FAULT DETECTION AND CLASSIFICATION OF AN ELECTROHYDROSTATIC ACTUATOR USING A NEURAL NETWORK TRAINED BY THE SMOOTH VARIABLE STRUCTURE FILTER

R. M. Ahmed, S. A. Gadsden, M. A. El Sayed, and S. R. Habibi
Department of Mechanical Engineering
McMaster University
Hamilton, ON, Canada
mohamerm@mcmaster.ca, gadsdesa@mcmaster.ca, abugabma@mcmaster.ca, and habibi@mcmaster.ca

ABSTRACT

A multilayered neural network is a multi-input, multi-output (MIMO) nonlinear system in which training can be regarded as a nonlinear parameter estimation problem by estimating the network weights. In this paper, the relatively new smooth variable structure filter (SVSF) is used for the training of a nonlinear multilayered feed forward network. The SVSF is a recursive sliding mode parameter and state estimator that has a predictor-corrector form. Using a switching gain, a corrective term is calculated to force the network weights to converge to within a neighbourhood of the optimal weight values. SVSF-based trained neural networks are used to classify faults on the input and output data of an electrohydrostatic actuator (EHA). Two faults are induced in the system: friction and leakage. Furthermore, a comparative study between the popular back propagation method, the extended Kalman filter (EKF), and the SVSF is presented.

INTRODUCTION

Back propagation (BP) is the most commonly used algorithm in the field of multilayer perceptron training [1]. It is a first-order stochastic gradient descent method that iteratively adjusts weights to minimize the output error in a supervised manner. However, since BP involves a constant learning rate, a slow speed of convergence is attained. In fact, several enhanced training algorithms have been developed through the literature to improve training performance, mapping accuracy and speed of convergence compared to the BP algorithm [2]. Most of these techniques like quasi-Newton and Levenburg-Marquardt demonstrate better performance as they involve second-order derivative information.

Since its inception in the 1960s, the Kalman filter (KF) remains the most popular state estimation tool. It provides statistically optimal estimations for linear systems in the presence of Gaussian white noise. In the case of nonlinear systems, the extended Kalman filter is applied by linearizing the system around the latest state estimate at each time interval. An EKF-based neural network training technique was first introduced by Singhal and Wu in 1989 [3]. The EKF provides a powerful neural network training capability compared to conventional first-order gradient algorithms, such as the BP [2]. In literature, the EKF has been extensively applied for training of both feed-forward [4] and recurrent networks [5,6] in both a global form (GEKF) or in a decoupled form (DEKF). Although the EKF demonstrates a close performance compared to a second-order derivative, batch-based method, it can avoid local minima problems by encoding second-order information in terms of a state error covariance matrix [2]. Accordingly, the EKF represents an efficient and practical alternative to second-order training methods.

In 2007, the smooth variable structure filter (SVSF) was proposed [7]. It provides a robust dynamic adaptation, high-rate of convergence, and can guarantee estimation stability for bounded uncertainties and noise levels [7]. The SVSF has been successfully applied for parameter and state estimation problems, showing robustness to modeling uncertainties [8,9]. In this paper, an overview of the combined SVSF and neural network methodology is provided. This methodology has only recently been developed and implemented for fault detection. The BP, EKF, and SVSF methodologies are applied on an electrohydrostatic actuator (EHA) for the purposes of fault detection. The paper then concludes with a summary of the results.

FEED-FORWARD MULTILAYERED NEURAL NETWORK

A multilayer feed forward network consists mainly of a set of sensory units (input source nodes) that constitutes the input layer, one or more hidden layers and an output layer. As shown in Fig. 1, each node is connected to all nodes in the adjacent layer by links (weights), and computes a weighted sum of the inputs. An offset (bias) is added to the resultant sum followed by a nonlinear activation function application. The input signal propagates through the network in a forward direction on a layer-by-layer basis. Consequently, the network represents a static mapping between inputs and outputs.

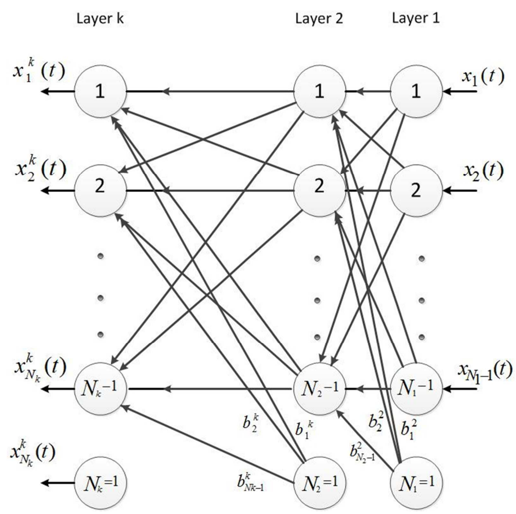

Fig. 1. Schematic diagram of a feed-forward multilayer perceptron network. [10]

Let k denote the total number of layers, including the input and output layers. $\mathrm{Node(n,i)}$ denotes the $\mathrm{i^{th}}$ node in the $\mathrm{n^{th}}$ layer, and $\mathrm{N_n}-1$ is the total number of nodes in the $\mathrm{n^{th}}$ layer. As shown in Fig. 2, the operation of $\mathrm{node(n+1,i)}$ is described by the following equation:

$$x_i^{n+1}(t) = \varphi\left(\sum_{j=1}^{N_n-1} w_{i,j}^n\, x_j^n(t) + b_i^{n+1}\right) \tag{1}$$

Where, $\mathrm{x_i^n(t)}$ denotes the output of $\mathrm{node(n,j)}$ for the t training pattern, $\mathrm{w_{i,j}^n}$ denotes the link weight from $\mathrm{node(n,j)}$ to the $\mathrm{node(n+1,i)}$. $\mathrm{b_i^n}$ is the node offset (bias) for $\mathrm{node(n,i)}$. The function $\varphi(.)$ is a nonlinear sigmoid activation function defined by:

$$\varphi(w) = \frac{1}{1+e^{-aw}} \qquad a > 0 \text{ and } -\infty < w < \infty \tag{2}$$

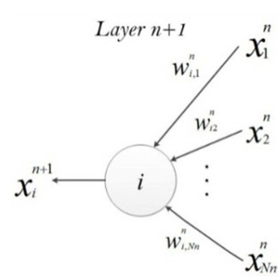

Fig. 2. Node (n+1, i) representation.

For simplicity, the node bias is considered as a link weight by setting the last input $\mathrm{N_n}$ to $\mathrm{node(n+1,i)}$ to value of one as follows:

$$\begin{aligned} x_{N_n}^n(t) &= 1,\;\; 1 \le n \le k \\ w_{i,N_n}^n &= b_i^{n+1},\;\; 1 \le n \le k-1 \end{aligned} \tag{3}$$

Consequently, (1) can be rewritten in the following form:

$$x_i^{n+1}(t) = \varphi\left(\sum_{j=1}^{N_n} w_{i,j}^n\, x_j^n(t)\right) \tag{4}$$

NEURAL NETWORK TRAINING USING THE SVSF

A. Introduction to the SVSF

In 2007, the smooth variable structure filter (SVSF) was introduced based on variable structure theory and sliding mode concepts [7]. It implements a switching gain to converge the estimates to within a boundary of the true states (i.e., existence subspace). In its present form, the SVSF has been shown to be stable and robust to modeling uncertainties and noise [11,12]. The SVSF method is model based and may be applied to differentiable linear or nonlinear dynamic equations. The original form of the SVSF as presented in [7] did not include covariance derivations. An augmented form of the SVSF was presented in [13], which includes a full derivation for the filter. The basic estimation concept of the SVSF is shown in the following figure.

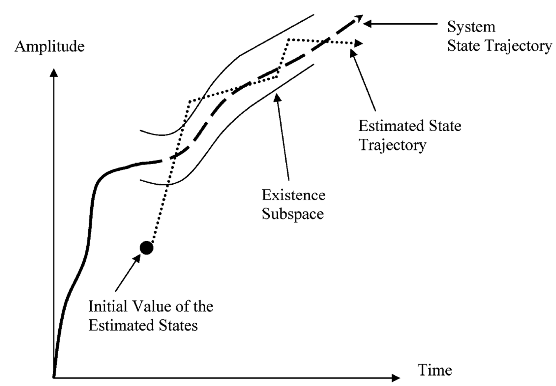

Fig. 3. SVSF estimation concept. [7]

Further to an initial condition, the estimated state is forced to within a region around the state trajectory referred to as the existence subspace. The width of the existence subspace β is a function of the uncertain dynamics associated with the inaccuracy of the internal model of the filter and varies with time [7]. Typically this value is not exactly known but an upper bound may be selected based on a priori knowledge. Once the estimate enters the existence subspace, it is forced into switching along the system state trajectory. A saturation term may be used in this region to reduce the magnitude of chattering and smooth out the result.

B. SVSF-Based Neural Network Training

The SVSF can be applied for training nonlinear feed-forward neural networks through estimation of the network weights. The SVSF has been adapted to train feed-forward neural networks by visualizing the network as a filtering problem. This process is iterative and may be summarized by the following set of equations. The predicted state (or neural network weight) estimates $\mathrm{\hat{w}_{k+1|k}}$ and state error covariances $\mathrm{P_{k+1|k}}$ are first calculated respectively as follows:

$$\widehat{w}_{k+1|k} = F\widehat{w}_{k|k} + Gu_k \quad (5)$$
$$P_{k+1|k} = HP_{k|k}H^T + Q_k \quad (6)$$

Utilizing the predicted state estimates $\widehat{w}_{k+1|k}$, the corresponding predicted measurements $\hat{z}_{k+1|k}$ and measurement errors $e_{z_{k+1|k}}$ may be calculated:

$$\hat{z}_{k+1|k} = H\widehat{w}_{k+1|k} \quad (7)$$
$$e_{z_{k+1|k}} = z_{k+1} - \hat{z}_{k+1|k} \quad (8)$$

The SVSF gain is a function of: the a priori and a posteriori measurement errors $e_{z_{k+1|k}}$ and $e_{z_{k|k}}$, the smoothing boundary layer widths ψ, the 'SVSF' memory or convergence rate γ, as well as the linear measurement matrix H. For the derivation of the SVSF gain K_{k+1}, refer to [7,13]. The SVSF gain is defined as a diagonal matrix such that:

$$K_{k+1} = H^{+}diag\left[\left(\left|e_{z_{k+1|k}}\right| + \gamma\left|e_{z_{k|k}}\right|\right) \circ sat\left(\frac{e_{z_{k+1|k}}}{\psi}\right)\right] diag\left(e_{z_{k+1|k}}\right)^{-1} \quad (9)$$

This gain is used to calculate the updated state estimates $\widehat{w}_{k+1|k+1}$ as well as the updated state error covariance matrix $P_{k+1|k+1}$:

$$\widehat{w}_{k+1|k+1} = \widehat{w}_{k+1|k} + K_{k+1}e_{z_{k+1|k}} \quad (10)$$
$$P_{k+1|k+1} = (I - K_{k+1}H)P_{k+1|k}(I - K_{k+1}H)^T + K_{k+1}R_{k+1}K_{k+1}^T \quad (11)$$

FAULT DETECTION PROBLEM

A. Electrohydrostatic Actuator Setup

An electrohydrostatic actuator (EHA) is an emerging type of actuator typically used in the aerospace industry. They are self-contained units comprised of their own pump, hydraulic circuit, and actuating cylinder [12]. The main components of an EHA include a variable speed motor, an external gear pump, an accumulator, inner circuitry check valves, a double-rod double-acting cylinder, and a bi-directional pressure relief mechanism. The schematic of the EHA circuitry is shown in the following figure.

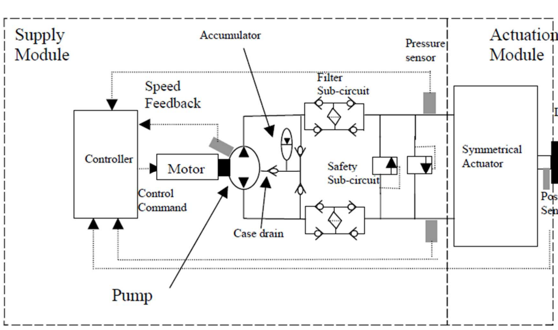

Fig. 4. A typical electrohydrostatic actuator (EHA) circuit diagram is shown above [12].

The EHA can be divided into two subsystems. The first is the inner circuit that includes the accumulator and its surrounding check valves. The second is the high pressure outer circuit which performs the actuation. The inner circuit prevents cavitation which occurs when the inlet pressure reaches near vacuum pressures and provides make-up fluid for any dynamic leakage [12]. This section is statically charged to 276 kPa (40 psi) which is enough pressure to avoid cavitation but it is also low enough to allow flow from the case drain back into the circuit. The inner circuit during normal operation is negligible in mathematical modeling. Mathematical modeling of the EHA has been performed and can be seen in detail in [12]. The two faults that were introduced to this system are increased friction and internal leakage. To incur these faults, one of the axes will be used as the driving mechanism while the other will act as a load. Mathematical models were derived for these two faults, as well as the normal case, based on the experimental setup.

B. Simulation Results

The models obtained from the experimental setup were used to derive simulated outputs for each mode, based on a pseudo-random binary signal (PRBS) acting as an input to the system. A cross-power spectral density (CPSD) was obtained for the three cases. This signal (containing 2,501 sample points) was used to train the network, with the following target signals: [1 0 0] represents a normal case, [0 1 0] represents friction, and [0 0 1] represents the leakage case. In total, 30 sets of signals were obtained for each operating mode. Of these signals, 20 were used for training the network weights, and 10 were used for testing the trained network.

The following three figures show the corresponding training and testing confusion matrices for each mode. The BP method performed the worst, as it was only able to correctly identify a fault 83.3% of the time (as shown in Fig. 5). It appeared to have difficulties distinguishing between the friction and leakage faults. Both the EKF and SVSF methods were able to train with 100% accuracy. However, during the testing phase, the EKF provided 93.3% accuracy, while the SVSF performed better with 96.7% accuracy. For further explanation on the confusion matrix, please refer to the Appendix.

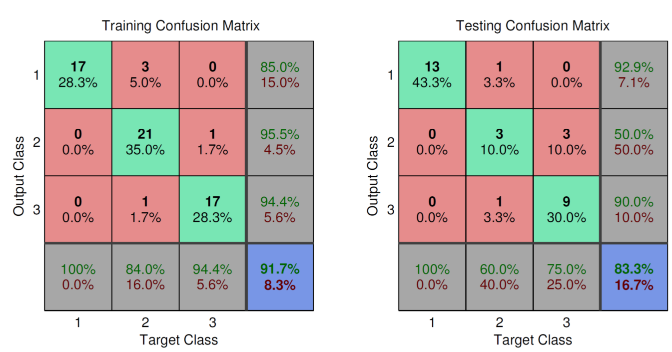

Fig. 5. BP Confusion Matrices

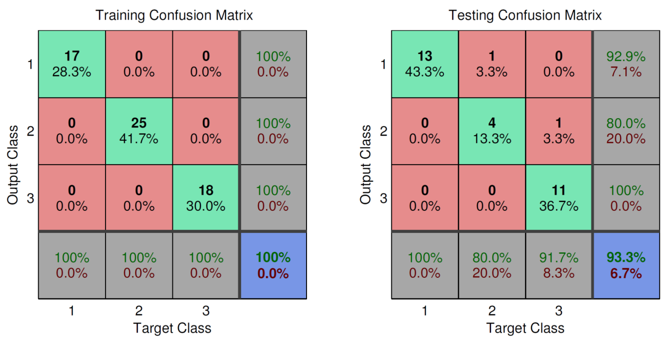

Fig. 6. EKF Confusion Matrices

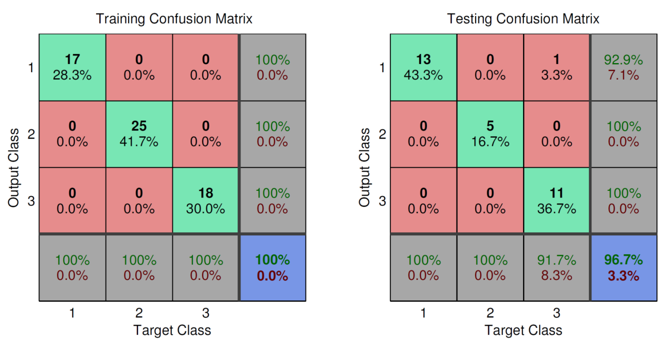

Fig. 7. SVSF Confusion Matrices

CONCLUSIONS

In this paper, the SVSF has been successfully applied to train feed-forward neural networks on an EHA for the purposes of fault detection and classification. The SVSF demonstrated guaranteed stability as well as excellent generalization capability, as illustrated by the confusion matrices. The training performance outperforms the BP algorithm, and was found to be comparable to the popular EKF strategy.

APPENDIX

The following figure helps to illustrate the concept of the general confusion matrix.

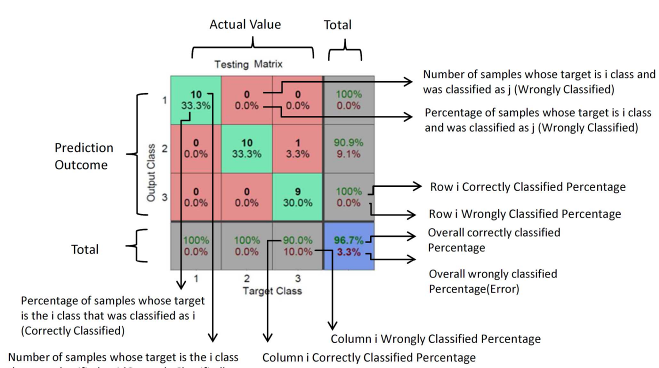

Fig. 8. Generalized Confusion Matrix

REFERENCES

[1] P. Werbos, "Backpropagation Through Time: What it Does and How to Do It," *Proceedings of the IEEE*, vol. 78, pp. 1550–1560, 1990.

[2] S. Haykin, *Kalman Filtering and Neural Networks*. New York, U.S.A.: John and Wiley Sons, Inc., 2001.

[3] S. Singhal and L. Wu, "Training Multilayer Perceptrons with the Extended Kalman Algorithm," *Advances in Neural Information Processing Systems 1*, pp. 133–140, 1989.

[4] G.V. Puskorius and L.A. Feldkamp, "Decoupled Extended Kalman Filter Training of Feedforward Layered Networks," in *International Joint Conference on Neural Networks*, Seattle, 1991, pp. 771-777.

[5] R. J. Williams, "Training Recurrent Networks Using the Extended Kalman Filter," *IJCNN International Joint Conference on Neural Networks*, vol. 4, pp. 241-246, 1992.

[6] P. Sun and M. Kenneth, "Training Recurrent Neural Networks for Very High Performance with the Extended Kalman Algorithm," *Intelligent Engineering Systems Through Artificial Neural Networks*, vol. 8, pp. 121-126, 1998.

[7] S. R. Habibi, "The Smooth Variable Structure Filter," *Proceedings of the IEEE*, vol. 95, no. 5, pp. 1026-1059, 2007.

[8] S. R. Habibi and R. Burton, "Parameter Identification for a High-Performance Hydrostatic Actuation System Using the Variable Structure Filter Concept," *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME*, vol. 129, no. 2, pp. 229-235.

[9] S. Wang, S. Habibi, and R. Burton, "A Smooth Variable Structure Filter for State Estimation," *Control and Intelligent Systems*, vol. 35, no. 4, pp. 386-393, 2007.

[10] Y. Iiguni, H. Sakai, and H. Tokumaru, "A Real-Time Learning Algorithm for a Multilayered Neural Network Based on the Extended Kalman Filter," *IEEE Transactions on Signal Processing*, pp. 959-966, April 1992.

[11] S. R. Habibi and R. Burton, "The Variable Structure Filter," *Journal of Dynamic Systems, Measurement, and Control (ASME)*, vol. 125, pp. 287-293, September 2003.

[12] S. R. Habibi and R. Burton, "Parameter Identification for a High Performance Hydrostatic Actuation System using the Variable Structure Filter Concept," *ASME Journal of Dynamic Systems, Measurement, and Control*, 2007.

[13] S. A. Gadsden and S. R. Habibi, "A New Form of the Smooth Variable Structure Filter with a Covariance Derivation," in *IEEE Conference on Decision and Control*, Atlanta, Georgia, 2010.