Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Evaluation of Two Segmentation Methods on MRI Brain Tissue Structures

X. Cai, Y. Hou, C. Li, J-H. Lee*, and W. G. Wee

Abstract—**In this paper, we evaluate two segmentation methods on 15 brain tissue structures. One is narrow band level set method and the other is pattern classification method based on maximum a posteriori (MAP) probability framework. Two sets of experiments are conducted on 18 verified MRI brain data sets. Dice Similarity Index (DSI) is used to evaluate the closeness between our segmentation results and the gold standards, which were provided by experienced radiologists. The results for comparison of two methods are given and their potential applicability is discussed. Tissue structures such as left and right lateral ventricle have achieved over 70% DSI, while other structures such as third ventricle, caudate nucleus, globus pallidus, putamen and thalamus have achieved above 60% DSI.**

I. Introduction

In this paper, evaluations of two segmentation methods, the narrow band level set method [1] and a classification method [2], on a set of 18 verifiable MRI brain tissues [3] are provided. Two sets of tests are conducted on a total of 15 brain tissue structures. The main purpose of this investigative work is to evaluate the applicability of these two automatic segmentation methods on the newly available manually-segmented brain structures. The Dice Similarity Index (DSI) [4, 5] is used as a measure of closeness between computer segmented tissue volumes and their corresponding published hand drawn results. Computer segmented results with mean DSI indices grater than 70% are considered to be acceptable, and such method is considered to be suitable for the application of segmenting specific tissue structures.

We first provide a brief summary of the two methods by introducing the basic concepts and their implemented functional flow charts in Section 2. A description of the test data is discussed in Section 3. Testing details together with parameter settings and results are presented in Section 4. Conclusions and remarks are in Section 5.

X. Cai is with the Department of Electrical and Computer Engineering and Computer Science, University of Cincinnati, Cincinnati, OH 45221 USA (phone: 513-556-4773; e-mail: caixg@ececs.uc.edu).

Y. Hou was with the Department of Electrical and Computer Engineering and Computer Science, University of Cincinnati, Cincinnati, OH 45221 USA. He is now with Professional Engineering Corporation, Kansas City, KS 66215 (e-mail: houyu78@yahoo.com).

C. Li is with the Department of Electrical and Computer Engineering and Computer Science, University of Cincinnati, Cincinnati, OH 45221 USA (e-mail: lic4@email.uc.edu).

J-H. Lee*, the corresponding author, is with the Department of Biomedical Engineering, and the Center for Imaging Research, University of Cincinnati, Cincinnati, OH 45221 USA (e-mail: leej8@uc.edu).

W. G. Wee is with the Department of Electrical and Computer Engineering and Computer Science, University of Cincinnati, Cincinnati, OH 45221 (e-mail: wwee@ececs.uc.edu).

II. Segmentation Methods

A. Narrow Band Level Set Method

Level set method [1] is a powerful numerical technique for tracking the evolution of interfaces moving under a variety of complex motions. It is a way to implement a deformable contour method with a velocity function. Unlike snakes or other deformable models, the level set method can overcome the problem of topology changes on the contour or surface during the evolution process.

The key idea of the level set method is to assign a named value level set to each pixel of the image after an initial zero level set (initial contour) is placed in the object needed to be segmented. Each pixel's level set value is updated based on the speed function F of that pixel. At any time, pixels with level set values equaled to zero are used to compose the contour. The speed function is generally related to the pixel intensity. Driven by the speed function, the contour (surface) will propagate to match the object's boundary. To make the propagating contour accurately reach the object's boundary, choosing an appropriate speed function is very important.

The following expression is the popular speed function:

$$F = k_I \left(F_0 + F_1 \left(K\right)\right) \tag{1}$$

Usually $F_0 = 1$ and $F_1(K) = -\varepsilon K$, where ε is a value that controls the smoothness of the propagating curve or surface and is normally set at a very small value. K is the curvature of the curve or surface, which can keep the propagating interface smooth, and the stopping term k_I is generally defined as an inverse function of the image or surface gradient and has values closed to zero at the boundary pixels where intensity gradient values are large and values close to unity in regions with a relatively constant intensity. This term is the primary factor to stop the propagation of the contour front at places with high image gradient. Specifically, different k_I will produce different results. From our test of the segmentation of different MRI brain tissues, we find that the following k_I [11, 12] provides us with the best results.

$$k_I = A \times B \tag{2}$$

where $A = 1 + \nabla h$ and $h = 1/(1 + |\nabla G_\sigma * I|^2)$. $B = e^{-\frac{|I_{(i,j)} - I_0|}{\omega}}$ and I_0 is the average intensity inside the contour. ∇h, the gradient of *h*, is often used as an attraction force that attracts the contour to locations with large image gradient. A large value of $\nabla h(i, j)$ indicates that point (i, j) is close to strong edge segments. G_σ is a Gaussian distribution with a zero mean and variance σ^2, *I(i,j)* is the intensity at pixel (i, j) and * is the convolution operator. ω is the dynamic range of the interior intensity. $A(i, j)$ has a large value when pixel (i, j) is close to a boundary. $B(i, j)$ is an exponential function, which can make the evolving surface stop at the objects' boundary, especially for some indistinct objects. So the advantage of speed function with $A \times B$ is that when the interface points are close to the boundary, the propagating interface can be attracted to the boundary with a large speed because of the large value of *A*. Similarly, when the interface reaches the object's boundary, the interface points will have a near zero speed because of the small value of *B*.

To save computation time, the narrow band level set method is used. In this method, only pixels within a narrow band around the contour rather than all pixels on the image are updated in each iteration. A functional flow chart is given in Figure 1 for extracting 3D boundary surface.

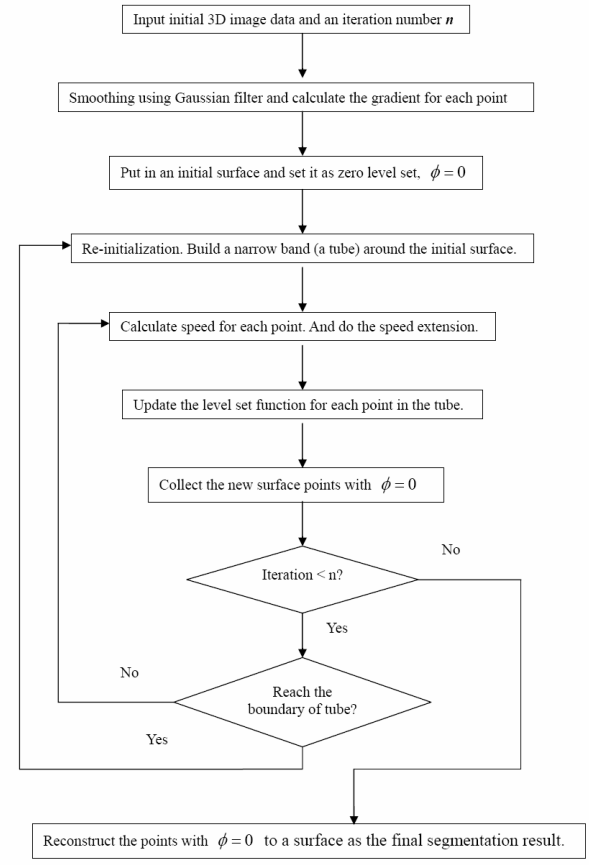

Fig. 1. Flow chart for narrow band level set method.

B. A Pattern Classification Method

In this method, Markov models are employed to derive a recursive algorithm for computing the maximum a posteriori (MAP) probability [2]. The a priori probability is modeled based on a Markov Random Field (MRF) [6, 7]. Spatial-Varying Gaussian Mixture (SVGM) model is used to model the intensity probability distribution of different tissue structures. A learning mechanism has been designed that can learn to classify tissue structures from a reference MR brain image, which has been segmented into different tissue structures by experienced radiologists and it is available in the internet [3]. The learning mechanism consists of two major parts: i) the learning process and ii) the verification process. The learning process uses a modified K-mean (MKM) algorithm to perform knowledge abstraction and representation. In order to assure that all learned cluster centers are located within the region of their respective representations, we employ the adaptive sample set construction (ASSC) algorithm [8] to increase the number of centers for each erroneously designated center until no additional center designation is needed. The verification process is to verify the validity of the learned knowledge using an MRF decision algorithm operating on the same image. Then, with the learned cluster center results and parameter settings, we apply a classification algorithm based on iterated conditional modes (ICM) [9] to the sequential MR brain image slices for structure segmentation. The parameter estimation and structure labeling algorithms are iteratively operated until the results converged.

The complete segmentation algorithm consists of the following steps:

1.) Use learning process to estimate the parameters of SVGM from the reference image.
2.) Use the ICM algorithm to perform structure segmentation according to the parameters obtained from step 1.
3.) Use the MKM algorithm to re-estimate the parameters according to the latest segmentation result.
4.) Repeat step 2 and 3 until the results converged.

This classification method is best suited for segmenting three major brain tissues: grey matter (GM), white matter (WM) and cerebral spinal fluid (CSF). To apply it to the structure segmentation problem and also to improve the speed, for each data set, we first obtain the region of interest that includes the ventricles and other deep GM structures. For each image slice, we use ventricle, one GM structure and the background to represent the three classes, CSF, GM, and WM, in the original algorithm respectively.

III. TESTING DATA SETS

In our experiments, 18 MRI brain volumetric data sets were used. The MR brain data sets and 15 manually

segmented tissue structures were provided by the Center for Morphometirc Analysis (CMA) at Massachusetts General Hospital and are available at http://www.cma.mgh.harvard.edu/ibsr/ [3]. Each data set is a T1-weighted volumetric data with a data size of 256×128×256. There are 7 data sets with an imaging resolution of 0.9375mm× 0.9375mm×1.5mm (X×Z×Slice), 6 data sets 1mm×1mm×1.5mm, and 4 data 0.837mm×0.837mm×1.5mm. The imaging resolution of data set number 6 was not available from the publication (i.e. the internet). Manual segmentation results for 43 brain tissue structures considered to be the gold standards were provided by the authors of the CMA. We selected only the tissue structures from right and left hippocampus, right and left amygdale (combined), right and left putamen, right and left thalamus, right and left caudate nucleus, right and left globus pallidus, nucleus accumbens, right and left lateral ventricle and the third ventricle. We selected these 15 ventricle and deep grey matter structures primarily because of their significance in clinical interests. The size and location of these structures as well as the applicability of our methods to these structures were also taken into consideration. All data sets have been 'geometrically normalized' to the Talaiarch coordination and have also been processed by the CMA 'autoseg', a biasfield correction routine, before they were published.

IV. Parameter Settings and Results

In this paper, we chose DICE similarity index [4] as our quantitative error measurement:

$$D(A,B) = 2|A \cap B| / (|A| + |B|) \times 100\% \quad (3)$$

where A denotes the area/volume of the gold standard and B denotes the area/volume of the segmentation result obtained from our algorithms. $|A \cap B|$ denotes the number of pixels/voxels in the overlay of the gold standard and our segmentation result, $|A|$ and $|B|$ denote the number of pixels/voxels in the gold standard and our segmentation results respectively. Segmentation results with DSI over 70% are considered to be excellently agreed with the gold standards [5, 10]. Both segmentation methods were apply to all 18 MRI brain data set to isolate the 15 tissue structures. For the narrow band level set method, seven parameters were selected after a rather careful experimental evaluation. Detailed information has been reported in elsewhere [12]. Two voxels were manually chosen as the initial centers and two 3x3x3 cubic surfaces centered at these two voxels were used as the initial surfaces. We found that further increasing the number of the initial surfaces did not substantially improve the performance. Total number of iterations n was set to 50 that was sufficient for our cases. Time-step size T was set to 1.6 unit time for all GM structures and 6 for ventricles. The average intensity of ventricle was significantly different from other GM structures and its boundary is rather clear, therefore larger value of T can increase the convergence speed without decreasing the accuracy. From our experience, we found that setting the bandwidth D to be 3 provided a good balance between speed and accuracy. Usually, the standard deviation with zero mean Gaussian filter, σ, is set between 0.5 and 1.2. In this work, we set σ to be 0.6. ε governs the smoothness of the propagating curve or surface and normally has a value between 0.005 and 0.5. A value of 0.025 was set for ε. The intensity dynamic range of the interior ω, which was data related and was tuned to 8 in our study. These parameters were applied for all data sets. For the classification method, two different sets of parameters were selected. One was the numbers of initial cluster centers for each of three classes. They were set to 400, 5, and 10 for the background, GM structure, and the ventricle, respectively, and remained the same for all data sets. The other was the window sizes for searching the neighborhood cluster centers of a given pixel while calculating the maximum a posterior probability. The values of these parameters were between 5 and 15 pixels and varied according to structure size.

Test results for both methods are summarized in table 1.

TABLE 1
DSI PERFORMACE FOR BOTH METHODS

Stru-cture	Narrow Band Level Set Method		Classification Method		Average Volume in Voxels
	Mean (%)	STD (%)	Mean (%)	STD (%)	
RLV	75.32	8.03	78.00	4.90	6375
LLV	71.0	12.73	75.54	5.41	5722
RCN	68.3	9.39	54.14	6.56	2611
LCN	71.3	6.87	59.09	7.93	2705
RGP	68.3	5.18	50.74	9.69	1307
LGP	62.0	8.44	54.64	6.99	1258
RPU	64.9	7.68	45.04	10.39	3804
LPU	61.9	7.71	49.94	6.11	3877
TV	63.29	7.34	64.50	7.56	1450
RTH	57.59	10.01	65.52	4.24	5905
LTH	61.09	6.26	65.55	3.97	5955
RHP	58.41	7.73	49.22	7.18	2794
LHP	56.99	7.96	54.24	3.96	2803
NA	54.93	7.32	47.77	9.08	791
AMG	52.76	8.69	53.02	7.12	2351

STD: standard derivation; RLV: right lateral ventricle; LLV: left lateral ventricle; RCN: right caudate nucleus; LCN: left caudate nucleus; RGP: right globus pallidus; LGP: left globus pallidus; RPU: right putamen; LPU: left putamen; RTH: right thalamus; LTH: left thalamus; RHP: right hippocampus; LHP: left hippocampus; NA: nucleus accumbens (left and right); AMG: amygdale (left and right); TV: third ventricle

V. Conclusions

The narrow band level set method achieves acceptable results with DSI values greater than 70% and a reasonable STD for the left and right lateral ventricle. Next to this performance are caudate nucleus (right: 68.3% and left:

71.3%), globus pallidus (right: 68.3% and left: 62.0%), putamen (right: 64.9% and left: 61.9%) and the third ventricle (63.29%). These results show that this method can be considered for the segmentation of these structures with care. The performances (i.e. DSI) for thalamus (right: 57.59% and left: 61.09%), hippocampus (right: 56.99% and left: 54.93%), nucleus accumbens (54.93%) and Amygdala (52.76%) are all below 60%. In the current state, this method is not recommended for segmenting out these structures. Since all parameters for the narrow band level set method are consistent for all data sets, after a large number of data sets have been successfully evaluated, this automatic method can be considered for the segmentation of the lateral ventricle in the clinical application which often requires efficient and accurate method due to extensive number of data sets.

The pattern classification method outperforms the narrow band level set method for the segmentation of the lateral ventricle. The results (a DSI of 78.00% with a STD of 4.90% and a DSI 75.54% with a STD of 5.41% for the right and left lateral ventricle, respectively) demonstrated that the segmentation of lateral ventricle using the classification method is more consistent and accurate than the other approach. Apparently, it also seems to achieve better results for thalamus (a DSI of 65.52% with a STD of 4.24% and a DSI of 65.55% with a STD of 3.97% for the right and left thalamus, respectively) and the third ventricle (a DSI of 64.50% with a STD of 7.56%). For other GM structures, the classification method fails to obtain acceptable results (all these DSIs were below 60%). The low intensity contrast among these GM structures is the major obstacle to the success of the classification method. The result can be substantially improved if the shape information of these structures can be incorporated into the decision making process. The drawback of the classification method is the requirement of varying parameter settings for different structures. In order to apply this method extensively in clinical usage, standardization of the parameters is strongly desired.

ACKNOWLEDGMENT

We would like to thank reviewers for their very thoughtful comments, which improve the quality and the presentation of this paper.

REFERENCES

[1] J. A. Sethian, *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision and Materials Science,* Cambridge University Press, Cambridge, UK 1999.

[2] Z. Peng, W. G. Wee, and J. H. Lee, "MR brain imaging segmentation based on spatial Gaussian mixture model and Markov random field", *IEEE Int. Conf. Image Processing(ICIP05)*, vol 1. pp. 313-316, Genoa, Italy, Sept. 11-14, 2005.

[3] The Internet Brain Segmentation Repository (IBSR), Center for Morphometric Analysis, Massachusetts General Hospital. Available: http://www.cma.mgh.harvard.edu/ibsr/.

[4] L. R. Dice, "Measures of the amount of ecologic association between speicis," Ecology, vol. 26, pp. 297-302, 1945.

[5] A. P. Zijdenbos, B. M. Dawant, R. A. Margolin, and A. C. Palmer, "Morphometric analysis of white matter lesions in MR images: methods and validation," *IEEE Trans. Med. Imag.*, vol. 13, no. 4, pp. 716-724, 1994.

[6] D. Geman, S. Geman, C. Graffigne, and P. Dong, "Boundary detection by constrained optimization," *IEEE Trans. Pattern Anal. Machine. Intell.*, vol. 12, no. 7, pp. 609-628, 1990.

[7] J. L. Marroquin, B. C. Vemuri, S. Botello, F. Calderon, and A. Fernandez-Bouzas, "An accurate and efficient Bayesian method for automatic segmentation of brain MRI," *IEEE Trans. Med. Imag.*, vol. 21, no. 8, pp. 934-945, 2002.

[8] G. S. Sebestyen, "Pattern recognition by an adaptive process of sample set construction," *IRE Trans. Information theory*, vol. IT-8, pp. s82-s91, 1962.

[9] J. Besag, "On the statistical analysis of dirty pictures," *J. Roy. Statist. Soc. B*, vol. 48, no. 3, pp. 259-302, 1986.

[10] J. J. Bartko, "Measurement and reliability: statistical thinking considerations," *Schizophrenia Bullet.*, vol. 17, no. 3, pp. 483-489, 1991.

[11] X. Wang, L. He, and W. G. Wee, "Deformable contour method, a constrained optimization approach," *Int. J. Comput. Vision*, vol. 59, no. 1, pp. 87-108, 2004

[12] Y. Hou, "Application of a 3D level set method in MRI surface segmentation," M.S. thesis, Dept. Elect. Comput. Eng. Comput. Sci. Univ. Cincinnati, Cincinnati, OH, 2005