Extended Fuzzy Clustering Algorithm Based on an Inclusion Concept

S.Nefti and K. Djouani

Salfort University, Computer Science Department
Newton Building Manchester, United Kingdom
Email s.nefti-meziani@salford.ac.uk

University of Paris Val de Marne
Computer Science and Robotics Lab – LIIA, France
Email djouani@univ-paris12.fr

***Abstract*-Fuzzy modeling of complex systems is a challenging topic. This paper proposes an effective approach to data-based fuzzy optimizing fuzzy system structure and parameters. For this purpose, we cope with fuzzy clustering based on inclusion concept where the rule-base has to be simplified. This simplification occurs in the sense that similar Membership Functions (MF) pertaining to the premise of fuzzy rule-base are merged and replaced by one common MF, capturing the meaning of the former. Reduction of the total number of fuzzy sets improves semantic interpretation and reduces the demand on memory in implementation context. So, we propose an extended algorithm based on the class of fuzzy clustering method and on an inclusion concept proposed by Nefti and al [11], which is characterized by an inclusion index. During the optimization, the redundant rules are deleted. Finally, interpretability of the fuzzy system is improved. To show the effectiveness of the proposed algorithm, a comparative study of the obtained simulation results with a conventional algorithm based on the class of fuzzy C-means method introduced by Bezdek FCM is presented by a numerical example, which computes a MISO architecture.**

Keywords: **Fuzzy systems, Clustering, Inclusion index, Complexity reduction.**

I. Introduction

The number of parameters is one of the main concerns for fuzzy systems control, especially when it is desired to increase the number of inputs and rules, since for standard fuzzy system the number of parameters increases when the number of inputs or rules is increased, and computational complexity increases accordingly. Thus, the rule-base will suffer from redundancy and conflicts of data, most of which are less useful. This redundancy is often present in the form of similar membership functions MF in the premise of the resulting rule-base. Such similarity within fuzzy sets render difficult to attach qualitatively meaningful linguistic labels to the different MF. The high number of MF makes difficult to obtain the meaning of the model, and thus the working of system at hand. A semantically unclear model is not easily verified after design phase for the model. Consequently, a simplification phase allowing the elimination of redundancy is required. Different approaches using fuzzy clustering algorithms have been proposed to solve the rule explosion problem. In the earlier works, rule reduction in fuzzy systems have been, mainly, attempted via a variety of clustering techniques [1][7], in an effort to select only those rules that contribute the most to the inference outcome. For this purpose, we cope with fuzzy clustering [14] based on inclusion concept [11] where the rule-base has to be simplified. This simplification occurs in the sense that similar MF pertaining to the premise of fuzzy rule-base are merged and replaced by one common MF, capturing the meaning of the former. Reduction of the total number of fuzzy sets improves semantic interpretation and reduces the demand on memory in implementation context. So, we propose an algorithm based on the class of fuzzy clustering method introduced by Bezdek [3] and on inclusion concept introduced by Nefti [11], which is characterized by an inclusion index Id [2][4][6]. In what follows, a Multi-Inputs/Single-Output (MISO) fuzzy system (FS) architecture is presented and the influence of the parameters number is discussed in Section 2 and 3, followed by presentation of the proposed fuzzy c-means algorithm for the optimization in Section 4. In Section 5, a comparative study of the proposed algorithm and the conventional FCM [3] is presented to show the effectiveness of the approach. Finally, the conclusion is presented in Section 6.

II. Fuzzy System Architecture and Parameters Size

In this section, the architecture of a FS is presented to illustrate the effects of the parameters number. The architecture illustrated on Fig.1 represents a single MISO module, which is considered as a FS built from four-layer feed-forward network. Each fuzzy module takes the antecedents X_k at its inputs and produces a new action y_k, where k = 1,.., n, with n, number of inputs.

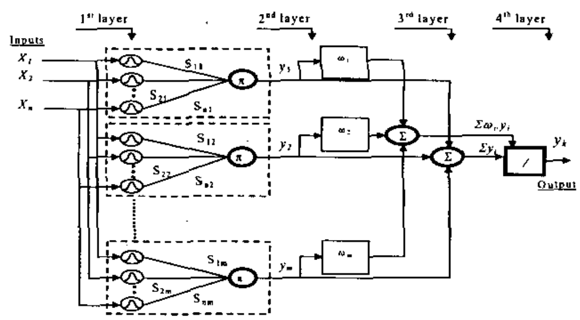


Fig. 1. MISO architecture of the fuzzy system.

As in ANFIS model, this architecture implements rules of the following form [15] :

R^j : *If* X_1 is $A_{1j}^{(1)}$ *and .. and* X_1 is $A_{nj}^{(j)}$ then $y_k = f(X_1, .., X_n)$

Where X_i and y_k represent respectively the input and the output variables, and $A^{(j)}_{ij}$ the fuzzy sets.

The FS design passes by the partitioning of each input/output variable space in several fuzzy sub-sets (grid or free partition). Firstly, the Gaussian MF are chosen and a free partition is applied to the universe of discourse. From some observations of X_i, the fuzzy inference consequence y_k obtained by simplified fuzzy reasoning method. The membership functions for each input variable, X_i, can be written as follows:

$$A_{ij}^{(k)} = S_{ij}^{(k)}(X_i) = exp\left(-\frac{1}{2}\left((X_i - m_{ij})/\sigma_{ij}\right)^2\right) \quad (1)$$

The inputs/output relationships are given by [14]:

$$y_k = \frac{\sum_{i=1}^{n}\left(w_j \cdot \prod_{j=1}^{k} S_{ij}^{(k)}(X_i)\right)}{\sum_{i=1}^{n}\left(\prod_{j=1}^{k} S_{ij}^{(k)}(X_i)\right)} \quad (2)$$

where ΔF_i is the input variable, m_{ij} and σ_{ij} the i^{th} mean and standard deviation respectively of the j^{th} rule.

When defining a FS, one can either use some of the input MF or all possible combinations of the input MF to construct the rule base. In this case, the number of parameters M_f grows exponentially by the growth in the number of inputs or number of rules. From (2) and Fig.1, the number of parameters, which can be, tuned is:

$$M_f = (l.n + 1).R \quad (3)$$

Where l is the number of MF parameters, n is the number of inputs, R is the number of MF for each input variable, which is equal, to the number of consequent parts. In this case and according to Fig. 1, the rules number R can be represented as:

$$R = \prod_{j=1}^{n} N_j \quad (4)$$

Where N_j is the number of MF in the j^{th} universe of discourse. Clearly, for either large n or R, M_f can be very large and there is an exponential increase in the number of parameters for additional inputs leading to the curse of dimensionality. This is usual the case in fuzzy control applications [12]. In the next section, we will focus on how to reduce the number of parameters needed to define the FS.

III. BACKGROUND

Let $P_{ij} = [m_{ij}, \sigma_{ij}]$ be the MF parameters associated to input variable x_i obtained after the parameter learning phase. The aim of the fuzzy clustering algorithm is to determine an optimal clusters set $\{v_k\}$, where $v_k = [m_k, \sigma_k]$, in order to replace the old fuzzy partition $\{P_{ij}\}$ by the new one $\{v_k\}$ according to the minimization of the following objective function [3]:

$$J_i^{(fcm)} = \sum_{i=1}^{M}\sum_{k=1}^{c} (\mu_{ki})^m . d(P_{ij}, v_k)^2 \quad (5)$$

With respect to constraints (6) and (7) :

$$\sum_{k=1}^{c} \mu_{ki} = 1, \forall i=1, ., n \quad (6)$$

$$\sum_{i=1}^{n} \mu_{ki} > 0, \forall k=1, ., c \quad (7)$$

Where M, c and m are respectively the number of data i=(1,.., M), the number of clusters initialized randomly and a weighting exponent which determines the fuzziness of the clusters ($m \in [1,+\infty[$). $d(P_{ij},v_k)$ represents the distance measure between P_{ij} and the prototype v_k, defined as:

$$d(P_{ij},v_k)^2 = (P_{ij} - v_k)^t . A_j . (P_{ij} - v_k) \quad (8)$$

Where A_j is a positive defined symmetric matrix. The choice of the matrix A_j induces a proper kind of distance interpretation, and consequently generates its own meaning of cluster shape. For instance, if A_j is the identity matrix, $d(P_{ij},v_k)$ corresponds to Euclidean distance and, roughly, it induces spherical clusters. The variants of objective-based fuzzy clustering are numerous. These generalizations deal with various shapes of clusters. Gustafson and Kessel [6] have focused on the case where the matrix A_j is fixed for each cluster j, while the determinant of each matrix A_j, which stands for the volume of the cluster, is globally preserved. This allows to detect more sophisticated shape clusters. Bezdek [2] has investigated the case where one of the eigenvectors of the matrix A_j, are maximized. This allows the detection of linear clusters like lines or hyper-plans. Dave [4] has investigated special formulation of the objective function $J_i^{(fcm)}$ that yields a better description of circular shape. Also, Krishnapuram and Keller [10] investigated another formulation of $J_i^{(fcm)}$ such that the membership matrix is seen as the distance from ideal prototype instead of being a degree of share of the unit value between all existing clusters, as the initial formulation does. This means that both the formulation of the matrix A_j and the objective function are not completely fixed and some flexibility is allowed. Kaymak [9] has studied the influence of the clusters number in the data and the problem of initialization, which must be appropriately for convergence to acceptable solutions and for finding all interesting clusters including the one small in size. For this purpose, he proposed an extended approach of fuzzy c-means to determinate an appropriate number of clusters by considering a volume prototype. In the present study, we are interested in a particular problem where some data are collapsed into others. The aim is to solve this collapse problem under another aspect as inclusion concept. The idea developed in this paper is somewhat similar to search for an inclusion concept hidden in the distance one. This allows for taking account for the inclusion only in global sense. A rational criteria that take account for an inclusion concept is used. In the second part, is to model how the distributions are included each other.

IV. EXTENDED FUZZY CLUSTERING ALGORITHM

A. *Inclusion Index Construction*

Let us denote by G the Gaussian distribution characterized by $[m,\sigma]$. In general, 97% of the information supplied by the distribution is concentrated into the interval $[m\text{-}3\sigma , m\text{+}3\sigma]$. $I_d(G_{ij},G_k)$ stands for the degree of inclusion of the Gaussian (m_{ij},σ_{ij}) in (m_k,σ_k), and characterized by the surface formed by the intersection of G_{ij} and G_k, $(G_{ij} \cap G_k)$ (Fig. 2).

Clearly, asserting that Id should depend on the following parameters, $(m_{ij}\text{-}m_k)$ and $3.(\sigma_{ij}+\sigma_k)$.

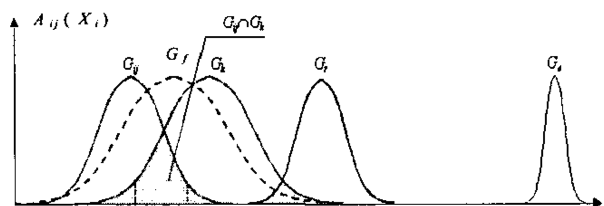


Fig. 2. Fuzzy set overlapping.

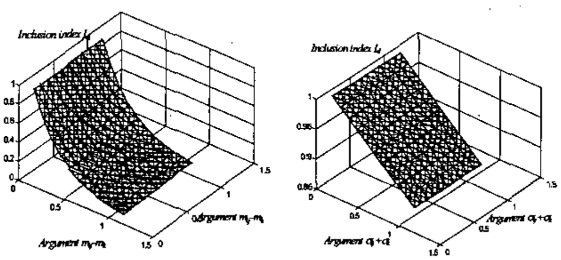
One can show that the inclusion index is given by [11][14]:

$$
I_d(G_{ij},G_k)=\begin{cases} 1-\left|m_{ij}-m_k\right|\cdot\exp\left(-3\cdot\left(\sigma_{ij}+\sigma_k\right)\right) & \text{si} \quad \sigma_{ij}\le\sigma_k \\ 0 & \text{si} \quad \sigma_{ij}>\sigma_k \end{cases} \qquad (9)
$$

It's easily checked that $I_d(G_{ij},G_k)$ is non-increasing with respect to $(m_k - m_{ij})$ (Fig.3 a) and to $3.(\sigma_k+\sigma_{ij})$ (Fig.3 b), one may determine its derivative with respect to $(m_k - m_{ij})$:

$$
\frac{\partial I_d}{\partial\left|m_k-m_{ij}\right|}=-\exp\left(-3\left(\sigma_{ij}+\sigma_k\right)\right) \le 0 \quad \forall\sigma_{ij} \ and \ \forall\sigma_k
$$

On the other hand, the factor 1 in $1-\left|m_{ij}-m_k\right|\cdot\exp\left(-3\cdot\left(\sigma_{ij}+\sigma_k\right)\right)$ allows to preserve the positive ness of the value of $I_d(G_{ij},G_k)$. The second term of I_d $(I_d(G_{ij},G_k) = 0)$ is used to avoid the inverse inclusion of the fuzzy sets.



(a) (b)

Fig. 3. Inclusion index evolution.

B. Formulation of the extended Fuzzy Clustering Algorithm

The new clustering algorithm is an extension of the proposed algorithm [11] where the distance concept defined by (10) can be considered as an inclusion concept [11][14]. It can be formulated as follows:

$$
\sum_{k=1}^{c} [d(P_{ij}, v_k)]^2 \cong \sum_{k=1}^{c} [I_d(G_{ij},G_k)]^2 \qquad (10)
$$

$\forall$ i$\in$[1,n] and j$\in$[1,M], from (9) and (10), we have :

$$
\sum_{k=1}^{c} (P_{ij}-v_k)^t A_j (P_{ij}-v_k) = \sum_{k=1}^{c} (1-|m_{ij}-m_k|.\ e^{(-3.(\sigma ij+\sigma k))})^2 \qquad (11)
$$

Consider a new representation of the Gaussian parameters, so that, Pij(1) = mij , Pij(2) = σij , vk(1) = mij and vk (2) = σij, and a new quadratic formulation of (11). Let us denote by B1 and B2 two matrix, where $B_1=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$ and $B_2=\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}$. Then, (11) can be formulated as follows:

$$
\sum_{k=1}^{c} (P_{ij}-v_k)^t.A_j.(P_{ij}-v_k) -
$$

$$
\sum_{k=1}^{c} (1-2.(P_{ij}^{(1)}-v_k^{(1)}).\ e(-3.(P_{ij}^{(2)}+v_k^{(2)})^t.B_2) -
$$

$$
\sum_{k=1}^{c} |P_{ij}^{(1)}-v_k^{(1)}|^t.B_1.|P_{ij}^{(1)}-v_k^{(1)}|) \times e(-6.(P_{ij}^{(2)}+v_k^{(2)})^t.B_2) = 0 \qquad (12)
$$

Leading to an equality constraint to the optimization problem, so the non-linear optimization problem (P_r) can be written as:

$$
P_r: \begin{cases} \text{Minimize } J_i^{(fcm)}(\mu_{ki},A_j,v_k), \text{under constraint:} \\ g_r(\mu_{ki},A_j,v_k)=0, \forall r\in I=(1,..,n) \end{cases} \qquad (13)
$$

In our case, two constraints are considered (r=1,2):

$$
g_1(\mu_{ki}): \sum_{k=1}^{c} \mu_{ki} - 1 = 0 \ , \forall i=1, 2, .., n \qquad (14)
$$

$$
g_2(\mu_{ki},A_j,v_k): \sum_{k=1}^{c} (P_{ij}-v_k)^t.A_j.(P_{ij}-v_k) -
$$

$$
\sum_{k=1}^{c} (1-2.(P_{ij}^{(1)}-v_k^{(1)}).e(-3.(P_{ij}^{(2)}+v_k^{(2)})^t.B_2) -
$$

$$
\sum_{k=1}^{c} |P_{ij}^{(1)}-v_k^{(1)}|^t.B_1.|P_{ij}^{(1)}-v_k^{(1)}|)\times e(-6.(P_{ij}^{(2)}+v_k^{(2)})^t.B_2)=0 \qquad (15)
$$

The second constraint g_2 can be considered as a bridge between the inclusion concept and the distance concept, which is necessary for the fuzzy clustering formulation. Using a dual methods, the (P_r) is transformed into a min-max problem by introducing the Lagrange parameter vector λ_r, so that:

$$
L(\mu_{ki},A_j,v_k,\lambda) = J_i^{(fcm)}(\mu_{ki},A_j,v_k) + \sum_{r=1}^{2} \lambda_r^T.g_r(\mu_{ki},A_j,v_k) \qquad (16)
$$

The optimal parameters are obtained by setting the derivative of Lagrangian according to each of its parameters $P_l = [\mu_{ki}, A_j, v_k, \lambda_1, \lambda_2]$ to zero, so that:

$$
\frac{\partial L(\mu_{ki},A_j,v_k,\lambda_1,\lambda_2)}{\partial P_l}=0 \qquad (17)
$$

For details on calculation refer to [5].

1) *Extended fuzzy C-means algorithm steps*: The proposed fuzzy clustering algorithm can be summarized by the following steps :

Given a data P_{ij}, choose the initial number of clusters $1<c^{(0)}<M$, the fuzziness parameter m (m=2) and the termination criterion $\xi>0$. Initialize $U^{(0)}=[\mu_{ki}^{(0)}]$ (eg. Random), and let $\alpha_1=1.05$ and $\alpha_2=0.95$.

Loop

1. *Compute a new clusters centers v_k:*

$$
v_k=\frac{\sum_{i=1}^{n}(\mu_{ki})^m.P_{ij}}{\sum_{i=1}^{n}(\mu_{ki})^m} \quad , \forall \text{ k=1, .., c}
$$

2. *Compute the matrix A_j, so that:*

$$
A_j=\frac{\sum_{k=1}^{c} 1-2.(P_{ij}-v_k)\left|P_{ij}^{(1)}-v_k^{(1)}\right|^t B_1\left|P_{ij}^{(1)}-v_k^{(1)}\right|.\exp\left(-6.\left(P_{ij}^{(2)}-v_k^{(2)}\right)\right)B_2}{\sum_{k=1}^{c}\left(P_{ij}-v_k\right)^t\left(P_{ij}-v_k\right)}
$$

3. *Update the partition matrix*:

$$\mu_{ki} = \frac{1}{\left(P_{ij}-v_k\right)^t A_j \left(P_{ij}-v_k\right)^{\frac{2}{m-1}} \cdot \sum_{i=1}^{n} \frac{1}{\left(P_{ij}-v_k\right)^t A_j \left(P_{ij}-v_k\right)^{\frac{2}{m-1}}}}$$

4. *Compute the criteria* $J_i^{(fcm)}$, *so that* :

$$J_i^{(fcm)} = \sum_{i=1}^{n}\sum_{k=1}^{c} \left((\mu_{ki})^m . (P_{ij}-v_k)^t . A_j . (P_{ij}-v_k)\right)$$

If $J_i^{(fcm)} \leq \xi$ *go to step* 5.
else go to step 1.

5. *Verify the rule-base property*:

$$\text{if } \sum_{k=1}^{c}(\mu_{ki}) < \alpha_1 \text{ , } \forall \text{ i=1, .., M}$$

c = c − 1 , *and go to step* 1.

$$\text{else if } \sum_{k=1}^{c}(\mu_{ki}) > \alpha_2 \text{ , } \forall \text{ i=1, .., M}$$

c = c + 1 , *and go to step* 1.
else STOP

End of Loop *(STOP)*

The algorithm proposed above is an extension of the fuzzy inclusion algorithm proposed by Nefti [11] and based on classical FCM, where the relationship between the cluster prototype and each datum is rather described in terms of inclusion relation instead of reasoning in terms of distance. The key issue in the proposal is to let the matrix A_j be fixed only for a given cluster. While its determination assumes that, for a given cluster, the amount of distances from the prototype to each datum is the same as the amount of inclusions of each datum into that cluster prototype according to the elaborated inclusion index. This method, as it is the case for all other fuzzy clustering algorithms, is an optimization-based approach.

V. Numerical Example

In order to illustrate the validity of the proposed algorithm, we consider a MISO FS architecture presented in Fig.1, with two inputs (X_1, X_2) and one output y_k. The set of MF associated to each input variable includes 8 MF, randomly initialized, and four initial clusters number. By ensuring a parameters learning algorithm, an optimal structure of the FS is generated. The contribution of each input variable in the overall fuzzy rules is represented by 8 fuzzy sets. Due to the MF overlapping, the proposed fuzzy c-means algorithm is used to reduce the MF number and optimize the fuzzy rule base. Fig.4 and Fig.5 give the optimal clusters within initial MF obtained after the application of the proposed algorithm (E-FCM) and the conventional one (C-FCM).

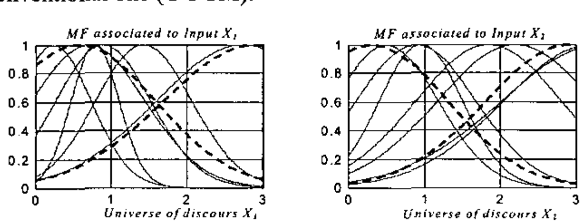


Fig. 4. Generated clusters with E-FCM algorithm.

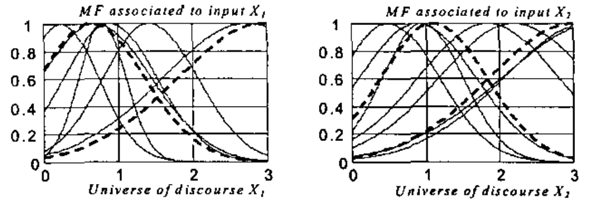


Fig. 5. Generated clusters with C-FCM algorithm.

According to the form of the rule presented above, the generated rules R^i are given by:

R^1 : **If** X_1 is (1.1387 , 0.5490) **and** X_2 is (0.0027 , 0.0009) **then** y_k is 0.1585

R^2 : **If** X_1 is (0.6365 , 1.0305) **and** X_2 is (0.0014 , 0.0022) **then** y_k is 0.0074

R^3 : **If** X_1 is (0.5021 , 1.1745) **and** X_2 is (0.0002 , 0.0157) **then** y_k is 0.0997

R^4 : **If** X_1 is (0.3214 , 1.1612) **and** X_2 is (0.0008 , 0.0032) **then** y_k is 0.4907

R^5 : **If** X_1 is (25.7626 , 3117.1) **and** X_2 is (3.1084 , 0.0033) **then** y_k is 0.1926

R^6 : **If** X_1 is (450.5010 , 0.9603) **and** X_2 is (0.0009 , 0.0018) **then** y_k is 0.5173

R^7 : **If** X_1 is (0.7556 , 0.6791) **and** X_2 is (0.0007 , 0.0005) **then** y_k is 0.1436

R^8 : **If** X_1 is (4.4532 , 0.7902) **and** X_2 is (0.3501 , 0.0009) **then** y_k is 0.8432

As an illustrations, Fig.6 and Fig.7 respectively show a plane distribution of the optimizing datum using the proposed clustering algorithm E-FCM and the conventional one C-FCM, where (•) represents the cluster center and (+) represents the optimized datum.

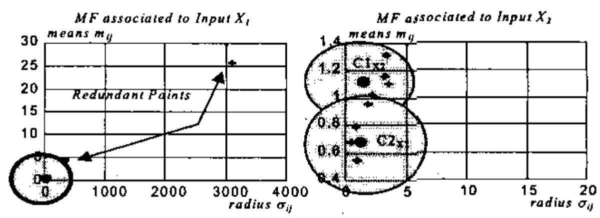


Fig. 6. MF plane distribution with E-FCM algorithm.

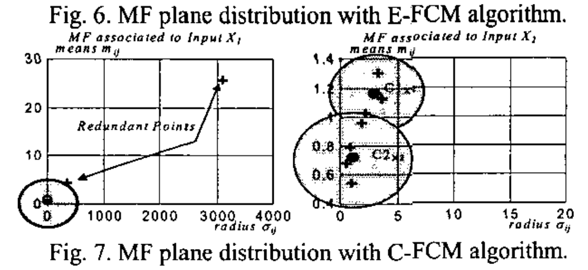


Fig. 7. MF plane distribution with C-FCM algorithm.

This representation gives a better idea about the clusters number to optimize. As we can we see, on Fig.6 and Fig.7, there exit two redundant points in the numerous optimizing data, which can be regarded as an inconsistency with the trends of the FS (Fig.1). These points are removed from the original data, because there does not exist a cluster center around any redundant datum. The circles define the clusters. They include fuzzy sub-sets, which have the highest membership degrees to the considered clusters centers. Thus, we see that the application of the inclusion based clustering algorithm and the conventional FCM give, for each input variable, four classes of MF that ensure the initial optimization problem. Such classes are given by:

<u>Clusters generated with E-FCM algorithm</u>

For X_1 : $C1_{X1} = [0.502, 0.225]$ and $C2_{X1} = [1.149, 2.895]$

For X_2 : $C1_{X2} = [1.096, 2.911]$ and $C2_{X2} = [0.690, 1.052]$

<u>Clusters generated with C-FCM algorithm</u>

For X_1 : $C1_{X1} = [0.72, 0.54]$ and $C2_{X1} = [1.15, 2.95]$

For X_2 : $C1_{X2} = [1.16, 3.01]$ and $C2_{X2} = [0.72, 1.10]$

The learning error during clustering for E-FCM and C-FCM algorithms, associated for each input variable of the FS, is shown in Fig.8 and Fig.9.

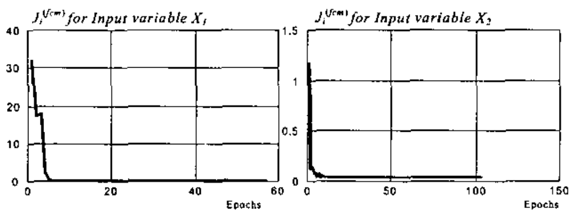


Fig. 8. Clustering error with E-FCM algorithm.

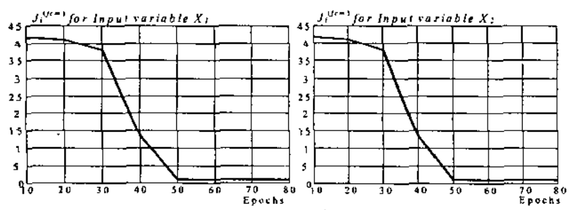


Fig. 9. Clustering error with C-FCM algorithm

As shown on Fig.8, using E-FCM algorithm, 02 clusters are generated after 5 iterations for the input variable X_1 and 10 iterations for the second one X_2, with convergence error respectively 0.0024 and 0.0040. In the case of the conventional FCM algorithm, starting with four initial clusters, 02 classes are created after 50 iterations for each input variable X_1 and X_2. The convergence error is 0.85 for the input variable X_1 and 0.911 for the second one (Fig.9).

Tab.1 and Tab.2 illustrate respectively the generated matrix $U=[\mu_{ki}]$, for each input variable, using the proposed algorithm E-FCM and the C-FCM algorithm. As we can see, only six fuzzy sub-sets are represented for the input variable X_1 because two redundant points are removed from the optimizing datum. These tables, ensure the inclusion of the initial MF set in the clusters set. For example, in Tab.1, the fuzzy sub-set $SF_1^{(4)}$ is closer to the first cluster, which is given by the highest degree (0.9992), than to the second one. For the second input variable X_2, the fuzzy sub-set $SF_2^{(5)}$, which is closer to the second cluster with the highest degree (0.9498), than the first one. Using the conventional clustering algorithm (Tab.2), the fuzzy sub-set $SF_1^{(6)}$ is closer to the second cluster, which is given by the highest degree (0.9681), than to the first one. For second input variable X_2, the fuzzy sub-set $SF_2^{(6)}$, which is closer to the second cluster with the highest degree (0.9999), than the first one. As we can see in Tab.2, the fuzzy rule-base consistency and completude property are not respected using the conventional FCM algorithm. As a matter of fact, the constraints defined in (6) and (7) are not assured and the obtained results, for example in the fuzzy sub-set $SF_1^{(3)}$, corresponding to the first input X_1, approve it.

TABLE I

GENERATED PARTITIONS WITH E-FCM ALGORITHM

Input variable X_1

	$\mu_{1i}(X_1)$	$\mu_{2i}(X_1)$
$SF_1^{(1)}$	1.0000	0.0000
$SF_1^{(2)}$	0.9996	0.0004
$SF_1^{(3)}$	0.9994	0.0006
$SF_1^{(4)}$	0.9992	0.0008
$SF_1^{(5)}$	0.0000	1.0000
$SF_1^{(6)}$	0.9996	0.0004

Input variable X_2

	$\mu_{1i}(X_2)$	$\mu_{2i}(X_2)$
$SF_2^{(1)}$	0.9539	0.0461
$SF_2^{(2)}$	0.4441	0.5559
$SF_2^{(3)}$	0.0167	0.9833
$SF_2^{(4)}$	0.0376	0.9624
$SF_2^{(5)}$	0.0502	0.9498
$SF_2^{(6)}$	0.7043	0.2957
$SF_2^{(7)}$	0.9904	0.0096
$SF_2^{(8)}$	0.9932	0.0068

TABLE II

GENERATED PARTITIONS WITH C-FCM ALGORITHM

Input variable X_1

	$\mu_{1i}(X_1)$	$\mu_{2i}(X_1)$
$SF_1^{(1)}$	0.9188	0.0025
$SF_1^{(2)}$	0.9982	0.0001
$SF_1^{(3)}$	0.9662	0.2077
$SF_1^{(4)}$	0.0158	0.8312
$SF_1^{(5)}$	0.0001	0.9951
$SF_1^{(6)}$	0.0017	0.9681

Input variable X_2

	$\mu_{1i}(X_2)$	$\mu_{2i}(X_2)$
$SF_2^{(1)}$	0.0128	0.9852
$SF_2^{(2)}$	0.0027	0.9969
$SF_2^{(3)}$	0.0.65	0.9734
$SF_2^{(4)}$	0.9883	0.0062
$SF_2^{(5)}$	0.3378	0.5070
$SF_2^{(6)}$	0.0021	0.9999
$SF_2^{(7)}$	0.9704	0.0260
$SF_2^{(8)}$	0.0376	0.9584

C. *Interpretability and Fuzzy Labeling*

In this section, we consider only the obtained results with E-FCM algorithm. Once the clusters are generated, the semantic interpretation and reducing the demand on memory in implementation is improved. In this case, labeling is an important phase. It consists of attributing for each generated cluster a linguistic label. Consider S and B, respectively the symbolic values designing Small and Big. If we attribute a label S to the clusters $C1_{X1}$ and $C1_{X2}$ and label B to the clusters $C2_{X1}$ and $C2_{X2}$, we obtain, for each input variable:

- 02 clusters for variable X_1: $C1_{X1} \rightarrow S$ and $C2_{X1} \rightarrow B$

- 02 clusters for variable X_2: $C1_{X2} \rightarrow S$ and $C2_{X2} \rightarrow B$

We note that the same labels attributed to each input variable X_1 and X_2 are not especially identical. For example, the label S associated to X_1 is not the same to the label S associated to X_2. Fig.10 illustrates the fuzzy labels attribution for each generated cluster. As an example, for a set of three fuzzy rules, the new rule Base will be written as follow:

R^1 : **If** X_1 is B **and** X_2 is S **then** y_k is 0.1585

R^2 : **If** X_1 is S **and** X_2 is B **then** y_k is 0.0074

R^3 : **If** X_1 is B **and** X_2 is B **then** y_k is 0.0997

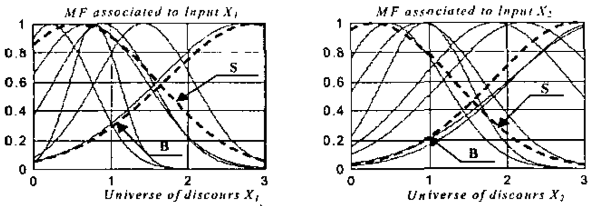


Fig. 10. Fuzzy labels attribution.

Tab.3 shows the performances between the proposed FCM algorithm (E-FCM) and the conventional algorithm (C-FCM) before and after clustering.

TABLE III
CLUSTERING COMPARATIVE RESULTS

		E-FCM algorithm	C-FCM algorithm
Fuzzy sets number	Before clustering	16	16
	After clustering	04	04
Parameters number	Before clustering	40	40
	After clustering	10	10
Number of rules	Before clustering	08	08
	After clustering	02	02
Convergence behavior	X_1	0.0024	0.85
	X_2	0.0040	0.911
Convergence Error in term of number of epochs	X_1	5	50
	X_2	10	50
Initial knowledge		Not required	Required
Parameters initialization (clusters number, centers, ..)		Randomly	Minimum knowledge
Rule-base completude property and consistency		Respected	Not respected

We can see that the proposed fuzzy clustering algorithm gives better performances than the conventional one. Using E-FCM algorithm with random initialization of classes number, 02 clusters are generated after 5 iterations for the input variable X_1 and 10 iterations for the second one X_2, with a convergence error respectively 0.0024 and 0.0040. In the case of the conventional FCM algorithm, starting with four initial clusters number, 02 clusters are created after 50 iterations for each input variable X_1 and X_2. The convergence error is 0.85 for the input variable X_1 and 0.911 for the second on. Thus, the gains values on convergence speed obtained by E-FCM algorithm compared to classical one C-FCM are 10 for input variable X_1 and 5 for the second variable X_2. The convergence error results show, also, the accuracy of E-FCM algorithm compared to C-FCM algorithm. Opposite to conventional fuzzy clustering algorithm, the proposed algorithm doesn't require any initial knowledge on clusters number to be identified and on the distribution of all the optimizing datum only, but also the fuzzy rule-base consistency and completude property are respected.

VI. CONCLUSION

In this paper an extended fuzzy c-means algorithm is proposed. The extension consists of the use of spherical cluster prototypes and considering a new clustering aspect based inclusion concept. An inclusion index, then, has been proposed to model the degree of inclusion between the different fuzzy distributions. The optimization problem is based on some relationship between the inclusion and the distance paradigms that takes account for the inclusion only in global sense. In this way, a compact and interpretable FS can be obtained for complex systems. Through structure optimization, the relationship between the inputs and the output can also be revealed, which is very important for understanding an unknown system. The effectiveness of the proposed algorithm is shown by a numerical example.

VII. ACKNOWLEDGEMENT

The authors would like to thank the reviewers for their useful remarks.

REFERENCES

[1] P.J. Antsaklis, " A brief introduction to the theory and applications of hybrid systems ", *IEEE, Special Issue on Hybrid Systems*: Theory and Applications, vol. 88, July 2000, pp. 879-887.

[2] J.C. Bezdek, C. Coray, and R. Gunderson, " Detection and characterization of cluster structure ", *SIAM J. Appl. Math.*, Part.1, 40 (2), 1981, pp. 339-357.

[3] J. C. Bezdek, " Pattern recognition with fuzzy objective function algorithms ", *Plenum Press*, New York. 1981.

[4] R.N. Dave, " Fuzzy shell-clustering and applications to circle detection in digital images ", *J. of General Systems*, vol. 16, 1990, pp. 343-355.

[5] K. Djouani, Y. Touati and Y. Amirat, " Extended clustering algorithm for fuzzy systems modeling an optimization ", paper under preparation.

[6] D.E. Gustafson and J. Keller, " A possibilistic approach to clustering ", *IEEE Trans. on Fuzzy Systems*, vol. 1, 1993, pp. 98-109.

[7] M.K. Güven and K.M. Passino, " Avoiding exponential parameter growth in fuzzy systems ", *IEEE Trans. on Fuzzy Systems*, vol. 9, Feb. 2001, pp. 194-199.

[8] J.S.R. Jang, " Neuro-fuzzy modeling: architecture, analyses and applications ", *Ph.D. Thesis*, Dep. of Electrical Engineering and Computer Science, Univ. of California, Berkeley USA, July 1993.

[9] U. Kaymak and M. Setnes, " Extended fuzzy clustering algorithms ", *ERIM Report Series*, N°. ERS-2000-51-LIS, Nov.2000.

[10] R. Krishnapuram and J. Keller, " A Possibilistic approach to clustering ", *IEEE Trans. on Fuzzy Systems*, vol. 1, 1993, pp. 98-109.

[11] S. Nefti, K. Djouani, and J. Pontnau, " A new clustering based approach for system modeling and optimization ", in revision for *Int. J. on Fuzzy Sets and Systems*.

[12] K.M. Passino and S. Yurkovitch, " Fuzzy Control ", *Menlo Park*, CA: Addison Wesley Longman, 1998.

[13] 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(1), 1987, pp. 116-132.

[14] Y. Touati, K. Djouani, and Y. Amirat, " Adaptive fuzzy logic controller for robots manipulators ", IEEE, *The 10th Artificial Neural Networks in Engineering Conference ANNIE 2001*, St Louis, MI, USA, vol. 11, Nov. 2001, pp. 247-254.

[15] L. X. Wang and R. Langari, " Complex systems modeling via fuzzy logic ", *IEEE Trans. on Sys, Man and Cybernetics*, part-B, vol. 26, Feb. 1996, pp. 100-106.