

Deep Searching for Parameter Estimation of the Linear Time Invariant (LTI) System by Using Quasi-ARX Neural Network

Mohammad Abu Jami'in, Imam Sutrisno, and Jinglu Hu

Abstract—This work exploits the idea on how to search parameter estimation and increase its convergence speed for the Linear Time Invariant (LTI) system. The convergence speed of parameter estimation is the one problem and plays an important role in the adaptive controller to increase performance. The well-known algorithm is the recursive least square algorithm. However, the speed of convergence is still low and is influenced by the number of sampling, which is represented by the limited availability for the information vector. We offer a new method to increase the convergence speed by applying Quasi-ARX model. Quasi-ARX model performs two steps identification process by presenting parameter estimation as a function over time. The first, parameters estimation of macro-part sub-model are searched by the least square error, and the second is to sharpen the searching by performing backpropagation learning of multi layer perceptron network.

I. INTRODUCTION

THE problem of identifying parameters of the linear system has been studied extensively [1]-[13]. Especially in the application such as controller design, some controller laws are created by utilizing parameters estimation as controller variable [14]-[21]. The accuracy in tracking parameter estimation, and the speed of convergence plays an important role to increase the response performance of the adaptive controller. The parameter estimations are also able to be applied to estimate state observers, and to design Kalman filtering [18]-[24].

The one problem to increase performance in the adaptive controller is the speed of convergence to search parameter estimation. The recursive least square algorithm is the well-known method to identify the parameter estimation. However, the speed of convergence is still low and is influenced by the number of sampling, which is represented by the limited availability for information vector. In this case, we offer a new method to increase the speed of convergence in the identification process by sharpening learning process to search true parameter estimation. We use two steps learning process. The first is the least square error algorithm to mapping the system input-output, then we have the parameter estimation, which will be sharpened by the next learning

Manuscript received February 26, 2013. This work was partly supported by the Indonesian Government Scholarship (Beasiswa Luar Negeri DIKTI - Kementerian Pendidikan dan Kebudayaan Republik Indonesia).

Mohammad Abu Jami'in and Imam Sutrisno are with the Graduate School of Information Production and Systems Waseda University, Kitakyushu 808-0135, Japan, and Politeknik Perkapalan Negeri Surabaya (PPNS), Jalan Teknik Kimia Kampus ITS Sukolilo Surabaya, 60111, Indonesia, (e-mail: mohammad@ruri.waseda.jp and imams3jpg@moegi.waseda.jp).

Jinglu Hu is with the Graduate School of Information Production and Systems Waseda University, Kitakyushu 808-0135, (e-mail: jinglu@waseda.jp).

process. The second learning is performed by using the core-part sub-model with the back propagation learning algorithms to update parameter estimation.

Quasi-ARX model decomposes the system into two sub-models, macro-part and core-part. The linear time invariant system is depicted by the macro-part. It represents the correlation between the input vector and its coefficients. The core-part represents the coefficient or parameter estimation. It parameterize the input space. The first learning of Quasi-ARX model is to searching parameter estimation by gradient learning of the least-squares error algorithm [25]-[32]. The second learning is to refining parameter estimation by core-part sub-model. The core-part has a specific structure, so it can be formed by using Multi-Layer Perceptron (MLP) neural network model, fuzzy, radial basis function or wavelet network. The structures of neural network in the core-part sub-model are able to be trained with different method and algorithm [25]. Generally, the improvement of Quasi-ARX model can be found by modification in learning or in structure of the core-part sub-model [25]-[29].

Quasi-ARX model is then applied to identifying linear time invariant system. To show the best performance in identification method, the performance of Quasi-ARX algorithm is compared with RLS algorithm. The index performances are measured to show the accuracy of the model, the accuracy of the estimated parameter, fast learning, and convergence speed of error. The proposed method needs a few times to estimate the parameter accurately. Through the theoretical analysis and simulation experiments, it offers best performance in both of accuracy and fast convergence time.

II. PROBLEM DESCRIPTION

The discrete model with single-input and single-output (SISO) of LTI system is shown in Fig. 1. The system, which

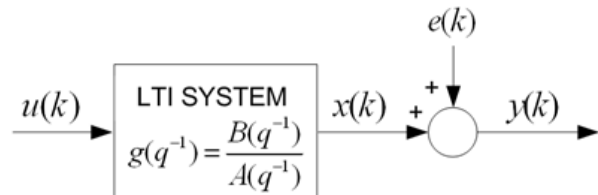


Fig. 1. The linear time invariant system is added with noise disturbance is depicted in Fig. 1, can be described into n -th order of difference equation with constant coefficients as,

$$x(k) + a_1x(k - 1) + \dots + a_{n_y}x(k - n_y) =$$

$$b_1 u(k-1) + \dots + b_{n_u} u(k-n_u) \quad (1)$$

$$y(k) = x(k) + e(k) \quad (2)$$

where $u(k) \in R$ is system input, $y(k) \in R$ is system output, $e(k) \in R$ is zero mean noise added to the system, and $k = 1, 2, \dots, N$ is the sampling sequences. The noise $e(t)$ is a stationary random process with zero mean, and it is uncorrelated to the system's input-output. The continuous time t is equal to kT in discrete time, where k is the number of sampling, and T is the constant time sampling. Suppose that the system's input-output are sampled at k -th, so the time at k -th is kT and the signal value at time kT can be written as $u(k)$ for the input, and $y(k)$ for the output.

We introduce the shifting operator such as $q^{-l}y(k)=y(k-l)$. By using this operator, The system also can be written as,

$$A(q^{-1})y(k) = B(q^{-1})x(k) + e(k) \quad (3)$$

$$A(q^{-1}) = 1 + a_1 q^{-1} + a_2 q^{-2} + \dots + a_{n_y} q^{-n_y}$$

$$B(q^{-1}) = b_1 q^{-1} + b_2 q^{-2} + \dots + b_{n_u} q^{-n_u}$$

$$\theta = [a_1, a_2, \dots, a_{n_y}, b_1, b_2, \dots, b_{n_u}].$$

we assumed that the degree of n_y and n_u are known. The initial condition at $k=0$ are known as $y(0)$, $u(0)$, and $e(0)$, and the θ is the parameter estimation of LTI system. In this case, the problem under study is to estimate parameter estimation of the system stated as θ . In matrix equation, if the initial condition at $k=0$, $y(0) = 0$, $u(0) = 0$, and $e(0) = 0$, then the output of the system can be described as,

$$y(k) = \phi(k)\theta^T + e(k) \quad (4)$$

where θ is the parameter estimation and $\phi(k)$ is the information vector. The information vector is the system input-output composed with time delay. Written in discrete is stated as,

$$\phi(k) = [y(k-1) \dots y(k-n_y) \ u(k-1) \dots u(k-n_u)]$$

III. QUASI-ARX MODEL FOR LTI SYSTEM IDENTIFICATION

Consider a function with single-input single-output (SISO) black-box model is described as,

$$y(t) = g(\phi(t)) + e(t) \quad (5)$$

where $y(t) \in R$, $e(t) \in R$, t is the continuous time. The $g(\cdot) \in R$ is the unknown function. The $\phi(t)$ is the information vector. The noise of the system $e(t)$ is the zero mean random signal added to the system. By using Taylor-expansion series, (5) can be rewritten into matrix equation as (6) [25]-[32],

$$y(t) = y_0 + \phi(t)\theta^T(\phi(t)) + e(t). \quad (6)$$

The initial condition is stated as y_0 . Based on (6), we propose to identification LTI system by two steps identification process. The first is to searching parameter estimation by LSE algorithm in the macro-part sub-model, and the second step is backpropagation algorithm in the core-part sub-model. The proposed identification technique is shown in Fig. 2.

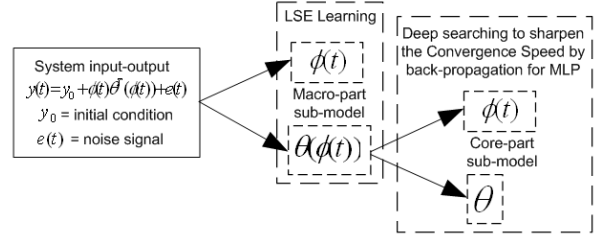


Fig. 2. Identification technique to sharpen the convergence speed by two steps identification process.

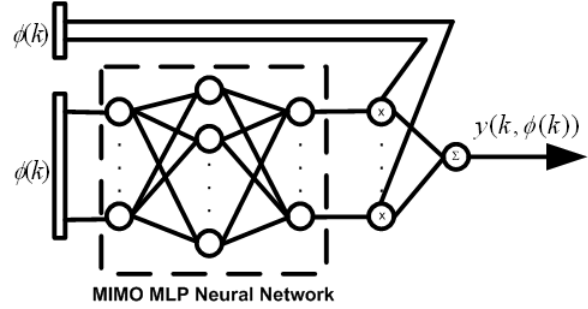


Fig. 3. Quasi-ARX neural network model.

See in (6), the information vector $\phi(t)$ is as variable of the parameter estimation $\theta(\phi(t))$. The output of the Quasi-ARX neural network model for $y_0=0$ in discrete time can be stated as,

$$y(k, \phi(k)) = \phi(k)\theta^T(\phi(k)). \quad (7)$$

where $\theta(\phi(k)) = [a_{(1,k)} \dots a_{(n_y,k)} \ b_{(1,k)} \dots b_{(n_u,k)}]$. It is clearly that $\theta(\phi(k))$ is a function over time. The macro-part sub-model is expressed by,

$$y(k, \phi(k)) = \phi(k)\mathcal{U}^T(k). \quad (8)$$

The information vector dimension of $\phi(k)$ is equal to $n_u + n_y$. The number of hidden node is m . The number of output node in the core-part sub-model is the same as the information vector dimensions n . The core-part sub-model is performed by the MLP neural network model expressed as,

$$\Delta\theta(\phi(k)) = W_2\Gamma(W_1(\phi(k)) + B). \quad (9)$$

The update of the parameter estimation is performed by,

$$\theta(\phi(k)) = \Delta\theta(\phi(k)) + \mathcal{U}(k) \quad (10)$$

where, $\Delta\theta = \{W_1, W_2, B\}$, $W_1 \in R^{n \times m}$, $W_2 \in R^{m \times n}$, $B \in R^n$ is the weight matrix at the first, the second layer, and the bias vector of input nodes. $\mathcal{U} \in R^n$ is the bias vector of the output nodes in the macro-part sub-model, and Γ is the diagonal nonlinear operator with the identical sigmoidal elements on hidden nodes. The best solution of the Quasi-ARX model incorporating neural network is shown on Fig. 3.

The prediction output from the Quasi-ARX neural network model can be determined if the problem in (7) satisfies to

mapping the input-output of the system. The optimal solution of parameter estimation can be applied to prediction for the next output stated as,

$$y_p(k+d|k, \phi(k)) = \phi(k+d)\hat{\theta}^T(\phi(k)) \quad (11)$$

where, $y_p(k+d|k, \phi(k))$ is the output of d ahead prediction, $\phi(k+d) = [y(k+d-1) y(k+d-2) \cdots y(k+d-n_y) u(k+d-1) u(k+d-2) \cdots u(k+d-n_u)]$ is the information vector for d ahead prediction, and $\hat{\theta}(\phi(k)) = [\hat{a}_{(1,k)} \cdots \hat{a}_{(n_y,k)} \hat{b}_{(1,k)} \cdots \hat{b}_{(n_u,k)}]$ is the estimated parameter estimation, and d is the time delay operator.

IV. LEARNING ALGORITHM FOR QUASI-ARX NEURAL NETWORK

The learning algorithm for Quasi-ARX model is performed by the backpropagation error algorithm for the core-part sub-model, and LSE algorithm for the macro-part sub-model. Let we introduce two sub-models $z_l(k) = y(k, \phi(k)) - \phi(k)\Delta\theta^T(\phi(k))$, and $z_n(k) = y(k, \phi(k)) - \phi(k)\mathcal{U}^T$ to become learning guide. The learning output guidance to train sub-models are expressed as,

$$SM1 \quad z_l(k) = \phi(k)\mathcal{U}^T(k). \quad (12)$$

$$SM2 \quad z_n(k) = \phi(k)\Delta\theta^T(\phi(k)). \quad (13)$$

The step of learning algorithm of Quasi-ARX neural network is described by,

- 1) set $\mathcal{U} = 0$; and small initial values to W_1 , W_2 , and B , set $i = 1$, where i is the learning number.
- 2) calculate $z_l(k)$, then estimate \mathcal{U} for sub-model $SM1$ by using a least-squares error algorithm.
- 3) calculate $z_n(k)$, then estimate W_1 , W_2 , and B for sub-model $SM2$ This is realized by using the well-known backpropagation (BP) algorithm.
- 4) use the (10) to update $\theta(\phi(k))$
- 5) stop if pre-specified conditions are met, otherwise go to Step 2, and repeat the estimation of \mathcal{U} , and W_1 , W_2 , and B , set $i = i + 1$.

V. RECURSIVE LEAST-SQUARES ALGORITHM

Recursive least square algorithm is very well known algorithm applied to estimate the parameter estimation of the LTI system by online. Many researchers have made improvement, modify and development of RLS algorithm to increase the accuracy of the estimated parameter estimation [3]-[13]. The RLS estimation techniques are the fundamental technique in adaptive signal processing applications. The equation (4) is the problem that will be solved by Quasi-ARX neural network, then the results are compared with RLS algorithm.

The Algorithm of RLS is performed by minimizing cost function stated as,

$$J(\theta) = \sum_{k=1}^{k_n} (y(k) - \phi(k)\theta^T(k)) \quad (14)$$

where k_n is the last sampling number, and $\phi(k)\phi^T(k)$ is nonsingular for all k . The estimation of $\hat{\theta}(k)$ recursively are described as,

$$\hat{\theta}(k) = \hat{\theta}(k-1) - K(k)(y(k) - \phi(k)\hat{\theta}^T(k-1)). \quad (15)$$

$$K(k) = \frac{P(k-1)\phi^T(k)}{1 + \phi(k)P(k-1)\phi^T(k)}. \quad (16)$$

$$P(k) = [1 - K(k)\phi^T(k)]P(k-1). \quad (17)$$

VI. EXPERIMENTAL STUDIES

The Quasi-ARX neural network model is applied to identification Linear Time Invariant (LTI) system. The Pseudo Random Binary Sequence Signal (PRBS) is as input. The system is added with zero mean noise. The proposed algorithm is tested to measure performance identification in switched mode power converters (SMPC) in [4]. The SMPC discrete transfer function is stated as,

$$G(q^{-1}) = \frac{0.226q^{-1} + 0.1118q^{-2}}{1 - 1.914q^{-1} + 0.949q^{-2}}. \quad (18)$$

The output of the system has a ripple caused by signal perturbation is approximately 10% or source to noise ratio (SNR) 20 dB of the system output. The SNR gaussian noise is expressed as,

$$SNR = 10 \log \left(\frac{\sum_{k=1}^N x(k)^2}{\sum_{k=1}^N e(k)^2} \right) dB. \quad (19)$$

The core-part sub-model is performed by MLP neural network. The number of input node n is the sum of $n_u=2$ and $n_y=2$, and the information vector is as the input, the number of hidden node and output node is the same as input. The result of system identification which represent the accuracy of input-output of the system by Quasi-ARX neural network is shown in Fig. 4 and Fig. 6. The accuracy of input-output of the system by the RLS algorithm is shown in Fig. 5 and Fig. 6. The MSE error over time is shown in Fig. 6, the red line represent the MSE error by RLS algorithm and blue line by Quasi-ARX model.

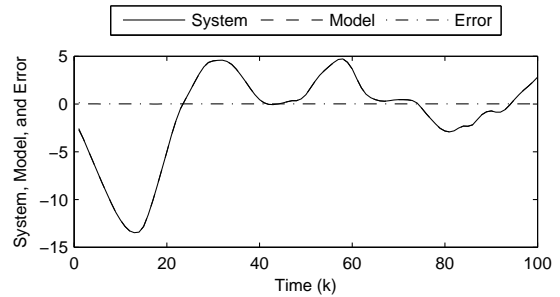


Fig. 4. System Output, Model Output and Error of the Quasi-ARX for step ahead prediction

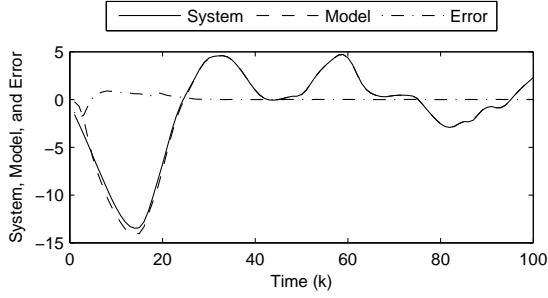


Fig. 5. System Output, Model Output and Error the RLS Algorithm for step ahead prediction

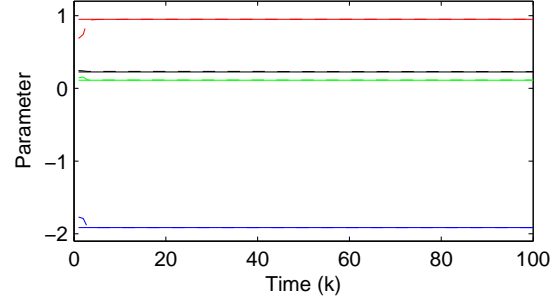


Fig. 7. The estimated parameter by using Quasi-ARX neural network

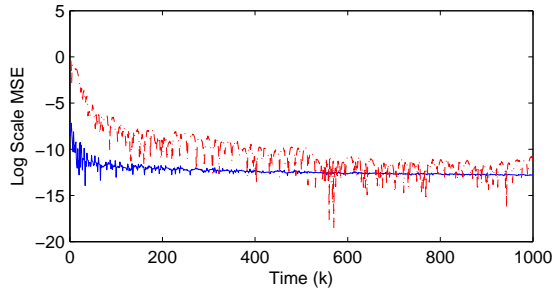


Fig. 6. The MSE Index Performance for step ahead prediction

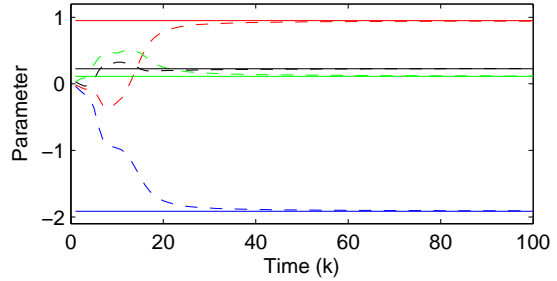


Fig. 8. The estimated parameter by using RLS algorithm

The performances of system identification are measured by Mean Squares Error (MSE) Index to show the accuracy of system input-output.

$$MSE = \frac{\sum_{k=1}^N (y_p(k) - y(k))^2}{N}. \quad (20)$$

The accuracies of the estimated parameter are measured by Root Mean squares of Parameter (RMSP) stated as,

$$RMSP = \sqrt{\frac{\sum_{k=1}^N \sum_{r=1}^{n_u+n_y} (\theta(r) - \hat{\theta}(r))^2}{N(n_u + n_y)}}. \quad (21)$$

The parameter estimation and its accuracy are shown in Fig. 7 and Fig. 9 by using Quasi-ARX neural network, and are shown in Fig. 8 and Fig. 9 by using RLS algorithm. The solid line is the true value, and the dashed line is the estimated value. Blue line, red line, black line and green line are the parameter estimation of a_1 , a_2 , b_1 and b_2 . The RMSP error is shown in Fig. 9, the red line represent the RMSP error by RLS algorithm and blue line by Quasi-ARX model.

The performance of estimated parameter (EMP) by Quasi-ARX is shown in Table I, and the EMP by RLS algorithm is shown in Table II. The average of RMSP in one hundred sampling is 0.00247 by Quasi-ARX, and 0.329 by RLS algorithm.

VII. RESULT AND DISCUSSION

In this system identification, we introduce to apply Quasi-ARX model to estimate a parameter for the LTI system. Therefore, we pay much attention in the accuracy of the

parameter estimation instead of the accuracy of the system input-output. The fast convergence to find the estimated parameter are to be sharpened by the core-part sub-model of Quasi-ARX neural network. We can see the accuracy of the model, which is shown in Fig. 4, Fig. 5 and Fig. 6. It indicates that the Quasi-ARX neural network model is more accurate model compared to the RLS model algorithms.

The performances in searching of the estimated parameter estimation are shown in Fig. 7, Fig. 8, Fig. 9, Table I, and Table II. We can see that the Quasi-ARX neural network is the more accurate model to estimate the parameter estimation of the LTI system. The average of RMS error of the parameter estimation (RMSP) in one hundred sampling are 0.00247 identified by the Quasi-ARX model, and is 0.329 identified by RLS algorithm. The convergence speed by using Quasi-ARX model is also fast. See Fig. 9, the RMSP error in log scale is the -2.134 or 0.74 percent can be achieved at first sampling identified by Quasi-ARX model. The same performance is reached by RLS algorithm in about forty three sampling. We can conclude that the tracking parameters' estimations are very slow by using RLS algorithm.

REFERENCES

- [1] J.Hu, K.Hirasawa and K.Kumamaru, "A Homotopy Approach to Improving PEM Identification of ARMAX Model," *Automatica*, Vol.37, No.9, pp.1323-1334, 2001.
- [2] J.Hu and K.Hirasawa, "A Homotopy Approach to Identification of ARMAX Systems," *Trans. of the Institute of Electrical Engineers of Japan*, Vol.119-C, No.2, pp.206-211, 1999.
- [3] F. Ding, P.X. Liu, G. Liu, "Multiinnovation Least-Squares Identification for System Modeling," *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics*, vol. 40, no. 3, pp. 767-778, June 2010.

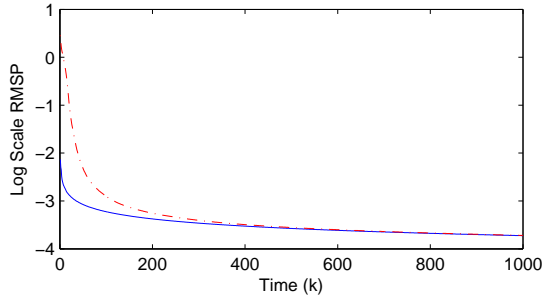


Fig. 9. The RMSP Index Performance for step ahead prediction

TABLE I

ESTIMATED PARAMETER ARE TRACKED BY USING QUASI-ARX NEURAL NETWORK MODEL.

	a_1	a_2	b_1	b_2
True	-1.914	0.949	0.226	0.1118
Mean	-1.912	0.946	0.231	0.1141
MSE	3.57E-06	2.07E-05	9.61E-09	2.19E-07
RMS	1.89E-03	4.55E-03	9.80E-05	4.68E-04

TABLE II

ESTIMATED PARAMETER ARE TRACKED BY USING RLS ALGORITHM.

	a_1	a_2	b_1	b_2
True	-1.914	0.949	0.226	0.1118
Mean	-1.697	0.732	0.208	0.1787
MSE	0.0475	0.0380	2.99E-05	0.0002
RMS	0.2179	0.1950	0.005468	0.0143

[4] M. Algreer, M. Armstrong, and D. Giaouris, "Active Online System Identification of Switch Mode DC-DC Power Converter Based on Efficient Recursive DCD-IIR Adaptive Filter," *IEEE Transactions on Power Electronics*, vol. 27, no. 11, pp. 4425-4435, Nov. 2012.

[5] J. Ding, F. Ding, X.P. Liu, and G. Liu, "Hierarchical least-squares Identification for Linear SISO Systems With Dual-Rate Sampled-Data," *IEEE Transactions on Automatic Control*, Vol.56, No.11, pp.2677-2683, November 2011.

[6] S. Bhama and H. Singh, "Single Layer Neural Networks for Linear System Identification Using Gradient Descent Technique," *IEEE Transactions on Neural Networks*, Vol.4, No.5, pp.884-888, September 1993.

[7] M. Schoukens, R. Pintelon, and Y. Rolain, "Parametric Identification of Parallel Hammerstein Systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 12, pp. 3931-3938, 2011.

[8] S.A. Fattah, W.-P.Zhu, and M.O. Ahmad, "Identification of autoregressive moving average systems based on noise compensation in the correlation domain," *IET Signal Process.*, vol. 5, iss. 3, pp. 292-305, 2011.

[9] D.Q. Wang and F. Ding, "Hierarchical least squares Estimation Algorithm for HammersteinWiener Systems," *IEEE Signal Processing Letters*, Vol.19, No.12, pp.825-828, Dec 2012.

[10] H.F. Chen, "New Approach to Recursive Identification for ARMAX Systems," *IEEE Transactions on Automatic Control.*, vol. 55, no. 4, pp. 868-879, 2010.

[11] D.Q. Wang, "least squares based recursive and iterative estimation for output error moving average systems using data filtering," *IET Control Theory Appl.*, vol. 5, iss. 11, pp. 1648-1657, 2011.

[12] A. Jamoos, E. Grivel, W. Bobillet, and E Guidorzi, "Errors-In-Variables-Based Approach for the Identification of AR Time-Varying fading Channels," *IEEE Signal Processing Letters*, vol. 14, no. 11, pp. 793-796, November 2007.

[13] L. Vanbeylen, R. Pintelon, J. Schoukens, "Blind Maximum-Likelihood Identification of Wiener Systems," *IEEE Transactions on Signal Processing*, vol. 57, no. 8, pp. 3017-3029, August 2009.

[14] S. Kelouwani, K. Adegnon, K. Ogbossou, and Y. Dube, "Online System Identification and Adaptive Control for PEM Fuel Cell Maximum Efficiency Tracking," *IEEE Transactions on Energy Conversion*, vol. 27, no. 3, pp. 580-592, 2012.

[15] F. Sun, Z. Sun, and P.Y. Woo, "Neural Network-Based Adaptive Controller Design of Robotic Manipulators with an Observer," *IEEE Transactions on Neural Networks*, Vol. 12, No. 1, pp. 54-67, January 2001.

[16] C. Wang and Y. Lin, "Multivariable Adaptive Backstepping Control: A Norm Estimation approach," *IEEE Transactions on Automatic Control*, vol. 57, no. 4, pp. 989-995, April 2012.

[17] Y.-B. Zhao, G.P. Liu, and D. Rees, "Improved predictive control approach to networked control systems," *IET Control Theory Appl.*, vol. 2, iss. 8, pp. 675-681, 2008.

[18] V. M. Becerra, F. R. Garces, S.J. Nasuto, and W. Holderbaum, "An Efficient Parameterization of Dynamic Neural Networks for Nonlinear System Identification," *IEEE Transactions on Neural Networks*, Vol.16, No.4, pp.983-988, July 2005.

[19] J.A. Riveros, A.G. Yepes, and F. Baarrero, "Parameter Identification of Multiphase Induction Machines With Distributed Windings-Part 2: Time-Domain Techniques," *IEEE Transactions on Energy Conversion*, vol. 27, no. 4, pp. 1067-1077, 2012.

[20] TC.L. Hwang, and C. Jan, "State-Estimator-Based Feedback Control for a Class of Piezoelectric Systems with Hysteretic Nonlinearity," *IEEE Transactions on System, Man, and Cybernetics-Part A: System and Humans*, vol. 35, no. 5, pp. 654-664, September 2005.

[21] M. Stocks and A. Medvedev, "Linear-quadratic optimal time-varying observers with time-varying guaranteed convergence rate," *IET Control Theory Appl.*, vol. 3, iss. 2, pp. 181-188, 2009.

[22] L.Y. Wang, C. Li, G.G. Yin, L. Guo, and C.Z. Xu, "State Observability and Observers of Linear-Time-Invariant Systems Under Irregular Sampling and Sensor Limitations," *IEEE Transactions on Automatic Control*, vol. 56, no. 11, pp. 2639-2654, November 2011.

[23] M. Pachter, "Kalman Filtering When the Large Bandwidth Control is not Known," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 1, pp. 542-551, January 2012.

[24] S.V. Vaerenbergh, M.L. Gredilla, Member, and I. Santamara, "Kernel Recursive Least-Squares Tracker for Time-Varying Regression," *IEEE Transactions on Neural Networks and Learning Systems*, Vol.23, No.8, pp.1313-1326, August 2012.

[25] M.A. Jami'in, I. Sutrisno and J. Hu, "Lyapunov Learning Algorithm for Quasi-ARX Neural Network to Identification of Nonlinear Dynamical System," *IEEE International Conference on Systems, Man, and Cybernetics (IEEE SMC'12)*, Seoul, Oct., 2012, pp.3141-3146.

[26] Y. Cheng, L. Wang and J. Hu, "Quasi-ARX Wavelet Network for SVR Based Nonlinear System Identification," *Nonlinear Theory and its Applications (NOLTA), IEICE*, Vol.2, No.2, pp.165-179, April 2011.

[27] Y. Cheng and J. Hu, "Nonlinear System Identification Based on SVR with Quasi-linear Kernel," *2012 IEEE International Joint Conference on Neural Networks*, Brisbane, June, 2012, pp.2368-2375.

[28] J.Hu, K.Kumamaru and K.Hirasawa, "A Quasi-ARMAX Approach to the Modeling of Nonlinear Systems," *International Journal of Control*, Vol.74, No.18, pp.1754-1766, 2001.

[29] J.Hu and K.Hirasawa, "An Embedded Sigmoidal Neural Network for Modeling of Nonlinear Systems," *2012 IEEE International Joint Conference on Neural Network*, Washington, vol.3, 7, 2001, pp.1698-1703.

[30] J.Hu and K.Hirasawa, "A Method for Applying Multilayer Perceptrons to Control of Nonlinear Systems," *9th International Conference on Neural Information Processing (ICONIP'02)*, Singapore, 11, 2002, pp.1267-1271.

[31] J.Hu, K.Hirasawa and K.Kumamaru, "A Quasi-ARX Model Incorporating Neural Networks for Control of Nonlinear Systems," *Proc. of the 15th IFAC World Congress*, Barcelona, 7, 2002.

[32] L. Wang, Y. Cheng and J. Hu, "A Quasi-ARX Neural Network with Switching Mechanism to Adaptive Control of Nonlinear Systems," *SICE Journal of Control, Measurement, and System Integration*, Vol.3, No.4, pp.246-252, July 2010.