Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Sound and Vibration 267 (2003) 178–186
JOURNAL OF SOUND AND VIBRATION
www.elsevier.com/locate/jsvi

Letter to the Editor

Chatter identification with mutual information

B. Berger*, C. Belai, D. Anand

Department of Mechanical Engineering, 2181 Glenn L. Martin Hall, University of Maryland, College Park, MD 20742-3035, USA

Received 9 December 2002; accepted 16 December 2002

1. Introduction

Automatic online control of the cutting process is essential for the efficient machining of metal parts. Cutting states compatible with surface finish and metal removal rate requirements must be maintained. To this end, recently developed signal processing methodologies have been employed including neural networks [1], autoregressive SVD [2], coherence analysis [3], fuzzy set theory [4] and course-grained entropy rates [5].

In the course of an analysis of metal cutting tool force and acceleration time series utilizing course-grained entropy rates [5,6], it was observed that mutual information, I_1 ([7,8], Eq. (6)), as a function of delay, differentiated between chatter and non-chatter cutting states. Mutual information, associated with two-dimensional delay co-ordinates, as a function of delay was determined for sequences of cutting experiments in which either depth of cut, turning frequency or feed rate was varied while all other cutting parameters were held constant. One sequence, s-1, of experiments with variable turning frequency, two sequences, s-2, s-3, with variable depth of cut and one sequence, s-4, with variable feed rate, a total of 19 cutting experiments, were studied. Cutting forces and cutting tool accelerations were measured along z- and x-axis which are, respectively, parallel and perpendicular to the axis of rotation of the work piece. The mutual information, I_1, was averaged over $50<\tau<150$, where $\tau=$ delay.

For a given data set, the averaged mutual information, AI_1, associated with the chatter case was found to be approximately the same for force and acceleration measurements independent of axes. The chatter case averaged mutual information, AI_1, was found to be 1.58, 1.56, 1.56 for sets s-1, s-2, s-3, respectively, and 1.27 for s-4.

AI_1 associated with non-chatter cutting states in s-1, s-2, s-3 satisfied the inequality $0.052<AI_1<0.40$. For the s-4 non-chatter case, the AI_1 satisfied $0.25<AI_1<0.626$. In general, AI_1 associated with acceleration measurements along the x-axis were non-decreasing as the variable cutting parameter monotonically approached its chatter value, taking values greater or

*Corresponding author. Tel.: +1-301-405-2410; fax: +1-301-314-9477.
E-mail address: berger@glue.umd.edu (B. Berger).

doi:10.1016/S0022-460X(03)00067-1

equal to 1.27 for chatter and less than or equal to 0.62 for non-chatter. The results reported here are drawn generally from Ref. [9].

2. Experimental apparatus

The experimental apparatus consisted of a Harding CNC lathe, a force dynamometer utilizing three Kistler 9068 force transducers, Kistler 8628B50 accelerometers with their associated electronics and a digital spectrum analyzer, Hewlett Packard 3566A. All experiments involved only right-handed orthogonal cutting. Positive rake tool inserts, Kennametal TPMR 322, were employed supported by Kennametal KT- GPR123B tool holders. The rake and clearance angles were 5° and 4°, respectively.

Cylindrical work pieces of 1020 steel were machined under a wide range of cutting conditions. Since all work pieces were stubby, work piece modal characteristics did not affect the turning dynamics. The sampling rate was 4096 Hz and the cut-off frequency 1100 Hz.

3. Mutual information

The following definitions and theorems [7,8,10,11] are included to provide a background in redundancies for subsequent application.

Given a partition of probabilities $p_i, \ldots, p_m$ with $\sum p_i = 1,\ i = 1, m$. The Shannon entropy, H, is defined by

$$H = -\sum_i p_i \log p_i. \tag{1}$$

For a time series $x(t)$, $t = 1, \ldots, N$ the Shannon entropy quantifies the average information gained with each finite precision measurement of $x(t)$. The p_i may be determined by a box counting approach with a partition size of δ. The time series $x(t)$ is then discretized by the integers $y = 1, \ldots, M$ depending on which bins its elements fall. Let $p(k)$ be the probability of an element falling into the yth bin. Then $H_1(x, \delta) \equiv H_1(y)$ and $H_1(y) = -\sum_y p(y) \log(p(y))$, where $y = 1, M$. For m variables, the entropy is given by

$$H_1(x_1, \ldots, x_m, \delta) \cong H_1(y_1, \ldots, y_m) \tag{2}$$

and

$$H_1(y_1, \ldots, y_m) = -\sum_{y_i} \cdots \sum_{y_m} p(y_1, \ldots, y_m) \cdot \log p(y_1, \ldots, y_m). \tag{3}$$

It is shown in Ref. [4] that if $x(t)$ is measured then the average uncertainty in a measurement of $x(t + \tau)$ is $H(x_2|x_1)$, where $x_1(t) \equiv x(t)$, $x_2(t) \equiv x(t + \tau)$ and

$$H(x_2|x_1) = H(x_1, x_2) - H_1(x_1). \tag{4}$$

It follows that the amount that a measurement of $x_1(t)$ reduces the uncertainty of $x_2(t)$, $I_1(x_1; x_2, \delta)$, is

$$I_1(x_1; x_2, \delta) = H_1(x_2, \delta) - H(x_2|x_1) \tag{5}$$

or

$$I_1(x_1; x_2, \delta) = H_1(x_1, \delta) + H_1(x_2, \delta) - H_1(x_1, x_2, \delta). \quad (6)$$

The mutual information, $I_1(x_1; x_2, \delta)$, may be interpreted as follows: Assume that the log functions in Eqs. (2) and (3) are taken to the base 2. Given a measurement of x_1. Then the mutual information is a measure of the number of bits which can be predicted about x_2 from a measurement of x_1. The mutual information vanishes if x_1 and x_2 are independent. $I_1(x_1; x_2)$ may be seen as a global measure of the variation of $p(x_1, x_2)$ [7].

In Ref. [7], x_1 and x_2 were identified with the delay embedding $x_1(t) \equiv x(t)$, $x_2(t) \equiv x(t+\tau)$. The mutual information, $I_1(x_1; x_2)$ (Eq. (6)), as a function of the delay τ, evaluates the redundancy of the $x_2(t)$-axis. It was shown that the first minimum in $I_1(x_1; x_2)$, as of function of τ, is a good criterion for the choice of τ in phase-portrait reconstruction from time series.

If $\delta \to 0$ as M increases, then Eq. (2) diverges. However, the integral for the continuous form of Eq. (1),

$$H = -\int p(x) \log(p(x)) \, dx, \quad (7)$$

does not diverge as the partition becomes finer. The continuous form of the mutual information $I_1(x_1, x_2)$,

$$I_1(x_1; x_2) = \int p(x_1, x_2) \log(p(x_1, x_2)/(p(x_1)p(x_2))) \, dx_1 \, dx_2, \quad (8)$$

follows from Eq. (6) [7,8].

A recursive, self-adapting algorithm for the evaluation of Eq. (8) was derived in Ref. [7]. The associated computer program, provided by H.L. Swinney, was utilized in all subsequent computations of $I_1(x_1; x_2)$.

Redundancy, R_1 [10], defined by

$$R_1(x_1, \ldots, x_m) = \sum_i H_1(x_i) - H_1(x_1, \ldots, x_m) \quad (9)$$

for $i = 1, m$, is the generalization of mutual information to m dimensions. The marginal redundancy R_1',

$$R_1'(x_1, \ldots, x_{m-1}; x_m) = R_1(x_1, \ldots, x_m) - R_1(x_1, \ldots, x_{m-1}), \quad (10)$$

quantifies the information regarding x_m contained in $x_1, x_2, \ldots, x_{m-1}$. The Kolmogorov–Sinai entropy, K_1, is a measure of the mean rate of information creation by the system. For a time delay embedding $x(t) = (x_1(t), \ldots, x_m(t))$, $x_j(t) = x(t-(j-1)\tau)$, $j = 1, m$ and small time delays then approximately

$$\lim_{m \to inf} R_1'(x_1, \ldots, x_{m-1}; x_m) = H_1(x_1 - \tau K_1). \quad (11)$$

Utilizing Eq. (11), a theory of coarse-grained entropy was derived in Ref. [6] and effectively applied to chatter detection in Ref. [5].

4. Chatter identification

Sequences of cutting experiments were performed in which one cutting parameter was varied while all others were held constant. The mutual information, $I_1(x_1; x_2)$ (Eq. (6)), was computed as a function of delay for one sequence of experiments, s-1, with variable turning frequency, two sequences s-2, s-3 with variable depth of cut and one sequence, s-4, with variable feed rate. Sequences s-2 and s-3 ended in chatter while s-1 and s-4 each contained at least one chatter state.

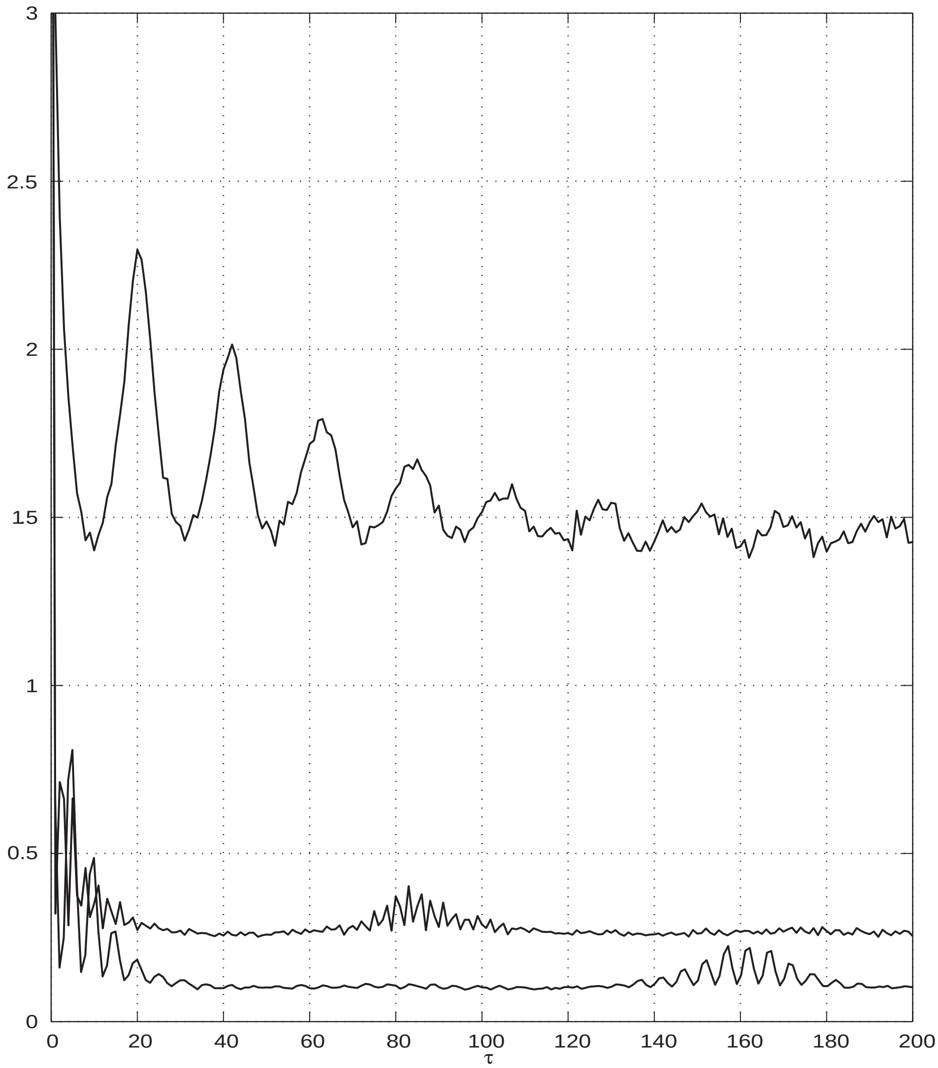

Fig. 1. Data set s-1 mutual information versus delay for acceleration in the x direction: *lower* = 360 r.p.m., *middle* = 380 r.p.m. and *upper* = 371 r.p.m.

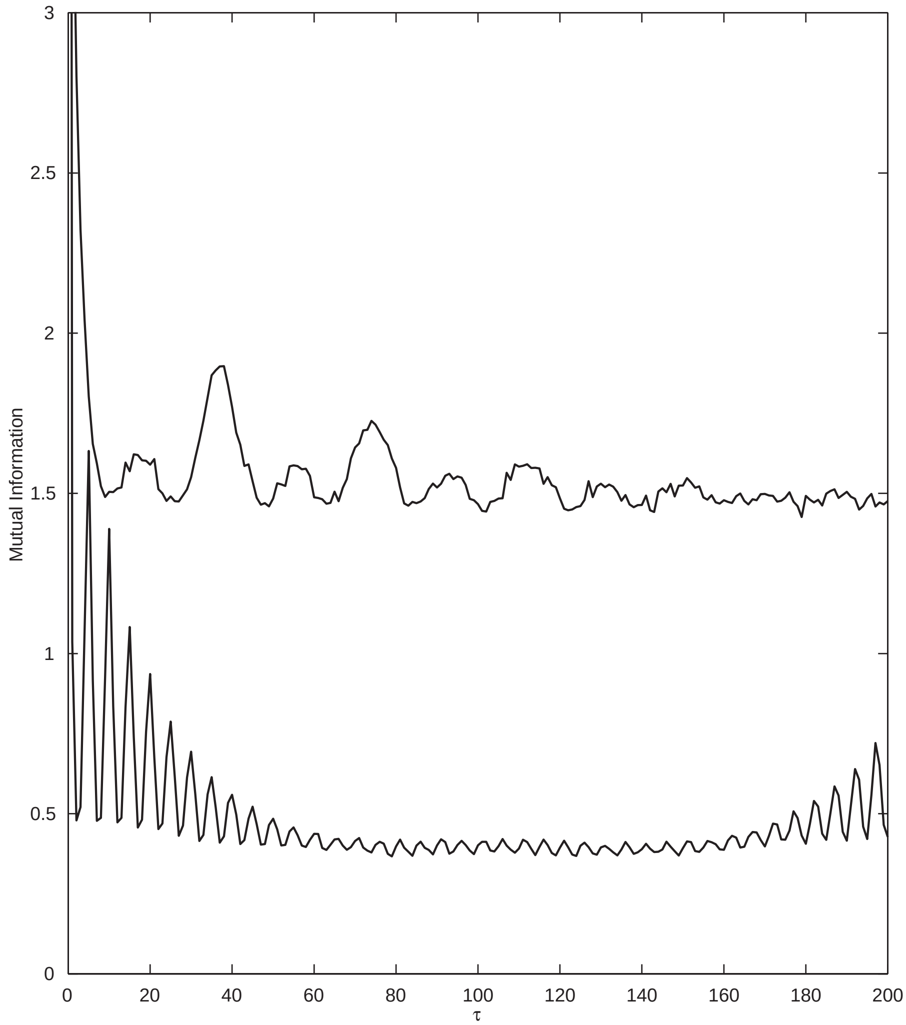

Fig. 2. Data set s-2 mutual information versus delay for acceleration in the x direction: *lower* = 2.7 mm, *upper* = 2.725 mm.

For sequence s-1, the following data applies: feed rate = 0.007 in/rev, surface speed = 90 m/min, depth of cut = 2.8 mm, resampling rate = 1024 Hz, 27 s < *time series length* < 45 s for non-chatter and 1 s for chatter, and turning frequency = 335, 360, 371, 380, 390 rpm.

Tool force and acceleration measurements were made in two orthogonal directions, x and z. In Fig. 1, I_1, computed from acceleration measurements in the x direction, versus delay is shown for the chatter state, 371 rpm and a pair of adjacent bracketing frequencies, 360 and 380 rpm. The characteristic difference in magnitude and frequency content between I_1 for the chatter and non-chatter states is clearly indicated. Since the values of I_1 for 335 and 390 rpm are less than those for 360 and 380 rpm they are omitted from the figure for clarity.

The chatter case may be identified by noting that the mutual information, I_1, for acceleration in the x direction averaged, AI_1, between $\tau = 50$ and 150, $AI_1 = 1.54$ for 371 Hz while $AI_1 = 0.088$,

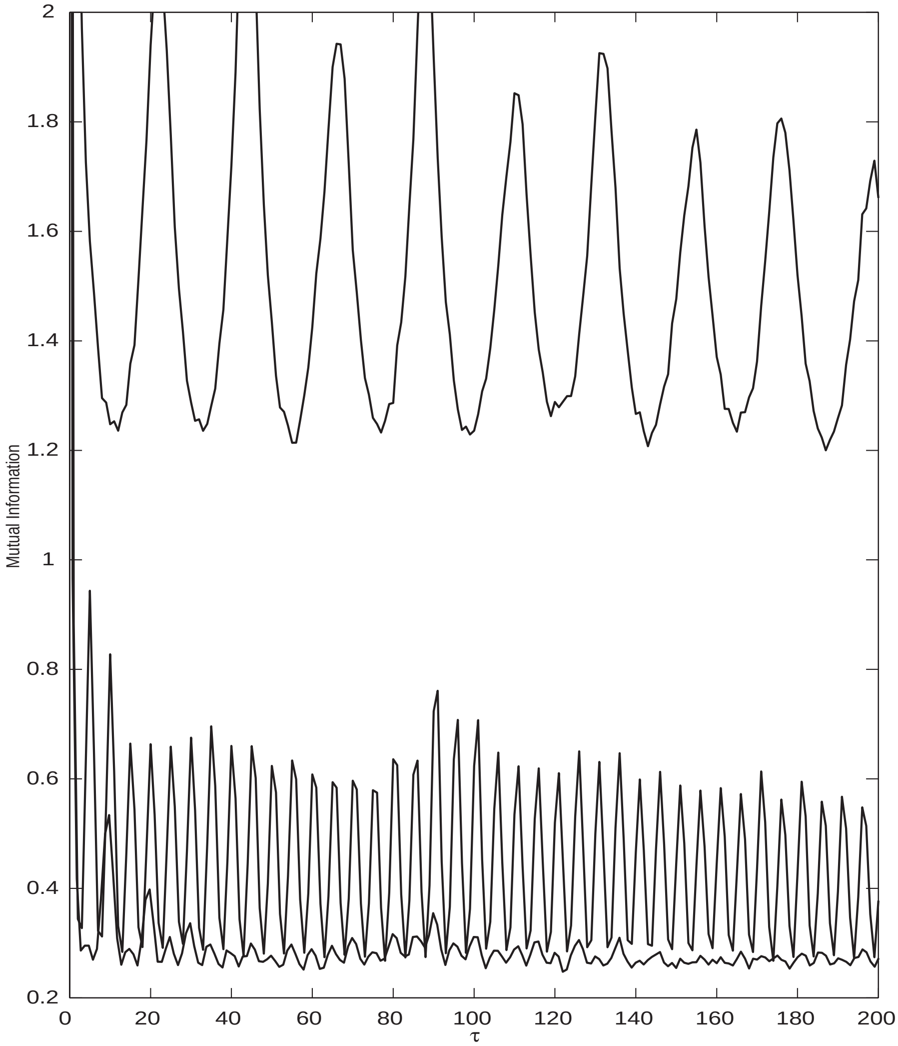

Fig. 3. Data set s-3 mutual information versus delay for acceleration in the x direction: *lower* = 2.3 mm, *middle* = 2.6 mm, *upper* = 2.8 mm.

0.092, 0.264 and 0.052 for 335, 360, 380, 390 Hz, respectively. The AI_1 for the tool forces and acceleration in the z direction was found to be 1.58 for the chatter case with a distribution, as a function of turning frequency, similar to the non-chatter AI_1 for the x direction acceleration.

The depth of cut was given values of 2.675, 2.70 and 2.725 mm in the experiments comprising set s-2. Feed rate, surface speed and resampling rate values were identical to those in s-1 while the turning frequency = 297 rpm. The time series lengths were 13 and 12 s for the non-chatter states, 2.675 and 2.70 mm, respectively and 5 s for the chatter state, 2.725 mm.

I_1 versus delay is given in Fig. 2 for the 2.70 mm non-chatter and the 2.725 mm chatter states. The 2.675 mm case has been omitted for clarity. The characteristic difference in magnitude and

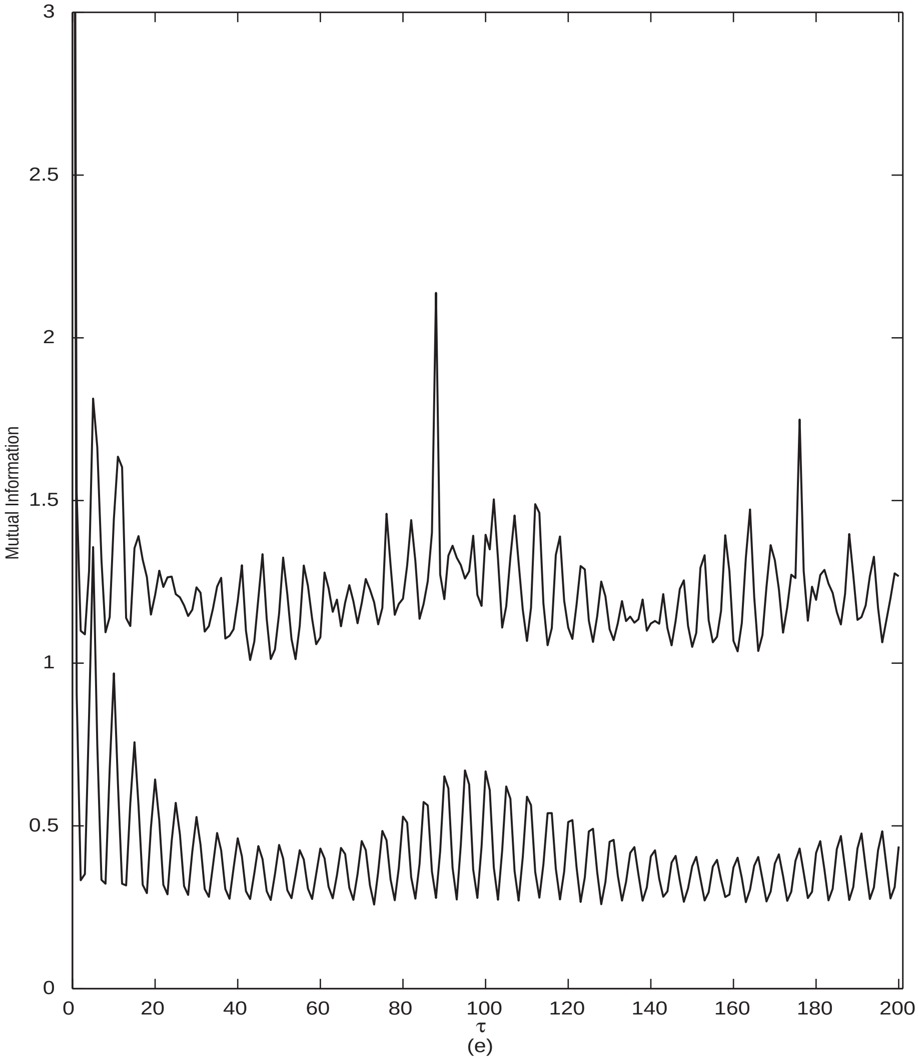

Fig. 4. Data set s-4 mutual information versus delay for acceleration in the x direction: *lower* $= 0.003$ in/rev, *upper* $=$ 0.005 in/rev.

frequency content between I_1 for the chatter and non-chatter states is evident. For the non-chatter state, 2.70 mm, $AI_1 = 0.40$ while $AI_1 = 1.56$ for the chatter state, 2.725 mm.

Set s-3 consists of five experiments, each of 5 s duration, in which the depth of cut took values of 2.3, 2.5, 2.6, 2.7 and 2.8 mm at which value chatter occurred. The turning frequency $= 708$ rpm while all other parameters were identical to those of set s-1.

Fig. 3 displays I_1 versus delay for the 2.3 and 2.6 mm non-chatter cases together with the 2.8 mm chatter case. The values of I_1 associated with the chatter and non-chatter cases display the characteristic differences in magnitude and frequency content. The I_1 for non-chatter cases exhibits an oscillatory behavior which for the 2.6 mm case has a frequency of 194 Hz.

This frequency, twice the first natural frequency of the cutting system, 97 Hz, does not correspond to any of the higher natural frequencies the smallest of which are known to be 135 and 260 Hz. It is probably associated with the phase coupling of the 97 Hz frequency with itself [9].

An explanation is suggested by the comments following Eq. (6). Consider the delay embedding of $x(t)$, $x_1(t) = x(t)$ and $x_2(t) = x(t+\tau)$. If $x_1(t)$ is independent of $x_2(t)$, then $I_1(x_1, x_2) = 0$. If $x(t)$ is periodic with period α, then with $\tau = n\alpha$, $n = 1, 2, \ldots, x_1(t) = x_2(t)$ and $I_1(x_1, x_2)$ takes a local maximum. If $x(t)$ is a symmetric, periodic function, e.g., $\sin(\omega t)$, with period α, then as before, $x_1(t) = x_2(t)$ with $\tau = n\alpha$. However, symmetry implies that $x_2(t) = C - x_1(t)$ for $\tau = \alpha/2, 3\alpha/2, \ldots$. It follows that $I_1(x_1, x_2)$ would take local maxima for $\tau = n\alpha/2$, $n = 1, 2, \ldots$. This phenomenon is observed in all of the non-chatter cases as well as the results for the periodic function given in Figs. 3(a) and 3(g) of Ref. [8].

In set s-4, the feed rate was given values between 0.003 and 0.008 in/rev. in steps of 0.001 in/rev., turning frequency of 700 r.p.m., surface speed of 90 m/min and depth of cut of 2.6 mm. I_1 versus delay is shown in Fig. 4 for the 0.003 in/rev. non-chatter case and 0.005 in/rev. chatter case, for which AI_1 are 0.366 and 1.27, respectively. Two light chatter cases were observed for which AI_1 was 0.62 and 1.04. As in the previous case of set s-3, the non-chatter cases of s-4 exhibit small values of AI_1 and oscillations which for the 0.003 in/rev. case have a frequency of 205 Hz.

5. Conclusion

Mutual information, I_1, had been shown to provide a good criterion for the choice of time delay [7]. The effects of noise on the determination of I_1 are discussed in Ref. [8]. Computational results presented here and in Ref. [9] indicated that the Fraser–Swinney algorithm [7] is sufficiently robust to yield good estimates of I_1, for noisy cutting force and tool acceleration time series. Averaged mutual information, AI_1, was relatively constant for the chatter states in sets s-1, s-2, and s-3 at 1.58, 1.56, and 1.56, respectively. For the corresponding non-chatter states, the value of AI_1 was reduced to 0.25 or less of the corresponding chatter AI_1 values.

An AI_1 value of 1.27 was associated with the chatter case in set s-4, light chatter occurred for $AI_1 = 1.04$ and 0.62 while $AI_1 = 0.25, 0.37$ and 0.62 for the non-chatter cases, respectively. The light chatter and non-chatter cases associated with $AI_1 = 0.62$ were differentiated by the frequency content of AI_1. For the data sets examined, the averaged value of the mutual information was found to provide a robust and readily computable means of distinguishing between chatter and non-chatter cutting states.

Acknowledgements

The author acknowledges the support of N.S.F. through GER-9354956 under the direction of H.R. Russell. The interest of P. Grootenhuis, Imperial College of Science and Technology, London, England, M. Rokni, Promatis Inc., H.L. Swinney, University of Texas, Austin and A.M. Fraser are appreciated.

References

[1] X.Q. Li, Y.S. Wong, A.Y.C. Nee, Comprehensive identification of tool failure and chatter using a parallel multi-ART 2 neural network, Journal of Manufacturing Science and Engineering, Transactions of the ASME 120 (2) (1998) 433–442.
[2] B. Berger, J.A. Manzari, D. Anand, C. Belai, Auto-regressive SVD algorithms, Journal of Sound and Vibration 248 (2) (2001) 351–370.
[3] W. Doug, Y.H. Joe Au, A. Mardapittas, Machine tool chatter monitoring by coherence analysis, International Journal of Research 30 (8) (1992) 1901–1924.
[4] F. Kong, J. Yu, X. Zhou, Analysis of fuzzy dynamic characteristics of machine cutting process: fuzzy stability analysis in regenerative-type-chatter, International Journal of Machine Tools and Manufacture 39 (8) (1999) 1299–1309.
[5] J. Gradisek, E. Govekar, I. Grabec, Using coarse-grained entropy rate to detect chatter in cutting, Journal of Sound and Vibration 214 (5) (1998) 941–952.
[6] M. Palus, Coarse-grained entropy rates for characterization of complex time series, Physica D 93 (1996) 64–77.
[7] A.M. Fraser, H.L. Swinney, Independent coordinates for strange attractors from mutual information, Physical Review A 33 (2) (1986) 1134–1140.
[8] A.M. Fraser, Information and entropy in strange attractors, IEEE Transactions on Information Theory 35 (2) (1989).
[9] C. Belai, Cutting State Identification, Ph.D. Thesis, University of Maryland, College Park, MD, 2002.
[10] D. Prichard, J.T. Theiler, Generalized redundancies for time series analysis, Physica D 84 (1995) 476–493.
[11] R.S. Shaw, Strange attractors, chaotic behavior and information flow, Zeitschrift für Naturforschung 36a (1981) 80–112.