Tool Condition Monitoring using the TSK Fuzzy Approach based on Subtractive Clustering Method

Qun Ren[1], Marek Balazinski[1], Luc Baron[1], Krzysztof Jemielniak[2]

[1] Mechanical Engineering Department, École Polytechnique de Montréal
C.P. 6079, succ. Centre-Ville, Montréal, Québec, Canada, H3C 3A7
{qun.ren, marek.balazinski, luc.baron}@polymtl.ca
[2] Faculty of Production Engineering, Warsaw University of Technology
Narbutta 86, 02-524 Warsaw, Poland
k.jemielniak@ wip.pw.edu.pl

Abstract. This paper presents a tool condition monitoring approach using Takagi-Sugeno-Kang (TSK) fuzzy logic incorporating a subtractive clustering method. The experimental results show its effectiveness and satisfactory comparisons with several other artificial intelligence methods.

Keywords: tool condition monitoring, TSK fuzzy logic, subtractive clustering

1 Introduction

Tool condition has a strong influence on the resulting surface finish and dimensional integrity of the workpiece, as well as vibration levels of the machine tool. The information obtained from tool wear monitoring can be used for several purposes that include: establishing tool change policy, economic optimization of machining operations, compensating for tool wear on-line and to some extent avoiding catastrophic tool failures [1]. Effective monitoring of a manufacturing process is essential for ensuring product quality and reducing production costs. Analysis, implementation and evaluation of machining processes present significant challenges to the manufacturing industry.

Cutting force measurement, currently the most reliable and accurate sensing method available in metal cutting, is one of the most commonly employed methods for on-line tool wear monitoring. It is frequently applied in turning processes because cutting force values are more sensitive to tool wear than other measurements such as vibration or acoustic emission [2].

In a tool condition monitoring system, real-time data are acquired from sensors located at different locations on the workpiece, tool and machine-tool, then a signal processing technique is used to extract valid data. A decision making system is then used to analyse the data and classify the results to make a reliable estimate of the state of the tool and consequently of the machined parts themselves [3].

Advanced signal processing techniques and artificial intelligence play a key role in the development of modern tool condition monitoring systems [4]. The most frequently chosen methods are neural network (NN) [5], Mamdani fuzzy logic (FL)

[6], [7], or a combination of NN with either Mamdani FL [8] or an automatic generating method, i.e., genetic algorithm (GA) [9]. All these methods have a similar objective – matching the estimate of average cutting tool wear with the directly measured wear value.

The aim of this paper is to present an effective tool wear monitoring method using the Takagi-Sugeno-Kang (TSK) fuzzy approach incorporating a subtractive clustering method [10] to accomplish the integration of multi-sensor information and tool wear information. It generates fuzzy rules directly from the input-output data acquired from sensors, and provides high accuracy and high reliability of the tool wear prediction over a range of cutting conditions.

This paper is divided into four sections. Section I contains tool wear monitoring development and some introductory remarks. Section II recalls the initial theoretical foundation: TSK fuzzy logic system (FLS), subtractive clustering method and least-square estimation. Section III is a specific turning case study. The experimental results show the effectiveness and advantages of the TSK fuzzy approach compared with other different artificial intelligence methods – NN, Mamdani FL and a neural network based fuzzy system (NF). Section IV contains concluding remarks and future research recommendations.

2 Theoretical Foundation

The proposed linguistic approach by Zadeh [11, 12], following the first "Fuzzy Sets" paper in 1965 [13], is effective and versatile in modeling ill-defined systems with fuzziness or fully-defined systems with realistic approximations. Later it is expanded into fuzzy systems modeling as a qualitative modeling approach. Qualitative modeling has the capability to model complex system behavior in such a qualitative way that the model is more effective and versatile in capturing the behavior of ill-defined systems with fuzziness or fully defined system with realistic approximation. In the literature, different modeling techniques can be found, and TSK FLS [14], [15] has attracted much attention.

2.1 TSK Fuzzy Logic System

TSK FLS was proposed in an effort to develop a systematic approach to generate fuzzy rules from a given input-output data set. This model consists of rules with fuzzy antecedents and a mathematical function in the consequent part. Usually the conclusion function is in the form of a dynamic linear equation [14], [15]. The antecedents divide the input space into a set of fuzzy regions, while consequents describe behaviours of the system in those regions. The main difference with more traditional [16] (Mamdani FL) fuzzy rules is that the consequents of the rules are a function of the values of the input variables. TSK FLSs are widely used for model-based control and model-based fault diagnosis. This is due to the model's properties; on one hand being a general non-linear approximator that can approximate every continuous mapping, and on the other hand being a piecewise linear model that is relatively easy to interpret [17] and whose linear sub-models can be exploited for control and fault detection [18].

A generalized type-1 TSK model can be described by fuzzy IF-THEN rules which represent input-output relations of a system. For an MISO first–order type-1 TSK model, its *k*th rule can be expressed as:

$$\textbf{IF}\ \ x_1\ is\ Q_{1k}\ \textbf{and}\ x_2\ is\ Q_{2k}\ \textbf{and} \ldots \textbf{and}\ x_n\ is\ Q_{nk},$$
$$\textbf{THEN}\ \ Z\ is\ w^k = p_0^k + p_1^k x_1 + p_2^k x_2 + \ldots + p_n^k x_n \quad \textbf{(1)}$$

where x_1, x_2 …, x_n and *Z* are linguistic variables; Q_1^k, Q_2^k, …, and Q_n^k are the fuzzy sets on universe of discourses *U*, *V*, …, and *W*, and p_0^k, p_1^k, p_2^k, …, p_n^k are regression parameters.

2.2 Subtractive Clustering Method

The structure of a fuzzy TSK model can be done manually based on knowledge about the target process or using data-driven techniques. Identification of the system using clustering involves formation of clusters in the data space and translation of these clusters into TSK rules such that the model obtained is close to the system to be identified.

The aim of Chiu's subtractive clustering identification algorithm [10] is to estimate both the number and initial location of cluster centers and extract the TSK fuzzy rules from input/output data. Subtractive clustering operates by finding the optimal data point to define a cluster centre based on the density of surrounding data points. This method is a fast clustering method designed for high dimension problems with a moderate number of data points. This is because its computation grows linearly with the data dimension and as the square of the number of data points. A brief description of Chiu's subtractive clustering method is as follows:

Consider a group of data points $\{x^1, x^2, \ldots, x^w\}$ for a specific class. The M dimensional feature space is normalized so that all data are bounded by a unit hypercube.

Then calculate potential P_i for each point as follows:

$$P_i = \sum_{j=1}^{w} e^{-\alpha \left\| x^i - x^j \right\|^2} \quad \textbf{(2)}$$

with $\alpha = 4/r_a^2$ and r_a is the hypersphere cluster radius. Data points outside r_a have little influence on the potential. $\|.\|$ denotes the Euclidean distance.

Thus, the measure of potential for a data point is a function of the distance to all other data points. A data point with many neighboring data points will have a high potential value. After the potential of every data point is computed, suppose $k = 1$ where k is a cluster counter. The data point with the maximum potential P_k^*

with $P_k^* = P_1^*$ is selected as the first cluster center $x_k^* = x_1^*$. Then the potential of each data point x^i is revised using the formula

$$P_i = P_i - P_k^* e^{-\beta\left\|x^i - x_1^*\right\|^2} \quad \textbf{(3)}$$

with $\beta = 4/r_b^2$ and r_b is the hypersphere penalty radius. Thus, an amount representing the potential of each data point is subtracted as a function of its distance from x_1^* .

More generally, when the *k*th cluster center x_k^* has been identified, the potential of all data is revised using the formula:

$$P_i = P_i - P_k^* e^{-\beta\left\|x^i - x_k^*\right\|^2} \quad \textbf{(4)}$$

When the potential of all data points has been revised using (4), the data point x^t with the highest remaining potential is chosen as the next cluster center. The process of acquiring a new cluster center and revising potentials uses the following criteria:

if $P_t > \bar{\varepsilon} P_1^*$ ($\bar{\varepsilon}$ is accept ratio)

Accept x^t as the next cluster center, cluster counter $k = k + 1$, and continue.

else if $P_t < \underline{\varepsilon} P_1^*$ ($\underline{\varepsilon}$ is reject ratio)

Reject x^t and end the clustering process

else

Let $d_{\min}$ = shortest of the distances between x^t and all previously found cluster centers.

if $\frac{d_{\min}}{r_a} + \frac{P_t}{P_1^*} \geq 1$

Accept x^t as the next cluster center. Cluster counter $k = k + 1$, and continue.

else

Reject x^t and set $P_t = 0$.

Select x^t with the next highest potential as the new candidate cluster center and retest.

end if

end if

The number of clusters obtained is the number of rules in the TSK FLS. Because Gaussian basis functions (GBFs) have the best approximation property [19], Gaussian functions are chosen as the MFs. A Gaussian MF can be expressed by the following formula for the *v*th variable:

$$Q_v^k = \exp\left[-\frac{1}{2}\left(\frac{x_v - x_v^{k*}}{\sigma}\right)^2\right] \quad \textbf{(5)}$$

where x_v^{k*} is the mean of the vth input feature in the kth rule for $v \in [0, n]$. The standard deviation of Gaussian MF σ is given as

$$\sigma = \sqrt{1/2\alpha} \quad \textbf{(6)}$$

2.3 Least Square Estimation

For the first order model presented in this paper, the consequent functions are linear. In the method of Sugeno and Kang [15], least-square estimation is used to identify the consequent parameters of the TSK model, where the premise structure, premise parameters, consequent structure, and consequent parameters were identified and adjusted recursively. In a TSK FLS, rule premises are represented by an exponential membership function. The optimal consequent parameters $p_0^k, p_1^k, p_2^k, \ldots, p_n^k$ (coefficients of the polynomial function) in (1) for a given set of clusters are obtained using the least-square estimation method.

When certain input values $x_1^0, x_2^0, \ldots, x_n^0$ are given to the input variables $x_1, x_2, \ldots, x_n$, the conclusion from the kth rule (1) in a TSK model is a crisp value w^k:

$$w^k = p_0^k + p_1^k x_1^0 + p_2^k x_2^0 + \ldots + p_n^k x_n^0 \quad \textbf{(7)}$$

with a certain rule firing strength (weight) defined as

$$\alpha^k = \mu_1^k(x_1^0) \cap \mu_2^k(x_2^0) \cap \ldots . \cap \mu_n^k(x_n^0) \quad \textbf{(8)}$$

where $\mu_1^k(x_1^0)$, $\mu_2^k(x_2^0)$, …, $\mu_n^k(x_n^0)$ are membership grades for fuzzy sets Q_1^k, Q_2^k, …, and Q_n^k in the kth rule. The symbol $\cap$ is a conjunction operator, which is a T-norm (the minimum operator $\wedge$ or the product operator $*$).

Moreover, the output of the model is computed (using *weighted average aggregation*) as

$$w = \frac{\sum_{k=1}^{m} \alpha^k w^k}{\sum_{k=1}^{m} \alpha^k} \quad \textbf{(9)}$$

Suppose

$$\beta^k = \frac{\alpha^k}{\sum_{k=1}^{m} \alpha^k} \quad \textbf{(10)}$$

Then, (9) can be converted into a linear least-square estimation problem, as

$$w = \sum_{k=1}^{m} \beta^k w^k \quad \textbf{(11)}$$

For a group of λ data vectors, the equations can be obtained as

$$
\begin{aligned}
w^1 &= \beta_1^1(p_0^1 + p_1^1 x_1 + p_2^1 x_2 + \ldots + p_n^1 x_n) + \beta_1^2(p_0^2 + p_1^2 x_1 + p_2^2 x_2 + \ldots + p_n^2 x_n) + \cdots \\
&\quad + \beta_1^m(p_0^m + p_1^m x_1 + p_2^m x_2 + \ldots + p_n^m x_n)) \\
w^2 &= \beta_2^1(p_0^1 + p_1^1 x_1 + p_2^1 x_2 + \ldots + p_n^1 x_n) + \beta_2^2(p_0^2 + p_1^2 x_1 + p_2^2 x_2 + \ldots + p_n^2 x_n) + \cdots \\
&\quad + \beta_2^m(p_0^m + p_1^m x_1 + p_2^m x_2 + \ldots + p_n^m x_n)) \\
&\vdots \\
w^\lambda &= \beta_\lambda^1(p_0^1 + p_1^1 x_1 + p_2^1 x_2 + \ldots + p_n^1 x_n) + \beta_\lambda^2(p_0^2 + p_1^2 x_1 + p_2^2 x_2 + \ldots + p_n^2 x_n) + \cdots \\
&\quad + \beta_\lambda^m(p_0^m + p_1^m x_1 + p_2^m x_2 + \ldots + p_n^m x_n))
\end{aligned}
\tag{12}
$$

These equations can be represented as:

$$
\begin{bmatrix}
\beta_1^1 x_1 & \beta_1^1 x_2 \cdots \beta_1^1 x_n & \beta_1^1 & \cdot & \cdot & \cdot & \beta_1^m x_1 & \beta_1^m x_2 \cdots \beta_1^m x_n & \beta_1^m \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\beta_\lambda^1 x_1 & \beta_\lambda^1 x_2 \cdots \beta_\lambda^1 \mu_{n1} & \beta_\lambda^1 & \cdot & \cdot & \cdot & \beta_\lambda^m x_1 & \beta_\lambda^m x_2 \cdots \beta_\lambda^m x_n & \beta_\lambda^m
\end{bmatrix}
\begin{bmatrix}
p_1^1 \\ p_2^1 \\ \vdots \\ p_n^1 \\ p_0^1 \\ \cdot \\ \cdot \\ p_1^m \\ p_2^m \\ \vdots \\ p_n^m \\ p_0^m
\end{bmatrix}
=
\begin{bmatrix}
w^1 \\ w^2 \\ \cdot \\ \cdot \\ w^3
\end{bmatrix}
\tag{13}
$$

Using the standard notation $AP = W$, this becomes a least square estimation problem where A is a constant matrix (known), W is a matrix of output values (known) and P is a matrix of parameters to be estimated. The well-known pseudo-inverse solution that minimizes $\|AP - W\|^2$ is given by

$$P = (A^T A)^{-1} A^T W \tag{14}$$

3 Case Study

The experiments described in this paper were conducted on a conventional lathe TUD-50. A CSRPR 2525 tool holder equipped with a TiN-Al_2O_3-TiCN coated sintered carbide insert SNUN 120408 was used in the test. To simulate factory floor conditions, six sets of cutting parameters were selected and applied in sequence as presented in Fig. 1. During machining, the feed force (F_f) and the cutting force (F_c) were recorded while the tool wear was manually measured after each test.

For our purposes tool wear (*VB*) was estimated from three input sources: f, F_f and F_c. The choice of input variables was based on the following two observations: F_f is independent of f, but rather depends on *VB* and the depth of cut, denoted a_p. Moreover, F_c depends on a_p and f, while being only weakly dependent on *VB*. So, in

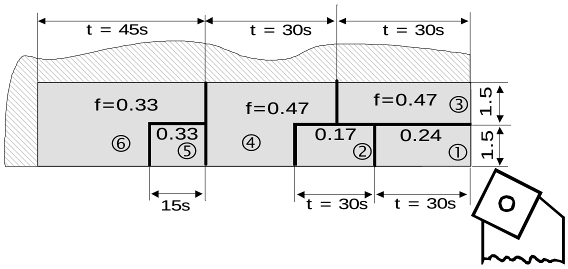

Fig. 1. Cutting parameters used in experiments, where f is feed rate and t is time.

this paper f and the measurement F_c are used to determine a_p , and the measurement F_f is used to determine *VB* without requiring a_p as an input variable.

Cutting speed of each cut was selected to ensure approximately the same share in the tool wear. *VB* was measured after carrying out each sequence. The value for F_f and F_c were measured corresponding to a single cut using a Kistlter 9263 dynamometer during 5-s intervals while the cut was executed. Recent research has attempted to investigate the application of multiple sensors with complementing characteristics to provide a robust estimate of tool wear condition. Since the inserts used in the experiments had a soft, cobalt-enriched layer of substrate under the coating, the tool life had a tendency to end suddenly after this coating wore through.

The experiments were carried out until a tool failure occurred. Two experiments were carried out until a tool failure occurred. In the first tests (designated W5) 10 cycles were performed until a sudden rise of the flank wear VB occurred, reaching approximately 0.5 mm. In the second test (designated W7) failure of the coating resulted in chipping of the cutting edge at the end of 9th cycle. W5 was devised for TSK fuzzy rule identification, while W7 was used to verify the performance of the different monitoring system. Fig. 2 presents the cutting force components F_c and F_f versus *VB* obtained in the experiments.

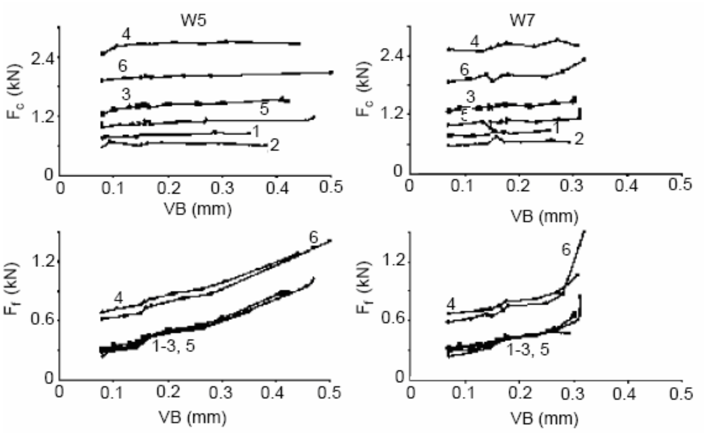

Fig. 2 Cutting force components F_c and F_f versus tool wear *VB* obtained in the experiments with six sets of cutting parameters shown in Fig.1

By using TSK fuzzy approach, a six rule TSK fuzzy model can be used to describe the tool wear condition with cutting feed, cutting force and feeding force as input variables. Table 1 lists the six cluster centers obtained by subtractive clustering method learning from the first experiment W5.

Table 1 Six cluster centers obtained by using subtractive clustering with the four parameters initialized as $r_a = 0.25$, $\underline{\varepsilon} = 0.60$, $\bar{\varepsilon} = 1$, $\eta = 0.25$

Cluster	f (mm)	F_c (N)	F_f (N)	*VB*(mm)
1	0.47	1397	389	0.145
2	0.47	1392	442	0.165
3	0.33	1044	405	0.158
4	0.33	1051	349	0.135
5	0.47	1332	347	0.102
6	0.47	1455	508	0.2

Figure 3 summarizes results of tool wear conditioning from W5 (learning) and W7 (testing) and compares them with several different artificial intelligence methods described in [4], [9] applied to the same experimental arrangements.

For these four AI methods, the quality of the tool wear estimation was evaluated using root-mean-square-error (rmse):

$$rmse = \sqrt{\sum (VB_m - VB_e)^2 / N} \qquad (15)$$

and maximum error (max):

$$\max = \max(VB_m - VB_e) \qquad (16)$$

where VB_m, VB_e are measured and estimated flank wear respectively, and N is the number of patterns in the set (N = 71 for the experiment W5 and N = 66 for the experiment W7).

From Table 2, the TSK fuzzy approach has the lowest root-mean-square-error and the smallest maximum error.

Table 2 Summary of root-mean-square-error (rmse) and maximum error (max) form the experimental results with different AI methods

AI method	Learning (W5)		Testing (W7)	
	rmse (mm)	max (mm)	rmse (mm)	max (mm)
Neural Network	0.015	0.036	0.029	0.081
FDSS (Mamdani FL)	0.024	0.068	0.034	0.056
NF	0.014	0.030	0.030	0.081
TSK FL	0.011	0.023	0.015	0.037

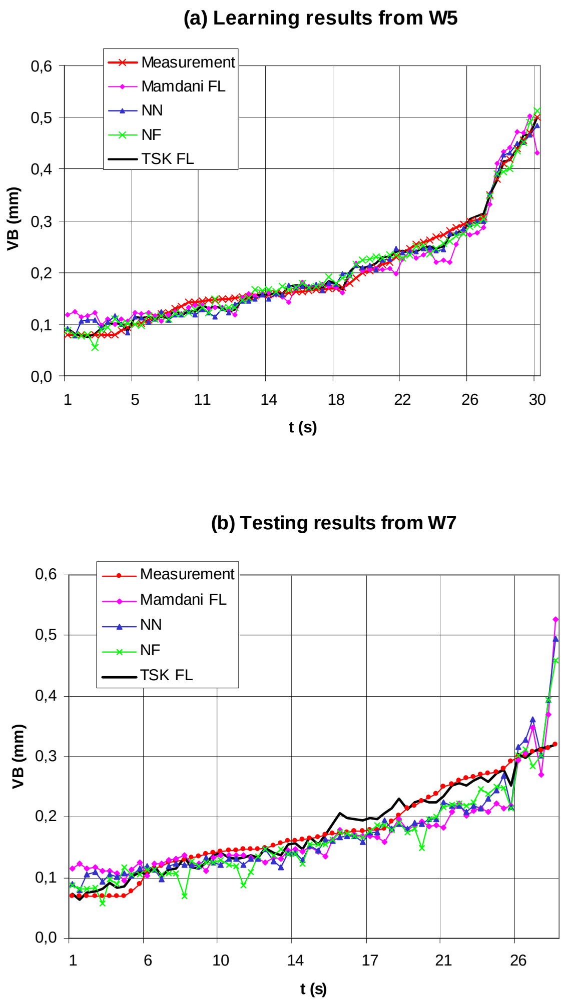

Fig. 3 Tool wear monitoring using different artificial intelligent methods: Mamdani FL, neural network (NN), neural network based fuzzy system (NF) and TSK fuzzy approach.

The TSK fuzzy modeling program used for tool condition monitoring in this paper was developed by Geno-flou development group in Lab CAE at École Polytechnique de Montréal.

4 Conclusion

A TSK fuzzy approach using subtractive clustering is described in detail in this article. It generates fuzzy rules directly from the input-output data acquired from sensors and provides high accuracy and high reliability of the tool wear prediction over a range of cutting conditions. The experimental results show its effectiveness and a satisfactory comparison with several other artificial intelligence methods.

References

1. El Gomayel, J. I., Bregger K. D.: On-line Tool, Trans. of ASME. J. of Eng. in Industry, 108 (1), 4447 (1988)
2. Du R., Elbestawi M. A., Wu S. M.: Automated monitoring of manufacturing processes, Part 1 and Part 2, J. of Eng. In Industry, 117, 121-141 (1995)
3. Balazinski M., Bellerose M., Czogala E.: Application of fuzzy logic techniques to the selection of cutting parameters in machining processes, Fuzzy Sets and Systems, 61, 301–317 (1994)
4. Balazinski M., Czogala E., Jemielniak K., Leski J.: Tool condition monitoring using artificial intelligence methods, Eng. App. of Art. Int., 15, 73–80 (2002)
5. Jemielniak K., Kwiatkowski L., Wrzosek P.: Diagnosis of tool wear based on cutting forces and acoustic emission measures as inputs to neural network, J. of Int. Man., 9, 447–455 (1998)
6. Balazinski M., Jemielniak K.: Tool conditions monitoring using fuzzy decision support system", VCIRP, AC'98 Miedzeszyn, 115–122 (1998)
7. Baron L., Archiche S., Balazinski M.: Fuzzy decisions system knowledge base generation using a genetic algorithm, Int. J. of App. Reasoning, 28 (2-3), 125-148 (2001)
8. Li, X. L., Li H. X., Guan X. P, Du R.: Fuzzy estimation of feed-cutting force from current measurement-a case study on intelligent tool wear condition monitoring, IEEE Trans. on Sys., Man and Cyb., Part C (Applications and Reviews), 34 (4), 506-512 (2004)
9. Achiche S., Balazinski M., Baron L., Jemielniak K.: Tool wear monitoring using genetically-generated fuzzy knowledge bases, Eng. App. of Art. Int., 15 (3-4), 303-314, (2002)
10. Chiu S. L.: Fuzzy Model Identification Based on Cluster Estimation, J. on Int. Fuzzy Sys., v 2, 267–278 (1994)
11. Zadeh L. A.: Fuzzy Algorithm, Inf. and Contr., 12, 94-102 (1968)
12. Zadeh L. A.: Outline of A New Approach to the Analysis of Complex Systems and Decision Processes, IEEE Trans. on Sys., Man, and Cyb., 3, 28-44 (1973)
13. Zadeh L. A., Fuzzy Sets, Inf. and Cont., 8, 338-353 (1965)
14. Takagi T., Sugeno M., Fuzzy Identification of Systems and Its Applications to Modelling and Control, IEEE Trans. on Sys., Man, and Cyb., 15 **(1),** 116-132 (1985)
15. Sugeno M. and Kang G.: Structure Identification of Fuzzy Model, Fuzzy Sets and Systems, 28 (1), 15-33 (1988)
16. Mamdani E. H., Assilian S., Applications of Fuzzy Algorithms for Control of Simple Dynamic Plant, Proc. Inst. Elec. Eng., 121, 1585-1588 (1974)
17. Johansen T., Foss B.: Identification of Non-Linear System Structure and Parameters Using Regime Decomposition, Automatic, 31 (2), 321–326 (1995)
18. Füssel D., Ballé P., Isermann R., Closed Loop Fault Diagnosis Based on A Nonlinear Process Model and Automatic Fuzzy Rule Generation, 4th IFAC Sym. on Fault Det., Sup. and Safety for Tech.l Proc., (1997)
19. Wang L. –X and Mendel J. M.: Fuzzy basis functions, universal approximation, and orthogonal least-squares learning, IEEE Trans. on Neural Networks, 3, 807-813 (1992)