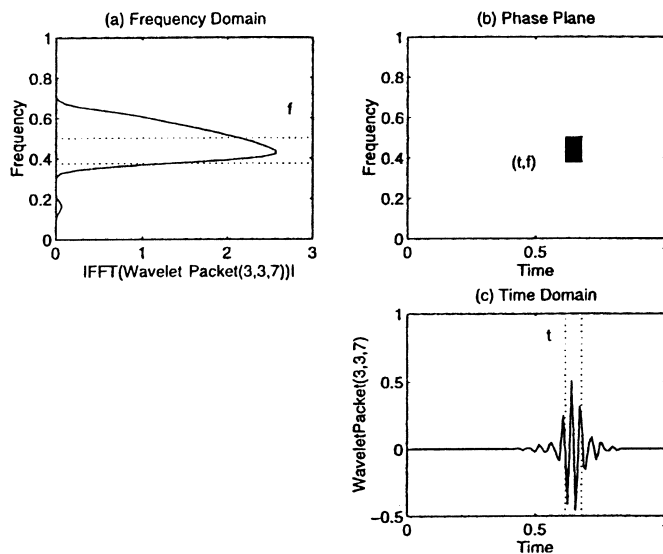


FIGURE 4.8 Phase plane of a wavelet packet.

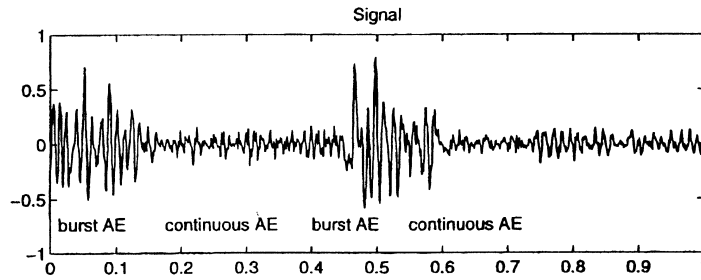


FIGURE 4.9 Continuous- and burst-type AE signals.

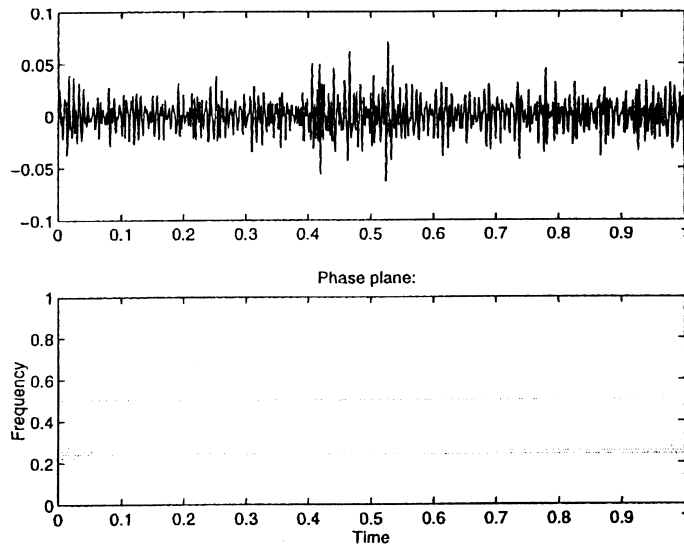


FIGURE 4.10 Normal state of tool.

The acoustic emission signal is usually of two distinct types, continuous and discontinuous or burst (see Fig. 4.9). In the machining process, the continuous-type AE signals are generated in the shear zone, at the tool-chip interface, and at the tool-flank-workpiece interface while the discontinuous burst-type signals are generated due to tool fracture, chipping, and chip breakage [11]. The wavelet packet method has been applied to the AE signals from the turning process to obtain a comprehensive set of time-frequency feature patterns (in the form of phase plane) corresponding to important tool conditions [29, 16, 17]. Figures 4.10, 4.11, 4.12, 4.13, and 4.14 show the time-frequency feature patterns in the phase plane representations after the best-basis wavelet packet transform of the corresponding AE signals obtained for different tool conditions.

Vibration (Acceleration)

Another possible approach to monitoring tool wear in turning is to measure the vibration of the tool shank using an accelerometer. For example, there exists a peak at the natural frequency f_0 of the tool overhang in spectra of the acceleration signal. Sometimes, another peak can also be found at half of the resonant frequency $f_{1/2}$ as shown in Fig. 4.15(a). However, $f_{1/2}$ does not always exist, as can be seen in Fig. 4.15(b) for the case of the coated tool under the same machining conditions. Studies on the tool shank vibration indicate that although there exists a characteristic trend in the amplitude of the natural

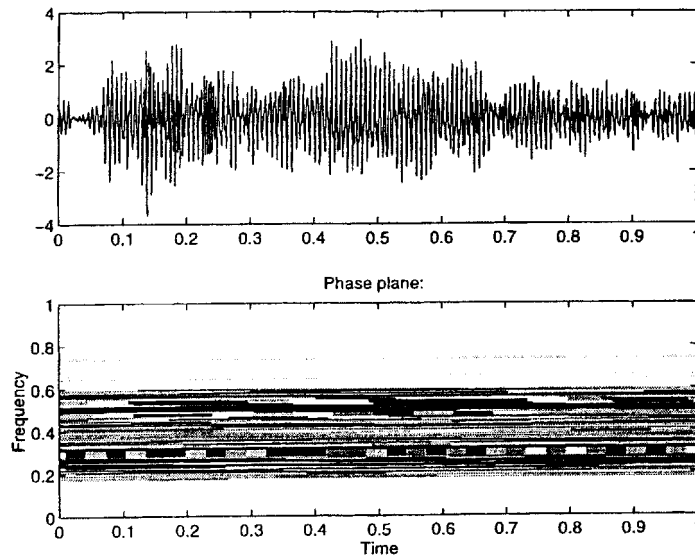


FIGURE 4.11 Intensive tool wear.

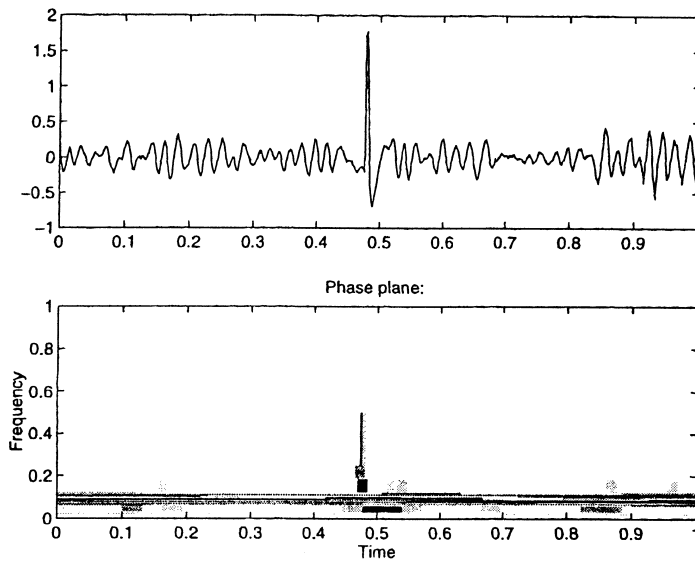


FIGURE 4.12 Tool chipping.

frequency f_0 , this trend is not consistent [13, 24]. A more reliable approach is to employ the coherence function of two cross vibration signals.

Coherence Function of Cross Vibration Signals

The coherence function between two cross accelerations from the bending vibration of the tool shank has been found to be suitable for the identification of both tool wear and chatter in turning [23]. The cross accelerations in the horizontal (X) and vertical (Z) directions are measured by two piezoelectric

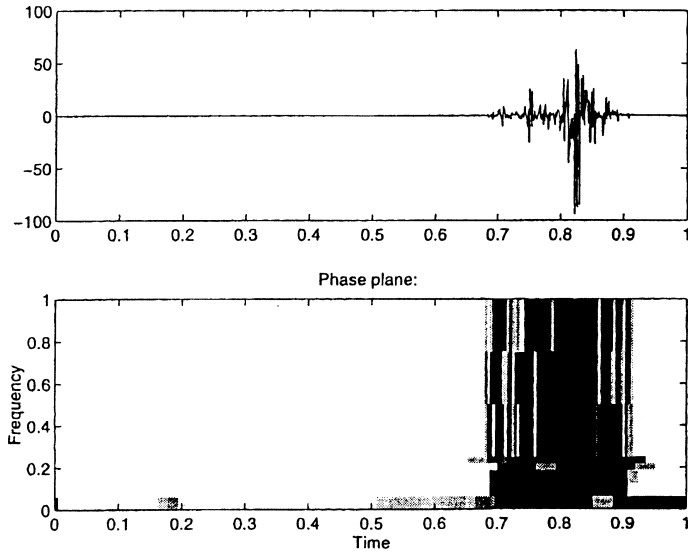


FIGURE 4.13 Tool fracture.

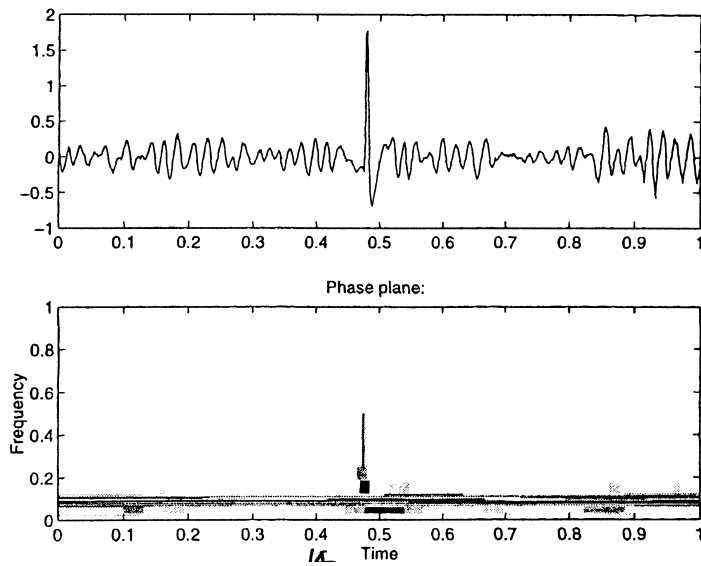


FIGURE 4.14 Chip breakage.

sensors mounted on the tool shank in a CNC lathe, as shown in Fig. 4.16. The coherence function of the two acceleration signals $x(t)$ and $z(t)$ is defined as

$$\gamma^2(f) = \frac{|G_{xz}(f)|^2}{G_x(f)G_z(f)}$$

$$0 \leq \gamma^2(f) \leq 1$$

where $G_x(f)$ and $G_z(f)$ are the respective auto-spectra of $x(t)$ and $z(t)$, and $G_{xz}(f)$ is the cross-spectrum between $x(t)$ and $z(t)$. $\gamma^2 = 0$ when $x(t)$ and $z(t)$ are uncorrelated over the range of frequency f . $\gamma^2 = 1$

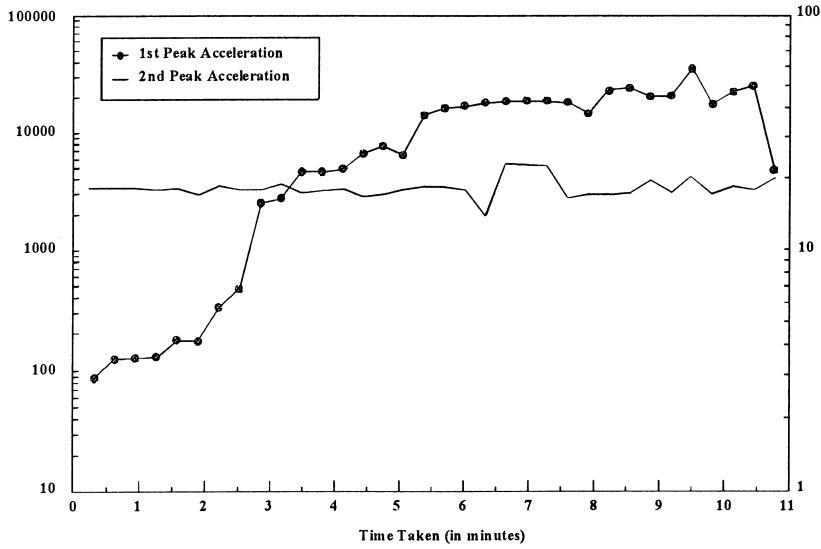


FIGURE 4.15(a) Acceleration of tool overhang (uncoated turning insert).

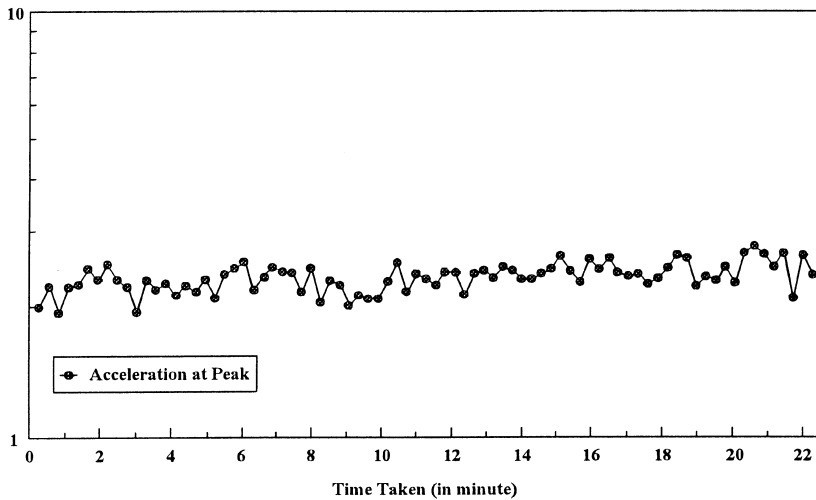


FIGURE 4.15(b) Acceleration of tool overhang (coated turning insert).

when $x(t)$ and $z(t)$ are completely correlated for all f . Otherwise, $0 \leq \gamma^2 \leq 1$ according to the degree of correlation between $x(t)$ and $z(t)$.

Figures 4.17(a) and (b) show the coherence function of the acceleration signals measured during turning of a nickel-based super alloy [4]. As can be seen from the figures below, the value of the coherence function at the chatter frequency reaches unity at the onset of chatter. Its values at the first natural frequency of the tool shank approach unity in the severe tool wear stage. The advantage of using this method is that the thresholds for detecting severe tool wear and chatter can be easily set for the following two reasons: first, the values of coherence function are normalized to a range of between zero to unity and secondly, they are also not so susceptible to changing cutting conditions because the value of coherence function is close to unity at the onset of chatter and severe tool wear.

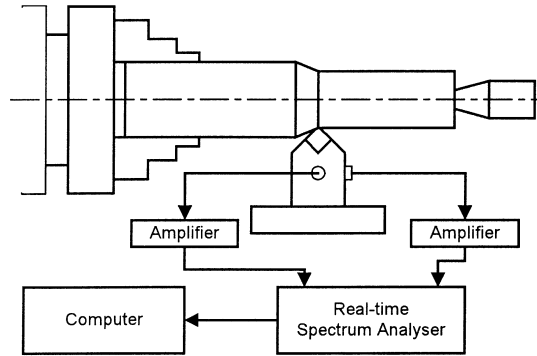


FIGURE 4.16 Setup for measurement of X- and Z-accelerations of tool shank.

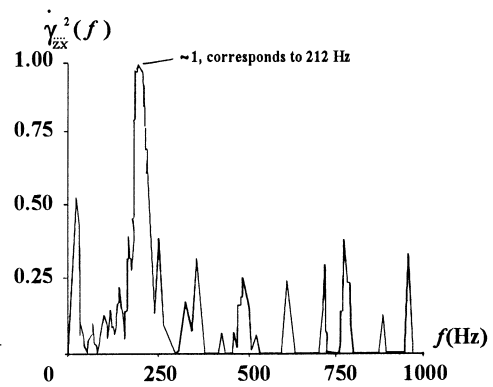


FIGURE 4.17(a) Coherence function of X- and Z-accelerations at onset of chatter.

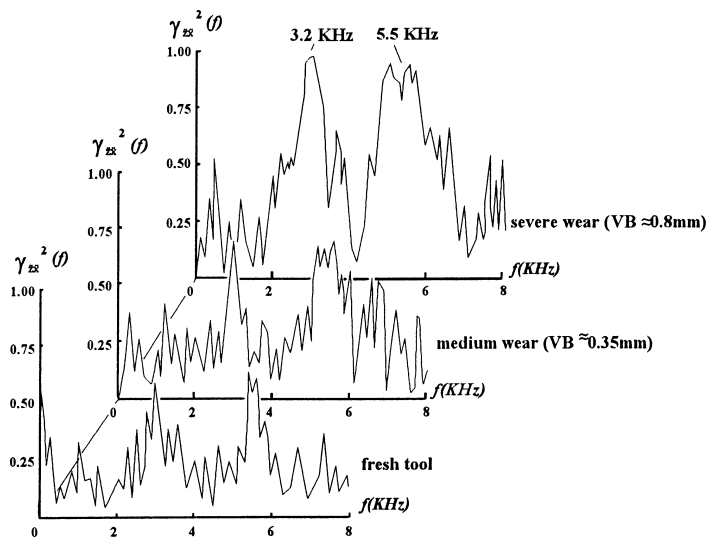


FIGURE 4.17(b) Coherence function of X- and Z-accelerations at different tool wear.

4.4 Feature Extraction

Once a preprocessed signal exhibits characteristics of features that can be correlated with the tool conditions, it is necessary to develop suitable techniques to extract those features. The successful identification of the various tool conditions depends on how the features can be reliably extracted from the raw or preprocessed sensor signals for the sensor fusion stage. These features can be expressed in the form of a feature vector from different signal sources. The selected features should comprehensively characterize the different tool conditions. Although redundant features can reinforce the decision process in the determination of the tool condition, a minimum possible set of complementary features may have to be used to reduce computation requirements. For reliable and generalized identification of tool conditions, features used should not be sensitive to machining parameters, such as speed/feed, tool/work-piece materials, tool geometry, etc. In practice, this requirement is very difficult to meet. One method of achieving this is to have some means of normalizing the preprocessed output, such as the use of the aforementioned coherence function. By using a normalized output, threshold setting is confined to a range between zero and unity and is simpler to set. Another is to employ some form of pattern recognition to deduce the state of the tool from characteristic patterns of the processed outputs rather than by some single-value threshold setting.

4.5 Neural Network Architectures

After the feature extraction, an inference process is needed to associate each feature vector to a corresponding tool condition. This classification process usually involves an inference engine to map groups of feature vectors to their associated tool conditions. Such process, especially when crisp logic rule based methodology is employed, can be extremely tedious for feature vectors of large dimension. As discussed in the previous section, more extracted feature components are needed to reinforce the identification decision and complement each other in their limited operational range. A more practical approach is to use an intelligent system which can “learn” itself to correct inference decision via past examples. A solution to such problem is to use an artificial neural network as a classifier [8, 16, 17, 29, 33, 34].

An artificial neural network is an information-processing system inspired by the performance of the human brain. It consists of a collection of very simple processing element called the neuron as shown in Fig. 4.18. The neuron embodies a node (soma) with multiple inputs (dendrite) and an output (axon). The output of the node can be described by

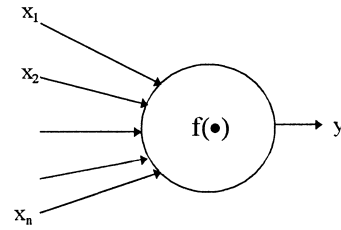


FIGURE 4.18 A simple processing element.

$$y = f(w_1, w_2, \dots, w_n, x_1, x_2, \dots, x_n) \quad (4.3)$$

where $w_i, \forall i \in [1, n]$ is the weight of the node associated with each input and $f(\cdot): \mathbf{R}^n \rightarrow \mathbf{R}$ is called a transfer function of the neuron.

The collection of such neurons forms a neural network. A type of neural network is characterized by three fundamental factors as follows:

- Transfer function—the transfer function defines the way the output associated with the inputs and their corresponding weights, and the type of activation function used.
- Architecture—the architecture defines the pattern of connections between neurons and the number of layers of connection.
- Learning algorithm—the method of determining the weights on the connections.

There are numerous types of neural networks developed to suit the needs of their applications [12, 14, 18, 45]. This chapter does not intend to show all types of neural networks but the three common types of neural networks that are used in the case study have examples in the following sections. There are Multi-layer Perceptron (MLP), Kohonen Network, and Adaptive Resonance Theory 2 (ART2).

Multi-Layer Perceptron (MLP)

This is the most common neural network model used by many researchers and is better known as the back-propagation network due to the way it is trained [14]. A simple M-layer MLP is illustrated in Fig. 4.19 with each neuron represented by a circle and each interconnection, with its associated weight, by an arrow. For each neuron, the i^{th} neuron in p^{th} layer, its output, can be described by the equations

$$h_i^{(p)} = \sum_{j=1}^{N_{p-1}} w_{ij}^{(p)} v_j^{(p-1)}, \quad \text{for } p \in [1, M] \quad (4.4)$$

$$v_i^{(p)} = f_p(h_i^{(p)}) \quad (4.5)$$

where $f_p(\bullet)$ is any monotonically increasing activation function. A common choice is to use the sigmoid function for all the hidden layers and a linear function for the output layer. That is:

$$f_p(h) = \begin{cases} \frac{1}{1 + e^{-h}}, & \text{for } p \neq M \\ h, & \text{for } p = M \end{cases} \quad (4.6)$$

The number of hidden layers in the network and the number of nodes for each layer are arbitrary. Hecht-Nielsen [15] has proved that one hidden layer is sufficient to approximate a bounded continuous function to an arbitrary accuracy. However, one hidden layer may result in an excessive number of neurons used in the hidden layer. Thus, for practical implementation, one hidden layer is used for simple functions and two hidden layers are used for functions that are more complex.

The learning law employed by this network is the gradient decent based delta rule:

$$w_{ij}^{(\text{new})} = w_{ij}^{(\text{old})} + \Delta w_{ij} \quad (4.7)$$

and

$$\Delta w_{ij} = -\eta \frac{\partial J}{\partial w_{ij}} \quad (4.8)$$

where J is the cost function usually defined as $\sum_{i=1}^{n_y} (\zeta_i - y_i)^2$ and ζ_i is the desired output of the i^{th} node.

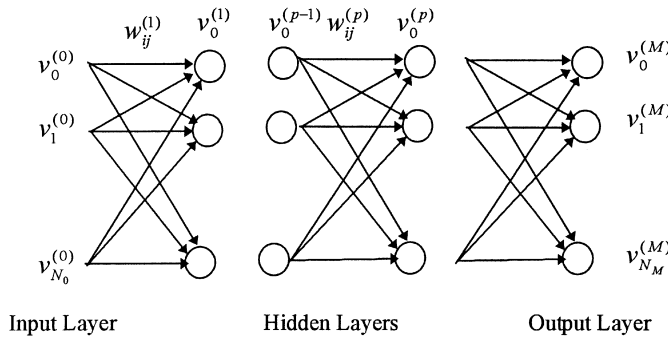


FIGURE 4.19 A M-layers multi-layer perceptron networks.

With the assumption of $f_x(\bullet)$ and $f_y(\bullet)$ as defined in Eq. 4.7 and Eq. 4.8, respectively, the training procedure can be summarized as follows:

- Step 1: Read in the input pattern and the desired output (x, ζ)
- Step 2: Propagate the signal forward by Eq. 4.5 and Eq. 4.6
- Step 3: Calculate the output layer error, δ as

$$\delta_i = \zeta_i - y_i \quad (4.9)$$

- Step 4: Back propagate this error to the hidden layer by

$$\delta_i^{(p-1)} = f'(h_i^{(p-1)}) \sum_j w_{ji}^{(p)} \delta_j^{(p)} \quad (4.10)$$

- Step 5: Calculate Δw_{ij} by

$$\Delta w_{ij}^{(p)} = \eta \delta_i^{(p)} y_j^{(p-1)} \quad (4.11)$$

- Step 6: Update the weights by

$$w_{ij}^{(p)} = w_{ij}^{(p)} + \Delta w_{ij}^{(p)} \quad (4.12)$$

- Step 7: Go to Step 1

Kohonen Networks

Kohonen Network [19] is comprised of a single layer network as shown in Fig. 4.20. This network uses a learning technique called competitive learning. Only one component of the output vector is activated (or ON) at a time. It is a network used for clustering operation.

For each output neuron, the output compete with one another by the equation below:

$$h_i = D(\mathbf{w}_i, \mathbf{u}) \geq 0 \quad (4.13)$$

$$y_i = \begin{cases} 1, & h_i = \min_j \{h_j\} \\ 0, & \text{otherwise} \end{cases} \quad (4.14)$$

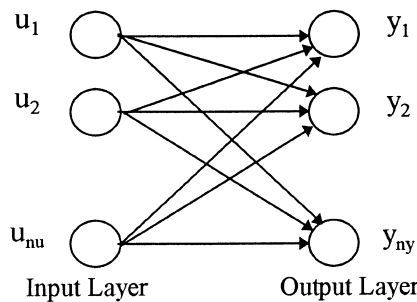


FIGURE 4.20 A Kohonen Network.

where $D(\cdot, \cdot) : \mathbf{R}^{n_u} \times \mathbf{R}^{n_u} \rightarrow \mathbf{R}$ is an metric function which measures the “distance” vectors and $\mathbf{w}_i \in \mathbf{R}^{n_u}$ is the weight vectors that connect the i^{th} out with the input vector \mathbf{u} . For this network, the activation function sets the output node whose weight vectors is closest to the input pattern to one and the rest of the output node to zero. The metric function $D(\cdot, \cdot)$ can be any distance measurement, say the Euclidean norm. However, for computational simplicity, the Hammin distance

$$D(\mathbf{w}_i, \mathbf{u}) = \sum_{j=1}^{n_u} |w_{ij} - u_j| \quad (4.15)$$

is more commonly used.

The learning process for the Kohonen network only involves the weights associated with the winning output nodes. Hence, Kohonen learning is sometime referred to as competitive learning. The Kohonen learning law can be presented as

$$\mathbf{w}_i^{(\text{new})} = \mathbf{w}_i^{(\text{old})} + \eta(\mathbf{u} - \mathbf{w}_i^{(\text{old})})y_i, \quad \forall i \in [1, n_y]. \quad (4.16)$$

As shown in Eq. 4.17 that the learning law employed by Kohonen network does not involve the desired output, it does not require past examples to train the network. It is sometime referred as self-organizing network.

ART2 Networks

The adaptive resonance theory (ART) was developed by Carpenter and Grossberg [2] in two forms. One form, ART1 is designed for handling binary pattern whereas ART2 [1] can accept patterns of continuous magnitude. As shown in Fig. 4.21, a typical ART2 network is composed of two successive stages or layers of cells—an input representation layer F_1 and a category representation layer F_2 . These layers are linked by feedforward and feedback weight connections (w_{ij}, ω_{ji}) that define a pattern specified by a corresponding F_2 cell. Each F_1 cell consists of three processing sub-layers with six nodes that enhance the

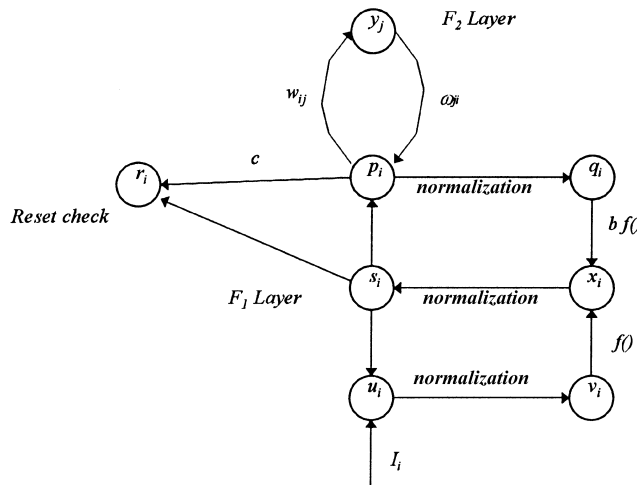


FIGURE 4.21 An ART2 network.

salient feature and suppress noise in the received signals, as shown in Fig. 4.21. Their activities can be characterized by the following equations:

$$u_i = I_i + as_i \quad (4.17)$$

$$v_i = \frac{u_i}{e + \|\mathbf{u}\|} \quad (4.18)$$

$$t_i = f(v_i) + bf(q_i) \quad (4.19)$$

$$s_i = \frac{t_i}{e + \|\mathbf{t}\|} \quad (4.20)$$

$$p_i = s_i + \sum_{j=1}^M g(y_j) \cdot \omega_{ji} \quad (4.21)$$

$$q_i = \frac{p_i}{e + \|\mathbf{p}\|} \quad (4.22)$$

where

$$f(x) = \begin{cases} x, & \text{if } x \geq \theta \\ 0, & \text{if } x < \theta \end{cases} \quad \text{for } (0 \leq \theta \leq 1) \quad (4.23)$$

Every cell in the $F2$ layer competes with the others by the rule denoted in Eq. 4.25 until the only winner remains active:

$$g(y_j) = \begin{cases} d, & \text{if } y_j = \max_{r \in [1, M]} \{y_r\}, \text{ i.e., } j^{\text{th}} F2 \text{ cell is active} \\ 0, & \text{otherwise} \end{cases} \quad (4.24)$$

in which

$$y_j = \sum_{i=1}^N p_i \cdot w_{ij} \quad (4.25)$$

The degree of match between an input pattern and a responded feedback pattern is measured by the combined normalized feedforward/feedback vector $\mathbf{r} = (r_1, r_2, \dots, r_N)$ with

$$r_i = \frac{s_i + cp_i}{e + \|\mathbf{s}\| + \|\mathbf{cp}\|} \quad (4.26)$$

where $0 < c < 1$ and $\frac{cd}{1-d} \leq 1$.

The vigilance parameter ρ , which has a value of between 0 and 1, determines how well an input pattern matches with the feedback pattern of the active $F2$ cell. The closer ρ is to 1, the more sensitive is the system to mismatches. The matching criterion is defined as shown in the table below.

Condition	Match Result
$\ \mathbf{r}\ < \rho - e$	mismatch
$\ \mathbf{r}\ \geq (\rho - e)$	proper match)

(4.27)

where

- a. For $\|\mathbf{r}\| \geq \rho - e$, $F2$ is reset so that the active cell is deactivated with an output value of zero. Meanwhile, the input pattern is considered to mismatch with the stored pattern.
- b. For $\|\mathbf{r}\| \geq \rho - e$, $F2$ is not reset. The active cell supposedly presents the proper class of input pattern during the classification operation, and is activated with an output value of one. Besides, if the active $F2$ cell has not been encoded with any pattern, such as during the initial learning, then a new category is established in the weighted connections with the cell, or the input pattern is merged into the weighted connections with the correctly encoded or stored pattern during the multi-sample learning cycle.

The feedforward and feedback weights are adjusted according to the following equations

$$w_{ij}(t+1) = [1 - g(y_j)][w_{ij}(t) - p_i] + p_i \quad (4.28)$$

$$w_{ij}(t+1) = \alpha g(y_i)s_i + \{1 + \alpha g(y_j)(g(y_j) - 1)\}w_{ij}(t)$$

$$\omega_{ij}(t+1) = \alpha g(y_j)s_i + \{1 + \alpha g(y_j)(g(y_j) - 1)\}\omega_{ij}(t) \quad (4.29)$$

Further details of the ART2 algorithm can be found in the description by Carpenter and Grossberg [1].

4.6 Tool Condition Identification Using Neural Networks

Features extracted from the sensor inputs are synthesized for more reliable and accurate estimation of the state of the process. Basically, tool condition diagnostics approaches can be divided into two categories: [3]

- Statistical Approach. In this approach, estimates of the state variables are obtained by evaluating process models based on the physics of the process. Statistical data concerning the physical properties of the materials are used. Bayesian estimator and Shafer-Dempster reasoning methods are examples of this approach.
- Training Approach. The synthesis occurs through a *mechanism* or a *mode* which first *learns* through a training phase on how the synthesis should occur. Methods based on this approach go through a training phase to capture the behavior or learn to synthesize the data.

This chapter presents the works based on the training approach, in which, neural networks are used as the learning mechanism.

The following sections demonstrate three different approaches to the application of the neural network for tool condition identification. The first example uses very simple preprocessing and an MLP network to identify various tool conditions for the turning process. The following three examples use a combination of ART2 networks. Two are based on a single-ART2 network and employ a two-step feature extraction strategy for the identification of tool wear state. The first single-ART2 network uses only

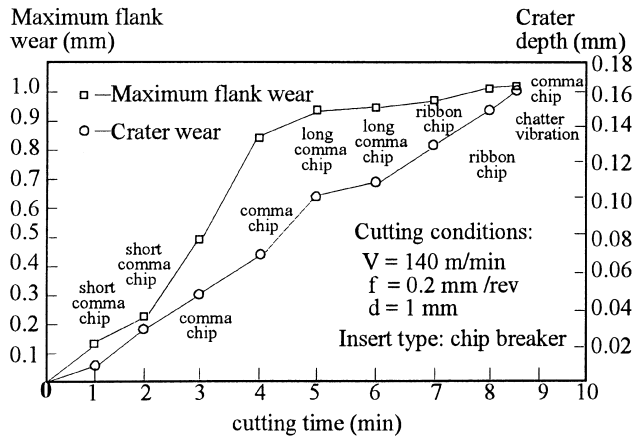


FIGURE 4.22 A typical tool wear experiment.

acoustic emission sensing and a wavelet transform technique in the signal processing. The other uses both force and acoustic sensing. The third neural network consists of multiple ART2 sub-networks and uses three sensors to measure acoustic emission and cross vibrations from the tool shank. The input feature information to the multi-ART2 network is obtained from the distribution of the coherence function between two cross vibration signals and the power spectrum of the acoustic emission signal in the effective frequency bandwidth of these signals.

MLP for Force Sensor with Simple Pre-processing [32]

In modern manufacturing processes, an essential part of a machining system is the ability to monitor and automatically diagnose the faults that occur in the machining process. In practice, such faults are not isolated and they are often co-existing. Figure 4.22 shows the coexistence of various cutting conditions in a typical turning process. It is therefore important to identify these faults from the detected abnormal phenomena during the machining process.

The occurrence of a fault can be due to either system characteristics' changes (true reflection of a fault) or propagation of other undetected faults. It is important that the root of the fault must be detected correctly. Otherwise, this may result in a wrong control action being taken. The process of fault detection in machining processes consists of the detection of abnormality, followed by the identification of the cutting state and then the recognition of the cause of the fault. This section focuses the discussion on the monitoring of tool wear, chatter vibration, and chip breaking in a turning process.

Feature Requirements

From the pattern recognition viewpoint, the necessity for process diagnostics is to extract enough independent governing features for the purpose of identification. In this experiment, the exact value of the dynamic cutting force is not important. It is the sudden variation from its normal zero-mean value that reflects the presence of chip breaking. This phenomenon is reflected as "spikes" in the dynamic cutting force. The magnitudes of the spikes are more distinctive in the worn tool than the fresh tool. These spikes usually have negative values of a few times the magnitude of the normal zero-mean values. Therefore, a subtraction from its maximum value will ensure all data to have positive value and high spikes are the point of chip breaking as suggested by Eq. (4.31). However, these spikes are usually of high magnitude. Therefore, the set of data is divided into ten data subsets defined in Eq. (4.30) to avoid these chip-breaking spikes from dominating the other inherent properties of the signal. Four features are

extracted from the cutting force in feed direction with the following operation:

1. The signal (2000 data points) is passed through a low-pass filter to eliminate the high frequency noise.
2. This signal is divided into ten data subsets S_i , for $i \in [0, 9]$ with

$$S_i = \{x(200i + 0), x(200i + 1), \dots, x(200i + 199)\} \quad (4.30)$$

where $x(i)$ is the filtered dynamic component of cutting force.

3. For each S_i , the data are transformed into an absolute valued data by

$$y_i = |x(200i + t) - x_{\max}^i|, \text{ for } t \in [0, 199], i \in [0, 9] \quad (4.31)$$

where

$$x_{\max}^i = \max[x(200i+0), x(200i+1), \dots, x(200i + 199)] \quad (4.32)$$

4. For each pre-processes data $y_i(\cdot)$, the following features are extracted.

- The mean value F_i of y_i is defined by

$$F_i = \frac{\alpha_1}{200} * \sum_{t=0}^{199} y_i(t) \quad (4.33)$$

where α_1 is a scaling factor.

- Similarly, the variance V_i of y_i is defined as

$$V_i = \frac{\alpha_2}{200} * \sum_{t=0}^{199} [y_i(t) - F_i]^2 \quad (4.34)$$

where α_2 is a scaling factor.

- To avoid the presence of too many high valued spikes, the ratio R_i of mean and variance is also introduced as

$$R_i = \alpha_3 * V_i / F_i \quad (4.35)$$

- Another dominating property is the magnitude D_i of the spike which is defined as

$$D_i = \alpha_4 * \max\{y_i(t)\}, \text{ for } t \in [0, 199] \quad (4.36)$$

- One important property is the frequency of chip breaking. This can be measured via its coherent coefficient P_i which is defined by

$$P_i = \left| \frac{\sum_{t=1}^{199} [y_i(t) - F_i] * [y_i(t-1) - F_i]}{\sum_{t=1}^{199} [y_i(t-1) - F_i]^2} \right| \quad (4.37)$$

These features form the input to the MLP for the diagnosis process.

An Integrated Fault Diagnosis Scheme

In this case, an intelligent faults diagnosis scheme (IGDS) is proposed. For a fixed interval, the IFDS collects 2000 points of the dynamic cutting force component measured in the feed direction which are sampled over a time span of 0.8 second. This set data is filtered to cut-off the high-frequency noise

component. The filtered data are divided into ten subsets of 200 points, in which five features are extracted from each subset to form a feature vector. These ten feature vectors are used in the MLP.

The MLP has an architecture of 5-8-3 and sigmoid function as its threshold function. The five input nodes of the MLP correspond to the five feature components of the extracted feature vector and these will be explained in the next section. The minimum of one hidden layer in a multi-layer perceptron is inevitable due to the linear separability properties of perceptron. This experiment shows that a hidden layer with eight nodes is sufficient for the purpose of identification. The output values of the MLP are between 0 and 1. When a binary decision is needed, output values of less than 0.5 are considered to be 0 or otherwise to be 1. The first, second, and third component of the output layer represent the level of tool wear, chatter vibration, and chip breaking, respectively and can distinguish typical patterns: (0,0,0), (0,0,1), (0,1,0), (0,1,1), (1,0,0), (1,0,1), (1,1,0), (1,1,1). Noting that, (0,1,1) and (1,1,1) are redundant in actual machining process. This is because long chips are not produced when chatter vibration occurs.

Experiment

A lathe was used in this experiment on chip breaking, chatter vibration, and tool wear in turning. Three workpiece of ASSAB 760 steel each with sizes of $\phi 60 \times 1000$, $\phi 130 \times 700$, and $\phi 130 \times 260$ mm. Two types of inserts, AC25 coated and G10E uncoated, were used. The cutting speeds were varied from 50–160 m/min, feed rates from 0.1–0.4 mm/rev. and depth of cut from 0.5–1.5 mm. Three kinds of chip breaking experiments has been carried out under the conditions mentioned above. The first was under cutting state of chatter vibration with fresh tool, and the size of the workpiece was $\phi 60 \times 1000$ mm and AC25 coated inserts were used in the experiments to guarantee that no tool wear occurred in one cutting pass. The second was under cutting state of chatter vibration with a worn tool and the size of the workpiece was $\phi 130 \times 700$ mm G10E uncoated inserts were used to ascertain at least the average flank wear land width V_{Bp} of 0.3 mm in one cutting pass and when AC25 coated inserts were used, no chatter vibration occurred in machining process. The third was under the cutting state of no chatter vibration with a worn tool. The size of the workpiece was $\phi 130 \times 260$ mm and G10E inserts were used.

Discussion of Results

A total of 136 measurement samples (200 data point per sample) corresponding to variance levels of chip breaking and the cutting states (tool wear and chatter vibration) were collected. For each measurement sample, the set of data are divided into ten data subsets in which a feature vector of five features is extracted from data sets. This forms a total of 1360 feature vectors for the experiment. Among these 1360 feature vectors, representative training samples belonging to types of measurement samples were chosen and used to train the MLP. The remaining sets of feature vectors were used for testing and validation.

During the training stage, the target states of the output nodes were fixed at 0 for patterns of fresh tool, no chatter vibration and short comma chips; and 1 for patterns of worn tool, chatter vibration and long chips, respectively. The learning rate of the FDNN is 0.2 and its momentum coefficient is 0.4. The weights were initialized to uniformly distributed random values between -0.1 and 0.1 . During the testing stage for each sample, the final output vector of the IFDS is the mean value of the MLP output vectors for the ten subgroup of feature vectors.

Figure 4.23 shows the relationship between each features and the six typical patterns. It is clear that the system will fail when insufficient feature components are used for the identification process. Table 4.1 shows the training results with different number of features used as the input vector. The results in Table 4.1 show that a minimum of four feature components (P, F, V and D) must be used for successful identification. However, experiments show that the addition of the mean ration (R) enhances the convergent rate in MLP training. Hence, all the five feature were used as the input vector to the MLP.

This system has been tested on a completely new set of data and shown an approximately 95% success rate.

TABLE 4.1 Training Results with Different Number of Features

Number of Features	1	2	3	4	5
Features	P; F; V; D; R	PF; PV; PR; PD; FV; FR; FD; VR; VD; RD	PFV; PFR; PFD; PVR; PVD; PRD; FVR; FVD; FRD; VRD	PFVD	PFVDR
Training Results	Failure	Failure	Failure	Success	Success

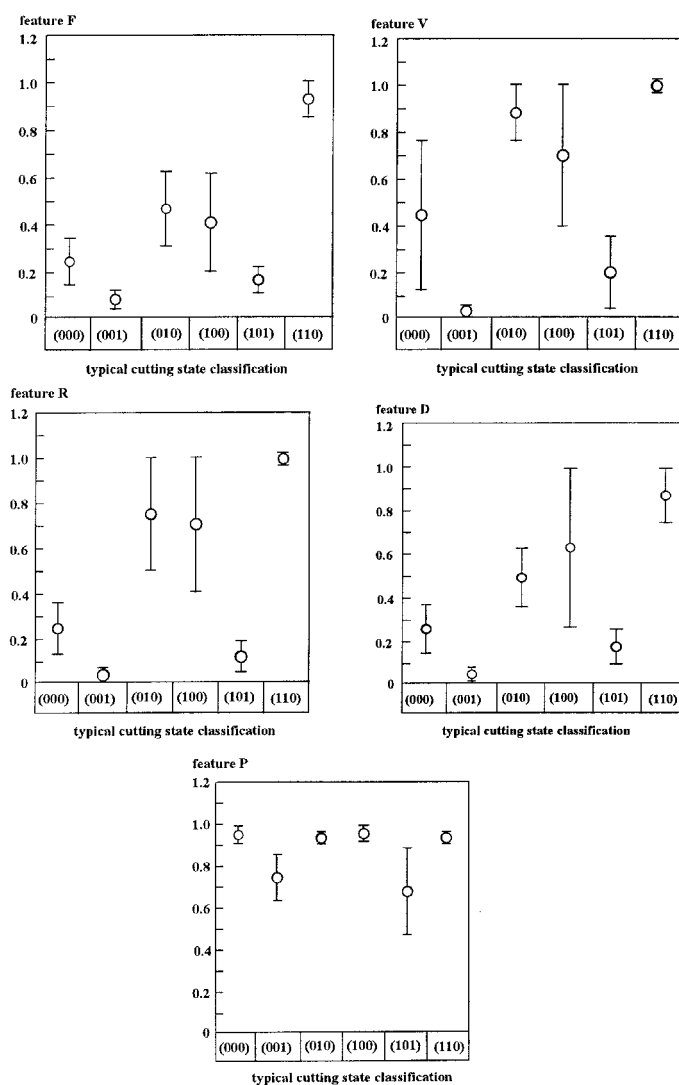


FIGURE 4.23 The relationship between features and typical patterns.

Single-ART2 Neural Network with Acoustic Emission Sensing

As shown in the section entitled “wavelet packet analysis of AE and force signals,” using the best basis wavelet packet transform and phase plane representation, a comprehensive set of time-frequency feature patterns corresponding to different tool conditions can be obtained from the AE signals generated during

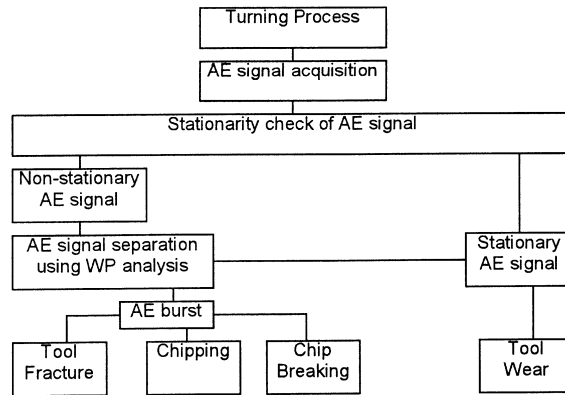


FIGURE 4.24 Multiple tool condition monitoring system.

turning. Based on the processed information, a new strategy for multiple tool condition identification has been established. An efficient separation scheme using the inverse wavelet transform for the separation of burst AE and continuous AE has also been developed. This method plays a key role in the progressive tool wear identification of a tool condition classification approach based on an ART2 neural network discussed in this section. Figure 4.24 shows a graphical presentation of this approach.

Given on AE signal from the turning process, a stationarity check is first conducted to judge the state of the signal (stationary or non-stationary). If the signal is stationary, then it will be used for tool wear monitoring. If the signal is non-stationary, it is first separated into burst and continuous signals by using a separation scheme. The burst signal is then used for the identification of transient tool conditions, such as the tool fracture, chipping, etc., and the continuous signal is used for tool wear monitoring.

The characteristics of this approach based on wavelet packet transform and ART2 neural network [29] are:

- Wavelet packet transform is employed for AE signal separation and burst AE signal feature extraction.
- For tool wear monitoring, a two-step feature extraction scheme is employed aiming at obtaining more meaningful feature vector related to tool flank wear process.
- An unsupervised ART2 neural network is employed for multi-category classification of tool conditions. Two separately trained ART2 networks are used for automatic classification of transient tool conditions (fracture, chipping, and chip breakage) and progressive tool flank wear.

Transient Tool Condition Identification

As presented earlier, the time-frequency feature of the tool fracture, chipping, and chip breakage each produces a set of characteristics components in a certain frequency range at the location of the burst. Tool fracture produces the largest magnitude in almost the entire frequency range. Chip breakage produces the smallest magnitude in a very narrow frequency range, and those of chipping have characteristics that are between the other two states. Based on this understanding, a transient tool condition identification approach is proposed as shown in Fig. 4.25.

A feature vector corresponding to frequency band value of the wavelet packet coefficients in the phase plane at the burst location is extracted. In the phase plane, the entire frequency range is divided into 16 equal zones (each zone equivalent to 62.5 kHz), and two time indices t_s and t_e which indicates the start and end point of the burst signal are determined from the separated AE burst signal. At each frequency division, all wavelet packet coefficients within the time interval t_s and t_e are summarized. Thus, a feature vector F_i

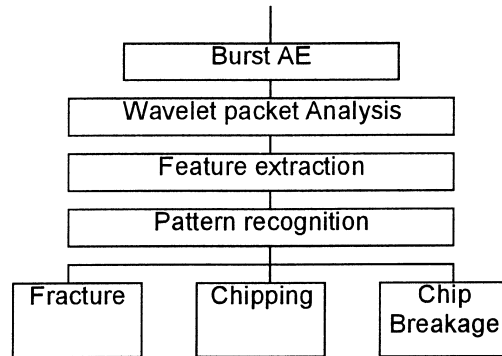


FIGURE 4.25 Transient tool condition identification approach.

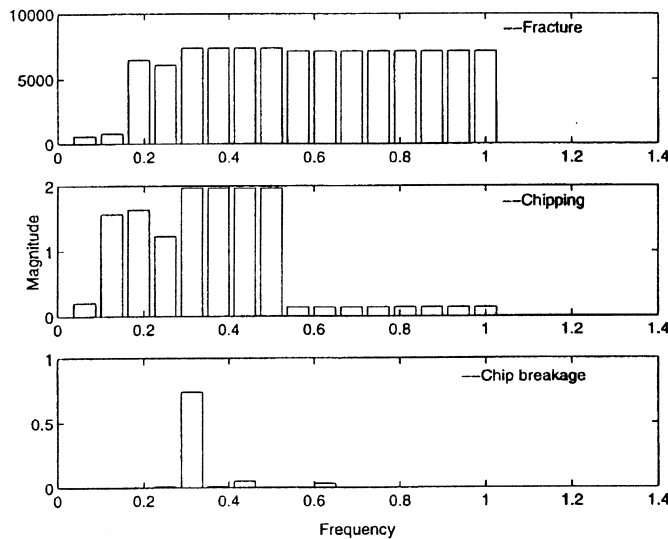


FIGURE 4.26 Feature patterns of different AE burst signals.

containing 16 elements is derived. Figure 4.26 shows the typical feature pattern of fracture, chipping, and chip breakage. The characteristics of the above three typical feature vectors can be summarized from Fig. 4.26:

- For tool fracture and chipping, the last 8 elements of the feature vector (corresponding to 0.5–1 MHz) are generally of the same value, while the first 8 (corresponding to 0–0.5 MHz) have varied values. The difference lies in that the sum of the last 8 elements is no less than the first 8 for tool fracture, while, for chipping, the case is just opposite.
- For chip breakage, there only exists a dominant element somewhere between 0–0.5 MHz while the remaining ones are nearly zero.

From the above feature extraction process, the advantage of wavelet packet analysis over FET is evident. Only with this efficient time frequency analysis can this kind of feature extraction in certain time location be made possible.

An ART2 network with 16-element input is trained to classify these three transient tool conditions. A total of 20 train patterns is employed to train the ART2 network (including 3 patterns from the fracture, 2 from the chipping, and 15 from chip breakage). Then the trained ART2 network is used for the transient tool condition

identification from sampled data of AE signal. Fracture, chipping, and chip breakage incidents that have occurred during the capture of the sampled AE signal during the machining test have been successfully identified.

Tool Wear Monitoring

As mentioned previously, the acoustic emission signal from the flank wear process originates mainly from rubbing between the workpiece and tool flank face. If this rubbing process is even and continuous in machining, we expect a consistent increase in the continuous AE signal. However, in practical machining, the tool flank wear is not a uniform process so that fluctuation in the AE signal is normally encountered. Nevertheless, the energy of the AE signal pertinent to the tool flank wear is generally related to severity of rubbing between the workpiece and tool flank. The experimental signal shows that although there is much variation in the AE signal energy during flank wear, there exists a progressive increase in the AE energy in a certain frequency range (300 kHz–600kHz) with tool flank wear, as shown in Fig. 4.27.

In view of the above reason, a special feature analysis approach has been developed. As shown in Fig. 4.28, this approach essentially applies a two-step feature extraction (the primary and secondary feature extraction) strategy, coupled with an ART2 neural network for condition identification. Figure 4.29 shows a sample of primary and secondary AE features. A feature vector F_w is derived from primary and secondary AE features and used as input to ART2 neural network. Preliminary results show that for un-grooved tool inserts, good classification results have been directly achieved. For grooved tool inserts, there is significant interference from chip breakage. As a result, the burst signal separation scheme has to be applied before a correct identification result can be obtained. Table 4.2 shows the results of the ART2 classification results of experimental test data obtained for different machining conditions involving coated and uncoated tools and different workpiece materials.

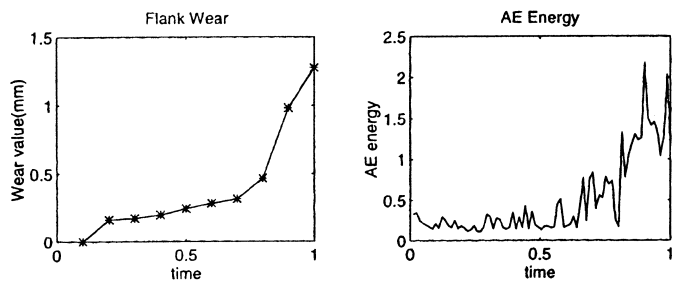


FIGURE 4.27 AE energy characteristics with tool flank wear.

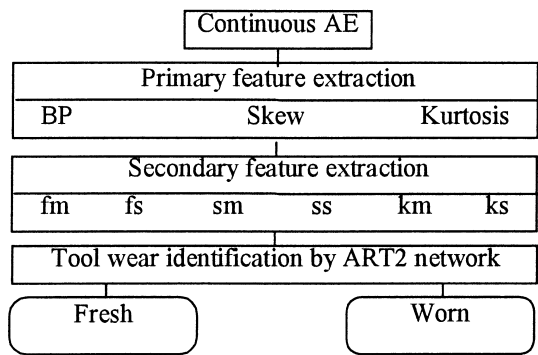


FIGURE 4.28 Tool wear identification from the continous AE signal.

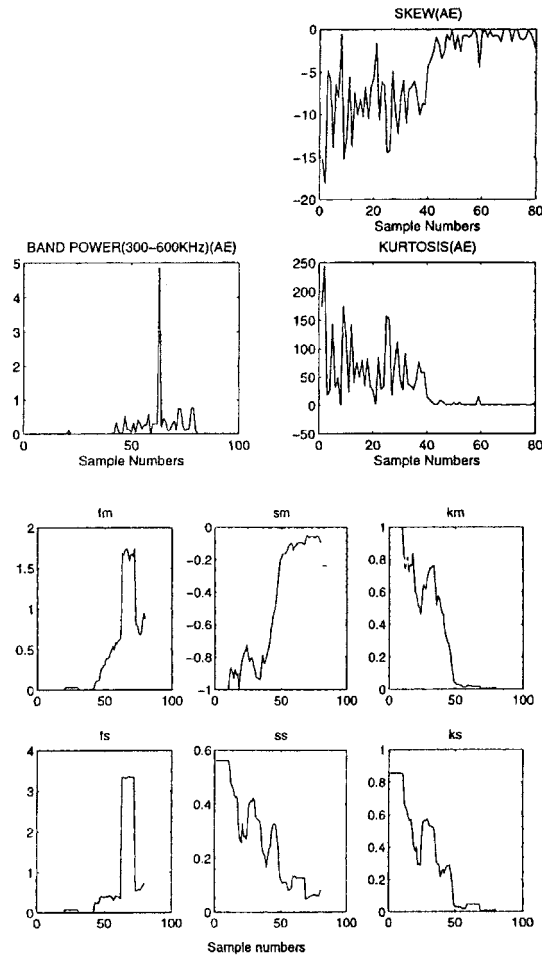


FIGURE 4.29 Primary and secondary AE features for tool wear identification.

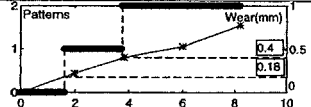

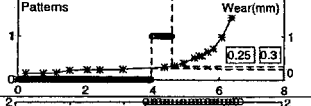
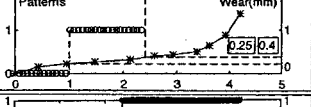
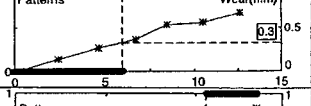
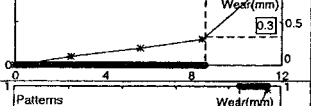
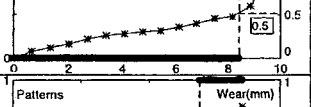
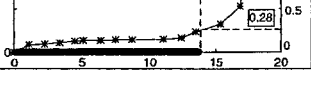
Single-ART2 Neural Network with Acoustic Emission and Force Sensing

Based on acoustic emission and force sensing, an intelligent sensor system [Fig. 4.30](#) has been developed to integrate multiple sensing, advanced feature extraction, and information fusion methodology [28]. Such a system employs

- More than one sensor to extend the effective range of the sensing system.
- Signal processing techniques to extract the compact feature vector sensitive to the monitored tool condition.
- Information fusion methodology to make correct decisions about the condition of cutting tool.

A two-step feature extraction strategy is proposed. Spectral, statistical, and dynamic analysis have been used to determine primary features from the sensor signals. In the primary feature determination, three features from the acoustic emission, the frequency band power (300 kHz–600 kHz) by FFT and the skew and the kurtosis by statistics, are obtained. One prominent feature from the tangential force signal, the natural frequency component resulting from the tool overhang [20], is also derived. [Figure 4.31](#) shows an example of the behaviour of these features with the tool flank wear. It can be seen that the frequency band power exhibits increasing activities while the skew and the kurtosis have decreasing activities as the

TABLE 4.2 Identification Results by ART2 Network for Different Cutting Conditions

Case	Speed (m/min)	Feed (mm/r)	DOC (mm)	Work	Tool	Identification Result
TEST 1	170	0.2	2.0	A	X	
TEST 2	170	0.25	2.0	A	X	
TEST 3	180	0.3	2.0	B	X	
TEST 4	180	0.4	2.0	B	X	
TEST 5	170	0.2	2.0	A	Y	
TEST 6	170	0.3	2.0	A	Y	
TEST 7	230	0.3	2.0	B	Y	
TEST 8	200	0.3	2.0	A	Z	

DOC-Depth of cut; A-ASSAB760; B-ASSAB705.

X-uncoated ungrooved; Y-coated ungrooved; Z-coated grooved

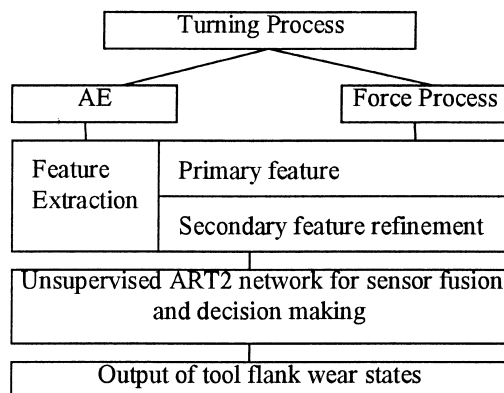


FIGURE 4.30 Intelligent sensor system for tool wear monitoring.

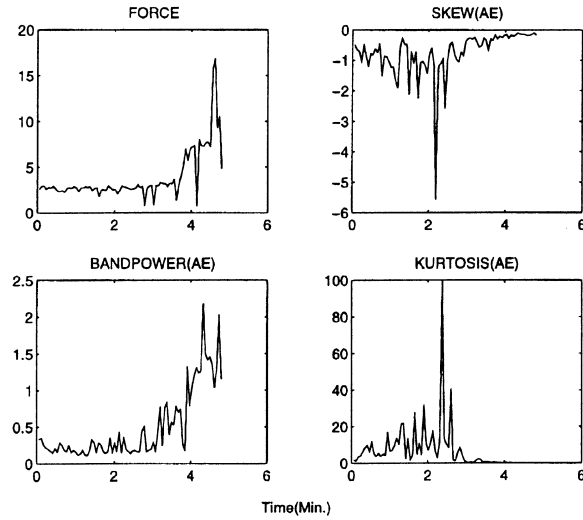


FIGURE 4.31 Primary features.

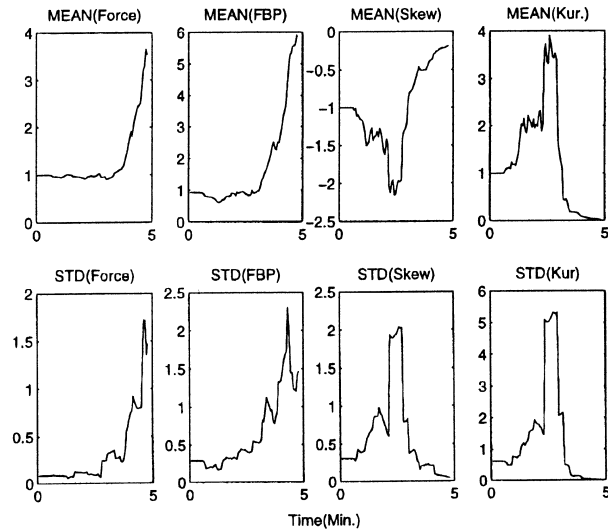


FIGURE 4.32 Secondary features.

flank wear progresses. As the values of the skew and the kurtosis are negative and positive, respectively, they provide complementary functions that have the effect of enhancing feature contrast. The tangential force components at the natural frequency of the tool overhang display an accelerated increase trend with respect to the tool wear before falling rapidly preceding the onset of tool failure. Another important characteristics of these four features, as experimental results show, is that they are nearly independent of the cutting conditions. Although changes in the cutting conditions may expedite or impede the tool wear process, the nature of the aforementioned features generally does not change.

Although the trends of the four primary features correlate well with the tool flank wear, they cannot be reliably used, either as an individual monitoring index or collectively as a feature vector for the decision making process. This is because of the severe variation of the features during the tool wear process (as

shown in Fig. 4.31). Experimental evidence shows that even for the least varied force feature, there are still occasional fluctuations along the process of the tool flank wear, let alone the other three greatly fluctuated AE features. For the purpose of getting more meaningful features, further refinement of the primary features is necessary for reliable tool flank wear identification.

Therefore, a secondary feature refinement is further applied to the primary features in order to obtain more correlated feature vector to the tool flank wear process. For each of the primary features, two refined features are extracted, namely the mean and the standard deviation within a moving window (Fig. 4.32). For a feature series (it can be frequency band power, skew, kurtosis, or tangential force component at the natural frequency of tool overhang) $P(n)$ ($n = 1, 2, \dots, N$), where N is the number of data sample, the secondary features in the moving window ($n, n + l - 1$) can be represented as follows (l is the window size):

1. Mean value

$$\text{mean} = \frac{1}{l} \sum_{i=n}^{n+l-1} P(i) \quad i = n, \dots, n + l - 1 \quad (4.38)$$

2. Standard deviation

$$\text{std} = \sqrt{\frac{1}{l-1} \sum_{i=n}^{n+l-1} (P(i) - \text{mean})^2} \quad i = n, \dots, n + l - 1 \quad (4.39)$$

From the above process, a feature vector \mathbf{F} is obtained:

$$\mathbf{F} = \{fm, fs, pm, ps, sm, ss, km, ks\}$$

- fm, pm, sm, km correspond to the mean values of the four primary features (the tangential force component, the frequency band power, the skew, and the kurtosis);
- fs, ps, ss, ks are the standard deviation values of the four primary features.

Figure 4.33 shows a typical representation of the above feature vector. It can be seen from Fig. 4.33 that the refined features are more meaningful to the tool flank wear than the primary features. The general feature vector patterns which reflect the fresh and worn states are given in Fig. 4.34. It can be

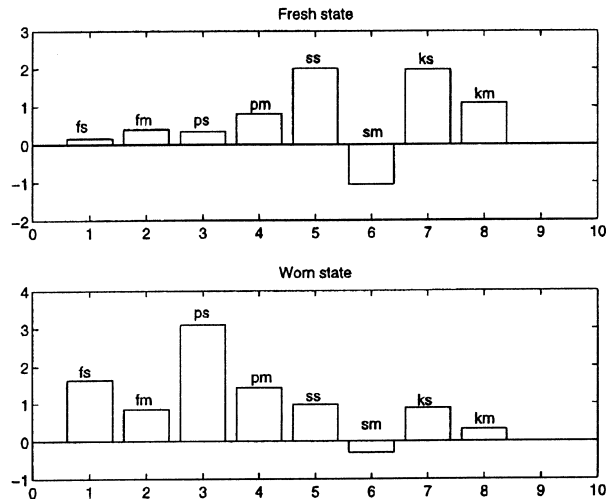


FIGURE 4.33 Pattern of feature vectors for fresh and worn tools.

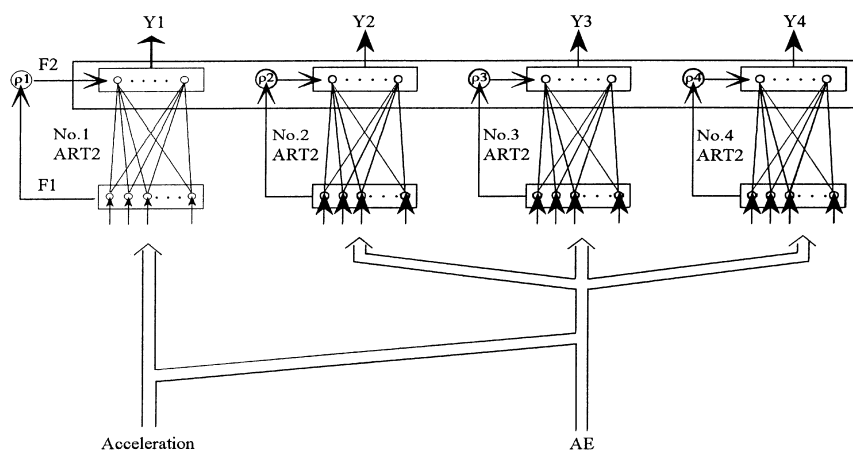


FIGURE 4.34 Parallel multi-ART2 neural network for tool condition monitoring.

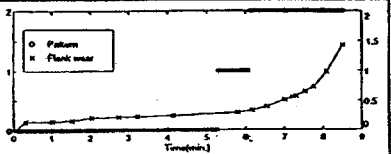
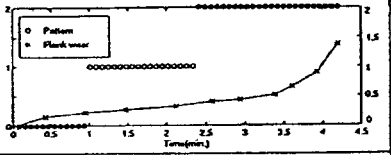
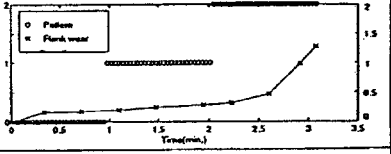
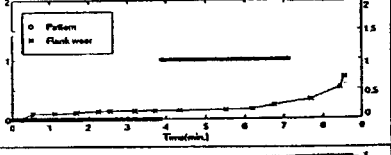
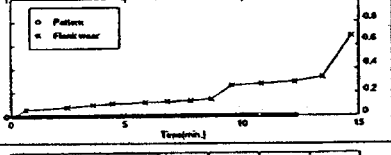
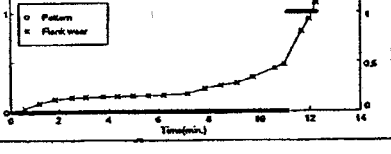
seen from Fig. 4.34 that the fresh state feature pattern is quite different from that of the worn state. In the fresh state, the standard deviation and mean of the force feature and frequency band power are smaller than those of the skew and kurtosis. While in the worn state, the case is just the opposite. One thing worth noting is that the standard deviations from the four primary features are more prominent than their corresponding mean values. Therefore, the feature vector thus obtained can distinctively reflect the tool wear state. For a neural network, the more meaningful features we have, the more reliable and faster the classification can be achieved.

Note that the window size l is determined experimentally and the first l samples are used to calculate the first feature vector. Therefore, it is assumed that tool is fresh within the first l sample. This assumption is practically acceptable because the time used to obtain the first l sample is very short. The flank wear of the tool is very unlikely to occur in such a short time.

After obtaining a meaningful feature vector F , it is then used as input to an ART2 neural network for the fusion of AE and force information and decision-making of the tool flank wear state. As shown in Table 4.3, the experimental results confirm that the developed intelligent sensor system can be reliably used to recognize the tool flank wear state over a range of cutting conditions. The cutting conditions of the turning tests used to verify the proposed system include various types of inserts and workpiece materials under different machining conditions. The ART2 recognition results are listed in the last column of Table 4.3 (the X-axis represent cutting time in minutes). Table 4.4 shows the flank wear values at identified intermediate tool wear states (uncoated tool only) and worn state of tool (both the uncoated tool and coated tool). From the results obtained, it can be seen that:

1. For both the uncoated and coated tools, the fresh and worn states of the turning tool are all successfully recognized. The worn tools are identified around 0.3 mm of the measured flank wear value which is very close to those used in industry.
2. For uncoated tools inserts, the tool flank wear states have been classified into three categories: the fresh state (0), intermediate wear state (1) and worn state (2). The worn state occurs around the tool flank wear threshold value of 0.3 mm. The existence of the intermediate tool wear state may be caused by the fuzziness of the feature information between the fresh state and the worn state of the tool, since there is no clear cut information between the fresh and worn state in the wearing process of the uncoated insert tool. The intermediate state (1) can be used as a warning index in tool wear monitoring of the uncoated tools.
3. For coated tool inserts, the tool flank wear states have been correctly classified into two categories: the fresh (0) and the worn (1) states. Unlike uncoated tools, no intermediate flank wear state has been produced. This is because the wear process of the coated tool inserts is quite different. The

TABLE 4.3 Cutting Conditions and Identification Results of Test Data

Process	Speed	Feed	Depth of Cut	Workpiece	Tool	Identification Result
TEST 1	180	0.3	2.0	ASSAB705	A	
TEST 2	180	0.4	2.0	ASSAB705	A	
TEST 3	200	0.4	2.0	ASSAB705	A	
TEST 4	200	0.3	2.5	ASSAB760	A	
TEST 5	180	0.4	2.0	ASSAB760	B	
TEST 6	230	0.3	2.5	ASSAB760	B	

Note: A-SNMN-120408; B-SNMG120408

TABLE 4.4 Tool Flank Wear Value Identified by ART2

	Intermediate Wear States (mm)	Worn States (mm)
TEST1	0.30	0.45
TEST2	0.25	0.40
TEST3	0.20	0.35
TEST4	0.15	0.30
TEST5	—	0.30
TEST6	—	0.48

tool flank wear increases very slowly until it reaches a point where the coated materials near the tool edge are worn off. Then the tool flank wear increases significantly because in this case the tool insert is similar to an uncoated one. So, distinct information between the fresh and worn states has been produced.

4. The proposed approach is less dependent on cutting conditions, since the features selected are only those relevant to the nature of the tool wear. The change of cutting conditions may expedite or slow down the tool flank wear. However, when tool wear occurs the characteristic of the primary features are fixed. Therefore, cutting conditions have little influence on this approach, as the above results indicate (Table 4.5).

Multi-ART2 Neural Network with Force and Vibration Sensing

The value of ρ determines the granularity in which the input patterns are classified by the network. For a given set of patterns with different classes, a larger value of ρ may result in finer discrimination between patterns of the same class and a smaller value of ρ may lead to the merging of some of the classes into a single category. In a complex situation, the input patterns can be presented in different classes (such as tool failure and chatter), with each class having several cases (e.g., the various types of tool condition, the different frequencies or different magnitudes of chatter as well as the different cutting conditions). Due to the varied boundaries and dispersions of the various classes of input patterns, it is difficult to choose an accurate value of ρ for a single-ART2 network to classify all the patterns correctly, which is even less possible in a complex situation. In view of the varying cases, the network may not effectively represent each class of input patterns with only one cell in the $F2$ layer of the ART2 network. It then becomes necessary to encode or enroll more cells in the $F2$ and $F1$ layers in order to handle the different classes as well as their cases in complex patterns. The corresponding increase in the computation time for the expanded single-ART2 network then makes it less feasible for on-line use.

Due to the aforementioned limitations of the single-ART2 network, a novel parallel multi-ART2 neural network (as shown in Fig. 4.34), has also been developed to recognize tool conditions and machining chatter in turning operations [22]. With the multi-ART2, each subnetwork is individually designed with its own network parameters and is fed with the appropriate feature vector(s) relevant to a targeted machining condition (that is, only one of the input categories) so that the setting of the threshold is easier, compared with that for the single-ART2 network. In Figure 4.34, the input vector I consists of multiple features extracted from both vibration and AE signals. The input vector γ , also with multiple elements, contains features based on the vibration signals only. The neural network consists of four parallel ART2 subnetworks with the binary outputs denoted by $Y1$, $Y2$, $Y3$, and $Y4$, respectively. These sub networks are employed to identify the following four categories of tool states: machining chatter, tool failure, simultaneous severe tool wear, and chatter. According to the simultaneous outputs of the multi-ART2 network (i.e., $Y1$, $Y2$, $Y3$ and $Y4$), the patterns of the four machining states can be indicated by the following:

1. *chatter*: if $Y1 = 1$, $Y2 = 0$, $Y3 = 0$, $Y4 = 0$.
2. *tool failure (severe wear and breakage)*: if $Y1 = 0$, $Y2 = 1$, $Y3 = 0$, $Y4 = 0$.
3. *simultaneous chatter and severe tool wear*: if $Y1 = 0$, $Y2 = 0$, $Y3 = 1$, $Y4 = 0$.
4. *normal*: if $Y1 = 0$, $Y2 = 0$, $Y3 = 0$, $Y4 = 1$.

During the learning process, each of the four ART2 sub-networks is trained to represent one category of input patterns through arranged sample learning by grouping the test samples according to the respective categories. The $F2$ cells within an ART2 subnetwork characterize the typical cases of each category. Hence, if needed, more categories can be achieved by adding more ART2 subnetworks while more cases of each category can be handled by enrolling more $F2$ cells in the respective category representation sub-layer. To minimize the possibility that an input pattern does not match the existing patterns, the multi-ART2 network can be retrained to represent a new pattern according to the following two cases:

1. The input pattern does not belong to any of the existing categories and needs to be oriented. In this case, an additional category can be incorporated by adding a new ART2 subnetwork with the new input data, without any changes in the other ART2 subnetworks.

2. The input pattern is considered to be a specific case of one of the existing categories. In this case, the new pattern can be merged into the weighted connections with a newly added cell by the new input data, with no changes in the other cells.

Furthermore, four finer vigilance thresholds ρ_1, ρ_2, ρ_3 , and ρ_4 for the four ART2 subnetworks are employed in place of a single vigilance threshold ρ which is set to classify all categories for single-ART2 network. By appropriate setting of their threshold values, ρ_1, ρ_2, ρ_3 , and ρ_4 are directed at the respective four categories of input patterns (a)–(d) so that four specialized classifiers are formed on their feedforward/ feedback weight connections. The values of the other parameters, such as a, b, c, d , and θ , are also chosen for every F1 cell of each ART2 sub-network so as to enhance the contrast of the information that is correlated with the corresponding category.

When a new input pattern is presented for classification, the multi-ART2 network carries out parallel processing and fast searching among the four ART2 subnetworks. The operation of the parallel multi-ART2 neural network

1. *Competition within each subnetwork*: each of the four ART2 subnetworks operates for the single-ART2 network. Hence, only one cell in an F2 sub-layer of a subnetwork is activated, which represents one of the four patterns (a)–(d);
2. *Competition among the subnetworks*: when two or more F2 cells from different subnetworks happen to be activated after Step 1, the F2 cell with the maximum degree of match is the one selected to represent the corresponding one of the four patterns (a)–(d) as the final result classified by the multi-ART2 network.

Feature Information

With reference the coherence function for the acceleration signals:

$$\gamma(f) = |G_{xz}(f)| / \sqrt{G_x(f) \cdot G_z(f)} \quad (4.40)$$

If the concerned frequency range is divided into $N1$ appropriate bands, $\Delta f_i = f_i - f_{i-1} (i = 1, 2, \dots, N1)$, the maximum value of the coherence function in each band is

$$\gamma_i = \max[\gamma(f), f \in (f_{i-1}, f_i)] (i = 1, 2, \dots, N1) \quad (4.44)$$

An $N1$ -feature vector is thereby extracted from the acceleration signals as

$$\gamma = (\gamma_1, \gamma_2, \dots, \gamma_{N1}) \quad (4.45)$$

Ten frequency bands are used, (i.e., $N1 = 10$). Five are distributed over the chatter range from 0 to 500 Hz with $\Delta f_i = 100$ Hz. One is from 500 Hz to 3 kHz. The remaining four ranges are around the first natural frequency of the cantilever shank, from 3 kHz to 5 kHz with $\Delta f_i = 500$ Hz, for monitoring the tool condition. Figure 4.35 presents an example of the increase in the maximum frequency-band value pertaining to the coherence function of the two acceleration signals corresponding to the onset of chatter. Different distribution characteristics have been observed for the severe tool wear as well as the simultaneous onset of chatter and severe tool wear [23].

It has also been found from a careful study of the power spectrum that the AE signal emitted during the turning operation is more sensitive to chipping, breakage and severe wear of the cutting tool in the characteristic frequency range of 200 kHz–500 kHz. Hence, the spectrum in this band embodies the input patterns corresponding to the conditions of the tool. Meanwhile, the AE spectrum in the same frequency band does not exhibit a consistent relation to the occurrence of chatter. If the power spectrum

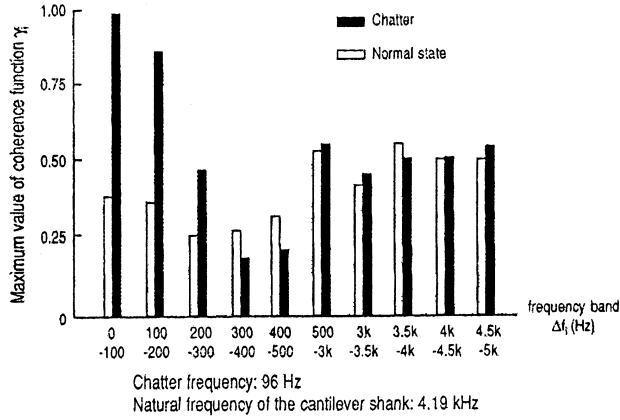


FIGURE 4.35 Frequency band distribution of maximum coherence function (Chatter) Depth of Cut - 2 mm; Feed-45 mm/min; Speed-72 m/min; Workpiece-S45C; Tool-P10.

of AE signal is $G(f)$ and the chosen frequency range is separated into $N2$ equivalent frequency bands, $\Delta f_j = f_j - f_{j-1}$ ($j = 1, 2, \dots, N2$), then the relative weighted power ratio in each band is

$$\delta_j = \int_{f_{j-1}}^{f_j} \lambda G(f) df / \int_{f_{j-1}}^{f_j} G(f) df \quad (j = 1, 2, \dots, N2) \quad (4.41)$$

with

$$\lambda = \begin{cases} 1.0, & G(f) > T_2 \\ 0.5, & T_1 < G(f) \leq T_2 \\ 0.25, & G(f) \leq T_1 \end{cases} \quad (4.42)$$

λ is the weight, T_2 and T_1 ($T_2 > T_1$) are two threshold constants which can be estimated during sample runs. T_2 and T_1 indicate the lower levels of the AE spectrum in normal and abnormal tool conditions, respectively. Thus, the $N2$ -feature vector extracted from the AE signal is

$$\delta = (\delta_1, \delta_2, \dots, \delta_{N2}) \quad (4.43)$$

$N2$ is set to 10 for the frequency range of 200 kHz–500 kHz, so that $\Delta f_j = 30$ kHz. Based on the aforementioned vectors of γ and δ , the N -input feature vector I from both acceleration and AE signals is assigned as follows:

$$\begin{aligned} I &= (I_1, I_2, \dots, I_N) \\ &= (\gamma, \delta) \\ &= (\gamma_1, \gamma_2, \dots, \gamma_{N1}; \delta_1, \delta_2, \dots, \delta_{N2}) \end{aligned}$$

where $N = N1 + N2$ is chosen to be 20 in this study.

With the above input feature information from the maximum coherence function of two acceleration signals and the relative weighted frequency-band power ratio of an acoustic emission signal, the multi-ART2 network has been found to identify various tool failure and chatter states in turning operations with more than 93% success rate over a wide range of cutting conditions, compared to that of 71.4%–89.3% obtainable with the single-ART2 neural network (Tables 4.6 and 4.7).

TABLE 4.5 Machining Conditions

	Cutting Speed (mm/min)	Feed Rate (mm/min)	Depth of Cut (mm)	Workpiece Material	Tool Type (Insert)
C1	134	50	1.7	S45C	P10
C2	134	70	2.5		
C3	95	70	1.7		
C4	95	50	2.5		
C5	95	100	1.7		
C6	134	70	2.0		
C7	95	70	2.5		
C8	64	50	3.1		
D1	85	27	1.0	Aluminum alloy	K6
D2	85	35	2.0		
D3	70	35	1.0		
D4	70	27	2.0		
D5	70	50	1.0		
D6	85	27	1.5		
D7	70	35	1.5		
D8	62	35	2.0		

TABLE 4.6 Experimental Results based on Multi-ART2 Network

Pattern Categories	(Sample Size/Machining Conditions*)		Recognized Rate (%)
	Training	Classification	
Chatter	2/C2, C4	16/C2, C4, C7, C8	100
Tool Breakage	4/D1-D4	8/D1-D8	100
Severe Tool Wear	4/C1-C4	8/C1-C8	87.5
Chatter & Tool Wear	2/C2, C4	8/C2, C4, C7, C8	87.5
Normal State	8/C1-C4, D1-D4	16/C1-C8, D1-D8	100
Total (Σ)	20/C1-C4, D1-D4	56/C1-C8, D1-D8	96.4

TABLE 4.7 Experimental Results Based on Single-ART2 Network

Pattern Categories	(Sample Size/Machining Conditions*)		Recognized Rate (%)
	Training	Classification	
Chatter	2/C2, C4	16/C2, C4, C7, C8	93.8
Tool Breakage	4/D1-D4	8/D1-D8	87.5
Severe Tool Wear	4/C1-C4	8/C1-C8	62.5
Chatter & Tool Wear	2/C2, C4	8/C2, C4, C7, C8	62.5
Normal State	8/C1-C4, D1-D4	16/C1-C8, D1-D8	81.3
Total (Σ)	20/C1-C4, D1-D4	56/C1-C8, D1-D8	80.4

(*Refer to Table 4.1 for specific machining conditions.)

4.7 Conclusions

Tool conditions monitoring is an important component in modern manufacturing environment, in particular, the machining processes. The chapter presented some simple neural networks architectures for the application of tool condition monitoring. The approaches presented in the examples illustrates the use of neural network to identify the tool condition from the input feature vectors. The chapter also illustrates the important of sufficiently rich feature data to be extracted for successful identification.

References

1. Carpenter, G.A. and Grossberg, S., "ART2: Self-organization of stable category recognition codes for analog input patterns," *Applied Optics*, Vol. 26, No. 23, (1987) pp. 4919–4930.
2. Carpenter, G.A. and Grossberg, S., "A massively parallel architecture for a self-organizing neural recognition machine," *Computer Vision, Graphics, and Image Processing*, 37, 54–115, 1987.
3. Chryssolouris, G. and Domroese, M., "Sensor synthesis for control of manufacturing process," *J. Engineering for Industry, Trans. ASME*, May 1992, pp. 158–174.
4. Cohen, L., "Time-frequency distributions—A review," *Proceedings of the IEEE*, Vol. 77, 1989, pp. 941–981.
5. Coifman, R., Meyers, Y., Quake, S., and Wickerhauser, M.V., "Signal processing and compression with wavelet packets," *Progress in Wavelet Analysis and Applications*, Meyer and Roques Ed., Editions Frontieres, Toulouse, France, 1992, pp. 77–93.
6. Colgan, J., Chin, H., Danai, K., and Hayashi, S. R., "On-line tool breakage detection in turning: A multi-sensor method," *Trans. ASME Journal of Engrg. for Industry*, Vol. 116, February 1994, pp. 117–123.
7. Dornfeld, D. A., "In-process recognition of cutting states," *Int. Journal JSME, Series C*, Vol. 37, No. 4, 1994, pp. 638–650.
8. Dornfeld, D. A., "Neural network sensor fusion for tool condition monitoring," *Annals of CIRP*, 1990, pp. 101–105.
9. Du, R., Elbestawi, M. A., and Yan, D., "Time-frequency distribution of acoustic emission for tool wear detection in turning," *Proc. of 4th World Conference on Acoustic Emission*, Boston, M.A., 1991, pp. 269–285.
10. Editors, "Future view: Tomorrow's manufacturing technologies," *Manufacturing Engineering*, January 1992, pp. 76–88.
11. Emel, E., and Kannatey-Asibu, E., "Tool failure monitoring in turning by pattern recognition analysis of AE signals," *Trans. ASME J. Eng. Industry*, Vol. 110, 1988, pp. 137–145.
12. Fausett, A., "Fundamentals of neural networks, architectures, algorithms, and applications," Prentice-Hall International Ltd., 1994.
13. Gui, J.S., "On-line tool failure monitoring of carbide inserts," B. Eng. Thesis, Department of Mechanical and Production Engineering, National University of Singapore, 1994/95.
14. Hecht-Nielsen, R., "Neurocomputing," Addison-Wesley Publishing Company, 1989.
15. Hecht-Nielsen, R., "Theory of backpropagation neural network," *Proc of the Int. Conf on Neural Networks*, I, 593–611, IEEE Press, New York, June 1989.
16. Hong, G.S., Rahman, M., and Zhou, Q., "Using neural network for tool condition monitoring based on wavelet decomposition," *International Journal of Machine Tools & Manufacture: Design, Research and Application*, 36, no. 5 (1996): 551–556.
17. Hong, G.S., Rahman, M., and Zhou, Q., "Tool condition monitoring using neural network." In *Proceedings of the 26th International Symposium on Industrial Robots*, pp. 455–460. Singapore: International Federation of Robotics, 4 October 1995.
18. Khanna, T., "Foundations of neural networks," Addison-Wesley Publishing Company 1990.
19. Kohonen, T., "Self-organization and associative memory" (3rd Edition), Berlin: Springer-Verlag.
20. Lee, K.S., and Gan, C.S., "On the correlation between dynamic cutting force and tool wear," *International Journal of Machine Tools & Manufacture: Design, Research and Application*, Vol. 29, pp. 295, 1989.
21. Lee, K.S., Lee, L.C., and Teo, C.S., "On-line tool wear monitoring using a PC," *Journal of Materials Processing Technology*, Vol. 29, 1992, pp. 3.
22. Li, X.Q., Wong, Y.S., and Nee, A.Y.C., "A comprehensive identification of tool failure and chatter using a multi-ART2 neural network," To be published in *ASME Journal of Manufacturing Science and Engineering*.

23. Li, X.Q., Wong, Y.S., and Nee, A.Y.C., "Tool wear and chatter detection using the coherence function of two crossed acceleration signals," *International Journal of Machine Tools & Manufacture: Design, Research and Application*, Vol 4, 1997, pp. 425–435.
24. Lim, C., "Tool wear in milling," B., Eng. Thesis, Department of Mechanical and Production Engineering, National University of Singapore, 1995/96.
25. Lim, Y.H.C., "Surface coating for cutting tools," Ph.D. Thesis, Department of Mechanical and Production Engineering, National University of Singapore, 1996/97.
26. Matsushima, K., Bertok, P., and Sata, T., "In-process detection of tool breakage by monitoring the spindle motor current of a machine tool," *ASME Winter Annual Meeting*, Phoenix, Arizona, Nov. 1982, pp. 145–153.
27. Nee, A.Y.C., Wong, Y.S., and Chan, K.Y., "Force pulsations in milling," *Society of Manufacturing Engineers, Technical Paper*, No. MRR78–09, U.S.A. (1978).
28. Niu, Y.M., Wong, Y.S. and Hong, G.S., "An intelligent sensor system approach to reliable tool flank wear recognition." To be published in *The International Journal of Advanced Manufacturing Technology*.
29. Niu, Y.M., Wong, Y.S., Hong, G.S., and Liu, T.I., "Neural-based multi-category classification of tool conditions using wavelet packets and arts network." To be published in *ASME Journal of Manufacturing Science and Engineering*.
30. Niu, Y.M., Wong, Y.S., and Hong, G.S., "A comprehensive review on tool condition monitoring techniques," TRME-002-CON96, Dept of Mech & Prod Eng, National University of Singapore, 1996.
31. Owen, J.V., "Feedback from the cutting edge," *Manufacturing Engineering*, January 1993, pp. 39–45.
32. Rahman, M., Hong, G.S., and Zhou, Q., "On-line cutting state recognition using neural network." *The International Journal of Advanced Manufacturing Technology* (October 1995): 87–92.
33. Rangwala, S., Dornfeld, D. A., "Sensor integration using neural network for intelligent tool condition monitoring," *J. Engineering for Industry, Trans. ASME*, Vol 112, Aug., 1990, pp. 219–228.
34. Tansel I. N., Wagiman, A., and Tziranis, A., "Recognition of chatter with neural networks," *International Journal of Machine Tools & Manufacture*, Vol. 31(4), pp. 539–552, (1991).
35. Tonshoff, H.K., "Development and trends in monitoring and control of machining processes," *Annals of CIRP*, 1988, pp. 611–622.
36. Trent, E.M., "Metal Cutting," 3rd Edition, Butterworth-Heinemann, Oxford, 1991.
37. Trent, E.M., "Tool wear and machinability," *Journal of the Institute of Production Engineers*, Vol. 38, pp. 105–130.
38. Venkatesh, V.C., and Sathithanandam, M., "A discussion on tool life criteria and total failure causes," *Annals of the CIRP* Vol. 29/1/1980, pp. 19–22.
39. Waschkies, E., Sklarczyk, C., and Hepp, K., "Tool wear monitoring in turning," *Trans. ASME Journal of Engrg. for Industry*, Vol. 116, Nov. 1994, pp. 521–524.
40. Wright, P.K. and Bagchi, A., "Wear mechanisms that dominate tool-life in machining," *Journal of Applied Metal Working*, Vol. 1, No. 4, pp. 15–23, 1981.
41. Wright, P.K. and Trent, E.M., "Metallurgical appraisal of wear mechanisms and processes on high-speed-steel cutting tools," *Metals Technology*, Vol. 1, pp. 12–23, 1974.
42. Chrysolouris, Domroese, M., and Beaulieu, P., "Sensor integration for tool wear estimation in machining," *Sensors and Controls for Manufacturing*, Vol. 33, *ASME Winter Annual Meeting*, pp. 115–123.
43. Zhou, Q., Hong, G.S., and Rahman, M., "New tool life criterion for tool condition monitoring using neural network." *Engineering Applications of Artificial Intelligence*, 8, no. 5 (1995): 579–588.
44. Zhou, Q., Hong, G. S., and Rahman, M., "A neural network approach for the on-line diagnosing of tool wear." In *Proceedings of the IEEE International Conference on Neural Network Applications to Signal Processing*, 17–20 August 1993, pp. 132–137. Singapore, 1993.
45. Zurada, J.M., "Introduction too artificial neural systems," Info access Distribution Pte Ltd., 1992.