

Automatic MR Brain Tumor Image Segmentation

Yisu Lu^{1,2}

1. Electronic Engineering Department,
South China Institute of Software Engineering
Guangzhou, China

2. Key Lab for Medical Image Processing,
Southern Medical University,
Guangzhou, China
E-mail: Yisu_lu@126.com

Wufan Chen²

2. Key Lab for Medical Image Processing,
Southern Medical University,
Guangzhou, China

Abstract—Traditional Dirichlet process mixture (MDP) models has the characteristic that the image segmentation can be done without initialization of clustering numbers. For the computing speed of the classical MDP segmentation is jogging, a new kind of nonparametric segmentation (DMMDP algorithm) combined with anisotropic diffusion and Markov Random Fields (MRF) prior was inferred in this paper. The experiment results of meningioma MR images segmentation showed that the properties, such as accuracy and computing speed, of the DMMDP algorithm were significantly greater than the classical MDP model segmentation.

Keywords—image segmentation; Dirichlet process mixtures; anisotropic diffusion; MRF

I. INTRODUCTION

Traditional Dirichlet process mixture (MDP) models have been studied in nonparametric Bayesian statistics for more than three decades. Originally introduced by Ferguson [1] and Antoniak [2], interest in these models has increased since efficient sampling algorithms became available [3-5].

MDP models provide a Bayesian framework for clustering problems with an unknown number of groups. They support a range of prior choices for the number of classes; the different resulting models are then scored by the likelihood according to the observed data. The number of clusters as an input constant is substituted by a random variable with a control parameter. Instead of specifying a constant number of clusters, the user specifies a level of cluster resolution by adjusting the parameter.

For the property of nonparametric method, the MDP model algorithm can be used to segment some images whose class number can not be set at the very start. So in the tumor segmentation field, the nonparametric method represents important and challenging work. The segmentation of brain tumor images segmentation was studied in this paper[6,7,8]. For the computing speed of the classical MDP segmentation is slow, anisotropic diffusion and Markov Random Fields (MRF) were combined to the Dirichlet process mixture models, and a kind of faster nonparametric algorithm (DMMDP algorithm) was inferred in this paper.

The DMMDP algorithm was used to segmentation to meningioma MR images. The experiment results show that

DMMDP method was accurate extraordinarily, and particularly the computing time was significantly faster than the traditional MDP segmentation.

II. DIRICHLET PROCESS MIXTURE MODELS

A. Dirichlet Process Mixture Models

Dirichlet processes are most commonly used in the form of so-called Dirichlet process mixture models, introduced in [5]. These models employ a Dirichlet process to choose a prior at random (rather than a likelihood, as the standard DP model does). The initial motivation for considering MDP models is an inherent restriction of the DP: Regarded as a measure on the set of Borel probability measures, the DP can be shown to be degenerate on the set of discrete measures [4]. In other words, a distribution drawn from a DP is discrete with probability 1, even when the base measure G_0 is continuous. MDP models avoid this restriction by drawing a random prior G from a DP and by combining it with a parametric likelihood $F(x|\theta)$. Sampling from the model is conducted by sampling:

$$\begin{aligned} x_i &\sim F(\cdot|\theta_i) \\ \theta_i &\sim G \\ G &\sim \text{DP}(\alpha G_0) \end{aligned} \quad (1)$$

If the parametric distribution F is continuous, the model can be regarded as a convolution of the degenerate density G with a continuous function, resulting in a continuous distribution of the data x_i . In the MDP case these are the parameters θ_i . Since the parameters are not observed (but rather the data x_i), they have to be integrated out to obtain a sampling formula conditional on the actual data.

B. MDP Models and Data Clustering

The principal motivation for the application of MDP models in machine learning is their connection with both

clustering problems and model selection: MDP models may be interpreted as mixture models. The number of mixture components of these mixtures is a random variable, and may be estimated from the data. The term clustering algorithm is used here to describe an unsupervised learning algorithm which groups a set x_1, \dots, x_n of input data into distinct classes. The number of classes will be denoted by N_C , and the class assignment of each data value x_i is stored by means of an indicator variable $S_i \in \{1, \dots, N_C\}$.

A standard modeling class in data clustering is finite mixture models with distributions of the form:

$$p(x) = \sum_{k=1}^{N_C} c_k p_k(x) \quad (2)$$

Subject to $c_k \in R \geq 0$ and $\sum_k c_k = 1$, the vector (c_1, \dots, c_{N_C}) defines a finite probability distribution, with $c_k = \Pr\{S = k\}$ for a data value drawn at random from the model. Each individual cluster is represented by a single probability distribution p_k . The model assumes a two-stage generative process for \mathcal{X} :

$$\begin{aligned} \mathcal{X} &\sim p_s \\ S &\sim (c_1, \dots, c_{N_C}) \end{aligned} \quad (3)$$

If the component distributions are parametric models $p_k(x) = p_k(x|\theta_k)$, the distribution (2) is a parametric model as well, with parameters c_1, \dots, c_{N_C} and $\theta_1, \dots, \theta_{N_C}$.

Now we consider the MDP model (1). For a given set of values $\theta_1, \dots, \theta_{N_C}$, which are already drawn from the random measure G , the measure can be integrated out to obtain a conditional prior:

$$\begin{aligned} p(\theta_{n+1}|\theta_1, \dots, \theta_n) = \\ \frac{1}{n + \alpha} \sum_{i=1}^n \delta_{\theta_i}(\theta_{n+1}) + \frac{\alpha}{n + \alpha} G_0(\theta_{n+1}) \end{aligned} \quad (4)$$

Due to the clustering property of the DP, $\theta_1, \dots, \theta_n$ will accumulate in $N_C \leq n$ groups of identical values. Each of these classes is represented by its associated parameter value, denoted θ_k^* for class $k \in \{1, \dots, N_C\}$.

(That is, $\theta_i = \theta_k^*$ for all parameters θ_i in class k) The sum over sites in the first term of (4) may be expressed as a sum over classes:

$$\sum_{i=1}^n \delta_{\theta_i}(\theta_{n+1}) = \sum_{k=1}^{N_C} n_k \delta_{\theta_k^*}(\theta_{n+1}) \quad (5)$$

Where n_k denotes the number of values accumulated in group k . The conditional distribution (4) may be rewritten as

$$\begin{aligned} p(\theta_{n+1}|\theta_1, \dots, \theta_n) = \\ \sum_{k=1}^{N_C} \frac{n_k}{n + \alpha} \delta_{\theta_k^*}(\theta_{n+1}) + \frac{\alpha}{n + \alpha} G_0(\theta_{n+1}) \end{aligned} \quad (6)$$

Combination with a parametric likelihood F as in (1) results in a single, fixed likelihood function $F(\cdot|\theta_k^*)$ for each class k . Therefore, the model may be regarded as a mixture model consisting of N_C parametric components $F(\cdot|\theta_k^*)$ and a “zero” component (the base measure term), responsible for the creation of new classes.

For a parametric mixture model applied to a clustering problem, the number of clusters is determined by the (fixed) number of parameters. Changing the number of clusters therefore requires substitution of one parametric mixture model by another one. MDP models provide a description of clustering problems that is capable of adjusting the number of classes without switching models. This property is, in particular, a necessary prerequisite for Bayesian inference of the number of classes, which requires N_C be a random variable within the model framework, rather than a constant of the model.

III. DMMDP SEGMENTATION METHOD

A. anisotropic diffusion

Magnetic resonance imaging technology has been widely applied to medical diagnosis systems and the accuracy of many diagnosis systems are mainly based on the quality of the images acquired [9][10][11]. However, the images obtained by magnetic resonance imaging usually contain heavy noise and the noise in the MR images degrades the quality of the images and makes the post-processing of the images such as segmentation, classification and detection difficult [9][10]. The noise in the images sometimes even affects the evaluation of the human segmentation [9]. Thus in order to alleviate the affection of the noise, noise reduction is generally used to remove or reduce the noise before segmentation, classification and detection.

B. Markov Random Fields

This work combines nonparametric Dirichlet process mixture models with Markov random field (MRF) models to enforce spatial constraints. MRF models have been widely used in computer vision [12,13], the following brief exposition is intended to define notation and specify the type

of models considered. Markov random fields provide an approach to the difficult problem of modeling systems of dependent random variables. To reduce the complexity of the problem, interactions are restricted to occur only within small groups of variables

C. DMMDP algorithm

To reduce the effect noise to the MR image, anisotropic diffusion was applied to the MR images firstly. Then the MDP model was constrained by Markov Random Fields. Compared with the traditional MDP model segmentation algorithm, the new segmentation algorithm was called as DMMDP algorithm.

The new algorithm is designed as following:

1. Anisotropic diffusion process is used to the original MR image.
2. Generate a single cluster containing all points of the anisotropic diffused, and the parameter is :

$$\theta_1^* \sim G_0(\theta_1^*) \prod_{i=1}^n F(x_i | \theta_1^*) \quad (7)$$

3. Generate a random permutation σ of the data indices.
4. Assignment step. For $i = \sigma(1), \dots, \sigma(n)$

(1) If x_i is the only observation assigned to its cluster $k = S_i$, remove this cluster.

(2) Compute the cluster probabilities

$$q_{i0} \propto \alpha \int_{\Omega_\theta} F(x_i | \theta) G_0(\theta) d\theta$$

$$q_{ik} \propto n_k^{-i} \exp(-H(\theta_k^* | \theta_{-i})) F(x_i | \theta_k^*) ,$$

For $k = 1, \dots, N_C$ (8)

(3) Draw a random index k according to the finite distribution $(q_{i0}, \dots, q_{iN_C})$

(4) Assignment:

If $k = \{1, \dots, N_C\}$, assign x_i to cluster k .

If $k = 0$, create a new cluster for x_i .

5. Parameter update step. For each cluster $k = 1, \dots, N_C$;

Update the cluster parameters

$$\theta_k^* \sim G_0(\theta_k^*) \prod_{i|S_i=k} F(x_i | \theta_k^*) \quad (9)$$

6. Estimate assignment mode: For each point, choose the cluster it was assigned to most frequently during a given number of iterations.

IV. EXPERIMENTAL RESULTS

A. segmentation results

To compare the effect of the segmentation results of the modified algorithm, two different algorithm are applied to segment the of meningioma MR images independently, the

results are showed in Figure1. Four original clinical meningioma MR images were showed in the first line of Figure1. The classical MDP algorithm segmentation results were showed in the second line of Figure1. The third line showed DMMDP segmentation results. The results showed that the segmentation class numbers of MDP results are all more than those of the DMMDP results. With more class numbers, the segmentation of classical MDP can not segment the meningioma of the MR images correctly.

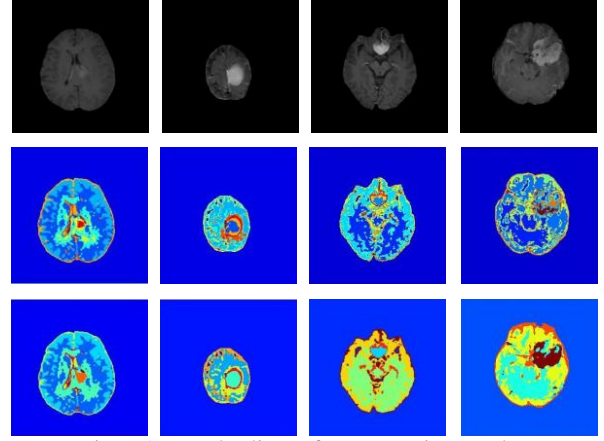


Figure1. sample slices of segmentation results.

Figure2 showed the clustering class number of fifteen MRI samples segmentation results, which were segment by the MDP results and the DMMDP results respectively. The results approved that all the class number of DMMDP results were less than that of the traditional MDP segmentation.

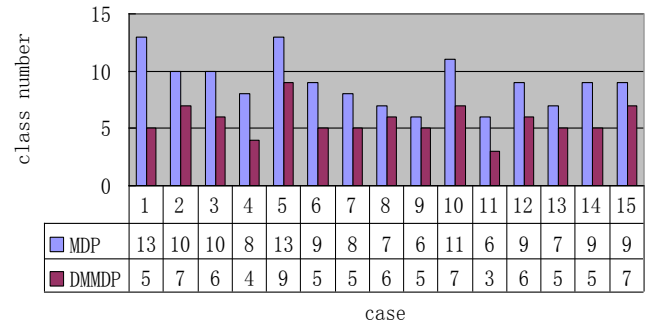


Figure2. Clustering numbers

B. Accuracy of DMMDP segmentation

To demonstrate the accuracy of modified algorithm segmentation quantitatively, MRI data segmentation results were analyzed in Table 1. The results were demonstrated at pixel level, together with there validation parameters. DMMDP: Proposed DMMDP results; GT: Ground Truth; FP: False Positive; FN: False Negative; Dice Similarity Coefficients (DSC) = $2 * NOO / (NOM + NOMDP)$. In TABLE 1, the average DSC of the all the segmentation results exceed 91.27%, which showed that the accuracy and

robustness of the modified segmentation algorithm was very great.

Table1. Accuracy of DMMDP segmentation

Case	DMMDP	GT	FP	FN	DSC
1	7344	7316	467	439	0.9382
2	11045	10937	579	471	0.9522
3	8288	10724	257	2693	0.8448
4	4953	8356	128	3532	0.7250
5	1274	1236	49	11	0.9761
6	2804	2437	399	32	0.9178
7	5680	5168	569	57	0.9423
8	7080	6798	422	140	0.9595
9	7615	7106	545	36	0.9606
10	6020	5818	342	140	0.9593
11	3336	2767	588	19	0.9005
12	1307	1144	189	26	0.9123
13	2006	2529	107	630	0.8375
14	1745	1679	121	55	0.9486
15	593	555	67	29	0.9164

C. Convergence speed analysis

The results in the Figure 3 showed the computing time of fifteen MR menigioma images segmented by MDP and DMMDP algorithms, which showed that the computing time of DMMDP algorithm was much shorter than that of the classical MDP segmentation result. This result showed the advantage at convergence speed of the new algorithm.

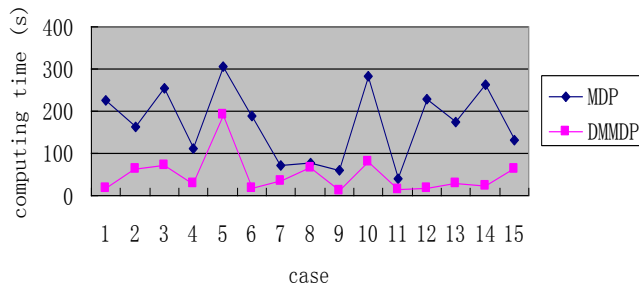


Figure3. Convergence speed analysis

V. CONCLUSION

We developed a kind of new nonparametric image segmentation method for the MR brain tumor images, and applied it to menigioma data. Based on actual clinical data, our algorithm demonstrates a very impressive performance. The particular advantage is that the new algorithm could segment menigioma images correctly and the computing time was significantly shorter than the classical MDP model algorithm. So this method can be used in complicated brain tumor images with high noise and biasing field effect. The further work is to extend our algorithm to 3D and to explore further in the studies for more clinical applications.

ACKNOWLEDGMENT

This work was partly supported by the National Basic Research Program of China (Grant No. 2010CB732505) and the Research Program of Bureau of Education of Guangzhou Municipality of China (Project No. 2012A072). We would like to thank the Radiology Department of Tianjin General Hospital for clinical assistance and providing all the magnetic resonance images in our experiments.

REFERENCES

- [1] Ferguson, T. S., "A Bayesian analysis of some nonparametric problems," *Annals of Statistics*, 1(2), pp. 209-230, 1973.
- [2] Antoniak, C. E., "Mixtures of Dirichlet processes with applications to Bayesian nonparametric estimation," *Annals of Statistics*, 2(6), pp. 1152-1174, 1974.
- [3] Escobar, M. D., "Estimating normal means with a Dirichlet process prior," *Journal of the American Statistical Association*, 89(425), pp. 268-277, 1994.
- [4] MacEachern, S. N., "Estimating normal means with a conjugate style Dirichlet process prior," *Communications in Statistics: Simulation and Computation*, 23, pp. 727-741, 1994.
- [5] Neal, R. M., "Markov chain sampling methods for Dirichlet process mixture models," *Journal of Computational and Graphical Statistics*, 9, pp. 249-265, 2000.
- [6] Balafar MA, Ramli AR, Saripan MI, Mashohor S, "Review of brain MRI imagesegmentation methods," *Artif Intell Rev*, 33, pp. 261-274, 2010.
- [7] Xiea K, Yanga J, Zhanga ZG, Zhub YM, "Semi-automated brain tumor and edema segmentation using MRI," *Eur J Radiol*, 56:12-19, 2005.
- [8] Wafa M, Zagrouba E, "Tumor extraction from multimodal MRI," *In Computer Recognition* 3(57), pp. 415-422, 2009.
- [9] Jinshan Tang, Garry Choy, Peter Choyke, "Dynamic Contrast Enhanced MRI Predicts Early Response to Tumor Necrosis Factor in a Colon Carcinoma Model," *Investigative radiology*, 41(9), 2006.
- [10] Jinshan Tang, Steven Millington, Scott Acton, Jeff Crandall and Shepard Hurwitz, "Surface Extraction and Thickness measurement of the articular Cartilage from MR images using Directional Gradient Vector Flow Snake," *IEEE Tr. On Biomedical Engineering*, 52(5), 2006.
- [11] G. Gerig, O. Kubler, R. Kikini, and F.A. Jolez, "Nonlinear anisotropic filtering of MRI data," *IEEE Trans. Medical Imaging*, 11, pp. 221-232, 1992.
- [12] J.Besag, P.Green, D. Higdon, and K. Mengersen, "Bayesian computation and stochastic systems," *Bayesian computation and stochastic systems. Statistical Science*, 10(1), pp. 3-66, 1995
- [13] Li Y, Shen Y, "Fuzzy c-means clustering based on spatial neighborhoodinformation for image segmentation," *J Syst Eng Electron*, 21, pp. 323-328, 2010.