International Journal of Molecular Medicine and Advance Sciences 2 (2): 190-198, 2006


# A New Hybrid RMN Image Segmentation Alogrithm

Abdelouahab Moussaoui, Nabila Farahta and Victor Chen
Department of Computer Science, University Ferhat, Abbas of Setif, Algeria

**Abstract:** The development of aid's systems for the medical diagnosis is not easy thing because of presence of inhomogeneities in the MRI, the variability of the data from a sequence to the other as well as of other different source distortions that accentuate this difficulty. A new automatic, contextual, adaptive and robust segmentation procedure by MRI brain tissue classification is described in this article. A first phase consists in estimating the density of probability of the data by the Parzen-Rozenblatt method. The classification procedure is completely automatic and doesn't make any assumptions nor on the clusters number nor on the prototypes of these clusters since these last are detected in an automatic manner by an operator of mathematical morphology called Skeleton by Influence Zones detection (SKIZ). The problem of initialization of the prototypes as well as their number is transformed in an optimization problem; in more the procedure is adaptive since it takes in consideration the contextual information presents in every voxel by an adaptive and robust non parametric model by the Markov Fields (MF). The number of bad classifications is reduced by the use of the criteria of MPM minimization (Maximum Posterior Marginal).

**Key words:** Clustering, automatic classification, SKIZ, markov fields, image segmentation, maximum posterior marginal (MPM).

## INTRODUCTION

The survey of the human brain, that it is on the anatomical or functional plan, is currently a domain of research in full expansion. The main factors that contribute to make possible these studies are mainly the evolution of the materials and the imaging techniques, the capacity of calculation, constantly increasing, of the computers and finally the definition and the implementation of complex systems of data processing.

Such systems are not limited to the roles of visualization of the data; they rather have for ambition to be able to help the practitioner in his choices of diagnosis, to signal him the possible pathological risks and even to guide his surgical gesture.

The implementation of such systems is a difficult problem and has been studied extensively all over the world by numerous teams of research[1,2]. The difficulties are located to different levels as for example the choice of acquisition's parameters of the images that will permit us to get a best data in order to automate the treatments, the complexity of the human brain whose anatomy is extremely variable of an individual to another and this requires a developing of other very adaptive methods for the treatments , or to act on the type of information that it will be necessary to provide to the systems to get a good working of these.

The existing methods of segmentation are generally founded on very different principles as for example the process of classification, the use of deformable contours models or knowledge models.

The work presented in this article has for goal the implementation of a segmentation system of cerebral MRI that permits us to combine between several complementary approaches in answer to the complexity of the posed problem. More precisely, the objective is to be able to integrate the approach proposed in a more global system dedicated to the segmentation of the cerebral tissues by the implementation of cooperation between heterogeneous approaches. The interest of such an approach is to be able to exploit the complementary information that results from the application of several methods in order to propose a complete segmentation system.

**Medical image segmentation (state of the art):** The segmentation of the cerebral RMI is a difficult problem due to the big number of present organs and to the complexity of their structure. The use of operators of low level type of gradient detecting doesn't permit to provide the sufficiently structured information. In answer to this difficulty, many models have been proposed[3]. The more used models are the energizing models and the parametric models. Among those using an energizing model, we mention the method of Chiou and Hwang[4] that propose an approach using training by a neural network to define a specific knowledge[5]. Kapkur *et al.*[1] propose an algorithm of brain detection in 3 stages to improve the initialization of the contour and the determination of the energy bound to the image. Davatzikos and Prince[6] propose a model of specific active contour to the cortex segmentation. The internal energy is limited to the use of

**Corresponding Author:** Abdelouahab Moussaoui, Department of Computer Science, University Ferhat, Abbas of Setif, Algeria

elastic strengths in order to allow the contour to follow the convolutions of the cortical furrows. In[7], the contours of the furrows are initialized to the surface of a brain already segmented, to the points of minimal curvature. Barillo *et al.*[8] propose a similar approach for the numeric representation of the cortical furrows. In an extension of the active contours in 3D is used for the representation of deformable surfaces.

Other methods (also named parametric methods) permit to modeling better the objects. Two sub-families of methods have been developed: those using some knowledge and those using not any knowledge. Among those that use some knowledge for the parameterization we mention the approach of Staib *et al.*[10] that proposes a parameterization founded on the Fourier's coefficients to describe present elements in biomedical images. In[11] the authors proposed a parameterization in 3D to segment cardiac structures. However, other parametric approaches[12] introduce some priori knowledge from a set of training to force better the space of variation of the parameters used for the modeling. Cootes *et al.*[13] proposed an approach to basis of deformable statistical models founded on the existence of a training set. Duta and Sonka[14] propose an extension of the approach of Cootes that adds the priori knowledge (contours forces, average localization of a contour, relations between contours) to improve the process of localization as well as the detection of the non valid shapes.

To the same moment, other methods were developed, known as knowledge basis methods. Contrary to the approaches using processes of classification, the knowledge used in these methods is described explicitly. It permits to use better the priori knowledge before the data processing. Gong and Kulikowski[15] propose a system to do the automatic generation of plans. Li *et al.*[16] propose a model of the brain integrating the symbolic and numeric knowledge to guide a process of interpretation. In[17] the proposed approach is guided by the knowledge founded on the generation of a fuzzy modeling of the attributes (size, mean gray level ...) of the different organs searched for in the RMI cuts. Sonka *et al.*[18] propose a basis knowledge approach founded on the genetic algorithms. The interest of the basis knowledge approaches that we presented amounts is their capacity to use a model of knowledge to control the segmentation processes.

## APPROACHES BY CLASSIFICATION

The approaches by classification are used extensively for the segmentation in cerebral RMI. Several methods have been developed then; each among them leans on a very definite principle. We can differentiate them by two main features: the existence of a modeling and its global or local character as well as the nature of the knowledge that can be implicit or explicit.

Some approaches based on the Bayesian classification exploit this principle while introducing some priori knowledge to segment 3D images of the brain in 4 classes (gray matter, white matter, cerebro-spinal liquid and other)[19]. Wells *et al.*[20] propose an approach by Bayesian classification to segment in an adaptive manner the brain images. The objective is to integrate in the process of classification an evaluation of the bias that reflects the presence of inhomogeneities in the RM Images.

Other approaches use the Markov fields; these last permits to associate the process of segmentation to the minimization of an energy function defined on the image to segment. Rajapakse *et al.*[21] propose to apply the classification by Markov fields to MRI where the brain has been isolated previously of the rest of the image. The approach consists in classifying the 3 classes remaining after isolation of the brain (gray matter, white matter and cerebro-spinal liquid). Jaggi[22] proposes another approach of classification by Markov fields in 2 stages. The first stage permits a classification in 5 classes (gray matter, white matter, cerebro-spinal liquid, mixtures of gray and white matter and mixtures of gray matter and liquid). The result of the first classification is used to operate a second classification more refined which permits to extract the 3 present tissues in image (gray substance, white substance and cerebro-spinal liquid). Mangin[23] proposes to introduce a priori local knowledge in the process of segmentation. The use of stochastic processes of optimization permits to get the classifications corresponding to global minima of energy[24,25]. The adaptability is also an important factor intervening in results quality of the segmentation that the approaches by classification can produce. The approaches proceeding by classification permit to get efficiently and quickly, in 2D as in 3D, very interesting segmentation results[26]. However, the knowledge that permits to guide these processes remains in the majority of cases less explicit in comparison with the approaches using models.

## PROPOSED METHOD

The approach that we proposed for the classification of the data is founded essentially on an evaluation of the Probability Density Function (p.d.f). The evaluation of this function constitutes one of the bases of the statistical data processing[27] and several methods permit to estimate the density are described in[28].

The Bayesian approach takes in account a model of p.d.f; the model that maximizes a likelihood criterion is kept to estimate the density. We can notice that the proposed models make some hypotheses implicitly on the number of classes or on the shape of these classes. The density of probability will be estimated by the kernel method called Parzen method[29] which is not based on any previous model. In what follows, we propose a study of the estimation of the probability density by the kernel method[30]. This estimation depends on a parameter of smoothing whose importance will be underlined. The multi-dimensional data of which we arrange are considered like a sample of a population, the population being constituted by the set of all potentially possible data. The density of probability is the function describing the distribution of this population.

**Estimation of the probability density function:** The data to classify are in a space $\mathbb{R}^n$ of dimension n. Let's consider a sample of N data $(X_i)_{1\le i\le N}$ represented by N points with: $xi = (x_{i,1}, x_{i,2}, \ldots, x_{i,n})$

The probability density function is defined on $\mathbb{R}^n$ by a function f satisfying at least to the following conditions:

$$\forall x \in \mathbb{R}^n, f(x) \ge 0 \text{ and } \int_{\mathbb{R}^n} f(x)\,dx = 1 \qquad (1)$$

In the practice, this function, f is unknown and several methods exist to estimate it. The most widespread method consists in approaching the function f by a histogram, but the gotten evaluation doesn't present the properties of continuity required by most applications. Besides, the widths of the classes (or elements) of the histogram are delicate to determine. Among the methods of evaluation of the density by a continuous function, one of the most known is the method of the nearest neighbors. The evaluation of the density function is defined by:

$$\hat{f}(x) = \frac{k-1}{2Nd_k(x)} \qquad (2)$$

Where k is the fixed number of the nearest neighbors (typically $k = \sqrt{N}$ according to Silverman[2]) and $d_k(x)$ is the distance of x to its $k^{ieme}$ nearest neighbors (while classifying by growing order the distances of x to each of the N data, $d_k(x)$ is the $k^{ieme}$ distance).

The estimation by the method of the k nearest neighbors doesn't permit to get a derivable function and besides it will give peaks corresponding to data situated in the densest regions. In the literature, we consider that these continuous evaluations but non derivable of the density are not smoothed enough and remain very approximate[28].

The methods solely based on the maximum of likelihood permit to drift the estimated function, but the interpolation between the data is merely heuristic. The likelihood of the estimated f of density for the samples $(X_i)_{1\le i\le N}$ is defined:

$$L\left(\hat{f}/(X_i)_{1\le i\le N}\right) = \prod_{i=1}^{N} \hat{f}(x_i) \qquad (3)$$

Maximize the likelihood doesn't permit to estimate the density outside of the sample $(X_i)_{1\le i\le N}$ (i.e. $\hat{f}(x)$ for $x \ne X_i$). It is necessary to impose some restrictions to $\hat{f}$ permitting an acceptable interpolation of $\hat{f}$ between the data. These restrictions or penalties have for goal to smooth the estimation of the function. In dimension 1, we most often use a criterion of the following type:

$$\ell_\alpha(f) = \sum_{i=1}^{N} \log f(x_i) - \alpha \int_{-\infty}^{\infty} (f'')^2 \qquad (4)$$

The first term $\sum_{i=1}^{N} \log f(x_i)$ is the logarithm of the likelihood that it is necessary to maximize, the second term $\alpha \int_{-\infty}^{\infty} (f'')^2$ is a penalty that it is necessary to minimize and that is controlled by the parameter $\alpha$. This parameter is more or less adjusted to smooth the estimation of the density[28]. In dimension 2 or more, this method of the penalties is more delicate to put in work, it is necessary to find the good compromised between the parameter of smoothing and the likelihood. In this type of approach, the part of heuristic is important. Because of this rigor's lack, it seemed preferable to us to estimate the density by the classic method of the kernel (even named Parzen-Rosenblatt method) described below.

**Parzen-rosenblatt method:** In Parzen-Rosenblatt method, we estimate the density of probability while using a convolution kernel. The kernel is a function K that is generally itself a function of probability density. In this description, we take like kernel the multi-normal function (centered and reduced) definite by:

$$\forall x \in \mathbb{R}^n, \quad \hat{f}(x) = \frac{1}{Nh^n} \sum_{i=1}^{N} K\left(\frac{x - X_i}{h}\right) \qquad (5)$$

h being the parameter of smoothing of the estimation.

In this density approach, each data $X_i$ contribute in the same way to the calculation of $\hat{f}$ and this contribution depends on h. The kernel K being a unimodal and positive function, the contribution of every data to $\hat{f}$ be added and is worth at more $\frac{1}{Nh^n}$ ( the maximal contribution of a data $X_i$ to $\hat{f}(x)$ is gotten when $x = X_i$).

This estimation of the probability density corresponds to a convolution of the function K with the function $\Delta$ definite by:

$$\Delta(x)=\frac{1}{N}\sum_{i=1}^{N}\delta_i(x) \quad \text{with } \delta_i(x)=\begin{cases}0 & \text{if } x=X_i\\ 1 & \text{otherwise}\end{cases} \tag{6}$$

The parameter h corresponds to the square root of variance of the kernel K. More h, will be small, more the kernel will be "narrow" and $\hat{f}$ will present some "peaks" of probabilities to the points $X_i$. In this setting, h be called the window of the estimation or window of smoothing. This type of estimation of the probability density function depends on the choice of the smoothing window.

**Evaluation criteria:** The density approach of f, by the Parzen method, suppose that we can determine an optimal value of the window h.

The estimation $\hat{f}$ of the density f can be valued by a criteria that measures a distance between $\hat{f}$ and f. Such a criteria supposes the existence of a functional space in which is defined a distance between functions. Several functional spaces are used in the literature[28].

The most known functional space is the space $\Im^2$ for which we define a distance between two functions that we will note $d_1$. The distance $d_1(f,g)$ between two functions f and g can be defined by: $d_1(f,g)=\int_{\Omega^n}(f-g)^2$

In the discreet space D, this distance is more known as criteria of the least squares noted LS (Least Squares). This criteria is defined by:

$$LS(\hat{f})=\sum_{x\in D}\left(f(x)-\hat{f}(x)\right)^2 \tag{7}$$

To determine the optimal value of h, it is necessary to minimize this criteria $LS(\hat{f})$.

In the functional space $\Im^1$, we define the distance $d_2(f,g)=\int_{\Omega^n}|f-g|$.

In the discreet space D, this distance corresponds to the sum of the absolute values of the differences between the estimated $\hat{f}$ and f. This distance is noted SAVD (Sum of Absolute Values of Differences) and is defined by:

$$SAVD(\hat{f})=\sum_{x\in D}\left|f(x)-\hat{f}(x)\right| \tag{8}$$

There again, this criteria will be minimized to determine the optimal value of h. This last criteria is more robust than LS in relation to possible aberrant data.

We have therefore two criteria LS and SAVD hat it is necessary to minimize to determine the optimal value of the window h. We note that these two criteria reach their minimum for the same value of if the discretization of the space is sufficient to really separate the different data.

**Unsupervised classification by iterative SKIZ:** The probability density f is defined on a space D to included values between 0 and a maximal value noted max f. To explain our procedure, we are going to consider a threshold s included between 0 and f. The thresholded part $P_s$ of the space D is defined by:

$$P_s=\{x\in D\setminus f(x)\geq s\} \tag{9}$$

This thresholded part contains related components (i.e. components in only one piece). More precisely, a set is connected when two any points of this set can be related by a path who is himself on the considered set.

We have L levels for a f.d.p going from 0 to L-1. To every level e ($0 \leq e < L$) we associate $S_e$ the set of points of which the density level is superior to e (i.e. $S_e \subseteq S$). we use this set to define the temporary class of S to every level of density. First, some voluminous classes are defined. Then the classification is refined while incrementing the every time the level e. This sophistication is obtained by the division of the previous classes while using the procedure of the SKIZ that we are going to describe in the following paragraphs.

In this paper, we stand in the setting of a discreet topology or the connection is established while fixing the neighbors of every pavement of the discreet space. This kind of topology is known well in image processing where, in dimension 2, we also use the 4-connexity or the 8-connexity.

While using the neighborhood network of dimension n of the points of the S, the set $S_e$ is composed of the connected regions (i.e. sets of the connected components). $C(S_e)$ definite the number of the connected regions. The algorithm of classification is described by the use of an example (Fig. 1) who shows the classification process for 4 levels (L = 4) in a data space of dimension 2(n = 2). Every point is the prototype of a segment of the partition domain, the gray and black zones show the neighborhoods of the different points (i.e. the neighborhood network). The following paragraph illustrates the different stages of the classification algorithm by iterative SKIZ.

**Stages of the classification algorithm by iterative:** SKIZ

- Stage 0 (Level L = 0)
  L = 0⇒ $S_0$ = S, therefore only one connected region ($c(S_0) = 1$)(only one connected region); e = e+1
- Stage 1 (Level L = 1)
  L = 1⇒ $S_1 \subseteq S_0 \subseteq S$, and ($c(S_0) = 2$) (2 connected regions); e= e+1;
- Stage e (Level L = e)
  L = e ⇒ $S_e \subseteq S$ and ($c(S_0) = C$) (to the level e there is C classes); e = e+1;
- Stage e+1( Level l = e+1)
  L = e+1 ⇒ $S_{e+1} \subseteq S_e \subseteq S_e$ and ($c(C \cap S_{e+1})>1$).

**Regulation by markov random field:** The Markov random fields are extensively used in images analysis[31]. The Markovian modeling allows us to introduce contextual information and therefore to segment our image.

The Markovian random fields or simply Markovian Fields (MF) permit the modeling of the global properties while using local constraints.

We note S(card(s) = N) the network of the image sites and $V = V_s | \ s \in S$ the neighborhood system ($V_s$ is the neighborhood of the site s).

We have $X = \{X_1,...,X_N\}$ a family of random variables defined on the network S. Each random variable $X_i$ takes the value xi ($x_i \in \Omega$). In the same way we note $X = x, x \in \Omega^N$, to mean the following event: $X = x_1,...,x_N = x_n$. x, called configuration, corresponds to one realization of the random field X. This configuration has a certain probability that we note P(X = x). is called Markov field in relation to the system of neighborhood if it verifies the 2 properties follow:

$$- \text{ (positivity) } \forall x \in \Omega^N, P(X = x) > 0 \tag{10}$$

$$\begin{aligned} &- \text{ (markovian propriety)} \\ &\forall s \in S \text{ and } \forall x \in \Omega^N, \\ &P(X_s = x_s \mid X_r = x_r, r \in S - \{s\}) = \\ &P(X_s = x_s \mid X_r = x_r, r \in V_s) \end{aligned} \tag{11}$$

In the continuation, to facilitate the notations, we confound Xand x. P(X = x) is noted then P(x).

The theorem of Hammersley-Clifford[31] makes the relation with the Gibbs fields. This theorem permits to make the relation between the property of locality of the Markovian fields (Markovian property) and the property of globality of the Gibbs fields (Gibbs distribution). X is a Markovian field in relation to the neighborhood system V if and only if its distribution P(X) is a Gibbs distribution, i.e. :

$$P(X) = \frac{\exp(-E(X))}{Z} \tag{12}$$

$Z = \sum_{X \in \Omega^N} E(X)$ is a constant of normalization named function of partition.

E(X) is a energy function defined like follows:

$$E(X) = \sum_{cl \in CL} V_{cl}(X) \tag{13}$$

cl is a clique, i.e. a set of two to two neighboring sites or a singleton site.

CL is the set of the cliques related to the considered neighborhood system.

$V_{cl}(X)$ is a function of potential.

The global energy E(X) can be written therefore like a sum of local energies (or sum of potentials).

We used the Markovian modeling to regularize our image while introducing contextual information by the criteria of the MPM (Maximum Posterior Marginal)[32] that minimizes the number of pixels badly classified. It is therefore more adapted to applications of classification-segmentation that the criteria of the MAP (Maximum A Posteriori)[33].

We note X the random field relative to the observed image where its values are in $\Omega_N$.

We also note L the random field relative to the labeled image to values in $\Lambda^N$.

X being known, we look for the configuration L that maximizes the following expression:

$$P(L \mid X) = \frac{P(X \mid L) P(L)}{P(X)} \tag{14}$$

what leads us, X being known and therefore P(X) constant, to maximize the following expression:

$$P(X \mid L) P(L) \tag{15}$$

The data are modeled (model of texture, model of noise[34]…) what allows us to define the law P(X|L) on $\Omega^N$. We model the field L by a Markovian field what allows us to define a priori law P(L) on $\Lambda^N$. The priori model that we consider in the following chapter is the model of Potts.

The approach of the MPM consist in maximizing the posteriori marginal probabilities. Thus, the label $\hat{l}_s$ retained for the site is the one that maximizes the following probability:

$$\begin{cases} P(L_s = \hat{l}_s \mid X = x) \geq P(L_s = l_s \mid X = x) \\ \forall l_s \in \Lambda \end{cases} \tag{16}$$

The estimator of the MPM is obtained by Gibbs sampler[31]. The algorithm consists in generating a Markov chain on $\Lambda^N$. Once the convergence reached, the posteriori marginal probabilities at s are estimated while counting the number of time where each value of the set $\Lambda$ appears at s for a set of configurations. The label $\hat{l}_s$ takes the value that appears the biggest number of time at s.

We considered that the data follow a Gaussian law of which the parameters (mean and variance of the classes) are already estimated from the classification relatively precise (but that doesn't hold account of no contextual information), result of the SKIZ algorithm. The attachment law of considered data is the following:

$$P(X|L)=\prod_{i\in S}\frac{1}{\sqrt{2\pi\sigma_{l_i}}}\exp\left(\frac{(x_i-\mu_{l_i})^2}{2\sigma_{l_i}^2}\right) \quad (17)$$

$\sigma^2_j$ and $\mu_j$ are respectively the variance and the mean of the class j.

The priori considered model is the model of Potts. The priori probability is the following:

$$P(L)=\frac{1}{Z_\beta}\exp\left(-\beta\sum_{cl=\{l_i,l_j\},cl\in CL}\delta(l_i,l_j)\right)$$

$$\delta(l_i,l_j)=1, \text{ if } l_i=l_j$$

$$\delta(l_i,l_j)=0, \text{ if } l_i\neq l_j$$

## SEGMENTATION ALGORITHM

In this part we are going to describe in detail the different stages of our hybrid algorithm that combine at a time a non parametric method of estimation of the probability density function, a mathematical morphology method for the classification of these data and another method based on the Markov fields for a final contextual segmentation.

- Stage 1: (estimation of the p.d.f)
  - Fix the kernel K,
  - $\forall x \in \mathbb{R}^n$, compute $\hat{f}(x)$ the estimation of f(x) by the formula ( 5).
  - Under the constraint of the minimization of the two criteria LS and SAVD (formulas (7) and (8) respectively) optimize the size of the window h.
- Stage 2: (classification by skiz)
  - Let $m=\max_{0\le i\le S}(p.d.f(x_i))$
  - Let ta threshold between 0 and m ($o\le t < m$)
  - Apply the algorithm describes in (4.5) while incrementing every time t by $\delta$ ($t = t+\delta$)
- Stage 3: (Segmentation by Markov random fields)
  - Maximize the posteriori marginal probabilities (MPM) according to the formula (14) with as conditions the constraints described respectively by the formulas (17) and (18).

## RESULTS

We applied the algorithm on different types of images. First of all, we have validated our classification algorithm on synthesis images (Fig. 1 A-E) that we generated by hand. Each of these images is composed respectively of 400 points on a 2D space. We were able to noted that for the separation, by influence zones, the classification thus obtained is very obvious and the

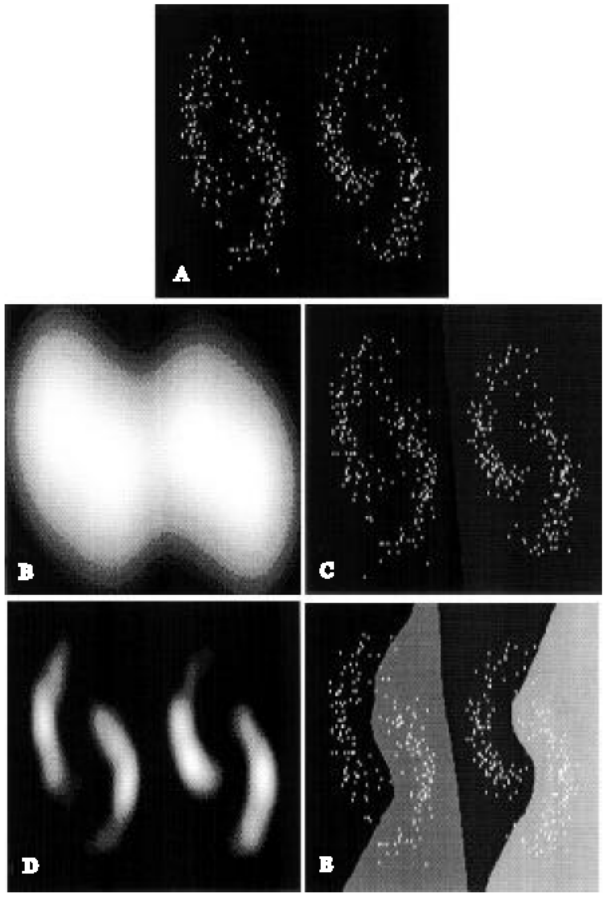


Fig. 1: Two estimations of the p.d.f and classification of simulated data. A: 400 points of simulated data, B: Estimation of the p.d.f with h=20, C: Detection of 2 influence zones, D: Estimation of the p.d.f with h=7, E: Detection of 4 influence zones

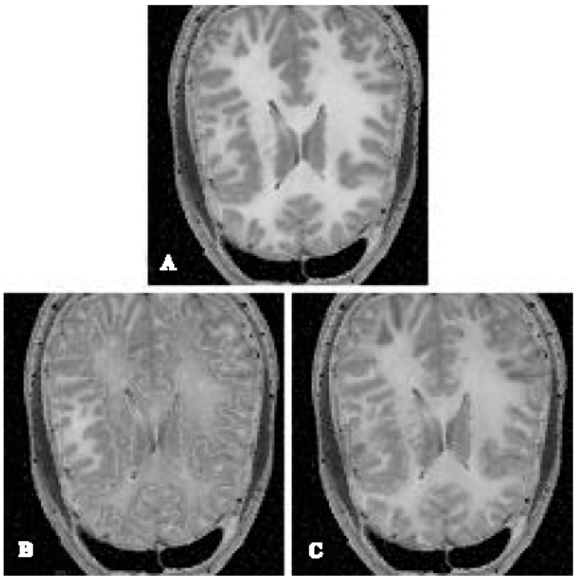


Fig. 2: Classification of the white matter. A: T1-weighted image (axial cut), B: skiz (h=20) (blue) +MRF (red), C: skiz (h=7) (blue) +RMF (red)

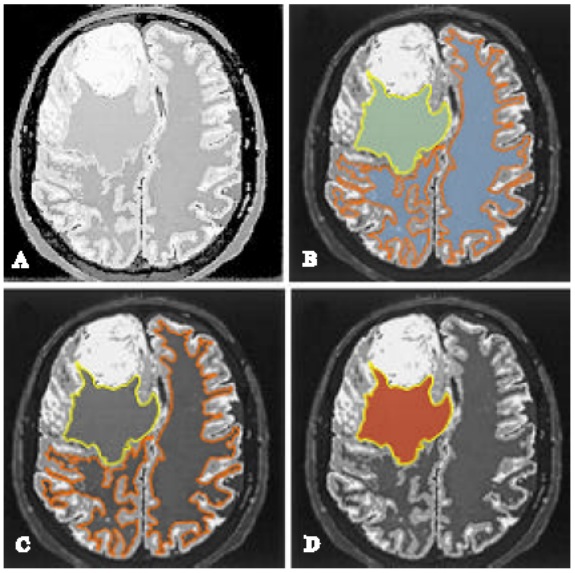


Fig. 3: Extraction of a tumor (meningioma). A: PD-weighted image (axial cut), B: Data classification by SKIZ (h=7), C: SKIZ segmentation (h=7)+MRF Regulation, D: Final segmentation and detection of the tumor (in red)

number of classes by skiz, who ensues of this last, is very close to the number of real classes. It is while necessary to underline that these results depend on the size of smoothing window h (Fig. 1). The parameter h has an effect of smoothing on the data. Indeed, if h is too big the density is consequently smoothed and the number of classes is less important. With a reduced size of h, we notice the apparition of supplementary peaks and therefore of supplementary classes.

After having proven the efficiency of the classification algorithm, we proceeded to the RMI segmentation. The image 2(A) represent an axial cut of a human brain T1-weighted, the image 2(B) represent a classification by skiz, whereas the image 2(C) represent a segmentation after Markovian regulation by using the law of data attachment described by the Eq. 17 and 18. The introduction of the contextual information by the Markovian fields seems improved the results of the segmentation (Fig. 2 B and C).

The algorithm is used for the extraction of a cerebral pathology of a picture so-called RMI in density of proton (image 1 A). After the evaluation of the probability density, a classification by iterative SKIZ is operated (Fig. 3 B), the contextual information of the texture and to the distribution of the pathology are introduced by the Markovian random fields, what improved the classification considerably (Fig. 3 C) and that permitted to extract the pathology thereafter (Fig. 3 D).

## CONCLUSION

Our method of data classification of medical images rests on the iterative SKIZ with successive thresholding of the density function. Our original approach of the unsupervised classification requires no previous hypothesis nor on the shape nor on the number of classes. For this reason, this method of classification is a qualitative progress in relation to most classic methods that suppose implicitly or explicitly that the classes are topologically convex. However it remains again to compare the different methods of classification while using quantitative criteria. Cooperation between several techniques of segmented seems to be efficient. Indeed, the regulation by Markov fields pushes the segmentation farther while introducing contextual information. However, this technique will depend on several parameters to know the kernel for the estimation of the probability density , the size of the smoothing window and the law of data attachment.

## REFERENCES

1. Kapur, T. *et al*., Segmentation of brain tissue from magnetic resonance images, Medical Image Analysis, pp: 109-127.
2. Bezdek, J.C. and N.R. Pal, 1998. Some new indexes of cluster validity. IEEE Trans. Systems Man. Cybern., pp: 301-315.

3. Liew, A.W.-C., Hong Yan, 2003. Adaptive fuzzy segmentation of 3D MR brain images, Fuzzy Systems, The 12th IEEE International Conference on Fuzzy System, pp: 25-28.
4. Chiou, G. and J.N. Hwang, 1995. A Neural Network-Based Stochastic Active Contour Model (NNS-SNAKE) for Contour Finding of Distinct Features, IEEE Transactions on Image Processing, pp: 1407-1416.
5. Moussaoui, A. and V. Chen, 2004. Segmentation neuro-floue des tissus cérébraux pathologiques en IRM 3D, In Proc. of International Symposium on Medical Image Processing, Blida, Algeria, pp: 81-86.
6. Davatzikos, C. and J.L. Prince, 1995. An active contour model for mapping the cortex, IEEE Transactions on Medical Imaging, pp: 65-80.
7. Vaillant, M. and C. Davatzikos, 1997. Finding parametric representations of the cortical sulci using an active contour model, Medical Image Analysis, pp: 295-315, 1997.
8. Barillot C., *et al*., 1998. Représentation mixte numérique/symbolique des sillons corticaux. In Proc. of 11$^{e}$ congrès Reconnaissance des Formes et Intelligence Artificielle. Paris, France, pp: 165-174.
9. Sandor, S. and R. Leahy, 1997. Surface-Based Labeling of Cortical Anatomy Using a Deformable Atlas, IEEE Transactions on Medical Imaging, pp: 41-54.
10. Staib, L.H. and J.S. Duncan, 1992. Boundary Finding with Parametrically Deformable Models, IEEE Transactions on Pattern Analysis and Machine Intelligence. pp: 1061-1075.
11. Bardinet E., L. Cohen and N. Ayache, 1995. A parametric deformable model to fit unstructured 3D data. Technical Report 2617, INRIA Sophia Antipolis.
12. Soltanian-Zadeh, H. *et al*., 2004. Model-independent method for fMRI analysis , EEE Transactions on Medical Imaging. pp: 285-296.
13. Cootes, T.F. *et al*., 1995. Active Shape Models- Their Training and Application, Computer Vision and Image Understanding, pp: 38-59.
14. Duta, N. and M. Sonka, 1998. Segmentation and Interpretation of MR Brain Images : An Improved Active Shape Model, EEE Transactions on Medical Imaging, pp: 1049-1062.
15. Gong, L. and A. Kulikowski, 1995. Composition of image Analysis Processes Through Object-centered Hierarchical Planning, IEEE Transactions on Pattern Analysis and Machine Intelligence, pp: 997-1009.
16. Li, H., *et al*., Object-recognition in brain CTscans: knowledge-based fusion of data from multiple feature extractors, IEEE Transactions on Medical Imaging, pp: 212-229.
17. Dellepiane, S., G. Venturi and G. Vernazza, 1991. A fuzzy model for the processing and recognition of MR pathological images, In Lecture notes in computer science, IPMI proceedings, pp: 444-457.
18. Sonka, M., S. Tadikonda and A.M. Collins, 1996. Knowledge-Based Interpretation of MR brain Images, IEEE transactions on Medical Imaging. pp: 15.
19. Friston, K. *et al*., 1995. Statistical Parametric Maps in Functional Imaging : a general linear approach, Human Brain Mapping, pp: 189-210.
20. Wells, W.M. *et al*., 1996. Adaptative Segmentation of MRI data, IEEE Transactions on Medical Imaging, pp: 429-442.
21. Rajapakse, J., J. Giedd, 1997. Rapoport and J., Statistical Approach to segmentation of single-channel cerebral MR images, IEEE Transactions on Medical Imaging, pp: 176-186.
22. Jaggi, C., 1998. Segmentation par méthode markovienne de l'encéphale humain par résonance magnétique: Théorie, mise en oeuvre et évaluation, Thèse de doctorat de l'université de Caen.
23. Mangin, J.F. *et al*., 1995. A MRF based random graph modeling the human cortical topography, In First International Conference CVRMed, Nice, pp: 177-183.
24. Held, K. *et al*., 1997. Markov Random Field Segmentation of Brain MR Images. IEEE Transactions on Medical Imaging, pp: 878-886.
25. Géraud, T., 1998. Segmentation des structures internes du cerveau en imagerie par résonance magnétique 3D, Thèse de doctorat, Télécom Paris.
26. Moussaoui, A. and V. Chen, 2004. Fuzzy automatic classification without a prior knowledge-mean shift application to MRN brain images segmentation, In Symposium of Electronics and Telecommunications, Timisoara, Romania, pp: 213-219.
27. Herbin, M., N. Bonnet and P. Vautrot, 2001. Estimation of the number of clusters by influence zones, Pattern Recognition Letters, pp: 1557-1568.
28. Silverman, B.W., 1986. Density Estimation for Statistics and Data Analysis, Chapman and Hall, London.
29. Parzen, E., 1962. On estimation of a probability density function and mode. Ann. Math. Statist., pp: 1065-1076.
30. Moussaoui, A. and V. Chen, 2004. Segmentation d'images médicales par coopération de classifieurs, In Proc. of International Symposium on Medical Image Processing, Blida, Algeria, pp: 57-62.
31. Descombes, X., M. Sigelle and F. Préteux, 1998. Estimating Gaussian Markov random field parameters in a non-stationary framework, IEEE Transaction On Image Processing.

32. Palubinskas, G., S. Miter and T. Poggio, 1998. An unsupervised clustering method using the entropy minimisation. In ICPR. Austria.
33. Descombes, X., R. Moris and J. Zerubia, 1997. Quelques améliorations à la segmentation d'images bayesienne, Traitement du signal, pp: 373-382.
34. Geman, D., G. Renolds, 1992. Constrained restoration and the recovery of discontinuities, IEEE Transaction On Pattern Analysis and Machine Intelligence, pp: 367-383.
35. Sigele, M. and R. Ronfard, 1992. Modèles de potts et relaxation d'images de labels par champs de Markov, Traitement du signal, pp: 449-458.