LOCAL LINEAR PID CONTROLLERS FOR NONLINEAR CONTROL

Jing Lan[1], Jeongho Cho[1], Deniz Erdogmus[2], José C. Principe[1], Mark A. Motter[3], Jianwu Xu[1]

[1]CNEL, ECE Department, University of Florida, Gainesville, FL 32611

[2]CSE & BME Department, Oregon Graduate Institute, OHSU, Portland, OR 97006

[3]Dynamics and Control Branch, NASA Langley Research Center, Hampton, VA 23681

{jinglan, jeongho, principe, jianwu}@cnel.ufl.edu, derdogmus@ieee.org, m.a.motter@larc.nasa.gov

Abstract

Nonlinear PID design is difficult if one approaches the problem from a global design perspective. In this paper, we propose coalescing the adaptive local linear modeling approach with traditional linear PID controller design techniques to arrive at a principled, intuitive, and simple nonlinear PID controller design methodology. The paper, in particular, discusses two local linear modeling approaches, one based on competitive linear models attached to a self-organizing map, and one based on cooperative linear models attached to the modes of a Gaussian mixture model. The system identification and closed-loop tracking performances of the proposed PID controller design methodology is demonstrated to be better than inverse and PID controllers for a global nonlinear system identifier and a global linear model respectively.

Index Terms

PID controller, local linear model, self-organizing maps, Gaussian mixture.

I. INTRODUCTION

Identification and control of nonlinear, time-varying dynamical systems is of great practical importance, since virtually all real-world systems belong to this class. Typically, the control system design procedure follows the system identification and controller design steps. While the system identification step is traditionally achieved by data-driven coefficient estimation for physically conceived parametric models [1], [2], contemporary approaches that are driven by recent developments in adaptive system and neural network theories undertake the unconventional black-box modeling route [3], [4], [5], [6], [7]. Sometimes, both techniques are united in the so-called gray-box modeling approach, where parts of the model might be parametrically described based on the understanding of system physics and a neural network approximates the remaining portion. Once the system identification stage is complete, the controller design stage starts, usually with certain linearization assumptions and gain scheduling [8]. Modern adaptive control techniques could also be utilized in conjunction with linear or nonlinear system models to close the control system loop [3], [7], [9].

In this paper we will focus on the class of nonlinear, time invariant systems. In some cases, the physics underlying the dynamics of the system is well understood (as in the aircraft designs), however, there are many systems where this traditional modeling-from-first-principles approach is simply unthinkable. On the other hand, the success of a global nonlinear black-box modeling approach using neural networks or fuzzy systems is hindered by the scalability problems in network approximation capabilities and increased training difficulty levels. Typically, large nonlinear networks will be difficult to train and will most likely have generalization problems. Even after much difficulty, if a physics-based mathematical or a neural network model is obtained, designing global nonlinear controllers becomes difficult. Traditional methods based on linear models will be, in general, difficult to extend. Nonlinear and adaptive controllers, on the other hand, have major drawbacks in terms of meeting performance expectations.

In linear control theory, PID controllers are extensively studied by researchers and well understood by practitioners [9]. Extensions of the PID control methodology to nonlinear systems, however, are not trivial. Usually, global nonlinear models are, as always, linearized and the parameters of the PID controller are scheduled according to the regime. Alternatively, neuro-fuzzy PID controller design techniques have been investigated [10]. All these difficulties, in fact stem from the fact that the system model is constructed from a nonlinear dynamical equation.

This difficulty could be eliminated by a piece-wise linear approximation of the nonlinear dynamics (as one does in gain scheduling). However, gain scheduling methods are not flexible for model uncertainties and they cannot be scaled up to regimes that are not described in the initial system identification stage by a linear model. It also has the problem of either inefficient use of system identification data due to a large number of arbitrarily selected operating (linearization) points or inaccurate modeling due to the small number of linearization points. The first problem could be alleviated by using an adaptive system identification mechanism to determine the linearization points. The second problem can be addressed by a flexible piece-wise linearization scheme that allows the addition of new linear patches to the system model whenever data not accurately explained by the existing model is encountered. In this paper, we will not address this second problem. But, in earlier work, we demonstrated how this could be addressed [11].

The nonlinear PID control design method described here is rooted in the principle of using local linear models (LLM) to construct a globally nonlinear (piece-wise linear) system model that is determined completely from the input-output data collected from the actual plant [12]. The local models could be associated with the PEs (processing elements) of a self-organizing map (SOM) [13] or the modes of a Gaussian mixture model (GMM) [14]. Whatever LLM modality is utilized, the design of a piece-wise linear PID controller for the overall nonlinear system becomes trivial because of the simplicity of LLM. Using any standard linear PID design method one determines the desired PID coefficients for every linear model. Depending on the modality (SOM or GMM), in practice, these local linear PID controllers are either used competitively or cooperatively. Since SOM or GMM can smoothly represents the dynamics using multiple models, the proposed controller follows the same switching rule here. These two modalities will be discussed in the following sections, as well as how to utilize the designed multiple PID controllers to produce a single overall control input to the actual plant.

A problem in system identification is the selection of state variables. Especially in the case of black-box modeling
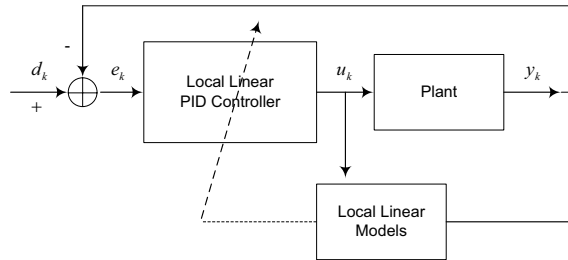
Fig. 1. Overall schematic diagram of the nonlinear PID closed-loop control mechanism using local linear models and controllers.

where no or negligible physical understanding of the system exists, we can resort to the time-delay embedding method to reconstruct a replica of the state space dynamics of the nonlinear system from its input output data [15], [16]. It is well known from nonlinear dynamical system theory that, in the small noise regime, using a suitable length of delay-line one can reconstruct a state vector whose state transition function is isomorphic to that of the original system [17]. Even if one knows what state variables to choose, the delay embedding method can be employed. The LLM approaches discussed in the following sections will make use of this principle.

In summary, the goal of this paper is to illustrate how the local linear modeling technique can be incorporated with the well-known linear PID controller design techniques to obtain a principled and simple nonlinear (and possibly adaptive) PID controller design methodology. The proposed closed loop scheme is illustrated in fig. 1 where the active local linear models also determine which linear PID controllers contribute to the instantaneous control input.

II. ADAPTIVE LOCAL LINEAR MODELING

In the 60s and 70s Magill [18] and Lainiotis [19] studied Kalman filter-based models to improve the accuracy of state estimation. In the context of adaptive control, switching was first proposed by Martensson [20]. The concept of using multiple models with switching according to the operating regime has been explored by several researchers previously [12], [21], [22], [23]. In [24], [25] the use of multiple fixed models for set point control was studied. [24] also proposed an adaptive control framework where multiple fixed models and one or two adaptive model(s) were used to construct the system model and a certain local model's active region is pre-defined.

Our laboratory proposed the use of self-organizing maps (SOM) to smooth out the switching between the models. The SOM represents the topological relation in the input space, and facilitates the design of LLMs from local neighborhoods with overlaps to guarantee smoothness. An identification-controller pair designed based these principles outperformed an expert human operator and the existing classical controller in the NASA Langley Wind Tunnel [26]. The modeling of the NASA LoFlyte waverider aircraft was also addressed using LLM by the authors in previous work [27], [28]. The technique is flexible, and provides a good compromise between modeling accuracy and computational simplicity. In this section, we will particularly discuss two LLM approaches: SOM-based competitive

LLM and GMM-based cooperative LLM. In the rest of the paper, we will assume that no a priori knowledge about the system structure is available and that mathematical modeling and controller design needs to be carried out utilizing only the available input-output data collected from the plant. For simplicity, we will provide simulations for a single-input single-output (SISO) system. However, the extension of these principles to multi-input multi-output (MIMO) systems is possible.

A. SOM-Based Competitive LLM

Suppose that input-output data pairs of the form $\{(u_1, y_1), \ldots, (u_N, y_N)\}$, where u is the input signal and y is the output signal, is available from a SISO system for system identification. It is assumed that the system that generated this data sequence is of the following general nonlinear time-invariant form (assuming that there is no instantaneous effect of the input on the output):[1]

$$
\begin{aligned}
\tilde{\mathbf{x}}_{k+1} &= f(\tilde{\mathbf{x}}_k, u_k) \\
y_k &= h(\tilde{\mathbf{x}}_k)
\end{aligned}
\tag{1}
$$

In state space,[2] this nonlinear dynamical equation can be approximated by an ordinary linear differential equation of the NARMAX-similar form [29]:

$$
y_k = a_1 y_{k-1} + \cdots + a_n y_{k-n} + b_1 u_{k-1} + \cdots + b_m u_{k-m}
\tag{2}
$$

The coefficient vectors **a** and **b** depend on the local regime; consequently system identification using multiple local linear models of this form divide the state space in a set of local areas where the linear approximation is valid. The SOM-based local linear modeling is defined in the schematic diagram of the overall system identification topology in fig. 2. The clustering can be adaptively achieved by training a SOM on the combined state vector **x** defined as

$$
\mathbf{x}_k = [y_{k-1}, \ldots, y_{k-n}, u_{k-1}, \ldots, u_{k-m}]^T
\tag{3}
$$

The SOM consists of a multidimensional array of PEs with weight vectors $\mathbf{w}_i$, which are trained competitively on the input vector samples $\mathbf{x}_k$. While the samples are presented to the SOM one at a time, in a series of epochs, the weight vector of the winner PE $\mathbf{w}_w$ for the particular sample (i.e. the weight vector that minimizes the Mahalanobis distance $d(\mathbf{w}, \mathbf{x}_k) = (\mathbf{w} - \mathbf{x}_k)^T \Sigma^{-1} (\mathbf{w} - \mathbf{x}_k)$ between the two vectors) and its topological neighbors $\mathbf{w}_n$ are updated according to the following rules [13].

$$
\begin{aligned}
\mathbf{w}_w(t+1) &\leftarrow \mathbf{w}_w(t) + \eta(t) \Sigma^{-1} (\mathbf{x}_k - \mathbf{w}_w(t)) \\
\mathbf{w}_n(t+1) &\leftarrow \mathbf{w}_w(t) + \eta(t) h(\mathbf{w}_n(t), \mathbf{w}_w(t), \sigma(t)) \\
&\quad \cdot \Sigma^{-1} (\mathbf{x}_k - \mathbf{w}_w(t))
\end{aligned}
\tag{4}
$$

[1]This assumption can be relaxed easily.

[2]Here, we use the phrase state space in a broader sense than traditional; it refers to the vector space formed by the delayed values of the input and output signals as they appear in (2).
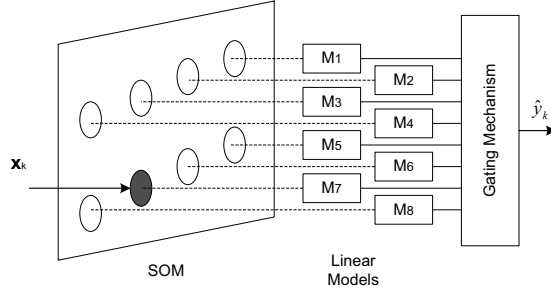
Fig. 2. Schematic diagram of the SOM-based LLM approach for system identification. The winner PE is illustrated by the full circle.

In (4), the (Gaussian) neighborhood function $h(., \sigma)$ is annealed in time such that initially the neighborhood radius σ covers most of the network and towards the end of training it is narrow enough to include only the winner PE. Typically linear or exponential annealing schemes are used for both the neighborhood radius and the step size η.

Once the SOM is trained, the available data is partitioned into smaller sets; the samples associated with weight vector $\mathbf{w}_i$ are $\{(\mathbf{x}_{i1}, y_{i1}), \ldots, (\mathbf{x}_{iN_i}, y_{iN_i})\}$, where N_i is the number of samples clustered to PE i. In order to approximate the local dynamics, each PE has a linear models attached to each it, whose coefficient vectors $\mathbf{a}_i$ and $\mathbf{b}_i$ can be optimized using least squares with this input-output training data that is clustered to this PE (and its neighbors) in the SOM. The overall LLM output can be produced either using hard competition or soft competition between the linear models associated with each PE. In particular, the model output is determined as follows:

$$\hat{y}_k = \sum_{i=1}^{M} p_{ik} [\mathbf{a}_i^T, \mathbf{b}_i^T] \mathbf{x}_k \tag{5}$$

For hard competition, the scaling factor p_{ik} is either 1 or 0 depending on the ith PE being the winner or not for input $\mathbf{x}_k$. For soft competition, which also allows a smoother transition between the linear models at switching boundaries, the model output can be determined by a weighted average. In this case, one can use, for example, a weighting based on the Mahalanobis distance again:

$$p_{ik} = \frac{d(\mathbf{w}_i, \mathbf{x}_k)}{\sum_{j=1}^{M} d(\mathbf{w}_j, \mathbf{x}_k)} \tag{6}$$

Alternatively, similar weighting schemes where only a few of the nearest models are activated could be employed. From a probabilistic point-of-view, such a weighting mechanism can be regarded as the probability of the corresponding linear model being responsible for the sample under consideration. The GMM approach described in the following section makes explicit use of this probabilistic view for the data.

Another modification on SOM-based modeling is to choose the winner based on the minimum Mahalanobis distance of $[y_{k-1}, \ldots, y_{k-n}]$ part of $\mathbf{x}_k$ in (3) instead of the whole $\mathbf{x}_k$. we show that the system identification results from this new criterion has smaller modeling error.
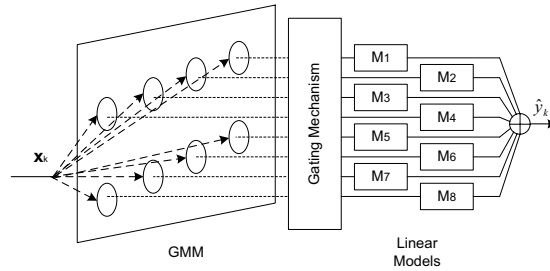
Fig. 3.   Schematic diagram of the GMM-based LLM approach for system identification.

B. GMM-Based Cooperative LLM

Suppose that input-output data pairs of the form $\{(u_1, y_1), \ldots, (u_N, y_N)\}$ are available from a SISO system as described in (1). The linear approximation in (2) still holds locally. The representative state vector is again described as in (3). In the GMM approach of fig. 3, instead of partitioning this vector into disjoint sets using a SOM, it is assumed that the data is generated by a multi-dimensional joint mixture of Gaussian distribution.[3] In other words, the distribution of vector $\mathbf{x}$ is given by

$$p(\mathbf{x}_k) = \sum_{i=1}^{M} \alpha_i G(\mathbf{x}_k; \mu_i, \mathbf{\Sigma}_i) \tag{7}$$

where $G(\mathbf{x}_k; \mu, \mathbf{\Sigma})$ is a multivariate Gaussian density with mean μ and covariance $\mathbf{\Sigma}$. The coefficient α_i denotes the probability of occurrence of this mode in the GMM. Given the training data, once the vector samples $\mathbf{x}_k$ are constructed, the maximum likelihood solution for the parameters α_i, μ_i, and $\mathbf{\Sigma}_i$ can be determined using the expectation maximization (EM) algorithm [14]. Alternative statistical model selection approaches could also be employed. As in the SOM-based model, here it is assumed that a linear model is associated with each of the Gaussian model in the GMM. Each linear model has weight vectors $\mathbf{a}_i$ and $\mathbf{b}_i$ that can be optimized using least squares once the parameters of the GMM are determined using the EM algorithm. The overall model output is a weighted average of individual linear models

$$\hat{y}_k = \sum_{i=1}^{M} p_{ik}[\mathbf{a}_i^T, \mathbf{b}_i^T]\mathbf{x}_k = \theta^T \mathbf{z}_k \tag{8}$$

where $p_{ik} = \alpha_i G(\mathbf{x}_k; \mu_i, \mathbf{\Sigma}_i)$. The overall linear model weight vector is defined as $\theta = [\ \mathbf{a}_1^T \quad \mathbf{b}_1^T \quad \ldots \quad \mathbf{a}_M^T \quad \mathbf{b}_M^T \ ]$, and the corresponding overall modified input vector is $\mathbf{z}_k = [\ p_{1k}\mathbf{x}_k^T \quad \ldots \quad p_{Mk}\mathbf{x}_k^T \ ]$. The linear models can be optimized simultaneously using least squares considering the available output measurements y_k as desired outputs corresponding to the modified input sample $\mathbf{z}_k$.

---

[3]Asymptotically (in the number of training samples and the number of modes), both piece-wise uniform approximation by the SOM and the GMM can perfectly reconstruct the input distribution.
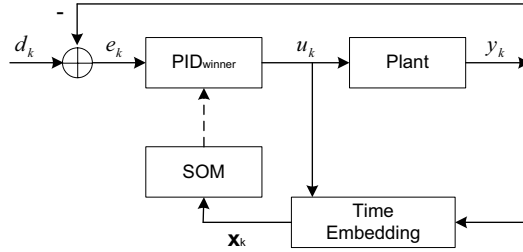
Fig. 4.   Schematic diagram of the switching linear PID controllers based on the SOM decision for the winner model.
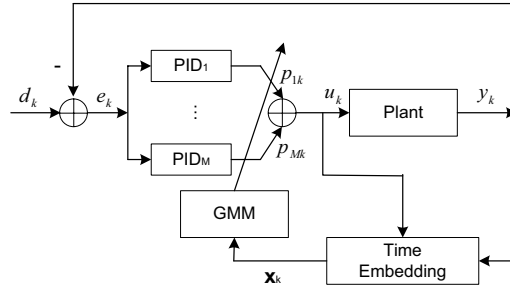

Fig. 5.   Schematic diagram of the weighted average PID controller based on the GMM assessment of contributions.

III. LOCAL LINEAR PID CONTROLLER DESIGN

Once system identification is complete, the design of a globally piece-wise linear PID controller system can easily be accomplished using standard techniques. The literature has an abundance of PID design methodologies for linear SISO systems including direct pole-placement techniques and optimal coefficient adjustment according to some criteria [30], [33]. In the case of MIMO systems, where the system identification phase is carried out as described in the previous sections, one can resort to MIMO PID design techniques [34].

Given the PID design and local linear models, the overall closed loop nonlinear PID design reduces to determining the coefficients of the individual local linear PID controllers using their respective linear plant model transfer functions. In the competitive SOM approach (fig. 4), the model output depends only on a single linear model at a given time; therefore, the PID coefficients are set to those values determined for the instantaneous winner model. In the GMM approach (fig. 5), since the models are averaged with probabilistic weights, the same weighting must be done with the PID controllers. An alternative look at (8) shows that the GMM-based LLM output is also equivalently expressed as

$$\hat{y}_k = \mathbf{x}_k^T \left(\sum_{i=1}^{M} p_{ik} \begin{bmatrix} \mathbf{a}_i \\ \mathbf{b}_i \end{bmatrix} \right) = \mathbf{x}_k^T \begin{bmatrix} \bar{\mathbf{a}}_{ik} \\ \bar{\mathbf{b}}_{ik} \end{bmatrix} \tag{9}$$

where $\bar{\mathbf{a}}_{ik}$ and $\bar{\mathbf{b}}_{ik}$ denote the weighted average linear model at time instant k. Consecutively, The instantaneous

PID controller will be characterized by a weighted average of the PID coefficients of all linear models as follows:

$$PID_k = \sum_{i=1}^{M} p_{ik} PID_i \tag{10}$$

where PID_i is the transfer function of the PID controller for the ith local model. In the case of a soft combination SOM-based LLM as described by (6), the formulation in (10) can be employed for the PID controller.

The PID controller design is carried out using the standard pole-placement technique [30], where the poles of the tracking error obey the equation

$$e_{k+1} + \lambda_1 e_k + \lambda_2 e_{k-1} = 0 \tag{11}$$

Due to the modeling error (and also possibly measurement noise, etc.), the equation (11) should be changed to

$$\tilde{e}_{k+1} + \lambda_1 \tilde{e}_k + \lambda_2 \tilde{e}_{k-1} = 0 \tag{12}$$

where $\tilde{e}_k = e_k + e_{M,k}$ is the combination of tracking error e_k and modeling error $e_{M,k}$ from LLM. Under the assumptions that $[1, \lambda_1, \lambda_2]^T$ represents a dynamical function with two different poles p_1 and p_2 within the unit circle and the modeling error is a wide sense stationary random variable with zero mean and finite power, the asymptotic tracking error is given by

$$e_k = \frac{\Delta}{(1-p_1)(1-p_2)} - \frac{\Delta}{(1-p_1)(p_1-p_2)} p_1^{k+1}$$
$$- \frac{\Delta}{(1-p_2)(p_2-p_1)} p_2^{k+1} \tag{13}$$

where Δ is a certain linear combination of modeling error. As it can be expected, the modeling error cannot be compensated completely, but (13) shows that the tracking error power is proportional to the modeling error power. Note that the pole's position control the gain and dynamic of the error. A proof of stability is in the Appendix.

IV. SIMULATIONS

The performance of the proposed local linear PID design approach will be demonstrated on a simplified SISO nonlinear missile dynamic model where only the yaw dynamics are considered [5]. The simplified two-state dynamics are described by:

$$\dot{x}_1 = x_2 - 0.1cos(x_1)(5x_1 - 4x_1^3 + x_1^5)$$
$$- 0.5cos(x_1)u$$
$$\dot{x}_2 = -65x_1 + 50x_1^3 - 15x_1^5 - x_2 - 100u \tag{14}$$
$$y = x_1$$

Since the input is the rudder deflection, it is limited by 0.5 radians in practice. The simulations obey this limit on the control input value although the PID design will not take this into consideration.

Two models were identified using 6000 samples of input-output pairs obtained from the dynamical equation (14) using a discretization time step of T_s=0.05s (which corresponds to 300s of flight time). In order to excite a rich
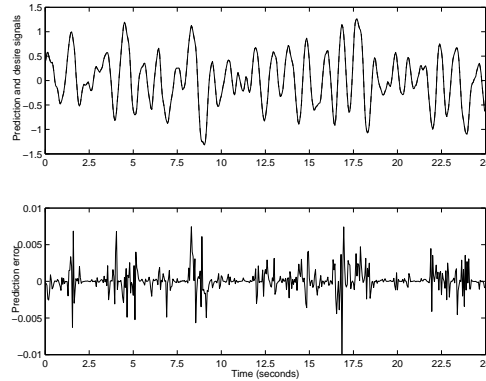
Fig. 6. System identification performance is shown at the top for the SOM-based LLM approach. The bottom plot shows the system identification error
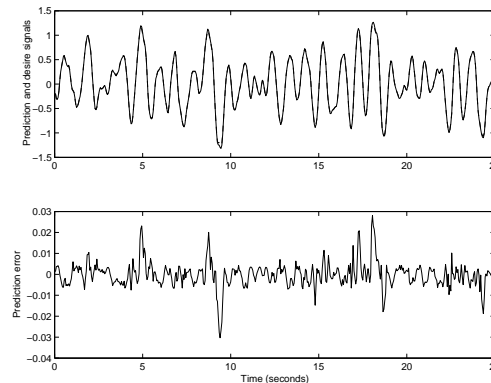

Fig. 7. System identification performance is shown at the top for the GMM-based LLM approach. The bottom plot shows the system identification error

variety of dynamical modes in the plant, the system identification data was generated using a uniformly random input signal within the specified limits. The embedding delays for the input and the output were both selected to be 2, resulting in 4-coefficient linear models. This embedding delay was obtained experimentally with Lipschitz quotients [32]. Increasing the embedding delay did not result in system identification performance improvement. Through experiments, the SOM-based LLM approach utilized a 5×5 rectangular grid of PEs, amounting to 25 competitive local linear models, whereas the GMM-based LLM approach assumed to five-mode mixture model, resulting in five cooperative linear models. The identified models were tested on 50s-length data (1000 samples) with new sequence of random input, generated using a new sequence of random input. The actual plant output, the predicted model output and the estimation error for the SOM-based LLM and GMM-based LLM models are provided in fig. 6 and fig. 7. The signal-to-error ratio (SER) for these two models on the testing set are found to be 37.4dB and 31.7dB, respectively. Notice that the SOM system identification is a little better providing smaller errors. However the GMM has much fewer local models.
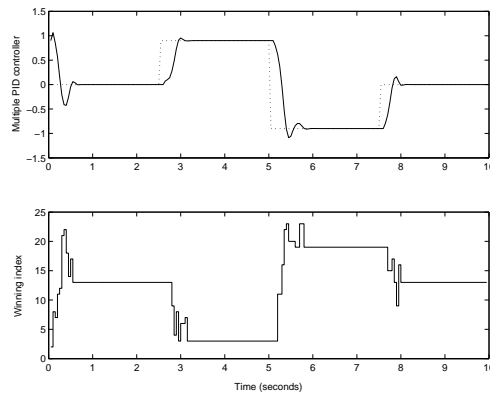
Fig. 8. Response of the closed loop nonlinear PID control system to various step changes in the desired response is shown at the top for the SOM-based LLM approach. The bottom plot shows the index of the winner linear model changing in time as determined by the SOM based on the instantaneous input.
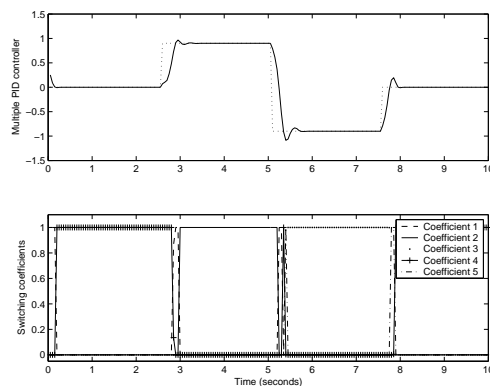

Fig. 9. Response of the closed loop nonlinear PID control system to various step changes in the desired response is shown at the top for the GMM-based LLM approach. The bottom plot shows the contribution coefficients of the linear models changing in time as determined by the Gaussian mixture model based on the instantaneous input.

For both modeling methodologies, the corresponding PID coefficients are determined to bring the closed-loop response poles from the plant output to the desired output to $(0, 0, 0.05 + i0.3, 0.05 - i0.3)$. Those poles make the system stable in the case of the nonlinear uncertainty exists. This closed-loop nonlinear PID controller is tested in two cases: one with various step changes to the desired output in fig. 8 and fig. 9 and one with a smoothly changing desired output, which is the sum of three sinusoids with different amplitudes, phases and frequencies in fig. 10 and fig. 11. As can be observed, the system tracks very well (small overshoot and fast settling times) the desired trajectory. During transients the models switch smoothly.

Set-point tracking by the proposed multiple PID (MPID) control scheme (both SOM and GMM-based MPID) showed better performance when compare with other approaches. Two types of controllers were utilized in the comparison with different system identification methods.
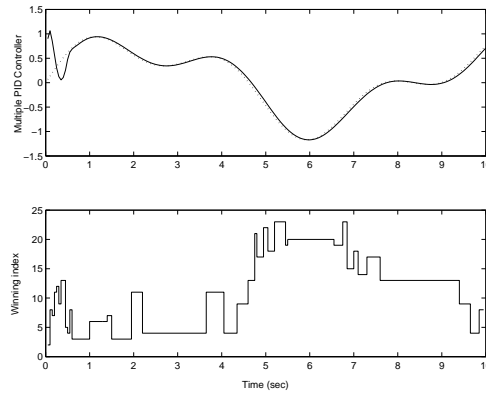
Fig. 10. Response of the closed loop nonlinear PID control system to smoothly varying desired response is shown at the top for the SOM-based LLM approach. The bottom plot shows the index of the winner linear model changing in time as determined by the SOM based on the instantaneous input.
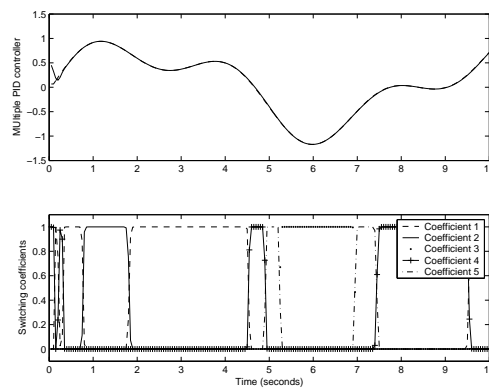

Fig. 11. Response of the closed loop nonlinear PID control system to smoothly varying desired response is shown at the top for the GMM-based LLM approach. The bottom plot shows the contribution coefficients of the linear models changing in time as determined by the Gaussian mixture model based on the instantaneous input.

Set point tracking (0.5 to -0.5) with an inverse controller is compared with multiple system identification methods based on the SOM and GMM (MIC-SOM and MIC-GMM) and a TDNN to demonstrate the transient response. As can be observed from fig. 12 and Table I the falloff time is very similar among the three models, but the TDNN controller generates an offset error because of the different characteristics of the modeling (global versus local). This is to be expected since local modeling generates more accurate models. When the inverse controller was changed to a PID controller we noticed that the fall time of the transient response was increased (set point 0.2 to -1.5, fig. 13 and Table II). We also included here a global MA (moving average) model trained with the Wiener solution. The MPID-SOM, MPID-GMM and the PID-MA model have comparable fall times among themselves, but the PID-MA model provides a longer settling time. When the proposed MPIDs are compared with the inverse controller and TDNN, we can observe that the proposed methods are slower in the fall time, but the overshoot and
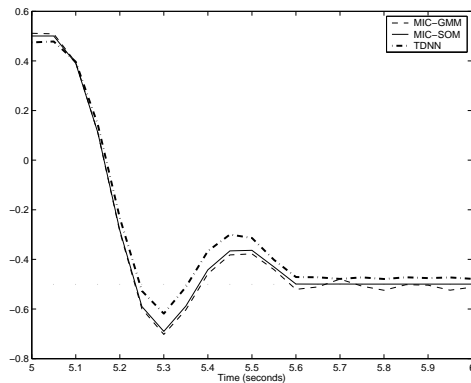
Fig. 12. Comparison of transient response of inverse control systems from SOM-based LLM approach and GMM-based LLM approach to the closed loop nonlinear adaptive TDNN control system.

TABLE I

PERFORMANCE COMPARISON OF THE MULTIPLE MODEL-BASED INVERSE CONTROLLERS WITH TDNN

	Falling time(s)	Settling time(s)	Overshot(%)
MIC-GMM	0.2340	0.5683	20.24
MIC-SOM	0.2353	0.5789	19.01
TDNN	0.2454	0.5992	11.85

settling time are much better. The TDNN controller generates larger transient error and made large tracking errors. The fall time difference arises from the fact that the inverse controller is only sensitive to $1st$ order error dynamics (one sample memory), while the PID controller is designed based on $3rd$ order error (three sample memory). We conclude that the MPID controller offer a different compromise in tracking performance that may by useful in practical applications.

Further experiments with the multiple PID design method proposed here using various other pole assignments and desired response profiles not shown in this paper have revealed similar performance. We also evaluated the performance of alternative PID controller design methodologies [33] which boil down to selecting the desired pole locations based on an optimality principle. Since such criteria are highly case dependent, they are not considered in this paper.

V. CONCLUSIONS

The design of PID controllers for nonlinear systems based on global modeling approaches is in general a difficult task because of global stability as well as performance degradation due to tuning mismatches between the current PID coefficients and the local dynamics of the nonlinear system. In this paper, we proposed a completely adaptive strategy to design nonlinear PID controllers on a piece-wise linear basis using only input-output data from the unknown system. if desired, however, *a priori* information about the system could be incorporated into the design scheme (not studied in this paper). The competitive local linear modeling scheme as well as the mixture of linear
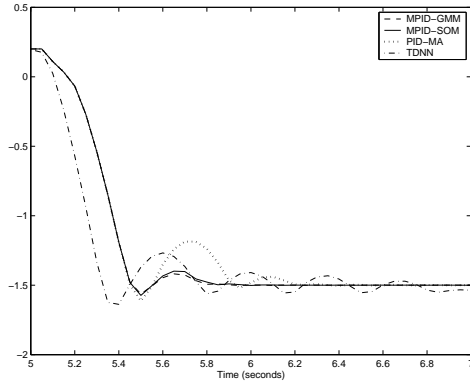
Fig. 13. Comparison of transient response of PID control systems from SOM-based LLM approach and GMM-based LLM approach to the PID control from global linear model and the closed loop nonlinear adaptive TDNN control system.

TABLE II

PERFORMANCE COMPARISON OF THE PROPOSED CONTROLLERS WITH A SINGLE PID CONTROLLER AND TDNN

	Falling time(s)	Settling time(s)	Overshot(%)
MPID-GMM	0.4640	0.7286	3.73
MPID-SOM	0.4609	0.7454	4.34
PID-MA	0.4522	1.1204	6.66
TDNN	0.3286	1.3873	8.14

model experts allow the designer to break down the overall nonlinear controller design task into easier linear controller design tasks (as commonly done in the control literature, such as in gain scheduling). From the results we can conclude that wild switching, one of the fundamental problems of the multiple model controllers, is alleviated in our design. The reason is associated with the smooth variation of the SOM-LLM and GMM-LLM. In fact, in the SOM-LLM neighboring models can share coefficients due to the topological neighborhood preservation of the SOM. In the GMM-LLM the cooperation among the mixture models ensure also smooth transactions. Therefore the PID controllers, which are derived from these system identification blocks, share the same property. One may think that designing LLM in the 2-D shape of the SOM or the low dimensional shape of GMM may not capture the rich dynamics of the input that may be of high dimensions. The results show that locally access to the full multidimensional input is not necessary when the goal is to design local controllers.

In addition, the global convergence and stability properties of the closed-loop control system is secured by the universal approximation capabilities of the described local linear models [21], [35] and the local stability of the PID controllers within the operation region that they are designed for. The SOM and GMM based local modeling do achieve good system identification performance, and the closed-loop tracking performance of the LLM based nonlinear PID controller designed according to this prescription outperformed a traditional adaptive globally nonlinear controller (based on a TDNN) as well as a global MA model.

A possible drawback is generalization, since the controller is completely designed based on experimental input-output data. therefore, its performance is guaranteed for state-space regions from which training data for system identification was available. Clearly, as in any adaptive solution, the performance will degrade if the system identification data is not sufficiently rich to get a complete model representation of the system. This problem, however, is not worse than the similar problem in gain scheduling where the performance of the local linear controllers cannot be guaranteed outside their effective linearization regions.

Future work will investigate the optimal selection of PID coefficients for robust closed-loop performance under system identification and operation under noisy data. In addition, extensions to MIMO nonlinear PID design will be studied in more detail to be presented in future publications.

APPENDIX: PROOF OF CONTROL STABILITY

From (12), λ_is are coefficients to be designed such that two poles p_is are within the unit circle where $p_1 + p_2 = -\lambda_1$ and $p_1 \times p_2 = \lambda_2$. Since the LLMs are trained based on MSE criterion, the modeling errors at each step can be deemed bounded i.i.d. random variable with zero mean. It can also be assumed that $E[e_{M,k}^2] < \infty$, which is reasonable if the system identification is done properly, and the measurement noise levels are finite. From the experiment results in the simulation section, it has been shown that the modeling error power is finite. Under those conditions and assumptions about nonlinear uncertainties on e_ks, the error dynamic function is changed to

$$e_{k+1} + \lambda_1 e_k + \lambda_2 e_{k-1} = -(e_{M,k+1} + \lambda_1 e_{M,k} + \lambda_2 e_{M,k-1}) \tag{15}$$

Define the right side of equation (15) as Δ. Considering the worst case when modeling errors reach their bound, Δ is set to its upper bound Δ_u ($|\Delta| \leq |\Delta_u|$ and Δ_u is a constant). Then we do Z transform and anti-Z transform to the function to get expression of e_k as

$$
\begin{aligned}
E(z) &= \frac{\Delta_u}{1 - z^{-1}} \cdot \frac{z^{-1}}{1 + \lambda_1 z^{-1} + \lambda_2 z^{-2}} \\
e_k &= \frac{\Delta_u}{(1 - p_1)(1 - p_2)} - \frac{\Delta_u}{(1 - p_1)(p_1 - p_2)} p_1^{k+1} \\
&\quad - \frac{\Delta_u}{(1 - p_2)(p_2 - p_1)} p_2^{k+1}
\end{aligned}
\tag{16}
$$

As $k \to \infty$, the mean of the tracking error is obtained as

$$
\begin{aligned}
\lim_{k \to \infty} E[e_k] &= \lim_{k \to \infty} E\left[\frac{\Delta_u}{(1 - p_1)(1 - p_2)}\right] \\
&= \lim_{k \to \infty} \frac{1 + \lambda_1 + \lambda_2}{(1 - p_1)(1 - p_2)} E[(-e_{M,k})] \\
&= E[-e_{M,\infty}]
\end{aligned}
\tag{17}
$$

As for the power of tracking error, the expression can be written from (13) and (17) as

$$
\begin{aligned}
\lim_{k \to \infty} E[e_k^2] &= \lim_{k \to \infty} E\left[\frac{\Delta_u^2}{(1-p_1)^2(1-p_2)^2}\right] \\
&= \lim_{k \to \infty} \frac{1+\lambda_1^2+\lambda_2^2}{(1-p_1)^2(1-p_2)^2} E[e_{M,k}^2] \\
&= \frac{1+\lambda_1^2+\lambda_2^2}{(1+\lambda_1+\lambda_2)^2} E[e_{M,\infty}^2]
\end{aligned}
\tag{18}
$$

The modeling error (and noise) contributions $e_{M,k}$ are wide sense stationary, and we can compute the asymptotic (finite) tracking error in terms of the power spectral density of $e_{M,k}$ and the coefficients λ_i. As discussed before, the mean of modeling error is zero or small value. Consequently, the modeling error will take effect on tracking performance and the pole placement problem must be solved considering the trade off between convergence speed (faster poles, wider pass-band) versus better disturbance rejection (slower poles, narrower pass-band).

REFERENCES

[1] O. Nelles, *Nonlinear System Identification* (New York: Springer, 2001).

[2] I.J. Leontaritis, S.A. Billings, Input-Output Parametric Models for Nonlinear Systems Part I: Deterministic Nonlinear Systems, *International Journal of Control*, 41(2), 1985, 303-328.

[3] K.S. Narendra, K. Parthasarathy, Identification and Control of Dynamical Systems Using Neural Networks, *IEEE Trans. Neural Networks*, 1(1), 1990, 4-27.

[4] T.A. Johansen, B.A. Foss, Constructing NARMAX Models Using ARMAX Models, *International Journal of Control*, 58(5), 1993, 1125-1153.

[5] X. Ni, M. Verhaegen, A.J. Krijgsman, H.B. Verbruggen, New Method for Identification and Control of Nonlinear Dynamic Systems, *Engineering Applications of Artificial Intelligence*, 9(3), 1996, 231-243.

[6] D.M. Walker, N.B. Tufillaro, P. Gross, Radial-Basis Models for Feedback Systems with Fading Memory, *IEEE Transactions on Circuits and Systems*, 48(9), 2001, 1147-1151.

[7] B.S. Kim, A.J. Calise, Nonlinear Flight Control Using Neural Networks, *Journal of Guidance, Control, and Dynamics*, 20(1), 1997, 26-33.

[8] C.H. Lee and M.J. Chung, Gain-Scheduled State Feedback Control Design Technique for Flight Vehicles, *IEEE Transactions on Aerospace and Electronic Systems*, 37(1), 2001, 173-182.

[9] R.C. Dorf, R.H. Bishop, Modern Control Systems, 8th ed., (New York: Addison, Wesley, 1998).

[10] J.S.R. Jang, ANFIS: Adaptive-Network-Based Fuzzy Inference System, *IEEE Transactions on Systems, Man and Cybernetics, Part B*, 23(3), 1993, 665-685.

[11] J. Lan, J.C. Principe, M.A. Motter, Identification of Dynamical Systems Using GMM with VQ Initialization, *Proceedings of IJCNN'03*, 1, 2003, 764-768.

[12] J.C. Principe, L. Wang and M.A. Motter, Local Dynamic Modeling with Self-Organizing Maps and Applications to Nonlinear System Identification and Control, *Proceedings of IEEE*, 86(11), 1998, 2240-2258.

[13] T. Kohonen, Self-Organizing Maps, (New York: Springer, 1995).

[14] G.J. Mclachlan, D. Peel, *Finite Mixture Models*, (New York: Wiley, 2001).

[15] J. Stark, D.S. Broomhead, M.E. Davies, J. Huke, Takens Embedding Theorems for Forced and Stochastic Systems, *Nonlinear Analysis, Theory Methods, and Applications*, 30(8), 1997, 5303-5314.

[16] M. Casdagli, Nonlinear Prediction of Chaotic Time Series, *Physica. D*, 35(3), 1989, 335-356.

[17] F. Takens, On Numerical Determination of the Dimension of an Attractor, in (D. Rand, L.S. Young eds.), *Dynamical Systems and Turbulance*, Warwick 1980, Lecture Notes in Mathematics, vol. 898, (Berlin: Springer-Verlag, 1981) 366-381.

[18] D.T. Magill, Optimal adaptive estimations of sampled stochastic processes, *IEEE Trans. Automat. Contro.*, AC-10, 1965, 434-439.

[19] D.G. Lainiotis, Partitioning: A unifying framework for adaptive systems–I: Estimation–II: Control, *Proc. IEEE*, 64, 1976, 1126-1143, 1182-1197.

[20] B. Martensson, *Adaptive stabilization*, doctoral diss., Lund Institute of Technology, Lund, Sweden, 1986.

[21] K.S. Narendra and J. Balakrishnan, Adaptive Control Using Multiple Models, *IEEE Transactions on Automatic Control*, 4(2), 1997, 171-187.

[22] R.A. Jacobs, M.I. Jordan, S.J. Nowlan, G.E. Hinton, Adaptive Mixtures of Local Experts, *Neural Computation*, 3, 1991, 79-87.

[23] R. Murray-Smith, T.A. Johansen, *Multiple Model Approaches to Modeling and Control*, (New York: Taylor & Francis, 1997).

[24] K.S. Narendra, C. Xiang, Adaptive Control of Discrete-Time Systems Using Multiple Models, *IEEE Transactions on Automatic Control*, 45(9), 2000, 1669-1686.

[25] A.S. Morse, Suporvisory control of families of linear set-point controller–Part 1: Exact matching, *IEEE Transactions on Automatic Control*, 42, 1996, 1413-1431.

[26] M.A. Motter, *Control of the NASA Langley 16-foot Transonic Tunne; with the Self-Organizing Feature Map*, doctoral diss., University of Florida, Gainesville, FL, 1997.

[27] G. Thampi, J.C. Principe, M.A. Motter, J. Cho, J. Lan, Multiple Model Based Flight Control Design, *Proceedings of MWSCAS'02*, 3, Tulsa, OK, 2002, 133-136.

[28] J. Cho, J. Lan, G. Thampi, J.C. Principe, M.A. Motter, Identification of Aircraft Dynamics Using a SOM and Local Linear Models, *Proceedings of MWSCAS'02*, 2, Tulsa, OK, 2002, 148-151.

[29] S. Chen, S. Billings, Representations of Non-Linear Systems: the NARMAX Model, *International Journal of Control*, 49(3), 1989, 1013-1032.

[30] R.E. Brown, G.N. Maliotis, J.A. Gibby, PID Self-Tuning Controller for Aluminum Rolling Mill, *IEEE Transactions on Industry Applications*, 29(3), 1993, 578-583.

[31] B. Widrow, S.D. Stearns, *Adaptive Signal Processing* (NJ: Prentice Hall, 1985).

[32] X. He, H. Asada, A New Method for Identifying Orders of Input-Output Models for Nonlinear Dynamic Systems, *Proc. ACC*, 1993, 2520-2524.

[33] K.M. Vu, Optimal Setting for Discrete PID Controllers, *IEE Proceedings-D*, 139(1), 1992, 31-40.

[34] J. Bao, J.F. Forbes, P.J. McLellan, Robust Multiloop PID Controller Design: A Successive Semidefinite Programming Approach, *Industrial and Engineering Chemistry Research*, 38, 1999, 3407-3419.

[35] S. Haykin, *Neural Networks A Comprehensive Foundation* (NJ: Prentice Hall, 1998).
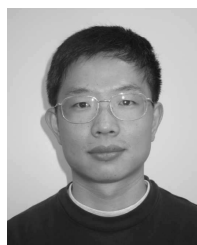
Jing Lan received his B.S. degrees in automatic control in 1996 from Beihang University, China, and the M.S. degree in automation, in 1999, from Tsinghua University, China. He is currently a Ph.D. candidate in electrical engineering at University of Florida under the supervision of Dr. Principe. His current research interests include adaptive control using multiple models, neural networks, and adaptive signal processing.
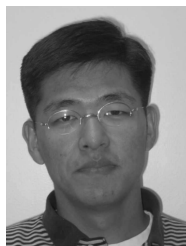
Jeongho Cho received his B.S. degree from Soonchunhyang University, Korea, majoring in control and instrumentation engineering in 1995. And he earned M.S. degrees in electrical engineering from Dongguk University, Korea, and University of Florida in 1997 and 2001, respectively. Since 2001, he has been with the Computational NeuroEngineering Laboratory at the University of Florida to pursue his Ph.D. degree. He is currently involved in the research project to develop an algorithm for modeling and control of unmanned aerial vehicle.

Deniz Erdogmus received his B.S. degrees in electrical engineering and mathematics in 1997, and his M.S. degree in electrical engineering, with emphasis on systems and control, in 1999, all from the Middle East Technical University, Turkey. He received his Ph.D. in electrical engineering from the University of Florida, Gainesville, in 2002. From 1999 to 2004, he has been with the Computational NeuroEngineering Laboratory at the University of Florida, working with Jose Principe. Currently, he is an assistant professor of computer science and biomedical engineering at the Oregon Graduate Institute at the Oregon Health and Science University. His current research interests include information theoretic aspects of adaptive signal processing and machine learning, with applications to problems in biomedical signal processing, communications, and controls. He is the recipient of the IEEE-SPS 2003 Young Author Best Paper Award, and the INNS 2004 Young Investigator Award. He is a member of IEEE, Tau Beta Pi, and Eta Kappa Nu. He has coauthored 3 book chapters, and more than 25 papers in refereed journals, and 70 papers in international conferences.

José C. Principe is Distinguished Professor of Electrical and Biomedical Engineering at the University of Florida, Gainesville, where he teaches advanced signal processing and artificial neural networks (ANN's) modeling. He is BellSouth Professor and Founder and Director of the University of Florida Computational NeuroEngineering Laboratory (CNEL). He has been involved in biomedical signal processing, in particular the electroencephalogram (EEG), and the modeling and applications of adaptive systems. Dr. Principe is Editor in Chief IEEE Transactions on Biomedical Engineering, President Elect of the International Neural Network Society, and formal Secretary of the Technical Committee on Neural Networks of the IEEE Signal Processing Society. He is also a member of the Scientific Board of the Food and Drug Administration, and a member of the Advisory Board of the University of Florida Brain Institute. He has more than 90 publications in refereed journals, 10 book chapters, and over 190 conference papers. He has directed 39 Ph.D. degree dissertations and 57 master degree theses.
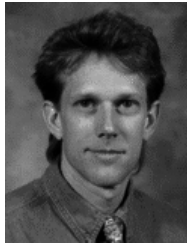
Mark A. Motter was born in Columbia, Pennsylvania, on September 3, 1955. He served in the United States Navy from 1973 until 1979, and was honorably discharged at the rank of Electronics Technician First Class. He then began his formal engineering education at Old Dominion University in Norfolk, Virginia, receiving his BSEE, magna cum laude, and MSEE in 1983 and 1985, respectively. Dr. Motter received his Ph.D. in Electrical and Computer Engineering from the University of Florida in 1998.

Since 1985 he has been employed at NASA Langley Research Center, initially involved in the modeling and control of wind tunnels and associated experimental equipment. Currently, he is a controls research engineer in the Electronics Systems Branch. His current research project is investigating the implementation of self-organizing controllers as well as other biologically inspired control approaches, using a fully autonomous unmanned aerial vehicle, which has successfully completed over 500 test flights. He is a senior member of the IEEE, a member of the Academy of Model Aeronautics, and a registered Professional Engineer.
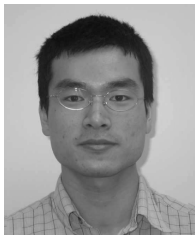
Jianwu Xu received his B.S. degree in Electrical Engineering in 2002 from Zhejiang University, Hangzhou, China. Since 2002, he has been with Computational NeuroEngineering Laboratory, University of Florida, under the supervision of Dr. Principe. His current research interests include information theoretic learning, adaptive signal processing, control and machine learning.