Cortical Sulci Recognition And Spatial Normalization

Matthieu Perrot[*,a,b,c], Denis Rivière[a,c], Jean-François Mangin[a,b,c]

[a] *LNAO, Neurospin, CEA, Saclay, France*
[b] *INSERM U.797, Orsay, France*
[c] *IFR 49, Paris, France*

Abstract

Brain mapping techniques pair similar anatomical information across individuals. In this context, spatial normalization is mainly used to reduce intersubject differences to improve comparisons. These techniques may benefit from anatomically identified landmarks useful to drive the registration. Automatic labeling, classification or segmentation techniques provide such labels. Most of these approaches depend strongly on normalization, as much as normalization depends on landmark accuracy. We propose in this paper a coherent Bayesian framework to automatically identify approximately 60 sulcal labels per hemisphere based on a probabilistic atlas (a mixture of SPAM models: Statistical Probabilistic Anatomy Map) estimating simultaneously normalization parameters. This way, the labelization method provides also with no extra computational costs a new automatically constrained registration of sulcal structures. We have limited our study to global affine and piecewise affine registration. The suggested global affine approach outperforms significantly standard affine intensity-based normalization techniques in term of sulci alignments. Further, by combining global and local joint labeling, a final mean recognition rate of 86% has been obtained with much more reliable labeling posterior probabilities. The different methods described in this paper have been integrated since the release version 3.2.1 of the BrainVISA software platform (Rivière et al., 2009).

Key words: cortical folds labeling, sulci, normalization, registration, SPAM, Bayesian, Expectation Maximization
PACS: 87.57.nj, 42.30.Sy, 02.50.-r
2010 MSC: 62F15

1. Introduction

In neuroimaging, many challenging methodological or neuroscientific works regard group studies that summarize, repeat or generalize individual results at group levels. Brain mapping techniques have been widely explored in this way and aim to identify and pair similar structures among individuals. They cover a wide range of applications, from detection of reliable landmarks to understanding of brain functions, through characterization of neurodevelopmental disorders. The interpretation of group studies depends widely on the quality of such mapping.

In this respect, two popular approaches are commonly considered: 1) normalization and 2) identification techniques (including labeling or segmentation). The former one warps studied subjects to a common space allowing comparisons of similar anatomical structures or functional patterns, whereas the latter directly recognizes the structures from their own specific «signature» (intensities, morphometry or their mutual relations). The two methods are complementary and are often combined in the literature. However, this combination has been, thus far, limited to sequential concatenation of the two operations (both orders appear in the literature). To exploit at full the synergy of the two techniques, we propose in this study to use them simultaneously.

Historically, registration-based approaches consist in normalizing all studied data into a common space through anatomically sensible transformations (affine, constrained non-linear or diffeomorphic ones). For a detailed survey on registration techniques, the reader is referred to Maintz and Viergever (1998); Gholipour et al. (2007). Herein, we focus on studies dedicated to sulcal data.

Initially, most registration algorithms were iconic, driven by voxelwise intensities of MRI data. Afterwards, surface-based approaches appeared (Fischl et al., 2004; Tosun and Prince, 2008) favouring multiscale features like the curvature or the cortical depth (distance to brain hull). Since these approaches can not distinguish neighbouring folds with similar geometric characteristics, group-identified structures were added as constraints. The latter became popular recently. Regarding sulci, the landmark contraints can be defined as point-based (Caunce and Taylor, 1998; Chui et al., 1999; Chui and Rangarajan, 2000; Lohmann and von Cramon., 2000), surface-based (Joshi et al., 2005; Lui et al., 2007; Thompson et al., 2000; Van Essen, 2005) (popular representations are spherical and flat mapping) or volumetric (Auzias et al., 2009; Cachier et al., 2001; Collins et al., 1995, 1998). Typically, landmarks are initially labeled manually, which is time-consuming and also introduces human-based errors. On top, the increasing amount of data renders manual identification difficult in practice.

[*]Corresponding author. Tel.: +33-1-6908-7836; fax: +33-1-6908-7980.
Email address: matthieu.perrot@cea.fr (Matthieu Perrot)
URL: http://www.lnao.fr (Matthieu Perrot)

Automatic identification of anatomical structures results from segmentation and classification methods. In the case of cortical structures as opposed to subcortical ones, these methods are generally used one after another: the cortical sulci are first extracted, then labeled. These approaches comprise graphical modeling of structured anatomical data (Rivière et al., 2002; Shi et al., 2008, 2009; Vivodtzev et al., 2005; Yang and Kruggel, 2007, 2009), spatial localization of structures from a common space (Le Goualher et al., 1998; Lohmann and von Cramon., 2000; Perrot et al., 2008) and surface-based methods taking into account the geometry of the cortex, especially through curvature information (Fischl et al., 2004; Tosun et al., 2003; Yeo et al., 2010).

Another family of methods, intermediate between normalization and identification of structures, warp anatomical atlases towards individual MRIs, to transfer the labels to each subject's image. There has been in the last decade an increasing interest for probabilistic atlases (Allassonnière et al., 2007; Ashburner and Friston, 2005; De Craene et al., 2005; Joshi et al., 2004; Mazziotta et al., 1995; Shattuck et al., 2008; Thompson and Toga, 1997). These atlases explain better the inter-individual variability, and pay a special attention to unbiased atlas building (independent of chosen reference subject).

The atlas accuracy depends greatly on coregistration quality (Lyttelton et al., 2007), and better structural alignments are obtained onto more reliable templates with more anatomical landmarks. Very few papers took advantage of the latter. In the context of sulci labeling, Vaillant and Davatzikos (1999) explore iterative strategies to reconsider and hierarchically refine structure identification using registration. Recently, Ashburner and Friston (2005); Pohl et al. (2006); Yeo et al. (2008) have fully addressed these issues within Bayesian frameworks and proposed joint registration and classification (segmentation or labeling) estimation. The current paper follows this general Bayesian formulation but introduces novelties regarding anatomical variability modeling, offers techniques to estimate the latter, and before all puts forth the notion of registration priors. Indeed, Ashburner and Friston (2005) constrained the transformation parameters with a Gaussian prior with covariance parameters derived from the Laplacian of the deformation field. Yeo et al. (2008) used edgewise Gaussian priors on each mesh triangles, scaled accordingly by edge length, whereas Pohl et al. (2006) empirically estimated Gaussian priors from a training database. In this paper, we follow this latter approach, refining registration parameters to limit the range of possible transformations.

This paper is an extension of our previous work (Perrot et al., 2008; Rivière et al., 2002) where sulci recognition has been considered from a structural angles through Markovian modeling of cortical folding neighbourhoods. Here, we propose to focus on the use of localization features.

This paper is organized as follows. Section 2 describes a database of 62 subjects designed for the current study. Then, section 3 refers to an introductory model of the sulcal variability, which is refined in section 4 with a joint probabilistic sulcal-based atlas estimation, and joint sulci labeling and normalization. Lastly, in section 5, cross-validated results are presented and demonstrate enhancements achieved with such approaches.

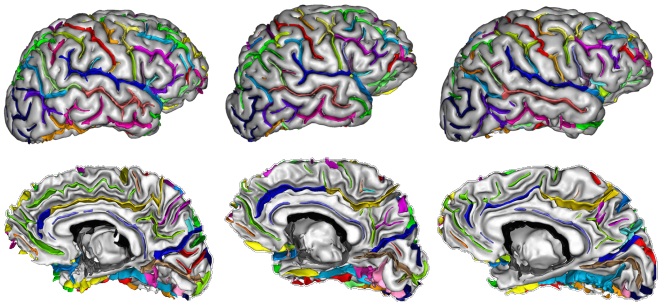


Figure 1: Sulcal patterns variability: 3 manually labeled brains from our new database of 62 subjects. Each color represents a label: 63 on the left hemisphere, 62 on the right one. Even 3 brains allow to picture the sulcal variability extent according to many criterions: size, shape, branches, orientation, topology of sulcal components...

2. Database

Human cortices are highly convoluted by series of intricated folds and vary strongly from one individual to another. Only the largest and deepest folds have led to a clear consensus amongst experts of sulcal anatomy regarding their characteristic pattern in terms of shape or spatial relations. It refers to the primary folds, appearing at the early stage of brain maturation. These elementary folds (Dubois et al., 2008), namely the seeds which will become the future sulci are rather well localized. Beyond, in the matter of secondary and tertiary folds, no ground truth is available: several theories (Lohmann and von Cramon., 2000; Ono et al., 1990; Régis et al., 2005) with many shared principles coexist. Basically, they claim that the cortical sulci are made of elementary folds with different orientations depending on the subject considered. We believe that the study of brain maturation in regard to white fiber structuration, as well as functional activity in regards to cortical patterns will help to refine and merge theses theories.

From the analysis of previous works (Perrot et al., 2008) we point out that sulcal variability modeling provides some ways to measure and give insights into the reliability of sulci definitions. In the latter work, we used SPAM-based localization models (see 3.3 for details on the inference principles and 4.5 for details on the localization model) to analyze the quality of the manual labeling of our training database of 26 subjects. It allows us to detect missing or erroneous labels. Besides, the database was not large enough to claim to be representative of the high inter-subjects variability of sulcal patterns. In fact, our previous studies have resulted in a quite good understanding of the largest folds, but a limited or even a poor comprehension of the more variable ones. In this respect, the sulci nomenclature has been slightly reworked and refined (adding, deleting, splitting or merging labels) in the face of 62 subjects (left and right hemispheres). Indeed, some previous labels were found to be unreliable, hence they have been removed and merged with other labels. Conversely, we identified reliable sub-structures in some labels which we decided to split in several parts. As a result, the labeling consistency has been improved across subjects and

between the two hemispheres, even for the smaller folds. Nevertheless, some limited uncertainties remain and will need the study of many more subjects. However, the new data amount we have at our disposal is enough to suggest some enhancements to sulci modeling and, in consequence, to automatic labeling.

During several months of work, these 62 subjects (including the previous 26 ones) have been labeled or relabeled by two people trained to this task and working in tandem. Brains have been compared to each other several times in groups of a dozen of subjects. For that purpose, we have extended Anatomist, an interactive visualization software provided by the BrainVISA (Rivière et al., 2009) software platform, to perform dedicated convenient group visualizations. Thus, each labeling choice has been discussed at length in order to enhance the consistency. Left and right hemispheres have been labeled one after the other: the whole right hemisphere followed by the left one (and interestingly, we noticed that the resulting models present strong symmetry properties in figure 3). Thus, we hope to have avoided some bias risks. During the labeling process, we have noticed some known results about brain asymmetries: the Wernicke and Broca areas of the left hemisphere, involved in language skills are more variable than their corresponding area on the other hemisphere. Besides, basal folds (collateral fissure, lingual sulcus, occipito temporal sulci...) seem to be less variable and so, easier to recognize on the left hemispheres.

The sulci nomenclature and the labeling principles used to build this database are based on the sulcal roots theory (Régis et al., 2005). These sulcal roots are the locations of the first signs of folding for each sulcus during brain growth and may be linked to white fibers maturation. We believe that these elementary folds are reliable for brain matching. Beyond genetic factors, the remaining complexity and variability of adult brain folds could be rooted in the diversity of environmental stimuli experienced during brain maturation. Thus, its development would depend on the partially chaotic structuration of complex neural networks with many feedback loops at microscopic scales which induces huge differences at macroscopic scales.

Because of our current limited understanding of sulcal anatomy and segmentation limitations, the set of labels we have used has been limited to the most reliable sulcal roots (for instance 5 labels are used to model the Pre-Central sulcus), and groups of neighbouring folds for the more variable and less understood areas (for instance one label only is used for the occipital folds). That way, we defined a set of 63 labels on the left hemisphere and 62 on the right one (see table 3 and figure 12 for label locations and definitions). The unpaired label is localized between the posterior branch of the Sylvian fissure and the bottom of the post-central sulcus. Sulcal roots surely have some strong connections with the concept of sulcal pits (Lohmann et al., 2008) that cover the deepest parts of cortical folds, except for those which have been pushed outward by the development of neighbouring folds.

Whatever the chosen labeling rules, there are many ways to represent folding data. In our representation, we consider only the sulcal parts of the cortical folds, which are better defined and bounded than the gyral ones. In fact, the gyri are reputed to be even more difficult to identify than the sulci, and on top of that, they are delineated by the latter. The cortical folds are represented by a collection of sulcal pieces each made up of a voxels set. These structures are obtained through the BrainVISA (Rivière et al., 2009) anatomical segmentation pipeline (Mangin et al., 1995) (see also (Rivière et al., 2002) and (Mangin et al., 2004) for more recent descriptions) which extracts the cortical folds from T1 MRI. First the image inhomogeneities are corrected thanks to a smooth multiplicative field that minimizes the entropy of the resulting intensity distribution. Then, a scale-space analysis of gray and white matters histograms, helped by Markovian regularization and morphological techniques, is used to segment the brain, gray/white matters from Cerebrospinal fluid (CSF), separate both hemispheres, and ensure that they have a spherical topology. Lastly, a watershed algorithm is used to skeletonize the CSF, thus preserving the initial topology. According to depth, curvature and topological criterions, the resulting structure is over-segmented in elementary folds, a kind of 3D ribbon without any breaks or branches, to assure, as far as possible, that at most one label is expressed on a sulcal piece. Thus, we call *sulcus* a set of such sulcal pieces with the same label. As shown further, the main issue about sulci recognition is to group the sulcal pieces with each other to find sulci. This representation allows to reduce greatly the number of structures to be labeled compared to most surface-based or voxel-based approaches.

To build this larger sulcal-based dataset, several heterogeneous databases: our former database, a diffusion-dedicated database (Poupon et al., 2006), a twins database (Pinel et al., 2007) and some subjects from the ICBM database were grouped from several sites: the SHFJ neuroimaging laboratory (*Service Hospitalier Fréderic Joliot*, France), *La Pitié Salpêtrière* Hospital (France), *La Timone* Hospital (France), McGill University (Canada). Four different 1.5T scanners with various spatial resolutions (about $1\text{mm}^3 \pm 0.16$) and imaging protocols were used to prevent most of the possible bias and insure better generalization properties. Most of the subjects are right-handed men, between 25 and 35 years old. The original goal of these databases are not necessarily linked to sulcal studies, but all contained the needed T1 MRI. Besides, the additional data also includes functional and diffusion data which should be put in correlation with each other in future works. We think that the data heterogeneity has contributed to the robustness of our resulting models.

All the preparation done on this database (including extending, cleaning, matching, comparing...) has been supported by the BrainVISA software (Rivière et al., 2009) initiative. For the circumstance, we have developed efficient ways to manually label cortical folds, from scratch or from a first automatic labeling (an operating mode which is faster but requires more attention to avoid bias in favor of the model previously used) based on only 3 actions: selection of folding structures, copy and paste of labels. Lastly, we have built a tool to certify that each subject appears at most once in the database, otherwise the robust leave-one-out scheme of evaluation used in our results (section 5) would be strongly biased.

3. Standard Labeling Based on Localization

In the following, we present a basic simplified model using only localization information defined from a fixed common space. Subjects are normalized to the latter with registration techniques (section 3.1). Then, the remaining localization variability is modeled thanks to SPAM models (section 3.2). In this context, we describe how to estimate such models and label sulci on new subjects (section 3.3).

3.1. Normalization to Find a Common Space

Many normalization techniques could be considered to define a common space to study sulcal data. These methods proliferate in the literature with various degrees of complexity (from linear to diffeomorphic ones) or with various information (from raw image to labeled anatomical landmarks). Thus, we intentionally restrict our study to standard MRI intensity-based affine techniques, easy to compare with those proposed in this paper since they have the same degrees of freedom.

First, we have used a variant of the Talairach coordinate system (Talairach and Tournoux, 1988) as defined in the BrainVISA (Rivière et al., 2009) anatomical segmentation pipeline. The transformation from the subject space to the Talairach space is affine. Its orientation and origin are defined by the manual alignment of the anterior and posterior commissures (AC-PC). Then the interhemispheric plane and scaling factors are automatically found by matching brain boundaries along the three cardinal axes with the Talairach atlas. This coordinate system was originally designed to align the deepest anatomical structures (thalamus, putamen, caudate nucleus...) which are rather invariant spatially. With respect to cortical structures, their variability is much larger, but potentially overestimated from such an unadapted referential.

We have also tried the Baladin software (Ourselin et al., 2000) dedicated to multimodal image registration, restricted to an affine transformation here. This method is quite reliable because it is based on a multi-scale block matching strategy designed to capture both small and huge displacements. Moreover, an optimal similarity measure, the correlation coefficient, is used to match blocks in pairs between the two images. Lastly, the affine transformation results from the optimization of robust estimators (least trimmed squares) that discard outliers from the process. This method was originally invented to deal with volume reconstruction from unregistered sequential slices, but it is also rather efficient for whole volume registration.

For comparison, we have tested the widely used affine registration provided by the SPM8b software (Ashburner et al., 1997). Its main idea is to find the registration parameters that minimize the sum of squared differences between the 2 images to be matched in a Bayesian framework with priors taking into consideration shape and size variability of standard human heads (distinctly visible on the considered template).

The co-registration of more than 2 training subjects with Baladin or SPM methods is handled by the registration to a template (or a mean subject) as destination image. To that end, we choose the widely used MNI template named ICBM152 (Mazziotta et al., 2001) which results of the averaging of 152 subjects registered with affine transformation on the MNI305 template (Evans et al., 1993). This template and others built from affine transformations are rather similar: only a raw shape of each hemisphere and skull can be discerned because of the high inter-individual cortical variability which blurs the folds.

Lastly, we have looked at the affine version of the ANIMAL (Collins et al., 1994, 1995) software (used to build the MNI305 template), but we did not succeed in getting improvements on the quality of sulcal alignments so detailed results have not been specified further.

3.2. Localization Model: SPAM

After standard affine normalization to a standard template or to a known common space, cortical sulci of a group of subjects are more or less aligned with each other. That is the reason why sulcal variability can be modeled to some extent through estimating the spatial presence probability of finding a sulcus at a given voxel. Namely, we are interested in a voxelwise measure that assesses the probability of being owned by a specific sulcus. To that end, we explore the use of the SPAM model (Statistical Probabilistic Anatomy Map (Evans et al., 1994)), firstly used in the context of sulci modeling in past studies (Le Goualher et al., 1998). For each label, the principle is to count how many subjects have a sulcus going through a given location. These voxelwise frequencies are denoted for a given label l, by a function $f_l(x)$ depending on the location parameter x. These measures are reliable only if many subjects are used. With the database presented previously (section 2), we obtain rather well-defined models for most sulcal labels. For the most variable sulci, many more subjects may be needed. A good way to overcome this difficulty is to model spatial uncertainties during frequency estimations. To do so we use a soft isotropic Gaussian kernel K_σ with a standard deviation σ equal to 2 mm to blur the raw frequency map. The kernel width σ has been set arbitrarily and should be discussed in future refinements. This value is not too large to keep sharp SPAM models on reliable sulcal labels, and not too small to reduce the expression of individual evidences in frequency maps. Lastly, the probability of finding one given structure with label l at the spatial location x writes:

$$P_{spam}(x|L{=}l) \;=\; \frac{(K_\sigma * f_l)(x)}{\sum\limits_{x\in\Omega} (K_\sigma * f_l)(x)} \qquad (1)$$

with $K_\sigma * f$ standing for the convolution of f and K_σ. Here Ω denotes the whole space over which the expression is normalized. In theory Ω is $\mathbb{R}^3$ the whole spatial 3D space. In pratical way, we use one bounding box per label to limit the definition domain of these likelihoods and save memory. SPAM models are encoded as a 3D volume, centered on the distribution core and bounded by negligible values, where each voxel stores the related likelihood. The normalization factor denotes that we consider each SPAM model, dedicated to a given sulcus, as a generative model of voxels positions, so the probability integrates to 1 over Ω.

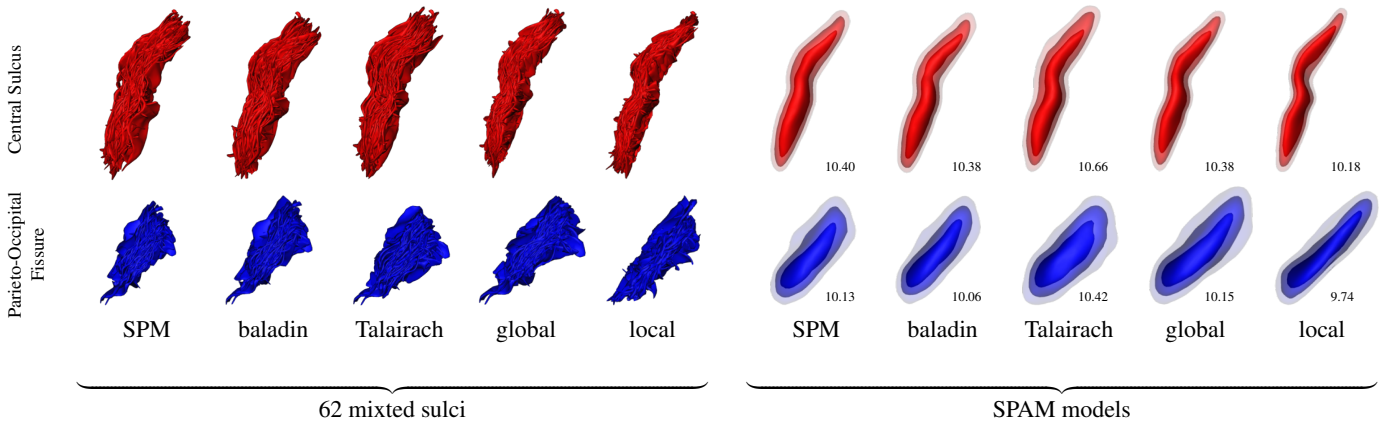


Figure 2: Comparison between standard affine normalization (SPM, Talairach, baladin, all in affine mode) and proposed sulci-based registration (with global and local sulcuswise co-registration) on two major sulci, top: the central sulcus, bottom: parieto-occipital fissure. Left: 62 registered sulci. Right: 3D convenient representations of related SPAM models with their related entropy measure: a lower entropy means a sharper model and less uncertainty. We used 3 nested isosurfaces corresponding to 30, 60 and 80% of the whole probability mass, computed by integrating the probability from the highest likelihoods to the lowest. The registration template used for SPM and baladin registration is the ICBM152, so the comparison is straighforward. About the three other methods, each defines its own space but the variation means are quite close because the global registration refines the Talairach one and the local one refines the global one: so they can also be compared between each other. Of course, the non-linear sulcuswise local registration gives always the best match. About affine approaches, SPM and Talairach give the worst central sulcus alignements and our global method the best. For the parieto-occipital, Talairach is the worst, whereas baladin and our global approach are the best.

The above formula (equation 1) is also known from density estimation field as the Parzen window method (Parzen, 1962) and more generally known as kernel density estimation. According to this point of view, better estimators can be considered. This study focuses only on this formulation, others could be part of future improvements.

Our representation of cortical sulci is based on sulcal pieces (section 2) which are elementary segments with contiguous voxels, weak curvature, without any hole or branch. Each sulcal piece can take at most one label, so we have focused on the probability of having such a label on structures in their wholeness rather than on single voxels. Thus, the achievement of a full generative model relies on the modeling of the joint probability of a set of locations. The true answer to this question is rather difficult and may need complex model design which may be hard to estimate. Actually, the generative aspect is useless here since our method does not need any sampling: it is restricted to fixed real folding structures where only a way to compare likelihoods from one label to another is needed. Thus, approximations can be explored: a sensible voxels independency assumption has been considered in the following. Thus, the probability of finding the label l on one sulcal piece made up of the voxels set $\mathcal{X}$ is given by:

$$P_{spam}(\mathcal{X}|l) \;=\; \prod_{x\in\mathcal{X}} P_{spam}(x|l)^{\frac{1}{|\mathcal{X}|}} \qquad (2)$$

with $|\mathcal{X}|$ the number of voxels of the sulci which behaves like a normalization factor. To explain why we need to normalize geometrically the voxelwise likelikoods, let us consider two sulcal pieces, one twice the size of the other, and a spurious voxelwise multinomial distribution of labels in such way that labels frequencies are invariant spatially. In this situation, each likelihood computed on the larger sulcal piece is the square of the smallest one, so labels compare quite differently (only labels sorting is preserved). The larger the sulcal piece is, the more contrasted the posterior probabilities are. Without this normalization factor, all posteriors would be null except for one label which would have a posterior equal to 1. In other words, the model and so the labeling would be sensitive to scale which is not satisfactory. Moreover, further studied models need comparable posteriors over sulcal pieces because they appear as weights during the registration step (see equation 10).

At sulcal piece scale, the proposed summarized probability does not refer any more to a generative model. In fact, breaks, inherent in sulcal pieces are not modeled here. Therefore, our model is quite sensible to the quality of folding segmentation (section 2 for details on segmentation).

Whether previously described standard normalization methods or further described sulci-based ones are considered, all transformations are affine or piecewise affine. Thus, SPAM model estimation from this common space must take into account sampling issues since sulci are made of voxels of which positions are expressed from subjects coordinate system. Therefore, once sulci are transformed, then linear interpolation is used to assess the effective contribution to its respective frequency map. On the other hand, when likelihoods are estimated we merely round the transformed coordinates using the affine transformation. The visual effects of various co-registration of 62 subjects on sulci alignments can be seen in details for 5 different methods on two selected sulci in figure 2 and for the 3 main methods on all sulci in figure 3. In the subsequent sections, to use SPAM as a concrete model in place of the generic sulcal piece probabilities, consider the following definition which takes into account the affine registration parameters A (linear part) and t (the translation vector):

$$P\left(D_i{=}\mathcal{X}_i|L_i,\Theta{=}(A,t)\right) \;\triangleq\; P_{spam}\left(\{A\cdot x+t\}_{x\in\mathcal{X}_i}\right) \qquad (3)$$

where D_i denotes the parameter for localization data related to

the sulcal piece i. The coordinates of the voxels set $\mathcal{X}_i$ are expressed in the subject space, so $\{Ax + t\}_{x \in \mathcal{X}_i}$ is expressed in the chosen normalization space. Here, Θ is a metaparameter for all possible registration parameters, which does not depend on the label parameter L_i. This definition introduces how to use SPAM with affine registration.

3.3. Sulci Modeling and Labeling from a Common Space

When a suitable common space has been defined, where sulci alignments are rather good and their variations reduced, simple and effective models can be defined to recognize sulci labels. Previously, we introduced the SPAM model which is able to provide the likelihood of having a sulcal piece at a given location given a label. From this model (or similar ones) based on strongly localized information, a straightforward and larger model can be designed making the most of the strong assumption introduced earlier: the statistical independency between sulcal pieces.

The principal goal we are after is the labeling of sulcal pieces from an unseen subject from a test database. Thus, if the normalization parameters Θ are known from a previous step, we seek to infer a labeling L from localized information D. Unknown labels and measured structural data are expressed over the set $\mathcal{E}$ of sulcal pieces in such way that $L = \{L_i\}_{i \in \mathcal{E}}$ and $D = \{D_i\}_{i \in \mathcal{E}}$, each sulcal piece with its own local data D_i (in our case, the set of the 3D coordinates of the voxels of the structure i from the MRI data, but other information could be considered) and its single label L_i. The labeling inference is based on an empirical model $\mathcal{M}$ comprised of all information which can help to model the inter-individual cortical variability, including labels prior $P(L)$ and likelihood $P(D|L, \Theta)$ (in our case SPAM models). From a Bayesian framework, this inference is based on a mere application of the Bayes rule (to lighten probabilistic formula, random variables and their respective concrete values will be confused when no ambiguity is encountered):

$$\begin{aligned} P(\textbf{Labeling}|\textbf{Data}) &= P(L|D, \Theta; \mathcal{M}) \\ &\propto P(D|L, \Theta; \mathcal{M}) P(L|\Theta; \mathcal{M}) \\ &\propto \prod_{i \in \mathcal{E}} P(D_i|L_i, \Theta; \mathcal{M}) P(L_i|\mathcal{M}) \end{aligned} \tag{4}$$

Here, $P(L|\Theta; \mathcal{M})$ represents the labeling prior, namely the chance we have to face the labeling L before any localization data D is observed. Thus, we assumed that this prior is independent of parameters Θ since only D is sensitive to subject referential. Since all sulcal pieces are independent, this formula stands equally for MAP (Maximum A Posteriori) or MPM (marginal posterior mode) Bayesian risks. This means that each local independent labeling problem has the same importance in the eyes of the model.

Thanks to the assumed independencies, the labeling consists in a set of local labelings where each sulcal model is questioned to assess the likelihood of having rather one label or another. Lastly, the most probable label is kept. Moreover, local posterior probabilities can also be computed by normalizing $P(D_i|L_i=l,\Theta;\mathcal{M})P(L_i=l|\mathcal{M})$ over all labels $l \in \mathcal{L}$, which qualifies to what extent the model is confident in its labeling choice.

Sulcal data are rather complex and not fully understood, that is the reason why the sulcal model can not be set a priori but must be estimated empirically from a training database $\mathcal{A}$. From a Bayesian point of view, the estimation of model parameters (SPAMS voxelwise likelihood values or labels prior) is expressed as following:

$$\begin{aligned} &P(\textbf{Model}|\textbf{Training Set}) \\ &\quad = P(\mathcal{M}|\{D_a, L_a, \Theta_a\}_{a \in \mathcal{A}}) \\ &\quad \propto \prod_{a \in \mathcal{A}} \underbrace{P(D_a|L_a, \Theta_a; \mathcal{M}_l)}_{\text{localization models}} \underbrace{P(L_a|\Theta_a; \mathcal{M}_p)}_{\text{labels prior}} \underbrace{P(\mathcal{M}|\Theta_a)}_{\text{model prior}} \end{aligned} \tag{5}$$

where each subscript a stands for data specific to the subject a, $\mathcal{M}_l$ and $\mathcal{M}_p$ stands respectively for parameters specific to localization models and labels prior. Here each transformation Θ_a moves a subject a from its own input space to a common space and is considered to be known, so the model parameters $\mathcal{M}$ are estimated under this hypothesis. Thus, this optimization is broken up in three parts: the optimization of localization models, labels prior and the model prior. The only worthwhile model prior we used is a Dirach prior on the kernel width σ set to 2 mm during SPAM estimation. Thus, $\mathcal{M}_l$ and $\mathcal{M}_p$ can be estimated independently. In the same way, model parameters specific to a given sulcal label can be optimized on their own side (section 3.2 for details on SPAM estimation). About the labeling prior, we consider that each local labeling is independent from others since most structural information is contained in D_a, except the model structure: the segmentation in sulcal pieces which is laid aside in the current paper. Details on how these probabilities are estimated can be found in the appendix A.

This labeling scheme takes advantage of strong and reliable localization information modeled by SPAM. Moreover, shape information is modeled to some extent for the largest sulci. So far, we consider the normalization as a first step before any further analysis, so its choice is determinant for the quality of the model and so the quality of the labeling (see MRI *registration* column of table 2 and section 5 for detailed explanations).

4. Joint Labeling and Spatial Normalization

None of the 3 standard normalization methods seen previously (Talairach, SPM, Baladin) can claim to provide an optimal affine transformation according to sulci alignment, because they may not discriminate neighbouring folds from each other. Therefore, the introduction of anatomical information, through sulcal models like SPAMS, in the registration process should remove most ambiguities. Conversely, the referential choice is essential to reduce as far as possible the anatomical variability and make the comparison of a subject to a sulcal model easier. Both considered information types (transformation to a common space and identification of anatomical structure) are of interest for each other. Bayesian formulation allows us to answer elegantly to this question, but registration parameters optimization must be reworked to be included in this framework. About the registration part, the main idea is to replace the MRI template by the sulcal probabilistic model (section 3.3) which can be seen as a sulcuswise refined template.

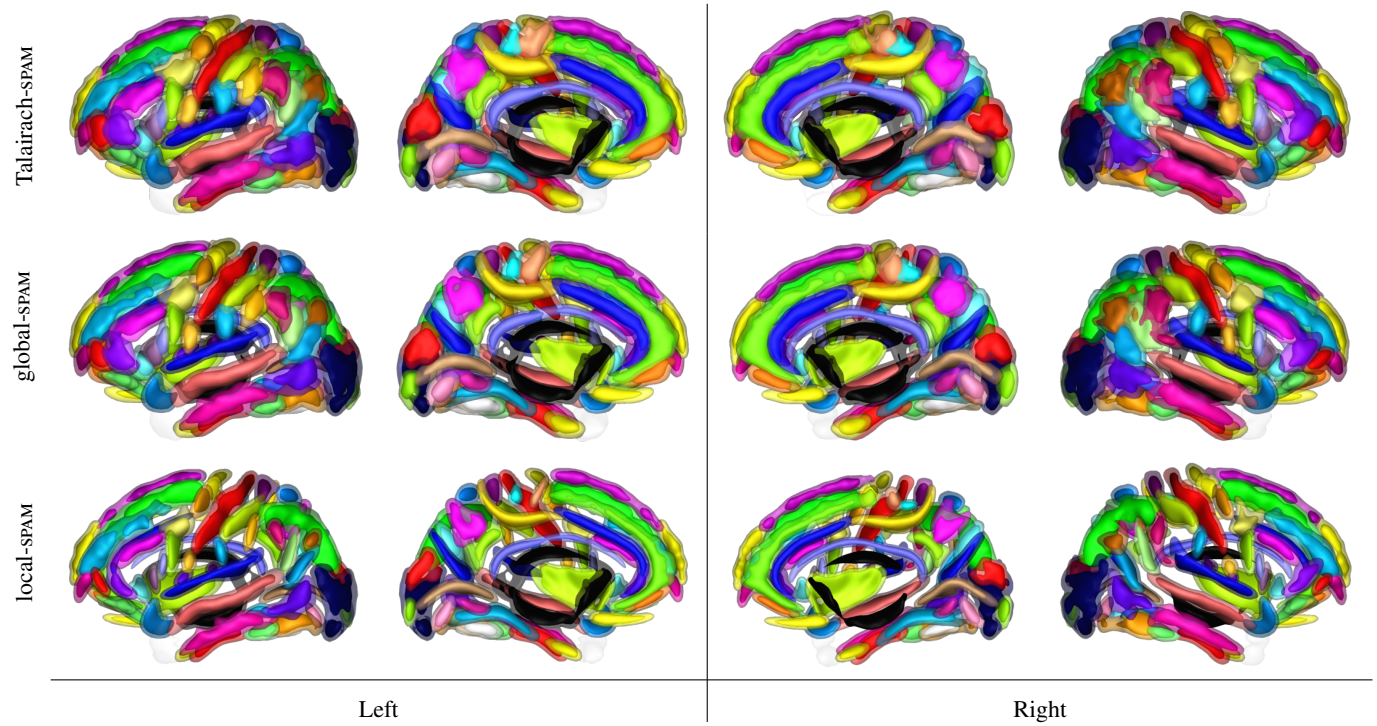


Figure 3: Lateral and medial views of SPAM-based localization models for left and right hemispheres, built from 62 subjects (see section 2 for details on the database). We used 3D convenient representations for each SPAM models based on 2 nested isosurfaces corresponding to 30 and 60% of the whole probability mass, computed by integrating the probability from the highest likelihoods to the lowest. *Talairach*-SPAM: basic SPAM learned from Talairach space. *global*-SPAM: SPAM learned jointly with a global rigid co-registration of subjects from Talairach. *local*-SPAM: SPAM learned jointly with a local rigid co-registration of sulci after applying first the *global*-SPAM model. These models are further described at the begining of section 5.

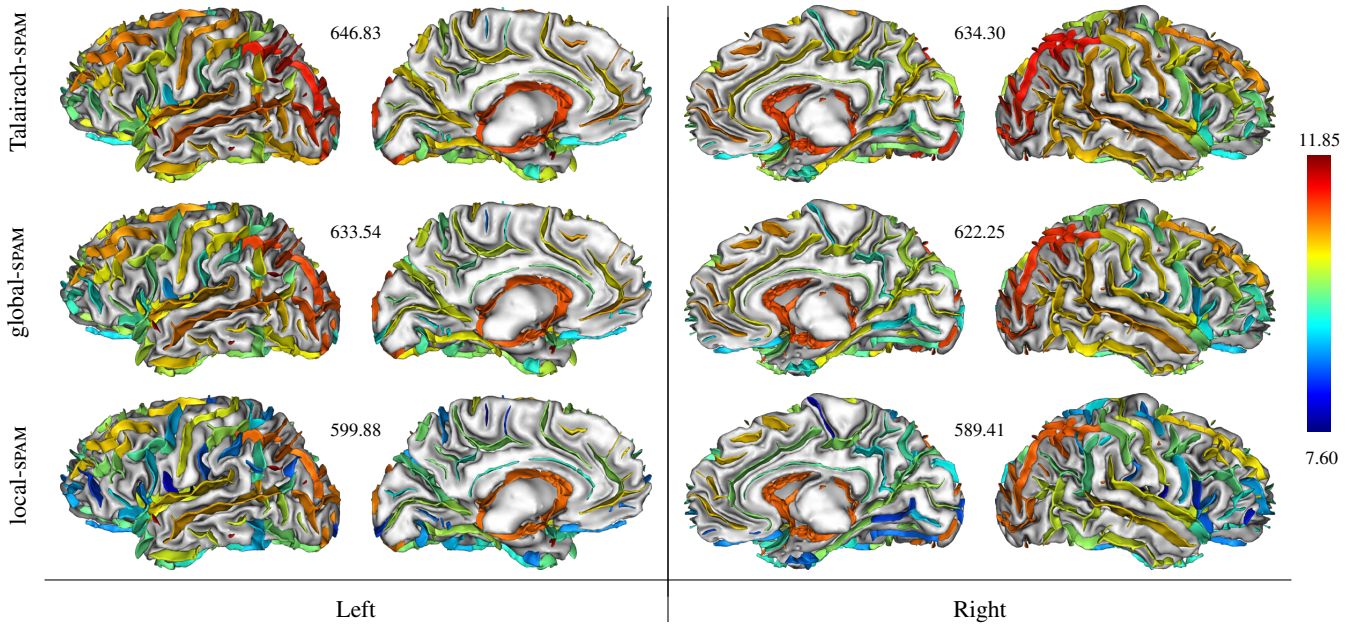


Figure 4: Lateral and medial views of local entropies of SPAM models: $\int_x p_l(x) log\,(p_l(x))\,dx$ with $p_l(x)$ the density of probability for sulcus l at location x, mapped on their respective SPAM-based models for left and right hemispheres (see above figure 3 for details on the models). Their related cumulative entropy (sum of sulcuswise entropy) is specified between lateral and medial views of the same hemisphere. A lower entropy means sharper sulcal models and less localization uncertainty. Entropies of global-SPAM are lower than Talairach-SPAM. The local-SPAM has even lower entropies.

Following these ideas, we will re-consider the arbitrary and a priori chosen normalization space made in the previous section. Basically, we are looking for the common space given rise to the best sulcal variability reduction. Thus, we propose a way to refine an initial common space according to the proposed sulci model to improve both sulcal models and labeling through a Bayesian framework close to the one defined in Pohl et al. (2006) or Yeo et al. (2008) for joint registration and segmentation purposes.

In the following, two main registration methods have been considered in section 4.1: the first with one affine or rigid transformation (called **global** transformation), the second with sulcuswise rigid transformation (called **local** transformation). In this context, the labeling process is presented in section 4.2 and the modeling estimation in section 4.3. Finally, additional details are given on priors used to constrain the normalization steps (section 4.4) and on specificities with regard to SPAM models in a normalization context (section 4.5).

4.1. Transformations

4.1.1. One Global Transformation: Rigid or Affine

The global transformation is defined by only two parameters $(A_g, t_g) = \theta_g$, where A_g is a nonsingular matrix and t_g a translation vector. Here, g subscripts stand for *global* approach. A_g can be broken down into three intelligible components thanks to singular value decomposition (SVD): $A_g = U_g D_g V_g^t$, with U_g and V_g two rotation matrices and D_g a diagonal matrix of scaling components. V_g^t defines the plane along which the scaling factors are applied (which are expected to align principal axes of the brain along the canonical axes after optimization assuming the largest scaling deformations followed them). $U_g V_g^t$ represents the rigid part of the transformation. Affine transformations are controlled by 12 parameters (3 for each rotation, 3 for scaling and 3 for translation): it includes the composition of translation, rotation, scaling and shearing. Our study also focuses on registrations restricted to rigid transformations (keeping aside the scaling matrix D_g), defined by only 6 parameters (3 for rotation, 3 for translation). This leads to more constrained transformations, that are much easier and faster to optimize.

4.1.2. Sulcuswise Local Rigid Transformations

In this article, we explore a natural extension of the previous global registration (see differences between equation 11 and 13). Rather than one global affine or rigid transformation, a collection of local rigid transformations is considered: one per sulcus, denoted by the following parameters $\theta_s = \{\theta_{s,l}\}_{l\in\mathcal{L}} = \{(R_{s,l}, t_{s,l})\}_{l\in\mathcal{L}}$. Here, s subscripts stand for *sulcuswise* method and l for the respective labels. We limit our study to local rigid transformations because, locally, affine ones may provide negligible profits with additional computing costs.

The proposed sulcuswise constrained registration, be it global or local, can be controlled by additional priors over available transformations (section 4.4). Obviously, in the case of the non-linear method these priors are vital as detailed further.

4.2. Joint Registration and Labeling

We assume in this section that we have already learned a model $\mathcal{M}$ of sulcal localization variability from a training database. Either, it has been estimated from a given a priori common space (section 3.1 for standard normalization techniques and section 3.3 for sulci model estimation from a given space) or the common space and other model parameters have been optimized together. In this section, the reference to the model $\mathcal{M}$ is implicit and so not specified, since all probabilities are defined given $\mathcal{M}$.

Our goal here is to find the best sulcal labeling l of an unseen subject from a test database (namely the subject has not been used during the model estimation step). In the same time, we are looking for the best registration θ (limited in our experiments to global affine and sulcuswise local affine ones) in order to maximize the matching between the model and the unlabeled subject, which following a Bayesian formulation writes:

$$l^*, \theta^* = \underset{l,\theta}{\operatorname{argmax}}\ P(L{=}l, \Theta{=}\theta | D) \tag{6}$$

Solving straightforwardly this equation is quite hard since all parameters are optimized at the same time whereas the optimized formula may present many local maxima. However, this probability can be integrated along some parameters, so as to take the fullest account of their uncertainty rather than their principal mode. This leads to consider the following marginal probabilities: $P(L|D)$ or $P(\Theta|D)$ which, once optimized, should give close solutions since the studied probabilities are rather sharp. Ideally, we should prefer $P(L|D)$ since it favours the labeling point of view taking into account all measured uncertainties over the unknown transformation θ:

$$P(L|D) \propto \int P(D|L, \Theta{=}\theta) P(L|\Theta{=}\theta) P(\theta) d\theta \tag{7}$$

The transformation parameters θ live in a continus space since we deal with the compositions of rotations, scaling factors and translation vectors. Therefore, we should compute labeling for all possible θ values to estimate posteriors $P(\Theta{=}\theta|L, D)$, essentials for the optimization, which is impossible as it is since the considered space is continuous. Thus additional modeling and reduction would be needed (using Markov chain monte Carlo MCMC approaches for instance).

Therefore, the other marginal probability $P(\Theta|D)$ seems easier to handle since it provides some mathematical reductions. This time, this measure is dedicated to the optimization of registration parameters θ taking all labeling uncertainties into account. Therefore, this method leads to a robust registration technique constrained by hidden sulcal labels (unknown but inferred from sulcal models). Then, the labeling is optimized under the best θ^* previously obtained:

$$\begin{cases} \theta^* &= \underset{\theta}{\operatorname{argmax}}\ P(\Theta{=}\theta|D) \\ &= \underset{\theta}{\operatorname{argmax}}\ P(\theta) \sum_{L=l} P(L{=}l|\theta) P(D|L{=}l, \theta) \\ l^* &= \underset{l}{\operatorname{argmax}}\ P(L{=}l|D, \Theta{=}\theta^*) \\ &= \underset{l}{\operatorname{argmax}}\ P(D|L{=}l, \Theta{=}\theta^*) P(L{=}l|\Theta{=}\theta^*) \end{cases} \tag{8}$$

The optimization of parameters under hidden variables can be assessed through the Expectation Maximization (EM) algorithm (Dempster et al., 1977). This method starts from a suitable initialization of the parameters to be optimized. In our case, the parameters are given by a first registration step as those described in section 3.1 for instance. Moreover, our joint approach with global registration can be used to initialize a second one with local registration. The EM method is iterative and improves the considered probability at each step, by first computing the posteriors probabilities $P(L{=}l|D,\Theta)$ for each possible labeling l. That is $|\mathcal{L}|^{|\mathcal{E}|}$ combinations, with $|\mathcal{L}|$ the number of labels and $|\mathcal{E}|$ the number of sulcal pieces on the considered subject, which represents about 10^{1000} for 62 labels and 250 sulcal pieces. Secondly, these posteriors are used as weights during parameters refinement. Lastly, the process leads parameters to a local maximum of the probability of interest. The scheme is the following:

$$\theta^{(n+1)} = \underset{\theta}{\operatorname{argmax}}\ Q\left(\theta\middle|\theta^{(n)}\right) \tag{9}$$

with

$$\begin{aligned} Q\left(\theta\middle|\theta^{(n)}\right) &= \mathbb{E}_L\left[\log(P(D,L,\theta))\middle|D,\theta^{(n)}\right] \\ &= \sum_{L=l} w_l^{(n)} \log\left[P(D|L{=}l,\theta)P(L{=}l|\theta)P(\theta)\right] \end{aligned} \tag{10}$$

where $\mathbb{E}_L[\cdot|\cdot]$ stands for the expectation over L given known information, and $w_l^{(n)}{=}P\left(L{=}l,\middle|D,\theta^{(n)}\right)$. The summation is over all possible labelings l.

4.2.1. Global Approach

Now, we will propose a global affine or rigid registration derived from this framework. Rather than assuming strong statistical dependencies between observed data D, we used the rational assumption of conditionnal independence of data $\{D_i\}_{i\in\mathcal{E}}$ given the labeling L. Besides, as detailed before, the labels prior does not depend on the transformation θ and their statistical independence over sulcal pieces is also assumed. Then, the EM version for global registration writes:

$$\begin{aligned} Q_g\left(\theta_g\middle|\theta_g^{(n)}\right) = \log\left[P(\theta_g)\right] + \\ \sum_{i\in\mathcal{E}}\sum_{L_i=l} P\left(L_i{=}l\middle|D_i,\theta_g^{(n)}\right)\log\left[P(D_i|L_i{=}l,\theta_g)P(L_i{=}l)\right] \end{aligned} \tag{11}$$

Here, the expectation step (E) constists in computing local posteriors $P\left(L_i{=}l\middle|D_i,\theta_g^{(n)}\right)$ for each sulcal piece and each label under a known transformation $\theta_g^{(n)}$. Namely, we have to apply the scheme described in section 3.3. These posteriors measure how the model is confident on a given labeling and weight the contribution of each sulcal model during the maximization step (M). Since sulcal models are quite localized, on each sulcal piece most posterior probabilities are null, so only a few labels drive the registration. These weights behave like force constants of many springs linking sulcal pieces and sulcal models together.

Readers familiar with mixture models have surely recognized an almost standard parameters optimization within this framework. Here, sulcal models (in this paper, we used SPAM models, see section 3.2) are mixed together. The only subtlety is that all model parameters are fixed and defined up to an affine registration.

4.2.2. Local Approach

Starting again from equation 10, if we consider that the whole registration θ_s is made up of independent transformations $\theta_{s,l}$ (where s subscript for *sulcuswise*), one dedicated to each label (in our case one rigid transformation per label is considered), the EM algorithm writes:

$$\begin{aligned} Q_s\left(\theta_s\middle|\theta_s^{(n)}\right) = \log\left[P(\theta_{s,l})\right] + \\ \sum_{l\in\mathcal{L}}\sum_{i\in\mathcal{E}} P\left(L_i{=}l\middle|D_i,\theta_{s,l}^{(n)}\right)\log\left[P(D_i|L_i{=}l,\theta_{s,l})P(L_i{=}l)\right] \end{aligned} \tag{12}$$

with $P(\theta_s) = \prod_{l\in\mathcal{L}} P(\theta_{s,l})$.

This time, labelwise assessment can be uncorrelated and so the optimization can be done separately:

$$\theta_s^* = \left\{\underset{\theta_{s,l}}{\operatorname{argmax}} \sum_{i\in\mathcal{E}} P\left(L_i{=}l\middle|D_i,\theta_{s,l}^{(n)}\right)\log[P(D_i|L_i,\theta_{s,l})P(\theta_{s,l})]\right\}_{l\in\mathcal{L}} \tag{13}$$

Since labelwise local registrations are supposed to be independent whereas in reality they are not, the details of the optimization process may be counterintuitive. The process starts from a common space where all local transformations are equal. One more time, posterior probabilities provide weights to drive the registration step, but this time each optimization yields a different transformation. Virtually, we obtained as much transformed brains (with the related rigid transformation) as labels since labels are unknown. In the following E step, likelihoods of a given label is obtained by mixing each transformed brain with its related sulcal model. Posteriors are obtained after normalization as following:

$$P(L_i{=}l|D_i,\theta_s^{(n)}) = \frac{P\left(D_i|L_i{=}l,\theta_{s,l}^{(n)}\right)P(L_i{=}l)}{\sum_{l_i\in\mathcal{L}} P\left(D_i|L_i{=}l_i,\theta_{s,l_i}^{(n)}\right)P(L_i{=}l_i)} \tag{14}$$

since $P\left(D_i|L_i{=}l_i,\left\{\theta_{s,l_i}^{(n)}\right\}_{l\in\mathcal{L}}\right) = P\left(D_i|L_i{=}l_i,\theta_{s,l_i}^{(n)}\right)$. Namely the likelihood of a sulcal piece i of having the label l depends only on the transformation dedicated to this label.

4.2.3. Final Labeling

Be it for global or local approach, once an optimal value θ^* has been reached for registration parameters, a labeling is inferred as previously (section 3.3 for details) but this time, after applying the transformation induced by the parameter θ^*:

$$\begin{aligned} l^* &= \{l_i^*\}_{i\in\mathcal{E}} \\ &= \left\{\underset{l_i}{\operatorname{argmax}}\ P\left(D_i|L_i{=}l_i,\Theta{=}\theta^*\right)P\left(L_i{=}l_i|\Theta{=}\theta^*\right)\right\}_{i\in\mathcal{E}} \end{aligned} \tag{15}$$

4.3. Joint Registration and Model Estimation

The previous section gives some ways to correct remaining localization uncertainties following a normalization step while

inferring sulcal labels. Similar uncertainties may exist within the model estimation process, so we believe similar improvements are reachable. Now, the estimation of the sulcal model $\mathcal{M}$, presented earlier in section 3.3 (see in particular the equation 5), will be refined to also optimize the choice of subjects common space. From a training database $\mathcal{A}$ (in this work, we use the database described in section 2), this space is described by a set of unknown registration parameters θ_a (that moves the subject a from its own space to this common space). In this context, for each training subject a, labels $L_{a,i}$ of each sulcal piece i are known. Thus, the ideal Bayesian formulation of this problem is to find the best model parameters m^* and registration parameters θ_a for each training subject as below:

$$m^*, \{\theta_a^*\}_{a\in\mathcal{A}} = \underset{m,\{\theta_a\}_{a\in\mathcal{A}}}{\operatorname{argmax}} P(\mathcal{M}{=}m, \{\Theta_a{=}\theta_a\}_{a\in\mathcal{A}} | \{D_a, L_a\}_{a\in\mathcal{A}}) \tag{16}$$

A common approximation used to create *MRI* templates consists in first starting from an initialization and then alternating between building a template from registered subjects and subjectwise registration to this template, and so on. This idea is easily adaptable to our needs. From a raw referential space (provided by a first normalization step, see section 3.1 for details, represented by a set of transformations $\left\{\theta_a^{(0)}\right\}_{a\in\mathcal{A}}$), a first sulcal model (SPAM in our case) is computed. Then each subject is registered on it in a Bayesian sense (see below) which determines a refined common space from which the model is assessed again. This approach suggests to separate and alternate optimization of $\mathcal{M}$ and $\{\theta_a\}_{a\in\mathcal{A}}$:

$$\begin{cases} m^{(n)} = \underset{m}{\operatorname{argmax}}\ P\left(\mathcal{M}{=}m \middle| \{D_a, L_a\}_{a\in\mathcal{A}}, \left\{\Theta_a{=}\theta_a^{(n)}\right\}_{a\in\mathcal{A}}\right) \\ \left\{\theta_a^{(n+1)}\right\}_{a\in\mathcal{A}} = \underset{\{\theta_a\}_{a\in\mathcal{A}}}{\operatorname{argmax}}\ P\left(\{\Theta_a{=}\theta_a\}_{a\in\mathcal{A}} \middle| \{D_a, L_a\}_{a\in\mathcal{A}}; \mathcal{M}{=}m^{(n)}\right) \end{cases} \tag{17}$$

The considered localization model $\mathcal{M}$ is made up of one submodel $\mathcal{M}_l$ per label l where each can be estimated separately. Thus, for a given submodel $\mathcal{M}_l$ only sulcal pieces related to the same label are considered (denoted by $L_{a,i}{=}l$):

$$\begin{cases} m_l^{(n)} = \underset{m_l}{\operatorname{argmax}} \prod_{\substack{a\in\mathcal{A} \\ i\in\mathcal{E}_{a,l}}} P\left(D_{a,i} \middle| L_{a,i}{=}l, \Theta_a{=}\theta_a^{(n)}; \mathcal{M}{=}m\right) P\left(L_{a,i}{=}l \middle| \mathcal{M}{=}m\right) \\ \theta_a^{(n+1)} = \underset{\theta_a}{\operatorname{argmax}}\ P\left(D_a \middle| L_a, \Theta_a{=}\theta_a; \mathcal{M}{=}m^{(n)}\right) P\left(\Theta_a{=}\theta_a \middle| \mathcal{M}{=}m^{(n)}\right) \end{cases} \tag{18}$$

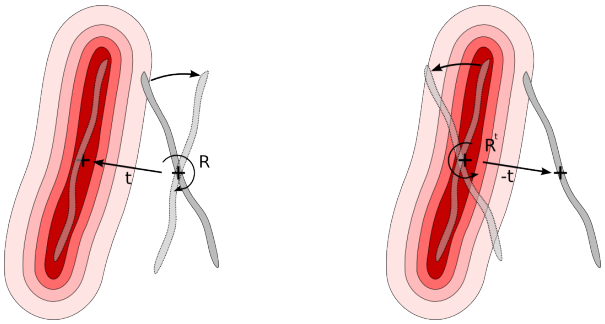
where $\mathcal{E}_{a,l}$ stands for the set of sulcal pieces of subject a restricted to those with the label l. We have omitted the model prior $P(\mathcal{M}{=}m)$ which only fixed internal parameters of the chosen localization model (see spam estimation in section 3.2). Once again we consider that our label prior $P(L_{a,i}{=}l|\mathcal{M}{=}m)$ is independent from registration parameters, so it can be estimated during the precomputing phase as described in appendix A.

Now, let us consider the optimization of a given subject a (the subscript a, θ_a and $m^{(n)}$ references are omitted here since they are not ambiguous) and take the logarithm of the second line of equation 18:

$$\begin{aligned} & \log\left[P\left(D|L,\Theta;\mathcal{M}\right)P\left(\Theta|\mathcal{M}\right)\right] \\ = & \sum_{i\in\mathcal{E}} \log\left[P\left(D_i|L_i,\Theta;\mathcal{M}\right)P\left(\Theta|\mathcal{M}\right)\right] \\ = & \sum_{i\in\mathcal{E}} \sum_{L_i=l} \delta_{l,l_i} \log\left[P\left(D_i|L_i{=}l,\Theta;\mathcal{M}\right)P\left(\Theta|\mathcal{M}\right)\right] \end{aligned} \tag{19}$$

where δ_{l,l_i} stands for the Kronecker delta which equals 0 if $l \neq l_i$ (labels are different) and 1 if $l = l_i$ (labels are the same). Here, l denotes the sum iterator whereas l_i is the true label (remember that labels are known on the training database). After rewriting the optimized expression we recognize the M step of equation 10 for joint labeling and registration. In other words, with suitable weights (posterior probabilities modeling uncertainties or strong knowledge of a known labeling) the same algorithm can be used to compute both.

Lastly, this iterative scheme gives a local maximum of the probability expressed in equation 16, close to the optimum if the initialization is good enough, which seems to be our situation.



subject/input referential model/destination referential

Figure 5: Two optimal referentials for the considered sulci with their respective minimal translation. Here, the red to white gradient shapes suggest an artificial 2D SPAM sulcal model, whereas the grey shapes represent a sulcus at several steps of a rigid registration. Left: the sulcus is first rotated by R around its center and then translated with t. Right: this shows the inverse transformation from the previous referential translated by t (which is more or less at the model center); so the sulcus starts from the model, then it is rotated around its center (this center is defined in the body of section 4.4) by R^t and then translated with $-t$.

4.4. Registration Priors

The joint approaches proposed for cortical folds labeling (section 4.2) and model refinement (section 4.3) include registration steps. Both cases are constrained by priors on registration parameters θ which limit the range of available transformations. In this paper, we focus on affine transformations and sulcuswise affine ones. The prior choice needs to fit the special nature of these measures. As previously stated, such transformations are made up of an invertible matrix A and a translation vector t. Difficulties are quite different for the two proposed joint normalization techniques and will be detailed below.

First, about the *global* approach only one affine transformation controls the registration, so only few parameters are tunable. Let us begin by the rigid restricted case for a better understanding. The process starts from an initialization close to the

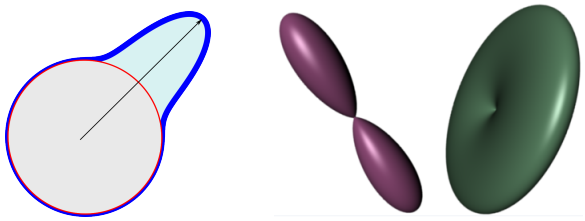


Figure 6: Left: Von Mises distribution (Mardia and Jupp, 2000) where the vector is the mean angle and the blue curve its likelihood function up to a scaling factor. Right: bi-modal Bingham (Bingham, 1974) distributions: the likelihood is proportional to the radial distance between the displayed surface and an unitary sphere. (a) is rather directional whereas (b) is uniform along a given plane.

solution, namely almost in adequation with the sulcal model likelihoods, therefore huge displacements are alreaday penalized, thus registration priors are not essential.

Now, the argument does not make sense anymore for affine transformations since subject and model adequation is measured up to scaling factors. To avoid divergence phenomena and numerical instabilities, these components are constrained around the original scale of data (see appendix B for details and formula) thanks to a gamma distribution.

From now on, we will focus on the proposed *local* registration scheme with one rigid transformation per sulcal label rather than affine transformation for computing facilities. If the necessity arises, additional scaling priors can be used as seen above. Previously, we have presented the sulcal models as a way to globally constrain the registrations. We study up to 60 labels, so most of the smallest ones are rather similar in shape, since they refer to elementary folds. Without any constraint, a fold may be moved to and get mixed up with matching neighbours. In this paper, our study will be limited to transformation priors satisfying the following independency assumption so as to benefit from mathematical reductions: $P(\theta) = \prod_{l \in \mathcal{L}} P(\theta_l)$.

Registration priors appear during model estimation from a training database and labeling of a test subject. For the first, the priors must be set a priori since no more accurate information is available, so priors have to be rather flat to prevent impossible registrations (local translation of 10 cm, rotation of π) but tolerate a large range of transformations that may appear on the training data.

For the second, registration priors can be either set a priori or learned empirically. We chose the second since the sulcal variability is not homogeneous and may depend on the considered dataset (if the considered subjects have specific common anatomical features), so a training database is used to estimate it, and model it. Next, their role will be to constrain parameters inference on test data during labeling. Thus, in reality, the learned priors depend on training data, so they write $P(\theta_l | \{\theta_{a,l}\}_{a \in \mathcal{A}})$ with $\theta_{a,l}$ the transformation parameters of a given subject a on the sulcus l, but now that this precision has been done, this dependency is made implicit to simplify formula notations. Lastly, during the modeling estimation step, each training subject has been registered to a common space. These transformations are used to define the range of available transformations to further register test data.

To optimize model design of such priors, a suitable formulation is needed. Indeed, the rigid transformation y of a point x writes: $y = R \cdot (x - g) + (t + R \cdot g - g) + g = R \cdot x_g + t_g + g$ where g denotes a referential point (the center of the rotation). Thus, depending on this choice the variability of the translation component t_g may change hugely. To mimize such a criterion the referential must be centered in the subject space (left of figure 5) or in the model space (right of figure 5). We derive in appendix C a way to estimate the optimal g in order to reduce translations variability.

Since suitable referentials g_l have been chosen for each label l, now prior choices will be discussed. From such referentials, sulcuswise translations t_l scatters spatially favouring some axes (in both directions since each referential is centered around the distribution of related translations). In this way, each translation prior can be modeled by a full-covariance 3D Gaussian:

$$t_l \sim \mathcal{N}(\mu_l, \Sigma_l) \tag{20}$$

with μ_l the Gaussian mean and Σ_l the Gaussian covariance matrix. The Gaussian parameters are set a priori at the model estimation stage: $\mu_l = \vec{0}$ and $\Sigma_l = (10 * I)^2$, namely an isotropic standard deviation of 1 cm. Then, during labeling the Gaussians are learned from translations measured over the training data.

The 3D rotation matrices group constitutes a curve algebraic subspace of the more general set of 3×3 matrices. Thus, a Gaussian distribution can not be straightforwardly derived from the matrix elements as can be done with translations data. The best way to do so is to use Gaussians generalization dedicated to Riemannian spaces (Boisvert et al., 2008; Pennec and Thirion, 1997) that straighten up locally curve spaces working in local tangential planes through standard log and exponential maps from differential geometry. To begin, an approximate formulation has been chosen in order to give some visual insights and computing facilities by simplifying rotations through a suitable parametrization: the 3D vector-rotation w (see appendix D for details) related to such Riemannian considerations. Lastly, the set of 3D rotations is a 3 dimensional manifold in the 9 dimensional space of 3×3 matrices. This parametrization is easy to understand since w direction ($\frac{w}{\|w\|}$) represents the rotation axis and $\alpha = \|w\|$ the angle of rotation in radians. This representation is unique as long as α is different from π or $-\pi$ (the cut locus of the considered manifold). Therefore, for a given label l, the rotation prior over R_l can be split into 2 parts:

$$P(R_l) = P(w_l) \approx P\left(\frac{w_l}{\alpha_l}, \alpha_l\right) \approx P\left(\frac{w_l}{\alpha_l}\right) P(\alpha_l) \tag{21}$$

where $\alpha_l = \|w_l\|$ and w_l is the vector-rotation parametrization of R_l.

Gaussian approximations dedicated to these particular data are now studied. Since α equals a norm, all rotations are counter-clockwise around their vector w which encodes the direction of the rotation. For a given label, we expect the studied set of rotations to be close from each other, so the same goes for their rotation axes, but their directions are grouped only up to a

	MRI registration			Global registration			Local registration	
	SPM	Baladin	**Talairach**	global basic	**global**	global affine	local Talairach	**local**
joint labeling and registration				✓	✓	✓	✓	✓
joint model estimation and registration					✓	✓	✓	✓
transformations type	affine	affine	affine	rigid	rigid	affine	sulcuswise rigid	sulcuswise rigid
initialization				Talairach	Talairach	Talairach	Talairach	global

Table 1: Description of the main differences of all SPAM-based models presented in this paper. The 3 main models are enlighten by a light gray backround. The initialization stage (last row) describes the relations between the models: the output of some models becomes the input of others.

sign since opposite rotations have opposite rotation-vector representation because α is always positive. Thus, only the axial part of rotations are modeled, namely w and $-w$ are considered equally, and in the other hand, the rotation angle is modeled up to a sign. Lastly, the axes are modeled by a Bingham distribution (Bingham, 1974) and the angle by a Von Mises distribution (Mardia and Jupp, 2000) (see appendix E for details, and figure 6 to see what the distributions look like).

4.5. SPAM and Registration

A SPAM model is non-parametric since it is currently represented as a 3D-volume of probabilities. Thus, energy derivatives-based optimization schemes could be limited by discrete gradient approximations. Therefore, a model-blind but efficient method has been preferred, which needs only a way to assess the optimized function at each tested value. To that end, we chose to use the well-known Powell method (Powell, 1964) to cover the registration parameters space. Be it affine or rigid, the rotations are represented through their vector-rotation parametrizations. Thus, the Powell method suggests transformations covering the parameters space along a set of axes (linear combination of canonical axes) which are refined iteratively during the optimization process according to the current value improvements.

5. Results

In the following, 8 models divided in 3 categories will be discussed, the following names are used throughout this paper. For each considered model, Left and Right hemispheres have been processed independently in hopes of better normalization adjustments.

The first category groups *MRI-based normalization* techniques (section 3.1) followed by sulcal model estimation (SPAM models in this study, see section 3.2) from this common space. The following names will be used to represent these models according to the applied normalization: **SPM-SPAM**, **Baladin-SPAM** and **Talairach-SPAM**. In the second category named *global registration*, labeling is done jointly with *global registration* (see equation 11): **global basic-SPAM** is based on SPAM models built from the Talairach space (the same used by the **Talairach-SPAM** model), **global-SPAM** and **global affine-SPAM** are based on SPAM models refined by global registration (using the method defined in equation 18). **global basic-SPAM** and **global-SPAM** models optimized one global rigid transformation (1 rotation and 1 translation) whereas **global affine-SPAM** optimizes an affine one (2 rotations, 1 diagonal scaling matrix and 1 translation). These 3 studied global approaches are based on the **Talairach-SPAM** one as an initialization since it is the default approach in the BrainVISA anatomical pipeline, which gives a good starting point for further rigid refinements. The last category deals with *local registration*: **local Talairach-SPAM** uses the **Talairach-SPAM** model as an initialization whereas **local-SPAM** is based on the **global-SPAM** model. Both models are built with local registration refinements (sulcuswise rigid transformations: see equation 18) and labeling inference is coupled with local registration optimization (see equation 13) with sulcuswise registration referential points (defining center of rotations and how to apply translation a priori, see section A) determined from subject space. The features of these models are summarized in the table 4.4.

5.1. Error Measures

Since many different error measures are used in other studies, several measures are suggested here to make comparisons easier with these works.

The more reliable the models are, the harder their comparaison is since they do much less errors, so some local improvements become exceptional. Then it could be hard to determine if the result reaches the significance level. That is the reason why we have looked for more sensitive measures to assess more and more accurately the quality of our models.

The following measures rate a single subject and assess the quality of its labeling (automatic labels) inferred from a sulcal model, comparing it to the known ground truth (manual labels). First of all, let us introduce some notations: $FP(l)$, $FN(l)$, $TP(l)$ stands respectively for false postive (sulcal pieces automatically labeled with l whereas their true labels are different), false negative (sulcal pieces manually labeled with l whereas their automatic labels are different) and true positive (sulcal pieces where both manual and automatic labels are l) measures of label l. For each of these measures, the retained value is the cumulative size (number of voxels in the subject space) of sulcal pieces verifying the specified conditions.

Historically (used in our previous works (Rivière et al., 2002)), E_{mass} measured only the proportion of false positive (or false negative) errors:

$$E_{\text{mass}} = \frac{\sum_{l\in\mathcal{L}} FP(l)}{\sum_{l\in\mathcal{L}} FP(l) + TP(l)} = \frac{\sum_{l\in\mathcal{L}} FN(l)}{\sum_{l\in\mathcal{L}} FP(l) + TP(l)} \quad (22)$$

Next, we derive sulcuswise local measures which take into

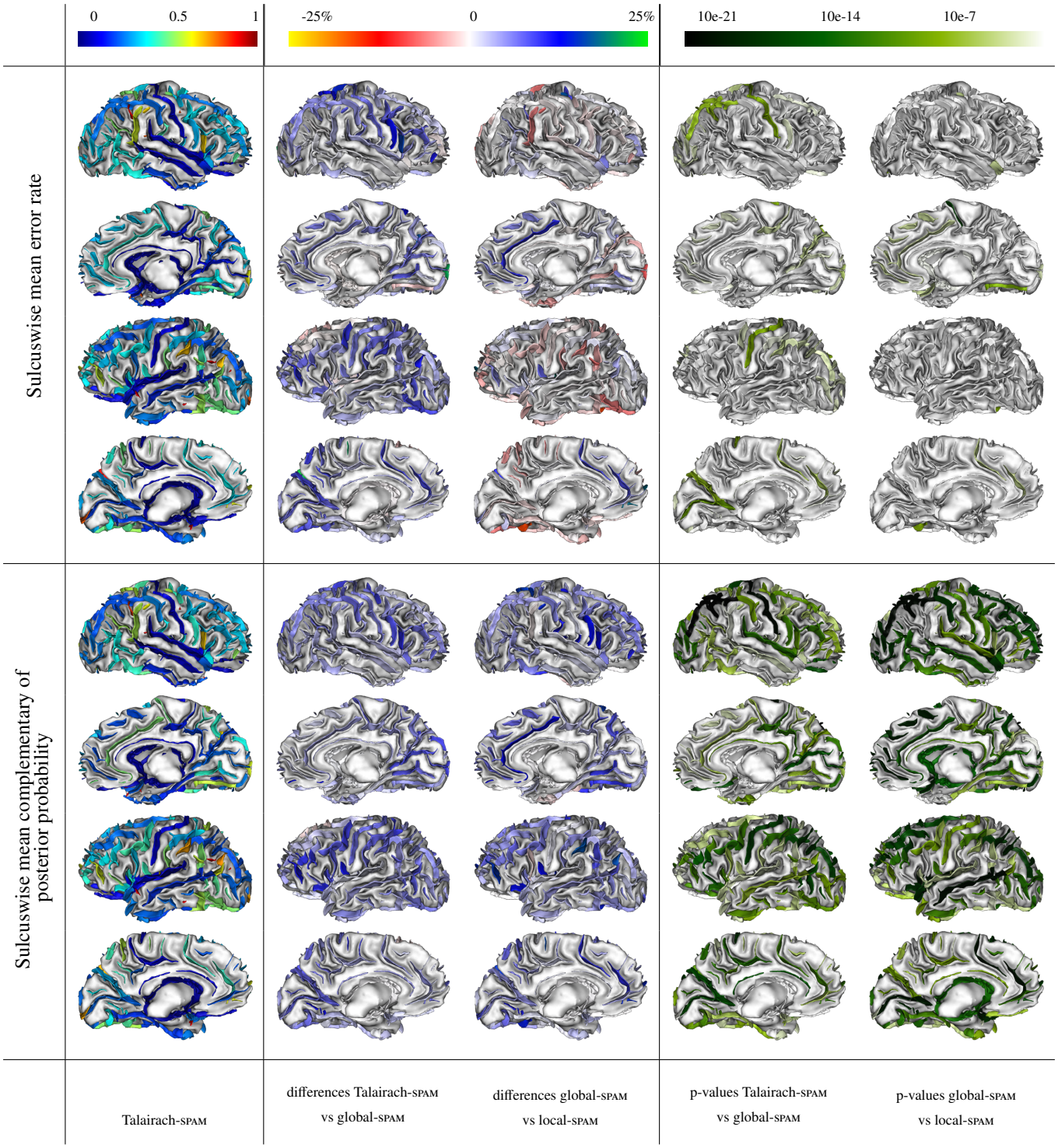


Figure 7: Local improvements of global and local approaches and their significance. Top: sulcuswise leave-one-out mean error rates (see equation 23). Bottom: sulcuswise leave-one-out mean complementary of posterior probabilities (see equation 25). All results are mapped on manual labels of only 2 subjects (one for left hemispheres and another one for the right ones) chosen because they present most of the studied sulci. 3 columns: left: error measures from the *Talairach*-SPAM model. Middle: local improvements (error measure differences) of *global*-SPAM versus *Talairach*-SPAM (left) and *local*-SPAM versus *global*-SPAM. Right: logarithmic mapping of two tailed corrected p-values (using the Bonferroni correction, dividing uncorrected p-values by the number of considered sulci) computed thanks to a Wilcoxon test on the respective compared models. It measures the statistical significance of both increase or decrease of error rates. Besides, in the above figures, most decreases are not significant. See the begining of section 5 for details on compared models. The posterior-based measure (bottom) gives rather similar improvements in mean than the label-based one (top) for the *global*-SPAM model, and slightly better improvements for the *local*-SPAM model. In the other hand, results are much more significant in the case of posterior-based measure (see p-values: light green color already means significant results with p-values inferior than $5.10e^{-2}$, whereas most are green or dark green). Thus, this measure is more sensitive to slight improvements.

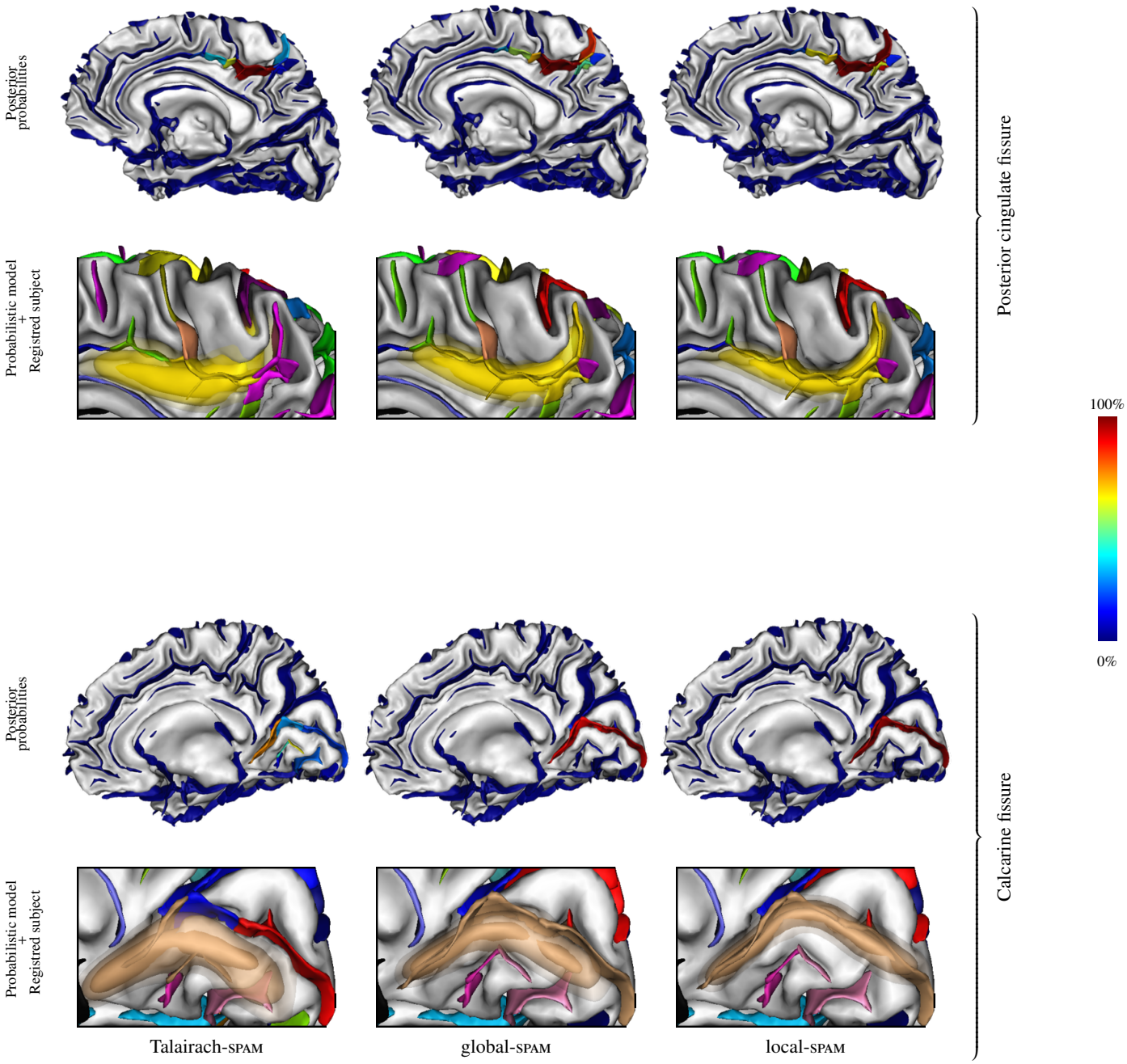


Figure 8: Effects of joint labeling and registration (global and local) on 2 primary sulci: the posterior cingulate fissure (F.C.M.post.) and the calcarine fissure (F.Cal.ant.-Sc.Cal.). The illustrated phenomena have been established on all 62 subjects but we focused here on those which give the best improvement for the concerned sulci. Top: local posterior probabilities of F.C.M.post. (Left) and F.Cal.ant.-Sc.Cal. (Right), mapped on each sulcal piece of the respective subjects. Bottom: zoomed and cropped details of studied sulci and their neighbourhoods, sulcal pieces are colored according to labels inferred from several SPAM models (left to right: *Talairach*-SPAM, *global*-SPAM and *local*-SPAM), and mixed with visual representation of the sulcal distribution of interest. Both illustrations emphasize the interest of the joint approach for labeling (for *global*-SPAM approach only since *local*-SPAM one gives rather similar labels) and in term of local posterior probabilities (for *global*-SPAM and *local*-SPAM approaches, since subjects sulcal data and sulcal models are better and better superimposed). Details of considered models are given at the begining of section 5.

Testing results:

		MRI registration			Global registration			local registration		Previous
		SPM	Baladin	**Talairach**	global basic	**global**	global affine	local Talairach	**local**	Brainvisa model
E_{SI}	Left	18.06 (5.31)	17.52 (4.59)	17.55 (5.93)	15.27 (2.74)	14.59 (2.88)	14.31 (2.87)	16.79 (4.35)	14.22 (2.96)	21.37 (5.34)
	Right	16.65 (4.61)	15.82 (3.79)	16.83 (3.77)	14.70 (3.09)	13.97 (2.91)	13.80 (2.62)	15.61 (3.72)	13.48 (3.13)	20.12 (5.05)
E_{mass}	Left	17.24 (4.95)	16.84 (4.21)	16.64 (3.69)	14.56 (2.54)	14.01 (2.62)	13.75 (2.57)	16.61 (4.17)	14.11 (2.79)	23.63 (5.42)
	Right	15.87 (4.37)	15.17 (3.63)	15.96 (3.66)	13.95 (2.98)	13.37 (2.82)	13.17 (2.52)	15.29 (3.51)	13.30 (2.99)	22.12 (5.05)

Training results:

		MRI registration			Global registration			local registration		Previous
		SPM	Baladin	**Talairach**	global basic	**global**	global affine	local Talairach	**local**	Brainvisa model
E_{SI}	Left	15.09 (5.30)	14.78 (4.34)	14.34 (3.51)	13.11 (2.78)	11.61 (2.38)	11.64 (2.49)	14.71 (4.23)	11.97 (2.30)	17.11 (4.32)
	Right	14.29 (4.58)	13.43 (3.55)	13.64 (3.13)	12.66 (2.68)	11.22 (2.39)	11.66 (2.26)	13.93 (3.66)	11.96 (2.77)	16.86 (4.78)
E_{mass}	Left	14.44 (4.95)	14.19 (4.10)	13.65 (3.26)	12.55 (2.46)	11.23 (2.27)	11.26 (2.32)	14.63 (4.16)	11.89 (2.05)	19.26 (4.54)
	Right	13.66 (4.34)	12.88 (3.20)	12.87 (2.87)	12.08 (2.57)	10.80 (2.27)	11.19 (2.15)	13.72 (3.46)	11.81 (2.62)	19.03 (4.89)

SPAM-based models from this paper

Markovian, SVM-based model from Rivière et al. (2002)

Table 2: Top: Leave-one-out mean (over 62 subjects) labeling error measures (in %): SI and mass error rates and their standard deviations between parenthesis for all studied models (see main text, section 5 for detailed descriptions). Bottom: same measures but only one model has been learned from the whole database and all the subjects have been labeled from it. Left: results related to all the SPAM-based models used and described in this paper (new methods proposed by the Brainvisa software). The results of the 3 main models are enlightened by a light gray backround. Right: best results obtained previously by the standard model of the Brainvisa software (Markovian, Support Vector Machine-based model as described in Perrot et al. (2008); Rivière et al. (2002)) in the same experimental conditions (see section 6 for details and comments). The differences between training and testing results are in the order of 3% which proves that the localization models are rather reliable and generalize well to unknown data. Notice that training results between both hemispheres are more balanced than testing ones which suggest that left models are slightly less robust since left hemispheres are more variable.

account all errors involving a given label l (missing and overmuch sulcal pieces):

$$E_{\text{local}}(l) = \frac{FP(l) + FN(l)}{FP(l) + FN(l) + TP(l)} \quad (23)$$

This measure allows to see local improvements which could be drown into noise at the whole brain scale.

The last labeling error proposed here synthetizes above local errors to a single measure which refines the historical global error E_{mass}. It draws inspiration from the measure called similarity index: SI (Yang and Kruggel, 2007):

$$E_{\text{SI}} = \sum_{l \in \mathcal{L}} w_l \frac{FP(l) + FN(l)}{FP(l) + FN(l) + 2 * TP(l)} \quad (24)$$

where $w_l = \frac{s_l}{\sum_{l \in \mathcal{L}} s_l}$ with $s_l = FN(l) + TP(l)$ is the true size of the sulcus l.

Each component of the sum over labels differs on two points compared to the local measure. First, true positive measures $TP(l)$ count twice as false positive $FP(l)$ and negative ones $FN(l)$, in order to remove errors shared by several labels, since each extra sulcal piece for a given label is a missing part for another label. Second, each component (ranging between 0 and 1) is weighted according to the sulcus true size so that each local component count as much as its size. In a practical way, this measure seems more draconian than the historical one since the first is always higher than the second on our results.

All these measures depend strongly on the choice of automatic labels. Sometimes, several labels are closely in competition with their label posteriors $P(L_i|D_i)$ almost equal on a given sulcal piece i. In this case, the infered label is rather arbitrary. Besides, two models can provide labeling very alike from each other, whereas their posteriors are quite different which means that both models agree but one may be more confident than the other since its posteriors are more uneven. Thus, we propose a new local measure based on local posteriors of labeling, namely we measure the mean posterior of a studied label l over sulcal pieces which are really labeled with this same label:

$$E_{\text{post}}(l) = \frac{\sum_{i \in \mathcal{E}_l} s_i P(L_i = l|D_i)}{\sum_{i \in \mathcal{E}_l} s_i} \quad (25)$$

where $\mathcal{E}_l$ is the set of sulcal pieces labeled manually with label l and s_i the number of voxels of the sulcal piece i.

*5.2. Study **1**: on 62 subjects*

5.2.1. Assessment Strategy

Previously, in past works (Rivière et al., 2002), we split our former database (26 subjects) in a training (21 subjects) and a testing one (5 subjects). This scheme was not satisfactory since not enough data were used to assess the model quality or to build the model. Moreover, in this context, the results were quite sensitive to the choice of subjects for the split.

To answer to this issue, we used a leave-one-out validation scheme. Namely, for each subject we derive a training

database including all subject except the studied one, and a testing database composed of the subject laid aside. Thus, for each hemisphere, 62 models (as many as subjects) are estimated (computing of localization models: SPAM, of registration parameters prior models, label priors) from each training database. Then 62 leave-one-out testing subjects are labeled each from its respective model. Thereby, all these results are much more reliable than our previous ones.

The estimation of all parameters of the most advanced proposed models last about 3 hours, so the full computation for both hemispheres and for all 8 studied models reaches more than 125 days cumulating. Therefore, we developped grid computing strategies to make this study possible and reduce computation time by a factor of about 50.

5.2.2. Labeling and Registration

For all the following results, the reader is referred to table 2 for global error measures, figure 2 for visual details on central sulcus and parieto occipital fissure alignments, figure 7 for local improvements.

The first comparison refers to affine registration and labeling. Using first an affine registration set a priori and then dealing with sulcal labels gives rather similar results for the 3 tested approaches (Talairach, SPM and Baladin) with a certain preference for Baladin which gives even better results than the manual registration to Talairach. With these 3 techniques, common and severe errors are observed even for the largest sulci like misplaced chunks of sulci or double delineations. Besides, the quality of sulcal pieces segmentation may considerably affect the emergence of such errors.

Previously, we have tested Markov-based approaches (Perrot et al., 2008; Rivière et al., 2002) that prevent such errors to some extent, but their study is beyond the scope of the current paper. We noticed that many of these issues are induced by erroneous normalizations resulting in a bad adequacy between subjects compared to a sulcal model. Thus, compared to sulci labeling from an a priori defined common space (*Talairach*-SPAM), our joint global approaches (rigid or affine) (*global*-SPAM or *global affine*-SPAM) bring significant improvements in sulcal alignments and so sulcal labeling. The significance has been assessed by the study of signed differences (decrease of errors) of subjectwise global mean errors through a Wilcoxon test. Left hemisphere-based models (*Talairach*-SPAM versus *global*-SPAM for instance) give a p-value of $6.5e^{-8}$ and right ones $6.2e^{-9}$. In table 2, the columns titled *Global registration* show the interest of using global normalization refinement during labeling (*global basic*-SPAM) and model estimation (*global*-SPAM and *global affine*-SPAM), each step reduces uncertainties and improves labeling. The global affine version is slightly better than the rigid one but is harder and longer to optimize. Besides, we need constraints on the rotation defining the axes along which affine scaling factors are applied to add robustness and efficiency to the process. We also measured thanks to an entropy differences criterion (see figure 4) that each SPAM-based sulcal model is sharper with global refinement (*global*-SPAM) than without (*Talairach*-SPAM) so the sulcal variability has been reduced in this optimized common space.

Whereas most sulci are rather-well aligned onto this probabilistic template, others may be shifted locally. The use of local registration from a well-chosen common space allows to enhance or even fix local fits. The Talairach space is not good enough to define suitable priors (section 4.4 for details on priors) as hard constraints. Their labeling results (*local Talairach*-SPAM) improve SI error according to those from *Talairach*-SPAM model, but worsen them compared to *global*-SPAM one. Lastly, first starting from the global optimized space (*global*-SPAM) and then locally optimizing the matching gives some improvements (*local*-SPAM). In fact, the labeling results are rather close and the mean error rates improvements (see table 2) reach the significance limits (p-values of $3.7e^{-2}$ for the left hemispheres and $4.1e^{-2}$ for the right ones with a one sided Wilcoxon test). Only a few sulci present quite significant improvements (see Bonferroni corrected p-values on top-rows of figure 7). Nevertheless, during the same time, local posterior probabilities significantly increase (bottom-rows of figure 7) which means a better adequacy between each labeled subject and this model. However, these results are partly explained by sharper SPAM models which give higher posterior probabilities, but on condition that the alignment is correct which is performed by the local registration. To realize concretely these effects, the reader is referred to the figure 8 where the enhancements are detailed on two sulci of two specific subjects.

Lastly the best labeling results are obtained with *local*-SPAM with a leave-one-out global mean error rate (E_{SI}) of 14.22% for the left hemispheres and 13.48% for the right ones, that is, respectively, a recognition rates of 85.78% and 86.52%.

Some training results are presented in table 2 to highlight possible robustness issues. Basically, the differences of error rates between testing and training results are comparable from one model to another or even decrease for local registration on right hemispheres, which states that models complexity does not seem to harm the robustness. On the other hand, training results are more balanced between both hemispheres than testing ones, which reveals a lack of robustness in the face of the higher variability of left hemispheres.

*5.3. Study **2**: varying database size*

The previous study gives good insights on the quality of the proposed sulci identification models in regard to a database of 62 manually labeled subjects. Most certainly, this amount of data should be quite representative of the true anatomical variability of most well-understood and less-variable folds, but more subjects may be needed for secondary ones. Lastly, these observations only reflect a statistical estimation issue. In the following, we suggest a way to point out which parts of our models lack information and limit their current performances.

5.3.1. Description

The main idea of this study is to assess the variation of labeling error rates (global and sulcuswise results) in regard to the number of subjects used to estimate the proposed models in order to extrapolate results behaviour beyond the current database.

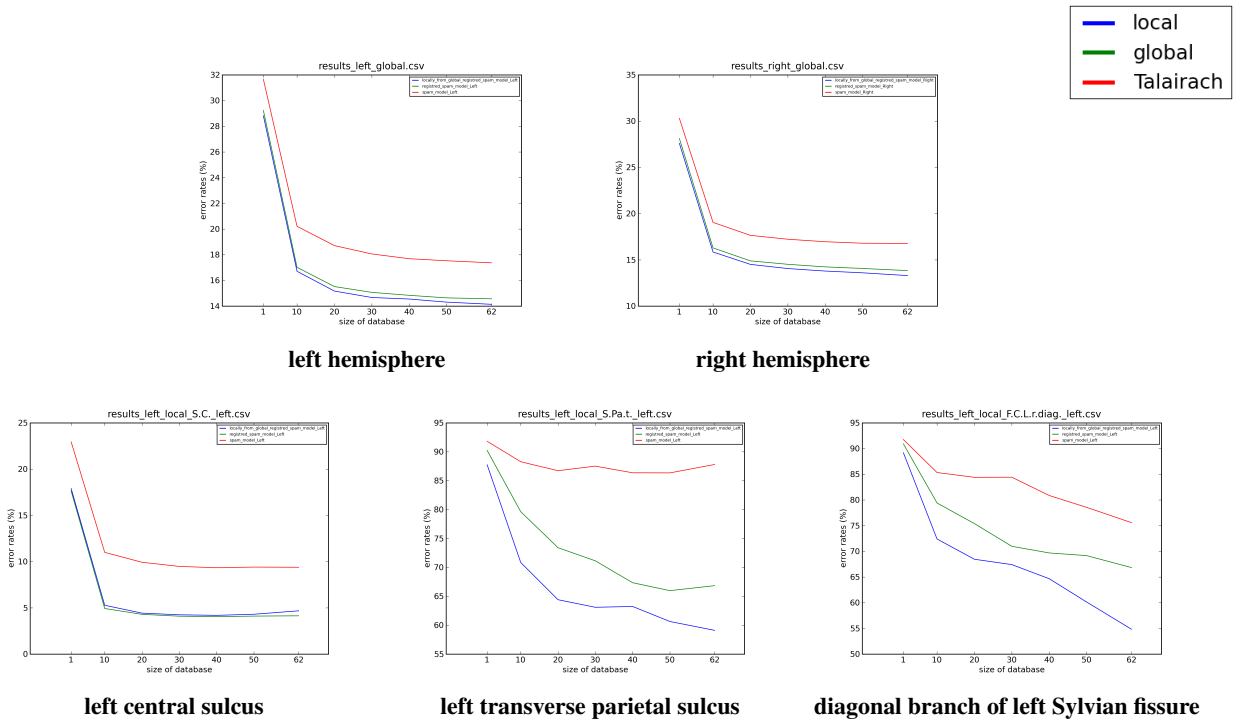


Figure 9: Error rates changes while database size is increased on the following models: Talairach-SPAM (red), global-SPAM (green), local-SPAM (blue). **Top**: global results. **Bottom**: typical local results: all with decreasing exponential shape, but varying convergence rates.

The followed strategy consists in selecting randomly n subjects from a full database of 62 subjects to make up a training database used to estimate each studied model (Talairach-SPAM, global-SPAM and local-SPAM) and the $(n - 62)$ remaining subjects are used to test these models through labeling error rate assessments. For each studied database size, this operation is repeated 100 times to sample the possible combinations of drawing n subjects among the full database. In a practical way, the sampling may artificially favour the drawing of some subjects. This biases averaged measures without any caution. A standard solution is to first compute one sampling mean over all trials per subject and then compute the mean of these subjectwise measures, which is unbiased since each subject has the same weight. Special cases are followed when n equals 1 or 62 since the sampling step is not needed anymore. Actually, a standard leave-one-out scheme is used instead, over training or testing data depending on the matter.

5.3.2. Comments

The estimation of SPAM-based models relies on a convolution kernel. In this study its width has been specified empirically for the whole database to 2 mm and remains fixed while the number of subjects varies. Actually, its optimal value strongly depends on the number of subjects: the less subjects are considered, the larger the kernel width becomes and conversely. This aspect goes beyond the bounds of this paper and will not be explored further since it should only affect the details of the presented results but not the main message.

5.3.3. Results

The results of the current study are depicted in figure 9. Top and bottom rows show respectively global and local sulcuswise error rates changes in regards to the number of subjects of the training database. At first glance, for all models, performances order is overally preserved. In greater details, error rates follow a natural decreasing exponential curve since each new subject or manually labeled sulcus yield all the less additional information as the database size gets larger. The actual statistical analysis of our results is rather difficult to cope with because of considerable dependencies between boostraped training and testing sets whose degree of freedoms is affected in an unknown measure. Besides, this influence most probably varies with the database size. So, the following results must be seen as tendencies whose statistical interpretation is limited for the time being.

Global results seem to reveal that the current number of subjects is almost sufficient to reach the limit of the proposed models since asymptotes of these curves are almost reached. On the other hand, local results reveal a range of specific behaviours which leaves room for improvement. Indeed, whereas some local models are quite well estimated with only few subjects (for instance, for central sulcus, 10 subjects seem to be enough, see figure 9 bottom left), other models still benefit from additional information with the full database (for instance, the curve of the diagonal branch of the left Sylvian fissure has not even reached its asymptote with 62 subjects, see figure 9 bottom right). The behaviour related to a given cortical fold depends on its inter-subjects variability and the complexity (degree of freedom) of the considered model. Indeed, if we consider, for instance, the case of the left transverse parietal sulcus, about 20 subjects seem enough to reach the rather limited maximum capacity (error rate of 90%) of the Talairach-SPAM model, whereas about 60 subjects are needed by the global-SPAM approach to converge to

an error rate of 70%. Lastly, the full database seems not even sufficient to estimate the best possible model based on the local-SPAM technique.

Rather than blindly increasing the number of manually labeled database, this study yields a way to limit the tedious manual labeling only to the few sulci models which may benefit from a better estimation. Lastly, current results could be efficiently improved with such targeted additions.

6. Discussion

6.1. Labeling Independence Assumptions: Known Limitations

A sulcus is generally made of several sulcal pieces, layed close to each other. During the brain development and afterwards even in the adulthood these structures can touch, push or shift their neighbourhood. This implies strong dependence between distinct neighbouring sulcal pieces. Thus, the independance assumption in the proposed models is somewhat transgressed. Actually, the approximation remains fully valid only for the main parts of the primary sulci (exceptions exist even on healthy subjects: see figure 11), but less valid for quite variable secondary folds like branches. It is clear that this approach does not give definitive answers to sulci labeling. Nonetheless, starting from an extremely simplified initial model (section 3.3), we have developped in this paper new and refined models (section 4.2 and 4.3) that overcome the limitations of the latter.

In fact, human experts identified sulcal branches step by step thanks to their neighbouring folds from the most reliable to the least reliable. This suggests the use of a graphical model such as a Markov field, as proposed in our previous works (Perrot et al., 2008; Rivière et al., 2002). Loopy Belief Propagation algorithm could mimic the knowledge diffusion process (as showed by Shi et al. (2009) in the context of sulci labeling) and efficiently replace a standard optimization strategy scheme. Indeed, to couple such models with our joint labeling and registration framework, new approximations still have to be defined to keep moderate computing times.

6.2. Comparison with Previous BrainVISA Models

The interest of neighbouring information to sulcal labeling needs no further proof, that is why we paid attention to it in previous works (Rivière et al., 2002) and tried to make use of it in the context of SPAM models (Perrot et al., 2008) but with moderate success (it is without doubt owing to difficulties we encounter when mixing descriptive and discriminative modeling aspects). These previous approaches are valuable but they do not take into account local information. The difference between the two methods turns to be large and statistically significant.

Let us consider the Markovian and SVM[1] (Support Vector Machine)-based model from Perrot et al. (2008). The last column of table 2 shows global errors (E_{SI} and E_{mass}, section 5.1) under the same experimental conditions as SPAM-based models using a leave-one-out evaluation scheme. Compared to the previous model proposed in BrainVISA, even the basic SPAM models that learned from the Talairach space (MRI registration/Talairach column) performed significantly better (see figure 10 for details). Even with this simplest spam model that used basic shape matching and location information, the error decreased by 3% (with a one-sided Wilcoxon-based p-value of $1.6e^{-7}$ for the left hemisphere and $2.3e^{-6}$ for the right one) compared to the previously published model (Rivière et al., 2002), even though the latter made use of much richer information (various location, high level shapes and topological neighbourhood descriptions). In fact, the main flaw of the previous model lies in its complexity and the difficulties to interpret model parameters. That is why we have explored new perspectives in this paper with the intention of reintroducing the complex high level features but in an improved basis.

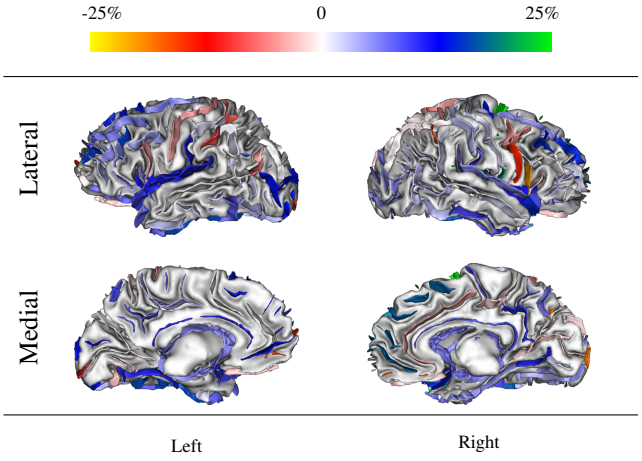


Figure 10: Sulcuswise leave-one-out mean error rates differences highlighting local improvements of Talairach-SPAM model against the previous Brainvisa model (based on a Markov field of sulcuswise SVM experts). Most of the improvements are even visible in the comparison with the Talairach-SPAM model. Results from global-SPAM and local-SPAM models give similar tendencies. Locally, only a few sulci errors increases with the use of SPAM models. In these areas the neighbourhood modeling seems to surpass location informations.

6.3. Comparison with Other Models

The comparison with other labeling strategies from the literature should provide valuable insights to the brain mapping community. There are several issues to deal with in order to reach this aim. A combined effort from the community should be required to standardize the anatomical definition of sulci and build bridges across the various mathematical representations : sulcal lines (Hurdal et al., 2008; Lohmann, 1998; Kao et al., 2007; Seong et al., 2010; Shattuck et al., 2009; Shi et al., 2008; Tao et al., 2002; Thompson and Toga, 1997; Tosun and Prince, 2008; Tu et al., 2007), surfaces (Goualher et al., 1997; Vaillant and Davatzikos, 1997; Zhou et al., 1999), voxels (Klein and Hirsch, 2005; Le Goualher et al., 1998; Perrot et al., 2008; Rivière et al., 2002), with or without branches or breaks. This im-

[1] SVM-based models from Perrot et al. (2008) are, from a software point of view, technically easier to handle than MLP (Multi-Layer Perceptron)-based ones from Rivière et al. (2002), particularly in order to extend it for cross-validation purpose. Besides, they reach the same performances under various conditions which allow the comparison with both models. The computation of all models estimation and labelings required more than 5500 CPU hours and the use of cluster computing.

plies various kind of error measures. It mainly concerns metric-based approaches in the case of shape inference and classification error rates in the case of labeling.

6.4. Registration: Known Limitations

Even in the case of the non-linear extension we proposed in this article, there is another well-known situation where the independence assumptions we have made upon the labeling or between local transformations are no longer confirmed. It concerns outlier subjects with huge consistent shifts. Such shifts up to 2 cm exist on healthy subjects (for instance, see the central sulci of the subject in figure 11), but they happen too rarely to be estimated accurately from one hundred subjects. Further prior modeling would be needed to allow this kind of transformation.

The proposed piecewise rigid transformations model would benefit from coupling all the local estimations into a global deformation field. In this respect, polyaffine diffeomorphic transformations (Arsigny et al., 2005) accounts for a good candidate to yield a continuous and smooth normalization from a set of local landmarks with locally affine moves. The current known variability of the largest sulci suggests to use more than one control point on some of them, which should be defined from reliable structures like sulcal roots or sulcal basins. Further non-linear extensions which have already been used with the proposed labeling framework (Auzias et al., 2009) are also envisaged like discrete measures or currents (Glaunès et al., 2004; Glaunès and Joshi, 2006).

More generally, any state of the art diffeomorphic registration techniques (Ashburner, 2007; Avants and Gee, 2004; Joshi et al., 2004; Yeo et al., 2010), which provides smooth, invertible and non-linear transformations, should give better normalization results than those obtained here in cases where neighbouring folds ambiguity is limited. But even better results could be expected using the proposed general framework with the integration of anatomical knowledge through joint labeling estimation. However, a balance needs to be respected between the computing overhead of such advanced transformations and the actual needs of such precision, which may be excessive, in the context of sulci labeling. Finally, such integration should be carried out with caution, with the introduction of suitable priors over transformation parameters.

To conclude, the proposed non-linear registration remains basic but rather reliable since it is based on implicit sulcal constraints, without requiring any manual delineation which are quite difficult to define. It would be interesting to assess the impact of the suggested normalization techniques on functional activation detection or white fiber bundles clustering.

7. Conclusion and Perspectives

A very important advantage of the models proposed in this paper is that they use only few assumptions to characterize sulcal data, especially the fact that they can be well localized up to an unknown normalization. Indeed, even priors used for the labeling step are learned empirically from a database. No specific case has been assumed for any sulcus. Thus, these models are generic and can adapt easily to other sulci definitions or even to other anatomical structures (e.g., internal structures like deep gray matter nuclei or largest well-known white matter fiber bundles).

We have explored several original methods to label sulcal pieces from localization information. Sulci variability has been modeled through a mixture of SPAM models, inspired by the work of Le Goualher et al. (1998), that offer visual information on the mean position and shape of cortical folds. First, we have proposed a method based on a normalization technique working either on labeled or unlabeled subjects. To provide straightforward comparisons with the proposed models, we have focused only on standard affine normalization techniques: Talairach, SPM and Baladin. In future work, we will compare the suggested sulcuswise registration and labeling approach to standard non-linear registration techniques.

Secondly, we have proposed a refinement of this approach by reducing the uncertainty of affine normalization with either model estimation or labeling of new subjects. This method gives significant improvements compared to those of the first category, both regarding sulcal labeling and the quality of affine normalization. Indeed, sulci alignment is considered to be a good criterion to validate normalization techniques (Fischl et al., 1999; Hellier et al., 2003; Van Essen, 2005; Eckstein et al., 2007; Tosun and Prince, 2008). The proposed global approach maximizes these alignments even if labels are unknown, and outperforms all other tested affine registration methods. Thus, the generalization capacities of SPAM models behave like an indicator of the quality of the common space upon which the model has been estimated.

Lastly, we have suggested a non-linear extension based on sulcuswise locally rigid transformations. They proved fast and efficient thanks to sensible model approximations and suitable priors. In most cases, this approach results in labelings very similar to the results of global approach, but their posterior probabilities (how much the model is confident in the proposed labeling) of real labels are significantly improved. This indicates that uncertainties have been reduced. This could be useful to detect abnormal folding. Nevertheless, significant improvements have been assessed in frontal medial areas.

From now on, as stated before, we we have practically reached the limit of current experts knowledge. For this purpose, first, we will continue our efforts to ensure better normalization of sulcal data and enhance the adequation between tested subjects and models through the integration of smooth and continuous transformations. Second, we will extend these models with the key features used by the previous BrainVISA's sulcal labeling model (Rivière et al., 2002), namely, high level sulcal shape and neighbourhood features.

Acknowledgements

The database used in this paper could not have existed without the worthwhile help of Alan Tucholka, a neuroimaging scientist involved in anatomo-functional studies, for his support for the manual labeling.

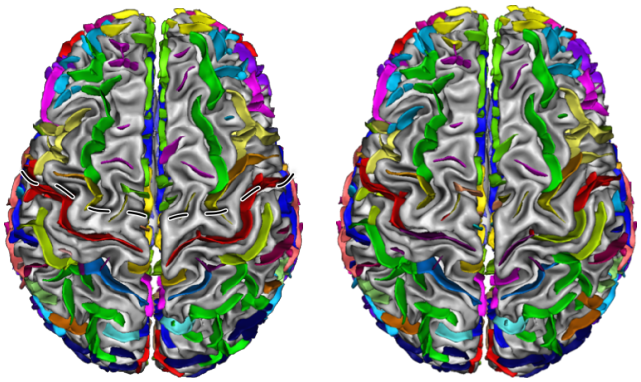


Figure 11: This subject, represented with its true manual labels on the left and with labels infered from local-SPAM method on the right, is the best illustration of the limitations of the presented joint labeling and normalization techniques, induced by local transformations independency. In fact, the central sulci (red color at the middle of the image) and therefore their neighborhood (pre-central and post-central sulci, posterior parts of medial frontal sulci and anterior part of parietal sulci) are shifted posteriorward up to 2 cm, compared to standard locations represented here by dashed lines. These huge displacements need coherent translation moves along the brain and maybe with constraints linking both hemispheres as suggested by this subject.

A. Labeling Prior Optimization

Let us consider the optimization of labels priors from the training database. We have, for each subject: $P(L_a|\Theta_a;\mathcal{M}_p) = \prod_{i\in\mathcal{E}_a} P(L_{a,i}|\Theta_a;\mathcal{M}_p)$ with $\mathcal{E}_a$ the set of sulcal pieces of the subject a and $L_{a,i}$ the label given to the sulcal piece i of the subject a. Since labels prior modeled the chance to have one labeling rather than another without any consideration from subject data (be it structural, size, shape, or position), the prior values are limited by the range of possible combination of labels. The labels are considered independent and the number of labels is limited, so are the possible values taken by label priors. So, we rewrite $P(L_{a,i}{=}l|\Theta_a;\mathcal{M}_p) = \pi_l$ for whatever a and i, with $\sum_{l\in\mathcal{L}} \pi_l = 1$, where $\mathcal{L}$ is the set of all considered sulcal labels. Lastly, the optimization of labels priors writes:

$$\begin{aligned} \pi^* &= \underset{\pi}{\operatorname{argmax}} \prod_{a\in\mathcal{A}} \prod_{i\in\mathcal{E}_a} P(L_{a,i}|\Theta_a;\mathcal{M}_p) \\ &= \underset{\pi}{\operatorname{argmax}} \prod_{l\in\mathcal{L}} {\pi_l}^{|l|} \qquad \text{with } |l| \text{ the number of sulcal pieces with the label } l \end{aligned} \tag{26}$$

Using the logarithm of the previous equation, the optimization problem is:

$$\begin{cases} \pi^* = \underset{\pi}{\operatorname{argmax}} \sum_{l\in\mathcal{L}} |l| \log(\pi_l) \\ \text{s.t} \quad \sum_{l\in\mathcal{L}} \pi_l = 1 \end{cases} \tag{27}$$

where $\pi = \{\pi_l\}_{l\in\mathcal{L}}$.

We now introduce the lagrangian parameter λ to incorporate the constraint into the maximization problem:

$$E = \sum_{l\in\mathcal{L}} |l| \log(\pi_l) + \lambda \left(1 - \sum_{l\in\mathcal{L}} \pi_l\right) \tag{28}$$

The optimal parameter values cancel partial derivatives:

$$\begin{aligned} \frac{\partial E}{\partial \pi_l} &= \frac{|l|}{\pi_l} - \lambda = 0 \\ \text{then,} \quad \pi_l &= \frac{|l|}{\lambda} = \frac{|l|}{\sum_{l\in\mathcal{L}} |l|} \end{aligned} \tag{29}$$

since π_l is normalized.

Thus, according to the independence assumption made over the labeling prior, the optimal prior of having the label l on a given sulcal piece is the frequency of finding this label on sulcal pieces over the training database.

B. Affine Registration: Scaling constraints

The linear part of an affine transformation writes: $A_g = U_g D_g V_g^t$ with U_g and V_g two rotation matrices and D_g a diagonal matrix of scaling components. We do not have any a priori on the directions (defined by the rotation V_g^t) on which scaling factors are applied. So as to satisfy the constraints of positivity of the scaling factors and keep the prior as simple as possible, a standard gamma $\Gamma(\cdot,\cdot)$ distribution defined by its shape parameter k and its scaling parameter s has been choosen. Its mean complies with the constraint $ks = 1$, namely the chosen arbitrary but natural scale is the original scale of the data.

$$\begin{aligned} &P(\theta_g) = P(D_g) = \prod_{i\in\{1,2,3\}} P(D_{g,i}) \quad \text{where each } D_{g,i} \sim \Gamma\left(k, \frac{1}{k}\right) \\ &\text{so,} \quad \log\left[P(\theta_g)\right] = (k-1) \sum_{i\in\{1,2,3\}} \log\left[D_{g,i}\right] - kD_{g,i} + C(k) \end{aligned} \tag{30}$$

where $D_{g,i}$ is the scaling parameter of the ith direction and $C(k)$ a constant depending only on the fixed parameter k (so it may be ignored during optimization). In our test, we used $k = 1600$ which gives a distribution close to a Gaussian centered around 1 with its standard deviation equal to $\frac{1}{\sqrt{k}}$, here 0.025.

C. Local Registration Referential Optimization

For a given subject a and a given label l, the local rigid transformation of a point x to the point y, expressed in the referential g is $y = R_a \cdot (x - g) + (R_a \cdot g + t_a - g) + g$. So the inverse transformation is $x = R_a^t \cdot (y - g) + (R_a^t \cdot (g - t_a) - g) + g$ with $(R_a^t \cdot (g - t_a) - g)$ corresponding to the local translation from the referential g. The ideal g is the one which reduces the variability of these translations so we suggest to minimize the following energy:

$$E = \sum_{a\in\mathcal{A}} \|R_a^t \cdot (g - t_a) - g\|^2 \tag{31}$$

First, if all $R_a = Id$, E does not depend on g anymore, so all referentials are equivalent. For other cases, we have to study the derivatives of the energy:

$$\begin{aligned} &\frac{\partial E}{\partial g} = 4|\mathcal{A}|g - 2 \sum_{a\in\mathcal{A}} \left[(R_a + R_a^t) \cdot g + (R_a - Id)^t \cdot t_a\right] = \vec{0} \\ &\text{Thus,} \quad g = \left[\sum_{a\in\mathcal{A}} (S_a + S_a^t)\right]^{-1} \left[\sum_{a\in\mathcal{A}} S_a^t \cdot t_a\right] \end{aligned} \tag{32}$$

where $S_a = Id - R_a$, and $|\mathcal{A}|$ is the number of subjects.

In fact g is also the solution of minimizing the following energy:

$$E = \sum_{a \in \mathcal{A}} \|R_a \cdot g + t_a - g\|^2 \tag{33}$$

where each component of the sum is the square norm of a local translation from the referential space to the common space (the opposite direction of the previous energy 31). Thus, both directions give the same optimal point of reference g. But at the optimum, the first proposed energy is lower than the second and so the variations are more constrained and should give sharper priors.

D. Vector-Rotation Parametrization

A rotation can be described as follows:

$$\begin{array}{rcl} R & = & \exp(U) \\ & = & I + \frac{s}{\alpha}U + \frac{(1-c)}{\alpha^2}U^2 \end{array} \quad \text{where} \quad \left\{ \begin{array}{rcl} s & = & \sin(\alpha) \\ c & = & \cos(\alpha) \end{array} \right. \tag{34}$$

where $\alpha = \|w\|$, $\exp(\cdot)$ stands for standard matrix exponential (or $\exp_I$: the exponential map at point I, the 3×3 identity matrix), U is a skew-symmetric ($U^t = -U$) matrix:

$$U = \begin{pmatrix} 0 & -w_z & w_y \\ w_z & 0 & -w_x \\ -w_y & w_x & 0 \end{pmatrix} \quad \text{with } w = (w_x, w_y, w_z) \tag{35}$$

related to cross product since $U \cdot v = w \times v$.

The second line of equation 34 is known as the Rodrigues formula and yields tractable ways to do conversion between matrix representations and vector-rotation ones.

E. Circular and Directional Statistics

Circular and directional statistics cover a wide range of data related to spherical data (data living in circles, spheres or their p-dimensional generalizations). Various generalizations and approximations of Gaussians have appeared to deal with normed vectors, lines directed by those vectors, angles data or orthogonal axial sets.

E.1. Distribution details

In our context, we have modeled each rotation by an axis w and an angle α following the distribution below:

$$\left\{ \begin{array}{rcll} \frac{w_l}{\alpha_l} & \sim & \mathcal{B}(A_l) & \text{a Bingham distribution} \\ \alpha_l & \sim & \mathcal{VM}(\nu_l, \kappa_l) & \text{a Von Mises distribution} \end{array} \right. \tag{36}$$

with A_l a 3×3 positive-definite symmetric matrix (the eigen vector related to the largest eigen value is the mean axial direction whereas the others are the scattering directions along the orthogonal plane), ν_l the angle mean and κ_l controlling the concentration. See figure 6 to see what the distributions look like.

A Von Mises (Mardia and Jupp, 2000) (see left of figure 6) distribution is used for angles data:

$$P(\alpha|\kappa, \nu) = \frac{\exp(\kappa \cos(\alpha - \nu))}{2\pi I_0(\kappa)} \tag{37}$$

where $I_0(\kappa)$ is the modified Bessel function of order 0. A maximum likelihood estimator (MLE) exists to compute κ and ν parameters from training data.

A Bingham distribution (Bingham, 1974) (see right of figure 6) is used for the axis part. This distribution is bimodal since the axis direction does not matter:

$$P(w|A) = P(-w|A) = \frac{\exp(\|w\|_A^2)}{{}_1F_1(0.5, 1.5, Z)} \tag{38}$$

where $\|w\|_A^2 = w^t A w$ is the canonical dot product defined by A and $A = MZM^t$ (M orthogonal and Z diagonal) with ${}_1F_1$ is the confluent hypergeometric function of matrix argument Z. The maximum likelihood estimator (MLE) of M is the matrix of eigen vectors of the scattering matrix of training data. In the other hand, Z is optimized numerically.

E.2. A priori values for model estimation

As previously, suitable parameter values are set a priori during model estimation and estimated in regard to training data variability for the labeling step. The angle mean ν_l is set to zero and its concentration κ_l is set to approximate a standard deviation of $\frac{\pi}{4}$. About A_l, the axis mean is set to be roughly orthogonal to a mean skull because local sulci distortions needed to match the model are expected to move along the brain main curves. Lastly, the dispersion directions are set to represent a solid angle of $\frac{2}{3}$ sr (steradians), namely an isotropic dispersion of about $\frac{\pi}{4}$ radians from the axis.

F.C.L.a.	anterior lateral fissure
F.C.L.r.ant.	anterior ramus of the lateral fissure
F.C.L.r.diag.	diagonal ramus of the lateral fissure
F.C.L.r.sc.ant.	anterior sub-central ramus of the lateral fissure
F.C.M.ant.	calloso-marginal anterior fissure
F.Cal.ant.-Sc.Cal.	calcarine fissure
F.I.P.Po.C.inf.	superior postcentral intraparietal superior sulcus
F.I.P.r.int.1	primary intermediate ramus of the intraparietal sulcus
F.P.O.	parieto-occipital fissure
OCCIPITAL	lobe occipital
S.C.	central sulcus
S.Call.	subcallosal sulcus
S.F.inf.	inferior frontal sulcus
S.F.int.	internal frontal sulcus
S.F.marginal.	marginal frontal sulcus
S.F.orbitaire.	orbital frontal sulcus
S.F.sup.	superior frontal sulcus
S.Li.ant.	anterior intralingual sulcus
S.O.T.lat.ant.	anterior occipito-temporal lateral sulcus
S.O.T.lat.med.	median occipito-temporal lateral sulcus
S.O.p.	occipito-polar sulcus
S.Or.	orbital sulcus
S.Pa.sup.	superior parietal sulcus
S.Pe.C.inf.	inferior precentral sulcus
S.Pe.C.marginal.	marginal precentral sulcus
S.Pe.C.sup.	superior precentral sulcus
S.R.inf.	inferior rostral sulcus
S.T.i.ant.	anterior inferior temporal sulcus
S.T.pol.	polar temporal sulcus
S.T.s.ter.asc.ant.	anterior terminal ascending branch of the superior temporal sulcus
S.p.C.	paracentral sulcus
ventricle	ventricle

F.C.L.p.	posterior lateral fissure
F.C.L.r.asc.	ascending ramus of the lateral fissure
F.C.L.r.retroC.tr.	retro central transverse ramus of the lateral fissure
F.C.L.r.sc.post.	posterior sub-central ramus of the lateral fissure
F.C.M.post.	calloso-marginal posterior fissure
F.Coll.	collateral fissure
F.I.P.	intraparietal sulcus
F.I.P.r.int.2	secondary intermediate ramus of the intraparietal sulcus
INSULA	insula
S.C.LPC.	paracentral lobule central sulcus
S.C.sylvian.	central sylvian sulcus
S.Cu.	cuneal sulcus
S.F.inf.ant.	anterior inferior frontal sulcus
S.F.inter.	intermediate frontal sulcus
S.F.median.	median frontal sulcus
S.F.polaire.tr.	polar frontal sulcus
S.GSM.	sulcus of the supra-marginal gyrus
S.Li.post.	posterior intra-lingual sulcus
S.O.T.lat.int.	internal occipito-temporal lateral sulcus
S.O.T.lat.post.	posterior occipito-temporal lateral sulcus
S.Olf.	olfactory sulcus
S.Pa.int.	internal parietal sulcus
S.Pa.t.	transverse parietal sulcus
S.Pe.C.inter.	intermediate precentral sulcus
S.Pe.C.median.	median precentral sulcus
S.Po.C.sup.	superior postcentral sulcus
S.Rh.	rhinal sulcus
S.T.i.post.	posterior inferior temporal sulcus
S.T.s.	superior temporal sulcus
S.T.s.ter.asc.post.	posterior terminal ascending branch of the superior temporal sulcus
S.s.P.	sub-parietal sulcus
unknown	unknown cortical folds

Table 3: List of sulcal labels. Left column : acronyms used by Brainvisa software and report on figure 12. Right column : anatomical names. All Labels are defined on both hemispheres except *S.GSM.* defined on the left hemisphere only. The *unknown* label is a null label used for unknown anatomical folds.

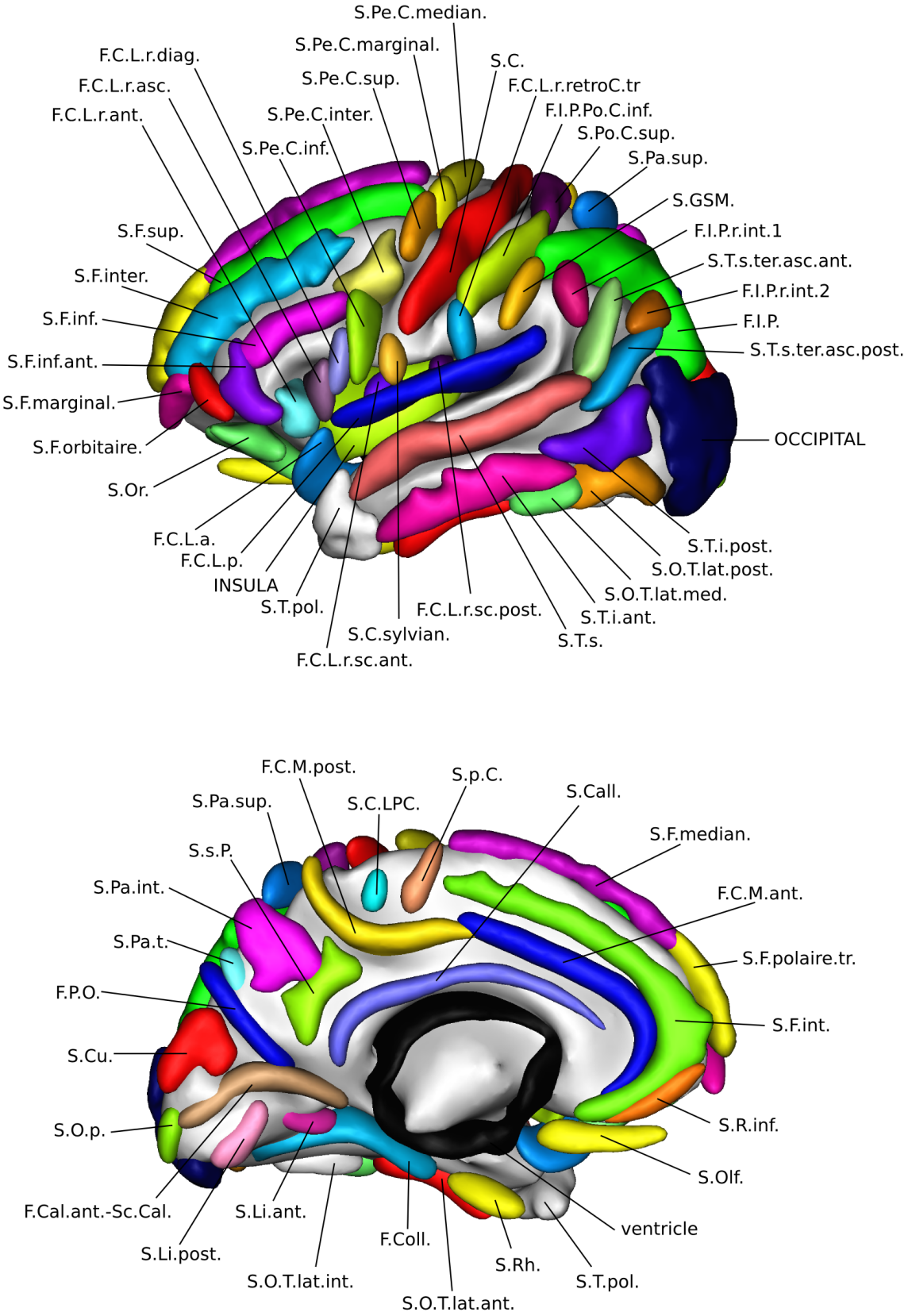


Figure 12: Lateral and medial view of the template optimized with local rigid registrations. Standard location of the sulcal labels (named according to their relative acronym from Brainvisa software) are precised. See table 3 to get the definition of the acronyms.

References

Allassonnière, S., Amit, Y., Trouvé, A., 2007. Towards a coherent statistical framework for dense deformable template estimation. Journal Of The Royal Statistical Society Series B 69 (1), 3–29.

Arsigny, V., Pennec, X., Ayache, N., December 2005. Polyrigid and polyaffine transformations: a novel geometrical tool to deal with non-rigid deformations - application to the registration of histological slices. Medical Image Analysis 9 (6), 507–523.

Ashburner, J., 2007. A fast diffeomorphic image registration algorithm. NeuroImage 38 (1), 95–113.

Ashburner, J., Friston, K., 2005. Unified segmentation. NeuroImage 26, 839–851.

Ashburner, J., Neelin, P., Collins, D. L., Evans, A., Friston, K., November 1997. Incorporating prior knowledge into image registration. NeuroImage 6 (4), 344–352.

Auzias, G., Glaunés, J., Colliot, O., Perrot, M., Mangin, J.-F., Trouvé, A., Baillet, S., 2009. DISCO: a Coherent Diffeomorphic Framework for Brain Registration Under Exhaustive Sulcal Constraints. Springer.

Avants, B., Gee, J., 2004. Geodesic estimation for large deformation anatomical shape averaging and interpolation. NeuroImage 23, S139–S150.

Bingham, C., 1974. An antipodally symmetric distribution on the sphere. Annals of Statistics 2 (6), 1201–1225.

Boisvert, J., Cheriet, F., Pennec, X., Labelle, H., Ayache, N., April 2008. Geometric variability of the scoliotic spine using statistics on articulated shape models. IEEE Transactions on Medical Imaging 27 (4), 557–568.

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: MICCAI '01: Proceedings of the 4th International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer-Verlag, London, UK, pp. 734–742.

Caunce, A., Taylor, C. J., 1998. 3d point distribution models of the cortical sulci. In: ICCV '98: Proceedings of the Sixth International Conference on Computer Vision. IEEE Computer Society, Washington, DC, USA, p. 402.

Chui, H., Rambo, J., Duncan, J. S., Schultz, R., Rangarajan, A., 1999. Registration of cortical anatomical structures via robust 3d point matching. In: IPMI '99: Proceedings of the 16th International Conference on Information Processing in Medical Imaging. Springer-Verlag, London, UK, pp. 168–181.

Chui, H., Rangarajan, A., 2000. A feature registration framework using mixture models. In: MMBIA '00: Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis. IEEE Computer Society, Washington, DC, USA, pp. 190–197.

Collins, D., Holmes, C., Peters, T., Evans, A., 1995. Automatic 3d model-based neuroanatomical segmentation. Human Brain Mapping 3 (3), 190–208.

Collins, D. L., Goualher, G. L., Evans, A. C., 1998. Non-linear cerebral registration with sulcal constraints. In: MICCAI '98: Proceedings of the First International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer-Verlag, London, UK, pp. 974–984.

Collins, L. D., Neelin, P., Peters, T. M., Evans, A. C., March 1994. Automatic 3d intersubject registration of mr volumetric data in standardized talairach space. Journal of Computer Assisted Tomography 18, 192–205.

De Craene, M., du Bois d'Aische, A., Macq, B., Warfield, S., Sept. 2005. Multi-subject variational registration for probabilistic unbiased atlas generation. Image Processing, 2005. ICIP 2005. IEEE International Conference on 3, III–601–4.

Dempster, A. P., Laird, N. M., Rubin, D. B., 1977. Maximum likelihood from incomplete data via the em algorithm. Journal of the Royal Statistical Society. Series B (Methodological) 39 (1), 1–38.

Dubois, J., Benders, M., Cachia, A., Lazeyras, F., Leuchter, R. H.-V., Sizonenko, S., Borradori-Tolsa, C., Mangin, J.-F., Hüppi, P., 2008. Mapping the early cortical folding process in the preterm newborn brain. Cereb. Cortex 18 (6), 1444–54.

Eckstein, I., Joshi, A. A., Kuo, C. C. J., Leahy, R. M., Desbrun, M., 2007. Generalized surface flows for deformable registration and cortical matching. In: MICCAI (1). pp. 692–700.

Evans, A., Collins, D., Mills, S., Brown, E., Kelly, R., Peters, T., 1993. 3d statistical neuroanatomical models from 305 mri volumes. IEEE Nuclear Science Symposium and Medical Imaging Conference 3, 1813–1817.

Evans, A., Collins, D., Neelin, P., Kamber, M., Marrett, T., 1994. Three-Dimensional Correlative Imaging: Applications in Human Brain Mapping. Functional Neuroimaging, Ch. 14, pp. 145–161.

Fischl, B., Sereno, M. I., Tootell, R. B. H., Dale, A. M., 1999. High-resolution intersubject averaging and a coordinate system for the cortical surface. Hum. Brain Mapp 8, 272–284.

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., January 2004. Automatically parcellating the human cerebral cortex. Cereb Cortex 14 (1), 11–22.

Gholipour, A., Kehtarnavaz, N. D., Briggs, R., Devous, M., Gopinath, K. S., 2007. Brain functional localization: A survey of image registration techniques. IEEE Trans. Med. Imaging 26 (4), 427–451.

Glaunès, J., Joshi, S., 1st of October 2006. Template estimation form unlabeled point set data and surfaces for computational anatomy. In: Pennec, X., Joshi, S. (Eds.), Proc. of the International Workshop on the Mathematical Foundations of Computational Anatomy (MFCA-2006). pp. 29–39.

Glaunès, J., Trouvé, A., Younes, L., 2004. Diffeomorphic matching of distributions: A new approach for unlabelled point-sets and sub-manifolds matching. cvpr 02, 712–718.

Goualher, G. L., Barillot, C., Bizais, Y., 1997. Modeling cortical sulci with active ribbons. IJPRAI 11 (8), 1295–1315.

Hellier, P., Barillot, C., Corouge, I., Gibaud, B., Goualher, G. L., Collins, L., Evans, A., Mal, G., Ayache, N., 2003. Retrospective evaluation of intersubject brain registration. IEEE Transactions on Medical Imaging 22, 1120–1130.

Hurdal, M. K., Gutierrez, J. B., Laing, C., Smith, D. A., 2008. Shape analysis for automated sulcal classification and parcellation of mri data. J. Comb. Optim. 15 (3), 257–275.

Joshi, A. A., Shattuck, D. W., Thompson, P. M., Leahy, R. M., 2005. A framework for registration, statistical characterization and classification of cortically constrained functional imaging data. In: IPMI. pp. 186–196.

Joshi, S., Davis, B., Jomier, B. M., B, G. G., 2004. Unbiased diffeomorphic atlas construction for computational anatomy. Neuroimage 23, 151–160.

Kao, C.-Y., Hofer, M., Sapiro, G., Stern, J., Rehm, K., Rottenberg, D. A., 2007. A geometric method for automatic extraction of sulcal fundi. IEEE Trans. Medical Imaging 26 (4), 530–540.

Klein, A., Hirsch, J., January 2005. Mindboggle: a scatterbrained approach to automate brain labeling. Neuroimage 24 (2), 261–280.

Le Goualher, G., Collins, D., Barillot, C., Evans, A., 1998. Automatic identification of cortical sulci using a 3d probabilistic atlas. In: MICCAI '98: Proceedings of the First International Conference on Medical Image Computing and Computer-Assisted Intervention. Vol. 1496. pp. 509–518.

Lohmann, G., 1998. Extracting line representations of sulcal and gyral patterns in mr images of the human brain. IEEE Trans Med Imaging 17 (6).

Lohmann, G., von Cramon, D., Colchester, A., 2008. Deep sulcal landmarks provide an organizing framework for human cortical folding. Cerebral Cortex 18 (6), 1415–1420.

Lohmann, G., von Cramon., D. Y., 2000. Automatic labelling of the human cortical surface using sulcal basins. Medical Image Analysis 4 (3), 179–188.

Lui, L. M., Wang, Y., Chan, T. F., Thompson, P., 2007. Landmark constrained genus zero surface conformal mapping and its application to brain mapping research. Appl. Numer. Math. 57 (5-7), 847–858.

Lyttelton, O., Boucher, M., Robbins, S., Evans, A., 2007. An unbiased iterative group registration template for cortical surface analysis. Neuroimage 34 (4), 1535–1544.

Maintz, J. B., Viergever, M. A., March 1998. A survey of medical image registration. Med Image Anal 2 (1), 1–36.

Mangin, J.-F., Frouin, V., Bloch, I., Régis, J., Lopez-Krahe, J., 1995. From 3D magnetic resonance images to structural representations of the cortex topography using topology preserving deformations. Journal of Mathematical Imaging and Vision 5, 297–318.

Mangin, J.-F., Rivière, D., Cachia, A., Duchesnay, E., Cointepas, Y., Papadopoulos-Orfanos, D., Scifo, P., Ochiai, T., Brunelle, F., Régis, J., 2004. A framework to study the cortical folding patterns. Neuroimage 23, S129–S138.

Mardia, K. V., Jupp, P. E., 2000. Directional Statistics, 2nd Edition. Wiley.

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

356 (1412), 1293–1322.

Mazziotta, J. C., Toga, A. W., Evans, A., Fox, P., Lancaster, J., 1995. A probabilistic atlas of the human brain: Theory and rationale for its development : The international consortium for brain mapping (icbm). NeuroImage 2 (2, Part 1), 89 – 101.

Ono, M., Kubik, S., Abarnathey, C. D., 1990. Atlas of the Cerebral Sulci. Thieme Medical Publishers.

Ourselin, S., Roche, A., Prima, S., Ayache, N., October 2000. Block matching: a general framework to improve robustness of rigid registration of medical images. In: Delp, S., DiGioia, A., Jaramaz, B. (Eds.), 3rd International Conference on Medical Image Computing and Computer-Assisted Intervention, MICCAI'2000. Vol. 1935 of Lecture Notes in Computer Science. Springer, Pittsburgh, USA, pp. 557–566.

Parzen, E., 1962. On estimation of a probability density function and mode. The Annals of Mathematical Statistics 33 (3), 1065–1076.

Pennec, X., Thirion, J.-P., December 1997. A framework for uncertainty and validation of 3D registration methods based on points and frames. Int. Journal of Computer Vision 25 (3), 203–229.

Perrot, M., Rivière, D., Mangin, J.-F., 2008. Identifying cortical sulci from localization, shape and local organization. In: ISBI. pp. 420–423.

Pinel, P., Thirion, B., Mériaux, S., Jobert, A., Serres, J., Le Bihan, D., Poline, J.-B., Dehaene, S., Oct 2007. Fast reproducible identification and large-scale databasing of individual functional cognitive networks. BMC Neurosci 8 (1), 91.

Pohl, K. M., Fisher, J., Grimson, Kikinis, R., Wells, W. M., May 2006. A bayesian model for joint segmentation and registration. NeuroImage 31 (1), 228–239.

Poupon, C., Poupon, F., Allirol, L., Mangin, J.-F., 2006. A database dedicated to anatomo-functional study of human brain connectivity. In: 12th HBM Neuroimage. No. 646. Florence, Italie.

Powell, M. J. D., February 1964. An efficient method for finding the minimum of a function of several variables without calculating derivatives. The Computer Journal 7 (2), 155–162.

Régis, J., Mangin, J.-F., Ochiai, T., Frouin, V., Rivière, D., Cachia, A., Tamura, M., Samson, Y., 2005. "sulcal root" generic model: a hypothesis to overcome the variability of the human cortex folding patterns. Neurol Med Chir (Tokyo) 45, 1–17.

Rivière, D., Mangin, J.-F., Papadopoulos-Orfanos, D., Martinez, J.-M., Frouin, V., Régis, J., 2002. Automatic recognition of cortical sulci of the human brain using a congregation of neural networks. Medical Image Analysis 6 (2), 77–92.

Rivière, D., Geffroy, D., Denghien, I., Souedet, N., Cointepas, Y., 2009. BrainVISA: an extensible software environment for sharing multimodal neuroimaging data and processing tools. In: Proc. 15th HBM.
URL `http://brainvisa.info`

Seong, J., Im, K., Yoo, S., Seo, S., Na, D., Lee, J., 2010. Automatic extraction of sulcal lines on cortical surfaces based on anisotropic geodesic distance. NeuroImage 49 (1), 293–302.

Shattuck, D., Joshi, A., Pantazis, D., Kan, E., Dutton, R., Sowell, E., Thompson, P., Toga, A., Leahy, R., 2009. Semi-automated method for delineation of landmarks on models of the cerebral cortex. Journal of Neuroscience Methods 178 (2), 385–392.

Shattuck, D. W., Mirza, M., Adisetiyo, V., Hojatkashani, C., Salamon, G., Narr, K. L., Poldrack, R. A., Bilder, R. M., Toga, A. W., February 2008. Construction of a 3d probabilistic atlas of human cortical structures. Neuroimage 39 (3), 1064–1080.

Shi, Y., Thompson, P., Dinov, I., Toga, A., May 2008. Hamilton-jacobi skeleton on cortical surfaces. MedImg 27 (5), 664–673.

Shi, Y., Tu, Z., Reiss, A., Dutton, R., Lee, A., Galaburda, A., Dinov, I., Thompson, P., Toga, A., March 2009. Joint sulcal detection on cortical surfaces with graphical models and boosted priors. Medical Imaging, IEEE Transactions on 28 (3), 361–373.

Talairach, J., Tournoux, P., 1988. Co-planar Stereotaxic Atlas of the Human Brain.

Tao, X., Prince, J. L., Davatzikos, C., 2002. Using a statistical shape model to extract sulcal curves on the outer cortex of the human brain. IEEE Trans. Med. Imaging 21 (5), 513–524.

Thompson, P. M., Toga, A. W., Sep 1997. Detection, visualization and animation of abnormal anatomic structure with a deformable probabilistic brain atlas based on random vector field transformations. Med Image Anal 1, 271–294.

Thompson, P. M., Woods, R. P., Mega, M. S., Toga, A. W., February 2000. Mathematical/computational challenges in creating deformable and probabilistic atlases of the human brain. Hum Brain Mapp 9 (2), 81–92.

Tosun, D., Prince, J., December 2008. A geometry-driven optical flow warping for spatial normalization of cortical surfaces. IEEE Transactions on Medical Imaging 27 (12), 1739–1753.

Tosun, D., Rettmann, M. E., Prince, J. L., 2003. Mapping techniques for aligning sulci across multiple brains. In: Ellis, R. E., Peters, T. M. (Eds.), MICCAI (2). Vol. 2879 of Lecture Notes in Computer Science. Springer, pp. 862–869.

Tu, Z., Zheng, S., Yuille, A. L., Reiss, A. L., Dutton, R. A., Lee, A. D., Galaburda, A. M., Dinov, I. D., Thompson, P. M., Toga, A. W., 2007. Automated extraction of the cortical sulci based on a supervised learning approach. IEEE Trans. Med. Imaging 26 (4), 541–552.

Vaillant, M., Davatzikos, C., 1997. Finding parametric representations of the cortical sulci using an active contour model. Medical Image Analysis 1, 295–315.

Vaillant, M., Davatzikos, C., 1999. Hierarchical matching of cortical features for deformable brain image registration. In: IPMI. Springer-Verlag, pp. 182–195.

Van Essen, D. C., November 2005. A population-average, landmark- and surface-based (pals) atlas of human cerebral cortex. Neuroimage 28 (3), 635–662.

Vivodtzev, F., Linsen, L., Hamann, B., Joy, K. I., Olshausen, B. A., 2005. Brain mapping using topology graphs obtained by surface segmentation.

Yang, F., Kruggel, F., 2007. Optimization algorithms for labeling brain sulci based on graph matching. In: ICCV. pp. 1–7.

Yang, F., Kruggel, F., 2009. A graph matching approach for labeling brain sulci using location, orientation, and shape. Neurocomputing, in print.

Yeo, B., Sabuncu, M., Desikan, R., Fischl, B., Golland, P., 2008. Effects of registration regularization and atlas sharpness on segmentation accuracy. Medical Image Analysis 12, 603–615.

Yeo, B., Sabuncu, M., Vercauteren, T., Ayache, N., Fischl, B., Golland, P., 2010. Spherical demons: Fast diffeomorphic landmark-free surface registration. Medical Imaging, IEEE Transactions on 29 (3), 650–668.

Zhou, Y., Thompson, P. M., Toga, A. W., 1999. Extracting and representing the cortical sulci. IEEE Comput. Graph. Appl. 19 (3), 49–55.