Improved Fuzzy Clustering Approach: Application to Medical Image MRI

Nour-eddine El harchaoui*, Samir Bara*, Mounir Ait Kerroum*+, Ahmed Hammouch*▲
Mohamed Ouadou* and Driss Aboutajdine*
* LRIT-CNRST URAC 29, Faculty of Science, Mohamed V-Agdal University, Rabat, Morocco
+ LARIT Equipe Imagerie et Multimedia, Ibn Tofail University, ENCG Kénitra, Morocco
▲LRGE, Mohamed V-Souissi University, ENSET, Rabat, Morocco
noureddine.ieee@yahoo.com

Abstract—**Currently, the MRI brain image processing is a vast area of research, several methods and approaches have been used to segment these images (thresholding, region, contour, clustering). In this work, we propose a novel segmentation approach, which is based on fuzzy clustering and also it allows to combine cooperatively expectation maximization algorithms and possibilist c-means. To validate our approach, we have tested successfully on several databases of real images MRI. Thus, to show the performance of our method, we compared our results with different segmentation algorithms: k-means, fuzzy c-means, possibilist c-means and expectation maximization.**

Index Terms— MRI, clustering, k-means, FCM, PCM, EM

I. Introduction

Image segmentation is a technique that allows to partition the image into homogeneous regions. It is used in several areas : pattern recognition, artificial intelligence, medicine [1].

The segmentation of medical images is an essential diagnostic tool for doctors, there are several types of medical images, such as: Radiography (X-Ray), Ultrasound, Magnetic Resonance Image (MRI). Indeed MRI is a technique that is based on nuclear magnetic resonance, it is widely used for brain images, because it is a non-invasive technique and it also provides anatomical images of high resolution (1 mm) and excellent accuracy. Thus, MRI allows for high contrast images of different sections of the brain: sagittal (side view of the brain), coronal (frontal view of the brain) and axial (a top view of the brain) [1][2].

The segmentation of brain MRI can detect tumors and to know the evolution of various pathologies and also analyze and study the different brain tissues: gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF).

In recent years, several segmentation methods and approaches have been developed to segment the internal tissues of the brain (MG, MB, LCR), examples:

- *Thresholding approach:* The basic principle of this method is to find the optimal threshold value (or optimal values, where there are multiple levels)[3][4].
- *Region approach:* This technique involves extracting a region of interest of the image by growing a region from one or more elements constituting a subset of the searched area[5][6][7].
- *Approach based on contours:* This technique is to identify the discontinuities that separate different regions of the image, this approach seeks dissimilarities [8-11].
- *Clustering approach:* Technique which assigns each pixel of the image to a cluster and group pixels of the same property [12][15]. In this work, we are interested in clustering segmentation. The quality of the segmented image depends on the cluster used, and also the selection of the relevant parameters of the segmentation [13][14].

We propose in this article, in the sixth section, a new approach to segmentation (MFCM), which is based on fuzzy clustering and combining with both algorithms, expectation maximization (EM) and possibilist c-means (PCM). To validate our segmentation approach, we applied our method on several real image data bases MRI, and we also compared our results with other segmentation algorithms: K-means, Fuzzy c-means (FCM), PCM and EM. Indeed, the comparative results are presented in section seven. Thus, in sections two, three, four and five, we present the segmentation methods in the order k-means, FCM, PCM, EM. Finally, in the last section, we conclude with a conclusion.

II. Algorithme K-means

A. Presentation

The k-means algorithm is the most widely used in the clustering, because of its simplicity in implementation. It can group the image pixels into K clusters. Each cluster of the partition is defined by its objects (pixels) and its centroid [16].

The k-means is an iterative algorithm that minimizes the sum of the distances between each object (pixel) and centroid of the cluster [16][17]. It changes the objects cluster until the sum can not decrease more. The result is a set of clearly separated and compact clusters, provided that we have chosen the correct value of K number of clusters.

978-1-4673-4766-2/12/$31.00 ©2012 IEEE

B. Basic principle

The principle of K-means algorithm is to minimize the objective function *J* :

$$J(X,V) = \sum_{i=1}^{K} \sum_{j=1}^{L_i} d^2(X_j^i, V_i) \qquad (1)$$

where :

$X_j^{(i)}$: Pixel j of cluster i
V_i: Centroid of cluster i
L_i: Number of elements in the cluster i
$d^2(X_j^{(i)}, Y_i)$: Distance between $X_j^{(i)}$ and Y_i

K-means Algorithm

- Step 1:
  - *Choose a number of classes*
  - *Define the random K centroids*
- Step 2:
  while intra-class inertia in (1) is not stable
  - *Assign each level of gray to the cluser whose center is nearest*
  - *Calculate the cluster centers of gravity of the new classification C'*
  - *C ← C'*

  end while
- Step 3:
  - *View the result of the clustering*

We see that among the limitations and criticisms that gives the k-means algorithm, it takes the value of K fixed in the input parameters, this implies that the number of classes is known a priori. Thus, the initialization of the centers of inertia of a random influence the clustering process, each point of diversion causes an immediate change of the corresponding center. Indeed, to overcome the limitations of k-means, other clustering algorithms have been developed, such as FCM and ISODATA which are based on the principle of k-means.

III. Fuzzy C-means Algorithm

FCM (fuzzy C-means) is an iterative algorithm that based on the principle of fuzzy clustering, it allows classify each data element (pixel) in several classes in a degree of membership [18-20]. The principle of FCM is to minimize the objective function *J* :

$$J(X,Y,U) = \sum_{j=1}^{K} \sum_{i=1}^{N} U_{ij}^m d^2(X_i, Y_j) \qquad (2)$$

Where:

$X=(X_i$, $i=1...N)$
K : Number of cluster
N : Total number of pixels
Y_j : Centre of cluster i
$d^2(X_i, Y_j)$: Distance between Y_j and the pixel X_i
U_{ij}^m : Degree of membership and
m : the fuzzy degree

The matrix *U* satisfies the conditions in (3) and (4) :

$$0 \le U_{ij} \le 1, \ \forall i \in \{1,...,N\} \ and \ \forall j \in \{1,...,K\} \qquad (3)$$

$$\sum_{j=1}^{K} U_{ij} = 1 \quad \forall i \in \{1,...,N\} \qquad (4)$$

$$U_{ij} = \left(\sum_{l=1}^{k} \left(\frac{d^2(X_i, Y_j)}{d^2(X_i, Y_l)} \right)^{\frac{1}{m-1}} \right)^{-1} \quad \forall i \in \{1,...,N\} \qquad (5)$$

$$Y_j = \frac{\sum_{i=1}^{K} U_{ij}^m X_i}{\sum_{i=1}^{K} U_{ij}^m} \quad \forall j \in \{1,...,K\} \qquad (6)$$

FCM Algorithm

- Step 1:
  Initialize the parameters:
  - $X=(X_i$, $i=1...N)$
  - K : number of cluster
  - m: degree of fuzzy
  - ε : threshold representing the convergence error
- Step 2:
  - Initialize the matrix U by membership degrees random values in the interval [0,1] and it also satisfies the condition in (4).
- Step 3:
  Repeat
  - Update the matrix Y cluster centers in (6).
  - Update the matrix U degree of membership in (5).
  To obtain the stability of the matrix Y
  $\|Y^{new} - Y^{old}\| < \varepsilon$

We see that the parameter *m*, which was introduced by Bezdek [19], represents the degree of fuzziness of the partition. Indeed, the choice of this parameter influences the process of FCM algorithm, according Besdek [19][20] the parameter *m* must be strictly greater than 1.

IV. Possibilistic C-means Algorithm

The main advantage of the PCM algorithm is to eliminate the probabilistic constraint is imposed by the FCM algorithm and define degrees of membership in a relative manner . These membership degrees represent measures of similarity absolute between individuals and cluster centers [21][22]. The function objective is *J* :

$$J(U,V,\eta) = \sum_{j=1}^{K} \sum_{i=1}^{N} u_{ij}^m d_{ij}^2 + \sum_{j=1}^{K} \eta_j \sum_{i=1}^{N} (1-u_{ij})^m \qquad (7)$$

Where:

d_{ij} : *Distance between the j^{th} cluster center V_j and the i^{th} pixel X_i*

V_j : *Centroid of cluster j*

U_{ij} : *Degree of membership of the i^{th} pixel to the j^{th} cluster.*

m : Degree of fuzzy.

η_j : *Positive number which determines the distance that the degree of membership of a vector belongs to the j^{th} cluster.*

K : Number of the clusters.

N : Total number of pixels.

The matrix *U* satisfies the conditions *:*

$$0 \le U_{ij} \le 1,\ \forall i \in \{1,...,N\}\ \ and\ \forall j \in \{1,...,K\} \quad (8)$$

$$u_{ij} = \left(\left(1 + \left(d^2(x_i, v_j) \middle/ \eta_j\right)\right)\right)^{\frac{-1}{m-1}} \quad (9)$$

$$V_j = \sum_{i=1}^{N} u_{ij}^m x_j \Big/ \sum_{i=1}^{N} u_{ij}^m \quad \forall j \in \{1..K\} \quad (10)$$

PCM algorithm

- Step 1:
  Initialize the parameters:
  - *$X=(X_i$, i=1...N)*
  - *K : Number of cluster*
  - *m: Degree of fuzzy*
  - η_{ij}: Degree of weight
  - *ε : Threshold representing the convergence error*
- Step 2:
  - *Initialize the matrix U by membership degrees random values in the interval [0,1].*
- Step 3:
  Repeat
  - *Update the matrix V cluster centers in (10).*
  - *Update the matrix U degree of membership in (9).*
  To obtain the stability of the matrix V ;
  $$\|V^{new} - V^{old}\| < \varepsilon$$

V. Expectation Maximization: EM

The Expectation Maximization is an algorithm proposed by Dempster [23][24]. It is an algorithm to clustering and also to estimate the parameters of a mixture model. It combines two steps:

- Expectation (E): This step determines the expectation of the likelihood based on the past observed variables.
- Maximization (M): this step is to maximize the likelihood of step (E).

EM Algorithm

Input : *$X=\{x_i$, i=1..N\}*

x_i *: Intensity of pixel i*

π_i^k *: A priori probability of pixel i; belonging to cluster k*

Output :

μ_k, Σ_k *Parameters of the Gaussian associated with each cluster.*

γ_i^k *A posterior probability of pixel i; belonging to cluster k*

- Step 1 :
  - *Initialize the value* γ_i^k *by* π_i^k
- Step 2 :
  - *parameter calculation :* μ_k, Σ_k

$$\mu_k = \frac{\sum_{i=1}^{N} \gamma_i^k x_i}{\sum_{i=1}^{N} \gamma_i^k} \quad (11)$$

$$\Sigma_k = \frac{\sum_{i=1}^{N} \gamma_i^k (x_i - \mu_k)(x_i - \mu_k)^T}{\sum_{i=1}^{N} \gamma_i^k} \quad (12)$$

- Step 3 :
  - *Updating the posterior probabilities* γ_i^k

$$\gamma_i^k = \frac{\pi_i^k G_{\mu_k, \Sigma_k}(x_i)}{\sum_{l=1}^{K} \pi_i^l G_{\mu_l, \Sigma_k}(x_i)} \quad (13)$$

- Step 4 :
  - *Iterate* (Step1) and (Step2) until convergence.

Indeed, EM algorithm can make the classification by maximizing the likelihood of observed variables, this allows to find the optimal partitions. In our approach, we have integrated this algorithm.

VI. Improved FCM Algorithm : MFCM

In the previous sections, each algorithm was presented has its advantages and limitations. To improve the performance of segmentation, we propose a new approach to segmentation (MFCM), which is based on fuzzy classification and combination of two algorithms: EM and PCM. Thus, our approach MFCM reduces errors classification and also estimate the number of clusters, which overcomes the limitations of FCM algorithm for initializing clusters number. Indeed, we present our approach in detail in the following steps:

MFCM Algorithm

- Step 1:
  Initialize the parameters
  - *Maximum iteration*
  - *Initial number of clusters (nC=2 is the initial value)*

- *Maximum number of clusters nCMax*
- *Initialize V_j (cluster centers)*
- *Initialize $U=[u_{ij}]$ matrix, $U^{(0)}$*
- *Degree of fuzzy (m=1.5 is the initial value)*

• Step 2 :
- *FCM*
- *At i-step:*
  - *Calculate the vectors U,V*
  - *Calculate the objective function (F_obj)*
  - *EM*
  - *Optimum matrix (M_opt[i])*
- *$i \leftarrow i+1$*
- *if nC is less than nCMax go to step(2) else go to step(3)*

• Step 3 :
- *opt=min(M_opt[i])*
- *nC=opt*

• Step 4 :
- *FCM*
- *Calculate the vectors U,V*
- *Calculate the objective function (F_obj)*

• Step 5 :
- *PCM*

• End step

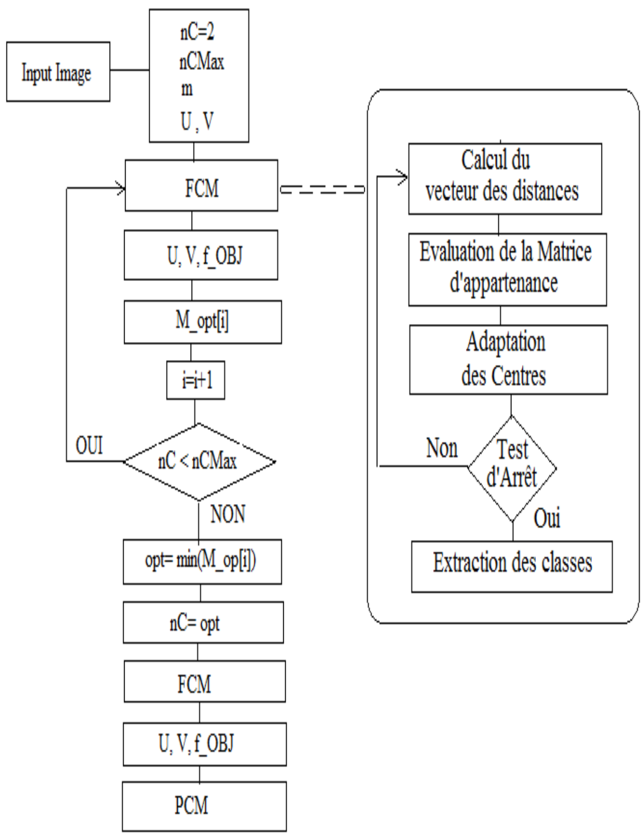
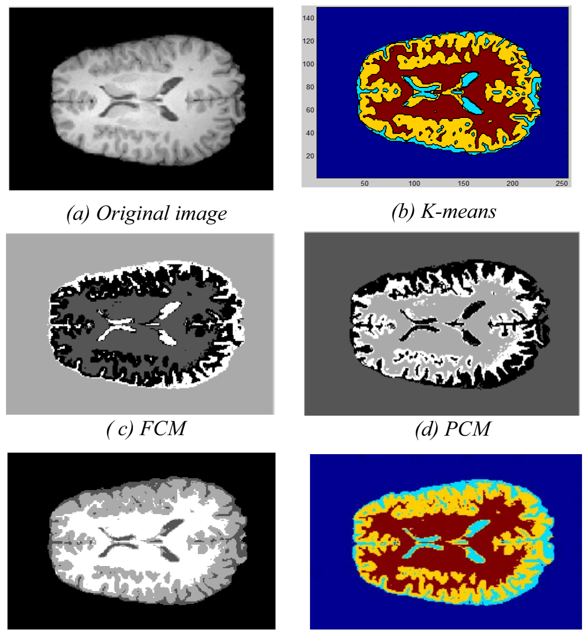


Figure 1: Algorithm MFCM

VII. Experimental Results

In our study, the validation performance of our algorithm has been tested on real MRI brain images. The images are acquired: T1-weighted, T2-weighted and PD (proton density) with size *150x256x256* voxels (and *1mmx1mmx1mm* for each voxel).

Figure 2 shows the results of classification of different brain tissues (WM, GM, CSF): Images (b), (c), (d), (e) and (f) are the results of segmentation by k-means, FCM, PCM, EM and MFCM succession. Figure 3 shows the histogram for the different brain tissues (MG, MB, CSF) using EM algorithm. Indeed, the evaluation of the performance of the methods was done by calculating the accuracy in the classification (in %). The table 1 shows the comparative results of classification for the five algorithms, we tested the algorithms on nine of datasets MRI brain images.



(a) Original image *(b) K-means*
(c) FCM *(d) PCM*
(e) EM *(f) MFCM*

Figure 2 : clustering of the MRI image with algorithms

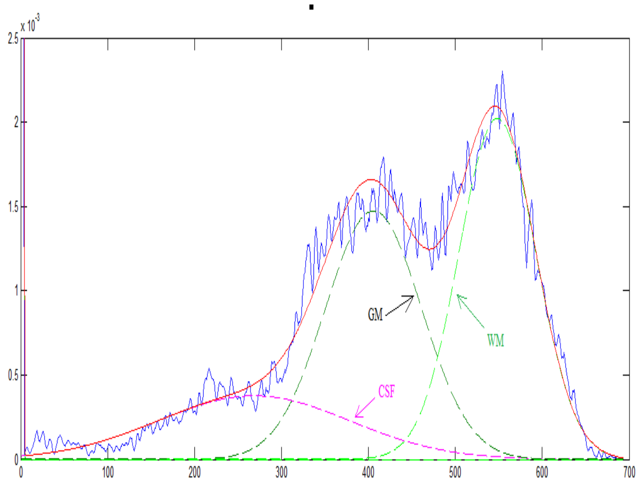


Figure 3 : probability distribution with EM

Table 1 : Performance of the MFCM algorithm

Test Datasets	Segmentation Algorithms				
	K-means (%)	FCM (%)	PCM (%)	EM (%)	MFCM (%)
Dataset1	61,54	85,44	82,90	83,20	**94,79**
Dataset2	63,67	85,03	83,55	84,32	**98,22**
Dataset3	67,34	91,83	87,87	90,95	**95,21**
Dataset4	62,25	87,40	85,98	87,57	**95,74**
Dataset5	67,10	83,85	81,48	84,91	**97,75**
Dataset6	69,94	90,77	87,81	89,17	**96,98**
Dataset7	72,43	86,69	85,92	87,69	**97,99**
Dataset8	64,79	88,05	86,63	88,17	**96,98**
Dataset9	60,77	89,76	87,34	88,58	**95,74**
Average	**65,54**	**87,65**	**85,50**	**87,17**	**96,60**

According the results in the table 1, we see that our method MFCM gives a better classification rate than the other algorithms. Indeed, we can also view these results on a graph, figure 4, where we can clearly see the stability of our approach and also see the excellent classification accuracy.

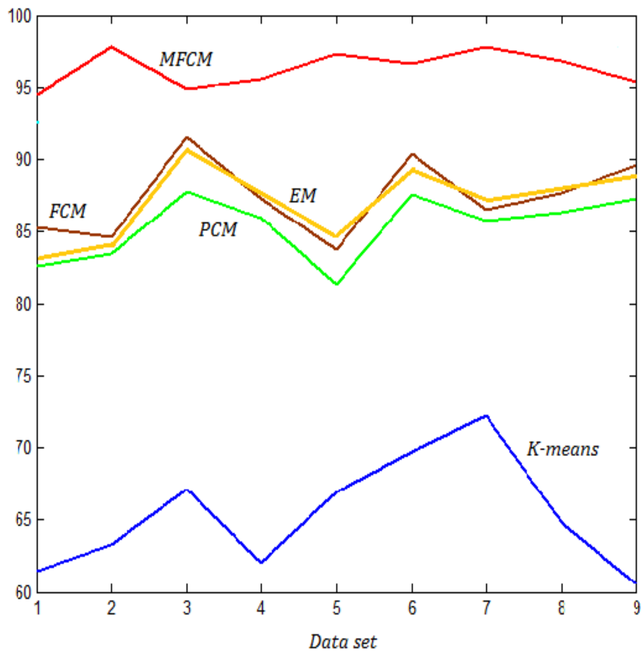


Figure 4: Comparative Performance of Accuray Classification (%)

VIII. Conclusion

In this article, we presented some classification algorithms, as well as common principles to them. Hence the interest to combine these algorithms to overcome the limitations of each one. Indeed, our approach is based on the principle of fuzzy clustering and also it combines two algorithms: EM and PCM. The performance evaluation of our method has been successfully tested on real MRI brain images. In future work, we will use optimization techniques in the selection of input variables to improve the quality and performance of our approach.

References

[1] C.S. Drapaca, V. Cardenas, C. Studholme, "Segmentation of tissue boundary evolution from brain MR image sequences using multiphase level sets," Computer Vision and Image Understanding, vol.100, 2005, pp. 312-329.

[2] M. Kamber, R. Shinghal, D. L. Collins, G. S. Francis, A. C. Evans, "Model-based segmentation of multiple sclerosis lesions in magnetic resonance brain images," *IEEE Trans. on Med. Imaging*, vol. 14, no. 3, pp. 442-453, 2000.

[3] Y. Qiao, Q. Hu, G. Qian, S. Luo, W. L. Nowinski, "Thresholding based on variance and intensity contrast," Pattern Recognition, vol. 40, 2007, pp. 596-608.

[4] Jun Zhang, and Jinglu Hu, "Image Segmentation Based on 2D Otsu Method with Histogram Analy-sis," 2008 International Conference on Computer Science and Software Engineering, IEEE 2008, P.105-108.

[5] M. Tabb and N. Ahuja, "Unsupervised multiscale image segmentation by integrated edge and region detection," IEEE Transactions on Image Processing, Vol. 6, No. 5, 642–655, 1997.

[6] J. Fan, D. Yau, A. Elmagarmid, and W. Aref, "Automatic image segmentation by integrating color-edge extraction and seeded region growing," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1454–1466, Oct. 2001.

[7] S. Wan and W. Higgins, "Symmetric region growing," *IEEE Trans. Image Process.*, vol. 12, no. 9, pp. 1007–1015, Sep. 2003.

[8] T. Pavlidis and Y.-T. Liow, "Integrating region growing and edge detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, pp. 225–233, 1990.

[9] L. Cohen and I. Cohen, "Finite-element methods for active contour models and balloons for 2-D and 3-D images," *IEE Trans. Patt. Anal. Mach. Intell*, pp.1131-1147, 1993.

[10] T.F. Chan and L.A. Vese, "Active Contours Without Edges," *IEEE Transaction on Image Processing*, vol.10, n°7, pp.266-277, 2001.

[11] L. Najiman and M. Schmitt, "Geodesic saliency of watershed contours and hierarchical segmentation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, pp. 1163–1173, 1996.

[12] M. Ait-Kerroum, A. Hammouch, and D. Aboutajdine, "Textural feature selection by joint mutual information based on Gaussian mixture model for multispectral image classification," Pattern Recognition Letters, Volume 31, Issue 10, 15 July 2010, Pages:1168-1174 .

[13] M. Ait kerroum, A. Hammouch and D. Aboutajdine, "Input Textural Feature Selection By Mutual Information For Multispectral Image Classification," International Journal of Signal Processing 6:1 2010.

[14] S. Le Hegarat-Mascle, I. Bloch, D. Vidal-Madjar. "Application of Dempster-Shafer Evidence Theory to Unsupervised Classication in Multisource Remote Sensing," IEEE Transactions on 70 Geoscience and Remote Sensing, 35(4) : 1018-1031, 1997.

[15] R Xu, D Wunsch, "Survey of Clustering Algorithms," IEEE Transactions on Neural Networks. 2005,16(3): 645-678.

[16] K. Wagsta, C. Cardie, S. Rogers, and S. Schroedl. "Constrained k-means clustering with background knowledge," In International Conference on Machine Learning, pages 557 584,2001.

[17] L. Jing, M. K. Ng and J. Z. Huang , "An Entropy weighting k-Means Algorithm for subspace clustering of high dimensional sparse data," IEEE Transaction on knowledge and Data Engineering Vol 19, No 8, August 2007.

[18] F. Hppner, F. Klawonn, R. Kruse, and T. Runkler, Fuzzy Cluster Analysis-Methods for Classication, Data Analysis and Image Recognition. John Wiley & Sons, LTD, 1999.

[19] Bezdek,J.C.,1974. Cluster validity with fuzzy sets J.Cybernit.3,58 72.

[20] J. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms. New York:Plenum,198..

[21] R.Krishnapuram and J.M. Keller, "A Possibilistic Approach to Clustering," IEEE Transactions on Fuzzy Systems, 1(2):98-110, 1993.

[22] K.P. Detroja et al., "A Possibilistic Clustering Approach to Novel Fault Detection and Isolation," Journal of Process Control,16(10):1055-1073, 2006.

[23] A. Dempster, N. Laird, and D. Rubin. "Maximum likelihood for incomplete data via the EM algorithm." Journal of the Royal Statistical Society, 39(1) :1–38, 1977.

[24] Biernacki, C., Celeux, G. and Govaert G, Assessing a mixture model for clustering with the Integrated Completed Likelihood, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22 (7), 719–725, 2000.