Multi-Atlas Skull-Stripping

Jimit Doshi, MSc, Guray Erus, PhD, Yangming Ou, PhD, Bilwaj Gaonkar, MSc, Christos Davatzikos, PhD

Rationale and Objectives: We present a new method for automatic brain extraction on structural magnetic resonance images, based on a multi-atlas registration framework.

Materials and Methods: Our method addresses fundamental challenges of multi-atlas approaches. To overcome the difficulties arising from the variability of imaging characteristics between studies, we propose a study-specific template selection strategy, by which we select a set of templates that best represent the anatomical variations within the data set. Against the difficulties of registering brain images with skull, we use a particularly adapted registration algorithm that is more robust to large variations between images, as it adaptively aligns different regions of the two images based not only on their similarity but also on the reliability of the matching between images. Finally, a spatially adaptive weighted voting strategy, which uses the ranking of Jacobian determinant values to measure the local similarity between the template and the target images, is applied for combining coregistered template masks.

Results: The method is validated on three different public data sets and obtained a higher accuracy than recent state-of-the-art brain extraction methods. Also, the proposed method is successfully applied on several recent imaging studies, each containing thousands of magnetic resonance images, thus reducing the manual correction time significantly.

Conclusions: The new method, available as a stand-alone software package for public use, provides a robust and accurate brain extraction tool applicable for both clinical use and large population studies.

Key Words: Brain extraction; registration; multi-atlas; label fusion; Jacobian determinant.

Brain extraction, or skull stripping is a very important preprocessing step preceding almost all automated brain magnetic resonance (MR) imaging (MRI) applications. It consists of the removal of the skull and the extracerebral tissues (e.g., scalp and dura) on brain MR images. An illustrative example of extraction on a T1-weighted image is shown in Figure 1. Brain extraction is known to be a difficult task, as the boundaries between brain and nonbrain tissues, especially those between the gray matter and the dura matter, might not be clear on MR images. Also, it is more prone to requiring manual intervention as the errors in this step propagate to most subsequent analysis steps, such as registration to a common space, tissue segmentation, cortical thickness and atrophy estimation, etc (1).

Particularly in large population studies, it is desirable to set up a fully automated processing pipeline with no or minimal manual intervention, to reduce processing time and to prevent any kind of human bias in the results. Thus, an automated brain extraction method should be accurate as well as robust, so that thousands of images within a study, potentially having different imaging characteristics and significant anatomical variations, could be successfully segmented.

Several brain extraction methods have been proposed since the early years of MRI research. Region- or boundary-based approaches (2–7) are simple, fast, and general. An example is the Brain Extraction Tool (BET) (5), a widely popular publicly available method, which is based on a deformable surface model that evolves to the boundaries of the brain. However, boundary-based methods might fail especially when the initial assumption of a clear separation between the brain and the nonbrain tissues is not completely satisfied.

Atlas-based (i.e., template-based) methods use an expert-defined segmentation on the atlas space as a prior for extracting the brain on the target image. In recent years, several multi-atlas–based methods have been proposed, producing very accurate, state-of-the-art segmentations (8–11). The main premise is that multiple atlases cover much wider anatomical variations and, when registered to the target image, they can correct errors among each other, thus providing an increased accuracy and robustness. The multi-atlas framework generally consists of three main modules: atlas selection based on similarity to the target scan, registration of the atlases to the target space, and fusion of the registered ground truth masks using various label fusion algorithms (11). In Carass et al (8), a probabilistic brain mask obtained using multi-atlas registration is combined with a fuzzy segmentation of the subject brain using topologically constrained morphological operators. In Iglesias et al (10), a hybrid (discriminative/generative) learning-based approach, named Robust Brain Extraction (ROBEX), is proposed: The discriminative component consists of a binary

Acad Radiol 2013; 20:1566–1576

From the Section of Biomedical Image Analysis, Department of Radiology, 3600 Market St. Suite 380, University of Pennsylvania, Philadelphia, PA, USA. Received May 16, 2013; accepted September 6, 2013. Grants supporting the research: 5R01AG014971-11, Computational Neuroanatomy of Aging and AD via Pattern Analysis, and 5R01EB009234-04, Computer Analysis of Brain Vascular Lesions in MRI: Evaluating Longitudinal Change. **Address correspondence to:** G.E. e-mail: guray.erus@uphs.upenn.edu

http://dx.doi.org/10.1016/j.acra.2013.09.010

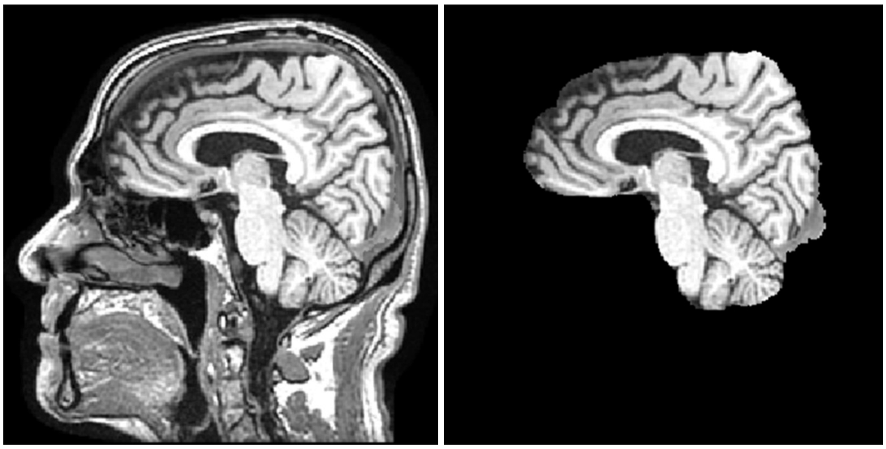

Figure 1. Brain extraction example.

classifier that is trained on aligned and intensity-normalized atlases with brain masks. The initial classification is refined using a generative model of the brain shape that is learned from the atlases by applying an active shape model approach. The method is validated on three publicly available data sets against six popular, publicly available methods [BET (5), Brain Surface Extractor (BSE) (3), FreeSurfer (12), 3dSkullStrip in AFNI package (13), BridgeBurner (14), and GraphCuts (GCUT) (15)] and obtained a higher accuracy for almost every method/data set combination.

Our approach follows a multi-atlas–based brain extraction strategy, because of its demonstrated strength and promise in the aforementioned studies. In addition, our approach makes contributions in dealing with two fundamental challenges in multi-atlas approaches:

The first challenge is the variability of imaging characteristics between studies. We proposed a study-specific template selection strategy. This is different from the approach used in recent multi-atlas–based methods, such as in Eskildsen et al (9) and in Leung et al (11), where final templates were selected from a predefined template library with manual ground truth masks. In our approach, a set of representative templates are automatically selected within the study, corresponding ground truth masks are created semiautomatically, and these templates are used for brain extraction on all images in the study.

The second challenge is the quality of individual registrations, particularly because the registration is performed on raw images (i.e., unprocessed images that contain both brain and the skull). Extracerebral tissues and organs have an intersubject variation much higher than that of the brain, which may misguide the registration and cause significant errors. In this report, we propose to use a registration algorithm that is particularly adapted to this task, as it adaptively aligns different regions of the two images based on not only their similarity but also the reliability of the matching between images, thus assigning lower importance to areas with missing correspondences and reducing the negative impact of outlier regions.

Several studies (16–18) have shown that a weighted voting (WV) strategy, which assigns higher weights to templates that are more similar to the target image, improves the segmentation quality. Following this approach, we applied a spatially adaptive WV strategy for combining coregistered template masks. However, instead of the commonly used intensity-based similarity metric, we used ranking of Jacobian determinant values to calculate local weights for each template. The underlying intuition is that a low volumetric change in registration between the template and the target images, as encoded by the Jacobian determinant value at each voxel, indicates higher local similarity between the two images.

Our method, which we call Multi-Atlas Skull-Stripping (MASS), is applied on three popular yet challenging public data sets, which have been widely used for validation in many recent brain extraction methods. The segmentations are compared to manual ground truth masks using Dice score and Hausdorff distance. The quantitative results are compared to the results reported by Iglesias et al (10).

METHODS

A general overview of the proposed method is given in Figure 2. Our framework consists of three components: template selection, registration, and label fusion. Each of these components is presented in the following subsections.

Template Selection

The quality of a registration is directly related to the similarity between the template and the target images. Due to either differences between populations (e.g., age, disease, etc) or changes in scanner type, technology, and protocol (e.g., 1.5 T to 3 T), images from two different studies might be significantly different. To increase the template–subject similarity, and hence to improve the registration accuracy, we select within-study templates, instead of the commonly used strategy of selecting templates from a predefined external template library. To limit the work required for the preparation of the ground truth brain masks for the selected