Multiresolutional Graph Cuts for Brain Extraction from MR Images

Yong-Sheng Chen[a,b], Li-Fen Chen[c], and Yi-Ting Wang[a]

[a]Department of Computer Science, National Chiao Tung University, Hsinchu, Taiwan
[b]Institute of Biomedical Engineering, National Chiao Tung University, Hsinchu, Taiwan
[c]Institute of Brain Science, National Yang-Ming University, Taipei, Taiwan

ABSTRACT

This paper presents a multiresolutional brain extraction framework which utilizes graph cuts technique to classify head magnetic resonance (MR) images into brain and non-brain regions. Starting with an over-extracted brain region, we refine the segmentation result by trimming non-brain regions in a coarse-to-fine manner. The extracted brain at the coarser level will be propagated to the finer level to estimate foreground/background seeds as constraints. The short-cut problem of graph cuts is reduced by the proposed pre-determined foreground from the coarser level. In order to consider the impact of the intensity inhomogeneities, we estimate the intensity distribution locally by partitioning volume images of each resolution into different numbers of smaller cubes. The graph cuts method is individually applied for each cube. Compared with four existing methods, the proposed method performs well in terms of sensitivity and specificity in our experiments for performance evaluation.

Keywords: MRI, brain extraction, graph cuts

1. INTRODUCTION

Brain extraction is the work of excluding regions of non-brain tissues such as skull, muscle, and fat from the head magnetic resonance (MR) images. Accurate brain extraction is important to structural analysis procedures, such as intensity inhomogeneity correction, tissue segmentation, registration, voxel-based morphology (VBM),[1] and cortical thickness analysis,[2] for non-invasive and quantitative investigation of human brain structure in vivo.

Many methods have been proposed for the brain extraction problem. These methods can be classified into four categories: intensity-based, edge-based, deformable, and hybrid approaches. Intensity-based approaches use intensity thresholds to segment brain and non-brain regions and are sensitive to noise or intensity inhomogeneity. Watershed (WAT) method[3] belongs to this category and it may encounter over-segmentation problem. Edge-based approaches use gradients to locate the boundary between brain and non-brain regions. As an example, Brain Surface Extractor (BSE)[4] uses anisotropic diffusion filtering to smooth noise followed by Marr-Hildreth edge detector for locating closed contours and mathematical morphology for refinement. Among the deformable approaches, BET[5] first finds a brain/non-brain threshold to initialize a triangular tessellation of a spherical surface and then iteratively deforms the triangular model to the brain surface. As another example, ISTRIP[6] uses Wendland's radial basis function (RBFs) as the implicit surface model and iteratively moves the locations of RBFs toward the brain contours. As an example of hybrid approaches, Hybrid Watershed (HWA) combines pre-flooded WAT and deformable surface model for better accuracy of brain extraction.

The pursuits of high sensitivity and high specificity are generally difficult to achieve simultaneously. Among the brain extraction algorithms, HWA method has the advantage of high sensitivity, which means that most of the true brain regions can be extracted. However, its specificity is low, indicating that many non-brain regions are wrongly classified as brain regions. In this paper, we propose an accurate brain extraction algorithm which uses HWA to obtain the initial brain regions and then trims non-brain parts from the initials with multiresolutional graph cuts. In graph cuts method, each vertex has an edge (t-link) connected to the source S, an edge (t-link) connected to the sink T, and an edge (n-link) connected to each of its neighboring vertices. By assigning the weight of n-link according to the difference of voxel values between neighboring vertices and assigning the weight of t-link according to the similarity to the foreground/background seed, all the vertices can be partitioned into foreground and background groups by applying max-flow min-cut theorem. To reduce the memory requirement, the graph cuts method is applied from coarse to fine level of images in a multiresolution framework.[7,8]

Further author information: Yong-Sheng Chen (corresponding author, E-mail: yschen@cs.nctu.edu.tw, Telephone: +886-3-5131316), Li-Fen Chen (E-mail: lfchen@ym.edu.tw), and Yi-Ting Wang (E-mail: wyt328@gmail.com). This work was supported in part by the National Science Council, Taiwan, under Grants NSC-101-2220-E-009-056 and NSC-101-2628-E-010-001.

Fifth International Conference on Digital Image Processing (ICDIP 2013), edited by Yulin Wang, Xie Yi, Proc. of SPIE
Vol. 8878, 88783B · © 2013 SPIE · CCC code: 0277-786X/13/$18 · doi: 10.1117/12.2031540

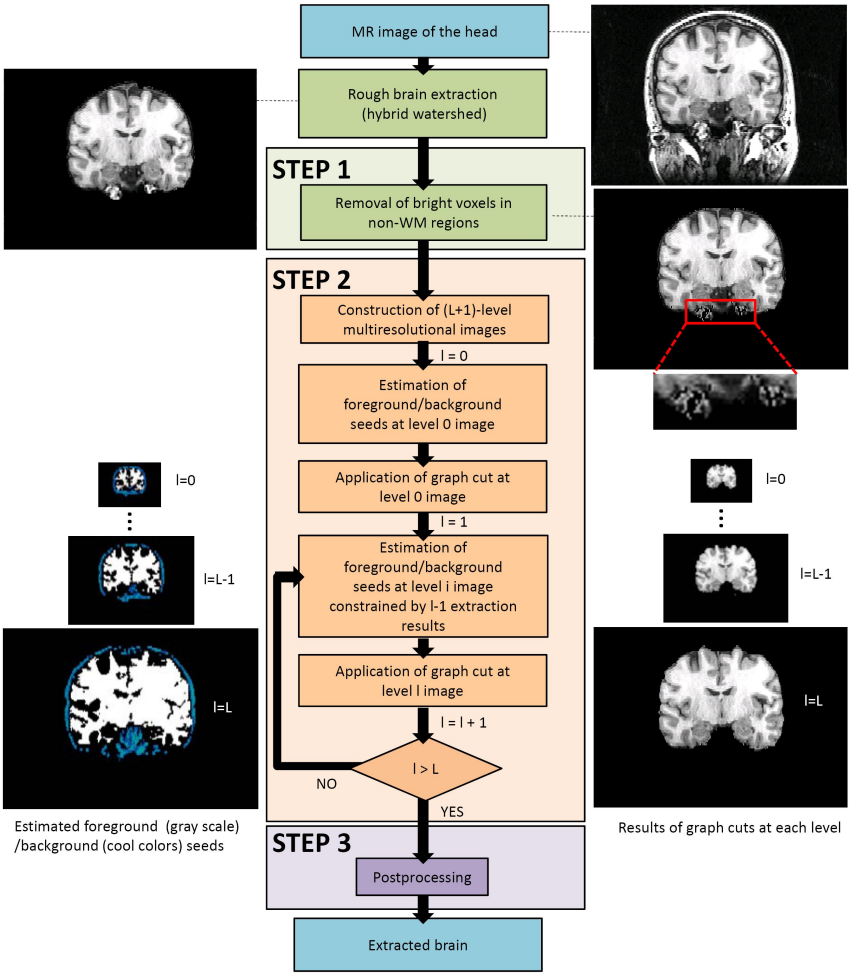

Figure 1. Flowchart of the proposed brain extraction method.

2. METHOD

The proposed brain extraction method is composed of three major steps, as shown in Figure 1. First, we apply the HWA method to initialize the brain region as the roughly extracted brain. In the image of the roughly extracted brain, we remove non-WM voxels with high intensities by segmenting the WM tissue with region growing and excluding those voxels brighter than the segmented WM tissue. This volume is denoted by V^L, where V^L represents that this volume is at level L. The deleted voxels are usually non-brain voxels, such as fat. The foreground seeds are determined by intensity distribution, which may be affected by the voxels with high intensities. Therefore, before application of graph cuts algorithm, we delete those bright voxels according to the estimated WM intensity distribution. We use 3-D region growing and Otsu's[9] method to estimate the WM tissue in the way similar to.[10] The other reason of removing bright voxels in non-WM regions is that they may be classified as foreground by graph cuts algorithm, because the intensities in non-WM brain regions should not be larger than those of WM regions.

The second step is the multiresolutional graph cuts and the advantage of multiresolutional framework to this work is that we can use the result of coarser levels as constraints to finer levels. We downsample V^L image into lower-resolution ones, and from fine to coarse levels are $V^{L-1}, V^{L-2}, \ldots, V^0$. The eight voxels in $L-l$ level (a $2*2*2$ cube) are combined into one voxel in $L-(l+1)$ level, $l = 0, 1, \ldots, L-1$. We start the multiresolutional graph cuts from the coarsest level. The shape of brain in this level is much easier to obtain, because of the brighter voxels not belonging to brain are blurred. To prevent short-cut problem in graph cut, the best foreground seeds should cover the object as much as possible. Hence,

we take the WM as our foreground seeds at the level 0. And the estimation of foreground/background seeds at finer levels are constrained by upper level extraction results.

At the final step, we fill holes, which are background areas surrounded by foreground areas in coronal view. Since the darker voxels in the ventricles might be classified as background by the graph cuts algorithm and the deleted voxels in the first step might include some voxels of WM, and resulting in scattered holes.

The proposed method constructs a graph at each level and apply the graph cuts algorithm to extract the brain region. In the constructed graph, neighboring voxels have a connected edge, which is denoted by n-link. Besides, each voxel has edges, denoted by t-link, connected to two specific vertices: foreground terminal S and background terminal T sas shown in Figure 2. There are two kinds of n-links that need to be assigned with weights, one is the inter-n-link and the other is the intra-n-link. Inter-n-link connects voxels in consecutive slices, shown as black lines in Figure 2, and the intra-n-link connects voxels in the same slice, shown as orange lines in Figure 2. For segmentation of foreground and background, we need to assign edge weight for each edge in the constructed graph and apply graph cuts algorithm. According to the weights of t-link and n-link, the graph cuts algorithm will find a global optimal path to classify foreground and background by minimizing the energy function: $E = \sum_{(x_i,x_j)\epsilon N} A(x_i, x_j) + \sum_{(x_i,S)\epsilon N, or (x_i,T)\epsilon N} B(x_i)$. The x_i is a voxel in the image. N is a set of the neighboring voxels x_i and x_j in the path of the graph cut, and $A(x_i, x_j)/B(x_i)$ is the function to assign the weight of n-link/t-link, respectively.

The finest level has different equation of edge weights from other levels. To assign the weight of n-link at coarser levels, the edge weight is assigned as: $A(x_i, x_j) = w_n \times \exp(-\frac{(I_{x_i}-I_{x_j})^2}{2\alpha^2})$, where w_n is a parameter representing the importance of n-link, I_{x_i} and I_{x_j} are the intensity of two adjacent voxels, x_i and x_j, connected by this link. If the difference between I_{x_i} and I_{x_j} is large, then the weight of this link becomes small. Thus, the cut path is more likely to go through this link.

The weight of n-link at the finest level is assigned as: $A(x_i, x_j) = w_n \times \exp(-\frac{(I_{x_i}-I_{x_j})^2}{2\alpha^2}) + w_d \times (\frac{D(x_i)+D(x_j)}{2})$. In the finest level, we add a distance term $D(x)$ to Eq. (2). The $D(x)$ is defined as the least-squares distance from x to the contour of the extracted brain which is upsampled from the coarser level to the finest level. In this work we use a library which is based on kd-trees and box-decomposition trees for approximate nearest neighbor searching (ANN)[11] to estimate the distance term. The parameter value of w_n and w_d for inter slice is half of the value in intra slice, in order to keep the relationship of different slices but not too strong.

For the weight of t-link connected to foreground terminal is defined as: $B(x_i) = w_t \times \frac{P_F(x_i)}{P_F(x_i)+P_B(x_i)}$, where w_t stands for the importance of this term in contrast to n-link. For each voxel x, $P_F(x_i)$ and $P_B(x_i)$ are its likelihood functions to foreground and background seeds, respectively: $P_F(x_i) = \frac{1}{\sqrt{2\pi}\sigma_f} \exp(\frac{-(I_{x_i}-\mu_f)^2}{2\sigma_f^2})$, $P_B(x_i) = \frac{1}{\sqrt{2\pi}\sigma_b} \exp(\frac{-(I_{x_i}-\mu_b)^2}{2\sigma_b^2})$. The σ_f is the standard deviation of foreground seeds, whereas and σ_b is set as twice the standard deviation of background seeds for increasing the separability of foreground and background by increasing the background range. The μ_f and μ_b are the average of foreground seeds and background seeds. If voxel I_{x_i} is more likely a background seed, then the t-link connected to the foreground terminal is much easier to be cut, then x has more chance to be classified as a background. The weights of foreground seeds connecting to foreground terminal are set to be infinity and they will be classified as foreground. Similarly, t-link connected to background terminal is defined as: $B(x_i) = w_t \times \frac{P_B(x_i)}{P_F(x_i)+P_B(x_i)}$.

The weights of background seeds connecting to background terminal are set to be infinity and they will be classified as background. These edge weight calculation functions were modified from those in the work of Boykov and Jolly.[12] We added a distance term at the original resolution level. In this work we use the graph cuts algorithm proposed in the work of Boykov and Kolmogorov.[13]

3. RESULTS

We applied three data sets to compare the accuracy of brain region extracted by using the existing methods, HWA,[14] BSE,[4] BET,[5] ISTRIP,[6] and the proposed method. The three data sets are the first and the second Internet Brain Segmentation Repository (IBSR) data set from the Montreal Neurological Institute and the data set acquired at Taipei Veterans General Hospital (VGHTPE). This work used five coefficients to compare the accuracy of the brain regions extracted from existing methods, including the Jaccard similarity coefficient (JSC), sensitivity S_e, specificity S_p, and two probabilities of miss rate, p_m and p_f, as described in our previous work.[4]

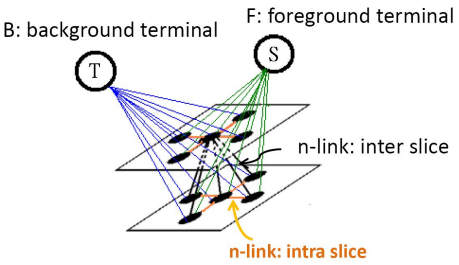

Figure 2. Demonstration of n-link and t-link. Blue lines are the t-link connecting to background terminal. Green lines are the t-link connecting to foreground terminal. Orange lines are the intra slice n-link (link in the same slice). Black lines are the inter slice n-link (link in different slices).

Table 1. Performance evaluation using the first IBSR data set and excluding the failure results of each brain extraction method.

Method	JSC	S_e	S_p	P_m	P_f
BET	0.878 (0.017)	0.983 (0.023)	0.987 (0.004)	0.016 (0.022)	0.107 (0.021)
BSE[4]	0.900 (0.025)	0.954 (0.035)	**0.993 (0.003)**	0.044 (0.034)	**0.055 (0.017)**
ISTRIP	**0.910 (0.018)**	0.986 (0.013)	0.991 (0.005)	0.013 (0.013)	0.077 (0.027)
HWA[3]	0.752 (0.037)	0.974 (0.068)	0.970 (0.008)	0.022 (0.056)	0.226 (0.026)
Our method[5]	**0.910 (0.021)**	**0.989 (0.012)**	0.991 (0.003)	**0.010 (0.011)**	0.079 (0.020)

The superscripts in the first column indicate the excluded cases.

Tables 1, 2, and 3 list the performance evaluation results of BET, BSE, HWA, ISTRIP and the proposed method for different data sets. In addition to Table 1, in other cases, although HWA has the highest S_e, its poor performance of S_p results in its worse performance of JSC. Since some of the brain volumes in the first IBSR data set has apparent intensity inhomogeneity and artifacts, therefore, not all brain volumes have the satisfied result. Table 1 excludes the extracted brains which the JSC value is below 0.6 or the result is blank. The number of excluded brains are four, three, and five for the BSE, HWA, and the proposed method, respectively.

The proposed method outperforms others for the second IBSR set, in terms of JSC, S_e, and S_p. As listed in Table 2, compared to HWA, our method maintains the high sensitivity of HWA as much as possible and improves the low specificity of HWA to the highest one. Our method achieves the highest JSC, the second highest S_e, and the highest S_p. In this data set, the total amount of voxels in an MR volume is 8388168, and the brain region is about one-eighth of the total voxels. Therefore, 0.01 voxels S_e of the value means there are about 10000 of the extracted brains overlapped with the ground truth. The 0.01 voxels S_p of the value means there are about 70000 of the extracted non-brain regions overlapped with the non-brain regions of the ground truth. The 0.01 voxels JSC of the value means there are more than 10000 of the extracted brains overlapped with the ground truth.

When using the VGHTPE data set, although BSE has the highest S_p and JSC values, it has the lowest S_e. This means BSE may exclude many brain voxles. Comparison of BET, ISTRIP, and the proposed method, BET has the highest S_p and the lowest S_e. On the contrary, ISTRIP has the highest S_e and the lowest S_p. The proposed method has the intermediate values of S_p and S_e. The results of performance evaluation for VGHTPE data set are listed in Table 3. The performance evaluation using the BrainWeb phantom data set and only considering WM and GM is listed in Table **??**. Although BSE has the highest S_p, but has the lowest S_e. The proposed method has the best performance compared with BET and ISTRIP, in terms of JSC, S_e, and S_p.

4. CONCLUSION

We have proposed an automatic method based on multiresolutional graph cuts technique to segment the head MR images into brain and non-brain regions. The contributions of this work are 1) using distance term at the highest level to constrain the brain contour, 2) using pre-determined foreground from the coarser level to avoid excluding the brain voxels, 3) and applying graph cuts algorithm individually for smaller cubes to tackle the intensity inhomogeneity problem. We have

Table 2. Performance evaluation using the second IBSR data set.

Method	JSC	S_e	S_p	P_m	P_f
BET	0.891 (0.052)	0.959 (0.042)	0.989 (0.005)	0.038 (0.038)	0.071 (0.031)
BSE	0.838 (0.083)	0.957 (0.042)	0.973 (0.030)	0.041 (0.041)	0.119 (0.104)
ISTRIP	0.915 (0.018)	0.978 (0.011)	0.990 (0.003)	0.021 (0.011)	0.064 (0.022)
HWA	0.814 (0.040)	**0.9997 (0.0003)**	0.965 (0.016)	**0.0002 (0.0002)**	0.186 (0.040)
Our method	**0.930 (0.011)**	0.981 (0.016)	**0.991 (0.005)**	0.018 (0.015)	**0.052 (0.017)**

Table 3. Performance evaluation using the VGHTPE data set.

Method	JSC	S_e	S_p	P_m	P_f
BET	0.921 (0.009)	0.971 (0.013)	0.994 (0.001)	0.028 (0.012)	0.051 (0.011)
BSE	**0.933 (0.008)**	0.958 (0.011)	**0.997 (0.001)**	0.041 (0.011)	**0.026 (0.008)**
ISTRIP	0.915 (0.011)	0.984 (0.012)	0.991 (0.002)	0.015 (0.011)	0.07 (0.015)
HWA	0.847 (0.019)	**0.998 (0.002)**	0.979 (0.003)	**0.002 (0.002)**	0.151 (0.019)
Our method	0.920 (0.010)	0.982 (0.009)	0.992 (0.002)	0.017 (0.009)	0.063 (0.017)

compared to the existing methods and compared with our method by using four data sets. Regarding the performance evaluation for brain extraction, our method outperforms others for the first/second IBSR data set and BrainWeb phantom data set, and comparably with the BET and ISTRIP methods for the VGHTPE data sets.

REFERENCES

[1] J. Ashburner and K. J. Friston, "Voxel-based morphometry-the methods," *NeuroImage* **11**(6), pp. 805–821, 2000.

[2] P. M. Thompson, A. D. Lee, R. A. Dutton, J. A. Geaga, K. M. Hayashi, M. A. Eckert, U. Bellugi, A. M. Galaburda, J. R. Korenberg, D. L. Mills, A. W. Toga, and A. L. Reiss, "Abnormal cortical complexity and thickness profiles mapped in williams syndrome," *Journal of Neuroscience* **25**(16), pp. 4146–4158, 2005.

[3] H. K. Hahn and H.-O. Peitgen, "The skull stripping problem in mri solved by a single 3d watershed transform," *MICCAI 2000, Lecture Notes in Computer Science* **1935**, pp. 134–143, 2000.

[4] D. W. Shattuck, S. R. Sandor-Leahy, K. A. Schaper, D. A. Rottenberg, and R. M. Leahy, "Magnetic resonance image tissue classification using a partial volume model," *NeuroImage* **13**(5), pp. 856–876, 2001.

[5] S. M. Smith, "Fast robust automated brain extraction," *Human Brain Mapping* **17**(3), pp. 143–155, 2002.

[6] J.-X. Liu, Y.-S. Chen, and L.-F. Chen, "Accurate and robust extraction of brain regions using a deformable model based on radial basis functions," *Journal of Neuroscience Methods* **183**(2), pp. 255–266, 2009.

[7] H. Lombaert, Y. Sun, L. Grady, and C. Xu, "A multilevel banded graph cuts method for fast image segmentation," in *Proceedings of International Conference on Computer Vision*, **1**, pp. 259–265, 2005.

[8] A. K. Sinop and L. Grady, "Accurate banded graph cut segmentation of thin structures using Laplacian pyramids," *MICCAI 2006, Lecture Notes in Computer Science* **4191**, pp. 896–903, 2006.

[9] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics* **SMC-9**(1), pp. 62–66, 1979.

[10] S. A. Sadananthan, W. Zheng, M. W. Chee, and V. Zagorodnov, "Skull stripping using graph cuts," *NeuroImage* **49**(1), pp. 225–239, 2010.

[11] S. Arya and D. Mount, "Approximate nearest neighbor queries in fixed dimensions," in *Proceedings of the fourth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 271–280, 1993.

[12] Y. Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images," in *Proceedings of International Conference on Computer Vision*, **1**, pp. 105–112, 2001.

[13] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26**(9), pp. 1124–1137, 2004.

[14] F. Ségonne, M. G. E. Busa, A. D. D. Salat, H. Hahn, and B. Fischl, "A hybrid approach to the skull stripping problem in MRI," *NeuroImage* **22**(3), pp. 1060–1075, 2004.