ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 24 (2005) 261–280

Mindboggle: a scatterbrained approach to automate brain labeling

Arno Klein* and Joy Hirsch

fMRI Research Center, Columbia University, New York, USA

Received 14 November 2003; revised 16 September 2004; accepted 17 September 2004
Available online 24 November 2004

Mindboggle (http://www.binarybottle.com/mindboggle.html) is a fully automated, feature matching approach to label cortical structures and activity anatomically in human brain MRI data. This approach does not assume that the existence of component structures and their relative spatial relationship is preserved from brain to brain, but instead disassembles a labeled atlas and reassembles its pieces to match corresponding pieces in an unlabeled subject brain before labeling. Mindboggle: (1) converts linearly coregistered subject and atlas MRI data into sulcus pieces, (2) matches each atlas piece with a combination of subject pieces by minimizing a cost function, (3) transforms atlas label boundaries to the matching subject pieces, (4) warps atlas labels to their transformed boundaries, and (5) propagates labels to fill remaining gaps in a mask derived from the subject brain. We compared Mindboggle with four registration methods: linear registration, and nonlinear registration using SPM2, AIR, and ANIMAL. Automated labeling by all of the nonlinear methods was found to be at least comparable with linear registration. Mindboggle outperformed every other method, as measured by the agreement between overlapping atlas labels and manually assigned subject labels, with respect to the union or the intersection of voxels. After applying the same procedure that Mindboggle uses to fill a subject's segmented gray matter mask with labels (step 5), the results of the other methods improved. However, after performing a one-way ANOVA (and Tukey's honestly significant difference criterion) in a multiple comparison between the results obtained by the different methods, Mindboggle was still found to be the only nonlinear method whose labeling performance was significant better than that of linear registration or SPM2. Further advantages to Mindboggle include a high degree of robustness against image artifacts, poor image quality, and incomplete brain data. We tested the latter hypothesis by conducting all of the tests again, this time registering the atlas to an artificially lesioned version of itself, and found that Mindboggle was the only method whose performance did not degrade significantly as the lesion size increased.
© 2004 Elsevier Inc. All rights reserved.

Keywords: Mindboggle; Feature-matching; MRI; Image registration; Warping

Introduction

Current directions in neuroscience depend upon understanding the relationship between human brain structures and mental functions. It is common practice to associate functional imaging data with anatomical regions and to compare structural and functional data from corresponding regions across brains. Manually assigning anatomical labels from a standard atlas, however, requires expertise and is both tedious and time-consuming. Furthermore, labeling is particularly sensitive to primary sources of error that are independent of imaging hardware or software, such as morphological variation across brains (Jouandet et al., 1989; Le Goualher et al., 2000; Ono et al., 1990; Rademacher et al., 1993; Steinmetz and Seitz, 1991; Steinmetz et al., 1994; Thompson et al., 2000; Wright et al., 2002; Zilles et al., 1997, 2001) and inconsistency within the labelings of one labeler and between labelers (Caviness et al., 1996; Fiez et al., 2000; Lancaster et al., 2000; Towle et al., 2003). Ideally, we would like to automate labeling so that it is consistent, accurate, and efficient. Although automated methods exist that mitigate errors due to human inconsistency, the accuracy with which they handle the morphological variation from brain to brain has rarely been rigorously assessed. In this paper, we address the need for more accurate fully automated labeling tools by presenting our feature-based labeling method called Mindboggle (Klein and Hirsch, 2001, 2002, 2003, http://www.binarybottle.com/mindboggle.html) and compare Mindboggle's performance with that of several methods currently in use.

Each approach to labeling the anatomy of an imaged brain has its assumptions and weaknesses. A common assumption is that topography is preserved from brain to brain, where topography refers to the existence of component structures (such as sulci) and their relative spatial relationship along the cortical surface. This assumption may not be justifiable, as there may be missing or interrupted sulci in some brains and the sequence of the sulci may be different between brains (Ono et al., 1990), requiring cuts in one brain's surface to map onto another brain's surface. Linear coregistration of a brain and an atlas is the simplest method for assigning correspondences. Linear (affine) registration may involve up to 12 parameters representing translation, rotation, scaling, and shearing about the x, y, and z axes. Linear registration

* Corresponding author. fMRI Research Center, Columbia University, Neurological Institute-B1, 710 West 168th Street, New York, NY 10032, USA. Fax: +1 212 342 0851.
E-mail address: arno@binarybottle.com (A. Klein).
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2004.09.016

assumes a bijective mapping that preserves topography and because it is an operation performed on the whole brain volume, it cannot be expected to coregister independent points well. Linear coregistration therefore serves as a starting point for nonlinear registration, where brain-to-brain correspondences are determined in a non-uniform manner that may or may not be spatially continuous.

Talairach registration

The forerunner of modern nonlinear registration methods is a 13 parameter, piecewise linear registration method using the original Talairach coordinate referencing system (Talairach and Szikla, 1967) and the revision of this system (Talairach and Tournoux, 1988). This system prescribes orientation of a brain by registering two medial landmarks to a pair of points in the coordinate system (six parameters: three rotations, three translations), division of the brain by four orthogonal planes intersecting these two points, and uniform scaling of brain matter within the resulting 12 boxes so that the box dimensions equal those of a postmortem elderly woman's brain (seven scaling parameters about the two landmarks: one between them and six in anterior, posterior, inferior, superior, left, and right directions). The system is well suited to labeling regions proximal to these landmarks (Grachev et al., 1998), but assumes a bijective map that is simply a uniform scaling within each box and is discontinuous across these boxes. It therefore does not deal adequately with nonlinear morphological differences, especially when applied to the highly variable cortex (Grachev et al., 1999; Mandl et al., 2000; Roland et al., 1997; Xiong et al., 2000).

Nonlinear registration

Truly nonlinear registration is usually performed by warping (nonlinearly deforming) one brain image to appear similar to another. Although there is a vast array of nonlinear registration techniques (reviewed in Maintz, 1996–1997; Maintz and Viergever, 1998; Toga, 1999), much of present day work is based on the fundamental ideas of Broit (1981) and Bajcsy and Kovacic (1989). Traditionally, a brain image or mathematical representation of brain contours, surfaces, or topological features are treated as fluid or elastic bodies that are subjected to deforming forces balanced by regularizing or smoothing forces or constrained by a cost function (Christensen, 1999; Christensen et al., 1994, 1996; D'Agostino et al., 2002, 2003a,b; Davatzikos, 1996, 1998; Gaens et al., 1998; Gee, 1999; Gee et al., 1993, 1995; Guimond et al., 2001; Liu et al., 2003; Miller et al., 1993; Sandor and Leahy, 1997; Schormann and Zilles, 1998; Shen and Davatzikos, 2002; Thompson and Toga, 1996; Vaillant and Davatzikos, 1997). The smoothness constraint enforces continuity and preserves topology, and usually topography as well. The primary problem with relying solely on warping to nonlinearly register one brain to another is that without sufficient constraints, there are many ways to reshape a brain to look like another without regard for anatomical borders. Because point correspondence from one cortex to another is ill-defined, some degree of manual intervention is often used to initially assign corresponding points or curves about which deformations are to be performed (Bookstein, 1989; Chui et al., 1999; Evans et al., 1991; Hartkens et al., 2002; Johnson and Christensen, 2001; Magnotta et al., 2003). It is also questionable to assume that intervening points between two well-defined landmarks correspond to intervening points between matching landmarks. The point correspondence problem is simply revisited at a smaller scale; if different anatomical structures happen to exist between the corresponding pairs of landmarks in two brains, there may be no point-to-point correspondence. Moreover, the efficacy of many nonlinear registration strategies is difficult to assess, as they are usually demonstrated with restricted label sets or sparse landmarks and evaluated under artificial conditions or most commonly by visual inspection, where image correspondence is mistaken for anatomic correspondence (Crum et al., 2003; Rogelj et al., 2002). For the remaining sections, anatomic correspondence will be considered at the scale of gross morphological structures (primary gyrii and sulci); these structures are defined manually in each brain primarily by their relative position, curvature, and landmarks.

Intensity-based registration software

Some of the most popular software packages that perform nonlinear registration are based on image intensity, rather than on landmarks or features or parametric representations derived from image intensity: SPM2 (http://www.fil.ion.ucl.ac.uk/spm/spm2.html), AIR (http://bishopw.loni.ucla.edu/AIR5/index.html), and AutoReg's mritotal and ANIMAL (http://www.bic.mni.mcgill.ca/software). SPM2, a popular software suite for functional brain mapping analysis, includes a nonlinear registration algorithm that models deformations as a linear combination of (3-D discrete cosine transform) basis functions; parameters represent coefficients of the deformations in three orthogonal directions (Ashburner and Friston, 1999; Friston et al., 1995). The algorithm simultaneously minimizes the sum of squared differences between the voxel (volume element) intensities of two brain images, and the bending energies of the deformation fields (by estimating an inverse covariance matrix for the parameters). The nonlinear registration program has demonstrated fully automated, continuous, and detailed labeling of the entire cortical volume (Tzourio-Mazoyer et al., 2002). AIR's align_warp is a polynomial warping algorithm that also minimizes summed squared differences between voxel intensities of two brain images (Woods, 1999; Woods et al., 1998). ANIMAL (Collins et al., 1995) is essentially a local, smooth, piecewise implementation of the global linear algorithm from which it was derived (Collins et al., 1994), where the cross-correlation of intensity values is maximized between the atlas and brain volumes at increasing spatial resolutions.

Hybrid approaches

ANIMAL + sulci (Collins and Evans, 1999; Collins et al., 1998) and ANIMAL + INSECT (Collins and Evans, 1999; Collins et al., 1999) are hybrid extensions of ANIMAL that incorporate image-based features to improve registration. ANIMAL + INSECT incorporates tissue classification; a probabilistic atlas is transformed to the subject space, masking the tissue-classified data to segment regions of the MRI volume. Another method that has incorporated tissue-classification is octree spatial normalization (OSN, Kochunov et al., 2000). OSN nonlinearly registers tissue-classified anatomical templates (instead of voxel intensities) by fast cross-correlation in larger octants and centroid feature matching in

smaller octants; the original gray matter images are then warped according to the resulting deformation field. Parameter space warping (PSW) is another hybrid approach that uses local features to guide spherical harmonic basis functions that warp a complete parametric representation of the cortical surface (Meier and Fisher, 2002). All of these methods enforce continuous deformation and therefore make inherent assumptions about preserving topography. SPM2, AIR, ANIMAL, and most of the other nonlinear registration methods are based on voxel intensities and thus are sensitive to image quality and artifacts, and require template masks for cropped or lesioned regions.

Bayesian-based labeling

A subset of the above methods incorporates Bayesian statistics that stress reliability or symmetry of the solution to the nonlinear registration problem (Ashburner et al., 1997, 1999; Gee, 1999; Gee et al., 1995) and are currently under active research. Freesurfer is an alternative method that does not presently rely on warping, and uses an anisotropic nonstationary Markov random field (MRF) to model the spatial relationships between neighboring labeled structures (Fischl et al., 2002, 2004). The label of each voxel is determined using a Bayesian approach that incorporates the prior probabilities that (1) a given label is assigned to that voxel in a probabilistic atlas and that (2) a given label has a spatial relationship with its neighborhood of voxels after linear registration of a subject volume to an atlas space. Unlike feature-based approaches such as Mindboggle, due to computational constraints, spatial relationships are considered only within the neighborhood of six flanking voxels along the coordinate axes about a given voxel in voxel-based coordinates (Fischl et al., 2002) or within the neighborhood along the two principal directions on the surface of a sphere at each spherically transformed location (Fischl et al., 2004). The resulting cortical labeling is constrained to a surface representation of the cortex rather than extending through the brain volume.

Feature matching

In addition to linear and nonlinear procedures that register whole volumes, there have been attempts to use feature-based methods that match or identify sulci based on size (Jaume et al., 2002) or sulcus substructures based on point distribution models and/or relational graphs or combinations of features (Chui et al., 2001). Substructures include parametric surfaces (Le Goualher et al., 1998), sulcus basins (Lohmann and von Cramon, 2000; Lohmann and Yves von Cramon, 1998, and manually identified in Rettman et al., 2002), crest lines (Declerck et al., 1995), hull-projected skeletons (Caunce and Taylor, 1999), and 3-D sulcus skeletons (Lohmann, 1998; Mangin et al., 1995; Rivière et al., 2000). The relational graphs obtained by Rivière's aggregate of neural networks are analogous to those of Le Goualher, Lohmann, Mangin, and Declerck. A hybrid extension of Rivière's work adds warping to regularize the sulcus feature matching with intensity matching and emphasizes point-to-point correspondence and smoothness in its transformations (Cachier et al., 2001). A recent addition to this body of work parcellates the cortical surface by constructing a cortical mesh and filling the mesh with gyrus labels in three steps (Cachia et al., 2003a). First, the bottoms of identified pairs of parallel sulci are projected onto the cortical mesh. Then, a Voronoi diagram of the projected sulci provides "gyrus seeds" between the sulci. Finally, a second Voronoi diagram based on the gyrus seeds defines the boundaries between competing gyrii.

Techniques of evaluation

Most of the above methods were evaluated by their authors using only visual inspection and very few of the approaches were actually used to demonstrate cortical labeling. Some studies used a small set of landmarks to infer how well different brain volumes were matched. For example, Salmond et al. (2002) found that constrained registration in SPM99 (smaller number of nonlinear basis functions and smaller degree of regularization) resulted in an optimal colocalization of eight homologous landmarks in different brains. Woods et al. (1998) demonstrated that AIR's higher-dimensional warps decreased the average distance between corresponding sulcus landmarks in different brains. To evaluate OSN, Kochunov et al. (2000) calculated the spatial variance of individual sulcus tracings around their mean tracing for several sulci. ANIMAL achieved an average result of 87% for basal ganglia structures according to overlap with manual labels (Collins et al., 1995, and see "Evaluation of Mindboggle"), but typically obtains results of 40% to 50% for cortical structures (Collins et al., 1999). ANIMAL + sulci was evaluated by computing the root-mean-squared minimum distances between sulcus points in a target brain and nearest neighbors in 16 sulci in each of 10 transformed brains (Collins and Evans, 1999). ANIMAL + INSECT was evaluated by computing the average Dice similarity coefficient (0.657 ± 0.037, Dice, 1945) between automated and manual labels for the prefrontal cortices of 20 brains (Collins et al., 1999b). Freesurfer (Fischl et al., 2004) was evaluated by comparing the surface areas of each automatically and manually labeled region, and by calculating the overall point-by-point agreement between their labels (see ''Atlas labels and a probabilistic database''). A comparison with this method would require representing the cortex as a 2-D surface. As for the feature-based methods, comparing results is impossible when features are identified but the cortical volume is not labeled. To evaluate Mindboggle, we compared its output with output from linear (12-parameter affine) registration alone and in combination with the most popular software packages to nonlinearly register brains: SPM2, AIR, and ANIMAL.

Mindboggle

Our method, Mindboggle, labels brains based on similarities between corresponding sulcus pieces and is therefore a feature-based method. However, Mindboggle differs from the above-mentioned feature-based approaches in the way that it defines and extracts sulcus substructures, and in that it does not ascribe any anatomical significance to the substructures but merely uses them to set up correspondences between two brains for further processing. Furthermore, as a feature-based method, Mindboggle compares individual anatomical structures derived from images as opposed to the image intensities themselves, and is therefore quite robust to poor image quality, intensity inhomogeneity artifacts, and missing regions. It also avoids some of the continuity assumptions of the standard nonlinear registration algorithms by allowing for secondary and tertiary sulci to exist in one brain but not in another, and therefore has the potential to more flexibly deal with greater local morphological variability due to intersubject differences and/or pathology. Continuity is assumed, however, at the level of

primary sulci (defined by the atlas), since a matching structure in the subject brain is sought for each primary sulcus in the atlas, although even this assumption may be relaxed by constraining unlikely matches (see below). Mindboggle is also deterministic, it does not require any random numbers or seed values and gives identical results for a given input. This combination of qualities makes Mindboggle a widely adaptable system for automated, nonlinear labeling of human brain images.

Materials and methods

Mindboggle is a software package written in Matlab (version 6, release 13, with the Image Processing Toolbox, The Mathworks Inc., USA) and has been tested on different models of desktop and laptop PCs and workstations running different distributions of Linux and Unix. The general system requirements are the basic requirements of the Matlab environment. The system used to conduct the following tests consists of a Sun Fire 6800 with a 750-MHz, 64-bit processor with 32 GB memory running Sun Solaris 8.

In the following subsections, we first describe (1) atlas selection, (2) MR image acquisition, and (3) image processing. We then explain how Mindboggle (4) prepares subject sulcus pieces, (5) matches each atlas piece with a combination of subject pieces, (6) transforms atlas label boundaries to the matching subject pieces, and (7) warps atlas labels to their transformed boundaries and propagates these labels to fill a mask derived from the subject brain. Mindboggle optionally (8) resets atlas-specific, feature-derived planar boundaries for frontal and temporal poles as well as the occipital lobes, if the atlas itself is labeled using these planar boundaries. We evaluate Mindboggle by (9) comparing its results with those obtained by one linear and three nonlinear methods. Fig. 1 portrays Mindboggle's processing pipeline, and all subsequent figures follow a single subject through the pipeline. The final step in the flowchart, labeling activity data, is discussed in Discussion.

Mindboggle uses a single set of default parameters, described in the relevant sections below. Since it is impossible to theoretically determine an optimal parameter set for this feature matching approach, those parameters that were not determined theoretically

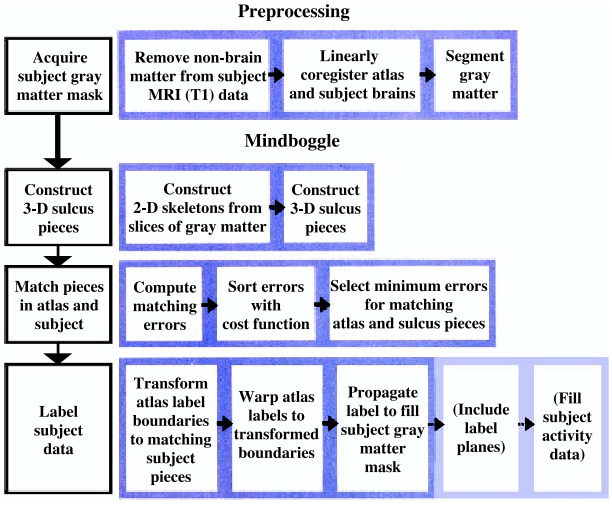

Fig. 1. Mindboggle flowchart.

were determined empirically using different atlases and subjects than were used in the present study. We systematically searched the parameter space associated with the matching, warping and label filling stages (5 and 7 above), and selected values for these parameters that produced minimum labeling errors according to the measures described below.

Atlas selection

Any atlas with anatomical labels and corresponding T1-weighted volume may be used with Mindboggle (atlas reviews: Mazziotta, 1997; Toga et al., 1998: http://www.loni.ucla.edu/~thompson/whole_atlas.html). A desirable format for labeling MRI data should (1) be of high resolution, (2) be acquired in the same manner as the subject data, (3) include most of the brain, (4) be representative of the subpopulation under study, and (5) should have a high spatial correlation between parcellation boundaries and sulci. The most commonly used atlas, the Talairach Atlas (Talairach and Tournoux, 1988), fulfills none of the first four criteria, since (1) the coarse 27 horizontal figures represent (2) postmortem sections of (3) one cerebral hemisphere of (4) a 60-year-old woman. Electronic versions of the atlas have corrected for inconsistencies (Nowinski et al., 1997), but mistakes still remain (Maldjian et al., in press). We have used several brain atlases with Mindboggle, including the Harvard Brain Atlas (Kikinis et al., 1996) and the MNI1 Atlas, Montreal Neurological Institute's single-subject atlas (Tzourio-Mazoyer et al., 2002).

To convert an atlas to one that may be used directly by Mindboggle, the labeled T1 volume is preprocessed in the same manner as a subject T1 volume (see "Image processing before applying Mindboggle"), and is converted to (primarily gyrus) labels and 3-dimensional label boundaries. The label boundaries are the subset of labeled voxels that have two different labels within each labeled voxel's $3 \times 3 \times 3$ neighborhood.

To validate Mindboggle, we compared automated labels with manual labels assigned to each subject brain. We increased the probability of label consistency between the atlas and manually labeled subject brains by having the subject brains manually labeled by the same labeler using the same parcellation scheme. We then randomly selected one of the subject brains to further process and serve as the atlas to perform automated labeling of these same brains. Ten subject brains were parcellated by Jason Tourville using a software tool developed by Satrajit Ghosh at the Department of Cognitive and Neural Systems, Boston University. The labeling scheme is modified from that implemented in Cardviews software (Caviness et al., 1996; Tourville and Guenther, 2003).

In addition to the labeling atlas, we have selected for coregistration the MNI152 Atlas, MNI's intensity average of 152 T1-weighted volumes in Talairach space (Evans et al., 1992). All registration and labeling is performed in this space (resolution of $1 \times 1 \times 1$ mm and dimensions of $181 \times 217 \times 181$ voxels). We also use the MNI1 Atlas for cropping the cerebellum, and made use of both the MNI1 and Harvard Brain Atlases to initially derive parameter settings for the algorithms.

Image acquisition

The T1-weighted MRI data were acquired at the MGH/MIT/HMS Athinoula A. Martinos Center for Biomedical Imaging using a 3T Siemens scanner and standard head coil (TE: 2.9 ms, TR: 6.6

s, flip angle: 8°). The in-plane resolution was approximately 1 × 1 mm, the slice thickness was 1.33 mm, and the dimensions and field of view were 256 × 256 voxels. The subjects in this study are four men and six women between the ages of 22 and 29 years old (mean of 25.3). All are right-handed. Using SPM2 software, the data were bias-corrected,[1] affine-registered to the MNI152 template, and segmented. The gray-white matter boundaries of the segmentations appear rather crude, but as described below, the extent of gray matter plays an insignificant role in the feature matching and labeling steps.

Image processing before applying Mindboggle

Mindboggle calls on third-party software to perform four preliminary steps on a subject brain image: (1) crop non-brain matter, (2) linearly coregister with the MNI152 Atlas, (3) segment cortical gray matter, and (4) crop subcortical structures, cerebellum, and exposed brain surface. Mindboggle has been used with different software suites and presently defaults to (1) BET (Smith, 2000), (2) FLIRT (Jenkinson and Smith, 2001), (3) FAST (Zhang et al., 2001), and (4) avwmaths, all developed by Oxford University's FMRIB group (http://www.fmrib.ox.ac.uk/fsl). FLIRT uses correlation ratios to find optimal 12-parameter affine transformations and registers using trilinear interpolation (linear interpolation of points within, for example, a cube given values at the vertices of the cube).

To ensure that only the cortical gray matter on the banks of sulci, not on the exposed surface of the cortex, contributes to the construction of 3-D sulcus pieces (see below), the exposed surface is cropped by erosion (3 voxels deep). The subcortex and cerebellum are also cropped with a mask constructed from the MNI1 and MNI152 Atlases. In the present study, T1-weighted MRI data were already segmented using SPM2 (see "Image acquisition" above), so the third step above (FAST segmentation) was unnecessary.

Preparation of 3-D sulcus pieces

The simplest representation of an atlas or atlas-labeled subject volume is a set of boundaries dividing adjacent anatomical regions. Sulci are prominent structures in MR images and usually serve as anatomical boundaries. Skeletal surfaces can capture the essential geometry of sulci and provide a sparse representation of an intensity volume, lessening the computational burden of comparing atlas and subject brains (examples of skeletonization procedures include Malandain and Fernández-Vidal, 1998, and those mentioned above in "Feature matching"). More importantly, a geometric representation of the cortex such as a skeleton may be easily fragmented, enabling combinatoric comparisons and matching (see below). Piecewise matching of component structures, unlike smooth deformations of cortex, can in principle resolve the above-mentioned topographical deviations encountered between brains, though this assertion has not yet been tested.

[1] From the SPM2 website: "The [bias] correction model is non-parametric, and is based on minimizing a function related to the entropy of the image intensity histogram…the cost function is based on histograms of the original intensities, but includes a correction so that the solution is not biased towards a uniform scaling of all zeros."

To create this skeletal representation (Fig. 2), Mindboggle performs three two-dimensional morphological operations to embedded gray matter in the horizontal slices of a T1 volume. For the first two operations, Matlab's bwmorph.m applies a majority-surround 2-D filter to remove isolated pixels in every 3 × 3 neighborhood, and then applies successive thinning (erosion that preserves 1-D extent and Euler number) to reduce binarized slices to one-pixel-width skeletons. For the third operation, each slice of skeleton is segmented into 2-D sets of 8-connected pixels (pixels adjacent in the horizontal, vertical, or diagonal direction) using Matlab's bwlabel.m. The second (thinning) operation iteratively replaces non-skeletal with skeletal arrangements of pixels in every 3 × 3 neighborhood. Non-skeletal arrangements acted upon by the thinning operation (48 of the 512 possibilities) are defined by a center pixel with three to five surrounding pixels adjacent to each other (sharing an edge), or by a center and two adjacent pixels forming a right angle with up to three adjacent corner pixels. The thinning operation is repeated until no change takes place, resulting in a 2-D skeleton that contains any arrangement of pixels except that no four adjacent pixels form a square and no pixel branches into only one right angle without at least one pixel diagonal to it. The result of these three morphological operations is a stack of 2-D cross-sections of all embedded sulci and the interhemispheric plane. These three steps are performed in two dimensions in order to facilitate 3-D fragmentation of the skeleton (see below).

The first division of the brain in 3-D is to partition the skeleton into left and right hemispheres. A sagittal plane is positioned in the general orientation of the interhemispheric plane. A type of unsupervised neural network called a Self-Organizing Map (SOM, Kohonen, 1997) was modified to warp the plane in one dimension, along its normal vectors. The target of the warp is the set of nearest points in a medial slab of the skeleton. The warped plane acts as a continuous representation of the interhemispheric plane and divides the skeleton into two hemispheres of sulcus points. The SOM acts on points of the plane as follows. Each point, P_i, is pulled at time t toward its nearest point W_i in the target (nearest in the original space, $t = 0$) according to the distance between them and a neighborhood function h. The neighborhood function consists of a learning rate parameter that determines the fraction of this distance that P_i will be translated toward W_i at each iteration (set to 0.1) and a neighborhood centered on P_i whose points also converge toward W_i according to a function of their distance from P_i (in this case a Gaussian function of $\sigma = 10$ mm, clipped at a radius of 20 mm):

$$P_i(t+1) = P_i(t) + h(P,t)|W_i - P_i(t)|$$

(Self-Organizing Map equation)

To match a finer set of corresponding structures than hemispheres between brains, we had to redefine the skeleton as a collection of 3-D "sulcus pieces" (Fig. 3), performed in three steps. In the first two independent steps, points are clustered across horizontal slices from the bottommost slice up and from the topmost slice down. For the third step, a finer clustering is achieved by intersecting the resulting clusters from the first two steps. The first and second steps are conducted as follows. Any set of 8-connected pixels in the first slice is considered a separate piece. In subsequent slices, each pixel is grouped with the nearest piece of the previous slice. The maximum in-plane distance for grouping a pixel to its nearest piece is set to three pixels, found to

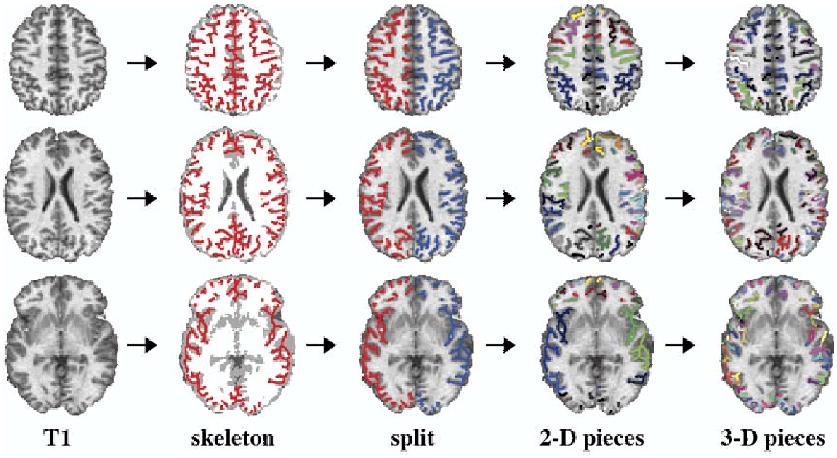

Fig. 2. Preparing 3-D sulcus pieces from subject MRI data. Mindboggle input consists of either T1-weighted MRI data or gray matter masks segmented from the MRI data. Third-party software skull-strips, linearly coregisters, and segments non-white matter from MRI data (see text). Mindboggle then thins the non-white matter to a skeleton for each slice (1st arrow). The sulcus skeleton is split into left and right hemispheres by a plane that is warped to the inter-hemispheric plane (2nd arrow). Each slice of the skeleton is segmented into 2-D pieces of contiguous sets of pixels (3rd arrow). Finally, 3-D pieces are constructed from the 2-D cross-sections (4th arrow), shown here in cross-section and in 3-D in Fig. 3. A single horizontal slice of each brain volume is shown for clarity with arbitrary colors (frontal lobe toward the top, left hemisphere on the left).

be roughly half the average width of horizontal cross-sections of segmented sulcus gray matter in preliminary tests with pilot subjects. Three pixels would therefore approximate the average in-plane distance between skeletons constructed from two identical sulcus cross-sections offset by half of its width. Pixels without a nearby piece (within three in-plane pixels) are grouped with 8-connected pixels of the same slice that are similarly ungrouped; this new group defines a new piece.

The number of pieces should be a compromise allowing for computational ease yet accurately reflecting the geometric and label resolution of the problem. To ensure that there are a reasonable number of discernibly shaped pieces, small sulcus pieces (less than 50 voxels) are grouped with nearby pieces in an extended neighborhood (within five voxels); if after this step there are any small, isolated pieces remaining, they are eliminated to reduce computational demands on subsequent steps. Setting the minimum size of a piece to 50 voxels was empirically found to result in a consistent number of pieces across ten pilot subjects; this number is a compromise between two competing concerns: minimizing the number of pieces to reduce processing time versus maximizing the number of pieces to generate high resolution label boundaries. Defining a nearby piece to be within five voxels allows for a distance along a right angle (in- and out-of-plane) of three and four pixels, where three defines the maximum separation between points of the same 2-D piece and four defines the minimum separation between two different 2-D pieces.

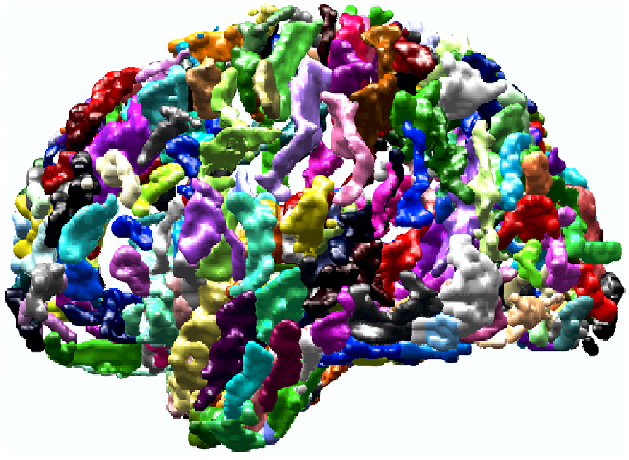

Fig. 3. 3-D sulcus pieces. Mindboggle constructs 3-D sulcus pieces from subject MRI data (Fig. 2), rendered here with surfaces and arbitrary colors (left view). The pieces do not have to correspond accurately to anatomical divisions because Mindboggle uses them not to label a brain directly, but to compute a transformation to deform atlas label boundaries.

The above three parameters for constructing pieces have not been evaluated for coregistration spaces of different size or resolution. However, they should scale according to the expected average piece dimensions. For example, three in-plane pixels defining adjacency and five voxels defining nearness would instead be based on half the expected piece width as above, and 50 voxels defining minimum size could instead be the size of pieces two standard deviations below the mean size of pieces after preliminary construction.

The resulting pieces may be of appropriate size and number, but in order to define compact shapes with high surface-to-volume ratios, Mindboggle performs two more steps to fragment and then regroup clusters into their final pieces. First, a k-means algorithm fragments each sulcus piece into smaller pieces, using as its initial means a set of 27 points distributed through the sulcus piece's bounding box: one in the center, one per corner (8), one at the middle of each edge (12), and one at the center of each plane (6). A second algorithm recombines each pair of resulting clusters if the two clusters share extensive borders. ("Extensive" was defined as a border-to-surface voxel ratio of at least one-tenth, in which a border voxel neighbors both clusters and a surface voxel has fewer than six neighbors, where the neighborhood consists of the six adjacent voxels, one to each side. This ratio is equivalent to the area of a circle, formed by the intersection of two spheres of equal radius, divided by their surface area, where the distance between the two spheres was chosen to be equal to 1.2 times the radius). For example,

a sphere that happens to have been fragmented by the k-means algorithm would recombine after applying the second algorithm, whereas a dumbbell split in the center would remain fragmented. Using these clustering procedures, the left and right hemispheres of the 10 subjects automatically divided into the same number of pieces on average (left hemisphere: μ = 165 pieces, σ = 13.5; right hemisphere: μ = 165 pieces, σ = 12.3).

Matching pieces

After Mindboggle has deconstructed a subject brain image into 3-D sulcus pieces, it searches for matching pieces in the atlas (left half of Fig. 4). As any pair of brain images will not fragment in the same manner, comparisons should be made between combinations of their pieces. However, matching combinations of atlas pieces with combinations of subject pieces would lead to a combinatorial explosion. The solution space can be reduced significantly by introducing the following constraints that also allow for some combinatoric matching: (1) each combination of subject pieces may contain a maximum of three pieces, (2) the mean location of each piece must be within 20 voxels of the mean of at least one of the other pieces in the combination (four times the maximum allowable distance between elements of the same piece, defined above), and (3) each combination may be compared with only one atlas piece at a time. The problem of identifying all of the most reasonable tentative matches for each atlas piece then reduces to one that may be computed much more rapidly—in a fraction of a second for the brain images in this study. Selecting the best of these tentative matches for each atlas piece is still not trivial, as they may contain subject pieces already matched with other atlas pieces, while some subject pieces may not find a match. These deviations from complete correspondence are not necessarily a disadvantage; sulcus structures may exist in one brain and not in another, and interruptions and branching of sulcus structures require discontinuous labeling.

To select the best set of matches for each atlas piece, Mindboggle orders the tentative matches by a sum of four weighted differences serving as a cost function. Three of the four costs are differences between quantities computed from subject and atlas skeleton pieces: the number of points (N), number of subvolumes (V), and mean position (P). The fourth cost is the degree of non-overlap (O) of subvolumes occupied by the pieces. Each of these costs is normalized by the mean value across all tentative matches. Each term of the cost function is therefore dimensionless, but the cost function has not been evaluated for coregistration spaces of different size or resolution. The relative degree of subsampling of the coregistration space (by breaking up the space into subvolumes) should account for such differences. Mindboggle uses a subvolume size of five times that of a voxel (5 mm^3 boxes = 5 $voxel^3$). The number of subvolumes occupied by a piece is computed by tallying the number of boxes containing at least one sulcus piece point. The subvolume dimensions are set large enough to defray computational cost and to increase the probability of overlap between tentative matches. If each subvolume was instead equal to one voxel, then two tentative matches with exactly the same distribution of points but offset from one another by a single voxel could result in no overlap. As with the numbers of points and mean positions, the numbers of subvolumes are precomputed for each piece before combinations are formed. Non-overlap of two pieces, P1 and P2, is equal to the fraction of subvolumes of P1 that do not overlap P2 added to the fraction of subvolumes of P2 that do not overlap P1. We avoided using non-intersection divided by the union to distinguish between cases such as the following: (a) P1 = 4 subvolumes, P2 = 5, with intersection = 3, and case (b) P1 = 4, P2 = 2, with intersection = 2. Our non-overlap measure would result in 1/4 + 2/5 = 13/20 for case (a) and 2/4 + 0/2 = 1/2 for

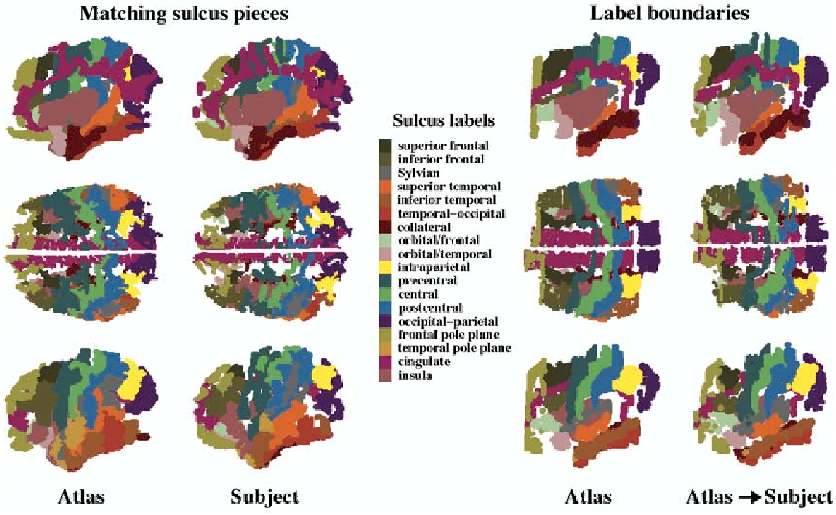

Fig. 4. Matching 3-D sulcus pieces and transforming atlas label boundaries. Mindboggle matches each atlas piece with a combination of subject pieces by minimizing a cost function. The atlas piece assigns its sulcus label to the matching subject pieces. Each colored region in the left half of the figure represents multiple pieces with a single sulcus label. Each atlas piece is paired with a patch of the atlas label boundaries (that correspond for the most part with sulcus surfaces). Mindboggle translates each atlas patch to the matching subject pieces, resulting in piece-wise linearly deformed atlas label boundaries in the coregistered subject brain. Each colored region in the right half of the figure represents multiple patches with a single sulcus label. All of the images in this figure are left views of atlas and subject sulcus labels, with a medial view of the right hemisphere at the top, an overhead view in the middle row, and a lateral view of the left hemisphere at the bottom.

case (b), whereas non-intersection over union would lead to 1/2 for both cases. The default weights w_N, w_V, w_P, and w_O were empirically determined by examining resulting matches between the MNI1 and Harvard Brain Atlases, and have been set to 1, 1, 10, and 10, respectively. The ratio between the first two and last two weights is rather robust, giving comparable results for ratios of 1:2 to less than 1:20, but the cost rises dramatically if any term is omitted. Matching is also robust to inaccurate or inconsistent piece construction because slight variations in the size, volume, position, and shape of the pieces have an insignificant impact on the cost function (after each unweighted cost term is normalized).

$$E_M = w_N N + w_V V + w_P P + w_O O \text{ (matching cost function)}$$

To search for a low cost set of matches, Mindboggle constructs a matrix with each column containing the tentative matches (subject piece combinations) for each atlas piece, sorted by cost. The least cost combinations (top row) are selected as winning matches for all the atlas pieces. For each subsequent row, unused subject pieces are appended to the above matches. For duplicate pieces within a row, the one with the least cost is chosen. The matching procedure is terminated when all pieces have been appended or when the desired number of rows has been searched. A threshold may be set when computing the cost function to reduce the chance of unlikely matches when, for example, a subject brain is missing a primary sulcus (no threshold was set in this study).

Transforming and propagating labels

Even after Mindboggle has matched subject pieces to each atlas piece, labeling the subject space between these pieces is not trivial. The pieces constitute a miniscule fraction of the brain volume and have a limited extent that is not sufficient to divide the volume. Even worse is the fact that the skeletal pieces do not necessarily follow the original, manually defined atlas label boundaries (see "Atlas selection"). For these reasons, Mindboggle does not rely on the atlas pieces to label the brain directly, but instead relies on the original, manually defined atlas label boundaries (corresponding roughly to sulcus surfaces) nearest to these pieces.

There are three steps to labeling the subject brain after the above matching has been performed: (1) transforming atlas label boundaries to the subject brain, (2) warping atlas labels to these new boundaries, and (3) propagating labels to fill the remaining gaps. First, the atlas label boundary surfaces are transformed to the subject brain by fragmenting them into patches and translating

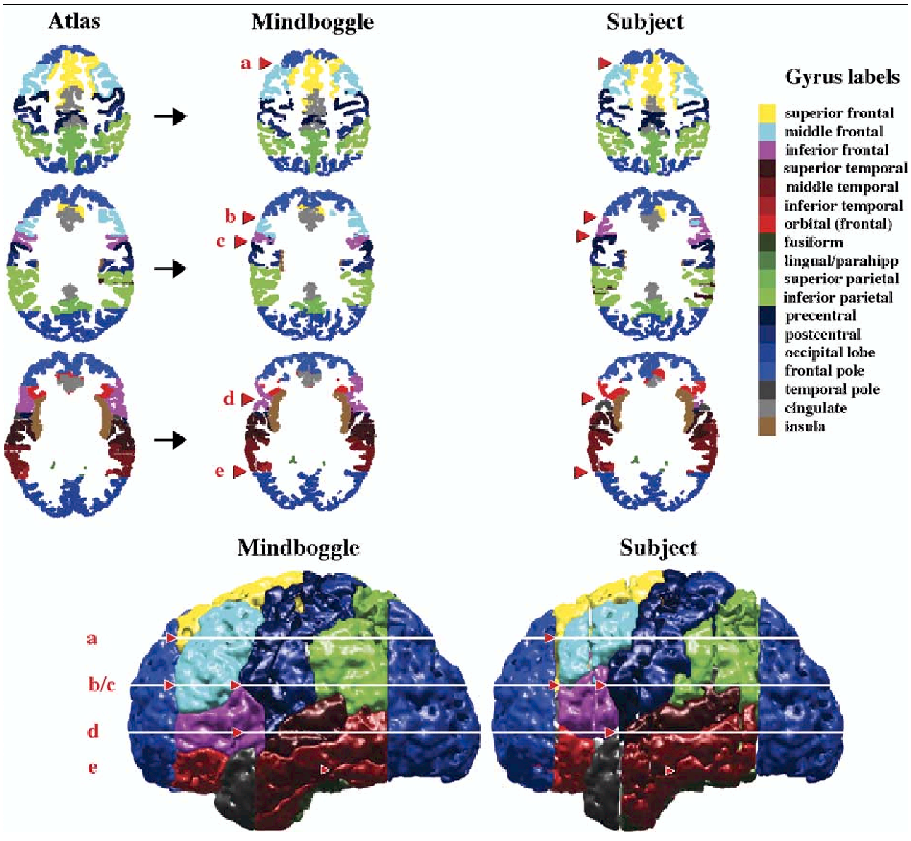

Fig. 5. Warping atlas labels to their transformed boundaries. Atlas gyrus, lobe, and pole labels are warped in 3-D (with a modified Self-Organizing Map) to smoothly line the atlas label boundaries that have been broken up into patches and transformed to the subject sulcus pieces. These labels are then propagated in 3-D to fill a mask of the subject's gray matter. Shown from left to right is the atlas, the atlas after label warping and propagation, and the manual labels of the target subject brain for evaluation. Red tags on the slices (oriented with the frontal lobe toward the top, left hemisphere on the left) and the surface rendered volumes (left views) denote the largest discrepancies between Mindboggle-assigned and manually assigned labels for this subject: (a) the plane defining the frontal pole, (b) the boundary between middle and inferior frontal gyrii, (c) the anterior boundary of the precentral gyrus, (d) the boundary of the temporal lobe, and (e) the boundary between middle and inferior temporal gyrii.

them (right half of Fig. 4). Each patch is the set of atlas label boundary points nearest to an atlas piece. Each patch is translated by the difference between the mean positions of the matching subject and atlas pieces. Because the mean position of a piece is robust to variations of its shape, Mindboggle's labeling algorithm is robust to inaccurate or inconsistent skeletal piece construction.

Second, the original atlas labels (enclosed by the atlas label boundaries and corresponding primarily to gyrii) are warped to the translated boundaries to apply a smooth coat of labels to these translated boundaries (Fig. 5). For each translated boundary point (voxel), the atlas label nearest to the point's original position translates the entire distance to the point. The label's neighbors (within a radius of 5 mm to reduce the number of computations) translate different fractions of this distance toward the point, with the fraction defined by the (Gaussian, σ = 1 mm) function of their original distances from the label. All relevant quantities other than distances between matching pieces are precomputed to increase computational speed.

Third, after warping, labels are cropped by a mask of the subject's gray matter. Each unlabeled voxel in the mask is assigned the majority label in a new neighborhood about the voxel ($5 \times 5 \times 5$ voxels, a neighborhood size empirically observed to give the best results in preliminary tests with 10 pilot brains). Ties are broken by selecting the label with the lowest index rather than by random selection to increase the algorithm's speed. The last step is repeated several times to label all remaining unlabeled voxels.[2]

Resetting planar boundaries

Many atlas-labeling schemes set planar boundaries between some adjacent regions. Though these planes do not follow true anatomical boundaries, they can be useful in separating regions where there is no clear or consistent boundary or no well-defined topographical landmarks, such as the juncture of the occipital, temporal, and parietal lobes.

The manually assigned label boundaries in our 10 subject brains incorporate planar boundaries in addition to label boundaries that follow the appropriate sulci. These planar boundaries constrain the frontal and temporal poles and occipital lobes. The plane marking the posterior boundary of the temporal pole is set where a clearly visible white matter tract connects the frontal and temporal lobes. The plane marking the posterior boundary of the lateral surface of the frontal pole is set at the most anterior point of the anterior horizontal ramus of the Sylvian Fissure, and for the medial surface is set at the anterior tip of the cingulate gyrus. The plane marking the anterior boundary of the lateral surface of the occipital cortex is set at the "opercularization" of the intraparietal sulcus (where it becomes highly curved and often splits into two sulci), and for the medial surface is set at the most anterior point of the parietal-occipital fissure (Tourville and Guenther, 2003).

Mindboggle may optionally refine the label boundaries of the temporal poles and the lateral portions of the frontal poles and occipital lobes (at least 15 mm (voxels) from the interhemispheric plane) to better correspond with the above definitions. This postlabeling step is important because Mindboggle relies on structural similarities to match and label, not on contrived planar boundaries; this step is optional depending on whether planar boundaries are included in the label definitions. Mindboggle will automatically set the temporal pole plane at the most anterior pair of adjacent voxels labeled as superior temporal gyrus and frontal lobe. It will set the lateral frontal pole plane at the most anterior voxel of the transformed sulcus boundary between the frontal-orbital and temporal regions. It will set the lateral occipital lobe plane at the most anterior voxel with an occipital lobe label; this plane also marks the anterior boundary for the lower portion of the medial occipital lobe, below the most inferior voxel with abutting occipital lobe and inferior parietal lobe labels. When Mindboggle resets these planar boundaries, some voxels become unlabeled. These are relabeled by the same label propagation step as above (assign each unlabeled voxel the majority label in its neighborhood).

Evaluation of Mindboggle

We evaluated the labels that Mindboggle assigned to each subject by defining and applying five measures that quantify the degree of similarity and dissimilarity between manually and automatically assigned labels: overlap, mask overlap, and filled mask label agreements, and type I and type II errors. To put these numbers in some context, we used these measures to also evaluate linear registration and linear registration followed by the three other nonlinear methods: SPM2, AIR, and ANIMAL.

We selected label agreements as evaluation measures because they are intuitive and may be applied to whole brain volumes, slices, or regions of interest. The above methods were used to automatically register the entire set of voxels with automatically assigned atlas labels, A, to the entire set of voxels with manually assigned labels, M, for each subject. Comparisons were then made between the two sets of voxels for each label of index i corresponding to a particular region of interest. The total number of voxels (volume) with a given label assigned either automatically or manually is $|A_i \cup M_I|$, corresponding to the union of label sets A_i and M_i. The number of voxels that have been assigned this label both automatically and manually is $|A_i \cap M_i|$, corresponding to the intersection of label sets A_i and M_i. The overlap between label sets A_i and M_i is defined as the volume of intersection divided by the volume of union (Gee et al., 1993; Jaccard, 1912), and may be extended to multiple labeled regions by summing over a set of labels of index i:

$$\text{Overlap} = \frac{\sum_i |A_i \cap M_i|}{\sum_i |A_i \cup M_i|}$$

The mask overlap penalizes for discrepancies in overlap but not in size. Mask overlap quantifies a comparison between label sets A_i and M_i only within the set M_i, here the labels in a subject's gray matter mask. The measure thus normalizes by $|M_i|$, the number of voxels manually assigned that label (Collins et al., 1995), and again may be extended to multiple labeled regions by summing over a set of labels of index i:

$$\text{Mask overlap} = \frac{\sum_i |A_i \cap M_i|}{\sum_i |M_i|}$$

For the third label agreement measure, filled mask, we filled a subject's gray matter mask with automated labels, A^F, using

[2] A similar procedure to this iterative propagation of majority labels through successive neighborhoods was discussed in the context of Markov random field models by Besag (1986).

Mindboggle's label propagation algorithm. This consists of assigning each unlabeled voxel within the mask the majority label in its (5 × 5 × 5) neighborhood in multiple steps.

$$\text{Filled mask} = \frac{\sum_i |A_i^F \cap M_i|}{\sum_i |M_i|}$$

To refine our evaluation, we also compute type I and type II statistical errors. Both of these errors assume that the manual labels are the correct labels. A type I error for a given region refers to the condition where incorrect labels are assigned to voxels within the region (e.g., voxels in the superior frontal gyrus are labeled as middle frontal gyrus). It is computed as the volume of a manually labeled region, $|M_I|$ that is outside the corresponding automatically labeled region, $|A_I|$, divided by the volume of the manually labeled region. Summed over a set of labels or regions of index i:

$$\text{Type I error} = \frac{\sum_i |M_i/A_i|}{\sum_i |M_i|}$$

where M_i/A_i is the set (theoretic complement) of elements in M_i but not in A_i. A type II error refers to the condition where a region's label is assigned to other regions (e.g., voxels outside of the superior frontal gyrus are labeled as superior frontal gyrus). It is computed as the volume of an automatically labeled region, $|A_i|$, that is outside the corresponding manually labeled region, $|M_i|$, divided by the volume of the automatically labeled region. Summed over a set of labels or regions of index i:

$$\text{Type II error} = \frac{\sum_i |A_i/M_i|}{\sum_i |A_i|}$$

Both the type I and type II errors can range from zero to one, so that perfect overlap between automatically and manually labeled voxels for a given label results in zero for that label; if the labeled voxels were completely disjoint, then the errors for that region would equal one.

Cortical gray matter in each of the brains was manually labeled in the same manner as the atlas (see "Atlas selection" and "Resetting planar boundaries" above). Of the 96 labels (Caviness et al., 1996), 74 were selected and merged to give 36 labels (18 per hemisphere): superior, middle, and inferior frontal and temporal gyrii, frontal and temporal poles, pre- and postcentral gyrii, superior and inferior parietal lobules, occipital lobe, fusiform, lingual/parahippocampal, and orbital (frontal) gyrii, insula, and cingulate gyrus. This gross parcellation of the cortex was chosen to include only distinct anatomical boundaries that are relatively consistent across brains. It would be difficult to over-emphasize this latter point, as brain researchers are familiar with different atlases with different anatomy and inconsistent labeling schemes. Subcortical structures, though less variable in morphology across brains than cortical structures, were not found to consistently segment well in the preprocessing stage, and were not included in the labeling. Mindboggle, SPM2, AIR, and ANIMAL were used to label the same subject volumes with the same 36 labels to compare with the manual labels.

Before performing the five labeling methods under comparison, the subject images were skull-stripped, then linearly registered to MNI152 space using a 12-parameter affine transformation (see "Image processing before applying Mindboggle"). The linear transform was also used to register the manually labeled voxels to MNI152 space.

In the linear registration condition, we calculated the agreement between the coregistered atlas and manual labels. We used the nonlinear registration methods SPM2, AIR, and ANIMAL to compute nonlinear transforms from the atlas T1-weighted image to each subject T1 image and applied the transforms to the atlas labels with nearest-neighbor interpolation (so as not to change the label values). For SPM2's "Normalisation," each subject image served as the template to which the atlas image was registered; a template bounding box was used with a voxel size of 1 × 1 × 1 mm. ANIMAL and AIR were used to transform each subject image to the atlas image and then an inverse of this transform was used to register atlas labels to the subject labels for comparison. The inverse transform was used because this is the way that ANIMAL was originally evaluated (Collins et al., 1995) and because AIR was found not to register as well when computing the forward transform. ANIMAL was invoked by AutoReg's "mritotal," a multi-scale method that fits a subject image to an atlas at finer resolutions, by successively blurring the two less and less (16, 8, 4, and 2 mm FWHM Gaussian kernels). AIR (v5.25) registered each subject to the atlas linearly with "alignlinear" then nonlinearly with a 5th degree polynomial transform computed by "align_warp."

In early tests, nonlinear registration of an atlas brain image to a subject brain image that did not span the same extent of the brain resulted in very poor registration and labeling by SPM2, AIR, and ANIMAL (but not Mindboggle). We therefore used whole brain acquisitions for these tests. But even with whole brain acquisitions, none of the methods result in perfect overlap of gray matter between the atlas and subject brains after warping; for our nine subjects, the mean overlap for gray matter equaled 36% (SPM2), 36% (AIR), and 40% (ANIMAL). Mindboggle warps labels to coat transformed boundaries, not to deform gray matter shapes; after warping, Mindboggle labels all voxels in the subject's gray matter mask and none outside the mask, so the resulting overlap is 100%. But because of Mindboggle's application of the gray matter mask, it is unreasonable to directly compare its results with those of the other methods under comparison. In order to compare the other four methods with Mindboggle, we first used the subject gray matter masks to mask the labeled volumes (mask overlap in Tables 1a and 1b and Fig. 6), then applied the same label propagation step that Mindboggle uses to the labels assigned by each of the other methods (filled mask in Tables 1a and 1b and Fig. 6). The percent correct of intersecting voxels for AIR, SPM, and ANIMAL were within one standard deviation of the values obtained by Crivello (Table 3 in Crivello et al., 2002).

Finally, to test the hypothesis that Mindboggle, a feature-based method, should be more robust to incomplete data than intensity-based methods, we used each of the methods (using default parameters) to register the labeled T1-weighted atlas to an artificially lesioned version of itself (or a lesioned version of its segmented image in the case of Mindboggle). Registered labels were then compared with lesioned labels. The self-labeling approach was taken to control for morphological variability and to directly compare methods without applying a label-filling step. Artificial lesions were used for careful control over their placement and extent. Also, a real lesion will distort a brain's shape. By

Table 1a
Comparison between registration methods: percent label agreements

	Linear	SPM2	AIR	ANIMAL	Mindboggle
Overlap	27.66% (1.75)	27.43% (1.78)	27.97% (2.49)	31.63% (3.73)	
Mask overlap	43.50% (1.44)	42.85% (1.50)	45.29% (1.58)	48.00% (3.39)	
Filled mask	74.29% (1.73)	74.51% (1.86)	75.56% (1.63)	76.41% (1.83)	76.88% (1.33)

This table compares different registration methods. One of the 10 subjects was randomly selected to serve as an atlas, and its T1-weighted MR volume was registered to the other 9 subjects' T1 images using the five different methods: Linear (12-parameter affine) registration alone and in combination with nonlinear warping by *SPM2*, AIR (5th degree polynomial transform), and ANIMAL (with multi-scale resolution steps of 16, 8, 4, and 2 mm FWHM Gaussian kernels), and with Mindboggle. The atlas labels were then transformed to each subject's manual labels in MNI152 space. Percentages were calculated from the overlap of (1 × 1 × 1 mm) labeled voxels in MNI152 space. Every entry is in percentage of voxels averaged across the nine subjects, with standard deviations in parentheses.

There was a mean of 582,715 manually labeled voxels for the nine subjects and a mean of 654,007 automatically labeled voxels for the nine subjects and the five registration methods. The average intersection between manually and automatically assigned labels, equal to the gray matter overlap, was 37% for Linear, 36% for SPM2, 36% for AIR, 40% for ANIMAL, and 100% for Mindboggle. Overlap is the number of intersecting voxels that have identical labels divided by the total number of labeled voxels. There is a substantial improvement in percent label agreement for these methods according to the less conservative mask overlap metric, where overlap is calculated only within the subject's gray matter mask. Overlap and mask overlap results are the final output of Linear, SPM2, AIR, and ANIMAL. There is greater than a two-fold improvement in these methods when applying the filled mask condition, where subject gray matter masks are then filled with automated labels using Mindboggle's label propagation algorithm (where every unlabeled gray matter voxel is assigned the majority label in the voxel's neighborhood). The data from the filled mask condition will therefore be used as the basis of comparison between the methods. Overlap and mask overlap results are absent for Mindboggle because these intermediate results are not particularly relevant. Mindboggle's warping algorithm was designed to simply coat label boundaries with labels, not to register gray matter to gray matter.

Table 1b
Type I and Type II errors and run times

	Linear		SPM2		AIR		ANIMAL		Mindboggle	
	I	II	I	II	I	II	I	II	I	II
Overlap	0.56 (0.01)	0.42 (0.03)	0.57 (0.02)	0.41 (0.03)	0.55 (0.01)	0.43 (0.04)	0.52 (0.03)	0.39 (0.04)		
Mask overlap	0.56 (0.01)	0.12 (0.01)	0.57 (0.02)	0.12 (0.01)	0.55 (0.01)	0.12 (0.01)	0.52 (0.03)	0.12 (0.01)		
Filled mask	0.26 (0.02)	0.20 (0.01)	0.25 (0.02)	0.20 (0.01)	0.25 (0.02)	0.20 (0.01)	0.23 (0.02)	0.19 (0.01)	0.23 (0.01)	0.19 (0.01)
Time	0		3 min		13 min		2 h		45 min*	

This table compares the average type I and type II errors under each of the conditions of Table 1a (with standard deviations in parentheses). The Time row gives the approximate run time for each method, after linear registration.
* Mindboggle takes about 45 min after gray matter segmentation.

excising a portion of normal brain data to serve as an artificial lesion, we could control for distortion and ensure that the original manual labels would give the correct result. One of the pair of lesions was centered on Broca's region (left inferior frontal gyrus, Talairach coordinates [35, −15, 55]) and the other on the right motor strip (precentral gyrus, Talairach coordinates [−55, 15, 20]). The shape of the artificial lesion was a cube and was increased from 1 to 5 cm^3 (10 to 50 $voxel^3$) (Fig. 7).

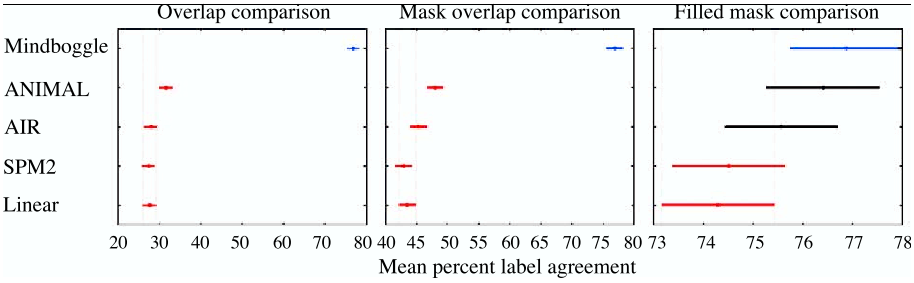

Fig. 6. Comparison between registration methods. A one-way ANOVA was performed to test if the means are the same for the overlap label agreements obtained by each of the methods (linear, SPM2, AIR, ANIMAL, Mindboggle). A multiple comparison test was then performed using Tukey's honestly significant difference criterion to determine which pairs of means are significantly different. The graphs display the mean for each method with a 95% confidence interval around the mean, based on the Studentized range distribution. If intervals are disjoint, their means are considered significantly different. The same test was conducted again for the mask overlap and filled mask overlap label agreements. The filled mask overlap was measured after Mindboggle's label propagation procedure was applied to each of the other methods to fill the subject gray matter masks with labels. Mindboggle's result is in blue and any significant different result is in red. Mindboggle obtained a significantly higher mean union label agreement than any other method ($P < 0.05$). After applying Mindboggle's label propagation procedure, only Mindboggle obtained a significantly higher mean label agreement than did linear registration or SPM2 ($P < 0.05$). See Tables 2a, 2b, and 2c for corresponding ANOVA tables. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Results

Tables 1a and 1b and Fig. 6 present a comparison between the registration methods used to transform the atlas labels to the nine subjects in MNI152 space. Table 1a contains label agreements between automated and manual labels averaged across the nine subjects. The mask overlap results for ANIMAL do not differ from an independent evaluation of the algorithm (Collins et al., 1999). Table 1b compares the type I and type II errors under the same conditions. The Time row gives the approximate run time for each method after performing linear registration (and after tissue segmentation for Mindboggle). SPM2 was the fastest and ANIMAL the slowest of the methods (3 min and over 2 h, respectively).

To evaluate the results obtained by each of the measures (overlap, mask overlap, filled mask overlap, type I error, type II error) applied to each of the methods (linear, SPM2, AIR, ANIMAL, Mindboggle), we used a one-way ANOVA followed by a multiple comparison test. The one-way ANOVA was performed for each measure to test the null hypothesis that the means of the results are equal for the different methods. The multiple comparison test was performed using Tukey's honestly significant difference criterion to determine which pairs of means are significantly different. Two means are considered significantly different if their intervals are disjoint. The 95% confidence intervals are based on the Studentized range distribution.

Mindboggle's filled mask results obtained a significantly higher mean label agreement and significantly lower mean type I and type II errors than the overlap or masked overlap results obtained by any other method ($P < 0.05$). After applying Mindboggle's label propagation procedure to all the other methods (the filled mask condition), their results dramatically improved, but only Mindboggle obtained a significantly higher mean filled mask label agreement and significantly lower mean type I and type II errors than did linear registration or SPM2 ($P < 0.05$). See Tables 2a–2c for the corresponding ANOVA table.

In another experiment, the atlas was registered to itself with each of the registration/labeling methods to determine their "reflexive accuracy" (the degree to which the most trivial condition is consistently satisfied). Table 3 presents the label agreements between the atlas and the atlas registered to itself. The overlap and

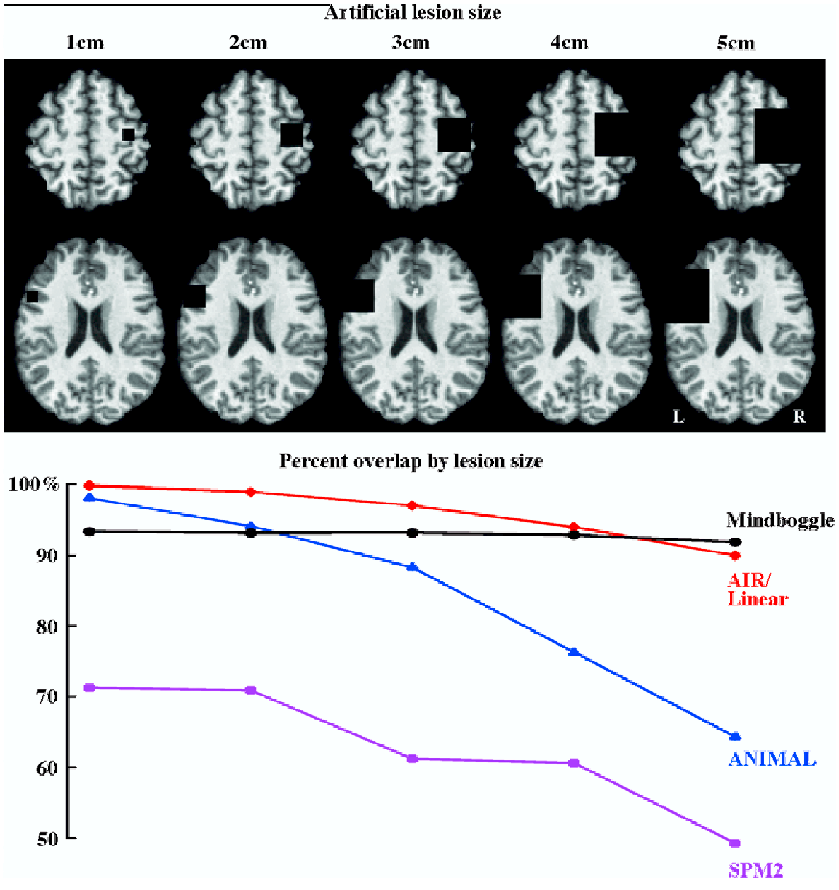

Fig. 7. Registering to artificially lesioned brain image data. Each of the methods (using default parameters) was used to register the labeled T1 atlas to an artificially lesioned version of itself (or a lesioned version of its segmented image in the case of Mindboggle). One of the two lesions was centered on Broca's region (left inferior frontal gyrus, Talairach coordinates [35, −15, 55]), pictured in the top row of images, and the other on the right motor strip (precentral gyrus, Talairach coordinates [−55, 15, 20]), pictured in the second row of images. The shape of the artificial lesion was a cube and was increased from 1 to 5 cm^3 (left to right in the figure). The accompanying graph shows the results of comparing registered labels with lesioned labels for each of the five methods. Without lesions, the union label agreements between the atlas and the atlas registered to itself are equal to 71.57% (SPM2), 100% (AIR), 98.72% (ANIMAL), and 93.50% (Mindboggle). AIR perfectly aligns the atlas with itself, but in the lesioning tests, AIR warps labels within lesioned areas, producing identical results as linear registration. Mindboggle's labeling performance hardly degrades as the lesion size increases, whereas the other methods appear to be more susceptible to discrepancies in the target volume. Moreover, Mindboggle's labeling performance did not degrade with increasing lesion size when the same tests were conducted with lesioned subject brains rather than lesioned atlas brains.

mask overlap label agreements for SPM2 were low, and are a result of slight differences along the edges of the gray mask before and after registration (the edges of the gray matter mask make up 46% of the gray matter itself, or 350,607 of 759,616 voxels). AIR produced the best results of the nonlinear methods, suggesting that it would be more reliable in situations where subject brains are very similar to the atlas. Mindboggle's 6% deviation from perfect reflexive accuracy occurs primarily at the label warping stage and to a lesser extent at the label propagation stage and is less than the number of voxels making up the boundary surfaces between the different labeled regions (8%, or 53,494 of 668,175 labeled voxels within the two hemispheres, not including the interhemispheric plane). It is important to note that the reflexive accuracy test does not measure accuracy and robustness across brains. An algorithm that always performs perfectly when comparing test data with itself may fail when comparing across different data sets. For example, although AIR could theoretically achieve 100% accuracy based on these results, in practice, this is never achieved when one brain is registered to another using this algorithm; therefore, the reflexive accuracy test is not a good indicator of how well one brain will register to another.

Fig. 7 presents the effect of artificial lesions on each labeling method's self-labeling performance. Mindboggle's labeling performance hardly degrades even as the size of each lesion exceeds 50 mm, whereas the performance of the other methods declines as the lesion size increases. The difference is primarily due to the fact that Mindboggle fills a subject's gray matter mask with labels, and the gray matter does not extend into the lesions, whereas the other nonlinear methods warp atlas labels within lesions or to differently labeled regions outside of the lesions. AIR, for example, continues to have perfect atlas-to-atlas label registration, but only in non-lesioned areas. When the same tests were conducted with lesioned subject brains rather than the lesioned atlas brain, Mindboggle's labeling performance still did not degrade with increasing lesion size.

Table 2a
ANOVA table for overlap results

Sources of variability	Sum of squares	Degrees of freedom	Mean squares	F statistic	P value
Columns	16836	4	4208.99	748.66	$<10^{-12}$
Error	224.9	40	5.62		
Total	17060.9	44			

Table 2b
ANOVA table for mask overlap results

Sources of variability	Sum of squares	Degrees of freedom	Mean squares	F statistic	P value
Columns	7500.78	4	1875.19	466	$<10^{-12}$
Error	160.96	40	4.02		
Total	7661.74	44			

This table corresponds to the three (left to right) graphs in Fig. 6.

Table 2c
ANOVA table for filled mask results

Sources of variability	Sum of squares	Degrees of freedom	Mean squares	F statistic	P value
Columns	46.378	4	11.5944	4.07	0.0073
Error	113.921	40	2.848		
Total	160.299	44			

This table corresponds to the three (left to right) graphs in Fig. 6.

Table 3
Comparison between registration methods: atlas labeling itself

	Linear	SPM2	AIR	ANIMAL	Mindboggle
Overlap	100%	71.57%	100%	98.72%	83.76%*
Mask overlap	100%	81.53%	100%	99.20%	84.79%*
Filled mask	100%	96.08%	100%	99.93%	94.21%

Table 3 presents label agreements as in Tables 1a and 1b, except, in this case, for comparisons between the atlas registered to itself.
* See captions to Tables 1a and 1b for explanation.

Discussion

There are numerous methods in the scientific literature that outline different ways to make one brain conform to the shape of another or to identify corresponding points between two brains. Few have been applied in a completely automated way to label brain anatomy (gyrii or sulci) and of these fewer still have been evaluated beyond visual inspection, where image correspondence is mistaken for anatomic correspondence (Rogelj et al., 2002). There exist methods that are readily available for assessing and comparing the accuracy of nonlinear registration methods (Dinov and Sumners, 2001; Grachev et al., 1999; Robbins et al., 2003), but the authors are aware of only a few studies that have compared different nonlinear registration algorithms (Crivello et al., 2002; Dinov et al., 2002; Gee et al., 1997; Hellier et al., 2001, 2002, 2003). In one such study (Hellier et al., 2002), SPM99 matched cortical sulci of different brains better than did ANIMAL, the opposite result of our comparison between SPM2 and ANIMAL, but this was with ANIMAL's finest resolution set to 4 mm (versus 2 mm in this study). As in this example, representing different algorithms fairly, running them with optimal parameter settings, and achieving a consensus on a set of satisfactory metrics have proven to be very difficult hurdles in cross-laboratory comparisons (Paul Thompson, personal communication). The evaluation measures we used in our study are voxel-based overlap label agreements (fractions of manually and automatically labeled voxels that share the same label) coupled with type I and type II statistical error measures. These measures are appropriate for comparing labeled volumes and may be supplemented with other measures. For example, point- or curve-based distance metrics could provide measures more sensitive to discrepancies in position between structural homologues or label boundaries.

In addition to the problem of comparing different automated methods is the problem of comparing automated and manual labeling. Because automated methods have higher consistency but are expected to have lower accuracy than manual labeling, it would be reasonable to compare the original atlas labels with labels assigned to the atlas by automated methods and by manual labeling by several humans. In this way, we could determine how Mindboggle's error rate, for example, compares to within and between labeler inconsistency. Although we cannot make this direct comparison without multiple sets of manual labels, we can recalculate our label agreement results to compare with the interlabeler agreement values measured by Caviness et al. (1996) for a label set

very similar to ours. They found the percent accord between two expert human labelers to be 80.23% (σ = 11.50%[3]) averaged across all 96 labels in four brains, where percent accord is the percent of voxels assigned the same label divided by the mean number of voxels assigned that label. When recalculated for the 74 labeled regions that coincide with the 36 labeled regions used in our study, their average percent accord was 80.00% (σ = 7.91%). This number would be somewhat higher if recalculated for the fewer and larger regions used in our study, since Caviness et al. found a weak correlation between percent accord and region size. The percent accord between Mindboggle and manual labels averaged across the 36 labels in nine brains was 73.06% (σ = 4.94%).

Because of the difference in accuracy and reliability estimates, the authors recommend that automated labeling should precede a round of manual corrections, especially in pathological cases. Having said this, Mindboggle performed better than any of the other registration/labeling methods. After improving the results of the other methods by applying Mindboggle's label propagation algorithm to fill subject gray matter masks with automated labels, Mindboggle was the only nonlinear method tested that gave significantly better results than linear registration or SPM2. Mindboggle is based on a completely different strategy than these other methods and does not make the assumptions they do about preserving topography. The resulting flexibility is advantageous in conditions where brain morphology varies considerably, and may be well suited to labeling pathological brains. Mindboggle is under further development to improve each stage of the algorithm. In this section, we will address areas of development in the same sequence as the subsections of Materials and methods.

Atlas labels and a probabilistic database

We have tested several different individual atlases with Mindboggle and are now incorporating a probabilistic atlas database. In this scenario, Mindboggle labels a subject brain with multiple atlases and each subject voxel is assigned a set of labels, one from each atlas. The number of atlas brains that assigned each label to the voxel provides a confidence measure for the voxel's label. Labeling with multiple atlases ensures that labels assigned to a subject brain are not entirely dependent on the idiosyncrasies of a single atlas brain's morphology. When Mindboggle uses multiple atlases, the label agreement is significantly higher than when using a single atlas with Mindboggle (to appear in a separate publication) or with any of the other methods compared in this study. The label agreement also compares favorably with results obtained by Freesurfer's surface-based cortical parcellation program (Fischl et al., 2004). However, Mindboggle's labels would have to be represented in 2-D for direct comparison with Freesurfer's results. Freesurfer has obtained a median label agreement of 78% for 36 subjects labeled with 50 labels from the Caviness et al. parcellation scheme (Caviness et al., 1996).

With regard to the atlas space itself, there is no reason to restrict the representation of the atlas or subject to the native, 3-D space. Mindboggle could be modified to perform the preparation, matching, and labeling stages in another (for example flat or spherical) coordinate frame if, for instance, distance along the cortical surface were a preferred distance metric to Euclidean distance through the native volumes.

[3] The mean of the standard deviations (Table 4, Caviness et al., 1996) is actually 8.08%.

Alternative image acquisitions

Brain image acquisition does not have to be limited to T1-weighted MR data. Multispectral approaches could help to delineate different and finer structural boundaries in the atlas and subject brains. Mindboggle-assigned labels could be transferred to any of the multiple acquisitions, to label particular imaged or segmented structures that are spatially distinct from the label boundaries, in the same way that Mindboggle labels activity data (see "Labeling activity data" below). For that matter, Mindboggle could be adapted to label other atlas-referenced data, from the histological to the physiological. If it were found that some structural, physiological, or functional subdivision of cortex held a close spatial relationship with another subdivision of cortex, for example, some morphological features such as certain sulci were to have a close correspondence to certain Brodmann's areas, receptor density maps, or activity data, then one set of labels could be used to infer another set of labels. There are bound to be more appropriate subdivisions to infer functional propensity than those provided by a macrostructural atlas, such as microstructural or physiological subdivisions, if not a database of functional mappings.

Improvements in image processing

Mindboggle is a modular program running a sequence of algorithms, beginning with third party software for preprocessing of acquired data. Different preprocessing algorithms could be used (Boesen et al., 2003; Schaper et al., 2003; Yoon et al., 2003); we selected FMRIB's software package because of its ease of installation, ease of use, and its speed. The present algorithms give results that appear robust (this would need to be confirmed by an intra- and intersite evaluation (Styner et al., 2002): we scanned one subject in two different scanners months apart, resulting in different segmentations but a difference in label agreement of less than 1%. The nature of the study or the subject may raise concerns about particular preprocessing steps. For example, segmentation results would be of particular concern if conducting morphometric studies on Mindboggle-labeled brains, since Mindboggle labels a gray matter mask, whose construction is sensitive to the skull-stripping and segmentation algorithms used. A second example would be if one were to label clinical cases where lesions can affect the labeling; for these cases, more sophisticated brain extraction and registration algorithms would probably need to be implemented (Brett et al., 2001; Itti et al., 1997; Periaswamy and Farid, 2003). Likewise, in any subsequent steps after preprocessing, Mindboggle could adopt more accurate and efficient algorithms. For example, the partitioning of a skeletonized brain into left and right hemispheres could benefit from a new algorithm for automating extraction of the midsagittal plane (Hu and Nowinski, 2003).

Another area of improvement is run time reduction. A single subject's run from the preprocessing through the final labeling stage presently takes about 45 min after preprocessing (linear registration and gray matter segmentation) on a Sun Fire 6800 with a 750-MHz, 64-bit processor running Sun Solaris 8: 5 min to construct a sulcus skeleton, 7 min to divide the skeleton with an interhemispheric plane, 13 min to construct and tally data on sulcus pieces, 10 min to transform matching pieces from the atlas to the subject brain, and the remaining 10 min to warp and propagate labels through the gray matter mask. The run time would reduce

significantly with faster preprocessing algorithms and optimized code rewritten in a lower-level language such as C as opposed to Matlab.

Improvements in matching

What is important for matching and labeling structures is that there is a tight correspondence between the structures used in matching and the parcellation boundaries used in labeling. Presently, neither sulcus skeletons nor atlas parcellation boundaries necessarily lie along the midlines of sulci due to noise and thresholding artifacts, and labeling conventions, respectively. Mindboggle matches and moves atlas to subject sulcus pieces, carrying along patches of atlas label boundaries to divide the subject brain. If the spatial correlation between sulci and parcellation boundaries were low, then Mindboggle would not be expected to adequately label the subject brain.

The matching stage may in the future exploit multiple constraints, such as statistical attributes of sulci (Le Goualher et al., 1999, 2000; Mangin et al., 2003; Royackkers et al., 1999; Wang and Staib, 1998), anatomical or functional connectivity information provided by diffusion-weighted imaging (Poupon et al., 2001), or developmental morphogenesis considerations (Cachia et al., 2003b; Jouandet and Deck, 1993). Previous matches could also refine the search for further matching; for example, identification of the central sulcus could constrain which sulcus is most likely to be the precentral sulcus (Corouge and Barillot, 2002; Vaillant and Davatzikos, 1999) and even the identification of distal structures could constrain further sulcus identification (Naidich et al., 2001). Proper matching of corresponding structures would not only aid in labeling a brain, but could refine the spatial localization of activity data (Corouge et al., 2003; Crivello et al., 2002; Flandin et al., 2002; Gee et al., 1997; Hellier et al., 2001; Kochunov et al., 2003). Conversely, it may also be fruitful to use functional mapping as a means of constraining the set of possible matches. For example, if a subject were to undergo a set of tasks that elicit well-characterized activity data, the locations of the data could be used to identify and match features to constrain Mindboggle's matching program. Functional mapping could also help to define anatomically ill-defined boundaries. For example, retinotopic mapping could be used to define the extent of the occipital lobe (Dougherty et al., 2003). And although Mindboggle was designed without concern for topographical preservation, it would be possible to enforce topographical preservation in a modified matching cost function that were take into account global and local neighborhood relationships between features.

Matching sulcus pieces is a significant combinatoric problem, and Mindboggle uses a simple, non-exhaustive, non-probabilistic, non-evolutionary strategy. The primary advantages of employing these characteristics are faster computation and elimination of intermediate results. An alternative would be to perform probabilistic matching to search the solution space more broadly. Genetic algorithms and simulated annealing are two such approaches; parallel recombinative simulated annealing (PRSA), a combination of the two, allows for parallelism and convergence to a global solution (Mahfoud and Goldberg, 1995). We attempted an implementation of PRSA for the matching stage and it failed to find adequate solutions when more than 10 of a reduced set of 25 sulcus pieces were used. This failure could be due to our energy function, but iterative observations of the tentative matches seemed to indicate that there were simply too many possible permutations for mutation and permutation crossovers to evolve toward a solution in reasonable time.

Improvements in label transformation and propagation

After an atlas piece is matched to a combination of sulcus pieces, its corresponding patch of label boundary is simply translated to the sulcus pieces by the difference between their mean and the mean of the atlas piece. We applied an iterative closest-point algorithm (Besl and McKay, 1992) to rotate as well translate each atlas piece to matching subject pieces and applied the transform to the atlas patch, but this did not improve results. Perhaps applying a nonlinear transformation to these boundaries might improve results. Alternatively, a multi-scale approach performed during the matching stage might lead to better registration, first by matching pieces, then by matching the smaller fragments that make up the two matching pieces.

After transforming the label boundaries in the subject space, the labels themselves are warped to coat these new boundaries. Mindboggle precomputes a neighborhood for each boundary point and distances within each neighborhood to the point for computational speed and perfect repeatability of results. However, there is an order effect on the spatial distribution of labeled voxels because voxels warped to one boundary point could overwrite voxels previously warped to a neighboring boundary point. And because convergence is performed in a single step for each neighborhood independently, the resulting warp is not continuous.

Like the Self-Organizing Map (SOM) used earlier to divide the skeleton into hemispheres, this simple warping algorithm preserves neighborhood relationships. This would suggest that topography is preserved, but warping is performed after matching. Although difference in mean position is penalized by the matching cost function, it is possible for the arrangement of the matching pieces to be different in the atlas and subject brains such that they have a different topography. Any labels that are warped to these rearranged boundaries may not preserve topography either.

Our present label propagation strategy to fill in these gaps is crude, simply assigning an unlabeled voxel with the majority label in its cubic neighborhood. A slightly more sophisticated strategy, incrementally increasing the radius of the neighborhood, and taking the majority label of a minimum number of enclosed points, did not improve results. These label propagation strategies and preceding linear boundary transformations fail to properly label about convoluted boundaries. For example, a region that protrudes more in the subject than in the atlas could be overwritten by the label of the surrounding region if it happens to lie between a boundary point and its transformed location. Parametric surface patches might aid registration at this local level, particularly for convoluted boundaries.

Evaluation

In order to improve and validate Mindboggle and other techniques, we need to use reasonable and consistent evaluation measures. We chose overlap, mask overlap, and filled mask label agreements as flexible and intuitive evaluation measures that may be applied to volumes, slices, or regions of interest. Variants of these measures could differentiate between variable and less variable structures, weight regions by size or emphasize border offsets between two brains. Presently, the voxel-based evaluation measures

used in this paper do not take into account misregistration within a labeled region (parcellation unit). The parcellation units may be subdivided fine enough to reduce this problem. However, as the parcellation units become smaller, our confidence in the correspondence between their registered boundaries degrades. The problem is particularly acute when dealing with clinical conditions. Real lesions and brain pathologies can affect the spatial positions of structural and functional regions. In cases where lesions may disrupt label boundaries, more sophisticated evaluation techniques (Fiez et al., 2000) should be employed than those used in this paper to evaluate registration of artificially lesioned brains. Evaluation of morphometric studies may also benefit from the use of different measures that are sensitive to the difference between registered shapes rather than volumes (voxel counts within a parcellation unit) (Gerig et al., 2001).

As long as some reasonable measure is adopted then there is the possibility of comparing different methods (for concerns regarding evaluation see Crum et al., 2003; Rogelj et al., 2002). This assumes, however, that the different methods are performing the same functions. Methods exist that do not perform certain preprocessing stages, nonlinearly register and label only discontinuous structures, and/or involve manual intervention at different stages. A good strategy to evaluate different methods would therefore be to break up the labeling problem into modular components (preprocessing, linear registration, nonlinear registration, label transfer, and label propagation) and evaluate every combination of the different components taken from different established methods. Such a comparison could be extremely helpful when deciding which method to use in order to label particular regions of interest. In this study, we selected only prominent methods that are amenable to whole brain nonlinear registration after a linear registration step. Unfortunately, because Mindboggle relies on a postprocessing step to fill a subject's gray matter mask with labels, its label warping algorithm is not suited to a modular break in processing after warping. Perhaps, employing another warping strategy, such as one of the other methods under comparison, would enable greater modularity of the Mindboggle processing pipeline. Within the nonlinear registration module, one could test different subcomponents such as regularization techniques, additional constraints, and different similarity or distance metrics (Fischer and Modersitzki, 2003).

A more fundamental concern regarding evaluation is the validity of comparing automated labels with manual labels. As stated in Introduction, manual labeling is inconsistent and can never provide an inarguable "gold standard." To ameliorate this situation somewhat, we may turn to probabilistic rather than individual manual labels as a reference with which to compare automated labels (see Atlas labels and a probabilistic database above). Not only could probabilistic labels represent the variability in a population of brains, but could even take into account interlabeler variability as well.

Labeling activity data

The label set tested in this study is not complete. Presently, cerebellums are crudely removed, and subcortical structures and ventricles are not included. We are currently testing other ways to extract cerebellums, since their extraction should improve registration (Rehm et al., 2003) as well as labeling. Subcortical structures are better conserved across brains than cortical structures, so Mindboggle could employ an automated version of the original Talairach approach (Lancaster et al., 2000) or other algorithms dedicated to labeling ventricles and subcortical structures (for example: Duta and Sonka, 1998; Fischl et al., 2002; Iosifescu et al., 1997; Poupon et al., 1998; Schnack et al., 2001).

Anatomical labels assigned by Mindboggle may be transferred to any functional activity data. Because Mindboggle is a postanalysis procedure it labels functional data independently of the manner in which the activity data are acquired, analyzed, or transformed, in the same way that labels that Mindboggle assigns to a gray matter mask are independent of the gray matter segmentation method used. As an example, Mindboggle uses third-party registration software to label fMRI BOLD activity data in the following manner. T2-weighted data are registered to T1-weighted data (voxel intensities may be compared using mutual information in FLIRT, for example). Anatomical labels in MNI152 coregistered space are inverse-transformed to T1-weighted data in the original subject space and then to the subject space where the activity data resides. Mindboggle's label propagation algorithm could then be applied if there are activity voxels that do not overlap the inverse-transformed labels. This approach results in one label assigned to each active voxel and at least one label assigned to each cluster of active voxels. If desired, the labels within a given cluster could then be reassigned as prescribed by a postprocessing algorithm.

In the future, we intend to evaluate Mindboggle using functional mapping in restricted but well understood areas to validate automated labels in those areas. As a preliminary test, we used Mindboggle to label activity data acquired from five subjects undertaking five standard tasks. The standard tasks, acquisition protocol, regions expected to be activated during each task, and the analysis procedure were those outlined by Hirsch et al. (2000) as a standard battery for neurosurgical planning. We determined that Mindboggle's labels of the fMRI BOLD data included regions that we expect to be active during these tasks.

Acknowledgments

This work was submitted in partial fulfillment for the requirements of the Phd degree granted by Cornell University, New York, NY, 2004. Much gratitude to Satrajit Ghosh at Boston University for making available the parcellated brains used in this study; Steve Smith's FMRIB group at Oxford University for the use of their software; Steve Chun and Baldwin Sung for computer assistance; Stephen Dashnaw for scanning early MR subjects; Jack Grinband and Kim E. Lumbard for engaging discussions; Jack, Dan Engber, Barrett Klein, Dr. Norman Relkin, Dr. Jonathon Victor, my thesis advisor Dr. Joy Hirsch, and my four reviewers for proofreading this manuscript; and my wife Deepanjana for indulging my warped logic about this scatterbrained approach.

References

Ashburner, J., Friston, K.J., 1999. Nonlinear spatial normalization using basis functions. Hum. Brain Mapp. 7, 254–266.

Ashburner, J., Neelin, P., Collins, D.L., Evans, A., Friston, K., 1997. Incorporating prior knowledge into image registration. NeuroImage 6, 344–352.

Ashburner, J., Andersson, J.L.R., Friston, K.J., 1999. High-dimensional image registration using symmetric priors. NeuroImage 9, 619–628.

Bajcsy, R., Kovacic, S., 1989. Multiresolution elastic matching. Comput. Vis. Graph. Image Process. 46, 1–21.

Besag, J.E., 1986. On the statistical analysis of dirty pictures. J. R. Stat. Soc., Ser. B 48, 259–302.

Besl, P.J., McKay, N.D., 1992. A method for registration of 3-D shapes. IEEE Trans. Pattern Anal. Mach. Intell. 14 (2), 239–256.

Boesen, K., Rehm, K., Schaper, K., Stoltzner, S., Woods, R., Rottenberg, D., 2003. Quantitative comparison of three brain extraction algorithms. 9th Annual Meeting of the Organization for Human Brain Mapping, New York City, USA.

Bookstein, F.L., 1989. Principal warps: thin plate splines and the decomposition of deformations. IEEE Trans. Pattern Anal. Mach. Intell. 11 (6), 567–585.

Brett, M., Leff, A.P., Rorden, C., Ashburner, J., 2001. Spatial normalization of brain images with focal lesions using cost function masking. NeuroImage 14, 486–500.

Broit, C., 1981. Optimal registration of deformed images. PhD thesis, Department of Computer and Information Science. University of Pennsylvania, Philadelphia.

Cachia, A., Mangin, J.-F., Rivière, D., Papadopoulos-Orfanos, D., Kherif, F., Bloch, I., Régis, J., 2003a. A generic framework for parcellation of the cortical surface into gyri using geodesic Voronoï diagrams. Med. Image Anal. 7 (4), 403–416.

Cachia, A., Mangin, J.-F., Rivière, D., Kherif, F., Boddaert, N., Andrade, A., Papadopoulos-Orfanos, D., Poline, J.-B., Bloch, I., Zilbovicius, M., Sonigo, P., Brunelle, F., Régis, J., 2003b. A primal sketch of the cortex mean curvature: a morphogenesis based approach to study the variability of the folding patterns. IEEE Trans. Med. Imaging 22 (6), 754–765.

Cachier, P., Mangin, J.-F., Pennec, X., Rivière, D., Papadopoulos-Orfanos, D., Régis, J., Ayache, N., 2001. Multisubject non-rigid registration of brain MRI using intensity and geometric features. In: Niessen, W.J., Viergever, M.A. (Eds.), Proc. 4th Medical Image Computing and Computer-Assisted Intervention, Utrecht, Netherlands, Lect. Notes Comput. Sci., vol. 2208. Springer-Verlag, Heidelberg, pp. 734–742.

Caunce, A., Taylor, C.J., 1999. Using local geometry to build 3D sulcal models. In: Kuba, A., Sámal, M., Todd-Pokropek, A. (Eds.), Proc. 16th Information Processing in Medical Imaging, Visegrád, Hungary, Lect. Notes Comput. Sci., vol. 1613. Springer-Verlag, Heidelberg, pp. 196–209.

Caviness Jr., V.S., Meyer, J., Makris, N., Kennedy, D.N., 1996. MRI-based topographic parcellation of human neocortex: an anatomically specified method with estimate of reliability. J. Cogn. Neurosci. 8 (6), 566–587.

Christensen, G.E., 1999. Consistent linear-elastic transformations for image matching. In: Kuba, A., Sámal, M., Todd-Pokropek, A. (Eds.), Proc. 16th Information Processing in Medical Imaging, Visegrád, Hungary, Lect. Notes Comput. Sci., vol. 1613. Springer-Verlag, Heidelberg, pp. 224–237.

Christensen, G.E., Rabbitt, R.D., Miller, M.I., 1994. 3D brain mapping using a deformable neuroanatomy. Phys. Med. Biol. 39, 609–618.

Christensen, G.E., Rabbitt, R.D., Miller, M.I., 1996. Deformable templates using large deformation kinematics. IEEE Trans. Med. Process. 5 (10), 1435–1447.

Chui, H., Rambo, J., Duncan, J., Schultz, R., Rangarajan, A., 1999. Registration of cortical anatomical structures via robust 3D point matching. In: Kuba, A., Sámal, M., Todd-Pokropek, A. (Eds.), Proc. 16th Information Processing in Medical Imaging, Visegrád, Hungary, Lect. Notes Comput. Sci., vol. 1613. Springer-Verlag, Heidelberg, pp. 168–181.

Chui, H., Win, L., Schultz, R., Duncan, J., Rangarajan, A., 2001. A unified feature registration method for brain mapping. In: Insana, M.F., Leahy, R.M. (Eds.), Proc. 17th Information Processing in Medical Imaging, Davis, CA, USA, Lect. Notes Comput. Sci., vol. 2082. Springer-Verlag, Heidelberg, pp. 300–314.

Collins, D.L., Neelin, P., Peters, T.M., Evans, A.C., 1994. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. J. Comput. Assist. Tomogr. 18, 192–205.

Collins, D.L., Holmes, C.J., Peters, T.M., Evans, A.C., 1995. Automatic 3-D model-based neuroanatomical segmentation. Hum. Brain Mapp. 3, 190–208.

Collins, D.L., Le Goualher, G., Evans, A.C., 1998. Non-linear cerebral registration with sulcal constraints. Proc. 1st Medical Image Computing and Computer-Assisted Intervention, Cambridge, MA, USA, Lect. Notes Comput. Sci., vol. 1496. Springer-Verlag, Heidelberg, pp. 974–984.

Collins, D.L., Evans, A.C., 1999. Animal: automatic non-linear image matching and anatomical labeling. In: Toga, A.W. (Ed.), Brain Warping. Academic Press, San Diego, pp. 123–142.

Collins, D.L., Zijdenbos, A.P., Baare, F.C., Evans, A.C., 1999. ANIMAL+INSECT: improved cortical structure segmentation. In: Kuba, A., Sámal, M., Todd-Pokropek, A. (Eds.), Proc. 16th Information Processing in Medical Imaging, Visegrád, Hungary, Lect. Notes Comput. Sci., vol. 1613. Springer-Verlag, Heidelberg, pp. 210–223.

Corouge, I., Barillot, C., 2002. Statistical modeling of pairs of sulci in the context of neuroimaging probabilistic atlas. In: Dohi, T., Kikinis, R. (Eds.), Proc. 5th Medical Image Computing and Computer-Assisted Intervention, Tokyo, Japan, Lect. Notes Comput. Sci., vol. 2489. Springer-Verlag, Heidelberg, pp. 655–662.

Corouge, I., Hellier, P., Gibaud, B., Barillot, C., 2003. Interindividual functional mapping: a nonlinear local approach. NeuroImage 19, 1337–1348.

Crivello, F., Schormann, T., Tzourio-Mazoyer, N., Roland, P.E., Zilles, K., Mazoyer, B.M., 2002. Comparison of spatial normalization procedures and their impact on functional maps. Hum. Brain Mapp. 16, 228–250.

Crum, W.R., Griffin, L.D., Hill, D.L.G., Hawkes, D.J., 2003. Zen and the art of medical image registration: correspondence, homology, and quality. NeuroImage 20, 1425–1437.

D'Agostino, E., Maes, F., Vandermeulen, D., Suetens, P., 2002. A viscous fluid model for multimodal non-rigid image registration using mutual information. In: Dohi, T., Kikinis, R. (Eds.), Proc. 5th Medical Image Computing and Computer-Assisted Intervention, Tokyo, Japan, Lect. Notes Comput. Sci., vol. 2489. Springer-Verlag, Heidelberg, pp. 541–548.

D'Agostino, E., Maes, F., Vandermeulen, D., Suetens, P., 2003a. An information theoretic approach for non-rigid image registration using voxel class probabilities. Biomedical Image Registration (Second International Workshop), Philadelphia, PA, USA, Lect. Notes Comput. Sci., vol. 2717. Springer-Verlag, Heidelberg, pp. 122–131.

D'Agostino, E., Modersitzki, J., Maes, F., Vandermeulen, D., Fischer, B., Suetens, P., 2003b. Free-form registration using mutual information and curvature regularization. Biomedical Image Registration (Second International Workshop), Philadelphia, PA, USA, Lect. Notes Comput. Sci., vol. 2717. Springer-Verlag, Heidelberg, pp. 11–20.

Davatzikos, C., 1996. Spatial normalization of 3D brain images using deformable models. J. Comput. Assist. Tomogr. 20, 656–665.

Davatzikos, C., 1998. Mapping image data to stereotaxic spaces: applications to brain mapping. Hum. Brain Mapp. 6, 334–338.

Declerck, J., Subsol, G., Thirion, J.-P., Ayache, N., 1995. Automatic retrieval of anatomical structures in 3D medical images. In: Ayache, N. (Ed.), Proc. 1st Computer Vision, Virtual Reality and Robotics in Medicine, Nice, France. Springer-Verlag, Berlin, pp. 153–162.

Dice, L.R., 1945. Measures of the amount of ecologic association between species. Ecology 26 (3), 297–302.

Dinov, I.D., Sumners, D.W.L., 2001. Applications of frequency dependent wavelet shrinkage to analyzing quality of image registration. SIAM J. Appl. Math. 62 (2), 367–384.

Dinov, I.D., Mega, M.S., Thompson, P.M., Woods, R., Sumners, D.W., Sowell, E.R., Toga, A.W., 2002. Quantitative comparison and analysis of brain image registration using frequency-adaptive wavelet shrinkage. IEEE Trans. Inf. Technol. Biomed. 6 (1), 73–85 (Enclosed in Appendix II).

Dougherty, R.F., Koch, V.M., Brewer, A.A., Fischer, B., Modersitzki, J., Wandell, B.A., 2003. Visual field representations and locations of visual areas V1/2/3 in human visual cortex. J. Vis. 3, 586–598.

Duta, N., Sonka, M., 1998. Segmentation and interpretation of MR brain images: an improved active shape model. IEEE Trans. Med. Imaging 17, 1049–1062.

Evans, A.C., Dai, W., Collins, D.L., Neelin, P., Marrett, S., 1991. Warping of a computerized 3-D atlas to match brain image volumes for quantitative neuroanatomical and functional analysis. In: Loew, M.H. (Ed.), Medical Imaging V: Image Processing, Proc. SPIE, vol. 1445. pp. 236–246. San Jose, CA.

Evans, A.C., Collins, D.L., Milner, B., 1992. An MRI-based stereotactic brain atlas from 300 young normal subjects. Proc. of the 22nd Symposium of the Society for Neuroscience, vol. 408. Anaheim.

Fiez, J.A., Damasio, H., Grabowski, T.J., 2000. Lesion segmentation and manual warping to a reference brain: intra- and interobserver reliability. Hum. Brain Mapp. 9, 192–211.

Fischer, B., Modersitzki, J., 2003. FLIRT: a flexible image registration toolbox. Biomedical Image Registration (Second International Workshop), Philadelphia, PA, USA, Lect. Notes Comput. Sci., vol. 2717. Springer-Verlag, Heidelberg, pp. 261–270.

Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M., 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 33, 341–355.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., Dale, A.M., 2004. Automatically parcellating the human cerebral cortex. Cereb. Cortex 14, 11–22.

Flandin, G., Kherif, F., Pennec, X., Malandain, G., Ayache, N., Poline, J.-B., 2002. Improved detection sensitivity in functional MRI data using a brain parcelling technique. In: Dohi, T., Kikinis, R. (Eds.), Proc. 5th Medical Image Computing and Computer-Assisted Intervention, Tokyo, Japan, Lect. Notes Comput. Sci., vol. 2488. Springer-Verlag, Heidelberg, pp. 467–474.

Friston, K.J., Ashburner, J., Poline, J.B., Frith, C.D., Heather, J.D., Frackowiak, R.S.J., 1995. Spatial registration and normalization of images. Hum. Brain Mapp. 2, 165–189.

Gaens, T., Maes, F., Vandermeulen, D., Suetens, P., 1998. Non-rigid multimodal image registration using mutual information. In: Wells, W.M., Colchester, A., Delp, S. (Eds.), Proc. 1st Medical Image Computing and Computer-Assisted Intervention, Cambridge, MA, USA, Lect. Notes Comput. Sci., vol. 1496. Springer-Verlag, Heidelberg, pp. 1099–1106.

Gee, J.C., 1999. On matching brain volumes. Pattern Recognit. 32, 99–111.

Gee, J.C., Reivich, M., Bajcsy, R., 1993. Elastically deforming 3D atlas to match anatomical brain images. J. Comput. Assist. Tomogr. 17, 225–236.

Gee, J.C., LeBriquer, L., Barillot, C., Haynor, D.R., Bajcsy, R., 1995. Bayesian approach to the brain image matching problem. In: Loew, M.H. (Ed.), Medical Imaging 1995: Image Processing, Proc. SPIE, vol. 2434. pp. 145–156. Bellingham, WA.

Gee, J.C., Alsop, D.C., Aguirre, G.K., 1997. Effect of spatial normalization on analysis of functional data. In: Hanson, K.M. (Ed.), Medical Imaging 1997: Image Processing, Proc. SPIE, vol. 3034. pp. 550–560. Bellingham, WA.

Gerig, G., Styner, M., Shenton, M.E., Lieberman, J.A., 2001. Shape versus size: improved understanding of the morphology of brain structures. In: Niessen, W.J., Viergever, M.A. (Eds.), Proc. 4th Medical Image Computing and Computer-Assisted Intervention, Utrecht, Netherlands, Lect. Notes Comput. Sci., vol. 2208. pp. 24–32.

Grachev, I.D., Berdichevsky, D., Rauch, S.L., Heckers, S., Alpert, N.M., 1998. Anatomic landmark-based method for assessment of intersubject image registration techniques: Woods vs. Talairach. 8th Annual Meeting of the Organization for Human Brain Mapping, Brighton, England.

Grachev, I.D., Berdichevsky, D., Rauch, S.L., Heckers, S., Kennedy, D.N., Caviness, V.S., Alpert, N.M., 1999. A method for assessing the accuracy of intersubject registration of the human brain using anatomic landmarks. NeuroImage 9, 250–268.

Guimond, A., Roche, A., Ayache, N., Meunier, J., 2001. Three-dimensional multimodal brain warping using the demons algorithm and adaptive intensity corrections. IEEE Trans. Med. Imaging 20 (1), 58–69.

Hartkens, T., Hill, D.L.G., Castellano-Smith, A.D., Hawkes, D.J., Maurer Jr., C.R., Martin, A.J., Hall, W.A., Liu, H., Truwit, C.L., 2002. Using points and surfaces to improve voxel-based non-rigid registration. In: Dohi, T., Kikinis, R. (Eds.), Proc. 5th Medical Image Computing and Computer-Assisted Intervention, Tokyo, Japan, Lect. Notes Comput. Sci., vol. 2489. Springer-Verlag, Heidelberg, pp. 565–572.

Hellier, P., Barillot, C., Corouge, I., Gibaud, B., Le Goualher, G., Collins, L., Evans, A., Malandain, G., Ayache, N., 2001. Retrospective evaluation of inter-subject brain registration. In: Niessen, W.J., Viergever, M.A. (Eds.), Proc. 4th Medical Image Computing and Computer-Assisted Intervention, Utrecht, Netherlands, Lect. Notes Comput. Sci., vol. 2208. Springer-Verlag, Heidelberg, pp. 258–265.

Hellier, P., Ashburner, J., Corouge, I., Barillot, C., Friston, K.J., 2002. Inter subject registration of functional and anatomical data using SPM. In: Dohi, T., Kikinis, R. (Eds.), Proc. 5th Medical Image Computing and Computer-Assisted Intervention, Tokyo, Japan, Lect. Notes Comput. Sci., vol. 2489. Springer-Verlag, Heidelberg, pp. 590–597.

Hellier, P., Barillot, C., Corouge, I., Gibaud, B., Le Goualher, G., Collins, D.L., Evans, A., Malandain, G., Ayache, N., Christensen, G.E., Johnson, H.J., 2003. Retrospective evaluation of inter-subject brain registration. IEEE Trans. Med. Imaging 22 (9), 1120–1130.

Hirsch, J., Ruge, M.I., Kim, K.H.S., Correa, D.D., Victor, J.D., Relkin, N.R., Labar, D.R., Krol, G., Bilsky, M.H., Souweidane, M.M., DeAngelis, L.M., Gutin, P.H., 2000. An integrated functional magnetic resonance imaging procedure for preoperative mapping of cortical areas associated with tactile, motor, language, and visual functions. Neurosurgery 47, 711–721.

Hu, Q., Nowinski, W.L., 2003. A rapid algorithm for robust and automatic extraction of the midsagittal plane of the human cerebrum from neuroimages based on local symmetry and outlier removal. NeuroImage 20, 2153–2165.

Iosifescu, D.V., Shenton, M.E., Warfield, S.K., Kikinis, R., Dengler, J., Jolesz, F.A., McCarley, R.W., 1997. An automated registration algorithm for measuring MRI subcortical brain structures. NeuroImage 6, 13–25.

Itti, L., Chang, L., Mangin, J.-F., Darcourt, J., Ernst, T., 1997. Robust multimodality registration for brain mapping. Hum. Brain Mapp. 5, 3–17.

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

Jaume, S., Macq, B., Warfield, S.K., 2002. Labeling the brain surface using a deformable multiresolution mesh. In: Dohi, T., Kikinis, R. (Eds.), Proc. 5th Medical Image Computing and Computer-Assisted Intervention, Tokyo, Japan, Lect. Notes Comput. Sci., vol. 2488. Springer-Verlag, Heidelberg, pp. 451–458.

Jenkinson, M., Smith, S., 2001. A global optimisation method for robust affine registration of brain images. Med. Image Anal. 5, 143–156.

Johnson, H.J., Christensen, G.E., 2001. Landmark and intensity-based, consistent thin-plate spline image registration. In: Insana, M.F., Leahy, R.M. (Eds.), Proc. 17th Information Processing in Medical Imaging, Davis, CA, USA, Lect. Notes Comput. Sci., vol. 2082. Springer-Verlag, Heidelberg, pp. 329–343.

Jouandet, M., Deck, M., 1993. The prenatal growth of the human cerebral cortex—A brainprint analysis. Radiology 188, 765–774.

Jouandet, M.L., Tramo, M.J., Herron, D.M., Hermann, A., Loftus, W.C., Bazell, J., Gazzaniga, M.S., 1989. Brainprints: computer-generated two-dimensional maps of the human cerebral cortex in vivo. J. Cogn. Neurosci. 1, 88–117.

Kikinis, R., Shenton, M.E., Iosifescu, D.V., McCarley, R.W., Saiviroonporn, P., Hokama, H.H., Robatino, A., Metcalf, D., Wible, C.G., Portas, C.M., Donnino, R., Jolesz, F.A., 1996. A digital brain atlas for surgical planning, model driven segmentation and teaching. IEEE Trans. Visual. Comput. Graphics 2 (3), 232–241.

Klein, A., Hirsch, J., 2001. Automatic labeling of brain anatomy and fMRI

brain activity. 7th Annual Meeting of the Organization for Human Brain Mapping, Brighton, England.
Klein, A., Hirsch, J., 2002. Fully-automated nonlinear labeling of human brain activity. 8th Annual Meeting of the Organization for Human Brain Mapping, Sendai, Japan.
Klein, A., Hirsch, J., 2003. Mindboggle: new developments in automated brain labeling. 9th Annual Meeting for the Organization of Human Brain Mapping, New York City, USA.
Kochunov, P., Lancaster, J., Thompson, P., Boyer, A., Hardies, J., Fox, P., 2000. Evaluation of octree regional spatial normalization method for regional anatomical matching. Hum. Brain Mapp. 11, 193–206.
Kochunov, P., Hasnain, M., Lancaster, J., Grabowski, T., Fox, P., 2003. Improvement in variability of the horizontal meridian of the primary visual area following high-resolution spatial normalization. Hum. Brain Mapp. 18 (2), 123–134.
Kohonen, T., 1997. Self-Organizing Maps, second ed. Springer-Verlag, New York.
Lancaster, J.L., Woldorff, M.G., Parsons, L.M., Liotti, M., Freitas, C.S., Rainey, L., Kochunov, P.V., Nickerson, D., Mikiten, S.A., Fox, P.T., 2000. Automated Talairach Atlas labels for functional brain mapping. Hum. Brain Mapp. 10, 120–131.
Le Goualher, G., Collins, L., Barillot, C., Evans, A., 1998. Automatic identification of cortical sulci using a 3D probabilistic atlas. Proc. 1st Medical Image Computing and Computer-Assisted Intervention, Cambridge, MA, USA, Lect. Notes Comput. Sci., vol. 1496. Springer-Verlag, Heidelberg, pp. 509–518.
Le Goualher, G., Procyk, E., Collins, L., Venegopal, R., Barillot, C., Evans, A., 1999. Automated extraction and variability analysis of sulcal neuroanatomy. IEEE Trans. Med. Imaging 18 (3), 206–217.
Le Goualher, G., Argenti, A.M., Duyme, M., Baaré, W.F.C., Hulshoff Pol, H.E., Boomsma, D.I., Zouaoui, A., Barillot, C., Evans, A., 2000. Statistical sulcal shape comparisons: application to the detection of genetic encoding of the central sulcus shape. NeuroImage 11, 564–574.
Liu, T., Shen, D., Davatzikos, C., 2003. Deformable registration of cortical structures via hybrid volumetric and surface warping. Medical Image Computing and Computer-Assisted Intervention, Lect. Notes Comput. Sci., vol. 2879. Springer-Verlag, Heidelberg, pp. 780–787.
Lohmann, G., 1998. Extracting line representations of sulcal and gyral patterns in MR images of the human brain. IEEE Trans. Med. Imaging 17 (6), 1040–1048.
Lohmann, G., von Cramon, D.Y., 2000. Automatic labelling of the human cortical surface using sulcal basins. Med. Image Anal. 4, 179–188.
Lohmann, G., Yves von Cramon, D., 1998. Automatic detection and labelling of the human cortical folds in magnetic resonance data sets. In: Burkhardt, H., Neumann, B. (Eds.), Proc. 5th European Conference on Computer Vision, Freiburg, Germany. Springer-Verlag, Berlin, pp. 369–381.
Magnotta, V.A., Bockholt, H.J., Johnson, H.J., Christensen, G.E., Andreasen, N.C., 2003. Subcortical, cerebellar, and magnetic resonance based consistent brain image registration. NeuroImage 19, 233–245.
Mahfoud, S.W., Goldberg, D.E., 1995. Parallel recombinative simulated annealing: a genetic algorithm. Parallel Comput. 21, 1–28.
Maintz, J.B.A., 1996–1997. An overview of medical image registration methods. Symposium of the Belgian hospital physicists association (SBPH/BVZF) 12 (V), 1–22.
Maintz, J.B.A., Viergever, M.A., 1998. A survey of medical image registration. Med. Image Anal. 2 (1), 1–36.
Malandain, G., Fernández-Vidal, S., 1998. Euclidean Skeletons. Image Vis. Comput. 16 (5), 317–327.
Maldjian, J.A., Laurienti, P.J., Burdette, J.H., 2004. Precentral gyrus discrepancy in electronic versions of the Talairach atlas. NeuroImage 21 (1), 450–455.
Mandl, R.C.W., Jansma, J.M., Slagter, H., Collins, D.L., Ramsey, N.F., 2000. On the validity of associating stereotactic coordinates with anatomical nomenclature. 6th Annual Meeting of the Organization for Human Brain Mapping, San Antonio, USA.
Mangin, J.-F., Frouin, V., Bloch, I., Régis, J., López-Krahe, J., 1995. From 3-D magnetic resonance images to structural representations of the cortex topography using topology preserving deformations. J. Math. Imaging Vis. 5 (4), 297–318.
Mangin, J.-F., Rivière, D., Coulon, O., Poupon, C., Cachia, A., Cointepas, Y., Poline, J.-B., Le Bihan, D., Régis, J., Papadopoulos-Orfanos, D., 2003. Coordinate-based versus structural approaches to brain image analysis. Artif. Intell. Med. (in press).
Mazziotta, J., 1997. Atlases and anatomies. Proc. 3rd International Conference on Functional Mapping of the Human Brain, Copenhagen.
Meier, D., Fisher, E., 2002. Parameter space warping: shape-based correspondence between morphologically different objects. IEEE Trans. Med. Imaging 21 (1), 31–47.
Miller, M.I., Christensen, G.E., Amit, Y., Grenander, U., 1993. Mathematical textbook of deformable neuroanatomies. Proc. Natl. Acad. Sci. 90, 11944–11948.
Naidich, T.P., Blum, J.T., Firestone, M.I., 2001. The parasagittal line: an anatomic landmark for axial imaging. Am. J. Neuroradiol. 22, 885–895.
Nowinski, W.L., Fang, A., Nguyen, B.T., Raphel, J.K., Jagannathan, L., Raghavan, R., Bryan, R.N., Miller, G.A., 1997. Multiple brain atlas database and atlas-based neuroimaging system. Comput.-Aided Surg. 2, 42–66.
Ono, M., Kubik, S., Abernathey, C.D., 1990. Atlas of the Cerebral Sulci. Georg Thieme Verlag, Stuttgart.
Periaswamy, S., Farid, H., 2003. Elastic registration with partial data. Biomedical Image Registration (Second International Workshop), Philadelphia, PA, USA, Lect. Notes Comput. Sci., vol. 2717. Springer-Verlag, Heidelberg, pp. 102–111.
Poupon, F., Mangin, J.-F., Hasboun, D., Poupon, C., Magnin, I., Frouin, V., 1998. Multi-object deformable templates dedicated to the segmentation of brain deep structures. Proc. 1st Medical Image Computing and Computer-Assisted Intervention, Cambridge, MA, USA, Lect. Notes Comput. Sci., vol. 1496. Springer-Verlag, Heidelberg, pp. 1134–1143.
Poupon, C., Mangin, J.-F., Clark, C.A., Régis, J., Le Bihan, D., Bloch, I., 2001. Towards inference of human brain connectivity from MR diffusion tensor data. Med. Image Anal. 5, 1–15.
Rademacher, J., Caviness Jr., V.S., Steinmetz, H., Galaburda, A.M., 1993. Topographical variation of the human primary cortices: implications for neuroimaging, brain mapping, and neurobiology. Cereb. Cortex 3, 313–329.
Rehm, K., Anderson, J., Woods, R., Rottenberg, D., 2003. Inter-subject spatial registration of the human cerebellum. 9th Annual Meeting for the Organization of Human Brain Mapping, New York City, USA.
Rettman, M.E., Han, X., Xu, C., Prince, J.L., 2002. Automated sulcal segmentation using watersheds on the cortical surface. NeuroImage 15, 329–344.
Rivière, D., Mangin, J.-F., Papadopoulos, D., Martinez, J.-M., Frouin, V., Régis, J., 2000. Automatic recognition of cortical sulci using a congregation of neural networks. In: Delp, S.L., DiGioia, A.M., Jaramaz, B. (Eds.), Proc. 3rd Medical Image Computing and Computer-Assisted Intervention, Pittsburgh, PA, USA, Lect. Notes Comput. Sci., vol. 1935. Springer-Verlag, Heidelberg, pp. 40–49.
Robbins, S., Evans, A.C., Collins, D.L., Whitesides, S., 2003. Tuning and comparing spatial normalization methods. Medical Image Computing and Computer-Assisted Intervention, Lect. Notes Comput. Sci., vol. 2879. pp. 910–917.
Rogelj, P., Kovačič, S., Gee, J.C., 2002. Validation of a non-rigid registration algorithm for multi-modal data. In: Sonka, M., Fitzpatrick, M.J. (Eds.), Medical Imaging 2002: Image Processing, Proc. SPIE, vol. 4684. pp. 23–28. San Diego, CA.
Roland, P.E., Geyer, S., Amunts, K., Schormann, T., Schleicher, A., Malikovic, A., Zilles, K., 1997. Cytoarchitectural maps of the human brain in standard anatomical space. Hum. Brain Mapp. 5, 222–227.
Royackkers, N., Desvignes, M., Fawal, H., Revenu, M., 1999. Detection and statistical analysis of human cortical sulci. NeuroImage 10, 625–641.

Salmond, C.H., Ashburner, J., Vargha-Khadem, F., Connelly, A., Gadian, D.G., Friston, K.J., 2002. The precision of anatomical normalization in the medial temporal lobe using spatial basis functions. NeuroImage 17, 507–512.

Sandor, S., Leahy, R., 1997. Surface-based labeling of cortical anatomy using a deformable atlas. IEEE Trans. Med. Imaging 16, 41–54.

Schaper, K., Rehm, K., Stern, J., Rottenberg, D.A., 2003. Recent Trends in MRI Brain-Tissue Segmentation. 9th Annual Meeting of the Organization for Human Brain Mapping, New York City, USA.

Schnack, H.G., Hulshoff Pol, H.E., Baaré, W.F.C., Viergever, M.A., Kahn, R.S., 2001. Automatic segmentation of the ventricular system from MR images of the human brain. NeuroImage 14, 95–104.

Schormann, T., Zilles, K., 1998. Three-dimensional linear and nonlinear transformations: an integration of light microscopical and MRI data. Hum. Brain Mapp. 6, 339–347.

Shen, D., Davatzikos, C., 2002. HAMMER: Hierarchical Attribute Matching Mechanism for Elastic Registration. Trans. Med. Imaging 21 (11), 1421–1439.

Smith, S., 2000. Robust automated brain extraction. 6th International Conference on Functional Mapping of the Human Brain. pp. 625.

Steinmetz, H., Seitz, R.J., 1991. Functional anatomy of language processing: neuroimaging and the problem of individual variability. Neuropsychologia 29 (12), 1149–1161.

Steinmetz, H., Herzog, A., Huang, Y., Hacklander, T., 1994. Discordant brain-surface anatomy in monozygotic twins [correspondence]. N. Engl. J. Med. 331 (14), 952–953.

Styner, M.A., Charles, H.C., Park, J., Gerig, G., 2002. Multi-site validation of image analysis methods—Assessing intra and inter-site variability. In: Sonka, M., Fitzpatrick, J.M. (Eds.), Medical Imaging 2002: Image Processing, Proc. SPIE, vol. 4684. pp. 278–286. San Diego, CA.

Talairach, J., Szikla, G., 1967. Atlas D'anatomie Stereotaxique Du Telencephale: Etudes Anatomo-Radiologiques. Masson & Cie, Paris.

Talairach, J., Tournoux, P., 1988. Co-Planar Stereotaxic Atlas of the Human Brain. Thieme Medical Publishers, New York.

Thompson, P.M., Toga, A.W., 1996. A surface-based technique for warping three-dimensional images of the brain. IEEE Trans. Med. Imaging 15, 402–417.

Thompson, P.M., Woods, R.P., Mega, M.S., Toga, A.W., 2000. Mathematical/computational challenges in creating deformable and probabilistic atlases of the human brain. Hum. Brain Mapp. 9, 81–92.

Toga, A.W. (Ed.), 1999. Brain Warping. Academic Press, San Diego.

Tourville, J.A., Guenther, F.H., 2003. A cortical parcellation scheme for speech studies. Boston University Technical Report CAS/CNS-03-022. Boston University, Boston, MA.

Towle, V.L., Khorasani, L., Uftring, S., Pelizzari, C., Erickson, R.K., Spire, J.-P., Hoffmann, K., Chu, D., Scherg, M., 2003. Noninvasive identification of human central sulcus: a comparison of gyral morphology, functional MRI, dipole localization, and direct cortical mapping. NeuroImage 19, 684–697.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labelling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single subject brain. 8th Annual Meeting of the Organization for Human Brain Mapping, Sendai, Japan.

Vaillant, M., Davatzikos, C., 1997. Finding parametric representations of the cortical sulci using an active contour model. Med. Image Anal. 1 (4), 295–315.

Vaillant, M., Davatzikos, C., 1999. Hierarchical matching of cortical features for deformable brain image registration. In: Kuba, A., Sámal, M., Todd-Pokropek, A. (Eds.), Proc. 16th Information Processing in Medical Imaging, Visegrád, Hungary, Lect. Notes Comput. Sci., vol. 1613. Springer-Verlag, Heidelberg, pp. 182–195.

Wang, Y., Staib, L.H., 1998. Elastic model based non-rigid registration incorporating statistical shape information. Proc. 1st Medical Image Computing and Computer-Assisted Intervention, Cambridge, MA, USA, Lect. Notes Comput. Sci. vol. 1496. Springer-Verlag, Heidelberg, pp. 1162–1173.

Woods, R.P., 1999. Automated global polynomial warping. In: Toga, A.W. (Ed.), Brain Warping. Academic Press, San Diego, pp. 365–376.

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. Comput. Assist. Tomogr. 22, 153–165.

Wright, I.C., Sham, P., Murray, R.M., Weinberger, D.R., Bullmore, E.T., 2002. Genetic contributions to regional variability in human brain structure: methods and preliminary results. NeuroImage 17, 256–271.

Xiong, J., Rao, S., Jerabek, P., Zamarripa, F., Woldorff, M., Lancaster, J., Fox, P.T., 2000. Intersubject variability in cortical activations during a complex language task. NeuroImage 12, 326–339.

Yoon, U., Lee, J.-M., Kim, J.-H., Kim, I.Y., Kim, S.I., 2003. A comparison of automated and semi-automated skull-stripping algorithms: simulated phantom and real data. 9th Annual Meeting of the Organization for Human Brain Mapping, New York City, USA.

Zhang, Y., Brady, M., Smith, S., 2001. Segmentation of brain MR images through a hidden Markov random field model and the expectation maximization algorithm. IEEE Trans. Med. Imaging 20 (1), 45–57.

Zilles, K., Shleicher, A., Langemann, C., Amunts, K., Morosan, P., Palomero-Gallagher, N., Schormann, T., Mohlberg, H., Bürgel, U., Steinmetz, H., Schlaug, G., Roland, P.E., 1997. Quantitative analysis of sulci in the human cerebral cortex: development, regional heterogeneity, gender difference, asymmetry, intersubject variability and cortical architecture. Hum. Brain Mapp. 5, 218–221.

Zilles, K., Kawashima, R., Dabringhaus, A., Fukuda, H., Schormann, T., 2001. Hemispheric shape of European and Japanese brains: 3-D MRI analysis of intersubject variability, ethnical, and gender differences. NeuroImage 13, 262–271.