ISSN 1870-4069

Recurrent Trainable Neural Networks for Complex Systems Identification: A Hybrid System Approach

Juan-Eduardo Velázquez-Velázquez[1], Rosalba Galván-Guerra[1], Ieroham Baruch[2], Silvestre Garcia-Sanchez[1]

[1] UPIIH-IPN, Pachuca de Soto, Hgo., Mexico

[2] CINVESTAV-IPN, Department of Automatic Control, Mexico

{jvelazquezv,rgalvang,sgarcias}@ipn.mx, baruch@ctrl.cinvestav.mx

Abstract. This paper is devoted to the development of an Identification Framework for unknown Complex Systems. The proposal is based on Recurrent Trainable Neural Networks following a Hybrid System approach. The complex system is identified by using hybrid input-output data defined by a given set of switching hypersurfaces. The effectiveness of the proposed approach is shown using a commutable pendulum with chaotic behavior.

Keywords: Complex systems identification, hybrid systems, recurrent trainable neural networks.

1 Introduction

It is well known that interpreting and predicting the behavior of complex dynamical systems is challenging, mainly because the causes and effects are not obviously related. One way to establish the direct and indirect relationships of causes and effects in a complex system is via an identification framework. The identification problem consists of obtaining a model that allow us to infer how the system will respond to other inputs that we have not yet measured by approximating the output trajectory of the system. In the complex systems context, there are several approaches with quite different viewpoints on system modeling: dynamical systems, discrete event systems, cellular automata models, neural network models, finite state machines, cognitive maps, multi-agent models and Hybrid Systems (HS) (see [3] and the references there in). These paradigms differ, rather, by concepts and views on the problems and approaches to solve them, than the applications areas [1, 7, 10, 13, 16, 17]. In this contribution, we follow the HS approach.

The identification theory for continuous state systems is well developed in the literature [18]. However, HS add extra complexity due to the interaction of continuous and discrete dynamics. The identification of complex systems following a HS approach has been devoted to the identification of switched affine and piecewise affine models. The main issues and difficulties connected with HS identification are discussed in [22].

Normally the identification is based on statistical techniques that need a reasonable amount of data and assume stationarity: they require that the underlying system does not change its parameters over time. If the parameters of the system are drifting or externally switched from time to time, the statistical algorithms can be applied to short segments of the data, thereby monitoring changes in the characteristic quantities. However, such methods may suffer from the curse of dimensionality and other statistical problems that arise when estimating from few data points.

Instead of a statistical technique, in this contribution we use a Neural Network (NN) technique that is very effective for identification and control of nonlinear systems when partial or null information about systems is available [20]. It is worth to mention, that the idea to incorporate NN for identification of complex systems following a HS approach has been previously addressed. In one side, some proposals incorporate the NN identifier as a global model, losing the nature of HS [2, 15, 19]. In the other side, it is assumed that the sequence of locations is known or that it is well established by considering the discrete-time dynamics at the higher level of a hierarchical framework [12, 14]. These kinds of assumptions could not be valid because the dynamic behavior of a HS is strongly influenced by discontinuities in its system trajectories [9, 23].

In this paper an identification methodology based on Recurrent Trainable NN (RTNN) to model complex systems, in a HS approach, is presented; following a gray box retaining the characteristic behavior of HS and introducing the Hybrid RTNN (HRTNN).

The outline of the paper is as follows: In Section 2 we describe the hybrid complex systems investigated in this contribution and the identification problem is formulated. In Section 3 we present the HRTNN description. In order to show the validity of the proposed HRTNN several simulations of a commutable ideal pendulum, presenting chaos behavior, are developed in Section 4. Section 5 concludes the paper.

2 Modeling Framework and Problem Formulation

In this contribution, a complex system is considered as a system composed by simple subsystems; where the subsystems are active in a set of locations. Then the modeling framework for the complex system identification is based on Hybrid Dynamical Systems (HDS) structures. The HDS are characterized by a set of subsystems interconnected in a discrete manner. The applicability of this approach can be seen in the fact that several types of hybrid systems are used in chemical, bioengineering, aerospace and electronics industries, among others to model complex control systems [7, 9, 10, 17]. The Nonlinear HDS (NHDS) discussed in this paper are represented by its hybrid state equation

$$\dot{x}(t)=\sum_{i=1}^{r} \beta_{[t_{i-1},t_i)}(t) f_{q_i}\big(t,x(t),u(t)\big) \text{ a.e. on } t \in [0,t_f], \qquad (1)$$

and by the hybrid output vector

$$\bar{y}(t)=\sum_{i=1}^{r}\beta_{[t_{i-1},t_i)}(t)h_{q_i}\big(t,x(t),u(t)\big)\text{ a.e. on, } t\in[0,t_f] \qquad (2)$$

with available sample-data outputs y(k). $\beta_{[t_{i-1},t_i)}(\cdot)$, is the characteristic function of the interval $[t_{i-1},t_i)$ for i=1,…,r, i.e.

$$\beta_{[t_{i-1},t_i)}(t)=\begin{cases}1, & \text{if } t\in[t_{i-1},t_i), i=1,\dots,r\\ 0, & \text{otherwise}\end{cases}.$$

Here $q_i\in Q$ represents a location, Q is a finite set of discrete states (called *locations*), $x(\cdot)\in X$ is an admissible state trajectory, $X=\{X_q\}$ with $X_q\subseteq R^n$; $u(\cdot)\in U$ is an admissible input signal and $U\subseteq R^m$ is a set of admissible measurable bounded functions; $f_q:[0,t_f]\times X_q\times U\to R^n$ defines a family of velocity vector fields $F=\{f_q\}$;$h_q:[0,t_f]\times X_q\times U\to R^p$ are diffeomorphisms which define a family of vector fields $H=\{h_q\}$; and $\bar{y}(\cdot)\in Y$ is an admissible output trajectory, $Y=\{Y_q\}$ is a collection of output sets with $Y\subseteq R^p$. The interest of considering NHDS with sample-data outputs is twofold: these models can represent a wide range of systems of practical interest and sample-data outputs can be interpreted as a result of application of a quantified procedure to an original continuous model. Note that, in difference to $\bar{y}(t)$ the stepwise value y(k) is a measurable output of the system under consideration. We assume that the dynamic transitions between two subsystems are characterized by the assembly of switching pairwise disjoint hypersurfaces $M_{q,q'}:=\{x\in R^n:m_{q,q'}(x)=0\}$, where $m_{q,q'}:R^n\to R$, $q,q'\in Q$ are smooth functions with nonzero gradients. The given hypersurfaces $M_{q,q'}$ represents the *switching sets* at which a switch from location *q* to location *q'* can take place. We say that a location switching from *q* to *q'* occurs at a *switching time* $t_{sw}\in[0,t_f]$. We consider NHDS with $r\in R$ switching times: $0<t_0<t_1<\cdots<t_{r-1}<t_r=t_f$. Note that the sequence of switching times $\{t_i\}$ is not defined a priory, neither the sequence of locations: $\{q_i\}$. In order to state the problem, that we are interested on, let us introduce the following definition.

Definition 1: A *hybrid trajectory* of NHDS is a 3-tuple $X=\{x(\cdot),q_i,t_i\}$ such that for each i=1,…,r and every admissible input $u(\cdot)\in U$ we have

- $x(0)=x_0\notin\bigcap_{q,q'\in Q}M_{q,q'}$ and $x_i(\cdot)=x(\cdot)|_{(t_{i-1},t_i)}$ is an absolutely continuous function on (t_{i-1},t_i) continuously prolongable to $[t_{i-1},t_i]$, i=1,…,r
- $x(t_i)\in M_{q,q'}$ for i=1,…,r;
- $\dot{x}(t)=f_{q_i}(t,x_i(t),u(t))$ for almost all $t\in[t_{i-1},t_i]$

Observe that given an admissible signal input u(·), the physical attributes of a NHDS governed by its state equation (1) are transformed by (2) into responses as system outputs. Here, we suppose that the hybrid output retains the state behavior, i.e., the transition from one location to another in the space state cause a transition in the output response. We assume that the transitions at the output are characterized by the assembly of switching pairwise disjoint hypersurfaces

$$N_{q,q'}:=\{y\in R^p:n_{q,q'}(y)=0\}, \qquad (3)$$

where $n_{q,q'}: R^p \rightarrow R$, $q,q' \in Q$ are smooth functions with nonzero gradients defined as: $n_{q,q'}(y) = m_{q,q'}\left(h_{q'}^{-1}(x)\right)$. We now could formulate the identification problem as:

Problem 1: Assume that the NHDS (1)-(2) is unknown. The identification problem consists of obtaining a model that allow us to infer how the NHDS will respond to other inputs that we have not yet measured by approximating the output hybrid trajectory of the system NHDS. That is, for an experiment with length t_f, we want to determine a model with a hybrid output $\hat{Y}=\{\hat{y}(\cdot),\{\hat{q}_i\},\{t_i\}\}$ which approximate the NHDS response, visited during the experiment, by using only the observed data $\{u(t),y(k),t\in[0,t_f]\}$.

3 HTRNN Description

During the last decade considerable research has been devoted towards developing RNN models applied for identification and control of complex nonlinear plants [4, 5]. Under the premise of Problem 1, we follow the structure of NHDS to approach complex systems; the dynamic transitions from one location to another observed in the output response are described by the equations of mathematical physics, but the response of the subsystems is modeled by neural networks. Then, to describe the global complex system's behavior, we use an arbitrary interconnection between subsystems with discrete event, characterized by switching hypersurfaces. It means that the transition locations $\{q_i\}$ and the sequences of switching times $\{t_i\}$ are not previously defined, only the hypersurface (3) are considered. It is worth to mention, that to develop a computational identification it is necessary to take some measurements in a time interval. Then, it is pertinent to consider a discrete RTNN to performs this task. It is well known that during the sampling process it is possible to lose a location transition. However, if the direction of the trajectory is considered it is possible to detect when a hypersurface $n_{q,q'}$ has been triggered.

The RTNN topology and its associated BP learning rule are described in vector-matrix form, [4, 5, 6], as:

$$\begin{gathered}
\hat{x}(k+1) = A\hat{x}(k) + B\hat{u}(k); \\
z(k) = G[\hat{x}(k)]; \;\; v(k) = Cz(k); \;\; \hat{y}(k) = F[v(k)]; \\
w(k+1) = w(k) + \eta\Delta w(k) + \alpha\Delta w(k-1); \\
e(k) = y(k) - \hat{y}(k); \;\; F'[\hat{y}(k)] = [1-\hat{y}^2(k)]; \\
e_3(k) = G'[z(k)]e_2(k); \;\; e_2(k) = C^T(k)e_1(k); \\
G'[z(k)] = [1-z^2(k)]; \;\; \Delta B(k) = e_3(k)\hat{u}^T(k); \\
\Delta A(k) = e_3(k)\hat{x}^T(k); \;\; \Delta C(k) = e_1(k)z^T(k);
\end{gathered} \tag{4}$$

where $\hat{y}(k)$, $\hat{x}(k)$ and $\hat{u}(k)$ are output, state and input vectors of the RTNN with dimension p, N, (m+1); here $u^T = [u; u_0]$, where u is the real plant input vector with dimension m and u_0=-1 is a threshold entry; y is the plant output vector with dimension p, considered as a RTNN reference; A is N×N block-diagonal matrix, defined by $A =$

$\mathrm{block-diag}(A_i)$; $B = [B_1; B_0]$ and $C = [C_1; C_0]$ are N×(m+1) and L×(N+1) augmented weight matrices; B_0 and C_0 are N×1 and L×1 threshold weights of the hidden and output layers; F[·] and G[·] are vector valued activation functions of type tanh(·); $F'[\cdot]$ and $G'[\cdot]$ are the derivatives of the activation functions; W is a general weight denoting each weight matrix (C, A, B) in the RTNN model, to be update; Δw (ΔC,ΔA,ΔB), is the weight correction of W (C, A, B); η,α are learning rate parameters; e, e_1, e_2, e_3 are error vector with appropriate dimensions, predicted by the adjoint RTNN model. It is well known that the stability of the RTNN model is assured by the activation functions (-1,1) bounds and by the local stability weight bound conditions, $|A_i| < 1$.

3.1 Gray-Box Approach

Following a gray-box approach we only assume that r location transitions in the hybrid output can take place. These transitions are characterized by the assembly of the given switching hypersurfaces (3). In this case we propose the use of the RTNN defined by (4) together with the switching hyper surfaces (3) used like a supervisor layer. This supervisor layer defines the switching instants between the RTNN_{q_i} to RTNN_{q_j}, i.e., when condition $\eta_{q_i,q_j}(y(k)) = 0$ is fulfilled a transition between the neural networks occurs. In Fig. 1 we introduce the HRTNN model.

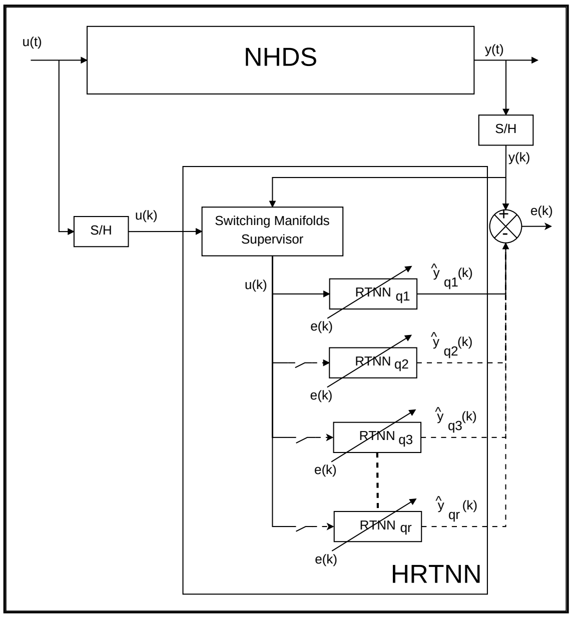

Fig. 1 HRTNN Identifier of the Complex System: NHDS approach.

Remark 1: Note that in an arbitrary switching time, we are not updating the initial weights neither the initial state of the new located RTNN_{q_i}. Then to avoid any impulse (error) in the updating of the weights and the state of the actual RTNN_{q_j}, we propose to add an Auxiliary Neural Network (ANN), showed in Fig. 2.

This ANN can be of any suitable topology. For simulation purposes we propose the use of a RTNN with the same topology as the RTNN_{q_j}. Observe that the convergence of this approach in a specific location is assured by the convergence of the RTNN training algorithm [4], [5], [6]. However, this convergence is subject to an enough amount of training data, and the absence of the Zeno behavior. Also note that the switching times can be computed analytically as it is shown in [11]. However, it is important to remark that the switching manifolds rule these switching times, so with the knowledge of these hypersurfaces we could know when a transition happens. The knowledge of the manifolds at the hybrid output along with an appropriated weight actualization strategy at the switching times makes possible to achieve a finite time convergence or at least a global convergence as in [12].

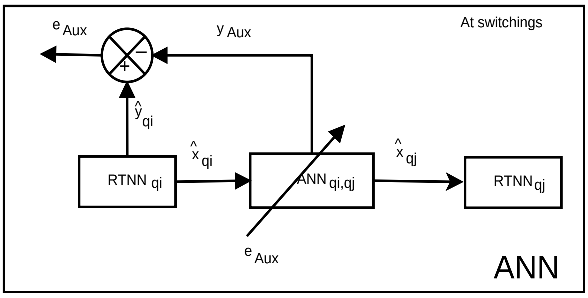

Fig. 2 Auxiliary Neural Network Topology.

Remark 2: The gray-box strategy is actually a hybrid strategy, which allows us to identify the nonlinear system like a hybrid one. To complete the hybrid strategy as a black-box approach, we need structures that allow us to identify the switching manifolds (see [22] for details).

4 Simulation Results: Chaotic Behavior of an Ideal Commutable Pendulum

Here, the applicability of our approach is illustrated by a commutable pendulum example, taken from [21]. In an ideal form, the general mathematical model of this system is given by the following equations:

$$\Sigma = \begin{cases} \Sigma_1: \ \ddot{x} + x = 0 & x, \ddot{x} \epsilon R, \\ \Sigma_2: \ \ddot{x} + \omega^2 x = 0 & x, \ddot{x} \epsilon R, \end{cases}$$

where ω is the pendulum frequency and the switching manifolds are defined by:

$$M_{1,2}: x < 0, \qquad \dot{x} = 0,$$
$$M_{2,1}: x = 0, \qquad \dot{x} > 0.$$

Under some parameter conditions, the commutable pendulum presents a chaotic complex behavior, when the pendulum frequency is greater than 1[rad/s], [21]. In particular, the system commutes from a stable behavior to an unstable one. However, this chaotic nature does not mean that the whole system is unstable (see [8] for more details). With the aim to show the effectiveness of the HRTNN in the presence of some specific complex behavior, we identify the commutable pendulum with the frequency $\omega = 3.5\big(1 + x(t_i)\big)$, i=1,3,… Note that under this condition the system presents two limit cycles and it is highly unstabilizable (see the phase plane of Fig. 3). To identify the complex behavior of this specific commutable pendulum, we applied the two identification strategies, using the oscillatory input: $u(t) = \frac{1}{10}\sin\left(\frac{2\pi t}{5}\right)$.

Following the Gray box strategy without any actualization at the switching times, with the RTNN with topology (2, 4, 2), and learning parameters: η=0.6; α=0.001, we train this network for 200 seconds with a sampling rate of 0.01, obtaining the results shown in Fig. 4, where $\frac{1}{2}e^2(t)$ is the mean square error. Note that due to the discontinuity between the states of both neural networks, there are impulses at the switching times as we have expected previously (recall Remark 1).

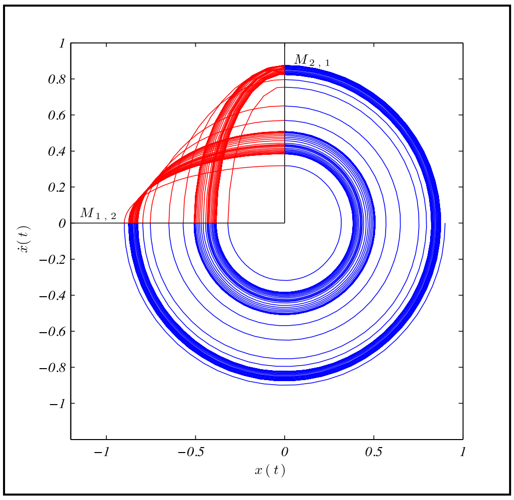

Fig. 3 Phase Plane: $\omega = 3.5\big(1 - x(t_i)\big)$, $i\,mod2 = 1$.

Finally, we follow the Gray Box Strategy with weight and state actualization at the switching times. We use the same topology and learning parameters as in the latter case. As a result, we obtain the graphics, shown in Fig. 5. Note that there are not impulses at the switching times and that the error is reduced. Observe that even when the common sense says that we can only identify stable systems; it is possible to identify unstable subsystems if the global system remains stable. The advantage of using the Gray Box Strategy is that we could identify the chaotic behavior of the system without changing the nature of the original system. If we compare these two strategies, we can see the following:

- The Gray Box Strategy without any actualization at the switching times retains the hybrid nature of the system, but it has discontinuities.
- The Gray Box Strategy with actualization, retains the hybrid nature of the system, and eliminates the discontinuities presented in the Gray Box Strategy without actualization.

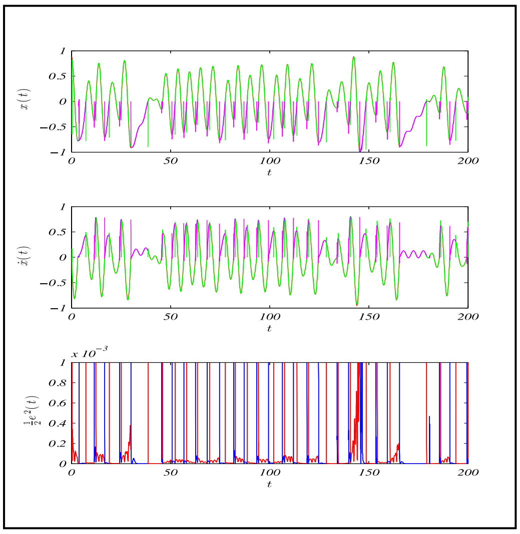

Fig. 4 Gray Box Strategy Identification ($\sum_1$: red line, $\sum_1$ identification: green line, $\sum_2$: blue line, $\sum_2$ identification: magenta line).

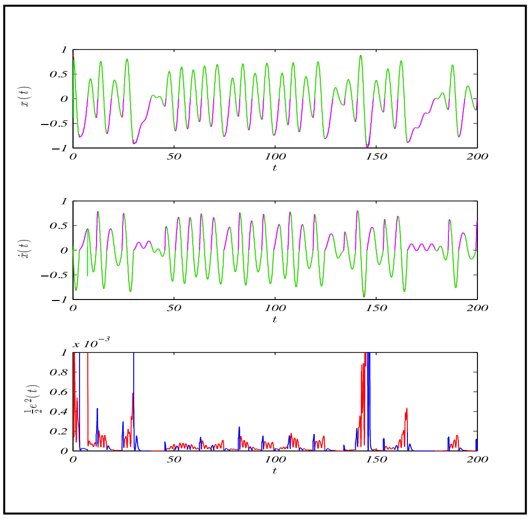

Fig. 5 Gray Box Strategy with ANN Identification ($\sum_1$: red line, $\sum_1$ identification: green line, $\sum_2$: blue line, $\sum_2$ identification: magenta line).

5 Conclusion

This paper presents an intelligent approach to identify complex systems following a hybrid structure approach. We introduced the topology of the Hybrid Trainable Recurrent Neural Network. Two strategies are presented: Gray Box and Gray Box strategy with actualization of the neural network state and weights. The Gray Box strategy without actualization can identify the system into the locations and preserve the complex nature of the system, but it has several errors at the switching times. The Gray Box strategy with weight and state actualization overcomes the disadvantages of the Gray Box strategy without actualization, making it a suitable option to identify a complex system with a hybrid nature even in the presence of chaotic behavior.

References

1. P. J. Antsaklis, X. Koutsoukos, J. Zaytoon: On Hybrid Control of Complex Systems: A Survey. Eur. J. Automat., Vol. 32, No. 9, pp. 1023–1045 (1998)
2. D. Back,T.P. Chen: Approximation of hybrid systems by neural networks. In: Proceedings of Int. Conf. on Neural Inf. Proc. and Intelligent Inf. Syst., Springer-Verlag, pp. 326–329 (1997)
3. A. Bagdasaryan: Systems Theoretic Techniques for Modeling, Control, and Decision Support in Complex Dynamic Systems. In: CoRR abs/1008.0775 (2010)
4. I. S. Baruch, C.R. Mariaca-Gaspar, J. Barrera-Cortes: Recurrent Neural Network Identification and Adaptive Neural Control of Hydrocarbon Biodegradation Processes. In: Recurrent Neural Networks, Xiaolin Hu and P. Balasubramaniam, (Eds.), Croatia: I-Tech/ARS Press, ch. 4, pp.61–88 (2008)
5. I.S. Baruch, C.R. Mariaca-Gaspar: A Levenberg-Marquardt Learning Algorithm Applied for Recurrent Neural Identification and Control of a Wastewater Treatment Bioprocess. Int. Journal of Intelligent Systems, vol. 24, pp. 1094–1114 (2009)
6. I. S. Baruch, R. Galvan-Guerra, S. M. Hernandez: Distributed Parameter Bioprocess Plant Identification and I-Term Control Using Decentralized Fuzzy-Neural Multi-Models. In Advances in Reinforcement Learning, Abdelhamid Mellouk (Ed.), Croatia: InTech, ch. 23, pp. 421–450 (2011)
7. M.S. Branicky, V.S Borkar, S.K. Mitter: A Unified Framework for Hybrid Control: Model and Optimal Control Theory. IEEE Trans. on Automatic Control, vol. 43, pp. 31–45 (1998)
8. M.S Branicky: Multiple Lyapunov functions and other analysis tools for switched and hybrid systems. IEEE Trans. on Automatic Control, Vol. 43, No. 4, pp. 475–482 (1998)
9. M. Buss, M. Glocker, M. Hardt, O. von Stryk, R. Bulirsh, G. Schmidt: Nonlinear Hybrid Dynamical Systems: Modeling, Optimal Control, and Applications. In: Lecture Notes in Control and Information Sciences, vol. 279, Springer-Verlag (2002)

10. C. Cassandras, D.L. Pepyne, Y. Wardi: Optimal Control of a Class of Hybrid Systems. IEEE Transaction on Automatic Control, vol. 46, pp. 398–415 (1998)
11. R. Galvan-Guerra, J.E. Velázquez-Velázquez, V. Azhmyakov: Explicit Switching Times and Location Detection for Linear Hybrid Systems. In: Proc. of the 7th Int. Conf. on Electrical Engineering, Computing Science and Automatic Control, Tuxtla Gutiérrez, Mexico, pp. 134–139 (2010)
12. A. García, I. Chairez, A. Poznyak: Hybrid DIfferential Neural Network Identifier for Partially Uncertain Hybrid Systems. In: Recent Advances in Intelligent Control Systems, Springer-Verlag (2009)
13. A. Hassibi, S. P. Boyd, J. P. How: Control of Asynchronous Dynamical Systems with Rate Constraints on Events. In: Proceedings of the IEEE Conference on Decision and Control, 2:1345–1351 (1999)
14. T. Hayakawa, W. M. Haddad, K. Y. Volyanskyy: Neural network hybrid adaptive control for nonlinear uncertain impulsive dynamical systems. Nonlinear Analysis: Hybrid Systems, vol. 2, pp. 862–874 (2008)
15. W. Holderbaum: Application of Neural Network to Hybrid Systems with Binary Inputs. IEEE Trans. on Neural Networks, vol. 18(4), pp. 1254–1261 (2005)
16. S. Kamalasadan: A Novel Multi-Agent Controller for Dynamic Systems based on Supervisory Loop Approach. Engineering Letters, EL_14_2_10 (2007)
17. A.B. Kurzhanskii, P. Varaiya: New Directions in Applications Control Theory. In: Lecture Notes Control Inform. Sci., Berlin: Springer, vol. 321, pp. 193–205 (2005)
18. L. Ljung: System Identification: Theory for the User. Prentice Hall, New Jersey, (1987)
19. N. Messai, B. Riera, J. Zaytoon: Identification of a class of hybrid dynamic systems with feed-forward neural networks: About the validity of the global model. Nonlinear Analysis: Hybrid Systems, vol. 2, pp. 773–785 (2008)
20. K.S. Narendra, K. Parthasarathy: Identification and Control of Dynamic Systems Using Neural Networks. IEEE Trans. Neural Networks, vol. 1, pp. 4–27 (2008)
21. V.R. Nosov: Complex Dynamics in an Ideal Commutable Pendulum. In: Proceedings of the 5th International Conference on Electrical Engineering, Computing Science and Automatic Control, Mexico, Mexico (2008)
22. S. Paoletti, A. Juloski, G. Ferrari-Trecate, R. Vidal: Identification of Hybrid Systems: A Tutorial. European Journal of Control, Vol. 13/2-3, pp. 242–260 (2007)
23. A. Van der Schaft, H. Schumacher: An Introduction to Hybrid Dynamical Systems. Berlin: Springer (2000)