5th WSEAS Int. Conf. on WAVELET ANALYSIS and MULTIRATE SYSTEMS, Sofia, Bulgaria, October 27-29, 2005 (pp24-30)

On the definition of functioning conditions of a mechanical system by means of orthogonal processing

VINCENZO NIOLA
Department of Mechanical Engineering for Energetics
University of Naples "Federico II"
Via Claudio, 21 - 80125 Napoli
ITALY

ANTONIO DONNARUMMA
Faculty of Engineering
University of Salerno
84084 Fisciano (SA)
ITALY

GIUSEPPE QUAREMBA
University of Naples "Federico II"
Via S. Pansini, 5 – 80131 Napoli
ITALY

Abstract: - In the present paper several statistical and fuzzy methods for identifying anomalies in mechanical systems have been proposed. In order to perform the study, more than 1000 tests, each with predefined characteristics and goals, have been carried out by means of dynamical test-bed based on a two circular-arc cam-follower mechanism. The signals, composed of acceleration of the follower and the applied torque, sampled electronically, were grouped into 4 main families and, for each family, into 4 groups depending to their features. All the signals were decomposed, by means of discrete wavelet transform, into 10 orthogonal components. The energetic content and many statistical parameters were calculated for each component.
Afterward, the result of their classification performed by a multivariate statistical analysis (i.e., discriminant analysis) was compared to the one obtained by applying a fuzzy algorithm.

Key-Words: - Fuzzy logic, fuzzy classification, robust statistics, complex signal processing, wavelet analysis, diagnostics.

1 Introduction

A cam is mechanical element, which is used to transmit a desired motion to another mechanical element by direct contact. Specifically, the purpose of the cam is the transmission of power, motion or information. Usually, a cam is composed of three different parts: a driving element called itself cam, a driven element called follower and a fixed frame. Cam mechanisms are usually used in most modern applications, especially in automatic machines and instruments, internal combustion engines and control systems. Generally, the design of cam profile is based on well note simple regular curves such as circles, parabolas cycloids, sinusoidal or trapezoidal curves, polynomial functions and Fourier series curves.

In the recent literature, many studies have been addressed to circular-arc cams [1]. The motion equation of an equivalent system model of an automotive valve train was studied [2].

On the other side, the Wavelet Transformation (WT) represents a time-scale analysis of the smoothness of a signal or, more in general, a time series of a curve profile [3]. The Wavelet analysis, unlike the Fourier one, is very useful when the goal is the analysis and decomposition of a signal showing a not constant frequency [4]. Let us consider the simple case in which we want to find the Fourier expansion of a signal, defined from 0 to 2, that assumes a linear form from 0 to 1 and it is

sinusoidal from 1 to 2. In this case, in order to obtain an appraisable approximation of the signal, we must evaluate many coefficients of the Fourier expansion.
Qualitatively, the difference between the usual sine wave and a wavelet can be described from the localization property: the sine wave is localized in frequency domain, but not in time domain, while a wavelet is localized both in the frequency and time domain. Furthermore, the duration of its maximum oscillation is relatively small. One can regard a wavelet as a shape of wave of limited duration and zero moments of a given order. The choice of a wavelet and of signal decomposition level depends on the shape of signals and on the experience of the analyst.
For its versatility, the wavelet analysis is diffused in many fields, such as Acoustics, Electrodynamics [5], Finance [6], Medicine and Statistics [7]. Furthermore, in order to investigate the anomalies in a vibrating system, the methodology of wavelet analysis was proposed in [8]. In the present paper, we study the acceleration of both the follower and torque, sampled by a specific electronic instrumentation, in several conditions of functioning.
Consider that the response is also due to the smoothness of the cam profile, which is composed of subsets of circular arcs as explained, in more details, in the following paragraph.

2 The two-circular cam profile

Referring to Fig.1, a cam profile can be composed by the following curves. The first two curves are the circle Γ_α, ($\alpha \in \{1, 2\}$), whose radius and center are, respectively, ρ_α and C_α. The third and the four circle, named respectively Γ_3 and Γ_4, are centered on the cam rotation axis O; their radiuses are, respectively, r and $r + h_1$. If one assumes a fixed frame OXY, three characteristic points can be identified: A, which joins Γ_2 with Γ_3; F, which is the point joining Γ_1 with Γ_2; D which joins Γ_1 with Γ_4. In these points, the relative circles have the same tangential vector [9]

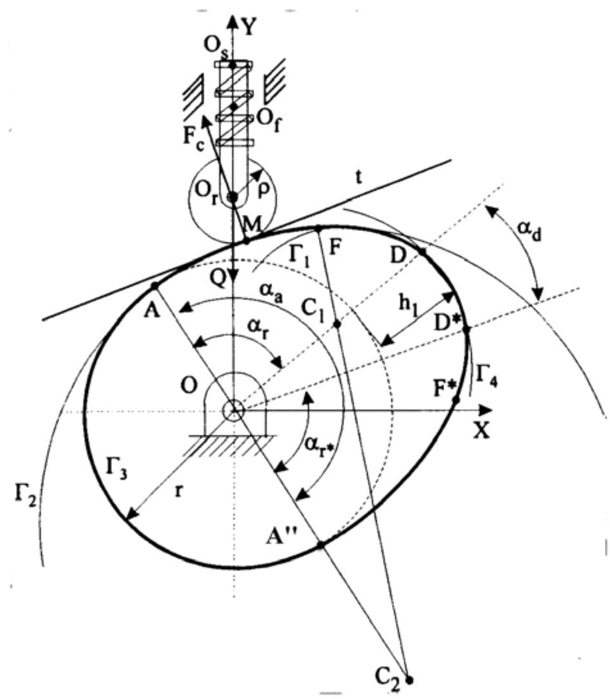

Fig.1. A roller follower two circular-arc cam

3 Methods and Mathematical background

3.1 Discrete Wavelet Transform

Mother wavelets are special functions, whose first h moments are zero. Note that, if ψ is a wavelet whose all moments are zero, also the function ψ_{ik} is a wavelet, where

$$\psi_{jk}(x) = 2^{-j/2}\psi(2^j x - k) \qquad (1)$$

Wavelets, like sinusoidal functions in Fourier analysis, are used for representing signals. In fact, consider a wavelet ψ and a function φ (father wavelet) such that $\left\{\{\varphi_{j_0 k}\},\{\psi_{jk}\}, k \in Z, j = 0,1,2,\ldots\right\}$ is a complete orthonormal system. By Parseval theorem, for every signal $s \in L^2(\mathbf{R})$, it follows that:

$$s(t) = \sum_k a_{j_0 k}\varphi_{j_0 k}(t) + \sum_{j=j_0}^{j_1}\sum_k d_{jk}\psi_{jk}(t) \,. \qquad (2)$$

In particular, the decomposition of a signal $s(t)$ by the Discrete Wavelet Transform (DWT) is represented by detail function coefficients $d_{jk} = <s,\psi_{ik}>$ and by approximating scaling coefficients $a_{j_0 k} = < s, \varphi_{j_0 k} >$.

Observe that d_{jk} can be regarded, for any j, as a function of k. Consequently, it is constant if the signal $s(t)$ is a smooth function, if we consider that a wavelet has zero moments.
Lemma 5.4 in [10] implies the recursive relations:

$$a_{jk} = \sum_{m \in Z} h_{m-2k} a_{j+1,m} \quad \text{and} \quad d_{jk} = \sum_{m \in Z} \lambda_{m-2k} d_{j+1,m},$$

where $\lambda = (-1)^{k+1} h_{1-k}$; $\{h_k,\ k \in \mathbf{Z}\}$ are real-valued coefficients such that only a finite number is not zero and that they satisfy the relations

$$\sum_{k \in Z} h_{k+2m} \overline{h_k} = \delta_{0m}$$

$$\frac{1}{\sqrt{2}} \sum_{k \in Z} h_k = 1.$$

For evaluating the features of the signal, a parameter (entropy) was defined [8]. Given a set S: $= \{x_i,\ I \in \{1,2,\ldots,n\}\}$ and a function c: $x_i \in S \rightarrow c(x_i) \in \mathbf{R}$, the entropy $H(c)$ of c is defined as follows:

$$H(c) := -\sum_{c(x_i) \neq m} \frac{1}{s} \cdot \frac{c(x_i) - m}{M - m} \cdot \ln\left(\frac{1}{s} \cdot \frac{c(x_i) - m}{M - m}\right) \tag{3}$$

where

$$s = \sum_{i \in I} \frac{c(x_i) - m}{M - m},$$

$M: = \max \{c(x_i), i \in \{1,2,\ldots,n\}\}$ and
$m: = \min \{c(x_i), i \in \{1,2,\ldots,n\}\}$.

The entropy measures the best ratio between the maximum dynamic showed by signal and the smallest uniformity of signal. Given $|S| = n$, the entropy, as before defined, riches its maximum value at $ln(n)$ iff, for any $i \in S$, $c(x_i) = const$. Finally $H(c) = 0$ iff, for any $i \in \{1, 2,\ldots,n\}$, $c(x_i) = S$ and, for any $j \in \{1, 2,\ldots,n\}-\{i\}$, $c(x_j) = 0$.
Finally, given a random sample $X_1, X_2,\ldots, X_n$, denote with $X_{(1)}, X_{(2)},\ldots, X_{(n)}$ the corresponding ordered statistics. In this case, we define the sample median as

$$Med(X_1,\ldots,X_n) = \begin{cases} X_{((n+1)/2)} & \textit{if } n \textit{ is odd} \\ \dfrac{X_{(n/2)} + X_{((n/2)+1)}}{2} & \textit{if } n \textit{ is even} \end{cases}$$

Furthermore we define the sample median absolute deviation as

$$Mad(X_1,\ldots, X_n) = Med(|\ X_i - Med(X_1,\ldots, X_n)\ |,\ i \in \{1, 2,\ldots, n\}).$$

On the other side, if X is a random variable, we define the population median $Med(X)$ as the value such that

$$Pr\ (X \leq Med(X)) \geq 1/2 \quad \text{and} \quad Pr\ (X \geq Med(X)) \geq 1/2.$$

The median absolute deviation of the population is defined as follows

$$Mad(X) = Med(|\ X - Med(X)|).$$

3.2 Multivariate analysis

Discriminant analysis with a stepwise elimination was performed [11]. The variables included into the model were the entropic values showed by the signal after the wavelet decomposition into 10 levels. The amount of explained variance was calculated by collinearity diagnostics and multivariate methods [12]. All the analyses were carried out by means of statistical software and statistical significance was accepted at $pr < 0.05$.
For each type of sample a "Group" was created by repeating the experiment.
The resulting vector (i.e., entropic measurements) was normalized to length 1 to compensate for arbitrary scaling differences. In order to identify the most related variable to the characteristics of Group the Spearman correlation coefficients were calculated for each measurement and Group. Discriminant analysis was carried out on all Groups. The Wilks' lambda method was used for selecting the test set to assess the success of the discriminant function, and also for choosing the discriminant variables [13].
The classification functions are appropriate when it can be assumed that the populations under study have both normal multivariate distribution and equal variance-covariance matrices. To test the last assumption the Box's Test of Equality of Covariance Matrices was performed.
To explain the identification process more precisely, let p be an observed signal and (W, ρ) be a specified representation/metric pair. The closest candidate index k* (i.e., the index of the representation in the database that is closest to the observed representation in the sense of ρ) is

$$k^* = \underset{k}{\arg\min}\ \rho\left(avg\left(W_{pk},.\right), W_p\right),$$

where avg is an averaging operator, and

$$\rho : W(\mathrm{H}) \times W(\mathrm{H}) \to (0, \infty)$$

is a metric.
To determine if this candidate is indeed the signal's Group a threshold-based decision function may be formulated. Such a decision function is specified with a closeness threshold δ for which candidates with distances greater than the threshold are deemed outside the database. More precisely, we define a decision function d_δ as:

$$d_\delta(W_p) = \begin{cases} k^*, \rho(x_{k^*}, W_p) < \delta, \\ NEW, else. \end{cases}$$

If all members from the same Group generate sufficiently close representations and members from different Groups generate sufficiently separated representations then this decision will provide perfect identification [14]. Calculations were made using multivariate statistical software (SPSS 10.0 for Windows).

3.3 Fuzzy analysis

A fuzzy clustering algorithm such as the Fuzzy C-Means (FCM) algorithm has been used to find "compact" or "filled" clusters.
In the last ten years the notion of fuzzy classification has been employed in statistics, but already various methods have been proposed for grouping a data set. A fuzzy classification of a data set consists in the subdivision of the initial data set into groups in order that each unit is assigned partially both to a group and more than one group.
Therefore the main difference between classic and fuzzy classification consists in the fact that in the classic theory each unit is assigned for entire to a group, while in the fuzzy theory a function membership is assigned to each unit which measures how much the unit belongs to the group (or the groups) to which it is assigned; that value is in the range [0,1] [15].
The FCM algorithm [16], [17] was used to find clusters that resemble filled hyper spheres or filled hyper ellipsoids.
Let $X = \{x_j, j = 1, \cdots, N\}$ be a set of feature vectors in **n**-dimensional feature space with coordinate-axis labels $[x_1, x_2, \cdots, x_n]$, where $x_j = [x_{j1}, x_{j2}, \cdots, x_{jn}]^T$. Let $B = (\beta_1, \cdots, \beta_C)$ represent a C-tuple of prototypes each of which characterizes one of the C clusters. Each β_i consist of a set of parameters. In the following, we use β_i to denote both cluster i and its prototype. Let u_{ij} represents the grade of membership feature point x_j in cluster β_i. The CxN matrix $U = [u_{ij}]$ is called a constrained fuzzy C-partition matrix if it satisfies the following conditions:

$$u_{ij} \in [0,1] \text{ for all } i,$$
$$0 < \sum_{j=1}^{N} u_{ij} < N \quad \text{for all } i, j \qquad (5)$$
$$\sum_{i=1}^{C} u_{ij} = 1 \quad \text{for all } j.$$

The problem of fuzzily partitioning the feature vectors into C clusters can be formulated as the minimization of the objective function [18]:

$$F(B,U;X) = \sum_{i=1}^{C} \sum_{j=1}^{N} (u_{ij})^m d^2(x_j, \beta_i).$$

In the above equation, $m \in [1, \infty)$ is a weighting exponent called the fuzzifier, and $d^2(x_j, \beta_i)$ represents the distance from a feature point x_j to the prototype β_i. Minimization of the objective function with respect to U subject to the constraints in (5) gives us:

$$\left.\begin{array}{l} u_{ij} = \dfrac{1}{\sum_{k=1}^{C} \left(\dfrac{d^2(x_j, \beta_i)}{d^2(x_j, \beta_k)} \right)^{\frac{1}{m-1}}} \quad if\ I_j = 0 \\ \left.\begin{array}{l} u_{ij} = 0 \quad i \notin_I j \\ \sum_{i \in Ij} u_{ij} = 1\ i \in I_j \end{array}\right\} \quad if\ I_j \neq 0 \end{array}\right\}$$

where $I_j = \{i, 1 \le i \le C, d^2(x_j, \beta_j) = 0\}$.
Minimization of $F(B,U;X)$ with respect to B varies according to the choice of the prototypes and the distance measure [17]. For example, in the FCM algorithm, the clusters are usually assumed to be compact and spherical in shape, and each of the

prototypes is described by the cluster center c_i. If the distance measure is Euclidean or an inner product norm metric, these centers may be updated in each iteration using:

$$c_i = \frac{1}{N_i}\sum_{j=1}^{N}\left(u_{ij}\right)^m x_j$$

where

$$N_i = \sum_{j=1}^{N}\left(u_{ij}\right)^m .$$

Calculations were made using fuzzy logic Toolbox (MATLAB 5.3, The Math Works, Inc, Natick, Mass).

4 Test-bed description

Referring to Fig. 2, one accelerometer S_1 [19], has been installed on the free extremity of the follower for monitoring the acceleration of the follower motion. In addition, dynamical properties can be experimentally evaluated by using a dynamic torsion meter S_2 [20], which was mounted on the actuator shaft of the motor. In order to provide suitable power supply to S_2 and to reduce the noise in the measured signal a signal conditioner and amplifier U2 has been used. One tachymeter S_3 [21] and one encoder S_4, [22] have been installed also on the cam shaft. In particular, the encoder gives the possibility to monitor the angle of the cam shaft, whereas the tachymeter is used to monitor the angular velocity of the cam shaft. Finally, in order to provide different input voltage for the sensors S_1, S_3 and S_4 and motor M three different power supply sources A_1, A_2 and A_3 have been used.

The cam-follower mechanical system was assembled on a frame fixed to the test-bed plate. The radius of the base circle of tested cam is equal to 40 mm. The diameter of the roller is 24 mm. The roller follower moves horizontally along a fixed grooved shaft. The roller is maintained in contact with the profile of the cam by using a suitable spring. In addition, Lab View software [23], and AT-MIO-16F-5 Acquisition Card [24], have been used to acquire and manipulate the data from the accelerometer.

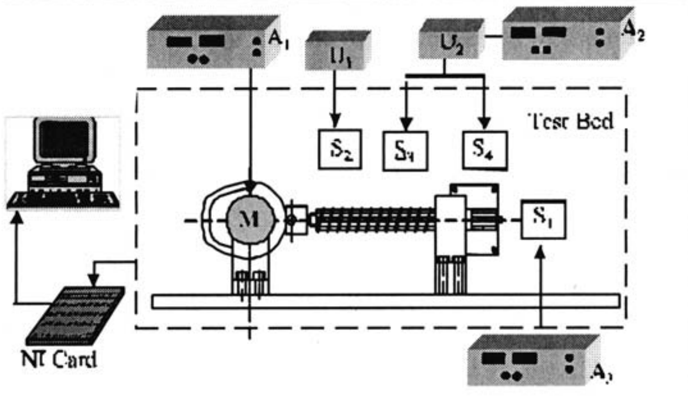

Fig. 2. A general scheme of test-bed

5 Results

Since our study was performed by applying the DWT, we concentrate our analysis on the point where the profile of the cam changes.

The statistical analysis performed by discriminant analysis confirmed the existence of 16 clusters/groups (belonging to 4 main families).

The 82.5% of original grouped cases was correctly classified. Tests performed on each group are reported in Tab.1. They influenced cam angular velocity, sense of rotation, tribological conditions and cam deviations.

The eigenvalues, the percentage of explained variance and canonical correlations referred to each canonical discriminant functions used in the analysis are shown in Tab. 2. They confirmed a good level of classification performed by 2 canonical discriminant functions..

As expected, not all the variables (i.e., 10 orthogonal entropies) employed were used for the best data classification for assigning each signal to the belonging group/family.

Family	**Group signals**
1	**3, 4, 11, 12**
2	**7, 8, 15, 16**
3	**5, 6, 13, 14**
4	**1, 2, 9, 10**

Group	Characteristics
3	OF0 L0 V5-
4	OF0 L0 V5+
11	OF0 L0 V6-
12	OF0 L0 V6+
Group	**Characteristics**
7	OF0 L1 V5-
8	OF0 L1 V5+
15	OF0 L1 V6-
16	OF0 L1 V6+
Group	**Characteristics**
5	OF1 L0 V5-
6	OF1 L0 V5+
13	OF1 L0 V6-
14	OF1 L0 V6+
Group	**Characteristics**
1	OF1 L1 V5-
2	OF1 L1 V5+
9	OF1 L1 V6-
10	OF1 L1 V6+

Legend: OF0 = no off-centre; **OF1** = presence off-centre.
L0 = no lubrication; **L1** = presence of lubrication.
V5+ = 60 rpm and clockwise; **V5-** = 60 rpm and anticlockwise.
V6+ = 80 rpm and clockwise; **V6-** = 80 rpm and anticlockwise.

Tab.1 Characteristics of tests performed by means of test-bed

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	19,609a	81,6	81,6	,975
2	4,408a	18,4	100,0	,903

a. First 2 canonical discriminant functions were used in the analysis.

Tab. 2 Statistical values of each canonical discriminant function

The results of fuzzy clustering was very interesting: the 98% of groups was correctly classified. In particular, the group numbered as 11, belonging to the family 1, was put between the families 1 and 4. Low electrical tension, occurred during the test, reduced the rotation speed of cam.
The Fig. 3 shows the scatter-plot of all groups obtained by the application of fuzzy analysis. It is easy to see the existence of 4 well defined families and, for each of them, the existence of 4 groups or clusters.

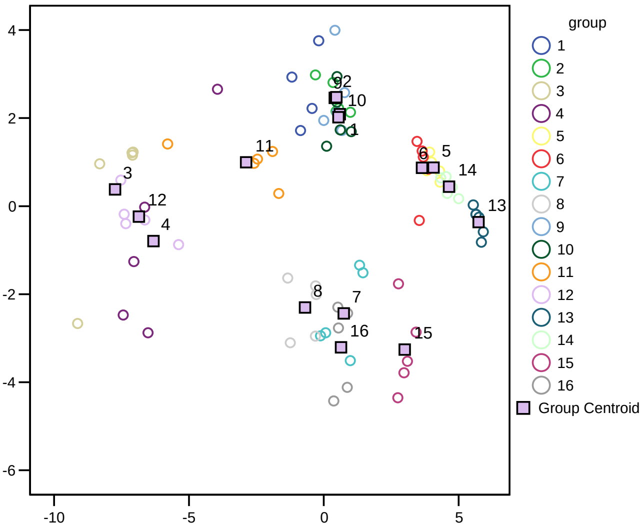

Fig. 3 A scatter-plot of all groups

Since the fuzziness generates also an overlapping of groups, it could provide a more complex classification. Such a complexity, in part can easily be limited by using simple options during the selection of results, on the other hand it constitutes the real wealth of these methods that supply an amount of information more advanced with respect to the classical statistical methods.
Moreover these algorithms concur to accept the real structure of the data by limiting to the minimum the forcing during the creation of the groups: probably an 'imprecise' model (in the sense of *fuzzy*) of the reality is a better representation of it instead of a *precise* model (in the mathematical sense of term).
Finally, the introduction of medians shows improved performance in clustering data sets generated by heavy-tailed distributions like the Cauchy distribution.
In fact, by processing the data-set modified by introducing the features calculated by medians as defined in the paragraph 3.1, we obtain a 99.3% of correct signal classification.

6 Conclusions

This work demonstrates that a powerful discrimination level is obtainable by means of orthogonal decomposition of signals with the application of DWT in conjunction with statistical and fuzzy analysis. It is more relevant if we consider that in such assessments the co-presence of several stochastic factors can influence the performance and the response of experimental models [25][26].

In this paper we proposed a new approach to feature selection.

An extension of two key robust statistics to data set is presented and applied to the fuzzy C-means clustering algorithm.

Acknowledgements

We would like to thank Prof. Marco Ceccarelli, for his contribution to this study for providing test-bed used for the present work.

References:

[1] Scotto Lavina G., *Applicazioni di Meccanica Applicata alle macchine*, Ed. Sistema, 1971.

[2] Mahyuddin A.I., Midha A, Influence of Varying cam profile and Follower Motion Event Types on Parametric Vibration and stability of Flexible Cam-Follower Systems, *Journal of Mechanical design*, Vol. 116, 1994, pp. 298-305.

[3] Daubechies I., *Ten Lectures on Wavelets*, SIAM, 1992.

[4] Donnarumma A., Liverani A., Persiani F., Podda G., A wavelet transform application to the stereo imaging process, *Intl. conf. on Computer Integrated Manufacturing*, Zakopane, Marzo 2001, pp. 98-105.

[5] Kaiser G., *A Friendly Guide to Wavelets*, Birkhäuser, 1999.

[6] Härdle W. et al. *Lecture Notes in Statistics - Wavelets, Approximation, and Statistical Applications*, Springer, 1998.

[7] Antoniadis A., Oppenheim G., *Lecture notes in Statistics - Wavelets and Statistics*, Springer, 1995.

[8] Niola, V., Quaremba, G. and Oliviero, R**.,** The location of anomalous spikes on a working rotor: a theoretical-experimental test. *Int. J. of Applied Mechanics and Engineering,*, Vol.9, No.3, 2004, pp.547-556.

[9] Lanni C., Ceccarelli M., Figliolini G., An Analytical design for three circular-arc cams, *Mechanism and Machine Theory*, Vol 37, 2002, pp. 915-924.

[10] Härdle, W, Kerkyacharian, G., Picard, D. and Tsybakov, A**.,** *Wavelets, Approximation, and Statistical Applications*, Springer, Berlin, 1998.

[11] Hand, D.J**.,** *Discrimination and Classification*, John Wiley & Sons, New York, 1981.

[12] Kendall, M., Stuart, **A.,** *The advanced Theory of Statistics*, MacMillan, New York, 1977.

[13] Van der Geer, J.P**.,** *Introduction to Multivariate Analysis for the Social Sciences*, W.H. Freeman, San Francisco, 1971.

[14] Anthony, T**.,** *Computational Signal Processing with Wavelets,* Birkhäuser, Boston, 1998.

[15] Zadeh L.A., Fuzzy sets. *Information and Control*, Vol.8, No.338, 1965.

[16] Chiu S., Fuzzy model identification based on cluster estimation. *J. of Intelligent & Fuzzy Systems*, Vol.3, 1994.

[17] Bezdek J.C., Pattern recognition with fuzzy objective function algorithms. *New York: Plenum*, 1981.

[18] Krishnapuram, R., Frigui H. and Nasraoui O., Fuzzy and possibilistic shell clustering algorithms and their application to boundary detection and surface approximation, *Part I, IEEE Tran. On Fuzzy Systems*, Vol. 3, No. 1, Feb 1995.

[19] Analog Devices, Analog Accelerometer ADXL105JQC, 1999.

[20] FGP Instr. Dynamic Torsionmeter mod. CD 1050, User's Manual, 1995.

[21] Hewlett Packard, Incremental Encoder mod. HEDS-5701 G00, 2001.

[22] Baumer Electric, Incremetal Encoder mod. BHG 06.25W5000-B2-5,2001.

[23] National Instruments, Lab View 3.01, 1994.

[24] National Instruments, AT-MIO-16F-5, Acquisition Card, 1995.

[25]DonnarummaA., Persiani, F., Podda, G., Planar contour recognition: compression via wavelet filters and classification via neural networks and fuzzy systems. Int. Conf. on Computer Integrated Manufacturing, Zakopane, 1999, pp.126-135.

[26]B.Kosko: Neural Network and fuzzy systems. A dynamical approach to machine intelligence. Prentice Hall, EnglewoodsCliff, NO7632.