

Fig. 4. Learning curve with different learning gains ($\alpha = 0.5, 0.9, 1, 1.1,$ and 1.5) and fixed feedback gain ($\beta = 1$).

The proposed method can guarantee convergence of the learning process even under the influence of unlearnable dynamics.

REFERENCES

- [1] S. Arimoto, S. Kawamura, and F. Miyazaki, "Bettering operation of robots by learning," *J. Robot. Syst.*, vol. 1, pp. 123–140, 1984.
- [2] D. H. Owens, "Iterative learning control—convergence using high gain feedback," in *Proc. IEEE 31th Conf. Decision and Control*, Tucson, AZ, Dec. 1992, pp. 2545–2546.
- [3] G. Tao and P. V. Kokotovic, "Discrete-time adaptive control of plane with unknown output dead-zones," *Automatica*, vol. 31, no. 2, pp. 287–291, 1995.

Multi-Scale Statistical Process Monitoring in Machining

Xiaoli Li and Xin Yao

Abstract—Most practical industrial process data contain contributions at multiple scales in time and frequency. Unfortunately, conventional statistical process control approaches often detect events at only one scale. This paper addresses a new method, called multiscale statistical process monitoring, for tool condition monitoring in a machining process, which integrates discrete wavelet transform (WT) and statistical process control. Firstly, discrete WT is applied to decompose the collected data from the manufacturing system into uncorrelated components. Next, the detection limits are formed for each decomposed component by using Shewhart control charts. A case study, i.e., tool condition monitoring in turning using an acoustic emission signal, demonstrates that the new method is able to detect abnormal events (serious tool wear or breakage) in the machining process.

Index Terms—Condition monitoring, machining processes, statistical process control (SPC), wavelet transform (WT).

I. INTRODUCTION

Statistical Process Control (SPC) has been widely applied for the detection of abnormal process operations based on the process variables such as dimensions, forces, vibrations, and so on; thus, SPC is becoming a common practice in modern manufacturing [1]. Traditional SPC methods include Shewhart chart, moving average (MA) method, cumulative sum (CUSUM) chart, principal component analysis, and partial least-squares regression [1]. Different methods can detect the different malfunctions in industrial processes. The disadvantage of the traditional SPC is that it is only applicable to certain problems as they work only on a fixed time scale. We also see that for most practical processes some information about the representation of the abnormal operation is not known beforehand. In short, the traditional SPC methods are very hard to deal with time-varying variables for monitoring in machining processes.

So far, SPC methods usually assume that the machining process contains contributions from events at a single scale. In fact, a machining process is inherently multiscale in nature due to events occurring with different localizations in time and frequency. A typical example of acoustic emission (AE) data from a machining process is shown in Fig. 1. The upper line in Fig. 1 shows original data from normal to abnormal condition (serious wear) of the cutting tool; the segments of these two conditions are marked as A and B, respectively. Their time-frequency distributions using continuous wavelet transform (WT) also were plotted, respectively. As can be seen from Fig. 1 (upper pane), AE data with different cutting conditions in a turning process are plotted. At 1.5 s, the cutting process begins. At 7 s, the cutting tool experiences light wear. Before 7 s, the wear of the cutting tool is acceptable. However, from 7 s, the wear of the cutting tool increases considerably, from 10 s, it increases sharply, and then the cutting tool fails completely and stops cutting. "A" and "B" in Fig. 1 plot the time-frequency distributions of two different cutting tool conditions. The malfunction of AE data mainly localizes in the high-frequency band by comparison with

Manuscript received April 11, 2003; revised August 23, 2004. Abstract published on the Internet March 14, 2005. This work was supported by the Alexander von Humboldt Foundation of Germany and by Advantage West Midlands (AWM), U.K.

The authors are with the Centre of Excellence for Research in Computational Intelligence and Applications (CERCIA), School of Computer Science, The University of Birmingham, Birmingham, B15 2TT, U.K. (e-mail: xiaoli.avh@gmail.com).

Digital Object Identifier 10.1109/TIE.2005.847580

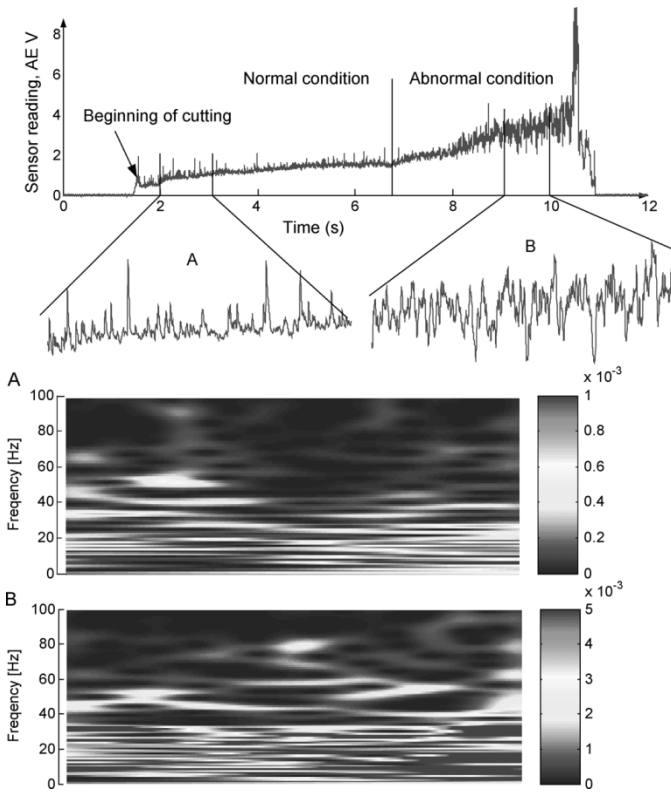


Fig. 1. Typical AE data in turning process that spans from normal to abnormal tool condition. The time-frequency distributions of two different conditions with continuous WT are plotted.

the two time-frequency distributions. Compared with the energy distribution at the different scales, the difference between normal and abnormal condition can be highlighted. Ideally, techniques for detecting the tool condition change via AE data in the cutting process should adapt automatically to the features of different scales.

In response to the need of the machining process monitoring, this paper proposes a new method for overcoming the single-scale nature of SPC. This new method integrates discrete WT, SPC, filter, and inverse WT. Although continuous WT is better to describe the time-frequency distribution of AE data, here, we only consider the application of discrete WT because the tool condition monitoring system is required to work in real time. A case study demonstrated that this new method is better in detecting the tool condition in turning than a traditional SPC method.

II. MULTISCALE STATISTICAL PROCESS MONITORING

Currently, WT is a great tool for signal processing [2], [3]. The basic idea of WT is to decompose a signal into a number of components, called the subsignals. WT can be viewed as a multiscale process. At the fine scale, wavelets may be long capturing low-frequency components of the signal. On the other hand, at low scales, wavelets are usually short, capturing the high-frequency components of the signal. Combining positions and scales of the wavelets, the local features of a signal can thereby be identified.

For discrete signals, $x(n)$ the discrete WT (DWT) is defined as

$$\begin{aligned} c_{j,k} &= \sum_n x(n)h_j(n - 2^j k) \\ d_{j,k} &= \sum_n x(n)g_j(n - 2^j k) \end{aligned} \quad (1)$$

where $c_{j,k}$ is the wavelet coefficient and $d_{j,k}$ is the scaling coefficients ($j = 1, 2, \dots, J$, $k = 1, 2, \dots, N$), J is the scales, N is the length of discrete signal. h is the discrete wavelets correspondent to scaling function $\psi(t)$, and g represents the scaling sequence correspondent to wavelet function $\phi(t)$. It can be shown that, for resolution $j > 0$, the approximations and the details can be computed using the following recursive formula:

$$\begin{aligned} c_{j+1,k} &= \sum_n g(n - 2k)d_{j,k} \\ d_{j+1,k} &= \sum_n h(n - 2k)d_{j,k}. \end{aligned} \quad (2)$$

Often, the terms g and h can be taken as high-pass and low-pass filters, respectively, which are derived from the wavelet $\psi(t)$ and the scaling function $\phi(t)$. The shifted and dilated basic transform described are applied to all the consecutive pairs of signals. In this letter, we will present fast Haar WT to analyze AE signals. The fast Haar WT begins with the initialization of an array with $2n$ entries, and then proceeds with n iterations of the basic transform. The procedure of fast Haar WT is as follows.

Step 1) *Initialization*.—The initialization consists of establishing a one-dimensional array $\mathbf{a}^{-(n)}$

$$\mathbf{a}^{-(n)} = (a_0^{(n)}, \dots, a_j^{(n)}, \dots, a_{2^n-1}^{(n)}) = (s_0, \dots, s_j, \dots, s_{2^n-1})$$

with a total number of sample values integral power of two $2n$, as indicated by the superscript (n) .

Step 2) *Sweep*—The l th sweep of the basic transform begins with an array of $2n - (l - 1)$ values

$$\mathbf{a}^{-(n-[l-1])} = (a_0^{(n-[l-1])}, \dots, a_{2^{n-(l-1)}-1}^{(n-[l-1])})$$

and applies the basic transform to each pair $(a_{2k}^{(n-[l-1])}, a_{2k+1}^{(n-[l-1])})$, which gives two new wavelet coefficients

$$\begin{aligned} a_k^{(n-1)} &= \frac{a_{2k}^{(n-[l-1])} + a_{2k+1}^{(n-[l-1])}}{2} \\ c_k^{(n-1)} &= \frac{a_{2k}^{(n-[l-1])} - a_{2k+1}^{(n-[l-1])}}{2}. \end{aligned}$$

These $2(n - l)$ pairs of new coefficients represent the result of the l th sweep, and the result can be reassembled into two arrays

$$\begin{aligned} \mathbf{a}^{-(n-l)} &= (a_0^{(n-l)}, \dots, a_{2^{n-l}-1}^{(n-l)}) \\ \mathbf{c}^{-(n-l)} &= (c_0^{(n-l)}, \dots, c_{2^{n-l}-1}^{(n-l)}). \end{aligned}$$

Therefore, the presentation of each step of the fast Haar WT requires additional arrays at each sweep. In addition, some applications require real-time processing, and they do not allow sufficient space for additional arrays at each sweep. The in-place fast Haar WT [4] is able to solve the two problems: lack of space or time. For each pair $(a_{2k}^{(n-[l-1])}, a_{2k+1}^{(n-[l-1])})$ in in-place fast WT, instead of placing its results in two additional arrays, the l th sweep of the in-place transform merely replaces the pair $(a_{2k}^{(n-[l-1])}, a_{2k+1}^{(n-[l-1])})$ by the new entries $(a_k^{(n-l)}, c_k^{(n-l)})$. Clearly, the in-place fast Haar transform differs from the fast Haar transform only in its indexing scheme, but it does not require additional arrays

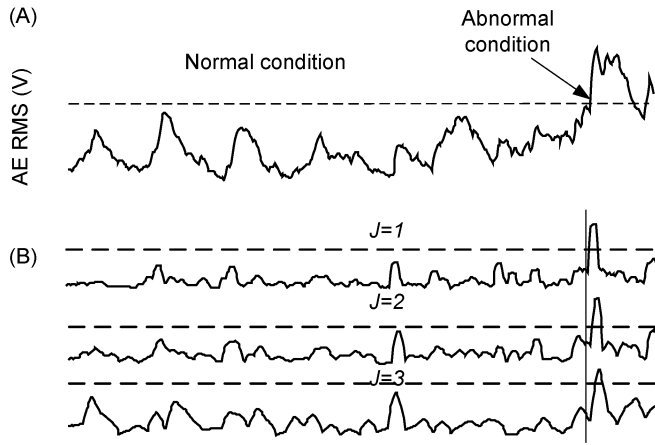


Fig. 2. (a) Detection of abnormal condition with Shewhart. (b) Detection of abnormal condition at the different scale.

at each sweep, so the lack of space and time problems have a solution in the in-place algorithm.

A typical Shewhart chart for a monitoring system contains a center-line that represents the mean value of the measured variable, and the other two horizontal lines represent the upper monitoring limit (UML) and the lower monitoring limit (LML). As long as the measured variable falls within the limits, the process is considered a normal condition. Otherwise, it is interpreted as evidence of an abnormal condition. Let $x(n)$ be a measured variable, and its mean and standard deviation are μ_x and σ_x , respectively. Then, the UML and LML are

$$\begin{aligned} \text{UML} &= \mu_x + k\sigma_x \\ \text{LML} &= \mu_x - k\sigma_x \end{aligned} \quad (3)$$

where μ_x is the center line, and k is the constant associated with the confidence coefficient in statistics. Specifying the monitoring limits is one of the critical decisions that must be made when designing the Shewhart chart. As shown in [1] and [5], the Shewhart chart has a couple of advantages and disadvantages. The Shewhart chart can detect large changes quickly, but it is difficult for it to detect small shifts in the mean. The method requires the measurements to be uncorrelated, whereas, in practice, correlated measurements are very common. The disadvantages of the Shewhart chart may be overcome by using a multiscale approach. This approach consists of decomposing each measured variable into multiscales by using a Haar wavelet basis function. The decomposition permits identification of signal features at various scales.

Using a simple example, as shown in Fig. 2, reveals the new method and its performance. At the end of the data, there occurs an abnormal event. The detection limits are determined from the normal data. The results of the Shewhart chart and the new method for this example are shown in Fig. 2(a) and (b), respectively. Although the Shewhart chart can detect the events, it is very different from the new method. In the new method, the representation of AE data after discrete WT mainly contains the features, which just represent the abnormal condition [see Fig. 2(b)]. Meanwhile, the detection limits change according to the nature of the signal and the scale at which the features are present. From Fig. 2(b), it is seen that the shift first occurred at the scale $J = 1$, then at $J = 2$ and 3. The complete procedure of this new monitoring method is plotted through an example, as shown in Fig. 3.

III. CASE STUDY: TOOL FRACTURE DETECTION IN TURNING

AE signals result from the stress waves generated by the sudden release of energy as the workpiece deforms, which is currently one of

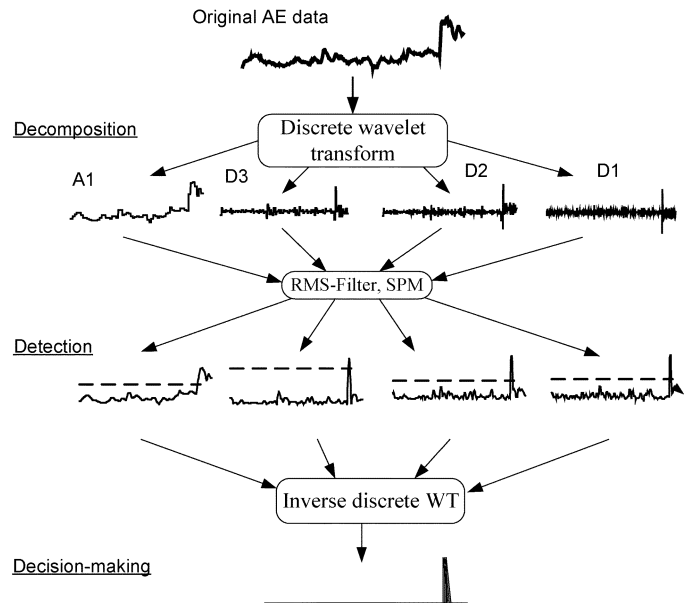


Fig. 3. Illustration of the proposed multiscale statistical process monitoring. A typical cutting condition is as follows: cutting speed $v = 200$ m/min, feed $f = 0.1$ mm/rev, and depth of cut $d = 1$ mm; workpiece material is mid-grade steel, and cutting tool is CNMG-12-04-04-QF.

the most effective ways of detecting tool condition in machining [6]. A piezoelectric AE sensor was mounted onto the side of the tool shank. The signal was firstly preamplified by 40 dB, and then went through a bandpass filter. The frequency band of the filter is set at 100 kHz–1 MHz and, hence, the low-frequency noises can be eliminated. The AE signal was also rectified and averaged with a time constant of 1.2 ms, which gave the AE rms. Finally, the AE rms was sampled by a PC at the sample frequency of 2500 Hz. Normally, the AE rms signal contains much information on cutting processing, including tool condition, cutting conditions, chip breakage, and tool condition. From Fig. 1, the AE signal in machining is nonstationary and often comprises many overlapping transient components.

Fig. 3 shows the complete detection procedure of tool fracture by using the AE rms signal during machining. A tool fracture occurs at the end of the data, causing a larger fluctuation of the signal. The proposed method can identify the event in the cutting process. As a first step, fast Haar WT is applied to decompose the AE rms signals, the approximation A_3 , and details D_3 , D_2 , and D_1 are obtained, respectively. A_3 represents the energy distribution of the AE rms signal. When a tool fracture occurs, a large peak is presented. However, this information cannot ascertain the tool fracture because other changes (such as the increase of the depth of cut) may also cause an increase of the energy. The effect of cutting conditions (e.g., the increase of the depth of cut) is rather small in D_1 of the AE rms signal, though the effect of the tool fracture is significant. D_1 corresponds to the high-frequency information of the AE rms. From a mechanical point of view, the strength of the tool material is higher than that of the workpiece. Hence, when a tool fracture occurs, a higher frequency AE rms signal is produced. Therefore, D_1 is effective in detecting the tool fracture during turning. D_2 has a similar effect as that of D_1 . Nevertheless, in D_3 , the information about the tool fracture is considerably less. Following the idea of multiscale statistics, we set alarm thresholds separately based on the Shewhart chart. In particular, for D_3 and A_3 , the thresholds are set high to avoid a false alarm. In contrast, for D_1 and D_2 , the thresholds are set low to avoid missing the alarm. Then, by combining A_3 , D_1 , D_2 , and D_3 , the accuracy of the condition monitoring can be improved. In order to minimize the effect of the noise, such as the chip-breaking noise and

the environmental noise, an rms filter is applied to treat the subsignals. Then, a Shewhart control chart is applied to the filtered subsignals to detect the abnormality (i.e., the threshold crossing) of each subsignal. In Fig. 3, it is seen that the tool fracture is detected successfully in each subsignal, though it is not a clear cut. Finally, combining the detection results at each scale and taking the inverse WT are to improve the accuracy of the detection approach, that is, taking all the points that cross the limit and setting the other points at zero, then applying the inverse WT. The tool fracture can easily be detected.

To test the performance of the new method, a total of 25 experiments under different cutting conditions were performed. It may be argued that tool condition monitoring with fixed cutting conditions is relatively easy. Hence, we also tested some cases in which the cutting conditions are varying. The performance of the new method with a comparison to the Shewhart chart for the tests is presented below. The new method was 100% correct in detecting tool fractures (the true-positive ratio, $TP = 100\%$). Nevertheless, out of the five negative events (normal condition), there is a false alarm (the false-negative ratio, $FN = 80\%$ and the false-positive ratio $FP = 20\%$). In contrast, the Shewhart chart cannot completely detect the tool fracture during turning: the TP ratio is about 68%, the FN ratio is 21%, and the FP ratio is about 26%. The success of the new method may be attributed to the fact that the Haar WT can effectively decompose the AE rms signal and extract the information about the tool fracture. Inverse DWT synthesizes the detection results at the different scales, resulting in the improvement in detecting tool fracture in machining. The main advantages of this new method were also discussed in [7]–[9].

IV. CONCLUSION

In a practical machining process, AE data contain deterministic changes with different temporal durations, and stochastic variation with changing intensity over frequency or time. These features are not known beforehand. Fortunately, a new tool, WT, is able to compress deterministic features at all scales into a small number of relatively large coefficients, thus, it adapts to features at different scales by

focusing only on those scales that contain coefficients outside the detection limits. Based on the experiments of tool fractures in turning, the new approach has excellent performance. Its success may be attributed to the fact that it examines the sensor signal in multiple scales (DWT), and highlights the variation using the threshold crossing and inverse DWT. Differently from traditional SPC methods, the new method is adaptive in general and detects very well the different types and magnitudes of changes. Although the nature of the change is unknown for most industrial processes, the new method seems to be ideal. Also, the new method can easily deal with autocorrelated measurements.

This letter has only focused on the simplest SPC, the Shewhart chart; in fact, the new method can also specialize in other existing SPC methods, such as MA, CUSUM, and EWMA charts, depending on the scale of the abnormal feature and the selected wavelet. Selection of the depth of decomposition and type of wavelet may be automated in future work.

REFERENCES

- [1] D. C. Montgomery, *Introduction to Statistical Quality Control*. New York: Wiley, 1996.
- [2] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 961–1005, Sep. 1990.
- [3] —, "Orthogonal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. XLI, pp. 909–996, 1988.
- [4] H. K. Tönshoff, X. Li, and C. Lapp, "Fast Haar transform and concurrent learning for manufacturing process monitoring," *IEEE/ASME Trans. Mechatron.*, vol. 8, no. 3, pp. 414–418, Sep. 2003.
- [5] B. R. Bakshi and S. Top, "Multiscale statistical process monitoring and diagnosis of univariate and multivariate processes," presented at the 1998 AIChE Annu. Meeting, Miami Beach, FL, 1998.
- [6] X. Li, "A brief review: Acoustic emission for tool wear monitoring during turning," *Int. J. Mach. Tools Manuf.*, vol. 42, pp. 157–165, 2002.
- [7] B. R. Bakshi, "Multiscale PCA with application to multivariate statistical process monitoring," *AIChE J.*, vol. 44, no. 7, pp. 1596–1610, 1998.
- [8] —, "Multiscale analysis and modeling using wavelets," *J. Chemometrics*, vol. 13, no. 3–4, pp. 415–434, 1999.
- [9] H. B. Aradhye, B. R. Bakshi, R. A. Strauss, and J. F. Davis, "Multiscale SPC using wavelets—theoretical analysis and properties," *AIChE J.*, vol. 49, no. 4, pp. 939–958, 2003.