Intelligent Techniques in Condition Monitoring based on Forecasting of Vibrational Signals

A. Lotfi, J. B. Hull, E. Lai and M. Howarth

Department of Mechanical and Manufacturing Engineering
The Nottingham Trent University
Burton Street, Nottingham, NG1 4BU, U.K.
Email: ahmad.lotfia@ntu.ac.uk

Dr Ahmad Lotfi *was born in Iran, in 1964. He received his B.Sc. and M.Tech. in control systems from Isfahan University of Technology, Iran and Indian Institute of Technology, India respectively. He received the Ph.D. degree in Learning Fuzzy Systems from the University of Queensland, Australia in 1995. He joined the Department of Mechanical and Manufacturing Engineering, The Nottingham Trent University as a post-doctoral research fellow in 1995 where he is currently a senior research fellow. Dr Lotfi is the author of about 40 scientific papers in the area of control and fuzzy systems. His main scientific research interest include intelligent control, learning and adaptive systems, neural networks, fuzzy logic and systems, intelligent data analysis, ultrasound technology and surface mount technology.*

Prof J B Hull *is the Head of the Department of Mechanical and Manufacturing Engineering at the Nottingham Trent University. He is the coordinator of the Ultrasound Technology Innovation Group. Prof. Hull has gained international recognition for research excellence in developing novel applications of ultrasound technology, control systems and instrumentation. He has promoted revolutionary developments in a number of research areas of significant interest to industry, including Ultrasonic Abrasive Flow Machining for Surface Finishing of Moulds and Dies and On-line Monitoring and Control of Fluid Levels in Food Containers.*

Dr Eugene Lai *is a Lecturer in the Department of Mechanical and Manufacturing Engineering at the Nottingham Trent University. Before joining the Department in 1986, he worked for Rolls Royce Limited on powerplant installation and design. He received his B.Sc. from the University of London (Queen Mary College) and PhD from the University of Cambridge (Queens College). His research interest include intelligent adaptive control, autonomous guided vehicles, condition monitoring and fluid dynamics.*

Dr Martin Howarth BSc PhD *is a Senior Lecturer in the Department of Mechanical and Manufacturing Engineering at The Nottingham Trent University. He was awarded his PhD by The Nottingham Trent University in the field of intelligent control of robotic assembly. He is a member of the department's Manufacturing Automation Research Group and leads research in the field of intelligent control using artificial neural networks and fuzzy logic techniques applied to manufacturing processes including robotic assembly and electronics manufacturing, particularly screen and stencil printing. Before joining the university he spent 6 years working in the electronics manufacturing industry and conducting research at the University of Nottingham.*

A. Lotfi

J. B. Hull

Eugene Lai

M. Howarth

SSN 1363-7681

Abstract
This paper provides an overview of the method employed to solve the problems introduced for the Second International Competition of Data Analysis by Intelligent Techniques organized by the Technology Transfer Committee of ERUDIT. The method was awarded the first prize (out of 12 different techniques) in achieving the minimum error for both categories of condition monitoring and customer classification for business application. The method employs fuzzy c-means clustering algorithm coupled with an intelligent data filtering technique for both applications. In this contribution only the problem of vibration diagnosis model for condition monitoring is addressed.

1. Introduction

The most important part in the realization of modern concepts of *predictive maintenance* is to establish a suitable diagnosis method for early detection of faults. This method of maintenance is scheduled dependent on the condition of the machine. It is essential that an early warning of the development of a fault is provided so that appropriate remedial actions are planned in advance. If predictive maintenance systems are carefully applied, industries will posses a powerful maintenance tool that can anticipate and plan for most machinery problems long before they reach the point where catastrophic failure is imminent [1].

The ability to predict failure, sufficiently in advance of the event is the fundamental premise upon which the philosophy of *condition monitoring* is based. It describes a range of equipment techniques used to acquire data related to the mechanical and operational condition of a machine. Subsequent analysis and interpretation of this data can provide information to advise on current and future maintenance.

Safety or production relevant plant components are already subject to monitoring at present and the analysis of *vibrational signals* has been established as a suitable diagnosis method for the early detection of faults. An initial vibration level taken from a machine tends to remain unchanged for a long period of time. As components start to wear or fail, the vibration level rises.

However, most commercially available machine monitoring systems are not sufficiently well suited for performing diagnostic functions, because they use only one, or very few diagnostic parameters and are not able to cope with severe fluctuations in measured signals. A significant improvement in diagnostic reliability may be achieved by simultaneous consideration of several vibrational features and supplementary processing of as many process parameters as possible, such as pressure, temperature, power etc. No competitive alternative is currently available for the application of classification and forecasting methods for automatic estimation of a technical quantity.

The process parameters and the vibration characterization are related. Thus the development of a model, *vibration diagnosis model*, relating these two quantities can be used to predict the process fault in advance. Establishing this model is not an easy task, mainly because of high dimensional feature vectors and a small data set. It is well known that these two characteristics usually result in an organized model with poor generalization power [2].

2. Problem Statement

To address the problem of data analysis for condition monitoring, based on forecasting of vibration signals, the Technology Transfer Committee of ERUDIT (European Network of Excellence in Uncertainty Techniques Developments for Use in Information Technology), Aachen, Germany, on September 1998, organized a competition (*Second International Competition of Data Analysis by Intelligent Techniques*) to compare different techniques. 2900 sets of vibration data (1400 for training and 1500 for evaluation)were given and they were collected from a turbine. Each data set consists of 23 values where the first 11 values are process parameters and the last 12 values are vibration parameters measured at different parts of the turbine.

The objective is to provide a vibration diagnosis model on the basis of the training data. The model is then applied to the evaluation of data. In this paper, a new way to cope with such tasks is presented, which has already won the first prize in the above competition by achieving the minimum error in the prediction of the vibration parameters for the evaluation data.

3. Vibration Diagnosis Model

The vibration of a mechanical system $v(t)$ is a time varying signal. It is a function of process parameters $p(t)$ and machine's condition $c(t)$. It is very likely that the vibration parameters are affected by noise $n(t)$. The above expression can be formulated as follows [3]:

$$v(t) = f(p(t), c(t), n(t), t) \qquad (1)$$

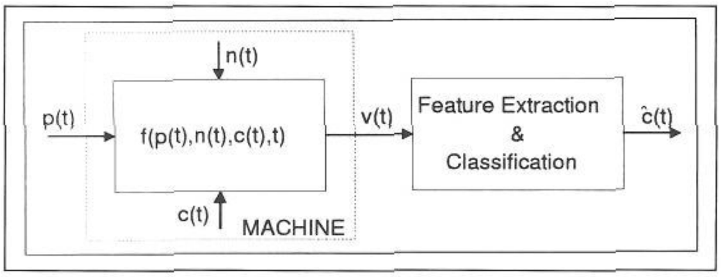

Figure 1:* *Framework for condition monitoring.

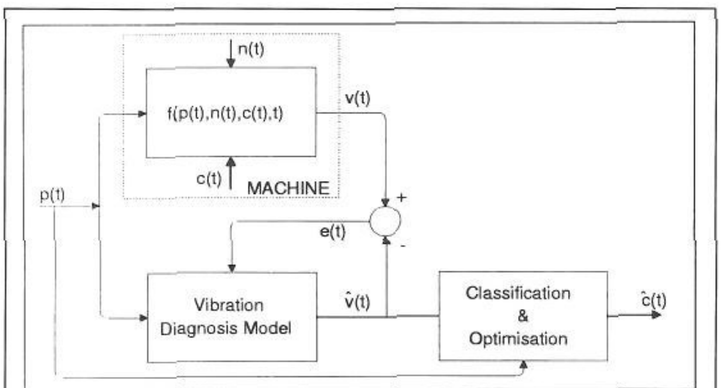

Figure 2:* *Block diagram of vibration diagnosis model.

In general the function $f(.)$ is unknown although in some cases it may be possible to estimate it. For condition monitoring purpose, the machine's condition $c(t)$ is determined from the vibration signal $v(t)$ filtering out the effect of noise. It should be noted that conditions vary at a much slower rate than the vibration i.e. $c(t)$ corresponds to a large number of vibration data points.

Therefore, to obtain the predicted condition $\hat{c}(t)$, an inverse of the function $f(.)$ is required. This problem can be split into feature extraction and classification phases [3]. Figure 1 illustrates the relationship between the condition signals, process parameters, vibration and predicted condition.

In general, the process parameters and vibration features are vectors of many elements. High dimensional feature and process parameters, restrict the use of a model representing their relationship. Establishing a *vibration diagnosis model* would help to predict the condition $\hat{c}(t)$. A block diagram of the proposed model is shown in Figure 2.

To predict the vibration features $\hat{v}(t)$, we need to establish a model which represents the relationship between the process parameters and the vibration features from a set of training data, subject to minimizing the error $e(t)$ defined below.

$$e(t) = (v(t) - \hat{v}(t))^2 \tag{2}$$

Fuzzy c-means (FCM) algorithm is employed to classify the process parameters into M classes. The classification is preceded by filtering the training data for any redundant information. The filtering technique is explained in the next section followed by the fuzzy c-means classification [4].

4. Intelligent Historical Data Generation

Consider a set of process parameter vectors:

$$\bar{p} = \{p_1(t), ..., p_k(t), ..., p_N(t)\}$$

and vibration features vector:

$$\bar{v} = \{v_1(t), ..., v_k(t), ..., v_N(t)\}$$

are given. N is the total number of data sets for training. Individual vectors representing the process parameters and vibration features are:

$$p_k(t)=[p_k^1(t), p_k^2(t), ..., p_k^Q(t)]^T$$

and

$$v_k(t)=[v_k^1(t), v_k^2(t), ..., v_k^R(t)]^T$$

respectively. Prior to establishing a model representing the relationship between the process parameters and vibration, the training data has been pre-filtered using a technique known as Intelligent Historical Data (IHD) generation method [5].

It is important to make sure that there are no similar process parameter vectors with different vibration feature vectors, as contradiction in the decision made by two similar parameters can cause an unpredictable behaviour from the model. The first step is to compare each process parameter vector with the rest of the data set. If there is any similarity between two or more vectors, their vibration vectors need to be compared. The process parameters are presented as $p_k(t)$ and the vibration vector as $v_k(t)$ where k=1, 2,... N for the training data where N is the total number of training data.

The similarity between the process parameters can be defined as the second norm of the process parameter vectors. This expression can be presented in the following form.

$$||p_k(t) - p_h(t)|| \geq \varepsilon, \quad h=1, 2, ..., N \tag{3}$$

where ε is a small number specifying the acceptable neighborhood for process parameters. If the above condition is met, it means there is no redundant or confusing data. If there are two similar process parameter vectors with similar vibration vectors, then there is redundant information in the data set and one of the process parameter vectors must be removed from the training data set. Otherwise, having two similar process parameter vectors with different vibration vectors can cause a problem.

If the condition (3) is not fulfilled for one of the data say $h = h^*$, which means there are two similar process parameters, their vibration $v_k(t)$ and the $v_h{}^*(t)$ must be compared.

$$||v_k(t) - v_h{}^*(t)|| \leq \delta \tag{4}$$

The constant δ is a small real number representing the acceptable neighborhood of the vibration feature vectors.

If the above condition is not satisfied, i.e. there are two similar process parameter vectors with completely different vibration vectors, only one of them should be used (the choice is important). However, if a combination of both is to be used, they could be combined as follows:

$$p_h{}^*(t) = (1 - \alpha)\, p_h{}^*(t) + \alpha p_k(t) \tag{5}$$

$$v_h{}^*(t) = (1 - \alpha)\, v_h{}^*(t) + \alpha v_k(t) \tag{6}$$

The *forgetting factor*, $0 < \alpha < 1$, is introduced to give less significance to the latest process parameters and vibration feature vectors.

5. Fuzzy C-Means Clustering

The pioneering work of Bezdek [6] on clustering has been shown to be a great success. The clustering algorithm so called fuzzy c-means is briefly reviewed in this section. For a good review of fuzzy c-means clustring, the reader is referred to [6,7].

The clustering of $\bar{p}$ into M clusters is a process of assigning a grade of membership to each object p_k for any cluster [6,8,9]. Fuzzy c-means assigns objects, which are described by several features, to different classes with different degrees of membership.

The fuzzy clusters can be characterized by a M x N Class Membership Function (CMF) matrix $\bar{U} = [u_k^m]$, whose entries satisfy the following conditions:

$$\sum_{m=1}^{M} u_k^m = 1, \quad k=1, 2, ..., N \tag{7}$$

$$0 < \sum_{k=1}^{N} u_k^m < Q, \quad m=1, 2, ..., M \tag{8}$$

where u_k^m is the grade of membership for p_k object in the mth cluster.

In FCM, cluster centers are determined first at the learning stage, and then the classification is made by the comparison of distances between the incoming feature and individual cluster centers. At the learning stage, cluster centers are obtained by the minimization of a cost function given below:

$$J(\bar{U},\bar{C}) = \sum_{k=1}^{N}\sum_{m=1}^{M} (u_k^m)^2 \; \|p_k - C_m\|^2 \tag{9}$$

where $\bar{C}$ = {C_1, C_2, ..., C_m,..., C_M} are M vectors of cluster centers with $C_m=[c_m^1, c_m^2, ..., c_m^Q]^T$ representing Q process parameters for the center of the mth cluster.

The following algorithm is used [10] to determine the CMFs of each object to a cluster.

1. Estimate the CMF matrix $\bar{U}$.
2. Calculate cluster centers {C_1, C_2, ...C_m, ..., C_M} using the following equation:

$$C_m = \frac{\sum_{k=1}^{N}(u_k^m)^2 p_k}{\sum_{k=1}^{N}(u_k^N)^2} \quad m = 1, 2, ..., M \tag{10}$$

3. Update the CMF matrix, $\bar{U}$ to $\bar{U}^*$ with:

$$u_k^m = \frac{1}{\sum_{r=1}^{M}\left(\frac{\|p_k - C_m\|}{\|p_k - C_r\|}\right)^2} \tag{11}$$

where

k=1, 2, ..., N; m=1, 2, ..., M

4. If control error, i.e. $\max_{k,m} |u_k^{m*} - u_k^m| \leq \eta$, stop. Otherwise substitute $\bar{U} \leftarrow \bar{U}^*$ and return to step 2.

After a number of iterations, cluster centers satisfy the minimization of the cost function, $J(\bar{U}, \bar{C})$, to a local minimum.

6. Condition Monitoring Based on Forecasting of Vibrational Signals

After pre-filtering the training data set, FCM algorithm is employed to classify the process parameters into a maximum number of inputs i.e. M=N. In this case the class centers are identical with the data set itself. The CMF for each data set is 1 for its own class and 0 for other classes. The resulting CMF is a N x N identity matrix.

On presentation of the evaluation data sets, the CMF of each evaluation data set is calculated with respect to the class centers produced from the training data. Since the class centers are identical to the training data set, the classification procedure explained here is an attempt to find out the degree of closeness of each evaluation data set to the training data set.

To predict the vibration feature vectors, $\hat{\upsilon}(t)$ based on the evaluation data, the weighted average of CMFs and their corresponding vibration feature vectors from the training data are calculated.

$$\hat{\upsilon}(t)_{evaluation} = \frac{\sum_{i=1}^{\beta} u_i \times \upsilon_i(t)}{\sum_{i=1}^{\beta} u_i} \tag{12}$$

where u_i is the CMF of evaluation data to the ith training data and is calculated using the following equation.

$$u_i = \frac{1}{\sum_{r=1}^{M}\left(\frac{\|p_{evaluation} - C_i\|}{\|p_{evaluation} - C_r\|}\right)^2} \tag{13}$$

After calculating the CMFs for evaluation data set, they are sorted in accordance to their degree of membership. It is found that using only a limited number of CMFs can produce an optimum solution. Hence, CMFs are sorted and, the vibration features for each evaluation data set is calculated using only β=15 process parameters with the highest values of CMF from the training data. The method presented here is a very fast and accurate method for a small data set in comparison with the number of inputs.

7. Illustrative Example

To illustrate the method described in this paper, a simple diagnosis model with two inputs $p=[p^1, p^2]^T$ and one output $\upsilon=[\upsilon^1]$ is considered. Five training data and one evaluation data are considered. Training data are given in Table 1 and they are depicted in Figure 3. The

p_1 (volts)	p_2 (volts)	υ_1 (volts)
0	0	0
0	1	1
1	0	1
1	1	0.02
1.05	0.95	0.2

Table 1: Training data for illustrative example.

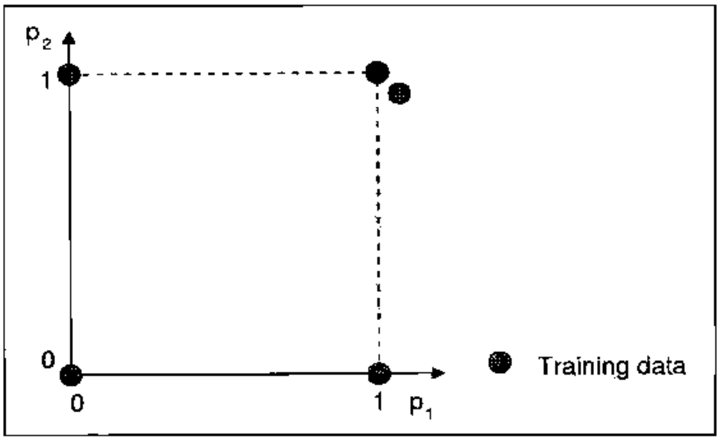

Figure 3: Training data for illustrative example.

first four rows of training data in Table 1 represents a XOR gate. The fifth training data is slightly different with the fourth row. Although they have more or less the same input values. Employing the IHD method explained in section 4 allows these two values to be merged into a new single value. The new value can be calculated using expressions (5) and (6) where α=0.9.

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 0.1 \begin{bmatrix} 1.05 \\ 0.95 \end{bmatrix} + 0.9 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (14)$$

$$0.02 = 0.1 \times 0.2 + 0.9 \times 0 \quad (15)$$

Therefore, the training data is reduced to four values as given in Table 2.

$p_1(volts)$	$p_2(volts)$	$v_1(volts)$
0	0	0
0	1	1
1	0	1
1	1	0.02

Table 2: ***Training data after employing IHD filtering.***

The next stage is to classify the training data using equations (12) and (13), the predicted value for the vibration can then be calculated. For example, given that $p_{evaluation} = [0.75, 0.8]^T$, then $u_1 = 0.06$, $u_2 = 0.10$, $u_3 = 0.12$, $u_4 = 0.71$. The predicted vibration $\hat{v}_1$ can be calculated as follows when β=3. Figure 4 shows the training and evaluation data. The predicted vibration for the evaluation data is given in Table 3.

$$\begin{aligned} \hat{v}_1 &= \frac{\sum_{i=1}^{3} u_i \times v_i}{\sum_{i=1}^{3} u_i} \quad (16) \\ &= \frac{0.71 \times 0.02 + 0.10 \times 1 + 0.12 \times 1}{0.71 + 0.10 + 0.12} \\ &= 0.25 \end{aligned}$$

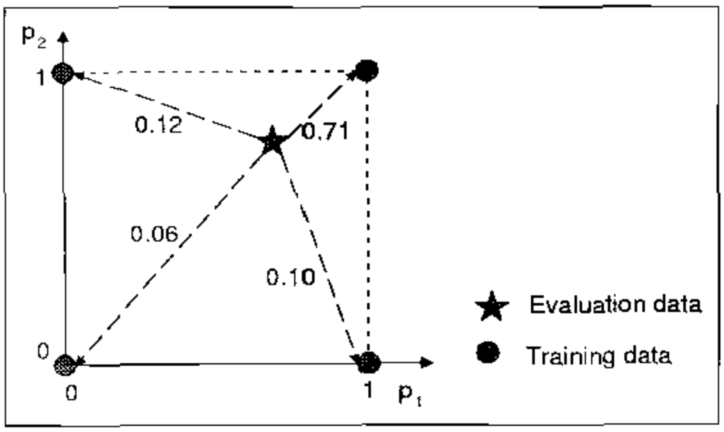

Figure 4: ***Training and evaluation data and their CMF values.***

$p_1(volts)$	$p_2(volts)$	$\hat{v}_1(volts)$
0.75	0.80	0.25

Table 3: ***Training data after employing IHD filtering.***

8. Competition Results

At a turbine, 2900 data sets were measured during a period of time covering all relevant operational modes. Each data set consists of 23 values. The first 11 values of each data set are process parameters and are probably relevant for vibration monitoring purposes. The last 12 values of each data set are vibration parameters that were measured at different parts of the turbine. The first 1400 complete data sets are used for training the model and the last 1500 data sets containing only the process parameters, are used for the evaluation of the model.

The objective is to provide a vibration diagnosis model on the basis of the training data. The evaluation data is then applied to this model. The results obtained in this way are then compared with the actual results of the evaluation data (which are known only to the organizer of the competition). The criterion used to evaluate the results is given below.

$$Error = \frac{1}{12} \sum_{j=1}^{12} \sum_{i=1401}^{2900} (v_i^j(t) - \hat{v}_i^j(t))^2 \quad (17)$$

The distribution of percentage error for the first and twelfth sets of evaluation data are depicted in Figures 5 and 6. The maximum deviation is less than 12% for all valuation data.

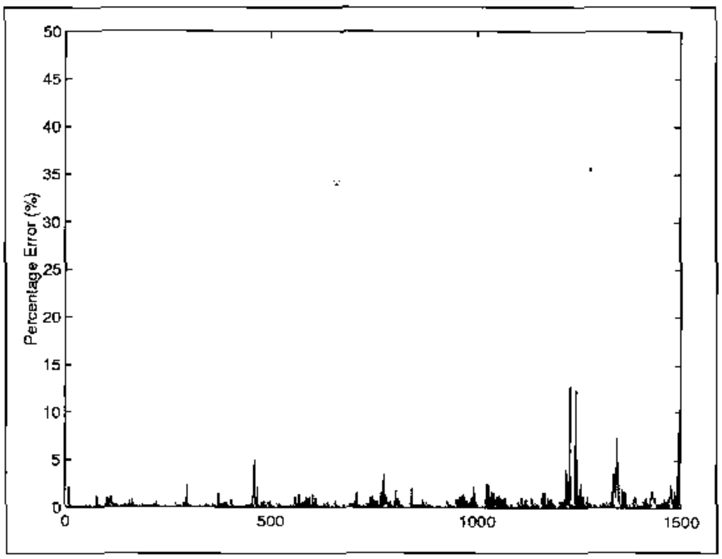

Figure 5: ***Percentage error for the first set of evaluation data.***

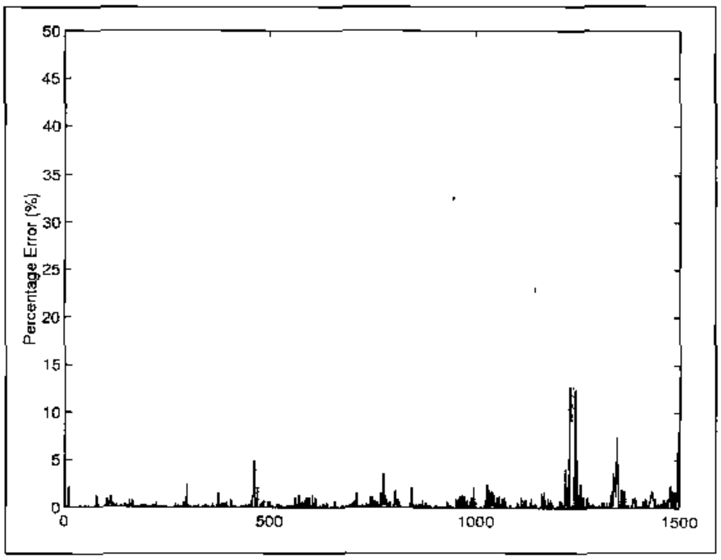

Figure 6: ***Percentage error for the twelfth set of evaluation data.***

9. Acknowledgment
The first author would like to thank the Technology Transfer Committee of ERUDIT for organizing the Second International Competition of Data Analysis by Intelligent Techniques.

References

[1] B. K. N. Rao. *Handbook of Condition Monitoring*. Elsevier Advanced Technology, Oxford, 1st edition, 1996.

[2] X. Wenhua and F. Kai. An intelligent diagnostic system for reciprocating machine. In *Proceedings of IEEE International Conference on Intelligent Processing Systems*, volume 2, pages 1520—1522, China, October 1997.

[3] A. C. McCormick, A. K. Nandi and L. Jack. Digital signal processing algorithms in condition monitoring. *Int. Journal of Condition Monitoring and Diagnostic Engineering Management*, 1(3):5—14, 1998.

[4] S. Marsili-Libelli. Adaptive fuzzy monitoring and fault detection. *Int Journal of Condition Monitoring and Diagnostic Engineering Management*, 1(3):31—37, 1998.

[5] A.Lotfi and M.Howarth. Intelligent learning control of stencil printing. *Proc. Of World Automation Congress*, Vol. 8, 323—328, Anchorage, Alaska, USA., 1998.

[6] J. C. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum, New York, 1981.

[7] J. S. R. Jang, C. T. Sun, and E. Mizutani. Neuro-Fuzzy and Soft Computing, A Computational Approach to Learning and Machine Intelligence, *Prentice Hall*, NJ, 1997.

[8] S. K. Pal and S. Mitra. Multilayer perceptron, fuzzy sets, and classification. *IEEE Trans. on Neural Networks*, 3(5):683—697, 1992.

[9] L. X. Wang. Training of fuzzy logic systems using nearest neighborhood clustering. *Proc. of the Second IEEE Int. Conf. on Fuzzy Systems*, pages 13—17, 1993.

[10] Z. Wang, W. Lowrenz, Raj B. K. N. Rao, and T. Hope. Feature-filtered fuzzy clustering for condition monitoring of tool wear. *Journal of Intelligent Manufacturing*, 7:13—22, 1996.