



MODULAR AND SELF-ORGANIZING CONNECTIONIST SYSTEMS: TOWARD HIGHER LEVEL INTELLIGENT FUNCTIONS

Kurosh Madani

Image, Signal and Intelligent Systems Laboratory (LISSI / EA 3956), Senart Institute of Technology, University PARIS XII, Av. Pierre Point, F-77127 Lieusaint, France
madani@univ-paris12.fr, <http://www.univ-paris12.fr>

Abstract: *Recent advances in “neurobiology” allowed highlighting some of key mechanisms of animal intelligence. Among them one can emphasize brain’s “modular” structure and its “self-organizing” capabilities. The main goal of this paper is to show how these primary supplies could be exploited and combined in the frame of “soft-computing” issued techniques in order to design intelligent artificial systems emerging higher level intelligent behavior than conventional Artificial Neural Networks (ANN) based structures.*

Keywords: *Modularity, Self-Organization, Artificial Intelligent systems, Real-World applications, Implementation.*

1. INTRODUCTION

Much is still unknown about how the brain trains and self-organizes itself to process so complex information. However, the recent advances in “neurobiology” allowed highlighting some of key mechanisms of animal (and human) intelligence. In fact, our simple and inappropriate binary technology remains too primitive to reproduce the biological complexity of these marvelous mechanisms, but a number of those highlighted points could already be sources of inspiration for higher level intelligent artificial systems. Among interesting features of animal’s and human’s brain, one can emphasize its “modular” structure and its “self-organizing” capabilities. If it is still early to state on “concurrent” or “cooperative” nature of ways that these complex features interact, they could already be considered as basic features in emergence of higher level artificial intelligent behavior.

On the other hand, overcoming limitations of conventional approaches thanks to their learning and generalization capabilities, Artificial Neural Networks (ANN) made appear a number of expectations to design “intelligent” information processing systems. If learning and generalization capabilities of these bio-inspired connectionist models appear as central requirements in intelligent systems’ design, nowadays, it is well admitted that intelligent behavior requires more sophisticated mechanisms than those performed by these “simple” models.

The main goal of this paper is to show how these

primary supplies could be exploited and combined in the frame of “soft-computing” issued techniques in order to design intelligent artificial systems emerging higher level intelligent behavior than conventional Artificial Neural Networks (ANN) based structures. These foremost features have inspired a set of implementations dealing with real-world applications and covering several different areas as: robotics, image processing and pattern recognition, classification and dynamic nonlinear behavior modeling (identification and prediction).

The present paper is organized in following way: the next section will briefly introduce the general frame of modular modeling. Section 3 will describe a first applicative implementation dealing with “biometric face recognition” dilemma in the challenging frame of “mass biometry”. In section 4, a different self-organizing tree-like modular system, taking advantage from a “complexity estimation” loop, will be described. Section 5 will present a modular Fuzzy-CMAC architecture dealing with fully autonomous biped robot’s walking dilemma. Section 6 will give an additional applicative example of a self-organizing modular connectionist system dealing with nonlinear dynamic systems’ behaviour identification. Finally, the last section will conclude the present article and discuss a number of perspectives.

2. GENERAL FRAME OF MODULAR MODELING

From a general point of view, a multi-model is composed of several models each of which is valid

in a well defined interval which corresponds to a part of the operation range of the system or covers a part of the whole feature space of the problem to be solved. The local validity of a model in a well defined interval is specified by using functions with limited supports which tend to significantly increase the contribution of the local models in that zone and tend to decrease it elsewhere. The combination of all local models allows description of the whole system's behavior. The local models participations in the multi-model's output are determined by "activation degree" associated to each local model. The action of "activation degrees" on multi-model's response could be seen as some kind of local models responses weighting fashioning its response in order to approximate the modeled behavior. In a general way, modular (or multiple models) structures could be classified according to two main classes: "predefined" structures and "adaptive" (evolutionary) architectures. In the first class of such systems, as well the numbers of modules as the whole system response's criterion (modules' responses aggregation criterion) are defined regarding the treated problem and remain unchanging. While in the second class of such systems as well the numbers of modules as the whole system response's criterion (modules' responses aggregation criterion) could evaluate regarding the treated dilemma adapting the system to the treated task.

Consider a system described by the general equation (or transfer function), expressed by relation (1), where $F(\cdot)$ represents a global unknown model (complex task to be performed, complex system to be identified, complex behavior to be described, etc...) and $\varphi(t)$ is a feature vector (characteristic vector composed by a number of features related to data to be processed, regression vector composed by a number of delayed system's inputs and outputs, etc...). β is a parameter vector.

$$y(t) = F(\varphi(t), \beta) \tag{1}$$

The associated multi-model, composed by M local models (or processing units) approximating (describing) the system's complex behavior (expressed by relation (1)) is defined by relation (2) with condition given by relation (3) where $f_i(\varphi(t))$ represents the i -th local model (or local processing unit), β is a parameter vector and $S(\cdot)$ represents a fusion operator or a selection function. ε is the approximation error which is supposed to be as small as possible.

$$\hat{y}(t) = S(\varphi(t), \beta) \tag{2}$$

$$y(t) = \hat{y}(t) + \varepsilon \tag{3}$$

A number of works dealing with modular computing and issued architectures have been proposed since 1993 associating a set of ANN in a modular structure in order to process a complex task by dividing it into several simpler sub-tasks. One can mention neural networks ensemble concept proposed by [1], intelligent hybrid systems [2], Mixture of experts concept proposed by [3] and [4] or structures based on dynamic cells [5]. In the same years, a number of authors proposed multi-modeling concept for nonlinear systems modeling ([6] to [12]) in order to avoid difficulties (modeling complexity). In fact, taking advantage from "modularity", multi-modeling concept reduces considerably modeling or processing complexity by dividing the initial complex problem (or task) into a set of local models (or local processing modules). It is important to remind that the most of proposed works (excepting those described in the four latest references) remain essentially theoretical and limited to a general frame work. In fact, if a relatively consequent number of different structures have been proposed, a very few of them have been applied to real-world dilemmas solution.

The main problems related to such kind of architectures, are linked, on the one hand to structure construction (appropriated number of modules determination) criteria, modules' nature determination and on the other hand to modules' interactions (organization) criteria determination. These major points could be defeated by adding self-organizing skill to a multi-model (or to a modular architecture) leading to powerful structure, especially if local models (or local modules) are ANN based units. A number of such self-organizing multiple ANN structures have been proposed and described in [9], [10], [12] and [13]. These works could be divided into two main directions.

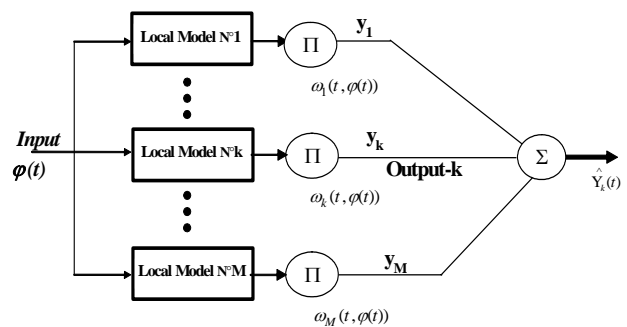


Fig. 1 - General bloc diagram of weighted contribution based modular connectionist system

The first one, described in [9] and [10], takes advantage from a weighted contribution of several local models. Figure 1 gives the general bloc

diagram of this kind of modular connectionist systems. The fusion operator is the weighted sum function. In this case, the associated multi-model, composed by M local models (or processing units) and their weights $\rho_i(\varphi(t), \beta_i)$, is defined by the weighted average expressed in the relation (4), with $\rho_i(\varphi(t), \beta_i) \geq 0$ (for all i) and $\sum_{j=1}^M \rho_j(\varphi(t), \beta_j) > 0$ (for all $\varphi(t)$). In this relation β_i is a parameter related to the validity function ρ_i .

$$\hat{y}(t) = \frac{\sum_{j=1}^M \rho_j(\varphi(t), \beta_j) f_j(\varphi(t))}{\sum_{j=1}^M \rho_j(\varphi(t), \beta_j)} \quad (4)$$

In fact, in this approach, the number of local models (modules), the “contribution rate” (called also “activation degree”, setting the contribution degree of each local model in whole modular system’s output), local models’ natures and other parameters are adjusted in order to minimize either the whole modular system’s output error (global learning criterion) or the local models’ output errors (local learning criterion). Different tuning mechanisms, based essentially on feature space partitioning (splitting) techniques, have been described and implemented in [10]. Two of the most promising of them are “output error based decision tree” and “output error based Fuzzy partitioning” mechanisms (described and implemented in [10]).

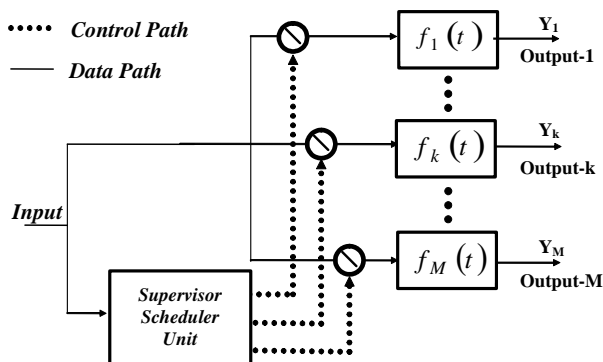


Fig. 2 - General bloc diagram of expert selector based modular connectionist system

The second one, described in [12] and [13], represented by the general bloc-diagram of figure 2, takes advantage from a “complexity estimation” unit acting as supervisor in multi-model’s structure construction. In this kind of connectionist modular architecture, the output is computed on the basis of a selection functions (expressed by (5)), which depends on $\varphi(t)$ and on some parameters p and/or

conditions ξ . p_k represents some particular values of parameter p and ξ_k denotes some particular value of condition ξ , respectively.

$$\begin{cases} \hat{y}(t) = S(\varphi(t), p, \xi) = Y_k & \text{if } p = p_k \text{ and } \xi = \xi_k \\ \hat{y}(t) = 0 & \text{else} \end{cases} \quad (5)$$

In fact, the modular structure is fashioned according to a “decision tree” based criterion regarding the data complexity, estimated from the problem’s (task’s) representative data. At the end of tuning process (adaptation process) an “experts’ selector” (called also “scheduler unit”) and a number of specialized models (experts), called also “sub-models” are constructed. In operation phase (after tuning process), a new data (unlearned data) is processed according to following principle: first the data (pattern, image, etc...) is classified by the scheduler unit regarding the available (generated) experts; then, the most appropriated expert processes that data.

However, the evolutionary natures of both of them are based on “data driven” criteria, meaning that the structure, number of modules, local models’ parameters and other adjustable parameters features are tuned from data (learning database) representing the treated problem or processed task.

3. MODULAR FACIAL RECOGNITION SYSTEM USING KERNEL FUNCTIONS ANN AS LOCAL UNITS

Contrary to “individual biometry” where both authentication and identification operations assume a precise biometrical characterization of concerned individuals, the main goal in “mass biometry” is to authenticate or identify an unusual (suspect) behavior within a flow of mass customary behaviors. That’s why, in “mass biometry” the chief requirements concern on the one hand, the ability of handling patterns containing relatively poor information and on the other hand, the skill of high speed processing in order to treat a mass number of patterns in a reasonably acceptable delay (real-time). The solution we propose [14] includes three main stages. The two firsts are a video (image flow) acquisition device, which could be a standard digital video camera and an image processing stage performing a set of image pre-processing operations and extracting a number of facial biometric features. The last stage is a modular stage composed by a set of kernel functions based ANN ([15] to [18]) units carrying out classification and decision operations. Figure 3 shows the block diagram of the proposed modular architecture.

A prototype of such modular facial recognition

system has been realized using three ANNs (figures 3 and 5). Each ANN is specialized in processing of a specific kind of biometric feature extracted from the input image. Then a decision logic based procedure performs (on the basis of classification results relative to each biometric feature) the identification of the concerned individual. The implementation has been done on the basis of ZISC-036 neuro-processor based board composed by 16 chips, each one including 36 neurons ([19] to [21]).

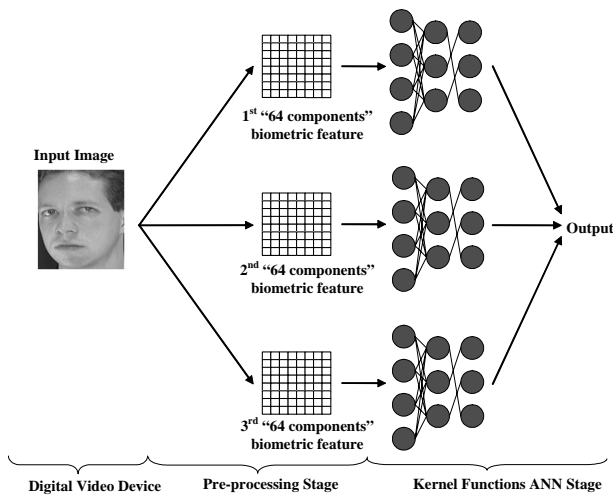


Fig. 3 - Bloc diagram of the implemented modular face recognition system

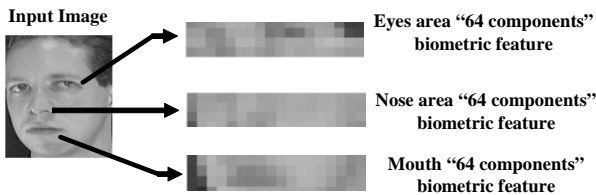


Fig. 4 - Example of "localized biometric features" processed by each module composing the classification-decision stage

The proposed solution takes advantage at the same time from kernel functions based ANN's image processing ability implemented by ZISC-036 and from the massively parallel architecture of this neuro-processor allowing very high processing speed. The obtained promising results show feasibility and effectiveness of the proposed solution reaching 85% correct identification involving a relatively weak number of learned samples (5 samples per face). The experimental validation has been done using the ORL (Olivetti Research Laboratory, Cambridge) faces database. This database is composed of 400 images representing 40 individuals. In other words, the database offers 10 different pictures of a set of 40 faces (corresponding to different individuals), each one representing a different situation: different mimics, with and without glasses, different degrees of rotation, etc....Figure 6 gives an example of images set corresponding to a given individual (face) of ORL database. The database has been divided into two equal parts. The first one including 5 pictures of the whole individuals (40 individuals) has been used for learning phase. The other 5 images (unlearned pictures) of each individual have been used for testing phase.

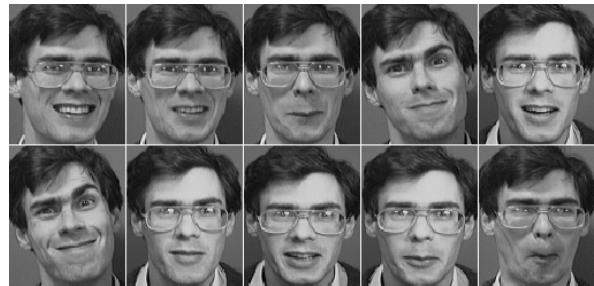


Fig. 6 - Example of images set offered by ORL database for a same face (individual).



Fig. 5 - Photographs, showing the implemented system (upper-left picture), the ZISC-036 neuro-processor based board (upper-right), and the screen of the implemented modular face recognition system (lower-left and lower-right pictures)

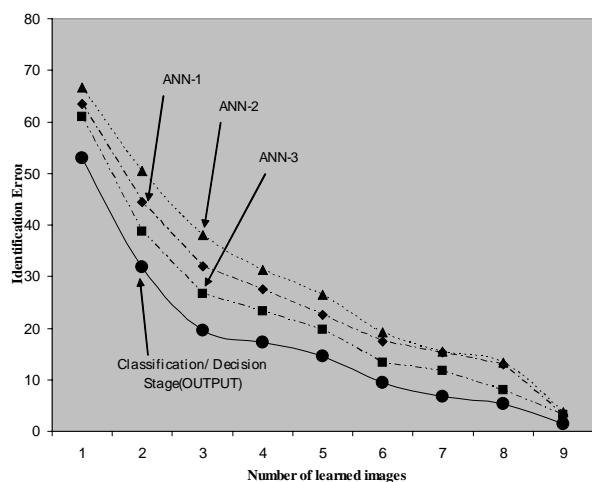


Fig. 7 - Obtained "identification error" versus the number of learned patterns (images) for the whole system and for each of the three neural nets composing the "Classification / Decision" stage

Results (shown in figure 7) concern identification using the KNN mode of ZISC-036 neuro-processor. They show the obtained “identification error” as well for the whole system as for each of three neural nets composing the “Classification / Decision” stage. The experimental validation protocol consists in progressive enlargement of the number of learned samples. The first comment concerns the significantly improved identification rate, reaching 85% for 5 learned samples (of each face) and more than 98% (e.g. less than 2% “identification error”) for 9 learned samples. The second interesting remark related to this figure concerns the comparison of whole system’s “identification errors” with those obtained for each neural net composing the system. In fact, as it could be remarked from figure 6, the “identification error” of the whole system remains lower than “identification errors” affecting each individual neural network composing the system: especially (and even) when the number of learned samples remains relatively small (3 to 5 learned samples). For example, in the case of 4 learned samples of each face, first, second and third (ANN-1, ANN-2 and ANN-3) commit 27%, 31% and 23% “identification errors”, respectively, while the whole system’s “identification error” remains about 17% (e.g. 83% of correct identification). Finally, it is pertinent to remain the relative poorness of the processed information characteristic of the “mass-biometry” context. These promising results open a number of auspicious perspectives concerning as well the proposed solution as the “mass biometry” related applications in general. We are working now on two directions. On the one hand we are investigating new “64 components biometric features” (representations), and on the other hand we develop more sophisticated learning strategies on ZISC-036.

4. TREE-LIKE MULTIPLE NEURAL NETWORKS MODELS GENERATOR WITH A COMPLEXITY ESTIMATION BASED DECOMPOSER

In a very large number of cases dealing with real world dilemmas and applications (system identification, industrial processes, manufacturing regulation, optimization, decision, pattern recognition, systems, plants safety, etc), information is available as data stored in files (databases etc.). So, the efficient data processing becomes a chief condition to solve problems related to above-mentioned areas. In the most of those cases, processing efficiency is closely related to several issues among which are:

- Data nature: including data complexity, data quality and data representative features.

- Processing technique related issues: including model choice, processing complexity and intrinsic processing delay.

One of the key points on which one can act is the complexity reduction. It concerns not only the problem solution level but also appears at processing procedure level. An issue could be model complexity reduction by splitting a complex problem into a set of simpler problems: multi-modelling where a set of simple models is used to sculpt a complex behaviour ([9] and [10]). Another promising approach to reduce complexity takes advantage from hybridization [22].

We designed and implemented an ANN based data driven treelike Multiple Model generator, that we called T-DTS (Treelike Divide To Simplify), able to reduce complexity on both data and processing chain levels ([9], [10], [23]). T-DTS and associated algorithm construct a tree-like evolutionary neural architecture automatically where nodes, called also “Splitting Units” (SU), are decision units, and leafs, called also “Neural Network based Models” (NNM), correspond to neural based processing units.

The T-DTS includes two main operation modes. The first is the learning phase, when T-DTS system decomposes the input data and provides processing sub-structures and tools for decomposed sets of data. The second phase is the operation phase (usage the system to process unlearned data). There could be also a pre-processing phase at the beginning, which arranges (prepare) data to be processed. Pre-processing phase could include several steps (conventional or neural stages). Figure 8 gives the general bloc diagram of T-DTS operational steps. As shows this figure, T-DTS could be characterized by three main operations: “data pre-processing”, “learning process” and “generalization process” (or “working process”).

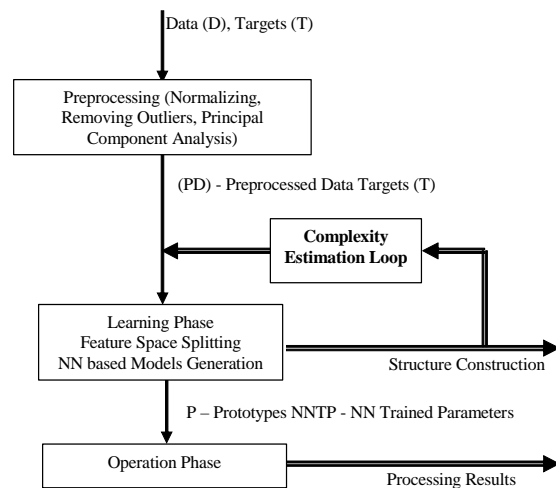


Fig. 8 - General bloc diagram of DTS, presenting main operation levels

The complexity estimation based splitting could be performed according to two general strategies: “static” splitting strategy and “adaptive” (dynamic) one. In both cases, the issued could be binary or multiple branches tree-like structure. The main difference between two strategies remains in nature of the complexity estimation indicator and the splitting decision operator used in splitting process.

The learning phase is an important phase during which T-DTS performs several key operations: splitting the learning database into several sub-databases, constructing (dynamically) a treelike Supervision/Scheduling Unit (SSU) and building a set of sub-models (NNM) corresponding to each sub-database. Figure 9 represents the division and NNM construction process bloc diagrams. As this figure shows, after the learning phase, a set of neural network based models (trained from sub-databases) are available and cover (model) the behaviour region-by-region in the problem’s feature space. In this way, a complex problem is decomposed recursively into a set of simpler sub-problems: the initial feature space is divided into M sub-spaces. For each subspace k , T-DTS constructs a neural based model describing the relations between inputs and outputs. If a neural based model cannot be built for an obtained sub-database, then, a new decomposition will be performed on the concerned sub-space, dividing it into several other sub-spaces.

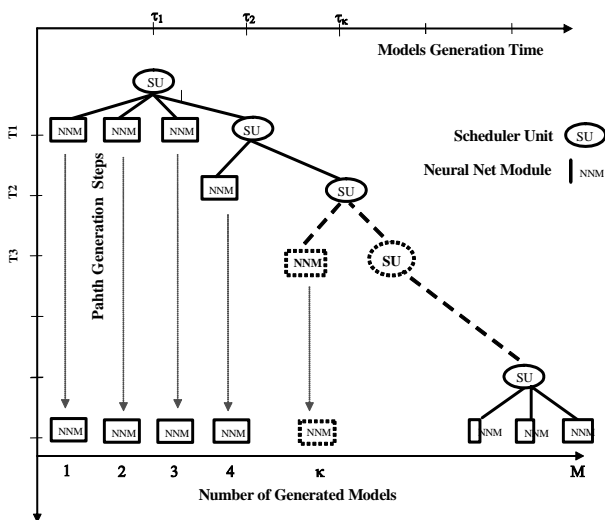


Fig. 9 - General bloc diagram of T-DTS learning phase and its tree-like splitting process

A software implementation of T-DTS including different splitting strategies, different complexity estimation methods and different ANN models has been achieved. The implementation has been performed under MathLab environment. Very promising results, obtained for different areas: classification problems, industrial process identification and prediction, pattern (biomedical

signal) recognition, etc... show efficiency of such self-organizing multiple model structure.

5. BIPED ROBOT'S ADAPTIVE WALKING USING INTUITIVE HYBRID MODULAR CONTROLLER

One of the most challenging topics, over the recent decades, in the field of robotics concerned the design and the control of biped robots. Several potentialities make this foremost research area particularly appealing in the frame of middle and long term projection. On the fundamental side, advances in this research area can lead to a better comprehension of the human locomotion mechanisms. From the applicative point of view, it could concern a wide spectrum of applications among which: the design of more efficient prosthesis and the construction of more sophisticated humanoid robots for interventions in hostile environments.

Two main control strategies are generally used in the field of biped robots' locomotion: one is based on a kinematics and dynamic modeling of the whole robot's mechanical structure, and another takes advantage from soft-computing techniques (fuzzy logic, neural networks, genetic algorithm, etc...) and heuristically established rules resulting from the expertise of the walking human. Additionally to requirements related to high precision measurement and to a fine interaction forces' evaluation, the first strategy needs the modeling of whole biped robot's real environment remaining a very complex task. That is why the computing of the on-line trajectories are generally performed using simplified models ([22] to [27]), making this first strategy not always well adapted when biped robot moves in real environment. Taking advantages from soft-computing skills, the second solution doesn't need the aforementioned requirements: firstly, it is not necessary to know perfectly the mechanical structure and secondly, this category of techniques takes advantage from learning capabilities ([24] to [28]).

Investigating soft-computing based fully autonomous biped robot's walking, we proposed a new approach taking advantage simultaneously from local and global generalization. Our approach [29] is based on a modular Fuzzy-CMAC architecture: a set of CMAC ANN (see [30] to [32]) based modules and a fusion stage. The fusion is carried out by using Takagi-Sugeno FIS (Fuzzy Inference System). The main task of Fuzzy-CMAC based modular part of the system is to compute the swing leg's trajectory (using a Fuzzy Inference System fusion of several CMAC neural networks' outputs). The second one allows regulating the average velocity from a modification of the desired pitch angle at each new step. Figure 10 gives the bloc diagram of the

proposed hybrid architecture.

Figure 10 shows the bloc diagram of the training strategy. The trajectories of the swing leg (in terms of joint positions and velocities) are learned by four "single-input/single-output" $CMAC_k$ with $k=1,\dots,4$ neural networks (four trajectories to learn). The learned trajectories are joint angles q_{i1} and q_{i2} , and the two corresponding angular velocities \dot{q}_{i1} and \dot{q}_{i2} . q_{i1} and q_{i2} are respectively the measured angles at the hip and the knee of the leg i .

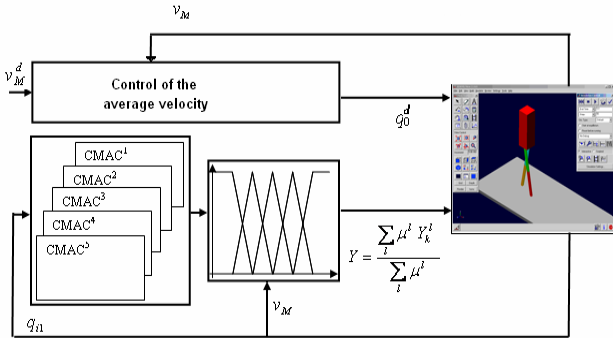


Fig. 10 - Bloc- diagram of the Fuzzy-CMAC based hybrid control strategy

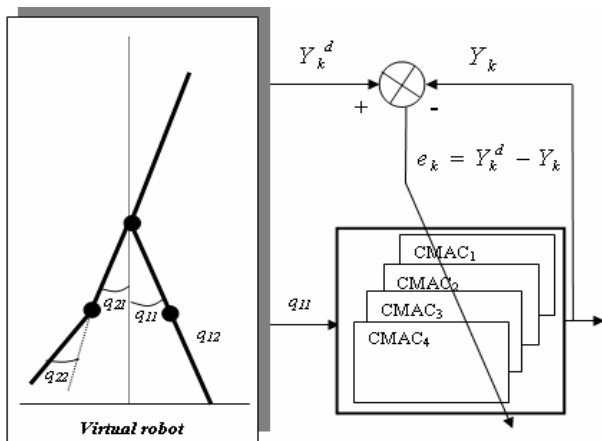


Fig. 11 - Learning strategy principle's bloc diagram

In the same way, \dot{q}_{i1} and \dot{q}_{i2} are respectively the measured angular velocities at the hip and the knee of the leg i (see figure 11). During the training stage, five trajectories corresponding to five different average velocity values (V_M measured in m/s) included in $[0.4, 0.8]$ interval are learned by five CMAC based modules. Each module (labelled $CMAC^l$, with $l \in \{1, 2, 3, 4, 5\}$) includes four $CMAC_k$ neural networks (corresponding to the four above-mentioned robot's trajectories). V_M is computed by using relation (6) where L_{step} is the distance between the two feet at the moment of double impact and t_{step} is the duration of the step (from takeoff to landing of the same leg).

$$V_M = \frac{L_{step}}{t_{step}} \quad (6)$$

The Fuzzy Inference System is obtained from the five following rules, where Y^l corresponds to the output of $CMAC^l$ with $l \in \{1, 2, 3, 4, 5\}$:

- IF V_M IS VerySmall THEN $Y = Y^1$
- IF V_M IS Small THEN $Y = Y^2$
- IF V_M IS Medium THEN $Y = Y^3$
- IF V_M IS Big THEN $Y = Y^4$
- IF V_M IS VeryBig THEN $Y = Y^5$

Figure 12 gives the membership functions corresponding to the upper-indicated FIS rules. The average velocity is modelled by five fuzzy sets ("VerySmall", "Small", "Medium", "Big", "VeryBig").

The validation of proposed approach has been done on an under-actuated robot: RABBIT [33], [34]. This robot constitutes the central point of a project, within the framework of CNRS (Centre Nationale de la Recherche Scientifique) ROBEA (ROBOTique et Entité Artificielle) program [35], concerning the control of walking and running biped robots, involving several French laboratories. This robot is composed of two legs and a trunk and has no foot as shown on figure 13. The characteristics (masses and lengths of the limbs) of this biped robot are summarized in table 1.

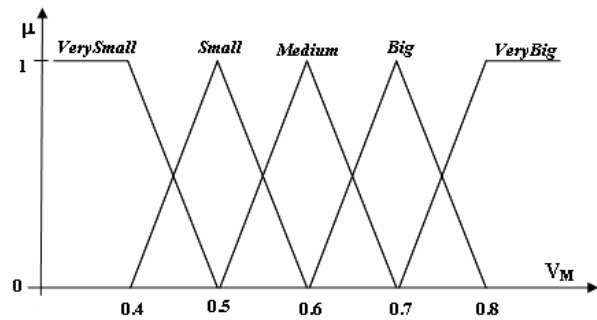


Fig. 12 - Membership functions used by Fuzzy Inference stage of Fuzzy-CMAC

Table 1. Masses and lengths of the robot's limbs

Limb	Weight (kg)	Length (m)
Trunk	12	0.2
Thigh	6.8	0.4
Shin	3.2	0.4

If it is true, from design point of view, that RABBIT is simpler compared to a robot with feet,

from the control theory point of view, the control of this robot is a more challenging task, particularly because, in phase of single support, the robot is under-actuated. A numerical model of the previously described robot has been implemented within the ADAMS software. This software is able to simulate RABBIT's dynamic behavior and namely to calculate the absolute motions of the platform and the relative motions of the limbs when torques are applied on the joints by the virtual actuators.

The model used to simulate the interaction between feet and ground is exposed in [36]. Figure 14 gives the stick diagram of the biped robot's walking sequence when the desired average velocity increases. It must be noticed that the control strategy allows adapting automatically the pitch angle and the step length as the human being.



Fig. 13 - RABBIT prototype's photograph

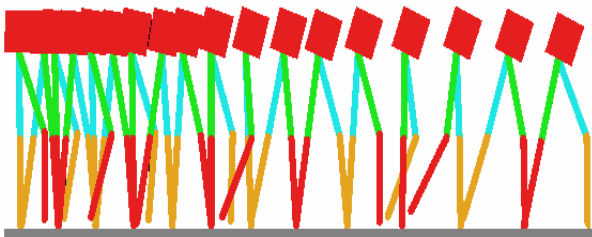


Fig. 14 - Stick diagram showing a walking sequence of the biped robot with increasing average velocity increases

The main interest of this approach is to proffer to the walking robot autonomy and robustness. The obtained results show the adaptability of the walking step length. Furthermore, the Fuzzy-CMAC approach allows decreasing the memory size in comparison to the traditional multi-input CMAC ANN. Future works will focus firstly on the extension of the Fuzzy-CMAC approach in order to

increase the autonomy of the walking robot according to the nature of the environment (get up and down stairs for instance), avoidance and dynamic crossing obstacles and secondly on the experimental validation of our approach.

6. IDENTIFICATION OF NONLINEAR DYNAMIC SYSTEMS' BEHAVIOR USING SELF-ORGANIZING MODULAR STRUCTURES

Identification of nonlinear systems behavior is an important task in a large number of areas dealing with real world requirements and issued applications. Among numerous areas concerned by this task, one can mention model based control and regulation, systems design, complex systems simulation, complex systems' behavior prediction, fault diagnosis, etc... The identification task involves two essential steps: structure selection and parameter estimation. These two steps are linked and generally have to be performed in order to achieve the best compromise between the identification (or prediction) error minimization and the number of parameters increase in the issued model. In real world applications (real world situations), strong nonlinearity and large number of related parameters make the realization of those steps challenging, and so, the identification task difficult.

To overcome the above-mentioned difficulties, we propose to take advantage simultaneously from multi-modeling concept's modularity (described in section 2) and self-organizing clusters construction, making the proposed solution self-adaptive regarding the system's (nonlinear system to be identified) nonlinearity. Concerning the self-organization, the proposed identifier benefits from a self-organizing clusters construction, based on concurrent minimization of both identification error and number of local models. Regarding partitioning strategy, two promising partitioning strategies have been investigated: "decision tree construction" (DTC - a deterministic partitioning approach) and "fuzzy clustering" (FC - a fuzzy based partitioning approach [37]).

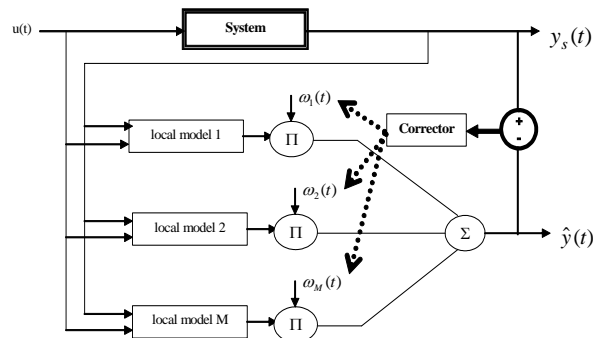


Fig. 15 - Learning bloc diagrams of EE multi-model

The identification could be performed using an “Equation Error” (EE) multi-model, known also as NARX (Nonlinear Autoregressive with eXogenous Inputs) multi-model, with “decision tree construction” or “fuzzy clustering” partitioning (to split the system’s feature space in a number of operating ranges) [38]. Figure 15 shows the bloc diagram of an EE multi-model based identifier. As one could remark from this figure, the EE multi-model based identifier identifies the system by using both system’s inputs and outputs.

In the case of a deterministic partitioning strategy, the “activation degree” of the i -th local model is defined conformably to the relation (7), where $\rho_i(\cdot)$, called the “validity function” of the i -th local model, is defined by the relation (8). In relation (8), $\mu_k(\cdot)$ represents the “membership function” defined for the k -th variable of the regression vector $\varphi(t)$ and Q is the number of variables in the regression vector. In our approach, we use Gaussian membership functions expressed in (9), where: $z_{k_i}(t)$ is the value of the k -th variable of the regression vector $\varphi(t)$ involved in the i -th local model, c_{k_i} is the center of the partition corresponding to the $z_{k_i}(t)$ and σ_k is the dispersion of the Gaussians for all partitions of the k -th variable. It is interesting to note that the parameters vector β_i contains all the dispersion of the Gaussians.

$$\omega_i(\varphi(t), \beta) = \frac{\rho_i(\varphi(t), \beta_i)}{\sum_{j=1}^M \rho_j(\varphi(t), \beta_j)} \quad (7)$$

$$\rho_i(\varphi(t), \beta_i) = \prod_{k=1}^Q \mu_k(z_{k_i}(t)) \quad (8)$$

$$\mu_k(z_{k_i}(t)) = \exp\left(-\frac{(z_{k_i}(t) - c_{k_i})^2}{2\sigma_k^2}\right) \quad (9)$$

The FC partitioning strategy uses the “fuzzy-c-mean” clustering algorithm. Conformably to the fuzzy nature of the clustering, the issued intervals (operating ranges) could share some overlapping region (with different membership degree). Feature space decomposition is performed in each dimension (for each input variable) according to concurrent minimization of both identification error and “intra-clusters” error defined by relation (10), where d_{ij} expressed by relation (11) denotes the distance

between the j -th value of the variable z (which could take Q different values) and the center c_i of the i -th cluster (among M possible clusters). μ_{ij} in relation (12) represents the membership degree relative to the variable z regarding the i -th cluster (among M possible clusters), defined by relation (12). The “activation degree” is the given by the values of μ_{ij} .

The center c_i of the i -th cluster is defined conformably to the relation (13). Finally, the parameter m , known as “fuzzy exponent”, is a parameter representing overlapping shapes between clusters. Generally, this parameter is set to $m=2$. But in our solution the value of this parameter will be optimized during the multi-model’s self-organization process (learning process).

The learning procedure (identification process) operates as follows: an observation matrix, including inputs and outputs (EE based multi-model’s estimated outputs) is updated progressively until the output’s value stabilizes.

$$J(c_1, c_2, \dots, c_M) = \sum_{i=1}^M \sum_{j=1}^Q \mu_{ij}^m d_{ij}^2 \quad (10)$$

$$d_{ij} = \|z_j - c_i\| \quad (11)$$

$$\mu_{ij} = \frac{1}{\sum_{k=1}^P \left(\frac{d_{ij}}{d_{kj}}\right)^{2/(m-1)}} \quad (12)$$

$$c_i = \frac{\sum_{j=1}^Q \mu_{ij}^m x_j}{\sum_{j=1}^Q \mu_{ij}^m} \quad (13)$$

The self-organization could either be done according to “decision tree construction” (DTC) or be performed using “fuzzy clustering” (FC) partitioning strategies. The FC partitioning has been based on FCM (Fuzzy Centers Mean) algorithm. The feature space partitioning is accepted if it reduces the global error. The learning process stops if the feature partitioning process doesn’t lead to a new lower error between estimated and system’s outputs. The EE based multi-model uses the linear criterion.

An example concerning a dynamic non linear system’s behavior prediction, described by relation (14), is reported in figure 16. $u_i(t)$ are system’s inputs and $Y_s(t)$ its output. $u_i(t)$ take random value in

interval $[0, 1]$: $u_1 \in [0, 0.990]$, $u_2 \in [0, 0.985]$ and $y_s \in [-0.703, 0.294]$.

$$Y_s(t) = \frac{Y_s(t-1) [0.5 u_1(t-1) - 0.3 u_2(t-1)]}{1 + Y_s^2(t-1)} + 0.3 u_1^2(t-1) - 0.5 u_2^2(t-1) \quad (14)$$

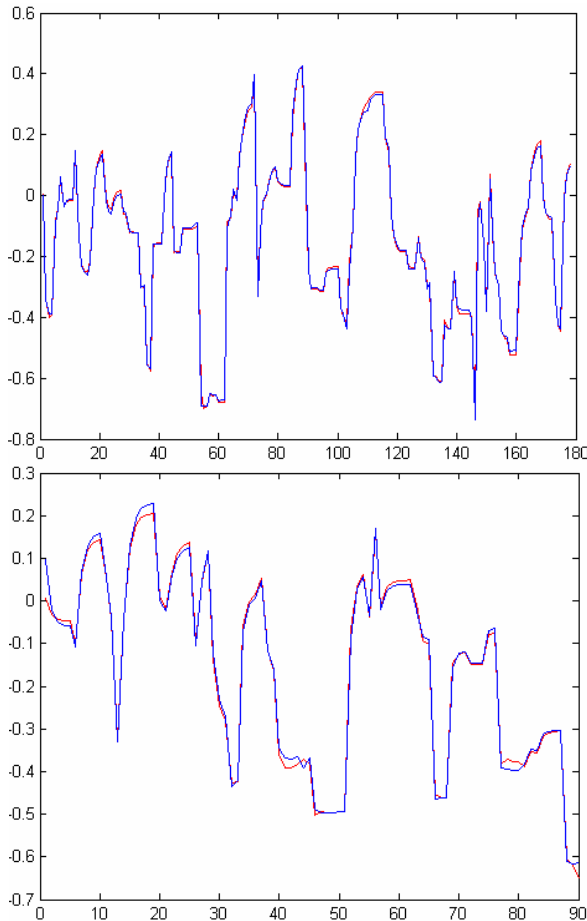


Fig. 16 - Identification (upper) and prediction (lower) results obtained for EE based modular “identifier / predictor” in the case of the nonlinear dynamic behavior described by relation (14)

A database, including systems inputs and outputs, is generated using the relation (14). The obtained database has been divided into two databases. One of them, considered as known, is used for learning phase (system’s behaviour identification) and the other one (supposed as unknown) is used for prediction.

7. CONCLUSION

If learning and generalization capabilities of ANN models appear as central requirements in intelligent systems’ design, nowadays, it is well admitted that intelligent behavior requires more sophisticated mechanisms than those performed by

these “simple” models.

On the other hand, a number of appealing features of animal’s and human’s brain, as its “modular” structure and its “self-organizing” capabilities, could be sources of inspiration in emergence of higher level artificial intelligent behavior. The main goal of this paper was to show how these primary supplies could be either exploited individually or combined in the frame of “soft-computing” in order to design intelligent artificial systems emerging higher level intelligent behavior than conventional ANN. These foremost features have inspired a set of implementations dealing with real-world applications and covering several different areas as: robotics, image processing and pattern recognition, classification and dynamic nonlinear behavior modeling (identification and prediction). The presented examples and issued results show the significant potentiality of modular connectionist architectures for designing higher level intelligent functions.

8. ACKNOWLEDGEMENTS

I would like to express my gratitude to Dr; Veronique Amarger, Dr. Abdennasser Chebira, Dr. Amine Chohra and Dr. Christophe Sabourin, working as staff members of my research team (and members of LISSI Lab.) involved in presented works. I also would like to thank Mr. Lamine Thiaw and Mr. Damien Langlois my Ph.D. and Masters Degree students, respectively for valuable discussions. Finally, I would express my gratitude to Dr. M. Rybnik, who worked (first as Ph.D. student, then as Post-Doc beside Lamine Thiaw) in my research team on aspects related to multi-modeling and self-organizing systems. Finally I would like to acknowledge Dr. R. Malti from Automation, Productic, Signal and Image Laboratory (LAPS / UMR5131) of BORDEAUX I University for active collaboration on systems identification aspects.

9. REFERENCES

- [1] A. Hanibal: VLSI building bloc for neural networks with on chip back learning. *Neurocomputing*, Vol. 5, pp. 25-37, (1993).
- [2] A. Krogh and J. Vedelsby: Neural Network Ensembles, Cross Validation and Active Learning, *Advances in Neural Information Processing Systems7*, pp. 231-238 (1995).
- [3] J. Bruske and G. Sommer: Dynamic cell structure. *Advances in Neural Information Processing Systems7*, pp. 497-504 (1995)
- [4] K. K. Sung and P. Niyogi: Active Learning for Function Approximation. *Advances in Neural Information Processing Systems7*, pp. 593-600 (1995).

- [5] K. J. Lang and M. J. Witbrock: Learning to tell two spirals apart. Connectionist Models Summer School, pp. 52-59 (1998).
- [6] M. Mayoubi, M. Schafer, S. Sinsel: Dynamic Neural Units for Non-linear Dynamic Systems Identification. *LNCS Vol. 930, Springer Verlag*, pp.1045-1051, (1995).
- [7] Multiple Model Approaches to Modeling and Control, edited by R. Murray-Smith and T.A. Johansen, Taylor & Francis Publishers, ISBN 0-7484-0595-X (1997).
- [8] S. Ernst, Hinging hyper-plane trees for approximation and identification, 37th IEEE Conf. on Decision and Control, Tampa, Florida, USA, (1998).
- [9] K. Madani, M. Rybnik, A. Chebira: Data Driven Multiple Neural Network Models Generator Based on a Tree-like Scheduler, *LNCS series, Edited by: J. Mira, A. Prieto - Springer Verlag*, ISBN 3-540-40210-1, pp. 382-389 (2003).
- [10] K. Madani, M. Rybnik, A. Chebira, Non Linear Process Identification Using a Neural Network Based Multiple Models Generator, *LNCS series, Edited by: J. Mira, A. Prieto - Springer Verlag*. ISBN 3-540-40211-X, pp. 647-654 (2003).
- [11] Ning Li, S. Y. Li, Y. G. Xi, Multi-model predictive control based on the Takagi-Sugeno fuzzy models: a case study, *Information Sciences* 165, pp. 247-263 (2004).
- [12] K. Madani, L. Thiaw, R. Malti, G. Sow, Multi-Modeling: a Different Way to Design Intelligent Predictors, *Lecture Notes in Computer Science (LNCS 3512): "Computational Intelligence and Bio-inspired Systems"*, Ed.: J. Cabestany, A. Prieto, and F. Sandoval, Springer Verlag Berlin Heidelberg, ISBN 3-540-26208-3, pp. 976 – 984, June (2005).
- [13] K. Madani, L. Thiaw, Self-Organizing Multi-Modeling: a Different Way to Design Intelligent Predictors. *Neurocomputing* ISSN 0925-2312, 2006. (Under press, to be published at the end of 2006 or at the beginning of 2007).
- [14] K. Madani, A. Chebira, D. Langlois, An Artificial Neural Network Based Approach to Mass Biometry Dilemma Taking advantage from IBM ZISC-036 Neuro-Processor Based Massively Parallel Implementation, *International Conference on Neural Networks and Artificial Intelligence (ICNNAI 2006)*, 31 May - 2 June 2006, Brest, Byelorussia, conference proceedings pp. 84-92..
- [15] L.M. Reyneri, Weighted Radial Basis Functions for Improved Pattern Recognition and Signal Processing, *Neural Processing Let., Vol. 2, No. 3, pp 2-6, May (1995)*.
- [16] G. Trémiolles (de), K. Madani, P. Tannhof, A New Approach to Radial Basis Function's like Artificial Neural Networks, *NeuroFuzzy'96, IEEE European Workshop, Vol. 6 N° 2, pp 735-745, April 16 to 18, Prague, Czech Republic, (1996)*.
- [17] Haykin S., Neural nets. A comprehensive foundation, 2on edition. Ed. Prentice Hall (1999).
- [18] M.A. Arbib (ed.), Handbook of Brain Theory and Neural Networks, 2ed. *M.I.T. Press. (2003)*.
- [19] ZISC/ISA ACCELERATOR card for PC, User Manual, IBM France, February (1995).
- [20] G. De Tremiolles, "Contribution to the theoretical study of neuro-mimetic models and to their experimental validation: a panel of industrial applications", *Ph.D. Report, University of PARIS XII, March 1998 (in French)*.
- [21] G. De Trémiolles, P. Tannhof, B. Plougonven, C. Demarigny, K. Madani, "Visual Probe Mark Inspection, using Hardware Implementation of Artificial Neural Networks, in VLSI Production", *LNCS - Biological and Artificial Computation: From Neuroscience to Technology, Ed.: J. Mira, R. M. Diaz and J. Cabestany, Springer Verlag Berlin Heidelberg, pp. 1374-1383, (1997)*.
- [22] S. Goonatilake and S. Khebbal, "Intelligent Hybrid Systems: Issues, Classification and Future Directions", in *Intelligent Hybrid Systems*, John Wiley & Sons, pp 1-20, ISBN 0 471 94242 1.
- [23] Madani K., Chebira A., "A Data Analysis Approach Based on a Neural Networks Data Sets Decomposition and it's Hardware Implementation", *PKDD 2000, Lyon, France, 2000*.
- [24] M. Vukobratovic, B. Borovac. Zero moment point – thirty five years of its live. *International Journal of Humanoid Robotics*, 2004, Vol.1 N°1, pp. 157-173.
- [25] S. Kajita, F. Kaneniro, K. Kaneko, K. Fujiwara, K. Harada, K. Yokoi and H. Hirukawa. Biped walking pattern generation by using preview control of Zero-Moment Point. *Proc. IEEE Conf. on Robotics and Automation*, 2003, pp. 1620 - 1626.
- [26] Q. Huang, K. Yokoi, S. Kajita, K. Kaneko, H. Arai, N. Koyachi, K. Tanie. Planning walking patterns for a biped robot. *IEEE Transactions on Robotics and Automation*, 2001, Vol.17, N°3, pp. 280-289.
- [27] K. Hirai, M. Hirose, Y. Haikawa, T. Takenaka. The development of honda humanoid robot. *Proc. IEEE Conf. on Robotics and Automation*, 1998, pp. 1321-1326.
- [28] C. Sabourin, O. Bruneau. Robustness of the dynamic walk of a biped robot subjected to

disturbing external forces by using CMAC neural networks. *Robotics and Autonomous Systems*, 2005, Vol.23, pp. 81-99.

- [29] C. Sabourin, K. Madani, O. Bruneau, A Fuzzy-CMAC Based Hybrid Intuitive Approach for Biped Robot's Adaptive Dynamic Walking, ICNNAI 2006 conference proceedings (to be published in June 2006).
- [30] J. S. Albus. A new approach to manipulator control: the Cerebellar Model Articulation Controller (CMAC). *Journal of Dynamic Systems, Measurement and Control*, (1975), pp. 220--227.
- [31] J. S. Albus, Data storage in the cerebellar model articulation controller (CMAC), *Journal of Dynamic Systems, Measurement and Control*, 1975, pp. 228-233.
- [32] W. T. Miller, F. H. Glanz, L. G. Kraft, CMAC: An associative neural network alternative to backpropagation, *Proceedings of the IEEE, Special Issue on Neural Networks*, vol.78, N°10, 1990, pp. 1561-1567.
- [33] C. Chevallereau, G. Abba, Y. Aoustin, F. Plestan, E.R. Westervelt, C. Canudas-de-Wit, J.W. Grizzle. RABBIT: A testbed for advanced control theory. *IEEE Control Systems Magazine*, 2003, Vol.23, N°5, pp. 57-79.
- [34] <http://robot-rabbit.lag.ensieg.inpg.fr/>.
- [35] <http://www.laas.fr/robea/>
- [36] O. Bruneau, F.B. Ouezdou. Distributed ground/walking robot interactions. *Robotica*, Cambridge University Press, 1999, Vol.17, N°3, pp. 313-323.
- [37] Bezdek, J.C. *Pattern Recognition with Fuzzy Objective Functions*. Plenum Press, N.Y., 1981.
- [38] K. Madani, L. Thiaw, Multi-Model based Identification: Application to Nonlinear Dynamic Behavior Prediction, in "*Image Analysis, Computer Graphics, Security Systems and Artificial Intelligence Applications*", Ed.: K. Saeed, R. Mosdorf, J. Pejas, O-P. Hilmola and Z. Sosnowski, ISBN 83-87256-86-2, pp. 365-375.



Prof. Kurosh Madani Received his Ph.D. degree in Electrical Engineering and Computer Sciences from University PARIS XI (PARIS-SUD), Orsay, France, in 1990. From 1989 to 1990, he worked as assistant professor at Institute of Fundamental Electronics of PARIS XI University. In 1990, he joined Creteil-Senart Institute of Technology of University PARIS XII – Val de Marne, Lieusaint, France, where he worked from 1990 to 1998 as assistant professor. In 1995, he received the DHDR Doctor Habilitate degree (senior research Dr. Hab. degree) from University PARIS XII – Val de Marne. Since 1998 he works as Chair Professor in Electrical Engineering of Senart Institute of Technology of University PARIS XII – Val de Marne. From 1992 to 2004 he has been head of Intelligence in Instrumentation and Systems Laboratory of PARIS XII – Val de Marne University located at Senart Institute of Technology. Since 2005, he is head of one of the three research teams of Image, Signal and Intelligent Systems Laboratory (LISSI / EA 3956) of PARIS XII University. He has worked on both digital and analog implementation of processors arrays for image processing by stochastic relaxation, electro-optical random number generation, and both analog and digital Artificial Neural Networks (ANN) implementation. His current research interests include large ANN structures behavior modeling and implementation, hybrid neural based information processing systems and their software and hardware implementations, design and implementation of real-time neuro-control and neural based fault detection and diagnosis systems. Since 1996 he is a permanent member (elected Academician) of International Informatization Academy. In 1997, he was also elected as Academician of International Academy of Technological Cybernetics.