Multimodal Elastic Matching of Brain Images

Alexis Roche[1], Alexandre Guimond[1,2], Nicholas Ayache[1], and Jean Meunier[2]

[1] INRIA Sophia Antipolis - Epidaure Project
2004 Route des Lucioles BP 93
06902 Sophia Antipolis Cedex, France
[2] Département d'Informatique et recherche opérationnelle,
Université de Montréal, CP 6128 succ Centre-Ville,
Montréal QC H3C 3J7, Canada

Abstract. This paper presents an original method for three-dimensional elastic registration of multimodal images. We propose to make use of a scheme that iterates between correcting for intensity differences between images and performing standard monomodal registration. The core of our contribution resides in providing a method that finds the transformation that maps the intensities of one image to those of another. It makes the assumption that there are at most two functional dependences between the intensities of structures present in the images to register, and relies on robust estimation techniques to evaluate these functions. We provide results showing successful registration between several imaging modalities involving segmentations, T1 magnetic resonance (MR), T2 MR, proton density (PD) MR and computed tomography (CT).
keywords: Multimodality, Elastic registration, Intensity correction, Robust estimation, Medical imaging.

1 Introduction

Over the last decade, automatic registration techniques of medical images of the head have been developed following two main trends: 1) registration of multimodal images using low degree transformations (rigid or affine), and 2) registration of monomodal images using high-dimensional volumetric maps (elastic or fluid deformations). The first category mainly addresses the fusion of complementary information obtained from different imaging modalities. The second category's predominant purpose is the evaluation of either the anatomical evolution process present in a particular subject or of anatomical variations between different subjects.

These two trends have evolved separately mainly because the combined problem of identifying complex intensity correspondences along with a high-dimensional geometrical transformation defines a search space arduous to traverse. Recently, three groups have imposed different constraints on the search space, enabling them to develop automatic multimodal non-affine registration techniques. All three methods make use of block matching techniques to evaluate local translations. Two of them use mutual information (MI) [30,17] as a similarity measure and the other employs the correlation ratio [23].

D. Vernon (Ed.): ECCV 2000, LNCS 1843, pp. 511–527, 2000.

An important aspect when using MI as a registration measure is to compute the conditional probabilities of one image's intensities with respect to those of the other. To do so, Maintz *et al.* [18] proposed to use conditional probabilities after rigid matching of the images as an estimate of the real conditional probabilities after local transformations. Hence, the probabilities are evaluated only once before fluid registration. However, Gaens *et al.* [11] argued that the assumption that probabilities computed after affine registration are good approximations of the same probabilities after fluid matching, is unsuitable. They also proposed a method in which local displacements are found so that the global MI increases at each iteration, permitting incremental changes of the probabilities during registration. Their method necessitates the computation of conditional probabilities over the whole image for every voxel displacement. To alleviate themselves from such computations owing to the fact that MI requires many samples to estimate probabilities, Lau *et al.* [16] have chosen a different similarity measure. Due to the robustness of the correlation ratio with regards to sparse data [23], they employed it to assess the similarity of neighbouring blocks. Hence no global computation is required when moving subregions of the image.

Our method distinguishes itself by looking at the problem from a different angle. In the last years, our group has had some success with monomodal image registration using the demons method [27,28], an optical flow variant when dealing with monomodal volumetric images. If we were able to model the imaging processes that created the images to register, and assuming these processes are invertible, one could transform one of the images so that they are both represented in the same modality. Then, we could use our monomodal registration algorithm to register them. We have thus developed a completely automatic method to transform the different structures intensities in one image so that they match the intensities of the corresponding structures in another image, and this without resorting to any segmentation method.

The rational behind our formulation is that there is a functional relationship between the intensity of a majority of structures when imaged with different modalities. This assumption is partly justified by the fact that the Woods criterion [31] as well as the correlation ratio [23], which evaluate a functional dependence between the intensities of the images to match, have been used with success in the past, and sometimes lead to better results than MI, which assumes a more general relation [22,21].

The idea of estimating an intensity transformation during registration is not new in itself. For example, Feldmar *et al.* [10] as well as Barber [1] have both published methods in which intensity corrections are proposed. These methods restrict themselves to affine intensity corrections in a monomodal registration context. We propose here a procedure based on one or two higher degree polynomials found using a robust regression technique to enable the registration of images from different modalities.

The remaining sections of this paper are organized in the following manner. First, we detail our multimodal elastic registration method. We then describe what kind of images were used to test the method and how they were acqui-

red. Next, results obtained by registering different images obtained from several modalities are presented and discussed. We conclude this paper with a brief discussion on future research tracks.

2 Method

Our registration algorithm is iterative and each iteration consists of two parts. The first one transforms the intensities of anatomical structures of a source image S so that they match the corresponding structures intensities of a target image T. The second part regards the registration of S (after intensity transformation) with T using an elastic registration algorithm.

In the following, we first describe the three-dimensional geometrical transformation computation and then the intensity transformation computation. We believe this ordering is more convenient since it is easier to see what result must provide the intensity transformation once the geometrical transformation procedure is clarified.

2.1 Geometrical Transformation

Many methods have been developed to deform one brain so its shape matches that of another [29]. The one used in the present work is an adaptation of the demons algorithm [27,28]. Adjustments were performed based on empirical observations as well as on theoretical grounds which are discussed below. For each voxel with position x in T, we hope to find the displacement $v(x)$ so that x matches its corresponding anatomical location in S. In our algorithm, the displacements are computed using the following iterative scheme,

$$v_{n+1}(x) = G_\sigma \otimes \left(v_n + \frac{S \circ h_n(x) - T(x)}{||(\nabla S \circ h_n)(x)||^2 + [S \circ h_n(x) - T(x)]^2} (\nabla S \circ h_n)(x) \right), \tag{1}$$

where G_σ is a Gaussian kernel, $\otimes$ denotes the three-dimensional convolution, $\circ$ denotes the composition and the transformation $h(x)$ is related to the displacement by $h(x) = x + v(x)$. As is common with registration methods, we also make use of multilevel techniques to accelerate convergence. Details about the number of levels and iterations as well as filter implementation issues are addressed in Section 4. We here show how our method can be related to other registration methods, notably the minimization of the sum of squared difference (SSD) criterion, optical flow and the demons algorithm.

Relation with SSD Minimization In this framework, we find the transformation h that minimizes the sum of squared differences between the transformed source image and the target image. The SSD between the two images for a given transformation h applied to the source is defined as

$$SSD(h) = \frac{1}{2} \sum_{x=1}^{N} [S \circ h(x) - T(x)]^2. \tag{2}$$

The minimization of Equation (2) may be performed using a gradient descent algorithm. Thus, differentiating the above equation we get

$$\nabla SSD(h) = -[S \circ h(x) - T(x)](\nabla S \circ h)(x).$$

The iterative scheme is then of the form,

$$h_{n+1} = h_n + \alpha[S \circ h_n(x) - T(x)](\nabla S \circ h_n)(x),$$

where α is the step length. This last equation implies,

$$v_{n+1} = v_n + \alpha[S \circ h_n(x) - T(x)](\nabla S \circ h_n)(x). \tag{3}$$

If we set α to a constant value, this method corresponds to a steepest gradient descent. By comparing Equation (3) to Equation (1), one sees that our method sets

$$\alpha = \frac{1}{||(\nabla S \circ h_n)(x)||^2 + [T(x) - S \circ h_n(x)]^2} \tag{4}$$

and applies a Gaussian filter to provide a smooth displacement field. Cachier *et al.* [6,20] have shown that using Equation (4) closely relates Equation (1) with a second order gradient descent of the SSD criterion, in which each iteration n sets h_{n+1} to the minimum of the SSD quadratic approximation at h_n. We refer the reader to these articles for a more technical discussion on this subject as well as for the formula corresponding to the true second order gradient descent.

Relation with Optical Flow T and S are considered as successive time samples of an image 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. The displacements are computed by constraining the brightness of brain structures to be constant in time, so that the following equality holds [14]:

$$\frac{\partial I}{\partial t} + v \cdot \nabla_x I = 0. \tag{5}$$

Equation (5) is however not sufficient to provide a unique displacement for each voxel. By constraining the displacements to always lie in the direction of the brightness gradient $\nabla_x I$, we get:

$$v(x) = -\frac{\partial I(x,t)/\partial t}{\|\nabla_x I(x,t)\|^2} \nabla_x I(x,t). \tag{6}$$

In general, the resulting displacement field does not have suitable smoothness properties. Many regularization methods have been proposed to fill this purpose [2]. One that can be computed very efficiently was proposed by Thirion [28] in his description of the demons registration method using a complete grid of demons. It consists of smoothing each dimension of the vector field with a Gaussian filter. He also proposed to add $[\partial I(x,t)/\partial t]^2$ to the denominator of Equation (6) for numerical stability when $\nabla_x I(x,t)$ is close to zero, a term which

serves the same purpose as α^2 in the original optical flow formulation of Horn and Schunck [14]. As is presented by Bro-Nielsen and Gramkow [5], this kind of regularization approximates a linear elasticity transformation model.

With this in mind, the displacement that maps a voxel position in T to its position in S is found using an iterative method,

$$v_{n+1}(x) = G_\sigma \otimes \left(v_n - \frac{\partial I(x,t)/\partial t}{\|\nabla_x I(x,t)\|^2 + [\partial I(x,t)/\partial t]^2} \nabla_x I(x,t) \right). \qquad (7)$$

Spatial derivatives may be computed in several ways [14,4,26]. We have observed from practical experience that our method performs best when they are computed from the resampled source image of the current iteration. As shown in Section 2.1, this is in agreement with the SSD minimization. Temporal derivatives are obtained by subtracting the target images from the resampled source image of the current iteration. These considerations relate Equation (7) to Equation (1). The reader should note that the major difference between this method and other optical flow strategies is that regularization is performed *after* the calculation of the displacements in the gradient direction instead of using an explicit regularization potential in a minimization framework.

Relation with the Demons Algorithm Our algorithm actually is a small variation of the demons method [27,28] using a complete grid of demons, itself closely related to optical flow as described in the previous section. The demons algorithm finds the displacements using the following formula,

$$v_{n+1}(x) = G_\sigma \otimes \left(v_n + \frac{S \circ h_n(x) - T(x)}{\|\nabla T(x)\|^2 + [S \circ h_n(x) - T(x)]^2} \nabla T(x) \right).$$

As can be seen from the last equation, the only difference between our formulation (Equation (1)) and the demons method is that derivatives are computed on the resampled source image of the current iteration. This modification was performed following the observations on the minimization of the SSD criterion.

2.2 Intensity Transformation

Previous to each iteration of the geometrical transformation, an intensity correction is performed on S so that the intensities of its structures match those in T. The displacement field is then updated by replacing S with its intensity corrected version in Equation (1).

The intensity correction process starts by defining the set C of intensity couples from corresponding voxels of T and of the current resampled source image $S \circ h$, which will be designated by S in this section for simplicity. Hence, the set C is defined as

$$C = \Big\{ \big(S(i), T(i)\big); 1 \leq i \leq N \Big\},$$

where N is the number of voxels in the images. $S(i)$ and $T(i)$ correspond to the intensity value of the i^{th} voxel of S and T respectively when adopting the customary convention of considering images as one-dimensional arrays. From there, we show how to perform intensity correction if one or two functional dependences can be assumed between the structures intensities.

Monofunctional Dependence Assumption Our goal is to model the transformation that characterizes the mapping from voxel intensities in S to those in T, knowing that some elements of C are erroneous, i.e. that would not be present in C if S and T were perfectly matched. If we can assume a monofunctional dependence of the intensities of T with regards to those of S as well as additive stationary Gaussian white noise η on the intensity values of T, then we can adopt the model

$$T(i) = f(S(i)) + \eta(i), \tag{8}$$

where f is an unknown function to be estimated. This is exactly the model employed in [22,21] which leads to the correlation ratio as the measure to be maximized for registration. In that approach, for a given transformation, one seeks the function that best describes T in terms of S. It is shown that, in a maximum likelihood context, the intensity function $\hat{f}$ that best approximates f is a least squares (LS) fit of T in terms of S.

Here the major difference is that we seek a high-dimensional geometrical transformation. As opposed to affine registration where the transformation is governed by the majority of good matches, we have seen in Section 2.1 that using the elastic registration model, displacements are found using mainly local information (i.e. gradients, local averages, etc.). Hence, we can not expect good displacements in one structure to correct for bad ones in another; we have to make certain each voxel is moved properly during each iteration. For this, since the geometrical transformation is found using intensity similarity, the most precise intensity transformation is required. Consequently, instead of performing a standard least squares regression, we have opted for a robust linear regression estimator which will remove outlying elements of C during the estimation of the intensity transformation. To estimate f we use the least trimmed squares (LTS) method followed by a binary reweighted least squares (RLS) estimation [25]. The combination of these two methods provides a very robust regression technique with outliers detection, while ensuring that a maximum of pertinent points are used for the final estimation.

Least Trimmed Squares Computation For our particular problem, we will constrain the unknown function f to be a polynomial function with degree p:

$$f(s) = \theta_0 + \theta_1 s + \theta_2 s^2 + \cdots + \theta_p s^p,$$

where we need to estimate the polynomial coefficients $\boldsymbol{\theta} = [\theta_0, \ldots, \theta_p]$. A regression estimator will provide a $\hat{\boldsymbol{\theta}} = [\hat{\theta}_0, \ldots, \hat{\theta}_p]$ which can be used to predict the value of $T(i)$ from $S(i)$, $\hat{T}(i) = \hat{\theta}_0 + \hat{\theta}_1 S(i) + \hat{\theta}_2 S(i)^2 + \cdots + \hat{\theta}_p S(i)^p$, as well

as the residual errors $r(i) = T(i) - \hat{T}(i)$. A popular method to obtain $\hat{\boldsymbol{\theta}}$ is to minimize the sum of squared residual errors,

$$\hat{\theta} = \arg\min_{\theta} \sum_{i=1}^{N} r(i)^2,$$

which leads to the standard LS solution. It is found by solving a linear system using the Singular Value Decomposition (SVD) method. This method is known to be very sensitive to outliers and thus is expected to provide a poor estimate of the monofunctional mapping from S to T. The LTS method solves this problem by minimizing the same sum on a subset of all residual errors, thus rejecting large ones corresponding to outliers,

$$\hat{\theta} = \arg\min_{\theta} \sum_{i=1}^{h} \rho(i),$$

where $\rho(i)$ is the i^{th} smallest value of the set $\{r(1)^2, \ldots, r(N)^2\}$. This corresponds to a standard LS on the c values that best approximates the function we are looking for. Essentially, c/N represents the percentage of "good" points in C and must be at least 50%. A lesser value would allow to estimate parameters that model a minority of point which could then all be outliers. The value of c will vary according to the modalities used during registration. Assigning actual values to c is postponed to Section 4.

Our method for LTS minimization is a simple iterative technique. First, we randomly pick c/N points from C. We then iterate between calculating $\hat{\boldsymbol{\theta}}$ using the standard LS technique on the selected points and choosing the h/N closest points from C. This process is carried until convergence, usually requiring less than 5 iterations and is guaranteed to find at least a local minimum of the LTS criterion [24].

Reweighted Least Squares Computation As discussed in [25], the LTS method is very robust, but it tends to provide an estimate $\hat{\boldsymbol{\theta}}$ that is notably less accurate than that we would obtain with a standard LS in the absence of outliers. The solution may be refined by considering all the points that relate well to the LTS estimate, not only the best $c/N \times 100\%$. An efficient technique to achieve this is the so-called RLS regression [25], which minimizes the sum of squared residuals over all the points that are not "too far" from the LTS estimate,

$$\hat{\theta} = \arg\min_{\theta} \sum_{i=1}^{N} w_i r(i), \quad \text{where} \quad w_i = \begin{cases} 1 & \text{if } r(i) \leq 3\hat{\sigma}, \\ 0 & \text{otherwise,} \end{cases}$$

where $\hat{\sigma}$ is a scale parameter which actually estimates the standard deviation of the Gaussian noise η introduced in Equation (8). Such an estimate can be computed directly from the final value of the LTS criterion,

$$\hat{\sigma} = \sqrt{\frac{K}{c} \sum_{i=1}^{c} \rho(i)}, \qquad \text{with} \quad \frac{1}{K} = \int_{-\alpha}^{\alpha} x^2 g(x)\, dx, \tag{9}$$

where $g(x)$ is the Gaussian distribution $N(0,1)$ and α is the $(0.5 + c/2N)^{\text{th}}$ quantile of $g(x)$. In Equation (9), K is a normalization factor introduced because the LTS criterion is not a consistent estimator of σ when the $r(i)$ are distributed like $N(0,\sigma^2)$, except when $c = N$.

Bifunctional Dependence Assumption Functional dependence as expressed in Equation (8) implicitly assumes that two structures having similar intensity ranges in S should also have similar intensity ranges in T. With some combinations of multimodal images, this is a crude approximation. For example, ventricles and bones generally give similar response values in a MR T1 weighted image while they appear with very distinct values in a CT scan. Conversely, white and black matter are well contrasted in a T1 image while corresponding to similar intensities in a CT.

To circumvent this difficulty, we have developed a strategy that enables the mapping of an intensity value in S to not only one, but two possible intensity values in T. This method is a natural extension of the previous method. Instead of computing a single function that maps the intensities of S to those of T, two functions are estimated and the mapping becomes a weighted sum of these two functions.

We start with the assumption that if a point has an intensity s in S, the corresponding point in T has an intensity t that is normally distributed around two possible values depending on s, $f_{\boldsymbol{\theta}}(s)$ and $f_{\boldsymbol{\psi}}(s)$. In statistical terms, this means that, given s, t is drawn from a mixture of Gaussian distribution,

$$P(t|s) = \pi_1(s)N(f_{\boldsymbol{\theta}}(s),\sigma^2) + \pi_2(s)N(f_{\boldsymbol{\psi}}(s),\sigma^2), \tag{10}$$

where $\pi_1(s)$ and $\pi_2(s) = 1 - \pi_1(s)$ are mixing proportions that depend on the intensity in the source image, and σ^2 represents the variance of the noise in the target image. Consistently with the functional case, we will restrict ourselves to polynomial intensity functions, i.e. $f_{\boldsymbol{\theta}}(s) = \theta_0 + \theta_1 s + \theta_2 s^2 + \cdots + \theta_p s^p$, and $f_{\boldsymbol{\psi}}(s) = \psi_0 + \psi_1 s + \psi_2 s^2 + \cdots + \psi_p s^p$.

An intuitive way to interpret this modelling is to state that for any voxel, there is a binary "selector" variable $\epsilon = \{1,2\}$ that would tell us, if it was observed, which of the two functions $f_{\boldsymbol{\theta}}$ or $f_{\boldsymbol{\psi}}$ actually serves to map s to t. Without knowledge of ϵ, the best intensity correction to apply to S (in the sense of the conditional expectation [19]) is seen to be a weighted sum of the two functions,

$$f(s,t) = P(\epsilon = 1|s,t)f_{\boldsymbol{\theta}}(s) + P(\epsilon = 2|s,t)f_{\boldsymbol{\psi}}(s), \tag{11}$$

in which the weights correspond to the probability that the point be mapped according to either the first or the second function. We see that the intensity correction is now a function of both s and t. Applying Bayes' law, we find that for $\epsilon = \{1,2\}$:

$$P(\epsilon|s,t) = \frac{P(\epsilon|s)P(t|\epsilon,s)}{P(t|s)},$$

and thus, using the fact that $P(\epsilon|s) = \pi_\epsilon(s)$ and $P(t|\epsilon, s) = G_\sigma(t - f_\epsilon(s))$, the weights are determined by

$$P(\epsilon|s,t) = \frac{\pi_\epsilon(s)\, G_\sigma(t - f_\epsilon(s))}{\pi_1(s)\, G_\sigma(t - f_{\boldsymbol{\theta}}(s)) + \pi_2(s)\, G_\sigma(t - f_{\boldsymbol{\psi}}(s))}, \tag{12}$$

where it should be clear from the context that $f_\epsilon \equiv f_{\boldsymbol{\theta}}$ if $\epsilon = 1$, and $f_\epsilon \equiv f_{\boldsymbol{\psi}}$ if $\epsilon = 2$.

In order to estimate the parameters of the model, we employ an ad hoc strategy that proceeds as follows. First, $\boldsymbol{\theta}$ is estimated using the LTS/RLS method described in section 2.2. The points not used to compute $\boldsymbol{\theta}$, in a number between 0 and $N - c$, are used to estimate $\boldsymbol{\psi}$ still using the same method. Note that if this number is less than $10 \times p$, p being the polynomial degree, functional dependence is assumed and we fall back to the monofunctional assumption.

This provides a natural estimation of the "selector" variable for each voxel: the n_1 points that were used to build $f_{\boldsymbol{\theta}}$ are likely to correspond to $\epsilon = 1$, while the n_2 points used to build f_ψ are likely to correspond to $\epsilon = 2$. Finally, the points that are rejected while estimating $\boldsymbol{\psi}$ are considered as bad intensity matches. A natural estimator for the variance σ^2 is then

$$\hat{\sigma}^2 = \frac{n_1}{n_1 + n_2}\hat{\sigma}_1^2 + \frac{n_2}{n_1 + n_2}\hat{\sigma}_2^2,$$

where $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ are the variances found respectively for $f_{\boldsymbol{\theta}}$ and $f_{\boldsymbol{\psi}}$ during the RLS regression (See Section 2.2.). Similarly, the mixing proportions are computed according to

$$\hat{\pi}_\epsilon(s) = \frac{n_\epsilon(s)}{n_1(s) + n_2(s)}, \qquad \epsilon = \{1, 2\},$$

in which $n_\epsilon(s)$ is the number of voxels having an intensity s and used to build the function f_ϵ. Notice that in the case where $n_1(s) = n_2(s) = 0$ (i.e. no voxel corresponding to the intensity class s has been taken into account in the computation of $f_{\boldsymbol{\theta}}$ or $f_{\boldsymbol{\psi}}$), then we arbitrarily set the mixing proportions to $\hat{\pi}_1(s) = \hat{\pi}_2(s) = 0.5$.

The intensity correction of S can now be performed by reinjecting the estimated parameters in Equations (12) and (11).

3 Data

Most of the data used in the following experiments were obtained from BrainWeb [3,8,15,9]. This tool uses an atlas with a resolution of $1 \times 1 \times 1\text{mm}^3$ comprising nine segmented regions from which T1, T2 and PD images can be generated. Three images, one of each modality, were generated with the same resolution as the atlas, 5% noise and no intensity non-uniformity. Since they are generated from the same atlas, they represent the same underlying anatomy and are all perfectly matched.

We also made use of a T1 MR image and a CT image, both from different subjects and having a resolution of $1 \times 1 \times 1\text{mm}^3$. Both these images were affinely registered with the atlas using the correlation ratio method [23]. To differentiate the T1 image obtained with the atlas from the other T1 image, the latter will be referenced as SCH.

The images all respect the neurological convention, i.e. on coronal and axial slices, the patient's left is on the left side of the image.

4 Results and Discussion

In the following section we present registration results involving images obtained from several different kinds of modalities. First, we show a typical example where monofunctional dependence can be assumed: registration of an atlas with an MR image. Then, more practical examples are shown where images from different modalities are registered and where bifunctional dependence may be assumed.

The multilevel process was performed at three resolution levels, namely 4mm, 2mm and 1mm per voxel. Displacement fields at one level are initialized from the result of the previous level. The initial displacement field v_0 is set to a zero. The Gaussian filter G_σ used to smooth the displacement field has a standard deviation of 1mm. 128 iterations are performed at 4mm/voxel, 32 at 2mm/voxel and 8 at 1mm/voxel. We believe that making use of a better stopping criterion, such as the difference of the SSD values between iterations, would probably improve the results shown below.

4.1 Monofunctional Dependence

We present here the result of registering the atlas with SCH. Since the atlas can be used to generate realistic MR images, it is safe to assume a functional dependence from the intensity of the atlas to that of SCH. Also, since SCH and the atlas are well aligned due to the affine registration, we have roughly estimated that the number of points already well matched are at least $0.80 \times N$, to which we have set the value of c. Since 10 classes are present in the atlas, the polynomial degree chosen was set to 9.

The result of registration is presented in Figure 1. For lack of space, we only show one set of corresponding slices extracted from the 3D images. However, we wish to make clear to the reader that the registration was performed in 3D, *not* slice by slice. More illustrations will be found in [12]. From left to right, the first picture shows an axial slice of the atlas. The second one presents the corresponding slice of SCH (which was chosen as the target image). The third and fourth pictures show the deformed atlas after elastic registration, respectively without and with intensity correction.

As can be seen, large morphometric differences have been corrected. Still, the matching is not perfect which may be observed by comparing the shape of several structures between SCH and the deformed atlas, e.g. the ventricles and the white matter. Registration imperfections are reflected in the intensity corrected image

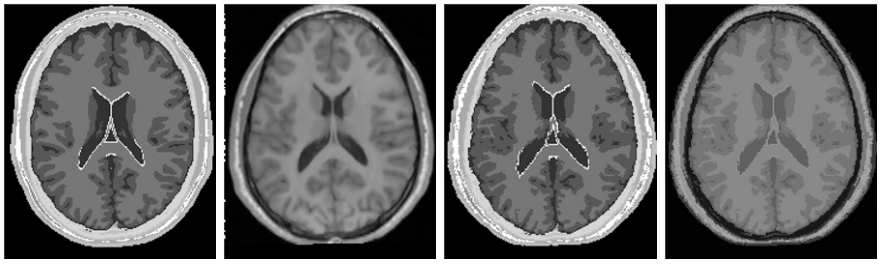

Fig. 1. Axial slices of the atlas to SCH registration result. From left to right: Atlas; SCH; atlas without intensity correction after registration with SCH; atlas with intensity correction after registration with SCH.

(right picture), where one may notice that the CSF intensity is slightly brighter than that in SCH (as can be seen in the ventricles and around the cortex). This problem can also be observed by looking at the intensity transformation function presented in Figure 5. The intensity level corresponding to the CSF is overestimated due to an overlap of the CSF in the atlas with the gray and white matter in SCH, especially around the cortical area which is known to present large variations between subjects.

This is probably an inherent limitation of elastic models when used in the context of inter-subject registration. The strong smoothness constraints imposed by the Gaussian regularization (or related regularization techniques) may prevent the assessment of large and uneven displacements required to match the anatomical structures of different subjects. To allow for larger displacements, another regularization strategy should be used, such as that based on a fluid model [7] or on a non-quadratic potential energy [13].

4.2 Bifunctional Dependence

When registering images from different modalities, monofunctional dependence may not necessarily be assumed. Here, we applied the method described in Section 2.2 where two polynomial functions of degree 12 are estimated. This number was set arbitrarily to a relatively high value to enable important intensity transformations.

Figure 2 presents the result of registering T1 with CT. Using these last two modalities, most intensities should be mapped to gray and only the skull, representing a small portion of the image data, should be mapped to white. After affine registration almost all voxels are well matched. Hence, in this particular case, we have chosen a high value for c set to $0.90 \times N$.

As we can see in Figure 2, the skull, shown in black in the MR image and in white in the CT scan, is well registered and the intensity transformation adequate. The top right graph of Figure 5 presents the functions $f_{\boldsymbol{\theta}}$ and $f_{\boldsymbol{\psi}}$ found during the registration process. The red line corresponds to $f_{\boldsymbol{\theta}}$ and the blue one to $f_{\boldsymbol{\psi}}$. The line width for a given intensity s is proportional to the value

of the corresponding $\pi_\epsilon(s)$. The gray values represent the joint histogram after registration. As can be observed on this graph, the polynomials found fit well with the high density clusters of the joint histogram. Still, some points need to be addressed.

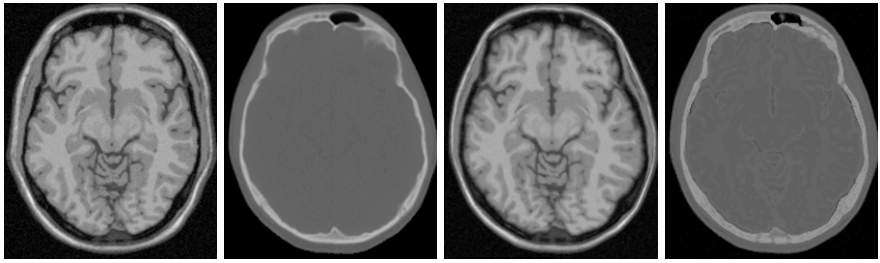

Fig. 2. Axial slices of the T1 to CT registration result. From left to right: T1; CT; T1 without intensity correction after registration with T1; T1 with intensity correction after registration with T1.

We can observe that due to the restricted polynomial degree, $f_{\boldsymbol{\theta}}$, (shown in red) oscillates around the CT gray value instead of fitting a strait line. This is reflected in the intensity corrected image, where the underlying anatomy can still be observed by small intensity variations inside the skull. This artifact has insubstantial consequences during the registration process since the difference between most of the voxel intensities is zero, resulting in null displacements. The displacements driving the deformation will be those of the skull and the skin contours, and will be propagated in the rest of the image as an effect of smoothing the displacement field.

We also notice that $f_{\boldsymbol{\psi}}$ (shown in blue), which is mainly responsible for the mapping of the skull, does not properly model the cluster it represents for intensities smaller than 5. The mapping for these intensities is slightly underestimated. This may have two causes. First, as in the previous case, it might be due to the restricted polynomial degree. Second, we can notice that some of the background values in T1 that have an intensity close to 0 are mapped to gray values in the CT which correspond to soft tissues. This means that some of the background in the T1 is matched with the skin in the CT. This has the effect of "pulling" $f_{\boldsymbol{\psi}}$ closer to the small cluster positioned around (2,65). If the underestimation of $f_{\boldsymbol{\psi}}$ arises because of the second reason, letting the algorithm iterate longer might provide a better result.

In Figures 3 and 4, we present the result of registering T2 and PD respectively with SCH. The bottom graphs of Figure 5 show the corresponding intensity transformations. For these experiments, c was set to $0.60 \times N$, a value we have found to be effective for these types of modalities after affine registration.

One observation that can be made by looking at the graphs of Figure 5 is that the estimated functions $f_{\boldsymbol{\theta}}$ and $f_{\boldsymbol{\psi}}$ are quite similar in both cases. This suggests

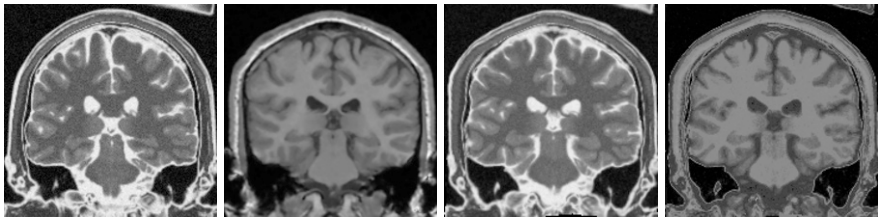

Fig. 3. Coronal slices of the T2 to SCH registration result. From left to right: T2; SCH; T2 without intensity correction after registration with SCH; T2 with intensity correction after registration with SCH.

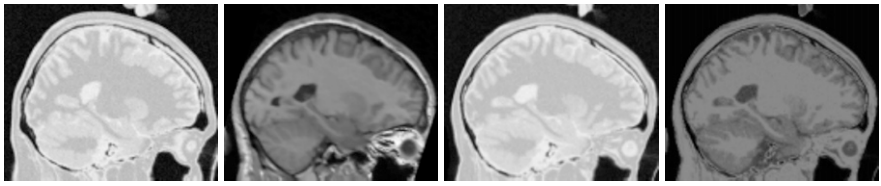

Fig. 4. Sagittal slices of the PD to SCH registration result. From left to right: PD; SCH; PD without intensity correction after registration with SCH; PD with intensity correction after registration with SCH.

that assuming a monofunctional dependence would be relevant. However, the results we obtained when registering T2 with SCH, and PD with SCH, using the monofunctional model were less convincing than when using the bifunctional model [12].

This may be explained by a closer look at our bifunctional intensity modelling. Equation 10 reflects the assumption that if an anatomical point has an intensity s in S, the corresponding point has an intensity t in T that is distributed normally around two possible values depending on s. But it makes no assumption about how the intensities in S are distributed. This models the intensities of S without noise, which may not necessarily be well justified, but enables the use linear regression to estimate the intensity transformation.

The effect of noise in S is reflected in the joint histograms by enlarging clusters along the x axis. This, added to bad matches and partial volume effect, creates many outliers in C and makes the assessment of the true intensity transformation more difficult and more resistant to our robust regression technique. Preprocessing of S using for example anisotropic diffusion may narrow the clusters and provide better results [22].

Adding the estimation of a second function in the bifunctional model helps counter the effect of noise on S. For example, the CSF in the PD image has intensity values ranging from about 200 to 240 and gray matter from about 175 to 210. In SCH, these ranges are about 30 to 70 and 55 to 80 respectively. As can be seen in Figure 5, $f_{\boldsymbol{\theta}}$ models well the gray matter cluster but fails to reflect the

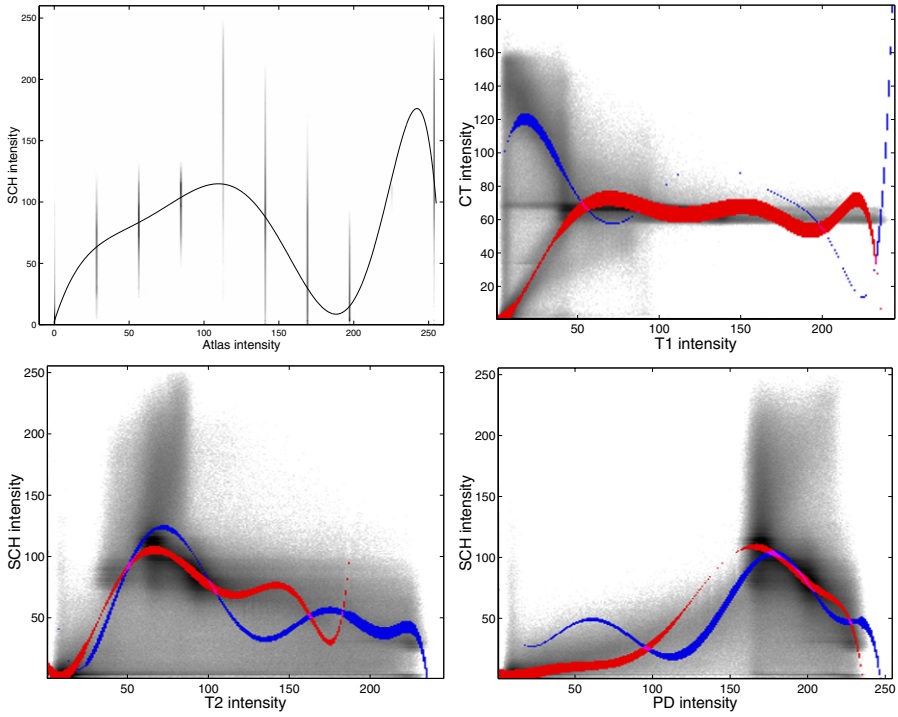

Fig. 5. Graphs of the intensity corrections found in our experiments. From left to right, top to bottom: Atlas to SCH, T1 to CT, T2 to SCH, PD to SCH. In the last three graphs, which correspond to bifunctional models, the red (bright) line represents $f_{\boldsymbol{\theta}}$ and the blue (dark) one $f_{\boldsymbol{\psi}}$. The line width for a given intensity value s in the source image corresponds to the value of the corresponding proportion, $\pi_\epsilon(s)$. The gray values represent the joint histogram after registration.

CSF transformation. Estimating the second polynomial $f_{\boldsymbol{\psi}}$ solves this problem by considering the CSF cluster.

4.3 Displacement Field Comparison

Since the atlas, the T1, the T2 and the PD images have all been registered with SCH, it is relevant to compare some statistics of the resulting displacement fields to assess if our algorithm provides consistent results across modalities.

We computed statistics regarding the difference between any two of these displacement fields. The length of the vectors of the resulting difference fields were calculated. Each cell of Table 1 presents, for each combination of displacement fields, the median length, the average length with the corresponding standard deviation and the maximum length of the difference field.

The two largest average errors are 1.58 mm and 1.76 mm, and were found when comparing the Atlas-SCH registration with T1-SCH and PD-SCH, respec-

Table 1. Statistics regarding the displacements difference between each type of registration. Each cell presents the median length, the average length with the corresponding standard deviation and the maximum length. All measures are in millimeters.

Difference (mm)	Atlas-SCH T1-SCH	Atlas-SCH T2-SCH	Atlas-SCH PD-SCH	T1-SCH T2-SCH	T1-SCH PD-SCH	T2-SCH PD-SCH
median	1.46	1.13	1.67	1.00	1.01	1.32
average	1.58	1.23	1.76	1.18	1.16	1.40
std. dev.	0.84	0.63	0.79	0.78	0.71	0.68
maximum	6.99	5.14	7.10	7.17	8.08	6.86

tively. This may be explained by the intensity correction bias for the CSF that would tend to attenuate displacements and produce larger errors, a problem invoked in Section 4.1. Aside from these, the average error length varies between 0.97mm and 1.40mm and the median error is between 0.85mm and 1.32mm. These are values in the range of the image resolution of 1.0mm. Note also that all the standard deviations are below this value.

Also, we observe that the results obtained when registering images from different modalities (Atlas-SCH, T2-SCH, and PD-SCH) seem to be consistent with the monomodal registration result (T1-SCH), in which no intensity correction was performed. This suggests that the intensity correction may not cause a sensible degradation of the registration when compared to the monomodal case. We point out, however, that these are global measures that are presented to provide an idea of the differences between the displacement fields. They do not strictly provide a validation of the method, but do show a certain coherence between the different results we obtained.

5 Conclusion

In this paper, we introduced an original method to perform non-rigid registration of multimodal images. This iterative algorithm is composed of two sections: the geometrical transformation and the intensity transformation. We have related the geometrical transformation computation to several popular registration concepts: SSD, optical flow and the demons method. Two intensity transformation models were described which assume either monofunctional or bifunctional dependence between the intensities of the images to match. Both of these models are built using robust estimators to enable precise and accurate transformation solutions. Results of registration were presented and showed that the algorithm performs well for several kinds of modalities including T1 MR, T2 MR, PD MR, CT and segmentations, and provides consistent results across modalities.

A current limitation of the method is that it uses Gaussian filtering to regularize the displacement field. This technique was chosen for its computational efficiency rather than for its physical relevance. In the context of inter-subject registration, other regularization strategies need to be investigated to better account for morphological differences.

References

1. D. C. Barber. Registration of low resolution medical images. *Physics in Medecine and Biology*, 37(7):1485–1498, 1992.
2. 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.
3. Simulated brain database. http://www.bic.mni.mcgill.ca/brainweb/.
4. J. W. Brandt. Improved accuracy in gradient-based optical flow estimation. *International Journal of Computer Vision*, 25(1):5–22, 1997.
5. M. Bro-Nielsen and C. Gramkow. Fast fluid registration of medical images. In K. H. Höhne and R. Kikinis, editors, *Proc. VBC'96*, volume 1131 of *Lecture Notes in Computer Science*, pages 267–276. Springer-Verlag, 1996.
6. P. Cachier, X. Pennec, and N. Ayache. Fast non rigid matching by gradient descent: Study and improvements of the "demons" algorithm. Technical Report 3706, INRIA, June 1999.
7. G. E. Christensen, R. D. Rabbitt, and M. I. Miller. Deformable templates using large deformation kinematics. *IEEE Transactions in Medical Imaging*, 5(10):1435–1447, October 1996.
8. C. A. Cocosco, V. Kollokian, R. K.-S. Kwan, and A. C. Evans. Brainweb: Online interface to a 3D MRI simulated brain database. *NeuroImage, Proc. HBM'97*, 5(4):S425, May 1997.
9. D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans. Design and construction of a realistic digital brain phantom. *IEEE Transactions in Medical Imaging*, 17(3):463–468, June 1998.
10. J. Feldmar, J. Declerck, G. Malandain, and N. Ayache. Extension of the ICP algorithm to non-rigid intensity-based registration of 3D volumes. *Computer Vision and Image Understanding*, 66(2):193–206, May 1997.
11. T. Gaens, F. Maes, D. Vandermeulen, and P. Suetens. Non-rigid multimodal image registration using mutual information. In W. M. Wells, A. Colchester, and S. Delp, editors, *Proc. MICCAI'98*, volume 1496 of *Lecture Notes in Computer Science*, pages 1099–1106. Springer-Verlag, 1998.
12. A. Guimond, A. Roche, N. Ayache, and J. Meunier. Multimodal Brain Warping Using the Demons Algorithm and Adaptative Intensity Corrections. Technical Report 3796, INRIA, November 1999.
13. P. Hellier, C. Barillot, E. Mémin, and P. Pérez. Medical image registration with robust multigrid techniques. In *Proc. MICCAI'99*, volume 1679 of *Lecture Notes in Computer Science*, pages 680–687, Cambridge, England, October 1999.
14. B. K. P. Horn and B. G. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, August 1981.
15. R. K.-S. Kwan, A. C. Evans, and G. B. Pike. An extensible MRI simulator for post-processing evaluation. In K. H. Höhne and R. Kikinis, editors, *Proc. VBC'96*, volume 1131 of *LNCS*, pages 135–140. Springer-Verlag, 1996.
16. Y. H. Lau, M. Braun, and B. F. Hutton. Non-rigid 3d image registration using regionally constrainted matching and the correlation ratio. In F. Pernus, S. Kovacic, H.S. Stiehl, and M.A. Viergever, editors, *Proc. WBIR'99*, pages 137–148, 1999.
17. F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions in Medical Imaging*, 16(2):187–198, 1997.
18. J. B. A. Maintz, E. H. W. Meijering, and M. A. Viergever. General multimodal elastic registration based on mutual information. In K. M. Hanson, editor, *Medical*

Imaging 1998: Image Processing (MI'98), volume 3338 of *SPIE Proceedings*, pages 144–154, Bellingham (WA), USA, April 1998.
19. A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, Inc., third edition, 1991.
20. X. Pennec, P. Cachier, and N. Ayache. Understanding the "demon's algorithm": 3d non-rigid registration by gradient descent. In *Proc. MICCAI'99*, volume 1679 of *Lecture Notes in Computer Science*, pages 597–605. Springer-Verlag, 1999.
21. A. Roche, G. Malandain, and N. Ayache. Unifying maximum likelihood approaches in medical image registration. *International Journal of Imaging Systems and Technology: Special Issue on 3D Imaging*, 2000. In press.
22. A. Roche, G. Malandain, N. Ayache, and S. Prima. Towards a better comprehension of similarity measures used in medical image registration. In *Proc. MICCAI'99*, volume 1679 of *LNCS*, pages 555–566. Springer-Verlag, September 1999.
23. A. Roche, G. Malandain, X. Pennec, and N. Ayache. The correlation ratio as a new similarity measure for multimodal image registration. In *Proc. MICCAI'98*, volume 1496 of *LNCS*, pages 1115–1124. Springer-Verlag, October 1998.
24. P. J. Rousseeuw and K. Van Driessen. Computing LTS Regression for Large Data Sets. Technical report, Statistics Group, University of Antwerp, 1999.
25. Peter J. Rousseeuw and Annick M. Leroy. *Robust Regression and Outlier Detection*. Wiley series in probability and mathematical statistics. John Wiley & Sons, 1987.
26. E. P. Simoncelli. Design of multi-dimensional derivative filters. In *International Conference on Image Processing*, Austin, USA, November 1994. IEEE.
27. J.-P. Thirion. Fast non-rigid matching of 3D medical images. Technical Report 2547, INRIA, Sophia-Antipolis, 1995.
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. Arthur W. Toga. *Brain Warping*. Academic Press, 1999.
30. P. Viola and W. M. Wells. Alignment by maximization of mutual information. *International Journal of Computer Vision*, 24(2):137–154, 1997.
31. R. P. Woods, J. C. Mazziotta, and S. R. Cherry. MRI-PET registration with automated algorithm. *Journal of Comp. Assist. Tomography*, 17(4):536–546, 1993.