TOOL WEAR PREDICTION BASED ON WAVELET TRANSFORM AND SUPPORT VECTOR MACHINES

Dongfeng Shi[1] and Nabil N. Gindy[2]
[1]*Optimized Systems and Solutions, Rolls-Royce Group, Derby, U.K.*
[2]*School of Mechanical, Materials and Manufacturing Engineering*
The University of Nottingham, NG7 2RD Nottingham, U.K.

Keywords: Support Vector Machine, Wavelet Transform, Machining Process Monitoring.

Abstract: The machining quality and efficiency may be improved significantly by using appropriate tool wear prediction techniques. A new approach based on wavelet transform and support vector machine is proposed to improve the accuracy of tool wear prediction in this paper. Firstly, the wavelet transform is introduced to decompose sensory signals into different scales to reduce the dimensionality of original signals and extract features associated with different tool wear condition. Secondly, the least square support vector machine is further presented to construct predictive model due to its high convergence rate and powerful generalization ability. Thirdly, the possibility to employ power sensor rather than delicate dynamometer for the tool wear monitoring is explored. Finally, the effectiveness of proposed tool wear prediction approach is demonstrated by extensive experimental turning trials.

1 INTRODUCTION

Tool wear will progress with the proceeding of the machining process due to the involvement of fracturing, abrasion, plastic deformation, diffusion and grain-pullout. The dimensional accuracy and surface quality of machined component may be deteriorated by excessive worn tool. Consequently, the online tool wear monitoring is required within aero-engine manufacturing industry to improve the machining quality of critical components made of Titanium or Nickel alloys. Due to high corrosion resistance associated with those super alloys, the wear of machining tool deteriorates rapidly. Through the utilization of tool wear prediction technique, the worn tool can be detected and replaced in time to avoid scrapping critical components. Moreover, common industrial practice by replacing or regrinding tools according to a conservative schedule is not cost-effective. By implementation of tool wear prediction technique, the tooling cost may be reduced and tool life may be prolonged significantly.

Several indirect tool wear predictive approaches have been investigated by modelling the correlation between tool wear and sensory signals, namely force, vibration and acoustic emission, acquired in machining processes (Sick, 2002). However, further efforts are still required in the following aspects despite the fact that several achievements have been made in tool wear prediction so far. Firstly, although several different types of sensor, e.g. accelerometer, dynamometer, acoustic emission and motor current sensor have been employed to measure the responses in machining processes, the overall performance of these sensors in terms of accuracy, robustness and cost-effectiveness is still not satisfaction. In general, the cutting force acquired from dynamometers is regarded as one of significant variables in the machining processes due to its direct relation with tool wear. However, the implementation of dynamometers in shop floor is restricted due to high cost, negative impact on machining system rigidity, the requirement for a wiring harness and extra space for installation (Shi et al., 2006). Recently, indirect sensing cutting force through the feed or spindle motor current of a machining tool has been investigated extensively due to the ease of installation and low cost (Stein and Wang, 1990, Altintas, 1992, Lee et al., 1995). However, this indirect approach has been reported not sensitive and accurate enough to measure the cutting force in machining process due to limited frequency range

Shi D. and N. Gindy N..
TOOL WEAR PREDICTION BASED ON WAVELET TRANSFORM AND SUPPORT VECTOR MACHINES.
DOI: 10.5220/0003647304790485
In *Proceedings of the 8th International Conference on Informatics in Control, Automation and Robotics* (MSIE-2011), pages 479-485
ISBN: 978-989-8425-75-1

(Altintas, 1992). As a result, for the purpose of implementation of tool wear monitoring system in industrial environment, alternative sensing solutions have to be investigated to strike the balance between effectiveness and cost. Secondly, feature extraction plays crucial role in the improvement of accuracy and robustness of tool wear predictive model since the original sensory signals usually are interfered with noise, disturbance and redundant information. Normally, statistical moments based features, i.e. mean value, standard deviation, extracted from sensory signal have been always employed to predict tool wear. However, this feature extraction technique is not effective enough to explore the instinct features associated with tool wear. Consequently, a more advanced feature extraction technique is required to filter out the noise component and reduce the dimensionality of the original data to improve prediction accuracy. Finally, neural network has been extensively used to model the correlation between sensory signals and tool wear. However, the prediction results were not satisfied due to some disadvantages, i.e. low convergence rate, obvious 'over-fitting' and especially poor generalization when few samples are available. Support Vector Machines (SVM) based on statistical learning theory is a new achievement in the field of data-driven modelling and implemented successfully in classification, regression and function estimation (Kwok, 1999, Cao and Tay, 2003, Goethals and Pelckmans, 2005). SVM has been proved less vulnerable to overfitting problem and higher generalization ability since SVM is designed to minimize structural risk whereas previous neural networks techniques, i.e. MLP, are usually based on minimization of empirical risk (Kwok, 1999). Consequently, the applicability of SVM in the tool wear modeling will be explored in this paper.

The objective of this paper is to develop a new monitoring approach to predict tool wear using sensory signals acquired in machining processes. The organization of the work is as follows. In Section 2, wavelet transform is explored to extract features from sensory signals. The SVM is further introduced to model the correlation between tool wear and extracted features in Section 3. The performance of proposed approach is demonstrated by experimental data acquired from turning processes in Section 4. The conclusions are given in last Section.

2 WAVELET TRANSFORM BASED FEATURE EXTRACTION

The sensory signals acquired in machining process are typical non-stationary multi-componential signals caused by uneven material removing process. Different tool malfunctions, i.e. tool wear, tool chipping and tool breakage, may possess different frequency characteristics in sensory signals. For instance, the cutting force will increase gradually with the increase of tool wear and will be obviously reflected in the lower frequency band or so-called static component of sensory signals. On the contrary, tool chipping or breakage will cause cutting force changed suddenly and may be observed in higher frequency band or so-called dynamic component of sensory signals. As a result, the features associated with different tool malfunctions may be extracted from either static or dynamic component of sensory signals. Several techniques, i.e. band-pass filtering, resample and wavelet transform, may be employed to decompose sensory signals. From the point of view of filter design, wavelet transform is a typical cascade band-pass filter with a varying bandwidth. The sensory signals can be decomposed into different frequency bands or scales to capture localized features i.e. abrupt or gradual changes within the sensory signals by analysis corresponding wavelet coefficients. Wavelet transform provides an efficient way to identify the location and possible root cause of the malfunction within the machining processes because of powerful decomposition ability. Additionally, by implementation wavelet transform at specified scale, the sensory signal can be descried as few wavelet coefficients and the dimensionality of sensory signals can be dramatically reduced. Hence, in comparison with other two decomposition techniques, wavelet transform is more powerful and flexible due to its multi-resolution capability and hence explored to obtain static component for feature extractions. The wavelet transform of signal *s(t)* is defined as the inner product in the Hilbert space of L2 norm as follows (Mallat, 1997):

$$C(a,b) = |a|^{-1/2} \int_{-\infty}^{+\infty} s(t)\psi_{a,b}^{*}(t)dt \qquad (1)$$

where $\psi_{a,b}^{*}(t)$ is the complex conjugate of $\psi_{a,b}(t)$ generated by scaling and shifting from so-called a 'mother wavelet' function expressed as

$$\psi_{a,b}(t) = |a|^{-1/2} \psi\left(\frac{t-b}{a}\right) \qquad (2)$$

where a is a scale factor and b is a translation or time shift parameter. The factor $|a|^{-1/2}$ is used to ensure energy preservation. A family of scaled and shifted wavelets can be produced through varying the parameters a and b. Therefore, the time-scale characteristics of the signal $s(t)$ can be analyzed by the inner product to the series of scaled and shifted wavelets. In order to obtain the numerical result of wavelet transform, the parameter of scale a and shift b must be discretized. Discrete wavelet transform normally is conducted by dyadic discretization, $a=2^j$, $b=k2^j$, $(i, j) \in Z^2$. Additionally, regarding the possibility of time-frequency localization, the mother wavelet must be compactly supported and satisfied with the admissibility condition:

$$C_\Psi = \int_{-\infty}^{+\infty} \left|\Psi(\omega)\right|^2 / \omega d\omega < \infty \tag{3}$$

where $\Psi(\omega)$ is the Fourier transform of $\psi(t)$. Then, the discrete synthesis of wavelet transform is expressed as

$$s(t) = \sum_{j\in Z}\sum_{k\in Z} C(j,k)\psi_{j,k}(t) \tag{4}$$

At specified scale J, the discrete synthesis can be further rewritten as

$$s(t) = A_J(t) + \sum_{j\le J} D_j(t) \tag{5}$$

where $D_j(t)$ is called the detail of the signal $s(t)$ at scale j and expressed as

$$D_j(t) = \sum_{k\in Z} C(j,k)\psi_{j,k}(t) \tag{6}$$

and $A_J(t)$ is called an approximation of the signal $s(t)$ at scale J and expressed as

$$A_J(t) = \sum_{j>J} D_j(t) \tag{7}$$

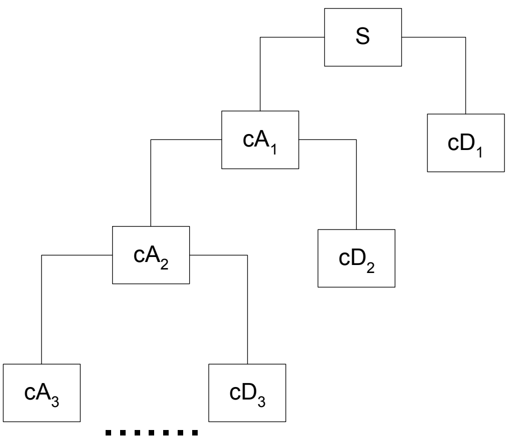

Figure 1: Illustration of decomposition tree of wavelet transform.

As a result, a decomposition tree is formed where the signal is decomposed to a number of details and one approximation as shown in Figure 1. The approximation captures the low frequency content which corresponds to static component of the signal and details reflect the high frequency contents which correspond to dynamic components of the signal. As described earlier, wavelet transform is a typical set of cascade band-pass filters with varying bandwidth. The central frequency and bandwidth of the wavelet-based cascade filter depends on the choice of scale. If Daubechies-wavelet, i.e. *db5*, is selected as a mother wavelet, the wavelet-based band-pass filter at scale J will be centred at the quotient between sampling frequency and 2^J. In this paper, the decomposition scale of sensory signals is specified as $J=8$ since the highest frequency of static component of sensory signal (sampled at 1000Hz) is found less than 4Hz. Additionally, the dimensionality of sensory signal can be reduced significantly since the length of the static component is only $1/2^J$ times of the length of original sensory signal. Hence, the corresponding wavelet coefficients at specified scale J can be formed as feature vectors to feed into SVM-based tool wear predictive model as introduced in Section 3.

3 LS-SVM BASED TOOL WEAR PREDICTIVE MODEL

SVM is a novel machine-learning tool and especially useful for the classification and prediction with small-sample cases (Vapnik, 1999). This novel approach motivated by statistical learning theory led to a class of algorithms characterized by the use of nonlinear kernels, high generalization ability and the sparseness of the solution. Unlike the classical neural networks approach the SVM formulation of the learning problem leads to quadratic programming (QP) with linear constraint. However, the size of matrix involved in the QP problem is directly proportional to the number of training points. Hence, to reduce the complexity of optimization processes, a modified version, called LS-SVM is proposed by taking with equality instead of inequality constraints to obtain a linear set of equations instead of a QP problem in the dual space (Suykens et al., 2002, Suykens and Vandewalle, 1999). Instead of solving a quadratic programming problem as in SVM, LS-SVM can obtain the solutions of a set of linear equations. The formulation of LS-SVM is introduced as follows.

Consider a given training set $\{x_k, y_k\}_{k=1,\cdots N}$ with input data $x_k \in \Re^n$ and output data $y_k \in \Re$. The following regression model can be constructed by using nonlinear mapping function $\varphi(\cdot)$

$$y(x) = w^T \varphi(x) + b \tag{8}$$

where w is the weight vector and b is the bias term. By mapping the original input data into a high-dimensional space, the nonlinear separable problem becomes linearly separable in space. Then, the following cost function is formulated in the framework of empirical risk minimization.

$$\min J(w,e) = \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{k=1}^{N} e_k^2 \; ; \tag{9}$$

subject to equality constraints

$$y_k = w^T \phi(x_k) + b + e_k \qquad k = 1, \cdots, N \tag{10}$$

where e_k is the random errors and γ is a regularization parameter in determining the trade-off between minimizing the training errors and minimizing the model complexity. To solve this optimization problem, Lagrange function is constructed as

$$L(w,b,e;\alpha) = J(w,e) - \sum_{k=1}^{N} \alpha_k \{ w^T \phi(x_k) + b + e_k - y_k \} \tag{11}$$

where a_k are Lagrange multipliers. The solution of Equation (11) can be obtained by partially differentiating with respect to w, b, e_k and a_k

$$\frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{k=1}^{N} \alpha_k \phi(x_k) \tag{12}$$

$$\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{k=1}^{N} \alpha_k = 0 \tag{13}$$

$$\frac{\partial L}{\partial e_k} = 0 \rightarrow \alpha_k = \gamma e_k \quad k = 1, \cdots N \tag{14}$$

$$\frac{\partial L}{\partial \alpha_k} = 0 \rightarrow w^T \phi(x_k) + b + e_k - y_k = 0, k = 1, \cdots N \tag{15}$$

The Equations (12)-(15) can be rewritten as

$$\begin{bmatrix} 0 & \vec{1}^T \\ \vec{1} & \Omega + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{16}$$

Where

$y = [y_1 \cdots y_N]$

$\vec{1} = [1 \cdots 1]$

$\alpha = [\alpha_1 \cdots \alpha_N]$

$\Omega_{kl} = \phi(x_k)^T \phi(x_l) \cdots k, \quad l = 1 \cdots N$

Finally, b and a_k can be obtained by the solution to the linear system

$$\hat{b} = \frac{\vec{1}^T (\Omega + \gamma^{-1} I_n)^{-1} y}{\vec{1}^T (\Omega + \gamma^{-1} I_n)^{-1} \vec{1}} \tag{17}$$

$$\hat{\alpha} = (\Omega + \gamma^{-1} I)^{-1} (y - \vec{1} \hat{b}) \tag{18}$$

According to Mercer's theorem, the resulting LS-SVM model can be expressed as:

$$y(x) = \sum_{k=1}^{N} \alpha_k K(x, x_k) + b \tag{19}$$

where $K(x, x_k)$ is the nonlinear kernel function. In comparison with some other feasible kernel functions, the RBF function is a more compact supported kernel and able to reduce computational complexity of the training process and improve generalization performance of LS-SVM. As a result, RBF kernel was selected as kernel function as

$$K(x, x_k) = \exp(\|x - x_k\|_2^2 \cdot \sigma^{-2}) \; , \tag{20}$$

where σ is the scale factor for tuning.

To achieve a high level of performance with LS-SVM models, some parameters have to be tuned, including the regularization parameter γ and the kernel parameter corresponding to the kernel type, i.e. σ. Finally, the features extracted in Section 2 and actual tool wear measured by optical scan microscope can be employed to construct input-output pairs to train LS-SVM. In the training stage, the correlation between sensory signals and tool wear is learned by LS-SVM. Once the training stage is accomplished, the trained LS-SVM is used to predict tool wear by using the features extracted from wavelet transform.

4 EXPERIMENTAL RESULTS AND DISCUSSIONS

4.1 Experimental Configuration

Two types of sensors, namely, dynamometers (Kistler 9257B) and power sensor (Load control LC-PH-3A-10V) are employed to conduct experiments in turning processes. The possibility of the utilization of power sensor rather than delicate dynamometer will be investigated based on critical analysis of experimental results. The power sensor was installed with spindle motor to measure the machining power. The power is estimated by vector multiplications between current and voltage samples sensed by Hall-effect sensors. In comparison with well-known motor current sensor, the power sensor is more accurate and appropriate to measure power consuming in machining process due to the con-

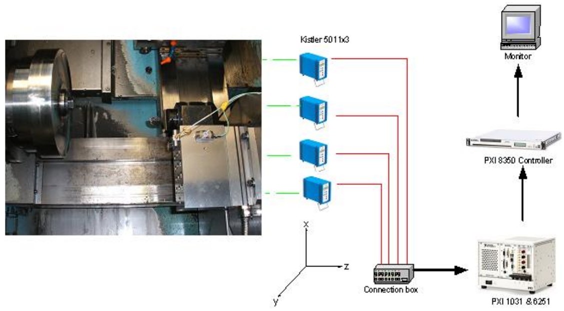

Figure 2: Schematic diagram of the online turning monitoring system.

sideration of power factor variation with the changing load.
National Instruments PXI modules, namely, NI PXI-1031 chassis, 3.0GHz Pentium 4 Rack-mount PXI controller and 16-Bit NI PXI-6251 with 16 analog inputs and 24 digital I/Os, have been specified as the hardware platform to construct DAQ package. LabVIEW has been selected as software platform to develop the whole package due to its powerful performance in data acquisition, graphical user interface (GUI) design, and hardware connectivity. The developed process monitoring software is capable to acquire, analyze and present the data simultaneously due to the utilization of multithread programming techniques i.e. queue technique. For the purpose of the reduction the manual interference, data can be automatically stored in specified file and the name of file can be stamped according to the starting time of sampling. Moreover, the power sensory signal has been selected as the triggering source to conduct self-triggering by using the impulse generated by the starting of spindle motor. The corresponding software has been developed to run in re-triggerable manner to acquire signals successively without manual interferences. By the implementation of self-triggering
technique, the acquired signals are started at exact same moment without the requirement for further alignment. The whole online machining process monitoring system is shown schematically in Figure 2.

4.2 Tool Wear Prediction in Turning Process

A Swedturn 4-axes CNC twin lathe was employed to manufacture Inconel 718 disc. Ceramic tools were used in the experimental trials due to the performance in terms of high melting point, excellent hardness and wear resistance for the machining of hard materials. Ceramic insert RCGX 35T-0320 with constant tool edge preparation (clearance angle 1° and rake angle 13°) and different tooling conditions were employed to conduct turning trials. To meet industrial requirements, the Inconel 718 disc with complicated profile as shown in Figure 3 was specified to manufacture.

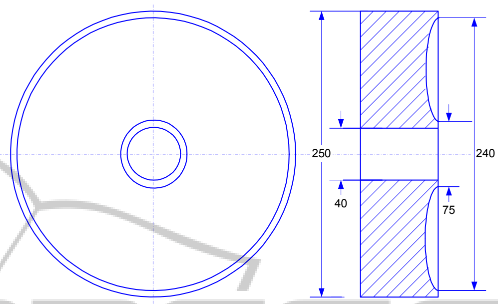

Figure 3: Geometrical parameter of Inconel 718 disc for turning.

Additionally, the dynamometer Kistler 9257B and power sensor Load control LC-PH-3A-10V were installed to acquire force and power signals respectively. To demonstrate the effectiveness of proposed prediction approach based on wavelet transform and SVM, several turning trials have been performed to acquire sensory signals under different tool wear conditions. The tool wear in terms of VB was measured by optical scan microscope after each cutting as shown in Figure 4.

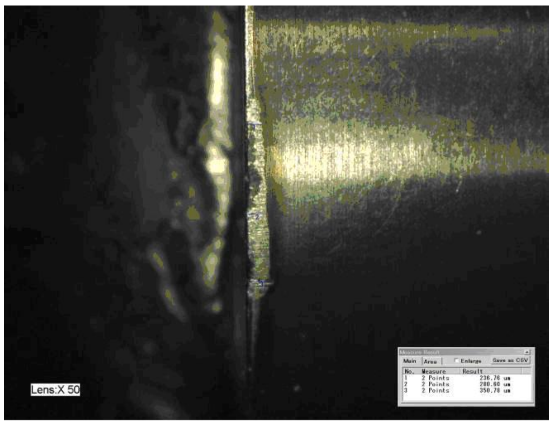

Figure 4: Photo of turning tool wear taken by optical scan microscope.

The original force and power signals acquired from initial fresh tool toward to excessive tool wear are shown in Figure 5 and 6 respectively. It can be seen that the power signals have the same pattern as force signals acquired from dynamometer. Both signals

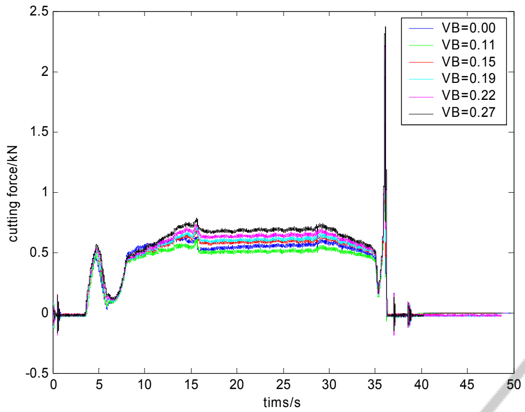

Figure 5: Original force acquired from dynamometer with different wear.

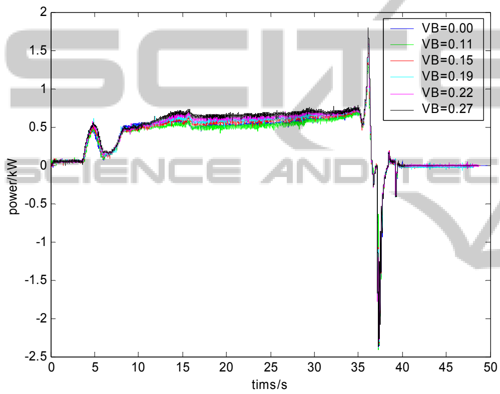

Figure 6: Original power signals with different tool wear.

possess different characteristics at different segment of the profile caused by the variation of effective cutting length between insert and workpiece. The power sensor is recognized as an appropriate alternative sensor for machining process monitoring due to the ease of installation and low cost. However, it seems that the power signal is less sensitive than force signal in the detection of tool wear due to the interference from dynamic components. Hence, the wavelet transform is further employed to decompose power signals into static and dynamic components. It can be seen that amplitude of static components of power signals increased with the proceeding of tool wear as shown in Figure 7.

Additionally, for the purpose of feature extraction, the dimensionality or length of sensory signal can be reduced significantly by the utilization of wavelet transform. Finally, the data sets composed features extracted by wavelet transform and corresponding tool wear measured by optical scan microscope were obtained. The desired output of the LS-SVM represents wear states of the cutting tool in terms of VB. Then all features were normalized against their respective standard deviations. The whole data sets can be further divide into two sub-sets, i.e. training sets and validation sets. Then, the SVM-based tool wear model was trained by training sets and two turning parameters γ and σ was selected as 10 and 0.3 respectively. By application of training algorithm for training sets, the b and a_k can be obtained and stored to construct predictive model. Once the training stage is accomplished, the SVM-based tool wear model was validated by validation sets. The predicted tool wear by using SVM model and actual tool wear measured by optical scan microscope is compared in the Figure 8. A good agreement between them can be found at each level of tool wear. The experimental results show that SVM-based model is effective to predict tool wear by using features extracted from wavelet transform.

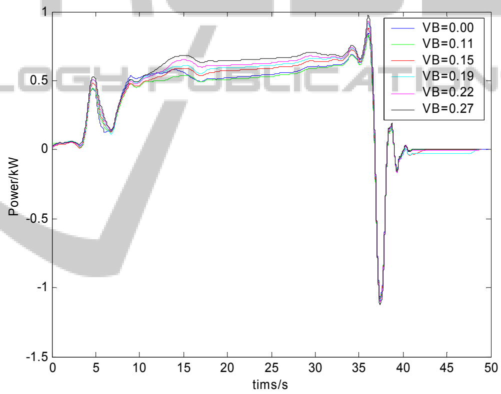

Figure 7: Static components of power signals extracted by wavelet transform.

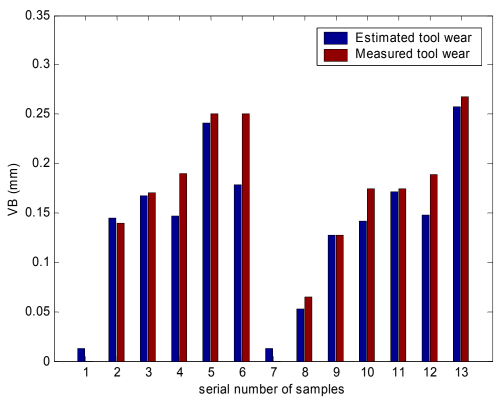

Figure 8: Comparisons between predicted and actual tool wear measured by optical scan microscope.

5 CONCLUSIONS

A new tool wear prediction approach based on wavelet transform and LS-SVM has been developed and demonstrated in turning trials. The major contributions of this work can be summarized as follows:

1. Wavelet transform has been implemented in dimensionally reduction and feature extraction for sensory signals acquired in machining processes. In comparison with conventional feature extraction approaches, wavelet transform technique is capable of exploring the instinct correlation between the sensory signals and tool wear due to its powerful multi-scale decomposition capability.

2. LS-SVM technique has been developed to predict tool wear by using extracted features from wavelet transform. Due to the utilization of statistical learning theory, LS-SVM can overcome several disadvantages with traditional machine learning techniques, e.g. local optimal solution, low convergence rate and poor generalization ability when few samples are available.

3. It has been proved that the sensory signal measured by alternative sensors, i.e. power sensor, correlate with dynamometer signal very well and is sensitive enough to detect tool wear. As a result, the power signals have been selected to conduct feature extraction due to the cost-effectiveness and the ease of installation.

4. The effectiveness of proposed prediction approach has been demonstrated in experimental turning trials. A good agreement can be found between predicted tool wear obtained by LS-SVM and actual tool wear measured by optical scan microscope.

ACKNOWLEDGEMENTS

The financial sponsorship from EPSRC and technical supports from industrial partners, namely, Rolls-Royce (Colin Sage, Jamie McGourlay and John Burkinshaw), Siemens (Julian Timothy and Gordon Lanes), Kistler (Eddie Jackson) and TBG Solution (Paul Rawlinson) are gratefully acknowledged.

REFERENCES

B. Sick, On-line and indirect tool wear monitoring in turning with artificial neural networks: a review of more than a decade of research, *Mechanical Systems and Signal Processing* 16 (2002) 487–546

D. F. Shi, D. A Axinte and N. N Gindy 'Development of an online machining process monitoring system: A case study of broaching process', *International Journal of Advanced Manufacturing Technology,* 2006, (in press)

J. L. Stein, C. H. Wang, "Analysis of power monitoring in AC induction drive systems", *ASME Trans. on Journal of Dynamic Systems, Measurement and Control* Vol. 112, pp239–248, 1990

Y. Altintas, "Prediction of cutting forces and tool breakage in milling from feed drive current measurements", *ASME Trans. on Journal of Engineering for Industry,* Vol. 114, pp386–392, 1992

J. M. Lee, D. K. Choi, J. Kim, and C. N. Chu, "Real-time tool breakage monitoring for NC milling process," *Ann. CIRP, Vol. 44,* No. 1, pp 59–62, 1995.

J. Kwok, Moderating the outputs of support vector machine classifier. *IEEE Trans. Neural Networks,* 10(1999) 1018–1031

L. J. Cao and FEH Tay, Support vector machine with adaptive parameters in financial time series forecasting, *IEEE Trans. Neural Networks,* 14(6) (2003) 1506–1518

I. Goethals, K. Pelckmans, JAK Suykens and Bart De Moor, Subspace identification of Hammerstein systems using least squares support vector machines, *IEEE Trans. on Automatical Control,* 50(10) (2005) 1509-1519

S. Mallat, *A Wavelet Tour of Signal Processing.* London, Academic Press Limited, 1997

V. N. Vapnik, *The nature of statistical learning theory,* Springer, New York, 1999

J. A. K. Suykens, T Van Gestel, J De Brabanter, B De Moor and J Vandewalle, *Least squares support vector machines,* World Scientific, Singapore, 2002

J. A. K. Suykens, J Vandewalle, Least squares support vector machine classifiers, *Neural processing letters,* 9 (1999),293-300