The Fuzzy Crystallization Algorithm: A New Approach to Complex Systems Modeling

Chi-Cheng Cheng, *Member, IEEE,* and Wen-Hsiung Hsieh

***Abstract*—A new identification method for fuzzy modeling is introduced. Since the method has some analogy with the process of material crystallization in nature, the name of fuzzy crystallization algorithm (FCA) is given to this novel approach. This method accomplishes structure identification and parameter identification at the same time, and possesses the properties of simplicity, flexibility, and high calculation speed. Compared with other modeling strategies, it is easier to construct a model with a specific accuracy. Numerical examples are provided to demonstrate the performance of this approach.**

***Index Terms*—Fuzzy logic, fuzzy systems, identification, knowledge acquisition, modeling, parameter estimation.**

I. Introduction

IT IS ALWAYS an important task to establish models of complex processes in the real world. However, in general, it is not easy to construct a proper mathematical model for engineering purposes. This is the main reason why fuzzy modeling is popular. In recent years, researchers have proposed a number of fuzzy modeling techniques to deal with complex, ill-defined, and uncertain systems. Their studies can be classified into two directions [1]: the direct approach and the identification approach.

The direct approach is an implementation from Zadeh's idea [2] of extracting the fuzzy model directly from an expert's knowledge [3]–[6]. Since these fuzzy modeling techniques are based solely upon the expert's description of the system, and quantitative observations are not specifically used, some inherent limitations inevitably exist. For example, if the expert's knowledge about the system is faulty or incomplete, a poor model could be obtained. Furthermore, if it is difficult to acquire the expert's knowledge directly, model discrepancy due to indirect approaches may not be easily removed through these techniques.

The second direction of the fuzzy modeling techniques, namely the identification approach, is based on an examination of input–output (I/O) data. In other words, it is the method of extracting fuzzy rules directly from the quantitative observations of the system. The identification approach consists of two major phases. The first phase is identification of the structure of the fuzzy model (structure identification), and the second phase is the estimation of parameter values of the fuzzy model (parameter identification). There are several methods [7]–[14] to manage the parameter identification and some other schemes [15]–[19] to handle the structure identification. There have also been a number of strategies that can model a system without a prior expert's knowledge by successfully combining both the structure identification and the parameter identification methods [20]–[24]. However, it is of interest to note that the structure identification phases always come before the parameter identification phases in these strategies. Essentially, in these methods, the second phase (parameter identification) cannot be started until the first phase (structure identification) is accomplished. Therefore, this class of methods belongs to the category of structure-then-parameter identification methods.

In principle, the structure and parameters should be identified simultaneously, since they have a mutual relationship to, and influence on, the accuracy of the identified model. Therefore, the major disadvantage of structure-then-parameter identification methods is that the most important rules identified from the I/O data may not work properly especially when accuracy of the model is required.

In this paper, a new approach called the fuzzy crystallization algorithm (FCA) is developed for identification of a fuzzy system model. The main reason why this name is chosen is that the procedures of this method have some analogy with the processes of crystallization in nature. In general, crystallization describes the solidification process of materials from their liquid state [25], [26]. When the liquid gradually cools, crystallization begins with the formation of solid nuclei, which then grow by consuming the melt. The processes of nucleation and crystal growth in material crystallization will be adopted in the proposed modeling algorithm by establishing "virtual crystals," based on specific fuzzy relations, in data space.

To the authors' knowledge, this method is the first one dealing with the structure identification and the parameter identification of a fuzzy system model simultaneously. Hence, this new identification approach of the fuzzy system model can be treated as a structure-and-parameter identification method. As compared with structure-then-parameter identification methods, an advantage of the proposed approach is that it is easier to make sure that the model will provide a specified accuracy. It is worth mentioning that this novel identification method has the virtues of simplicity, flexible adaptability to complex or multidimensional systems, and high modeling speed. Moreover, extracting fuzzy rules from numerical data can be automatically achieved.

Manuscript received April 5, 1998; revised June 26, 2001. This work was supported by the National Science Council of Taiwan, R.O.C., under Contract NSC 85 2212 E 110 001. This paper was recommended by Associate Editor A. Kandel.

C.-C. Cheng is with the Department of Mechanical and Electro-Mechanical Engineering, National Sun Yat-Sen University, Kaohsiung, Taiwan, R.O.C. (e-mail: chengcc@mail.nsysu.edu.tw).

W.-H. Hsieh is with Alcatel Taiwan, Kaohsiung, Taiwan, R.O.C.

Publisher Item Identifier S 1083-4419(01)08638-1.

II. Concept of the Hypercrystal

The operation of the FCA is based on the formation of hypercrystals. A hypercrystal can be viewed as a volume set in data space, whose elements possess similar data properties. The following concepts establish the development of a formed hypercrystal.

A. Unit Lattice

Without loss of generality, only the multi-input single-output system is considered here. Suppose we are given a collection of n data points, which come from an s-input one-output system. The data points are in the $s+1$ dimensional space $\mathbf{R}^s \times \mathbf{R}$: $(X^{(1)}, y^{(1)})$, $(X^{(2)}, y^{(2)})$, …, $(X^{(n)}, y^{(n)})$ where $X^{(i)} = (x_1^{(i)}, x_2^{(i)}, \ldots, x_s^{(i)})$, and $x_j^{(i)}$ and $y^{(i)}$ denote the jth input and the output of the ith data point, respectively. We shall restrict the s dimensional input space $\mathbf{R}^s$ to an s dimensional hypercube $I_1 \times I_2 \times \cdots \times I_s$where the intervals I_j, $j = 1, 2, \ldots, s$ are defined by the ranges of $x_j^{(i)}$, i.e., $I_j = [x_j^{(-)}, x_j^{(+)}]$ and

$$\forall i = 1, 2, \ldots, n \qquad x_j^{(-)} < x_j^{(i)} < x_j^{(+)}. \tag{1}$$

The hypercube contains the input parts of all the data points. Furthermore, each interval I_j is discretized into s equidistant segments with a small constant length Δx_j. Such a discretization establishes an s dimensional grid which separates the hypercube into a number of small s-dimensional cubic regions, namely unit lattices. A unit lattice L can be represented by $L = (l_1, l_2, \ldots, l_s)$ where l_j takes a value from the set $\{1, 2, \ldots, (x_j^{(+)} - x_j^{(-)})/\Delta x_j\}$. There are two important pieces of information associated with each unit lattice, i.e., a central data point and a set of directional properties.

A central data point of a unit lattice L is defined as the data pair $(X^{[L]}, y^{[L]})$, where $X^{[L]} = (x_1^{[L]}, x_2^{[L]}, \ldots, x_s^{[L]})$, i.e., the central inputs of the unit lattice L. In addition, for any data point $(X^{(i)}, y^{(i)})$ there exists a specific unit lattice L that meets the following condition:

$$\forall j = 1, 2, \ldots, s \quad x_j^{[L]} - \tfrac{1}{2}\Delta x_j < x_j^{(i)} \le x_j^{[L]} + \tfrac{1}{2}\Delta x_j. \tag{2}$$

We say that the data point $(X^{(i)}, y^{(i)})$ is contained in the unit lattice L. If the unit lattice consists of m data points, the central output of the unit lattice $y^{[L]}$ is set to be the average of the output of these data points, i.e.,

$$y^{[L]} = \left\{ \frac{\sum y^{(i)}}{m} \,\middle|\, \left(X^{(i)}, y^{(i)}\right) \text{ is contained in } L,\ m \neq 0 \right\}. \tag{3}$$

If m equals zero then $y^{[L]} = \phi$ where ϕ is a null set. In the FCA, a central data point of a unit lattice is used to represent all the data points within the lattice. For all data points $(X^{(i)}, y^{(i)})$ contained in a unit lattice L the following cost function should reach the minimum:

$$\sum_i \left\{ y^{(i)} - y^{[L]} \right\}^2. \tag{4}$$

B. Directional Properties

The other information in a unit lattice is a set of directional properties, which describes the data trends for the output. For the sake of simplicity, a special condition that the central output is a null set is not examined here, and will be carefully studied in Section III.

The directional property of a unit lattice L, $P^{[L]}$, can be represented by $P^{[L]} = (p_1^{[L]}, p_2^{[L]}, \ldots, p_s^{[L]})$ where $p_j^{[L]}$ is L's directional property in the direction of x_j, and is defined as follows.

For a unit lattice L and its central output $y^{[L]}$, considering U_j as a unit vector in the jth dimension of the hypercube

If $y^{[L-U_j]} < y^{[L]} < y^{[L+U_j]}$ then $p_j^{[L]} = (\rightarrow)$
If $y^{[L-U_j]} > y^{[L]} > y^{[L+U_j]}$ then $p_j^{[L]} = (\leftarrow)$
otherwise $p_j^{[L]} = (\leftrightarrow)$.

Hence, a unit lattice's directional property in the direction of x_j is to take an element from the set $\{(\rightarrow), (\leftarrow), (\leftrightarrow)\}$ to represent the relationships among $y^{[L-U_j]}$, $y^{[L]}$, and $y^{[L+U_j]}$. Actually, the directional property presents a qualitative gradient, which classifies tendency of data samples into three basic categories:

1) increasing;
2) decreasing;
3) constant features.

C. Hypercrystal

Generally speaking, the key to success in complex systems modeling is to let the data completely and honestly represent all their characteristics. Using the FCA, the nature of hypercrystals allows the data to represent information in terms of directional properties. Given a collection of unit lattices with their central data points and directional properties, a subset taken from this collection—namely the hypercrystal—is the basic element for the proposed modeling approach. The hypercrystal is defined as follows.

A hypercrystal is an s-dimensional cubic space of a number of s-dimensional unit lattices. For all directions x_j, $j = 1, 2, \ldots, s$, the directional properties p_j of these unit lattices should meet one of the following conditions.

1) All of the directional properties p_j are $(\rightarrow)$ or some are $(\rightarrow)$ and all the others are $(\leftrightarrow)$.
2) All of the directional properties p_j are $(\leftarrow)$ or some are $(\leftarrow)$ and all the others are $(\leftrightarrow)$.
3) All directional properties are $(\leftrightarrow)$.

Therefore, for a hypercrystal H, its directional property can be denoted by $P^{\{H\}} = (p_1^{\{H\}}, p_2^{\{H\}}, \ldots, p_s^{\{H\}})$ where $p_j^{\{H\}}$ is Hs directional property in the direction of x_j and should be equal to $(\rightarrow)$ or $(\leftarrow)$ or $(\leftrightarrow)$.

An s-dimensional cubic space has a hypercubic shape with 2^s corners. To have a complete description of the hypercrystal region, at least two of the unit lattices, located at corners of a hypercrystal, are sufficient for fully characterizing the region of the hypercrystal. If $C = [c_1, c_2, \ldots, c_s]$ is a corner unit lattice, its complementary lattice $\overline{C} = [\overline{c}_1, \overline{c}_2, \ldots, \overline{c}_s]$ denotes the other corner unit lattice in the same hypercrystal with the condition: $\forall j = 1, 2, \ldots, s \quad \overline{c}_j \neq c_j$. Therefore, the symbol $\langle C; \overline{C} \rangle$ can be utilized as a representation of the hypercrystal region.

Theorem 1: Given a hypercrystal H and its region $\langle C; \overline{C}\rangle$ where $c_j < \overline{c}_j$ for any unit lattice L belonging to H.

If $p_j^{\{H\}} = (\rightarrow)$, then $y^{[L+(c_j-l_j)U_j]} \le \ldots \le y^{[L+(-1)U_j]} \le y^{[L]} \le y^{[L+U_j]} \le \ldots \le y^{[L+(\overline{c}_j-l_j)U_j]}$

If $p_j^{\{H\}} = (\leftarrow)$, then $y^{[L+(c_j-l_j)U_j]} \ge \ldots \ge y^{[L+(-1)U_j]} \ge y^{[L]} \ge y^{[L+U_j]} \ge \ldots \ge y^{[L+(\overline{c}_j-l_j)U_j]}$

If $p_j^{\{H\}} = (\leftrightarrow)$, then $y^{[L+(c_j-l_j)U_j]} = \ldots = y^{[L+(-1)U_j]} = y^{[L]} = y^{[L+U_j]} = \ldots = y^{[L+(\overline{c}_j-l_j)U_j]}$.

Based on the definitions of directional properties of a unit lattice and a hypercrystal above, it is obvious that the preceding statements are true.

Theorem 2: For all the corner unit lattices of a hypercrystal, there is at least one whose central output is a global minimum and there also exists a global maximum.

Proof: Given a hypercrystal H and its region $\langle C; \overline{C}\rangle$ for any unit lattice L belonging to H then from Theorem 1

$$\min\left\{y^{[(c_1,l_2,\ldots,l_s)]}, y^{[(\overline{c}_1,l_2,\ldots,l_s)]}\right\} \le y^{[(l_1,l_2,\ldots,l_s)]} \le \max\left\{y^{[(c_1,l_2,\ldots,l_s)]}, y^{[(\overline{c}_1,l_2,\ldots,l_s)]}\right\}$$

$$\min\left\{y^{[(c_1,c_2,l_3,\ldots,l_s)]}, y^{[(c_1,\overline{c}_2,l_3,\ldots,l_s)]}\right\} \le y^{[(c_1,l_2,\ldots,l_s)]} \le \max\left\{y^{[(c_1,c_2,\ldots,l_s)]}, y^{[(c_1,\overline{c}_2,\ldots,l_s)]}\right\}$$

$$\vdots$$

$$\min\left\{y^{[(\overline{c}_1,\ldots,\overline{c}_{s-1},c_s)]}, y^{[(\overline{c}_1,\ldots,\overline{c}_{s-1},\overline{c}_s)]}\right\} \le y^{[(\overline{c}_1,\ldots,\overline{c}_{s-1},l_s)]} \le \max\left\{y^{[(\overline{c}_1,\ldots,\overline{c}_{s-1},c_s)]}, y^{[(\overline{c}_1,\ldots,\overline{c}_{s-1},\overline{c}_s)]}\right\}$$

therefore

$$y^{[L]} \le \max\left\{y^{[C]} \,\middle|\, C \in \text{corner unit lattices of } H\right\}$$
$$y^{[L]} \ge \min\left\{y^{[C]} \,\middle|\, C \in \text{corner unit lattices of } H\right\}.$$

Q.E.D.

Theorem 3: Given a hypercrystal H, if there exists a corner unit lattice C whose central output is a global minimum, then the central point of C's complement $\overline{C}$ is a global maximum, and vice versa.

Proof: If $y^{[C]}$ is a global minimum and from Theorem 1

$$y^{[(c_1,c_2,\ldots,c_s)]} \le y^{[(\overline{c}_1,c_2,\ldots,c_s)]} \le y^{[(\overline{c}_1,\overline{c}_2,\ldots,c_s)]} \le \ldots \le y^{[(\overline{c}_1,\overline{c}_2,\ldots,\overline{c}_s)]}$$

$$y^{[(c_1,c_2,\ldots,c_s)]} \le y^{[(\overline{c}_1,c_2,\ldots,c_s)]} \le y^{[(\overline{c}_1,c_2,\overline{c}_3\ldots,c_s)]} \le \ldots \le y^{[(\overline{c}_1,\overline{c}_2,\ldots,\overline{c}_s)]}$$

$$\vdots$$

$$y^{[(c_1,c_2,\ldots,c_s)]} \le y^{[(c_1,c_2,\ldots,\overline{c}_s)]} \le y^{[(\overline{c}_1,c_2,\ldots,\overline{c}_s)]} \le \ldots \le y^{[(\overline{c}_1,\overline{c}_2,\ldots,\overline{c}_s)]}$$

therefore $y^{[C]} \le \{y^{[V]} | V \subset H\} \le y^{[\overline{C}]}$. Q.E.D.

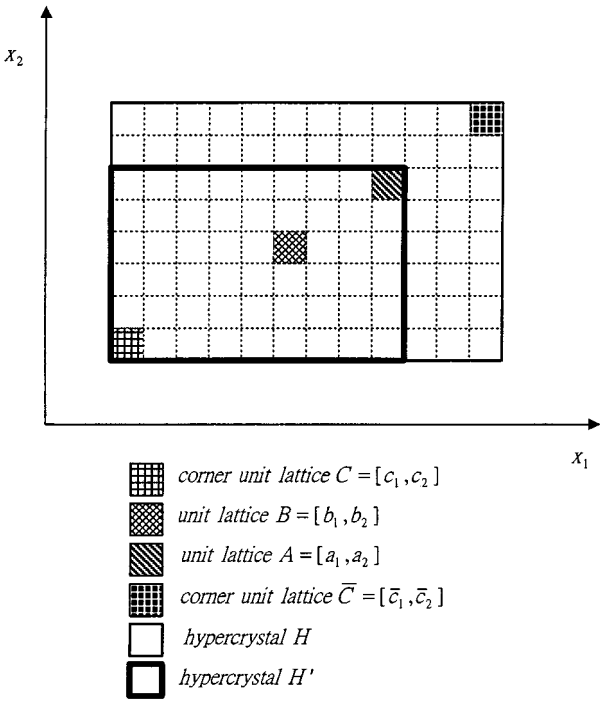
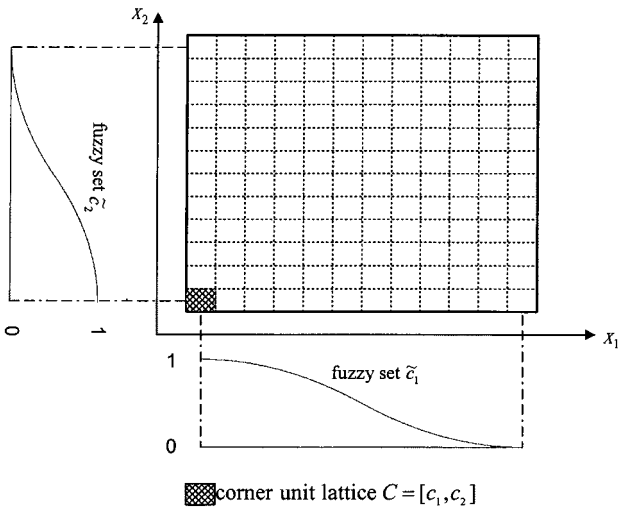


Fig. 1. Theorem 4.



R_1: if x_1 is $\tilde{c}_1$ and x_2 is $\tilde{c}_2$ then y is $y^{[1]} \cong y^{[c_1,c_2]}$

Fig. 2. Fuzzy rule of a corner unit lattice.

Theorem 4: Given a hypercrystal H and C is a corner unit lattice of H, if $y^{[C]}$ is a global minimum, then for any unit lattice A, $B \in H$ and $B \in \langle C; A\rangle$, $y^{[C]} \le y^{[B]} \le y^{[A]}$.

Proof: As illustrated in Fig. 1, since $A \in H$, $\langle C; A\rangle$ is also a region of a hypercrystal, namely H'. Because $y^{[C]}$ is a global minimum in H', $y^{[A]}$ is a global maximum in H'. From Theorem 3

$$\because B \subset \langle C; A\rangle \qquad \therefore y^{[C]} \le y^{[B]} \le y^{[A]}.$$

Q.E.D.

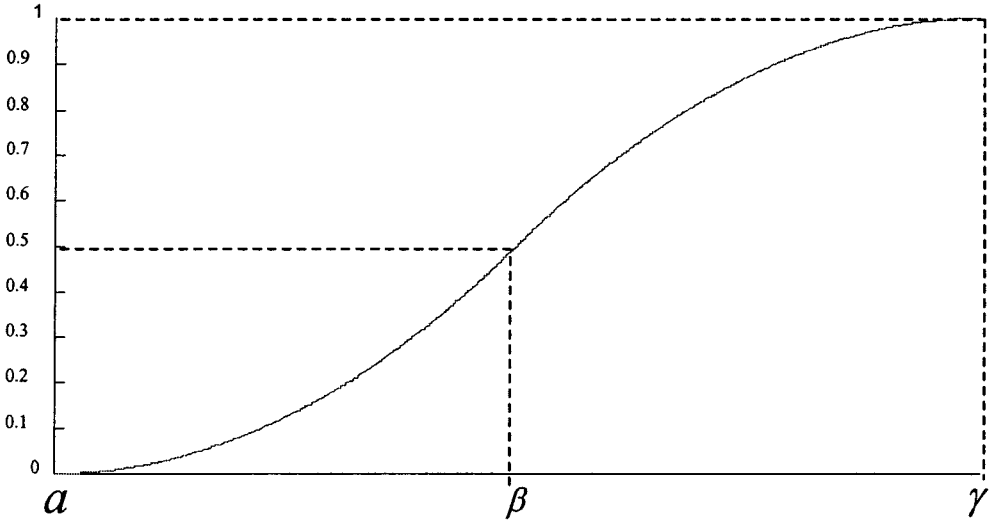


Fig. 3. Zadeh's S membership function.

D. Formed Hypercrystals

Instead of modeling all the data points at one time, the FCA tries to get a number of subsets, called hypercrystals, of all of the data, and attempts to model the data in each hypercrystal. Since all unit lattices in a hypercrystal own consistent directional properties, inclinations of data samples in the whole hypercrystal region behave similarly. Therefore, all the data in a hypercrystal can be modeled by approximate interpolation between the central outputs of all the corner unit lattices. As shown in Fig. 2, each corner unit lattice can be transformed into an s-input one-output fuzzy rule. Hence, 2^s fuzzy rules with $2s$ input fuzzy sets are used as the modeling structure of an s-dimensional hypercrystal.

$$
\begin{aligned}
R_1\text{:}\quad & \text{If } x_1 \text{ is } \tilde{c}_1 \text{ and } x_2 \text{ is } \tilde{c}_2 \text{ and } \cdots \text{ and } x_s \text{ is } \tilde{c}_s \\
& \text{then} \quad y \text{ is } y^{[1]} \cong y^{[(c_1, c_2, \ldots, c_s)]} \\
R_2\text{:}\quad & \text{If } x_1 \text{ is } \tilde{c}_1 \text{ and } x_2 \text{ is } \tilde{c}_2 \text{ and } \cdots \text{ and } x_s \text{ is } \tilde{\overline{c}}_s \\
& \text{then} \quad y \text{ is } y^{[2]} \cong y^{[(c_1, c_2, \ldots, \overline{c}_s)]} \\
& \vdots \\
R_{2^s}\text{:}\quad & \text{If } x_1 \text{ is } \tilde{\overline{c}}_1 \text{ and } x_2 \text{ is } \tilde{\overline{c}}_2 \text{ and } \cdots \text{ and } x_s \text{ is } \tilde{\overline{c}}_s \\
& \text{then} \quad y \text{ is } y^{[2^s]} \cong y^{[(\overline{c}_1, \overline{c}_2, \ldots, \overline{c}_s)]}
\end{aligned}
$$

where

R_i ith rule;
$\tilde{c}_j, \tilde{\overline{c}}_j$ fuzzy sets;
$y^{[i]}$ fuzzy singleton.

The central points of corner unit lattices are chosen as candidates of the fuzzy rules' outputs.

There are some differences between the fuzzy rules described above and the others that are broadly used in the field of fuzzy modeling and control. At first, the input fuzzy sets are in the form of Zadeh's S-function [27] instead of a simple triangular function, as depicted in Fig. 3. The main reason why Zadeh's S-function is selected is that whenever a hypercrystal is given, one more variable for each input fuzzy set is required for modeling adjustment. As a result, Zadeh's S-function is applied here and has the following formulation:

$$
S(x; \alpha, \beta, \gamma) = \begin{cases} 0, & x \le \alpha \\ \dfrac{1}{2}\left(\dfrac{x-\alpha}{\beta-\alpha}\right)^2, & \alpha < x \le \beta \\ 1 - \dfrac{1}{2}\left(\dfrac{x-\gamma}{\beta-\gamma}\right)^2, & \beta \le x \le \gamma \\ 0, & x > \gamma \end{cases} \tag{5}
$$

$$
\overline{S}(x; \alpha, \beta, \gamma) = \begin{cases} 0, & x \le \alpha \\ 1 - \dfrac{1}{2}, \left(\dfrac{x-\alpha}{\beta-\alpha}\right)^2, & \alpha < x \le \beta \\ \dfrac{1}{2}\left(\dfrac{x-\gamma}{\beta-\gamma}\right)^2, & \beta \le x \le \gamma \\ 0, & x > \gamma. \end{cases} \tag{6}
$$

Secondly, it is well known that many kinds of intersection operations exist for the fuzzy rules. In this paper, only one is chosen, for the sake of clarity. However, a single type of intersection is not enough for this proposed fuzzy modeling method. There are two kinds of corner unit lattice, differentiated by whether its central output is a global extreme or not. From Theorem 4, it can be concluded that the corner unit lattice whose central output is a global extreme should be more important than others. Therefore, two different intersections for these two kinds of fuzzy rules should be employed. In this method, an intersection operation (t-norm) is adopted from [28]. The intersection operation can be described by

$$
i_\omega(a, b) = 1 - \min\left\{1, [(1-a)^\omega + (1-b)^\omega]^{1/\omega}\right\} \tag{7}
$$

where both a and b are fuzzy sets. From the equation above, two t-norms, $i_{\omega 1}$ and $i_{\omega 2}$ with the condition $\omega 1 > \omega 2$, are selected for global extremes and nonextremes, respectively, to assign different weights based on their degrees of importance. The

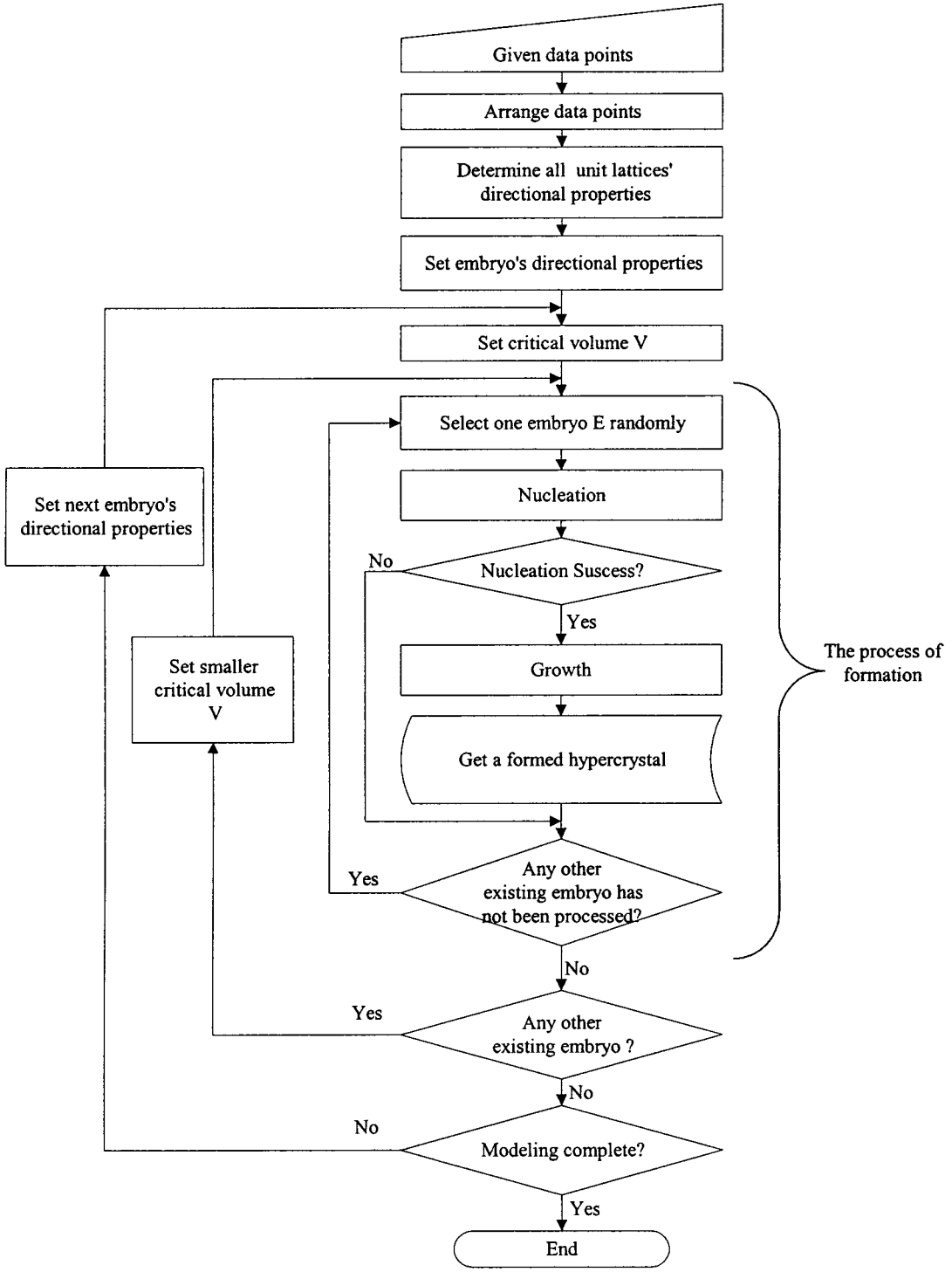


Fig. 4. Flowchart of the FCA process.

following two specific t-norms will be applied for numerical examples:

$$i_{\omega 1} = \min(a, b) \qquad \omega 1 \to \infty$$
$$i_{\omega 2} = \max(0, a + b - 1) \qquad \omega 2 = 1.$$

In the FCA, a hypercrystal H can be modeled by a set of fuzzy rules with a specific modeling accuracy, Θ. If there are m unit lattices whose central outputs are not null sets in the hypercrystal, and the following inequality is satisfied, then the hypercrystal becomes a formed hypercrystal.

$$\left\{ \frac{1}{m} \sum \left(y^{[L]} - \tilde{f}\left(X^{[L]} \right) \right)^2 \middle| \, L \in H \text{ and } y^{[L]} \neq \phi \right\} \leq \Theta \tag{8}$$

where $\tilde{f}(X^{[L]})$ is the output inferred with the input $X^{[L]}$. The goal of the FCA is to find a collection of formed hypercrystals.

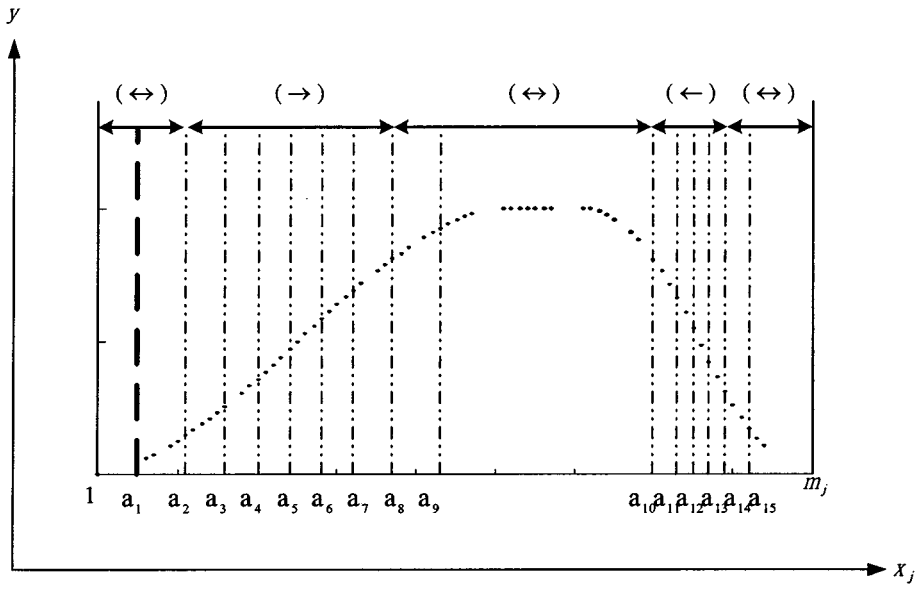


Fig. 5. Directional properties.

III. Procedures of the FCA

The procedures of the proposed modeling method, FCA, generate a collection of formed hypercrystals from an s-dimensional hypercube. All the procedures are shown in the flowchart of Fig. 4. The details of each step will be described below.

A. Data Arrangement

The first step of the FCA is to arrange the given data points into a set of orderly central data points of unit lattices. Given a set of n data points $(X^{(1)}, y^{(1)}), (X^{(2)}, y^{(2)}) \ldots (X^{(n)}, y^{(n)})$ where $X = (x_1, x_2, \ldots, x_s)$ determine an s-dimensional hypercube $I_1 \times I_2 \times \cdots \times I_s$ which contains the input parts of all the data points. Then, discretize each interval I_j into small equidistant segments with a fixed length Δx_j thus obtaining a collection of unit lattices. For each unit lattice, the central data point is determined by (3).

In the discretization of the s-dimensional hypercube, the smaller $\Delta X = (\Delta x_1, \Delta x_2, \ldots, \Delta x_s)$ becomes, the more unit lattices are generated. Consequently, there is a trade-off between the model's precision and computational efficiency.

B. Directional Properties of Unit Lattices

In general, there may be some unit lattices without any data points, and there will be noise contained in the available data points. Therefore, for practical complex systems modeling, some extensions of the method of directional properties of unit lattices have to be developed.

As shown in Fig. 5, instead of considering the relations among $y^{[L-U_j]}$, $y^{[L]}$, and $y^{[L+U_j]}$ an extended method using the sequence $(y^{[L+(-l_j)U_j]}, y^{[L+(1-l_j)U_j]}, \ldots, y^{[L+(m_j-l_j)U_j]})$ is proposed for global consideration where m_j is equal to $(x_j^{(+)} - x_j^{(-)})/\Delta x_j$. Given a specific value h depending on the required modeling accuracy, search through the series $(y^{[L+(-l_j)U_j]}, y^{[L+(1-l_j)U_j]}, \ldots, y^{[L+(m_j-l_j)U_j]})$ sequentially. At the beginning, find the first value that is not a null set in the sequence, for example $y^{[L+(a_1-l_j)U_j]}$, and then continue looking for a series $(a_2, a_3, \ldots)$ with the following condition:

$$\left| y^{[L+(a_i-l_j)U_j]} - y^{[L+(a_{i-1}-l_j)U_j]} \right| \geq h \quad \text{where} \quad i = 2, 3, \ldots. \tag{9}$$

Whenever the search is completed, a series of numbers $a_1, a_2, a_3, \ldots$ are obtained. Hence, the directional properties can be found by the following algorithm.

If $y^{[L+(a_i-1-l_j)U_j]} < y^{[L+(a_i-l_j)U_j]} < y^{[L+(a_{i+1}-l_j)U_j]} < y^{[L+(a_{i+2}-l_j)U_j]}$
then $p_j^{[L+(a_i-I_j)U_j]}, \ldots, p_j^{[L+(a_{i+1}-I_j)U_j]} = (\rightarrow)$

If $y^{[L+(a_i-1-l_j)U_j]} > y^{[L+(a_i-l_j)U_j]} > y^{[L+(a_{i+1}-l_j)U_j]} > y^{[L+(a_{i+2}-l_j)U_j]}$
then $p_j^{[L+(a_i-I_j)U_j]}, \ldots, p_j^{[L+(a_{i+1}-I_j)U_j]} = (\leftarrow)$

otherwise $p_j^{[L+(a_i-I_j)U_j]}, \ldots, p_j^{[L+(a_{i+1}-I_j)U_j]} = (\leftrightarrow)$.

There are always some regions of $(\leftrightarrow)$ between $(\rightarrow)$ regions and $(\leftarrow)$ regions and so the problem of crisp boundaries between hypercrystals is avoided.

C. Directional Properties of Embryos

In the FCA, the formation process of a hypercrystal starts from a specific unit lattice, called an embryo. The directional properties of a hypercrystal are determined by the directional properties of its embryos. Therefore, if a hypercrystal H grows from an embryo E then its directional properties should satisfy

$$\forall j = 1, 2, \ldots, s \qquad p_j^{\{H\}} = p_j^{[E]}. \tag{10}$$

It is obvious that the directional properties of a hypercrystal should be determined before the formation. From the definition of hypercrystals, it is clear that the unit lattice's directional properties $(\rightarrow)$ and $(\leftarrow)$ play more important roles than $(\leftrightarrow)$ in the

composition of a hypercrystal. Hence, the procedure of setting the embryos should start from unit lattices without any $(\leftrightarrow)$ to the ones that are all $(\leftrightarrow)$. Once the formed hypercrystals are sufficient to cover all the data points, the procedure stops and the end of the FCA is reached.

D. Critical Volume of a Hypercrystal

Since there are many embryos that could be formed into hypercrystals, the method of obtaining the better ones becomes substantial. The concept of the critical volume therefore needs to be introduced.

In the FCA, the volume of a hypercrystal indicates the number of the unit lattices in the hypercrystal. The larger volume a formed hypercrystal grows, the better the hypercrystal should be. The critical volume is the allowable minimum volume of a hypercrystal in the formation process. If a developing hypercrystal cannot build its volume larger than the critical value, then the hypercrystal fails. Furthermore, in order to have the hypercrystals as large as possible, the demanding size of the critical volume is appointed from the larger to the smaller. Hence, the minimum number of formed hypercrystals required for modeling will be obtained. The iterations of setting critical volume stop when there are no embryos left for formation.

E. Process of Formation

The process of formation of a hypercrystal is the most important task in the procedures of FCA. A number of variables are involved in each iteration of formation. They should be initialized at the beginning and include the following.

- Demanded directional properties of embryo $P^{[E]}$.
- Demanded critical volume V.
- Growing flag $\Psi = (\psi_1, \psi_2, \ldots, \psi_{2s})$ where ψ_i is the growing flag in the ith extending direction.
- Each fuzzy rule's output $(y^{[1]}, y^{[2]}, \ldots, y^{[2^s]})$.
- Middle point $(\beta_1, \overline{\beta}_1, \beta_2, \overline{\beta}_2, \ldots, \beta_s, \overline{\beta}_s)$ where β_j is the second parameter of the fuzzy set $\tilde{c}_j$ defined in (5) and (6).

The first two parameters are specified before formation and will not be changed during the formation process. However, the other three variables should be initialized and modified during the procedure of formation. A growing flag is assigned to each extending direction to indicate whether the extension of the hypercrystal in this direction is allowable. If an extension is permitted, we set its corresponding flag to one, otherwise the flag remains zero. In addition, the last two parameters represent the resulting fuzzy rules of the formed hypercrystal.

1) Search for an Embryo: An embryo is a unit lattice with the following properties.

- Its central output is not a null set.
- It does not belong to any other formed hypercrystals.
- Its directional property is $P^{[E]}$.

From a series of existing embryos, we randomly select one for the next step of formation. This embryo will not be utilized again during the iterations for the same demanded critical volume. Whenever a hypercrystal is formed, the number of the existing embryos will be reduced. Even if the formation fails, the number of the existing nontried embryos will still be decreased by one.

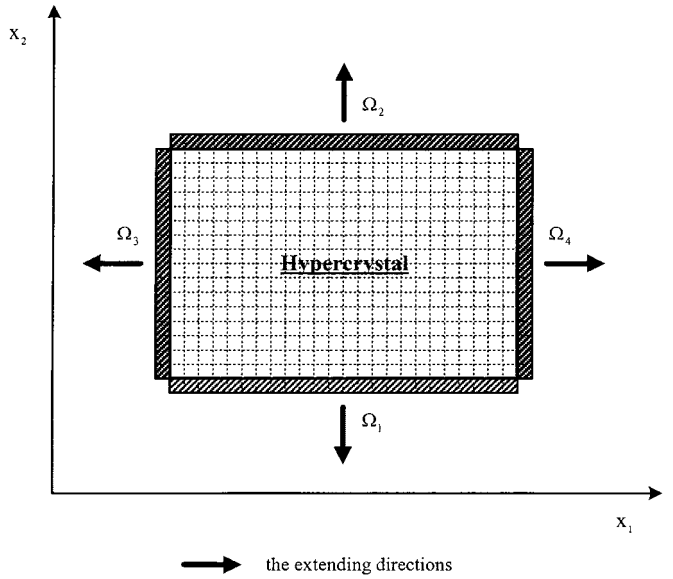


Fig. 6. Growing operation.

The iteration of searching embryos will be stopped when the number of remaining unused embryos becomes zero.

2) Nucleation: Once an embryo is selected, the next procedure, nucleation, is designed for allowing the embryo to be developed into a nucleus. A nucleus is a formed hypercrystal whose volume is just larger than the critical value, i.e., the smallest acceptable formed hypercrystal with the volume V. The nucleation process is accomplished by combining two important operations: developing and tuning.

Assume that the region of an s-dimensional hypercrystal H is denoted by $\langle C; \overline{C} \rangle$ where $c_j \leq \overline{c}_j$, $j = 1, 2, \ldots, s$. As shown in Fig. 6, there are $2s$ extending directions. For each extending direction, there is a layer of unit lattices waiting for the extension. The layer of unit lattices in the ith extending direction is described by Ω_i where $i = 1, 2, \ldots, 2s$. The developing operation of H is to determine an extending direction via the growing performance index. However, if all the growing performance indexes are zeros, the developing process is aborted. The growing performance index in the ith extending direction g_i is determined by examining the following conditions.

1) $p_j^{[L]} = p_j^{[E]}$ or $p_j^{[L]} = (\leftrightarrow)$.
2) The unit lattice L has not be modeled by other hypercrystals.
3) $y^{[L]} \neq \phi$.

The growing performance index can be evaluated by justifying how good the above conditions are met. If all three situations are satisfied, the maximum index value should be assigned. Finally, if the maximum nonzero growing performance index is g_k, then Ω_k is added to the hypercrystal H. Therefore, after a successful operation of developing, the hypercrystal's region will extend by one layer.

The tuning process is to adjust the parameters to achieve the aim of (8). In other words, whenever a hypercrystal is obtained, the fuzzy modeling rules for the hypercrystal are determined except for the input fuzzy sets' second parameters $(\beta_1, \overline{\beta}_1, \beta_2, \overline{\beta}_2, \ldots, \beta_s, \overline{\beta}_s)$ and the output singletons $(y^{[1]}, y^{[2]}, \ldots, y^{[2^s]})$ of the fuzzy rules. This is a problem of

multidimensional minimization. The tuning of an s-dimensional hypercrystal minimizes the model error by adjusting the parameters in the constrained space

$$\alpha_j < \beta_j < \gamma_j \quad \text{and} \quad \overline{\alpha}_j < \overline{\beta}_j < \overline{\gamma}_j \tag{11}$$

$$\left(y^{[C_i]} - \Delta y\right) \le y^{[i]} \le \left(y^{[C_i]} + \Delta y\right) \tag{12}$$

where C_i is the ith corner unit lattice and Δy is selected by the user. The downhill simplex method [29] is adopted for this optimization application due to its high efficiency. If the statement of (8) cannot be achieved after many iterations or the simplexes (all the βs and output singletons) leave the constrained space, then the tuning process stops. In the procedures of the FCA, the tuning operation consumes most of the calculation time. Fortunately, the tuning speed will be efficiently fast because the number of parameters is small and the space for adjusting the parameters is also limited.

The developing of a hypercrystal starting from an embryo cannot proceed until the hypercrystal is just larger than the critical volume V. If the nucleation stops due to failure in either the developing or the tuning, a search for a new embryo is executed. If the nucleation is successful, a nucleus will be formed and the next step, growth, will proceed.

3) Growth: The nucleation procedure is followed by the growth procedure, which is to expand the formed hypercrystal as much as possible. Actually, the growth of a hypercrystal is a series of iterations between the operations of growing and tuning. The growing process is similar to the process of developing, except that the growing process specifically indicates the expanding procedure from a formed nucleus. The region of the nucleus can therefore increase based on similar approaches described in nucleation. This is the last procedure of the formation of a hypercrystal and a formed hypercrystal is now obtained. The modeling process will be fully accomplished as long as sufficient number of hypercrystals can include all the data points.

It should be noted that the growing operation (structure identification) and the tuning action (parameter identification) are performed alternately in the formation process of a hypercrystal. Hence, the proposed modeling method is capable of performing structure identification and parameter identification concurrently. This is the reason why this approach is defined as a structure-and-parameter identification method.

4) Ending: When all the given data points are modeled, the FCA comes to an end. As a result, a collection of formed hypercrystals with some overlay regions along boundaries of adjacent formed hypercrystals are obtained. The set of formed hypercrystals can therefore be used as a model for those (I/O) data pairs as shown in Fig. 7 where $\widehat{y}_i$ is the output inferred from the fuzzy rules of the ith formed hypercrystals and κ_i is a value defined by

$$\text{if} \quad X \in H_i \quad \text{then} \quad \kappa_i = 1 \quad \text{else} \quad \kappa_i = 0.$$

IV. Numerical Examples

A. Box and Jenkins's Gas Furnace Data

Consider the modeling of a dynamic process given in series J of Box and Jenkins [30]. This process describes a gas furnace

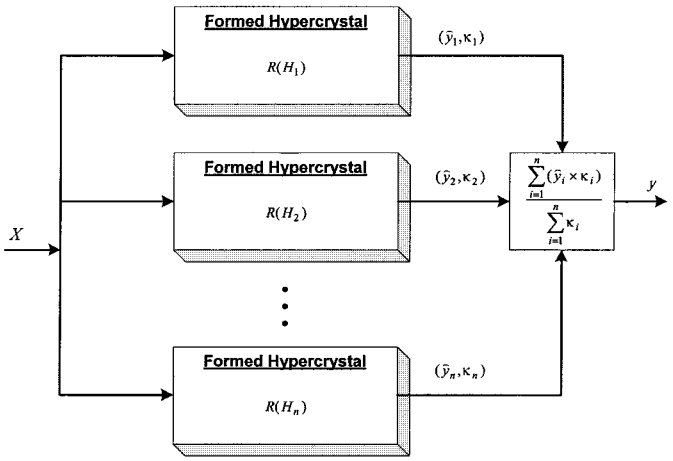


Fig. 7. Resulting fuzzy model.

TABLE I
Fuzzy Model of the Gas Furnace System

Formed hypercrystal H_1, $P^{\{H_1\}} = \{(\rightarrow),(\leftrightarrow)\}$,			
	$y(t-1)$	$u(t-4)$	$y(t)$
R_1	$\overline{S}(44.05, 56.39, 61.85)$	$\overline{S}(-2.975, -1.055, 2.925)$	57.04
R_2	$\overline{S}(44.05, 56.39, 61.85)$	$S(-2.975, -1.161, 2.925)$	46.01
R_3	$S(44.05, 55.70, 61.85)$	$\overline{S}(-2.975, -1.055, 2.925)$	60.73
R_4	$S(44.05, 55.70, 61.85)$	$S(-2.975, -1.161, 2.925)$	60.20

with a single input $u(t)$ of the gas flow rate and a single output $y(t)$ of the CO_2 concentration. There are 296 I/O data pairs $\{u(t), y(t) \,|t = 1, 2, \ldots, 296\}$ which are well known and frequently used as a benchmark example for testing of identification algorithms. $y(t-1)$ and $u(t-4)$ are taken as input variables to produce the output $y(t)$. From these data pairs, an I/O data set $\{(X(5), y(5)), (X(6), y(6)), \ldots, (X(296), y(296))\}$ is established where $X(i) = \{u(i-4), y(i-1)\}$.

The result of the FCA under the demanded accuracy $\Theta = 0.145$ is a single formed hypercrystal, as shown in Table I. The identified model's mean squared error (MSE) of lead one forecast can be calculated from

$$J = \frac{1}{292} \sum_{k=5}^{296} [y(k) - \hat{y}(k)]^2 \tag{13}$$

where $y(k)$ is the output of the furnace at the kth sampling instant and $\hat{y}(k)$ is the output of the model. Simulation results of the lead one forecast using the identified model are illustrated in Fig. 8. Since the demanded accuracy Θ defined in (8) is based on the central output of unit lattices $y^{[L]}$ instead of the genuine output y used in (13), the resulting modeling MSE may appear to be slightly larger than the demanded accuracy. Table II compares our fuzzy model's performance with other existing models identified from the same data. It can be seen that our model presents encouraging performance. It should also be noted that the modeling speed and the required intervention from humans

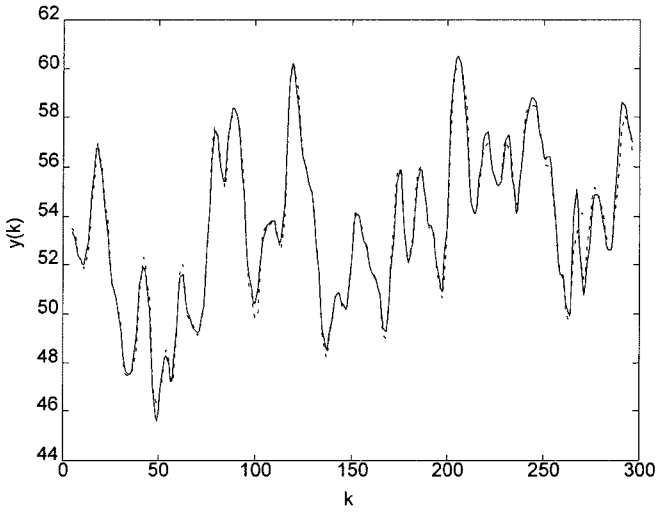


Fig. 8. Lead-one forecast of the gas furnace model. (Solid: Actual output. Dashed: Lead-one forecast).

TABLE II
COMPARISON OF OUR MODEL WITH OTHER MODELS

Model Name	Inputs	Number of Rules	Mean Squared Error
Tong's model [9]	$y(t-1)$ $u(t-4)$	19	0.469
Pedrycz's model [13]	$y(t-1)$ $u(t-4)$	81	0.320
Xu's model [14]	$y(t-1)$ $u(t-4)$	25	0.328
Box's model [30]	$y(t-1)$ $y(t-2)$ $u(t-3)$ $u(t-4)$ $u(t-5)$	---	0.202
Sugeno's model [21]	$y(t-1)$ $y(t-2)$ $y(t-3)$ $u(t-1)$ $u(t-2)$ $u(t-3)$	2	0.068
Sugeno's model [22]	$y(t-1)$ $u(t-3)$ $u(t-4)$	6	0.190
Wang's model [24]	$y(t-1)$ $u(t-4)$	5	0.158
Our model (FCA)	$y(t-1)$ $u(t-4)$	4	0.146

employing the proposed FCA method is less than that of other modeling approaches.

B. Box and Jenkin's Yields from a Batch Chemical Process

The series F in [30] is also considered as a typical demonstration example. It is a time series of 70 consecutive yields from a batch chemical process. An AR(2) model has been identified for this time series by Box and Jenkins. Two state space models were presented in [31] and [32], known as SP1 and SP2, respectively. There was also a recently proposed fuzzy model developed by Wang and Langari [24].

It is well known that this time series has a high level of stochasticity which increases the modeling difficulty. However, using the proposed FCA approach can effectively eliminate this problem. One of the advantages of using the FCA modeling method is that the modeling accuracy can be assigned as required. Modeling accuracies of most other modeling techniques can only be obtained after the modeling processes are fully accomplished.

TABLE III
FUZZY MODEL OF YIELDS FROM THE BATCH CHEMICAL PROCESS

Formed hypercrystal H_1, $P^{\{H_1\}} = \{(\leftrightarrow),(\leftarrow)\}$			
	x_1	x_2	y
R_1	$\bar{S}(23, 44.82, 80)$	$\bar{S}(23, 74.11, 80)$	67.05
R_2	$\bar{S}(23, 44.82, 80)$	$S(23, 26.92, 80)$	52.12
R_3	$S(23, 56.31, 80)$	$\bar{S}(23, 74.11, 80)$	51.02
R_4	$S(23, 56.31, 80)$	$S(23, 26.92, 80)$	35.46

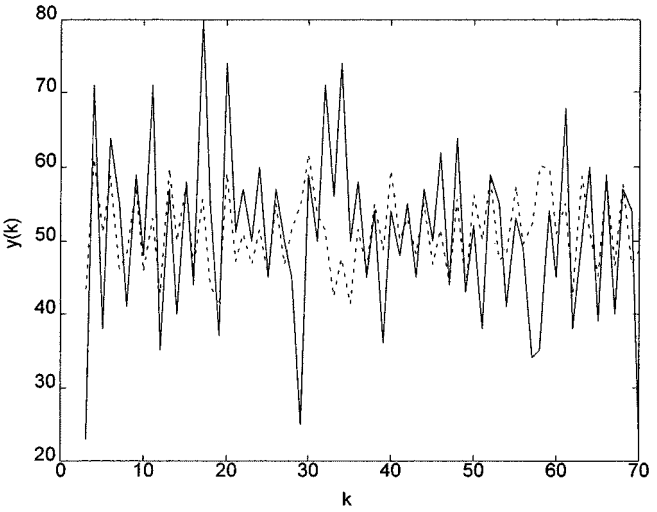


Fig. 9. Lead-one forecast of the model of yields for the batch chemical process. (Solid: Actual output. Dashed: Lead-one forecast).

TABLE IV
ERROR COMPARISON IN MODELING SERIES F

	AR2	SP1	SP2	Fuzzy model	Our model n* = 1	Our model n = 5
Mean squared error	113.80	114.44	114.30	96.75	114.30	72.25

* n : number of formed hypercrystals

A fuzzy model of one formed hypercrystal, identified by the proposed FCA approach, is shown in Table III. Satisfactory modeling performance is demonstrated by the lead one forecast in comparison with the actual process illustrated in Fig. 9. Table IV summarizes mean squared errors (MSEs) for Box and Jenkins' AR(2) model, Aoki's SP1 model, Libert, Wang, and Liu's SP2 model, Wang and Langari's fuzzy model, and our models with different numbers of hypercrystals. It is also confirmed that the more accuracy is desired for the model, the more formed hypercrystals will be obtained by using the FCA.

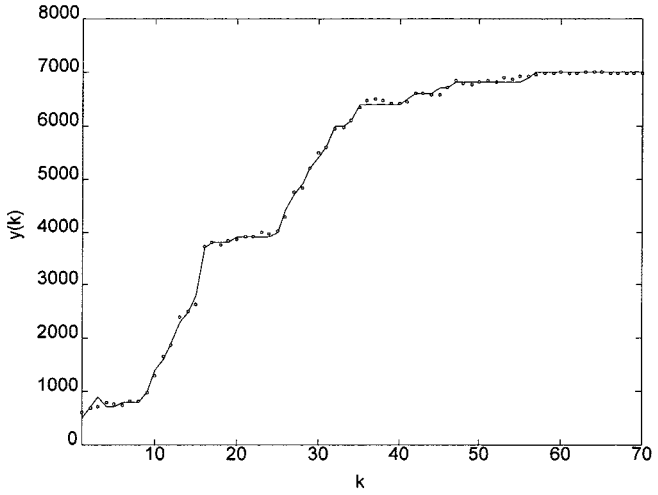


Fig. 10. Human operation model for a chemical plant. (Solid: Actual control action of the human operator. Dashed: Simulated control action).

C. Sugeno and Yasukawa's Human Operation at a Chemical Plant

A chemical plant produces a polymer by the polymerization of some monomers. Since the startup of the plant is very complicated, a human operator has to manually operate the system. The purpose of this example is to show that the proposed modeling method has the ability to build a model of a human operator's control actions.

The strategy of the human operation is explained in [22]. There are five inputs to which a human operator may refer. They include the following.

1) Monomer concentration u_1.
2) Change of monomer concentration u_2.
3) Monomer flow rate u_3.
4) Local temperatures inside the plant u_4 and u_5.

The output y is the set point for monomer flow rate.

There are 70 data points for each of the above six variables (five inputs and one output) obtained from [21]. The data came from an operator's actions in determining the set point for the monomer flow rate, which will be fed into the plant controlled by a PID controller. u_1, u_2, and u_3 are selected as the input variables for the FCA method. Consequently, through the proposed modeling method, the resulting fuzzy model consists of two formed hypercrystals and the performance of our model is demonstrated in Fig. 10.

V. Conclusion

A new approach for complex systems modeling, the FCA, has been developed. As all the formed hypercrystals are obtained independently, and the space for parameter identification is small for the formation of each hypercrystal, the proposed method has the advantage of high processing efficiency. Furthermore, this unique modeling strategy, simulating the nature of material crystallization, provides a strong autonomous modeling feature. The required user intervention is, therefore, very limited. Besides, the FCA method can guarantee the accuracy of the model to be within a predetermined tolerance. Remarkable modeling performance is successfully demonstrated by a number of numerical benchmark examples.

Traditionally, from Yager and Filev's point of view, "In general, structure identification is a difficult and extremely ill-defined process, more an art than science, and not readily amenable to automated techniques [7]." However, the proposed method in this paper provides a novel fuzzy modeling technique which incorporates a built-in automated structure identification scheme.

References

[1] R. R. Yager and D. P. Filev, *Essentials of Fuzzy Modeling and Control*. New York: Wiley, 1994.
[2] L. A. Zadeh, "Outline of a new approach to the analysis of complex systems and decision processes," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, pp. 28–44, Jan. 1973.
[3] E. H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *Int. J. Man–Mach. Stud.*, vol. 7, no. 1, pp. 1–13, 1975.
[4] E. H. Mamdani, "Application of fuzzy algorithms for control of simple dynamic plant," *Proc. Inst. Elect. Eng.*, vol. 121, no. 12, pp. 1585–1588, 1974.
[5] E. H. Mamdani and N. Baaklini, "Prescriptive method for deriving control policy in a fuzzy logic control," *Electron. Lett.*, vol. 11, no. 25/26, pp. 625–626, 1975.
[6] L. P. Holmblad and J.-J. Ostergaard, "Control of a cement kiln by fuzzy logic," in *Fuzzy Information and Decision Processes*, M. M. Gupta and E. Sanchez, Eds. Amsterdam, The Netherlands: North-Holland, 1982, pp. 389–399.
[7] R. R. Yager and D. P. Filev, "Template based fuzzy systems modeling," *J. Intell. Syst.*, vol. 2, pp. 39–54, 1994.
[8] R. M. Tong, "Synthesis of fuzzy models for industrial processes—Some recent results," *Int. J. Gen. Syst.*, vol. 4, no. 3, pp. 143–163, 1978.
[9] ——, "The construction and evaluation of fuzzy models," in *Advances in Fuzzy Set Theory and Applications*, M. M. Gupta, R. K. Ragade, and R. R. Yager, Eds. Amsterdam, The Netherlands: North-Holland, 1979, pp. 559–576.
[10] B. Kosko, *Neural Networks and Fuzzy Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1992.
[11] ——, "Fuzzy associative memory systems," in *Fuzzy Expert Systems*, A. Kandel, Ed. Boca Raton, FL: CRC, 1991, pp. 135–162.
[12] W. Pedrycz, "Identification in fuzzy systems," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-14, pp. 361–366, Mar./Apr. 1984.
[13] ——, "An identification algorithm in fuzzy relational systems," *Fuzzy Sets Syst.*, vol. 13, no. 2, pp. 153–167, 1984.
[14] C. W. Xu and Y. Z. Lu, "Fuzzy model identification and self-learning for dynamic systems," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-17, pp. 683–689, July/Aug. 1987.
[15] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York: Plenum, 1981.
[16] E. H. Ruspini, "Recent developments in fuzzy clustering," in *Fuzzy Set and Possibility Theory*, R. R. Yager, Ed. New York: Pergamon, 1982, pp. 133–147.
[17] D. Dubois, H. Prade, and C. Testemale, "Weighted fuzzy pattern matching," *Fuzzy Sets Syst.*, vol. 28, no. 3, pp. 313–331, 1988.
[18] M. Sugeno and G. T. Kang, "Structure identification of fuzzy model," *Fuzzy Sets Syst.*, vol. 28, no. 1, pp. 15–33, 1988.
[19] R. R. Yager and D. P. Filev, "Approximate clustering via the mountain method," *IEEE Trans. Syst., Man, Cybern.*, vol. 24, pp. 1279–1284, Aug. 1994.
[20] ——, "Unified structure and parameter identification of fuzzy models," *IEEE Trans. Syst., Man, Cybern.*, vol. 23, pp. 1198–1205, July/Aug. 1993.
[21] M. Sugeno and K. Tanaka, "Successive identification of a fuzzy model and its applications to prediction of a complex system," *Fuzzy Sets Syst.*, vol. 42, no. 3, pp. 315–334, 1991.
[22] M. Sugeno and T. Yasukawa, "A fuzzy-logic-based approach to qualitative modeling," *IEEE Trans. Fuzzy Syst.*, vol. 1, pp. 7–31, Feb. 1993.
[23] D. P. Filev, "Fuzzy modeling of complex systems," *Int. J. Approx. Reas.*, vol. 5, no. 3, pp. 281–290, 1991.
[24] L. Wang and R. Langari, "Complex systems modeling via fuzzy logic," *IEEE Trans. Syst., Man, Cybern. B*, vol. 26, pp. 100–106, Feb. 1996.

[25] D. A. Brandt, *Metallurgy Fundamentals*. South Holland, IL: Goodheart-Willcox, 1985, pp. 122–131.
[26] P. Haasen, *Physical Metallurgy*, 3rd ed. Cambridge, U.K.: Cambridge Univ. Press, 1996, pp. 60–64.
[27] L. A. Zadeh *et al.*, *Fuzzy Sets and Their Application to Cognitive and Decision Processes*. New York: Academic, 1975.
[28] R. R. Yager, "On a general class of fuzzy connectives," *Fuzzy Sets Syst.*, vol. 4, no. 3, pp. 235–242, 1980.
[29] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Comput. J.*, vol. 7, pp. 308–313, Jan. 1965.
[30] G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*, 2nd ed. San Francisco, CA: Holden-Day, 1976.
[31] M. Aoki, *State Space Modeling of Time Series*. New York: Springer-Verlag, 1983.
[32] G. Libert, L. Wang, and B. Liu, "An innovation state space approach for time series forecasting," *J. Time Series Anal.*, vol. 14, no. 6, pp. 589–601, 1993.

Chi-Cheng Cheng (M'87) was born in Taipei, Taiwan, R.O.C., in 1959. He received the B.S. and M.S. degrees in power mechanical engineering from National Tsing Hua University, Hsinchu, Taiwan, in 1981 and 1983, respectively, and the Sc.D. degree in mechanical engineering from Massachusetts Institute of Technology, Cambridge, MA, in 1991.

He is currently an Associate Professor with the Department of Mechanical and Electro-Mechanical Engineering, National Sun Yat-Sen University, Kaohsiung, Taiwan, R.O.C. His research interests are in the areas of system dynamics and control, visual servo, fuzzy modeling and control, and telerobotics.

Wen-Hsiung Hsieh was born in Taiwan, R.O.C., in 1972. He received the B.S. and M.S. degrees from National Sun Yat-Sen University, Kaohsiung, Taiwan, R.O.C., both in mechanical engineering, in 1994 and 1996, respectively.

Since 1998, he has been an Engineer with Alcatel Taiwan, Kaohsiung.