NONLINEAR MIMO ADAPTIVE PREDICTIVE CONTROL BASED ON WAVELET NETWORK MODEL[1]

Dexian Huang, Yuhong Wang and Yihui Jin

Department Automation , Tsinghua University, Beijing, 100084, China

ABSTRACT: A MIMO nonlinear adaptive predictive control strategy is presented in which the wavelet neural network based on a set of orthogonal wavelet functions is adopted. A nonlinear mapping from the network-input space to the wavelons output space in the hidden layer is performed firstly. Then, the output layer uses a linear structure. Its weight coefficients can be estimated by a linear least-squares estimation method. The excellent statistic properties of the weight parameter estimation can be obtained. Based on developed recursive algorithm, a MIMO nonlinear adaptive predictive control strategy is implemented. A simulated MIMO nonlinear process example shows that the control scheme is effective.

Keywords: Non-linear MIMO adaptive control, predictive control, wavelet, neural network.

1. INTRODUCTION[1]

During the past twenty years, model-predictive control algorithms (MPC), based on linear process models, have been widely studied and applied in the chemical process industries. However, many processes are highly non-linear, uncertain and MPC algorithms based on linear process models may result in poor control performance and as a result, MPC techniques have recently been extended to these processes during the last decade(Keerthi,1990, Proll, 1994). However, generic nonlinear models is difficult to get and apply. Neural networks hold the promise of solving the problems. Feed forward neural networks provide a connectionist model that performs a mapping from an input space to an output space. Such networks can approximate any non-linear functions to an arbitrary accuracy. However, some network training problems, such as undesirable local minimum, of multi-layer perceptrons preclude their wide applications to on-line nonlinear system identification in adaptive control.

Wavelet is a powerful tool for function approximation (Daubechies, 1988). Under some mild conditions, the universal approximation of wavelet networks is guaranteed (Zhang ,1995). Based on a set of orthogonal wavelet functions, a least-squares learning algorithm is adopted to train the wavelet network in contrast to the non-linear gradient optimization used in standard feed forward networks (Bakshi and Stephanopoulos, 1993). In addition, wavelet neural networks have advantages in their structure, which is easy to specify in model identification. SISO first-order and high order non-linear dynamic processes have been successfully identified by wavelet neural networks and SISO non-linear predictive controllers have been realized. The control performances superior to a standard PID controller were achieved (Huang, 1997, 1999).

In this paper, a nonlinear adaptive predictive control strategy based orthogonal wavelet network model is presented. Based on a set of orthogonal wavelet functions, wavelet neural network performs a nonlinear mapping from the network-input space to the wavelons output space in hidden layer firstly. Since almost all dynamic processes in the chemical industries are lowpass systems, they can be approximated only by scale function terms $\sum_n \langle f, \varphi_{M,n} \rangle \varphi_{M,n}(x)$ at any accuracy. Therefore, we only use scale function in wavelons. This will simplifies wavelet network and decreases network size in online training obviously. Then, the output layer adopts linear structure. Its weight coefficients, i.e. $\langle f, \varphi_{M,n} \rangle$, can be estimated by a linear least-squares estimation algorithm.

Because the solid theory basis and special structure of wavelet neural network, the wavelet neural network holds the advantages superior to other neural network. First, its network structure is easy to specify based on its theory analysis and intuition. Secondly, network training do not rely on stochastic gradient type techniques such as the "back propagation" and can avoids the problem of poor convergence or undesirable local minimum, which is more serious for other neural networks when training data is contaminated seriously by noise.

The excellent statistic properties of the weight parameters of wavelet network as linear least-squares estimation algorithm in system identification can be proved. In intuition, it can been seen that the wavelet network is a ideal lowpass filter which passes true

[1] The project is supported by National High Tech R&D Project(NO. 863/CIMSAA413130) of China and by National key technologies R&D program of China(NO. 2001 BA201A04)

dynamic signal of the system identified and sorts the noise out as excellent frequency property of wavelet. The theory results are showed by simulation results. Both the wavelet network and the usual feedforward neural network are compared in a simulated CSTR system with serious noise. The long-range predictions based on trained wavelet network for testing data have obviously better prediction accuracy than that of he usual feedforward neural network. The prediction is very close to true output without noise. Both theory analysis and simulation study show that the identification method based on wavelet network is a robust and reliable identification method for nonlinear system. In addition, it is also generic method and is easy to use, instead of a method based on trial and error.

For online application in adaptive predictive control strategy, a recursive algorithm is given. The properties similar to that of recursive linear least-squares algorithm can be obtained as the recursive algorithm is completely same as recursive linear least squares algorithm. In addition, the closed loop-identifiability can be guaranteed. This is because the different wavelon outputs in hidden layer are irrelevant each other as orthogonal wavelet functions are adopted.

With developed recursive algorithm, a single input – single output nonlinear adaptive predictive control strategy is implemented. A same simulated CSTR process as above illustrates the application of the control scheme. Two methods to start adaptive controller are realized. Simulation results show two methods have good control results and expected performances are attained. When the parameter of controlled system is changed, online identification algorithm can track the parameter changing rapidly and then, still give good control results. The nonlinear adaptive predictive control strategy based on wavelet network is superior to the standard PID controller.

2. APPROXIMATION PRINCIPLE OF WAVELET NEURAL NETWORKS

According to approximate theory of wavelet network(Huang, 1997), two schemes for decomposing the function f(x) in $L^2(R)$ can be obtained. They are:

$$f(x) = \sum_{m,n} \left\langle f, \psi_{m,n} \right\rangle \psi_{m,n}(x) \quad (1)$$

and

$$f(x) = \sum_{n} \left\langle f, \varphi_{m_0,n} \right\rangle \varphi_{m_0,n}(x) + \sum_{m \geq m_0, n} \left\langle f, \psi_{m,n} \right\rangle \psi_{m,n}(x) \quad (2)$$

where m_0 is an arbitrary integer and represents the lowest resolution, i.e. scale, in the decomposition. Comparing equation (2) with equation (1), the former is more useful in dynamic process modeling. This is because most dynamic processes in the process industries are low-pass systems and, therefore, using scaling function can obviously decrease wavelet function terms. Furthermore, it is noted that f(x) can be closed arbitrarily only in V_M for some integer M. As long as the wavelet basis satisfies the Frame, there exists an M sufficiently large for any $\varepsilon > 0$ (Zhang et al., 1995), such that

$$\left\| f(x) - \sum_{n} \left\langle f, \varphi_{M,n} \right\rangle \varphi_{M,n}(x) \right\| < \varepsilon \quad (3)$$

Therefore, it is realistic that a dynamic process can be approximated only by the scale function terms $\sum_{n} \langle f, \varphi_{M,n} \rangle \varphi_{M,n}(x)$ in permitted approximating accuracy. This will decrease approximating function terms and therefore decrease network size. In addition, this will also simplify wavelet network application.

The structure of a wavelet neural network is similar to that of an RBF network. However its structure can be decided by using wavelet frames. Because only scale function is used, then

$$\begin{aligned} f(x) &= \sum_{n} \langle f, \varphi_{M,n} \rangle \varphi_{M,n}(x) = \sum_{n} \langle f, \varphi_{M,n} \rangle 2^{M/2} \varphi(2^M x - n) \\ &= \sum_{n} \theta_n \varphi(Ra(x - b_n)) \end{aligned} \quad (4)$$

For multi-input systems $\varphi_{M,n}(X) = \varphi_{M,n_1}(x_1) \varphi_{M,n_2}(x_2) \cdots \varphi_{M,n_r}(x_r)$.

When the variation domains of the network inputs are defined, the neuron centers, i.e. b_n, are fixed on grids that are divided equally between each input domain. R_a is an adjustable parameter that changes the width of the frequency band of V_M .For the wavelet network studied in this paper, we use Shanonn wavelet function because it is an analytic function and is easy to use. Its scale function form is: $\varphi(x) = \frac{\sin \pi x}{\pi x}$. It is an orthonormal wavelet function. Other wavelet functions including non-orthonormal wavelet function can also be used.

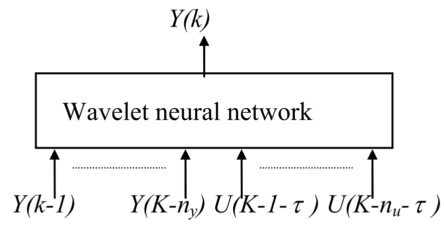

Figure 2. Wavelet network structure for dynamic system modeling

Consider a MIMO non-linear dynamic system denoted by the following equation:

$$\begin{aligned} Y(k) = f(&Y(k-1), \ldots\ldots, Y(k-n_y), U(k-1-\tau), \\ &\ldots\ldots, U(k-n_u-\tau)) \end{aligned} \quad (5)$$

where τ is the model input-output time delay, $(Y,U) \rightarrow f(Y,U)$: $R^l \times R^r \rightarrow R^l$. The network structure proposed by Narendra is adopted (Narendra and Parthasarathy, 1990). It is shown in Figure 2.

The network input and output dimensions are $\sum_{i=1}^{l} n_{yi} + \sum_{i=1}^{m} n_{ui}$ and l respectively.

The network weights are identified by linear least squares algorithm as following.

Firstly, we denote the optimum values of all $\langle f, \varphi_{M,n} \rangle$ as θ and all $\varphi_{M,n}(x)$ values at time k as $h(k)$. Then

$$Y(k) = h^T(k)\theta + n(k) \tag{6}$$

where both vector Y(k) and n(k) have dimension $1 \times l$, the dimension of h(k) is $N \times 1$ and the dimension of θ is $N \times l$. N is the number of neurons.

For $k=1,2,\ldots, L$, the above equation constructs a linear equation group. It can be expressed in matrix form as following.

$$Y_L = H_L\theta + n_L \tag{7}$$

where

$$Y_L = [Y(1), Y(2), \cdots, Y(L)]^T$$

$$n_L = [n(1), n(2), \cdots, n(L)]^T$$

$$H_L = [H(1), H(2), \cdots, H(L)]^T$$

By linear least squares estimation, we can get the estimates of the weight parameters of the wavelet network as:

$$\hat{\theta}_{LS} = (H_L^T H_L)^{-1} H_L^T Y_L \tag{8}$$

The appropriate network structure was found through cross validation. The data for training neural network models were partitioned into the training set and validation set. A neural network was trained on the training set and tested on the validation set. A number of network structures were tried and the one giving the least error on the validation data set is adopted.

To select a proper network structure and network parameters *Ra* and *bn,* the following method is used:

(1) Determine the variation domains of the network inputs, x_{imin} and x_{imax}, $i=1,\ldots,r$;

(2) Select neuron number and centres as following: divide each input domain equally by n_i (start from a small value) and put neurons on grids. The number of neurons is: $N = n_1.n_2.\cdots.n_r$;

(3) Start with a small *Ra* initial value, estimate weight parameters with training set by linear least squares estimation and test the network with validation set. Increase *Ra* value until a satisfied result is obtained. Of cause, the knowledge about frequency property of process identified is helpful to estimate *Ra.*;

(4) Increase n_i and repeat steps (2) and (3) until the best network structure is obtained.

As soon as we get the network structure parameters, we can train the wavelet network. Because network structure parameters have a wide adapted ability, we do not need to search network structure parameters again in general cases. Afterward, it is only a linear least squares estimation problem. This will simplify implementation of wavelet networks and decrease training time especially in on-line model identification.

For online application in adaptive predictive control strategy, a recursive LS algorithm with exponential forgetting algorithm as following is adopted.

$$\begin{aligned}
&\hat{\theta}(t) = \hat{\theta}(t-1) + a(t)K(t)e(t) \\
&e(t) = y(t) - \varphi^T(t)\hat{\theta}(t-1) \\
&K(t) = \frac{P(t-1)\varphi(t)}{1+\varphi^T(t)P(t-1)\varphi(t)} \\
&P(t) = \frac{P(t-1)}{\lambda} - \alpha\frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)}{I+\varphi^T(t)P(t-1)\varphi(t)} \\
&\qquad + \beta I - \delta P(t-1)^2
\end{aligned} \tag{9}$$

$$K(t) = \frac{P(t-1)\varphi(t)}{\alpha + \varphi^T(t)P(t-1)\varphi(t)} \tag{10}$$

$$P(t) = \frac{1}{\alpha}[P(t-1) - \frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)}{\alpha + \varphi^T(t)P(t-1)\varphi(t)}] \tag{11}$$

Because the solid theory basis and special structure of wavelet neural network, wavelet neural network holds some advantages superior to other types of neural networks. First, its network structure is easy to specify based on its theory analysis and intuition. Secondly, network training do not rely on stochastic gradient type techniques such as the "back propagation" and avoids the problem of poor convergence or undesirable local minimum, which is more serious for other types of neural networks when training data is seriously contaminated by noise.

The properties similar to that of recursive linear least-squares algorithm can be obtained as the recursive algorithm is completely same as recursive linear least squares algorithm. In addition, the closed loop-identifiability can be guaranteed. This is because the different wavelon outputs in hidden layer are irrelevant each other as orthogonal wavelet functions are adopted

As soon as we can select appropriate value for M, there exists $\langle f, \varphi_{M,n} \rangle$ making $\sum_n \langle f, \varphi_{M,n} \rangle \varphi_{M,n}(X)$ approximating $f(Y_k, U_k)$ with expected accuracy. Then, we can prove the statistic properties of the weight parameter estimation of wavelet network as linear LS.

3. WAVELET NETWORK MODELING OF A MIMO NONLINEAR PROCESS

The wavelet neural network is used to model a MIMO nonlinear process as following.

$$X_{k+1} = AX_k + f(X_k)U_k \tag{12}$$

$$A = \begin{bmatrix} 0.5 & -0.35 \\ -0.15 & 0.4 \end{bmatrix} \tag{13}$$

$$f(X_k) = \begin{bmatrix} 0.15(x_k(1)+0.1) & 0.05x_k(2) \\ 0.1x_k(1) & 0.1(x_k(2)+0.03) \end{bmatrix} \tag{14}$$

The 1000 data point length's simulation data are produced by the system described by equation (12),(13) and (14). They are split into two set. The first 500 data points were used as training data while the remaining 500 data points were used as testing data.

Throughout the simulation experiment, we will follow the guidelines listed below:
The trained neural network is evaluated only by long range prediction for both the training and 'unseen' testing data. This is because both have good accuracy for one-step-ahead predictions and the dynamic model for control purpose needs to have better long-rang prediction accuracy.
We select $\tau = 0$, $n_y = [1,1]$, $n_u = [1,1]$ and used 16 hidden neurons according to the experiment. The simulation results are shown in Figure 5 and Figure 6. Figure 5 is the prediction result of trained wavelet model for training data. Figure 6 is the prediction result of trained wavelet model for testing data. both prediction result for training data. and testing data is are very good. It is able to satisfy the requirement for dynamic control completely.

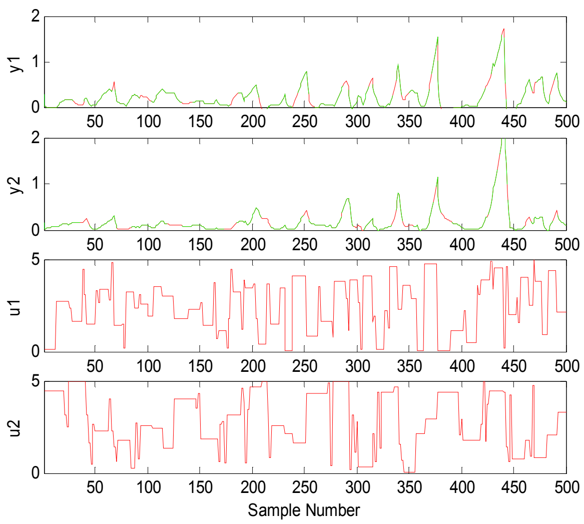

Figure 5. Predictions on training data

In the figures of output predictions, the true process output data is plotted as a solid line, the prediction output data is plotted as a dashed line.

From the simulation result, we observe that, the long range predictions based on wavelet network for training data and testing data have obviously very high prediction accuracy and the curves of both are almost superposition. Both theoretical analysis and simulation studies show that the identification method based on wavelet network is a robust and reliable identification method for non-linear systems.

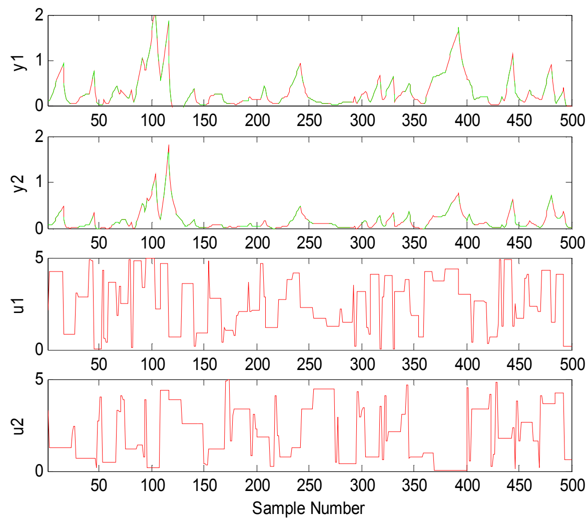

Figure 6. Predictions on testing data

4. Adaptive predictive control based wavelet network

Model predictive control is widely accepted, primarily due to its ability in real-time prediction, real-time optimisation and real-time feedback correction.

In the non-linear adaptive predictive control scheme shown in Figure 7, a process model, i.e., a wavelet network , is explicitly used to predict future process behavior. The same process model is also implicitly adopted to calculate control actions in such a way as to optimise the controller specifications at each sampling step. Furthermore, the difference between the current-time predicted output and the measured current-time process output is used to correct the model error and disturbances so as to improve its robustness. While predictive control is processed, the process model is updated by on-line recursive identification algorithm to enhance its robustness ulteriorly.

Consider MIMO non-linear dynamic system denoted by equation (5).

The selection of the control law is based on a quadratic performance index with a finite time horizon, resulting in the following quadratic programming (QP) problem at time k

$$\min_{\Delta u(k),\Delta u(k+1),\cdots,\Delta u(k+L-1)} J(k)$$

(15)

where J (k) is

$$J(k)=\sum_{i=1}^{P}\left\|Y^{S}(k+i)-\hat{Y}(k+i|k)-Y(k)+\hat{Y}(k|k-P-1)\right\|_{Q}^{2} \quad (16)$$
$$+\sum_{j=1}^{L}\left\|\Delta u(k+j-1|k)\right\|_{R}^{2}$$

where P is the prediction horizon, L is the control horizon, Q and R are weighting matrices.

The process model parameters, i.e. weight parameters of wavelet network, are updated by recursive identification algorithm with forgetting factor in Equation (9),(10) and (11).

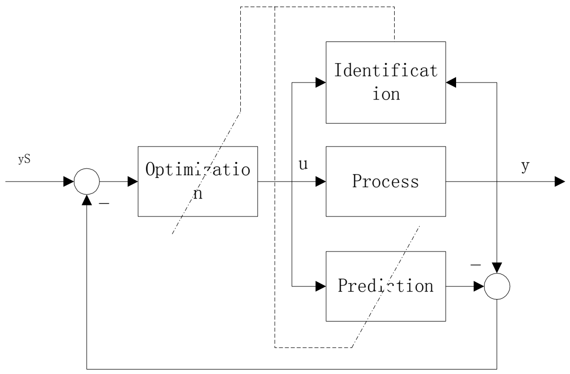

Figure 7. Adaptive Predictive Control Scheme based on Wavelet Network

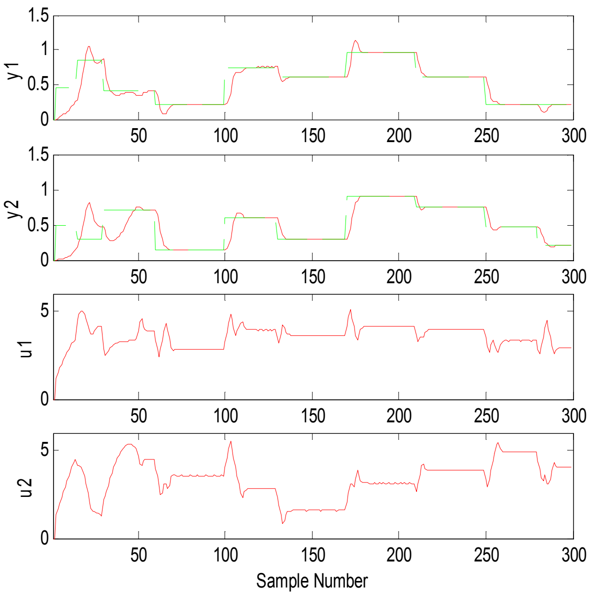

Figure 8. Adaptive Predictive Control based on Wavelet Network

A same simulated MIMO nonlinear process as above illustrates the application of the control scheme. Firstly, the control system uses PID control strategy (in this case, PID control strategy is used in first 50 steps) and then, the adaptive controller based on wavelet network model is closed after a crude model is obtained during PID control. The control result is shown in Figure 8. Simulation results show that expected performances are attained. The nonlinear adaptive predictive control strategy based on wavelet network is superior to the standard PID controller (The control result is shown in Figure 9. PID parameter is optimized by minimizing the integral squared error).

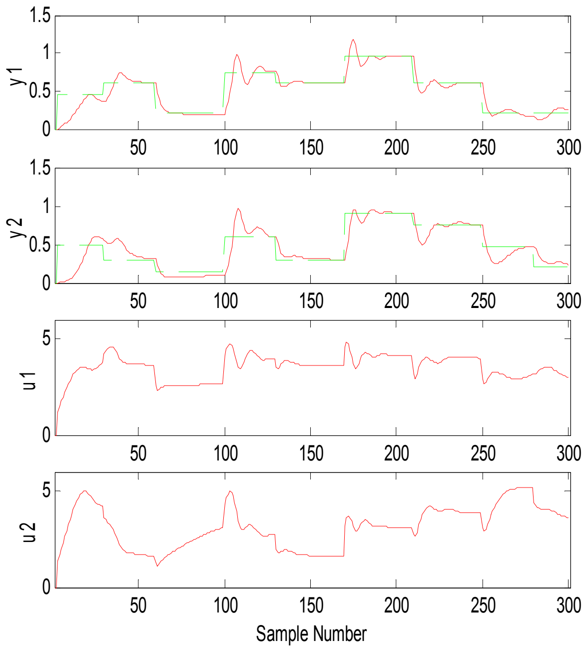

Figure 9. PID control

5. Conclusions

In this paper, a nonlinear MIMO adaptive predictive control strategy based orthogonal wavelet network model is realized. By both theory analysis and simulation study, The following conclusions can be educed.

(1) Wavelet network model only by scale function simplified wavelet network and decreased network size in online training obviously. Its weight coefficients can be estimated by a linear least-squares estimation algorithm. It is different from RBF network and other feed-forward neural networks, because its structure parameters are determined according to wavelet network reconstructing theory, instead of trial and error. In addition, its weight coefficients are estimated by a linear least-squares estimation algorithm, instead of non-linear optimization search method. Therefore, it can be proven that excellent statistic properties of its weight parameters as the

linear least-squares estimation algorithm in system identification has can be obtained (Huang, 2002). The identification method based on wavelet network is a robust and reliable identification method for nonlinear system. In addition, it is also generic method and is easy to use, instead of a method based on trial and error. The long-range predictions based on trained wavelet network for testing data with serious noise have obviously better prediction accuracy than that of the usual feed-forward neural network. The prediction is very close to true output without noise.

(2)For online application in adaptive predictive control strategy, a recursive algorithm is given. The properties similar to that of recursive linear least-squares algorithm can be obtained as the recursive algorithm is completely same as recursive linear least squares algorithm. In addition, the closed-loop identifiability can be guaranteed. This is because the different wavelon outputs in hidden layer are irrelevant each other as orthogonal wavelet functions are adopted. With developed recursive algorithm, a nonlinear MIMO adaptive predictive control strategy is implemented. A same simulated nonlinear process as above illustrates the application of the control scheme. The nonlinear MIMO adaptive predictive control strategy based on wavelet network is superior to the standard PID controller. Even if the optimal PID parameter is used, the control result for PID controller still has larger overshoot for controlled variable in high operation point and has week regulation action in low operation point as the simulated nonlinear system has a serious nonlinear character. In practice, it is difficult to get the optimal PID parameter for a large operation region. In contract, the nonlinear MIMO adaptive predictive controller can identify the control model on-line and achieve a satisfactory control effect by self. Because the controller does not need to be trained before it starts running, it is able to handle the any operating region. The nonlinear MIMO adaptive predictive control strategy is superior to nonlinear controllers in that it does not need to build non-linear control model by user and superior to the nonlinear adaptive controllers based conventional feed-forward neural networks in that it only need finite fix time to on-line updating network model in each control period because it is a recursive linear least squares problem. Besides, it is generic method for both model identification algorithm and control algorithm.

REFERENCES

I. Daubechies, “Orthonomal bases of compactly supported wavelets”, Comm. Pure Applied Math, Vol.91, pp.909-996, 1988.

D. Huang,, and Jin Y., “The application of wavelet neural networks to nonlinear predictive control”, ICNN97 IEEE, Houston, Texas, US, Vol.2, 1997.

D. Huang,, J. Wang and Y. Jin, “Application research of wavelet neural networks in process predictive control”, Journal of Tsinghua University, Vol.39, No. 1, pp. 91-94, 1999.

D. Huang, Y. Jin, J. ZHANG and A. J. Morris Nonlinear “Chemical Process Modelling and Application in Epichlorhydrine Production Plant Using Wavelet Networks”. Chinese Journal of Chemical Engineering, Vol 10, No.4, 2002.

Maner, R. B., Doyle. Ш, F. J., Ogunnaike, B. A. and Pearson, R. K., “ Nonlinear model predictive control of a simulated multivariable polymerization reactor using second-order volterra models” Automatica Vol.32, No.9, pp. 1285-1301, 1996

Bakshi, B. R. and Stephanopoulos, G., “Wave-net: a multiresolution, hierarchical neural network with localised learning”, AICHE J., Vol.39, No.1, pp.57-81, 1993.

T. Proll and M. N. Karim, "Real-time design of an adaptive nonlinear predictive controller", INT. J., Vol. 59, No. 3, pp. 863-889, 1994

J. Zhang,, G. Walter, G. G.,Miao, Y. and Lee, W. N. W., “Wavelet neural networks for function learning” IEEE Signal Processing, Vol.43, No.6, pp.1485-1497,1995.