Intelligent Signal Estimation Using Cosine Neural Networks with Variable Structure Systems Based Training Procedure

M. Onder Efe, Okyay Kaynak and Feza Kerestecioglu
Bogazici University, Electrical and Electronic Engineering Department
Bebek, 80815 Istanbul, Turkey
{efemond, kaynak, kerestec}@boun.edu.tr

Abstract

This paper demonstrates the estimation of signals by using a neural network structure composed of cosine neurons. The building blocks of the architecture are cosine components with adjustable amplitude, frequency and phase. The training procedure is based on the mixture of gradient descent with a method utilizing sliding mode control philosophy. The proposed use of mixed information in training dynamics leads to the minimization of the cost of estimation as well as the cost of stability. Two application examples are presented in the paper. The first example considers the reconstruction of a time signal having finite frequency components in the spectrum. The second example shows the reconstruction of the frequency plot of a FIR filter.

1. Introduction

Signal estimation is an interdisciplinary area, the techniques of which are required in most engineering practice. The main issue in estimation theory is to construct an appropriate estimator, which gives descriptive information about the signal or system to be identified. Some approaches dealing with signal estimation adopt methods in the least squares sense [1].

Artificial neural networks have been studied extensively in the context of least squares [2]. Various applications mentioned in the literature utilize the least squares based training procedures. A common property of the neural systems is their mathematical tractability in the sense of evaluating the sensitivity derivatives easily. In this study, a simple way for the estimating the spectra of unknown signals is proposed with the introduction of a new neuron model, which we call *cosine neurons*. The proposed estimation scheme is carried out in time domain and operates on-line. The use of cosine neurons provides an important information about the shape of the frequency spectrum of the input signal. In general, if there is no closed-form time-domain equation of the input signal, it is considerably difficult to obtain its frequency domain representation. Therefore a need arises for the use of an intelligent technique. The proposed approach has this characteristic and the fact that it can operate on-line makes it especially attractive.

At this point, the mechanism introducing the intelligence determines the overall performance of the estimator. Typically, the methods using the sensitivity derivatives of the cost function suffer from the shape of the cost surface, which is a multidimensional surface. In tuning the parameters, the information is extracted from the partial derivative of the cost function with respect to an adjustable parameter, along which the derivative may assume large values. Therefore, imposing a stabilizing force on training dynamics can eliminate the local unstability of gradient based techniques. This could be achieved by the use of Sliding Mode Control (SMC) philosophy in the training phase.

Earliest notion of SMC strategy was constructed on a second order system in the late 1960s by Emelyanov [3]. The work stipulated that a special line could be defined on the phase plane, such that any initial state vector can be driven towards the plane and then be maintained on it, while forcing the error dynamics towards the origin. The concept introduced by Emelyanov has first been applied to simple systems of order two. Since then, the theory has greatly been improved and a well defined design framework has been established. The sliding line has taken the form of a multidimensional surface, called the sliding surface and the function defining it is called the switching function. The main advantage introduced by the use of SMC approach is its robustness to unmodeled dynamics of the system under control. In this paper the system corresponds to the training dynamics.

Latest studies consider SMC approach with adjustable design parameters [4-5]. In [4], Kaynak *et al* demonstrate that the redesign of sliding surface can improve the performance of the overall control mechanism. The use of such techniques can therefore offer a practical alternative for stable training of intelligent systems. In [6-8], Efe and Kaynak demonstrate the distinguished performance introduced by using Variable Structure Systems (VSS) analogy in the training of neuro-fuzzy systems.

The organization of this paper is as follows. The second section describes the problem and the conventional solution. The third section demonstrates the modified training procedure. In the fourth section simulation examples are presented. Conclusions constitute the last part of the paper.

2. Estimation by Cosine Neural Networks

The approach presented in this section is based on the matching of two signals in time-domain. A natural consequence of this is the similarity in the frequency views. For this reason, an estimator composed of finite number of cosine components is constructed as described by (1).

$$x_s(t) = \sum_{i=1}^{neurons} C_i \cos(w_i t + p_i) \quad (1)$$

where, x_s is the response of the estimator. C_i, w_i and p_i denote the amplitude, frequency and the phase of the i^th^ unit respectively. Let x_d and e denote the desired signal and estimation error respectively. In order to minimize the instantaneous cost defined by (2), gradient descent based update formula given by (3) is employed.

$$J = \frac{1}{2}(x_d - x_s)^2 = \frac{1}{2}e^2 \quad (2)$$

$$\Delta\phi_{GD} = -\eta_{GD}\frac{\partial J}{\partial \phi} \quad (3)$$

In (3), ϕ is a generic parameter of the neuroestimator and η_{GD} is the learning rate from the interval (0,1). The subscript *GD* denotes the gradient descent. Having this in mind, the update formulas for i^{th} unit can be described in (4) through (6).

$$\Delta C_i = \eta e \cos(w_i t + p_i) \quad (4)$$

$$\Delta w_i = -\eta e C_i \, rem\left(t, \frac{2\pi}{w_i}\right)\sin(w_i t + p_i) \quad (5)$$

$$\Delta p_i = -\eta e C_i \sin(w_i t + p_i) \quad (6)$$

3. Variable Structure Systems Based Extraction of Stabilizing Information

The methods of computational intelligence frequently utilize the gradient based training procedures for parameter tuning. As mentioned earlier, the most crucial point in training of a neuro-fuzzy system is the fact that the training procedure tries to minimize the cost, which is a function of the realization error. However, during the training phase, there is no force ensuring the parametric stability or convergence. The approach analyzed here demonstrates how a VSS based stabiliznig information could be incorporated into the learning strategy adopted.

The parameter tuning formula given by (3) can be approximated by a first order system given by (7). In (7), T_s denotes the sampling period. If the parametric displacement $\Delta\phi$ is defined as the sliding line for parameter ϕ, adopting the reaching law in (8) and equating (7) and (8) yields the solution in (9).

$$\dot{\Delta\phi} = -\frac{1}{T_s}\Delta\phi + \frac{\eta e}{T_s}\frac{\partial x_s}{\partial \phi} \quad (7)$$

$$\dot{\Delta\phi} = -Q_\phi \tanh(\Delta\phi) - K_\phi \Delta\phi \quad (8)$$

$$\Delta\phi = \eta e \frac{\partial x_s}{\partial \phi} + Q_\phi \tanh(\Delta\phi) + K_\phi \Delta\phi \quad (9)$$

For the stability of the proposed solution, (10) is chosen as a Lyapunov function. Parametric stability is ensured if the inequality in (11) holds true.

$$V_\phi = \frac{1}{2}(\Delta\phi)^2 \quad (10)$$

$$\dot{V}_\phi = \left(\dot{\Delta\phi}\right)\Delta\phi < 0 \quad (11)$$

If (7) and (9) are substituted into (11), the following selection of learning rate satisfies the negative definiteness of the time derivative of the Lyapunov function in (10). This selection of η is given by (12).

$$\eta_\phi = \beta \min\left(\left|\frac{\Delta\phi}{N_\phi}\right|, \left|\frac{A_\phi}{N_\phi}\right|\right) \quad (12)$$

where,

$$N_\phi = e\frac{\partial x_s}{\partial \phi} \quad (13)$$

Define the following quantity;

$$A_\phi = Q_\phi \tanh(\Delta\phi) + K_\phi \Delta\phi \quad (14)$$

with β being a constant from the interval (0,1).

If this choice of η_ϕ is used in the approximate model of training dynamics (7), the stabilizing component of parameter update formula, which is given by (15), is obtained.

$$\Delta\phi_{VSS} = \beta \min(|\Delta\phi|, |A_\phi|)\operatorname{sgn}(N_\phi) + A_\phi \quad (15)$$

At this point, one might argue that whether this rule leads to the minimization of realization error or not. Clearly, the rule described above will enforce the adjustable parameters to settle down but will not minimize the cost in (2). An appropriate combination of this rule and gradient technique can result in the minimization of both the realization error and the displacement magnitude of the relevant parameter. This mixture can be performed by utilizing a weighted average as described in (16).

$$\Delta\phi = \frac{\alpha_1 \Delta\phi_{VSS} + \alpha_2 \Delta\phi_{GD}}{\alpha_1 + \alpha_2} \quad (16)$$

In (16), α_1 and α_2 are positive weights and determine the influence of each approach in the final value of the parameter change vector.

4. Simulation Results

Two simulation examples are presented. In the first example, a signal having finite frequency components in the spectra is reconstructed. The second example reveals the performance of the proposed scheme in extracting the frequency plot of a FIR filter.

4.1. Reconstruction of a Signal

In this part the signal described by (17) is reconstructed by the proposed structure and the training procedure. The desired and estimated signals are illustrated in Fig. 1.

$$x_d(t) = \sum_{j=1}^{4} \frac{\sin[(2j+1)t]}{2j+1} \quad (17)$$

In Fig. 2, the discrepancy between desired and estimated signals is depicted. For each period of the desired signal, a 628-point Fast Fourier Transform (FFT) is evaluated and corresponding to each period, results are given in Fig. 3 where the upper limit of the horizontal axis is adjusted such that the nonzero frequency components are easily seen. In this figure, the top row is the FFT of the desired signal (17). Subsequent plots in Fig. 3 illustrate the 628-point FFT of the estimated signal, which corresponds to 2π seconds in time-domain. Starting from the second period, the estimated spectral view and the desired one become nearly indistinguishable.

In (18), the cost of stability is formulated. In Fig. 4, the time behavior of this quantity is illustrated whereas Figs. 5 and 6 demonstrate the realization error and the cost of stability without using parameter stabilizing information derived in the third section.

$$J_s = \sum_{i=1}^{n} \Delta C_i^2 + \Delta w_i^2 + \Delta p_i^2 \quad (18)$$

During the simulations of signal reconstruction example, following parameters are used as the simulation settings.

Table 1. Simulation Settings for Signal Reconstruction

#Neurons	5
β	0.99
η_{GD}	0.01
Q	0.10
K	0.10
α_1	3.00
α_2	1.00
T_s	10 msec

4.2. Extraction of Frequency Plot of a FIR Filter

In this section, the average power spectral density (PSD) plot of a FIR filter is extracted by the use of the algorithm presented. The filter has the following transfer function.

$$H(z) = 1 + 0.999z^{-2} \quad (19)$$

In Fig. 7, the averaged power spectral density graph is illustrated for an intermediate step. As the input signal, a white noise sequence generated by Matlab is used. For this case following parameters are used as the simulation settings.

Table 1. Simulation Settings for Frequency Plot Extraction

#Neurons	3
β	0.01
η_{GD}	0.90
Q	0.10
K	0.10
α_1	1.00
α_2	1.00
T_s	10 msec

5. Conclusions

The method reported in this paper demonstrates that the conventional training procedures for computationally intelligent systems, such as neural networks, fuzzy systems or methods adapted from artificial intelligence can be incorporated with methodologies leading to parametric stability.

Variable Structure Systems technique or Sliding Mode Control philosophy is one of the methods which is well-known with its robustness to unmodeled internal dynamics of the system under investigation. A suitable combination of traditional training methods with VSS technique can offer much preferable solutions in the sense of safety. This is apparent from the comparison of estimation error trends illustrated in Figs. 2 and 5, and the cost of stability plots depicted in Figs. 4 and 6.

In this study, the performance of the proposed approach has been demonstrated on the estimation of signals and extraction of frequency plot of a FIR filter. The strategy is based on the matching of signals in time-domain. Naturally, if two signals have similar views in time-domain, their frequency plots are similar. The estimator studied in this study comprises cosine-like basis functions with adjustable parameters. The developed method tunes the parameters on-line.

Various identification and control oriented applications of the proposed approach clearly stipulate that the method is capable of eliminating the locally divergent behavior of gradient based training approaches. This study demonstrates that the method can also be used in signal processing applications.

6. References

[1] T. Söderström, Discrete-time Stochastic Systems, Prentice Hall, New Jersey, 1994.

[2] S. Haykin, *Neural Networks*, Macmillan College Printing Company, New Jersey, 1994.

[3] S. V. Emelyanov, *Variable Structure Control Systems*, Moscow, Nauka, 1967.

[4] O. Kaynak, F. Harashima and H. Hashimoto, "Variable Structure Systems Theory, as Applied to Sub-time Optimal Position Control with an Invariant Trajectory," *Trans. IEE of Japan*, Sec. E, v.104, no.3/4, pp.47-52, 1984.

[5] N. Bekiroglu, "Adaptive Sliding Surface Design for Sliding Mode Control Systems," Ph.D. Thesis, Bogazici University, 1996.

[6] M. O. Efe and O. Kaynak, " A Novel Optimization Procedure for Training of Fuzzy Inference Systems By Combining Variable Structure Systems Technique and Levenberg-Marquardt Algorithm," *Fuzzy Sets and Systems* (Submitted for publication).

[7] M. O. Efe and O. Kaynak, " Stable Training of Computationally Intelligent Systems By Using Variable Structure Systems Technique," *IEEE Transactions on Industrial Electronics* (Submitted for publication)

[8] M. O. Efe and O. Kaynak, " A Hybrid Training Procedure for Artificial Neural Networks Leading to Parametric Stability and Cost Minimization," The 7th Int. Conf. on Emerging Technologies and Factory Automation, ETFA'99, Spain (Submitted for publication)

6. Acknowledgments

This work is supported by the Bogazici University Research Fund (Grant No: 97A0202) and Foundation for Promotion of Advanced Automation Technology, FANUC.

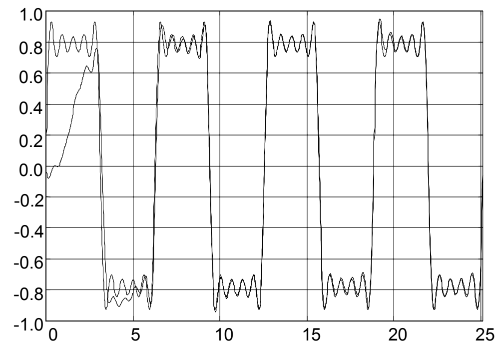

Figure 1. Desired and Estimated Signals

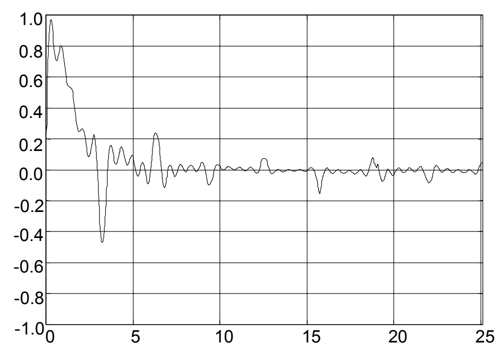

Figure 2. Estimation Error

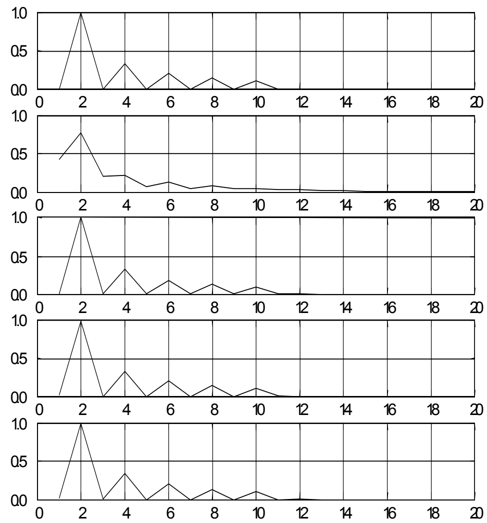

Figure 3. 628-Point FFT of the Desired Signal (top row) and Time Evolution of 628-Point FFT of Estimated Signal Corresponding to Each 2π Seconds

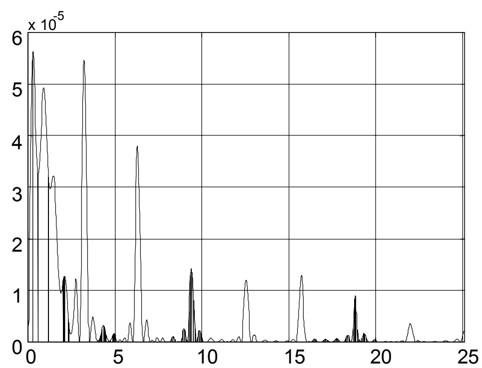

Figure 4. Cost of Stability with VSS and Gradient Descent Based Training Information Mixture

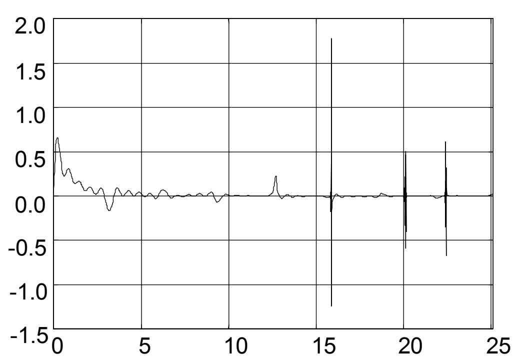

Figure 5. Estimation Error without VSS Based Component in Training Information

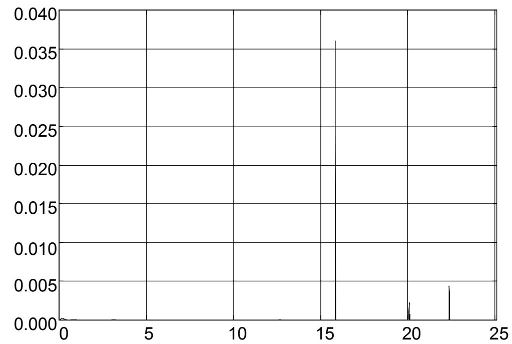

Figure 6. Cost of Stability without VSS Based Component in Training Information

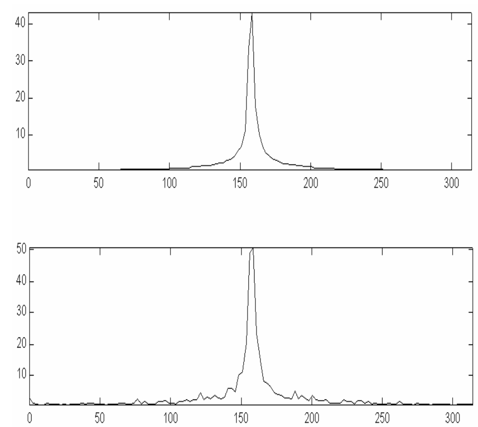

Figure 7. Desired Average Power Spectral Density and Estimated Average Power Spectral Density