

Generating Crisp-Type Fuzzy Models from Operating Data

Cheng-Liang Chen* , Shuo-Huan Hsu, Chung-Tyan Hsieh, Wen-Kuei Lin

Department of Chemical Engineering,
National Taiwan University,
10617 Taipei, TAIWAN
R.O.C.

Abstract— A simple method for generating a multiple-input/single-output crisp-type fuzzy model from observed data is presented. The crisp-type fuzzy model is composed of triangular fuzzy partitions for inputs and singletons for the output. Explicit input-output equation of the crisp-type fuzzy model is supplied. For a set of observed input/output data pairs, the structure and the parameters of a suitable fuzzy model with required accuracy can be determined by a series of algebraic computations. Numerical examples are given to demonstrate the effectiveness of the proposed identification method for modeling static and dynamic data.

I. INTRODUCTION

System Identification is concerned with setting up mathematical models to represent system input-output relationships, and with the choice of a specific model for the class of models. Though numerical data are abundant in most real world problems, the linguistic models are the one that people are accustomed to. This is because people very often make decisions basing on qualitative information. Zadeh's *fuzzy sets* theory was thus proposed to enable people to describe and formulate the linguistic mental models apparent in daily life behaviour [11].

Building model from numerical data by fuzzy technology has several advantages over the other competing modeling technology such as neural networks: (1) the procedures are easy to implement; (2) it is not necessary to apply cumbersome optimization calculation; (3) the structure of the model can be varied during the construction; (4) the resulting model has more physical meanings than others.

Two types of fuzzy models are popular in the literature: the pure fuzzy model and the Takagi-Sugeno-Kang (TSK) fuzzy model [7], [6]. The latter is appealing to practitioners since the consequent parts of the IF-THEN rules in a TSK fuzzy model are simple linear functions of input variables, and thus the model output value is some kind of weighted average of those linear output equations [10]. Some researches [3], [4] have pointed out the convenience of the TSK model, and have presented a constructing algorithm without using optimizations.

The crisp-type fuzzy model is the simplest possible form of the TSK model. The crisp-type fuzzy model studied in this article is composed of triangular fuzzy partitions

for inputs and singletons for the output. Such a fuzzy model with crisp consequent parts can afford adequate approximation properties for a function by using less number of model parameters, when compared to the classical TSK model.

In this article, a simple yet effective method will be presented for generating a multiple-input/single-output crisp-type fuzzy model from observed input/output data. Explicit input-output equation of the crisp-type fuzzy model will be given. For a set of observed input/output data pairs, the structure and the parameters of a suitable crisp-type fuzzy model with required accuracy will be determined by a series of algebraic computations. One advantage of the proposed method is no time-consuming optimization procedure is required. Several numerical examples will be supplied to demonstrate the effectiveness of the proposed model and the identification method.

II. THE CRISP-TYPE FUZZY MODEL

The following equation depicts a typical crisp-type fuzzy model with p inputs and single output,

$$R_i: \text{ If } x_1 \text{ is } X_1^{(i_1)}, x_2 \text{ is } X_2^{(i_2)}, \dots, x_p \text{ is } X_p^{(i_p)} \\ \text{ then } y = y^{(i)} \quad (i = 1, \dots, n) \quad (1)$$

where x_j and y are the j^{th} input variable and the output variable, respectively; $X_j^{(i_j)}$ denotes the i_j^{th} linguistic term for x_j ; $\{X_j^{(1)}, X_j^{(2)}, \dots, X_j^{(J)}\}$ is the term set for x_j with triangular membership functions $X_j^{(i_j)}(x_j)$'s such as shown in Fig. 1; in the figure, $x_j^{(i_j)}$'s are arbitrarily distributed principal values; $y^{(i)}$ is the crisp output value of the i^{th} rule.

For an input pattern (x_1, x_2, \dots, x_p) , the degree of fulfillment of the i^{th} rule, ν_i , is the fuzzy intersection (t-norm) of the relevant input membership values, i.e., $X_1^{(i_1)}(x_1), X_2^{(i_2)}(x_2), \dots, X_p^{(i_p)}(x_p)$.

$$\nu_i = \mathbb{T} \left(X_1^{(i_1)}(x_1), X_2^{(i_2)}(x_2), \dots, X_p^{(i_p)}(x_p) \right) \\ = \prod_{j=1}^p X_j^{(i_j)}(x_j) \quad (\text{if } \mathbb{T}(a, b) = ab) \quad (2)$$

*Author to whom all correspondence should be addressed,
E-mail: ccl@ccms.ntu.edu.tw

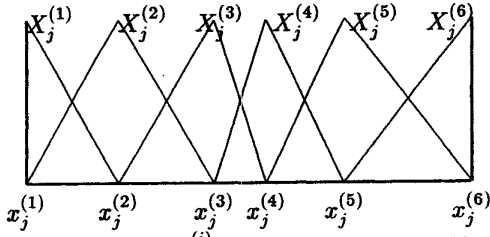


Fig. 1. Linguistic terms $X_j^{(i)}$'s and triangular membership functions $X_j^{(i)}(x_j)$'s for variable x_j , where $x_j^{(i)}$'s are principal values

Here, \mathbb{T} can be any t-norm. Without loss of generality, we adopt the algebraic product for \mathbb{T} in the following discussion. The corresponding model response is a weighted sum of crisp outputs of the n rules, $\{y^{(i)} \mid i = 1, \dots, n\}$,

$$y = \frac{\sum_{i=1}^n \nu_i y^{(i)}}{\sum_{i=1}^n \nu_i} = \sum_{i=1}^n \phi_i y^{(i)}$$

where

$$\phi_i = \frac{\nu_i}{\sum_{i=1}^n \nu_i} \quad (3)$$

Notably, only two input membership values are nonzero for each element of the input data, and the membership values sum to unity. Thus for any input pattern, there are 2^p non-zero elements in $\{\phi_1, \phi_2, \dots, \phi_n\}$. Therefore only 2^p rules make contribution to the final output for any input pattern.

III. GENERATING THE CRISP-TYPE FUZZY MODEL

The unknown parameters in a fuzzy model include structural parameters (number of rules or number of linguistic terms for each variable) and tuning parameters (locations of principal values $x_j^{(i)}$'s and crisp output of each rule $y^{(i)}$, i.e., the input/output membership functions). A simple yet effective identification method is presented in the following to determine both model structure and input/output membership functions in an integrated procedure. The proposed method comprises three computational steps: *initializing phase*, *growing phase*, and *refining phase*. Details of the three identification steps are provided in the following.

A. Initializing Phase

A simplest possible crisp-type fuzzy model is determined at first where two linguistic terms with the same support over the whole universe of discourse are used for each variable.

The target of this step is to build the initial model with two linguistic terms for each input. The initializing step will generate 2^p rules for a p-input-single-output system. The computational details include:

1. Find the universe of discourse for each input variable,

$[x_{j,min}, x_{j,max}]$, where

$$\begin{aligned} x_{j,min} &= \min \{x_{j,1}, x_{j,2}, \dots, x_{j,m}\} \quad j = 1, \dots, p \\ x_{j,max} &= \max \{x_{j,1}, x_{j,2}, \dots, x_{j,m}\} \quad j = 1, \dots, p \end{aligned}$$

2. Construct two linguistic terms, $X_j^{(1)}$ and $X_j^{(2)}$, with principal values at $x_{j,min}$ and $x_{j,max}$, respectively, for x_j with the universe of discourse as the support set. The p input variables, each with two linguistic terms, can establish 2^p rules with the following antecedent part,

$$\begin{aligned} R(i) : & \text{If } x_1 \text{ is } X_1^{(i_1)}, x_2 \text{ is } X_2^{(i_2)}, \dots, x_p \text{ is } X_p^{(i_p)} \\ & \text{then } y = y^{(i)} \quad i_j \in \{1, 2\} \quad i = 1, \dots, 2^p \end{aligned}$$

3. Determine the consequence part for each of the 2^p rules by the nearest data point. For example, if $(x_{1,\ell}, x_{2,\ell}, \dots, x_{p,\ell})$ in the training data set is the nearest point to $(x_{1,min}, x_{2,min}, \dots, x_{p,min})$, i.e., $\sum_{j=1}^p (x_{j,\ell} - x_{j,min})^2$ is minimum, the output measurement y_ℓ corresponding to $(x_{1,\ell}, x_{2,\ell}, \dots, x_{p,\ell})$ is set to be the crisp output of the rule. That is,

$$\begin{aligned} R(i=1) : & \text{If } x_1 \text{ is } X_1^{(i_1)}, x_2 \text{ is } X_2^{(i_2)}, \dots, x_p \text{ is } X_p^{(i_p)} \\ & \text{then } y = y^{(i=1)} = y_\ell, \quad i_j = 1 \quad \forall j \in \{1, \dots, p\} \end{aligned}$$

B. Growing Phase

Additional linguistic terms and inference rules should be appended on the existing fuzzy model if the latter can not afford required accuracy for modeling sampling data. It is reasonable to append new terms and rules on the existing model thus the point with maximum modeling error can be eliminated. The model outputs at sampling points can be found from Eq. (3),

$$\begin{aligned} \hat{y}_k &= \frac{\sum_{i=1}^n \nu_i y^{(i)}}{\sum_{i=1}^n \nu_i} = \sum_{i=1}^n \phi_i y^{(i)} \quad k = 1, \dots, m \\ \nu_i &= \prod_{j=1}^p X_j^{(i_j)}(x_{j,k}) \end{aligned}$$

Suppose maximum error of the current model outputs occurs at $(x_{1,\gamma}, x_{2,\gamma}, \dots, x_{p,\gamma})$. Then one should append new linguistic term for x_j with principal value $x_{j,\gamma}$. The crisp outputs of the additional rules can be found by the nearest-point method as mentioned in initializing phase. Notably the total number of inference rules becomes $L_1 \times \dots \times L_p$, where L_j denotes number of linguistic terms for x_j .

C. Refining Phase

One can enhance the performance of the fuzzy model by adjusting the crisp consequent part of the whole rules. With a set of new crisp outputs $y^{(i)}$'s of these rules, the model output at the k^{th} sampling point is,

$$\hat{y}_k = \sum_{i=1}^n \phi_{ki} y^{(i)} = [\phi_{k1} \quad \phi_{k2} \quad \dots \quad \phi_{kn}] \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(n)} \end{bmatrix} \quad (4)$$

For all sampling data, Eq. (4) becomes,

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_m \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} & \cdots & \phi_{1n} \\ \phi_{21} & \phi_{22} & \cdots & \phi_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{m1} & \phi_{m2} & \cdots & \phi_{mn} \end{bmatrix} \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(n)} \end{bmatrix} \quad (5)$$

or,

$$\hat{Y} = \Phi A \quad (6)$$

One can obtain a new A , the new crisp outputs, which gives minimized sum of squared errors between the model outputs, \hat{Y} , and the observed data, $Y \equiv [y_1 \ y_2 \ \cdots \ y_m]^T$,

$$A = (\Phi^T \Phi)^{-1} \Phi^T Y \quad (7)$$

Notably the matrix of firing levels, Φ , might be large and sparse. It may have some difficulty to calculate the pseudo-inverse. The recursive least-squares technique is recommended in such case [2], [4], [7].

D. Stopping Criteria

It is a trade-off problem between model complexity and accuracy. Two criteria are suggested for selecting a simplest possible model with required accuracy.

1. Maximum number of input linguistic terms

A realistic linguistic model should not consist of too many linguistic terms. The maximum number of linguistic terms of each variable can be predefined. Notably the number of inference rules is meanwhile limited. This criterion can prevent the fuzzy model from becoming unacceptable high dimension. The accuracy of the obtained model is not guaranteed, however.

2. Maximum mean of squared errors

The model complexity will be increased to proceed pre-specified accuracy. The mean of squared errors (MSE) is an adequate performance measure of model accuracy.

$$MSE = \frac{\sum_{k=1}^m (\hat{y}_k - y_k)^2}{m} \quad (8)$$

IV. NUMERICAL EXAMPLES

The proposed method gives a simple way to construct a fuzzy model without optimizations. Two examples are used to illustrate the effectiveness of the identification method.

Example 1: Modeling a static function

Consider a nonlinear function with two inputs and single output,

$$y = e^{x_1} + e^{x_2} \quad 0 \leq x_1, x_2 \leq 5 \quad (9)$$

There are 1,000 input/output data randomly distributed on the considered domain, such as shown in Fig. 2.

1. Initial phase

The upper and lower bounds for the two inputs, $x_{1,\min}$, $x_{1,\max}$, $x_{2,\min}$ and $x_{2,\max}$ are found at first. Two linguistic

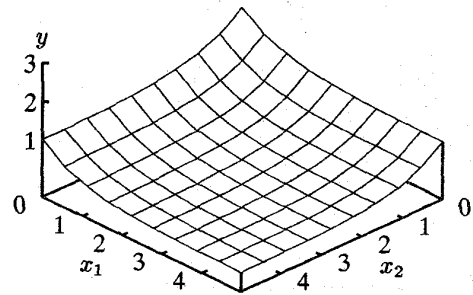


Fig. 2. Original output surface for example 1

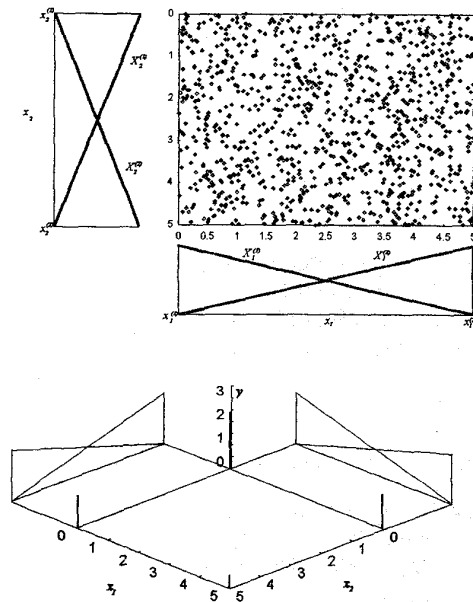


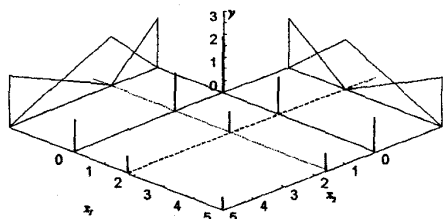
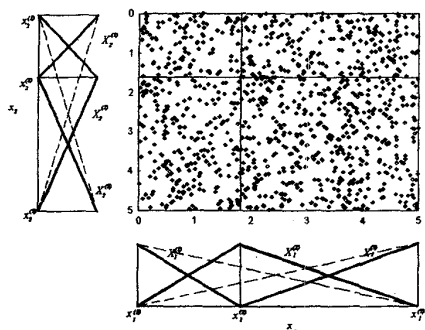
Fig. 3. Input membership functions and crisp-type fuzzy rules for the initializing phase of example 1

terms are defined for x_1 and x_2 , respectively, and the crisp outputs for the four initial rules are evaluated, such as shown in Fig. 3.

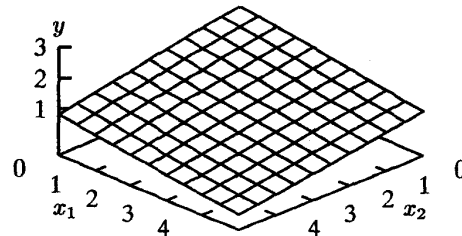
2. Growing phase

The maximum discrepancy between the model outputs and the observations is found. Additional linguistic terms are appended for both variables to eliminate the maximum error. The crisp consequent parts of the additional rules are determined subsequently such as shown in Fig. 4. The growing phase is continued until at least one stopping criterion is satisfied. Fig. 5 depicts results of the second iteration.

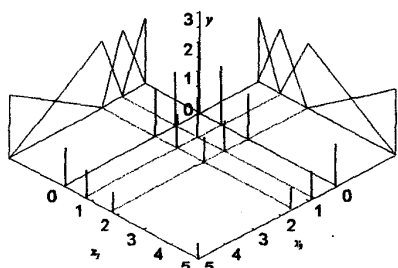
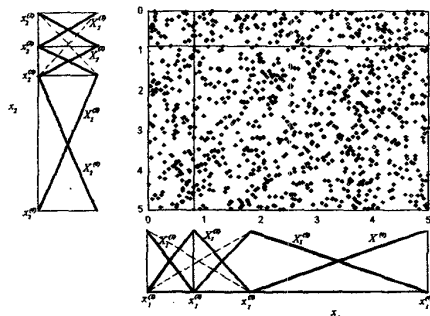
Fig. 6 illustrates the output surfaces of the resulting models. Fig. 7 are results of refining the model in the sense of least squared errors.



(a) 2×2 rules, MSE = 0.294



(b) 3×3 rules, MSE = 0.0124



(c) 4×4 rules, MSE = 0.00458

Fig. 5. Input membership functions and crisp-type fuzzy rules for the growing phase of example 1: second iteration

Fig. 6. The model output surfaces for example 1 during the development of the crisp-type fuzzy model

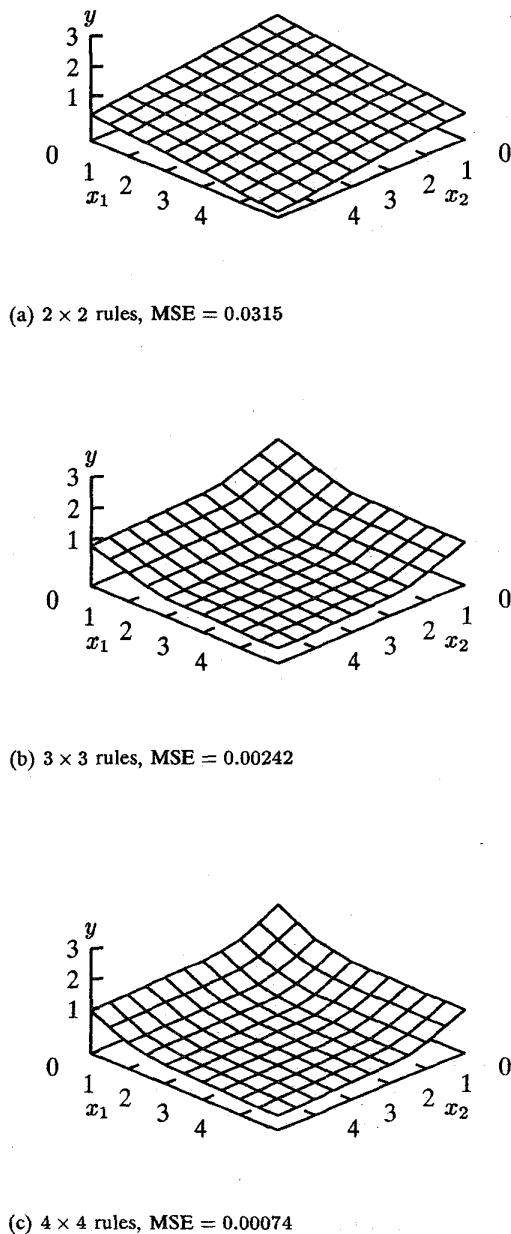


Fig. 7. The refined model output surfaces for example 1

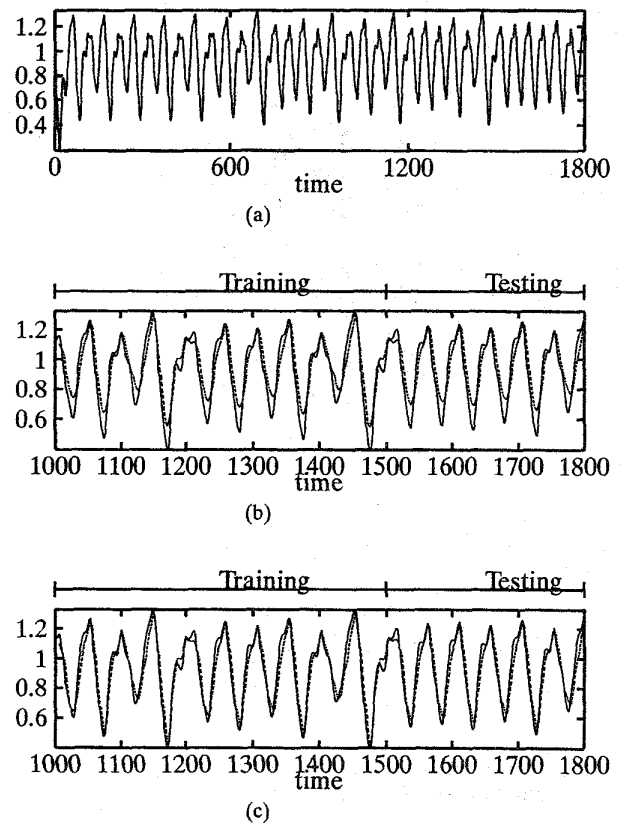


Fig. 8. Modeling Mackey-Glass time series: (a) original observations (-); (b) 3^4 rules, MSE = 0.0057; (c) 4^4 rules, MSE = 0.0023

Example 2: Mackey-Glass chaotic time series [5], [8]

Mackey-Glass delay differential equation is used as a benchmark problem in the neural network and fuzzy modeling communities:

$$\dot{x}(t) = \frac{0.2x(t-\tau)}{1+x^{10}(t-\tau)} - 0.1x(t), \quad \tau \geq 17 \quad (10)$$

Eq. (10) is integrated with initial conditions $x(0) = 1.2, \tau = 17$, and $x(t) = 0 \quad \forall t < 0$ [3]. The proposed model is a 4-input-single-output system. The inputs are $x_1 \equiv x(t-18), x_2 \equiv x(t-12), x_3 \equiv x(t-6), x_4 \equiv x(t)$. The output is $y \equiv x(t+6)$. One set of time series with 1,800 samples is shown in Fig. 8(a). In this modeling problem, the sampling data for $1,000 \leq t \leq 1,500$ are used for identification, and data in $1,500 \leq t \leq 1,800$ are applied to test the adequacy of the resulting model. The number of linguistic terms is limited within four for each input variable. The results of identification and testing are illustrated in Fig. 8 (b),(c).

V. EXTENSION TO TAKAGI-SUGENO-KANG MODEL

A Takagi-Sugeno-Kang (TSK) fuzzy model has linear equations as its consequent parts. For a p-input-single-

output system it is,

$$R^{(i)}: \text{ If } x_1 \text{ is } X_1^{(i_1)}, x_2 \text{ is } X_2^{(i_2)}, \dots, x_p \text{ is } X_p^{(i_p)} \\ \text{ then } y = a_{i0} + \sum_{j=1}^p a_{ij} x_j \quad (11)$$

Here a_{ij} 's are coefficients to be determined in the consequent part. The least squares technique [9] must be used in each growing phase for identifying a TSK model. For the k^{th} point in a set of m training data, the input-output relationship of a TSK model is,

$$\hat{y}_k = \sum_{i=1}^n \phi_{ki} y^{(i)} \\ = \phi_{k1}(a_{10} + a_{11}x_{1k} + a_{12}x_{2k} + \dots + a_{1p}x_{pk}) \\ + \phi_{k2}(a_{20} + a_{21}x_{1k} + a_{22}x_{2k} + \dots + a_{2p}x_{pk}) + \dots \\ + \phi_{kn}(a_{n0} + a_{n1}x_{1k} + a_{n2}x_{2k} + \dots + a_{np}x_{pk}) \\ = [\phi_{k1} \ \phi_{k2} \ \dots \ \phi_{kn}] \begin{bmatrix} a_{10} \\ a_{20} \\ \vdots \\ a_{n0} \end{bmatrix} \\ + [\phi_{k1}x_{1k} \ \phi_{k2}x_{1k} \ \dots \ \phi_{kn}x_{1k}] \begin{bmatrix} a_{11} \\ a_{22} \\ \vdots \\ a_{n1} \end{bmatrix} + \dots \\ + [\phi_{k1}x_{pk} \ \phi_{k2}x_{pk} \ \dots \ \phi_{kn}x_{pk}] \begin{bmatrix} a_{1p} \\ a_{2p} \\ \vdots \\ a_{np} \end{bmatrix}$$

Let

$$\mathbf{a}_j \equiv [a_{1j} \ a_{2j} \ \dots \ a_{nj}]^T \\ \varphi_0 \equiv \begin{bmatrix} \phi_{11} & \phi_{12} & \dots & \phi_{1n} \\ \phi_{21} & \phi_{22} & \dots & \phi_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{m1} & \phi_{m2} & \dots & \phi_{mn} \end{bmatrix} \\ \varphi_j \equiv \begin{bmatrix} \phi_{11}x_{j1} & \phi_{12}x_{j1} & \dots & \phi_{1n}x_{j1} \\ \phi_{21}x_{j2} & \phi_{22}x_{j2} & \dots & \phi_{2n}x_{j2} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{m1}x_{jm} & \phi_{m2}x_{jm} & \dots & \phi_{mn}x_{jm} \end{bmatrix} \\ j = 1, 2, \dots, p$$

Thus, if we consider all training data, the result is almost the same as Eq. (7):

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_m \end{bmatrix} = [\varphi_0 \ \varphi_1 \ \dots \ \varphi_p] \begin{bmatrix} \mathbf{a}_0 \\ \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_p \end{bmatrix} \quad (12)$$

or,

$$\hat{\mathbf{Y}} = \Phi \mathbf{A} \quad (13)$$

The least-squares solution of \mathbf{A} is the same as Eq. (7).

VI. CONCLUSIONS

A simple algorithm for identifying a crisp-type fuzzy model has been proposed in this article. One major advantage of using the crisp-type fuzzy model is that the input-output relation of the model can be derived explicitly. For a set of observed patterns, a simplest possible fuzzy model with required accuracy can be determined in three steps: the initializing phase, the growing phase, and the refining phase. No time-consuming optimization techniques are required in the identification of model structure and the estimation of model parameters. The proposed method has great potential for engineering applications since one can proceed the model identification without any searching procedure and thus no numerical converging problem. Two numerical examples are supplied to demonstrate the applicability of the proposed method in building a crisp-type fuzzy model from static or dynamic operating data.

References

- [1] P. Costa Branco and J. Dente, *Fuzzy Systems Modelling in Practice*, in Seventh IFSA World Congress, Prague, 1997, pp. 360-365.
- [2] T. C. Hsia, *System Identification*, Lexington, 1979.
- [3] Y. Nakoula, S. Galichet, and L. Foulloy, *Simultaneous Learning of Rules and Linguistic Terms*, in Fifth IEEE International Conference on Fuzzy Systems, New Orleans, 1996, pp. 1743-1749.
- [4] K. Nozaki, H. Ishibuchi, and H. Tanaka, *A Simple but Powerful Heuristic Method for Generating Fuzzy Rules from Numerical Data*, *Fuzzy Sets and Systems*, 86 (1997), pp. 251-270.
- [5] A. Shmilovici and J. Aguilar-Martin, *Fuzzy Systems Identification Using Data Transformations*, in Seventh IFSA World Congress, Prague, 1997, pp. 330-335.
- [6] M. Sugeno and G. T. Kang, *Fuzzy Modelling and Control of Multi-layer Incinerator*, *Fuzzy Sets and Systems*, 18 (1986), pp. 329-346.
- [7] T. Takagi and M. Sugeno, *Fuzzy Identification of Systems and Its Applications to Modeling and Control*, *IEEE Trans. on Systems, Man, and Cybernetics*, SMC-15 (1985), pp. 116-132.
- [8] L.-X. Wang and J. M. Mendel, *Generating Fuzzy Rules by Learning from Examples*, *IEEE Trans. on Systems, Man, and Cybernetics*, 22 (1992), pp. 1414-1427.
- [9] N. Watanabe and T. Imaizumi, *On Least Squares Methods in Fuzzy Modeling*, in Seventh IFSA World Congress, Prague, 1997, pp. 336-341.
- [10] R. R. Yager and D. P. Filev, *Essentials of Fuzzy Modeling and Control*, John Wiley & Sons, Inc., New York, 1994.
- [11] L. A. Zadeh, *Fuzzy sets*, *Information and Control*, 8 (1965), pp. 338-353.