International Journal of New Computer Architectures and their Applications (IJNCAA) 3(1): 35-44
The Society of Digital Information and Wireless Communications (SDIWC) 2013 (ISSN: 2220-9085)

Medical Image Segmentation on a Cluster of PCs using Markov Random Fields

El-Hachemi Guerrout*, Ramdane Mahiou, Samy Ait-Aoudia

ESI - Ecole nationale Supérieure en Informatique, BP 68M, 16270, Oued-Smar, Algiers, Algeria
e_guerrout@esi.dz, r_mahiou@esi.dz, s_ait_aoudia@esi.dz

Abstract. Medical imaging applications produce large sets of similar images. The huge amount of data makes the manual analysis and interpretation a fastidious task. Medical image segmentation is thus an important process in image processing used to partition the images into different regions (e.g. gray matter(GM), white matter(WM) and cerebrospinal fluid(CSF)). Hidden Markov Random Field (HMRF) Model and Gibbs distributions provide powerful tools for image modeling. In this paper, we use a HMRF model to perform segmentation of volumetric medical images. We have a problem with incomplete data. We seek the segmented images according to the MAP (Maximum A Posteriori) criterion. MAP estimation leads to the minimization of an energy function. This problem is computationally intractable. Therefore, optimizations techniques are used to compute a solution. We will evaluate the segmentation upon two major factors: the time of calculation and the quality of segmentation. Processing time is reduced by distributing the computation of segmentation on a powerful and inexpensive architecture that consists of a cluster of personal computers. Parallel programming was done by using the standard MPI (Message Passing Interface).

Keywords. Medical image segmentation, Hidden Markov Random Field, Gibbs distribution, Iterated Conditional Modes, Cluster of PCs, MPI (Message Passing Interface)

1. INTRODUCTION

Medical image segmentation is thus an important process in image processing used to partition the images into different regions (e.g. gray matter(GM), white matter(WM) and cerebrospinal fluid(CSF)). Various methods were used to perform the segmentation task. Among this wide variety, HMRF Model and Gibbs distributions provide powerful tools for image modelling [1,2,14,15,18]. Since the seminal paper of Geman and Geman[10], Markov Random Fields (MRF) models for image segmentation have been investigated by many other researchers [3,12,13,19,21]. In this paper, we use a Hidden Markov Random Field (HMRF) model to perform segmentation of volumetric medical images (3D segmentation). We have a problem with incomplete data. We seek the segmented images according to the MAP (Maximum A Posteriori) criterion [22]. MAP estimation leads to the minimization of an energy function. This problem is computationally intractable. Therefore, optimizations techniques [6,7,17,20] are used to compute a solution. Choosing a good optimization method is a crucial task. A poor optimization process can lead to "disastrous" results. We will use the well known method ICM (Iterated Conditional Modes). Quantization error and dmin, dmax and Dice coefficients (Kappa Indexes) are used for the quality measures of segmantation. The dmin shows the minimum distance between different classes means(error inter-classes) while dmax shows maximum internal error of classes(error intra-classe). The segmentation evaluation is made by calculating Dice coefficients (Kappa Indexes) only when we have the ground truth images (a priori segmented images known) that gives the similarity between segmented images and the a priori labeled images. Time processing will be evaluated by the use of a common network that is a cluster of PCs. Parallel programming was done by using the standard MPI (Message Passing Interface). The efficiency of parallelization is proved by good acceleration factors or speed-up.

This paper is organized as follows. We remind in section 2 basis of Markov Random Field model. In section 3, we give principles of Hidden Markov Field model in the context of image segmentation and describe the ICM technique. The parallel architecture used is described section 4. Experimental results on medical samples datasets are given in section 5. Section 6 gives conclusions.

2. MARKOV RANDOM FIELD MODEL

In this section we remind some important notions relative to Markov Random Field model and some terms used in the context of image analysis issues.

2.1. Neighborhood System

The pixels of the image are represented as a lattice S of $M=n*m$ sites.

$$S=\{s_1,s_2,\ldots,s_M\}$$

In an MRF, the sites (pixels in our case) in S are related by a neighborhood system V(S) having the following properties:

$$\forall s \in S,\ s \notin V_s(S)$$

$$\forall \{s,t\} \in S,\ s \in Vt(S) \Leftrightarrow t \in Vs(S)$$

The relationship V(S) expresses a neighborhood constraint between adjacent sites. An r-order neighborhood system noted $V^r(S)$ is given by the following formula :

$$Vrs(S)=\{t \in S \mid d(s,t)^2 \leq r,\ s \neq t\},\ s \in S$$

Where $d(s,t)$ is the Euclidean distance between s and t.

The first and second order neighborhood systems are the most commonly used. In these systems, a site has four and eight neighbors respectively. When a site has four or eight neighbors, we speak about a 4-neighborhood or an 8-neighborhood as shown in figure 1.

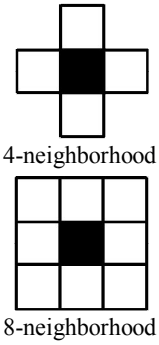


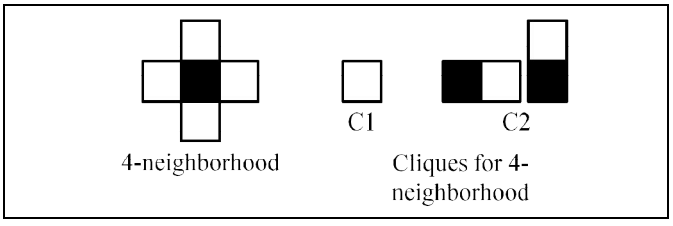
Figure 1. Neighborhood system

2.2. Clique

A clique c is a subset of sites in S relatively to a neighborhood system. c is a singleton or all the distinct sites of c are neighbors. For a non single-site clique we have:

$$\forall \{s,t\} \in c,\ t \in Vs(S)$$

A *p-order* clique noted c_p contains p sites i.e. p is the cardinal of the clique. Figure 2 shows some cliques given a neighborhood system.



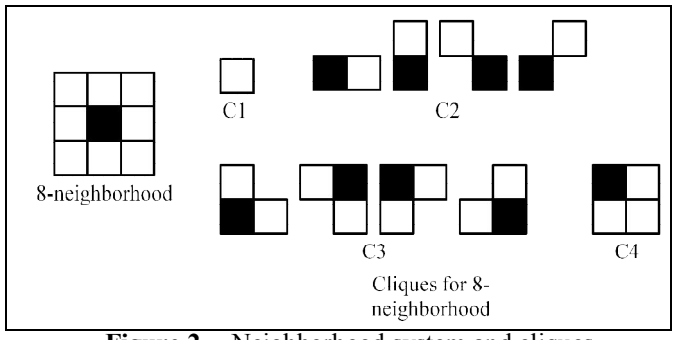


Figure 2. Neighborhood system and cliques

2.3. Markov Random Field

Let $X=\{X_1,X_2,...,X_M\}$ be a family of random variables on the lattice S. Each random variable taking values in the discrete space $\Lambda=\{1,2,...,K\}$. The family X is a random field with configuration set $\Omega = \Lambda^M$.

A random field X is said to be an MRF on S with respect to a neighborhood system $V(S)$ if and only if :

$$\forall x \in \Omega,\ P(x) > 0$$

$$\forall s \in S,\ \forall x \in \Omega, P(X_s=x_s/X_t=x_t, t \neq s)=P(X_s=x_s/X_t=x_t, t \in V_s(S))$$

The Hammersley-Clifford theorem establishes the equivalence between Gibbs fields and Markov fields. The Gibbs distribution is characterized by the following relation:

$$P(x)= Z^{-1} e^{-\frac{U(x)}{T}}$$

$$Z = \sum_{y \in \Omega} e^{-\frac{U(y)}{T}}$$

where T is a global control parameter called temperature and Z is a normalizing constant called the partition function. Calculating Z is prohibitive. $Card(\Omega)=2^{2097152}$ for a *512x512* gray level image. $U(x)$ is the energy function of the Gibbs field defined as :

$$U(x)=\sum_{c \in C} U_c(x)$$

$U(x)$ is defined as a sum of potentials over all the possible cliques C.

The local interactions between the neighbor sites properties (gray levels for example) can be expressed as a clique potential.

2.4. Standard Markov Random Field

We shortly review in this section standard Markov random fields used for image analysis purposes.

2.4.1. Ising Model

This model was proposed by Ernst Ising for ferromagnetism studies in statistical physics. The Ising model involves discrete variables s_i (spins) placed on a sampling grid. Each spin can take two values, $\Lambda=\{-1,1\}$, and the spins interact in pairs. The first order clique potential are defined by $-Bx_s$ and the second order clique potential are defined by:

$$U_{c=\{s,t\}}(x_s,x_t)=-\beta x_s x_t=\begin{cases}-\beta & if \quad x_s=x_t\\ \beta & if \quad x_s\neq x_t\end{cases}$$

The total energy is defined by :

$$U(x)=-\sum_{c=\{s,t\}}\beta x_s x_t+\sum_{s\in S}Bx_s$$

The coupling constant β between neighbor sites regularize the model and B represents an extern magnetic field.

2.4.2. Potts Model

The Potts model is a generalization of the Ising model. Instead of $\Lambda=\{-1,1\}$, each spin is assigned an integer value $\Lambda=\{1,2,...,K\}$. In the context of image segmentation, the integer values are gray levels or labels. The total energy is defined by :

$$U(x)=\beta\sum_{s,t\in C_2}(2\delta(x_s,x_t)-1)$$

where δ is the Kronecker's delta.

When β>0, the probable configurations correspond to neighbor sites with same gray level or label. This induces the constitution of large homogenous regions. The size of these regions is guided by the value of β.

3. HMRF MODEL

3.1. Hidden Markov Random Field

A strong model for image segmentation is to see the image to segment as a realization of a Markov Random Field $Y=\{Ys\}_{s\in S}$ defined on the lattice S. The random variables $\{Ys\}_{s\in S}$ have gray level values in the space $\Lambda obs=\{0..255\}$. The configuration set is Ω obs.

The segmented image is seen as the realization of another Markov Random Field X defined on the same lattice S, taking values in the discrete space $\Lambda=\{1,2,...,K\}$. K representing the number of classes or homogeneous regions in the image.

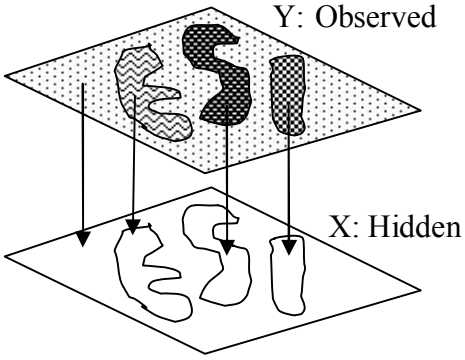


Figure 3. Observed and hidden image.

In the context of image segmentation we have a problem with incomplete data. To every site $i\in S$ is associated two different information. Observed information expressed by the random variable Y_i and a missed or hidden information expressed by the random variable X_i. The Random Field X is said Hidden Markov Random Field.

The segmentation process consists in finding a realization x of X by observing the data of the realization y representing the image to segment.

3.2. MAP Estimation

We seek a labeling $\hat{x}$ which is an estimate of the true labeling x^*, according to the MAP (Maximum A Posteriori) criterion (maximizing the probability $P(X=x|Y=y)$).

$$\hat{x}=arg\max_{x\in X}\{P(X=x|Y=y)\}$$

$$P(X=x|Y=y)=\frac{P(Y=y|X=x)P(X=x)}{P(Y=y)}$$

The first term of the numerator describe the probability to observe the image y knowing the labeling x. Based on the conditional independence assumption the pixels, the joint likelihood probability is given by :

$$P(Y=y|X=x)=\prod_{s\in S}P(Y_s=y_s|X_s=x_s)$$

The second term of the numerator describe the existence of the labeling x. The denominator is constant and independent of x. We have then :

$$P(X=x|Y=y)=KP(Y=y|X=x)P(X=x)$$

$$P(X=x|Y=y)=Ke^{ln(P(Y=y|X=x))-\frac{U(x)}{T}}$$

$$P(X=x|Y=y)=Ke^{-\Psi(x,y)}$$

The labeling $\hat{x}$ can be found by maximizing the probability $P(X=x|Y=y)$ or equivalently by minimizing the function $\Psi(x|y)$.

$$\Psi(x,y)=-ln(P(Y=y|X=x))+\frac{U(x)}{T}$$

$$\Psi(x,y)=-\sum_{s\in S} ln(P(Y_s=y_s|X_s=x_s))+\frac{1}{T}\sum_{c\in C} U_c(x)$$

$$\Psi(x,y)=\Psi_1(y|x)+\Psi_2(x)$$

$$\Psi_1(y|x)=-\sum_{s\in S} ln(P(Y_s=y_s|X_s=x_s))$$

$$\Psi_2(x)=\frac{1}{T}\sum_{c\in C} U_c(x)$$

$$\hat{x}=arg\min_{x\in X}\{\Psi(x,y)\}$$

The searched labeling $\hat{x}$ can be found using some optimization techniques.

Assuming that the pixel intensity follows a Gaussian distribution with parameters μ_k (mean) and σ_k^2 (variance) given the class label $x_s=k$, we have :

$$P(Y_s=y_s|X_s=k)=\frac{1}{\sqrt{2\pi\sigma_k^2}}e^{-\frac{(y_s-\mu_k)^2}{2\sigma_k^2}}$$

$$\Psi_2(x)=-\frac{\beta}{T}\sum_{s,t\in C_2}\phi(x_s,x_t)$$

The Potts model is often used in image segmentation to privilege large regions in the image. The energy is then :

$$\Psi_2(x)=-\frac{\beta}{T}\sum_{s,t\in C_2}(2\delta(x_s,x_t)-1)$$

$$\Psi(x,y)=\sum_{s\in S}\frac{(y_s-\mu_{x_s})^2}{2\sigma_{x_s}^2}+ln(\sqrt{2\pi}\sigma_{x_s})+\frac{\beta}{T}\sum_{s,t\in C_2}(1-2\delta(x_s,x_t))$$

3.3. ICM Method

The MAP estimation leads to the minimization of an energy function. This problem is computationally intractable. Therefore, optimizations techniques are used to compute a solution. We will use the well known method that is ICM.

The Iterated Conditional Modes (ICM) algorithm proposed by Besag[4], is a deterministic relaxation scheme with a constant temperature. Performances of the ICM algorithm tightly depend on the initialization process. It converges toward the local minimum close to the initialization. The following Algorithm summarizes the ICM technique.

ICM Algorithm:

1. Initialization: Start with an arbitrary labeling x^0 and let $n=0$.

2. At step n:

Visit all the sites according to a visiting scheme and in every site s:

$$x_s^{n+1}=\arg\min_{x\in\Omega^{card(S)}} U_s(x_s=\lambda), \quad \lambda\in\Omega .$$

3. Increment n. Goto 2, until a stopping criterion is satisfied.

4. PARALLEL ARCHITECTURE

4.1. Cluster of PCs

Clustering is a technique to configure multiple machines belonging to a general network for parallel purposes. A cluster consists therefore of a set of nodes interconnected by a fast LAN. The cluster becomes nowadays one of the most common parallel architectures for obvious reasons of cost. A cluster of PCs can match a supercomputer in terms of performance. The following figure describes a cluster.

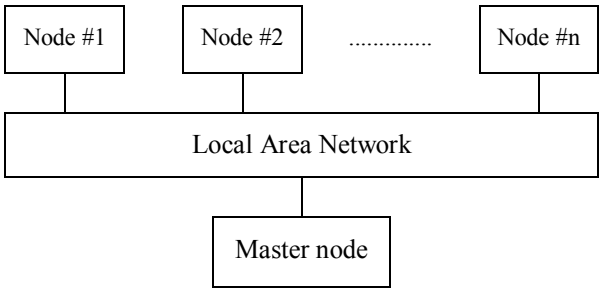


Figure 4. A cluster of PCs

4.2. Speed-up (Accélération)

Designing a parallel architecture to accelerate the calculations is a good option but quantifying its contribution to performance is crucial. The acceleration or commonly used under the term speed-up allows us to quantify the gain in terms of execution time. Let T (1) be the time required for a program to solve the problem A on a sequential machine and let T (p) be the time required for a program to solve the same problem on a parallel architecture containing p processors. The speed-up is given by the following relationship :

$$S(P) = \frac{T(1)}{T(P)} \Rightarrow \begin{cases} S(P) < 1 : Slow\ down\ !(bad\ parallelization) \\ 1 < S(P) < P : Normal \\ P < S(P) : Hyper - acceleration\ \ (analyze\ and\ justify) \end{cases}$$

The speedup is therefore the execution time gain of a parallel program over the same sequential program.

4.3. Parallel Program

The core parallel program of our application distributes the segmentation tasks over the PCs. The results are assembled by a specialized PC called the coordinator. The segmented images are then visualized. The following algorithm summarizes the parallel algorithm.

```
Distribute the calculation task between the PCs and for each task do :
  initialization:
  Initialize n = 0 and T = t_max a high enough temperature
  Optimizing the initial configuration x (0).
  Compute x (n +1) from x (n):
  * Explore all sites s (according to a strategy of site visit)
```

$$x(n+1) = argmin_{x_s \in S}\left(\sum_{s \in S}\left[\ln(\sigma_{x_s}) + \frac{(y_s - \mu_{x_s})^2}{2\sigma_{x_s}^2}\right] + \beta \sum_{c_t = (s,t)} (1 - 2\delta(x_s, x_t))\right)$$

```
  n= n + 1
  GOTO (*) until achieving a stopping criterion
Assemble the results of the tasks in the PC coordinator
```

5. EXPERIMENTAL RESULTS

The evaluation of the segmentation is made on volumetric medical data samples. All images were gray-level, and were scaled to 8 bits/pixel. The cluster of PCs that we used in our experiments consists of eight identical machines related by a switch (Catalyst 3560G). Its characteristics are listed in the table below.

Table 1. Architecture description

System Information	CPU Information	Memory Information	Network Information
Linux, the Ubuntu 11.04, gnome 2.32.1 (Ubuntu 2011-04-14), Kernel 2.6.38-8	Genuine Intel, Pentium(R) Dual-Core CPU E5300 @ 2.60GHz, number of CPUs 2, cache 2048 KB	Total memory 1978 MB	Ltd. RTL8111/8168B PCI Express Gigabit Ethernet controller (rev 03)

The Parallelization library used in our work is based on a High Performance Message Passing Library that is Open MPI library. The parallelization library follows the standard MPI2, which can be used with C, C + + and Fortran. Remote computers or multiprocessor can communicate by message passing.

We have used platform application framework Qt under linux system (ubuntu 11.04) for developing application software with a graphical user interface (GUI).

The implementation of the segmentation model requires estimating the parameters μ and σ for each class. Since the segmentation is unsupervised as in [9,16], the expectation-maximization (EM) algorithm [8,11,23] is used to compute μ and σ.

The segmentation model needs also the estimation of several parameters that are: the parameter β, the initial temperature T_0 for the simulated annealing process, the constant τ and the neighborhood system. This is a non trivial task. We have conducted several tests with different parameter choices. The choice of wrong parameters can have "dramatic" consequences on image segmentation quality and the time needed to do this segmentation.

We have used several benchmarks images in our tests. We can thus evaluate our work and compare the results. Table 2 gives some samples of images that served in the tests conducted.

Table 2. Medical Images samples.

benchmark	Name of benchmark	Dimension	Link
1	MRI Phantom 8Bits (t1_icbm_normal_1mm_pn0_rf0.rawb)	181 x 217 x 181	http://www9.informatik.uni-erlangen.de/External/vollib/
2	Head MRT Angiography 8Bits (mrt8_angio2.raw)	256 x 320 x 128	http://www9.informatik.uni-erlangen.de/External/vollib/
3	Head MRI CISS 8Bits (mri_ventricles.raw)	256 x 256 x 124	http://www9.informatik.uni-erlangen.de/External/vollib/
4	jeff orchard's brain	256x256 x129	https://cs.uwaterloo.ca/~jorchard/UWaterloo/My_Brain.html

The parallel program for achieving the segmentation of medical images on the cluster of PCs is summarized by the algorithm given below.

5.1. Measure of Quality

Different measures can be used to express the quality of segmentation algorithms. The most general measure of performance are Dice coefficients DC (Kappa Index), the quantization error Je, intra-distances dmax, inter-distances dmin. So we classify measure of quality in two parts, with ground truth and without ground truth.

5.1.1. Measure of quality with ground truth

Dice coefficients DC (Kappa Index)

Evaluating the quality of the segmentation can only be made on synthetic images where the a priori segmentation is known. The Dice Coefficient *DC* or Kappa Index given hereafter measures the quality of the segmentation.

$$DC=2\times\frac{TP}{2\times TP+FP+FN}$$

where *TP* stands for True positive, *FN* False Negative and *FP* False Positive.

The Dice coefficient equals 1 when the two segmentations are identical and 0 when no classified pixel matches the true segmentation.

Table 3 contains some Kappa Index results on images from benchmark 1. To have an idea about the results achieved, we have compared our method to well known thresholding methods [5]. These methods are the Otsu method, the Mixture of Gaussians (MoG) and Mixture of Generalized Gaussians (MoGG).

Table 3. Kappa Index values

Slice	WM	GM	CSF
90	0,924027	0,839919	0,637914
91	0,92521	0,832649	0,646517
92	0,927428	0,830795	0,652111
93	0,926435	0,825073	0,651829
94	0,925901	0,825975	0,645608
95	0,926292	0,8313	0,627347
96	0,921984	0,838924	0,622548
97	0,9216	0,840191	0,606709
98	0,920335	0,841836	0,594203
99	0,920711	0,84733	0,574987

The figures 5, 6 and 7 show some visual results of the segmentation process based on the improved implemented ICM algorithm.

Table 4 shows mean kappa index values obtained on the segmentation of slices 90-119 from benchmark 1 using the three methods cited that are Otsu, MoG and MoGG and our implemented method.

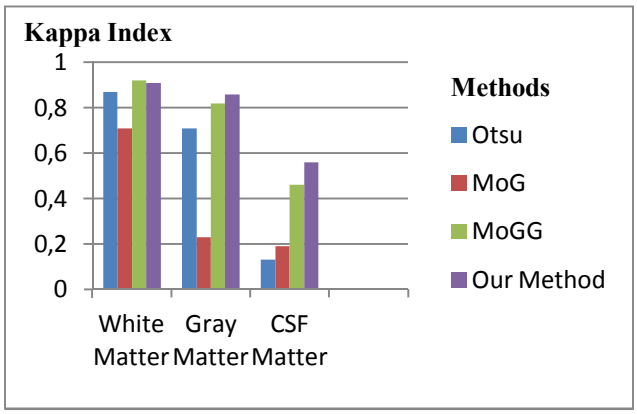


Figure 5. Mean Kappa Index values

5.1.2. Measure of quality without ground truth

Quantization error (Je), The maximum intra-distances (dmax), The minimum inter-distances (dmin) generally are used to analyze and take ideas on segmentation methods

Quantization error (Je)

Quantization error is used to express error of whole image, it is the sum of all internal error of each class. The Quantization error is given by:

$$\boldsymbol{J_e=\frac{\sum_{k=1}^{K}\left[\sum_{s\in S,x_s=k}(d(y_s,m_k))\right]/n_k}{K}}$$

Where $\boldsymbol{d}$ is the Euclidean distance, $\boldsymbol{n_k}$ number of pixels $\boldsymbol{s}$, so that $\boldsymbol{x_s=k}$

The maximum intra-distances (dmax):

dmax is used to express the maximum internal error of classes(error intra-classe). dmax is defined by:

$$\boldsymbol{dmax=max_{k=1\ldots K}\left[\sum_{s\in S,x_s=k}(d(y_s,m_k))\right]/n_k}$$

is the maximum average Euclidean distance of pixels to their associated clusters,

Where d is the Euclidean distance, n_k number of pixels s, so that $x_s = k$

The minimum inter-distances (dmin):

dmin is used to express the minimum distance between different classes means(error inter-classes). dmin is given by the following relationship:

$$dmin = \min_{\forall k,kk,k \neq kk} \{d(m_k, m_{kk})\}$$

is the minimum Euclidean distance between any pair of clusters.

Where d is the Euclidean distance, n_k number of pixels s, so that $x_s = k$

Table 4. Quantization error, the maximum intra-distances, the minimum inter-distances

Benchmark	slice	Je	dmax	dmin
1	94	6.77103	9.21012	41.3947
	120	8.57584	11.0951	44.6378
2	4	3.04897	4.60551	12.1346
	70	3.24359	5.59884	10.0092
3	3	3.04228	6.27288	7.64922
	98	7.56695	20.8302	11.8842
4	126	15.7733	32.9268	13.172
	127	14.3834	53.2261	2.13812

5.2. Processing Time

In the table below we will give the segmentation processing time and speedups in some conducted tests.

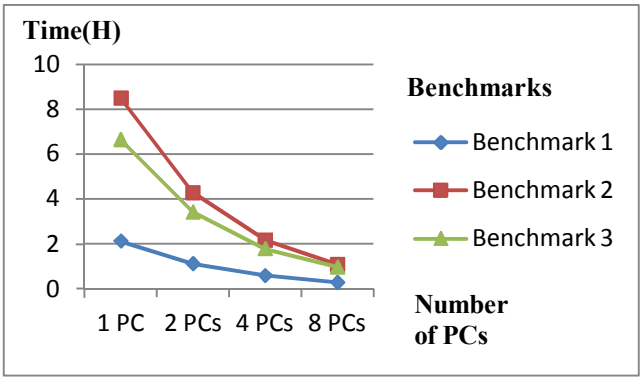


Figure 6. Time

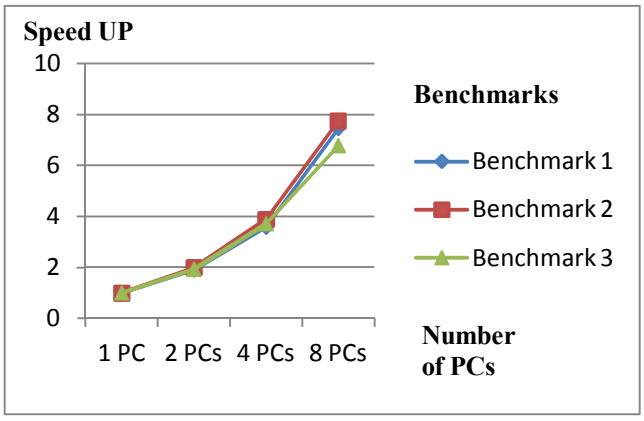


Figure 7. Speed-Up

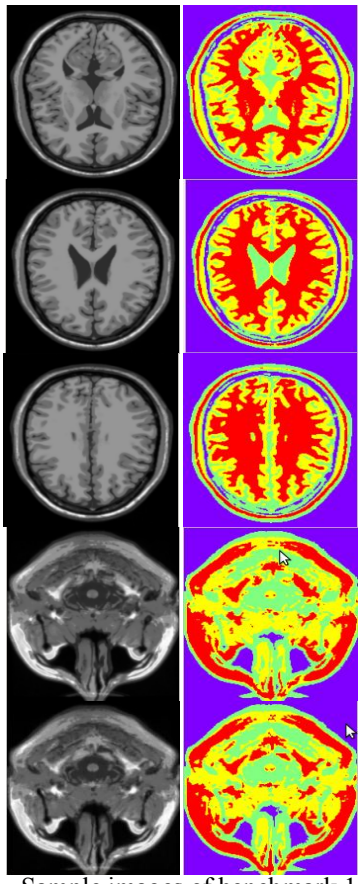


Figure 8. Sample images of benchmark 1 with their segmentation

From the results obtained, we remark that the processing time is improved almost linearly with the number of PCs in the network. The goal seems to be achieved with a small network.

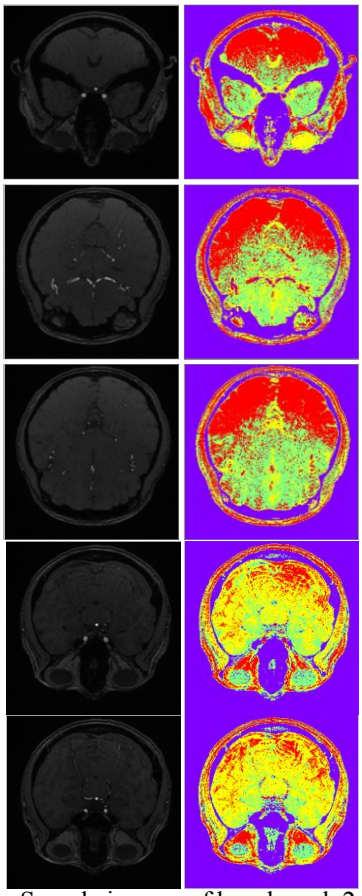


Figure 9. Sample images of benchmark 2 with their segmentation

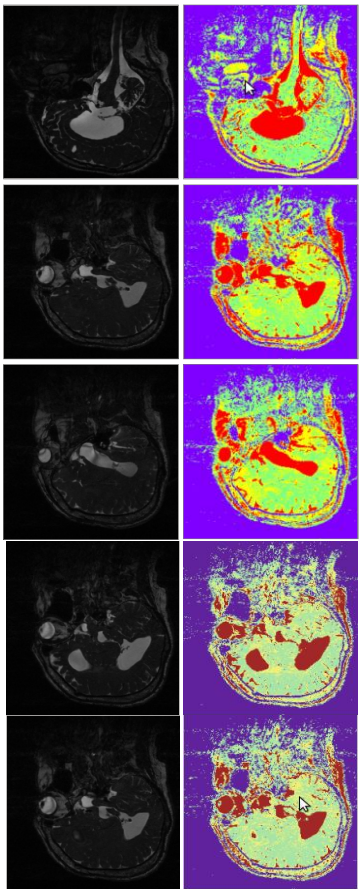


Figure 10. Sample images of benchmark 3 with their segmentation

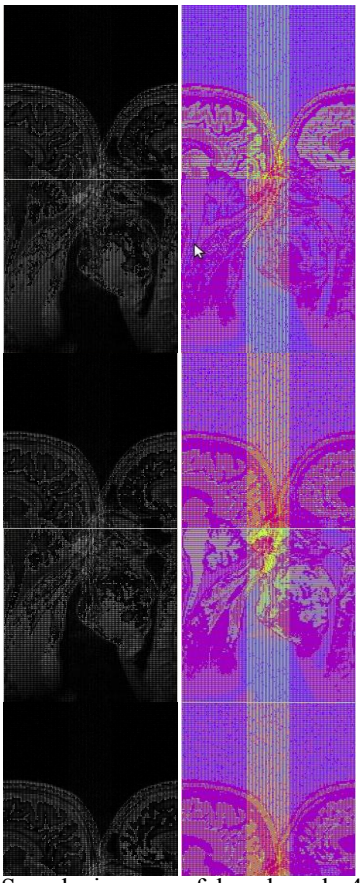


Figure 11. Sample images of benchmark 4 with their segmentation

6. CONCLUSION

This paper attempts to evaluate the segmentation of Magnetic Resonance Images using Hidden Markov Random Field Model on a common parallel architecture. When the ground truth is known, the segmented images are satisfactory (in similarity) to the a priori segmented images. The processing time is improved by the use of a cluster of PCs. We do not claim reaching the Grail but the implemented method seems to generally outperform the thresholding-based segmentation methods.

Nevertheless, further works must consider segmenting sets of MR brain images taken from other sources. The opinion of specialists must also be considered in the evaluation when no ground truth is available to have a more synthetic view of the whole segmentation process. The cluster of PCs must be incremented to see the limits of its contribution. The overheads induced by the communication process must be considered carefully.

7. REFERENCES

1. Ait-Aoudia, S., Belhadj, F., Meraihi-Naimi, A.: Segmentation of Volumetric Medical Data using Hidden Markov Random Field Model: 5th International Conference on Signal Image Technology & Internet Based Systems, pp. 65--72, IEEE, Maroc, (2009).

2. Ait-Aoudia, S., Mahiou, R., Guerrout, E.: Evaluation of Volumetric Medical Images Segmentation using Hidden Markov Random Field Model: 15th International Conference on Information Visualisation, pp. 513--518, IEEE, London, (2011).

3. Angelini, E.D., Song, T., Mensh, B.D., Laine, A.F.: Brain MRI Segmentation with Multiphase Minimal Partitioning A Comparative Study: Int. J. Biomed Imaging, vol. 2007, 15p.

4. Besag, J.: On the Statistical Analysis of Dirty Pictures with Discussion: Journal of the Royal Statistical Society. Series B (Methodological), vol. 48, pp. 259--302, (1986).

5. Boulmerka, A., Allili, M.S.: Thresholding-Based Segmentation Revisited Using Mixtures of Generalized Gaussian Distributions: 21st International Conference on Pattern Recognition, Japan, (2012).

6. Boykov, Y., Veksler, O., Zabih, R.: Fast Approximate Energy Minimization via Graph Cuts: IEEE Transactions on PAMI, vol. 23, no. 11, pp. 1222--1239, (2001).

7. Boykov, Y., Kolmogorov, V.: An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision: IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 26, no. 9, pp. 1124--1137, (2004).

8. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum Likelihood from Incomplete Data via the EM Algorithm: Journal of the Royal Statistical Society, Series B Methodological, vol. 39, no. 1, pp. 1--38, (1977).

9. Deng, H., Clausi, D.A.: Unsupervised Image Segmentation Using a Simple MRF Model with a New Implementation Scheme: Proceedings of the 17th International Conference on Pattern Recognition, pp. 691--694, (2004).

10. Geman, S., Geman, D.: Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images: IEEE Transaction on. Pattern Analysis Machine Intelligence, vol. 6, no. 6, pp. 721--741, (1984).

11. Gu, D.B., Sun, J.X.: EM Image Segmentation Algorithm Based on an Inhomogeneous Hidden MRF Model: Vision, Image and Signal Processing, IEEE Proceedings, vol. 152, no. 6, pp. 184--190, (2005).

12. Guerrout, E.: Segmentation of Volumetric Medical Images: A Distributed Approach, Editions European University, ISBN: 978-613-1-59164-8, (2011).

13. Held, K., Kops, E.R., Krause, B.J., Wells, W.M., Kikinis, R., Muller-Gartner, H.-W.: Markov Random Field Segmentation of Brain MR Images: IEEE Transactions on Medical Imaging, vol. 16, no. 6, pp. 878--886, (1997).

14. Huang, A., Abugharbieh, R., Tam, R.: Image Segmentation Using an Efficient Rotationally Invariant 3D Region-Based Hidden Markov Model: IEEE Computer Vision and Pattern Recognition Workshops, Anchorage, AK, USA, pp. 1--8, (2008).

15. Ibrahim, M., John, N., Kabuka M., Younis, A.: Hidden Markov Models-Based 3D MRI Brain Segmentation: Image and Vision Computing vol. 24, pp. 1065--1079, (2006).

16. Kato, Z., Zerubia, J., Berthod, M.: Unsupervised Parallel Image Classification Using Markovian Models: Pattern Recognition 32, pp. 591--604, (1999).

17. Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P.: Optimisation by Simulated Annealing", Science, vol. 220, no. 4598, pp. 671--680, (1983).

18. Li, S.Z.: Markov Random Field Modeling in Computer Vision: Springer-Verlag, New York, (2001).

19. Marroquin, J.L., Vemuri, B.C., Botello, S., Calderon, E., Fernandez-Bouzas, A.: An Accurate and Efficient Bayesian Method for Automatic Segmentation of Brain MRI: IEEE Transactions on Medical Imaging, vol. 21, no. 8, pp. 934--945, (2002).

20. Szeliski, R., Zabih, R., Scharstein, D., Veksler, O., Kolmogorov, V., Agarwala, A., Tappen, M., Rother, C.: A Comparative Study of Energy Minimization Methods for Markov Random Fields with Smoothness-Based Priors: IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 6, pp. 1068--1080, (2008).

21. Van Leemput, K., Maes, F., Vandermeulen, D., Suetens, P.: A Unifying Framework for Partial Volume Segmentation of Brain MR Images: IEEE Transactions on Medical Imaging, vol. 22, no. 1, pp. 105--119, (2003).

22. Wyatt, P., Noble, J.A.: MAP MRF Joint Segmentation and Registration of Medical Images: Medical Image Analysis, vol. 7, no. 4, pp. 539--552, (2003).

23. Zhang, Y., Brady, M., Smith, S.: Segmentation of Brain MR Images through a Hidden Markov Random Field Model and the Expectation-Maximization Algorithm: IEEE Transactions on Medical Imaging, vol. 20, no. 1, pp. 45--57, (2001).

24. Guerrout, E., Mahiou, R., Ait-Aoudia, S.: Medical Image Segmentation Using Hidden Markov Random FieldA Distributed Approach The Third International Conference on Digital Information Processing and Communications, pp. 423--430, Dubai, (2013).

25. Gonzalez, Rafael C.; Woods, Richard E. Digital Image Processing, 1992, Addison-Wesley. Publishing Company, Inc.

26. Aydın D., Adaptation Of Swarm Intelligence Approaches Into Color Image Segmentation And Their Implementations On Recognition Systems, PhD thesis in Computer Engineering, Ege University, 2011.

27. T. Kanungo, D. M. Mount, N. Netanyahu, C. Piatko, R. Silverman, & A. Y.Wu (2002) An efficient k-means clustering algorithm: Analysis and implementation Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp.881-892.

28. S. C. Chen, D. Q. Zhang. Robust Image Segmentation Using FCM With Spatial Constraints Based on New Kernel-Induced Distance Measure [J]. IEEE Transactions on Systems, Man and Cybernetics-part B: Cybernetics, 2004,34(4): 1907-1916.

29. Ramos V. and Almeida, F., 2000, Artificial Ant Colonies in Digital Image Habitats- A Mass Behavious Effect Study on Pattern Recognition, Second International Workshop on Ant Algorithms, 113-116pp