Average Brain Models: A Convergence Study

Alexandre Guimond, Jean Meunier, Jean-Philippe Thirion

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Average Brain Models: A Convergence Study

Alexandre Guimond — Jean Meunier — Jean-Philippe Thirion

N° 3731

Juillet 1999

THÈME 3

ISSN 0249-6399

INRIA
SOPHIA ANTIPOLIS

Average Brain Models: A Convergence Study

Alexandre Guimond* , Jean Meunier† , Jean-Philippe Thirion

Thème 3 — Interaction homme-machine,
images, données, connaissances
Projet Epidaure

Rapport de recherche n° 3731 — Juillet 1999 — 21 pages

Abstract: We present a completely automatic method to build "stable" average anatomical models of the human brain using a set of magnetic resonance (MR) images. The models computed present two important characteristics: an average intensity and an average shape, both in a single image. We provide results showing convergence toward the centroid of the image set used for the computation of the model. In particular, the RMS distances between the model and the MR images contained in the set stabilize in a range of 2.88mm to 3.36mm from a range of 4.62mm to 5.51mm initially after only one iteration. As for the influence of the reference image chosen for the model construction, this is minimal with differences of about 1.0mm, from approximately 3.5mm initially. These results ensure the usefulness of our approach.

Key-words: Anatomical Atlas, Average Model, Registration, Magnetic Resonance Imaging (MRI).

* Email: alexandre.guimond@sophia.inria.fr
† Université de Montréal, C.P. 6128, Succursale Centre-Ville, Montréal (Québec), H3C 3J7 Canada.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, B.P. 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : 04 92 38 77 77 - International : +33 4 92 38 77 77 —Fax : 04 92 38 77 65 - International : +33 4 92 38 77 65

Modèles moyens du cerveau: Une étude de convergence

Résumé : Nous présentons une méthode complètement automatique de construction de modèles moyens anatomiques "stables" du cerveau humain en utilisant un ensemble d'images de résonance magnétique (RM). Les modèles calculé ont deux caractéristiques importantes: une intensité moyenne et une forme moyenne, toutes deux dans une seule image. Nous présentons des résultats montrant la convergence du modèle vers le centroïde de l'ensemble d'images utilisées lors de la construction. En particulier, les distances RMS entre le modèle et les images de l'ensemble se stabilisent a l'interieur d'un intervalle de 2.88mm à 3.36mm après seulement une itération, ces valeurs ébtant dans un intervalle de 4.62mm à 5.51mm initialement. Quant à l'influence de l'image de référence choisie pour la construction modèle, elle est minime avec des différences d'environ 0.9mm, initialement a 3.5mm. Ces résultats assurent l'utilité de notre approche.

Mots-clés : Atlas anatomique, Modèle moyen, Recalage, Imagerie par résonance magnétique (IRM).

1 Introduction

An important tool used to diagnose abnormal anatomical variations are medical atlases [1]. Traditional ones, such as by Talairach & Tournoux [2] or Schaltenbrand & Wahren [3], are presented in textbooks, but computerized atlases comprising information in a more practical and quantitative manner are becoming available [4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16]. They usually include information obtained from a set of subjects, as opposed to a single individual in most paper atlases, making them more representative of a population. For example, the Montreal Neurological Institute (MNI) used three hundred and five (305) normal subjects to build an atlas comprising intensity variations after affine registration in the stereotactic space defined by Talairach & Tournoux [8]. These methods also enable the calculation of normal shape variations, such as in the work of Gee *et al.* [17] which present a statistical framework for the construction of upgradable statistical atlases, and Thompson & Toga [18] which presents a probabilistic atlas of the human brain based on random vector field transformations.

The following work aims to develop and validate a concept drafted in a previous paper [19] to build an average model of the human brain using a set of magnetic resonance (MR) images obtained from normal subjects. This model has two important characteristics: average tissue intensity and average tissue shape up to an affine transformation. We intend to demonstrate that the model construction converges toward the centroid of the MR image set.

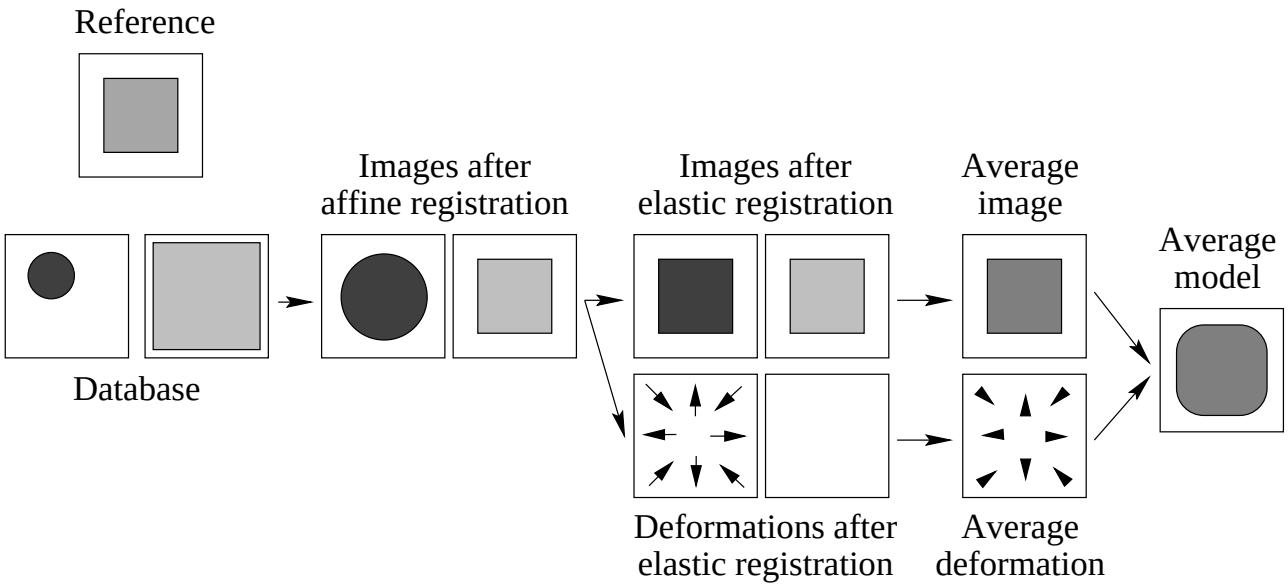


Figure 1: Average model construction method.

As depicted in Figure 1, our method can be summarized in the following manner. Affine registration between all the images of the set and a reference image corrects for positioning and global shape differences due to translation, rotation, scaling and shearing, as well as global linear intensity variations due to acquisition parameters or preprocessing. These are variations that are not of concern for our study. Elastic registration is then used to evaluate residual variations due to pure morphological differences and produce images having the same shape as the reference. Averaging the residual deformations and the locally registered images yields an average deformation and an average intensity image, respectively. The average deformation is then applied to the average image to produce the model. It presents in a single image an average intensity and shape modulo an affine transformation corresponding to the affine characteristics of the reference image.

Although similar in terminology, the average shape and the average intensity characteristics have different purposes. The average shape represents an average of *morphological variations*. This is generally a well understood concept. On the other hand, the average intensity is meant to average the *response of corresponding tissues to the acquisition parameters*. Also, assuming Gaussian noise in each MR scan, the average intensity increases the signal to noise ratio in the resulting model compared with individual scans.

The main contribution of this paper is the description of a fully automatic technique to obtain an *average intensity* and *shape* image, producing the *average model* M, and to show that this model, up to an affine

transformation, is stable with respect to the choice of the initial reference image and repeated applications of the algorithm (iterations).

The most similar work regarding average intensity atlases is that of Bookstein [20] who created from nine MR scans a two-dimensional image representing the average intensity of the mid-sagittal plane. Thirteen manually identified landmarks in the mid-sagittal plane of each scan where matched with a reference image using the thin-plate spline interpolant [21]. The nine resampled images where then averaged to result into a morphometric average atlas. Our method differs mainly by two aspects. First, as suggested by Bookstein [20], we make full use of the three-dimensionality of the scans to compute a three-dimensional average image. Second, our registration method is automatic and computes a dense deformation field instead of an interpolated function based on thirteen landmarks. This deformation identifies for each voxel of the reference the corresponding positions in the other scans. Within this process, every voxel of the reference can be though of as a landmark automatically determined in the other scans.

The work of the MNI group [8], where three hundred and five (305) three-dimensional MR scans were registered using translations, rotations and scalings, and averaged to build a statistical neuroanatomical model, also relates to our work. We enrich this idea by proceeding further in using a less constrained type of deformation after the affine match to accommodate for local shape variations.

In the same vein, Woods *et al.* [13] describe a method that finds from a set of images a common space that preserves the average orientation, size, and affine shape of the group by registering all possible pairs of images it contains. Averaging the images after affine mapping to this common space produces an average intensity brain atlas in the average affine space. Their method is computationally very intensive as it requires $n(n-1)/2$ registrations, n being the number of subjects in the group. The basic difference between this approach and the one by the MNI group is that this one finds the average affine space, whereas the MNI method uses Talairach space.

The average shape concept is most similar to the beautiful work of the Brown/Washington group [14, 22], who have put together a framework in which the construction of a template from a set of anatomies is proven to minimized the energy of the deformations required to map it onto all the elements of that set. Our work complements theirs in that we provide quantitative measurements confirming their formulation, though we do not compute small deformations as is required for their proof. It should be noted that our resulting model also includes average intensity information and that our respective groups use different registration methods.

Le Briquer and Gee [12] have also developed a method that provides, for a given group of subjects, the mean shape and the modes of principal variation along with their amplitude. Their approach is set in a statistical framework and aims at deriving a shape model. Our method differs in that we aim the analysis of local information rather than global patterns.

The work presented here also relates to the methodologies of Subsol *et al.* [16], Bookstein [23] and Kendall [24] who compute average shapes modulo similarity or affine transformations. We have not tried to strictly follow the theory developed in their works. Our intention was to conform to the idea of making abstraction of differences between images due to first order transformations, and analyze residual variations. Our main contribution resides in the characteristics used to build the average shape, that is the image intensities instead of landmarks or crestlines. Again, this enables the computation of dense deformations fields representing variations everywhere is the MR scan, as opposed to interpolating transformations found using landmarks, lines or surfaces. We believe this technique may find less accurate matches in the close surroundings of the landmarks, but provides better overall registration.

As will be shown, compared to these previous efforts, our method provides clearer images with higher contrasts and more sharp definitions of tissue boundaries. Most importantly, we provide numbers showing the convergence of the model towards the centroid of the image set.

The remaining sections of this paper are organized in the following manner. First, we detail the method used to construct the average model. We then present results showing the convergence of the method towards an average intensity and shape model, and show the effect of the choice of reference image. We conclude by a discussion on future research tracks.

2 Methodology

2.1 Registration

The work that follows assumes each point in one image has a corresponding equivalent in the others. It also assumes available a matching method able to find these correspondences and capable of providing a vector field

representing those relationships. In theory, neither of these conditions is realized. That is, at a microscopic scale, there is not a one to one relationship between the brain cells of two individuals, and assuming there was, to this day, no algorithm is able to find it. In practice however, deforming one brain so its shape matches the one of another is conceivable and many algorithms realizing this process have been developed [21, 4, 25, 26, 27, 17, 13]. The procedure used in the following work is the demons method [28] using a complete grid of demons. We briefly detail it here and refer the reader to the original article for more information.

2.1.1 Evaluating Shape Differences

When applied to MR images, the demons algorithm can be considered as an optical flow variant [29]. From this point of view, the 3D images to be registered are considered as a time sequence represented by $I(x,t)$ where $x = (x_1, x_2, x_3)$ is a voxel position in the image and t is time. It computes forces by constraining the brightness of brain structures to be constant in time so that

$$\frac{dI(x,t)}{dt} = 0.$$

This leads us to the basic optical flow formulation (See [29] for details about the derivation.),

$$v = -\nabla_x I(x,t) \frac{\partial I(x,t)/\partial t}{||\nabla_x I(x,t)||^2},$$

which is the movement component in the brightness gradient direction $\nabla_x I(x,t) = \left(\frac{\partial I(x,t)}{\partial x_1}, \frac{\partial I(x,t)}{\partial x_2}, \frac{\partial I(x,t)}{\partial x_3}\right)$. For numerical stability reasons when $\nabla_x I(x,t)$ is close to zero, the denominator of the above formula is modified to result in the basic displacement formulation for the demons algorithm using a complete grid of demons,

$$v = -\nabla_x I(x,t) \frac{\partial I(x,t)/\partial t}{||\nabla_x I(x,t)||^2 + |\partial I(x,t)/\partial t|^2}.$$

When $||\nabla_x I(x,t)|| = 0$ no displacement is computed.

As with all optical flow formulations based on differential techniques, the problem here resides in finding the components of the movement in the directions orthogonal to the gradient. Many regularization methods have been proposed [30] each with their strengths and weaknesses. The one proposed by Thirion is to apply a Gaussian filter to each of the three components of v. This provides a smooth displacement field in a time efficient way. It is interesting to know that Bro-Nielsen and Gramkow [31] have shown that regularizing the deformation field using a Gaussian filter approximates linear elasticity.

The method is iterative and makes use of a multi-scale scheme which resolved the problem of finding large deformations, a common problem with optical flow techniques and a basic assumption in the formulation and implementation of the derivative filters.

2.1.2 Relaxing the Intensity Constraint

We mentioned that the registration algorithm assumes the same intensity for corresponding brain structures in the images to be registered. For all sort of reasons, such as acquisition parameters or preprocessing, this may not be the case. To relax this constraint, a linear intensity correction is evaluated at each iteration of the registration procedure. It is obtained by finding the line that best fits the joint histogram of the two images (See Figure 2). This line is obtained using linear regression and outlier rejection. From experience, we know that such an intensity correction provides images in which boundary definitions are clearer and better matched.

2.1.3 Relevance of the Resulting Transformation

In the case of inter-subject non-rigid registration, quantifying the accuracy of a method is difficult. One could deform an image I using a known deformation D into I', register I with I' and compare the result of the registration with D, but this comparison is biased by the way D is generated. For example, since the demons algorithm produces a smooth deformation field, if the vectors of D were to be generated randomly, the method is expected to perform poorly. Another method is to place landmarks in the images to register and evaluate differences between landmarks after registration (See for example [13]). No such study has been performed using the demons algorithm.

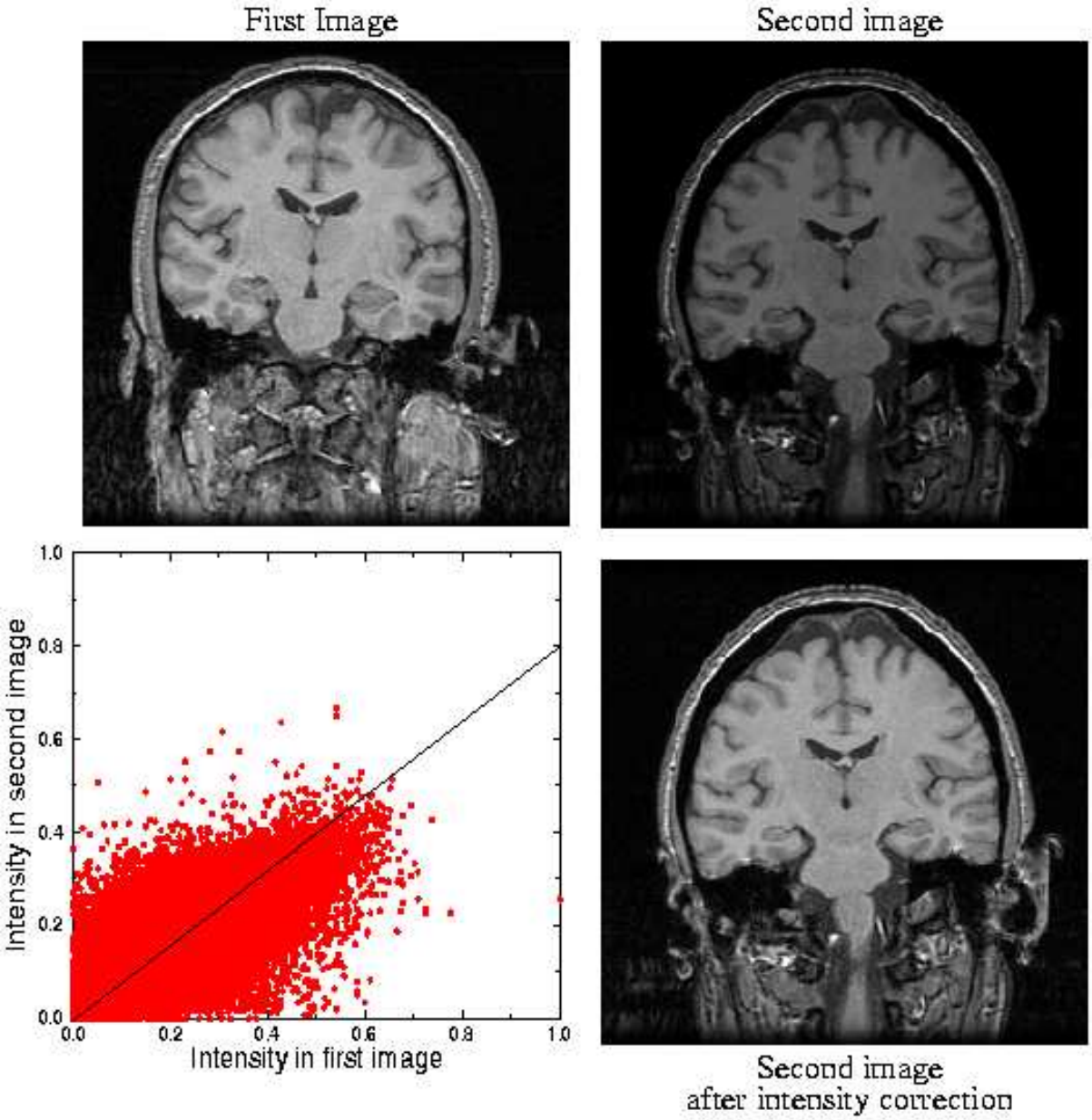


Figure 2: Intensity correction method.

An approach which is a generalization of the previous one is to compare manual and automatic segmentation using segmentation propagation, the manual result serving as ground truth (See for example [32]). We believe this approach may be better suited to evaluate the quality of high dimensional transformations such as the ones obtained using the previously described algorithm (typically $3 \times 200^3 = 24 \times 10^6$ degrees of freedom). Such a study has been perform previously by Dawant *et al.* [33] for the demons algorithm. To summarize their work, contours of different brain structures, large and small, have been segmented manually on nine MR images. One of these images was mapped to all eight other images and manual and propagated segmentations where compared. Their similarity index is defined as two times the area encircled by both contours divided by the sum of the areas encircled by each contour. This index ranges from zero to one, with zero indicating zero overlap and one indicating a perfect agreement between two contours. It is sensitive to both displacement and differences in shape and it is thus preferable to a simple area comparison. The average similarity indices between the manual and automatic segmentations was 0.96, 0.97 and 0.845 for the whole head, the cerebellum and the head of the caudate respectively. Compared with intra-rater results of 0.97, 0.97 and 0.88, the difference in the mean similarity indices between two manual delineations and between the manual delineations and the automatic segmentation method are statistically significant for the whole head and the caudate but not for the cerebellum. The authors put forth though that similarity indices over 0.85 correspond to contours that are virtually indistinguishable and that a more relevant comparison should be performed using inter-rater manual results which are known to have more discrepancies.

It should be pointed out that the demons algorithm does not explicitly track the transformation's Jacobian to make sure its determinant is positive, such that in theory it is possible to obtain a singular transformation. In our experience on MR data, this does not occur when using a sigma of 1 voxel to define the Gaussian filter for the smoothing operation. Also, the algorithm matches intensities and a *global* intensity correction is made over the whole image. Hence the transformed image is not an exact duplicate of the target. This is due to the smoothness constraint applied to the displacement field which establishes a compromise between intensity resemblance and uniform local deformations at each iteration and thus in the final result.

2.2 Average Model Construction

The average model construction needs as input a reference image I_R and a set S of N images $I_1, \ldots, I_N$ representing the group of subjects under consideration. The method can be divided in six steps as follows:

1. The first step regards the evaluation of global shape and intensity differences between the reference and each image of the set. Elastic registration between I_R and I_i provides vector fields D_i giving for each voxel x_R of I_R the analogous anatomical location x_i in I_i as well as an intensity transformation IT_i. An affine transformation A_i which best approximates, in a least squares sense, the corresponding D_i is computed. Since we have correspondences between anatomical points of the I_i and I_R that have the form $x_i = D_i(x_R)$, we compute the A_i by minimizing the distance $\sum_x ||x - A_i^{-1}(D_i(x))||^2$ (See for example [34] for a closed form), where the summation is performed on the voxel positions in I_R corresponding to cerebral tissues[1].

2. In the second step, residual variations due to pure morphological differences are evaluated. Elastic registration is performed between I_R and each I_i using the corresponding A_i and IT_i as initial transformation estimates. This provide the resulting matched images I_i' as well as the residual vector fields R_i.

3. The third step averages the I_i', producing a mean intensity image $\overline{I}$ with the shape of I_R.

4. The fourth step aims to produce the deformation presenting the shape variations between I_R and the average shape of the set elements after correction of affine differences. Since the residual deformations R_i are all defined in the same anatomical space, that of I_R, calculating their vectorwise average $\overline{R}(x) = 1/N \sum_i^N R_i(x)$ will provide the desired deformation.

5. The fifth and final step consists of applying this average residual deformation to the average intensity image to obtain an average intensity and shape image representing the anatomical average model M (see appendix A for details about the resampling process used in this step).

Considering numerical errors due the fact that automatic registration methods usually perform better when images are closer to each other, all these steps may be repeated by replacing I_R with M, thus constructing

[1] These positions are obtained using an automatic method for brain segmentation similar to that of Brummer *et al* [35]. From hereon, all summations over x are assumed to be on the voxel positions obtained using this algorithm.

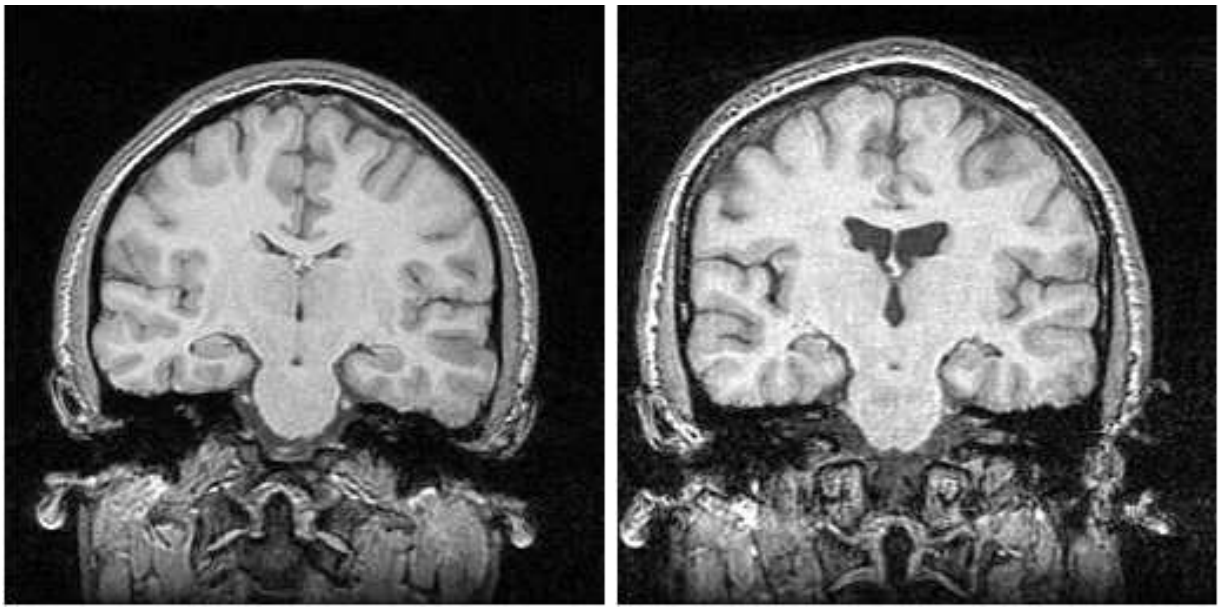
Model	Reference	Image Set
$\mathbf{M_{11}}$	I_{R_1}	S_1
$\mathbf{M_{21}}$	I_{R_2}	S_1
$\mathbf{M_{12}}$	I_{R_1}	S_2
$\mathbf{M_{22}}$	I_{R_2}	S_2

Table 1: References and image sets used to build the different models.

a model with a reference image closer to the centroid of our set. Intuitively, this should reduce the mean registration error and provide a new model M' closer to the theoretical solution.

In the next section we will further study this convergence with respect to the choice of the reference image I_R and the number of iterations needed to achieve convergence.

3 Results



(a) A coronal slice of I_{R_1}.

(b) A coronal slice of I_{R_2}.

Figure 3: coronal slices from the two reference images I_{R_1} and I_{R_2} respectively.

The method is tested by computing four models using two reference images I_{R_1} and I_{R_2} (see Figures 3(a) and 3(b)) and two image sets S_1 and S_2, each composed of five images (see Table 1).

The 3D MR protocol provides coronal images obtained using a 1.5 Tesla SIGNA (General Electric, Milwaukee, U.S.A.) whole body MR imaging system. One hundred and twenty four (124) coronal T1-weighted images were obtained using a spoiled gradient echo (SPGR) pulse sequence (TE=9 seconds, TR=34 seconds, flip angle=45^o). Two NEX acquisitions took 27 minutes and 52 seconds. The Field of View (FOV) of the images was 20 cm and each image refers to a contiguous section of tissue of 1.6 mm thickness. The two acquisitions, as opposed to one, gave increased contrast between gray and white matter, and therefore more ready definition of structure boundaries. The images showed no evidence of movement or chemical shift artifacts, and partial voluming effects were minimal. The acquisition time was well tolerated by all subjects. The $256 \times 256 \times 124$ voxels of size 0.78mm $\times$ 0.78mm $\times$ 1.6mm were trilinearly interpolated to $200 \times 200 \times 198$ to give cubic voxels of 1mm side.

We analyze our results with regards to two factors. First, the iteration process is investigated to see if convergence is achieved, and if so how fast is the convergence rate. Second, we study the effect of changing the

reference image. If the model is a veritable average of the image set, changing the reference should produce an identical model up to an affine transformation defined by the affine difference between references.

In our evaluation procedure, three metrics are used. The first determines the average distance (AD) from an image I to the elements of a set S,

$$\mathrm{AD}(I,S) = \sqrt{\frac{1}{n}\sum_x \frac{1}{N}\sum_{i=1}^{N} ||x - R_i(x)||^2},$$

where R_i is the residual deformation from I to the ith element of S, n is the number of voxels characterizing cerebral tissues and N represents the number of elements in S.

The second is the root mean square norm (RMSN) which supplies information regarding the shape variation expressed by a deformation field D,

$$\mathrm{RMSN}(D) = \sqrt{\frac{1}{n}\sum_x ||x - D(x)||^2},$$

where n is the number of voxels characterizing cerebral tissues in the reference from which D was obtained.

The third provides a measure of brightness disparity between two images I_i and I_j. It is the normalized intensity difference (NID) of the images intensities at corresponding locations,

$$\mathrm{NID}(I_i, I_j) = \sqrt{\frac{\sum_x (I_i(x) - I_j(x))^2}{\sum_x (I_i(x))^2}}.$$

An easy way to interpret this formula is to notice that if $I_j = I_i$, $\mathrm{NID}(I_i, I_j) = 0$; if $I_j = 2I_i$, $\mathrm{NID}(I_i, I_j) = 1$; if $I_j = \frac{1}{2}I_i$, $\mathrm{NID}(I_i, I_j) = 0.5$; and so on.

3.1 Effect of Iterating

To evaluate the effect of iterating, we construct the four models repeating the process five times and using the result of the previous iteration as the reference image. We will designate the model M_{jk}, obtained with reference image I_j and set S_k, computed at the ith iteration by $M_{jk}^{(i)}$. For convenience, $M_{jk}^{(0)}$ will be identified to the average intensity image having the shape of I_j. This represents a sort of iteration after applying only the three first steps described in section 2.2.

Four measures were computed:

AD$(M_{jk}^{(i)}, S_k)$ The average distance from the reference of the current iteration to all the elements of the set.

RMSN$(\overline{R}_{jk}^{(i)})$ The shape variation expressed by the residual deformation field $\overline{R}_{jk}^{(i)}$ when $M_{jk}^{(i)}$ is used as the reference.

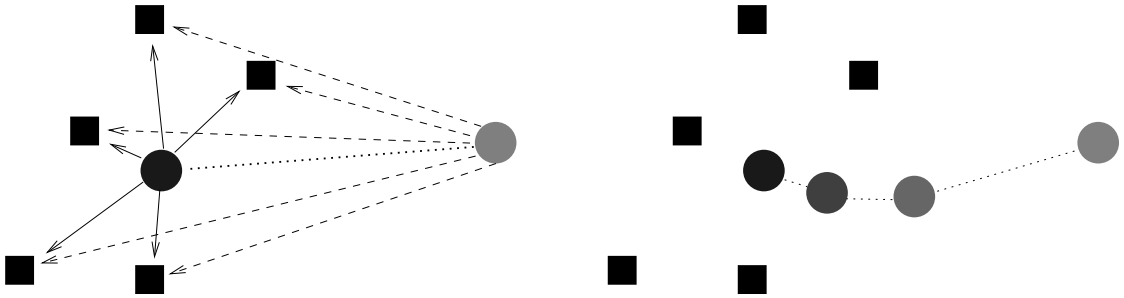
RMSN$(D_{jk}^{(i)})$ The shape difference between models computed at successive iterations. $D_{jk}^{(i)}$ is the deformation obtained by registering $M_{jk}^{(i)}$ with $M_{jk}^{(i+1)}$.

NID$(M_{jk}^{(i)}, M_{jk}^{(i+1)})$ The brightness disparity between models obtained at successive iterations.

If the models computed tend towards the centroid of the image set, the first measure should diminish. This process is depicted in Figure 4(a): as the model evolves towards the center (dotted line), the average distance to the image set elements decreases. The second and third measures, representing the shape evolution of the model (see Figure 4(b)), should tend towards zero. Finally, the fourth value should also decrease to zero since it represents the brightness differences between successive models.

The results of these calculations on the four models are presented in Figure 5. Note that the iterations range up to 4 and not 5 since we compare models computed at iterations i and $i+1$. We remind the reader that "models" $M_{jk}^{(0)}$, that is models before the first iteration, characterize only average intensities and not average shapes.

From Figure 5(a), we know the average distance from the references to the image set elements is between 4.62mm and 5.51mm and reduces to a range of 2.88mm to 3.36mm. Notice that the average distances for the models build using the same image set are very close (0.09mm for both models) while the distance for the



(a) The average distance of the model to the image set elements decreases.

(b) The distance between successive models diminishes.

Figure 4: Evolution of the model (circles) toward the center of the image set (squares).

models built using different image sets is much higher (in the range of 0.30 to 0.48mm, depending of which models are compared). This is easily explainable by the fact different small image sets will tend to have different average distance from their centroid. Compared to these values, the variation between successive models (see Figure 5(b) and 5(c)), which is in the range of 0.41mm to 0.52mm, seems minor. Figure 5(d) presents numbers showing the brightness difference between successive models diminishes rapidly to almost 0, increasing our belief that models do not evolve significantly after the first iteration.

3.2 Effect of the Reference

If the models computed are equal up to an affine transformation, changing the reference image should produce a model identical to the previous one after removing their affine differences. To verify this characteristic, we performed an *affine* registration between models built using the same image set. $M_{21}^{(i)}$ is registered with $M_{11}^{(i)}$ to provide the image $M'^{(i)}_{21}$ and $M_{22}^{(i)}$ with $M_{12}^{(i)}$ to result in $M'^{(i)}_{22}$.

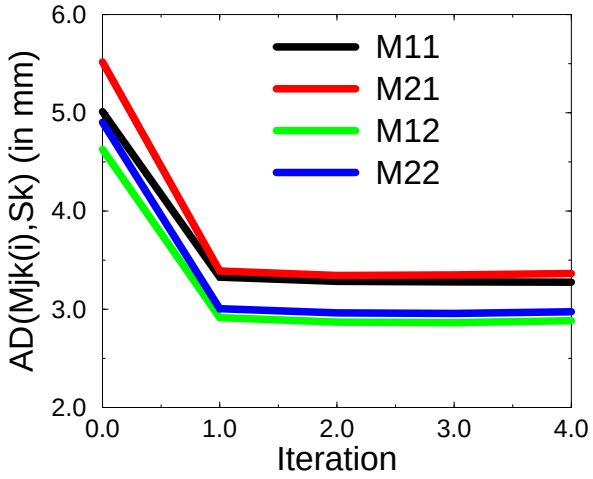
Two measure were used:

RMSN$(D_k^{(i)})$ The shape variation from $M_{1k}^{(i)}$ to $M'^{(i)}_{2k}$. $D_k^{(i)}$ is the deformation obtained by registering the two images.

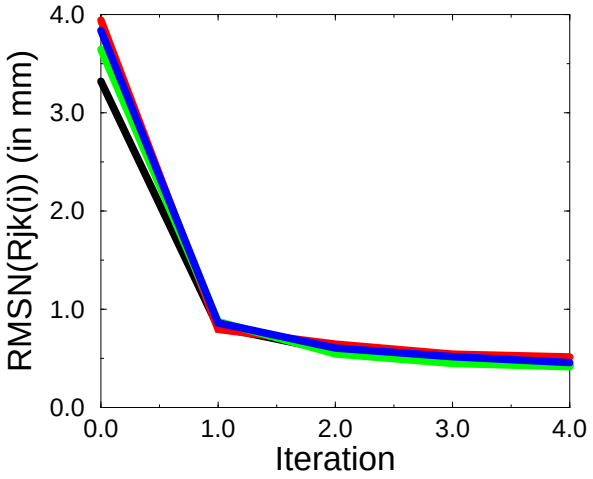
NID$(M_{1k}^{(i)}, M'^{(i)}_{2k})$ The brightness disparity between the two models.

Results are shown in Figures 6(a) and 6(b) respectively. We notice that shape variation between the models reduces from about 3.4mm to 0.9mm. This last value is close to the difference between successive models which we know from Figures 5(b) and 5(c) to be approximately 0.4mm. The brightness disparity also diminishes rapidly and does not change drastically after the first iteration. From these results, one can see that the models build using different reference image are very close but not identical. We believe the differences are due to errors in our registration method and to the resampling procedure applied to the models to put them in the same affine space, which tends to smooth the image. This last artifact is not present when comparing models obtained from successive iterations as we did for the results of Figure 5(d). Still, the average distance between models is less than the resolution of the image, and both the average distance and the normalized intensity difference evolve only slightly after the first iteration.

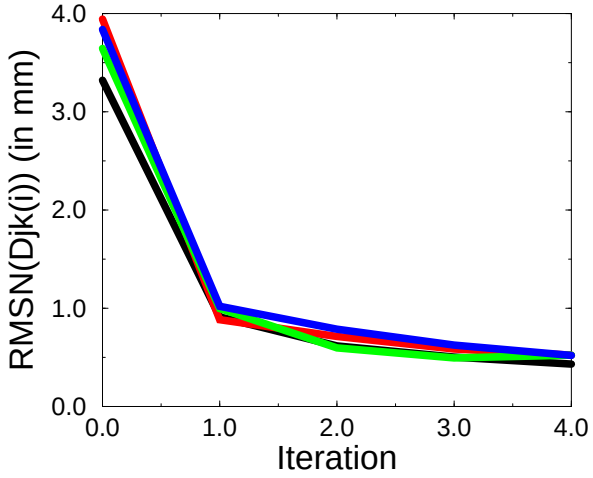
Figures 7 to 10 present some results obtained during this process. In Figure 7 the average intensity images corresponding to the reference images of Figure 3 are presented. They basically have the same shape as their respective reference image and their intensities corresponds to the average of the image set S_1. Notice how the signal to noise ratio is increased while contrast is preserved. These average intensity images are deformed using the corresponding average residual deformations to provide the average models of Figure 8. The average model $M_{21}^{(1)}$ is then registered with $M_{11}^{(1)}$ using an affine transformation to set it in the same affine space. This result is presented in Figures 9(a) and 9(b). The same procedure has been performed for all five iterations for both image sets S_1 and S_2. The resulting models for the fifth iteration using image set S_1 is also shown in Figures 9(c) and 9(d). In Figure 10, slices are taken where there is more variability in the cortex area and thus where our registration method finds correspondences in which we have less confidence. As can be seen, in these regions the model image is less clear and the contrast between gray and white matter is less pronounced.



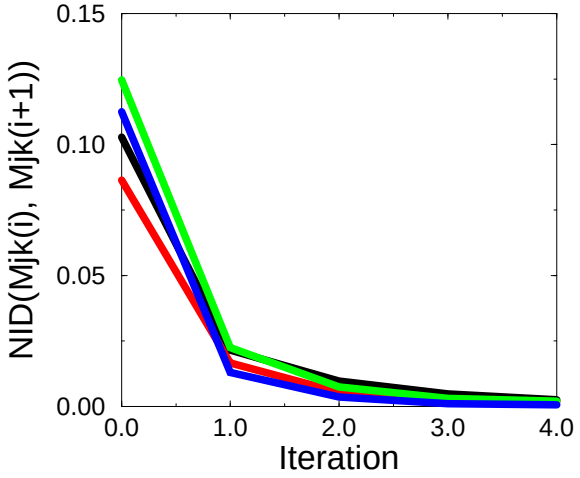
(a) Average distance to the reference of the current iteration.



(b) Shape variation of the reference for the current iteration.

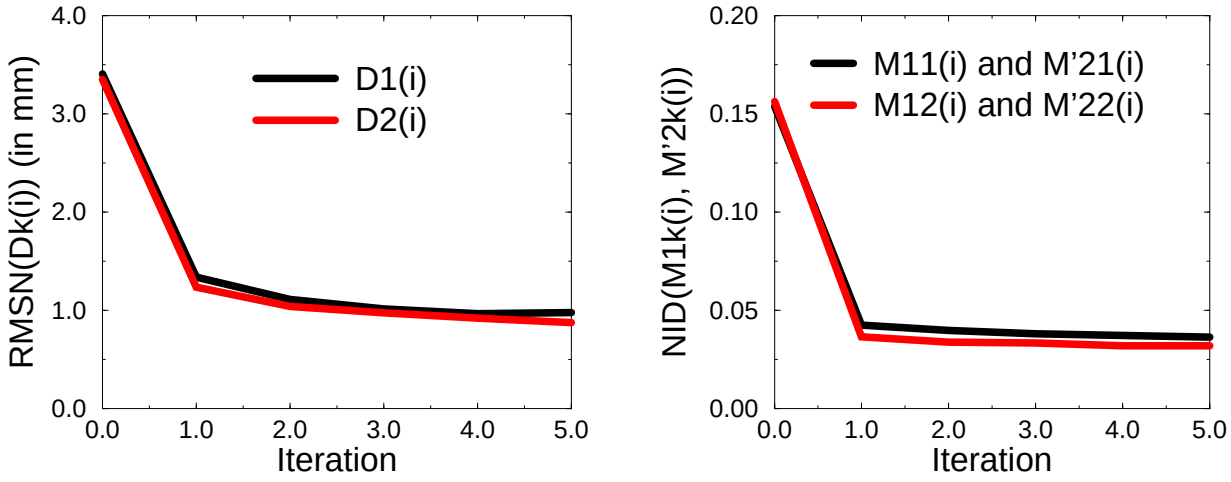


(c) Shape difference between models computed at successive iterations.



(d) Brightness disparity between models computed at successive iterations.

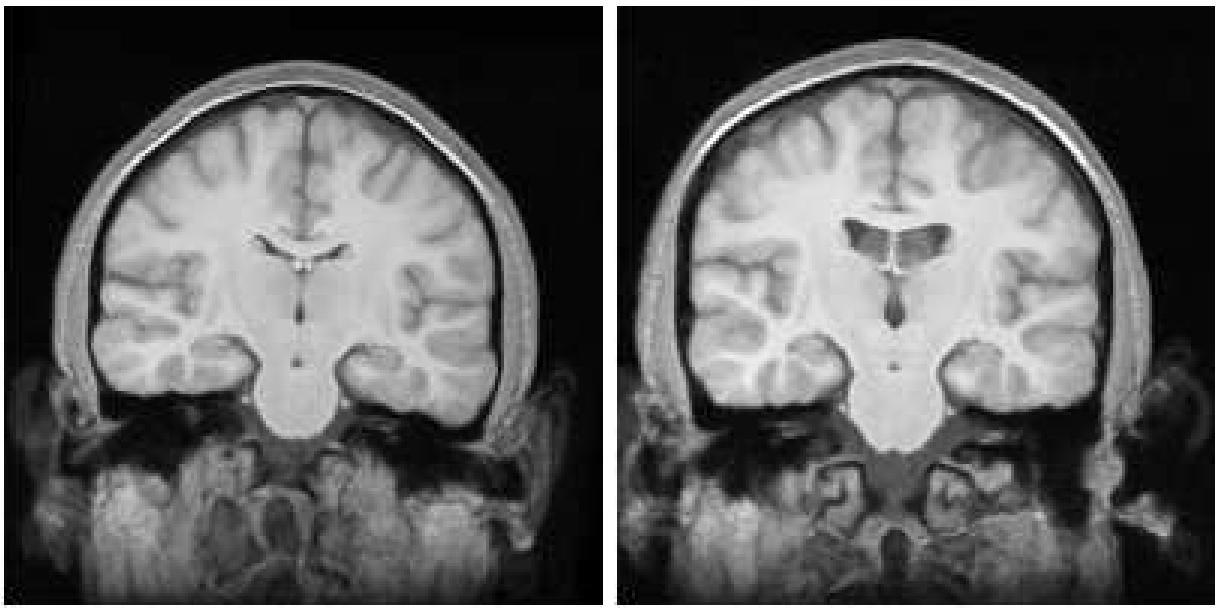
Figure 5: Impact of the iteration process when computing the models. Note that the iterations range up to 4 and not 5 since we compare models computed at iterations i and $i+1$. We remind the reader that "models" $M_{jk}^{(0)}$, that is models before the first iteration, characterize only average intensities and not average shapes.



(a) Shape variation between $M_{1k}^{(i)}$ and $M'^{(i)}_{2k}$.

(b) Brightness disparity between $M_{1k}^{(i)}$ and $M'^{(i)}_{2k}$.

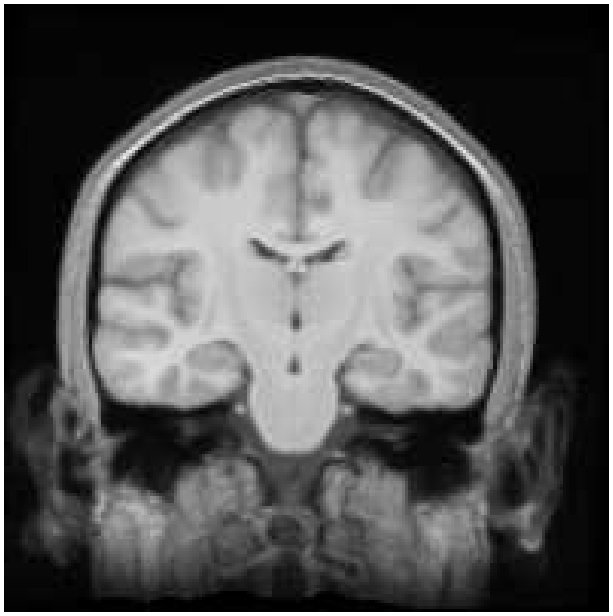
Figure 6: Influence of the reference on the model computed.



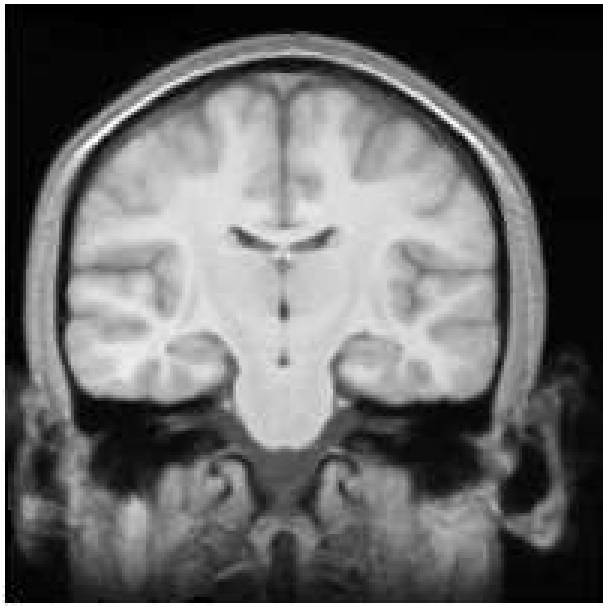
(a) Average intensity image $\overline{I}$ build using I_{R_1} and S_1 during first iteration.

(b) Average intensity image $\overline{I}$ build using I_{R_2} and S_1 during first iteration.

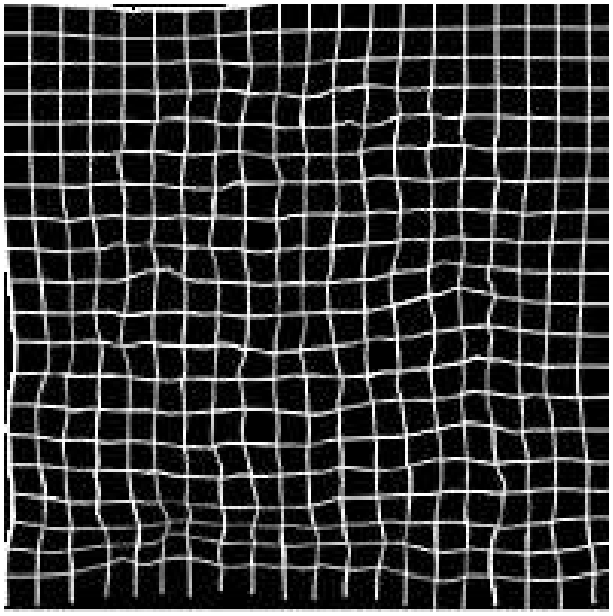
Figure 7: Average intensity images build using reference images I_{R_1} (Figure 3(a)) and I_{R_2} (Figure 3(b)) and the same image set S_1. They were obtained during the first iteration of the model construction method. Notice how the signal to noise ratio is increased while contrast is preserved.



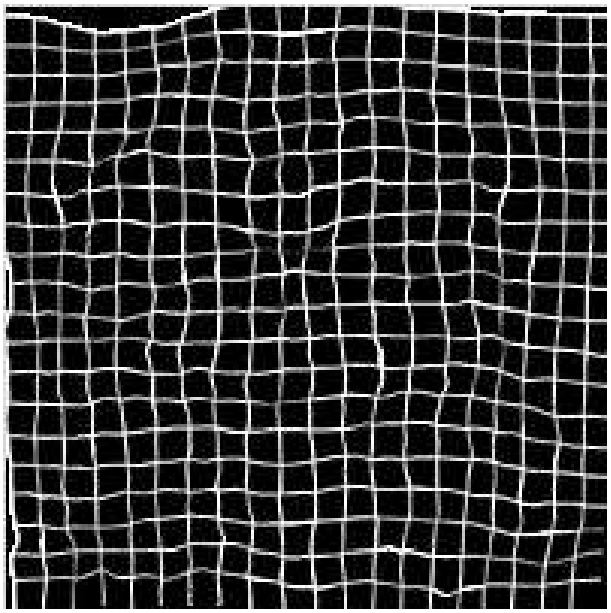
(a) A coronal slice of $M_{11}^{(1)}$.



(b) A coronal slice of $M_{21}^{(1)}$.

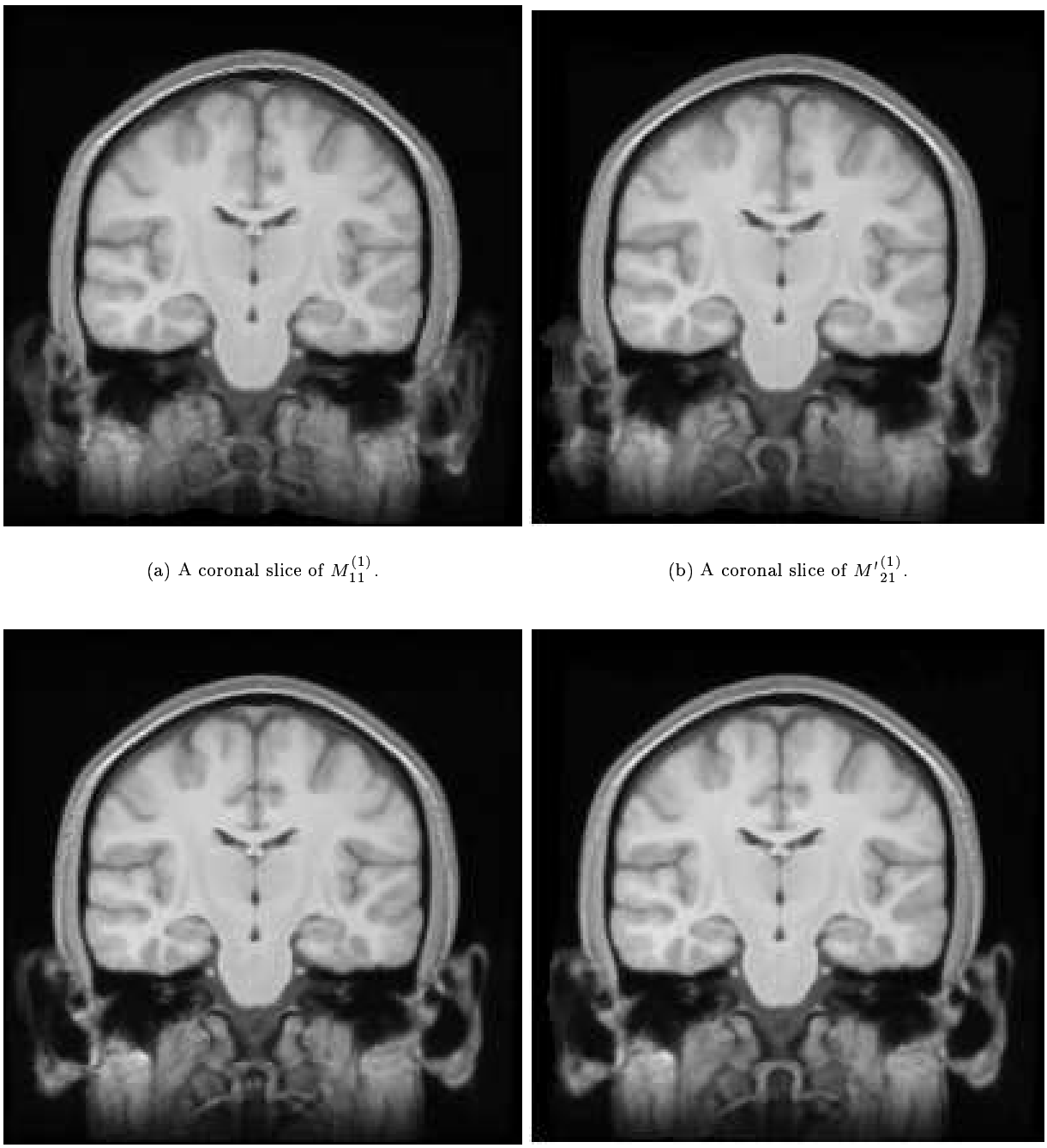


(c) A regular grid deformed by the average residual deformation $\overline{R}_{11}^{(0)}$.



(d) A regular grid deformed by the average residual deformation $\overline{R}_{21}^{(0)}$.

Figure 8: Average models $M_{11}^{(1)}$ (Figure 8(a)) and $M_{21}^{(1)}$ (Figure 8(b)) computed using reference images I_{R_1} (Figure 3(a)) and I_{R_2} (Figure 3(b)) and the same image set S_1. They were build from the average intensity images of Figure 7(a) and Figure 7(b) respectively. Figures 8(c) and 8(d) are the corresponding deformed grids.



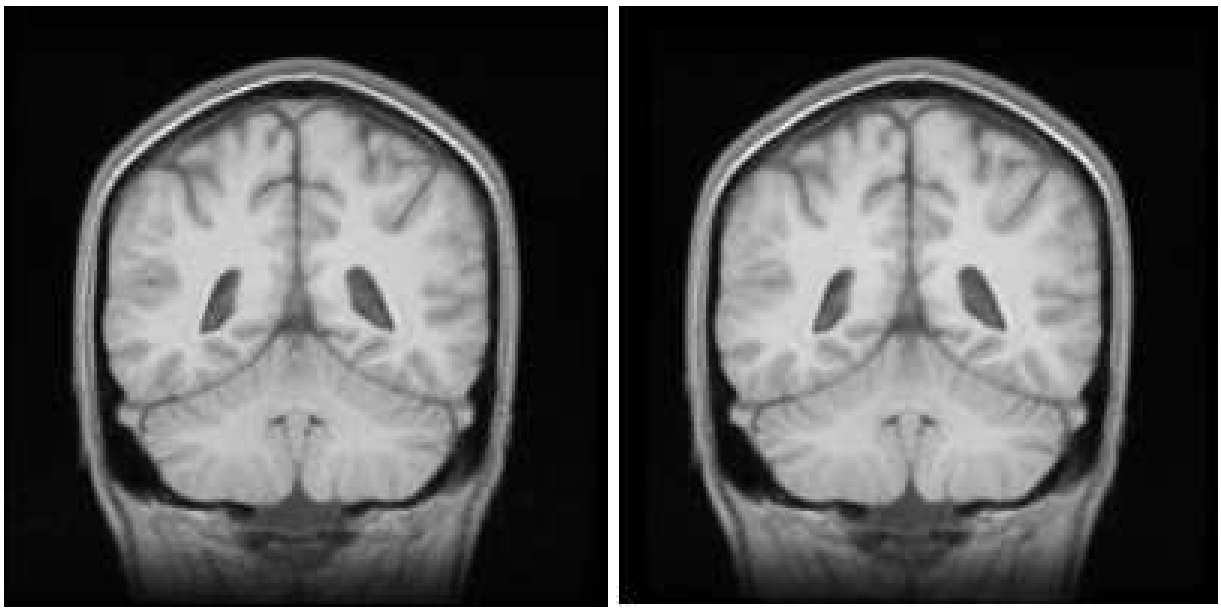
(a) A coronal slice of $M_{11}^{(1)}$.

(b) A coronal slice of $M'^{(1)}_{21}$.

(c) A coronal slice of $M_{11}^{(5)}$.

(d) A coronal slice of $M'^{(5)}_{21}$.

Figure 9: Average models computed using reference images I_{R_1} (Figure 3(a)) and I_{R_2} (Figure 3(b)) and the same image set S_1. Figures 9(b) and 9(d) have been transformed into the same affine space as Figures 9(a) and 9(c) respectively, and can therefore be compared directly.



(a) A coronal slice of $M_{11}^{(5)}$.

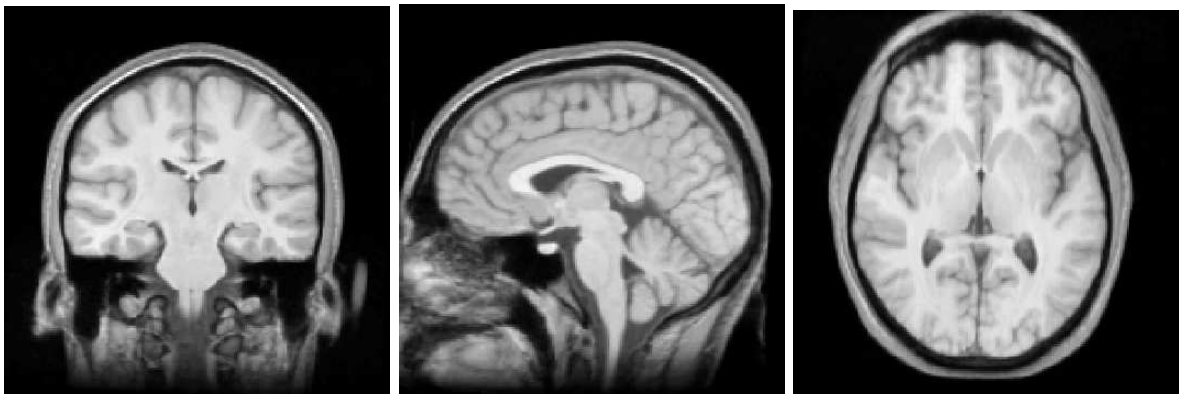
(b) A coronal slice of $M'^{(5)}_{21}$.

Figure 10: Average models computed using reference images I_{R_1} (Figure 3(a)) and I_{R_2} (Figure 3(b)) and the same image set S_1. Figure 10(b) has been transformed into the same affine space as Figure 9(c) and can therefore be compared directly. Compared with Figure 9, these slices are taken where there is more variability in the cortex area and thus where our registration method finds correspondences in which we have less confidence. As can be seen, in these regions the model image is less clear and the contrast between gray and white matter is less pronounced.

4 Discussion

Figure 5 presents numbers showing that our method constructs average models well representing the average intensity and shape of our image sets. In particular, Figure 5(a) shows that the average distance from one image to the set elements varies between 4.62mm and 5.51mm. This distance reduces and stays between 2.88mm to 3.36mm after the first iteration. Figure 5(b) and 5(c) illustrate a minor shape evolution of the models at each iteration. Furthermore, Figure 5(d) allows us to claim the visual aspect of the models changes only minimally. This leads us to the conclusion that models constructed are different, but equivalent from a practical point of view. That is, we believe the model slowly wanders around the optimum solution. Their intensity difference is practically null, and their shapes, although different, all have the same average distance to the other elements of the set. Hence, we believe one or two iterations are sufficient to build representative average models.

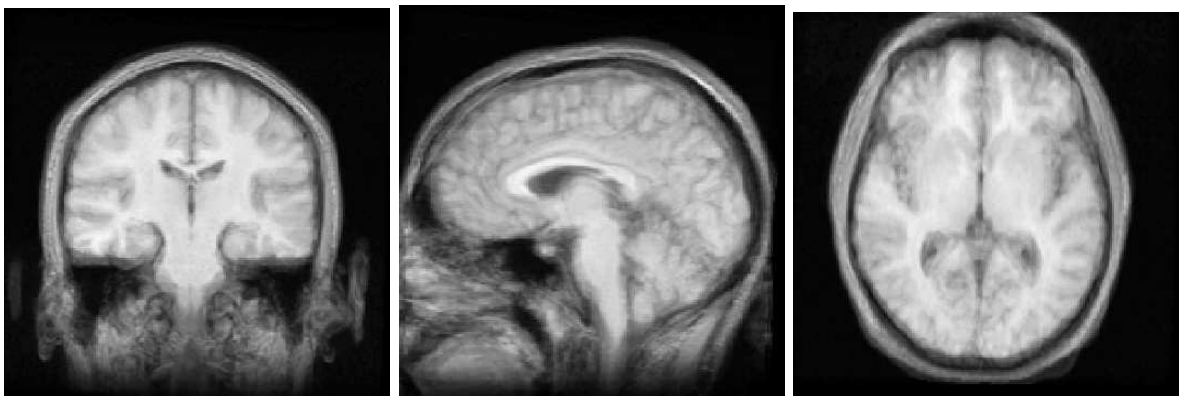
The reader should appreciate the low noise, high contrast and high quality of the models produced in Figures 8, 9 and 10. They present models built using different reference images and the same image set. The ventricular shape bias introduced using I_{R_2} is minimal if not null. Also, Figure 6 shows that the models built using different references seem to converge towards the same solution. Their shape difference presented in Figure 6(a) of about 0.9mm is low compared to the average distance between the models and the set elements, which is in the range of 2.88mm to 3.36mm, and just over the distance between successive average models which varies from 0.30 to 0.48mm. Figure 6(b) also presents a low disparity between the different models intensities.



(a) Coronal slice, elastic registration.

(b) Sagittal slice, elastic registration.

(c) Axial slice, elastic registration.



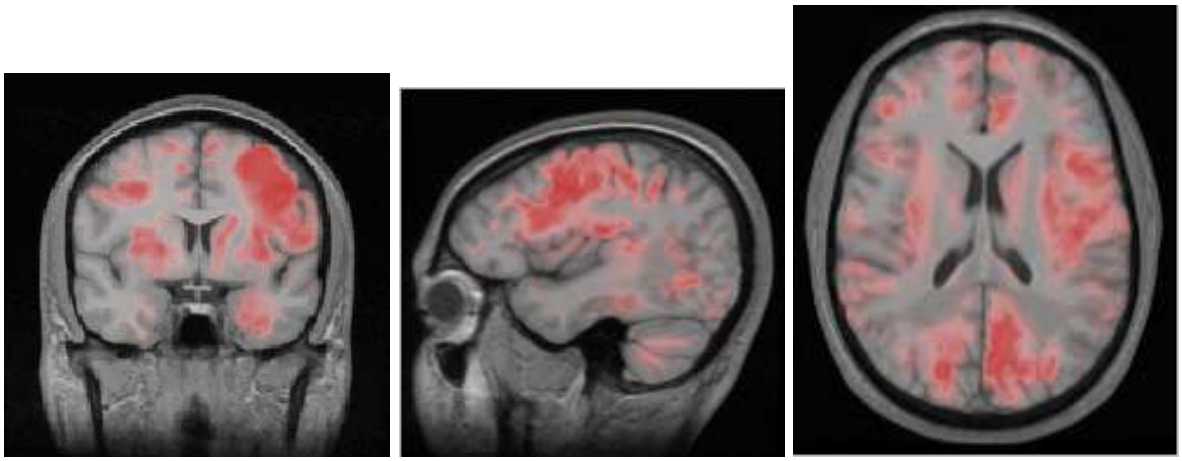
(d) Coronal slice, affine registration.

(e) Sagittal slice, affine registration.

(f) Axial slice, affine registration.

Figure 11: Corresponding slices of M_{11} using affine registration (Figures 11(d), 11(e) and 11(f)) instead of elastic registration (Figures 11(a), 11(b) and 11(c)).

If familiar with the work of Bookstein [20] or the MNI group [8], the reader will appreciate the high contrast and visual quality of the images produced, although due in part to the smaller number of images used. To better visually appreciate the gain in using high-dimensional volumetric maps ($200^2 \times 198 \times 3 = 2.376 \times 10^7$ degrees of freedom) instead of affine transformations (12 degrees of freedom) during registration, Figure 11 presents corresponding slices of the model M_{11} build using the method presented in this paper and the one build restraining registration to affine transformations. Note that Figures 11(a), 11(b) and 11(c), obtained using the method described above, provide higher contrasts and more sharp definitions of tissue boundaries than Figures 11(d), 11(e) and 11(f) which where obtained using affine registration only.



(a) Coronal section.

(b) Sagittal section in left hemisphere.

(c) Axial section.

Figure 12: Coronal, sagittal and axial sections of the probabilistic atlas. Shape variability is superimposed in red (see text for more details). Images are shown using the radiological convention, i.e. patient's left side is on the right side and vice-versa.

In the event that bad matches occur during registration, the model would reflect this by having larger variances in deformation and/or intensity. The image obtained and the corresponding statistics regarding deformations and intensities would not reflect normal variations, but to a certain extent reflect normal variations *within our registration framework*. It is our belief that the models obtained in this paper are in a great deal linked to our modeling of the registration problem. Though cross-validation of registration techniques results reports mutually coherent matches [36], average variations of 3mm to 4mm have been observed in this study. These figures are too large to give significance to our models' statistics using other registration methods.

Although beyond the scope of this article, we present in Figure 12 preliminary results of a probabilistic atlas built using the information gathered during the construction of M_{11}. This figure presents slices of the average model on which information about shape variability is superimposed in red. This variability is computed as follows: we calculate the voxelwise covariance matrix of the residual deformation fields R_i obtained from the registration of each element of S_1 with M_{11}. The determinant, or the product of the eigenvalues, of each of these matrices is computed. The resulting values can be interpreted as probabilistic volumes (See [37] for more detailed geometric interpretations of this measure). Hence, for a given voxel, a large volume corresponds to a large variability of the positions of corresponding voxels found in S_1. This volume has been normalized and converted to a red intensity value added to the MR image.

We remind the reader that this probabilistic atlas was built using only five normal subjects, and hence does not reflect the extent of normal variability that is to be found in a larger population. Nonetheless, we wish to point out that high variability was found in the cortical region, which is known to present large shape variations. There also seems to be more variability in the left hemisphere than in the right one.

We are presently working on the evaluation of the number of subjects needed to fully represent the extent of variations in a normal population. Further interpretations of the corresponding probabilistic atlas will follow this work.

5 Conclusion

We have presented a completely automatic method to build average anatomical models of the human brain using a set of MR images. To this end, brain shape variations between subjects were identified. Differences due to linear transformations were excluded, resulting in the quantification of pure morphological differences. The result is an average intensity and shape image representative of the characteristics of the image set elements used for the construction. Furthermore, we have shown that this model corresponds to the centroid of the image set and does not depend on the reference image used for its construction. This stability is very important to ensure the efficiency and usefulness of our method. The coupling of such a high quality model with statistical information regarding normal deformations, such as the one presented in Figure 12 or in the work of Thompson & Toga [18] or the Washington/Brown group [14], could enrich the significance of statistical tests by adding intensity information, useful for example in detecting gliosis in T2 MR images, and would supply an important tool in the analysis of normal anatomy.

6 Acknowledgments

Dr. Neil Roberts from the Magnetic Resonance & Image Analysis Research Center (MARIARC) of Liverpool University provided the MR images used in this study. Part of this work was funded by the Natural Sciences and Engineering and Research Council of Canada (NSERC) and the Ministère de l'Éducation du Québec (MEQ).

A Forward Resampling

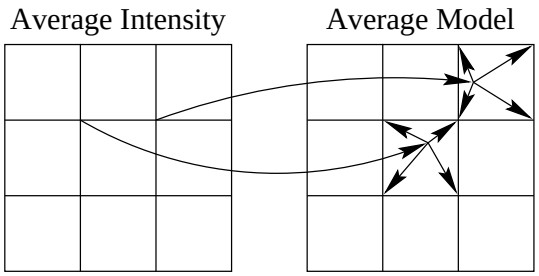


Figure 13: Forward resampling with bilinear distribution for two voxels.

Because of the way the average deformation is computed, we only have correspondences going from the average intensity to the average model, $M(\overline{R}(x)) = \overline{I}(x)$. To be able to resample the average intensity using a customary interpolation method, the inverse deformation $\overline{R}^{-1}$ is required in order to have a relation of the form $M(x) = \overline{I}(\overline{R}^{-1}(x))$. Instead of tackling the problem of inverting a vector field or of finding explicitly bijective deformation functions, we chose to use a different kind of resampling strategy. For this purpose, "forward resampling" involving trilinear distribution was used (see Figure 13). The intensity of each voxel of the average intensity image is distributed to the neighbors of its location in the average model according to the same weights found for trilinear interpolation. As in linear interpolation where the sum of weights sums to unity, each voxel of the average model needs to be normalized according to the sum of the weights.

This kind of resampling may produce holes which can be filled using gray-scale image mathematical morphology [38], a process intended for image reconstruction. This "hole" effect did not occur in our experiments and mathematical morphology was not used.

References

[1] J. C. Mazziotta, A. W. Toga, A. Evans, P. Fox, and J. Lancaster. A probabilistic atlas of the human brain: Theory and rationale for its development. *Neuroimage*, 2:89–101, 1995.

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

[3] G. Schaltenbrand and W. Wahren. *Atlas of Stereotaxy of the Human Brain.* Georg Thieme Verlag, Stuttgart, 1977.

[4] R. Bajcsy and S. Kovacic. Multiresolution elastic matching. *Computer Vision, Graphics and Image Processing*, 46(1):1–21, April 1989.

[5] T. Greitz, C. Bohm, S. Holte, and L. Eriksson. A computerized brain atlas: Construction, anatomical content, and some applications. *Journal of Computer Assisted Tomography*, 15(1):26–38, January 1991.

[6] D. Lemoine, C. Barillot, B. Gibaud, and E. Pasqualini. An anatomical-based 3D registration system of multimodality and atlas data in neurosurgery. *Lecture Notes in Computer Science*, 511:154–164, 1991.

[7] K. H. Höhne, M. Bomans, M. Riemer, R. Schubert, U. Tiede, and W. Lierse. A volume-based anatomical atlas. *IEEE Computer Graphics and Applications*, pages 72–78, July 1992.

[8] A. C. Evans, M. Kamber, D. L. Collins, and D. Macdonald. An MRI-based probabilistic atlas of neuroanatomy. In S. Shorvon, D. Fish, F. Andermann, G. M. Bydder, and H. Stefan, editors, *Magnetic Resonance Scanning and Epilepsy*, volume 264 of *NATO ASI Series A, Life Sciences*, pages 263–274. Plenum Press, 1994.

[9] F. L. Bookstein. *Functional Neuroimaging*, chapter 10 - Landmarks, Edges, Morphometrics, and the Brain Atlas Problem. Academic Press, 1994.

[10] G. Christensen, M. I. Miller, and M. W. Vannier. A 3D deformable magnetic resonance textbook based on elasticity. In *Spring Symposium: Applications of Computer Vision in Medical Image Processing*, Stanford, United-States, March 1994. American Association for Artificial Intelligence.

[11] R. Kikinis, M. E. Shenton, D. V. Iosifescu, R. W. McCarley, P. Saiviroonporn, H. H. Hokama, A. Robatino, D. Metcal, C. G. Wible, C. M. Portas, R. M. Donnino, and F. A. Jolesz. A digital brain atlas for surgical planning, model-driven segmentation, and teaching. *IEEE Transactions on Visualization and Computer Graphics*, 2(3):232–241, September 1996.

[12] L. Le Briquer and J. C. Gee. Design of a statistical model of brain shape. In J. S. Duncan and G. R. Gindi, editors, *Proceedings of the Information Processing in Medical Imaging Conference (IPMI'97)*, Vermont, United-States, 1997. Springer-Verlag.

[13] R. P. Woods, S. T. Grafton, J. D. G. Watson, N. L. Sicotte, and J. C. Mazziotta. Automated image registration: Ii. intersubject validation of linear and nonlinear models. *Journal of Computer Assisted Tomography*, 22(1):153–165, 1998.

[14] U. Grenander and M. I. Miller. Computational anatomy: An emerging discipline. *Quarterly of Applied Mathematics*, 56(4), 1998 1998.

[15] P. M. Thompson, D MacDonald, M. S. Mega, C. J. Holmes, A. C. Evans, and A. W. Toga. Detection and mapping of abnormal brain structure with a probabilistic atlas of cortical surfaces. *Journal of Computer Assisted Tomography*, 21(4):567–581, 1998.

[16] G. Subsol, J.-P. Thirion, and N. Ayache. A scheme for automatically building three-dimensional morphometric anatomical atlases: Application to a skull atlas. *Medical Image Analysis*, 2(1):37–60, 1998.

[17] J. C. Gee, D. R. Haynor, L. Le Briquer, and R. K. Bajcsy. Advances in elastic matching theory and its implementation. In P. Cinquin, R. Kikinis, and S. Lavallee, editors, *Conference on Computer Vision, Virtual Reality and Robotics in Medecine & Medical Robotics and Computer Aided Surgery (CVRMed-MRCAS'97)*, Heidelberg, 1997. Springer-Verlag.

[18] P. M. Thompson and A. W. Toga. Detection, visualization and animation of abnormal anatomic structure with a deformable probabilistic brain atlas based on random vector field transformations. *Medical Image Analysis*, 1(4):271–294, September 1997.

[19] A. Guimond, G. Subsol, and J.-P. Thirion. Automatic MRI database exploration and applications. *International Journal of Pattern Recognition and Artificial Intelligence*, 11(8):1345–1365, December 1997.

[20] F. L. Bookstein. Thin-plate splines and the atlas problem for biomedical images. In A. C. F. Colchester and C. J. Hawkes, editors, *Proceedings of the Information Processing in Medical Imaging Conference*, volume 511 of *Lecture Notes in Computer Science*, pages 326–342, Wye, United Kingdom, July 1991.

[21] F. L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, June 1989.

[22] M. I. Miller et al. Statistical methods in computational anatomy. *Statistical Methods in Medical Research*, 6, 1997.

[23] F. L. Bookstein. Shape and the information in medical images: A decade of the morphometric synthesis. *Computer Vision and Image Understanding*, 66(2):97–118, may 1997.

[24] D. G. Kendall. A survey of the statistical theory of shape. *Statistical Science*, 4(2):87–120, 1989.

[25] G. Christensen, R. D. Rabbitt, and M. I. Miller. 3D brain mapping using a deformable neuroanatomy. *Physics in Medecine and Biology*, 39:608–618, March 1994.

[26] D. L. Collins, T. M. Peters, and A. C. Evans. An automated 3D non-linear image deformation procedure for determination of gross morphometric variability in human brain. In R. A. Robb, editor, *Visualisation in Biomedical Computing*, volume 2359 of *SPIE Proceedings*, pages 180–190, Rochester, October 1994.

[27] P. Thompson and A. W. Toga. A surface-based technique for warping 3-dimensional images of the brain. *IEEE Transactions in Medical Imaging*, 15:1–16, 1996.

[28] J.-P. Thirion. Image matching as a diffusion process: an analogy with Maxwell's demons. *Medical Image Analysis*, 2(3):243–260, 1998.

[29] B. K. P. Horn and B. G. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, August 1981.

[30] J. L. Barron, D. J. Fleet, and S. S. Beauchemin. Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1):43–77, January 1994.

[31] M. Bro-Nielsen and C. Gramkow. Fast fluid registration of medical images. In K. H. Höhne and R. Kikinis, editors, *Visualisation in Biomedical Computing*, volume 1131 of *Lecture Notes in Computer Science*, pages 267–276, Hamburg, Germany, September 1996. Springer-Verlag.

[32] J. W. Haller, A. Banerjee, G. E. Christensen, M. Gado, S. C. Joshi, , M. I. Miller, Y. Sherline, M. W. Vannier, and J. G. Csernansky. Three-dimensional hippocampal MR morphometry with high-dimensional transformation of a neuroanatomic atlas. *Radiology*, 202:504–510, 1997.

[33] B. M. Dawant, J.-P. Thirion, F. Maes, D. Vandermeulen, and P. Demaerel. Automatic 3D segmentation of internal structures of the head in MR images using a combination of similarity and free form transformations. In K. M. Hanson, editor, *Medical Imaging 1998: Image Processing*, volume 3338 of *SPIE Proceedings*, pages 545–554, San Diego, United-States, February 1998. Internation Society for Optical Engineering.

[34] O. Faugeras. *Three-Dimensional Computer Vision: A Geometric Viewpoint.* MIT Press, Cambridge, Massachusetts, 1993.

[35] M. E. Brummer, R. M. Mersereau, R. L. Eisner, and R. R. J. Lewine. Automatic detection of brain contours in MRI data sets. *IEEE Transactions in Medical Imaging*, 12(2):153–166, June 1993.

[36] J.-P. Thirion, G. Subsol, and D. Dean. Cross validation of three inter-parients matching methods. In K. H. Höhne and R. Kikinis, editors, *Visualisation in Biomedical Computing (VBC'96)*, volume 1131 of *Lecture Notes in Computer Science*, pages 327–336, Hamburg, Germany, September 1996. Springer-Verlag.

[37] K. Fukunaga. *Introduction to Statistical Pattern Recognition*, chapter 2 – Random Vectors and Their Properties, pages 11–50. Computer science and scientific computing. Academic Press, London, England, 2nd edition, 1990.

[38] R. C. Gonzalez and R. E. Woods. *Digital Image Processing.* Addison-Wesley, 1992.

Contents

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - B.P. 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Lorraine : Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - B.P. 101 - 54602 Villers lès Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot St Martin (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, B.P. 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399