Proceedings - 2009 2nd IEEE International Conference on Computer Science and Information Technology, ICCSIT 2009
2009, Article number 5234475, Pages 85-89
DOI: 10.1109/ICCSIT.2009.5234475

PCA fused NN approach for drill wear prediction in drilling mild steel specimen

S.S. Panda
Department of Mechanical Engineering, Indian Institute of Technology Patna, Bihar-800013, India,
Tel No.: 91-9905647117(M);
Fax No.: 91-612- 2277383
Email:sspanda@iitp.ac.in

S.S. Mahapatra
Department of Mechanical Engineering, National Institute of Technology Rourkela,Orissa-769008, India 3, Tel No.: 91-661-2462512(O);
Fax No.: 91-661-2462999 , India,
Email: ssm@nitrkl.ac.in

Abstract: **The present paper describes use of principal components for drill wear prediction. It also makes a comparative analysis in using large sensor based technique in predicting drill wear. In order to reduce the redundancy of the network, principal component has been fused with artificial neural network (ANN) for prediction of drill wear. Large numbers of experiments have been conducted and sensor signals have been acquired using data acquisition system. Cutting force, torque, vibrations along with other process parameters such as spindle speed, feed rate, drill diameter, chip thickness and surface roughness have been used as indicative parameters for characterizing the progressive wear of drill. Principal component of these input parameters has been derived thereafter and has been used to predict the flank wear using BPNN**

Keywords-component: Neuron; sensor integration; signal analysis; design of experiment; flank wear; BPNN;PCA

I. INTRODUCTION

The reason for acquiring the drill wear state information is to enhance the predictive capability to allow the machine operator to schedule tool change or regrind just in time to avoid under use or overuse of tools, avoid shutdown of machines due to damage, and to minimize scrap or rework. Drill wear also affects the ability of the hole cutting system to satisfy specified performance characteristics, such as hole roundness, centering, burr formation at drill exit, and surface finish. Wearing action in the tool is an inevitable phenomenon leading to the loss of dimensional accuracy and possible damage to the work piece. Hence, online prediction of cutting tool wear becomes an important issue for today's manufacturing industries.

Many early works on deterioration and failure of drill have been reported in literature. Altogether different types of drill wear can be recognized as outer corner wear, flank wear, margin wear, crater wear, chisel wear and chipping at the lip as reported by Kanai and Kanda [1]. Some of the previous works (Chungchoo and Saini [2], Haili et al. [3]) reported that average width of crater and maximum depth of flank wear could be used as to assess tool failure. It is relatively common to measure only the thrust or axial component of force for TCM in drilling as suggested by Thangaraj and Wright [4]. The spindle and feed motor current are also used as inputs to the TCM system by (Ramamurthi and Hough [5]; Liu and Chen [6] used eight indices of thrust force and torque as input parameters for drill wear monitoring. Stiffness and damping properties of the drilling system has no effect on wear as suggested by Abu-Mahfouz [7]. Vibration monitoring techniques applied to the detection of drill failure have been reported by several investigators (Wang et al. [8]; Ko and Cho [9]; Li et al. [10]; Liu and Wu [11]). Liu and Anantharaman [12] used nine features representing drill condition as input to multi layer perceptron and found 100% accuracy for correct classification of drill wear. Sanjay et al. [13] used drill diameter, feed, cutting speed, time, force and torque as inputs and flank wear was estimated using different structures of artificial neural network (ANN).

Since wear of drill related to number of direct and indirect factors and hence can be considered as dimension reduction problem (Principal component analysis), a dimension reduction technique to identify the wear on drill. Rummel [14] used factor analysis using a mathematical model to study the human ability. Detailed description of factor analysis can be found in the relevant literature such as Kleinbaum et al. [15] approach. Fisher [16] had developed a linear classification algorithm to transform n dimensions pattern to 2 dimensions. Hong and Yang [17] have used a method for constructing a classifier on the optimal discernment plane by using minimal distance criterion for multi classification problems.

II. BACK PROPAGATION NEURAL NETWORK

The basic structure of a BPNN having input, hidden and output layers has been considered in the present work. The input layer receives information from external sources, and passes this information to the network for processing. The hidden layer receives information from the input layer, and does the information processing. The output layer receives processed information from the network, and sends the results to an external receptor. The input signals are modified by the interconnection weights, known as weight factor w_{ji} , which represents the interconnection of i^{th} node of the first layer to j^{th} node of the second layer. The sum of modified signals (total activation) is then modified by a sigmoid transfer function. Batch mode type of supervised learning has been used in the present case. During training, the predicted output has been compared with the desired output, and the mean square error has been calculated. If the mean square error is more than a prescribed limiting value, it is back propagated from output to input, and weights are further

978-1-4244-4520-2/09/$25.00 ©2009 IEEE

modified till the error or number of iteration are within a prescribed limit.

Mean square error, E_p for pattern p is defined as

$$E_p = \sum_{i=1}^{n} \frac{1}{2}(D_{pi} - O_{pi})^2 \qquad (1)$$

where, D_{pi} is the target output, and O_{pi} is the computed output for the i^{th} pattern.

Weight change at any iteration t, is given by

$$\Delta W(t) = -\eta E_p(t) + \alpha \times \Delta W(t-1) \qquad (2)$$

where η is learning rate, and α momentum parameter.

Entire experimental data set is divided into training set, testing set and validation set. The error on the testing set is monitored during the training process. The testing error normally decreases during the initial phase of training, as does the training set error. However, when the testing error starts increasing for a specified number of iterations, the training is stopped; and the weights at the minimum value of the testing error are returned. With the trained network, unseen data set (validation set) are verified and percentage of variation of predicted output (flank wear) with respect to the actual output is thus evaluated.

III. PRINCIPAL COMPONENT ANALYSIS

Principal component analysis is a statistical technique which generated uncorrelated linear combinations of original variables and account for the total variance and original data so that adequate information of original data can be extracted.

In principle, each of the principal components is a linear combination of the original Y values for the p variables;

$$PC_1 = c_{11}Y_1 + c_{12}Y_2 + c_{13}Y_3 + ... + c_{1p}Y_p \qquad (axis\ Z_1)$$

$$PC_2 = c_{21}Y_1 + c_{22}Y_2 + c_{23}Y_3 + ... + c_{2p}Y_p \qquad (axis\ Z_2)$$

$$..................................$$

$$PC_p = c_{p1}Y_1 + c_{p2}Y_2 + c_{p3}Y_3 + ... + c_{pp}Y_p \qquad (axis\ Z_p) \qquad (3)$$

Where, $c_{a,b}$ is the component score coefficient for variable b on PC axis Z_a, and Y_b is the Y score for variable b.

Principal Component Analysis transforms a multivariate set of variables (Y_1, Y_2, ... , Y_p) to new variables (Z_1, Z_2, ... , Z_p) , which are uncorrelated. The first principal component consists of a principal component coefficient (α_i) for each variable (p) such that there is maximal variance in the calculated score for each case (n); the factor score for each case is calculated as $\alpha_1.Y_1 + \alpha_2.Y_2 + \ldots + \alpha_i.Y_i + \ldots \alpha_p.Y_p$ where Y_i is the centered value for the i^{th} (Y_i - mean Y for the i^{th} variable). The first principal component axis for the raw variables Y_1 is the new axis Z_1. The second principal component consists of the next set of principal component coefficients (α_i) such that there is a maximal remaining variance in the calculated score for each case (n), and there is no correlation of the second principal component with the first. Z_2 is the second principal component axis for Y_2. Further sets of principal components (third, fourth, *etc*) can be calculated until no statistical significance can be attributed to that component (*e.g.* by χ^2 test) or all m principal components are computed.

IV. EXPERIMENTAL SET-UP

In the present work, a radial drilling machine (Batliboi Limited, BR618 model) has been used for the drilling operation. Uncoated HSS drills with different diameters have been used for drilling holes in mild steel work piece at different cutting conditions. Different sensory devices such as dynamometer and vibration analyzer are used for sensing thrust force signals, torque signals and vibration signals in the experiments is shown in fig. 1.

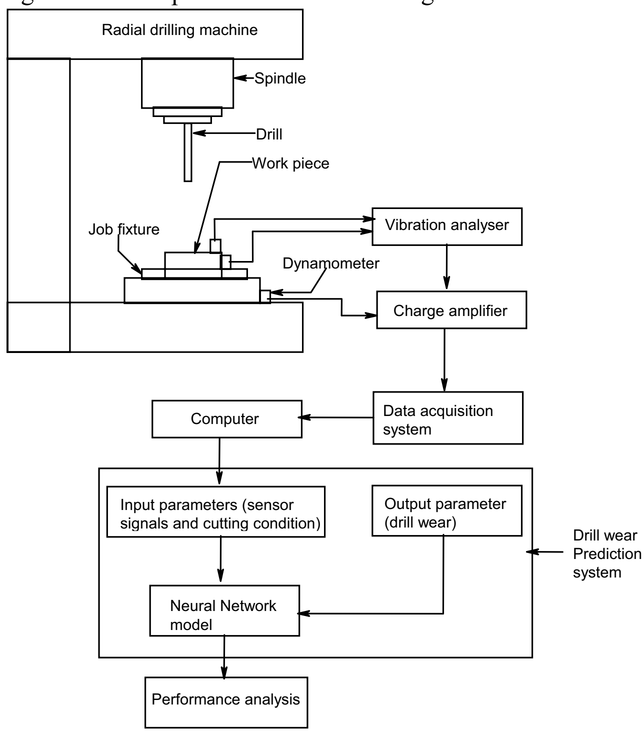

Figure 1 Schematic diagram of the experimental set-up

HSS drill bits with different diameters have been used for drilling in mild steel work-piece at different cutting conditions in dry state. Chips are collected during each cutting condition and average thickness is measured. Similarly, after each cutting operation surface roughness of drill hole is measured with help of pocket surface roughness tester (make Mahr) with maximum roughness depth 0.2 μm to 25.3 μm and maximum stylus force 15.0 *mN*. In all the drilling operations performed in the present work, no coolant has been used. Root mean square (RMS) values of thrust force and torque signal are recorded through a piezo- electric dynamometer (Kistler, 9272). Signals from the dynamometer were passed through low pass filter, amplified through charge amplifier (Kistler, type 5015 model), and stored in the computer through a data acquisition system (Advantech, PCL 818 HG, 16 channel analog to digital (A/D) converter in 16 bit digital time-discrete, and 100 *kHz* sampling rate). Two piezo electric accelerometers (model Bruel & Kjaer, type 4396) have been used to capture vibration signals. One accelerometer has been attached on the top surface of the mild steel specimen to extract feed vibration and other on the side surface of the mild steel specimen to extract radial vibration. Signals from accelerometer were passed through vibration analyzer (Bruel & Kjaer, type 3560 D) in the frequency range 7 *Hz*-25.6 *kHz*. RMS of maximum amplitude of vibration both in feed and radial direction were collected through Bruel & Kjaer pulse software (version 7) and is stored in the computer through data acquisition system through data recorder (type Bruel & Kjaer, Type 7701). The optical microscopes along with Carl-Zeiss software interfacing have been used to measure flank wear. The maximum flank wear has been used as the criterion to characterize the drill condition, and is obtained by measuring the wear at different points on either of the cutting edges.

V. RESULTS AND DISCUSSION

Drilling operations have been conducted over a wide a range of cutting condition. Spindle speed has been varied in the range 315 *rpm* to 1000 *rpm* in four steps. Feed rate has been varied from 0.13 to 0.36 *mm/rev* in four steps. High-speed steel (HSS) drills of four different diameters of (9*mm*, 10*mm, 11mm* and 12*mm*) have been used for drilling 64 numbers of through holes of 15*mm* depth for various combination of spindle speed, feed rate and drill diameter in mild steel plates.

VI. WEAR PREDICTION BY PCA FUSED NN

In order to extract the principal component, a test statistics of different variable of sample size of 64 have been studied. XLSTAT version 6 has been used for finding of principal components. Table 1 provides the Eigen values for all nine principal components, with the percentage of the variance which they explain and their cumulative percentage of variation explained. Cumulatively Eigen values one, two and three explain 98.11% of the total variance. Rest of the component explains only about 1.89% of cumulative variance. Total 3 principal component axes have been considered from the present work is gauged from the fig. 2 as these three components explain about 98.11% of the total variance. Taking consideration of these three principal component (PCA1, PCA2 and PCA3) which is having all the variables at different score coefficient. So these three principal components is now acting as input to neural network and output is drill flank wear. Now network has been trained at different condition of learning rate, momentum coefficient and number of neuron in hidden layer. Large numbers of runs were given for selecting the best architecture and table 2 shows only the best and the second best network in terms of the validation error and the number of iteration.

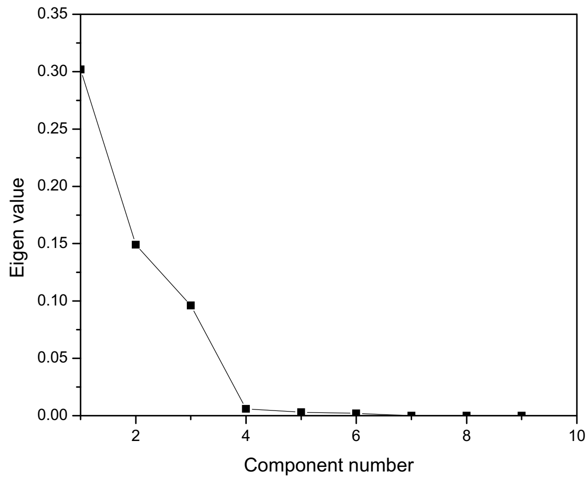

Figure 2 Screen plot

The optimum network architecture in the present case has been observed to be 3-3-1 with η = 0.1 and α = 0.3. Variation of mean square error in training and testing with number of iteration for the optimum network is shown in fig. 3. After the network has been trained, it has been validated with unknown data sample. Fig. 4 shows percentage of error between actual value and the predicted value and it can be observed that present neural network of 3-3-1 with η = 0.1 and α = 0.3 predicts the results within $\pm$6.42 % error.

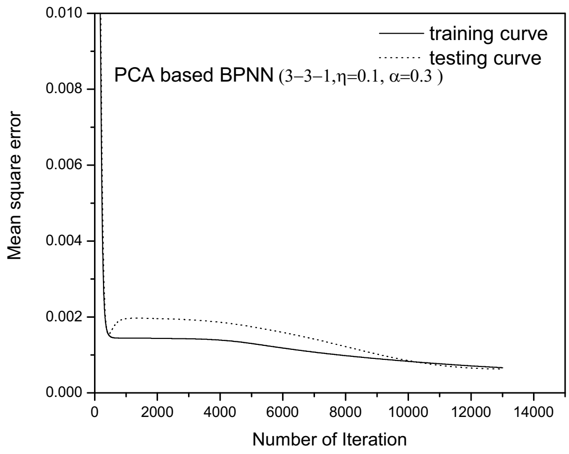

Figure 3 Variation of mean square error with number of iteration of 3-3-1 with η = 0.1 and α = 0.3

This optimum network is later on used for classifying the new set of pattern of low wear and high wear. Hence in the same experimental environment another 64 test data (32 data in low range of wear and 32 data in high range of

Table 1 Variance of Eigen values

Value	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7	PC 8	PC 9
Eigen value	0.302	0.149	0.096	0.006	0.003	0.002	0.000	0.000	0.000
Variance (%)	54.127	26.799	17.189	1.056	0.523	0.271	0.024	0.009	0.003
Cumulative variance (%)	54.127	80.926	98.115	99.170	99.694	99.964	99.988	99.997	100.00

Table 2 Optimum network architectures for PCA based BPNN

Architecture L-M-N	η	α	Iteration	MSE training	MSE testing	Maximum Validation error (%)
3-3-1	0.1	0.3	13010	0.000664	0.000631	6.42
	0.1	0.7	5578	0.000666	0.000631	6.44
	0.3	0.3	4336	0.000665	0.000631	6.43
	0.3	0.7	1868	0.000668	0.000631	6.46
	0.5	0.1	2426	0.000871	0.000944	8.51
	0.5	0.5	1861	0.000667	0.000631	6.45
3-5-1	0.1	0.1	13665	0.000469	0.000479	9.45
	0.1	0.5	7685	0.000465	0.000477	9.43
	0.1	0.9	1486	0.000506	0.000529	7.06
	0.3	0.1	4614	0.000464	0.000477	9.42
	0.3	0.3	3596	0.000469	0.000476	9.44
	0.5	0.7	938	0.000486	0.000498	8.03
	0.5	0.9	327	0.000516	0.000676	7.8
	0.7	0.5	1114	0.000493	0.000497	7.59
	0.7	0.7	564	0.000545	0.000591	6.78
3-8-1	0.1	0.1	11773	0.000406	0.000512	7.39
	0.1	0.7	3921	0.000407	0.000514	7.39
	0.1	0.9	1343	0.000399	0.000535	7.22
	0.3	0.1	3944	0.000402	0.000512	7.36
	0.3	0.3	3045	0.000408	0.000513	7.4
	0.5	0.1	2124	0.000525	0.000518	7.9
	0.5	0.3	1756	0.000518	0.00055	7.45
	0.7	0.3	1192	0.000452	0.000483	8.638785
	0.7	0.5	838	0.000531	0.000523	7.845272
	0.7	0.7	509	0.000529	0.000522	7.834434
	0.9	0.7	392	0.000381	0.00053	8.833469
3-12-1	0.1	0.1	11084	0.00045	0.000523	7.261126
	0.1	0.5	6151	0.000451	0.000522	7.262812
	0.1	0.7	3681	0.000454	0.000519	7.265948
	0.3	0.3	2868	0.000452	0.000517	7.230488
	0.3	0.5	2061	0.000449	0.000511	7.15226
	0.5	0.3	1613	0.000497	0.000454	7.957349
	0.5	0.5	1158	0.000534	0.00048	7.74159
	0.7	0.5	829	0.000515	0.000452	8.904967
	0.9	0.5	650	0.000484	0.000443	8.814713

wear) were selected. The classification accuracy of the network was calculated by taking the number of correctly classified sample by the network and divided by the total number of sample into the test data is represented as confusion matrix in table 3. Table 3 shows that classification accuracy of the network on average is approximately 90%.

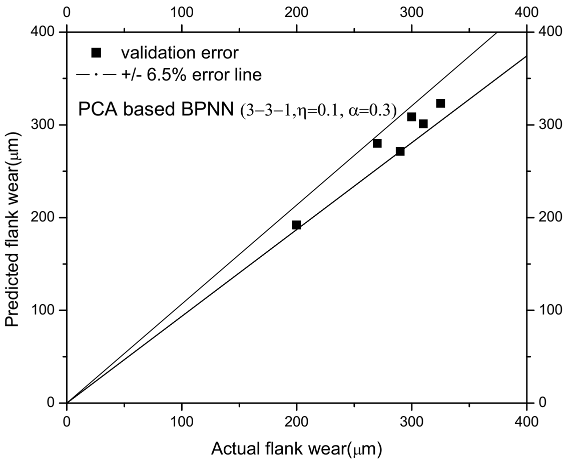

Figure 4 Comparison of Predicted value with actual value of 3-3-1 with $\eta = 0.1$ and $\alpha = 0.3$

Table 3 Confusion matrix for wear classification

	No of sample	Low wear	High wear	Accuracy
Low wear	32	26	6	81.25%
High wear	32	1	31	96.87%
	64			89.06%

VII. CONCLUSION

From the present study following conclusions have been made.

Generalized Conclusions:

(i) Direct process parameters such as spindle speed, feed rate, and drill diameter, has definitive effect on progressive flank wear of the drill.

(ii) Wear on the drill (condition of the drill) affects the sensor signals such as thrust force, torque, vibration and hence these could be used in the drill wear prediction,

(iii) Drill wear (condition of the drill) also affects the chip thickness as well as the surface roughness and hence thickness of the chip and the surface roughness could also be used as indicative parameters for prediction of drill condition.

Derived Conclusions:

(iv) Nine input variable could be better represented by three principal component

(v) PCA based BPNN architectures took approximately 13000, iteration to predict the drill flank wear within $\pm$ 6.5% error band.

(vi) Network could able to classify low wear and high wear within an accuracy of 90%.

REFERENCES

[1] Kanai, M. and Kanda, Y. (1978), Statistical characteristic of drill wear and drill life for standardized performance tests, Annals of CRIP, 27(1), 61-66

[2] Chungchoo, C. and Saini, D. (2002), A computer algorithm for flank and crater wear estimation in CNC turning operations, Int. J. of Mach. Tools & Manuf., 42, 1465–1477

[3] Haili, W., Hua, S., Ming, C. and Dejin, H. (2003), On-line tool breakage monitoring in turning, J. of Mater. Process. Technolo., 139, 237–242.

[4] Thangaraj, A and Wright, PK (1998), Computer assisted prediction of drill failure using in-process measurements of thrust force, J. of Eng. Ind., 110,192–200

[5] Ramamurthi, K. and Hough, CL (1993), Intelligent real-time predictive diagnostics for cutting tools and supervisory control of machining operations, ASME, J. of Eng. for Ind., 115, 268–277

[6] Liu, T. I. and Chen, W.Y. (1988), Intelligent detection of drill wear, Mech. Syst. & Signal Process., 12(6), 863-873

[7] Abu-Mahfouz, I. (2003), Drilling wear detection and classification using vibration signals and artificial neural network, Int. J. Mach. tools & Manuf., 43, 707–720

[8] Wong, S.V. and Hamouda, A.M.S. (2003), Machinability data representation with artificial neural network, J. of Mater. Process. Technolo., 138, 538–544

[9] Ko, T.J. and Cho. D.W. (1994), Cutting state monitoring in milling by a neural network, Int. J. of Mach. & Tools Manuf., 34(5), 659

[10] Li, X., Dong, S. and Venuvinod, P. K. (2000), Hybrid Learning for Tool Wear Monitoring, Int. J. of Adv. Manuf. Technolo., 16, 303–307

[11] Liu, T.I. and Wu, S.M. (1990), On-line detection of drill wear. J Eng. Ind., 112, 299–302

[12] Liu, T.I. and Anantharaman, K.S. (1994), Intelligent classification and measurement of drill wear, ASME, J. Eng. Ind., 116, 392

[13] Sanjay, C., Neema, M.L. and Chin, C.W. (2005), Modeling of tool wear in drilling by statistical analysis and artificial neural network, J. of Mater. Process. Technolo., 170, 494–500

[14] Rummel. R.J. (1988), Applied factor analysis, Evanston, Northwestern University Press.

[15] Kleinbaum, D., Kupper, L., and Vannelli, A. (1986), Applied regression analysis and other multi variable methods, California: International J. of Production Research, 24, 387-397

[16] Fisher, R. A. (1936), The use of multiple measurement in taxonomic problems, Annals of Eugenis, 7, 178-188

[17] Hong, Z. Q. and Yang, J. Y. (1991), Optimal discernment plane for a small number of samples and design method of classifier on the plane, Pattern Recognition, 24 (4), 317-324