

Retrospective Evaluation of Intersubject Brain Registration

P. Hellier*, C. Barillot, I. Corouge, B. Gibaud, G. Le Goualher, D. L. Collins, A. Evans, G. Malandain, N. Ayache, G. E. Christensen, and H. J. Johnson

Abstract—Although numerous methods to register brains of different individuals have been proposed, no work has been done, as far as we know, to evaluate and objectively compare the performances of different nonrigid (or elastic) registration methods on the same database of subjects. In this paper, we propose an evaluation framework, based on global and local measures of the relevance of the registration. We have chosen to focus more particularly on the matching of cortical areas, since intersubject registration methods are dedicated to anatomical and functional normalization, and also because other groups [7] have shown the relevance of such registration methods for deep brain structures. Experiments were conducted using 6 methods on a database of 18 subjects. The global measures used show that the quality of the registration is directly related to the transformation's degrees of freedom. More surprisingly, local measures based on the matching of cortical sulci did not show significant differences between rigid and non rigid methods.

Index Terms—Atlas matching, cortical sulci, evaluation, MRI, neuroanatomy, nonrigid registration.

I. INTRODUCTION

THE COMPARISON of brains of different individuals is an ancient objective in medicine. It has been traditionally treated by paper-based atlases with simple transformations. However, over the last ten years, electronic brain atlases [11], [17], [23], [30], [31], [42] have emerged and overcome some limitations of traditional atlases [14]. To build such an atlas, it is necessary to compare brains of different individuals so that each new subject contributes to the evolution and the relevance of the

atlas. The comparison of brains is accomplished by registering one brain image volume to another using a nonrigid transformation. The result of the registration – which in general is a dense deformation field – is used to project data from one subject to another one. This projected data can be of different types such as acquired or estimated, two-dimensional (2-D) or three-dimensional (3-D), dense or sparse, and anatomical or functional.

An increasing number of authors study this registration problem. As it would be a huge task to quote them all, we refer the reader to [12], [26], and [28] for a survey on that subject. These methods are generally divided into two groups: photometric, or intensity-based, methods, that exploit a relationship between voxels' luminance, and geometric methods that rely on the extraction and matching of sparse landmarks. Geometric methods dramatically depend on the extraction of features and are only relevant in a neighborhood of these features. On the contrary, photometric methods use the entire available information and make it possible to estimate transformations with high degrees of freedom (DOFs). This rapid comparison may explain the popularity of photometric methods which are used in rigid multimodal fusion [45], [46].

Nevertheless, the superiority of photometric methods has not been proved in the context of interindividual fusion. As a matter of fact, these methods usually rely on the minimization of an appropriate cost function that often relies on the assumption that two corresponding voxels have comparable luminance. The different methods mainly differ by the regularization scheme and optimization strategy which have a crucial influence on the registration process. Unfortunately, it is not straightforward to determine whether the formulation of the problem and the way it is solved leads to anatomically consistent transformations. Is it relevant to deform one subject toward another? Does a consistent anatomical transformation exist? What should be the “ideal” transformation between the brains of different subjects? Does it even exist? What can we expect from the different registration methods? These questions are the starting point and the motivation of our work. It is challenging to answer these questions, especially since the “ideal” transformation surely depends on the application (e.g., anatomical and functional normalization, automatic segmentation of deep structures, automatic segmentation of cortical structures).

In the context of this paper, the evaluation framework is based on the assumption that some anatomical structures are present in all individuals and registration algorithms should aim at aligning these features. In this paper, we have chosen to focus on the matching of cortical regions since the methods are dedicated to the anatomical and functional normalization. Furthermore,

Manuscript received November 22, 2002; revised March 20, 2003. This work was supported in part by the Brittany Country Council under a contribution to the student grant, in part by the GIS Project “cognition science.” (for the acquisition of the data), in part by the National Institutes of Health (NIH) under Grant NS35368, Grant DC03590, and Grant CA75371, and in part by a grant from the Whitaker Foundation. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was W. Niessen. *Asterisk indicates corresponding author.*

*P. Hellier is with the Projet Vista, IRISA/INRIA-CNRS Rennes, France (e-mail: pierre.hellier@irisa.fr).

C. Barillot and I. Corouge are with the Projet Vista, IRISA/INRIA-CNRS, F-35042 Rennes, France.

B. Gibaud is with the Laboratoire IDM, Faculté de Médecine, F-35042 Rennes, France.

G. Le Goualher is with the Laboratoire IDM, Faculté de Médecine, F-35042 Rennes, France, and also with the Montreal Neurological Institute, Mc Gill University, Montreal, QC H3A 2T5, Canada.

D. Collins and A. Evans are with the Montreal Neurological Institute, Mc Gill University, Montreal, QC H3A 2T5, Canada.

G. Malandain and N. Ayache are with the Projet Epidaure, INRIA, 06902 Sophia-Antipolis, France.

G. E. Christensen and H. J. Johnson are with the Department of Electrical and Computer Engineering, University of Iowa, Iowa City IA 52242 USA.

Digital Object Identifier 10.1109/TMI.2003.816961

other groups have already shown relevant evaluation on deep brain structures [7].

The evaluation project was first granted under a French national project and was extended to 6 registration methods on a database of 18 subjects. The Vista project (INRIA-CNRS, Rennes) gathered the registration results, i.e., the deformation fields that were used to deform specific anatomical landmarks. The goal of this project is to evaluate how well anatomical features are matched using different registration methods using global and local criteria.

This paper is organized as follows: Section II presents briefly the methods that were evaluated, Section III presents the data and the evaluation framework. Results for global measures (based on the matching of dense and global features) and local measures (based on the matching of sparse and local features) are presented in Sections IV and V, respectively. Conclusions are drawn in Section VI.

II. PARTICIPANTS

This evaluation project was inspired by the Vanderbilt evaluation project [45], [46] in which all participants downloaded the data and performed the registration processes in their laboratory without knowing the evaluation criteria. The results, i.e., the deformation fields, were sent to IRISA (INRIA-CNRS, Rennes, France) and evaluated using criteria unknown to the participants of the evaluation project. That means that the evaluation group did not reprogram the different methods. Consequently, it has not been feasible to compare the methods on the basis of computation time, sensitivity to parameters, difficulty to implement, etc.

Six methods have been evaluated so far. We do not describe extensively the different methods, but refer the reader to the appropriate references. We have adopted the following notation for the methods:

- Method A refers to the ANIMAL algorithm developed by L. Collins *et al.* at the MNI [7]. This registration is performed in two steps: a rigid transformation is first estimated to align the subject in the stereotaxic coordinate system. A nonrigid transformation is then sought through a multiresolution scheme. At each node of the grid, a translation is estimated that maximizes the correlation of image gradients. In the present experiments, the finest resolution of the method A is 4 mm.
- Method D refers to the Demons' algorithm developed by J.P. Thirion in the Epidaure Group at INRIA Sophia-Antipolis [41]. Two transformations are first estimated, one rigid and one affine. A dense grid of demons (that is to say, one demon per voxel) is then used to estimate a non-rigid transformation. The algorithm alternates between estimating forces for each demon and smoothing the deformation field by a Gaussian filtering. More recently, coworkers of the Epidaure group related the demon's algorithm to a second-order gradient descent of the displaced frame difference [33].
- Method I refers to the inverse consistent linear elastic image registration algorithm developed by G. Christensen and H. Johnson [5]. Given two subjects A and B , the

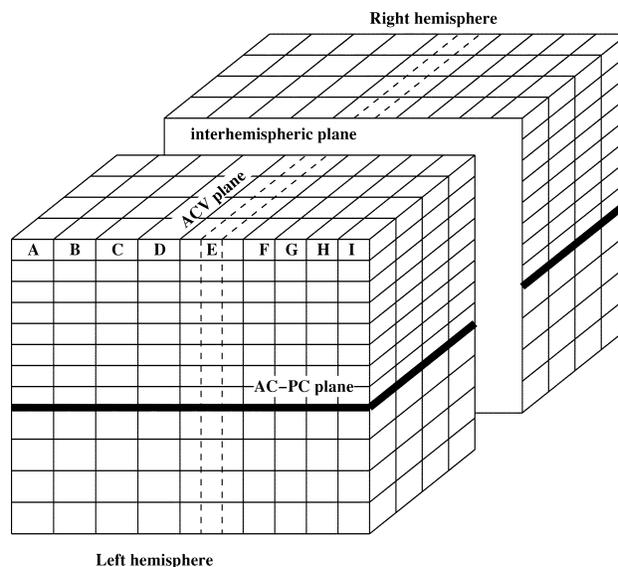


Fig. 1. Principle of the Talairach proportional squaring system.

method jointly estimates transformation from A to B and from B to A . The inverse consistency error is zero when the forward and reverse transformations are inverses of one another. Furthermore, the transformations obey the rules of continuum mechanics and are parameterized by Fourier series.

- Method M refers to a rigid transformation obtained by maximization of the mutual information [27], [44] with a Powell optimization scheme. This method does not appear to be adequate when registering brains of different subjects. However, it was included to serve as a comparison basis since we expect to perform worst in this study. This method has been implemented by IRISA (INRIA-CNRS, Rennes) based on the published work [27], [44].
- Method P refers to the proportional squaring of Talairach illustrated on Fig. 1. The method is based on the identification of the anterior commissure (AC) and posterior commissure (PC), as well as five brain extrema which makes it possible to specify a partition of the volume into 12 sub-volumes [39]. This defines a piecewise affine transformation. Although semi-automatic (AC-PC and five extremal points of the brain have to be located manually), the expertise needed to accomplish this task is limited. This method has been implemented by the IDM Laboratory (faculty of Medicine, Rennes, France).
- Method R refers to the algorithm developed at INRIA Rennes by P. Hellier *et al.* [19]. This method consists of a robust estimation of the optical flow with a multiresolution and multigrid minimization scheme. We note that there may be questions to the validity of the evaluation of the authors method since the evaluation criteria were known to us. Despite this, we hope that the reader believes that we acted in an unbiased manner.

III. DATA AND EVALUATION FRAMEWORK

For this evaluation, the underlying assumption is the following: the anatomical structures that we use can be retrieved

among all individuals and the registration algorithms should aim at aligning these features. In case of “dense” features (i.e., tissues), the evaluation criterion is an overlap measure after registration which is referred to as a “global” measure. In case of “sparse” features (i.e., cortical sulci), the evaluation criterion is a distance measure after registration which we refer to as a “local” measure.

We have acquired a database of 18 subjects (male, age 35 ± 10 , right handed and healthy). Each subject underwent a T1-magnetic resonance (MR) SPGR three-dimensional (3-D) study (GE 1.5T system, sagittal slices). The raw MR data have been linearly interpolated so that the voxel resolution is isotropic (the voxel resolution is 0.9375 mm, except for four subjects where the voxel resolution is 0.976562 mm). The dimension of an image slice is 256×256 for all subjects and the number of slices are subject specific.

We chose an arbitrary subject as the reference subject. For all methods, each subject (source image) was registered to the reference subject (target image) so that all the registration results could be compared in the same reference frame. Fig. 2 (bottom) presents different views of the reference subject. Although the image dimensions and resolutions are different, the deformation fields are expressed in the voxel coordinate of each subject. Convolutions that are necessary for the evaluation (i.e., computation of L_{vv}) take into account the voxel resolution (typically, to cope with Shannon’s criterion, the standard deviation of the Gaussian kernel was taken as twice the voxel resolution).

One could think of using simulated data based on a model for the intersubject variability to provide a gold standard deformation field to directly compare registration methods. However, no such data are currently available since the modeling of intersubject deformations is a highly challenging research area. Consequently, we use real image data and use independent evaluation measures derived from the data.

Features extracted from all MR image volumes were used to assess the quality of the registration processes. These features were selected to be anatomically meaningful since we wanted to determine how well the registration algorithms align anatomical structures. Furthermore, the features were selected so that they were not related to the similarity, or the “forces” used to drive the registration process. To be fair and objective, the evaluation was designed to be independent of the registration algorithms involved in this study.

We note that the extracted features on which the evaluation criteria are based might not be perfect and might even be erroneous because of acquisition noise, limited precision of the extracting algorithms, interpolation schemes, etc. This is inevitable despite the robustness and accuracy of the extracting procedures. Therefore, these results should only be considered as relative values and should not be taken as “absolute” values of the performances of the methods. Therefore, we consider this paper as an “evaluation” and not a “validation” of the methods. Despite this, we think that 1) the errors can be averaged out by the size of the database (assuming the errors are independent) and 2) the methods are treated equally with respect to these errors (assuming the evaluation features and registration methods are independent).

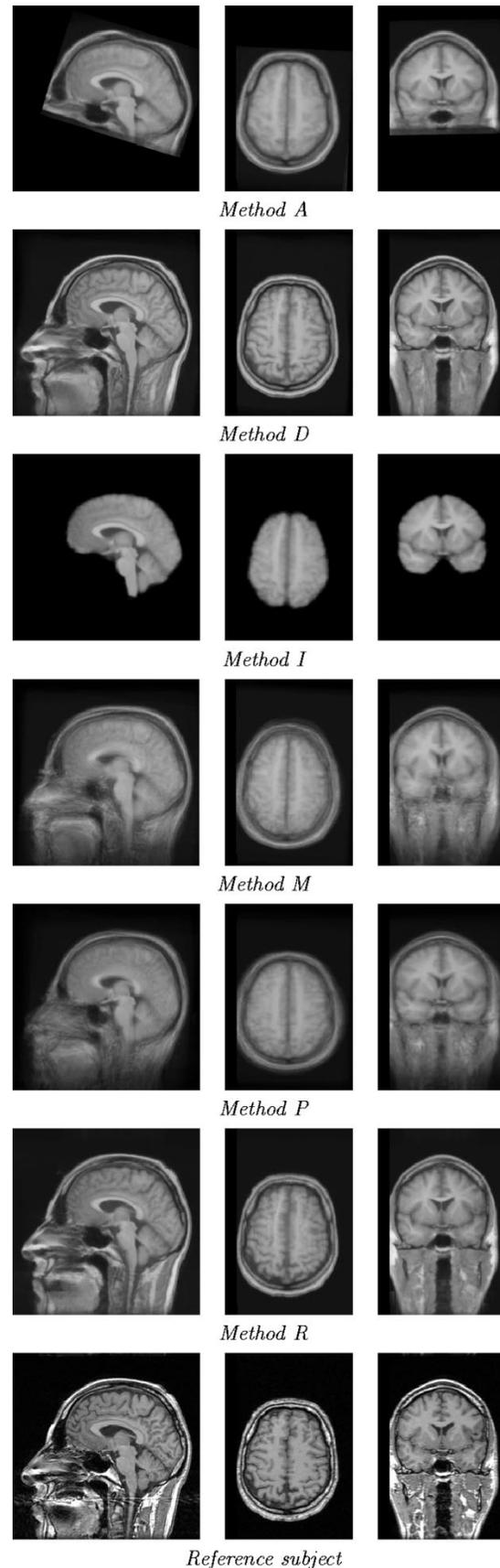


Fig. 2. For each method, the mean volume (sagittal, axial and coronal cut-planes) is obtained by averaging the deformed subjects (17 for all methods except method A) and can be compared to the corresponding view of the reference subject (bottom).

IV. GLOBAL MEASURES

A. Average Volume

For each method, each subject is deformed toward the reference subject using the transformation and trilinear interpolation. It thus becomes possible to compute, for each method, a mean volume by averaging the 17 deformed subjects. Cut-planes through the average volumes are presented in Fig. 2. For each view, the average volume can be visually compared to the reference subject. Furthermore, we compute the mean square error (mse) of intensities between the average volume and the reference volume (see Table I). Let R denote the reference volume, \tilde{S}_p denote the p^{th} registered and interpolated subject, N denote the number of subjects registered and $\#R$ denote the number of voxels where the error is computed. Then, the mse is defined as

$$\text{mse} = \frac{1}{\#R} \sum_{i,j,k} \left(R(i,j,k) - \left(\frac{1}{N} \sum_{p=1}^N \tilde{S}_p(i,j,k) \right) \right)^2.$$

The mse is not a good measure to evaluate the quality of the registration of one subject (first it is more or less related to similarity measure used to drive all registration methods but P, and second it assumes that the MR scanner calibration was identical for all acquisitions), but it is still a useful measure for evaluating average volumes. We only compute the mse for the voxels belonging to the brain's segmentation mask of the reference subject (see Section IV-B), since some methods limit the computation of the registration to a subset of voxels (e.g., stereotaxic space for method A, segmentation mask for method I).

It must be noted that the registration of the subject 9 failed for the method A. Therefore, and for all the experiments, subject 9 has been removed of method A's results (i.e., the evaluation is performed on 16 subjects for this method).

From the visual and numerical results, we are tempted to distinguish two classes of methods: D and R on one side and A, I, M and P on the other. These results are apparently directly related to the DOFs of each method and this point will be discussed further in Section IV-E.

B. Overlap of Grey and White Matter Tissues

The most straightforward way to assess the quality of the registration is to evaluate how the tissues are deformed from one subject to the other. Furthermore, if the evaluation was only based on mse, it would not be complete, as the mse is more or less related to the similarity used to drive the registration processes, at least for methods A, D, I, M, and R. We extract grey matter (GM) and white matter (WM) tissues from the MR volume using the method described in [24]. This segmentation algorithm performs a 3-D texture analysis using a clustering technique to give a rough classification and is refined using a Bayesian relaxation.

This classification was used to measure how well GM and WM overlapped after registration. For each subject, the GM and WM classes were deformed toward the reference subject using the deformation field (or the registration parameters, see Section II) and trilinear interpolation. The deformed classes were

TABLE I

VARIOUS SIMILARITY MEASURES BETWEEN THE AVERAGE VOLUME AND THE REFERENCE SUBJECT. THREE DIFFERENT MEASURES HAVE BEEN COMPUTED: MSE, MUTUAL INFORMATION (MI), AND CORRELATION. THESE MEASURES HAVE BEEN COMPUTED ONLY FOR VOXELS BELONGING TO THE SEGMENTATION MASK OF THE REFERENCE SUBJECT'S BRAIN. RESULTS FOR THE THREE MEASURES ARE ROUGHLY COMPARABLE AND ARE DISCUSSED IN SECTION IV-E

Method	MSE	MI	Correlation
A	987.9	0.55	0.79
D	491.1	1.12	0.94
I	1067.9	0.59	0.82
M	1389.9	0.47	0.78
P	1064.4	0.45	0.74
R	385.6	0.91	0.92

TABLE II

OVERLAP, COMPUTED BY THE TOTAL PERFORMANCE MEASURE, BETWEEN GM AND WM TISSUES AFTER REGISTRATION IN PERCENT. FOR EACH METHOD, THE MEAN AND STANDARD DEVIATION OF THE MEASURE IS COMPUTED OVER THE DATABASE OF SUBJECTS

Methods	Tissue	Mean	Standard deviation
A	grey	91.9	0.08
	white	89.6	0.07
D	grey	95.8	0.04
	white	96.7	0.04
I	grey	93.0	0.07
	white	95.0	0.03
M	grey	88.8	0.13
	white	87.5	0.17
P	grey	93.5	0.06
	white	95.1	0.04
R	grey	93.9	0.07
	white	95.2	0.07

compared to the classes of the reference subject by computing the total performance overlap measure [1]. Let T_1 denote the tissue (grey or white) of a given subject deformed toward the reference subject and T_2 denote the corresponding tissue of the reference subject. T_1 and T_2 are compared in the volume Ω of the reference subject. For any set A , let us note $A^c \triangleq \Omega \setminus A$, and $\#A$ the number of voxels of set A . For brevity, we only keep the total performance measure [1] defined by the ratio $(TP + TN)/(TP + FP + TN + FN)$ where TP is the number of true positives ($TP \triangleq \#(T_1 \cap T_2)$), TN is the number of true negatives ($TN \triangleq \#(T_1^c \cap T_2^c)$), FP is the number of false positives ($FP \triangleq \#(T_1 \cap T_2^c)$) and FN is the number of false negatives ($FN \triangleq \#(T_1^c \cap T_2)$). For each method, the mean and variance of this measure is computed over the database of 17 subjects, and the results are given in Table II.

These values must be interpreted carefully since we use only binary segmentation classes which were deformed using trilinear interpolation. Higher order interpolation [25], combined with fuzzy classes, would certainly give better results. We have previously mentioned that the quantitative measurements should not be taken as absolute values. Method M does not give very satisfactory results, whereas methods D, I, P, and R seem to give better and similar results. Method A appears to perform worse than the other methods, but is computed only down to 4 mm isotropic image grid resolution.

We have performed an analysis of variance (ANOVA) on these overlap measures results. Conceptually, the goal of ANOVA is to determine the amount of variability in groups of data and to see if the variability is greater between groups than within groups. For the ANOVA analysis, we have as many groups as methods (i.e., we have six groups). In each group, the number of samples is the number of subjects and each sample being the result of the overlap between tissues after registration. For brevity, we have combined results for GM and WM (i.e., the number of sample is twice the number of subjects). The strength of ANOVA is to provide a compact measure of the intergroup variance divided by the intragroup variance (more compact than the student t-test which would have been needed to be performed for each subject). A high F value of the ANOVA test indicates that the groups can be distinguished from the distribution of samples.

For the overlap of tissues, we have obtained $F = 225.8$, with $p = 0.0001$, indicating that there is a significant statistical difference between the methods. Fig. 3 presents the distribution of samples for each method, and the standard deviation. One can therefore visually appreciate in this figure the difference between the methods.

C. Correlation of L_{vv}

We extract differential characteristics from the subjects with the L_{vv} operator, introduced by Florack *et al.* [13]. This operator is related to the principal curvatures k_1, k_2 of the iso-intensity surface $I_0 = I(x, y, z)$. It can be computed by

$$L_{vv}(x, y, z) = -\frac{1}{2\|w\|^2} [(I_x^2(I_{yy} + I_{zz}) - 2I_y I_z I_{yz}) + (I_y^2(I_{xx} + I_{zz}) - 2I_x I_z I_{xz}) + (I_z^2(I_{xx} + I_{yy}) - 2I_x I_y I_{xy})],$$

where

$$I_{x^i y^j z^k} = \frac{\partial^u (I(x, y, z))}{\partial^i x \partial^j y \partial^k z} \text{ with } u = i + j + k$$

and

$$\|w\|^2 = (I_x^2 + I_y^2 + I_z^2)^{1/2}.$$

Partial derivatives are computed with a Gaussian filter with a scale-space parameter fixed at 2 mm. The sign of L_{vv} has a very precise interpretation: it can be demonstrated that when limited to the cortical region of interest (ROI) the crest of a gyrus corresponds to a negative value of the L_{vv} , while a deep fold like a sulcus corresponds to its positive part. Therefore, the sign of the mean curvature can be used to distinguish between sulci and gyri [15].

For each subject, we deform the corresponding L_{vv} according to the results of a given registration method using trilinear interpolation. We then compared the deformed L_{vv} with the L_{vv} volume of the reference subject by computing a simple correlation. As the L_{vv} is relevant on the brain only, we restrict the computation of the correlation coefficient on the brain's segmentation mask of the reference subject obtained with the method described in the Section IV-B. For each method, we compute the

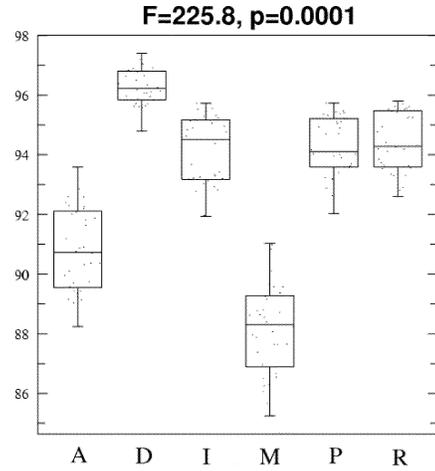


Fig. 3. Distribution of tissue overlap after registration for each method. The ANOVA has been performed on these groups.

TABLE III
MEAN AND STANDARD DEVIATION OF THE CORRELATION COEFFICIENT BETWEEN THE REFERENCE L_{vv} AND DEFORMED L_{vv}

Method	Mean	Standard deviation
A	0.17	0.003
D	0.43	0.005
I	0.16	0.003
M	0.01	0.001
P	0.16	0.003
R	0.32	0.008

mean and variance of this measure over the database of 18 subjects and results are given in Table III.

We observe that the mean value of the correlation coefficient is quite low for all the registration methods. This might indicate that the matching of cortical features is not very good, but that point will be studied more extensively in Section V. The difference between the rigid method M (mean value of 0.01) and the other methods is significant. Method D seems to give a slightly better result with a mean correlation of 0.43.

On this measure, we have also performed an ANOVA, as laid out in Section IV-B. We have obtained $F = 395.6$ with $p = 0.0001$ indicating that the methods provide significantly different results. As a matter of fact, the difference between rigid method M (mean value of 0.01) and the other methods is large. This difference can be also observed in Fig. 4 where the distribution of L_{vv} correlation after registration has been plotted for each method.

D. Consistency of the Deformation Field

A key problem in registering brains of different subjects is the consistency of the estimated deformation field. The various ways in which the registration problem is modeled lead to different properties of the deformation field: continuous, differentiable, one to one, onto, etc. In paper, we did not try to exhibit differential properties of the field (this study is not straightforward because the deformation fields are defined in discrete space), but we want to exhibit possible singularities of the field.

As noted in [5], the Jacobian of the transformation characterizes these singularities especially zero-crossings. Recall that the

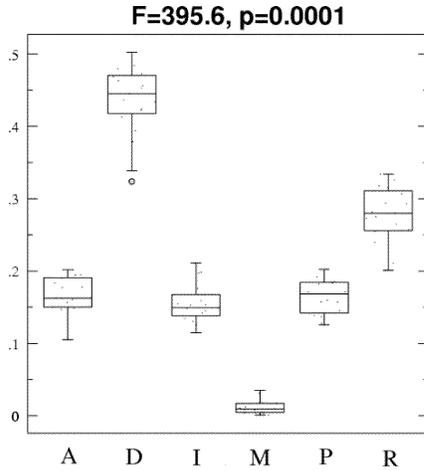


Fig. 4. For each method, distribution of L_{vv} correlation after registration. The ANOVA analysis has been performed on these groups.

Jacobian is the determinant of the first partial derivatives of the transformation, thus leading to the following expression:

$$\begin{aligned}
 J(i, j, k) = & \left(1 + \frac{\partial u}{\partial x}\right) \left(1 + \frac{\partial v}{\partial y}\right) \left(1 + \frac{\partial w}{\partial z}\right) + \frac{\partial u}{\partial y} \frac{\partial v}{\partial z} \frac{\partial w}{\partial x} \\
 & + \frac{\partial u}{\partial z} \frac{\partial v}{\partial x} \frac{\partial w}{\partial y} - \left[\left(1 + \frac{\partial u}{\partial x}\right) \frac{\partial v}{\partial z} \frac{\partial w}{\partial y} \right. \\
 & \left. + \left(1 + \frac{\partial v}{\partial y}\right) \frac{\partial u}{\partial z} \frac{\partial w}{\partial x} + \left(1 + \frac{\partial w}{\partial z}\right) \frac{\partial u}{\partial y} \frac{\partial v}{\partial x} \right]
 \end{aligned}$$

where (u, v, w) are the components of the deformation field ω (the transformation h is defined as $h(x) \triangleq x + \omega(x)$). The partial derivatives were computed using a convolution with a Gaussian derivative filter with standard deviation equal to twice the spatial resolution of the voxels.

First, we compute for each method the percentage of subjects for which we encounter voxels having a negative Jacobian. All methods have a positive Jacobian except for method D where 13 out of 17 (76.5%) transformations were subject to folding as detected by zero-crossings of the discrete Jacobian. Furthermore, we have computed, for each subject and only for method D, the percentage of voxels where the Jacobian is negative. We have found that 0.6% on average of voxels had a negative Jacobian which represents less than 10^5 voxels in a volume of size $256^2 \times 176$. This number represents the average number of singularities of the deformation field.

In addition to this, we would like to stress the following statements:

- We have computed the Jacobian in the discrete domain. The positivity of the Jacobian in the discrete domain does not necessarily ensure that the Jacobian is positive in the continuous domain. However, when the Jacobian goes negative in the discrete domain, it is likely to be true also in the continuous domain.
- The positivity of the Jacobian does not ensure that the transformation computed from volume A toward volume B is the inverse of the transformation computed from volume B toward volume A .

E. Partial Conclusion

In this section, global measures have been used to evaluate the registration algorithms. Except for the mse between the average volume and the reference volume, it must be noted that the measures rely on anatomical features (grey and WM tissues, L_{vv} volume) that are not linked to the similarity used to drive the registration process.

As expected, method M, which is only a rigid transformation, gives the poorest results for intersubject matching. These measures make it possible to distinguish rigid and non rigid methods. More precisely, we think that the results are very related to the DOFs of the estimated transformations. A gross classification of the methods, from the lowest DOF to the highest, would be: M(rigid), P (piecewise affine), A (grid of 4 mm) and I (the inverse consistency constraint reduces the DOF of the transformation), R and D (dense deformation field with similar DOF). The results for global measures follow more or less the same order.

However, one should be careful when increasing the DOFs of the estimated transformation. As noted in Section IV-D, the Jacobian of the transformation might go negative, expressing a singularity of the field. It is difficult to affirm which properties of the deformation field are realistic. One could imagine that the Jacobian would tend to zero in areas where anatomical structures undergo a modification of topology (for instance, a sulcus being present or interrupted for different subjects).

V. LOCAL MEASURES

A. Segmentation of Cortical Sulci

Descriptive anatomy of the cerebral cortex is based on its subdivision in a set of sulci and gyri. The sulci are of great interest in this paper since they are relevant anatomical and functional landmarks. Due to the interindividual cortical variability, the matching of sulci is crucial to evaluate different registration methods.

Different authors have studied the detection and segmentation of cortical sulci [15], [29], [35], [36], [38], [43], [47], but we will only describe briefly the method we use in this paper [15]. The sulci shapes are extracted with the help of differential geometry and more precisely the curvature information. The “active ribbon” method makes it possible to model sulcal patterns by a set of surfaces in 3-D space, defined as B-splines, representing the buried part of the cerebral cortex.

A compact numerical description of a sulcus can be obtained by modeling this sulcus with a surface representing its deep part. The method used is based on the active contour paradigm evolving from a 1D curve located at the external part of the brain to a 2D surface modeling the medial axis of the sulcus. Since a detailed description of this method can be found in [15], we only recall its main stages below.

1) *Segmentation of Cortical Regions*: Prior to the extraction of cortical features we use a cooperation between contour-based and region-based segmentation methods in order to extract the brain and to label the GM and WM and cerebrospinal fluid (CSF) regions [24]. From this brain tissue classification procedure, a mask representing the cortex with the CSF included in

its folds is computed [15]. Two different anatomical structures belong to this mask: the gyri and the sulci. This mask is called cortical ROI.

2) *Segmentation of Cortical Folds*: The goal is to characterize sulci and gyri within the cortical ROI. As long as we deal with these highly convoluted shapes, one natural way to characterize sulci from gyri is to analyze the curvature information of all iso-intensity surfaces belonging to the ROI. Differential geometry allows us to describe the shape of an iso-surface by its two principal curvatures and by a combination of them, namely the Gaussian and the mean curvatures. The operator used to compute curvature information is a 3-D extension of the L_{vv} operator introduced by Florack [13], which has already been described in Section IV-C.

3) *Numerical Modeling of Sulci Using "Active Ribbons"*: A compact and parametric description of a sulcus can be obtained by a medial surface representing the buried part of this sulcus. The method used here consists of modeling this surface by using an "active ribbon" which evolves in the three-dimensional space from a 1D curve at the learning stage to a 2D surface at the final step. "Active ribbons" are based on the active model paradigm and simulate the behavior of a physical object submitted to a set of forces. Forces are defined such that a curve evolves from its initial position at the surface of the brain to the bottom of the sulci.

For each subject of the database, we extract 12 major sulci with the method described above. The sulci used for the evaluation project are the central sulcus, precentral sulcus, postcentral sulcus, sylvian fissure, superior frontal sulcus and superior temporal sulcus. Fig. 5 shows a volume rendering of the left cortex of the reference subject with the segmented sulci. It is well known that cortical sulci may be interrupted [32]. The "active ribbon" method makes it possible to extract topologically varying sulci with each part of a sulcus being described by a B-spline.

We used these extracted sulci to assess locally the registration of the different methods. For each subject, each sulcus is deformed toward the reference subject using the transformation from the given registration method. Since the sulci are modeled by 3-D B-splines and their associated control points, we deform each control point of the spline using trilinear interpolation. The deformed control points naturally define the deformed sulcus which can be compared to the corresponding sulcus of the reference subject.

We visualize how sulci deform toward the corresponding sulcus of the reference subject in Section V-B. Beyond visualization, a numerical evaluation is performed in Section V-C by computing distances between deformed sulci and corresponding sulci of the reference subject.

B. Visualization of Deformed Sulci

We have chosen to visualize how the left central sulci of the 17 subjects deform toward the left central sulcus of the reference subject. Fig. 6 shows the left central sulcus of the reference subject in yellow, the left precentral sulcus of the reference subject in red, and the left postcentral sulcus of the reference subject in green. The left central sulci of the different subjects are de-

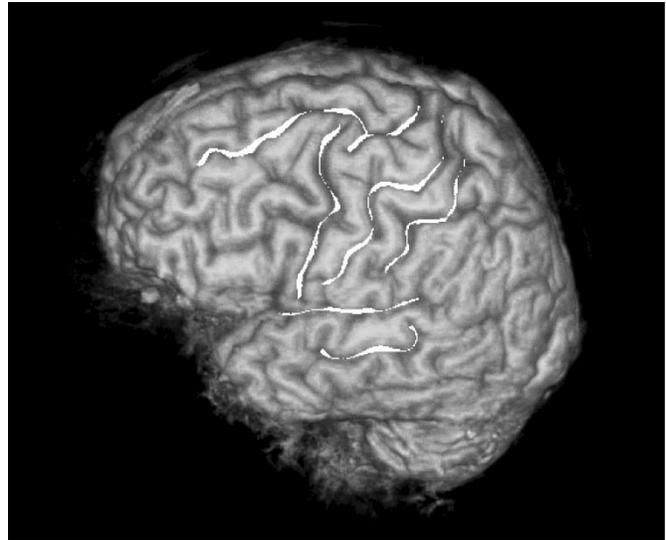


Fig. 5. Extracted sulci on the left hemisphere of one subject: central sulcus, precentral sulcus, postcentral sulcus, superior frontal sulcus, lateral sulcus and superior temporal sulcus. The deep parts of the sulci are also segmented but not visible on that figure due to the volume rendering of the subject's brain.

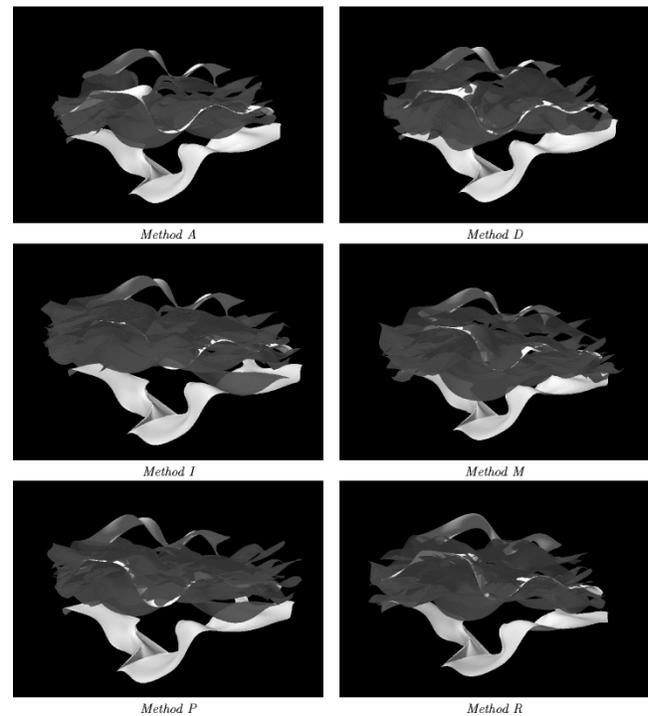


Fig. 6. Left central sulci of the database deformed toward the reference subject. The deformed sulci should ideally be superimposed to the left central sulcus of the reference subject (white). The left precentral sulcus (bottom, light gray) and postcentral sulcus (top, light gray) of the reference subject are also drawn. One should note that although deformed sulci are almost always in between the postcentral and the precentral sulci, the variability and dispersion of deformed sulci around the central sulcus is quite large.

formed toward the reference subject and are drawn in blue. Ideally, the blue sulci should be superimposed on the yellow sulcus.

It can be observed that (a) the different registration methods seem to give a significant dispersion around the reference sulcus and (b) it seems difficult to distinguish visually the performances of registration methods. The postcentral and

precentral sulci of the reference subject (in red and green) give the order of magnitude of the dispersion and indicate that deformed central sulci are located in between these two sulci. However, in most cases, a crude identification algorithm (i.e., labeling deformed sulci as central sulci) based on the position of deformed sulci with respect to the left central reference sulcus would probably not give satisfactory results.

C. Numerical Evaluation

Beyond visualization, a numerical evaluation is needed. In this section, we investigate two measures that reflect more or less two aspects: the global positioning of sulci on the one hand (Section V-C1) and the similarity of shapes on the other hand (Section V-C2).

1) *Distance Between Registered Sulci*: To assess how well sulci are matched, we compute the distance between a sulcus deformed toward the reference subject and the corresponding sulcus of the reference subject. It is an interesting but challenging question to know how sulci should be matched. Therefore, we compute four distances :

- As explained in Section V-A, sulci are modeled by B-splines, and may be resampled. It is possible to associate the distance between sulci to the distance between control points. This distance assumes that there is a one-to-one correspondence between sulci control points. There might be cases where homologous sulci are interrupted. In this case, we use the following procedure (see Fig. 7) : 1) If we note p_m as the maximum depth ($p_m = \max\{p_0, p_1, p_2\}$), the sulci are resampled along the depth direction in order to obtain the same number of control points (p_m) for all the sulci segments on their depth axis. 2) Note $l_m = \max\{l_1 + l_2, l_0\}$ (l_0 is the length of the uninterrupted sulcus and l_1 and l_2 and the length of the interrupted sulcus). If $l_m = l_1 + l_2$, the sulcus \mathcal{S}_0 is resampled by a factor of l_m/l_0 and sulci $\mathcal{S}_i, i \in \{1, 2\}$ remain identical. If $l_m = l_0$, then the sulci segments $\mathcal{S}_i, i \in \{1, 2\}$ are resampled by a factor of $l_m/l_1 + l_2$ while sulcus \mathcal{S}_0 remains identical. When two homologous sulci are both interrupted, we match each segment as if they were continuous. This is possible because we have a labeling of each piece, Inferior-Superior (for the precentral sulcus for instance) or Anterior-Posterior (for the superior temporal sulcus for instance). We have presented an approach for a sulcus described by two distinct segments, but this method can easily be extended if we have to deal with sulci having more segments. This distance between control points is referred to as distance D_1 .
- A distance between sulci can be given by the distance between centers of gravity. This distance between sulci gravity centers is referred to as distance D_2 .
- The assumption of control points correspondence is questionable. Consequently, we compute a distance based on the assumption that a point of a sulcus should correspond to the nearest point on the other sulcus. This distance between sulci closest points is referred to as distance D_3 .

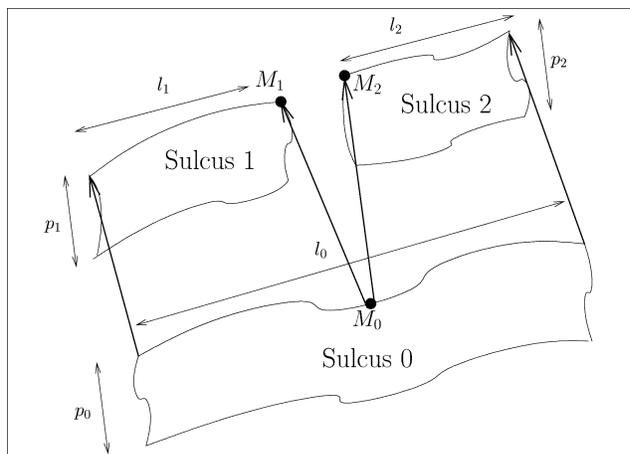


Fig. 7. Left: distances of two homologous sulci, one being interrupted.

TABLE IV
AVERAGE DISTANCE IN FRACTIONAL VOXELS BETWEEN REGISTERED SULCI AND CORRESPONDING SULCI OF THE REFERENCE SUBJECT. FOUR DISTANCES HAVE BEEN COMPUTED: D_1 DISTANCE BETWEEN SULCI CONTROL POINTS; D_2 DISTANCE BETWEEN SULCI GRAVITY CENTERS; D_3 DISTANCE BETWEEN SULCI CLOSEST POINTS; D_4 SYMMETRIC HAUSDORFF DISTANCE. FOR ALL DISTANCES, THE MEAN IS COMPUTED FOR ALL THE SUBJECTS AND ALL THE SULCI

Method	D_1	D_2	D_3	D_4
A	9.9	7.3	4.2	16.4
D	10.3	7.7	4.3	17.9
I	11.4	9.0	5.4	19.1
M	11.5	9.1	5.8	18.7
P	10.7	8.2	4.9	18.2
R	10.8	8.3	4.8	17.7

- A classical distance between shapes is the Hausdorff distance. Between two point sets A and B , the symmetric Hausdorff distance is defined as $D(A, B) = \max\{h(A, B), h(B, A)\}$, where $h(A, B) = \max_a \min_b \|a - b\|$. This measure is *a priori* sensitive to noise. This symmetric Hausdorff distance is referred to as distance D_4 .

To present a compact measure, the mean of each distance is computed for all the subjects and all the sulci (we have 12 sulci extracted for each of the 18 subjects). Results are presented in Table IV where the distances are expressed in fractional voxels (the resolution of the voxels is 0.93 mm, see Section III for more details).

We have performed an ANOVA on the distance measurement D_1 , as laid out in Section IV-B. The following results have been obtained: $F = 2.001$ with $p = 0.085$. Based on these results, it seems difficult to distinguish clearly the registration methods. In other words, the variability within groups (intramethod variability) is quite comparable to the variability between groups (intermethod variability). Both variances are shown in Fig. 8 where the distribution of mean sulci distance has been plotted for each method.

It is also interesting to compute the distance for particular sulci. We chose to restrict the computation of D_1 to the central sulci, superior frontal sulci, and sylvian sulci (left and right hemisphere). These results are presented in Table V. It can be seen that the distances for the central sulci are significantly

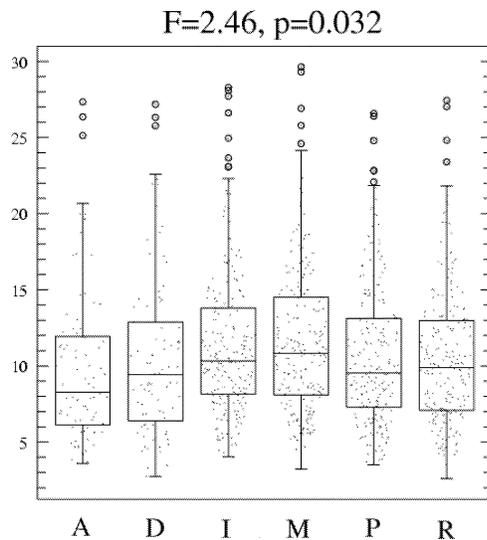


Fig. 8. Distribution of mean sulci distance D_1 after registration. Since the distances give the same classification, we have retained distance D_1 for the ANOVA.

lower which could be explained by the stability of the central sulcus over individuals [32]. A surprising result comes from the superior frontal sulcus for which the method M (rigid registration), gives a better matching than nonrigid methods.

2) *Statistical Study of Deformed Shapes*: The distance between registered sulci is not a sufficient measure to characterize how sulci deform. This is illustrated in Fig. 9, where two sulci 1 and 2 are deformed toward the reference sulcus. Sulcus 1 is probably the best in terms of positioning, but sulcus 2 may be better in terms of shape. To characterize shape similarity, we use the principal components analysis (PCA) technique [8].

For each method, we have a population of shapes that are computed from the sulci of the different subjects deformed toward the reference subject according to the transformation of the considered registration method. The purpose of PCA is to analyze the variations of each shape \mathbf{x} with respect to the reference shape \mathbf{x}_{ref} (the reference sulcus). The main idea of PCA is to decompose the displacement $\tilde{\mathbf{x}} = \mathbf{x} - \mathbf{x}_{ref}$ on the eigenvectors of the covariance matrix ($C = \mathbb{E}[(\mathbf{x} - \mathbf{x}_{ref})(\mathbf{x} - \mathbf{x}_{ref})^T]$). This decomposition induces a metric of the shape space.

For brevity, we have chosen to consider the trace of the covariance matrix. This measure reflects the entire variation of the population around the reference sulcus along all the axes of the decomposition. Furthermore, traces can be compared since it is invariant when the axes of the decomposition change (under the condition that the referential spaces are orthogonal). These results are given in Table VI for the left central sulci, left superior frontal sulci, and left sylvian sulci. For each method, the trace is normalized by the number of subjects involved.

In addition to the total variation along all axes, it is possible to compute the percentage of total variation that is explained by each mode. Fig. 10 presents the relative importance of variation modes for the first few modes (the next modes are not relevant since they explain less than 1% of variation). We observe that the importance of the first modes for nonrigid methods is greater (the two first modes explain more than 75% of the dispersion for

TABLE V
MEAN DISTANCE BETWEEN DEFORMED SULCI AND CORRESPONDING SULCI COMPUTED ON SUB-GROUPS OF EXTRACTED SULCI (CENTRAL SULCI, SUPERIOR FRONTAL SULCI, AND SYLVIAN FISSURE)

Method	Central	Superior frontal	Sylvian
A	6.5	10.4	9.5
D	6.9	11.1	9.9
I	8.3	12.0	10.8
M	8.5	10.5	11.8
P	7.1	11.9	10.2
R	7.4	10.8	9.6

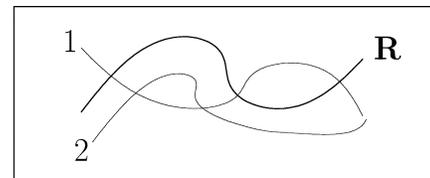


Fig. 9. The sulcus R is the reference sulcus, and sulci 1 and 2 are deformed sulci that should ideally match the reference sulcus. Which one is the best?

TABLE VI
FOR THREE DIFFERENT POPULATIONS OF SULCI, THE VARIATIONS OF DEFORMED SULCI CAN BE ANALYZED BY PCA. THE TRACE OF THE COVARIANCE MATRIX, NORMALIZED BY THE NUMBER OF SUBJECTS, EXPRESSES THE ENTIRE VARIATION OF DEFORMED SULCI AROUND THE REFERENCE SULCUS IN THE SHAPE SPACE. THE LOWER THE TRACE, THE LOWER THE VARIATION AROUND THE REFERENCE SULCUS

Method	central	superior frontal	sylvian
A	547	736	1172
D	675	767	1046
I	723	907	1121
M	621	622	1373
P	510	859	1233
R	735	741	1064

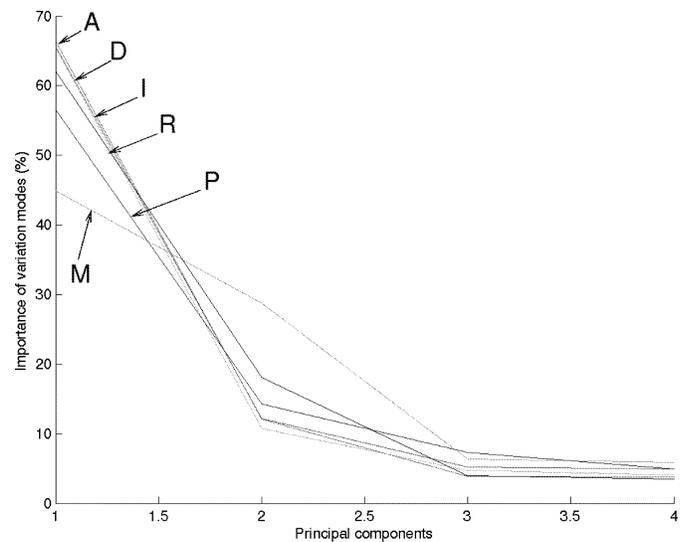


Fig. 10. Relative importance of variation modes for each registration method. Non-rigid methods A, D, I, and R are characterized by a strong first mode (almost 70% of total variation) while methods M and P have more equally distributed modes.

methods A, D, I, and R). The variation modes for rigid methods seem to be more equally distributed.

D. Partial Conclusion

In this section, we used local measures based on matching sulci to evaluate different registration methods. These local measures are based on matching cortical sulci extracted with the method described in [15]. We have provided visual inspection and numerical evaluation of the registration of sulcal patterns.

These measures do not make it possible to affirm that nonrigid methods used in this paper perform better in registering major cortical sulci from different subjects. This result is quite surprising.

In addition, average distances between deformed sulci and reference sulci should be compared to average distances between neighboring sulci of the reference subject. Let us note that for the reference subject, the distance between the precentral (respectively postcentral) sulcus and the central sulcus is 25 voxels (30 voxels, respectively). These distances are four times larger than the distances between deformed left central sulci and the reference left central sulcus displayed in Table V. In other words, the left central sulci of the subjects deformed toward the reference subject are on average four times closer to the central sulcus than they are to postcentral and precentral sulci. This result combined with topological knowledge about the sulci (i.e., the central sulcus is in between the precentral sulcus and the postcentral sulcus) could be used to automatically identify sulci on the basis of a nonrigid registration method.

VI. CONCLUSION

The questions that motivated our study (relevance, existence, and expectations of intersubject registration methods) are challenging issues. In this paper, we have proposed an evaluation framework for methods that aim at registering brains of different subjects. Global and local measures of the registration have been used to evaluate six registration methods on a database of 18 subjects. The relevance of various methods to register anatomical structures has been investigated. More precisely, we focused on the matching of cortical patterns which are relevant in the context of anatomical and functional normalization. On one hand, global measures show the efficiency of nonrigid methods and indicate that the quality of the registration increases with the DOFs of the estimated transformation. On the other hand, rigid and nonrigid methods give surprisingly similar results for local measures which are based on the matching of major cortical sulci.

We computed the Jacobian of the transformation to investigate whether singularities of the deformation field are present. This study has revealed that all methods except method D produce positive Jacobians of the transformation. We first state that the Jacobian measure has been computed in the discrete domain. In order to affirm that the transformation can be inverted, the transformation has to be continuous, piecewise differentiable and have a positive Jacobian in the continuous domain. The positivity of the Jacobian at discretization nodes does not necessarily ensure that this property is true in the continuous domain.

To explain the surprising results concerning the matching of cortical sulci, we must first keep in mind that the variability of cortical patterns between individuals is very high [32]. We could

be tempted to affirm that the modeling of the problem, based on the matching of voxels having comparable luminance, might not be adapted to address this huge variability. All methods would benefit by combining landmarks with intensity-based registration. Some methods have been proposed to explicitly introduce sparse constraints in the registration process [2], [4], [6], [20], [22].

A questionable issue is the choice of the reference subject. In this study, it is quite difficult to assess the impact of that choice on the results. However, we note that some studies have been conducted [18] indicating that this choice has a minimal influence on the result. Ideally, this study should have been conducted 18 times using a different reference subject each time. The major problem would then be the resource management (the data to be transferred would then be larger than 1300 GB).

Future work should focus on functional data. It would be interesting to know what the impact of nonrigid anatomical registration has on the variability (spatial dispersion) of the functional data since intersubject registration methods are also dedicated to the anatomo-functional mapping.

ACKNOWLEDGMENT

The authors would also like to thank the Radiology Department of Pontchaillou hospital for the data acquisition. They would also like to thank L. Hirschy for his help in preparing the results for the consistent image registration method I.

REFERENCES

- [1] J. Van Bemmelen and M. Musen, *Handbook of Medical Informatics*. Berlin, Germany: Springer-Verlag, 1997.
- [2] P. Cachier, J.-F. Mangin, X. Pennec, D. Rivière, D. Papadopoulos-Orfanos, J. Régis, and N. Ayache, "Multisubject nonrigid registration of brain MRI using intensity and geometric features," in *Lecture Notes in Computer Science*. Berlin, Germany: Springer-Verlag, 2001, vol. 2208, Proc. MICCAI, pp. 734–742.
- [3] A. Counce and C. Taylor, "Using local geometry to build 3D sulcal models," in *Lecture Notes in Computer Science*. Berlin, Germany: Springer-Verlag, 1999, vol. 1613, Proc. Information Processing in Medical Imaging, pp. 196–209.
- [4] G. E. Christensen, S. C. Joshi, and M. I. Miller, "Volumetric transformation of brain anatomy," *IEEE Trans. Med. Imag.*, vol. 16, pp. 864–877, June 1997.
- [5] G. E. Christensen and H. J. Johnson, "Consistent image registration," *IEEE Trans. Med. Imag.*, vol. 20, pp. 568–582, July 2001.
- [6] L. Collins, G. Le Goualher, R. Venugopal, A. Caramanos, A. Evans, and C. Barillot, "Cortical constraints for nonlinear cortical registration," in *Lecture Notes in Computer Science*. Berlin, Germany: Springer-Verlag, Sept. 1996, vol. 1131, Proc. Visualization in Biomedical Computing, pp. 307–316.
- [7] L. Collins and A. Evans, "Animal: Validation and applications of nonlinear registration-based segmentation," *Int. J. Pattern Recogn. Artif. Intell.*, vol. 8, no. 11, pp. 1271–1294, 1997.
- [8] T. Cootes, C. Taylor, D. Hooper, and J. Graham, "Active shape models—their training and application," *Comput. Vis. Image Understanding*, vol. 61, no. 1, pp. 31–59, 1995.
- [9] M. Desvignes, N. Royackkers, H. Fawal, and M. Revenu, "Detection and identification of sulci on 3D MRI," *Human Brain Mapping*, p. 410, 1997.
- [10] A. Evans, W. Dai, D. Collins, P. Neelin, and T. Marrett, "Warping of computerized 3D atlas to match brain image volumes for quantitative neuroanatomical and functional analysis," *Proc. SPIE*, vol. 1445, pp. 236–246, 1991.
- [11] A. Evans, L. Collins, and B. Milner, "A MRI-based stereotaxic atlas from 250 young normal subjects," *Soc. Neurosc. Abstr.*, vol. 18, p. 408, 1992.

- [12] J. Fitzpatrick, D. Hill, and C. Maurer, "Image registration," in *Handbook of Medical Imaging*, M. Sonka and J. Fitzpatrick, Eds. Bellingham, WA: SPIE, 2000, vol. 2, Medical Image Processing and Analysis, pp. 447–513.
- [13] L. Florack, B. Romeny, J. Koenderink, and M. Viergever, "Scale and the differential structure of images," *Image Vis. Computing*, vol. 10, pp. 376–388, 1992.
- [14] B. Gibaud, S. Garlatti, C. Barillot, and E. Faure, "Computerised brain atlases as decision support systems: A methodological approach," *Artif. Intell. Med.*, vol. 14, pp. 83–100, 1998.
- [15] G. Le Goualher, C. Barillot, and Y. Bizais, "Modeling cortical sulci with active ribbons," *Int. J. Pattern Recogn. Artif. Intell.*, vol. 8, no. 11, pp. 1295–1315, 1997.
- [16] G. Le Goualher, E. Procyk, L. Collins, R. Venegopal, C. Barillot, and A. Evans, "Automated extraction and variability analysis of sulcal neuroanatomy," *IEEE Trans. Med. Imag.*, vol. 18, pp. 206–217, Mar. 1999.
- [17] D. Graf Von Keyserlingk, K. Niemann, and J. Wasel, "A quantitative approach to spatial variation of human cerebral sulci," *Acta Anatomica*, vol. 131, pp. 127–131, 1988.
- [18] A. Guimond, J. Meunier, and J. Thirion, "Average brain models: A convergence study," *Computer Vision and Image Understanding*, vol. 77, pp. 192–210, 2000.
- [19] P. Hellier, C. Barillot, E. Mémin, and P. Pérez, "Hierarchical estimation of a dense deformation field for 3-D robust registration," *IEEE Trans. Med. Imag.*, vol. 20, no. 5, pp. 388–402, 2001.
- [20] P. Hellier and C. Barillot, "Coupling dense and landmark-based approaches for non rigid registration," *IEEE Trans. Med. Imag.*, vol. 22, pp. 217–227, Feb. 2003.
- [21] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the hausdorff distance," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, pp. 850–963, Sept. 1993.
- [22] H. J. Johnson and G. E. Christensen, "Consistent landmark and intensity-based image registration," *IEEE Trans. Med. Imag.*, vol. 21, pp. 450–461, May 2002.
- [23] R. Kikinis, M. Shenton, D. Iosifescu, R. McCarley, P. Saiviroonporn, H. Hokama, A. Robatino, D. Metcalf, C. Wible, C. Portas, R. Donnino, and F. Jolesz, "A digital brain atlas for surgical planning, model driven segmentation and teaching," *IEEE Trans. Visual. Comput. Graphics*, vol. 2, pp. 232–241, Sept. 1996.
- [24] F. Lachmann and C. Barillot, "Brain tissue classification from MRI data by means of texture analysis," *Proc. SPIE Conf. Medical Imaging VI: Image Processing*, vol. 1652, pp. 72–83, 1992.
- [25] T. Lehmann, C. Gonner, and K. Spitzer, "Survey: Interpolation methods in medical image processing," *IEEE Trans. Med. Imag.*, vol. 18, pp. 1049–1075, Nov. 1999.
- [26] H. Lester and S. Arridge, "A survey of hierarchical nonlinear medical image registration," *Pattern Recogn.*, vol. 32, no. 1, pp. 129–149, 1999.
- [27] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imag.*, vol. 16, pp. 187–198, Apr. 1997.
- [28] J. Maintz and M. Viergever, "A survey of medical image registration," *Med. Image Anal.*, vol. 2, no. 1, pp. 1–36, 1998.
- [29] J.-F. Mangin, V. Frouin, I. Bloch, J. Regis, and J. López-Krahe, "From 3d magnetic resonance images to structural representations of the cortex topography using topology preserving deformations," *J. Math. Imag. Vis.*, vol. 5, no. 4, pp. 297–318, 1995.
- [30] J. Mazziotta, A. Toga, A. Evans, P. Fox, and J. Lancaster, "A probabilistic atlas of the human brain: Theory and rationale for its development," *Neuroimage*, vol. 2, pp. 89–101, 1995.
- [31] M. I. Miller, G. E. Christensen, Y. Amit, and U. Grenander, "Mathematical textbook of deformable neuroanatomies," *Proc. Nat. Acad. Sci.*, vol. 90, pp. 11 944–48, 1993.
- [32] M. Ono, S. Kubik, and C. Abernathy, *Atlas of the Cerebral Sulci*. Berlin, Germany: Springer-Verlag, 1990.
- [33] X. Pennec, P. Cachier, and N. Ayache, "Understanding the demon's algorithm: 3D non rigid registration by gradient descent," in *Lecture Notes in Computer Science*, vol. 1679, Proc. MICCAI. Cambridge, U.K., Sept. 1999, pp. 597–605.
- [34] D. Riviere, J. Mangin, D. Papadopoulos-Orfanos, J. Martinez, V. Frouin, and J. Regis, "Automatic recognition of cerebral sulci using a congregation of neural networks," in *Lecture Notes in Computer Science*, vol. 1935, Proc. MICCAI, Berlin, Germany, 2000, pp. 40–50.
- [35] N. Royackkers, H. Fawal, M. Desvignes, M. Revenu, and J. Travere, "Morphometry and identification of brain sulci on three-dimensional MR images," in *Proc. Information Processing in Medical Imaging*, Brest, Juin 1995, pp. 379–380.
- [36] S. Sandor and R. Leahy, "Toward automated labeling of the cerebral cortex using a deformable atlas," in *Information Processing in Medical Imaging*. Norwell, MA: Kluwer Academic, 1995, pp. 127–138.
- [37] ———, "Surface-based labeling of cortical anatomy using a deformable atlas," *IEEE Trans. Med. Imag.*, vol. 16, no. 1, pp. 41–54, 1997.
- [38] G. Szekely, C. Brechbühler, O. Kübler, R. Ogniewicz, and T. Budinger, "Mapping the human cerebral cortex using 3D medial manifolds," in *Proc. SPIE Conf. Visualization in Biomedical Computing*, 1992, pp. 130–144.
- [39] J. Talairach and P. Tournoux, *Co-Planar Stereotaxic Atlas of the Human Brain*. Stuttgart, Germany: Georg Thieme Verlag, 1988.
- [40] J. P. Thirion and A. Gourdon, "The Marching Lines Algorithm: New Result and Proofs," INRIA, Rennes, France, Res. Rep. 1881, 1993.
- [41] J. Thirion, "Image matching as a diffusion process: An analogy with Maxwell's demons," *Med. Image Anal.*, vol. 2, no. 3, pp. 243–260, 1998.
- [42] P. Thompson, R. Woods, M. Mega, and A. Toga, "Mathematical/computational challenges in creating deformable and probabilistic atlases of the human brain," *Human Brain Mapping*, vol. 9, pp. 81–92, 2000.
- [43] M. Vaillant and C. Davatzikos, "Finding parametric representation of the cortical sulci using an active contour model," *Med. Image Anal.*, vol. 1, no. 4, pp. 295–315, 1996.
- [44] P. Viola and W. Wells, "Alignment by maximization of mutual information," *Int. J. Comput. Vis.*, vol. 24, no. 2, pp. 137–154, 1997.
- [45] J. West, J. Fitzpatrick, M. Wang, B. Dawant, C. Maurer, R. Kessler, R. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetend, D. Vandermeulen, P. Van Den Elsen, S. Napel, T. Sumanaweera, B. Harkness, P. Hemler, D. Hill, D. Hawkes, C. Studholme, J. Maintz, M. Viergever, G. Malandain, X. Pennec, M. E. Noz, G. Q. Maguire, M. Pollock, C. A. Pellizzari, R. A. Robb, D. Hanson, and R. Woods, "Comparison and evaluation of retrospective intermodality brain image registration techniques," *J. Comput. Assist. Tomogr.*, vol. 21, no. 4, pp. 554–566, 1997.
- [46] J. West, J. Fitzpatrick, M. Wang, B. Dawant, C. Maurer, R. Kessler, and R. Maciunas, "Retrospective intermodality registration techniques for images of the head: Surface-based versus volume-based," *IEEE Trans. Med. Imag.*, vol. 18, pp. 144–150, Feb 1999.
- [47] X. Zeng, L. H. Staib, R. T. Schultz, H. Tagare, L. Win, and J. S. Duncan, "A new approach to 3D sulcal ribbon finding from MR images," in *Proc. MICCAI*, Sept. 1999, pp. 148–157.