Title
Skull-stripping magnetic resonance brain images using a model-based level set

Journal
Neuroimage, 32(1)

ISSN
1053-8119

Authors
Zhuang, A H
Valentino, Daniel J
Toga, A W

Publication Date
2006-08-01

Peer reviewed

Author version for the University of California Digital Library
A.H. Zhuang, D.J. Valentino, and A.W. Toga / NeuroImage 32 (2006) 79-92

Skull-stripping Magnetic Resonance Brain Images Using a Model-based Level Set

Audrey H. Zhuang, [*] Daniel J. Valentino, [*†‡] and Arthur W. Toga[*]

[*]Laboratory of Neuro Imaging, Department of Neurology,
[†]Division of Interventional Neuroradiology, Department of Radiological Sciences
University of California at Los Angeles, CA 90095, USA

Abstract

The segmentation of brain tissue from nonbrain tissue in magnetic resonance (MR) images, commonly referred to as skull-stripping, is an important image processing step in many neuroimaging studies. A new mathematical algorithm, a model-based level set (MLS), was developed for controlling the evolution of the zero level curve that is implicitly embedded in the level set function. The evolution of the curve was controlled using two terms in the level set equation, whose values represented the forces that determined the speed of the evolving curve. The first force was derived from the mean curvature of the curve, and the second was designed to model the intensity characteristics of the cortex in MR images. The combination of these forces in a level set framework pushed or pulled the curve toward the brain surface. Quantitative evaluation of the MLS algorithm was performed by comparing the results of the MLS algorithm to those obtained using expert segmentation in 29 sets of pediatric brain MR images and 20 sets of young adult MR images. Another 48 sets of elderly adult MR images were used for qualitatively evaluating the algorithm. The MLS algorithm was also compared to two existing methods, the brain extraction tool (BET) and the brain surface extractor (BSE), using the data from the Internet brain segmentation repository (IBSR). The MLS algorithm provides robust skull-stripping results, making it a promising tool for use in large, multi-institutional, population-based neuroimaging studies.

I. INTRODUCTION

Intracranial segmentation, commonly referred to as skull stripping, aims to segment the brain tissue (cortex and cerebellum) from the skull and nonbrain intracranial tissues in magnetic resonance (MR) images of the brain. Skull stripping is an important preprocessing step in neuroimaging analyses because brain images must typically be skull stripped before other processing algorithms such as registration, tissue classification or bias field correction can be applied (Woods *et al.,* 1998, 1999; Van Horn et al., 1998; Shattuck *et al.,* 2001; Strother et al., 2004). In practice, skull stripping is widely used in

[‡] Corresponding author. Mail Stop 172115, Department of Radiological Sciences, UCLA, Los Angeles, CA 90095-1721, USA. E-mail: dvalentino@mednet.ucla.edu. Fax: (310)206-2967, Phone: (310)794-7131.

neuroimaging analyses such as multi-modality image fusion and intersubject image comparisons (Woods *et al.,* 1998, 1999; Toga, 1999); examination of the progression of brain disorders such as Alzheimer's Disease (Rusinek *et al.,* 1991; Thompson *et al.,* 2001), multiple sclerosis (Bermel *et al.,* 2003; Horsfield *et al.,* 2003; Zivadinov *et al.,* 2004; Sharma *et al.,* 2004) and schizophrenia (Narr *et al.,* 2004; Tanskanen *et al.,* 2004); monitoring the development or aging of the brain (Jernigan *et al.,* 2001; Blanton *et al.,* 2004); and creating probabilistic atlases from large groups of subjects (Mazziotta *et al.,* 2001).

Furthermore, larger numbers of subjects are needed to detect subtle differences between groups, and contemporary neuroimaging studies often achieve larger subject populations by pooling data from multiple institutions. However, different institutions use different scanners with different scan protocols and may utilize different imaging sequences (Vannier and Summers, 2003; Van Horn et al., 2004; Neu et al., 2005). The trend toward ever larger analyses in neuroimaging studies will continue because they are needed in order to obtain the statistical power required to detect subtle differences in brain structure and function. Thus, contemporary analyses require skull stripping algorithms that provide robust performance across different age groups, diseases, and image sequences, using images that are generated at multiple sites using different protocols.

A number of techniques have been proposed to automate the segmentation of brain structures in MR images (Kikinis et al., 1992; van Ginneken et al., 2002; Grau et al., 2004). Subsequent work focused on developing automated algorithms specifically for skull stripping brain MR images. A recent study published by Fennema-Notestine *et al.* (2006) compared the performance of the most commonly used skull stripping algorithms. This study concluded that existing algorithms had both strengths and weaknesses, but that no single algorithm was robust enough for large-scale analyses. This finding was in agreement with other recently published studies (Boesen *et al.,* 2004; Rehm *et al.,* 2004). As the number of subjects and participating institutions in a study increases, the need for more robust skull-stripping algorithms will become critical. In this paper, we examine new techniques for increasing the robustness of skull-stripping algorithms.

Skull stripping methods can generally be categorized into three types: intensity based, morphology based, and deformable model based. Intensity-based methods rely upon modeling the intensity distribution used for threshold classification. DeCarli *et al.* (1992) proposed a semi-automated classification method for brain tissue classification in brain MR images. This method used intensity distribution functions to identify major brain tissues (e.g., CSF, GM and WM). Each brain tissue was modeled using a modified log normal distribution function. One limitation of intensity-based methods is that they are frequently sensitive to intensity bias caused by magnetic field inhomogeneities, sequence variations, scanner drift, or random noise.

Morphology-based methods frequently combine connectivity-based morphological operations and thresholding or edge detection to extract image features and identify brain surfaces. Lee *et al.* (1998) proposed a 2D skull-stripping method applied to a midsagittal

slice, which was later extended by Huh *et al.* (2002) to all slices in a sagittal series. First, thresholds were used to separate dark pixels (e.g., background, skull and cavities, etc) from bright pixels (e.g., brain, skin, facial tissues, etc), then brain regions were identified using a connectivity-based algorithm. Shattuck *et al.* (2001) developed a tool called the brain surface extractor (BSE), which used a combination of edge detectors and morphological operators to skull strip the brain. A Marr Hildreth edge detector was first used to identify anatomic boundaries, then morphological operators were used to separate the tissues into component regions. Next, the largest central connected component was extracted as the brain region. Finally, nonbrain structures still attached to the brain region were removed. A potential disadvantage of these methods is that they are often dependent upon many parameters, and the parameters are often empirically generated and sensitive to small changes in the data.

Skull-stripping methods based upon deformable models typically evolve and deform an active contour to fit the brain surface, which is identified using selected image characteristics. Aboutanos *et al.* (1999) evolved a 2D contour by maximizing its corresponding 1D optimization problem, which was obtained via geometrical transformation from a 2D contour using dynamic programming techniques. The 1D optimization problem was described by a cost function that consisted of six terms including intensity value, morphology, gradient, moving speed of the contour, and smoothness of the contour. Zeng *et al.* (1999) proposed a system of two level set equations whose zero level curves represented their inner and outer boundaries of the gray matter of the cortex. Each level set equation was driven toward the inner or outer boundary by a force term determined by the intensity distribution of brain tissues (i.e., CSF, WM and GM). The two level set equations were further related to each other by constraining the distance between the inner and outer boundaries (i.e., the thickness of gray matter). Suri (2001) proposed an active contour algorithm that uses the level set methods to evolve the active contour. He used a fuzzy membership function to classify images into four components: WM, GM, CSF and background, then used a gradient detector and a deformable model to evolve an active contour to fit the surface between the CSF and GM. Baillard *et al.* (2001) registered brain data to an atlas and used the brain surface from the atlas as the initial contour. Then, an equation based on the level set method was used, in which the speed term was determined by the curvature of the evolving curve and by a sign function that signaled whether to include or exclude a pixel that the curve passed. Smith (2002) proposed an automated deformable model for skull-stripping, called the brain extraction tool (BET), in which a set of forces, including morphological and image-based forces, were applied in the tangential and normal directions of the evolving surface. In general, deformable models have the potential to produce more robust and accurate skull-stripping results than methods using edge detection and threshold classification (Duncan and Ayache, 2000). Hybrid schemes have also been proposed to combine multiple existing methods to compensate for problems encountered with individual methods (Rehm *et al.,* 2004; Rex *et al.,* 2004).

We developed a new, automatic algorithm called the model-based level set (MLS) method to remove the skull and intracranial tissues surrounding the brain in magnetic resonance (MR) images. The level set method was used to evolve an active curve defined

by the zero level set of the implicit function Φ. The velocity of the evolving curve was determined by the image data and the morphology of the brain surface. The image data function was derived from the intensity value and the contrast between brain and nonbrain tissues; the morphology term was derived from the mean curvature of the evolving curve.

Compared to other level set based models used for skull-stripping such as those of Zeng *et al.* (1999), Suri (2001), and Baillard *et al.* (2001), the MLS method is more automated and is locally adaptive to reduce sensitivity to a bias field. Compared to the deformable models proposed for skull-stripping by Dale *et al.* (1999) or Smith (2002), the MLS algorithm uses an active curve embedded in the higher dimensional function Φ, evolves the curve using an implicit method, and automatically deals with topological changes such as self-intersections. Lastly, the initialization algorithm used in the MLS is more robust across different data sets.

The MLS algorithm was tested using different imaging sequences, age groups, and diseases and was compared with two widely-used, automated skull-stripping methods. The remainder of this paper is organized as follows: the Methods section describes the implementation of the MLS algorithm; the Results and evaluation section summarizes the results, evaluation, and comparison of our method to other skull-stripping tools; and the Discussion section discusses the contribution of this research to solving the skull-stripping problem.

II. METHODS

A. *The Level Set Method*

The level set model (Osher and Sethian, 1988) implicitly represents a curve as a level set of the function $\phi(x, y): \Omega \to R$, where Ω is a subset of R^2. Thus, the k^{th} level curve C can be described as:

$$C = \{(x, y) \mid \phi(x, y) = k\} \tag{2.1}$$

The level set function, $\phi(x, y)$, is a signed distance function, that is, its value at (x, y) is equal to the distance from (x, y) to the nearest point (x', y') on the zero level set, and its sign is positive or negative if the point (x, y) is inside or outside the zero level curve.

A common approach to evolving the zero level curve is to let the function, $\phi(x, y)$, evolve with respect to time. A time parameter, t, is artificially added so that the function becomes

$$\Phi(t, x, y) = 0. \tag{2.2}$$

The original level set method described by Osher and Fedkiw (2003) used a partial differential equation to evolve the implicit function, $\phi(t, x, y)$:

$$\frac{\partial \phi(t,x,y)}{\partial t} + \nabla \phi(t,x,y) \cdot \frac{d\vec{x}}{dt} = 0 \tag{2.3}$$

Equation 2.3 was derived from Equation 2.2 by differentiating both sides with respect to time t and applying the chain rule. $\frac{d\vec{x}}{dt}$ is a speed term representing the movement of the point $\vec{x} = (x, y)$ over time, and ∇ is the gradient operator. In our numerical implementation, we chose the time step as $\Delta t = 0.1\Delta x$, where Δx is the spatial step size in the x direction.

B. Formulation of the Speed Function for Skull-Stripping

To solve the skull-stripping problem, we sought to produce an equation for evolving the zero level curve until it fit the boundary of the brain. The intensity values in each image were used to determine the boundary of the brain; however, the intensity values were often skewed by artifacts introduced by noise and inhomogeneities in the magnetic imaging system, so we added a curvature-based force to constrain the smoothness of the evolving curve. Thus the speed function is a combination of two forces: a brain surface attraction force, F_{img}, and a morphological smoothing force, F_{curv}, each applied in the normal direction of the iso-level set at point *(x, y)*:

$$F = a \cdot F_{img} \cdot \vec{N} + b \cdot F_{curv} \cdot \vec{N} \tag{2.4}$$

where a and b are scalars that weight each force separately. The values of a and b are initially set to the same value to equalize F_{img} and F_{curv}, then the value of b is dynamically adjusted to avoid leakage through weak boundaries (as discussed in the Methods section *K*).

C. The Formulation of F_{curv}

F_{curv} was determined by the mean curvature κ of the level set function:

$$\kappa = \nabla \cdot \vec{N} = \nabla \cdot \left(\frac{\nabla \phi}{|\nabla \phi|} \right) = \nabla \cdot \left(\frac{\phi_x}{|\nabla \phi|}, \frac{\phi_y}{|\nabla \phi|} \right) = \frac{\partial \left(\phi_x / |\nabla \phi| \right)}{\partial x} + \frac{\partial \left(\phi_y / |\nabla \phi| \right)}{\partial y} \tag{2.5}$$

The curvature-dependent term was first applied in computational fluid dynamics as an "artificial viscosity" term used to achieve stability when the central difference was used to approximate the gradient of Φ (Osher and Fedkiw, 2003). The brain has fairly smooth surfaces, so the curvature term was used to ensure that the evolving curve was similarly smooth.

D. The Formulation of F_{img}

We extended the brain surface detection algorithm that was originally introduced by Dale *et al.* (1999) and later simplified by Smith (2002). The image based force, F_{img}, was used to model the features of the brain surface as they appear in grayscale images. In T1-weighted MR images, the brain surface appears as a change in grey-scale intensity from relatively high intensity gray matter to relatively low intensity CSF. In this case, the image based force was described by Equation 2.6:

$$F_{img} = \frac{2(I_{\min} - h_l)}{I_{\max} - h_{2\%}} \tag{2.6}$$

where $h_{2\%}$ is the intensity threshold separating the lower 2% of the cumulative histogram; and h_l is the threshold separating the intensity of skull from the local maximum intensity $I_{\max}$, which thus divides the brain from non-brain tissue. The threshold, h_l, was computed as:

$$h_{l.} = (I_{\max} - h_{2\%}) \cdot T_h + h_{2\%} \tag{2.7}$$

where T_h is a constant equal to 0.5. F_{img} takes positive or negative values such that the effect of F_{img} is to "push" or "pull" the evolving curve toward the desired edge.

I_{max} and I_{min} are the local maximum and minimum intensities of pixels on a line that starts from the zero level curve and points inward to the brain in the normal direction. The model for computing I_{max} was modified in order to make the model more robust in the presence of inhomogeneities, as described in equations 2.8 and 2.9:

$$I_{min} = MAX(h_{2\%}, MIN(h_M, I(0), I(1), \ldots I(d_1))) \tag{2.8}$$
$$I_{max} = MIN(h_{98\%}, MAX(h_M, I(0), I(1), \ldots I(d_2))) \tag{2.9}$$

where $I(n)$ is the intensity of a pixel on the line n millimeters (mm) away from the origin; h_M is the median intensity calculated in the brain region within the initial zero level curve; and d_1 and d_2 are the lengths of the line used to find I_{min} and I_{max} respectively (Figure 1). I_{max} is the local maximum representing the intensity of the WM, and I_{min} is the local minimum intensity approximating the intensity of the CSF (or bone if CSF is thin) as the evolving curve approaches the brain surface.

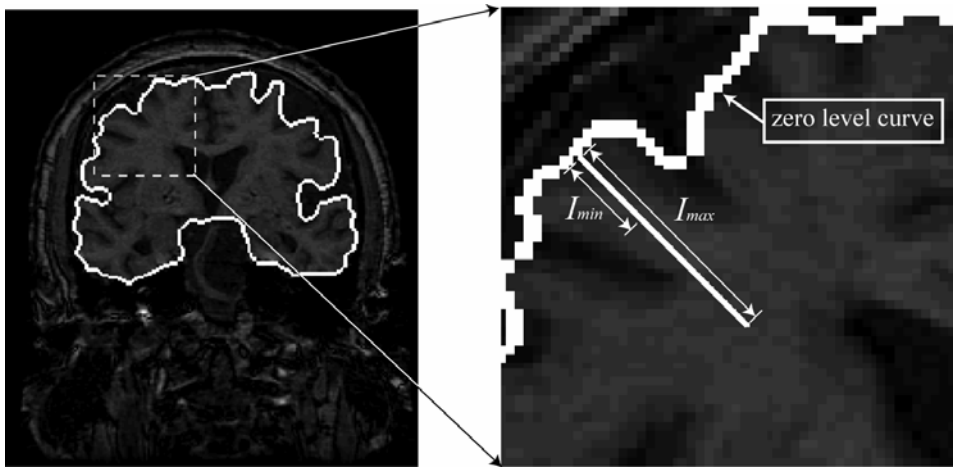


Figure 1 Heuristic search for I_{min} and I_{max}. (a) A representative coronal slice from an MR data set. (b) An enlarged region that shows the line profile normal to the zero level curve along which I_{max} and I_{min} are searched for.

In developing the image force term, F_{img}, for the MLS algorithm, a number of issues were considered. When searching in the normal direction of the zero level curve, the distance searched for I_{max} should be long enough to pass deep sulci and reach the WM (e.g., d_2 = 10mm); the distance searched for I_{min} should be long enough so that the sampling line reaches the CSF when the evolving contour passes the brain surface (e.g., d_1 = 0 ~ 20mm). Setting d_1 = 20mm often worked well with pediatric data or with data from young adults. However, d_1 = 20mm was too long to work well with elderly subjects' data because it over sampled the CSF in deep sulci. Figure 2a illustrates that the evolving contour with d_1 = 20mm stopped near the deep sulci before reaching the brain surface. When d_1 was reduced (e.g., to 6 mm), the evolving curve successfully stopped at the brain surface, as illustrated in Figure 2b.

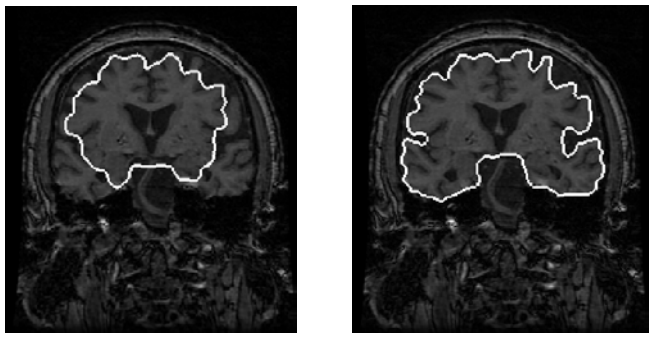


Figure 2 The results of skull stripping a coronal slice in a data set obtained from a 73-year-old subject whose images contain a large volume of CSF. (a) The evolving contour within the brain, using d_1 = 20mm. (b) The evolving contour at the brain surface, using d_1 = 6mm.

Therefore, d_1 should be adapted to the CSF volume of the brain. The CSF volume of the brain is hard to compute automatically but can be estimated by the subject's age. Courchesne *et al.* (2000) published a study reporting the changes in CSF volume in 116 volunteers aged 19 months to 80 years of age. They found that the volume of intracranial

CSF increased substantially and nearly linearly over time in healthy subjects. The data used in our study showed the same trend of increasing CSF and dilating sulci with brain aging. In the MLS algorithm, the parameter d_1 was adaptively changed from 20 to 0mm in proportion to age (from 1 year to 100 years old). In our experiments, the age of the subject was automatically obtained from the image metadata. If the metadata are incomplete, then the user can select the age group to which the subject belongs. If no age group is selected, then the algorithm defaults to the parameters for a normal adult.

E. Implementation of the Model-based Level Set Algorithm

The model-based level set skull-stripping equation is given in Equation 2.10:

$$\frac{\partial \phi}{\partial t} + \nabla \phi \cdot (a \cdot F_{img} + b \cdot F_{curv}) \cdot \vec{N} = 0 \tag{2.10}$$

Because $\vec{N} \cdot \nabla \phi = \frac{\nabla \phi}{|\nabla \phi|} \cdot \nabla \phi = \frac{|\nabla \phi|^2}{|\nabla \phi|} = |\nabla \phi|$ (2.11)

Equation 2.10 can be simplified as:

$$\frac{\partial \phi}{\partial t} + |\nabla \phi| \cdot (a \cdot F_{img} + b \cdot F_{curv}) = 0 \tag{2.12}$$

As the term $(a \cdot F_{img} + b \cdot F_{curv})$ approaches zero, the level set function goes to its steady state.

F. Replacing $|\nabla \phi|$ With $\delta(\phi)$

For the $|\nabla \phi|$ term in equation (2.12), Chan and Vese found that a nonmorphological approach was more effective than using $|\nabla \phi|$ so they replaced it with a delta function $\delta(\phi)$(Chan and Vese, 2001; Vese and Chan, 2002), which was defined as the derivative of the one-dimensional Heaviside function, described in Equations 2.13 and 2.14 respectively,

$$\delta(\phi) = H'(\phi) \tag{2.13}$$

$$H(\phi) = \begin{cases} 0 & if \phi \le 0, \\ 1 & if \phi > 0. \end{cases} \tag{2.14}$$

G. Narrow Band Extension

When the level set function Φ is discretized on an image grid, the zero level curve lies more often between grid points than on grid points. Thus, the zero level curve cannot be found by searching for grid points where Φ = 0. Instead, the first-order smeared-out approximations of *H(Φ)* and $\delta(\phi)$ are used to construct a narrow band around the zero

level curve, which prunes out other level sets. The smeared-out Heaviside function is written as:

$$H_\varepsilon(\phi)=\begin{cases}0 & if\phi<-\varepsilon,\\ \frac{1}{2}\left(1+\frac{\phi}{\varepsilon}+\frac{1}{\pi}\sin\left(\frac{\pi\phi}{\varepsilon}\right)\right) & if|\phi|\le\varepsilon,\\ 1 & if\phi>\varepsilon,\end{cases} \tag{2.15}$$

where ε determines the size of the band width of the smeared-out band. We chose $\varepsilon = 1.5\Delta x$ as recommended in (Osher and Fedkiw, 2003), and thus the bandwidth was 3 grid cells. Then the delta function $\delta(\phi)$ is defined as:

$$\delta_\varepsilon(\phi)=\begin{cases}0 & if\phi<-\varepsilon,\\ \frac{1}{2}\left(\frac{1}{\varepsilon}+\frac{1}{\varepsilon}\cos\left(\frac{\pi\phi}{\varepsilon}\right)\right) & if|\phi|\le\varepsilon,\\ 0 & if\phi>\varepsilon,\end{cases} \tag{2.16}$$

The velocity term is multiplied by $\delta_\varepsilon(\phi)$, which ensures that only level sets within the band are updated and that level sets outside the band remain stationary. This makes computation more efficient than updating all of the level sets on the image grid. However, this distorts the signed distance map of the level set function, which can be reinitialized as discussed in the next section (Sussman *et al.,* 1994).

H. Reinitialization

The level set function, Φ, was initialized as a signed distance function for computational efficiency (Osher and Fedkiw, 2003). As mentioned previously, the narrow band extension only evolved the level sets within the band and left other level sets unchanged, which caused the function, Φ, to no longer represent the distance to the zero level set. So the level set function, Φ, was reinitialized periodically to approximate the signed distance. Sussman *et al.* (1994) proposed a fast method to rectify Φ which is given by:

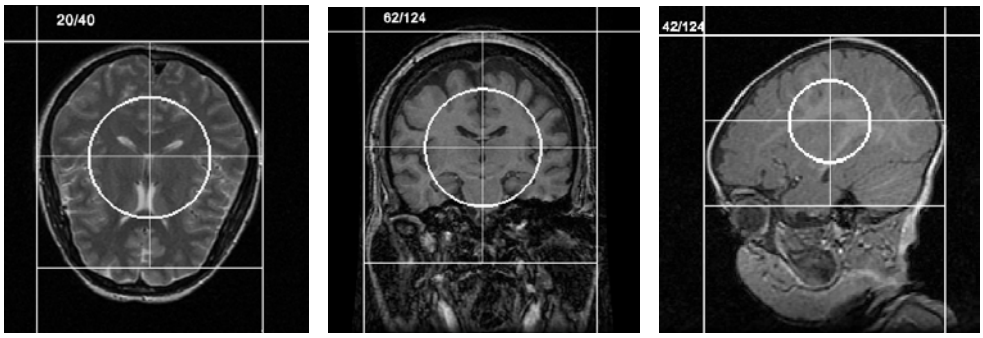
$$\phi_t = S_\varepsilon(\phi_0')(1-|\nabla\phi|) = S_\varepsilon(\phi_0')(1-\sqrt{\phi_x^2+\phi_y^2}) \tag{2.17}$$

$$S_\varepsilon(\phi_0') = \frac{\phi_0'}{\sqrt{\phi_0'^2+\varepsilon^2}} \tag{2.18}$$

where ϕ_0' is the level set function whose value drifted away from the signed distance. Sussman et al.'s method rectified ϕ_0' to a distance function while keeping its zero level curve. Our experiments showed that this method was fast; only one iteration was typically required for Φ to converge to a signed distance function.

I. Initialization

Automatic initialization (of the zero level curve) is required for fully automated skull-stripping and can affect the accuracy of the segmentation results. Some nonbrain tissues (i.e., skin and muscle tissue) have intensity characteristics similar to brain tissues and could produce errors in segmentation if the initial curve included them. Therefore, our goal was to initialize the zero level curve at the center of the brain, with a diameter just small enough so that the initial curve was completely within the brain surface. Existing initialization methods are frequently inaccurate or computationally expensive to implement. For example, Smith (2002) used the center of the head volume (which contains both brain and nonbrain regions) and the half radius of the head volume to initialize a sphere. The sphere often includes nonbrain tissue that is outside the brain surface, particularly in studies that include substantial muscle tissue. Baillard *et al.* (2001) used registration to initialize the curve, which is potentially very accurate, but is computationally expensive, particularly in large-scale analyses.



(a) (b) (c)

Figure 3 The initial zero level set curves used on axial (a), coronal (b) and sagittal (c) slices.

We sought a solution that would accurately initialize the zero level curve, be robust across data sets, and be computationally efficient. Our approach is fast, accurate, and independent of the data; however, it does incorporate some steps that are based upon empirical observations. We assume that the brain is head-up, so we provide a rotation function to enable a user to rotate the brain into the head-up position. The algorithm starts with the middle slice of the volume if the slice is in axial or coronal orientation, or with the slice located one third into the volume if it is in sagittal orientation. A rough mask of the head is obtained using a threshold that separates the lower 1/4 histogram between $h_{2\%}$ and $h_{98\%}$, where $h_{98\%}$ is the intensity threshold separating the lower 98% of the cumulative histogram and $h_{2\%}$ separating the lower 2%. The threshold was chosen empirically to separate the background and the skin or muscle tissue that covers the head, then the edges of the head, that is, the top, left, and right sides of the head, are automatically found (as illustrated by the orthogonal lines in Figure 3). The shape of the brain in axial or coronal orientation was approximated as a square, and the shape of the brain in sagittal orientation was approximated as a rectangle with a 3 : 2 ratio (as illustrated in Figure 3). Then the initial zero level set is positioned at the center of the square or rectangle, (x_c, y_c), with its radius, r, equal to half of the length of the square or width of the rectangle, which is described as:

$$\phi_0(x,y) = -\sqrt{(x-x_c)^2 + (y-y_c)^2} + r \tag{2.19}$$

where $\Phi_0(x, y)$ denotes $\Phi(t, x, y)$ at $t = 0$.

J. Segmentation of Subsequent Slices

Given the continuity of the brain surface, we used the resultant zero level curve from the current slice to initialize the contour in adjacent slices. This can save computation time and improve the accuracy of results. The zero level curve on the current slice was contracted inward before initializing the next slice in order to ensure that the new initial curve was located inside the brain (Figure 4). The distance contracted was proportional to the slice spacing, which can be found in the image metadata. For example, if the inter-slice distance was 2 mm, the resultant curve was shrunk inward 2 mm and used to initialize the following slice, and so on, until each 2D slice had been processed.

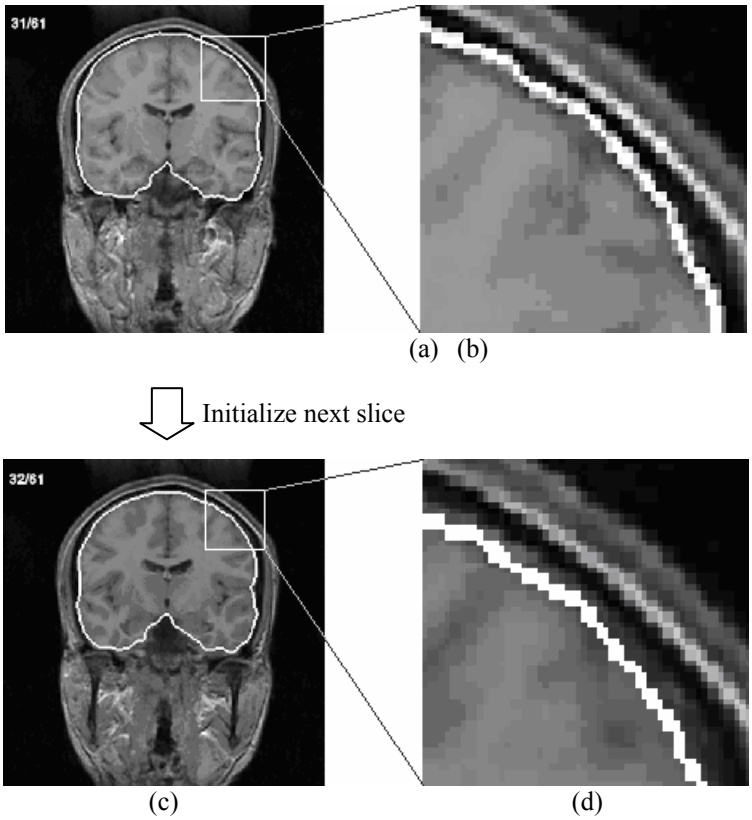


Figure 4 illustrates subsequent initialization. (a) shows the segmentation results of slice 31 (the total number of slices is 61). (b) shows an enlarged tile drawn in (a). (c) shows the initialization on slice 32, which is obtained by shrinking the segmentation results of slice 31 inward 3.1mm (3.1mm is equal to the interslice distance). (d) shows an enlarged tile drawn in (c).

K. Adjustment of F_{curv} for weak boundaries

In MR brain images, there are often areas where brain tissue and other tissues having similar intensities to brain tissue are adjacent to each other, resulting in a weak gradient

between the tissues. For example, in areas where there is a lack of CSF between the brain tissue and the muscle tissue surrounding the cerebellum, then the evolving curve tends to move past the desired brain/muscle boundary and evolve into the muscle tissue as illustrated in Figure 5. When the evolving curve progresses through a weak boundary, it usually forms a relatively narrow neck that has high curvature. If the high curvature was prohibited, then the formation of a narrow neck and the corresponding leakage though the weak boundary could be prevented. The MLS algorithm first checks the evolving curve on a slice for leakage, then increases the weight of F_{curv} to prevent high curvature, and then segments the same slice again.

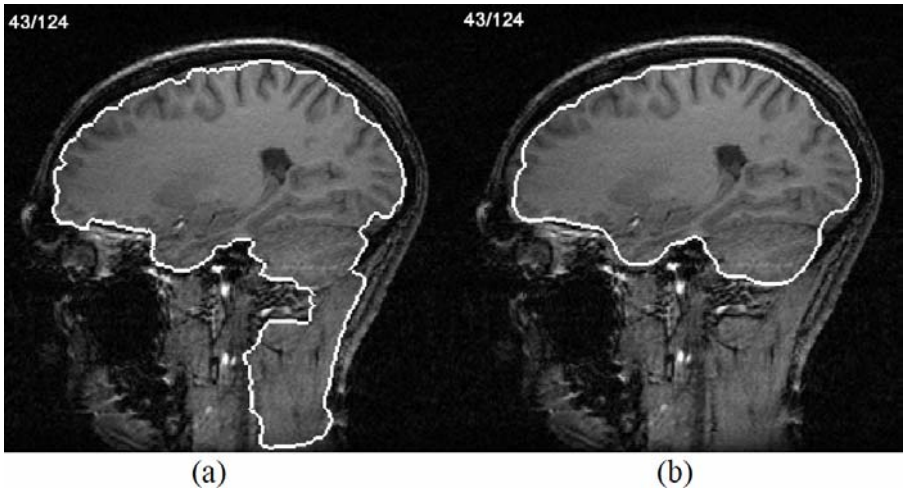


Figure 5 Correction of the leakage of the evolving contour through weak boundaries. (a) Leakage of the contour in a region with weak boundaries. (b) Shows the contour after increasing the curvature force by 5 times the original curvature force, which effectively eliminates the leakage.

Leakage through a weak boundary was detected by calculating the Jaccard (1912) coefficient between a new slice and its predecessor slice to determine their similarity. If the segmentation results for a new slice were similar to its predecessor (Jaccard coefficient ≥ 0.85), then the segmentation for that slice was accepted. However, if the segmentation results differed (Jaccard coefficient < 0.85), then the segmentation was rejected, the mean curvature force F_{curv} was increased by a factor of 5, and the segmentation was repeated. If the segmentation was rejected again, then the original force was increased by a factor of 10. If the segmentation failed a third time, then heuristics were used to improve the segmentation result. For example, the segmentation result form the previous slice was used to generate the segmentation for the current slice by uniformly expanding or shrinking the curve from the previous slice. If the current slice was closer to the middle plane than the previous slice, then the curve was expanded, otherwise it was contracted. The distance that the curve was expanded or shrunk depended upon the spacing between the slices; for example, if the interslice spacing was 2mm, then the curve was expanded or shrunk by 2mm.

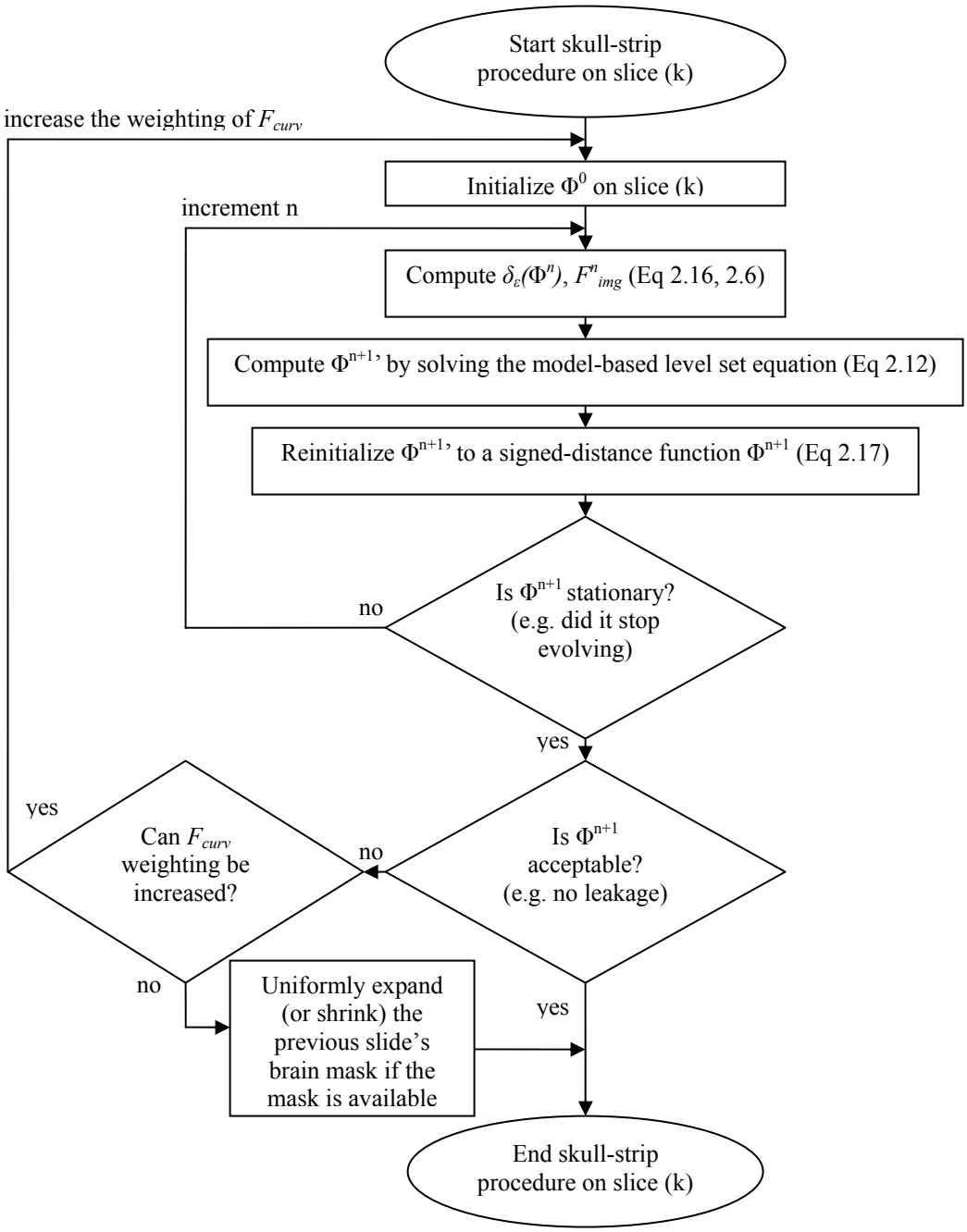


Figure 6 Flowchart of the model-based level set (MLS) algorithm for skull-stripping the kth slice of a volumetric data set.

L. Implementation

The model-based level set (MLS) algorithm illustrated in Figures 6 and 7 was implemented using the Java programming language to facilitate portability and sharing with other researchers. No effort was made to optimize the numerical techniques or the Java code. However, the code is well documented and available so that other researchers could evaluate and modify it. The MLS algorithm was implemented in two versions: an interactive version with a graphical user interface for exploring new data and a

noninteractive version for batch processing. The noninteractive version is fully automated and available as a LONI Pipeline module (Rex *et al.*, 2003). The use of the LONI Pipeline facilitates the use of the MLS algorithm with other algorithms, for example, registration modules, shape analysis modules, etc., to form automated pipelines that are suitable for sophisticated neuroimaging analyses.

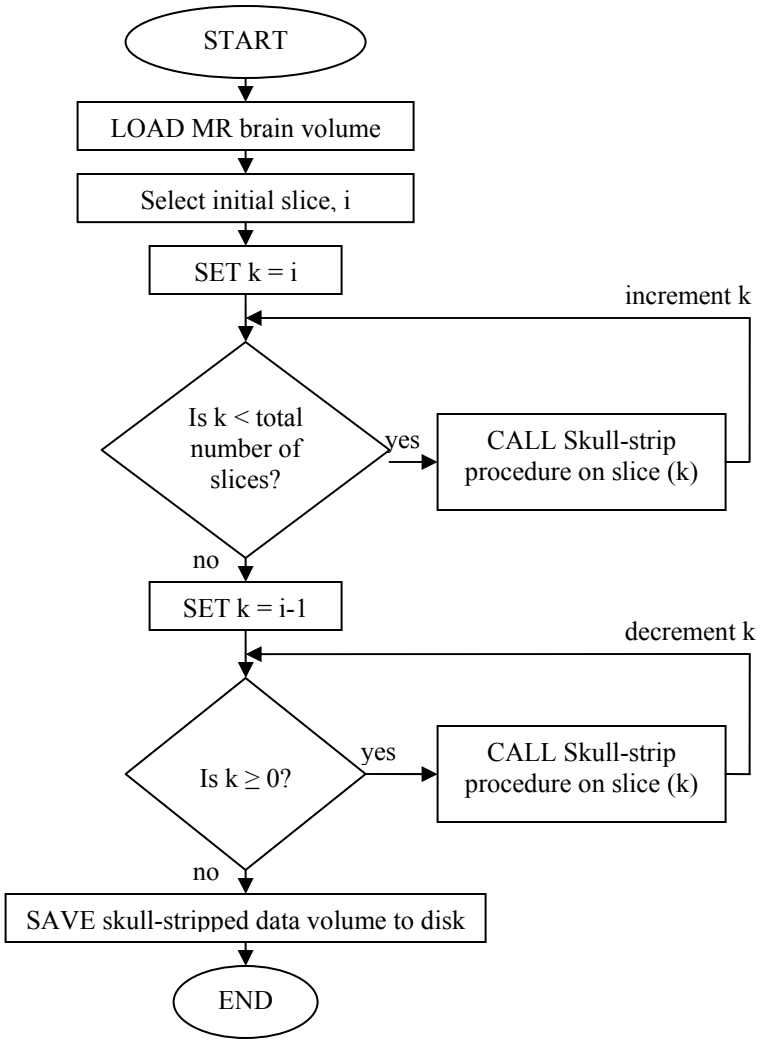


Figure 7 Flowchart of the MLS skull-stripping procedure.

Existing image file reader and writer plug-ins provided with the LONI Debabeler (Neu *et al.,* 2005) were used to decode and encode image file formats, to translate between file formats, and to anonymize and correct metadata in the header fields. The plug-ins support all of the common medical imaging file formats, including ANALYZE[1] (Robb *et al.,* 1989), NIFTI[2], Minc[3], and DICOM[4] (Bidgood and Horii, 1992; Bidgood *et al.,* 1997).

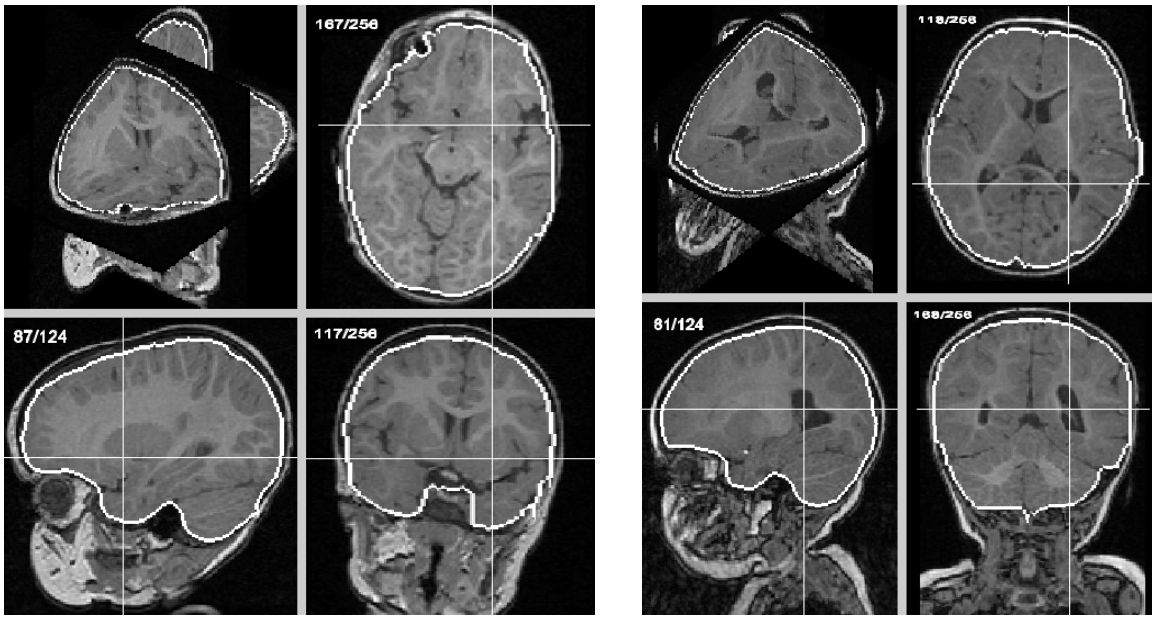
[1] http://www.mayo.edu/bir/PDF/ANALYZE75.pdf
[2] http://nifti.nimh.nih.gov/
[3] http://www.bic.mni.mcgill.ca/software/minc/minc.html
[4] http://medical.nema.org/dicom/2003.html

III. RESULTS AND EVALUATION

A. Comparison of methods

(a) Accuracy of pediatric skull-stripping

We evaluated the accuracy of our algorithm using 29 pediatric T1 MR brain images. Our goal was to identify the interface between the CSF and GM. The experiments were performed on a 1.8 GHz Pentium IV running Windows XP with 512 MB of RAM. Each volume consisted of 124 slices, with 256 × 256 pixels per slice. Each slice had a resolution of 0.98 × 0.98 mm^2 and the interslice spacing varied from 1.2 mm to 1.5 mm. The average elapsed time ("wall clock" time) was 5.3 s/slice. Figure 8 shows the final results on two sample volumes, displayed in three orientations.



(a) (b)

Figure 8 Results of skull-stripping representative pediatric data. The dimensions of the data volume in panels a and b are 256 x 256 x 124. The contours in white represent the zero level curve of the function Φ. The left-upper corner in each figure shows a 3D projection of the axial, coronal, and sagittal planes.

All of the neuroimaging data used in this study was obtained from multiple, collaborating institutions using procedures approved by the institutional review boards of all institutions. Informed consent was obtained from all subjects. The data was anonymized using the LONI Debabeler (Neu *et al.,* 2005) and stored in the Laboratory of Neuro Imaging (LONI) at UCLA. A subset of the data was manually segmented by an expert neuroanatomist to generate a "gold standard" data set that was used to measure the accuracy of the MLS algorithm.

The results of the MLS algorithm were compared to the 29 pediatric data sets that were manually skull-stripped by an expert. We computed the sensitivity and specificity coefficients of our segmentation results using the manually segmented brain masks. The

sensitivity is the percentage of brain voxels recognized by the algorithm (Equation 3.1). The specificity is the percentage of nonbrain voxels recognized by the algorithm (Equation 3.2).

$$Sensitivity = \frac{TP}{TP + FN} \tag{3.1}$$

$$Specificity = \frac{TN}{TN + FP} \tag{3.2}$$

where *TP* and *FP* stand for *true positive* and *false positive*, which were defined as the number of voxels correctly and incorrectly classified as brain tissue by the automated algorithm. *TN* and *FN* stand for *true negative* and *false negative*, which were defined as the number of voxels correctly and incorrectly classified as nonbrain tissue by the automated algorithm.

We also computed the Jaccard similarity (Jaccard, 1912) between our results and the manually stripped masks. The Jaccard similarity measures the intersection of the two sets divided by their union.

$$J(S_1, S_2) = \frac{|S_1 \bigcap S_2|}{|S_1 \bigcup S_2|} = \frac{TP}{TP + FP + FN} \tag{3.3}$$

where S_1 and S_2 refer to two data sets separately.

We also computed the Dice (Dice, 1945) coefficient to show the similarity level of our segmentation to manual segmentation. The Dice coefficient is defined as

$$\kappa(S_1, S_2) = \frac{2|S_1 \bigcap S_2|}{|S_1| + |S_2|} = \frac{2TP}{2TP + FP + FN} \tag{3.4}$$

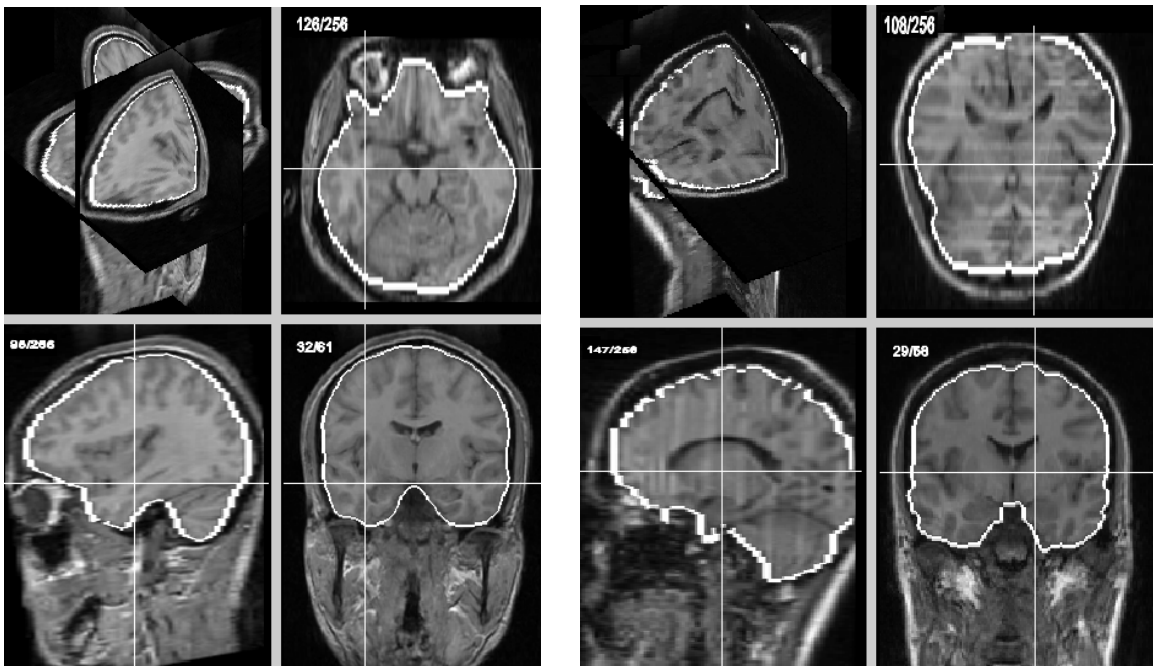
The {FP} rate is the number of voxels incorrectly classified as brain tissue by the automated algorithm divided by manually segmented brain masks (Table 1).

$$FP\_Rate = \frac{FP}{TP + FN} \tag{3.5}$$

The average sensitivity of our algorithm and its standard deviation on the 29 pediatric data sets is 0.98 ± 0.01; the specificity is 0.997 ± 0.001. The Jaccard coefficient is 0.95 ± 0.01, and the Dice coefficient is 0.97 ± 0.01.

(b) Accuracy of Skull-stripping Normal Adult Data

We also tested the MLS algorithm on 20 normal MRI brain data sets and the corresponding manual segmentations from the Internet brain segmentation repository[5] (IBSR) developed by the Center for Morphometric Analysis (CMA) at Massachusetts General Hospital. The IBSR data sets were used to enable objective comparison between skull-stripping algorithms. Each volume had approximately 60 slices, with 256 × 256 pixels per slice. The slice resolution varied from 1.02 × 1.04 mm^2 to 1.26 × 1.51 mm^2, and the interslice spacing varied from 3.05 mm to 3.37 mm. Figure 9 shows the final results of segmenting two sample volumes.



(a) (b)

Figure 9 Results of skull-stripping representative normal adult data from the IBSR. The dimensions of the data volume in panel a are 256 x 256 x 61 and in panel b are is 256 x 256 x 58. The contours in white represent the zero level curves of the function Φ. Each volume is displayed in four views: a 3D orthogonal projection, and the axial, sagittal, and coronal planes.

Table 1 Summary of multiple evaluation indexes (their mean and standard deviation) on 20 normal data sets from the CMA (the best average results for each index is highlighted). Overall, MLS performed better than BSE and BET and was more consistent in generating accurate results for different subjects.

	Jaccard	Dice	FP Rate	Sensitivity	Specificity
BET	0.55 ± 0.19	0.69 ± 0.15	1.01 ± 0.62	**0.999** ± 0.001	0.90 ± 0.05
BSE	0.83 ± 0.22	0.88 ± 0.22	0.13 ± 0.23	0.92 ± 0.22	0.99 ± 0.02
MLS	**0.93** ± 0.03	**0.96** ± 0.02	**0.06** ± 0.03	0.98 ± 0.03	**0.99** ± 0.01

By computing the Jaccard index, Dice index, sensitivity, specificity, and the FP rate (Table 1), we compared the MLS segmentation results to the manual segmentation results provided by the IBSR.

[5] http://www.cma.mgh.harvard.edu/ibsr/

We also compared the MLS algorithm to widely used skull-stripping tools, including the brain surface extractor (BSE) and the brain extraction tool (BET), using the 20 normal data sets from the IBSR. BSE was run with its default parameters, which were (3, 5.0) for the Anisotropic Filter and 0.75 for the edge detector kernal. BET was also run with its default parameters, which were 0 for the vertical gradient and 0.5 for the fractional intensity threshold. The Jaccard coefficients were computed and graphed in Figure 10, which shows the results obtained using BSE, BET and MLS on each data set. The higher the Jaccard coefficient, the more accurate the segmentation results. Figure 10 indicates that the Jaccard coefficients for MLS were higher than BSE on 17 data sets and were higher than BET on all data sets. It also shows that MLS provided more consistent skull-stripping performance than BSE and BET; the standard deviations of Jaccard coefficient were 0.03, 0.22 and 0.19 for MLS, BSE and BET respectively. Table 1 summarizes multiple evaluation indexes computed for BSE, BET and MLS.

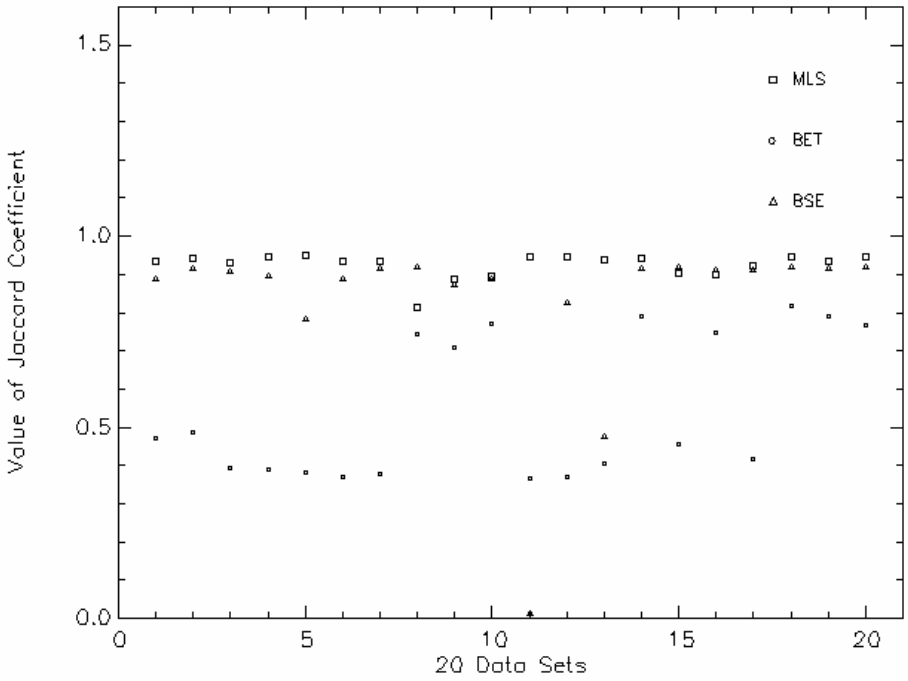


Figure 10 The Jaccard coefficient (also called similarity coefficient) obtained using the MLS, BET, and BSE algorithms on each of the 20 data sets from the IBSR. The MLS performed better than BET and BSE.

B. Sample Results of Different Sequences, Age groups and Diseases

The MLS algorithm was further tested on data obtained using different age and disease groups. Quantitative metrics (described in the Level set method section) were used to evaluate the results of using the MLS algorithm on T1-weighted MR images from pediatric and normal adult subjects. A qualitative metric (ranking by expert anatomists) was also used to subjectively evaluate the results of using the MLS algorithm on T1-weighted and T2-weighted MR images in normal elderly subjects. The quantitative and qualitative results indicated that the MLS algorithms accurately skull-stripped all of these data sets.

IV. DISCUSSION

The MLS algorithm extends the level set method introduced by Osher and Sethian (1988) to achieve robust skull-stripping of brain MR images. The level set method was used for several reasons, including its ability to handle complex topologies, incorporate image-based constraints, easily compute geometrical properties such as the mean curvature, and be easily visualized using isosurface rendering techniques. A model of the intensity distribution and curvature of the brain surface was incorporated into the level set equation to improve the robustness and accuracy of skull-stripping.

Theoretically, the MLS should be insensitive to the bias field introduced by inhomogeneities of the radio frequency (RF) coils, gradient coils, magnetic field (B_0) (Simmons *et al.,* 1994) or random noise. The bias field is a multiplicative noise that affects high-intensity pixels more than low-intensity pixels (Wells *et al.,* 1996; Leemput *et al.,* 1999). The bias field is also locally smooth, such that the contrast of anatomical structures is not significantly affected. For example, in T1-weighted images, the white matter (WM) has higher intensity value than the gray matter (GM) with or without the presence of the bias field. The MLS algorithm is only dependent upon differences in the local intensity contrast of brain structures and upon the curvature of the brain surface; thus, it is theoretically insensitive to the bias field. This was confirmed by experiments on the IBSR data, some of which contain perceivable bias.

The MLS algorithm was also insensitive to random noise scattered in the image because the curvature term limited artificially sharp changes in the evolving curve that might be caused by random noise. All of the results reported in this paper were obtained using the original, noisy MR data; no processing was used to reduce or remove the noise.

Another advantage of the MLS algorithm is that it required only one parameter, which was used to weight the curvature term. This parameter was easy to adjust because it influenced the segmentation results in a predictable way, for example, as the curvature term was set higher, the segmented surface was smoother. The default parameter value of 0.05 was found to work well for all of the data reported in this study.

Several skull-stripping algorithms have been proposed in the literature (Lee *et al.,* 2000; Huh *et al.,* 2000; Shattuck *et al.,* 2001; Aboutanos *et al.,* 1999; Zeng *et al.,* 1999; Dale *et al.,* 1999; Suri, 2001; Smith, 2002; Baillard *et al.,* 2001), and a number of recent studies have compared the accuracy and performance characteristics of these techniques (Fennema-Notestime *et al.,* 2006; Boesen *et al.,* 2004; Rehm *et al.,* 2004). Many of the existing techniques can be tuned to skull-strip a given data set, but none of these techniques can automatically skull-strip the variety of data sets contained in the large-scale studies that increasingly characterize contemporary neuroimage analyses. More recently, hybrid techniques have attempted to combine skull-stripping algorithms to achieve more robust skull-stripping results (Rex, *et al.,* 2004; Rehm *et al.,* 2004). Unfortunately, hybrid techniques often require extensive training to teach the program to accurately skull-strip the data. Thus, skull-stripping remains a time-consuming and rate-limiting step in the neuroimage analysis pipeline.

The MLS algorithm was designed to facilitate large-scale neuroimage analyses by addressing some of the limitations of existing techniques. For example, the algorithm automatically adjusts the appropriate segmentation parameters using available information; some parameters are adjusted based upon information in the image metadata, whereas others are adjusted based upon information in the image itself, such as the image histogram. The MLS is less sensitive to bias fields and random noise than other methods, which also facilitates automation in large scale studies. The MLS algorithm can be used as a standalone tool with an easy-to-use graphical user interface, which enables interactive evaluation of the skull-stripping results. However, it can also be used as a module in the LONI pipeline processing environment (Rex et al., 2003), which enables automated, parallel processing of multiple data sets as is typically needed in large studies.

The MLS algorithm achieves moderate computational efficiency relative to existing techniques. It does not contain computationally expensive procedures such as registration, which is used to initialize in some skull-stripping algorithms, but it is not as fast as some morphology-based algorithms. The advantage of MLS is that it requires only one parameter, which is relatively robust across data sets, while morphology-based algorithms often require multiple parameters that are sensitive to small changes in the data. Thus, in large scale analyses, MLS achieves greater efficiency by skull-stripping a large number of data sets automatically, whereas prior approaches require a user to train an algorithm or manually adjust multiple parameters for, at best, a large subset of the data set (and, potentially, for every individual data set). Thus, MLS trades performance for greater accuracy, robustness, and automation.

Additional performance could be obtained by utilizing more efficient numerical schemes, optimizing the Java code, or by implementing the algorithm in a language with a highly optimized compiler such as C or C++. No attempt has been made to optimize the software; however, the Java source code is available so that other researchers can evaluate the algorithm and modify or rewrite the current implementation of the algorithm.

The MLS algorithm was tested using data sets obtained from multiple institutions with subject populations containing different age groups and pathologies. It was evaluated using expert-generated, manual segmentations as a “gold standard” and using multiple, objective evaluation indices. Although the MLS algorithm has proven generally robust and accurate, there are some disadvantages of the current implementation. First, MLS was implemented as a 2D algorithm that obtains image intensity information from one image slice at a time rather than from adjacent slices in the 3D volume. A 3D algorithm (i.e., using image intensity information from neighboring voxels in adjacent slices) might be less sensitive to noise than the 2D algorithm and might therefore produce more robust and accurate results. However, in some cases, the 2D MLS algorithm achieves more accurate segmentation than some 3D algorithms by using the similarity between 2D cortical contours on adjacent slices. For example, after the first slice, the initialization of each subsequent slice is based upon the results obtained from the prior slice. The segmentation result from each of the following slices is evaluated during the automatic MLS segmentation by comparing the current segmentation to that of the previous slice

and by determining whether the results are acceptable. If the segmentation significantly differs from the previous segmentation (i.e., with a very low Jaccard coefficient), the result is rejected, and the MLS adjusts the weight of the curvature term and re-segments the slice.

A second disadvantage of the current algorithm is that it sometimes failed to automatically skull-strip data sets characterized by extremely high noise or poor contrast resolution (BSE and BET also failed to automatically skull-strip these data sets). Incorporating information from adjacent slices could further improve the ability of the algorithm to automatically skull-strip such data at the cost of increased computational complexity. Another potential disadvantage of the algorithm is that it is likely to fail on data sets containing abnormal anatomic structures, such as postoperative cases, because the algorithm currently models normal brain anatomy.

Additional research is needed to address the potential disadvantages and limitations of the current model. Future studies should focus on extending model-based level set by including a combination of frequency-selective noise filtering and a 3D model of anatomical features to enable automatic skull stripping of extremely noisy or low contrast data sets. Future studies should also include postoperative data to evaluate the sensitivity of the model to abnormal structure. Alternative models might be incorporated to account for postsurgical brain anatomy. Lastly, although the number of data sets used in this study was comparable to other published studies on skull stripping, larger sample sizes and different types of studies are needed in order to demonstrate that the MLS algorithm can achieve robust skull-stripping in large, multi-institutional neuroimage analyses.

V. ACKNOWLEDGEMENTS

This work was supported by the NIH/NIMH research grant R01 MH071940 and the NIH/NCRR resource grant P41 RR013642. Additional support was provided by the National Institutes of Health through the NIH Roadmap for Medical Research, Grant U54 RR021813 entitled Center for Computational Biology (CCB). Information on the National Centers for Biomedical Computing can be obtained from <http: //nihroadmap.nih.gov/bioinformatics>.

VI. REFERENCES

Aboutanos, G.B., Nikanne, J., Watkins, N., Dawant, B.M., 1999. Model creation and deformation for the automatic segmentation of the brain in MR images. IEEE Transactions on Biomedical Engineering 46(11), 1346-1356.

Baillard, C., Hellier, P., Barillot, C., 2001. Segmentation of brain 3D MR images using level sets and dense registration. Medical Image Analysis 5, 185-194.

Bermel, R.A., Sharma, J., Tjoa, C.W., Puli, S.R., Bakshi, R., 2003. A semiautomated measure of whole-brain atrophy in multiple sclerosis. Neurological Sciences 208, 57-65.

Bidgood, W.D., Horii, S.C., 1992. Introduction to the ACR-NEMA DICOM standard. Radiographics 12(2): 345-355.

Bidgood, W.D. Jr., Horii, S.C., Prior, F.W., Van Syckle, D.E., 1997. Understanding and using DICOM, the data interchange standard for biomedical imaging. Journal of American Medical Informatics Association 4(3): 199-212.

Blanton, R.E., Levitt, J.G., Peterson, J.R., Fadale, D., Sporty, M.L., Lee, M., To, D., Mormino, E.C., Thompson, P.M., McCracken, J.T., Toga, A.W., 2004. Gender differences in the left inferior frontal gyrus in normal children. NeuroImage 22(2), 626-636.

Boesen, K., Rehm, K., Schaper, K., Stoltzner, S., Woods, R., Lüders, E., Rottenberg, D., 2004. Quantitative comparison of four brain extraction algorithms. NeuroImage 22(3), July, 1255-1261.

Chan, T.F., Vese, L.A., 2001. Active Contours Without Edges. IEEE Transaction on Image Processing 10(2), 266-277.

Courchesne, E., Chisum, H.J., Townsend, J., Cowles, A., Covington, J., Egaas, B., Harwood, M., Hinds, S., Press, G.A., 2000. Normal brain development and aging: quantitative analysis at in vivo MR imaging in healthy volunteers. Radiology 216, 672-682.

Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis I: Segmentation and surface reconstruction. NeuroImage 9, 179-194.

DeCarli, C., Maisog, J., Murphy, D.G.M., Teichberg, D., Rapoport, S.I., Horwitz, B., 1992. Method for quantification of Brain, Ventricular, and Subarachnoid CSF Volumes from MR Images. Journal of Computer Assisted Tomography 16(2), 274-284.

Dice, L., 1945. Measures of the amount of ecologic association between species. Ecology 26, 297-302.

Duncan, J.S., Ayache, N., 2000. Medical image analysis: Progress over two decades and the challenges ahead. IEEE Transactions on Pattern Analysis and Machine Intelligence 22, 85-106.

Fennema-Notestine, C., Ozyurt, I.B., Clark, C.P., Morris, S., Bischoff-Grethe, A., Bondi, M.W., Jernigan, T.L., Fischl, B., Segonne, F., Shattuck, D.W., Leahy, R.M., Rex, D.E., Toga, A.W., Zou, K.H., Brown, G.G., 2006. Quantitative Evaluation of Automated Skull-stripping Method Applied to Contemporary and Legacy Images, Effects of Diagnosis, Bias Correction, and Slice Location. Human Brain Mapping 27 (2), 99 - 113.

Grau, V., Mewes, A.U.J., Alcaniz, M., Kikinis, R., and Warfield, S.K., 2004. Improved watershed transform for medical image segmentation using prior information. IEEE Trans. Med. Imag., 23(4): 447-458.

Horsfield, M.A., Rovaris, M., Rocca, M.A., Rossi, P., Benedict, R.H.B., Filippi, M., Bakshi, R., 2003. Whole-brain atrophy in multiple sclerosis measured by two segmentation processes from various MRI sequences. Neurological Sciences 216, 169-177.

Huh, S., Ketter, T.A., Sohn, K.H., Lee, C., 2002. Automated cerebrum segmentation from three-dimensional sagittal brain MR images. Computers in Biology and Medicine 32, 311-328.

Jaccard, P., 1912. The distribution of flora in the alpine zone. New Phytol. 11(2), 37 – 50.

Jernigan, T.L., Archibald, S.L., Fennema-Notestine, C., Gamst, A.C., Stout, J.C., Bonner, J., Hesselink, J.R., 2001. Effects of age on tissues and regions of the cerebrum and cerebellum. Neurobiology of Aging 22(4), 581-594.

Kikinis, R., Shenton, M.E., Gerig, G., Martin, J., Anderson, M., Metcalf, D., Guttmann, C.R., McCarley, R.W., Lorensen, W., Cline, H., et al., 1992. Routine quantitative analysis of brain and cerebrospinal fluid spaces with MR imaging. Journal of Magnetic Resonance Imaging, 2(6): 619-629.

Lee, C., Huh, S., Ketter, T.A., Unser, M., 1998. Unsupervised connectivity-based thresholding segmentation of midsagittal brain MR images. Computers in Biology and Medicine 28, 309-338.

Leemput, K.V., Maes, F., Vandermeulen, D., Suetens, P., 1999. Automated model-based bias field correction of MR images of the brain. IEEE Transactions on Medical Imaging 18(10), 885-896.

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., Holmes, C., Collins, L., Thompson, P., MacDonald, D., Iacoboni, M., Schormann, T., Amunts, K., Palomero-Gallagher, N., Geyer, S., Parsons, L., Narr, K., Kabani, N., Le Goualher, G., Boomsma, D., Cannon, T., Kawashima, R., Mazoyer, B., 2001. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). Philos. Trans. R. Soc. Lond., B Biol. Sci. 356 (1412), 1293 – 1322.

Narr, K.L., Thompson, P.M., Szeszko, P., Robinson, D., Jang, S., Woods, R.P., Kim, S., Hayashi, K.M., Asunction, D., Toga, A.W., Bilder, R.M., 2004. Regional specificity of hippocampal volume reductions in first-episode schizophrenia. NeuroImage, 21(4), 1563-1575.

Neu, S.C., Valentino, D.J., Toga, A.W., 2005. The LONI Debabeler: a mediator for neuroimaging software. NeuroImage 24(4), 1170-1179.

Osher, S., Sethian, J., 1988. Fronts Propagating with Curvature Dependent Speed: Algorithm Based on Hamilton-Jacobi Formulations. Journal of Computational Physics 79, 12-49.

Osher, S., Fedkiw, R., 2003. Level Set Methods and Dynamic Implicit Surfaces: Motion in the Normal Direction. Springer.

Rehm, K., Schaper, K., Anderson, J., Woods, R., Stoltzner, S., Rottenberg, D., 2004. Putting our heads together: a consensus approach to brain/non-brain segmentation in T1-weighted MR volumes. NeuroImage 22(3), 1262-1270.

Rex, D.E., Ma, J.Q., Toga, A.W., 2003. The LONI Pipeline Processing Environment. NeuroImage 19(3), 1033-1048.

Rex, D.E., Shattuck, D.W., Woods, R.P., Narr, K.L., Luders, E., Rehm, K., Stolzner, S.E., Rottenberg, D.A., Toga, A.W., 2004. A meta-algorithm for brain extraction in MRI, NeuroImage 23(2), 625-637.

Robb, R.A., Hanson, D.P., Karwoski, R.A., Larson, A.G., Workman, E.L., Stacy, M.C., 1989. Analyze: a comprehensive, operator-interactive software package for multidimensional medical image display and analysis. Comput. Med. Imaging Graph 13(6): 433-454.

Rusinek, H., de Leon, M.J., George, A.E., Stylopoulos, L.A., Chandra, R., Smith, G., Rand, T., Mourino, M., Kowalski, H., 1991. Alzheimer disease: measuring loss of cerebral gray matter with MR imaging. Radiology 178(1), 109-114.

Sharma, J., Sanfilipo, M.P., Benedict, R.H.B., Weinstock-Guttman, B., Munschauer, F.E.III, Bakshi, R., 2004. Whole-brain atrophy in multiple sclerosis measured by automated versus semiautomated MR imaging segmentation. AJNR Am J Neuroradiology 25, 985-996.

Shattuck, D.W., Sandor-Leahy, S.R., Schaper, K.A., Rottenberg, D.A., Leahy, R.M., 2001. Magnetic resonance image tissue classification using a partial volume model. NeuroImage 13, 856-876.

Simmons, A., Tofts, P.S., Barker, G.J., Arridge, S.R., 1994. Sources of intensity nonuniformity in spin echo images at 1.5T. Magn. Resonance Med. 32, pp. 121-128.

Smith, S.M., 2002. Fast Robust Automated Brain Extraction. Human Brain Mapping 17, 143-155.

Strother, S., La Conte, S., Kai Hansen, L., Anderson, J., Zhang, J., Pulapura, S., and Rottenberg, D., 2004. Optimizing the fMRI data-processing pipeline using prediction and reproducibility performance metrics: I. A preliminary group analysis. NeuroImage 23, S196-S207

Suri, J.S., 2001. Two-Dimensional Fast Magnetic Resonance Brain Segmentation. IEEE Engineering in Medicine and Biology, 84-95.

Sussman, M., Smereka, P., Osher, S., 1994. A Level Set Approach for Computing Solutions to Incompressible Two-Phase Flow. Journal of Computational Physics 114, 146-159.

Tanskanen, P., Veijola, J.M., Piippo, U.K., Haapea, M., Miettunen, J.A., Pyhtinen, J., Bullmore, E.T., Jones, P.B., Isohanni, M.K., 2005. Hippocampus and amygdala volumes in schizophrenia and other psychoses in the Northern Finland 1966 birth cohort. Schizophrenia Research 75, 283-294.

Thompson, P.M., Mega, M.S., Woods, R.P., Zoumalan, C.I., Lindshield, C.J., Blanton, R.E., Moussai, J., Holmes, C.J., Cummings, J.L., Toga, A.W., 2001. Cortical change in Alzheimer's disease in detected with a disease-specific population-based brain atlas. Cerebral Cortex 11(1), 1-16.

Toga, A.W., 1999. Brain Warping. Academic Press, 1-26.\

van Ginneken, B., Frangi, A.F., Staal, J.J., ter Haar Romeny, B.M., Viergever, and M.A., 2002. Active shape model segmentation with optimal features. IEEE Trans. Med. Imaging. 21(8): 924-933.

Van Horn, J., Ellmore, T.M., Esposito, G., and Berman, K.F., 1998. Mapping Voxel-based Statistical Power on Parametric Imaging. NeuroImage 7, 97-107.

Van Horn, J.D., Grafton, S.T., Rockmore, D., Gazzaniga, M.S., 2004. Sharing neuroimaging studies of human cognition. Nature Neuroscience 7(5), 473-481.

Vannier, M.W., and Summers, R.M., 2003. Sharing images. Radiology, 228(1): 23-25.

Vese, L., Chan, T., 2002. A Multiphase Level Set Framework for Image Segmentation Using the Mumford and Shah Model. International J. of Computer Vision 50(3), 271-293.

Wells, III, W.M., Grimson, W.E.L., Kikinis, R., Jolesz, F.A., 1996. Adaptive segmentation of MRI data. IEEE Transactions on Medical Imaging 15, 429-442.

Woods, R.P., Grafton, S.T., Watson, J.D.G., Sicotte, N.L., Mazziotta, J.C., 1998. Automated image registration: II. Intersubject validation of linear and nonlinear models. J. of Computer Assisted Tomography 22(1), 139-152.

Woods, R.P., Dapretto, M., Sicotte, N.L., Toga, A.W., Mazziotta, J.C., 1999. Creation and use of a Talairach-Compatible atlas for accurate, automated, nonlinear intersubject registration, and analysis of functional imaging data. Human Brain Mapping 8, 73-79.

Zeng, X., Staib, L.H., Schultz, R.T., Duncan, J.S., 1999. Segmentation and Measurement of the Cortex from 3-D MR Images Using Coupled-Surfaces Propagation. IEEE Transactions on Medical Imaging 18(10), 927-937.

Zivadinov, R., Bagnato, F., Nasuelli, D., Bastianello, S., Bratina, A., Locatelli, L., Watts, K., Finamore, L., Grop, A., Dwyer, M., Catalan, M., Clemenzi, A., Millefiorini, E., Bakshi, R., Zorzon, M., 2004. Short-term brain atrophy changes in relapsing-remitting multiple sclerosis. Neurological Sciences 223, 185-193.