Unsupervised Skull Stripping in MRI

by

Florent Ségonne

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of Master of Science in Electrical
Engineering and Computer Science
at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2002

© 2002 Florent Ségonne. All rights reserved.
The author hereby grants to MIT permission to reproduce and to distribute publicly paper and
electronic copies of this thesis document in whole or in part.

Author........................................................................................................
Department of Electrical Engineering and Computer Science

Certified by..............

Olivier Faugeras
Adjunct Professor of E.E.& C.S.
Thesis Supervisor

Certified by............................................................................................

Bruce Fischl
Assistant Professor of Radiology Harvard Medical School
Thesis Supervisor

Accepted by..............

Arthur C. Smith
Chairman, Departmental Committee on Graduate Students
Unsupervised Skull Stripping in MRI

by

Florent Ségonne

Submitted to the Department of Electrical Engineering and Computer Science on May 10, 2002, in partial fulfillment of the requirements for the degree of Master of Science in Electrical Engineering and Computer Science.

Abstract:

Whole brain segmentation, referred to as skull stripping, is an important technique in neuroimaging. Many applications, such as presurgical planning, cortical surface reconstruction and brain morphometry, depend on the ability to accurately segment brain from non-brain tissue, i.e. remove extra-cerebral tissue such as skull, sclera, orbital fat, skin, etc. However, despite the clear definition of this essential step, no universal solution has been developed that is robust to neuroanatomical variability and the types of noise present in standard structural MRI sequences. We approach the skull-stripping problem by combining watershed algorithms and deformable surface models. Our method takes advantage of the simplicity and robustness of the former, while using the accuracy and surface information available to the latter. A training set of accurately segmented brains validates the segmentation and eventually corrects it, resulting in a robust and automated procedure.

Thesis Supervisor: Olivier Faugeras
Title: Adjunct Professor of E.E.& C.S.

Thesis Supervisor: Bruce Fischl
Title: Assistant Professor of Radiology

Keywords: skull stripping, brain segmentation, watershed transformation, template deformation, atlas-based segmentation.
Contents

1 Introduction.................................................................................................................. 9
2 MR Brain Image Segmentation ............................................................................. 11
  2.1 Nuclear Magnetic Resonance Imaging.......................................................... 11
  2.2 Brain extraction techniques......................................................................... 12
  2.3 Motivation........................................................................................................ 14
3 Methods..................................................................................................................... 15
  3.1 Preprocessing.................................................................................................. 17
  3.2 The watershed algorithm ............................................................................ 19
      3.2.1 The different steps............................................................................ 20
      3.2.2 Numerical Issues and discussion................................................. 21
      3.2.3 The post-watershed analysis: correction of some errors............ 22
  3.3 Deformable surface algorithm .................................................................... 23
      3.3.1 The parametric deformable model.............................................. 24
      3.3.2 Coarse surface deformation ....................................................... 26
      3.3.3 Calculation of fitting parameters.............................................. 26
      3.3.4 Fine surface deformation........................................................... 29
  3.4 Brain atlas comparison, validation and final segmentation ..................... 33
      3.4.1 Mapping the brain surface onto a sphere.................................... 34
      3.4.2 Atlas Construction.......................................................................... 35
      3.4.3 Validation of the shape of the surface......................................... 36
      3.4.4 Correction of the geometry of the surface ..................... 38
      3.4.5 Semi-local parameters estimation and final deformation........... 41
4 Assessment of the results and Discussion......................................................... 43
  4.1 The validation problem.................................................................................. 43
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1.1</td>
<td>Presentation of different algorithms</td>
<td>43</td>
</tr>
<tr>
<td>4.1.2</td>
<td>Definition of a norm: risk evaluation</td>
<td>44</td>
</tr>
<tr>
<td>4.2</td>
<td>Results and Discussion</td>
<td>45</td>
</tr>
<tr>
<td>4.3</td>
<td>Conclusion</td>
<td>46</td>
</tr>
<tr>
<td>5</td>
<td>Integration of Tools</td>
<td>49</td>
</tr>
<tr>
<td>5.1</td>
<td>Corrective Options</td>
<td>49</td>
</tr>
<tr>
<td>5.2</td>
<td>Additional tools</td>
<td>51</td>
</tr>
<tr>
<td>6</td>
<td>References</td>
<td>53</td>
</tr>
</tbody>
</table>
### Tables of Figures

Figure 1: An example of the function $f$ is used to compute the intensity characteristics of the white matter. .................................................................18
Figure 2: A simple illustration of the merging process........................................21
Figure 3: The smoothness update fraction.......................................................25
Figure 4: Construction of the CSF and gray matter curve............................27
Figure 5: Estimation of the transition threshold...........................................29
Figure 6: An incorrect segmentation that has expanded around the right temporal lobe......................................................................................30
Figure 7: A super-sampled surface and the resulting surfaces ......................31
Figure 8: Different iterations of the fine surface deformation.......................32
Figure 9: Incorrect segmentation removing the left part of the cerebellum; the corresponding curvature map is projected onto the surface.................................................................33
Figure 10: Three views of the curvature field and its mapping onto a sphere........................................................................................................35
Figure 11: Three views of the mean curvature field (ventral, lateral, dorsal), the corresponding variance field, and distance field. The variance map of the distanced field is not sketched, as it is quite low and uniform. ........................................................................36
Figure 12: The curvature, distance and global error maps.............................38
Figure 13: The corrective deformation process with the curvature field of the canonical template projected for convenience of the visualization.................................................................40
Figure 14: The resulting corrected surface....................................................40
Figure 15: Error in skull strip performed by three algorithms vs manual skull strip in ten subjects.................................................................45
Figure 16: Four views of the outer skin surface..................................................52
Figure 17: An example of a labeled volume..........................................................52
1 Introduction

A few years ago, only a small number of non-invasive techniques were available to radiologists, and deep insight, even intuition, was required for clinical diagnostic imaging. In recent years, the improvement and development of many image acquisition techniques (magnetic resonance imaging, computed tomography, digital mammography...), the enhancement of the general quality of the acquired images, associated with the advances in image processing and the development of large computational capacities, have considerably eased radiological tasks, and have drastically boosted the expansion-interest of research in the interdisciplinary field of medical image processing. Nowadays, with the multiplication of medical images and the global enhancement of their quality, the use of image processing techniques has become necessary.

The delineation of anatomical structures and other regions of interests, known as image segmentation, is a common need in medical imaging tasks. In this work, we focus on the problem of extracting the brain from an MR Image. Many applications, such as presurgical planning, cortical surface reconstruction and brain morphometry, depend on the ability to accurately segment brain from non-brain tissue, e.g. remove extra-cerebral tissue such as skull, sclera, orbital fat, skin, etc. However, despite the clear definition of this essential step, no universal solution has been developed that is robust to neuroanatomical variability and the types of noise present in standard structural MRI sequences. We approach the skull-stripping problem by combining watershed algorithms, deformable surface models and Atlas-based techniques. Our method takes advantage of the simplicity and robustness of the former, while using the accuracy and surface information
available to the latter. At each step of the pipelined algorithm, the results are evaluated and eventually corrected, resulting in a robust and automated procedure.

The report is organized as follows. In Section 2, a brief review of MR brain image segmentation is completed and the main current techniques are described. We present our method in Section 3 and explain its different steps in details. In Section 4, we discuss some issues about the validation problem and our results. Finally, we describe some useful tools that we have integrated to the algorithm.
2 MR Brain Image Segmentation

Before describing some current brain segmentation techniques, we briefly review some general concepts on Nuclear Magnetic Resonance Imaging (NMR) and the main limitations in NMR image processing. A good analysis of the principles of Magnetic Resonance Imaging (MRI) can be found in [13] and [16] (for a quantum explanation of MR principles) or in [11] (for a classical interpretation), and an extensive survey of segmentation techniques in Medical Imaging is completed in [21].

2.1 Nuclear Magnetic Resonance Imaging

Nuclear Magnetic Resonance Imaging, is a non-invasive 3D-image acquisition technique, based on differences in induced magnetization in biological tissues in the presence of a magnetic field. MRI attempts to capture magnetic properties of the material, such as proton density and relaxation times T1 and T2. These intrinsic parameters differ from one tissue type to another and are therefore the source of contrast in the image. Different acquisition parameters (pulse repetition time TR, pulse echo time TE...) allow the generation of diverse types of images, with different contrasts between tissues (note that, for some types of acquisition, a forward model exists for estimating an acquired image given the tissue parameters; Bloch equations: Bloch 1946). T1-weighted MR images are commonly acquired for investigating structural aspects of the brain as they have good contrast between different tissue classes such as cerebrospinal fluid (CSF), gray matter and white matter. For these reasons, many different skull-stripping techniques have been specifically designed for this type of images, and we follow this choice for the implementation of our algorithm.
MR Image segmentation, which plays a crucial role in many medical imaging applications by automating or facilitating the delineation of anatomical structures and other regions of interest, is a difficult task due to a variety of limitations in the data (noise, partial voluming, tissue inhomogeneities, RF inhomogeneities, subject motion...). While image-processing techniques can reduce some of these difficulties, any robust segmentation technique must account for these unwanted sources of variation in the intensity distributions of a tissue class. In particular, intensity inhomogeneities, which cause a shading effect over the image, can significantly degrade the performance of methods that assume that the intensity value of a tissue class is constant over space. In T1-weighted MR images, this undesirable intensity variation can significantly perturb the intensity distributions, causing them to overlap, leading to misclassification, when solely intensity based models are used. Therefore, we will have to take into consideration these limitations for the design of our method.

### 2.2 Brain extraction techniques

The skull-stripping problem is a difficult task, which aims at extracting the brain from the skull, eliminating all non-brain tissue such as bones, eyes, skin, fat… This procedure requires a semi-global understanding of the image, as the brain is constituted of different structures with various geometric and intensity properties (cerebellum, cortex...), as well as a local understanding of it (accurate localization of the pial surface). In addition, the presence of imaging artifacts, anatomical variability, varying contrast properties and poor registration considerably increase the difficulty of this essential step in computational analysis of neuroimaging data.
Current approaches to automated skull stripping can be divided into different categories: Region-based [1],[15],[17],[18],[27], watershed-based [9],[26],[31] and template-based [3],[4],[7],[8],[9],[15],[27],[32].

- **Region-based approach:**
  These methods extracts regions that are connected based on some predefined criteria (for instance, intensity based): they are often used with some other techniques such as thresholding or clustering methods, which aim at determining some unknown parameters, or at approximately pre-segmenting the image into different classes. Most of the time, these techniques require an interaction with the user (some parameters are usually needed, as these methods have proved difficult to cope with different types of scans), and are highly sensitive to intensity inhomogeneities.

- **Watershed-based approach:**
  These methods are similar to the Region-based approach, in so far as they usually tend to segment an image into connected components (the gradient intensity is usually the criteria defining connectivity, but some intensity-based approaches are used in a similar way, with the advantage of being less noise-sensitive). One of the main drawbacks of this method is that it can easily suffer from over-segmentation, which is the reason why they are usually followed by a post-processing step to merge separate regions that belong to the same structure. Depending on the post-processing step, these methods can be quite robust to intensity inhomogeneities.

- **Template-based approach:**
  These techniques incorporate shape information into the segmentation process, modeling the brain surface by a smoothed deformed balloon. An iterative template deformation, usually driven by an intrinsic smoothing force and an image-based force, tends to fit the model to the correct part of the image. These methods seem
to be relatively robust, and less sensitive to image artifacts. However, even if they do not require extensive human interaction, some parameters still need to be adjusted, depending on the type of scanners used. In addition, some parts of the brain are difficult to correctly extract, and are subject to recurrent errors, such as the base of the cerebellum or the temporal lobes.

2.3 Motivation

Despite the clear definition of this essential process, no universal solution has been found yet, and the skull-stripping problem remains an area of active research. Our goal is to design a robust algorithm, capable of automatically extracting the brain from MR images, even in the presence of strong noise artifacts, and without requiring any user interaction. The proposed algorithm should be able to detect its own errors, and correct them, leading to a final brain surface that accurately reflects the outer pial surface. The validity of the proposed method will be evaluated, by comparing it to other segmentation techniques using a manually labeled set of images as a gold standard.
3 Methods

The skull-stripping technique is based on a few general concepts. Regarding brain anatomy, we exploit the connectivity of the white matter: the white matter is a wide bright and uniform region that connects all parts of the brain and that is surrounded by darker regions (gray matter, cerebrospinal fluid). The global shape of the brain can be modeled by a simple smoothed and deformed balloon, which should lie at the limit between the gray matter and the cerebrospinal fluid (CSF). Besides, as its topology is that of a sphere, it allows atlas information to be incorporated.

Therefore, the skull-stripping technique consists of a series of sequential steps. First, some relevant parameters are robustly estimated from the input image $I$. Next, a watershed algorithm is performed on the intensity image, with a global minimum initialized within the cerebral white matter. Finally, a deformable surface procedure is applied to the output of the watershed algorithm to recover parts of cortex that may have been erroneously removed, using smoothness constraints on the shape of the skull and atlas information. Each of these steps is discussed in detail below.
Preprocessing
Global Parameters Estimation
(centroid of the brain, brain/head radius, CSF – white matter parameters)
Localization of one seed point

Watershed Computation
Sorting
Watershed Processing
Post-Analysis

Template Deformations
First Template Smoothing
Parameters Estimation
(CSF, gray-matter, transition intensity)
Second Template Smoothing

Brain Atlas Comparison
Validation
Semi-local Parameters Estimation
Final Template Deformation

Shape Correction
Atlas-based Energy Functional

Output Image
3.1 Preprocessing

Following the work of Stephen Smith in BET [17], we first compute estimates of a set of parameters required for subsequent processing. These include initial estimates of the intensity of CSF, the coordinates of the centroid of the brain (COG), and an estimate of the average radius of the brain. We proceed as follows: ignoring very bright and dark voxels, which are defined as voxels whose intensity lies above 98% of the cumulative histogram of the image and below 2% respectively, the CSF threshold is set to lie 10% of the way between the minimum intensity and the maximum intensity in the image. This initial estimate is used to roughly distinguish between brain and non-brain tissues (cerebrospinal fluid, skull...). Then, the voxels that are classified as brain matter are used to estimate the centroid of the brain: every voxel, whose intensity is above the previously calculated CSF threshold, is used in a standard sum of position weighted by their intensity. Finally, a rough estimate of the brain/head radius is calculated.

In order to determine the white matter parameters, we construct a curve using the fact that, in T1-weighted MRI, the white matter can be defined as a uniform region localized in the center of the head, with a quasi-constant intensity. That is, we assume that the variance within the white matter is small relative to other brain regions. The white matter parameters are computed from a cubic region centered on the centroid of the brain volume. The length of each edge of the cube is half the previously estimated brain radius. For each voxel within the cube, a 27-neighbor mean and variance is computed. Then, for each intensity $i$ encountered in the image $I$ (with domain $\mathbb{R}$), we calculate the following ratio: $f(i)=n(i)^2/v(i)$, where $n(i)$ is the number of voxels having an intensity equal to $i$, and $v(i)$ is the sum of the variance of these voxels:

$$n(i) = \sum_{\tilde{x} \in R} \delta(I(\tilde{x}) - i) \quad \text{and} \quad v(i) = \sum_{\tilde{x} \in R} \delta(I(\tilde{x}) - i) \var(\tilde{x}).$$
The function $f$, indexed by the intensity $i$, is used to calculate class statistics of white matter (see Figure 1). We first localize the main lobe of the function $f$, defined as the region where $f$ reaches its maximum $f_{max}$ (over a 5-point window), and for which $f(i)$ is superior to $f_{max}/3$; The bounds of the main lobe region are $WM_{MIN}$ and $WM_{MAX}$. We can define this region mathematically by:

$$\{i \in [WM_{MIN},WM_{MAX}] \mid \forall i \in [WM_{MIN},WM_{MAX}] \; f(i) \geq f_{max}/3\}$$

Then, within this region, we calculate the average intensity of the white matter and its variance, using maximum likelihood estimates.

![Diagram of function f](image)

**Figure 1:** An example of the function $f$ is used to compute the intensity characteristics of the white matter.

Finally, we locate the white matter voxel with minimum variance in the cube of interest, with an intensity greater than $WM_{MIN}$. This location is then used to establish a global minimum, ensuring that the main basin of the watershed algorithm will represent the brain, and therefore preventing it from being merged with non-brain regions such as the eye sockets.
3.2 The watershed algorithm

Watershed algorithms are based on image intensities. They usually try to locate the local maximums of the gradient intensity to segment the image into different connected components. The image intensity is often interpreted as height information: they use voxel values as “heights” in a landscape in which the brightest points correspond to the hills, and the darkest points represent the valleys. One of the main drawbacks of these techniques is that they often result in an over-segmentation, and appropriate merging criterions are required to postprocess the images. An interesting approach has been proposed by H.K. Hahn in [12], in which a solely intensity-based watershed algorithm is described: a simple merging criterion is defined to overcome the over-segmentation problem resulting in a fast and robust segmentation technique when quite uniform regions are targeted.

Regarding brain anatomy, our basic assumption is the connectivity of the white matter. This approach is exactly the one described in [12]. Since darker gray matter and even darker CSF surround the connected white matter, this region can be interpreted as the top of a hill.

We apply to the MRI image a watershed transform based on the concept of preflooding in order to avoid over-segmentation, a common problem in watershed techniques. We consider the inverted gray level in the T1-weighted brain image: under this transformation, the $WM$ hill becomes a valley. Two points of the inverted image are connected if a path of adjacent voxels exists between them that is at least as dark as the brighter one of the two points. Under this strict definition of "connectivity", the transformation would result in an over-segmented brain. For that reason, we have to weaken our criterion for connectivity and utilize the concept of preflooding, as proposed by [12]. We do so by allowing the connectivity path to
show a lower intensity than the darker of the two connected points up to a maximum difference: the *preflooding height* $h_f$.

### 3.2.1 The different steps

The first step of the watershed algorithm is the sorting of all voxels (of the gray level inverted image) according to their intensity. Then, we process each voxel in ascending order:

- **If** the voxel doesn’t have any already processed neighbors in its *three-dimensional* 6-neighborhood (i.e. voxels of same or less intensity), a new basin is formed: this voxel represents a local intensity minimum.

- **Otherwise**, we merge the voxel with the deepest neighboring basin, which is the basin with the darkest bottom voxel: “*voxel-basin merging*”. If two or more neighbors have already been processed belonging to different basins, these are tested for “*basin-basin merging*”: All the neighboring basins whose depth relative to the current voxel intensity is less than or equal to the *preflooding height* $h_f$ will be merged with the same basin as the voxel itself, i.e. the deepest neighboring basin.

After the transform with an appropriate preflooding height, one basin should exist that represents the whole brain, and will include the previously identified white matter voxel in Section 3.1.
3.2.2 Numerical Issues and discussion

The complexity of the modified watershed transform is linear in the number of voxels $N$ (usually $N=256 \times 256 \times 256$) for the voxel processing, and in the worst case proportional to $N \log N$ for the sorting. However, taking into account the fact that the intensity of each voxel intensity can be represented with a single byte (range from 0 to 255), we can accomplish the sorting in $N$ operations, resulting in a watershed transform whose computational complexity is linear in the number of voxels.

The main objective of the skull-stripping algorithm is to be robust: this first step performs very well in the presence of intensity non-uniformity and noise, and does not fail as long as the connectivity of white matter is preserved. For any given image, the parameter $h_{\beta}$ can be varied over a certain range without significantly changing the output. For the current algorithm, we use a value of 25 (corresponding to 25% of the maximum intensity).
However, the resulting image is often inaccurate and non-smooth, since some extra-cerebral tissue and CSF still remains, and some artifacts, such as susceptibility artifacts, in the original image may cause an irrelevant segmentation. The brain may be split in two or more basins, and the resulting segmentation may give an inaccurate result. For these reasons, we first check the validity of the initial watershed segmentation, correct it where possible, and subsequently apply a surface-based technique in order to utilize geometric criteria (i.e. smoothness) to more accurately estimate the true brain/non-brain boundary.

3.2.3 The post-watershed analysis: correction of some errors

We first determine whether the size of the main basin seems to be correct. We use the brain radius previously calculated to estimate the volume of the brain, assuming spherical geometry. If the size of the main basin is significantly smaller than the estimated volume (4 times smaller), we assume that the watershed segmentation failed and needs to be corrected. This is accomplished by merging basins into the main basin representing the initial segmentation. The basin we merge is one that is adjacent to the main basin, localized in the white matter, and whose volume would result in the segmentation being the closest to the estimated brain volume. In the cases in which the watershed segmentation fails, few basins are found and the chosen basin is often the biggest. In our experience, this post-watershed basin merging frequently results in a correct segmentation.

Nevertheless, some artifacts result in the brain being segmented in many equally sized basins. In this case, further corrections are required. To correct this type of failure, we merge all the ambiguous basins with the main one. An ambiguous basin is defined as a basin, which contains a significant volume of white matter at its common border with the already segmented basin. A voxel is said to be ambiguous if it is a neighbor of the main basin, if its intensity is greater than $WM_{MIN}$ and smaller than $WM_{MAX}$, with a variance smaller than $WM_{VAR}$. If the number of
ambiguous voxels is superior to a specific threshold, then we label this basin as “ambiguous”, and merge it to the brain basin. The threshold we are using is equal to the cubic root of the considered basin: if the number of ambiguous voxels into an uncertain basin is greater than the cubic root of its size, then we merge it to the main one. We found that the use of a threshold depending on the size of the targeted basin leads to better results, in so far as it avoids merging some large basins with a comparatively small number of ambiguous voxels. We iterate this process until no new basin is merged with the main one. Usually, even with an over segmentation of the brain basin into many different smaller basins, this post watershed analysis merges most of them in a few passes: typically 2 or 3. However, most of the time, no new basin is found as the result of the watershed process is very likely correct.

3.3 Deformable surface algorithm

After the watershed computation, the segmented volume may contain some non-brain tissue such as CSF, or some parts of the skull. In some other infrequent cases, important parts of the brain may be removed, especially if the connectivity of the white matter is not preserved: the complete cerebellum may not be merged into the main basin, and be detached from the whole brain by the watershed segmentation. We first try to remove these extraneous parts, assuming that the watershed segmentation did not remove some important structures of the brain: we apply a deformable balloon-like template, with the segmented volume used as a mask. After the initial template deformation is completed, some global parameters regarding the brain/non-brain border are computed. Finally, the template deformation resumes, accurately matching the boundary of the brain.
3.3.1 The parametric deformable model

We follow the approach of Dale et al. in [3], adding to the curvature reducing force a smoothing term as in the work of Smith in BET [28]. The brain surface is modeled using a super-tessellated icosahedral surface tessellation: the initial model is an icosahedron, for which each triangle was iteratively subdivided into 4 smaller triangles. In this common tessellation, each vertex has five or six neighbors, according to its position relative to the original icosahedron.

We center the template at a recalculated center of gravity of the brain with an initial radius set to include the whole previous segmented brain, and then gradually deform it through a series of iterative steps. The deformation process is driven by two kinds of forces:
- an MRI-based force, designed to drive the template outward from the brain: $F_M(t)$,
- a curvature reducing force, enforcing a smoothness constraint on the deformed template: $F_S(t)$.

On each iteration $t$, we update the coordinate $x_k$ of each vertex according to the "forces" mentioned above: $x_k(t+1) = x_k(t) + F_S(t) + F_M(t)$.

The curvature reducing force $F_S$, which enforces a smoothness constraint on the deformed template, is based on the work of [28]. At each iteration, the mean position of all neighboring vertices is calculated to form a difference vector $\vec{s}$, which is decomposed into orthogonal components, normal and tangential to the normal surface: $\vec{s} = \vec{s}_n + \vec{s}_t$. The smoothness force used in [3] is given by $\vec{F}_S = \alpha_i \vec{s}_t + \alpha_n \vec{s}_n$, with $\alpha_i$ and $\alpha_n$ two constants (usually set to 0.5). This
The approach can be significantly improved by making $\alpha_n$ a nonlinear function of $\tilde{s}_n$. In this way, small departures from planarity are not penalized, while large ones are disallowed. More specifically, we proceed as follows.

We estimate the local radius of curvature $r$ at location $x_k$: 

$$r = l^2 / (2 \cdot |\tilde{s}|),$$

where $l$ is the mean distance from vertex to neighboring vertex. Then, the update fraction $\alpha_n$ is set to a sigmoid function of $r$:

$$\alpha_n = \frac{1 + \tanh(B \cdot (1/r - A))}{2},$$

with $A$ and $B$ equal to: 

$$A = \frac{1}{r_{\text{in}}} + \frac{1}{r_{\text{max}}},$$

$$B = 6 \cdot \frac{1/r_{\text{min}} - 1/r_{\text{max}}}. $$

The two parameters $r_{\text{min}}$ and $r_{\text{max}}$ reflect prior hypothesis about the allowable curvature of the brain/non-brain boundary and are set to $r_{\text{min}} = 3.33$ and $r_{\text{max}} = 10$. The sigmoid function of $r$ penalizes high local mean curvatures of the surface, where $1/3.33$ is assumed to be a superior bound to the expected local curvature. We note that future work will incorporate atlas-based information to locally set these two parameters.

![Figure 3: The smoothness update fraction](image)

The tangential update fraction $\alpha_t$, which aims at regularizing the tessellations, will be set to 0.8: this strong tangential force will lead to quite uniform tessellations, which could be used by other techniques that require equally tessellated surfaces (Section 5.2).
3.3.2 Coarse surface deformation

The surface deformation thus proceeds as follows. We first center the template at recalculated COG coordinates, with its radius set to include the whole previous segmented brain. Once the initial template has been positioned, we gradually deform it through a series of iterative steps, using the segmented brain found in the watershed process as a mask. Therefore the MRI-based force used here is simple: the force acting on each vertex is designed to drive the surface to regions out of the segmented brain (i.e. with 0 MRI intensity values), repelling the surface outwards from contiguous regions consistent with the segmented brain (i.e. with strictly positive MRI intensities). The smoothness force is updated at every step using $r_{\text{min}}$ and $r_{\text{max}}$ set to 3.33 and 10. The deformation terminates when the maximum vertex displacement falls below a predetermined value (usually between 50 and 70 iterations).

3.3.3 Calculation of fitting parameters

Before applying the final deformable template in order to accurately fit the outer boundary of the brain, we require more information regarding the intensity values at the brain/non-brain boundary, such as the mean intensity of CSF and gray matter. The contrast between gray matter and CSF is sufficiently large that using a predefined global transition intensity should suffice to distinguish the two tissues. We compute these parameters by examining the intensity values along the surface normals, extending from a few millimeters inward to a few millimeters outward from the previously calculated surface. Along this line segment, we construct two curves, which reach their maxima around the mean CSF and gray matter intensities.
**Construction of the CSF curve and extraction of some parameters:** As part of the watershed processing we have already determined a value for the CSF intensity. This value now needs to be refined. Our first estimation corresponds to a simple threshold that has been set to roughly distinguish between brain and non-brain tissues, based on the cumulative histogram of the image. It did not directly use CSF intensity information. However, the surface that is a result of the coarse template deformation described above mostly passes through “dark” regions in the T1-weighted image, corresponding to CSF or skull. For each vertex of the surface, looking 2-mm outwards and 2-mm inwards among 5 voxels, we find the voxel with the smallest intensity, and keep it for the construction of the CSF curve if its intensity is less than 3-times the previously estimated CSF intensity (We ignore bright areas where the surface passes, such as the eye sockets):

$$f_{CSF}(i) = \sum_{\text{all kept voxel } \bar{x}} \delta_i \cdot \text{intensity}(\bar{x})$$

Following the same approach as for the estimation of the white matter parameters, we localize the main lobe of the function $f$: we find the maximum of the curve, i.e.
the new CSF intensity value, and identify two inferior and superior limits of the main lobe: CSF_MIN and CSF_MAX.

Construction of the gray matter curve: The gray matter curve is obtained looking 20-mm along the inward normal. Along the normal of each vertex of the tessellation, progressing iteratively mm-by-mm inwards, we search for a uniform white matter area, that we defined as a 3*3*3-voxel region, in which the average intensity falls between WM_MIN and WM_MAX, and the variance is smaller than WM_VAR. As soon as we succeed in localizing this area, meaning that we have identified a uniform white matter region along the inward normal, we use the voxels located along the normal, between this area and the tessellated surface, to increment a gray matter curve, as we previously did for the CSF curve. These voxels, bounded by white matter voxels and CSF ones, are expected to represent gray matter. Then, we extract the gray matter parameters.

Extraction of the transition intensity: Finally, a transition threshold is easily found as an intersection point between the gray matter curve and the CSF curve. This intensity intersection determines a robust threshold to accurately fit the pial surface of the brain during the final template deformation. As we previously noticed, the contrast between CSF and gray matter is sufficiently large that using a global threshold is a simple and viable choice to distinguish between them.
3.3.4 Fine surface deformation

The brain surface is still modeled by a surface tessellation using connected triangles. This second tessellation uses a super-sampled icosahedron with 10242 vertices (resulting in a mean triangle edge size of approximately 2.5 mm). Before applying the template deformation process, we remove non-brain tissue with the previously calculated surface. This step may seem inappropriate, in so far as the final deformation will only be able to remove some parts of the brain that the first deformation failed to, but never add anything new. However, as we will describe it in the next section, a comparison with an atlas will validate or invalidate the segmentation and eventually correct it. The main argument for peeling the outside of the first segmentation surface is robustness: the design of a general MRI-based force that will lead to successful segmentation on any MRI scans is very difficult. Some regions of the brain are difficult to distinguish from non-brain regions, such as the temporal lobes or the base of the cerebellum; in these complex areas, where...
the boundary of the brain is not easy to find, a general MRI-based force will occasionally expand the surface away from the correct boundary into non-brain tissues. Figure 6 illustrates this kind of error. The peeling will prevent the second deformation balloon from expanding away in these intricate regions.

Figure 6: An incorrect segmentation that has expanded around the right temporal lobe

The deformation process is still driven by two kinds of forces, but with a different MRI-based force. The force used here is designed to push the surface outward if the local mean intensity is higher than the estimated threshold intensity. Denoting the volume intensity at position \( x_i \) by \( I(x_i) \), we define the following energy functional:

\[
J_I = \frac{1}{2V} \left( \sum_{\text{vertices } i} T - I(x_i) \right)^2,
\]

where \( T \) is the transition intensity previously calculated. The value of \( I(x_i) \) is computed on a subvoxel basis using trilinear interpolation. The MRI-based force of the \( i^{th} \) vertex is then defined by the following equations:

\[
f_i = \begin{cases} 
-n_i & \text{if } I(x_i) > WM_{MAX} + 2\sigma_{WM} \\
-\frac{\partial J_i}{\partial x_i} & \text{otherwise}
\end{cases}
\]
where $WM_{MAX}$ and $\sigma_{WM}$ are respectively $WM_{MAX}$ and $WM_{VAR}$. The first term is introduced to avoid bright regions, such as the eye sockets: in such regions, the surface is pushed inwards. In this series of iterative steps, which we limit to 40 (in case of an incorrect watershed segmentation, it will prevent the surface from expanding away in some areas difficult to segment), $r_{min}$ and $r_{max}$ are still set to 3.33 and 10 respectively (prior information on the main curvature of the brain). The resulting surface should closely follow the brain/non-brain boundary.

Figure 7: A super-sampled surface and the resulting surfaces
Figure 8: Different iterations of the fine surface deformation
3.4 Brain atlas comparison, validation and final segmentation

As we pointed out in the last section, the watershed computation may cause over-segmentation, and remove some important structures of the brain. In our experience, this kind of over-segmentation often leads to the removal of half of the cerebellum or of the whole cerebellum (see Figure 9). In order to detect and correct these errors, which are mainly due to a deterioration of the connectivity of the white matter, the surface is compared to an atlas containing geometric information compiled from successfully segmented brains. Based on its local curvature map and its distance to the center of gravity map, which we first normalize (zero mean and unit variance), the surface correctness is evaluated and the errors precisely localized. By normalizing the fields locally attached to the surface, we don’t incorporate prior information on the expected size of the brain, making the algorithm more general. If necessary, based on the information given by the brain atlas, the shape of the surface is iteratively corrected. Once the surface segmentation has been validated, we locally estimate a new transition threshold, and the deformation resumes, incorporating an atlas-based term, preserving the shape of the surface. The resulting surface should then accurately reflect the true brain surface.

Figure 9: Incorrect segmentation removing the left part of the cerebellum; the corresponding curvature map is projected onto the surface.
3.4.1 Mapping the brain surface onto a sphere

The correctness level of the segmentation is easily confirmed by visually inspecting the shape of the highly tessellated surface. However, this task becomes much more difficult when we require a computer to validate the segmentation without external help. We have decided to base our validation assessment on two fields that include much of the needed information: the mean curvature of the surface and the distance to its COG. Before calculating these two fields, we first smooth the surface by a series of iterative steps. The mean curvature map will be a measure of the convexity of the surface, and the distance to the COG will contain important information concerning the corrective deformation field (between the surface and the atlas). In order to assess the validity of our segmentation, we need to map the surface into a parameterizable shape, which will allow us to compare surfaces across subjects. The sphere is a natural choice as it allows the preservation of the topological structure of the original surface (no cuts need to be introduced preserving the local connectivity), and it retains much of the computational attractiveness of a flat space (which would not be the case with asymmetric surfaces such as ellipsoids). We note that some high-resolution cortical coordinate systems based on this shape have been proposed in the literature [3],[8]. In our case, the spherical surface will facilitate the rigid registration of our surface with an atlas, and will allow us to calculate the maximum likelihood at each iteration of the template deformation in a very simple manner. Therefore, we need to find a mapping from the brain surface onto a sphere. Fortunately, we can avoid unfolding our surface ([3],[7],[9]), by noticing that the initial tessellated surface is modelled using an icosahedron, meaning that our original surface is exactly a tessellated sphere. At each vertex of our initial tessellated sphere, we attach the two corresponding fields of the final surface: that will characterize our mapping. We note that the metric distortion induced by this method will be small, since the deformation process makes use of an important tangential force that tends to preserve the geodesic
distances and since the expected errors will not introduce big folds into the surfaces; as a consequence, we can neglect this distortion.

Figure 10: Three views of the curvature field and its mapping onto a sphere

### 3.4.2 Atlas Construction

The first step in validating the shape of a surface is the construction of an atlas that will contain the information necessary to detect and eventually correct the segmentation. We iteratively build this atlas from a training set of accurately segmented brains for which the curvature and distance to COG fields are computed and normalized (zero mean and unit variance). First, a non-aligned template is built, containing the mean and variance of each field at each coordinate. The parameterization of the sphere uses 512*256 images (latitude longitude coordinate system), weighted by the metric tensor. Then, each brain is rigidly aligned with the template and a new canonical surface is constructed; we note that the two fields are used to align each brain with the canonical template.
Figure 11: Three views of the mean curvature field (ventral, lateral, dorsal), the corresponding variance field, and distance field. The variance map of the distance field is not sketched, as it is quite low and uniform.

3.4.3 Validation of the shape of the surface

Once an accurate representation of the brain surface has been found, the curvature and COG distance fields are computed. Then each field is smoothed with a surface-based Gaussian filter $\sigma^2=20$ on the tessellation, which spatially corresponds to a 2-3 cm Gaussian smoothing, and the mapped sphere is rigidly aligned with the canonical surface. The alignment minimizes the following energy functional:
\[ J_R(\varphi, \theta) = \frac{1}{2 \cdot V} \sum_{i} \left[ (R(\varphi, \theta) \circ \mathbf{X}_i - \overline{\mathbf{X}}_i)^T \overline{V}_i^{-1} (R(\varphi, \theta) \circ \mathbf{X}_i - \overline{\mathbf{X}}_i) \right], \]

where \( \mathbf{X}_i \) is a 2-dimensional vector containing the curvature and distance fields, \( \overline{\mathbf{X}} \) is a mean vector obtained from the parameterized template, \( \overline{V} \) the corresponding covariance matrix and \( R \) represents a rotation operator. For simplifying the registration, we assume that the two fields are not correlated and we choose \( \overline{V} \) to be diagonal. Future work will incorporate the full covariance matrix.

After the alignment the energy functional \( J_R \) provides a measure of the alignment error. An incorrect segmentation will lead to large errors, as the result of aligning dissimilar field regions, contrary to correct segmentations that will be accurately aligned overall. Figure 12 shows the resulting error map of a scenario when the segmentation fails and removes one half of the cerebellum. The threshold that we are taking for assessing the validity of the segmentation based on the functional energy is \( J_R \geq 4^2 \), meaning that the mean maximum likelihood distortion over vertices is greater than \( 4 \sigma \).

We can go one step further and use the information available from the canonical surface to precisely localize the misaligned regions. Since the main drawbacks of the watershed segmentation is to occasionally remove parts of the cerebellum, we check the correctness of the segmentation in these regions by evaluating the same energy functional, but restricted to the targeted vertices (localized from the atlas). More precisely, we divide the surface into 8 connected components relatively to their position to the centroid of the brain (from left-front-top to right-back-bottom). The computation of the errors in each of these regions allows us to detect whether the segmentation is correct and to localize the errors. The threshold used in each of these areas is \( 10 \sigma \), leading to less misclassification than with a global threshold. Given the atlas-tessellated sphere, we could precisely identify other regions, but we do not need a better accuracy for the purpose of the algorithm.
3.4.4 Correction of the geometry of the surface

If the surface segmentation fails to give the expected contour, we try to correct its shape with a series of deformations that iteratively minimize an energy functional $J$. Given a tessellated surface, a natural coordinate system would be the intrinsic vertices coordinate system. In order to use this convenient coordinate system, we inversely rotate the canonical parameterized surface such that the two surfaces align. We then project the two fields of the canonical surface onto a new tessellated sphere, such that the two vertices coordinate systems are the same. Therefore, we avoid the projection of our surface onto a parameterized surface (a $512 \times 256$ image) at each iteration, and we can directly work in the natural vertices coordinate system of our tessellation.

The mean vectors and covariance matrices of the parameterized template, which are zero mean and unit variance, need to be scaled such that they match the means and variances of the fields of the incorrect surface. We process iteratively: first, an estimate of these values is computed from the whole surface and the four canonical fields (two mean fields and two variance fields) are scaled to these values. Then, we compute a second approximation of these values from the vertices having a squared difference error below $4 \sigma^2$. This process is iterated until the canonical surface converges (two iterations are usually enough).
Consequently, the energy that we minimize can be generally written:

$$J_i = \frac{1}{2V} \sum_{j \in \Omega} (\mathbf{X}_i - \overline{\mathbf{X}})^T \overline{\mathbf{V}}_i (\mathbf{X}_j - \overline{\mathbf{X}}),$$

where $\mathbf{X}$ and $\mathbf{V}$ are the previous projected-and-scaled mean vectors and covariance matrices, and $\mathbf{X}_i = \mathbf{X}_j, (i, j \in \Omega_i). \Omega_i$ represents a set of neighbors of $x_i$, necessary to calculate the local curvature at $i$-th vertex. Taking the derivative with respect to $x_i$, we obtain the update term for the $i$-th vertex:

$$\frac{1}{V} (\mathbf{X}_i - \overline{\mathbf{X}})^T \overline{\mathbf{V}}_i \cdot \partial \mathbf{X}_i + \frac{1}{V} \sum_{j \in \Omega_i} (\mathbf{X}_j - \overline{\mathbf{X}})^T \overline{\mathbf{V}}_i \cdot \partial \mathbf{X}_j,$$

where $\partial_i = \partial x_i$.

In our situation, we would take $\mathbf{X}_i = (\text{Curv}_i, \text{Dist}_i)^T$ with $\text{Curv}_i$ being the curvature at the $i$-th vertex and $\text{Dist}_i$ being the distance to the center of gravity of the surface at the $i$-th vertex. The derivative of the distance term is easy to calculate since $\partial_i \text{Dist}_i = e_i$ and $\partial_i \text{Dist}_j = 0$. On the other hand, the derivative of the curvature term is much more complicated and results in a gradient that is dependent on the square of the inverse of the vertex spacing, and is therefore somewhat numerically unstable. The use of an oriented area could avoid these problems as proposed by Fischl et al in [3],[8], but the shape correction of our surface will not require such developments.

In the regions labeled as incorrectly segmented, we deform the tessellated surface under two forces:

- a curvature reducing force, enforcing a smoothness constraint on the deformed template and preventing the tessellation from intersecting $F_S(t)$.
- an atlas-based force, designed to drive the template outward from the brain and to align it with the canonical surface $F_A(t)$.

The atlas-based force we choose is $F_A = 0.5 \tanh(d_{\text{COG}})$, where $d_{\text{COG}}$ represents the Mahalanobis distance for the distance field. A small number of iterative steps is usually enough to correct the surface (less than 30 iterations). Figure 13 illustrates this process.
Figure 13: The corrective deformation process with the curvature field of the canonical template projected for convenience of the visualization.

Figure 14: The resulting corrected surface.

Once the shape has been corrected, two solutions are offered to accurately find the boundary of the brain. On the one hand, the shape-corrected surface allows us to extract a significant number of voxels that were not merged during the watershed computation. Similarly to the work described in Section 3.1, we could first estimate the local white matter variance and localize one or more new seed points before reapplying the whole algorithm. To our knowledge, this correction is sufficient to provide a final valid segmentation in most cases. However, some low quality images, connecting parts of the brain with other “invalid” structures will still lead to incorrect watershed segmentations, which this approach will be unable to correct.

On the other hand, we could try to incorporate three energy terms into our deformation process adding an MRI-based force to the former. We note that an initial deformation, as the previous one, with the MRI-based force set to zero is
suggested in order to avoid local minima problems. However, this approach leads to the same problems that we have pointed out in Section 3.3.4, i.e. it is difficult to extract the correct boundaries in some parts of the brain. In addition, these areas usually have significant anatomical variability, resulting in an atlas-based force that does not contain much useful information. In such areas, the design of an appropriate MRI-based force becomes critical, and the question of the robustness is logically raised. One solution would be to register non-rigidly our tessellated surface onto a non-rigidly aligned atlas, accurately mapping the problematic regions onto known areas of the canonical template, and allowing the use of a much richer atlas-based force. Nevertheless, we avoided this computationally intense problem and chose another method that made use of the preceding steps: from the error maps, we precisely localize the regions that need to be corrected. In these regions, in which some important parts have been removed, the MRI-based force make use of the original volume; the remaining vertices use the MRI information available from the peeled volume. These incorrect regions are defined by the areas, in which the local mean error is greater than $10\sigma$ (local refers to second neighborhood areas and the error measure is still based on the Mahalanobis distance). The surface is able to recover the previously removed regions, such as the cerebellum, but is prevented from expanding far away in some intricate regions. We describe this last deformation in the next section.

### 3.4.5 Semi-local parameters estimation and final deformation

At that point, the obtained template should closely follow the pial surface of the brain. However, one last weakness has still to be taken into account: as described in Section 2.1, intensity inhomogeneities can significantly degrade the performance of methods that assume that the intensity value of a tissue class is constant over space. Indeed, the tissue intensity of the cerebellum is often different from the intensity of the cerebral cortex, due to differences in intrinsic tissue properties (e.g. T1),
resulting in a final surface that could cut off some parts of its base. In order to avoid this last drawback, from the location of the surface, we estimate the local intensity parameters, by computing the corresponding transition thresholds, and smooth these local parameters with a surface-based Gaussian filter $\sigma^2=5$. Then the deformation process resumes: the curvature reducing force and the MRI-based force are the same as previously described in Sections 3.3.1 and 3.3.4, and the atlas-based force is given by $F_4=0.5 \tanh(d_{\text{COG}})$, where $d_{\text{COG}}$ represents the Mahalanobis distance for the distance field. After a few iterations (usually less than 30), the final surface accurately matches the pial surface of the brain.
4 Assessment of the results and Discussion

Stripping the skull and other non-brain tissues from the structural images of the head is a challenging and critical component for a variety of post processing tasks. It involves complex local and regional analysis of the MR Image. Large anatomical variability among brains, different acquisition methods, and the presence of artifacts increase the difficulty of designing a robust algorithm, such that current techniques are often susceptible to problems and require manual intervention. In order to validate the proposed algorithm, we compare it to some other existing techniques, and suggest a way of characterizing the correctness of the segmentation for every algorithm. We will then discuss the results.

4.1 The validation problem

We compared our algorithm with two other existing automated skull-stripping programs: the FMRIB's Brain Extraction Tool (BET: Smith [28]) and Freesurfer's strip skull (Dale et al [3]). In order to assess the validity of each segmentation, we ran a study on a small population of T1 weighted brain volumes that had also been manually skull stripped. Hence, we were able to use the manually segmented brain images as a gold standard for comparison.

4.1.1 Presentation of different algorithms

The three selected skull-stripping programs are all rapid (<1min on a current machine) and require no preprocessing. They all use surface deformation but the forces driving the deformation are different. The Dale et al algorithm uses a linear smoothness constraint for the curvature reducing force and an MRI-based force described in [4]. The BET algorithm makes use of a non-linear smoothness
constraint, as the one discussed in Section 3.3.1, and of an MRI-based force described in [28]. The MRI-based force of BET's algorithm is based on regional properties of the image and the threshold to push the template outwards is computed locally at each vertex. A threshold of 0.5 was applied for Brain Extraction Tool. Our algorithm makes use of regional statistics (see Section 3.4.5) for the template deformation, but utilizes a pre-watershed segmentation, and an atlas-based validation of the segmentation.

4.1.2 Definition of a norm: risk evaluation

In order to compare techniques, we have to come up with a way of measuring the quality of the segmentation. The error function that we will define should be a measure of the risk taken when using one algorithm or another. The manually segmented brains are identified as our gold standard, and the automatically extracted brains are compared to them. In our view, it is much more important to obtain a segmentation that keeps non-brain extra-parts than one which removes some fractions of the brain, since no other tools will be able to recover the missing parts. For that reason, we define the following error function:

\[ E = \frac{(N_{\text{voxels}}(b|nb) + C \cdot N_{\text{voxels}}(nb|b))}{N_{\text{GoldStandardBrainVoxels}}}, \]

Where \( N_{\text{voxels}}(b|nb) \) is the number of erroneously brain classified voxels and \( N_{\text{voxels}}(nb|b) \) is the number of incorrectly non-brain classified voxels. The factor \( C \) is introduced to penalize the eliminated brain voxels more than the added non-brain voxels. We arbitrarily set that constant to 10, implying that we would equate retaining 10 non-brain voxels or removing 1 brain voxel.
4.2 Results and Discussion

The results of the study are reported in Figure 15. Our method performed the best overall, followed by the ones of Smith and Dale et al. respectively. In the ten cases, our algorithm was able to precisely determine the pial boundary of the brain, resulting in an accurate segmentation associated with a small error. More specifically, very few brain voxels were erroneously removed by our algorithm; on the contrary, the two other algorithms cut off parts of the cortex. We note that the algorithm of Dale et al. is associated with a high risk of incorrect segmentation, as shown by the following results.

Figure 15: Error in skull strip performed by three algorithms vs manual skull strip in ten subjects
To our knowledge, the proposed algorithm is quite robust and the validation-correction steps are seldom required. In more than 90% of the segmentations (the algorithm has been run on more than one hundred brains), the computation does not necessitate any correction steps and the final segmentation is accurate. In most of the remaining "failing" cases, sufficient corrections occur in the post-analyze process. In rare cases, the watershed computation removes some important parts of the brain, such as the cerebellum, but, in our experience, the atlas-based correction always leads to a valid segmentation. Especially, the atlas-based correction requires a more intensive testing on a broader database. We believe that some images will still lead to an incorrect segmentation, but we think that our algorithm achieves the lowest risk overall.

4.3 Conclusion

In this paper, we have presented a new skull-stripping algorithm, combining watershed algorithms, deformable surface models and atlas-based techniques. Our method takes advantage of the simplicity and robustness of watershed approaches, while using the accuracy and surface information available to the deformable models and atlas-based techniques. At each step of the pipelined algorithm, the results are evaluated and eventually corrected, resulting in a robust and automated procedure. Specifically, we extend the watershed approach proposed by H.K. Hahn in [12] and improve the robustness of the algorithm by introducing a post-watershed analysis that validates and eventually corrects the segmentation. Template-based approaches, which incorporate shape information into the segmentation process and model the brain by a smoothed deformed balloon, seem to be relatively robust and less sensitive to image artifacts. Recently, they have been successfully used to extract the brain from MR images by Dale et al. in [3] and Smith in [28]. However, some parts of the brain are difficult to correctly extract
and are subject to recurrent errors. In this work, we propose a new template-based approach that incorporates an atlas-based term, which preserves the shape of the surface during the deformation. The atlas, compiled from a set of successfully segmented brains, is used to validate the shape of the segmentation and to eventually correct it during the deformation. To the best of our knowledge, this is the first time such atlas-surface-based techniques have been applied to this problem.

The validation of the method has been achieved through the comparison of our results with ones provided by other methods (Dale et al. [3], Smith [28]), and several manually segmented brains. Our method performed best overall, leading to few misclassified voxels in any segmentation.

The proposed algorithm is part of the free surfer software distributed by the center for Functional Neuroimaging Technologies at the MGH NMR Center.
5 Integration of Tools

We have presented an automated procedure to segment out the brain from T1 weighted MR Images. On the one hand, some of the methods presented in this paper may provide other interesting tools leading to differently processed images that could be used for some other perspectives. For instance, the deformation process model may generate high-resolution tessellated surfaces of the brain that could be used in fMRI analysis; or the segmentation of the brain, associated with the estimation of local parameters, may provide a segmentation into different classes (skull, csf, gray and white matter). On the other hand, even if the procedure is supposed to be fully automatic, some low quality images, containing artifacts, may lead to an inaccurate segmentation, that the atlas-based force may not be able to correct. In that case, the availability of some options, to detect the reasons of the failing segmentation and to correct it manually, are necessary. For these reasons, we have integrated a number of options that allow the user to set some parameters, and modify the output images.

5.1 Corrective Options

As we have just noted, the integration of some options that allow the user to modify the behavior of the algorithm, enabling the detection of errors, is of interest. The skull stripping algorithm is a pipelined algorithm in which the output image of a module acts as the input to the next. We have added a number of options that permit the user to stop the current segmentation after any step of the algorithm, or to eventually avoid some validation procedures. We detail below the available corrective options of the algorithm:

- Option \texttt{--wat} : option that stops the algorithm after the initial watershed segmentation (the post-analyze segmentation is included).
• Option \(-h h_{pf}\), with \(h_{pf}\) an integer: modification of the preflooding height used in the watershed segmentation. A typical value for \(h_{pf}\) is 25 (corresponding to 25% of the maximal intensity); lower values tend to decrease the size of the main basin (which could be confirmed by the post-analyze process) and higher values increase its size. However, as discussed in Section 3.2.2, for any given image, the preflooding height \(h_{pf}\) can be varied over a certain range without significantly changing the output.

• Option \(-n\): the post-analyze process (See section 3.2.3) is not used with this option.

• Option \(-t t_{ps}\), with \(t_{ps}\) an integer: that option specifies the threshold used in the post-watershed analysis. The default value is 100 (100%, meaning that the threshold is exactly the cubic root of the size of the ambiguous basin; see Section 3.2.3 for more explanations). A value of 0 will result in a post-watershed analysis that will merge any ambiguous basin with the main one.

• Option \(-\text{wat+temp}\): with this option, only the watershed segmentation and the first template deformation is used. This option could be typically used when the image contains a lot of artifacts, resulting in a fine surface segmentation that shaves some part of the cortex; in that extreme case, the watershed segmentation may still lead to successful results and the first template deformation to a valid brain segmentation.

• Options \(-\text{less}\) or \(-\text{more}\): option designed to increase or decrease the transition threshold used in the fine surface deformation described in Section 3.3.4. The result is respectively a shrunk or expanded surface compared to the typical surface.

• Option \(-\text{surf debug}\): this option will write out, into the output volume, the localization of the different surfaces used during the algorithm; it will enable the user to visualize the results of the different template deformations, and eventually to detect, understand and correct some errors.
• Option `-s int_i int_j int_k`: this option allows the user to add some new seed points for the watershed process. This option is very useful when the cerebellum is removed by the watershed segmentation, since a seed point inside the removed cerebellum may be added, leading to the merging of the corresponding basin with the main one. However, the atlas-based validation-correction step described in Section 3.4 should be sufficient to correct the segmentation.

• Option `-man int_CSF int_transition int_GM`: this option allows the user to manually set the parameters used in the fine surface deformation described in section 3.3.4.

5.2 Additional tools

We have included a number of options that generate different outputs than the typical segmented brain volumes. We describe these options below:

• Option `-surf surface_name`: this option is used to generate tessellated surfaces of the brain, inner and outer skull, and outer scalp from the T1-weighted image. The brain surface is the one produced by the fine surface deformation described in Section 3.3.4. The outer scalp surface is generated by another surface deformation. However, the MRI-based force is quite simple, since the scalp appears bright on T1-weighted images. The deformation is initialized with a radius set to include the whole head, and a series of iterative deformation matches the surface onto the head scalp-skin. The outer and inner skull surfaces are much harder to find, since the skull and the cerebrospinal fluid look the same in these type of images (dark intensity). Besides, the skull may contain some fat tissue that makes the border between the scalp and the outer skull difficult to localize. For these reasons, the two surfaces, outer skull and inner skull, are respectively
interpolated from the outer scalp and brain surface. They are a smoothed version of the corresponding surface that has been respectively shrunk or expanded a few millimeters inwards or outwards (3 mm).

Figure 16: Four views of the outer skin surface

- Option `-brainsurf surface_name`: that option writes out the final tessellated brain surface.
- Option `-shk_br_surf h_{shk} surface_name`: this option writes out the final tessellated brain surface that has been shrunk $h_{shk}$ mm inwards.
- Option `-LABEL`: this option enables the user to obtain a labeled volume into 5 classes: scalp, skull, $c_{sf}$, gray and white matter. First the whole algorithm is run and the surfaces corresponding to the brain, outer scalp, outer skull, inner skull are estimated. Based on the localization of these surfaces, we label the volume into brain, scalp, skull, and $c_{sf}$; then, the brain segment is divided into gray and white matter based on the intensity statistics calculated during the template deformation process (see section 3.3.3).

Figure 17: An example of a labeled volume
6 References


