Middle-East Journal of Scientific Research 7 (6): 1024-1029, 2011
ISSN 1990-9233

Optimization Approaches in Markov Random Field Model: A Comparative Survey for MR Image Segmentation Case Study

[1]Sahar Yousefi, [1]Morteza Zahedi and [2]Reza Azmi

[1]Department of IT and Computer Shahrood University of Technology, Shahrood, Iran
[2]Alzahra University, Iran

Abstract: Magnetic resonance image (MRI) brain segmentation plays an increasingly important role in computer-aided detection and diagnosis (CAD) of abnormalities. This task can be done by marking the interest regions slice-by-slice but MRI segmentation manually is time consuming and consumes valuable human resources. Hence a great deal of efforts has been made to automate this process. MRF has been one of the most active research areas of MRI brain segmentation which seeks an optimal label field in a large space. The traditional optimization method is Simulated Annealing (SA) that could get the global optimal solution with heavy computation burden. Therefore great deal efforts have been made to obtain the optimal solution in a reasonable time. In this paper, we conduct a comparative study with the traditional minimization approach and two novel proposed methods for segmentation of MR images. The qualitative and quantitative results of each system are investigated as well.

Key words: Magnetic Resonance Image (MRI) · Markov Random Field (MRF) · Simulated Annealing (SA) · Improved Genetic Algorithm (IGA) · Ant Colony Optimization (ACO) · Gossiping Algorithm · Image Segmentation

INTRODUCTION

Automatic Magnetic Resonance Image (MRI) segmentation is one of the most important steps in computer-aided detection and diagnosis (CAD) of abnormalities such as lesions or neurodegenerative disorders such as Alzheimer disease, or movement's disorders such as Parkinson. Various segmentation methods have been proposed in the literature. Among them, there has been significant interest in Markov random field (MRF) based approaches in the past few years [1-4]. MRF is a statistic model which seeks the optimal label field of the image voxels [5]. Using local interaction between voxels by defining neighborhood system, MRF models spatial coherence constraints. This causes noise effect reduction intensively; instead, the segmentation algorithm needs heavy computation burden. In order to alleviate the computational load, many methods have been proposed which involve solving an energy function optimization [6-7].

This paper investigates three different algorithms for seeking the optimal solution. The first is classical MRF method which is based on Simulated Annealing (SA) [8]. This method converges to the global optima asymptotically but requires a great deal of computation. The second method applies a hybrid of simulated annealing (SA) and improved genetic algorithm (IGA) in order to optimize the problem which is formulated by MRF [9]. The third algorithm is a social algorithm which tackles the foraging behavior concept of ant colony combined with gossiping algorithm to seek the optimal solution.

This paper is organized as follows: in Section 2, MRF model is introduced briefly. Section 3 describes the optimization algorithms in order to seek optimal solution. Image segmentation experiment results are presented in Section 4 and the conclusion is obtained in the Section 5. Markov Random Field: MRF model poses image segmentation as a labeling problem in which a set of labels are assigned to the set of image voxels. Suppose that $S = \{s_1, s_2, ..., s_N\}$ is a two dimensional image lattice which for each site like $s \in S$ a gray level value is assigned. A neighborhood system with respect to S is defined as $N(s) = \{s, r \in S \mid 0 < \|s - r\| < d\}$, where d is a constant and $\|.\|$ means Euclidean distance. MRF seeks a label field like $Y = \{y_s \in \Gamma \mid s \in S\}$ where $\Gamma = (\lambda_1, \lambda_2, ..., \lambda_M)$ is the set of tissue labels. Label field Y is a MRF if it satisfies

Corresponding Author: Reza Azmi, Alzahra University, Iran. E-mail: azmi@alzahra.ac.ir.

Markovianity condition: $P(Y_s|Y_{Ss}) = P(Y_s \mid Y_{M(s)})$, which states label of each site given other labels in Y only depends on its neighbors. For determining this property, a theorem was ascribed to MRF by Hammersley and Clifford which stated the equivalence between MRFs and Gibbs distributions [10]. According to this theory Y is MRF if and only if

$$P(Y) = \frac{\exp(-E(Y)/T)}{\sum_{Y \in \Xi} \exp(-E(Y)/T)}, \quad (1)$$

Where T is the system temperature, Ξ is the set of all possible label fields or configurations, $E(Y)$ is an energy function which is a sum of clique potentials over all possible cliques. Clique means a subset of S, such as c, in which all $s \in c$ are neighbors of one another. We use 3 dimensional cliques. For two sites like s, $r \in S$ potential function is defined as (2), in which v_{xy} is in-plane-clique potential function and v_z is out-plane-clique potential function.

$$v_{xy} = \begin{cases} -1 & if\ y_s = y_r \\ 1 & if\ y_s \neq y_r \end{cases}, \quad v_z = \begin{cases} -0.5 & if\ y_s = y_s \\ 0.5 & if\ y_s \neq y_s \end{cases} \quad (2)$$

The optimal Y^* is defined by maximum a posterior (MAP) criterion: $Y^* = \arg\max_{Y \in \Xi}(P(x \mid y,\theta)P(y \mid \theta))$, In which $P(y \mid \theta)$ is Gibbs distribution according to Hammersley-Clifford theory and $P(x \mid y,\theta)$ adheres Gaussian distribution with mean μ and variance σ. Therefore the MAP minimizes energy function (3).

$$\begin{aligned} Y^* &= \arg\min_{Y \in \Xi}(\ln\sqrt{2\pi}\sigma + \frac{(x-\mu)^2}{2\sigma^2} + E(Y)) \\ &= \arg\min_{Y \in \Xi}(U(Y)) \end{aligned} \quad (3)$$

Optimization Algorithms

Simulated Annealing: One of the most prevalent approaches in order to find the optimal solution in MRF is SA [11, 12]. SA relies on thermodynamic concepts which its idea was motivated by an analogy to annealing in solids. SA is introduced by Metropolis *et al.* in 1953 [13] and is a stochastic relaxation algorithm which carries out by reducing the system temperature gradually. Cause of the stochastic behavior, SA gets rid of trapping sub-optima. In other words SA has strong capability to climb hills but slow convergence speed. Therefore the combination of SA with MRF is not used for real time processes.

Hybrid of IGA-SA: This method uses a hybrid of IGA and SA to seek the solution. Although SA is time consuming, it has a good capability of hill climbing. In comparison, GA is weak in hill climbing but converges rapidly toward the solution. Therefore an appropriate hybrid of these two algorithms can alleviate their shortcomings. As stated in [9], the algorithm begins with initialization and population generation. Then a loop start until the convergence criterion satisfied. The loop contains following step.

Selection: Two individuals selects using roulette wheel algorithm as parents.

Crossover: Two-point crossover is applied to generate two offspring.

Mutation: Some genes of each offspring are selected randomly and are mutated according to most frequent gene in their neighbours.

Replacing: Parents are replaced with their offspring if energy value of parent is bigger than energy value of offspring or $\xi > exp(-\Delta U/T)$, in which ξ is a random number between [0,1] and T is temperature and U is defined in (3).

Temperature Reduction: The system temperature is reduced.

Hybrid of ACO-Gossiping: We consider each site as an ant which seeks for food. The goal of colony is to find Y^* (3). At time $(t+1)$, each ant transits from node i to node j according to state transfer probability $P_{i,j}^{(t+1)}$ [14].

$$P_{i,j}^{(t+1)} = \frac{\left(\tau_{i,j}^{(t)}\right)^{\alpha}\left(\eta_{i,j}\right)^{\beta}}{\sum_{k \in \Omega}\left(\tau_{i,j}^{(t)}\right)^{\alpha}\left(\eta_{i,j}\right)^{\beta}} \quad (4)$$

In which $\tau_{i,j}^{(t)}$ is pheromone density between i and j, $\eta_{i,j}$ is heuristic information between i and j, Ω all possible destinations for ant, α and β represent the influence of pheromone information and heuristic information respectively. In this paper, heuristic information is determined as $\eta_{i,j} = -L_{U_{i,j}} + |\min(-L_{U_{i,j}})|$, In which $L_{U_{i,j}}$ is the local energy at the path (i, j) and is defined as

$$L_{U_{i,j}} = \frac{(x_i - \mu_j)^2}{2\pi{\sigma_j}^2} + \ln\sqrt{2\pi}\sigma_j + \frac{\sum_{c \in C_i} V_c}{T} \quad (5)$$

Where μ_j and σ_j are mean and variance of destination, V_c is potential function of clique c, C_i is a set of possible cliques and T is temperature.

The standard ant colony system [14] performs two update operations for updating the pheromone matrix. The first update is performed after the movement of each ant; the second update is performed after the move of all ants. These two steps are imposed in this paper. In addition, we proposed a new updating method based on gossiping algorithm which considers the coherence among the neighboring voxels.

Local Update: When an ant passes through a path, it reinforces pheromone trials. Therefore, the local update is performed after the movement of an ant according to,Where ñ is the evaporation rate, is fitness value of the destination and we define as Where x_i is intensity of source, i_j is the mean of the destination.

Global Update: The second update is carried out after the movement of all ants as $\tau^{(t+1)} = (1-\psi)\tau^{(t)} + \psi\left|\frac{\Delta U}{U_{old}}\right|$, Where ψ is the pheromone decay coefficient, U is defined in (3).

Gossiping Update: Gossiping or rumor spreading is one of the information dissemination algorithms in distributed systems. This algorithm inspired from spreading a piece of gossip in social life [15]. When someone tells person A a piece of gossip, he will try to tell other people who they don't know beforehand (called susceptible). If he tells person B who knows that (called infected), person A will lose interesting in telling more. Here, ants are individuals who interest to spread their pheromone information as a gossip. Here we use a push-pull gossip protocol in which A and B exchange gossips with each other. In order to simplify the algorithm we define neighborhood gossiping as follow:

Finding Neighbours: Ant *l* finds its neighbors. Here we define a 5×5 neighborhood window, called $N_g(l)$, which demonstrates all ants who are neighbors with l. Fig (1) shows the defined neighborhood window.

Push Operator: Ant *l* assists all ants in the set $N_g(l)$ to make a true decision in foraging process. In this process l broadcasts its best destination, b, such as a piece of gossip to r which $r \in N_g(l)$. Therefore the pheromone matrix is updated as (6), in which γ is the pheromone reinforcing coefficient, x_l and x_r are intensities of sites l and r respectively and $(x_l - x_r)$ imposes the fitness of path *r,b*.

$$\tau_{r,b} = \tau_{r,b} + \gamma[1 - (x_l - x_r)/255],\ r \in N_g(l) \qquad (6)$$

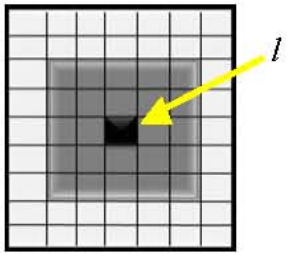

Fig. 1: Neighbourhoods in gossiping update

If some of ants in $N_g(l)$, know the gossip beforehand, ant l loses interesting in spreading the gossip. Therefore the fraction of nodes that will remain susceptible in next iteration satisfies the equation (7), in which $N_g^{u(t-1)}(l)$ is the infected neighbours and |.| means the size of set.

$$\tilde{s} = \left[1 - \left(\left|N_g^{u(t-1)}(l)\right| \Big/ \left|N_g(l)\right|\right)\right] \times \left|N_g(l)\right| \qquad (7)$$

Pull Operator: In this step ant l receives the pheromone information of its neighbours and updates the pheromone matrix as follows $\tau_{lk} = \tau_{lk} + \gamma.\phi(k)$, in which $r \in N_g(l)$, $k \in \Omega$, $\phi(k)$ is relative frequency of path *r,k* in neighborhood window, γ is the pheromone reinforcing coefficient and Ω is the set of all possible states which *l* can selects.

Experimental Results: We applied the methods to segment two datasets. The first is normal MRI brain in Internet Brain Segmentation Repository (IBSR) dataset and the second is multiple T1-wheighted and gadolinium enhanced coronal MRI scans of a patient with a tumour taken at roughly 6 month intervals over three and a half years in IBSR dataset. IBSR was provided by the Center for Morphometric Analysis (CMA) at Massachusetts General Hospital and is available at http://www.cma.mgh.harvard.edu/ibsr/. The intended dataset is in 536/ sub-folder and contains the result of semi-automated segmentation which is considered as ground truth (GT) for validating the automatic segmentation methods.

The methods are coded in MATLAB 7.2 running on an Intel 2.00 GHz CPU system with 3.00G Bytes memories. The volume overlap metric we use is the Dice similarity coefficient (DSC) [13]. For a given pair of segmentation volumes A and B, the measured overlap is.

$$\kappa = 2|A \cap B|/(|A| + |B|) \qquad (8)$$

In which |.| means size of sets.

Table 1: Algorithm Performance and Calculated Times

	STND-MRF		MRF-SA-IGA		MRF-ACO-Gossiping	
IBSR MR#	$\kappa_{average}$	Calculated time (sec)	$\kappa_{average}$	Calculated time (sec)	$\kappa_{average}$	Calculated time (sec)
1-24/12	0.666	1313.483	0.666	169.250	0.669	263.764
1-24/14	0.665	1152.762	0.667	111.866	0.668	522.564
1-24/16	0.731	1398.474	0.731	189.269	0.734	657.938
1-24/18	0.776	1517.988	0.779	106.376	0.775	639.585
1-24/20	0.711	1542.545	0.717	152.710	0.718	706.002
1-24/22	0.675	1477.189	0.676	134.743	0.676	804.798
1-24/24	0.717	1181.077	0.724	118.189	0.721	518.789
1-24/26	0.690	1139.004	0.696	140.371	0.694	279.614
1-24/28	0.659	989.672	0.660	144.519	0.663	482.722
1-24/30	0.699	1036.734	0.702	97.211	0.705	356.984
1-24/32	0.713	1107.103	0.720	122.741	0.720	634.450
Average	0.700	1259.639	0.703	135.204	0.704	533.383

Table 2: Algorithm Performance and Calculated Times for

	STND-MRF		MRF-SA-IGA		MRF-ACO-Gossiping	
IBSR MR#	κ_{tumour}	Calculated time (sec)	κ_{tumour}	Calculated time (sec)	κ_{tumour}	Calculated time (sec)
32i/25	0.823	3184.442	0.838	230.771	0.794	767.771
32i/26	0.719	2425.532	0.740	172.859	0.698	690.210
32i/27	0.532	2483.694	0.542	221.472	0.537	610.940
32i/28	0.670	2528.072	0.721	175.761	0.587	676.341
32i/29	0.565	2227.476	0.649	162.891	0.550	467.083
88i/26	0.811	2247.417	0.824	278.388	0.791	549.373
88i/27	0.783	2090.443	0.802	223.548	0.751	717.973
88i/28	0.791	2100.650	0.813	107.468	0.781	882.279
88i/29	0.798	2109.861	0.814	180.158	0.762	770.683
45i/25	0.671	2731.595	0.661	260.099	0.662	655.228
45i/26	0.759	2488.324	0.813	224.677	0.732	500.498
Average	0.720	2419.8	0.747	203.463	0.695	662.580

The average dice values for normal MRIs and tumour volume and calculation times of different algorithms are shown as table 1 and 2. As indicated in table 1, the standard MRF produces average similarity measures of 0.700. This value for MRF-SA-GA and MRF-ACO-Gossiping methods are equate to 0.703 and 0.704 respectively. Additionally, calculation time reduction percentage of MRF-SA-GA method compared to standard MRF in average is 89.13%. Also calculation time reduction percentage of MRF-ACO-Gossiping method compared to standard MRF is 57.83% averagely. Table 2 indicates that the standard MRF produces average similarity measures of 0.720. This value for MRF-SA-GA and MRF-ACO-Gossiping methods are equate to 0.747 and 0.695 respectively. Additionally, calculation time reduction percentage of MRF-SA-GA method compared to standard MRF in average is 91.56%. Also calculation time reduction percentage of MRF-ACO-Gossiping method compared to standard MRF is 72.15% averagely. Therefore MRF-SA-GA algorithm is the fastest method.

Fig. 2 represents a qualitative comparison between these methods on some MRI scans of IBSR. As it is observed in this figure, MRF-SA-GA can yield more robust segmentation results to noise.

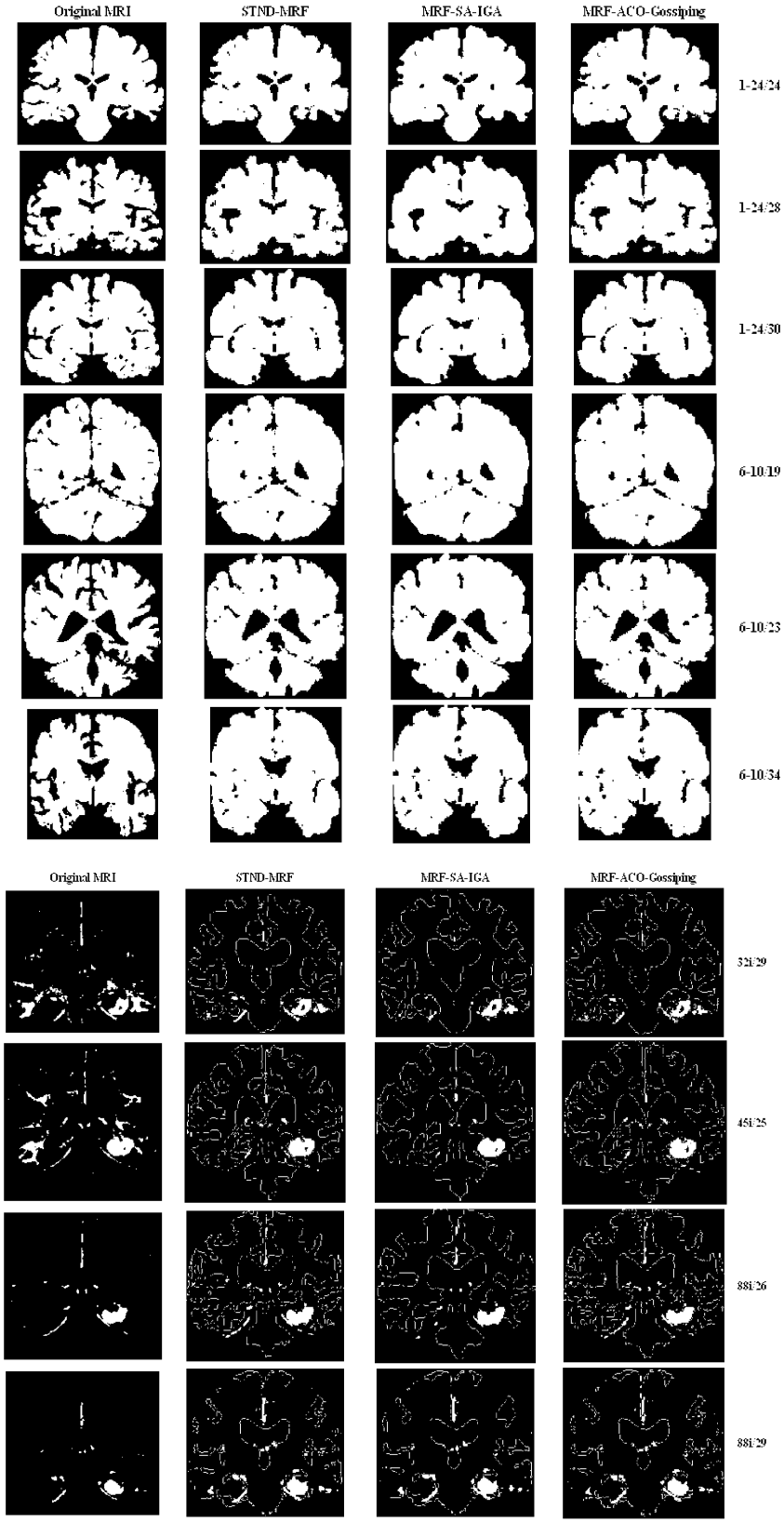

Fig. 2: Segmentation results of different methods

CONCLUSION

In this paper, a comprehensive comparison evaluation of three proposed optimization methods in Markov random field (MRF) model was presented. In this study, classical MRF that uses simulated annealing, MRF-SA-IGA which uses a hybrid of simulated annealing and improved genetic algorithm and MRF-ACO-Gossiping which uses a hybrid of social algorithm contains ant colony optimization and gossiping algorithm was examined for segmenting the brain into white matter and grey matter, vessel and cerebrospinal fluid (CSF). Results indicated that the MRF-SA-IGA approach outperforms two other methods in convergence speed and accuracy.

REFERENCES

1. Liang, Z., 1993. Tissue classification and segmentation of MR images. IEEE Eng. Med. Biol., 12(1): 81-85.
2. Caputo, B., E.L. Torre, S. Bouattour and G.E. Gigante, 2002. A new kernel method for microcalcification detection: Spin glass-Markov random fields, Stud. Health Technol. Inf., 90: 30-34.
3. Hongmei Sun and Tianfu Wang, 2008. LCG-MRF-Based Segmentation of MRI Brain Images, International Conference on Computer Science and Information Technology.
4. Hugh Gribben, Paul Miller, Gerard G. Hanna, Kathryn J. Carson and Alan R. Hounsell, 2009. Map-mrf segmentation of lung tumours in PET/CT Images.
5. Yousefi, S., M. Zahedi and R. Azmi, 2010. 3D MRI brain segmentation based on MRF and hybrid of SA and IGA, ICBME.
6. Kim, E.Y., S.H. Park and H. J. Kim, 2000. *A Genetic Algorithm-Based Segmentation of Markov Random Field Modeled Images"*, IEEE Signal Processing Letters, 7(11): 301-304.
7. Wang, X. and H. Wang, 2003. *Evolutionary Optimization in Markov Random Field Modeling*, Evolutionary Computation, IEEE Transactions on, 8(6): 567-579.
8. Geman, S. and D. Geman, 1984. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. IEEE Trans. Pattern Analysis and Machine Intelligence, 6: 721-741.
9. Yousefi, S., R. Azmi and M. Zahedi, 2010. *Brain Tumor Segmentation in 3D MRIs Using an Improved Markov Random Field Model*, International Conference on Signal and Information Processing.
10. Hammersley, J.M. and P. Clifford, 1971. Markov field on finite graphs and lattices.
11. Zhang, Y., M. Brady and S. Smith, 2001. Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm, IEEE Trans Med. Imag., 20: 45-57.
12. Barker, S.A. and P.J.W. Rayner, 2000. Unsupervised image segmentation using markov random field models, J. Pattern Recognition, 33: 587-602.
13. Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller and E. Teller, 1953. Equation of state calculations by fast computing machines, J. Chemical Physics, 21: 1087-1092.
14. Dorigo, M., M. Birattari and T. Stutzle, 2006. *Ant colony optimization*, IEEE Computational Intelligence Magazine, 1: 28-39.
15. Tanenbaum, A.S., 2006. Distributed systems, pp: 170-176.
16. Dice, L.R., 1945. Measures of the amount of ecologic association between species. Ecol., 26(3): 297-302.