Neural Networks 69 (2015) 51–59

Contents lists available at ScienceDirect

Neural Networks

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

Improving nonlinear modeling capabilities of functional link adaptive filters

Danilo Comminiello *, Michele Scarpiniti, Simone Scardapane, Raffaele Parisi, Aurelio Uncini

Department of Information Engineering, Electronics and Telecommunications (DIET), "Sapienza" University of Rome, Via Eudossiana 18, 00184 Rome, Italy

HIGHLIGHTS

- This paper proposes an improved split functional link adaptive filter (SFLAF).
- The proposed model is characterized by the adaptive combination of two APA filters.
- An advanced scheme is also proposed involving the combination of multiple filters.
- The adaptive combinations are performed for all the projections of the APA filters.
- The proposed models are assessed in three different nonlinear modeling problems.

ARTICLE INFO

Article history:
Received 19 March 2015
Received in revised form 14 May 2015
Accepted 17 May 2015
Available online 27 May 2015

Keywords:
Functional links
Nonlinear modeling
Adaptive combination of filters
Online learning algorithms

ABSTRACT

The *functional link adaptive filter* (FLAF) represents an effective solution for online nonlinear modeling problems. In this paper, we take into account a FLAF-based architecture, which separates the adaptation of linear and nonlinear elements, and we focus on the nonlinear branch to improve the modeling performance. In particular, we propose a new model that involves an adaptive combination of filters downstream of the nonlinear expansion. Such combination leads to a cooperative behavior of the whole architecture, thus yielding a performance improvement, particularly in the presence of strong nonlinearities. An advanced architecture is also proposed involving the adaptive combination of multiple filters on the nonlinear branch. The proposed models are assessed in different nonlinear modeling problems, in which their effectiveness and capabilities are shown.

1. Introduction

Nonlinear modeling problems have always aroused great interest in the research community. In particular, applications requiring an online modeling of nonlinearities have led to the development of many linear-in-the-parameter (LIP) nonlinear models, which consist in a nonlinear expansion of the input followed by a linear model. This approach derives from Cover's Theorem on the separability of patterns (see Haykin, 2008), which ensures universal approximation capabilities given a sufficiently large number of nonlinear elements.

Among the family of LIP nonlinear models for online learning, representative examples include adaptive Volterra models (Azpicueta-Ruiz, Zeller, Figueiras-Vidal, Kellermann, & Arenas-García, 2013; Zhao & Zhang, 2009a), regularized networks (Poggio & Girosi, 1990; Solazzi & Uncini, 2004), spline adaptive filters (Scarpiniti, Comminiello, Parisi, & Uncini, 2013; Vecci, Piazza, & Uncini, 1998), even mirror Fourier nonlinear filters (Carini & Sicuranza, 2014), kernel adaptive filters (Fan & Song, 2014; Zhu, Chen, & Príncipe, 2012), online extreme learning machines (Huang, Huang, Song, & You, 2015; Scardapane, Comminiello, Scarpiniti, & Uncini, in press). In this work, we focus on a class of LIP nonlinear filters based on the *functional links* (Pao, 1989; Pao & Beer, 1988). The functional link is a functional operator, which allows to represent an input pattern in a feature space where its processing turns out to be enhanced. The functional links have been widely used in single-layer feedforward neural networks, named *functional link artificial neural networks* (FLANNs), or also functional link networks (FLNs) (Amin, Savitha, Amin, & Murase, 2012; Patra, Pal, Chatterji, & Panda, 1999; Scardapane, Wang, Panella, & Uncini, 2015; Zhao & Zhang, 2009b). They have also been used in conjunction with adaptive filters for online learning applications, with the name of FLANN filters (Sicuranza & Carini, 2011) or also *functional link adaptive filters* (FLAFs) (Comminiello, Azpicueta-Ruiz, Scarpiniti,

* Corresponding author. Tel.: +39 06 44585495; fax: +39 06 4873300.
E-mail address: danilo.comminiello@uniroma1.it (D. Comminiello).

http://dx.doi.org/10.1016/j.neunet.2015.05.002

Uncini, & Arenas-Garcia, 2011; Comminiello, Scarpiniti, Azpicueta-Ruiz, Arenas-García, & Uncini, 2013).

In this paper, we take into account a functional link-based filter, the *split FLAF* (SFLAF) (Comminiello, Scarpiniti, Azpicueta-Ruiz et al., 2013), which separates the adaptation of linear and nonlinear elements, thus performing two distinct tasks in parallel: the estimation of the linear impulse response and the modeling of nonlinearities. With respect to prior works on FLAFs (Comminiello et al., 2011; Comminiello, Scarpiniti, Azpicueta-Ruiz et al., 2013), we propose a new architecture, called *combined SFLAF* (cS-FLAF), in which the nonlinear path is characterized by an adaptive combination of two filters downstream of the nonlinear expansion. Adaptive combination of filters exploits the diversity of parallel branches to improve the filtering performance when no much information is *a priori* provided on the model of signal to be processed (Arenas-García, Martínez-Ramón, Navia-Vázquez, & Figueiras-Vidal, 2006). In this regard, many efforts have been made in the linear case (Arenas-García et al., 2006; Comminiello, Scarpiniti, Parisi, & Uncini, 2013; Silva & Nascimento, 2008), but in the nonlinear case the combined output takes into account directly the joint effect of linear and nonlinear filtering (Azpicueta-Ruiz, Zeller, Figueiras-Vidal, Arenas-García, & Kellermann, 2011; Comminiello, Scarpiniti, Azpicueta-Ruiz et al., 2013).

Here, we investigate the effects of the combination of two purely nonlinear outputs that leads to an improvement of the task of nonlinear modeling. The two adaptive filters on the nonlinear branch are updated by using the same *affine projection algorithm* (APA) (Ozeki & Umeda, 1984) but with different projection orders. In particular, choosing a unitary projection order for one filter and a higher order for the other one, we provide the two filters with different adaptation rules, respectively a gradient-based one and a Hessian-based one. This gives robustness to the model, which is able to provide improved performance, especially when an unknown system introduces very strong nonlinearities. Another novel insight in this architecture is represented by the fact that the adaptive combination is performed involving not only the current projection, as in Arévalo, Apolinário, de Campos, and Sampaio-Neto (2013), Comminiello, Scarpiniti, Parisi et al. (2013) and Ferrer, de Diego, Gonzalez, and Piñero (2009), but all the available ones.

We also propose an advanced combined architecture involving the adaptive combination of three APAs downstream of the functional link expansion. This model further improves the nonlinear modeling performance by taking advantage of the capabilities of the individual filters. The proposed models are assessed in several nonlinear system identification problems showing the performance capabilities of the combined architectures that, according to the system to be identified, can just select the best performing filter or take advantage of all the filters, giving rise to an emerging learning behavior.

The rest of the paper is organized as follows: the nonlinear FLAF model is described in Section 2 and then, the proposed combined FLAF-based architecture is presented in Section 3. The advanced combined architecture is described in Section 4, while experimental results in Section 5 prove the effectiveness of the proposed architectures in different nonlinear modeling scenarios. Finally, in Section 6 our conclusions are drawn.

1.1. Notation

In this paper, matrices are represented by boldface capital letters and vectors are denoted by boldface lowercase letters. Time-varying vectors and matrices show discrete-time index as a subscript index, while in time-varying scalar elements the time index is denoted in square brackets. A regression vector is represented as $\mathbf{x}_n \in \mathbb{R}^M = \begin{bmatrix} x[n] & x[n-1] & \cdots & x[n-M+1] \end{bmatrix}^T$, where M is the overall vector length and $x[n-i]$ is individual

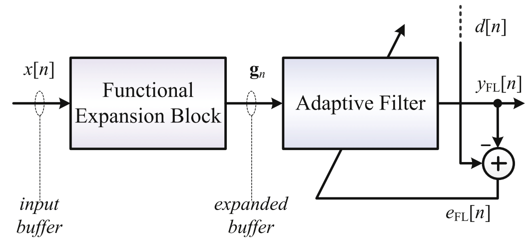

Fig. 1. The nonlinear functional link adaptive filter.

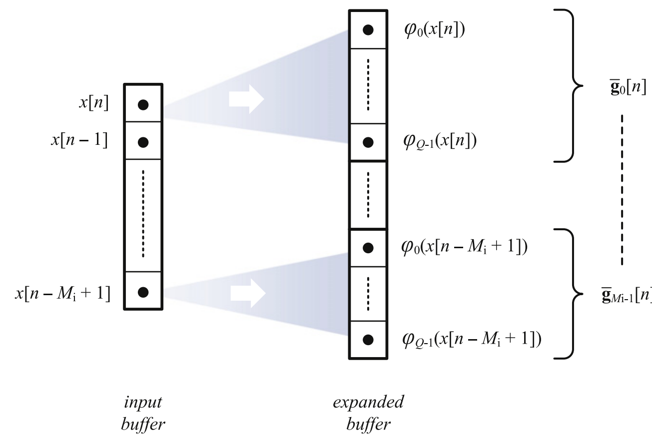

Fig. 2. Memoryless functional link expansion.

entry at the generic time instant $n-i$. A generic coefficient vector, in which all the elements depend on the same time instant, is denoted as $\mathbf{w}_n \in \mathbb{R}^M = \begin{bmatrix} w_0[n] & w_1[n] & \cdots & w_{M-1}[n] \end{bmatrix}^T$, where $w_i[n]$ is the generic ith individual entry at the n-th time instant. The index related to a generic jth filter is denoted as subscript, before the time index for vectors and matrices, e.g. $\mathbf{w}_{j,n}$. All vectors are represented as column vectors.

2. Nonlinear functional link adaptive filter

The FLAF model is based on the representation of the input signal in a higher-dimensional space (Pao, 1989), where an enhanced nonlinear modeling is allowed. Such approach derives from machine learning theory, more precisely from Cover's Theorem on the separability of patterns (see for example Haykin, 2008).

The purely nonlinear FLAF is composed of two main parts: a nonlinear *functional expansion block* (FEB) and a subsequent linear adaptive filter, as depicted in Fig. 1. The FEB consists of a series of functions, which might be a subset of a complete set of orthonormal basis functions satisfying universal approximation constraints. The term "functional links" actually refers to the functions contained in the chosen set $\Phi = \left\{\varphi_0(\cdot), \varphi_1(\cdot), \ldots, \varphi_{Q-1}(\cdot)\right\}$, where Q is the number of functional links. As depicted in Fig. 2, at the n-th time instant, the FEB receives the input sample $x[n]$, which is stored in an input buffer $\mathbf{x}_n \in \mathbb{R}^{M_{\mathrm{in}}} = \begin{bmatrix} x[n] & x[n-1] & \cdots & x[n-M_{\mathrm{in}}+1] \end{bmatrix}^T$, where M_{in} is defined as the input buffer length. Each element of $\mathbf{x}_n$ is passed as argument to the chosen set of functions Φ, thus yielding a subvector $\bar{\mathbf{g}}_{i,n} \in \mathbb{R}^Q$:

$$\bar{\mathbf{g}}_{i,n} = \begin{bmatrix} \varphi_0(x[n-i]) \\ \varphi_1(x[n-i]) \\ \vdots \\ \varphi_{Q-1}(x[n-i]) \end{bmatrix}. \tag{1}$$