A modified Kernelized Fuzzy C-Means algorithm for noisy images segmentation: Application to MRI images

Abdenour Mekhmoukh[1], Karim Mokrani[2] and Mohamed Cheriet[3]

[1] **LTII Laboratory, Electrical Engineering department, University of Bejaia, Algeria**

[2] **LTII Laboratory, Electrical Engineering department, University of Bejaia, Algeria**

[3] **Department of Automated Manufacturing Engineering , ETS, Montréal, Québec, Canada**

Abstract

Image segmentation is a low-level processing operation; it is the basis for many applications such as industrial vision and medical imaging. Segmentation provides a partition of the image by gathering pixels with similar grey levels in the same class. The main problem of this algorithm is that it does not take into account the image topology; it is based only on the pixels values. Thus, it is very sensitive to noise and inhomogeneities in the image moreover, it remains dependent on the initialization of the cluster centers.

In general the clustering algorithm chooses the initial centers in a random manner but the cluster centers initialization, using "Expectation Maximization" algorithm allows an optimal choice of these centers. To account for the topology of the image, the statistical parameters of a window around the pixel are considered. These a priori are used in the optimization of the cost function. The application to MRI images from the Brain Web public database, with different noise levels, shows the performance of the proposed approach.

Keywords: *image segmentation, FCM, KFCM, clustering, MRI images.*

1. Introduction

The magnetic resonance imaging (MRI) is one of the most significant advances in medicine; it allows, for the physician, the diagnosis and the monitoring of various diseases [1]. Given the size of the data to be analyzed, the use of image processing techniques can provide an aid to the diagnosis. Moreover, MRI images suffer from inhomogeneities and noise. Due to these problems, classical, or net, approaches do not succeed to correctly segment these images. Classification techniques, known from unsupervised artificial intelligence, fuzzy logic, probability and statistics, know a great success. In this paper, we propose to apply these techniques to segment MRI brain images.

2. Segmentation by unsupervised classifiers

2.1Net segmentation

Classification methods are used to group objects into clusters or classes of homogeneous objects. The objects sets share common features; they are similar but are clearly distinguished from objects of other classes. The items listed are the pixels of the image that allows us to have groups. Generally, the classification methods lead to more or less different results. K-Means algorithm [2] is used because of its implementation simplicity.

2.1.1 K-Means algorithm

The K-means algorithm aims to collect N points in C or primitive groups whose number is predetermined. Given a set of points $X = x_1, x_2, . . . , x_N$, each xj is characterized by n variables, and thus can be represented in an n-dimensional space $x_j \in R^n$.

As part of the clustering, we generally seek to partition the space into compact classes that are separated from each other. The K-Means algorithm aims to minimize the intra-class variance, which results in the minimization of the following objective function:

$$J(X,V) = \frac{1}{2}\sum_{i=1}^{C}\sum_{j=1}^{N}\left\| x_j - v_i \right\|^2 \qquad (1)$$

Where:

- v_i is the center of class i.

IJCSI
www.IJCSI.org

- $\|x_j - v_i\|^2$ represents the Euclidean norm (clusters) which measures the dissimilarity between a pixel and a center.

2.1.2 Choice of the number of classes

The classical K-Means algorithm leaves a free parameter, the number of classes which, in the case of image segmentation, is the number of pixel intensities used to represent the image. Usually the choice of K is made empirically by selecting the value that minimizes Eq 1. This segmentation method considers that the regions are subsets of the net set of pixels of the image. Reconstruction of the regions is made by assigning each pixel to a single region. This approach has limitations when we do not have an a priori knowledge on the pixels allocation.

Fuzzy segmentation methods assume that the pixels belong more or less to these subsets and consider that the pixels can be grouped into fuzzy sets of pixels. These methods are based on the concepts of fuzzy logic.

2.2 Fuzzy segmentation

Fuzzy segmentation is build on the concepts of fuzzy logic [3] by considering the pixels as fuzzy sets. Methods such as C-Means and Fuzzy C-Means (FCM) [4], Possibilistic C-Means (PCM) [5] and Possibilistic Fuzzy C-Means (PFCM) [6] are often used.

2.2.1C-Means algorithm

Let $E = \{x_1, x_2, \ldots, x_k\}$ be the set of samples of the training set, x_k is the k^{th} attribute vector x. The sets of prototypes or cluster centers are $\{V_1, V_2, ., V_c\}$ where c denotes the number of classes. The C-Means algorithms minimize any objective function J (sum of the distances intra-classes). The Euclidean distance d is calculated between the attribute data vector and the group prototypes. The objective function J has the form:

$$J = \sum_{i=1}^{c} \sum_{k=1}^{n} \mu_{jk} d^2(x_k, V_i) \tag{2}$$

Where:

- v_i is the center of class i.
- $\mu_{jk} = P(c_j/x_k)$: usually represents the membership probability of x_k to class c_j

2.2.2 Fuzzy C-Means algorithm

Let E a set of vectors of attributes, $\{V_1, V_2, ., V_c\}$ the centers of the classes, where c denotes the number of classes. The degree of membership of an element x_k in a group V_i is noted $\mu_i(x_k)$. In this approach, a given attribute vector can belong to several groups. U is the matrix of membership degrees (also called C-fuzzy partition matrix) whose size is $c*n$, where c is the number of classes and n the number of items to classify. The unsupervised clustering Fuzzy C Means (FCM) algorithm is a method of partitioning based on the total Picard iteration through necessary conditions for optimizing a sum of squared errors objective function (Jm) [7]. It is based on the sum of the distances between the vector and weighted classes center by the membership functions:

$$J_m(U, V) = \sum_{i=1}^{c} \sum_{k=1}^{n} \mu_{ik}^{m} d^2(x_k, V_i) \tag{3}$$

Where:

- $d = \|x_k - V_i\|$ is the Euclidean norm and V_i represents the center of class i.
- m is any real number greater than 1.

$$\forall i, j \quad \mu(x_k) \in \{0,1\} \quad U = [\mu_{ik}(x_k)] \tag{4}$$

$$U = \begin{pmatrix} \mu_{11} & \mu_{12} & . & . & \mu_{1n} \\ \mu_{21} & . & . & . & \mu_{2n} \\ . & . & . & . & . \\ . & . & . & . & . \\ \mu_{c1} & . & . & . & \mu_{cn} \end{pmatrix} \tag{5}$$

Three constraints are defined for the membership matrix of degrees U:

$$0 \le \mu_{ik} \le 1 \quad 1 \le i \le c \quad 1 \le k \le n \tag{6}$$

$$\sum_{i=1}^{c} \mu_{ik} = 1, \forall k \in [1, n] \tag{7}$$

$$0 \le \sum_{i=1}^{c} \mu_{ik} \le 1, \forall i \in [1, c] \tag{8}$$

The FCM is based on the update of the membership function. During the algorithm iterations, the FCM changes the partition (U matrix) by minimizing the objective function J_m (Algorithm 1).

$$\mu_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{jk}} \right)^{\frac{2}{m-1}}} \tag{9}$$

Copyright (c) 2012 International Journal of Computer Science Issues. All Rights Reserved.

where :

- μ_{ik} is the degree of membership of x_i in the cluster k

The cluster centers V_i are calculated by

$$V_i = \frac{\sum_{i=1}^{n} \mu_{ik}^{m} x_k}{\sum_{k=1}^{n} \mu_{ik}^{m}} \tag{10}$$

The process stops when $\left|J_m^{t+1} - J_m^{t}\right| < \varepsilon$ or a predefined number of iterations is reached.

Algorithm 1 FCM implementation

Step 1 :
– Set the number of class: c
– Set m ,
– Choose ε to stop the test.
– Initialize (t = 0) matrix $U^t = 0$ by random values.
– Choose the norm d_{ik} .

Step 2: Compute the new cluster centers using Eq (10)

Step 3 : Update U^t using :

$$\mu_{ik}^{t+1} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{jk}} \right)^{\frac{2}{m-1}}} \tag{11}$$

Compare J^{t+1} to J^t if $\left|J_m^{t+1} - J_m^{t}\right| < \varepsilon$ then stop the process, else t = t + 1 and return to step 2.

2.2.3 Kernelized Fuzzy C-Means algorithm

By using the popular 'kernel method', we construct a kernel version of FCM (KFCM), where the original Euclidian distance in FCM is replaced with the following kernel-induced distance measures [8] :

$$d(x, y) = \|\Phi(x) - \Phi(y)\| \tag{12}$$

$$d(x, y) = \sqrt{K(x, x) - 2K(x, y) + K(y, y)} \tag{13}$$

Here $\boldsymbol{\Phi}$ is a nonlinear function which maps $\boldsymbol{x_k}$ from the input space $\boldsymbol{X}$ to a new space $\boldsymbol{F}$ with higher or even infinite dimensions. $\boldsymbol{K(x,y)}$ is the kernel function which is defined as the inner product in the new space $\boldsymbol{F}$ with:

$$\boldsymbol{K(x,y)} = \boldsymbol{\Phi(x)^T \Phi(y)}, \text{ for } \boldsymbol{x}, \boldsymbol{y} \text{ in input space } \boldsymbol{X}.$$

An important fact about kernel function is that it can be directly constructed in the original input space without knowing the form of $\boldsymbol{\Phi}$. That is, a kernel function implicitly defines a nonlinear mapping function. There are several typical kernel functions, e.g. the Gaussian kernel:

$$\boldsymbol{K(x,y) = exp\left(-\frac{\|x-y\|^2}{\sigma^2}\right)} \tag{14}$$

By replacing the Euclidean distance in Eq (3) with the Gaussian kernel-induced distance, the new objective function is:

$$J_m(U, V) = \sum_{i=1}^{c} \sum_{k=1}^{n} \mu_{ik}^{m} \|\Phi(x_k) - \Phi(V_i)\|^2 \tag{15}$$

$$\boldsymbol{J_m(U,V) = 2\sum_{i=1}^{c}\sum_{k=1}^{n} \mu_{ik}^{m}(1 - K(x_k, V_i))} \tag{16}$$

The membership degrees terms become

$$\mu_{ik} = \frac{(1-K(x_k,V_i))^{\frac{-1}{m-1}}}{\sum_{j=1}^{c}(1-K(x_k,V_j))^{\frac{-1}{m-1}}} \tag{17}$$

And cluster centers

$$\boldsymbol{V_i = \frac{\sum_{k=1}^{n} \mu_{ik}^{m} K(x_k,V_i) x_k}{\sum_{k=1}^{n} \mu_{ik}^{m} K(x_k,V_i)}} \tag{18}$$

2.2.4 Modified Kernelized Fuzzy C-Means algorithm

Although KFCM can be directly applied to image segmentation like FCM, we propose to modify the algorithm by taking into account the image topology. For the Modified KFCM (MKFCM) we propose:

1. To initialize the cluster centers by using the "Expectation Maximization" (EM) algorithm [9] for an optimal choice of the centers.

2. To take into account the image topology; the statistical parameters of a window around the pixel are considered.

For an image y, of size $(N \times M)$, and a sliding window of size $(p \times p)$, the four features extracted from the window

Copyright (c) 2012 International Journal of Computer Science Issues. All Rights Reserved.

centered at pixel *(n, r)* are given by the following equations:

$$Me = \frac{1}{MN}\sum_{i=-\frac{p-1}{2}}^{\frac{p-1}{2}}\sum_{j=-\frac{p-1}{2}}^{\frac{p-1}{2}} y(n+i, r+j), \qquad (19)$$

$$V = \frac{1}{MN}\sum_{i=-\frac{p-1}{2}}^{\frac{p-1}{2}}\sum_{j=-\frac{p-1}{2}}^{\frac{p-1}{2}} \left(y(n+i, r+j) - Me\right)^2, \qquad (20)$$

$$Sk = \frac{1}{MN}\sum_{i=-\frac{p-1}{2}}^{\frac{p-1}{2}}\sum_{j=-\frac{p-1}{2}}^{\frac{p-1}{2}} \left(y(n+i, r+j) - Me\right)^3, \qquad (21)$$

$$Ku = \frac{1}{MN}\sum_{i=-\frac{p-1}{2}}^{\frac{p-1}{2}}\sum_{j=-\frac{p-1}{2}}^{\frac{p-1}{2}} \left(y(n+i, r+j) - Me\right)^4, \qquad (22)$$

Where, *Me, V, SK* and *Ku*, are the mean, the variance, the skewness, and the kurtosis respectively. Note that *p* must be odd for a window centered on each pixel. Thus, we obtain a matrix *H* that contains the extracted features. H is obtained as shown in the following figure, pixel by pixel, from left to right and top to bottom.

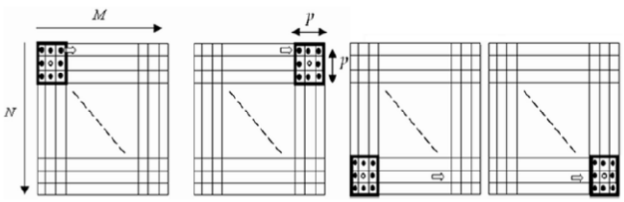


Fig 1. Extraction of statistical features.

So the equation (17) and (18) become

$$\mu_{ik} = \frac{(1-K(H_k,V_i))^{\frac{-1}{m-1}}}{\sum_{j=1}^{c}(1-K(H_k,V_j))^{\frac{-1}{m-1}}} \qquad (23)$$

$$V_i = \frac{\sum_{i=1}^{n}\mu_{ik}^{m}K(H_k,V_i)H_k}{\sum_{k=1}^{n}\mu_{ik}^{m}K(H_k,V_i)} \qquad (24)$$

The steps of the algorithm are summarized in the following diagram:

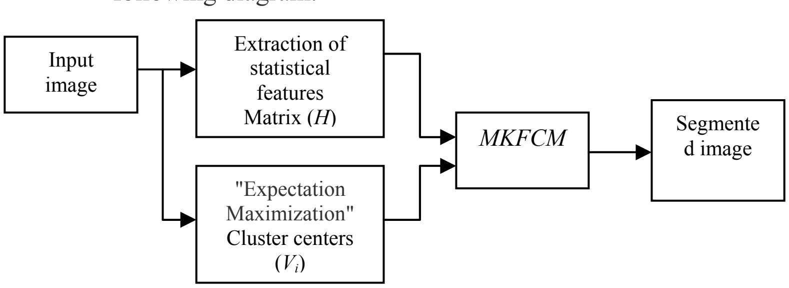


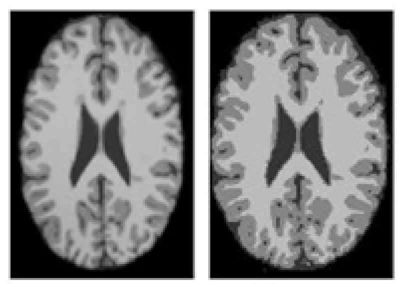
Fig 2. Proposed segmentation process.

3. Results and discussion

To apply the classification approach described above, we use the Brainweb database [10] of the Neurological Institute of McGill University in Montreal. The Brainweb website simulates brain MRI with different levels of noise and inhomogeneities. These simulations are generated from a ground truth which is the classification of a brain image (millimeter resolution) into different classes corresponding to different tissues and structures.

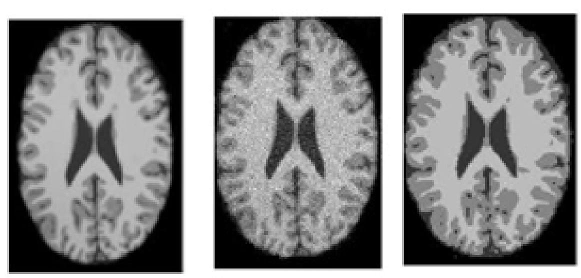
For our evaluations, we have thus made different data sets, varying noise levels. A particular set consists of T1-weighting for each noise level (0%, 5% and 9%). To segment brain tissues in MRI images, we must define the different parameters governing the algorithm, namely the values of m, the number of classes C, and the image pixels vectors form. Our goal is to segment the brain, we set the number of class *i* (c = 4) corresponding to the three brain tissue present in the brain (white matter, grey matter, cerebrospinal fluid and background).

The choice of the vectors form is crucial since their relevance will allow the pixels discrimination. The vector form x_j of a pixel *j* is its grey level.



Original image Segmented image (FCM)

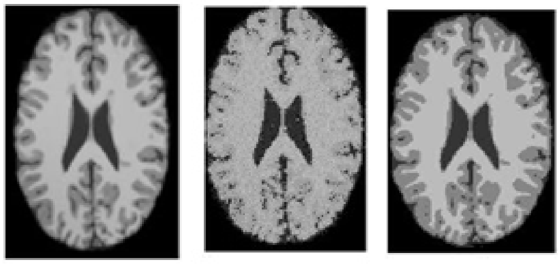
Fig 3. Results for an axial slice T_1 with 0% of noise with classical FCM



Original image FCM MKFCM

Fig 4. Results for an axial slice T_1 with 5% of noise

Copyright (c) 2012 International Journal of Computer Science Issues. All Rights Reserved.



Original image FCM MKFCM

Fig 5. Results for an axial slice T_1 with 9% of noise

Figure. 4 and 5 show a comparison of segmentation results between FCM and MKFCM, when applied on T1-weighted MR phantom corrupted with noise. From these images, we can see that traditional FCM is unable to correctly segment the images.

MKFCM segmented the image into three classes corresponding to background, grey matter (GM) and white matter (WM) and cerebrospinal fluid (CSF). MKFCM produced better results than FCM due to its ability to cope with noise.

Two types of cluster validity functions, fuzzy partition and feature structure, are often used to evaluate the performance of clustering. The representative functions for the fuzzy partition are partition coefficient *Vpc* [11] and partition entropy *Vpe* [12]. They are defined as:

$$V_{pc} = \frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^2 \qquad (25)$$

$$V_{pe} = \frac{-1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}\, log\, u_{ij} \qquad (26)$$

The idea of these validity functions is that the partition with less fuzziness means better performance. As a result, the best clustering is achieved when the value *Vpc* is maximum or *Vpe* minimum.

Table 1. Results for the image corrupted by 5% of noise

Algorithm	*Vpc*	*Vpe*
FCM	0.8520	0.2891
MKFCM	0.9693	0.0835

Table 2. Results for the image corrupted by 9% of noise

Algorithm	*Vpc*	*Vpe*
FCM	0.8293	0.2949
MKFCM	0.9484	0.0975

4. Conclusion

The results show that the MKFCM algorithm outperforms the FCM. Although the problem of image segmentation is a problem that is far from being resolved, segmentation result, as defined, is obviously not unique. The choice of a method is related to several factors and parameters adjustment that governs the algorithm deserves special attention. Finally, it seems interesting to consider the integration of other constraints on the pixels spatial arrangement and to combine several classification algorithms working in cooperation. Cooperation strategies should be considered in order to achieve faster convergence and improve the segmentation results.

References

[1] N.A. Mohamed, “Modified fuzzy C-mean algorithm for medical image segmentation,” M.Sc. thesis, Elect. Eng. Dept., Univ. Louisville, KY, 1999.

[2] J. B. MacQueen, "Some Methods for classification and Analysis of Multivariate Observations”, Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, 1971.

[3] L.A Zadeh , ”Fuzzy sets, Information and Control”, 8,1965 pp 338-353.

[4] J.P. Stitt, R.L.Tutwiler and A.S.Lewis,”Fuzzy C-Means image segmentation of side scan sonar images”, Proceeding of, International Symposium on Signal and Image Processing, IASTED 2001, vol. 1, pp. 27-32, 2001.

[5] R. Krishnapuram and J.M Keller, ”A possibilistic approach to clustering”. IEEE Transactions on Fuzzy Systems, 1(2), pp.98-110, 1993.

[6] Nikhil R. Pal, Kuhu Pal, James M. Keller, and James C. Bezdek, ”A possibilistic fuzzy c-means clustering algorithm”, IEEE Transactions on Fuzzy Systems, 13(4), pp.517–530, 1997.

[7] J.C. Bezdek, J. Hathaway, M.J. Sabin, and W. T. Tucker, ”Convergence Theory for Fuzzy C-Means: Counterexamples and Repairs”, IEEE Trans, 17, pp. 873-877, Sept./Oct. 1987

[8] D.Q. Zhang, and S.C. Chen, "A novel kernelized fuzzy c-means algorithm with application in medical image segmentation," Artif. Intell. Med, vol.32, pp.37–50, 2004.

[9] Geoffrey McLachlan and Thriyambakam Krishnan. The EM Algorithm and Extensions. John Wiley & Sons, New York, 1996.

[10] BrainWeb: “Simulated Brain Database”, McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill.

[11] J. C. Bezdek. Cluster validity with fuzzy sets. Cybernetics and Systems, 3(3), pp. 58–73, 1973

[12] J. C. Bezdek. Mathematical models for systematic and taxonomy. Proc. of Eight International Conference on Numerical Taxonomy, San Francisco, pp. 143–166, 1975.

Copyright (c) 2012 International Journal of Computer Science Issues. All Rights Reserved.