Understanding the "Demon's Algorithm": 3D Non-Rigid registration by Gradient Descent
Xavier PENNEC, Pascal CACHIER, Nicholas AYACHE
2nd Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI'99)
Pages 597–605, Cambridge, UK – Sept. 19-22, 1999

Understanding the "Demon's Algorithm": 3D Non-Rigid registration by Gradient Descent

Xavier Pennec, Pascal Cachier, and Nicholas Ayache

INRIA Sophia - Epidaure Project
2004 Route des Lucioles BP 93
06902 Sophia Antipolis Cedex, France
{Xavier.Pennec, Pascal.Cachier, Nicholas.Ayache}@sophia.inria.fr
http://www-sop.inria.fr/epidaure/Epidaure-eng.html

Abstract. The "Demons Algorithm" in increasingly used for non-rigid registration of 3D medical images. However, if it is fast and usually accurate, the algorithm is based on intuitive ideas about image registration and it is difficult to predict when it will fail and why. We show in this paper that this algorithm can be considered as an approximation of a second order gradient descent on the sum of square of intensity differences criterion. We also reformulate Gaussian and physical model regularizations as minimization problems. Experimental results on synthetic and 3D Ultrasound images show that this formalization helps identifying the weak points of the algorithm and offers new research openings.

1 Introduction

Over recent years, a number of non-rigid registration techniques has been proposed. In 1981, Broit [Bro81] used the linear correlation as a measure of similarity between the two images to match. Bajcsy [BK89] differentiated this criterion and used a fixed fraction of its gradient as an external force to interact with a linear elasticity model.

Christensen [CJM97] shows that the linear elasticity, valid for small displacements, cannot guaranty the conservation of the topology of the objects as the displacements become larger: the Jacobian of the transformation can become negative. Thus, he proposed a viscous fluid model of transformations as it can handle larger displacement. This model is also linearized in practice.

Bro-Nielsen [BN96] started from the fluid model of Christensen and used the linearity of partial derivative equations to establish a regularization filter, several order of magnitude faster than the previous finite element method. He also justified his forces as the differential of the sum of square intensity differences criterion, but he still used a fixed fraction of this gradient, and show that Gaussian smoothing is an approximation of the linear elastic model.

Some authors [MMV98] tried to apply to non-rigid registration some criterions developed for rigid or affine matching using bloc-matching techniques. However, the evaluation of such a criterion imposes a minimal window size, thus limiting the resolution of these algorithm. Moreover, the regularization of the displacement field is usually implicit, i.e. only due to the integration of the criterion over the window, which means that there is no control on the quality of the solution transformation.

Recently, Thirion [Thi98] proposed to consider non rigid registration as a diffusion process. He introduced in the images entities (demons) that push according to local characteristics of the images in a similar way Maxwell did for solving the Gibbs paradox in thermodynamics. The forces he proposed were inspired from the optical flow equations. This algorithm is increasingly used in several teams as it it remarkably fast [BFC98,WGR+99,PTSR98]. However, there is up to now no underlying theory that could predict when it fails and why.

In this paper, we investigate non-rigid registration as a minimization problem. In the next section, we differentiate the sum of square intensity differences criterion (SSD) and show that the demons forces are an approximation of a second order gradient descent on this criterion. In the third section, we investigate regularization and show that the Gaussian smoothing is a greedy optimization of the regularized criterion. The experiment section shows that really minimizing the SSD criterion can lead to improvement with respect to the original demons algorithm, but more importantly that regularization is the critical step for non-rigid registration.

2 Non-Rigid Registration by Gradient Descent

Simple transformations, like rigid or affine ones, can be represented by a small number of parameters (resp. 6 and 12 in 3D). When it comes to free-form deformations, we need to specify the coordinates $T(x)$ of each point x of the image after the transformation. Such a non-parametric transformation is usually represented by its displacement field $U(x) = T(x) - x$ (or $U = T - \mathrm{Id}$).

Now, the goal of the registration is to find the transformation T that maps each point x of an image I to the "homologous" point $y = T(x)$ in image J. In this paper, we consider the registration of images from the same modality and we assume that they differ only by a geometric transformation and an independent Gaussian noise on the intensities. With these hypotheses, the well known SSD criterion is perfectly adapted for registration. Taking into account the fact that the image I deformed by T is $I \circ T^{(-1)}$ (and not $I \circ T$), we have:

$$SSD_I(T) = \int \big((T \star I)(x) - J(x)\big)^2 .dx = \int \big(I \circ T^{(-1)} - J\big)^2 \tag{1}$$

Known (or even unknown) intensity transformations are often better hypotheses even within mono-modal registration but they deserve other registration criterions. They will be investigated in a forthcoming study.

We note that the SSD criterion is not symmetric. Indeed, registering J on I leads to the criterion

$$SSD_J(T) = \int \big(I - J \circ T\big)^2 = \int \big(I \circ T^{(-1)} - J\big)^2 .|\nabla T^{(-1)}| \tag{2}$$

which has generally a minimum at a (hopefully slightly) different location.

2.1 Taylor Expansion of the SSD_J Criterion

Let T be the current estimation of the transformation and $(\nabla_J \circ T)(x)$ (resp. $(\mathcal{H}_J \circ T)(x)$) be the transformed gradient (resp. Hessian) of the image J. A perturbation by a displacement field $u(x)$ gives the following Taylor expansion:

$$(J \circ (T+u))(x) = (J \circ T)(x) + (\nabla_J \circ T)^{\mathrm{T}}.u(x) + \tfrac{1}{2} u(x)^{\mathrm{T}}.(\mathcal{H}_J \circ T).u(x)$$

Thus, the Taylor expansion of the criterion is:

$$\begin{aligned} SSD_J(T+u) = SSD_J(T) &+ 2\int (J \circ T - I)\,.(\nabla_J \circ T)^{\mathrm{T}}.u \\ &+ \int ((\nabla_J \circ T)^{\mathrm{T}}.u)^2 + \int (J \circ T - I)\,.u^{\mathrm{T}}.(\mathcal{H}_J \circ T).u + O(||u||^2) \end{aligned}$$

where $||u||^2 = \int_x \|u(x)\|^2.dx$ is the $\mathcal{L}_2$ norm of the small perturbation. As, by definition, $\int_x f(x)^{\mathrm{T}}.u(x).dx$ is the dot product of f and u in the space of square-integrable functions, we get by identification:

$$\nabla_{SSD_J}(T) = 2(J \circ T - I).(\nabla_J \circ T) \tag{3}$$

$$\mathcal{H}_{SSD_J}(T) = 2(\nabla_J \circ T).(\nabla_J \circ T)^{\mathrm{T}} + 2(J \circ T - I).(\mathcal{H}_J \circ T) \tag{4}$$

2.2 Second Order Gradient Descent on SSD_J

Let us now approximate the criterion by its tangential quadratic form at the current transformation T. We get the following first order approximation of the criterion gradient: $\nabla_{SSD_J}(T+u) \simeq \nabla_{SSD_J}(T) + H_{SSD_J}(T).u$

Assuming that the Hessian matrix of the criterion is positive definite, the minimum is obtained for a null gradient, i.e.: $u = -\mathcal{H}^{(-1)}_{SSD_j}(T).\nabla_{SSD_J}(T)$. This formula require to invert the Hessian matrix $\mathcal{H}_{SSD_J}(T)$ at each point x of the image. To speed up the process, we approximate this matrix by the closest scalar matrix (for the $\mathcal{L}_2$ norm on the matrix vector space):

$$\mathcal{H}_{SSD_J}(T) \simeq \frac{\mathrm{Tr}\,(\mathcal{H}_{SSD_J}(T))}{n}.\mathrm{Id} = \frac{\|\nabla_J \circ T\|^2 + (J \circ T - I).(\Delta_J \circ T)}{3}.\mathrm{Id}$$

where n is the space dimension (3 for us). Using this approximation, we get the following adjustment vector field:

$$u \simeq \frac{-3.(J \circ T - I).(\nabla_J \circ T)}{||\nabla_J \circ T||^2 + (J \circ T - I).(\Delta_J \circ T)} \tag{5}$$

2.3 Second Order Gradient Descent on SSD_I

Let us now consider the criterion SSD_I (formula 1). We can follow the same calculi as above but replacing I by J and T with $T^{(-1)}$. Thus, the optimal adjustment field w is:

$$w \simeq \frac{-3.(I \circ T^{(-1)} - J).(\nabla_I \circ T^{(-1)})}{||\nabla_I \circ T^{(-1)}||^2 + (I \circ T^{(-1)} - J).(\Delta_I \circ T^{(-1)})}$$

Now, we can develop the expression $(T^{(-1)} + w)^{(-1)}$ using a first order Taylor expansion: $(T^{(-1)}+w)^{(-1)} = T\circ(\mathrm{Id}+w\circ T)^{(-1)} \simeq T\circ(\mathrm{Id}-w\circ T)$. Thus, $u' = -w\circ T$ is the optimal adjustment displacement field for SSD_I with this formulation:

$$\hat{T} = T\circ(\mathrm{Id}+u') \qquad \text{with} \qquad u' = \frac{3.(I - J\circ T).\nabla_I}{||\nabla_I||^2 + (I - J\circ T).\Delta_I} \tag{6}$$

2.4 The Demons Method

For monomodal registration of non-segmented images, Thirion [Thi98] proposed a force playing the same role as our adjustment displacement field. This force was inspired from the optical flow equations, but was renormalized to prevent instabilities for small image gradient values:

$$v = \frac{(I - J\circ T).\nabla_I}{||\nabla_I||^2 + \alpha.(I - J\circ T)^2} \tag{7}$$

2.5 Comparison

If we except the constant factor 3, absent from the demons equation, u' and v are almost equivalent. However, we get from the gradient descent an explicit and automatic normalization between the gradient norm and the intensities at the denominator: $\alpha = \Delta_I/(I - J\circ T)$ (Thirion used a constant value $\alpha = 1$).

Another important difference is that formula (6) is only valid when the Hessian matrix of the criterion is positive definite. With our approximation, this means $||\nabla_I||^2 + (I - J\circ T).\Delta_I > 0$. If this denominator becomes close to zero, we have to switch to another gradient descent method, such as Levenberg-Marquardt for instance. On the contrary, the demons forces (equ. 7) are always defined and even bounded for $\alpha > 0$ by $\|v\| \leq 1/(2\sqrt{\alpha})$. However, nothing imposes that the SSD criterion decreases with such displacements.

3 Constraining the Transformation: Regularization

One of the main problem with free form deformations is that we can theoretically map all points of the same intensity to one point having this intensity in the other image. Of course, when repeated for all intensities, such a transformation perfectly minimizes the SSD criterion, but is also perfectly irregular. To regularize the transformation, there are two main methods: either rely on physical models, such as elastic and viscous fluid ones, or rely on regularization theory. We will see in the experimental part that the choice of the transformation constraints appears to be critical for doing measurements on the deformation field whereas it has little influence on the resampled image appearance.

3.1 Physical Models

Within physical models, elasticity was first used [BK89]. It is based on the *Green-St. Venant strain tensor*: $E = \frac{1}{2}\,(\nabla_T^{\mathrm{T}}.\nabla_T - \mathrm{Id}) = \frac{1}{2}(\nabla_U + \nabla_U^{\mathrm{T}} + \nabla_U^{\mathrm{T}}.\nabla_U)$. Then the St. Venant Kirchoff elastic energy is given by $W = \frac{\lambda}{2}(\mathrm{Tr}(E))^2 + \mu.\mathrm{Tr}(E^2)$, where λ and μ are Lamé coefficients describing the material. For linear elasticity, one simply neglects the quadratic term in the strain tensor, and one gets the following evolution equation: $f + \mu \Delta_U + (\lambda + \mu)\nabla(\nabla \cdot U) = 0$

When [CRM94] showed the invalidity of linear elasticity for large deformations, most efforts turned on viscous fluid models [CJM97,BN96]. This time, equations are based on the velocity v through the *rate of deformation tensor* $D = \frac{1}{2}(\nabla_v + \nabla_v^{\mathrm{T}})$. The energy is once again $W = \frac{\lambda}{2}(\mathrm{Tr}(D))^2 + \mu.\mathrm{Tr}(D^2)$ and we get the same evolution equation as before, but with the velocity field: $f + \mu \Delta_v + (\lambda + \mu)\nabla(\nabla \cdot v) = 0$

3.2 Gaussian Smoothing

The principle is to add a smoothness functional $\phi(T)$ to our criterion in order to disadvantage oscillatory solutions. Thus, ϕ should measure what remains of T after a high pass filter in the frequency domain [GJP93]. Let $K = k.\mathrm{Id}$ be an isotropic (i.e. radial basis) high-pass filter, $\tilde{K} = \tilde{k}.\mathrm{Id}$ its Fourier transform (also a high-pass filter) and $*$ the convolution operator. Thanks to Parseval theorem, the above properties of the smoothness kernel suggest functionals of the form:

$$\phi(T) = \frac{1}{2}\int \|K * T\|^2 dx = \frac{1}{2}\int \|\tilde{K}.\tilde{T}\|^2 = \frac{1}{2}\int \tilde{k}^2.\|\tilde{T}\|^2 d\omega$$

One can show that (for well formed kernel k), the gradient is: $\nabla_\phi(T) = (k*k)*T$. The Hessian matrix is more difficult to compute.

Ideally, we should add the gradient and Hessian matrix of the smoothing kernel to the gradient and Hessian matrix of our SSD criterion. In the original Demon's algorithm of [Thi98], Thirion prefers to use a greedy algorithm where he performs alternatively a gradient descent step on the SSD and a smoothing step. Let us consider the high-pass filter $k * k = (\delta - G_\sigma)/2$. Its gradient is $\nabla_\phi(T) = T - G_\sigma * T$, and a single step of a first order gradient descent gives:

$$T_{n+1} = T_n - \nabla_\phi(T_n) = G_\sigma * T_n$$

We note that smoothing the global displacement field U (or equivalently the global transformation T) could be considered as a rough simulation of an elastic model. By the way, [BN96] reported very similar results for these two methods. Following the same principle, we could consider the adjustment field u or v (equations 5,6,7) as an approximation of the velocity field: smoothing it with a Gaussian should be a rough simulation of a fluid model.

3.3 Comparison of Methods

Physical models are interesting because they model materials we are used to. However, the resolution of such models is often very complex and requires simplifications to be computationally tractable, such as the linearization of the strain tensor for elasticity of neglecting the pressure for fluids. Moreover computation times are usually high even if new implementations of fluid models are much faster [BN96]. Basically, the important point point seems to be that elastic equations are performing a regularization on the displacement field whereas fluid equations perform it on the velocity field.

On the other hand, Gaussian smoothing appears as a greedy optimization which is fast and easy to compute. Smoothing the global transformation (as in the original demons) could be considered as a rough simulation of an elastic model whereas smoothing the adjustment field simulates a rough fluid model.

4 Experiments

In Section 2, we saw the demons are approximately minimizing the SSD criterion if the Hessian matrix of this criterion is positive definite. To verify if minimizing the real criterion really helps, we implemented a very simple version of the demons with a steepest descent instead of the newton like method: at each point of the image, we compute the gradient and look for the minimal difference of intensities along this line. Then we smooth the adjustment field (with standard deviation σ_{fl}), add it to the global transformation, and smooth the global transformation (with standard deviation σ_{el}).

4.1 Synthetic experiments

In this experiment, we want to register a small 2D disk on a "C". We made the experiment with the original demons method and our first order gradient descent simulating "elastic" ($\sigma_{el} \gg \sigma_{fl}$) and "fluid" ($\sigma_{el} \ll \sigma_{fl}$) deformations (Figure 1).

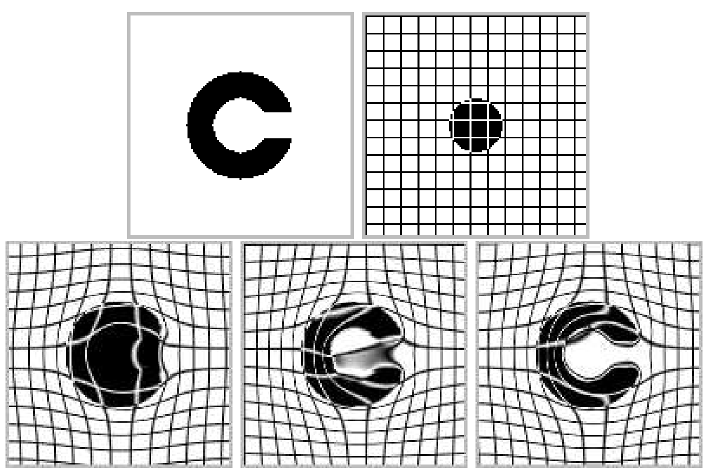

Fig. 1. Registration of a disk on a C. The grids are added after the registration to help visualizing the deformation. **Top:** reference image on the left, and image to register on the right. **Bottom:** from left to right, registration results with demons ($\sigma = 1.2$), with the new method simulating elastic deformations ($\sigma_{fl} = 0$, $\sigma_{el} = 1.2$) and with the new method simulating fluid deformations ($\sigma_{fl} = 3$, $\sigma_{el} = 0$).

Even using the same standard deviation for smoothing transformations, the demons (Figure 1, bottom left) and the minimization of the SSD criterion (bottom, middle) do not give the same results. This could be explained by demons forces where adding a high square intensity difference to a small gradient at the denominator lead to very small displacements. Smoothing is then sufficiently strong to prevent "white matter" to enter the interior of the "C". Thus, in this particular case, minimizing explicitly the SSD criterion gives better results.

There is surprisingly few differences between the structure of the "elastic" (Fig. 1, bottom middle) and the "fluid" (bottom right) deformations except in the interior of the "C". The "fluid" matches the edge much closer, but we can observe a singularity on the upper left part of the interior of the "C" (Jacobian going negative, visualized by an unexpected crossing of two grid lines). In fact, for such strong deformations, Gaussian smoothing is no more an approximation of physical models.

4.2 Registration of Ultrasound Images

In this experiment, we register two 3D US images from the European project ROBOSCOPE. A catheter was introduced in the brain of a (dead) pig and a balloon was inflated at different volumes to simulate the deformation that could happen during a tumor excision. Images presented in Figure (2) are 75x62x73 isotropic sub-images extracted from the original US images.

The result of the registration is visually correct both for the demons and for the new implementation. There is no significant difference on the difference image. Since we know that the deformation is physical and quite regular, we wanted to analyze the displacement field. We run the new implementation with $\sigma_{fl} = 3$ and $\sigma_{el} = 7$ whereas $\sigma (\simeq \sigma_{el})$ is limited to 2 in the original implementation of the demons. The visualization of the deformed grid on Figure (2) clearly shows that the new deformation field is much smoother.

To analyze the deformation field more carefully, we used the method developed in [RSDA99] to compute the Jacobian of the transformation that expresses the local volume variation due to the non-rigid transformation. We extracted an "edge surface" of this image of Jacobian and superimposed the results on the bottom of Figure (2). With the new method, we almost obtain the delineation of the balloon, except in the right part (the catheter in a fixed point and we segment on motion) and on the right lower part where the balloon is not visible in the images. More interestingly, the volumes measured inside our segmentation are respectively 0.982 and 0.875 cm^3 instead of 1.25 and 1 cm^3, which gives a ratio of 1.12 instead of 1.25. If this result is not perfect, it is encouraging. With the original demons, the deformation field is too noisy to measure something useful.

5 Discussion

Among non-rigid registration algorithms, the demon's method is fast but not well understood from a theoretical point of view. We showed in this article that the forces proposed by Thirion correspond to a second order gradient descent on the sum of square of intensity differences (SSD) criterion. Experiments showed that

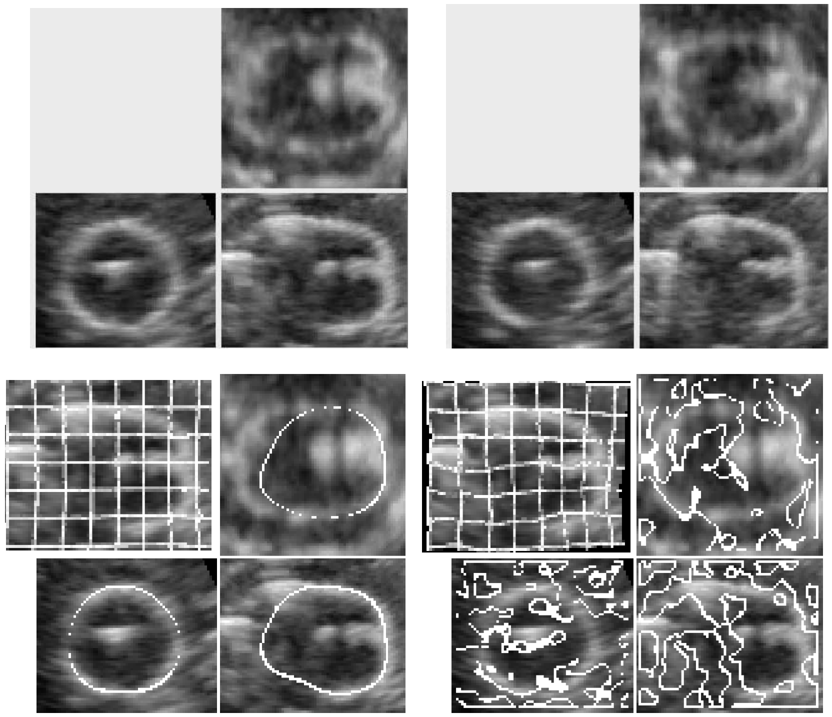

Fig. 2. Tracking deformations in 3D ultrasound images of a pig brain. Images were provided by Volker Paul, Fraunhofer (Institute) Society, Munich (Germany) as part of the EC-funded ROBOSCOPE project. **Top left:** three orthogonal views of the image to register (balloon volume: 1.25 cm^3). **Top right:** reference image (balloon volume: 1 cm^3). **Bottom:** "edges" of the Jacobian with new method (**left**) and with demons (**right**). We added on the upper left corner of each image triplet a slice (corresponding to the lower right image) after registration to visualize the corresponding deformation.

the real minimization of the SSD criterion could give a better results than the demons with synthetic data, but demons are performing well on real data. Both synthetic and real data experiments put into evidence that a good regularization is critical to obtain a useful deformation field.

The understanding of the demons algorithm and its weaknesses opens many new research avenues for non-rigid registration. Probably the most important one is regularization: the first work will be to minimize the criterion and the smoothing kernel in the same gradient descent step. The second work will be to look for families of smoothing kernels that better approximate physical models, the filtering methodology ensuring reasonable computation times.

Last but not least, the SSD criterion relies on imaging acquisition assumptions that are hardly verified in modalities such as US. We need to develop new criterions better adapted to the physics of the acquisition process. Then, the gradient descent scheme developed in this paper offers us a methodology to plug them into an operational registration algorithm.

Acknowledgments We would like to thank specially Volker Paul, Fraunhofer Institute Society, Munich (Germany) who provided us with the US images as part of the EC-funded ROBOSCOPE project HC 4018, a collaboration between The Fraunhofer Institute Society (Germany), Fokker Control System (Netherlands), Imperial College (UK), INRIA Sophia Antipolis (France), ISM Salzburg and Kretz Technik (Austria). Thanks also to Kretz and ISM for organizing / providing a high-end imaging system for these examinations. Many thanks also to David Rey, Alexandre Guimond and Jean-Philippe Thirion, for fruitful discussions and experiments on non rigid registration and analysis of the deformation field.

References

[BFC98] I. Bricault, G. Ferretti, and P. Cinquin. Registration of real and ct-derived virtual bronchoscopic images to assist transbronchial biopsy. *Transactions in Medical Imaging*, 17(5):703–714, 1998.

[BK89] R. Bajcsy and S. Kovačič. Multiresolution elastic matching. *Computer Vision, Graphics and Image Processing*, 46:1–21, 1989.

[BN96] M. Bro-Nielsen. *Medical image registration and surgery simulation.* PhD thesis, IMM-DTU, 1996.

[Bro81] C. Broit. Optimal registration of deformed images. Doctoral Dissertation, University of Pennsylvania, August 1981.

[CPA99] P. Cachier, X. Pennec and N. Ayache. Fast Non Rigid Matching by Gradient Descent: Study and Improvements of the "Demons" Algorithm. *INRIA Research Report 3706*, 1999.

[CJM97] G. E. Christensen, S. C. Joshi, and M. I. Miller. Volumetric transformation of brain anatomy. *IEEE Transactions on Medical Imaging*, 16(6):864–877, December 1997.

[CRM94] G.E. Christensen, R.D. Rabitt, and M.I. Miller. 3d brain mapping using a deformable neuroanatomy. *Physics in Medicine and Biology*, 39:609–618, 1994.

[GJP93] F. Girosi, M. Jones, and T. Poggio. Priors, stabilizers and basis functions: from regularization to radial, tensor and additive splines. A.I. Memo 1430, M.I.T., A.I. Laboratory, June 1993.

[MMV98] J. B. A. Maintz, E. H. W. Meijering, and M. A. Viergever. General multimodal elastic registration based on mutual information. *Image Processing*, 1998.

[PTSR98] S. Prima, J.-P. Thirion, G. Subsol, and N. Roberts. Automatic Analysis of Normal Brain Dissymmetry of Males and Females in MR Images. In *Proc. of MICCAI'98*, volume LNCS 1496, pages 770–779, 1998.

[RSDA99] D. Rey, G. Subsol, H. Delingette, and N. Ayache. Automatic Detection and Segmentation of Evolving Processes in 3D Medical Images: Application to Multiple Sclerosis. In *Proc. of IPMI'99*, LNCS, Visegrád, Hungary, June 1999. Springer. To appear.

[Thi98] J.-P. Thirion. Image matching as a diffusion process: an analogy with maxwell's demons. *Medical Image Analysis*, 2(3), 1998.

[WGR+99] J. Webb, A. Guimond, N. Roberts, P. Eldridge an D. Chadwick, J. Meunier, and J.-P. Thirion. Automatic detection of hippocampal atrophy on magnetic resonnance images. *Magnetic Resonance Imaging*, 17, 1999. To appear.