Fast Non-Rigid Matching of 3D Medical Images

Jean-Philippe Thirion

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Fast Non-Rigid Matching of 3D Medical Images

Jean-Philippe THIRION

N° 2547

May 1995

PROGRAMME 4

ISSN 0249-6399

Fast Non-Rigid Matching of 3D Medical Images

Jean-Philippe THIRION *

Programme 4 — Robotique, image et vision
Projet Epidaure **

Rapport de recherche n° 2547 — May 1995 — 37 pages

Abstract: We present a method to perform the non-rigid matching of two 3D medical images, which is fast and fully automated. We illustrate this method in several applications, such as inter-patient registration, cardiac image sequence analysis, and the automatic registration with anatomical atlases, which we identify as currently important challenges for medical image processing. Several interesting inter-patient matching methods have been proposed recently, showing the interest of such techniques, but whose computation time is still prohibitive for practical clinical applications. The main interest of our method is to perform a "satisfying" non-rigid matching of large image volumes within only several tens of minutes on a conventional workstation. We describe our method as based on "demons" (in the sense of Maxwell's demons), which locally deform the model to push it into a mold, as opposed to more traditional "attractor" based deformable model techniques. Finally, we present experiments on real data, and discuss two current limitations of our technique: the sensitivity to initial positioning, and to intensity variations between images.

Key-words: inter-patient registration, non-rigid matching, elastic matching, image sequence analysis, deformable model

(Résumé : tsvp)

*Email: jean-philippe.thirion@sophia.inria.fr
**http://www.inria.fr/Equipes/EPIDAURE-eng.html

Unité de recherche INRIA Sophia-Antipolis
2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex (France)
Téléphone : (33) 93 65 77 77 – Télécopie : (33) 93 65 77 65

Une méthode rapide de recalage non-rigide d'images médicales 3D

Résumé : Nous présentons une méthode de recalage non-rigide d'images médicales 3D, qui est rapide et entièrement automatique. Nous illustrons cette méthode par plusieurs applications, telles que le recalage entre patients différents, l'analyse d'images cardiaques, ou le recalage avec un atlas anatomique, que nous identifions comme étant des problèmes stratégiques actuels du traitement d'images médicales. Des méthodes intéressantes de recalage entre patients ont été proposées récemment, qui ont montré l'intérêt d'une telle technique, mais dont le temps de calcul reste prohibitif pour des applications cliniques pratiques. L'intérêt principal de notre méthode est de réaliser un recalage non-rigide "satisfaisant" de grands volumes d'images en quelques dizaines de minutes seulement, à l'aide d'une station de travail standard. Nous décrivons notre méthode comme étant basée sur des "démons" (dans un sens très similaire aux démons de Maxwell), qui déforment localement le modèle pour le faire tenir dans un moule, par opposition aux techniques plus traditionnelles basées sur des "attracteurs". Pour conclure, nous présentons des expériences sur des données réelles, et nous discutons deux limites actuelles de notre technique : la sensibilité au positionement initial, et aux variations d'intensité entre images.

Mots-clé : recalage inter-patient, recalage non-rigide, recalage élastique, analyse de séquence d'images, modèle déformable

1 Introduction

Contrary to rigid registration, the non-rigid matching of 3D medical images raises many unanswered questions. Rather than debating a single definition of non-rigid matching, we will discuss it through several very challenging medical applications.

We then present some existing non-rigid matching methods, and introduce an original distinction between two types of matching algorithms, that we call "attractor" based and "demon" based deformable models. This allows us to present our own technique, which belongs to the demon-based techniques.

We discuss the current limitations of our non-rigid matching technique: those limitations are shared by the majority of existing techniques, and are the sensitivity to original positioning, and to intensity changes. Finally, we display experimental results for image sequence analysis and inter-patient matching.

2 Non-rigid matching in general

2.1 Several possible definitions

A first natural application of 3D non-rigid matching is the tracking of deformable objects in sequences of 3D images, such as the heart motion in SPECT or in Magnetic Resonance (MR) gated images. The aim is the early diagnosis of cardiac disease, which are among the leading causes of death in western countries. Studying heart motion, with possible comparisons between stress and rest, is a current clinical application which needs quantitative motion analysis tools to help the diagnosis.

Fortunately in that case, we can give an objective definition of non-rigid matching, which is establishing the trajectory of each physical point (ideally the trajectory of atoms). However, and except for very specific acquisitions such as tagged MRI, image processing techniques are generally better in establishing the displacement normal to the object boundaries than along them (known as the "aperture problem"). For example, the twist motion of the heart is generally difficult to quantify.

Those lateral displacements can be partially recovered with feature-based registration techniques. In that case, what is matched are not physical points,

but similar object features between images, such as crest lines or feature points, defined, for example, using differential geometry. As the heart beat corresponds to the propagation of the nervous influx in the myocardium, waves of deformation can be tracked, and may reveal information just as useful for diagnosis than physical point motion. This is similar to following waves over the sea surface, which is basically a translation, instead of focusing on the trajectories of individual water molecules which are known to be circular.

Another large application field for 3D non-rigid matching is the matching of the 3D images of two different patients. Anatomy relies on the assumption that a large set of macroscopic structures is common between individuals: we generally have two eyes and ears, one nose, five fingers in each hand,..., and also microscopic differences such as the finger prints. Sometimes, differences are characteristic of a pathology, for example large brain ventricles for hydrocephaly. Automatic inter-patient registration tools can serve in building an average representation of patients, with associated variability measurements, with at least two broad classes of applications: the diagnosis of pathologies by the measurement of shape discrepancies, and the functional localization of disease, not only for diagnosis, but also for surgical planning and guidance.

Unfortunately, and despite the wide range of applications, no clear and objective definition of this kind of non-rigid matching can be given. We will refer to it as *anatomical matching*, which leads us to, at least, three distinct definitions of non-rigid matching:

- physical point matching
- geometric feature matching
- anatomical matching

2.2 Common features of non-rigid matching methods

Non-rigid matching techniques have some general common features which help in their classification. To define such a method, one has to define precisely at least two things:

- the type of features which are matched (for example, physical points, geometric features, anatomical features ...)

- the class of deformation allowed (rigid, affine, quadratic, spline, incompressible deformation ...)

Another general trend of non-rigid matching (except perhaps for physical point matching) is that a compromise must be devised between the proximity of similar features and the "amount" of deformation allowed. In the case of deformable surface models, or "snakes" as defined by Kass, Witkin and Terzopoulos in [13], this corresponds to the balance between internal forces[1] (the rigidity of the model), and external forces (the attraction of similar features). We will refer to *deformability* when referring to this balance.

This parameter is unavoidable, especially for anatomical matching, because if any deformability is allowed, one can always deform one object into any other object, without any anatomical significance (i.e. morphing a head into a teapot). We might consider using statistical information to establish global or local deformability. However, as automatic matching techniques are necessary to compute those statistics, this problem becomes a closed loop, difficult to solve.

2.3 Some existing non-rigid matching techniques

2.3.1 Deformable surfaces

Most (but not all) of the existing non-rigid matching methods in 3D are variations of deformable surface models, initially presented in [13] for the case of 2D curves in 2D images, and extended to 3D surfaces by Cohen and Cohen in [8].

Initially, those methods were designed to segment the boundary of an object. Edge extraction is performed to extract boundary points in the image, and an initial deformable model evolves, attracted by the contour points while maintaining an internal coherence (or rigidity). Three different types of information are introduced by the user: the deformability of the model; the topology

[1] In fact, those internal forces could be preferably named "reaction" forces, because they correspond to the restitution by the deformable model of the energy given to it by the external forces. Some deformable models can however possess an internal energy source, such as the inflation force proposed in [8].

of the object surface (generally spherical, sometimes toric); and the initial positioning, which can lead to very different solutions. Such techniques have then been extended to motion tracking, each frame serving to initialize the deformable model for the next frame. They are generally poor at retrieving the lateral motion (along the contour).

Feldmar and Ayache in [10] have shown that taking into account differential geometrical information about the contour helps to recover partially this kind of information, and have also shown an application to inter-patient registration.

Among other deformable surface techniques, we can cite also [3], where a parametric surface is fit to the data, and where the deformation is analyzed using modal decomposition. At last, very interesting studies about the relevance of surface deformable models to heart motion analysis can be found in [17], with clinical experiments with dogs.

Those techniques necessitate however a segmentation of the object surface, and the problem of extending the deformation to the whole volume has still to be solved.

2.3.2 3D grid deformable models

In the case of inter-patient matching, probably the most convincing results have been achieved using a complete 3D deformable grid. Among the approaches proposed, we can cite [1], which is based on multiscale correlation, and [6], [7] which are based on a modeling of elasticity, relying on fluid mechanics.

While giving positive answers about the feasibility of inter-patient matching, those works raise two problems: prohibitive computation time, and the relevance of the physical model to inter-patient matching.

As regards computation time, it is stated in [6] that the registration of two 3D $128 \times 128 \times 148$ volumes took 9 hours on a MASPAR (massively parallel machine), rating an impressive 0.8 Gigaflops performance. In [1], the computation time is not mentioned, but the method relies on correlation, which is known to be very time consuming, even in 2D.

As those two methods rely on elasticity modeling, we refer to them as *elastic* matching methods (non-rigid matching referring to a broader class of methods, including for example plastic deformations). However, there is no a-priori reason why matching based on fluid dynamics applies to model the

similarities between two different patients. Those similarities come from the sharing of a common phylogeny (the development of species) and the undergoing of a similar ontogeny (the development of the organism, from embrio to adult).

In the design of our non-rigid matching method, we do not try to adhere too closely to a physical modeling of elasticity: there is no need to model precisely the behavior of a dynamic fluid if the subject is not a dynamic fluid; it can help, however, to get a better intuition about the effect of the parameters of the model (if only for those who are used to them). The same consideration applies to heart motion studies: the myocardium is an interleaved set of contractible fibers, not a viscous fluid. We must confess that the relevance of our own model is no more justified than the careful application of fluid dynamics, but being much simpler, its implementation is easier and the resulting algorithm much faster.

2.3.3 Feature lines and feature points

Contrary to the above techniques, we have previously developed another approach, presented in [9]. We extract specific feature lines, the *crest lines*, from the 3D images, perform non-rigid matching of those 3D lines, and then apply a warping to extend the point-to-point correspondence between lines to the whole 3D space. The main advantage is that a larger number of solutions can be explored in a reduced amount of time, because the 3D line representation is compact, yet conveys a great deal of relevant geometric information (see [19]). This allows us to perform successfully inter-patient registration in a reasonable amount of time, namely, about 20 minutes on a conventional DEC alpha workstation. Of course, this technique is canonical if what is needed is feature points matching. It is also justified with respect to anatomy: for the example of the brain, a large subset of the crest lines corresponds closely to the gyri and sulci, and have long been identified as landmarks by anatomists.

Some progress is still needed to make this method fully automatic (an iso-intensity threshold has to be defined). Also, the matching is good in the neighborhood of the crest lines, but could be improved in the other regions by a more local, intensity based algorithm.

One promising approach is to combine the crest lines technique, which has an anatomical foundation, and can anchor very precisely a set of corresponding points between the 3D images, with intensity-based matching, which is able to find local deformations in textured areas. Such a collaboration between features and correlation has already been suggested by many authors for 2D image analysis, for example in [2] for stereo image matching. This is why we concentrate now on the design of an intensity-based non-rigid matching, which is the main subject of the present paper.

3 Deformable models: demons versus attractors

Among the varieties of deformable model techniques, we propose to distinguish between two kinds of methods: the *attractor*-based deformable models and the *demon*-based deformable models. This distinction, along with the presentation of an archetypical demon-based deformable model, is the main point of the present paper.

3.1 Attractors

The most intuitive methods are attractor-based, which explains why they are the most widespread. The deformable model is a contour sampled into a given number of points (see figure 1), but it might be also a regular grid, covering the whole image. The sampling points are attracted by points in the scene (for example contour points).

The pulling forces can be derived simply as a function of the distance to the closest point in the scene, such as in Iterative Closest Point (ICP) methods (see [5], [21]). The forces can also be derived from more evolved similarity measures such as similar curvatures (see [10], [4]), or, in the case of correlation techniques, similar intensities or textures. A general problem is how to reduce the algorithmic complexity of finding correspondences.

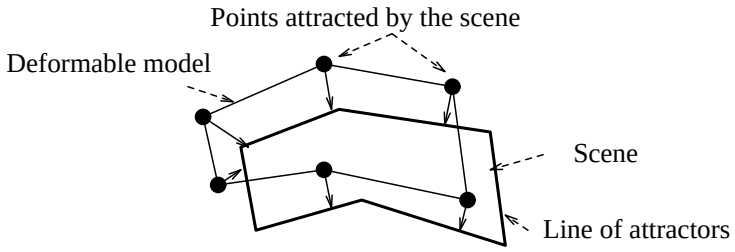

Figure 1: Deformable model with attractors

3.2 Demons

Let us now present the point of view of demons. We will suppose that we can determine the object boundary points, as in the case of attractors, but also the local orientation (inside-outside) of the contour for each point, both for the deformable model and for the scene. Those assumptions are easily fulfilled in the case of medical images, where an iso-contour generally suffices to define an object boundary and where the gradient defines the normal to that contour.

This time, it is the scene object contour which is sampled into a given number of points, with the associated orientation, as presented in figure 2. We will see later that alternatively we can sample the whole scene image with a regular grid of points, instead of sampling a single contour. A demon is associated to each of those contour points, in a sense very similar to Maxwell's demons[2].

Each demon acts locally, to push the deformable model in the normal direction of the scene contour, but the orientation of the push depends on the nature (inside or outside) of the current estimate of the model at that point.

[2]Maxwell's demons were introduced in physics to model the paradoxical idea that a semi-permeable membrane could violate the second law of thermodynamics (about entropy). Situated in the membrane, those demons are supposed to sort locally two different types of molecules of a mixed gas, to put one type on one side and the other type on the other side of the membrane, hence giving a decrease in entropy. As each demon needs to consume some energy to sort, the global entropy of the system is still increased, which suppresses the paradox.

If the point is inside the model, the push is inward, and if the point is outside, the push is outward.

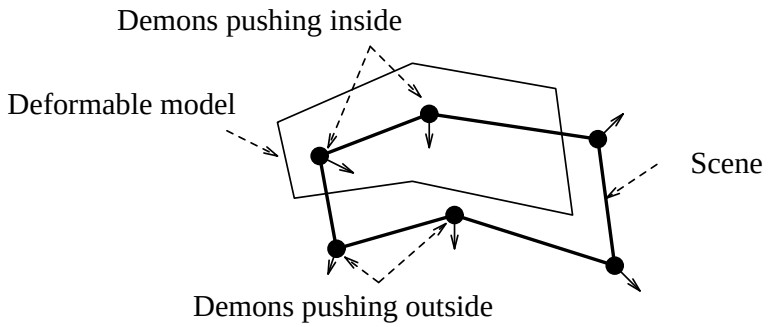

Figure 2: Deformable model with demons

Intuitively, this tends to push the content of the model inside the object shape, and reject the background of the model image outside the shape. In other words, this tends to sort inside and outside points of the model, with respect to the scene object interface, in the same way that Maxwell's demons sort two types of molecules of a mixed gas with respect to a semi-permeable membrane.

Figure 3 presents three iterations of a standard attractor-based method (uper line) and of the demon-based method (lower line) applied to the rigid matching of two disks. For the attractor case, forces originate from the model boundary (the moving disk), and are directed toward the closest point of the fix disk. For the demon case, forces originate from the scene boundary (the static disk), and are directed outward or inward this disk, depending on whether the corresponding points are inside or outside the model disk. For both methods, the force magnitudes decrease regularly. We note that the force directions are different, showing that the two methods are really different, and that this difference is not due to a change of the system of reference. We note also that the attractor method necessitates the computation of closest points, which is not required by the demon method, and that the demon method necessitates the estimation of the inside and the ouside, which is not required by the attractor method.

Figure 4 are three other iterations, and shows that the behaviors of the two methods become similar (in that particular case) when coming close to the final solution.

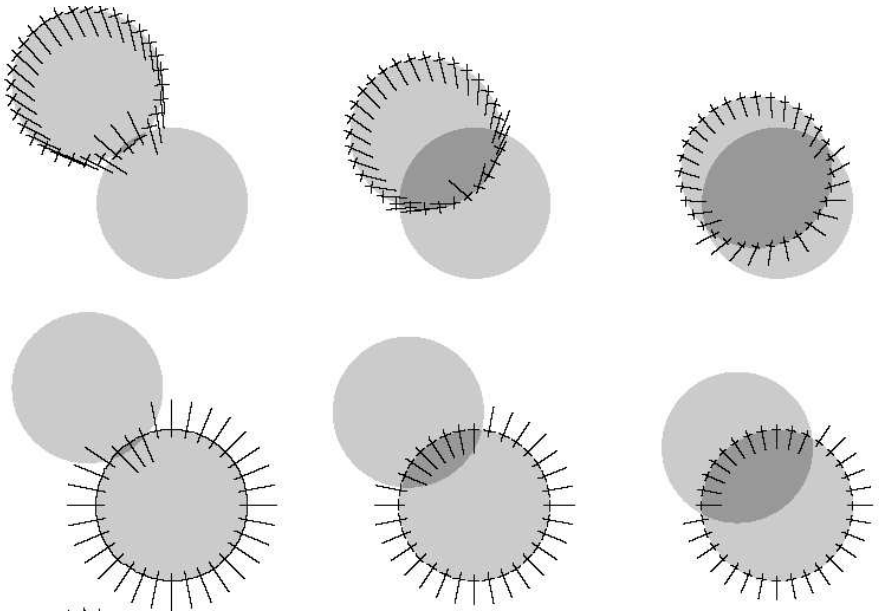

Figure 3: Three iterations of the attractor-based method (upper line) and of the demon-based method (lower line).

3.3 3D grids of demons

For medical images, iso-intensity contours are closely related to the shapes of the objects, because intensity represents density. Let f (resp. g) be the (3D) image intensity function in the scene (resp. in the model). We associate a demon to each voxel P of the scene image where the gradient norm $|\vec{\nabla} f|$ is not null: in that case, an iso-surface comes through that voxel P, whose implicit equation is $f = f(P)$, and whose oriented normal is $\vec{\nabla} f(P)$. The demon in P pushes the model image according to $\vec{\nabla} f(P)$ if $f(P) < g(P)$ and according to

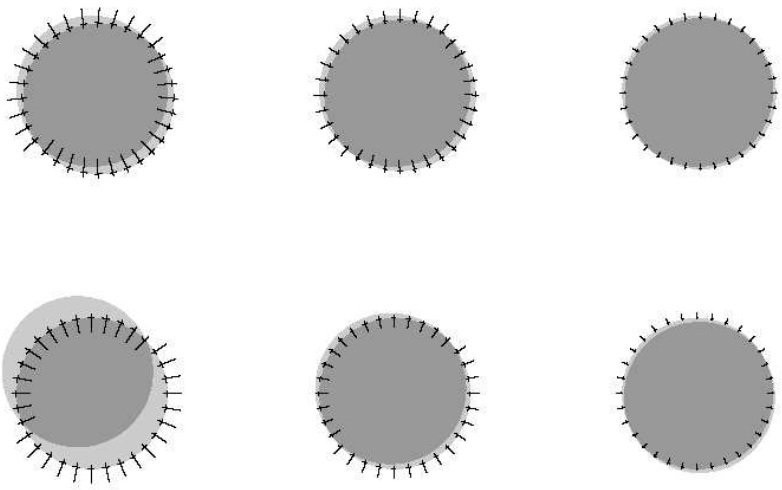

Figure 4: Another set of iterations of the attractor-based method (upper line) and of the demon-based method (lower line).

$-\vec{\nabla} f(P)$ if $f(P) > g(P)$ (see figure 5). Hence a whole 3D grid of demons acts to deform the model.

4 A simple implementation of a demon-based deformable model

We present now an iterative algorithm to perform the non-rigid matching, which can be considered to be an archetype for demon-based methods. We suggest also several opened slots in the basic scheme, which are:

- the deformation function
- the resampling method

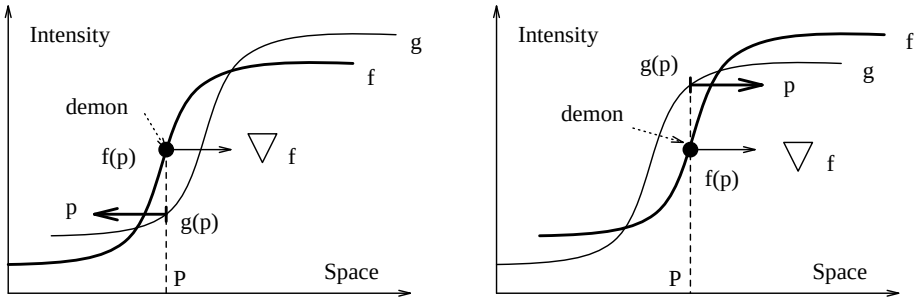

Figure 5: The model g is pushed ($\vec{p}$) by the demon according to $\vec{\nabla} f(P)$ if $f(P) < g(P)$ and according to $-\vec{\nabla} f(P)$ if $f(P) > g(P)$

- the expression of the demon force

Those slots can be filled with many different functions, which gives more flexibility to our algorithm. For each slot, we also indicate the specific choice that we have made for the implementation used in our current experiments, which was made to increase speed: when millions of voxels are to be processed, algorithms must generally be simple to be reasonably efficient, and/or easily parallelizable.

4.1 An iterative algorithm

We present here the case of a (3D) regular grid deformable model. We start from two (3D) images to be matched: the scene image f and the model image g.

At each iteration i of the process, g_i is the image g, deformed by the current transform T_i. We start with an initial transform $T_0 = identity$, and we look for the transform T_n which makes g_n most similar to f, n being the final iteration. Each iteration i is decomposed as follows:

- 1. For each demon P in f, compute the pushing force of the demon according to the local shape of g_i at P, which is g at $P_i = T_i^{-1}(P)$ and the local shape of f at P.

- 2. Compute an elementary deformation δ_T for the whole image, from the vector field of the pushing forces (or δ_T^{-1} from the reversed forces).
- 3. Apply the elementary deformation δ_T to obtain the new total deformation: $T_{i+1} = \delta_T \circ T_i$ (or $T_{i+1}^{-1} = T_i^{-1} \circ \delta_T^{-1}$).

Generally, it can be difficult to invert those transforms. In that case, it is easer to manipulate only the inverted transforms. If T_n is needed instead of T_n^{-1}, this can be done simply by exchanging f and g in the computation.

4.2 Several classes of deformation

Many types of transforms T can be used within this framework, ranging from rigid, affine, quadratic transforms to splines or totally free-form deformations. For example, a single rigid transform δT can be estimated at each step with a least square using all the vectors of the force field, with a quaternion representation of rigid transforms (see [5]). The choice of the model of deformation is one way to impose the rigidity constraint (or internal force) of the deformable model. On the other hand, the most permissive set of transforms is to store, for each voxel of the scene image f, a 3D floating-point vector that gives the corresponding position in g. For that totally free form case, a low pass filter of the displacement vector field (for example a Gaussian filtering, with a given σ) can be used to control the deformability of the model. We use this type of transform for the results presented here.

One alternative to that scheme is to merge step 2 and 3 in a single step, that is, to compute T_{i+1} from T_i without an intermediate δ_T computation. This can be useful in the case of transform classes with a locally-varying deformability. It is also useful for splines, if we want to end up with a single spline warp instead of a complicated composition of n spline functions.

4.3 Several types of resampling algorithms

$P_i = T_i^{-1}(P)$ is a floating point vector, but an image g is a sampled function. We can estimate $g(P_i)$ with tri-linear interpolation, as in our current implementation, but it can also be made more precise, for example by using a tri-spline interpolation.

This improvement may turn out to be very interesting for angiography or any functional imagery based on image subtraction. Small non-rigid deformations can exist between the images to subtract, due for example to the patient breathing. After correction of the deformation, the limitation of the performance is the smoothing of the deformed image induced by the tri-linear interpolation. However, this improvement is achieved to the detriment of speed, with a difference of more than one order of magnitude.

4.4 Several types of demons

So far, we left unstated the precise design of a demon. The general characteristic is that "it pushes the model outward when it is outside the model, and inward when it is inside".

In the case of the experiments presented here, we use the instantaneous optical flow equation, as presented in [12] for the 2D case. The hypothesis is that there is a conservation of the intensity of points under motion $i(x(t), y(t), z(t), t) = const.$ Differentiating this equation gives:

$$\frac{\partial i}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial i}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial i}{\partial z}\frac{\partial z}{\partial t} = -\frac{\partial i}{\partial t} \tag{1}$$

In our case, we have only two frames f and g to compare, and we are looking for a motion $\vec{v}$ which brings g closer to f, thus we consider that f and g are separated by one unit of time: $\partial i/\partial t = f - g$, an $\vec{v} = (dx/dt, dy/dt, dz/dt)$ is the instantaneous velocity from g to f, we have therefore:

$$\vec{v} \cdot \vec{\nabla} f = g - f \tag{2}$$

This equation, however, does not suffice to compute $\vec{v}$ locally, which is generally determined by a global regularization technique. Determining $\vec{v}$ can be made local in the following way (see figure 6). A first order approximation of the image intensity surface f at P is an hyperplane through $(P, f(P))$ with normal $(-\vec{\nabla} f(P), 1)$. The equation $\vec{v} \cdot \vec{\nabla} f = g - f$ states that the end point of $\vec{v}$ has to be at the intersection of this hyperplane with an horizontal hyperplane through $(P, g(P))$. Without any other information, it is natural to choose for $\vec{v}$ the closest point to P, whose equation is:

$$\vec{v} = \frac{(g-f)\vec{\nabla} f}{\vec{\nabla} f^2} \tag{3}$$

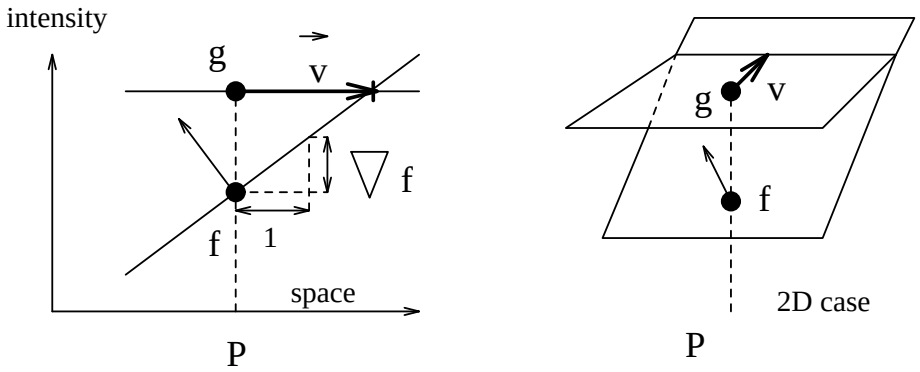

Figure 6: Instantaneous velocity from g to f

Intuitively, this is the shortest (spatial) translation that puts point $g(P)$ into the hyperplane $(P, f(P))$ with normal $(-\vec{\nabla} f, 1)$, that is, image g into accordance locally with image f. For our demons, we could take as pushing force p this velocity $\vec{v}$. This force pushes inward, that is, in the direction of $\vec{\nabla} f$ when $f < g$, and outward, that is, toward $-\vec{\nabla} f$, when $f > g$, which is the desired behavior of a demon.

However, this equation is unstable when the gradient norm is small. In that case, a small variation of the intensity can reject the end point of $\vec{v}$ to infinity in any direction, and it would be quite undesirable to have a system that can create almost infinite forces with random directions. Ideally, the forces should be given by equation 3 when the gradient is high, and should be close to zero when the gradient is low. For example, this can be achieved by multiplying the expression of equation 3 with $(\vec{\nabla} f^2/(\vec{\nabla} f^2 + k^2))$ where k is a constant depending on the contrast of the image. In fact, in order to avoid the confusing effect of a new parameter in the experiments, we chose to use $g - f$ instead of k, which is also a way of limiting the forces to take finite values.

Hence the pushing force $\vec{p}$ of the demon used in our experiments is given by 4:

$$\vec{p} = \frac{(g-f)\vec{\nabla} f}{\vec{\nabla} f^2 + (g-f)^2} \tag{4}$$

We note that this expression doesn't depend on the normalization of the intensities: multiplying those intensities by a constant doesn't change the value of $\vec{p}$. Of course, many other types of demon can be devised, but this one is especially easy to compute, which makes our algorithm very fast.

4.5 Some more evolved demons

In some places, there might exist isolated structures in the model image g that do not exist in the scene image f. The results are a large value of $|g-f|$ and a small value of $|\vec{\nabla} f|$, leading to local useless compression or dilation of the model. This phenomenon is limited by the expression of $\vec{p}$ itself (equation 4), where $\vec{p}$ tends toward 0 when $|\vec{\nabla} f|$ is small with respect to $|g-f|$.

Conversely, when an isolated structure exists in the scene image f, but not in the model image g, $|\vec{\nabla} g|$ is small, but both $|g-f|$ and $|\vec{\nabla} f|$ remain high, creating this time noticeable compression or dilation forces. This asymmetrical behavior comes from the asymmetry of equation 4 with respect to f and g. To restore the symmetry, we can replace $\vec{\nabla} f$ with the averaged gradient $(\vec{\nabla} f + \vec{\nabla} g)/2$, which gives a new formula for the motion $\vec{v_2}$:

$$\vec{v_2} = \frac{2(g-f)(\vec{\nabla} f + \vec{\nabla} g)}{(\vec{\nabla} f + \vec{\nabla} g)^2} \tag{5}$$

However, this expression tends toward 0 only when $|\vec{\nabla} f|$ and $|\vec{\nabla} g|$ are small simultaneously. We must therefore add a constraint that makes the expression for the force tend toward zero when one of the two gradient norms tends to 0, for example by multiplying the expression for $\vec{v_2}$ with a similarity coefficient s_2. We can set s_2 to 0 when $\vec{\nabla} f \cdot \vec{\nabla} g \leq 0$, and use the expression of equation 6 when $\vec{\nabla} f \cdot \vec{\nabla} g > 0$ ($\cdot$ is the scalar product):

$$s_2 = \frac{2\vec{\nabla} f \cdot \vec{\nabla} g}{\vec{\nabla} f^2 + \vec{\nabla} g^2} \tag{6}$$

This expression tends to zero when one of the gradient norms tends to zero, but also when the two gradient directions are too much dissimilar, which is an interesting behavior.

We must note, however, that these expressions require the computation of $\vec{\nabla} g$ for each point, at each iteration. This means that it is either necessary to precompute and store $\vec{\nabla} g$ for each point in g, which requires a large memory, or to recompute $\vec{\nabla} g$ at each iteration, which is time consuming. It also means that it is necessary to evaluate $\vec{\nabla} g$ for floating point positions, using for example tri-linear interpolation of the gradient vectors, which is even more expensive. Finally, if the transforms T_i induce a global or local rotation of the model, we have to find a way to apply the vectorial component of the transform to the gradient vectors $\vec{\nabla} g$, which is very difficult to realize if the transform is not rigid. It seems therefore difficult to use the vector $\vec{\nabla} g$ in the expression of the demon forces without losing one order of magnitude in the computation speed.

A more acceptable solution is to use only the gradient norm $|\vec{\nabla} g|$. In that case, we can reformulate the expression of the velocity into 7, and the expression of the similarity of the gradient norms into 8:

$$\vec{v_3} = \frac{2(g-f)\vec{\nabla} f}{\vec{\nabla} f^2 + \vec{\nabla} g^2} \tag{7}$$

$$s_3 = \frac{2|\vec{\nabla} f||\vec{\nabla} g|}{\vec{\nabla} f^2 + \vec{\nabla} g^2} \tag{8}$$

By combining those two expressions and by introducing a factor $k = g - f$ to limit the forces for small values of the gradient norms, we can propose a new expression of the demon force $\vec{p_3}$ defined by equation 9, which requires only about twice the computation time of $\vec{p}$, yet still limits the deforming effects of dissimilar structures.

$$\vec{p_3} = \frac{4(g-f)\vec{\nabla} f|\vec{\nabla} f||\vec{\nabla} g|}{(\vec{\nabla} f^2 + \vec{\nabla} g^2)(\vec{\nabla} f^2 + \vec{\nabla} g^2 + 2(g-f)^2)} \tag{9}$$

We can verify that when $|\vec{\nabla} f| = |\vec{\nabla} g|$, we have $\vec{p_3} = \vec{p}$. As we can precompute an image of $|\vec{\nabla} g|$, and as $|\vec{\nabla} g|$ are simple floating point values, the computation time using $\vec{p_3}$ is only about twice the computation time using $\vec{p}$.

This illustrates some of the many possible expressions for the demon forces, which can be arbitrarily complicated. It shows also the importance of defining first a general qualitative behavior of the demons, as presented in the previous section, before proposing examples of implementation, which can depend very much on the application.

4.6 A multiscale implementation

There are several advantages in using a multiscale implementation. By multiscale, we mean considering first the matching of undersampled images and then refining the solution from coarse-to-fine:

- it speeds up the computation. We can make many iterations at coarse scales during the computation time required by a single iteration at the finest scale.
- it speeds up the convergence. If the object has to be move from n pixels at the finest scale (scale 0), the number of iterations which are needed is about n, whereas at scale $Log_2(n)$, the displacement is a single voxel, which can be done within a single iteration.
- for human anatomy, macroscopic features are generally more stable than microscopic features, which justifies the use of coarse-to-fine schemes in general, for those kind of applications.

To conclude, a multiscale implementation is very well suited for the demon-based deformable model, because it makes the local influence of the demons much less local. Our current implementation is multiscale: we derive two image pyramids from the two images to register, up to a fixed scale, and then apply a set of iterations of the demon-based algorithm to the coarsest scale. Each final solution at one scale serves to initialise the matching at a finer scale. The number of iterations also decreases from coarse to fine, iterations at coarse scales being cheaper.

5 Comparisons between attractor and demon deformable models

Demon-Based deformable models share many features of attractor-based deformable models. They both search for a balance between internal and external forces. The deformable model is submitted to pushing forces (whereas in the case of attractors, it is pulling forces). It is displaced and deformed by those forces, but also tries to maintain an internal coherence.

As we saw, demon-based models can support as many classes of deformations as attractor-based models, ranging from rigid to totally free-form deformations, and there is also a deformability parameter: it can be the choice of the deformation class (rigid, affine) or it can be defined explicitly by a rigidity parameter (for example the σ of a Gaussian smoothing filter).

For rigid matching methods, using a sparse set of corresponding points, we are able to establish an objective quality of the registration, by way of the estimation of a covariance matrix associated with that transform (see [16]). It seems difficult to get similar results for non-rigid matching without a clearer definition of the problem.

5.1 A first classification into demon-based and attractor-based methods

Our method is certainly not the only one which is demon-based. However, because this method is much less intuitive than the attractor-based one, it is much less widespread. We describe here only a set of methods which are surely attractor-based. We hope that to distinguish between demons and attractors will help to have a better understanding of the inner mechanism of other matching methods.

As we said, methods using the distance to the closest point, such as [5] or [21], and illustrated in the experiments of figures 3 and 4, are definitely attractor-based. To reduce the complexity of finding corresponding points into a large set of possible correspondents, some elegant methods have been proposed to give fast solutions, such as using hierarchical data structures (the KD trees in [10], the octree splines in [18]), using adaptive grids (see [4]), choosing to process only a small subset of points (the crest lines of [9]), or to pre-compute

a distance map, such as in Chamfer matching methods (see [14] for the rigid case). Some kind of segmentation is always required in those methods, which is a big problem when trying to make the method fully automatic.

In the case of distance maps, a potential field can be pre-computed, and derived to give the forces at each point in the image, with only a local computation (similar to a gradient descent technique, searching for local extrema). Some methods, such as the original snakes technique of [13], use directly the gradient to indicate the direction to the closest contour point: those methods are still attractor based, forces being an approximation of the forces based on the actual closest point.

Correlation techniques are also attractor-based. The simpler implementation is an extensive search of the most correlated point within a given neighborhood. Each point of the model is attracted by the point in the scene which maximize the correlation function. For a given point P of the model g, the correlation function could be applied to the whole scene image, resulting in a potential field specific to $g(P)$ (which is of course never computed in practice). This potential field can also be derived to get a local estimation of the direction to the most correlated point, such as proposed in [1], which can be therefore classified as attractor-based.

The distinction between inside and outside, which is an essential feature of the demon-based technique, is not used in those methods.

5.2 Some current restrictions

There are two major restrictions of the method that we propose, which are shared by the majority of non-rigid matching techniques. The first is that the original positioning of the two objects to be registered is crucial for the success of the method; the second is that the intensities of two images must be similar.

For attractor-based methods, an initial snake placed outside the object can crush against its outer boundary, whereas, in the case of demons, the pushing forces will compensate to give no noticeable effects between the object and the model, but a local compression or dilation of the model. A partial solution to this local minima avoidance problem is that the rigidity of the model generally helps to overcome this situation. The multiscale approach is also another generic way to avoid local minima.

The second problem is more typical of inter-modality registration. When the responses according to the tissue types are different between the two acquisition devices, the simple implementation that we have presented here will not give "good" results. A general solution is to make the intensities of two images similar first.

Knowing the correspondence between the two device response functions may help to perform such a correction, but generally the correspondence is not bijective; several tissues having the same response in one modality and different responses in the other. In that case, it can be convenient to match the two images I_1 and I_2 with an intermediate atlas image J, where the tissue types are everywhere known (see figure 7).

With the two response functions, it is possible to derive from the atlas two superimposable pseudo-images J_1 and J_2, J_1 (resp. J_2) having intensities similar with I_1 (resp. I_2). I_1 is matched with J_1 and I_2 is matched with J_2 and as J_1 and J_2 are superimposable, we get the matching between I_1 and I_2. What is supposed here is that we know precisely what organ is represented in the two images (for example two heads), and also the exact response function of each acquisition device.

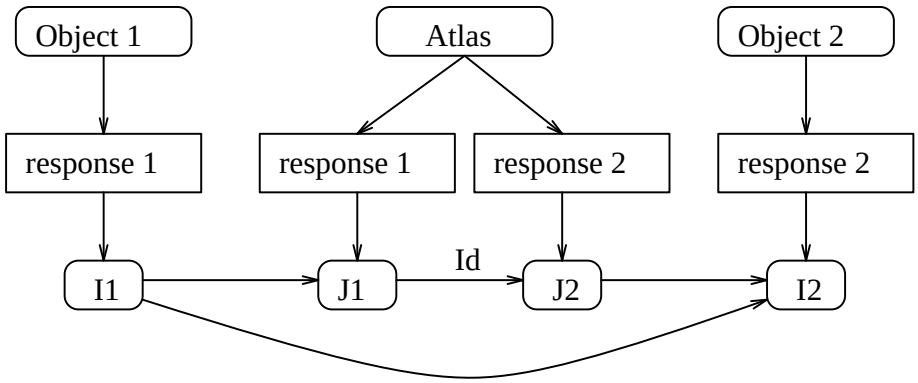

Figure 7: Non-rigid Matching of two images acquired with two different devices through an atlas

6 Experiments

6.1 The parameters of our current implementation

Our non-rigid matching software takes two 3D images as input, and generates as output either one image resampled to be superimposable to the other 3D image, or a description of the transform. For totally free-form deformations, the output is a 3D grid of 3D vectors, giving for a voxel in one image the corresponding floating point position in the other image. To summarize the choices made for the implementation used in the experiments:

- the transform is totally free-form
- the resampling is a tri-linear interpolation
- the demon force is given by the equation 4

The parameters which can be tuned are the following:

- n: the number of iterations at the finest scale
- l: the number of levels of the image pyramid which are used for the multiscale computation. If l is zero, the multiscale scheme is not used.
- σ: the rigidity of the deformable model.

In all our experiments to date, σ has been fixed to the same default value, l is used basically to speed up the computation (generally we use 2 or 3 pyramid levels) and n is a compromise between the computation time and the level of convergence. We generally used 4 iterations at the finest scale, but in some cases, we have also increased the number of iterations to test the effects of a better convergence.

6.2 3D image sequences displaying cardiac motions

The two major drawbacks of non-rigid matching are very reduced for image sequence analysis. The acquisition device being the same, the intensities of the

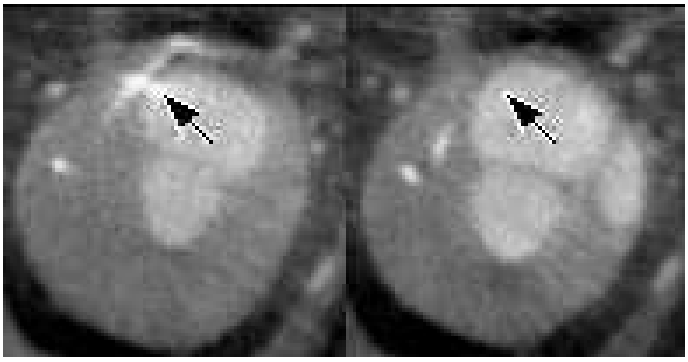

Figure 8: Corresponding diastolic and systolic slice without matching (dog)

scanned object do not vary much from frame to frame, and for small moves: the position of the object in one frame is a good initialization for the next one.

In the first experiment, we present an example of the analysis of a CT scan of the heart-beat of a dog, acquired with the Mayo Clinic DSR (courtesy of Dr. Richard Robb). To consider extreme conditions, we tried successfully to match directly the diastolic and systolic 3D images. In figure 8, we present two corresponding slices before elastic matching, and in figure 9, the same slices after 3D matching and resampling. With 3D tools, we verified the quality of the match in all the other parts of the images. To get an idea of the smoothness of the non-rigid transform, we artificially added a regular grid tag to the image before deformation (see figure 10). The computation time, for the matching of the two 100^3 voxels images is about 5 minutes on a DEC alpha workstation.

Figures 11 and 12 present the same kind of experiments for nuclear medicine 64^3 SPECT images (a human heart), with a CPU time of about 1 minute. Figure 13 is a 3D visualization of the same results.

For the analysis of the whole cardiac sequence (eight 3D images), and according to the physics of the SPECT acquisition, we have added two constraints: the cyclicity of the sequence, which is imposed by another set of iteration taking the whole sequence into account and performing temporal filtering, and the conservation of the radioactive matter. This last constraint states that, when the myocardium contracts, the density of radioactive material is increa-

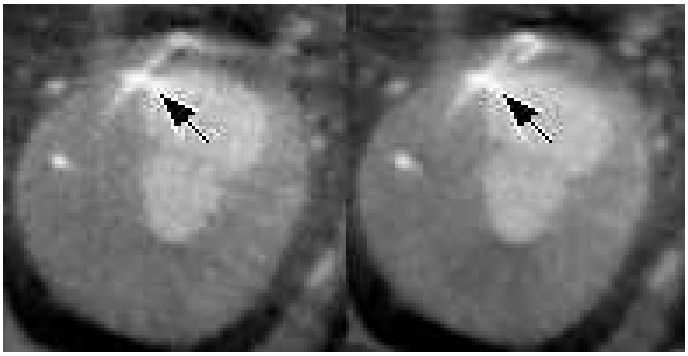

Figure 9: Corresponding diastolic and systolic slices after 3D matching and resampling

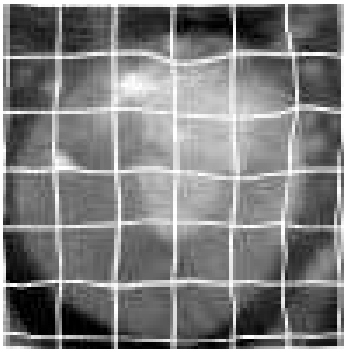

Figure 10: Artificially tagged image to show the deformation

sed, and thus the corresponding image intensity is enhanced. This can be easely modeled within the demons design, by multiplying $g(P_i)$ by the Jacobian of the transform, which represents the local variation of the density. An approximation of the Jacobian is to use the ratio of an elementary volume before and after transform, which is easy to compute when we have the displacement

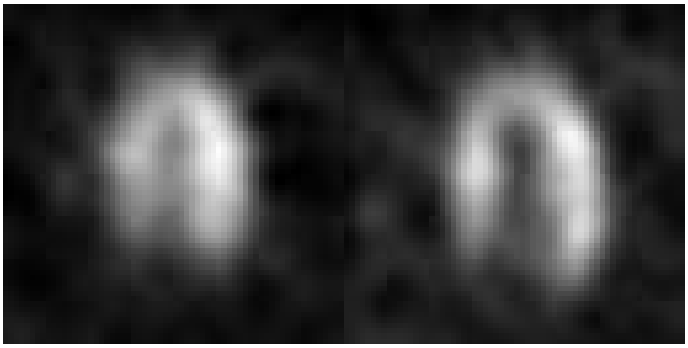

Figure 11: Corresponding diastolic and systolic slice without matching (SPECT,human)

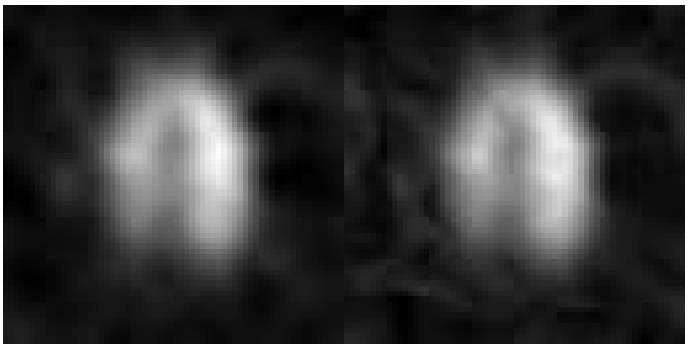

Figure 12: Corresponding diastolic and systolic slice after 3D matching and resampling (SPECT,human)

vector at each node of the grid. The analysis of the whole sequence takes only about 10 minutes.

To conclude, we have an automatic, fast algorithm to estimate a deformation field between frames in 3D image sequences, which does not necessitate a segmentation of the object, which is especially fortunate for SPECT images. A good point is that the physics of nuclear medicine acquisition can be easily

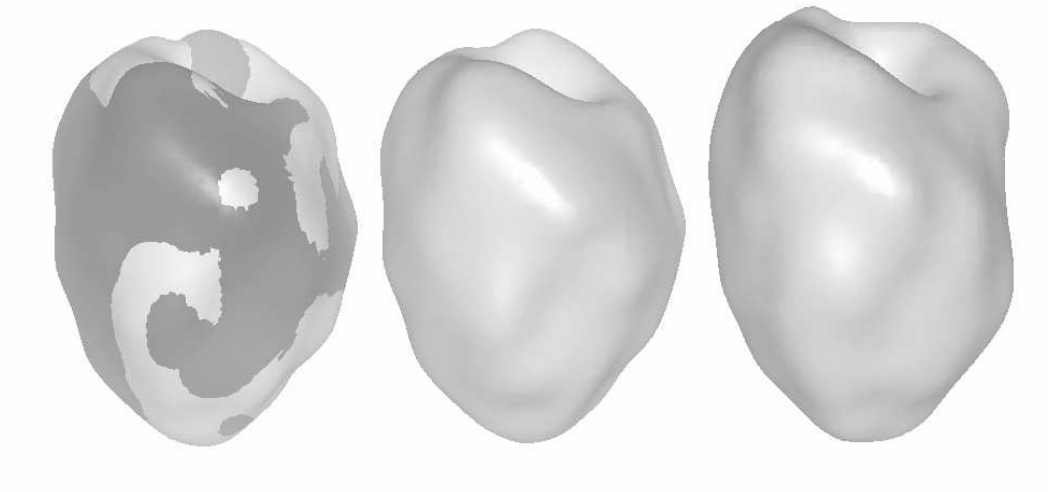

Figure 13: 3D visualization of the SPECT image of the heart. Right is the diastole, middle is the systole, left is the surface of the diastole deformed by the 3D vector field and superimposed to the systole, showing a good fit between the two surfaces.

taken into account in our scheme. The challenge now is to extract from this 4D flow field the parameters that precisely characterize a pathology, such as a risk of infarction, which is one of our current studies. For that purpose, we have started to apply modal decomposition techniques, such as described in [15].

6.3 Inter-Patient matching

A totally different problem is inter-patient registration. It took more than 10 hours by a skilled specialist, with up to date interactive graphic techniques, and 3D image processing tools such as mathematical morphology, to label the voxels of the MR image of a human brain (see figure 14, a $256 \times 256 \times 123$ voxels image).

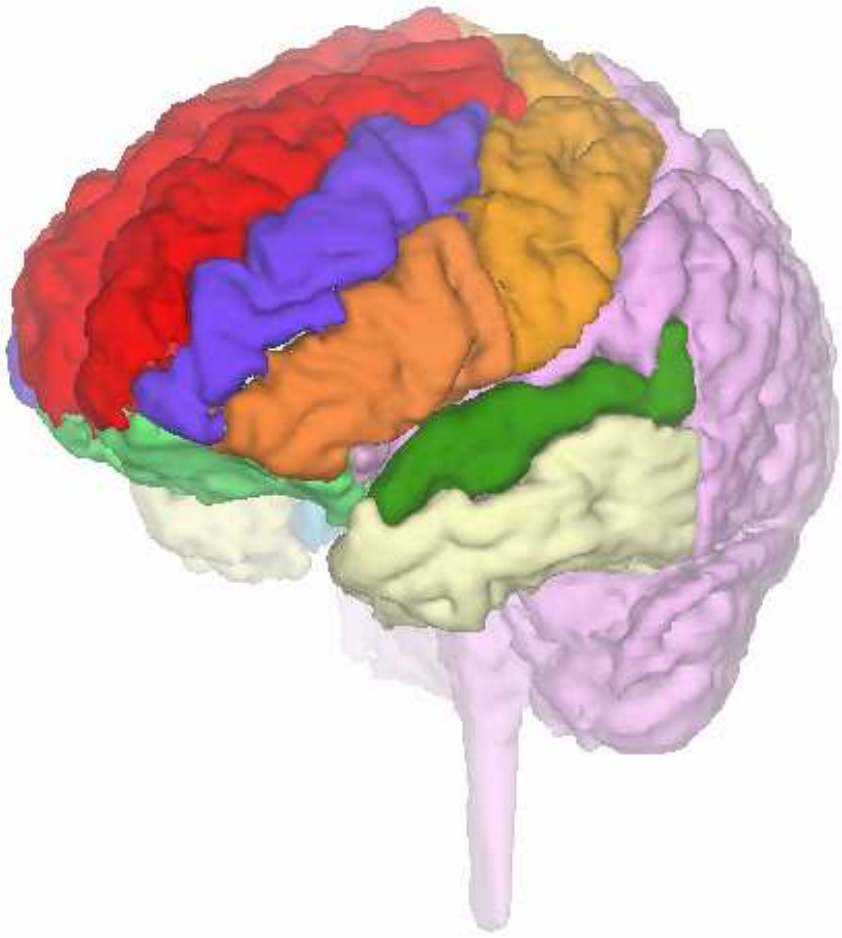

Figure 14: A 3D display of the segmented image of the brain, courtesy of Dr Ron Kikinis, and rendered by Gerard Subsol.

As dozens of such labeled images are needed to build a coarse numerical atlas, as each image still needs to be put into correspondence with the other cases to estimate variability measures, and as there is also a need to put

any patient into correspondence with the final atlas for diagnosis or surgical planing purposes, one might easily understand the strategic importance of a fully automatic, totally free-form elastic matching method, even if anatomical matching is not a well defined problem.

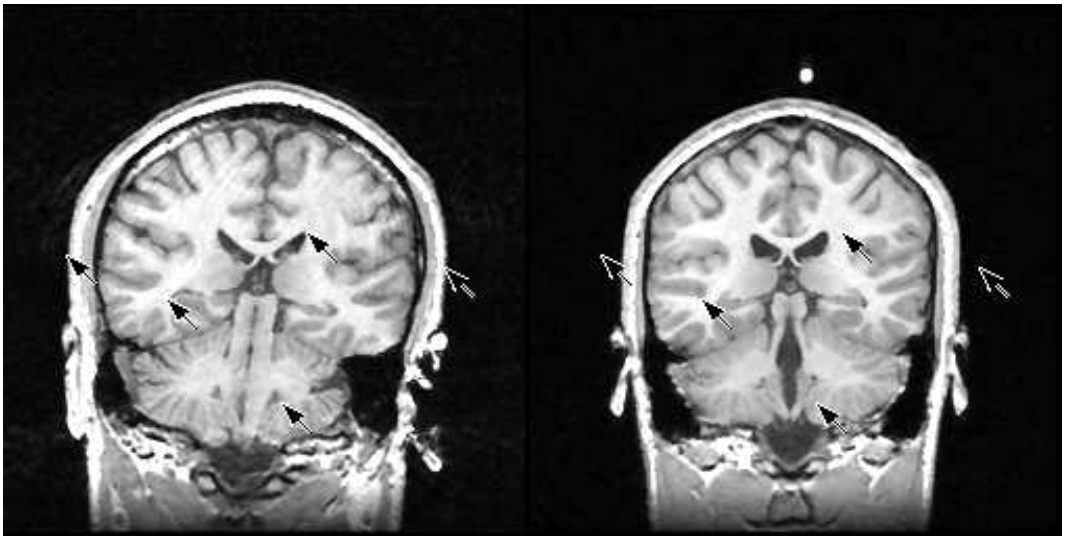

Figure 15: Corresponding MR slices of two different patients, without matching, with corresponding points (arrows)

In figure 15, we present corresponding slices, extracted from two 3D MR scans of two different patients, without matching. Corresponding points, marked with arrows, show the geometrical differences between the two patients (note for example the difference in shape of the ventricles, or that the right patient has a much smaller head than the left one). Figure 16 presents the same slices after 3D elastic matching and resampling. The images are so similar that it is hard to tell if these are two images of the same patient, or of two different patients, however, the underlying deformation is rather smooth. This means that, although the brain may present large geometrical variations between individuals, the macroscopic topological structures are mostly similar.

In fact, those results are qualitatively very similar with the results of elastic matching techniques, presented in [1] or [6]. What is original here is that we achieved those results with only 20 minutes CPU time on a conventional DEC

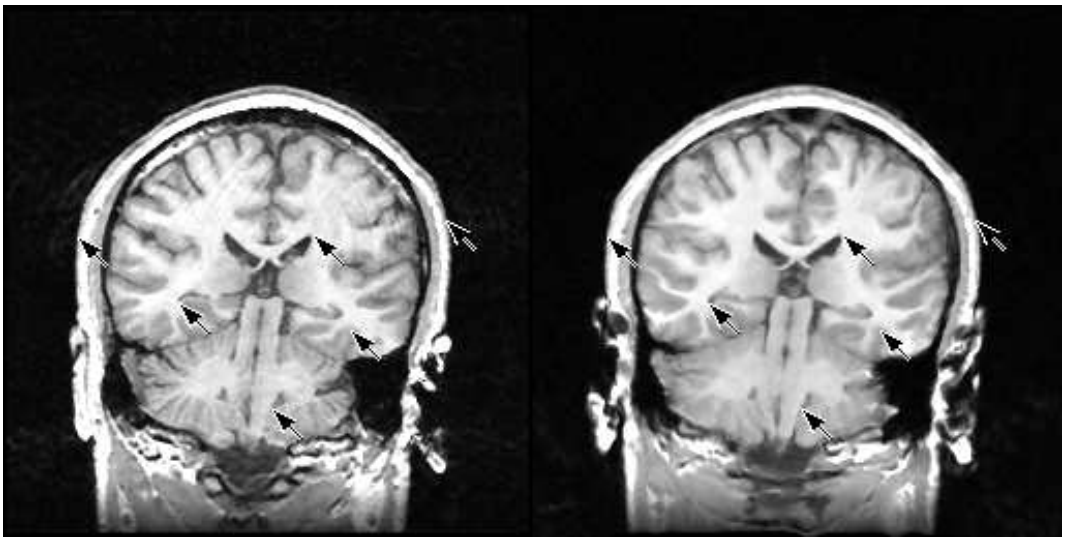

Figure 16: Corresponding MR slices of the two different patients, after matching, with corresponding points (arrows)

Alpha workstation, that is, one or two order of magnitude faster than what was previously reported (even more if compared with 9 hours of a MASPAR). Of course, qualitative similarities does not mean quantitative ones, and the methods presented in [1] or [6] might give better results than ours; however, such a quantitative comparison is hard to perform, because we still need a much clearer definition of non-rigid matching to be able to derive any type of quality measure of the matches. A first step in that direction can be found in [11], where what is proposed is to measure the percentage of overlap of manually segmented structures before and after matching.

We have started a quantitative evaluation of our matching with similar tools. Along with the right patient MR image I_1, we get a complete segmentation A_1 of the brain (courtesy of Dr Kikinis, Brigham and Women's Hospital, Boston). Using the same deformation field as obtained between the two MR images, we deformed A_1 into A_2, by using the deformation $I_1 \rightarrow I_2$ (one slice is presented figure 17 for (I_1, A_1) and (I_2, A_2)), which is the achievement of an automatic labeling of a 3D image.

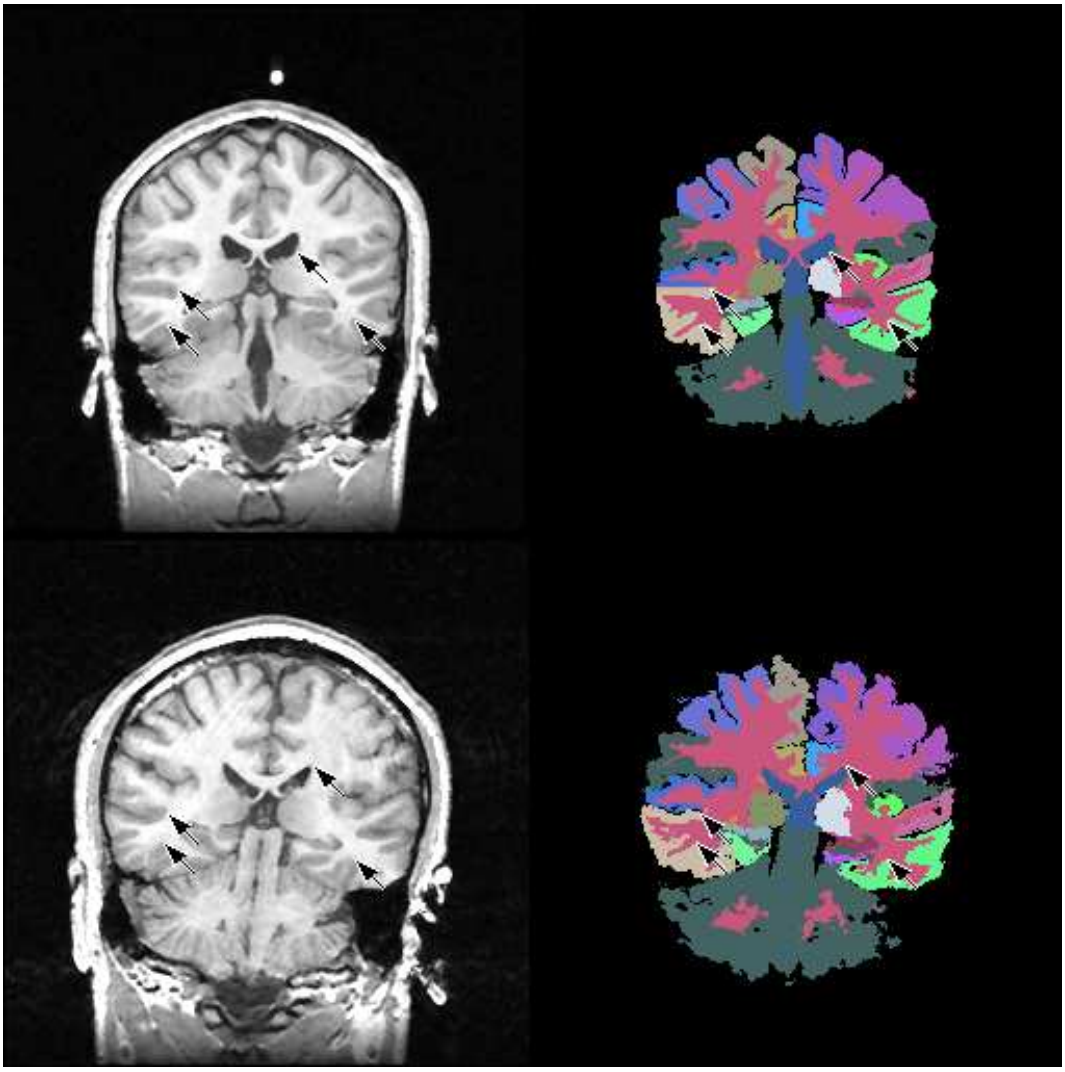

Figure 17: Up left, first patient (I_1), up right, the corresponding segmented image (A_1), down left (I_2) second patient, and down right the segmented image A_1 deformed into A_2 with the transform $I_1 \rightarrow I_2$

We then performed the matching the other way round, transforming back the segmented image A_2 into a segmented image A'_1, using the deformation

$I_2 \rightarrow I_1$. I_1' should be exactly superimposable to I_1, if the matching was perfect. In table 18 we present the percentage of points having a different label in both images (a slice of A_1 and A_1' is presented in figure 19). The statistics for the different structures of the brain, between A_1 and A_2, reflect the deformation or mis-alignment between the two patients. The results for the comparison of A_1 and A_1' should be ideally 0% of differences, but is about 25% for the external structures of the brain, and down to about 10% for internal structures such as the amygdala or the hippocampus. Note that many of those structures, such as the different gyri, cannot be distinguished by their image intensities, but only by their relative positioning, which explains partly why those results are not closer to zero.

In the future, we hope to improve the automatic labeling with a more local, final pass of segmentation, taking the tissue types into account, such as presented by Wells et al in [20]. We intend also to compare the automatically labeled images with several images segmented manually and independently, to get a better evaluation of our method.

7 Conclusion

In this paper, we have presented a very fast algorithm to perform automatically the non-rigid matching of 3D medical images. We have devised our method by identifying an original distinction between attractor-based and demon-based deformable model techniques (ours being demon-based). We intend to explore now the collaboration between feature based and intensity based matching to given a better anatomical justification of non-rigid matching in the case of inter-patient registration.

Acknowledgment

I wish to thank Nicholas Ayache, Michael Brady, Morten Bro Nielsen, Jérôme Declerck, Gérard Subsol, Xavier Pennec, and the Epidaure people in general for stimulating discussions and specific comments about the ideas presented in this paper. Part of this study has been supported by the Esprit Basic Research Action VIVA, and an External research program with Digital Equip-

Brain structure	left A_1, A_2	left A_1, A_1'	right A_1, A_2	right A_1, A_1'
White matter (all)	53.3 %	21.2 %		
Gray matter (all)	46.9 %	17.7 %		
Ventricles (all)	61.6 %	13.7 %		
Planum (only left)	88.3 %	24.1 %		
Inf. frontal gyrus	82.2 %	24.6 %	63.3 %	29.4 %
Mid. frontal gyrus	83.9 %	30.9 %	70.9 %	27.8 %
Sup. frontal gyrus	78.8 %	31.5 %	76.8 %	26.7 %
Insular cortex	90.9 %	20.7 %	68.7 %	22.6 %
Cingulate gyrus	69.0 %	22.6 %	70.2 %	16.1 %
Orbital cortex	63.9 %	22.4 %	72.3 %	18.8 %
Central gyrus	69.3 %	34.4 %	70.9 %	39.1 %
Temporal lobe gray	73.7 %	22.0 %	55.8 %	23.9 %
Sup temporal gyrus	92.1 %	20.8 %	82.9 %	27.7 %
Parahippocampal gyrus	69.8 %	18.7 %	35.1 %	18.9 %
Hippocampus	66.2 %	7.4 %	49.0 %	12.3 %
Amygdala	65.6 %	7.6 %	29.9 %	5.6 %
Caudate nucleus	99.3 %	10.4 %	78.0 %	12.1 %
Putamen	86.3 %	7.9 %	47.3 %	5.5 %
Globus palladus	83.7 %	11.8 %	64.3 %	18.2 %
Thalamus	33.2 %	5.7 %	54.2 %	4.9 %

Figure 18: Percentage of the points that do not correspond to the same structure, directly between the two patients I_1 and I_2 (A_1, A_2), and for the same patient I_1, after transformation into I_2 and back into I_1 (A_1, A_1')

ment. Dr Ron Kikinis of the Brigham and Women's Hospital, Harvard Medical School, Boston deserves special thanks for the original and segmented images of the brain.

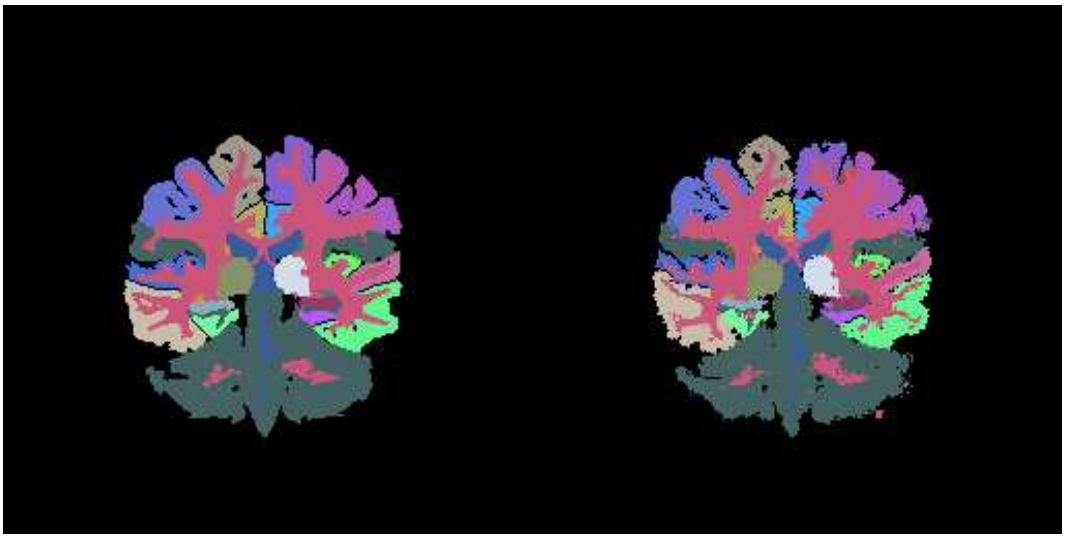

Figure 19: Left is the original segmentation A_1. This segmentation is transformed into A_2 using $I_1 \rightarrow I_2$ and back into A'_1, using $I_2 \rightarrow I_1$. A'_1 is the right image.

References

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

[2] H.H. Baker and T.O. Binford. Depth from edge- and intensity-based stereo. In *7th Joint Conference on Artificial Intelligence*, pages 631–636, August 1981.

[3] E. Bardinet, L.D. Cohen, and N. Ayache. Fitting of iso-surfaces using superquadrics and free-form deformations. In *Proceedings of the IEEE Workshop on Biomedical Images Analysis (WBIA'94)*, Seattle, Washington, June 1994.

[4] Serge Benayoun, Chahab Nastar, and Nicholas Ayache. Dense non-rigid motion estimation in sequences of 3d images using differential constraints. In *Computer Vision, Virtual Reality and Robotics in Medicine*, volume 905 of *Lecture Notes in Computer Science*, pages 309–318, Nice, France, April 1995. Springer.

[5] Paul J. Besl and Neil D. McKay. A method for registration of 3-d shapes. *IEEE PAMI*, 14(2):239–255, February 1992.

[6] Gary E. Christensen, Michael I. Miller, and Michael Vannier. A 3d deformable magnetic resonance textbook based on elasticity. In *AAAI Spring Symposium Series: Applications of Computer Vision in Medical Image Processing*, pages 153–156, Standford University, March 1994.

[7] Gary E. Christensen, Richard D. Rabbitt, and Michael I. Miller. 3d brain mapping using a deformable neuroanatomy. *Physics in Medecine and Biology*, 39:609–618, 1994.

[8] Isaac Cohen, Laurent Cohen, and Nicholas Ayache. Using deformable surfaces to segment 3d images and infer differential structures. *CVGIP : Image understanding '92*, September 1992.

[9] Jérôme Declerck, Gérard Subsol, Jean-Philippe Thirion, and Nicholas Ayache. Automatic retrieval of anatomical structures in 3d medical

images. In *Computer Vision, Virtual Reality and Robotics in Medicine*, volume 905 of *Lecture Notes in Computer Science*, pages 153–162, Nice, France, April 1995. Springer.

[10] J. Feldmar and N. Ayache. Rigid and affine registration of smooth surfaces using differential properties. In *Proceedings of the Third European Conference on Computer Vision 1994 (ECCV'94)*, Stockholm, Sweden, May 1994.

[11] Jim C. Gee, Martin Reivich, and Ruzena Bajcsy. Elastically deforming 3d atlas to match anatomical brain images. *Journal of Computer Assisted Tomography*, 17(2):225–236, March 1993.

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

[13] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1:312–331, 1987.

[14] G. Malandain, S. Fernández-Vidal, and J.M. Rocchisani. Improving registration of 3-D medical images using a mechanical based method. In *Proceedings of the Third European Conference on Computer Vision (ECCV '94)*, pages 131–136, Stockholm, Sweden, May 2–6 1994. Lecture Notes in Computer Science 801.

[15] Chahab Nastar. Vibration Modes for Nonrigid Motion Analysis in 3D Images. In *Proceedings of the Third European Conference on Computer Vision (ECCV '94)*, Stockholm, May 1994.

[16] X. Pennec and J.P. Thirion. Validation of 3-d registration methods based on points and frames. Research Report 2470, INRIA, 1995.

[17] P. Shi, G. Robinson, A. Chakraborty, L. Staib, R. Constable, A. Sinusas, and J. Duncan. A unified framework to assess myocardial function from 4d images. In *Computer Vision, Virtual Reality and Robotics in Medicine*, volume 905 of *Lecture Notes in Computer Science*, pages 327–337, Nice, France, April 1995. Springer.

[18] Richard Szeliski and Stéphane Lavallée. Matching 3-d anatomical surfaces with non-rigid volumetric deformations. In *Proceedings of the IEEE Workshop on Biomedical Images Analysis (WBIA'94)*, Seattle, Washington, June 1994. Also in AAAI 1994 Spring Symposium Series. Application of Computer Vision in Medical Image Processing, Stanford University, 1994.

[19] J-P Thirion. Steps toward the automatic labeling of 3d medical images. In *AAAI symposium: Application of Computer Vision in Medical Image Processing*, pages 18–21, Stanford, March 1994.

[20] W.M. Wells III, W.E.L. Grimson, R. Kikinis, and F.A. Jolesz. Adaptative segmentation of mri data. In *Computer Vision, Virtual Reality and Robotics in Medicine*, volume 905 of *Lecture Notes in Computer Science*, pages 59–69, Nice, France, April 1995. Springer.

[21] Zhengyou Zhang. Iterative point matching for registration of free-form curves. Technical Report 1658, INRIA, April 1992.

Unité de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 46 avenue Félix Viallet, 38031 GRENOBLE Cedex 1
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
ISSN 0249-6399