Neurocomputing 77 (2012) 12–19

Contents lists available at SciVerse ScienceDirect

Neurocomputing

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

An adaptive wavelet differential neural networks based identifier and its stability analysis

F. Jahangiri*, A. Doustmohammadi, M.B. Menhaj

Amirkabir University of Technology, Tehran, Iran

ARTICLE INFO

Article history:
Received 1 July 2010
Received in revised form
13 July 2011
Accepted 23 July 2011
Communicated by J. Zhang
Available online 3 September 2011

Keywords:
Adaptive wavelet differential neural network
Unknown dynamic nonlinear system
Lyapunov analysis
Nonlinear plant
Neuro-identifier

ABSTRACT

In this paper, identification problem of a general class of nonlinear dynamic systems is fully considered using adaptive wavelet differential neural networks. In these networks, the activation functions are described by wavelets where parameters are tuned adaptively. The stability analysis of such identifiers is performed by means of Lyapunov analysis. Asymptotic convergence of the error and boundedness of the parameters are proven. To validate the approach, the neuro-identifier is applied to both the Van der pole oscillator and the twin-tanks plant. The simulation results show that the proposed neuro-identifier outperforms the sigmoid based differential neural network identifier.

1. Introduction

In recent years, neural-networks (*NNs*) have been applied extensively to model nonlinear systems [1–7]. Although *NNs* are powerful tools for handling problems of large dimensions, nevertheless, the implementation of neural networks suffers from the lack of efficient constructive methods, both for determining the parameters of neurons and for choosing network structure. Most *NN* structures seen in literature use the sigmoid activation functions in neurons. The drawbacks of using such functions are that they are not orthogonal and their energies are infinite leading to a slow convergence rate. Also, *NNs* with sigmoid activation functions cannot accurately characterize local features that typically embody important information about the system such as discontinuities in curvature or jumps in objective function [8,9].

On the other hand, the wavelet decomposition can be used for approximation problems. Wavelets are local functions with limited durations. Although the wavelet theory has offered efficient algorithms for various purposes, their implementations are usually limited to wavelets of small dimensions. This is due to the fact that constructing and storing wavelet basis of large dimensions are of prohibitive cost [10]. To be able to handle problems of larger dimensions, it is necessary to develop algorithms that are scalable, i.e., less sensitive to dimensions.

Because of the similarity between wavelet decomposition and one-hidden-layer *NN*, the idea of combining both wavelets and *NNs* was proposed by Zhang and Benveniste [11]. Localization properties of wavelets together with learning abilities of *NN* can result in networks with efficient constructive methods that are capable of handling problems of moderately large dimensions [11–13]. Also, improved localized modeling can aid both data reduction and subsequent classification tasks that rely on accurate representation of local features [13,14]. Wavelet neural networks (*WNNs*) have the advantage over *NNs* in a sense that they can capture local information. Many researchers [15–19] have used such structures for solving approximation, classification, prediction, control and many other problems.

The existing *WNNs* are classified into two categories: fixed grid *WNNs* and adaptive *WNNs* [19]. In fixed grid *WNNs*, wavelets stem from the discrete wavelet transforms and the unknown dilation and translation factors of wavelets vary on some fixed discrete lattices. Hence, these parameters are fixed and only the networks weights are optimized in training phase. The number of candid wavelets in a fixed grid *WNN* often increases dramatically with the order of the model. Consequently, these *WNNs* are often limited to low dimensions; however, in adaptive *WNNs* wavelets stem from the continuous wavelet transform and the unknown continuous parameters (the weighting coefficients, the dilation, and the translation factors) are adjusted recursively. These networks have

* Corresponding author.
E-mail addresses: jahangiri_ftm@aut.ac.ir (F. Jahangiri), dad@aut.ac.ir (A. Doustmohammadi), menhaj@aut.ac.ir (M.B. Menhaj).

0925-2312/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2011.07.013

been successfully applied to nonlinear static function approximation and classification [19,20].

*NN*s with feedback have extensive applications since most systems to be modeled and controlled are indeed dynamic in nature. Therefore, differential neural networks (*DNN*s) were introduced [21]. In view of *DNN* continuous structure, more detailed techniques must be applied to answer important questions such as stability, convergence, etc. *Lyapunov*'s theorem is the main tool to prove stability of *DNNs* that are used to improve estimation problems or to control action design [22]. In [23], the wavelet theory over the *DNN* structures has been applied for state estimation and nonparametric identification. The activation functions used to approximate the uncertain nonlinear functions were orthogonal and non-redundant basis of wavelet functions. A learning law containing an adaptive adjustment rate is suggested there to imply the stability condition for the free parameters of the observer. *Lyapunov* theorem is used to obtain the upper bounds for both the weights dynamics and the mean squared estimation error.

To reduce the number of wavelet candidates required to identify the system, the adaptive wavelet differential neural networks (*AWDNN*s) are fully developed in the paper. In these networks, in addition to weights of the network, the dilation and translation factors of the wavelets are also tuned adaptively. This paper presents a new method, which is indeed an extension of previous one, in which the adaptive rule for tuning the identifier parameters derived through *Lyapunov* stability analysis. Asymptotic convergence of the identification error and boundedness of the parameters are further proven.

This paper is organized as follows. Section 2 describes approximation of any function belonging to L_2 space using adaptive *WNN* and it also describes *AWDNN* identifier structure. Stability analysis of *AWDNN* identifier is discussed in Section 3. Two illustrative examples are given in Section 4 that demonstrate the effectiveness of the *AWDNN*. Finally, conclusion is given in Section 5.

2. AWDNN identifier

In this section, the concept of an adaptive *WNN*, its capability to approximate nonlinear systems, and the identifier structure of *AWDNN* are described.

2.1. Adaptive wavelet neural network approximation

The wavelet analysis procedure is implemented with dilated and translated versions of a mother wavelet. Several kinds of mother wavelets have been developed. Examples of these wavelets are: *Daubechies*, *Morlet*, *Mexican Hat*, *Meyer*, etc. In theory, the dilation parameter of a wavelet can be any positive real value and the translation can be an arbitrary real number. This is referred to as the continuous wavelet transform. It is well known that any function $h(x)\in L_2(R^n)$, can be reconstructed by the inverse wavelet transform [12,13]. It is shown in [13] that the discrete representation of the inverse continuous wavelet transform can be written as

$$h(x)=\sum_i \omega_i\psi(b_i(x-a_i)) \tag{1}$$

In the above equation, $\psi(\cdot)$ is the wavelet function, b_i and a_i are the dilation and translation parameters, ω_i is the weight of ith wavelet function, and n is the dimension of x. For discrete version of the reconstruction of $h(x)$ to hold, some conditions must be satisfied; for example, one may look for some countable sets of a_i, and b_i, such that the corresponding families of dilated and translated wavelets given by the following:

$$\{b_i^{1/2}\psi(b_i(x-a_i)) : i\in Z\} \tag{2}$$

constitute an orthonormal basis of some functional space (typically, $L_2(R^n)$). In order to generate an orthonormal basis, the wavelet function has to satisfy strong restrictions and a compromise between the regularity and the compactness of the wavelet function is necessary. This may be considered as a first solution for providing the aforementioned conditions required.

The other solution consists of nonorthogonal wavelet families, in particular wavelet frames. By relaxing the orthogonality, much more freedom on the choice of wavelet function is achieved.

Definition. A sequence $\{\psi_i{:}i\in Z\}$ in a *Hilbert* space, H, is called a frame of H, if there exist two constants $A>0$ and $B<\infty$ such that for all $h\in H$ the following inequalities hold:

$$A\|h\|^2\le\sum_{i\in Z}|\langle h,\psi_i\rangle|^2\le B\|h\|^2 \tag{3}$$

where $\langle,\rangle$ is the inner product. Under the frame condition, h can be recovered. Hence, if family (2) constitutes a frame, then the discrete reconstruction formula (1) is ensured. However the applications of orthonormal wavelet bases and wavelet frames are usually limited to problems of small dimensions. For practical implementations, infinite wavelet bases and frames are always truncated that approximate h. The number of wavelets in a truncated basis or frame drastically increases with the dimension, therefore, constructing and storing wavelet bases or frames of large dimension are of prohibitive cost. If the inverse continuous wavelet transform is discretized according to the distribution of the data, we can expect to reduce the number of wavelets needed in function approximation. It is thus possible to handle problems of large dimensions with such adaptive discretization of the inverse wavelet transform. The goal of the adaptive discretization is to determine the parameters ω_i, a_i and b_i in (1) according to a data sample. This problem is very similar to neural-network training. As a matter of fact, (1) can be viewed as a one-hidden layer neural network with ψ as the activation function of the hidden neurons and with a linear neuron in the output layer. The adaptively discretized inverse wavelet transform is therefore referred to as adaptive *WNN*.

2.2. Adaptive wavelet differential neural network identifier

Consider an unknown dynamic nonlinear system described by the following equation:

$$\dot{x}=f(x,u,t) \tag{4}$$

where $x\in R^n$ and $u\in R^m$ $(m\le n)$ are the state vector and the input vector of the system, respectively.

In view of the adaptive *WNN* capabilities, as discussed in previous section, any function $f\in L_2$ could be approximated. The following approximation formula is used [23]:

$$f(x,u,t)\cong\beta x+W^{1*}\Delta(x)+W^{2*}\Phi(x)\gamma(u) \tag{5}$$

In the above equation, $\beta\in R^{n\times n}$, $W^{1*}\in R^{n\times N_1}$, $W^{2*}\in R^{n\times N_2}$, $\Delta(x)=[\delta_k(x)]_{N_1\times 1}$, $\Phi(x)=[\phi_{ij}(x)]_{N_2\times q}$, and $\gamma\in R^{q\times 1}$ where it is assumed that the number of states of the system is given by n. The symbol $\gamma(\cdot)$ is defined as the vector field from R^m to R^q. The activation function matrices, Δ and Φ are constructed with components defined by dilated and translated wavelet functions

$$\delta_k(x)=\delta(b_k^*(x-a_k^*));\quad k=1,2,\ldots,N_1$$

$$\phi_{ij}(x)=\phi(d_i^{j*}(x-t_i^{j*}));\quad i=1,2,\ldots,N_2,\quad j=1,2,\ldots,q \tag{6}$$

where a_k^*, $t_i^{j*}\in R^n$ and b_k^*, $d_i^{j*}\in R^+$. For the case $n>1$, one needs to use one of the available multidimensional wavelet functions [24]. The one that is used in this paper is *Euclidean* norm of the argument. The mother wavelet chosen in this paper, for both