Neurocomputing 74 (2010) 315–327

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

An MDL-based Hammerstein recurrent neural network for control applications

Jeen-Shing Wang *, Yu-Liang Hsu

Department of Electrical Engineering, National Cheng Kung University, Tainan 701, Taiwan, ROC

ARTICLE INFO

Article history:
Received 14 August 2009
Received in revised form
8 January 2010
Accepted 15 March 2010
Communicated by W. Yu
Available online 7 April 2010

Keywords:
Hammerstein model
Recurrent neural networks
Minimum description length
Nonlinearity eliminator

ABSTRACT

This paper presents an efficient control scheme using a Hammerstein recurrent neural network (HRNN) based on the minimum description length (MDL) principle for controlling nonlinear dynamic systems. In the proposed control approach, an unknown system is first identified by the MDL-based HRNN, which consists of a static nonlinear model cascaded by a dynamic linear model and can be expressed in a state-space representation. For high-accuracy system modeling, we have developed a self-construction algorithm that integrates the MDL principle and recursive recurrent learning algorithm for constructing a parsimonious HRNN in an efficient manner. To ease the control of the system, we have established a nonlinearity eliminator that functions as the inverse of the static nonlinear model to remove the global nonlinear behavior of the unknown system. If the system modeling and the inverse of the nonlinear model are accurate, the compound model, the unknown system cascaded with the nonlinearity eliminator, will behave like the linear dynamic model. This assumption turns the task of complex nonlinear control problems into a simple feedback linear controller design. Hence, well-developed linear controller design theories can be applied directly to achieve satisfactory control performance. Computer simulations on unknown nonlinear system control problems have successfully validated the effectiveness of the proposed MDL-based HRNN and its control scheme as well as its superiority in control performance.

1. Introduction

In most cases, the mathematical model of a plant and its linear approximation are required for the design of controllers [24,33]. However, due to uncertainties in parameters, unexpected disturbances and noises, practical systems are nonlinear and the acquisition of their mathematical expressions may become difficult if not impossible. In the past decades, many research efforts have been directed to addressing these issues in the areas of adaptive control [22,26], nonlinear mathematics [17,29], and robust control [2,5]. In response to the increasingly nonlinear and complex dynamic systems, many researchers have turned away from conventional control approaches to intelligent-based approaches [3,32]. Among these intelligent-based approaches, neural-network (NN) based controllers have been developed to compensate for the effects of nonlinearities and system uncertainties so that the stability, convergence, and robustness of the control system can be improved. Thorough reviews on neural-network-based control systems can be found in [9,13,35,36].

Among the diverse neural network structures, recurrent neural networks (RNNs) have been recognized as one of the most effective tools for modeling and controlling complex dynamic systems due to their learning capability and flexibility in incorporating "dynamics" into the structures. Among the successful applications [11,16,31,34], to name a few, Ku and Lee [16] proposed a partially connected diagonal recurrent neural network (DRNN) that has better learning capability than fully connected RNNs do. To control an unknown plant, they adopted a model reference adaptive control structure by using the DRNN as the plant identifier to provide a channel for propagating the error signal to fine-tune a DRNN-based controller. Huang and Lewis [11] developed an RNN-based predictive feedback control scheme for uncertain nonlinear time-delayed systems. In their approach, an RNN was used to estimate the dynamics of the delay-free nonlinear system, and a nonlinear compensator was extracted from the RNN to remove the system's nonlinear effects. This results in a so-called linearized local subsystem that facilitates the design of the remote predictive controller for handling time-delay problems. Recently, Zhu and Guo [34] designed a linear generalized minimum variance (GMV) controller based on the linearization of a nonlinear system around operating points. To reinforce the control performance, a nonlinear compensator implemented by an RNN was devised to provide compensation for the system nonlinearities.

From the above studies, a general agreement can be reached that a linearized model derived from an unknown nonlinear system can often capture the significant dynamics of the system

* Corresponding author.
E-mail address: jeenshin@mail.ncku.edu.tw (J.-S. Wang).

0925-2312/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2010.03.011

around operating points and thus can provide a good basis for controller design. In this study, we have developed an efficient control approach that integrates system modeling and controller design into a unified framework. That is, we first design a Hammerstein recurrent neural network (HRNN) consisting of a static nonlinear model and a dynamic linear model for modeling an unknown system into a state-space representation. Then, we develop a self-construction algorithm, integrating the MDL principle and recursive recurrent learning algorithm, to automate the construction of the proposed recurrent network. Based on the identified Hammerstein network, which models the dynamics of the unknown system, we generate a nonlinearity eliminator (NLE) to remove the nonlinear effects caused by the static nonlinear model of the Hammerstein network. This results in an unknown system with a trained NLE that behaves like the dynamic linear model of the Hammerstein network. With the linear model, we can design a linear controller by using well-developed linear control theories directly and effortlessly. The major contribution of this paper is as follows. In most controller design problems, identification and control are usually treated as two separable tasks. However, we regard these two tasks as integral and take full advantage of the procedures conducted for system modeling to design an effective but simple linear controller. The advantages of our approach include: (1) the proposed HRNN with the proposed identification algorithm can describe the nonlinear dynamics of a given unknown system as a state-space equation; (2) the proposed MDL algorithm can automatically determine a compact network structure for a system identification problem; and (3) the network controllability and observability can be analyzed by the state-space equation. However, the proposed MDL-based structure learning algorithm cannot be performed online. This is one of the limitations for using the MDL principle in the structure learning phase of recurrent network-based modeling problems.

The rest of this paper is organized as follows. In Section 2, we introduce the network architecture of the HRNN and the self-construction algorithm for establishing a parsimonious HRNN. The concept of the efficient control scheme, the design procedure of the nonlinear eliminator for removing the nonlinearity of the HRNN, and the relative control strategy are presented in Section 3. Section 4 provides the computer simulations to validate the effectiveness of the control approach, while Section 5 is devoted to conclusions.

2. Structure of MDL-based HRNN and its identification algorithm

According to [23], adaptive control can be classified into two major categories: direct and indirect approaches. In the direct approach, the controller is learned directly so as to minimize the difference between the desired outputs and plant outputs. In the indirect approach, the design of the controller requires two procedures: modeling of the unknown plant and training of the controller based on the identified model. Here, we propose a recurrent-network based indirect control scheme that unifies an MDL-based HRNN with a self-construction algorithm to systematize the designs of all necessary control components without a trial-and-error approach or any user manipulation. To explain our philosophy in developing such a control scheme, we first introduce the structure of our MDL-based HRNN and its relative identification algorithm. The overall control scheme and the methodology of controller design are provided in Section 3.

2.1. Structure of HRNN

The model identification of unknown systems (plants) is an important and integral part of control design methodology, since from the model we can get a physical understanding of the system and develop a simulation from which a control law is designed [19]. To obtain good model identification for the system requires the following considerations: (1) the selection of models for describing the system, (2) the construction of quality models to best fit the system, and (3) the accuracy of models in representing the system. These issues are integral parts for the development of effective system identification tools. Recent research studies have shown that the Hammerstein model, consisting of a nonlinear static subsystem cascaded with a dynamic linear subsystem, is one of the effective models for dealing with nonlinear problem [4,12,15]. In this paper, we present a HRNN that is capable of precisely capturing the dynamics of the true system with a transparent and concise network representation. The conventional Hammerstein model is shown in Fig. 1a. The novelty of this network is to incorporate dynamic elements in such a way that each state corresponds to a neuron that linearly combines its own output history and the outputs from other neurons. Such a deployment of dynamic elements enables the proposed structure to be mapped into a state-space equation from its internal structure. Fig. 2a shows the structure of the proposed HRNN. The structure consists of four layers, which can be expressed by the block diagram shown in Fig. 2b.

The whole HRNN can be classified into two major components: a static nonlinear model and a dynamic linear model. The static nonlinear model maps the input space into a state space via a nonlinear transformation and then the state space is transformed into the output space through a linear dynamic mapping. The state space equations of the proposed network are expressed as

$$\mathbf{x}(k+1)=\mathbf{A}\mathbf{x}(k)+\mathbf{B}\mathbf{N}(\mathbf{u}(k)),$$
$$\mathbf{y}(k)=\mathbf{C}\mathbf{x}(k), \tag{1}$$

where $A\in\mathbb{R}^{J\times J}$, $B\in\mathbb{R}^{J\times J}$, $C\in\mathbb{R}^{m\times J}$, $N\in\mathbb{R}^{J}$, $\mathbf{u}=[u_1,\ \ldots,\ u_p]^T$ is the input vector, $\mathbf{y}=[y_1,\ldots,y_m]^T$ is the output vector, and p and m are the dimensions of the input and output layers, respectively. The elements of matrix $\mathbf{A}$ stand for the degree of inter-correlation among the states. $\mathbf{B}$ is a diagonal matrix with diagonal elements $[b_{11},\ \ldots,\ b_{JJ}]^T$, representing the weights of the inputs of the dynamic layer. The elements of matrix $\mathbf{C}$ are the weights of the states. $\mathbf{N}=[n_1,\ldots,n_J]^T$ is the nonlinear function vector. $\mathbf{x}=[x_1,\ \ldots,\ x_J]^T$ is the state vector, where J is the total number of state variables and is equal to the number of neurons in the hidden layer and the dynamic layer. The current ith output $y_i(k)$ and the state variables $\mathbf{x}(k)$ are obtained by calculating the activities of all nodes on each layer, and the corresponding functions are

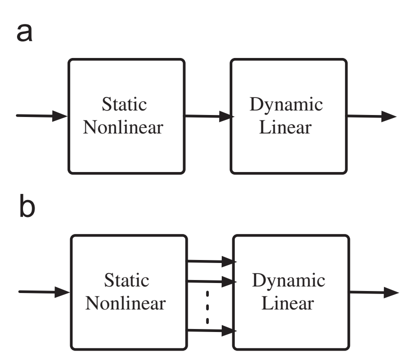

Fig. 1. The block diagram of Hammerstein models. (a) A general single-input-single-output Hammerstein model. (b) The proposed single-input-single-output recurrent neural network with J system order.