Control of Unknown Multivariable Systems Based on the Self-Organizing Maps

Jeongho Cho, Jose C. Principe and Mark A. Motter[§]
Computational NeuroEngineering Lab., University of Florida, Gainesville, FL, USA
[§]Dynamics and Control Branch, NASA Langley Research Center, Hampton, VA, USA
{jeongho,principe}@cnel.ufl.edu, m.a.motter@larc.nasa.gov

Abstract – *This paper is concerned with the development and analysis of a nonlinear approach to modeling and control of nonlinear complex systems. In particular, the problem of designing a mathematical model of a nonlinear plant using only observed data is considered. For the identification of the plants, the concept of multiple models with switching is employed in order to simplify both the modeling and the controller design. For this reason, a Self-Organizing Map (SOM) is utilized to divide the operating region into local regions as a modeling infrastructure to construct local models. Based on the identified multiple models, the problem of designing controllers is discussed. The effectiveness of the proposed approach is shown through experiments for modeling and control of complex nonlinear plant. Its comparison with neural networks-based alternatives, Time Delay Neural Network (TDNN), shows clear advantages of local modeling and control in terms of performance.*

Key words – MIMO Systems, Self-Organizing Map, Inverse Control, Sliding Mode Control

1 Introduction

The identification of nonlinear dynamical systems has received considerable attention since it is an indispensable step towards analysis, simulation, prediction, monitoring, diagnosis, and controller design for nonlinear systems [1]. In particular, the problem of designing a mathematical model of a nonlinear plant using only observed data has attracted much interest, both from an academic and an industrial point of view. During the past few years, neural networks as a global model have been suggested for nonlinear dynamical black-box modeling and successfully applied to the prediction and modeling of nonlinear processes [13].

Global models, however, have shown some difficulties in cases when the dynamical system characteristics vary considerably over the operating region, effectively bringing the issue of time varying parameters (or nonlinearity) into the design. On the other hand, local modeling derives a model based on neighboring samples in the operating space to characterize some operating point or similar feature [2,4,6,7]. If a function f to be modeled is complicated, there is no guarantee that any given global representation will approximate f equally across all space. Moreover, nonlinear models are too complex to be used for controller design [1]. Thus, nonlinear control methods cannot serve all needs of real industrial control problems.

In this case, the dependence on representation can be reduced using local approximation where the domain of f is divided into local regions and a separate model is used for each region [3,4,7]. In a number of local modeling applications, a Self-Organizing Map (SOM) has been utilized to divide the operating regions into local regions [3,4,12]. The SOM is particularly appropriate for switching, because it converts complex, nonlinear statistical relationships of high-dimensional data into simple geometric relationships that preserve the topology in the feature space [5]. Thus the role of the SOM is to discover patterns in the high dimensional state space and divide that space into a set of regions represented by the weights of each Processing Element (PE).

Linear models and associated techniques for linear control design are typically used to control the plant under certain specific operating conditions. This type of control is only valid in a small region around the operating point. For that reason, the concept of multiple models with switching, according to a change in dynamics, has been an area of interest in control theory in order to simplify both modeling and controller design [4,7]. The objective of this study, therefore, is to investigate if it is possible to obtain a better result in extending the formulation of the control problem from using just one global model to using several internal models. Thus a multiple modeling approach is presented and techniques to design controllers based on these model structures are developed.

2 Piece-wise linear models for control

The idea of multiple modeling is to approximate a nonlinear system with a set of relatively simple local models valid in certain operating regions [2,7]. The SOM is employed as a modeling infrastructure to construct the local models. It provides a codebook representation of the plant dynamics and organizes the different dynamic regimes in topological neighborhoods. Thus we can create a set of models that are local to the data in the Voronoi tessellation created by the SOM.

The SOM is trained to position the local models in the embedded output space ($\psi_{y,k} = [y_k, y_{k-1}, \cdots, y_{k-dy}]^T$). At any time instant, the model representing the plant dynamics is chosen by the SOM depending on the history of the plant and then incorporated with the embedded control input space ($\psi_{u,k} = [u_k, u_{k-1}, \cdots, u_{k-du}]^T$). After the operating regions are divided by the SOM the underlying dynamics f is then approximated as $f \approx \bigcup_{i=1}^{N} f_i$, where N is the number of operating regions. N local predictive ARX models $f_1, \cdots, f_N$ of the plant are described by

$$f_i(\psi_k) \approx \sum_{j=0}^{dy} a_{i,j} y_{k-j} + \sum_{j=0}^{du} b_{i,j} u_{k-j}, \quad i = 1, \cdots, N \tag{1}$$

where $a_{i,j}$ and $b_{i,j}$ are the parameters of the i^{th} model. Then, when each PE of the SOM is extended with a local model it can actually learn the mapping $\hat{y}_{k+1} = f_i(\psi_{y,k}, \psi_{u,k})$ in a supervised way (See [12] for details). Our proposed modeling methodology is summarized as follows: first, the delayed version of input-output joint space is decomposed into a set of operating regions that are assumed to cover the full operating space. Next, for each operating region we choose a simple linear ARX model to capture the dynamics of the region. Consequently, a nonlinear nonautonomous system is approximated by a concatenation of local linear models.

3 Multiple model based control

Researchers have been interested in control of nonlinear systems for a very long time. Progress in nonlinear control design, however, has been difficult because of the intrinsic complexity of the problem [2]. In order to overcome these difficulties in designing controllers for nonlinear systems, a simplified control method that keeps the advantage of the conventional approach is proposed, i.e., the SOM is explored as a modeling infrastructure, and the controllers are built based on SOM-based multiple models.

3.1 Multiple inverse control

Now we discuss the control problem for the local linear model using an inverse control framework [12]. The central advantage of such a framework is that an inverse model can be used directly to build a feed-forward controller. Thus, for the desired behavior, the controller just asks the model to predict the action needed. The system identification block has N predictive models denoted by $\{f_i\}_{i=1}^{N}$, in parallel. Corresponding to each model f_i, a controller C_i is designed such that C_i achieves the control objective for f_i. From (1), under the assumption that b_o is invertible, the control law of an inverse controller, C_{i^o}, for the winning model, f_{i^o}, can be directly calculated as $u_{i^o,k} = b_{i^o,0}^{-1}\left(d_{k+1} - \sum_{j=0}^{dy} a_{i^o,j} y_{k-j} - \sum_{j=1}^{du} b_{i^o,j} u_{k-j}\right)$. Therefore, at time instance *k*, the control $u_{i^o,k}$ can be obtained, if the future target of y_k, d_{k+1}, is known. One advantage of this scheme is its simplicity and fast convergence to get the desired response. Another advantage is that the dynamic space is decomposed in the appropriate switching among very simple linear models, which leads to accurate modeling and controls. On the other hand, creating a set of models by embedded input and output may cause serious problem in the presence of large noise or outliers since the wrong predictive model due to noise may cause poor control.

3.2 Multiple sliding mode control

While classical control techniques have produced many highly reliable and effective control systems, great attention has been devoted to the design of variable structure control systems (VSCS). The central advantage of the sliding mode control strategy is that it is an effective robust control strategy for incompletely modeled or uncertain systems [8,9,10]. Thus, the feature of the proposed control scheme is that the robustness for disturbances can be obtained by the simple control logic based on the linear model for each region. Another feature of the strategy is that it guarantees convergence of the system output to a vicinity of the predetermined, fixed plane in finite time, specified a priori by the designer. Consider one of the local l multi-input multi-output models f_i of the plant f,

$$
\begin{aligned}
y_{1,k+1} &= a_{1,1}y_{1,k} + \cdots + a_{1,m}y_{1,k-m+1} + b_{1,1,1}u_{1,k} + \cdots + b_{1,1,n}u_{1,k-n+1} \\
&\quad + \cdots + b_{1,l,1}u_{l,k} + \cdots + b_{1,l,n}u_{l,k-n+1} \\
&\vdots \\
y_{l,k+1} &= a_{l,1}y_{l,k} + \cdots + a_{l,m}y_{l,k-m+1} + b_{l,1,1}u_{1,k} + \cdots + b_{l,1,n}u_{1,k-n+1} \\
&\quad + \cdots + b_{l,l,1}u_{l,k} + \cdots + b_{l,l,n}u_{l,k-n+1}
\end{aligned} \tag{2}
$$

where available measurement $u, y \in \Re^l$ and the embedding dimensions of output and input are m and n, respectively. The state-space model of (2) can be written as

$$\vec{x}_{k+1} = \Phi\vec{x}_x + \Lambda_1\vec{u}_k + \Lambda_2\vec{u}_{k-1} + \cdots + \Lambda_n\vec{u}_{k-n+1} \tag{3}$$

where $\vec{x}_k = [y_{1,k-m+1}, \cdots, y_{1,k-1}, y_{1,k}, \cdots, y_{l,k-m+1}, \cdots, y_{l,k-1}, y_{l,k}]^T \in \Re^{m \cdot l}$ is the system state vector which is available for measurement, $\vec{u}_k = [u_{1,k}, \cdots, u_{l,k}]^T \in \Re^l$ is the control effort, $\Phi \in \Re^{m\cdot l \times m\cdot l}$ is a block diagonal matrix, $\Phi = diag(\Phi_i), i = 1, \cdots, l$, and $\Lambda_1, \cdots, \Lambda_n \in \Re^{m\cdot l \times l}$ have the following forms:

$$\Phi_i = \begin{bmatrix} 0 & 1 & & 0 \\ 0 & \ddots & \ddots & 0 \\ & & 0 & 1 \\ a_{i,m} & a_{i,m-1} & \cdots & a_{i,1} \end{bmatrix} \quad \Lambda_1 = \begin{bmatrix} \mathbf{0} \\ b_{1,1,1}, \cdots, b_{1,l,1} \\ \vdots \\ \mathbf{0} \\ b_{l,1,1}, \cdots, b_{l,l,1} \end{bmatrix} \quad \Lambda_2 = \begin{bmatrix} \mathbf{0} \\ b_{1,1,2}, \cdots, b_{1,l,2} \\ \vdots \\ \mathbf{0} \\ b_{l,1,2}, \cdots, b_{l,l,2} \end{bmatrix} \quad \cdots \quad \Lambda_n = \begin{bmatrix} \mathbf{0} \\ b_{1,1,n}, \cdots, b_{1,l,n} \\ \vdots \\ \mathbf{0} \\ b_{l,1,n}, \cdots, b_{l,l,n} \end{bmatrix}$$

Also defining the tracking error vector as $\vec{e}_{k+1} = \vec{r}_{k+1} - \vec{x}_{k+1}$, where $\vec{r}_{k+1} = [d_{1,k-m+2,} \cdots, d_{1,k+1}, \cdots, d_{l,k-m+2,} \cdots, d_{l,k+1}]^T \in \Re^{m \cdot l}$ represents a given desired signal vector assumed to be bounded, the sliding surface can be defined in the space of the tracking error vector given by

$$S_k = C^T \vec{e}_k \tag{4}$$

where

$$C^T = \begin{bmatrix} c_{1,1} & \cdots & c_{1,m} & & & \mathbf{0} \\ & & & \ddots & & \\ & \mathbf{0} & & c_{l,1} & \cdots & c_{l,m} \end{bmatrix}.$$

The ideal quasi sliding mode defined by Gao et al [9] occurs if there exists an integer k_q such that the trajectory crosses the sliding surface in succession keeping on within a specified band for all $k \ge k_q$ (k_q is the time when the trajectory hits the manifold). Substituting (3) in the ideal quasi-sliding mode condition, $S_{k+1} = S_k = \mathbf{0}$, yields the equivalent control

$$\vec{u}_k^{eq} = (C^T \Lambda_1)^{-1} \left\{ C^T (\vec{r}_{k+1} - \Phi \vec{x}_k) - C^T \Lambda_2 \vec{u}_{k-1} - \cdots - C^T \Lambda_n \vec{u}_{k-n+1} \right\} \tag{5}$$

and the equivalent control yields the following equivalent sliding system taking place on the sliding surface

$$\vec{x}_{k+1} = [I - \Lambda_1 (C^T \Lambda_1)^{-1} C^T] \Phi \vec{x}_k + \Lambda_1 (C^T \Lambda_1)^{-1} C^T \vec{r}_{k+1} \tag{6}$$

where it was assumed that $(C^T \Lambda_1)^{-1}$ exists. Equivalently, (6) can be represented in terms of the tracking error by the following equivalent linear system

$$\begin{bmatrix} e_{1,k-m+3} \\ \vdots \\ e_{1,k+1} \\ \vdots \\ e_{l,k-m+3} \\ \vdots \\ e_{l,k+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 & & 0 & & & & \\ & \ddots & \ddots & & & & & \\ 0 & & 0 & 1 & & & \mathbf{0} & \\ -\frac{c_{1,1}}{c_{1,m}} & & \cdots & -\frac{c_{1,m-1}}{c_{1,m}} & & & & \\ & & & & \ddots & & & \\ & & & & 0 & 1 & & 0 \\ & & & & & \ddots & \ddots & \\ & \mathbf{0} & & & 0 & & 0 & 1 \\ & & & & -\frac{c_{l,1}}{c_{l,m}} & & \cdots & -\frac{c_{l,m-1}}{c_{l,m}} \end{bmatrix} \begin{bmatrix} e_{1,k-m+2} \\ \vdots \\ e_{1,k} \\ \vdots \\ e_{l,k-m+2} \\ \vdots \\ e_{l,k} \end{bmatrix} = \overline{\Phi} \vec{e}_k \tag{7}$$

The characteristic polynomial of the equivalent system is $\pi_{\overline{\Phi}}(\lambda) = \det(\lambda I - \overline{\Phi})$. Given the desired characteristic polynomial as $\pi(\lambda) = \lambda^{2m-2} + \upsilon_1 \lambda^{2m-1} + \cdots + \upsilon_{2m-2}$, we can compute C^T by comparing $\pi(\lambda)$ with $\pi_{\overline{\Phi}}(\lambda)$. This condition guarantees an asymptotic convergence to the desired output. For a discrete-time system described by (3), the reaching law for the discrete-time sliding mode control is $S_{k+1} - S_k = -\alpha T S_k - \beta T \operatorname{sgn}(S_k)$ and the control law is derived by comparing $S_{k+1} - S_k = C^T \vec{e}_{k+1} - C^T \vec{e}_k$ with the reaching law, which yields,

$$\vec{u}_k = \left(C^T\Lambda_1\right)^{-1}\left[C^T\vec{r}_{k+1} - C^T\Phi\vec{x}_k - C^T\Lambda_2\vec{u}_{k-1} - \cdots - C^T\Lambda_n\vec{u}_{k-n+1} + (\alpha T - 1)S_k + \beta T\,\mathrm{sgn}(S_k)\right] \quad (8)$$

Salient feature of the multiple sliding mode controller is that one can obtain faster convergence to get the desired response due to multiple control scheme and one can employ variable structure system to control unknown nonlinear plants while gaining indemnity against noise and parameter variations.

4 Comparative Study

Flight vehicles, such as missiles and aircraft, are very complex systems that are typically non-minimum phase and have aerodynamic coefficients which vary over a wide dynamic range due to large Mach-altitude fluctuations [12]. Thus the proposed control algorithms were applied to the LoFLYTE® UAV designed by Accurate Automation Corporation (AAC) to examine the effectiveness [14]. Our objective is to design multiple controllers for unknown nonlinear MIMO plants that guarantees global stability and forces the output to asymptotically track the desired signal without any a priori knowledge of the plant. In this study, we wish to estimate and control the aircraft's lateral motion under the assumption that we can only access roll-rate (p) and yaw-rate (r), while the goal is to track the desired trajectories (p_d and r_d) during the course of the flight considering the case of an aircraft moving with a constant throttle.

Table 1. Comparison of modeling performance[1] for the lateral motion (p and r) of the LoFLYTE® UAV.

Methodology	*NRMSE*	
	roll-rate (p)	yaw-rate (r)
Multiple models (6×6)	4.7e-3	3.8e-3
TDNN (16:50:2)	7.9e-3	3.9e-3

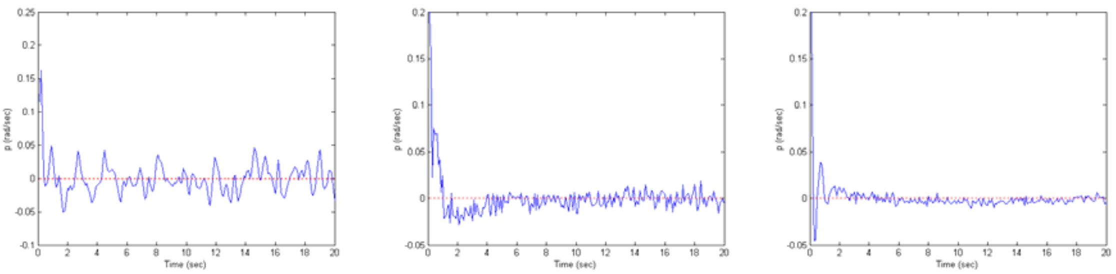
To model the aircraft lateral dynamics, a total of 2 SOMs were used for quantization of each embedded output as predictors, $\psi_{y,k} = \left[y_k, \cdots, y_{k-dy+1}\right]^T$. Thus the linear coupling between p and r is only implicitly modeled[2]. In this way, each output (either p or r) of the aircraft can be described by a dynamic model that takes into account the control input variables such as aileron (δ_a) and rudder (δ_r), $\hat{y}_{k+1} = f_i(\psi_{u,k}, \psi_{y,k}),\ i = 1, \cdots, N$ where $\psi_{u,k} = \left\{\delta_{a,k}, \cdots, \delta_{a,k-du+1}, \delta_{r,k}, \cdots, \delta_{r,k-du+1}\right\}^T$. By doing this, the complexity can be reduced and it helps to understand the raw data. We selected an embedding dimension based on the Lipschitz index (see [11]) as $dy = 2$ for each output and $du = 6$ for 2-D control inputs (aileron and

[1] Modeling performance was evaluated through Normalized Root Mean Square Error (*NRMSE*) $= 1/\max(y)\sqrt{1/L\sum_k^L (y_{k+1} - \tilde{y}_{k+1})^2}$

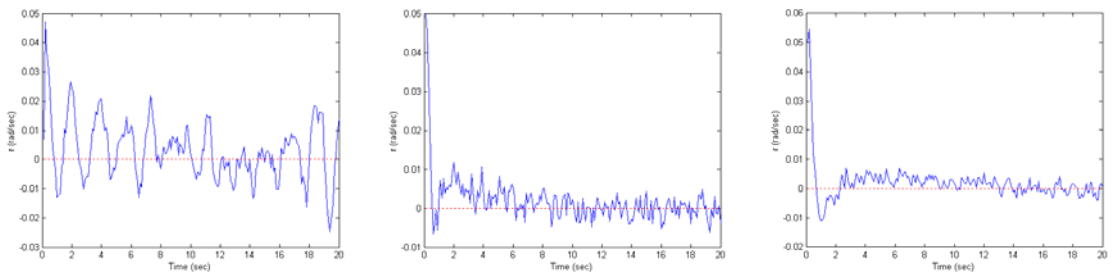
[2] Due to difficulties related with dynamic range normalization, multiple models that take state-coupling into account are not as accurate as this approach. Instead, we utilize the delayed outputs in order to compensate for the disregarded information due to the coupling.

rudder). Each SOM was trained with the embedded output, ψ_y whose dimension is 2, over 5000 samples and the reasonable result for identification of the roll-rate was obtained with a 6×6 grid map ($N = 36$).

After training the SOMs, 36 multiple models were constructed and the created models were tested by new sequence with 1,000 samples. Table 1 shows the identification performance of two dynamics of the system with 36 models and compares their performances with that from a TDNN model. Training conditions for a TDNN model, such as the embedding dimension was kept the same in this comparison between the local modeling and the global modeling. The best result with a TDNN model was obtained from 20 Monte-Carlo simulations with 50 PEs in the hidden layer. From the table, we can conclude that the constructed SOM-based network is a good model of the underlying dynamics because it provides smaller *NRMSE* for all dynamics than the TDNN model. Consequently, it turned out that the proposed strategy of finding proper location of fixed models depending on the prior information available to the designer for finding aircraft dynamics is superior to those using a single global nonlinear model. In addition, it should be noted that the proposed modeling scheme makes identification of the plant very compact and computationally efficient since the aircraft dynamics are captured in a compact lookup table of linear models.



(a)

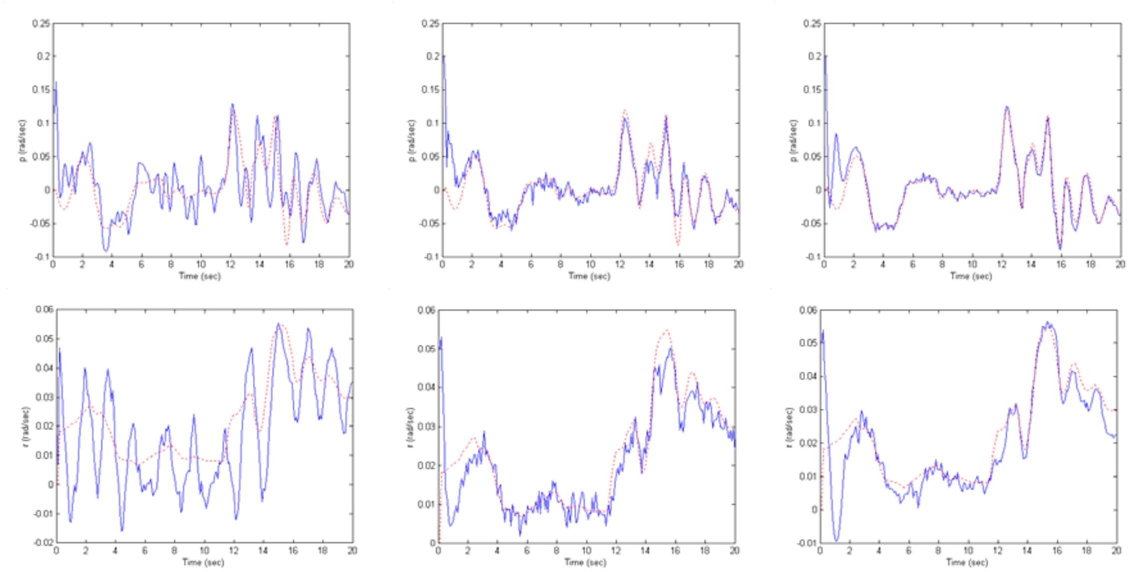


(b)

Figure 1. Comparison for controlling (a) roll-rate and (b) yaw-rate to track the set point in the presence of noise by TDNNC, MIC, and MSMC.

We now consider the control problem with the SOM-based local models created. Here, we performed a simulation to control the roll-rate (p) and yaw-rate (r) of the aircraft by aileron (δ_a) and rudder (δ_r), setting elevator to zero and throttle to constant. Thus, once we have the linear models for the roll-rate and the yaw-rate, and the desired values, $p_{d,k+1}$ and $r_{d,k+1}$, the inverse controller (inversion-based predictive model), $\delta_{a,k}$ and $\delta_{r,k}$, for the aircraft's roll-rate and yaw-rate tracking is obtained. And the sliding mode controller was designed such that the poles of the error dynamics are placed at 0.5 and other parameters are set to $\alpha T = 0.5$ and

$\beta T = 0.001$. Also, for performance comparisons, we applied the TDNN controller, which has 100 PEs in the hidden layer, for the same control problems.



(a) (b) (c)

Figure 2. Comparison for controlling roll-rate and yaw-rate to track an arbitrary trajectory with measurement noise by (a) TDNNC, (b) MIC, and (c) MSMC.

Figure 1 compares the set point tracking performance of the TDNNC, the MIC, and the MSMC in the presence of noise whose level is 20 dB of SNR. From the responses it can be seen that the multiple controller approach is very good except for the first few seconds. However, it shows poor transient response when the global control, the TDNNC, is utilized. Another performance test is to enforce the tracking of the roll-rate and yaw-rate to signals $p_{d,k+1}$ and $r_{d,k+1}$ which are given in real time during the course of the flight, while being subjected to unmeasured sensor disturbances. The 2 output measurements are corrupted by zero-mean random sequences with 20 dB of SNR. The results of a flight test with the proposed method are shown in Figure 2 where we also show the same with the TDNNC. It can be seen that the roll-rate and the yaw-rate track their command signals quite well even under the existence of measurement noise by the multiple controllers. The simulated flight test demonstrates that the proposed controller is capable of closely approximating the given mission by only looking at the past information. Also, it proves that the multiple controller framework indeed provides exceptional tracking.

5 Conclusions

The problem of nonlinear system identification and control system design was addressed under the divide-and-conquer principle. Especially in the case of unknown dynamics, where only input-output data from the plant are available, the proposed method is able to approximate the nonlinear dynamics of the plant using a piece-wise linear dynamical model that is optimized solely from the available data. An added advantage of the proposed local linear modeling approach is it greatly simplifies the design of control systems for nonlinear plants. In general, this is a daunting task and typically practical solutions involve linearization of the dynamics and then employing well-established controller design techniques from linear control systems theory. While designing globally stable nonlinear controllers with satisfactory performance at every point in the state space of the closed loop control system is extremely difficult, and perhaps impossible to achieve especially in the case of unknown plant dynamics, by using the local linear modeling technique presented, coupled with strong controller design techniques from linear control theory and recent theoretical results on switching control systems, it becomes possible to achieve this goal through the use of this much simpler approach of local modeling.

The problems that arise due to the uncertainties of the plant model and measurement noise are alleviated by incorporating the robustness provided by the sliding mode technique into the multiple modeling approach. The simulation results demonstrated that the algorithm proposed is able to compensate deficiencies caused by the imperfect observations of the state variables and complex plant dynamics, driving the tracking error vector to the sliding manifold and keeping it on the manifold. In addition, the proposed method shows better robustness against noise, faster transient response, and better steady-state accuracy of the controlled system by switching local controllers astutely through the SOM than other neural network-based alternatives.

Acknowledgement : This work was partially supported by NASA grant NAG-1-02068.

References

[1] O. Nelles (2001) *Nonlinear System Identification*. New York, Springer.
[2] R. Murray-Smith and T.A. Johansen (1997) *Multiple Model Approaches to Modeling and Control*. New York, Taylor & Francis Inc..
[3] J. Walter, H. Ritter and K. Schulten (1990) Nonlinear Prediction with Self-organizing Maps. *Proc. IJCNN*, vol.3, pp.589-594.
[4] J.C. Principe, L. Wang and M.A. Motter (1998) Local Dynamic Modeling with Self-Organizing Maps and Applications to Nonlinear System Identification and Control. *Proc. of IEEE*, 86(11), pp.2240-2258.
[5] T. Kohonen (1995) *Self-Organizing Maps*. New York, Springer.
[6] A. Packard (1994) Gain scheduling via linear fractional transformations. *Systems and Control Letters*, 22(1), pp.79-92.
[7] K.S. Narendra, J. Balakrishnan, and M.K. Ciliz (1995) Adaptation and learning using multiple models, switching, and tuning. *Control Systems Magazine*, 15(3), pp.37-51.
[8] K. Furuta (1993) VSS Type self-tuning control. *IEEE Trans. Ind. Electro.*, 40(1), pp.37-43.
[9] W. Gao, Y. Wang, and A. Homaifa (1995) Discrete-time variable structure control systems. *IEEE Trans. Ind. Electro.*, 42(2), pp.117-122.
[10] A. Bartoszewicz (1998) Discrete-time quasi-sliding-mode control strategies. *IEEE Trans. Ind. Electro.*, 45(4), pp.633-637.
[11] He and H. Asada (1993) A New Method for Identifying Orders of Input-Output Models for Nonlinear Dynamic Systems. *Proc. ACC*, pp.2520-2523.
[12] J. Cho, J.C. Principe, D. Erdogmus, and M.A. Motter (2005) Modeling and inverse controller design for an Unmanned Aerial Vehicle based on the self-organizing map. to appear in *IEEE Trans. Neural Networks*.
[13] K.S. Narendra and K. Parthasarathy (1990) Identification and Control of Dynamical Systems using Neural Networks. *IEEE Trans. Neural Networks*, 1(1), pp.4-27.
[14] http://www.accurate-automation.com/Technology/Loflyte/loflyte.html