

Multi-Atlas-based Segmentation with Hierarchical Max-Flow

Martin Rajchl^{1,2*}, John S.H. Baxter^{1,2}, Jing Yuan^{1,3}, Terry M. Peters^{1,2,3}, and Ali R. Khan^{1,3}

¹ Imaging Laboratories, Robarts Research Institute, London, ON

² Department of Biomedical Engineering, Western University, London, ON

³ Department of Medical Biophysics, Western, University, London, ON

Abstract. This study investigates a method for brain tissue segmentation from 3D T1 weighted (T1w) MR images via convex relaxation with a hierarchical ordering constraint. It employs a multi-atlas-based initialization from 5 training images and is tested on 12 T1w MR images provided by the MICCAI 2013 MRBrainS segmentation challenge. The registered atlas images, fully segmented into eight different brain structures, are used to formulate shape and intensity priors for a Maximum A-Posteriori (MAP) energy that is subsequently minimized with a dual hierarchical max-flow computation. The algorithm makes use of a hierarchical label ordering constraint to regularize label families individually and its inherently globally optimal results guarantee robust segmentations. Major parts of the image processing pipeline are implemented using General-Purpose Programming on Graphics Processing Units (GPGPU) for a substantial increase in computation speed.

1 Introduction

Markov Random Field modeling has been of increasing interest to the medical imaging community for the last decade [1–5]. Specifically for multi-region image segmentation, there exist several computationally inexpensive solvers approximating global optimality. A commonly studied model for representing multi-region segmentation is the convex relaxation to Potts model [6], minimizing the following energy functional:

$$E(u) = \sum_{\forall L} \int_{\Omega} (D_L(x)u_L(x) + \alpha S(x)|\nabla u_L(x)|)dx$$
$$\text{s.t. } u_L(x) \geq 0 \text{ and } \sum_{\forall L} u_L(x) = 1$$

where $u_L(x)$ represents a probabilistic segmentation of the image based on *a priori* data terms, $D_L(x)$, and regularization term, $S(x)$ and a parameter, α to weight the respective contributions.

* Send correspondence to mrajchl@robarts.ca

However, these models have been shown to have difficulty in managing multi-region segmentation problems in which several regions have individual regularization requirements not fully representative by a single smoothness term applied to all labels. In a former study [7], we proposed a method based on a partially ordered Potts model to group labels into families and regularize each family individually. This approach allows labels to be grouped and regularized together and can thus treat label groups with different smoothness constraints. In a dual formulation to the partially-ordered Potts model, *Hierarchical Max-Flow* (HMF) computation, the employed solver amounts to a rapid optimization technique, which can be readily implemented using General-Purpose Programming on Graphics Processing Units (GPGPU) and can be used with commercially available hardware configurations to achieve a substantial increase in computation speed. The application of this method towards brain segmentation is natural considering the inherent hierarchical nature of brain anatomy. Using fully-annotated training data we can take full advantage of differing requirements in smoothness that exist across the cortical gray matter, subcortical, and mid-brain structures.

2 Methods

In this study, we propose a fully-automatic multi-atlas initialized segmentation algorithm for brain tissue segmentation using T1 weighted (T1w) MR images. The method is fully automated and major components of the image processing pipeline are implemented using GPGPU. A general overview of the image processing pipeline is depicted in Figure 2.

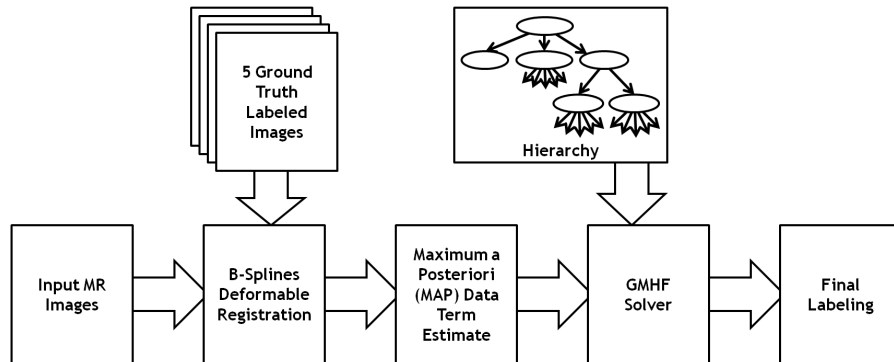


Fig. 1. Image processing pipeline for the proposed multi-atlas initialized HMF method

2.1 Materials

Data was provided by the MRBrainS challenge organizers and consisted of 17 3T MRI datasets: 5 for training, and 12 for testing. Each dataset included a thin-slice T1w image (1mm isotropic), and thick slice T1 inversion recovery (T1 IR) and T2 FLAIR sequences (0.958x0.958x3mm). The thick-slice scans were co-registered and all images were bias-corrected.

2.2 Pre-processing and Multi-Atlas initialization

Computations were performed in the T1w 1mm space, thus all images and labels were initially resampled to this 1mm isotropic resolution. Multi-atlas segmentation with the set of five training subjects was used to provide initial label estimates for the test data. Registration of the training data to the test subjects was carried out with the T1 images using an initial affine block-matching approach [8] (default parameters) followed by fast free-form deformable registration (spline spacing 8mm, bending energy 0.005) as implemented using GPGPU in the `nifty_reg` package [9]. Training labels (with white matter lesions labels combined with white matter) were then propagated to each test subject. The deformed training labels are used i) as seeds for sampling intensities for the intensity prior and ii) as probabilistic shape prior in the data term $D(x)$ (3).

2.3 Brain masking

The T1w images do not possess sufficient contrast between dural matter and CSF thus defining this boundary based solely on these images is difficult. Instead we make use of the T1 IR images to provide the brain mask using atlas-based registration. Deformation maps generated from the multi-atlas registration of T1w images were used to initialize an additional deformable registration update (spline spacing 5mm, bending energy 0.005) to align the training T1 inversion recovery (IR) images with the target T1 IR in the subject image space. A registration cost-function mask (based on the existing T1 registration) is used during this registration to avoid influence of the non-uniform background in the IR images.

2.4 Energy formulation and HMF optimization

The algorithm minimizes a maximum a-posteriori (MAP) energy functional via a dual convex relaxation approach using a hierarchical label ordering constraint. The data term $D(x)$ (3) is formulated as presented in [10] and the provided training labels (Background, brainstem, cerebellum, external CSF, white matter, cortical gray matter, basal ganglia, ventricles) are used as shape information in the energy functional (1): Note that this segmentation is performed without skull-stripping to avoid potential over-stripping of the brain. Given an isotropically resampled T1w image I with training labels T , an energy $E(u)$ (1) is minimized

such that hierarchical label ordering constraints (2).

The proposed energy functional minimized can be formulated as

$$E(u) = \sum_{\forall L} \int_{\Omega} (D_L(x)u_L(x) + \alpha_{L.P}S_{L.P}(x)|\nabla u_L(x)|) dx \quad (1)$$

s.t.

$$\forall L (u_L(x) \leq 0) \text{ and } \exists! S (\#S.P \wedge u_S(x) = 1) \text{ and } \forall L' \left(\sum_{L \in L|L.P=L'} u_L(x) = u_{L'}(x) \right). \quad (2)$$

The parameters, β_L , weight the contribution from the intensity prior and the labeling prior for each label, L in the data term $D(x)$:

$$D_L(x) = -\beta_L \log P(I(x)|L) - (1 - \beta_L) \log(P(T_L(x)) * G(x)), \quad \beta_L \in [0, 1]; \quad (3)$$

$$S(x) = \lambda_1 + \frac{\lambda_2}{1 + e^{\lambda_3 (|\nabla I(x)| - \lambda_4)}} \quad (4)$$

The T1w image was used to determine the intensity priors in form of log-likelihood data terms for each label in the training dataset. Histograms with 128 bins were used to calculate the probability $P(I(x)|L)$. The registered probabilistic labeling $P(T_L(x))$ was Gaussian smoothed with a standard deviation of 0.5 to account for the variability in the cortical folds. The function $S(x)$ was designed to locally weight regularization by gradient information from the input (4). The hierarchical label ordering is depicted in Figure 2.4 along with the corresponding regularization (α) and log-likelihood (β) weighting parameters.

A GPGPU-accelerated hierarchical max-flow optimizer was used to minimize the provided energy equation given the hierarchy. The product of this optimizer is a probabilistic labeling which combines the effects of the data terms, discouraging edges in regions with higher $\alpha S(x)$ values identified as being less likely to include boundaries between labels due to having low gradient magnitude.

All segmentation parameters used to generate the results were determined heuristically from segmentations on the provided 5 training datasets. First the minimum label agreement for the seeds of the intensity prior was determined without regularization, then the respective regularization weights α were adjusted to achieve a maximum Dice Similarity Coefficient on the training images. Lastly the contribution of the shape prior was added by lowering β for labels where leakage could be prevented with increased priors.

2.5 Post-processing

The images were resampled to the thick slice format required at submission by resampling the probabilistic labeling then voting to create the discrete labels into the required background, CSF, white matter, and grey matter regions.

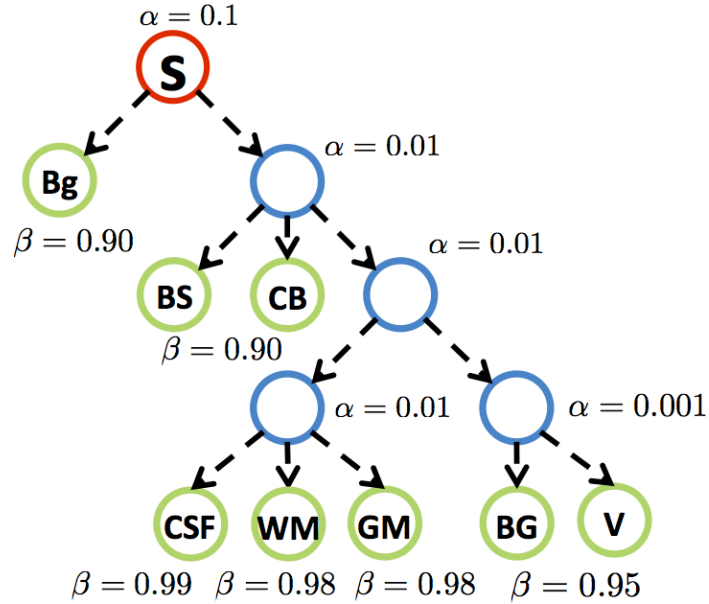


Fig. 2. Hierarchical graph with source (S) terminal, parent nodes (light blue) and leaf nodes (green): Background (Bg), corticospinal fluid (CSF), brain stem (BS), cerebellum (CB), cortical gray matter (GM), white matter (WM), basal ganglia (BG) and ventricles (V). All source flows (dashed arrows) are of unconstrained capacity. Each leaf and parent node state the used α (regularization) and β (data term) weighting parameters, respectively.

After discretization, the T1 IR-based brain masks were multiplied with the final segmentations to mask out dura and non-brain regions that may have been segmented using the T1 alone.

2.6 Computing Specifications and Run Times

Computations were carried out on a Tesla C2070 GPU (NVIDIA, St. Clara, CA) with Ubuntu 12.04 machine and 144 GB RAM, where each HMF computation took approximately 65 seconds. Each pairwise registration from training to target subject was made up of a rigid/affine registration (no GPU acceleration, 2.5 minutes), T1 deformable registration (GPU, 15 seconds) and T1 IR deformable registration for brain masking (GPU, 15 seconds). Post-processing took approximately 5 second per image. Thus with 5 atlases (training data) a given test subject requires approximately 16:10 to complete the entire pipeline.

3 Results & Discussion

The computed results were sent to the MRBrainS challenge organization team and evaluated in terms of label accuracy using the Dice Similarity Coefficient

(Dice), modified Hausdorff distance (Mod. HD) defined as 95 percentile HD and an absolute volume difference. Table 1 contains the numerical accuracy results and example segmentations can be found in Figure 3. The effect of the post-processing step is depicted in Figure 3 for the testing dataset 10.

The Dice overlaps for all three structures were generally high, with the white matter overlap being highest on average. However, the Hausdorff distance showed a different trend, with white matter having the highest distance error. We believe these lower scores to be related to over-smoothing of the anatomy in the cortical ribbon. Further tuning of the regularization parameters for the cortical regions (including white matter and external CSF) would be required to address this. Another issue with the accuracy results is the high degree of volume errors, most notably in the CSF. This can partly be attributed to the aforementioned issue of regularization in the cortex, however, another source of error could be the external boundary of the brain in the inferior regions. This is problematic because of reduced contrast in the T1w and apparent slice intensity artifacts in the T1 IR images. Future work could better couple the use of the T1 and T1 IR to avoid these issues and better resolve the brain mask in these inferior regions. Other issues could be related to registration errors, especially in cases where brain images have significant morphological differences such as enlarged ventricles. The use of atlas selection or weighting schemes in addition to use of a larger training set could address these issues and improve the shape priors.

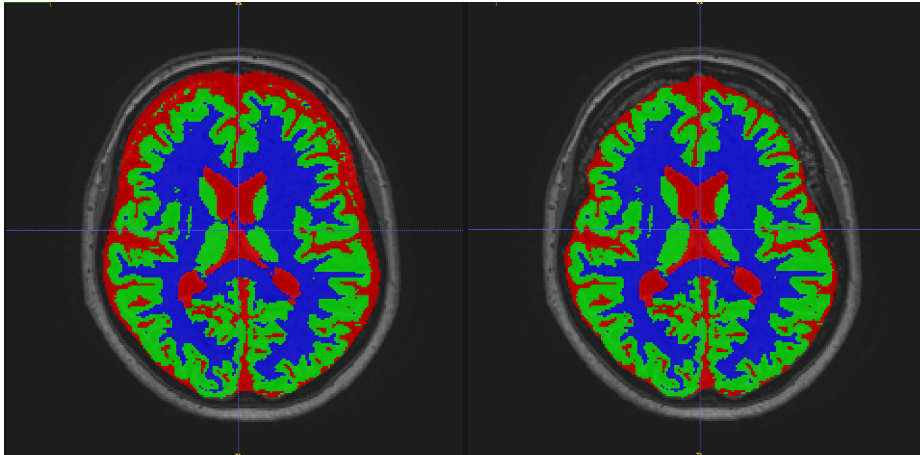


Fig. 3. Effects of brain masking on the CSF region: a) resulting labels after HMF optimization and b) after brain masking

Table 1. Accuracy results for the MRBrainS test database ($N_{TEST} = 12$): Dice Similarity coefficient (Dice), modified Hausdorff distance (Mod. HD), absolute volume differences (Abs. VD)

Structure	Dice Coeff. (%)	Mod. HD (mm)	Abs. VD (%)
	Mean \pm SD	Mean \pm SD	Mean \pm SD
Gray matter	80.49 \pm 1.93	1.95 \pm 0.09	7.26 \pm 5.62
White matter	85.88 \pm 1.03	3.19 \pm 0.37	7.64 \pm 6.37
Cerebrospinal fluid	81.15 \pm 3.42	2.53 \pm 0.41	16.22 \pm 9.06
Brain	93.83 \pm 0.57	2.66 \pm 0.40	6.51 \pm 2.48
All intracranial structures	97.92 \pm 0.28	2.57 \pm 0.43	0.74 \pm 0.42

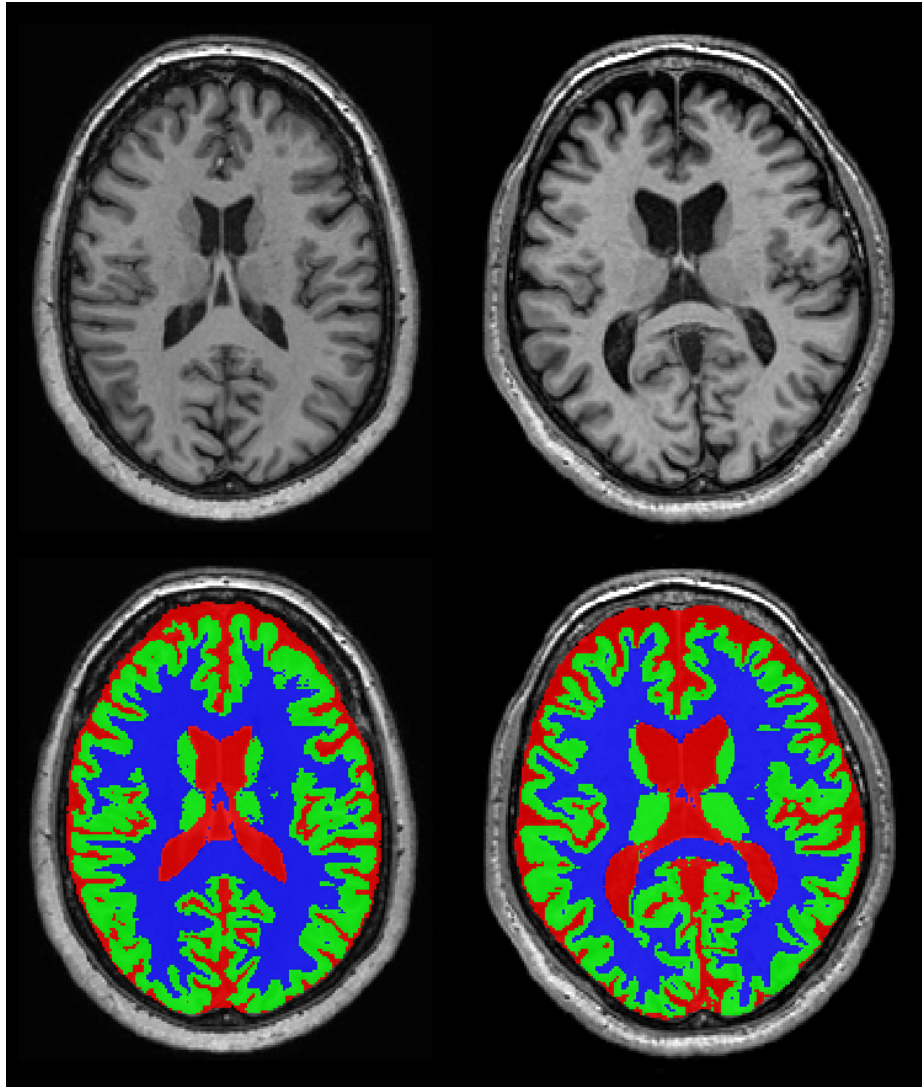


Fig. 4. Example axial results for worst (left, $N_{TEST}[1]$) and best case (right, $N_{TEST}[8]$) results determined by overall Dice Coefficient for CSF, GM and WM.

Segmentation methods that can take advantage of multi-sequence MRI data can potentially provide more accurate results because of the complementary information that these sequences provide. The T1 IR images for instance provide contrast critical to defining the outer boundary of the brain, however, exploiting this information with existing tools proved challenging because the background intensities overlapped with the brain intensities. Our approach used the IR images in a refinement registration step (with a brain-based cost-function mask) to generate an updated brain mask. Another option for including multi-sequence information would involve adding additional data log-likelihood terms to the segmentation cost function. Our initial experiments found this to be less accurate, likely due to the reduced contrast and sometimes blurred boundaries in the thick-slice images, thus we did not apply this approach in our final submission.

The generalized hierarchical framework we applied has several benefits with respect to flexibility in parameters for different label groups, however, with this added flexibility the challenge then becomes how to efficiently select, or tune, these parameters. The approach outlined here employed a heuristic determination of parameters, thus much improvement could be gained with automating this procedure, which we plan to address in future work.

4 Conclusion

We developed an automated method to segment brain tissue from T1w MR images. It makes use of a partially-ordered Potts model label configuration that allows to regularize label groups individually. The proposed pipeline is able to generate fast and approximate globally optimal results and can be implemented using commercially available hardware.

Acknowledgements

The authors would like to thank A.J. McLeod for valuable discussion on the topic. Graduate student funding for M. Rajchl and J.S.H. Baxter was provided by scholarships from the National Science and Engineering Research Council of Canada and by Western University. A.R. Khan was supported by a CIHR Fellowship.

References

1. Yongyue Zhang, Michael Brady, and Stephen Smith. Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm. *Medical Imaging, IEEE Transactions on*, 20(1):45–57, 2001.
2. Yuri Boykov and Marie-Pierre Jolly. Interactive organ segmentation using graph cuts. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2000*, pages 276–286. Springer, 2000.
3. Michael Wels, Gustavo Carneiro, Alexander Aplas, Martin Huber, Joachim Hornegger, and Dorin Comaniciu. A discriminative model-constrained graph cuts approach to fully automated pediatric brain tumor segmentation in 3-d mri. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008*, pages 67–75. Springer, 2008.
4. Zhuang Song, Nicholas Tustison, Brian Avants, and James C Gee. Integrated graph cuts for brain mri segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2006*, pages 831–838. Springer, 2006.
5. Lichen Liang, Kelly Rehm, Roger P Woods, David A Rottenberg, et al. Automatic segmentation of left and right cerebral hemispheres from mri brain volumes using the graph cuts algorithm. *NeuroImage*, 34(3):1160, 2007.
6. Jing Yuan, Egil Bae, Xue-Cheng Tai, and Yuri Boykov. A continuous max-flow approach to potts model. In *Computer Vision–ECCV 2010*, pages 379–392. Springer, 2010.
7. Martin Rajchl, Jing Yuan, James A White, Cyrus MS Nambakhsh, Eranga Ukwatta, Feng Li, John Stirrat, and Terry M Peters. A fast convex optimization approach to segmenting 3d scar tissue from delayed-enhancement cardiac mr images. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*, pages 659–666. Springer Berlin Heidelberg, 2012.
8. Sébastien Ourselin, Radu Stefanescu, and Xavier Pennec. Robust registration of multi-modal images: towards real-time clinical applications. In *Medical Image Computing and Computer-Assisted Intervention MICCAI 2002*, pages 140–147. Springer, 2002.
9. Marc Modat, Gerard R Ridgway, Zeike A Taylor, Manja Lehmann, Josephine Barnes, David J Hawkes, Nick C Fox, and Sébastien Ourselin. Fast free-form deformation using graphics processing units. *Computer methods and programs in biomedicine*, 98(3):278–284, 2010.
10. Fedde van der Lijn, Tom den Heijer, MM Breteler, and Wiro J Niessen. Hippocampus segmentation in mr images using atlas registration, voxel classification, and graph cuts. *Neuroimage*, 43(4):708, 2008.