On-line sensing of tool wear and breakage in machining has been a long standing goal of the manufacturing community. Wear and breakage detection systems are typically based on force, acoustic emission, current or temperature measurement. They are important for reliable unattended operation, and also for implementing optimization systems. This paper proposes a model-based approach to on-line tool wear and breakage detection under varying cutting conditions based on force measurement. The proposed model is used together with on-line parameter estimation to track flank wear during cutting. The proposed method is illustrated with a simulation study, the results of which confirm the feasibility of the model-based approach.

1. INTRODUCTION

A successful tool wear/breakage sensor has been a long standing goal of the manufacturing community [1-8]. In addition, a wear sensor has always been regarded as a necessary step toward the implementation of an adaptive control optimization system [9,10]. Wear and breakage detection systems are typically based on acoustic emission, motor current, temperature measurement, or cutting force.

Acoustic emission [11-16] and vibration [17] measurements are receiving a great deal of attention due to the non-obtrusive nature of the transducers, but important signal processing problems are still remain to be solved. Current monitoring is the simplest method for motor driven machine tools [18-19], but suffers from sensitivity and time lag problems. Temperature measurement is particularly important for wear [20-22], however, a practical on-line transducer has yet to be developed [23]. Among the indirect methods those based on force and torque sensing [25-28] are perhaps the most developed. Such commercial units have begun to appear on the market in recent years, and have been used with some success in certain limited production situations. With this recent introduction of the first commercial tool wear/breakage sensors [25-28] it is perhaps worthwhile to evaluate these sensors and to reassess their economic advantages [29-30]. For this purpose the machining operation has been divided into two direct and indirect cuts. In roughing operations one can consider two distinct types of production:

1. short production cycles (e.g., t < 5 min)
2. long production cycles (e.g., t > 10 min)

The first situation is frequently encountered in high volume production, whereas the second situation occurs in the production of large complex workpieces, or in the machining of hard materials such as titanium.

The importance of wear sensing, breakage detecting, and adaptive control is summarized according to the production mode in Table 1. As can be seen tool breakage detection is economically important in all machining operations except finishing operations. Wear sensing is a predictor of tool failure due to excessive wear. In long production cycle manufacturing operations, tool wear detection becomes important for surface roughness considerations. Tool wear sensing is also important in finishing cuts due to the effect on product quality.

Commericably available systems for wear and breakage detection are typically based on force or power limits. When the measured force falls outside these predetermined fixed limits the tool is assumed to have failed due to wear or breakage. The disadvantage of the fixed force limit method is that the cutting conditions must remain nearly identical throughout the entire cutting situation, and therefore this is applicable only in very simple cases.

To extend the force limit approach to various cutting conditions, force signature methods typically use "learning" or "averaging" strategies. These methods, however, have the following disadvantages:

1. many parts must be cut to allow the system to learn the force signature
2. it is not effective in long production cycle situations
3. it is not applicable in adaptive control systems [29]
4. the method is not effective in long production cycle situations
5. the method is not applicable in adaptive control systems [29]
6. the method is not effective in long production cycle situations
7. the method is not applicable in adaptive control systems [29]
8. the method is not effective in long production cycle situations
9. the method is not applicable in adaptive control systems [29]
10. the method is not effective in long production cycle situations

Table 1 also summarizes the importance of various types of adaptive control systems. Adaptive control is not particularly significant in high volume/short production cycle manufacturing. Geometric adaptive control (GAC) is important in finishing operations [31,32]. In manufacturing operations with a long production cycle, both adaptive control optimization (ACO) and adaptive control with constraint (AC) type systems are important.

The tool wear sensor is also a necessary component in ACO systems for turning and milling [8-10,35]. There is a fundamental difference, however, between using a tool wear monitoring system (which operates in open-loop) and applying some sensor in a closed-loop adaptive control system. ACO systems (e.g., the mid 1960's Bendix systems [9]) are based on maintaining the optimal cutting conditions by incrementing the feed and/or cutting speed in small steps. However, in practice an indirect measurement of the tool wear is used, and feeding back the sensor signals in order to close the ACO loop is not straightforward task.

For example, assume that the tool wear is monitored through measurements of the cutting force. Typically, increased tool wear causes a consequent increase in the cutting force. This is the direct effect. An increase in the feed shortens the tool life, namely increases the tool wear rate. As a consequence the force increases and the tool life becomes shorter. This is the indirect effect.

An ACO system should use only the latter effect. However, since both mechanisms affect the cutting force similarly, the problem of isolating the second effect from the first is not trivial. One might say that the direct (ACO) and adaptive control (AC) systems (theoretically after one revolution of the spindle) are the same, while the other mechanism affects the force after "some time". However, since the ACO convergence strategy is based upon incremental feed variations, and the signal-to-noise ratio in force measurements is low, it is difficult in practice to isolate the two phenomena by using simple electronics.

A solution to this problem might be obtained by programming a mathematical model of the cutting process and updating it in real time to obtain an accurate estimation of the two effects. The separation of the two effects has another significant outcome: it permits the separate detection of tool breakage and excessive tool wear. Tool breakage is detected by a sudden change in the force due to the first (direct) effect, whereas wear causes a steady change in the force due to the second (indirect) effect. By contrast to traditional force signature approaches, this proposed method does not require the cutting of many parts for the learning mode, and is also effective in long production cycles.
2. THE FORCE EQUATION

Before proceeding to the model-based approach of this paper, we would like to further clarify the problem. The relation between each component of the cutting force \( F \) and the flank wear is highlighted in the previous section by using a simplified mathematical analysis. The relation between each component of the cutting force \( F \) and the flank wear \( W \) is approximated by \[ F = F_0 + aC_W \] where \( F_0 \) is the initial cutting force (with a sharp tool), \( a \) is a constant for a certain tool and workpiece material, and \( C_W \) is a fixed set of cutting conditions (feed and speed). Equation (1) has been experimentally verified (see Fig. 1) for cases in which flank wear has the dominant effect on tool life. The initial cutting force is given by \[ F_0 = a\alpha_0^2, \] where \( a \) and \( \alpha \) are constants depending on the tool and workpiece material; typically \( 0.6 < \alpha < 1 \).

If the wear is assumed to be in its linear progression zone (see Fig. 2) it obeys the equation \[ W = W_0 - \frac{F(k) - F_0}{\alpha(t)} t + Y(t) \] where \( T \) is the tool life, \( W_0 \) is the corresponding wear at which the tool is replaced, and \( Y(t) \) is a slowly time varying function used to account for the non-constant slope of the flank wear curve. Combining Eq. (1) and (3) yields the following force equation:

\[ F = F_0 + \alpha C_W t \] (4)

where

\[ F_0 = Y(t)\alpha C_W/W_0 \] (5)

An acceptable model for \( T \) is the extended Taylor's tool life equation

\[ C_W \beta a^T = 1 \] (6)

substituting \( T \) from Eq. (6) into Eq. (5) yields

\[ F_1 = Y(t)\alpha a C_W \] (7)

where \( \beta = CC_W \), \( \alpha \) and is assumed to be constant in the linear progression zone of the wear curve. The signal that an indirect tool wear detector transmits is proportional to \( F \) in Eq. (4). However, only the term \( F_1 \) is proportional to the wear and therefore it should be separated from \( F \) (and be inserted as feedback to the ACO control loop). Thus, a wear sensor based on force measurements must resolve the following problems:

1. The real-time separation of the term \( F_1 \) in Eq. (4), particularly under continuously changing cutting conditions.

2. Since the coefficients \( F_1 \) and \( F_2 \) depend on the depth of cut, feed, and speed, any change in these process variables might be interpreted by the system as a change in \( W \).

3. The coefficients \( F_1 \) and \( F_2 \) in Eq. (4) depend not only on the cutting conditions, but also on the tool and workpiece material. They must be estimated accurately in real time in order to enable the identification of the level of wear increase.

In order to solve these problems, the coefficients \( \alpha, \beta \) and the exponents \( \alpha, \beta \) in Eqs. (2) and (7) must be known. The exponents \( \alpha \) and \( \beta \) of the Taylor's tool life equation have been traditionally determined from off-line experiments \[ T \]. These tool life estimates have not been sufficiently accurate due to variations in material properties and cutting conditions. On-line tool wear monitoring using force measurements was initially based on correlations between force and wear for various constant cutting conditions and known materials \[ 28-29, 36 \]. These results, although useful, require extensive off-line testing and are limited in applicability. Some of these methods have been extended even further by using "learning" or "averaging" techniques to account for process variability \[ 28, 29 \]. These approaches and others \[ 38-40 \] are all basically empirical, and utilize on-line or off-line estimation methods to develop simple relationships between the measured forces and wear.

Wear and breakage have also been investigated from a more fundamental viewpoint by researchers who attempt to identify and quantify the mechanisms of wear in metal cutting \[ 41-48 \]. Mechanically activated (e.g., abrasion and adhesion) and thermally activated mechanisms have been proposed for flank wear \[ 43 \]; and crater wear is generally attributed to thermally activated mechanisms \[ 48 \]. These mechanistic models have provided a better understanding of the problem of tool wear, but have been too complex to be of practical use.

Our proposed approach, described in the following section, builds on these previous studies to develop a model-based and breakage system which uses force sensing and is suitable for variable cutting conditions.

3. ESTIMATION OF PROCESS PARAMETERS

The estimation of the unknown coefficients \( \beta \) and \( \alpha \) and the unknown exponents \( \alpha \), \( \beta \) and \( \gamma \) in Eqs. (2), (6) and (7) can be accomplished by on-line (recursive) parameter estimation methods. In this paper the standard recursive least squares (RLS) algorithm is used \[ 49, 50 \].

For the process model presented in the previous section, we first consider the estimation of \( \alpha \) and \( \beta \) using Eq. (2). When the tool is sharp, the term \( F_1 \) in Eq. (4) is approximately zero, and one can write

\[ F = F_0 = a\alpha^2 \] (8)

This can be rewritten in the form of Eq. (10) by taking natural logarithms of both sides,

\[ \ln F(k) = \ln F_0 + \ln a + \ln \alpha + \ln \beta \] (11)

or,

\[ \ln F(k) = \ln F_0 + \ln a + \ln \alpha + \ln \beta \] (12)

So, initially the estimation algorithm in Eqs. (8) and (9) is used with,

\[ y(k) = \ln F(k) - \ln \alpha \] (13)

\[ y(k) = \ln F(k) - \ln \alpha \] (14)

\[ y(k) = \ln F(k) - \ln \alpha \] (15)

\[ y(k) = \ln F(k) - \ln \alpha \] (16)

\[ y(k) = \ln F(k) - \ln \alpha \] (17)

\[ y(k) = \ln F(k) - \ln \alpha \] (18)

\[ y(k) = \ln F(k) - \ln \alpha \] (19)

\[ y(k) = \ln F(k) - \ln \alpha \] (20)

\[ y(k) = \ln F(k) - \ln \alpha \] (21)

\[ y(k) = \ln F(k) - \ln \alpha \] (22)

\[ y(k) = \ln F(k) - \ln \alpha \] (23)

\[ y(k) = \ln F(k) - \ln \alpha \] (24)

\[ y(k) = \ln F(k) - \ln \alpha \] (25)

\[ y(k) = \ln F(k) - \ln \alpha \] (26)

\[ y(k) = \ln F(k) - \ln \alpha \] (27)

\[ y(k) = \ln F(k) - \ln \alpha \] (28)

\[ y(k) = \ln F(k) - \ln \alpha \] (29)

\[ y(k) = \ln F(k) - \ln \alpha \] (30)

\[ y(k) = \ln F(k) - \ln \alpha \] (31)

\[ y(k) = \ln F(k) - \ln \alpha \] (32)

\[ y(k) = \ln F(k) - \ln \alpha \] (33)

\[ y(k) = \ln F(k) - \ln \alpha \] (34)

\[ y(k) = \ln F(k) - \ln \alpha \] (35)

\[ y(k) = \ln F(k) - \ln \alpha \] (36)

\[ y(k) = \ln F(k) - \ln \alpha \] (37)

\[ y(k) = \ln F(k) - \ln \alpha \] (38)

\[ y(k) = \ln F(k) - \ln \alpha \] (39)

\[ y(k) = \ln F(k) - \ln \alpha \] (40)

\[ y(k) = \ln F(k) - \ln \alpha \] (41)

\[ y(k) = \ln F(k) - \ln \alpha \] (42)

\[ y(k) = \ln F(k) - \ln \alpha \] (43)

\[ y(k) = \ln F(k) - \ln \alpha \] (44)

\[ y(k) = \ln F(k) - \ln \alpha \] (45)

\[ y(k) = \ln F(k) - \ln \alpha \] (46)

\[ y(k) = \ln F(k) - \ln \alpha \] (47)

\[ y(k) = \ln F(k) - \ln \alpha \] (48)

\[ y(k) = \ln F(k) - \ln \alpha \] (49)

\[ y(k) = \ln F(k) - \ln \alpha \] (50)

\[ y(k) = \ln F(k) - \ln \alpha \] (51)

\[ y(k) = \ln F(k) - \ln \alpha \] (52)

\[ y(k) = \ln F(k) - \ln \alpha \] (53)

\[ y(k) = \ln F(k) - \ln \alpha \] (54)

\[ y(k) = \ln F(k) - \ln \alpha \] (55)

\[ y(k) = \ln F(k) - \ln \alpha \] (56)

\[ y(k) = \ln F(k) - \ln \alpha \] (57)

\[ y(k) = \ln F(k) - \ln \alpha \] (58)

\[ y(k) = \ln F(k) - \ln \alpha \] (59)

\[ y(k) = \ln F(k) - \ln \alpha \] (60)

\[ y(k) = \ln F(k) - \ln \alpha \] (61)

\[ y(k) = \ln F(k) - \ln \alpha \] (62)

\[ y(k) = \ln F(k) - \ln \alpha \] (63)

\[ y(k) = \ln F(k) - \ln \alpha \] (64)

\[ y(k) = \ln F(k) - \ln \alpha \] (65)

\[ y(k) = \ln F(k) - \ln \alpha \] (66)

\[ y(k) = \ln F(k) - \ln \alpha \] (67)

\[ y(k) = \ln F(k) - \ln \alpha \] (68)

\[ y(k) = \ln F(k) - \ln \alpha \] (69)

\[ y(k) = \ln F(k) - \ln \alpha \] (70)

\[ y(k) = \ln F(k) - \ln \alpha \] (71)

\[ y(k) = \ln F(k) - \ln \alpha \] (72)
to estimate \( \theta, \sigma, \) and \( \kappa \).

The RLS estimation algorithm requires that one select initial values for \( \theta(k) \). These selections can either be made from typical values of \( \theta, \sigma, \) and \( \kappa \) as published in the literature [36, 37, 38]. The value of the gain matrix \( P(k) \) is usually selected to be of the form,

\[
P(0) = \sigma I
\]

where \( I \) is the identity matrix, and the scalar constant \( \sigma \geq 0 \) is chosen by trial and error through simulation studies. The RLS algorithm is typically not too sensitive to the choice of \( P(0) \) and \( \sigma(0) \).

The choice of the weighting factor \( \lambda(k) \) in Eqs. (8) and (9) is more significant. The standard RLS algorithm is obtained when \( \lambda(k) = 1 \) when it is necessary to track parameters with a value that may jump or vary slowly with time, a value of \( 0 < \lambda(k) < 1 \) exponentially weights the data to ensure the "alertness" of the algorithm [49, 50]. In the simulation results presented in the next section a value of \( \lambda(k) = 1 \) was used. In those simulation studies scaling and factorization of the \( P(k) \) matrix into upper triangular and diagonal factors was also used. Such measures are often required in practice to eliminate numerical problems [49, 50]. For a process model with \( \phi(\tau) \) and \( \theta(\tau) \) the "adaptation algorithm" is based on the RLS algorithm presented in Eqs. (8) and (9). Fig. 1 shows the purpose of this simulation study the "cutting process" is represented by a model described previously in [48]. That model is more detailed than the simple process model presented here and accounts for crater wear as well as flank wear due to both thermally and mechanically activated mechanisms. Results presented in [48] show that the model gives good results for force, temperature, flank wear, and crater wear. The model equations as well as the cutting conditions used in the simulation are summarized in the Appendix. Under these cutting conditions flank wear dominates, and the effect of crater wear is negligibly small.

To provide sufficient input richness for parameter convergence using the RLS algorithm, the feed \( f(t) \) is varied and the tool wear was varied as shown in Fig. 5. The simulated flank wear and force versus time are shown in Figs. 5 and 6 respectively. It is seen that there is initially a rather high wear rate followed by a relatively constant wear rate for \( \tau > 10 \) minutes. For \( \tau > 10 \) minutes the wear rate again increases and the useful life of the tool is \( \tau = 11 \) minutes. The simulated force shows the effects of the variations in feed, and also an increase in level due to increasing flank wear.

The RLS algorithm in Eqs. (8) and (9) is used to estimate the parameters in the model presented in the previous sections. The parameter estimation is carried out during the first few sampling periods and leads to constant values of \( \alpha = 1520 \) and \( \kappa = 0.0787 \). Next, assuming that \( \alpha \) is constant, \( P(0) \) is calculated and used to estimate the \( \theta \) term in Eq. (7). The estimate of \( \theta(k) \) is shown in Fig. 7, and is quite effective at tracking \( \theta \). The effects of feed variations on force from those related to the flank wear are shown in Fig. 8, where we have found that \( \alpha(k) \) becomes large, and serves as a good indicator of tool failure. Thus, the simulation results shown in Fig. 4 -10 illustrate the potential usefulness of the model-based approach for separating the effects of feed and flank wear on force, and for predicting tool failure due to excessive flank wear.

5. SUMMARY AND CONCLUSIONS

This paper has presented a model-based approach to on-line tool wear and breakage detection in metal cutting. Such a model-based approach is considered important for machining under variable cutting conditions, and for use with adaptive control systems that automatically adjust feedrates. The basic approach has been developed, and illustrated with a simple simulation example. The simulation results confirm the feasibility of the proposed model-based approach, and indicate the need for further research to obtain experimental confirmation. Further research may also be desirable on process modeling, estimation algorithms, and on-line training of the model-based approach by using artificial intelligence methods.

6. REFERENCES


7. APPENDIX

The cutting process in Fig. 3 is simulated using the following model from [48],

\[ \begin{align*}
W_1 &= \left( v/K_1 \right) \left( W_2 + K_2 \cos \left( \alpha_1 / (9.0) \right) \right) \\
W_2 &= K_2 \sqrt{v} \exp \left( - \left( K_3 / (273 + v) \right) \right) \\
K_1 &= K_0 (1 + w) \\
K_2 &= K_3 (1 - K_4 v) \\
K_3 &= K_5 (1 - K_6 v) \\
K_4 &= K_7 (1 - K_8 v) \\
K_5 &= K_9 (1 - K_10 v) \\
\end{align*} \]

where the parameters \( k \) and exponents \( n \) are selected to be typical of a carbide tool [48]. The values used in the simulation study are,

\[ \begin{align*}
\alpha &= 2.5m \quad K_0 = 160 \\
\sigma &= 0.35 \text{ m/s} \quad K_1 = 0.57 \\
\tau &= 300 \text{ m/min} \quad K_2 = 86 \\
\end{align*} \]

The table below summarizes the comments on the importance of wear, breakage, and adaptive control.

<table>
<thead>
<tr>
<th>Commentary</th>
<th>Importance of Wear Breakage Adaptive Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short Productivity Cycle</td>
<td>Important</td>
</tr>
<tr>
<td>Long Productivity Cycle</td>
<td>Important</td>
</tr>
<tr>
<td>Finishing Cut</td>
<td>Important</td>
</tr>
<tr>
<td></td>
<td>Important</td>
</tr>
<tr>
<td></td>
<td>Important</td>
</tr>
</tbody>
</table>

Table 1. Summary of Comments on the Importance of Wear, Breakage, and Adaptive Control
Fig. 1 Cutting Force versus Wear [36]

Fig. 2 Typical Wear Curve

Fig. 3 Schematic of the Proposed Model Based Approach

Fig. 4 Feed Versus Time for the Simulation Example

Fig. 5 Flank Wear Versus Time for the Simulation

Fig. 6 Cutting Force Versus Time for the Simulation

Fig. 7 Estimated $F_i$ (i.e., $F_o$) Versus Time for the Simulation Example

Fig. 8 Estimated $a$ Versus Time for the Simulation Example
Fig. 9 Estimated $p$ versus Time for the Simulation Example

Fig. 10 Actual and Estimated $F_{1,t}$ (i.e., $F_{1,t}$) versus Time for the Simulation example