A Variational Approach to Multi-Modal Image Matching

Christophe Chefd'Hotel[1], Gerardo Hermosillo[2] and Olivier Faugeras[1,2]
[1]*MIT AI Laboratory*
Cambridge, MA, USA
{*chefdhot, faugeras*}*@ai.mit.edu*
[2]*INRIA*
Sophia-Antipolis, France
{*ghermosi, faugeras*}*@sophia.inria.fr*

Abstract

We address the problem of non-parametric multi-modal image matching. We propose a generic framework which relies on a global variational formulation and show its versatility through three different multi-modal registration methods : supervised registration by joint intensity learning, maximization of the mutual information and maximization of the correlation ratio. Regularization is performed by using a functional borrowed from linear elasticity theory. We also consider a geometry-driven regularization method. Experiments on synthetic images and preliminary results on the realignment of MRI datasets are presented.

1. Introduction

Image matching is one of the fundamental problems in computer vision, and a very important practical issue in medical image processing. One of its simplest forms is the optical-flow computation, which aims at recovering the displacement field between two frames of a video sequence. In medical imaging, registration techniques have also been developed to recover geometric distortions and misalignments between MRI datasets. In this case, large local and global deformations may occur and must be taken into account.

Most of the existing methods rely on a very strong assumption : images have been acquired through similar sensors and thus present comparable intensity maps, up to a certain amount of noise. However, there exist numerous situations where this assumption is hardly satisfied. For instance, one may consider the problem of face recognition and tracking under varying illumination conditions, or the stereo vision using cameras with different spectral sensitivities. This problem occurs also in medical imaging. There are several MRI acquisition modalities, and inter-modality registrations must be performed to allow an accurate fusion of complementary information. One must notice that the relation between the intensity maps of two modalities is not necessarily mono-functional, and is generally unknown.

To cope with this difficulty, some statistical and information-theoretic similarity measures have been proposed. These measures include the mutual information [24, 26], the correlation ratio [20, 21], and maximum likelihood criteria based on pre-registered datasets [15]. Registrations are usually performed by maximizing these criteria over a low-dimensional parametric class of deformations, e.g. rigid or affine transformations. The extension of this approach to larger sets of deformations is a challenging problem. The existing methods for non-rigid multi-modal registration are based on more complex parametric models [17, 22], block-matching strategies [16, 12, 11], or on the combination of a mono-modal algorithm with some parametric intensity corrections [19].

In this paper, we propose a generic viewpoint to cope with different modalities while offering the flexibility of some well-known variational techniques, widely used for optical-flow computations and mono-modal matching problems [4, 9, 2, 18] (among others). We derive the previous similarity measures in a variational setting, gaining access to infinite dimensional deformation spaces and advanced regularization techniques. The resulting algorithms allow the recovery of dense displacement fields.

This approach is presented as follows. Section 2 discusses the derivation of variational principles from multi-modal similarity measures. Three different methods are considered : supervised registration by joint intensity learning, maximization of the mutual information and maximization of the correlation ratio. Section 3 describes the choice of a suitable regularization functional. We consider a physically-based model borrowed from elasticity theory and a geometric approach relying on discontinuity-preserving diffusion tensors. Section 4 presents numerical experiments on synthetic images and preliminary results on MRI datasets. We conclude and sketch future developments of this work in Section 5.

2. Variational Principles and Multi-Modal Similarity Measures

First, we introduce some notations which will be used in the sequel. Given two images, represented by $f : \Omega \subset \mathbb{R}^n \longmapsto \mathbb{R}$ and $g : \Omega \subset \mathbb{R}^n \longmapsto \mathbb{R}$, a deformation is modeled by a mapping $\mathbf{v} : \Omega \longmapsto \Omega$. Arbitrarily, we set f as the *target* image and $\mathbf{v}$ acts on g by $g \circ \mathbf{v}$, where $\circ$ denotes the composition of applications. We assume that $\mathbf{v}$ belongs to a Hilbert space $\mathcal{F}$ equipped with its usual scalar product $\langle \mathbf{w_1}, \mathbf{w_2} \rangle = \int_\Omega \mathbf{w}_1(\mathbf{x}) \cdot \mathbf{w}_2(\mathbf{x})\, d\mathbf{x}$. It is often convenient to

0-7695-1278-X/01 $10.00 © 2001 IEEE

consider the decomposition $\mathbf{v} = \mathbf{Id} + \mathbf{u}$ where $\mathbf{u}$ is a displacement field.

We define the generic matching problem as the minimization of a cost functional

$$\mathcal{I}[\mathbf{u}] = \mathcal{J}[\mathbf{u}] + \alpha\mathcal{R}[\mathbf{u}],\ \alpha \in \mathbb{R}_+^*,$$

where $\mathcal{J}$ measures the disparity of f and $g \circ (\mathbf{Id}+\mathbf{u})$, while $\mathcal{R}$ defines regularizing constraints and smoothness assumptions on $\mathbf{u}$. For instance, one of the simplest choices for $\mathcal{J}$ would be the Sum of Squared Differences (**SSD**)

$$\mathcal{J}[\mathbf{u}] = \frac{1}{2}\int_\Omega (f(\mathbf{x}) - g(\mathbf{x}+\mathbf{u}(\mathbf{x})))^2\, d\mathbf{x}. \quad (1)$$

Of course, this criterion becomes meaningless when considering different modalities. Three alternatives will be detailed in the following. We propose to define the regularization term $\mathcal{R}$ from the Jacobian matrix $D\mathbf{u}$, such that

$$\mathcal{R}[\mathbf{u}] = \int_\Omega \varphi(D\mathbf{u}(\mathbf{x}))\, d\mathbf{x}, \quad (2)$$

for an arbitrary mapping $\varphi : \mathbb{R}^{n\times n} \mapsto \mathbb{R}_+$. When $\varphi(\cdot) = \frac{1}{2}||\cdot||^2$, where $||\cdot||$ denotes the Frobenius norm, $\mathcal{R}$ boils down to the so-called *Tikhonov* functional [23], often used to deal with ill-posed problems. In general, the selection of a mapping φ offers a simple and flexible way to design problem-specific regularization methods. This issue will be addressed in Section 3.

We presuppose the existence of a minimum $\hat{\mathbf{u}} = \text{argmin}_{\mathbf{u}\in\mathcal{F}}\mathcal{I}[\mathbf{u}]$. Using some classical tools of variational calculus (see [8] for an introduction to this topic), we consider the first variation (Gâteaux Variation) of $\mathcal{I}$

$$d\mathcal{I}[\mathbf{u},\mathbf{h}] = \left.\frac{\partial\mathcal{I}[\mathbf{u}+\epsilon\mathbf{h}]}{\partial\epsilon}\right|_{\epsilon=0},$$

for any suitable variation $\mathbf{h}$ in $\mathcal{F}$.

One can show that if $\mathcal{I}$ has a local extremum $\mathbf{u}$ in $\mathcal{F}$, its first variation at $\mathbf{u}$ must vanish. When $d\mathcal{I}[\mathbf{u},\mathbf{h}] = \langle\nabla_{\mathbf{u}}\mathcal{I},\mathbf{h}\rangle$, $\nabla_{\mathbf{u}}\mathcal{I}$ defines the gradient of the functional $\mathcal{I}$. Then, the necessary condition of optimality is readily equivalent to $\nabla_{\mathbf{u}}\mathcal{I} = \mathbf{0}$. This is the so-called Euler equation, which can be easily computed when $\mathcal{I}$ has a simple form. For instance, if

$$\mathcal{I}[\mathbf{u}] = \int_\Omega \phi(\mathbf{x},\mathbf{u}(\mathbf{x}),D\mathbf{u}(\mathbf{x}))\, d\mathbf{x}, \quad (3)$$

one can show that its Euler equation becomes

$$\phi_{\mathbf{u}}(\mathbf{x},\mathbf{u},D\mathbf{u}) - \mathbf{div}(\phi_{D\mathbf{u}}(\mathbf{x},\mathbf{u},D\mathbf{u})) = \mathbf{0}, \quad (4)$$

where $\mathbf{div}(\mathbf{T}) = [\nabla\cdot\mathbf{t}_1,\ldots,\nabla\cdot\mathbf{t}_n]^T$ for $\mathbf{T} = [\mathbf{t}_1,\ldots,\mathbf{t}_n]^T$. In our situation, we will not try to directly solve the Euler equation and prefer a suboptimal solution based on a gradient descent strategy (starting from a suitable initial guess $\mathbf{u}_0$). The direction of the steepest descent depends on the choice of the scalar product $\langle,\rangle$. With the previous definition, we get the classical gradient flow :

$$\begin{cases} \frac{\partial\mathbf{u}}{\partial t} &= -\nabla_{\mathbf{u}}\mathcal{I} \\ \mathbf{u}(0,\cdot) &= \mathbf{u}_0 \end{cases}$$

2.1. Supervised Registration and Joint Intensity Learning

In order to match images from two different modalities, several strategies are conceivable. They can be divided into two categories. One solution consists of selecting important geometric features, such as edges and corners, which could be used as landmarks in both modalities. A second way, which will be explored here, uses a probabilistic description of the dependence between different intensity maps.

One of the simplest probabilistic image models is based on the following assumption. Image pixels are independent samples of a given random variable, regardless of their relative positions. We must notice that this approach seems a priori incompatible with our previous modeling assumptions. In fact, we need to transpose this setting to the continuous case and consider strictly stationary processes whose time parameter becomes a vector defined on Ω. Then, our functions f and g become sample paths of two random processes.

We denote by p^f and $p^g_{\mathbf{u}}$ the intensity distributions estimated from f and $g \circ (\mathbf{Id} + \mathbf{u})$. $p^{f,g}_{\mathbf{u}}$ is an estimate of their joint intensity distribution. To simplify notations we introduce two random variables, X^f and $X^g_{\mathbf{u}}$, whose probability densities are respectively given by p^f and $p^g_{\mathbf{u}}$. Note that within this probabilistic framework, the link between two modalities is fully characterized by a joint density.

First, we assume that a learning set is available to estimate the "true" joint intensity distribution $p(i_1,i_2)$. In practice, one could use a Parzen-Rozenblatt estimator [5, 13] on sets of manually pre-registered images. From the knowledge of p, we derive a *supervised* registration principle (**SR**).

We borrow from information theory the notion of *uncertainty* or *information*, defined for an event with respect to a probability measure. In this case, an event is the co-occurence of two intensity values : $i_1 = f(\mathbf{x})$ and $i_2 = g(\mathbf{x}+\mathbf{u}(\mathbf{x}))$ (at any point $\mathbf{x}$ in Ω). The amount of information conveyed by this event is given by $-\log p(i_1,i_2)$. A global disparity measure follows by computing the total amount of information for a given displacement $\mathbf{u}$:

$$\mathcal{J}_{\mathbf{SR}}[\mathbf{u}] = -\int_\Omega \log p(f(\mathbf{x}),g(\mathbf{x}+\mathbf{u}(\mathbf{x})))\, d\mathbf{x}.$$

One may also give an interpretation of this expression as a continuous formulation of the maximum likelihood principle developed in [15]. Combined with (2), the corresponding cost functional $\mathcal{I}_{\mathbf{SR}}$ has a simple form given by equation (3) with

$$\phi(\mathbf{x},\mathbf{u},D\mathbf{u}) = -\log p(f(\mathbf{x}),g(\mathbf{x}+\mathbf{u}(\mathbf{x}))) + \alpha\varphi(D\mathbf{u}(\mathbf{x})).$$

With a differentiable and strictly positive estimate of p, and provided φ, f and g enjoy the appropriate regularity, ϕ is

continously differentiable with respect to each component of $\mathbf{u}$ and $D\mathbf{u}$, thus

$$\begin{aligned}\phi_{\mathbf{u}}(\mathbf{x},\mathbf{u},D\mathbf{u}) &= -\frac{p_{i_2}(f,g\circ(\mathbf{Id}+\mathbf{u}))}{p(f,g\circ(\mathbf{Id}+\mathbf{u}))}\nabla g\circ(\mathbf{Id}+\mathbf{u})\\ \phi_{D\mathbf{u}}(\mathbf{x},\mathbf{u},D\mathbf{u}) &= \alpha\,\varphi_{D\mathbf{u}}(D\mathbf{u})\end{aligned}$$

where p_{i_2} denotes the partial derivative of p with respect to its second variable. The corresponding Euler equation can not be solved directly. In effect, given an initial displacement field $\mathbf{u}_0$, we consider a suboptimal solution by building the following gradient flow :

$$\left\{\begin{array}{rcl}\mathbf{u}_t &=& \dfrac{p_{i_2}(f,g\circ(\mathbf{Id}+\mathbf{u}))}{p(f,g\circ(\mathbf{Id}+\mathbf{u}))}\nabla g\circ(\mathbf{Id}+\mathbf{u}) \\ & & \qquad +\alpha\,\mathbf{div}(\varphi_{D\mathbf{u}}(D\mathbf{u}))\\ \mathbf{u}(0,\cdot) &=& \mathbf{u}_0\end{array}\right. \tag{5}$$

It is particularly interesting to compare this flow with the one obtained from an **SSD** criterion (1). In this case, we would get

$$\left\{\begin{array}{rcl}\mathbf{u}_t &=& (f-g\circ(\mathbf{Id}+\mathbf{u}))\,\nabla g\circ(\mathbf{Id}+\mathbf{u}) \\ & & \qquad +\alpha\,\mathbf{div}(\varphi_{D\mathbf{u}}(D\mathbf{u}))\\ \mathbf{u}(0,\cdot) &=& \mathbf{u}_0\end{array}\right. \tag{6}$$

We immediatly notice the analogy between the function $\frac{p_{i_2}(i_1,i_2)}{p(i_1,i_2)}$ in equation (5), and the intensity comparison function i_1-i_2 of equation (6). They play the same role, but the function $\frac{p_{i_2}(i_1,i_2)}{p(i_1,i_2)}$ takes into account the existing relation between the intensity maps. In this case, there is no necessary assumption of a unique functional dependence. Any element of one intensity map may have several corresponding elements in the other map (and conversely). The knowledge of the nearest most likely intensity correspondence is implicitly taken into account in the comparison function.

Let us now assume we do not have any training set available for the learning process. One can easily develop the following argument : if the initial pose is close to the solution (small deformations), a sufficiently robust estimate of their joint intensity distribution may be a good approximation of the real joint density for these two modalities. The incorrectly matched image values are considered as noise. Subsequently, $\mathbf{u}$ could be somehow recovered from this initial estimate. In the following, we will see that this heuristic appears implicitly in the mutual information similarity measure.

2.2. Mutual Information and Correlation Ratio

When no training set is available, we propose to compensate for this lack of knowledge by using some statistical or information-theoretic similarity measures. Among numerous criteria, the mutual information and the correlation ratio have already been proven to be very effective in image matching. They are usually considered as global criteria optimized over low-dimensional classes of deformations.

Mutual information is borrowed from information theory. This intensity based similarity measure was introduced in the context of multi-modal registration in [24, 26]. Using our notation, the mutual information computed from f and $g\circ(\mathbf{Id}+\mathbf{u})$ is given by

$$\mathbf{MI}_{\mathbf{u}}^{f,g} = \int_{\mathbb{R}^2} p_{\mathbf{u}}^{f,g}(i_1,i_2)\log\frac{p_{\mathbf{u}}^{f,g}(i_1,i_2)}{p^f(i_1)p_{\mathbf{u}}^g(i_2)}di_1di_2\,,$$

which is nothing but the Kullback-Leibler divergence between $p_{\mathbf{u}}^{f,g}$ and $p^f\cdot p_{\mathbf{u}}^g$ [13]. From this viewpoint, this criterion indicates how the intensity distributions of two images fail to be independent.

On the other hand, the correlation ratio has been proposed in [20, 21] as a similarity measure for multi-modal affine registration. This criterion relies on a slightly different notion of similarity. From the definition

$$\mathbf{CR}_{\mathbf{u}}^{f,g} = \frac{\mathbf{Var}[\mathbf{E}[X_{\mathbf{u}}^g|X^f]]}{\mathbf{Var}[X_{\mathbf{u}}^g]}\,, \tag{7}$$

$\mathbf{CR}_{\mathbf{u}}^{f,g}$ can intuitively be described as the proportion of energy in $X_{\mathbf{u}}^g$ which is "*explained*" by X^f. More formally, this measure is bounded ($0\leq\mathbf{CR}_{\mathbf{u}}^{f,g}\leq 1$)) and expresses the level of functional dependence between X_u^g and X^f :

$$\left\{\begin{array}{lcl}\mathbf{CR}_{\mathbf{u}}^{f,g}=1 &\Leftrightarrow& \exists\phi\; X_{\mathbf{u}}^g=\phi(X^f)\\ \mathbf{CR}_{\mathbf{u}}^{f,g}=0 &\Leftrightarrow& \mathbf{E}[X_{\mathbf{u}}^g|X^f]=X_{\mathbf{u}}^g\end{array}\right.$$

Unlike $\mathbf{MI}$, $\mathbf{CR}$ is not symmetrical.

We propose to include these criteria in the definition of the functional $\mathcal{J}$. In both cases, the computation of their first variation is achieved by considering the variations of some joint density estimates according to $\mathbf{u}$. Formally, to ensure the compatibility of this approach with a variational setting, the continous form of Parzen-Rozenblatt estimators is used [5].

2.2.1 Mutual Information

We consider the opposite of the mutual information to transform a maximization problem into the minimization of the following cost functional :

$$\mathcal{I}_{\mathbf{MI}}[\mathbf{u}] = -\mathbf{MI}_{\mathbf{u}}^{f,g} + \alpha\int_\Omega\varphi(D\mathbf{u})\,d\mathbf{x}\,. \tag{8}$$

Due to the the complex form of $\mathcal{J}_{\mathbf{MI}}[\mathbf{u}] = -\mathbf{MI}_{\mathbf{u}}^{f,g}$, an explicit computation of its first variation must be carried out. After some simplifications (see [14] for details), we get the following expression :

$$\frac{\partial\mathcal{J}_{\mathbf{MI}}[\mathbf{u}+\epsilon\mathbf{h}]}{\partial\epsilon} = -\int_{\mathbb{R}^2}L^{\mathbf{u}+\epsilon\mathbf{h}}(i_1,i_2)\frac{\partial p_{\mathbf{u}+\epsilon\mathbf{h}}^{f,g}(i_1,i_2)}{\partial\epsilon}di_1di_2\,,$$

where

$$L^{\mathbf{u}+\epsilon\mathbf{h}}(i_1,i_2) = \left(1+\log\frac{p_{\mathbf{u}+\epsilon\mathbf{h}}^{f,g}(i_1,i_2)}{p^f(i_1)p_{\mathbf{u}+\epsilon\mathbf{h}}^g(i_2)}\right).$$

Generalizing the expression given in [5] for continous time random processes, the joint intensity distribution estimated from f and $g \circ (\mathbf{Id} + \mathbf{u})$ is provided by a nonparametric Parzen-Rozenblatt density model :

$$p_{\mathbf{u}}^{f,g}(i_1, i_2) = \frac{1}{Z} \int_\Omega \psi(f(\mathbf{x}) - i_1, g(\mathbf{x} + \mathbf{u}(\mathbf{x})) - i_2)\, d\mathbf{x}\,,$$

where Z is a normalizing constant and $\psi(\alpha, \beta) > 0$ is a smooth bidimensional density kernel. This estimate is used to compute $\frac{\partial p_{\mathbf{u}+\epsilon\mathbf{h}}^{f,g}(i_1,i_2)}{\partial \epsilon}$. Letting $\epsilon = 0$, the convolution operator, denoted $\star$, appears naturally [14]. It commutes with the derivation operator such that

$$d\mathcal{J}_{\mathbf{MI}}[\mathbf{u}, \mathbf{h}] = \left\langle -\frac{1}{Z}[\psi \star \frac{\partial L^{\mathbf{u}}}{\partial i_2}](f, g \circ (\mathbf{Id} + \mathbf{u}))\, \nabla g \circ (\mathbf{Id} + \mathbf{u}), \mathbf{h} \right\rangle ,$$

with

$$\frac{\partial L^{\mathbf{u}}}{\partial i_2} = \frac{1}{p_{\mathbf{u}}^{f,g}(i_1, i_2)} \frac{\partial p_{\mathbf{u}}^{f,g}(i_1, i_2)}{\partial i_2} - \frac{1}{p_{\mathbf{u}}^{g}(i_2)} \frac{\partial p_{\mathbf{u}}^{g}(i_2)}{\partial i_2}\,.$$

As we did previously, we actually consider the related steepest descent flow :

$$\left\{ \begin{array}{rcl} \mathbf{u}_t & = & \frac{1}{Z}[\psi \star \frac{\partial L^{\mathbf{u}}}{\partial i_2}](f, g \circ (\mathbf{Id} + \mathbf{u})) \\ & & \quad \nabla g \circ (\mathbf{Id} + \mathbf{u}) + \alpha\, \mathbf{div}(\varphi_{D\mathbf{u}}(D\mathbf{u})) \\ \mathbf{u}(0, \cdot) & = & \mathbf{u}_0 \end{array} \right. \tag{9}$$

We recognize the first term of $\frac{\partial L^{\mathbf{u}}}{\partial i_2}$ as the comparison function of the supervised registration method. This comparison function is re-evaluated continuously through time, combined with a similar expression of its marginal density, and smoothed by convolution. This formula can be considered as an extension of the heuristic introduced earlier. As a matter of fact, this method achieves registration by successive adjustments of its joint intensity model. Its first estimate is simply based on the initial pose.

2.2.2 Correlation Ratio

In the same setting, we now consider a functional of the form :

$$\mathcal{I}_{\mathbf{CR}}[\mathbf{u}] = \underbrace{-\mathbf{CR}_{\mathbf{u}}^{f,g}}_{\mathcal{J}_{\mathbf{CR}}[\mathbf{u}]} + \alpha \int_\Omega \varphi(D\mathbf{u})\, d\mathbf{x}\,. \tag{10}$$

Instead of using the original definition of $\mathbf{CR}_{\mathbf{u}}^{f,g}$ (7), we can use the total variance theorem to write

$$\mathbf{CR}_{\mathbf{u}}^{f,g} = 1 - \frac{\mathbf{E}[\mathbf{Var}[X_{\mathbf{u}}^g | X^f]]}{\mathbf{Var}[X_{\mathbf{u}}^g]}\,.$$

This transformation was suggested in [20, 21], and we notice that $\mathcal{J}_{\mathbf{CR}}$ might be directly replaced by the second term of this formula. Similar techniques apply to the study of the mutual information and the present situation. We show in [14] that the first variation of (10) is given by the expression :

$$d\mathcal{J}_{\mathbf{CR}}[\mathbf{u}, \mathbf{h}] = \left\langle \frac{1}{Z}[\psi \star \frac{\partial L^{\mathbf{u}}}{\partial i_2}](f, g \circ (\mathbf{Id} + \mathbf{u}))\nabla g \circ (\mathbf{Id} + \mathbf{u}), \mathbf{h} \right\rangle$$

such that

$$\frac{\partial L^{\mathbf{u}}}{\partial i_2} = \frac{2\left(i_2 - \mathbf{E}[X_{\mathbf{u}}^g | X^f = i_1] + (\mathbf{CR}_{\mathbf{u}}^{f,g} - 1)(i_2 - \mathbf{E}[X_{\mathbf{u}}^g])\right)}{\mathbf{Var}[X_{\mathbf{u}}^g]}\,.$$

Taking into account the usual regularization functional $\mathcal{R}$, this result leads to the following gradient flow :

$$\left\{ \begin{array}{rcl} \mathbf{u}_t & = & -\frac{1}{Z}[\psi \star \frac{\partial L^{\mathbf{u}}}{\partial i_2}](f, g \circ (\mathbf{Id} + \mathbf{u})) \\ & & \quad \nabla g \circ (\mathbf{Id} + \mathbf{u}) + \alpha\, \mathbf{div}(\varphi_{D\mathbf{u}}(D\mathbf{u})) \\ \mathbf{u}(0, \cdot) & = & \mathbf{u}_0 \end{array} \right. \tag{11}$$

One more time, we can observe that this resulting matching method relies on a dynamic re-evaluation of the joint intensity model. In this case, we notice that the expression of the correlation ratio is computed continuously through time. $\frac{\partial L^{\mathbf{u}}}{\partial i_2}$ provides a comparison of intensity levels based on some global estimate of the functional dependence between intensity maps.

2.3. Summary and comments

We assumed that joint intensity distributions fully characterize the relation between the intensity maps of different modalities. Either this density is estimated from a learning set, or it is also an unknown parameter. In the first case, a simple comparison function is computed, once for all, to establish a meaningful correspondence between intensity values. When the density is unknown, its estimate is built from the initial pose, and is continously re-estimated along with the gradient descent process. The mutual information and the correlation ratio illustrate this approach. From the expression of their respective comparison function, the choice between these two methods can be expected to be a classical trade-off between flexibility and robustness.

By letting h denote a generic comparison function, we can now define a global gradient flow formulation :

$$\left\{ \begin{array}{rcl} \mathbf{u}_t & = & h(f, g \circ (\mathbf{Id} + \mathbf{u}))\, \nabla g \circ (\mathbf{Id} + \mathbf{u}) \\ & & \quad + \alpha\, \mathbf{div}(\varphi_{D\mathbf{u}}(D\mathbf{u})) \\ \mathbf{u}(0, \cdot) & = & \mathbf{u}_0 \end{array} \right. \tag{12}$$

which can be specialized with Table 1.

In the case of supervised registration, h is fixed. The corresponding gradient flow is a nonlinear parabolic equation. If the expression of φ is sufficiently simple, we may expect to build a proof of existence and uniqueness of weak solutions by using the theoretical tools developed for reaction-diffusion equations [10]. Such a proof is provided in [3] for a similar mono-modal optical flow PDE.

Method	Intensity Comparison $h(i_1, i_2)$
SSD	$i_1 - i_2$
SR	$\frac{1}{p(i_1,i_2)} \frac{\partial p(i_1,i_2)}{\partial i_2}$
MI	$\frac{1}{Z}[\psi \star (\frac{1}{p^{f,g}_{\mathbf{u}}} \frac{\partial p^{f,g}_{\mathbf{u}}}{\partial i_2} - \frac{1}{p^{g}_{\mathbf{u}}} \frac{\partial p^{g}_{\mathbf{u}}}{\partial i_2})](i_1, i_2)$
CR	$\frac{-2\ \psi\star\left(i_2 - \mathbf{E}[X^g_{\mathbf{u}} \mid X^f = i_1] + (\mathbf{CR}^{f,g}_{\mathbf{u}} - 1)(i_2 - \mathbf{E}[X^g_{\mathbf{u}}])\right)}{\mathbf{Var}[X^g_{\mathbf{u}}]\ Z}$

Table 1. Comparison functions.

3. Regularization

We initially suggested a flexible regularization term (2) defined by a mapping $\varphi : \mathbb{R}^{n \times n} \mapsto \mathbb{R}_+$. In this framework, we can notice that a simple *Tikhonov* model with $\varphi : \mathbf{X} \mapsto \frac{1}{2}\mathrm{Tr}(\mathbf{X}^T\mathbf{X})$ yields a diffusion term, since $\mathbf{div}(\varphi_{D\mathbf{u}}(D\mathbf{u})) = \mathbf{div}(D\mathbf{u}) = \Delta\mathbf{u}$. It was one of the first regularizations proposed to solve the dense image matching problem [4]. This is a strong assumption which is often too restrictive on $\mathbf{u}$. In order to relax this constraint, several heuristics can be taken into account. They lead to different choices for φ.

3.1. Linear Elasticity

As in [6], where a linear elastic model is used for the realignment of mono-modal MRI, we can assume that the geometric deformation $\mathbf{v} = \mathbf{Id} + \mathbf{u}$ corresponds to the strain of an elastic, isotropic and materially uniform material. Following the elasticity theory formalism, we build a linear approximation of the strain tensor $\mathbf{E} = \frac{1}{2}(D\mathbf{v}^T D\mathbf{v} - \mathbf{I}) \simeq \frac{1}{2}(D\mathbf{u}^T + D\mathbf{u})$ (for small displacements $\mathbf{u}$). According to the Saint-Venant/Kirchoff model [7], the corresponding strain energy is given by the functional :

$$\mathcal{E}[\mathbf{u}] = \int_\Omega \frac{\lambda}{2}(\mathrm{Tr}(\mathbf{E}))^2 + \mu\mathrm{Tr}(\mathbf{E}^2)\, d\mathbf{x},$$

where λ and μ are the *Lamé* constants of a given elastic material. The key idea is to use the approximation $\mathcal{E}[\mathbf{u}] \simeq \int_\Omega \varphi(D\mathbf{u}(\mathbf{x}))\, d\mathbf{x}$, with $\varphi : \mathbf{X} \mapsto \frac{\lambda}{8}(\mathrm{Tr}(\mathbf{X}^T + \mathbf{X}))^2 + \frac{\mu}{2}\mathrm{Tr}((\mathbf{X}^T + \mathbf{X})^2)$, as a regularization functional $\mathcal{R}$. $\varphi_{D\mathbf{u}}(D\mathbf{u})$ corresponds to the so-called *stress tensor*, and we easily get

$$\mathbf{div}(\varphi_{D\mathbf{u}}(D\mathbf{u})) \quad = \quad \mu\Delta\mathbf{u} + (\lambda + \mu)\nabla(\nabla \cdot \mathbf{u}).$$

This expression is directly applied in combination with equation (12). A single parameter $\rho = \mu/\lambda$ is used in practice to adjust the elasticity level. In general, we will use this method with any type of datasets (not necessarily images of objects undergoing a "*real*" elastic deformation). Even when considering MRI volumes, we shall point out that this regularization principle has been selected for its trade-off between flexibility and simplicity, not in the intent of building a biomechanical deformation model.

3.2. Diffusion Tensors and Geometry-Driven Regularization

The previous smoothness constraint is not always relevant. Optical flows, for instance, present discontinuities at the boundaries of moving objects. Such sharp variations in a displacement field are strongly penalized in the previous model. In this case, a geometry-driven regularization can help. We consider the mapping φ given by $\varphi : \mathbf{X} \mapsto \frac{1}{2}\mathrm{Tr}(X^T\mathbf{T}_f X)$, where $\mathbf{T}_f$ is a symmetric tensor characterized by some features of the target image. By computing

$$\varphi_{D\mathbf{u}}(D\mathbf{u}) = \frac{1}{2}\frac{\partial \mathrm{Tr}(D\mathbf{u}^T\mathbf{T}_f D\mathbf{u})}{\partial D\mathbf{u}} = \mathbf{T}_f D\mathbf{u},$$

we get a modified diffusion term $\mathbf{div}(\mathbf{T}_f D\mathbf{u})$. We would like to find a tensor $\mathbf{T}_f$ allowing large local variations by reducing the isotropic regularization near edges (characterized by high gradients of f). As suggested in [1], we could use $\mathbf{T}_f = \psi(||\nabla f||)\mathbf{I}$, where ψ is a strictly positive, decreasing function. A slightly different choice was initially suggested by Nagel and Henkelman [18]. They consider an anisotropic smoothing of $\mathbf{u}$ along object contours, by choosing $\mathbf{T}_f$ as a modified projection matrix. First used with a linearized optical flow equation [18], this tensor has been recently applied in a nonlinear/scale-space setting [2, 3]. The original 2D tensor can be extended to higher dimensions. We propose here a generic n-dimensional formulation :

$$\mathbf{T}_f = \frac{(\lambda + ||\nabla f||^2)\mathbf{I} - \nabla f \nabla f^T}{(n-1)||\nabla f||^2 + n\lambda}.$$

The sensitivity of the edge detection is adjusted with the variable λ.

4. Numerical Experiments

The spatial discretization of all the gradient flows is induced by image samples. The corresponding flows are applied to a set of smoothed and subsampled images. The considered functionals are non convex and this coarse to fine strategy avoids irrelevant extrema. A larger class of deformations can be recovered and the computational cost of our algorithms is usually decreased. Parzen-Rozenblatt estimates are obtained by convolution of joint histograms with 2D Gaussian kernels.

4.1. Experiments

Figure 1. Supervised registration (**SR**) method. A synthetic learning set is produced by applying a *sin* function and adding Gaussian noise (zero mean with 0.01 variance) to 4 images (whose intensity has been normalized on $[0, 2\pi]$). We use the comparison function h estimated from this learning set to match two views of a face. The non-monotonic functional dependence between these two intensity maps is clearly recovered from the estimated comparison function, and a visually correct realignment is achieved. Some small

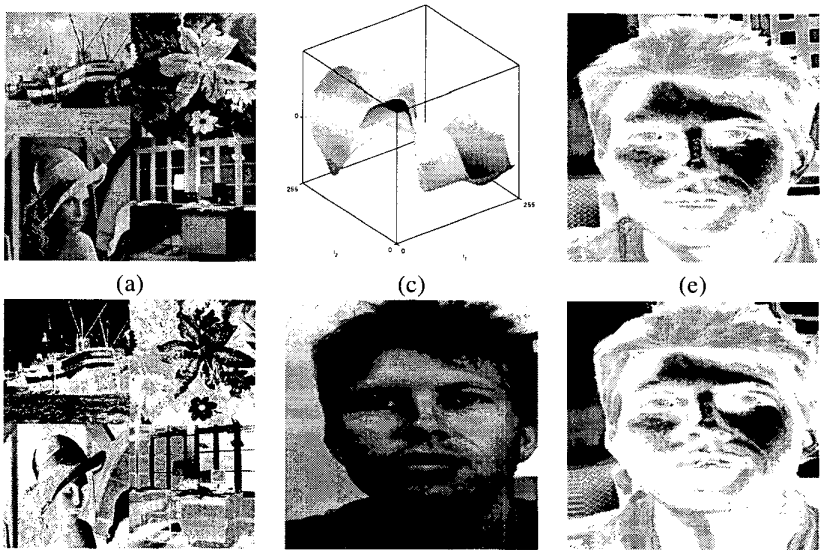

Figure 1. SR matching (with linear elasticity) : (a) and (b) learning sets : 4 pairs of 256x256 images, (c) estimated comparison function $h(i_1, i_2)$, (d) reference image (256x256), (e) image to register, (f) image (e) after realignment.

artefacts appear, mainly due to some elements in the background which are not present in both views. Regularization is provided by the linear elastic model.

Figure 2. We use the **MI** and **CR** criteria to realign two slices from a Proton Density (PD) and a T2 MRI volume (same patient). An artificial geometric distortion (based on a set of three Gaussian kernels) has been applied to the original pre-registered dataset. In order to evaluate the accuracy of the realignment, we superimposed some contours of the T2 image (initial and recovered pose) over the target image (PD). Most of the anatomical structures seem to be correctly realigned. In this case, the linear elastic regularization is used.

Figure 3. This geometric example is used to show the capacity of the **MI** criterion to cope with dependences which are not necessarily mono-functional. The algorithm succeeds in completely matching these two images with both linear elastic and geometry-driven regularizations. On this particular example, we clearly see the advantage of the last method. It preserves the natural discontinuities of the displacement field. This example underlines the role of the prior knowledge which is embedded in the regularization functional. Using such methods, we not only need to insure that two images are correctly realigned, but also that the recovered displacement field is meaningful with respect to the underlying cause of the deformation.

Figure 4. Application of the **MI** criterion to the realignment of 3D MRI datasets (using linear elasticity). The artificial change of intensity is identical to the one used in Figure 1. We applied a synthetic geometric distortion based on a set of eight Gaussian kernels (displacement with a maximum amplitude of 6.9 voxels). Some selected features (marked with black squares) suggest that the main structure of the underlying deformation has been recovered. Superimposed contours of the initial and final pose over an axial slice of the target volume confirm that the main anatomical structures are correctly realigned. A quantitative error study shows that 16.8 % of the displacement vectors whose norm was larger than one voxel have been recovered with a subvoxel accuracy. This percentage increases to 53.17 % for the displacement vectors whose norm was lower than 2 voxels. As we already noticed, such results are not fully relevant since different deformation fields can lead to meaningful realignments (see Figure 3.).

5. Conclusion

We developed a variational formulation for multi-modal image matching. The main contribution of this work lies in the derivation of local intensity comparison functions and dense displacement fields from statistical and information-theoretic similarity criteria. In this framework, one can use a large class of regularization methods to adapt the algorithms to problem-specific issues. In order to confirm the robustness and the accuracy of this approach for multi-modal MRI datasets, a more complete qualitative and quantitative experimental study must be carried out. It will form the next step of this research.

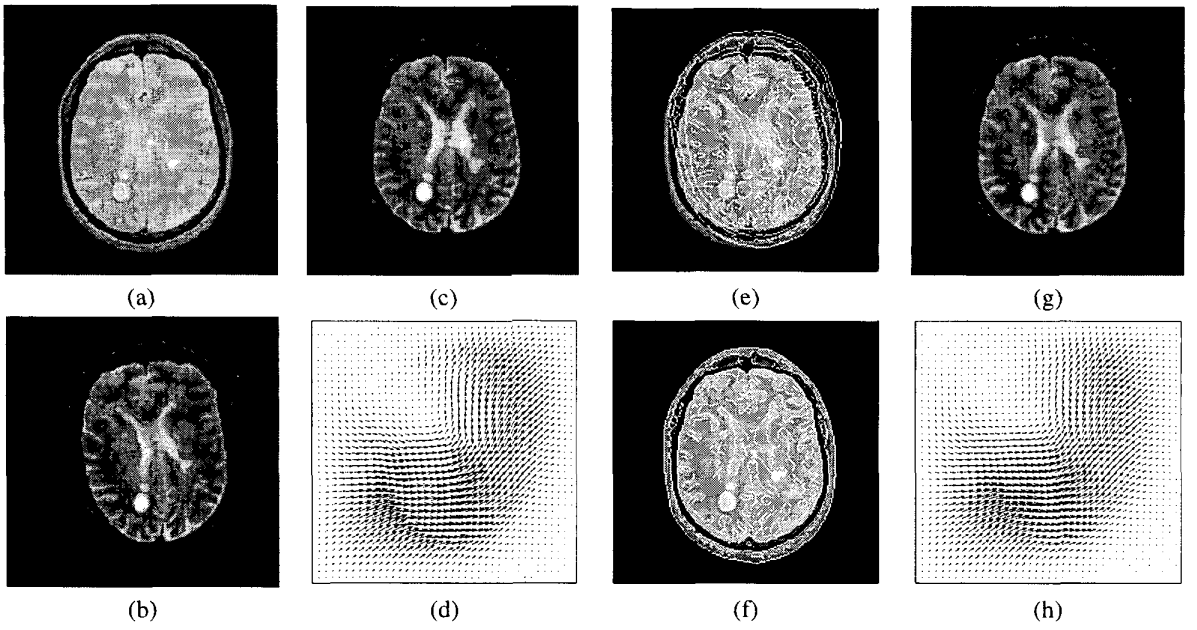

Figure 2. PD/T2 MRI realignment with MI and CR comparison functions (regularization using linear elasticity) : (a) reference image (PD 200x200), (b) image to register (T2 200x200), (c) image (b) after realignment (MI), (d) recovered displacement field (MI), (e) initial pose, (f) final pose, (g) image (b) after realignment (CR), (h) recovered displacement field (CR).

Acknowledgements. The authors would like to thank Javier Sánchez for fruitful discussions on geometry-driven diffusion tensors. The MRI data used in Figure 2 were provided by Dr. Charles R. G. Guttmann, Department of Radiology at Brigham and Women's Hospital. G. H. acknowledges the grant given by Conacyt. This research was partially supported by NSF (DMS 9872228) and the EC (QLG3-CT-2000-30161).

References

[1] L. Alvarez, J. Esclarín, M. Lefebure, and J. Sánchez. A pde model for computing the optical flow. In *Proceedings of CEDYA XVI*. Universidad de Las Palmas de Gran Canaria, 1999.

[2] L. Alvarez, J. Weickert, and J. Sánchez. A scale-space approach to nonlocal optical flow calculations. *Second International Conference on Scale Space, Lecture Notes in Computer Sciences*, 1682:235–246, 1999.

[3] L. Alvarez, J. Weickert, and J. Sánchez. Reliable Estimation of Dense Optical Flow Fields with Large Displacements. Technical report, Cuadernos del Instituto Universitario de Ciencias y Tecnologías Cibernéticas, 2000. A revised version of this paper will be published in the International Journal of Computer Vision.

[4] Y. Amit. A nonlinear variational problem for image matching. *SIAM Journal on Scientific Computing*, 15(1), Jan. 1994.

[5] D. Bosq. *Nonparametric Statistics for Stochastic Processes*, volume 110 of *Lecture Notes in Statistics*. Springer-Verlag, 2nd edition, 1998.

[6] G. Christensen, M. Miller, and M. Vannier. A 3d deformable magnetic resonance textbook based on elasticity. In *Proceedings of the AAAI, Symposium: Applications of Computer Vision in Medical Image Processing*, 1994.

[7] P. Ciarlet. *Mathematical Elasticity*, volume 1. North Holland, 1988.

[8] R. Courant. *Calculus of Variations*. New York, 1946.

[9] R. Deriche, P. Kornprobst., and G. Aubert. Optical flow estimation while preserving its discontinuities: a variational approach. In *Proceedings of the 2nd Asian Conference on Computer Vision*, volume 2/3, Singapore, Dec. 1995.

[10] L. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, 1998.

[11] T. Gaens, F. M. D. Vandermeulen, and P. Suetens. Non-rigid multimodal image registration using mutual information. In Wells et al. [25].

[12] N. Hata, T. Dohi, S. Warfield, W. W. III, R. Kikinis, and F. A. Jolesz. Multimodality deformable regristration of pre-and intra-operative images for mri-guided brain surgery. In Wells et al. [25].

[13] S. Haykin. *Neural Networks*. Prentice Hall, 1999.

[14] G. Hermosillo, C. Chefd'Hotel, and O. Faugeras. A variational approach to multi-modal image matching. RR 4117, INRIA, 2001.

[15] M. Leventon and W. Grimson. Multi-Modal Volume Registration Using Joint Intensity Distributions. In Wells et al. [25].

[16] J. Maintz, H. Meijering, and M. Viergever. General multimodal elastic registration based on mutual information. In *Medical Imaging 1998 - Image Processing*, volume 3338, pages 144–154. SPIE, 1998.

[17] C. Meyer, J. Boes, B. Kim, and P. Bland. Evaluation of control point selection in automatic, mutual information driven, 3d warping. In Wells et al. [25].

[18] H. Nagel and W. Enkelmann. An investigation of smoothness constraint for the estimation of displacement vector fields from images sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:565–593, 1986.

[19] A. Roche, A. Guimond, J. Meunier, and N. Ayache. Multimodal Elastic Matching of Brain Images. In *Proceedings of the 6th European Conference on Computer Vision*, Dublin, Ireland, June 2000.

[20] A. Roche, G. Malandain, X. Pennec, and N. Ayache. The correlation ratio as new similarity metric for multimodal image registration. In Wells et al. [25], pages 1115–1124.

[21] A. Roche, G. Malandain, X. Pennec, and N. Ayache. Multimodal image registration by maximization of the correlation ratio. Technical Report 3378, INRIA, Aug. 1998.

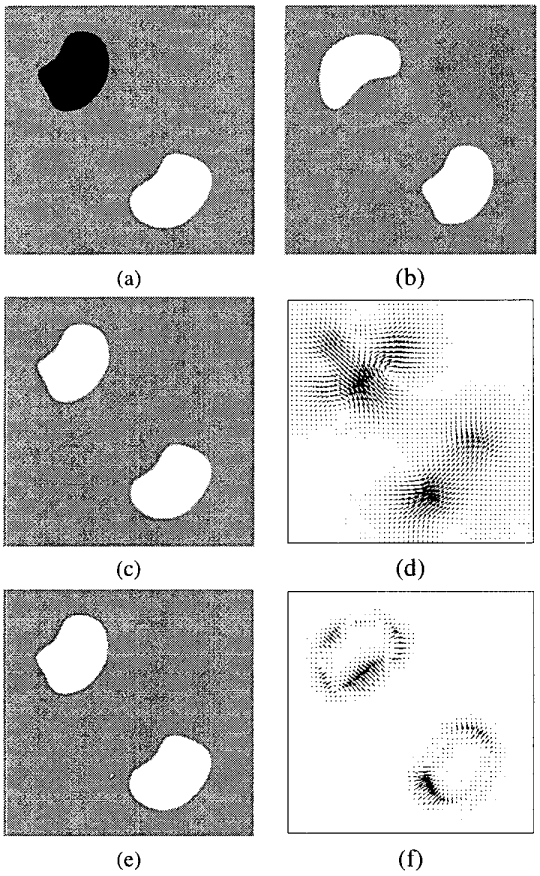

Figure 3. MI criterion with a synthetic example : (a) reference image (256x256), (b) image to register, (c) image (b) after realignment (with linear elasticity), (d) recovered displacement field (with linear elasticity), (e) image (b) after realignment (with geometry-driven regularization), (f) recovered displacement field (with geometry-driven regularization).

[22] D. Rückert, C. Hayes, C. Studholme, P. Summers, M. Leach, and D. Hawkes. Non-rigid registration of breast mr images using mutual information. In Wells et al. [25].

[23] A. Tikhonov and V. Arsenin. *Solutions of ill-posed problems*. Winston and Sons, Washington, D.C., 1977.

[24] P. Viola and W. M. Wells III. Alignement by maximization of mutual information. *The International Journal of Computer Vision*, 24(2):137–154, 1997.

[25] W. Wells, A. Colchester, and S. Delp, editors. Number 1496 in Lecture Notes in Computer Science, Cambridge, MA, USA, Oct. 1998. Springer.

[26] W. Wells III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis. Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis*, 1(1):35–51, 1996.

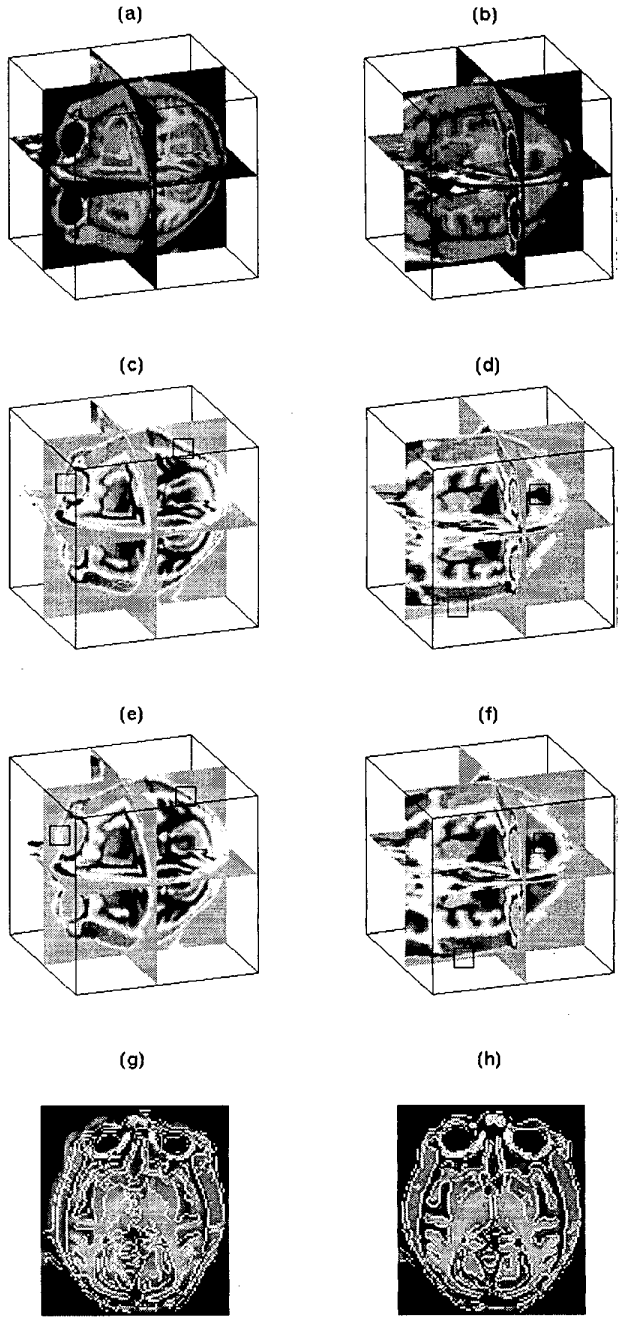

Figure 4. Realignment of MRI volumes (synthetic example with MI and linear elasticity) : (a) and (b) two views of the reference volume (T2 66x111x97), (c) and (d) volume to register (66x111x97), (e) and (f) volume after realignment, (g) and (h) contours of the deformed volume superimposed over an axial slice of the target volume (initial and final pose).