Hybrid Algorithm for Hammerstein System Identification Using Genetic Algorithm and Particle Swarm Optimization

Tomohiro Hachino, Kenji Shimoda, and Hitoshi Takata

Abstract—This paper presents a method of model selection and identification of Hammerstein systems by hybridization of the genetic algorithm (GA) and particle swarm optimization (PSO). An unknown nonlinear static part to be estimated is approximately represented by an automatic choosing function (ACF) model. The weighting parameters of the ACF and the system parameters of the linear dynamic part are estimated by the linear least-squares method. On the other hand, the adjusting parameters of the ACF model structure are properly selected by the hybrid algorithm of the GA and PSO, where the Akaike information criterion is utilized as the evaluation function. Simulation results are shown to demonstrate the effectiveness of the proposed hybrid algorithm.

Keywords—Hammerstein system, identification, automatic choosing function model, genetic algorithm, particle swarm optimization.

I. INTRODUCTION

With the advance of technology, the importance of nonlinear modeling in control engineering has been growing. Since most practical systems have inherently nonlinear characteristics such as saturation and dead-zone, the development of accurate nonlinear system identification algorithm is a key problem for precise analysis, prediction or control design. One of approaches for nonlinear system identification is use of the block-oriented models. The Hammerstein model is one of the block-oriented models and has been utilized to express nonlinear systems in many applications such as predistorters in wireless communication systems [1] and dc/dc converters [2]. The model consists of a nonlinear static part followed by a linear dynamic part. It has many advantages for control design or stability analysis owing to the simple model structure [3]. Several identification algorithms have been investigated for the Hammerstein model by using correlation theory [4], neural networks [5], orthogonal functions [6], polynomials [7], piecewise linear model [8], automatic choosing function (ACF) model [9], and so on.

In this paper a hybrid algorithm for model selection and identification of discrete-time Hammerstein systems is proposed by using the genetic algorithm (GA) [10] and particle swarm optimization (PAO) [11]. The unknown nonlinear static part to be estimated is represented by an ACF model. The weighting parameters of the ACF and the system parameters of the linear dynamic part are estimated by the linear least-squares (LS) method. The adjusting parameters of the ACF model structure, i.e., the number and widths of the subdomains and the shape of the ACF, are properly determined by the hybrid algorithm of the GA and PSO. The GA is a probabilistic search procedure based on the mechanics of natural selection and natural genetics [10]. It is well known that although the GA has a high potential for global optimization, its convergence is quite slow. On the other hand, PSO is a swarm intelligence optimization technique which was inspired by the social behavior of a flock of birds or a shoal of fish, and has been empirically shown to be very efficient for optimization [11]–[14]. Some theoretical researches such as stability analysis and parameter selection have been also reported for PSO [15], [16]. Similarly to the GA, PSO searches from not a single point but a population of points and only uses information about the objective function, not derivatives or other auxiliary knowledge. PSO is simpler than the GA, because the algorithm of PSO consists of only the basic arithmetic operations and does not require complicated coding and genetic operations such as crossovers and mutations. PSO has quite fast convergence property, but it is inferior to the GA in global optimization. Therefore, we combine the GA with PSO to optimize the adjusting parameters of the ACF model structure. At first the GA are utilized to serve for global optimization, then PSO is applied to work for local optimization, where the Akaike information criterion (AIC) [17] is used for the objective function (fitness value function).

The outline of this paper is as follows: In section II the problem is formulated. In section III the identification method is proposed in the case of fixed ACF model structure. In section IV the optimization method for the adjusting parameters of the ACF model structure is considered by hybridization of the GA and PSO. In section V simulation results are shown to examine the effectiveness of the proposed hybrid algorithm. Finally conclusions are given in section VI.

II. STATEMENT OF THE PROBLEM

Consider a discrete-time nonlinear system described by the Hammerstein model shown in Fig. 1:
\[
\begin{align*}
    A(q^{-1})y(k) &= B(q^{-1})x(k-1) + A(q^{-1})e(k) \\
    x(k) &= f(u(k)) \\
    A(q^{-1}) &= 1 + a_1 q^{-1} + \cdots + a_n q^{-n} \\
    B(q^{-1}) &= b_0 + b_1 q^{-1} + \cdots + b_r q^{-r}
\end{align*}
\]

where \( e(k) \) is the measurement noise. \( q^{-1} \) denotes the backward shift operator. \( n \) and \( r \) are the known degrees of polynomials \( A(q^{-1}) \) and \( B(q^{-1}) \), respectively. \( f(\cdot) \) is an unknown nonlinear function. The problem is to identify the system parameters \( \{ a_i \} \) and \( \{ b_j \} \) of the linear dynamic part and the nonlinear static function \( f(\cdot) \) from input and output measurements.

### III. Identification

In this section the identification algorithm in the case of fixed ACF model structure is presented. In the next section the adjusting parameters of the ACF model structure will be determined by the hybrid approach.

In order to represent the nonlinear function \( f(\cdot) \), the sigmoid type ACF [18] is introduced. Let a domain being a data region of \( u(k) \) be \( D = [u_{min}, u_{max}] \). The domain \( D \) is divided into some subdomains of \( D = \bigcup_{i=1}^{M} D_i \) where \( D_i = [\alpha_i, \beta_i] \). \( \alpha_i = u_{min}, \beta_M = u_{max}, \alpha_k = \beta_{k-1} \) (\( k = 2, 3, \ldots, M \)). Then the ACF is defined by

\[
I_i(u(k)) = 1 - \frac{1}{1 + \exp(H(u(k) - \alpha_i))}
\]

where \( H \) is positive real value. \( I_i(u(k)) \) is almost unity only on a subdomain \( D_i = [\alpha_i, \beta_i] \) and nearly equals to zero on \( D - D_i \), so it chooses \( D_i \) automatically. The ACFs are shown in Fig. 2 in the case of \( H = 6, 20 \) and 200.

Assume that \( f(u(k)) \) is well approximated linearly on each subdomain \( D_i \):

\[
f(u(k)) \simeq c_i + d_i u(k) \quad \text{on} \quad D_i.
\]

Then \( f(u(k)) \) is represented by using the ACF on the whole domain \( D \) as

\[
f(u(k)) = \sum_{i=1}^{M} (c_i + d_i u(k)) I_i(u(k)) + \epsilon(k)
\]

where \( \epsilon(k) \) is an approximation error. The schematic diagram of the ACF model is depicted in Fig. 3.

Substituting (4) into (1) yields

\[
A(q^{-1})y(k) = \sum_{i=1}^{M} c_i B(q^{-1}) I_i(u(k - 1)) + \sum_{i=1}^{M} d_i B(q^{-1}) I_i(u(k - 1)) + v(k)
\]

or in vector form,

\[
y(k) = \varphi^T(k) \theta + v(k)
\]

where \( v(k) = A(q^{-1})e(k) + B(q^{-1})\epsilon(k-1) \) is the equation error, and

\[
\theta = [\theta_a^T, \theta_{c_1}^T, \theta_{c_2}^T, \ldots, \theta_{c_M}^T, \theta_d^T, \theta_{d_1}^T, \theta_{d_2}^T, \ldots, \theta_{d_M}^T]^T
\]

\[
\varphi(k) = [\varphi_a^T(k), \varphi_{c_1}^T(k), \varphi_{c_2}^T(k), \ldots, \varphi_{c_M}^T(k), \varphi_d^T(k), \ldots, \varphi_{d_M}^T(k)]^T
\]

Thus, the unknown parameter vector \( \theta \) is easily evaluated by applying the linear LS method to (6):

\[
\hat{\theta} = \left[ \sum_{k=N}^{N+N} \varphi(k) \varphi^T(k) \right]^{-1} \left[ \sum_{k=N}^{N+N} \varphi(k) y(k) \right]
\]

where \( N \) is the number of input and output data.
Step 1-1: Initialization for GA

Generate an initial population of \( Q \) individuals for \( X \) randomly.

Set the generation number \( g \) to 0.

Step 1-2: Decoding

Decode \( Q \) strings into real values \( X_i^g \) \((i = 1, 2, \ldots, Q)\).

Step 1-3: Judgement of switch from GA to PSO

If the generation number \( g \) is greater than the prespecified \( g_{\text{max}} \), then go to Step 2-1 of the stage 2 (PSO).

Step 1-4: Construction of ACF model

Construct \( Q \) candidates of the ACF model using \( X_i^g \) \((i = 1, 2, \ldots, Q)\).

Step 1-5: Identification by the LS method

Identify \( \hat{\theta}_i \) and \( \hat{f}_i(u(k)) \) \((i = 1, 2, \ldots, Q)\) from (8)~(11), using each candidate of the ACF model.

Step 1-6: Fitness value calculation

Calculate the evaluation values (AIC):

\[
J_i(X_i^g) = N \log \left( \frac{1}{X} \sum_{k=N_i+1}^{N+N} (y(k) - \hat{y}_i(k))^2 \right) + 2P_i \quad (i = 1, 2, \ldots, Q) \quad (12)
\]

and the fitness values \( F_i^q(X_i^g) = \exp(-J_i(X_i^g)/300) \) using \( X_i^g \). \( P_i = n + M_i(r+1) \) is the number of the parameters in the identification model (6). \( \hat{y}_i(k) \) is the output of the \( i \)th candidate of the estimated model.

Step 1-7: Reproduction

Reproduce each of individual strings with the probability of \( F_i^q(X_i^g)/\sum_{j=1}^{Q} F_j \). Practically, the linear fitness scaling [10] is utilized to avoid undesirable premature convergence.

Step 1-8: Crossover

Pick up two strings randomly and decide whether or not to cross them over according to the crossover probability \( P_c \). Exchange strings at a crossing position if the crossover is required. The crossing position is chosen randomly.

Step 1-9: Mutation

After a bit of string (0 or 1) according to the mutation probability \( P_m \).

Step 1-10: Repetition for GA

Set the generation number to \( g = g + 1 \) and go to Step 1-2.

Step 2-1: Initialization for PSO

Set an initial population of \( Q \) particles with positions \( X_i^0 \) and velocities \( V_i^0 \) \((i = 1, 2, \ldots, Q)\). The positions \( X_i^0 \) are constructed by the final population of \( Q \) individuals in the GA. The velocities \( V_i^0 \) are set to be random values.

Set the iteration counter \( l \) to 0.

Step 2-2: Construction of ACF model

Construct \( Q \) candidates of the ACF model using \( X_i^l \) \((i = 1, 2, \ldots, Q)\).

Step 2-3: Identification by the LS method

Identify \( \hat{\theta}_i \) and \( \hat{f}_i(u(k)) \) \((i = 1, 2, \ldots, Q)\) from (8)~(11), using each candidate of the ACF model.

Step 2-4: Evaluation value calculation

Calculate the evaluation values (AIC):

The parameters of the linear dynamic part are estimated by

\[
[a_1, \ldots, a_n, b_0, \ldots, b_i]^T = \left[ I_{(n+r+1) \times (n+r+1)} : 0 \right] \hat{\theta},
\]

setting \( \hat{c}_1 = 1 \) without loss of generality. Namely the parameters of the linear dynamic part can be obtained by taking the first \( n + r + 1 \) elements of \( \theta \).

Next, the parameters for the nonlinear static part are obtained by using the linear LS technique again as

\[
\hat{c}_i = \sum_{j=1}^{r+1} \hat{\theta}_{c1}(j) \hat{\theta}_{c1}(j)/ \sum_{j=1}^{r+1} \hat{\theta}_{c1}(j) \quad (i = 2, 3, \ldots, M)
\]

\[
\hat{d}_i = \sum_{j=1}^{r+1} \hat{\theta}_{d1}(j) \hat{\theta}_{d1}(j)/ \sum_{j=1}^{r+1} \hat{\theta}_{d1}(j) \quad (i = 1, 2, \ldots, M).
\]

Thus the nonlinear static function is composed of the \( \hat{c}_i \) and \( \hat{d}_i \) in (10) as

\[
\hat{f}(u(k)) = \sum_{i=1}^{M} (\hat{c}_i + \hat{d}_i u(k)) L_i(u(k)).
\]
shown in Fig. 5.

Step 2-5: Update of the best positions $pbest_i$ and $gbest_i$. 
Update $pbest_i^l$, which is the personal best position, and $gbest_i$, which is the global best position among all particles as follows:

If $l = 0$ then

$$pbest_i^l = X_i^l \quad \text{and} \quad gbest_i = \arg \min_j J(X_j^l)$$

otherwise

$$pbest_i^l = \begin{cases} X_i^l \quad & (J(X_i^l) < J(pbest_i^{l-1})) \\
\text{pbest}_i^{l-1} \quad & \text{otherwise} \end{cases}$$

$$gbest_i = \arg \min_j J(pbest_i^l).$$

Step 2-6: Update of positions and velocities

Update the particle positions and velocities using (16):

$$V_i^{l+1} = w \cdot V_i^l + c_1 \cdot \text{rand}_1() \cdot (pbest_i^l - X_i^l) + c_2 \cdot \text{rand}_2() \cdot (gbest_i^l - X_i^l)$$

$$X_i^{l+1} = X_i^l + V_i^{l+1}$$

where $w$ is an inertia factor, $c_1$ and $c_2$ are constants representing acceleration coefficients, and $\text{rand}_1()$ and $\text{rand}_2()$ are uniformly distributed random numbers with amplitude in the range $[0, 1]$. Figure 4 shows the update of particle positions.

Step 2-7: Repetition for PSO

Set the iteration counter to $l = l + 1$ and go to Step 2-2 until the termination criterion is satisfied.

Finally, at the termination of this algorithm when $l = l_{max}$, the suboptimal adjusting parameter vector $X$ of the ACF model structure is determined by the best position $gbest^{l_{max}}$. Thus, the final estimated model is constructed from $X$ and the corresponding $\theta$ and $f(u(k))$.

For clarity, the flow chart of the proposed hybrid method is shown in Fig. 5.

V. NUMERICAL SIMULATIONS

Consider a system described by

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

$$x(k) = f(u(k))$$

with noise $\epsilon(k)$, which is a white Gaussian noise with distribution $N(0, 0.0015)$, where the noise-to-signal ratio is $5\%$. The number of input and output data is $N = 300$. The design parameters for the GA and PSO are chosen as follows:

GA parameters:
1) population size: $Q = 50$
2) search range of $\{\alpha_i\}$: $[\alpha_{\min}, \alpha_{\max}] = [-3.0, 3.0]$
3) search range of $H$: $[h_{\min}, h_{\max}] = [1.0, 200.0]$
4) crossover probability: $P_c = 0.8$
5) mutation probability: $P_m = 0.03$
6) maximum generation number: $g_{\max} = 80$

PSO parameters:
1) particle size: $Q = 50$
2) inertia factor: \( w = 0.7 \)
3) acceleration coefficients: \( c_1 = c_2 = 0.8 \)
4) maximum iteration number: \( l_{\text{max}} = 120 \)

Estimates of the system parameters of the linear dynamic part are shown in Table I. In this table, the estimate of \( b_0 \) is omitted because the final estimated model is normalized by \( b_0 \). Clearly estimates by the proposed hybrid method is very close to the true parameters. The estimated nonlinear static function \( f(u(k)) \) is shown in Fig. 6. We can confirm that \( f(u(k)) \) matches the true nonlinear function \( f(u(k)) \) well on the given data region. Figure 7 shows the true output \( y(k) \), the output of the estimated model \( \hat{y}(k) \) and \( |y(k) - \hat{y}(k)| \), where the outputs were generated by the test inputs. It is clear that \( \hat{y}(k) \) is very close to \( y(k) \).

For comparison, identification experiments by two other methods are carried out. One is the GA-based method in which the ACF model structure is determined by using only GA, and the other is the PSO-based method in which the ACF model structure is determined by using only PSO. Monte-Carlo simulations of 20 experiments are implemented for the proposed hybrid method, the GA-based method and the PSO-based method. Table II shows the best, mean and worst evaluation values (AIC) of the 20 experiments for the three methods. The means of the output errors \( \frac{1}{N} \sum_{k=N_s+1}^{N} |y(k) - \hat{y}(k)| \) for the 20 experiments using the test data are shown in Table III. The performance of the proposed hybrid method is superior to those of other two methods.

### VI. Conclusions

In this paper a method of model selection and identification of discrete-time Hammerstein systems by hybridization of the GA and PSO has been proposed. The nonlinear static part is approximately represented by the ACF model. Then the weighting parameters of the ACF and the system parameters of the linear dynamic part are estimated by the linear LS method. The adjusting parameters of the ACF model structure, i.e., the number and widths of the subdomains and the
shape of the ACF, are appropriately determined by the hybrid algorithm of the GA and PSO. Simulation results show that the identification by this method is easy in computation and can give accurate estimated models even in the presence of the measurement noises. Moreover the performance of the proposed hybrid algorithm is superior to those of the GA-based method and PSO-based method.

REFERENCES


Tomohiro Hachino received the B.S., M.S. and Dr. Eng. degrees in electrical engineering from Kyushu Institute of Technology in 1991, 1993 and 1996, respectively. He is currently an Associate Professor in the Department of Electrical and Electronics Engineering, Kagoshima University. His research interests include control and identification for nonlinear systems. Dr. Hachino is a member of IEEJ, SICE, and ISCIE.

Kenji Shimoda received the B.S. and M.S. degrees in electrical and electronics engineering from Kagoshima University in 2007 and 2009, respectively. In 2009, he joined Mitsubishi Electric Corporation. His research interests include nonlinear system identification and evolutionary computation.

Hitoshi Takata received the B.S. degree in electrical engineering from Kyushu Institute of Technology in 1968, and the M.S. and Dr. Eng. degrees in electrical engineering from Kyushu University, in 1970 and 1974, respectively. He is currently a Professor in the Department of Electrical and Electronics Engineering, Kagoshima University. His research interests include control, linearization, and identification for nonlinear systems. Dr. Takata is a member of IEEJ, IEEJ, SICE, RISP, and ISCIE.