

# A JOINT REGISTRATION AND SEGMENTATION APPROACH TO SKULL STRIPPING

Aaron Carass<sup>1</sup>, M. Bryan Wheeler<sup>1</sup>, Jennifer Cuzzocreo<sup>2</sup>, Pierre-Louis Bazin<sup>3</sup>, Susan S. Bassett<sup>2</sup> and Jerry L. Prince<sup>1</sup>

<sup>1</sup>Image Analysis and Communications Laboratory, Electrical and Computer Engineering,

<sup>2</sup>Division of Psychiatric Neuroimaging, Psychiatry and Behavioral Sciences,

<sup>3</sup>MedIC, Neuroradiology Division, Radiology and Radiological Science,

The Johns Hopkins University

<sup>1</sup>{aaron\_carass,bonus,prince}@jhu.edu      <sup>2,3</sup>{jlc,pbazin1,sbassett}@jhmi.edu

## ABSTRACT

Extraction of the cerebrum, cerebellum, and brain stem from structural magnetic resonance images (MRIs) is an important initial step in neuroimaging. We present an automated algorithm that solves this difficult problem, often referred to as skull stripping, which is novel for its use of registration, segmentation, and morphological operations. Our algorithm is also concerned with an accurate representation of the grey matter boundary, which is a unique feature. We also present results demonstrating the accuracy of this approach.

**Index Terms**—Biomedical image processing, brain, image registration, image segmentation.

## 1. INTRODUCTION

Isolation of brain from non-brain tissue in magnetic resonance images is still a topic of interest despite extensive work in the area [1–11]. This problem, known as skull stripping or brain extraction, is equivalent to a whole brain segmentation which correctly classifies cerebral grey matter (GM) and white matter (WM) from other tissues such as cerebrospinal fluid (CSF), skull, meninges, *etc.* Quantitative neurological image processing (NIP) is often based on structural magnetic resonance images (MRI) that use this preliminary step. In spite of the plethora of approaches, the gold standard for skull stripping remains that of a human rater. This is not desirable, as a human rater may spend several hours to manually segment one whole head MR image, and different raters may introduce biases or inconsistencies.

Automated skull stripping methods have encountered several difficulties. Brain tissues share similar intensities with neighboring extra-cerebral tissues and images are often heterogeneous in terms of intensity scales, patient orientation, or scanned portion of the head. Recent studies have compared the consistency of raters against each other and automated algorithms [12, 13] and observed that different methods will produce different types of segmentations, corresponding to

slightly different skull stripping requirements. The choice of a skull stripping technique should be motivated by the subsequent processing of the data. Unfortunately, there are no studies detailing how a raters or algorithms bias may influence a NIP pipeline.

In this paper, we present a new algorithm for skull stripping that combines elastic registration, tissue segmentation, and morphological techniques into a fast hybrid method that reliably produces accurate results in our experiments. The paper is arranged into a brief overview of existing methods and a discussion of the objectives for our skull stripping approach, follow by a description of our method. The algorithm's performance is then tested and validated against careful human delineation on a population of 18 T1-weighted MRI data sets.

## 2. BACKGROUND AND GOALS

There are four broad classes that cover the techniques used in existing skull stripping methods:

**1. Morphological Based** The most common approach is to apply successive morphological operations on the MRI volume. Generally, these methods first use operator input to determine certain thresholds, the region of interest or a seed for a region growing phase. Morphological dilation and erosion operations are repeated, in some hierarchically context, until a stopping criterion is satisfied or a user-specified end point is reached [1, 2].

**2. Atlas Based** Several groups have used atlases to identify the whole head segmentation. These methods vary in how many templates they use in distinguishing anatomical regions and also how they apply these atlases. Some methods use an affine registration to put the subject in a Talairach space and perform a segmentation within this space. Other algorithms use elastic registration techniques, multiple atlases, and what is termed a *voting scheme* to differentiate the tissue classes [3–5]. Most methods incur the risk of a bias toward the atlas images.

**3. Deformable Surface Based** These approaches evolve a surface using various forces to find the boundary of the brain

This work was supported by the NIH/NINDS under grants R01NS37747 and RO1AG016324.

from an initialization based on user input or other criteria such as the center of mass of the whole head. The brain extraction is thus mostly data-driven, but the definition of the initial guess can be critical [6, 7].

**4. Hybrid Methods** These methods are some combination of the classes listed above with additional elements [8–10]. Many of the algorithms that could be classified distinctly in one of the previous groups are now being refactored to incorporate some feature of the other methods. These methods combine the skull stripping results from different approaches to cancel out the bias or inaccuracies inherent to one of the other methods. Our method, like most modern methods, is best described as a hybrid method.

In most previous approaches, the goal of skull stripping was to segment out GM and WM from all other tissues, without any concern for an accurate representation of the GM boundary with the CSF and other non-WM tissue. As such, we wish to remove skull, meninges, and other tissues without excising any cortical surface GM or WM. The CSF surrounding the cortical surface is important to define the GM boundary, and we devised our method to consistently include a thin layer of CSF around the brain. This also influences how we evaluate our algorithm; we are primarily concerned with producing a mask that fully contains the mask of the human rater.

### 3. METHOD

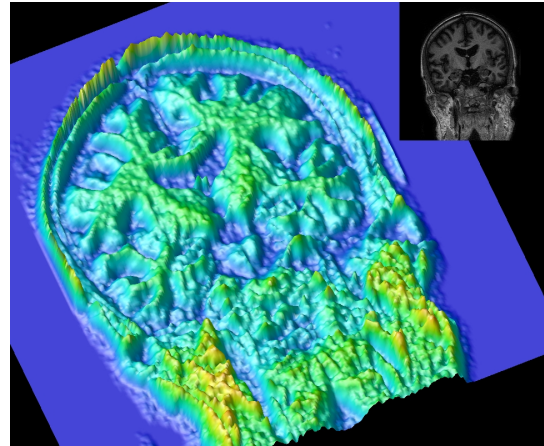
#### 3.1. FUNDAMENTALS

First we would like to outline a central assumption of our method. The GM and WM we are interested in are connected by the following scheme: *there is a path connecting any two GM/WM voxels such that from the point of highest intensity along the path to either end point the intensity is never increasing.*

This is readily understood by taking a slice of an MRI data set and using the intensities as height information. With this representation, voxels are considered to be connected if they are on the same *hill* and disconnected if they are separated by a *valley*, see Fig. 1 for an illustration. It should also be noted that this approach is heavily dependent on underlying properties inherent to T1-weighted MRIs.

With this concept it becomes apparent that if we can reasonably isolate the WM (*hill peaks*) we can perform a morphological dilation based on descending this *hill* to the *valley* represented by the CSF and image background. We will also add an appropriate stopping criterion in those instances when we are traversing a *plane*, with respect to image intensities.

To this end, our algorithm first performs a fast non-linear registration based on the Adaptive Bases Algorithm (ABA) [14] from multiple atlases to the subject. We generate separately a tissue classification on the subject using FANTASM [15]. The registration and segmentation are combined to ini-



**Fig. 1.** Inset we see a slice of a typical data set, while the main image shows a *hill/valley* representation of the same slice. The heights are derived from the image intensities, the color scheme is also derived from the intensities.

tialize a GM/WM region within the subject, two morphological operations follow to ensure we have not included any of the skull, dura or other tissue at this early stage. We then use the image intensity, segmentation and registration results as part of a dilation operation to descend the GM/WM *hill* to the CSF boundary.

#### 3.2. ALGORITHM

As the first step, we wish to identify the *crest of the hill* corresponding to the GM/WM connected component. To derive this initial state we perform an ABA registration [14] between the subject brain and the atlases. ABA registration maximizes mutual information with an underlying deformation field modeled on radially symmetric basis functions. Because we do not rely solely on the registration of an atlas, we perform the registration on the data downsampled by a factor of 4. The registration is thus very fast, and we compensate for the inherent bias of using only one atlas by performing multiple registrations using  $N$  different atlases, with typically  $N = 3$  or 4. We then transform a mask for each atlas into the subject space based on the output of the registration. The mask is a hand-delineated representation of the cerebral cortex, cerebellum, and brain stem. We then generate a probability mask given in the subject space at voxel  $i$  as,  $P_i = (\# \text{ Masks that include } i) / N$ , see Fig. 2(a).

In addition, we generate a tissue segmentation of the whole head MRI using FANTASM [15]. FANTASM is a robust tissue classification algorithm which is an extension of the adaptive fuzzy c-means algorithm. We use four classes within FANTASM, two of these classes ( $\Gamma_2$  and  $\Gamma_3$ ) are an approximate representation of WM and GM though dura, skull, and adipose tissue tend to also be included in these two classes. The other tissue classes break down into a class made up of

CSF, bone and background ( $\Gamma_1$ ) and another ( $\Gamma_4$ ) dominated by skin and adipose tissue and a small number of WM voxels. Fig. 2(b) shows all four tissue classes for a particular data set. We combine the registration and segmentation in order to have the best of what either has to offer. So we produce our initial mask  $\mathcal{M}$ , using the following criterion based on both the registration and segmentation:

$$\mathcal{M} = \left\{ i \mid i \in \Gamma_3, P_i \geq \frac{N-1}{N} \right\} \cup \{ i \mid i \in \Gamma_2, P_i = 1 \}.$$

This initial segmentation provides a representation of the WM, and subcortical GM with a high degree of confidence, see Fig. 2(d). There are two problems with this initial mask: tissue from the neck and around the base of the brain stem are included in the mask, secondly we can have gross errors in the registration that lead to portions of dura and skull being included. To deal with these problems, we perform an aggressive morphological erosion operation on the current mask with a structuring element of  $3 \times 3 \times 3$  voxels. This erosion is performed twice and then we find the largest 6-connected component that remains in the mask and excise all other connected components from the mask, an example is shown in Fig. 2(e).

The next major step is to perform our *hill decent* in a reasonable manner, using both the registration and segmentation. Starting from the eroded initial mask, we add voxel  $i$  on the boundary to our mask  $\mathcal{M}$  if it meets either of the two following criteria:

1.

$$i \in \Gamma_2 \cup \Gamma_3, \quad I_j \geq I_i, \quad P_i \geq \frac{1}{N},$$

for  $j \in \mathcal{M} \cup \Gamma_2 \cup \Gamma_3$  and  $i \in \mathcal{N}_{18}(j)$ ,

2.

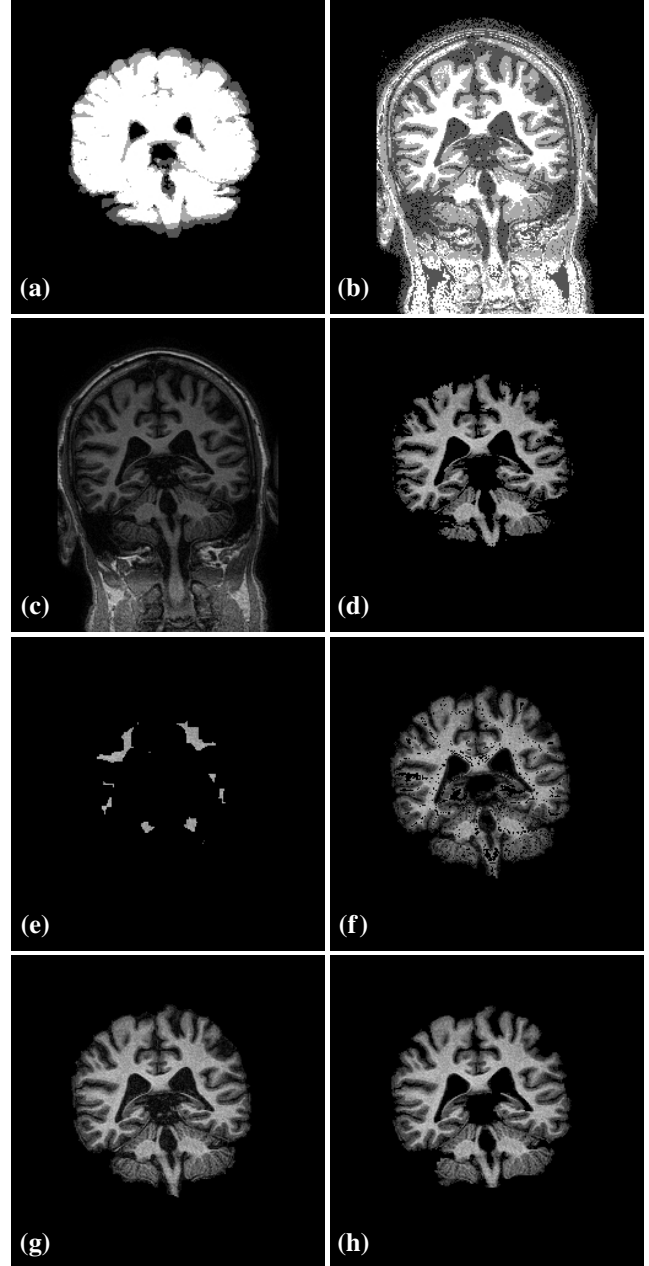
$$i \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3, \quad I_j > I_i$$

for  $j \in \mathcal{M} \cup \Gamma_2 \cup \Gamma_3$  and  $i \in \mathcal{N}_{18}(j)$ ,

where  $I_i$  denotes the intensity of voxel  $i$  and  $\mathcal{N}_{18}(j)$  is the 18-connected neighborhood of voxel  $j$ .

The first condition allows us to grow our mask in regions of non increasing intensity assuming the appropriate tissue classes and the probability mask is encouraging us to grow. While in the second condition we can add a CSF voxel to the mask only if its neighbor is GM/WM and we are strictly descending an image intensity, thus we can dilate beyond the GM boundary, but only by one voxel. We allow the mask to dilate to voxels meeting the above provisions until the number of voxels it has added in a single sweep through the mask is less than  $0 \cdot 333\%$  of the total volume of the MRI data set.

There are voxels that still may not be included in the mask during the dilation because they are local intensity peaks, even though they are areas with an appropriate tissue class and a high probability they would have us ascending an intensity hill and so are excluded at this stage, Fig. 2(f). To correct for these situations, we perform a one voxel dilation, a hole filling and a one voxel erosion to our current mask, we refer



**Fig. 2.** For a given data set we show (a) the probability mask, (b) the tissue classification, (c) the original image, (d) the initial mask, (e) the mask after erosion and retaining the largest connected component, (f) the mask prior to the modified morphological closing, (g)  $\mathcal{M}_A$  and (h)  $\mathcal{M}_H$ .

to this as a modified morphological closing. The hole filling finds the largest 6-connected component that is background and sets all other background components to be part of the mask. We perform the dilation and erosion, simply to avoid situations where the intensity peaks may be 6-connected to the

background.

#### 4. RESULTS AND DISCUSSION

Our validation study was carried out on 18 T1-weighted MRIs of the whole head [16]. Each image had been manually delineated by a human expert to create a mask containing only WM and GM tissues of the cerebrum, cerebellum, and brain stem cut at the lowest boundary of the cerebellum. For each subject three atlases were chosen from a larger population of similarly acquired and delineated data sets.

The primary metrics used for evaluating the algorithm mask,  $\mathcal{M}_A$ , in comparison to the rater mask,  $\mathcal{M}_H$ , are the Jaccard Similarity Coefficient (JSC), Dice Coefficient and containment index (CI). These metrics are defined as follows:

$$\text{JSC} = \frac{|\mathcal{M}_H \cap \mathcal{M}_A|}{|\mathcal{M}_H \cup \mathcal{M}_A|}, \quad \text{CI} = \frac{|\mathcal{M}_A \cap \mathcal{M}_H|}{|\mathcal{M}_H|}.$$

A JSC of 1 represents complete overlap, while a value of 0 denotes no overlap. The Dice coefficient is readily computed from the JSC, we report its value in keeping with previous studies. The Containment Index (CI), how much of  $\mathcal{M}_H$  is contained within  $\mathcal{M}_A$  for a given subject. Similar to the JSC, values of 1 for the CI are desired.

Table 1 shows the results from our study of 18 subjects, comparing  $\mathcal{M}_A$  with  $\mathcal{M}_H$ . The total run time for the algorithm was between 15 and 20 minutes. Our results are in keeping with recent articles [13] on the performance characteristics of various automated skull stripping algorithms. We also note that the CI, which is equivalent to Type 1 error, is higher than those reported by other approaches. From this we can say that our algorithm is identifying more true positives than other reported methods.

**Table 1.** The Mean, Standard Deviation (STD), Minimum, Median and Maximum for the JSC, Dice Coefficient, and the CI.

	Mean	STD	Min	Median	Max
Dice	0.9213	0.0185	0.8833	0.9245	0.9445
JSC	0.8547	0.0316	0.7911	0.8595	0.8950
CI	0.9960	0.0091	0.9597	0.9989	0.9999

We have presented a skull stripping algorithm that has been robust in our small study. In the future we plan to perform further evaluation with a larger population group and a broader comparison with multiple human raters. We also plan some refinements to our method to compensate for partial volume and image inhomogeneities. Another issue we would like to explore is the benefits of preselecting atlases to enhance performance.

#### 5. REFERENCES

- [1] B.D. Ward, "Intracranial segmentation", Milwaukee: Biophysics Research Institute, Medical College of Wisconsin. <http://afni.nimh.nih.gov/afni/>.
- [2] A.F. Goldszal, C. Davatzikos, D.L. Pham, M.X.H. Yan, R.N. Bryan and S.M. Resnick, "An image processing system for the qualitative and quantitative volumetric analysis of brain images", *J. Comput. Assist. Tomogr.*, Vol 22, pp. 827–837, 1998.
- [3] T. Kapur, W.E.L. Grimson, W.M. Wells and R. Kikinis, "Segmentation of brain tissue from magnetic resonance images", *IEEE Trans. Med. Imaging.*, Vol 16, pp. 41–54, 1997.
- [4] J. Ashburner and K.J. Friston, "Voxel-based morphometry: the methods", *Neuroimage*, Vol 11, pp. 805–821, 2000.
- [5] A.M. Dale, B. Fischl and M.I. Sereno, "Cortical Surface-Based Analysis I: Segmentation and Surface Reconstruction", *Neuroimage*, Vol 9, pp. 179–194, 1999.
- [6] S.M. Smith, "Fast robust automated brain extraction", *Human Brain Mapping*, Vol 17, pp. 143–155, 2002.
- [7] F. Ségonne, A.M. Dale, E. Busa, M. Glessner, D. Salat, H. K. Hahn, and B. Fischl, "A hybrid approach to the skull stripping problem in MRI", *Neuroimage*, Vol 22, pp. 1060–1075, 2002.
- [8] H. Hahn and H.O. Peitgen, "The skull stripping problem in MRI solved by a single 3d watershed transform", *MICCAI, LNCS 1935*:134–143, 2000.
- [9] D.E. Rex, D.W. Shattuck, R.P. Woods, K.L. Narr, E. Luders, K. Rehm, S.E. Stoltzner, D.A. Rottenberg and A.W. Toga, "A meta-algorithm for automated brain extraction in MRI", *Neuroimage*, Vol 23, pp. 625–637, 2004.
- [10] Y. Fan, D. Shen, R.C. Gur, R.E. Gur, and C. Davatzikos, "COMPARE: Classification Of Morphological Patterns using Adaptive Regional Elements", *IEEE Trans. on Medical Imaging*, to appear.
- [11] U.C. Yoon, J.S. Kim, J.S. Kim, I.Y. Kim and S.I. Kim, "Adaptable fuzzy c-means for improved classification as a preprocessing procedure of brain parcellation", *J. Digit. Imaging*, Vol 14, pp. 238–240, 2001.
- [12] J.M. Lee, U. Yoon, S.H. Nam, J.H. Kim, I.Y. Kim and S.I. Kim, "Evaluation of automated and semi-automated skull stripping algorithms using similarity index and segmentation error", *Comp. Biol. Med.*, Vol 33, pp. 495–507, 2003.
- [13] C. Fennema-Notestine, et al., "Quantitative evaluation of automated skull-stripping methods applied to contemporary and legacy images: Effects of diagnosis, bias correction, and slice location", *Human Brain Mapping*, Vol 27, pp. 99–113, 2006.
- [14] G.K. Rohde, A. Aldroubi and B.M. Dawant, "The Adaptive Bases Algorithm for intensity based nonrigid Image Registration", *IEEE Trans. Medical Imaging*, Vol 22, pp. 1470–1479, 2003.
- [15] D.L. Pham, "Robust Fuzzy Segmentation of Magnetic Resonance Images", *Proceedings of the Fourteenth IEEE Symposium on Computer-Based Medical Systems (CBMS2001)*, pp. 127–131, Bethesda, MD, 2001.
- [16] S.S. Bassett, D.M. Yousem, C. Crinziuzio, I. Kusevic, G. Yassa, B. Caffo, S. Zeger, "Familial risk for Alzheimer's disease alters fMRI activation patterns", *Brain*, Vol 129, pp. 1229–1239, 2006.