



A Variational Approach to Multi-Modal Image Matching

Gerardo Hermosillo, Christophe Chefd'Hotel, Olivier Faugeras

► To cite this version:

Gerardo Hermosillo, Christophe Chefd'Hotel, Olivier Faugeras. A Variational Approach to Multi-Modal Image Matching. RR-4117, INRIA. 2001. inria-00072513

HAL Id: inria-00072513

<https://inria.hal.science/inria-00072513>

Submitted on 24 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Variational Approach to Multi-Modal Image Matching

Gerardo Hermosillo Christophe Chef d'Hotel and Olivier Faugeras

N° 4117

February 16, 2001

_____ THÈME 3 _____



*rapport
de recherche*

A Variational Approach to Multi-Modal Image Matching

Gerardo Hermosillo Christophe Chéfd'Hotel and Olivier Faugeras

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

Rapport de recherche n° 4117 — February 16, 2001 — 34 pages

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.

Key-words: *Multi-modal Image Matching, Variational Methods, Registration, Optical Flow, Mutual Information, Correlation Ratio.*

A short version of this report has been submitted to the IEEE Workshop on Variational and Level Set Methods in Computer Vision (VLISM01).

This work was partially supported by EC contract No QLG3-CT-2000-30161, Mapawamo.

Une Approche Variationnelle pour l'Appariement Multimodal d'Images

Résumé : Nous considérons le problème du recalage non-rigide entre images de modalités différentes. Nous proposons un cadre général qui repose sur une formulation variationnelle, que nous appliquons sous la forme de trois algorithmes de recalage multimodal : recalage supervisé par apprentissage de la loi jointe, maximisation de l'information mutuelle, et maximisation du rapport de corrélation. Pour permettre une régularisation de la solution, nous utilisons un opérateur issu de la théorie de l'élasticité. Nous considérons aussi une méthode de régularisation avec préservation des contours. Des résultats expérimentaux préliminaires sur des images synthétiques et des données IRM sont présentés.

Mots-clés : *Mise en Correspondence Multimodale, Méthodes Variationnelles, Flot Optique, Information Mutuelle, Rapport de Corrélation.*

Contents

1	Introduction	4
2	Variational Principles and Multi-Modal Similarity Measures	6
2.1	Prolegomenon	6
2.2	Supervised Registration and Joint Intensity Learning	8
2.3	Information-Theoretic and Statistical Criteria: Gradient Flows of Mutual Information and Correlation Ratio	11
2.3.1	Mutual Information	13
2.3.2	Correlation Ratio	15
2.4	Summary and comments	17
3	Regularization	19
3.1	Linear Elasticity	19
3.2	Diffusion Tensors and Geometry-Driven Regularization	20
3.3	Remarks on Other Regularization Methods	21
4	Numerical Experiments	22
4.1	Implementation issues	22
4.2	Experiments	22
5	Conclusions	27

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 been developed in order 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 the same intensity range, up to a certain amount of noise. Consequently, the resulting matching algorithms are essentially based on the same principle. Given two images, one looks for a geometric transformation which minimizes the sum of the differences (or the squared differences) between their intensity values. In this setting, the main issue usually concerns the choice of a class of deformations, a finite or infinite-dimensional search space.

However, there are several situations where the considered images have different intensity maps. 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 sensibilities. The medical imaging community has been confronted very early to the same type of problem. There exist several MRI acquisition modalities, and intra-modality registrations must be performed to allow an accurate fusion and comparison of complementary information. One must notice that the relation between the intensity maps of two modalities is not necessarily functional, and is generally unknown (at least not explicitly).

To cope with this new difficulty, global statistical and information-theoretic similarity measures have been proposed, such as the mutual information [30, 31] and the correlation ratio [23, 24]. Multi-modal registration is usually performed by maximizing these criteria over a low-dimensional parametric class of deformations, e.g. rigid or affine transformations. In some cases, the relation between two modal-

ities can even be learned under reasonable assumptions, and a maximum likelihood principle leads to a very efficient multi-modal registration method [16].

Recent attempts have been made to extend these techniques to the non-rigid multi-modal image matching problem. They either rely on more complex parametric transformations [18, 25], block-matching strategies [17, 14, 13], or on the combination of a mono-modal algorithm with some parametric intensity corrections [22]. However, to our knowledge, no global variational formulation of this problem has been proposed, limiting the access to very useful classes of deformations.

Variational techniques have already been used for classical optical flow computations and mono-modal image matching problems [20, 5, 11, 3] (among others), showing that excellent results can be achieved through this approach. Prior knowledge and advanced regularization constraints can be easily integrated in a single cost functional. This is the natural approach to design gradient flows on infinite-dimensional deformation spaces. In this perspective, we propose a variational framework to recover dense displacement fields from global 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. Regularization is performed by using a physically-based functional relying on elasticity theory. We also consider a functional based on discontinuity-preserving diffusion tensors, first introduced by Nagel and Enkelman [20], and which has recently attracted a new interest in image restoration, optical flow and stereo [3, 2].

This paper is organized as follows. Section 2 discusses the variational formalism and its application to the definition of three matching algorithms. Section 3 describes the choice of a functional to perform a suitable regularization. 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

2.1 Prolegomenon

Let us first introduce some notations and basic assumptions which will be used all along this paper. We are mainly concerned by image matching problems (2D and 3D), but most the following developments will remain valid for any scalar signals defined on a bounded domain $\Omega \subset \mathbb{R}^n$ (with an arbitrary n). For instance, one may want to handle 1D matching problems, such as audio signals time warping. We consider here two images, respectively represented by the functions $f : \Omega \mapsto \mathbb{R}$ and $g : \Omega \mapsto \mathbb{R}$. 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. The deformation is modelled by a mapping $\mathbf{v} : \Omega \subset \mathbb{R}^n \mapsto \Omega$. We assume that \mathbf{v} belongs to a Hilbert space \mathcal{F} equipped with the usual scalar product $\langle \mathbf{v}, \mathbf{w} \rangle_{\mathcal{F}} = \int_{\Omega} \mathbf{v}(\mathbf{x}) \cdot \mathbf{w}(\mathbf{x}) \, d\mathbf{x}$. It is often simpler to decompose \mathbf{v} as $\mathbf{v} = \text{Id} + \mathbf{u}$ where \mathbf{u} is a displacement field.

The generic matching problem is defined as the minimization of a cost functional

$$\mathcal{I}[\mathbf{u}] = \mathcal{J}[\mathbf{u}] + \alpha \mathcal{R}[\mathbf{u}]$$

where \mathcal{J} measures the disparity of f and g given a displacement \mathbf{u} , while \mathcal{R} contains regularizing constraints and smoothness assumptions on \mathbf{u} . For mono-modal matchings, a typical choice of functional \mathcal{J} would be the Sum of Squared Differences (**SSD**)

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

If we measure the irregularity of \mathbf{u} in terms of the amplitude of its variations, a very simple example for \mathcal{R} is given by

$$\mathcal{R}[\mathbf{u}] = \frac{1}{2} \int_{\Omega} \|D\mathbf{u}(\mathbf{x})\|^2 \, d\mathbf{x} = \frac{1}{2} \int_{\Omega} \text{Tr}(D\mathbf{u}(\mathbf{x})^T D\mathbf{u}(\mathbf{x})) \, d\mathbf{x}, \quad (2)$$

where $D\mathbf{u}$ denotes the Jacobian matrix of \mathbf{u} . This type of regularization goes back to the so-called *Tikhonov* functional [28], first used to deal with ill-posed problems. The factor α is introduced to control the respective influence of \mathcal{J} and \mathcal{R} .

In the following, we propose to consider a more generic regularization term \mathcal{R} , such that

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

for an arbitrary mapping $\varphi : \mathcal{M}_{n \times n}(\mathbb{R}) \mapsto \mathbb{R}_+$, where $\mathcal{M}_{n \times n}(\mathbb{R})$ is the set of $n \times n$ matrices with real coefficients. Note that (2) corresponds to $\varphi(\cdot) = \frac{1}{2} \|\cdot\|^2$. 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.

In general, we assume the existence of a minimum $\hat{\mathbf{u}} = \operatorname{argmin}_{\mathbf{u} \in \mathcal{F}} \mathcal{I}[\mathbf{u}]$. Using some classical tools of variational calculus (see [10] for a complete introduction to this topic), we can introduce the first variation (Gâteaux Variation) of the functional \mathcal{I} , defined as

$$\text{for } \mathbf{h} \in \mathcal{F}, \quad d\mathcal{I}[\mathbf{u}, \mathbf{h}] = \lim_{\epsilon \rightarrow 0} \frac{\mathcal{I}[\mathbf{u} + \epsilon \mathbf{h}] - \mathcal{I}[\mathbf{u}]}{\epsilon} = \left. \frac{\partial \mathcal{I}[\mathbf{u} + \epsilon \mathbf{h}]}{\partial \epsilon} \right|_{\epsilon=0}$$

One can show that if \mathcal{I} has a local extremum \mathbf{u} in \mathcal{F} , its first variation at \mathbf{u} must vanish, that is

$$\forall \mathbf{h} \in \mathcal{F}, \quad d\mathcal{I}[\mathbf{u}, \mathbf{h}] = 0 \text{ must hold.}$$

Let us notice that if $d\mathcal{I}[\mathbf{u}, \mathbf{h}] = \langle \nabla_{\mathbf{u}} \mathcal{I}, \mathbf{h} \rangle_{\mathcal{F}}$, $\nabla_{\mathbf{u}} \mathcal{I}$ defines the gradient of the functional \mathcal{I} . Therefore, the previous necessary condition of optimality is readily equivalent to

$$\nabla_{\mathbf{u}} \mathcal{I} = \mathbf{0}$$

This formula is often referred to as the Euler or Euler-Lagrange equation. It can be easily computed when \mathcal{I} has a simple form. For instance, if

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

one can show that its Euler equation becomes

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

where we define a generalized divergence operator $\operatorname{div}(\mathbf{T}) = [\nabla \cdot \mathbf{t}_1, \dots, \nabla \cdot \mathbf{t}_n]^T$ for $\mathbf{T} = [\mathbf{t}_1, \dots, \mathbf{t}_n]^T$.

In our situation, we will not try directly to solve the Euler equation and prefer a suboptimal solution based on a gradient descent strategy. With non-convex criteria, one may of course get trapped in local extrema. Nevertheless, provided with a suitable initial guess \mathbf{u}_0 , numerical implementations based on multi-resolution methods will generally provide relevant results. The direction of the steepest descent depends on the choice of the scalar product $\langle \cdot, \cdot \rangle_{\mathcal{F}}$. With the previous definition, we get the classical gradient flow

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

With these tools, we can now focus our attention on the design of a matching functional \mathcal{I} allowing us to cope with different modalities. It requires the introduction of more complex disparity measures in the functional \mathcal{J} , since a simple **SSD** criterion becomes meaningless.

2.2 Supervised Registration and Joint Intensity Learning

In order to match images from two different modalities, several strategies are possible. They can be divided into two categories. Some solutions consist in using relevant geometric features, such as edges and corners, whose definition relies essentially on differences between intensity levels. Hopefully, these features are robust to global changes of intensity scale. A second way, which will be explored here, relies on a possible description of the relation between the intensity maps of two different modalities. A probabilistic formalism can be easily introduced in this framework.

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. Note that this approach seems a priori incompatible with our previous modelling assumptions. In fact, we need to transpose this setting to the continuous case, and consider strictly stationary processes whose time parameter is a vector defined on Ω . Consequently, f and g become some sample paths of two random processes. Stationarity is definitely a very strong assumption

but provides us with some simple and useful mathematical tools. Ultimately, successful experiments will support this questionable modelling assumption.

We denote by p^f the intensity distribution estimated from f . The image g undergoing a displacement field \mathbf{u} becomes $g \circ (\text{Id} + \mathbf{u})$ and leads to the estimated density $p_{\mathbf{u}}^g$. Finally, $p_{\mathbf{u}}^{f,g}$ stands for an estimate of their joint intensity distribution. By analogy with the discrete case, we introduce two random variables, X^f and $X_{\mathbf{u}}^g$, whose probability densities are respectively given by p^f and $p_{\mathbf{u}}^g$. Note that within this probabilistic framework, the link between two modalities is fully characterized by a joint density.

We first assume that reference templates, or sets of manually pre-registered images, are available to estimate the real joint intensity distribution p . In practice, a non-parametric Parzen-Rozenblatt estimator [6, 15] can be used. From the knowledge of p , we can 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 nothing but the co-occurrence 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 displacement \mathbf{u} :

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

One may also give an interpretation of this expression as a continuous formulation of the maximum likelihood principle developed in [16].

Combined with (3), these considerations are summarized by the following matching functional :

$$\mathcal{I}_{\text{SR}}[\mathbf{u}] = \int_{\Omega} -\log p(f(\mathbf{x}), g(\mathbf{x} + \mathbf{u}(\mathbf{x}))) d\mathbf{x} + \alpha \int_{\Omega} \varphi(D\mathbf{u}(\mathbf{x})) d\mathbf{x}$$

In this simple case, the explicit computation of the first variation is not required. Indeed, we can rewrite

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

With a differentiable and strictly positive estimate p of the joint intensity distribution, and provided φ , f and g enjoy the appropriate regularity, ϕ is continuously differentiable with respect to each component of \mathbf{u} and $D\mathbf{u}$, thus

$$\begin{aligned}\phi_{\mathbf{u}}(\mathbf{x}, \mathbf{u}, D\mathbf{u}) &= -\frac{\partial \log p(f, g \circ (\mathbf{Id} + \mathbf{u}))}{\partial \mathbf{u}} \\ &= -\frac{p_{i_2}(f, g \circ (\mathbf{Id} + \mathbf{u}))}{p(f, g \circ (\mathbf{Id} + \mathbf{u}))} \nabla g \circ (\mathbf{Id} + \mathbf{u}),\end{aligned}\quad (5)$$

where p_{i_2} denotes the partial derivative of p with respect to its second variable. Moreover,

$$\phi_{D\mathbf{u}}(\mathbf{x}, \mathbf{u}, D\mathbf{u}) = \varphi_{D\mathbf{u}}(D\mathbf{u}) \quad (6)$$

Finally, combining (5) and (6) with (4), we get the following Euler equation :

$$-\frac{p_{i_2}(f, g \circ (\mathbf{Id} + \mathbf{u}))}{p(f, g \circ (\mathbf{Id} + \mathbf{u}))} \nabla g \circ (\mathbf{Id} + \mathbf{u}) - \alpha \operatorname{div}(\varphi_{D\mathbf{u}}(D\mathbf{u})) = 0$$

In effect, given an initial displacement field \mathbf{u}_0 , we consider a suboptimal solution by building the following gradient flow :

$$\begin{cases} \mathbf{u}_t &= \frac{p_{i_2}(f, g \circ (\mathbf{Id} + \mathbf{u}))}{p(f, g \circ (\mathbf{Id} + \mathbf{u}))} \nabla g \circ (\mathbf{Id} + \mathbf{u}) + \alpha \operatorname{div}(\varphi_{D\mathbf{u}}(D\mathbf{u})) \\ \mathbf{u}(0, \cdot) &= \mathbf{u}_0 \end{cases} \quad (7)$$

Boundary conditions may apply. In practice, we assume either Neumann or Dirichlet conditions.

It is particularly interesting to compare this gradient flow with the one based on the **SSD** criterion (see equation (1)). In this case, we would get

$$\begin{cases} \mathbf{u}_t &= (f - g \circ (\mathbf{Id} + \mathbf{u})) \nabla g \circ (\mathbf{Id} + \mathbf{u}) + \alpha \operatorname{div}(\varphi_{D\mathbf{u}}(D\mathbf{u})) \\ \mathbf{u}(0, \cdot) &= \mathbf{u}_0 \end{cases} \quad (8)$$

We immediately notice the analogy between the function $\frac{p_{i_2}(i_1, i_2)}{p(i_1, i_2)}$ in equation (7), and the intensity comparison function $i_1 - i_2$ of equation (8). 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. Notice that 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 that we don't 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 recovered from this initial estimate. In the following, we will see that this heuristic appears implicitly in the mutual information similarity measure.

2.3 Information-Theoretic and Statistical Criteria: Gradient Flows of Mutual Information and Correlation Ratio

When no training set is available, we may want 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 the image matching context. 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 [30, 31]. Using our notation, the mutual information computed from f and g (given \mathbf{u}) is provided by

$$\mathbf{MI}_{\mathbf{u}}^{f,g} = \mathcal{H}[X_u^g] - \mathcal{H}[X_u^g | X^f] = \mathcal{H}[X^f] - \mathcal{H}[X^f | X_u^g] = \mathcal{H}[X_u^g] + \mathcal{H}[X^f] - \mathcal{H}[X^f, X_u^g]$$

where \mathcal{H} stands for the differential entropy. This similarity measure is positive and symmetric. We often find more convenient to write the mutual information in terms of a Kullback-Leibler divergence

$$\mathbf{MI}_{\mathbf{u}}^{f,g} = \mathcal{D}_{p_{\mathbf{u}}^{f,g} || p^f p_{\mathbf{u}}^g} = \int_{\mathbb{R}^2} p_{\mathbf{u}}^{f,g}(i_1, i_2) \log \frac{p_{\mathbf{u}}^{f,g}(i_1, i_2)}{p^f(i_1) p_{\mathbf{u}}^g(i_2)} di_1 di_2$$

$\mathcal{D}_{p_a || p_b}$ defines a non symmetric distance between two probability densities p_a and p_b [15]. From this viewpoint, the mutual information indicates how the intensity distributions of two images fail to be independent.

On the other hand, the correlation ratio has been proposed in [24] to complete multi-modal affine registrations. This criterion relies on a slightly different notion of similarity. From the definition

$$\mathbf{CR}_{\mathbf{u}}^{f,g} = \frac{\mathbf{Var}[\mathbf{E}[X_{\mathbf{u}}^g | X^f]]}{\mathbf{Var}[X_{\mathbf{u}}^g]}, \quad (9)$$

$\mathbf{CR}_{\mathbf{u}}^{f,g}$ 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 dependency between $X_{\mathbf{u}}^g$ and X^f :

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

We must notice that this measure is not symmetric. If the real dependence is given by a non monotonic function, the registration may fail if the suitable random variable order is not selected. Note that correlation ratio and mutual information are equivalent for a pair of gaussian random variables [24, 15].

We propose to include these criteria in the definition of the functional \mathcal{J} . In both cases, the computation of their first variation is carried out by considering the variations of some joint density estimates according to \mathbf{u} . To insure the compatibility of this approach with a variational setting, a continuous form of Parzen-Rozenblatt estimators [6] is used. It provides us with a simple way to quantify the contribution of infinitesimal variations of \mathbf{u} to these global similarity criteria.

2.3.1 Mutual Information

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

$$\mathcal{I}_{\text{MI}}[\mathbf{u}] = -\text{MI}_{\mathbf{u}}^{f,g} + \alpha \int_{\Omega} \varphi(D\mathbf{u}) \, d\mathbf{x} \quad (10)$$

Due to the the complex form of $\mathcal{J}_{\text{MI}}[\mathbf{u}] = -\text{MI}_{\mathbf{u}}^{f,g}$, an explicit computation of its first variation must be carried out. For this purpose, we recall the following definition

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

We have

$$\begin{aligned} \frac{\partial \mathcal{J}_{\text{MI}}[\mathbf{u} + \epsilon \mathbf{h}]}{\partial \epsilon} &= - \int_{\mathbb{R}^2} \frac{\partial}{\partial \epsilon} \left[p_{\mathbf{u}+\epsilon \mathbf{h}}^{f,g}(i_1, i_2) \log \frac{p_{\mathbf{u}+\epsilon \mathbf{h}}^{f,g}(i_1, i_2)}{p^f(i_1) p_{\mathbf{u}+\epsilon \mathbf{h}}^g(i_2)} \right] di_1 di_2 \\ &= - \int_{\mathbb{R}^2} \underbrace{\left(1 + \log \frac{p_{\mathbf{u}+\epsilon \mathbf{h}}^{f,g}(i_1, i_2)}{p^f(i_1) p_{\mathbf{u}+\epsilon \mathbf{h}}^g(i_2)} \right)}_{L^{\mathbf{u}+\epsilon \mathbf{h}}(i_1, i_2)} \frac{\partial p_{\mathbf{u}+\epsilon \mathbf{h}}^{f,g}(i_1, i_2)}{\partial \epsilon} di_1 di_2 \\ &\quad - \underbrace{\int_{\mathbb{R}^2} \frac{p_{\mathbf{u}+\epsilon \mathbf{h}}^{f,g}(i_1, i_2)}{p_{\mathbf{u}+\epsilon \mathbf{h}}^g(i_2)} \frac{\partial p_{\mathbf{u}+\epsilon \mathbf{h}}^g(i_2)}{\partial \epsilon} di_1 di_2}_P \end{aligned}$$

We first notice that

$$\begin{aligned} P &= \int_{\mathbb{R}} \frac{\partial p_{\mathbf{u}+\epsilon \mathbf{h}}^g(i_2)}{\partial \epsilon} \frac{1}{p_{\mathbf{u}+\epsilon \mathbf{h}}^g(i_2)} \underbrace{\left(\int_{\mathbb{R}} p_{\mathbf{u}+\epsilon \mathbf{h}}^{f,g}(i_1, i_2) di_1 \right)}_{p_{\mathbf{u}+\epsilon \mathbf{h}}^g(i_2)} di_2 = \\ &\quad \frac{\partial}{\partial \epsilon} \underbrace{\left[\int_{\mathbb{R}} p_{\mathbf{u}+\epsilon \mathbf{h}}^g(i_2) di_2 \right]}_1 = 0 \end{aligned}$$

Assuming a displacement field $\mathbf{u} + \epsilon \mathbf{h}$, the joint intensity distribution estimated from f and g is provided by a non parametric Parzen-Rozenblatt density model

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

where $\mu(\Omega)$ denotes the volume of Ω and $\psi(\alpha, \beta)$ is a bidimensional density kernel (strictly positive and differentiable). This is a straightforward generalization of the expression given in [6] for continuous time random processes. Derivatives of $p_{\mathbf{u} + \epsilon \mathbf{h}}^{f,g}(i_1, i_2)$ can be easily computed. In particular,

$$\begin{aligned} \frac{\partial p_{\mathbf{u} + \epsilon \mathbf{h}}^{f,g}(i_1, i_2)}{\partial \epsilon} &= \frac{1}{\mu(\Omega)} \int_{\Omega} \psi_{\beta}(f(\mathbf{x}) - i_1, g(\mathbf{x} + \mathbf{u}(\mathbf{x}) + \epsilon \mathbf{h}(\mathbf{x})) - i_2) \\ &\quad \nabla g(\mathbf{x} + \mathbf{u}(\mathbf{x}) + \epsilon \mathbf{h}(\mathbf{x})) \cdot \mathbf{h}(\mathbf{x}) d\mathbf{x} \end{aligned}$$

We now let $\epsilon = 0$,

$$\begin{aligned} d\mathcal{J}_{\text{MI}}[\mathbf{u}, \mathbf{h}] &= -\frac{1}{\mu(\Omega)} \int_{\mathbb{R}^2} \int_{\Omega} L^{\mathbf{u}}(i_1, i_2) \psi_{\beta}(f(\mathbf{x}) - i_1, g(\mathbf{x} + \mathbf{u}(\mathbf{x})) - i_2) \\ &\quad \nabla g(\mathbf{x} + \mathbf{u}(\mathbf{x})) \cdot \mathbf{h}(\mathbf{x}) d\mathbf{x} di_1 di_2 \end{aligned}$$

The convolution operator, denoted \star , appears naturally. It commutes with the derivation operator $\frac{\partial}{\partial i_2}$ and therefore

$$\begin{aligned} d\mathcal{J}_{\text{MI}}[\mathbf{u}, \mathbf{h}] &= -\frac{1}{\mu(\Omega)} \int_{\Omega} [\psi \star \frac{\partial L^{\mathbf{u}}}{\partial i_2}](f(\mathbf{x}), g(\mathbf{x} + \mathbf{u}(\mathbf{x}))) \nabla g(\mathbf{x} + \mathbf{u}(\mathbf{x})) \cdot \mathbf{h}(\mathbf{x}) d\mathbf{x} \\ &= \left\langle -\frac{1}{\mu(\Omega)} [\psi \star \frac{\partial L^{\mathbf{u}}}{\partial i_2}](f, g \circ (\text{Id} + \mathbf{u})) \nabla g \circ (\text{Id} + \mathbf{u}), \mathbf{h} \right\rangle_{\mathcal{F}} \end{aligned}$$

Finally, the Euler-Lagrange equation of (10) is found to be equal to

$$\begin{aligned} \forall \mathbf{x} \in \Omega, \quad &-\frac{1}{\mu(\Omega)} [\psi \star \frac{\partial L^{\mathbf{u}}}{\partial i_2}](f(\mathbf{x}), g(\mathbf{x} + \mathbf{u}(\mathbf{x}))) \\ &\nabla g(\mathbf{x} + \mathbf{u}(\mathbf{x})) - \alpha \mathbf{div}(\varphi_{D\mathbf{u}}(D\mathbf{u})) = \mathbf{0}, \end{aligned}$$

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

$$\begin{cases} \mathbf{u}_t &= \frac{1}{\mu(\Omega)} [\psi \star \frac{\partial L^{\mathbf{u}}}{\partial i_2}] (f, g \circ (\mathbf{Id} + \mathbf{u})) \nabla g \circ (\mathbf{Id} + \mathbf{u}) + \alpha \operatorname{div}(\varphi_{D\mathbf{u}}(D\mathbf{u})) \\ \mathbf{u}(0, \cdot) &= \mathbf{u}_0 \end{cases} \quad (11)$$

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 direct extension of the heuristic we introduced earlier. As a matter of fact, this method performs registration by successive adjustments of its joint intensity model. Its first estimate is simply based on the initial pose.

2.3.2 Correlation Ratio

In the same setting, we now consider the correlation ratio, with the functional

$$\mathcal{I}_{\text{CR}}[\mathbf{u}] = \underbrace{-\text{CR}_{\mathbf{u}}^{f,g}}_{\mathcal{J}_{\text{CR}}[\mathbf{u}]} + \alpha \int_{\Omega} \varphi(D\mathbf{u}) \, d\mathbf{x} \quad (12)$$

Instead of using the original definition (9) of $\text{CR}_{\mathbf{u}}^{f,g}$, we can use the total variance theorem to obtain

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

This transformation was suggested in [23, 24], and we can notice that \mathcal{J}_{CR} might be directly replaced by the second term of this formula. In order to compute the first variation, we can write

$$\frac{\partial \mathcal{J}_{\text{CR}}[\mathbf{u} + \epsilon \mathbf{h}]}{\partial \epsilon} = \frac{\left(\frac{\partial \mathbf{E}[\text{Var}[X_{\mathbf{u}+\epsilon \mathbf{h}}^g | X^f]]}{\partial \epsilon} + (\text{CR}_{\mathbf{u}+\epsilon \mathbf{h}}^{f,g} - 1) \frac{\partial \text{Var}[X_{\mathbf{u}+\epsilon \mathbf{h}}^g]}{\partial \epsilon} \right)}{\text{Var}[X_{\mathbf{u}+\epsilon \mathbf{h}}^g]}$$

Since

$$\mathbf{Var}[X_{\mathbf{u}+\epsilon\mathbf{h}}^g | X^f = i_1] = \int_{\mathbb{R}} i_2^2 \frac{p_{\mathbf{u}+\epsilon\mathbf{h}}^{f,g}(i_1, i_2)}{p^f(i_1)} di_2 - \left(\int_{\mathbb{R}} i_2 \frac{p_{\mathbf{u}+\epsilon\mathbf{h}}^{f,g}(i_1, i_2)}{p^f(i_1)} di_2 \right)^2,$$

we have

$$\begin{aligned} \frac{\partial \mathbf{E}[\mathbf{Var}[X_{\mathbf{u}+\epsilon\mathbf{h}}^g | X^f]]}{\partial \epsilon} &= \frac{\partial}{\partial \epsilon} \left[\int_{\mathbb{R}} \mathbf{Var}[X_{\mathbf{u}+\epsilon\mathbf{h}}^g | X^f = i_1] p^f(i_1) di_1 \right] \\ &= \int_{\mathbb{R}^2} i_2 (i_2 - 2\mathbf{E}[X_{\mathbf{u}+\epsilon\mathbf{h}}^g | X^f = i_1]) \frac{\partial p_{\mathbf{u}+\epsilon\mathbf{h}}^{f,g}(i_1, i_2)}{\partial \epsilon} di_1 di_2 \end{aligned}$$

Similarly,

$$\frac{\partial \mathbf{Var}[X_{\mathbf{u}+\epsilon\mathbf{h}}^g]}{\partial \epsilon} = \int_{\mathbb{R}^2} i_2 (i_2 - 2\mathbf{E}[X_{\mathbf{u}+\epsilon\mathbf{h}}^g]) \frac{\partial p_{\mathbf{u}+\epsilon\mathbf{h}}^{f,g}(i_1, i_2)}{\partial \epsilon} di_1 di_2$$

We finally get

$$\frac{\partial \mathcal{J}_{\mathbf{CR}}[\mathbf{u} + \epsilon\mathbf{h}]}{\partial \epsilon} = \int_{\mathbb{R}^2} L^{\mathbf{u}+\epsilon\mathbf{h}}(i_1, i_2) \frac{\partial p_{\mathbf{u}+\epsilon\mathbf{h}}^{f,g}(i_1, i_2)}{\partial \epsilon} di_1 di_2,$$

with

$$\begin{aligned} L^{\mathbf{u}+\epsilon\mathbf{h}}(i_1, i_2) &= \frac{i_2}{\mathbf{Var}[X_{\mathbf{u}+\epsilon\mathbf{h}}^g]} ((i_2 - 2\mathbf{E}[X_{\mathbf{u}+\epsilon\mathbf{h}}^g | X^f = i_1]) \\ &\quad + (\mathbf{CR}_{\mathbf{u}+\epsilon\mathbf{h}}^{f,g} - 1)(i_2 - 2\mathbf{E}[X_{\mathbf{u}+\epsilon\mathbf{h}}^g])) \end{aligned}$$

We now let $\epsilon = 0$. Similar remarks apply to the study of the mutual information and the present situation. Thus,

$$\begin{aligned} d\mathcal{J}_{\mathbf{CR}}[\mathbf{u}, \mathbf{h}] &= \int_{\Omega} \frac{1}{\mu(\Omega)} [\psi \star \frac{\partial L^{\mathbf{u}}}{\partial i_2}](f(\mathbf{x}), g(\mathbf{x} + \mathbf{u}(\mathbf{x}))) \nabla g(\mathbf{x} + \mathbf{u}(\mathbf{x})) \cdot \mathbf{h}(\mathbf{x}) d\mathbf{x} \\ &= \left\langle \frac{1}{\mu(\Omega)} [\psi \star \frac{\partial L^{\mathbf{u}}}{\partial i_2}](f, g \circ (\mathbf{Id} + \mathbf{u})) \nabla g \circ (\mathbf{Id} + \mathbf{u}), \mathbf{h} \right\rangle_{\mathcal{F}} \end{aligned}$$

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

$$\begin{cases} \mathbf{u}_t &= -\frac{1}{\mu(\Omega)}[\psi \star \frac{\partial L^{\mathbf{u}}}{\partial i_2}](f, g \circ (\mathbf{Id} + \mathbf{u})) \nabla g \circ (\mathbf{Id} + \mathbf{u}) + \alpha \operatorname{div}(\varphi_{D\mathbf{u}}(D\mathbf{u})) \\ \mathbf{u}(0, \cdot) &= \mathbf{u}_0 \end{cases} \quad (13)$$

with

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

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

2.4 Summary and comments

We assumed that joint intensity distributions fully characterize the relation between the intensity maps of different modalities. Either this density can be computed 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. This function offers a large flexibility by somehow integrating multiple correspondences for a single intensity value. In this case, given an intensity value in one modality, the comparison function may be non monotonic in terms of the intensities of the second modality. The related gradient flows defines the so-called supervised registration.

When the density is unknown, its estimate is built from the initial pose, and is continuously re-estimated along with the gradient descent process. The mutual information and the correlation ratio illustrate this approach. Their main difference lies on how each method builds its intensity comparison function from the estimated joint density. As for the supervised method, the mutual information depends on a directional derivative of the joint intensity distribution, while the correlation ratio

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	$\frac{1}{\mu(\Omega)} \left[\psi \star \left(\frac{1}{p_{\mathbf{u}}^{f,g}} \frac{\partial p_{\mathbf{u}}^{f,g}}{\partial i_2} - \frac{1}{p_{\mathbf{u}}^g} \frac{\partial p_{\mathbf{u}}^g}{\partial i_2} \right) \right] (i_1, i_2)$
CR	$\frac{-2 \psi \star \left(i_2 - \mathbf{E}[X_{\mathbf{u}}^g X^f = i_1] + (\mathbf{CR}_{\mathbf{u}}^{f,g} - 1)(i_2 - \mathbf{E}[X_{\mathbf{u}}^g]) \right)}{\mathbf{Var}[X_{\mathbf{u}}^g] \mu(\Omega)}$

Table 1: Comparison functions.

relies mainly on a unique functional dependence computed from the first conditional moments. These two different models suggest that the choice between these two methods is given by 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

$$\begin{cases} \mathbf{u}_t &= h(f, g \circ (\mathbf{Id} + \mathbf{u})) \nabla g \circ (\mathbf{Id} + \mathbf{u}) + \alpha \operatorname{div}(\varphi_{D\mathbf{u}}(D\mathbf{u})) \\ \mathbf{u}(0, \cdot) &= \mathbf{u}_0 \end{cases} \quad (14)$$

It can be specialized with the correspondences shown in Table 1.

In the case of supervised registration, h is fixed. Its gradient flow is a set of n coupled nonlinear parabolic equations. 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 [12]. Such a proof is provided in [4] for a similar mono-modal optical flow PDE. In the case of the mutual information and the correlation ratio, their form is much more complex. The continuous re-evaluation of joint density estimates in $L^{\mathbf{u}}$ leads to some sort

of functional equations. None of the previous tools is directly applicable. Nevertheless, we will see that their numerical implementation does not raise any major problems.

3 Regularization

We initially suggested a flexible regularization term (3) defined for a mapping $\varphi : \mathcal{M}_{n \times n}(\mathbb{R}) \mapsto \mathbb{R}_+$. In this framework, we can notice that a simple *Tikhonov* model with $\varphi : \mathbf{X} \mapsto \frac{1}{2} \text{Tr}(\mathbf{X}^T \mathbf{X})$ leads to a diffusion term, since $\text{div}(\varphi_{D\mathbf{u}}(D\mathbf{u})) = \text{div}(D\mathbf{u}) = \Delta \mathbf{u}$. It was one of the first regularizations proposed to solve the dense image matching problem [5]. 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 . In the following, we focus our attention on a simple physically-based model borrowed from linear elasticity. We also consider geometry-driven heuristics, relying on diffusion tensors.

3.1 Linear Elasticity

As in [7], 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 a 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 [9], the corresponding strain energy is given by the functional

$$\begin{aligned} \mathcal{E}[\mathbf{u}] &= \int_{\Omega} \frac{\lambda}{2} (\text{Tr}(\mathbf{E}))^2 + \mu \text{Tr}(\mathbf{E}^2) \, d\mathbf{x} \\ &\simeq \int_{\Omega} \frac{\lambda}{2} \left(\text{Tr} \left(\frac{1}{2} (D\mathbf{u}^T + D\mathbf{u}) \right) \right)^2 + \mu \text{Tr} \left(\left(\frac{1}{2} (D\mathbf{u}^T + D\mathbf{u}) \right)^2 \right) \, d\mathbf{x}, \end{aligned}$$

where λ and μ are the *Lamé* constants (known for a given elastic material). The key idea is to use this energy as a regularization functional \mathcal{R} with

$$\varphi : \mathbf{X} \mapsto \frac{\lambda}{8}(\text{Tr}(\mathbf{X}^T + \mathbf{X}))^2 + \frac{\mu}{2}\text{Tr}((\mathbf{X}^T + \mathbf{X})^2)$$

Thus, we easily get

$$\begin{aligned} \text{div}(\varphi_{D\mathbf{u}}(D\mathbf{u})) &= \text{div}(\lambda(\nabla \cdot \mathbf{u})\mathbf{I} + \mu(D\mathbf{u}^T + D\mathbf{u})) \\ &= \mu\Delta\mathbf{u} + (\lambda + \mu)\nabla(\nabla \cdot \mathbf{u}) \end{aligned}$$

One can notice that $\varphi_{D\mathbf{u}}(D\mathbf{u})$ corresponds to the so-called *stress tensor*, an important element of elasticity theory. Finally, we can combine this expression with our previous results to rewrite equation (14)

$$\begin{cases} \mathbf{u}_t &= h(f, g \circ (\mathbf{Id} + \mathbf{u})) \nabla g \circ (\mathbf{Id} + \mathbf{u}) \\ &\quad + \alpha (\mu\Delta\mathbf{u} + (\lambda + \mu)\nabla(\nabla \cdot \mathbf{u})) \\ \mathbf{u}(0, \cdot) &= \mathbf{u}_0 \end{cases}$$

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 should point out that this regularization principle has been selected for the trade-off it offers between flexibility and simplicity, not in the intent of building a bio-mechanical 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 $\varphi : \mathbf{X} \mapsto \frac{1}{2}\text{Tr}(\mathbf{X}^T \mathbf{T}_f \mathbf{X})$, where \mathbf{T}_f is a symmetric tensor characterized by some features of the target image. By computing

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

we obtain a modified diffusion term $\text{div}(\mathbf{T}_f D\mathbf{u})$. 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 [20] and found a theoretical justification in [26]. They consider an anisotropic smoothing of \mathbf{u} along objects contours, by choosing \mathbf{T}_f as a modified projection matrix. First used with a linearized optical flow equation [20], this tensor has been recently applied in a nonlinear/scale-space setting [3, 4]. 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 .

3.3 Remarks on Other Regularization Methods

Selecting a function φ is not the only way to perform a suitable regularization. Several other models can be found in previous works dedicated to mono-modal matching problems.

For instance, one can avoid the additive regularization term \mathcal{R} by applying, along with the gradient descent, a spatial smoothing of the gradient field (convolution by a Gaussian kernel or a few iterations of the heat equation). The resulting method, related to the so-called demon's algorithm [27], does not directly fit into a simple variational formulation. However, regarding a possible equivalence between this type of approach and the additive regularization, the interested reader can find some theoretical results in [21].

Another way to deal with the regularization problem was first introduced in [8], with a viscous fluid PDE registration model. It has been generalized to a more abstract and mathematical setting in [29] and is based on the following principle. The deformation \mathbf{v} is considered as a continuous composition of infinitesimal transformations. This decomposition is used to define a distance on a related infinite

dimensional group of diffeomorphisms. Then, this distance is used to measure the regularity of \mathbf{v} . The main advantage of this principle is to recover a much larger class of deformations, by relaxing in a suitable way the functional \mathcal{R} . Based on these ideas, a very generic matching framework is proposed in [19].

4 Numerical Experiments

4.1 Implementation issues

The spatial discretization of all the gradient flows is induced by image samples. Several numerical schemes are possible, and generally depend on the interpolation method. Our implementation combines a linear interpolation with a centered differences scheme. As usual for this type of problem, a multi-resolution approach is used. The gradient descent is applied to a set of smoothed and sub-sampled images. Since the considered functionals are non convex, this coarse to fine strategy helps avoiding 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 a 2D Gaussian kernel, implemented as a recursive filter.

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

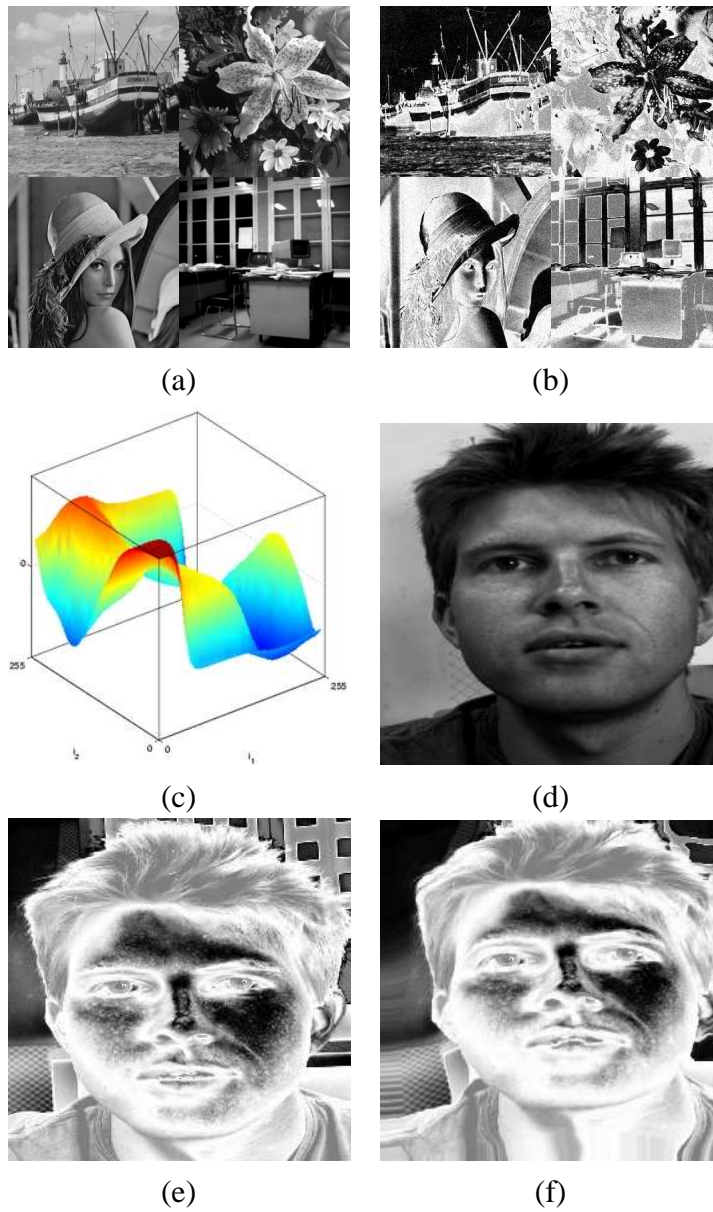


Figure 1: **SR** matching (with linear elasticity) : (a) and (b) learning sets : 4 pairs of 256x256 images, (c) estimated comparison function $h(i_1, i_2)$, (d) reference image (256x256), (e) image to register, (f) image (e) after realignment.

Figure 2. This example is interesting since the discrete joint histogram (i. e. before Gaussian smoothing) contains only five non-null entries. Registration was performed with both **MI** and **CR**, which yielded very similar results. Both algorithms succeeded in matching completely these two images with both linear elastic and geometry-driven regularizations (though the appropriate choice for the reference image was essential for **CR** to work). The results clearly emphasize the edge-preserving property of the latter regularization method, which recovers the “natural” discontinuities of the deformation field. A simple comparison of the two resulting fields underlines the importance of the prior knowledge embedded in the regularization functional. Even if such differences can be much less visible on real images, one not only needs 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. It also illustrates the fact that validating with synthetic deformation fields may not always be appropriate, since the solution may depend on the assumed properties of the solution.

Figure 3. 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). It gives a good qualitative indication of the quality of the registration. Most of the anatomical structures seem correctly realigned. In this case, the linear elastic regularization is used.

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). A first look at 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

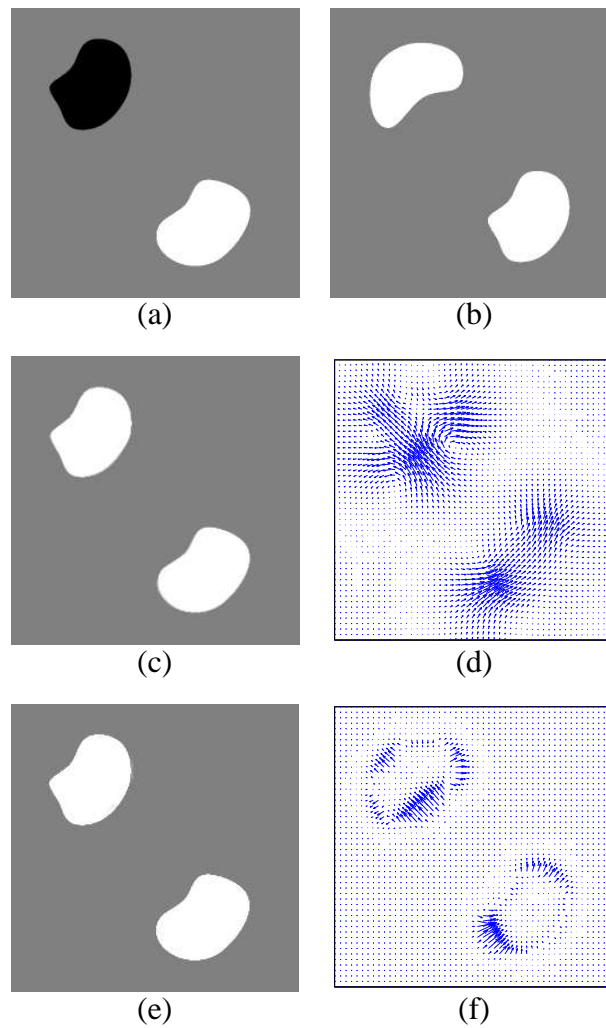


Figure 2: **MI** criterion with a synthetic example : (a) reference image (256x256), (b) image to register, (c) image (b) after realignment (with linear elasticity), (d) recovered displacement field (with linear elasticity), (e) image (b) after realignment (with geometry-driven regularization), (f) recovered displacement field (with geometry-driven regularization). The **CR** criterion yields similar results.

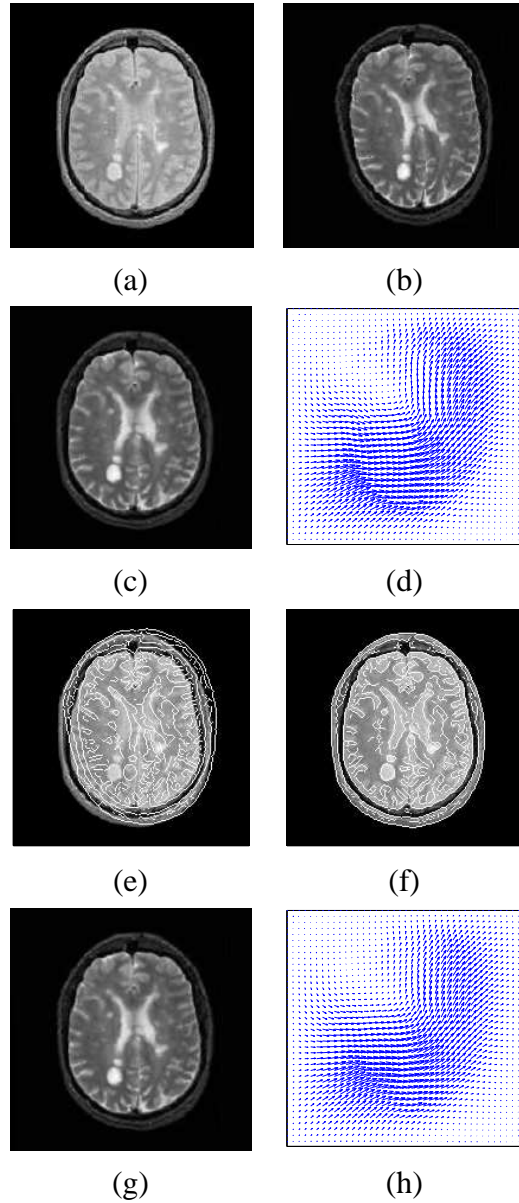


Figure 3: PD/T2 MRI realignment with **MI** and **CR** comparison functions (regularization using linear elasticity) : (a) reference image (PD 200x200), (b) image to register (T2 200x200), (c) image (b) after realignment (**MI**), (d) recovered displacement field (**MI**), (e) initial pose, (f) final pose, (g) image (b) after realignment (**CR**), (h) recovered displacement field (**CR**).

that the main anatomical structures seem correctly realigned. Let us now consider the elements of the initial displacement field whose norm was larger than one voxel. A quantitative error study shows that 16.8 % of these vectors have been recovered with a subvoxel accuracy. This percentage increases to 53.17 % for the displacement vectors whose norm was lower than 2 voxels. A direct interpretation of these objective measures is difficult and is not completely relevant to evaluate the quality of the registration. As we already noticed, different deformation fields can lead to meaningful realignments (see Figure 2). In this case, the linear elastic model offers a smoothness assumption which is slightly different from the artificial Gaussian-based deformation, explaining in part the differences between the applied and the recovered displacement fields.

Figure 5. In this last figure, we illustrate one of the alternative regularization methods mentioned in Section 3.3. We remove the term $\text{div}(\varphi_{D\mathbf{u}}(D\mathbf{u}))$ and apply a Gaussian smoothing to the gradient field at each iteration. This approach is somehow related to the isotropic regularization provided by $\varphi : \mathbf{X} \mapsto \frac{1}{2}\text{Tr}(\mathbf{X}^T \mathbf{X})$, and provides a robust regularization algorithm when no discontinuities are to be preserved. The example shows an artificially distorted image of a face. Both **CR** and **MI** gradient flows were tested.

5 Conclusions

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

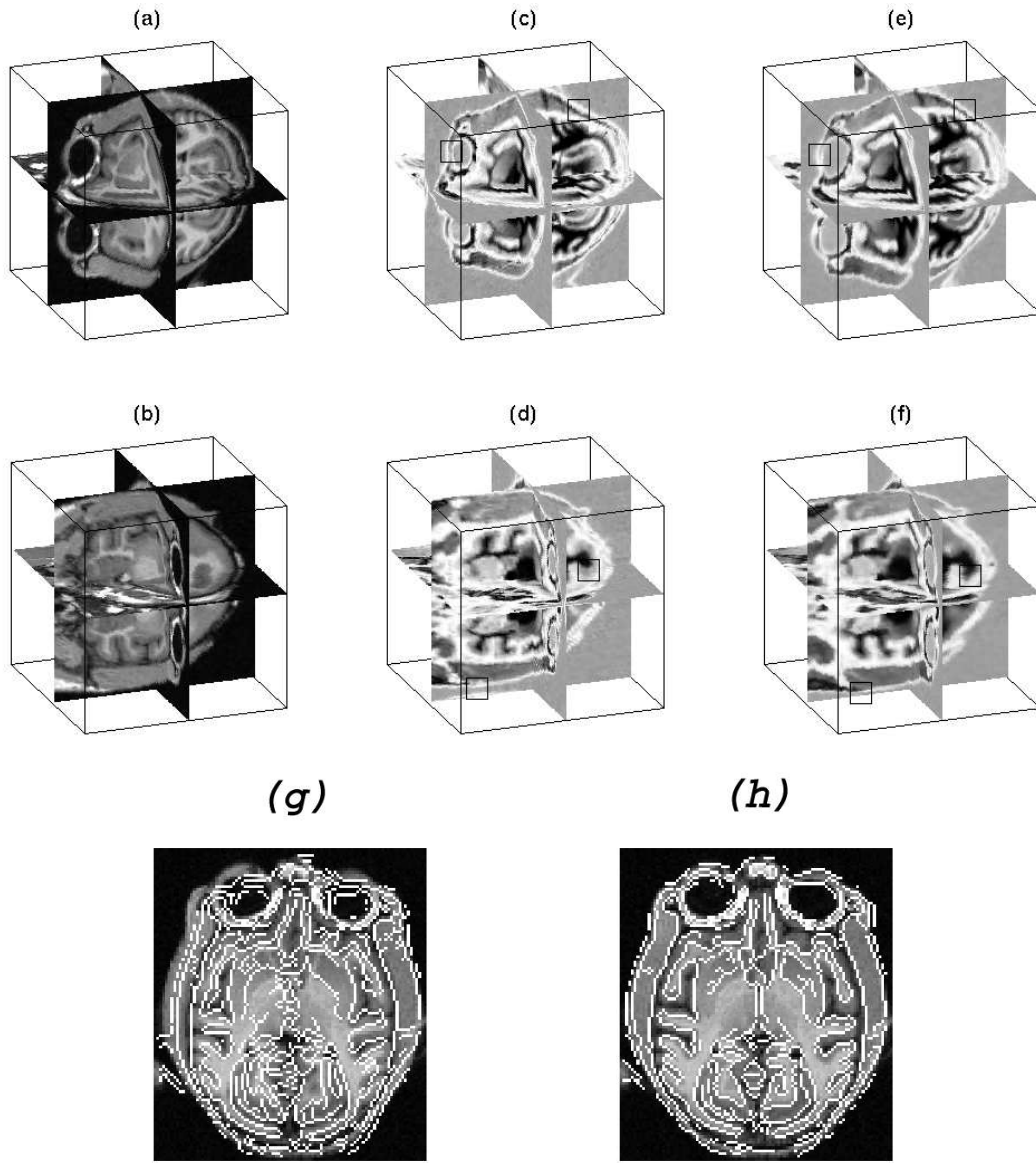


Figure 4: Realignment of MRI volumes (synthetic example with \mathbf{MI} and linear elasticity) : (a) and (b) two views of the reference volume (T2 66x111x97), (c) and (d) volume to register (66x111x97), (e) and (f) volume after realignment, (g) and (h) contours of the deformed volume superimposed over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

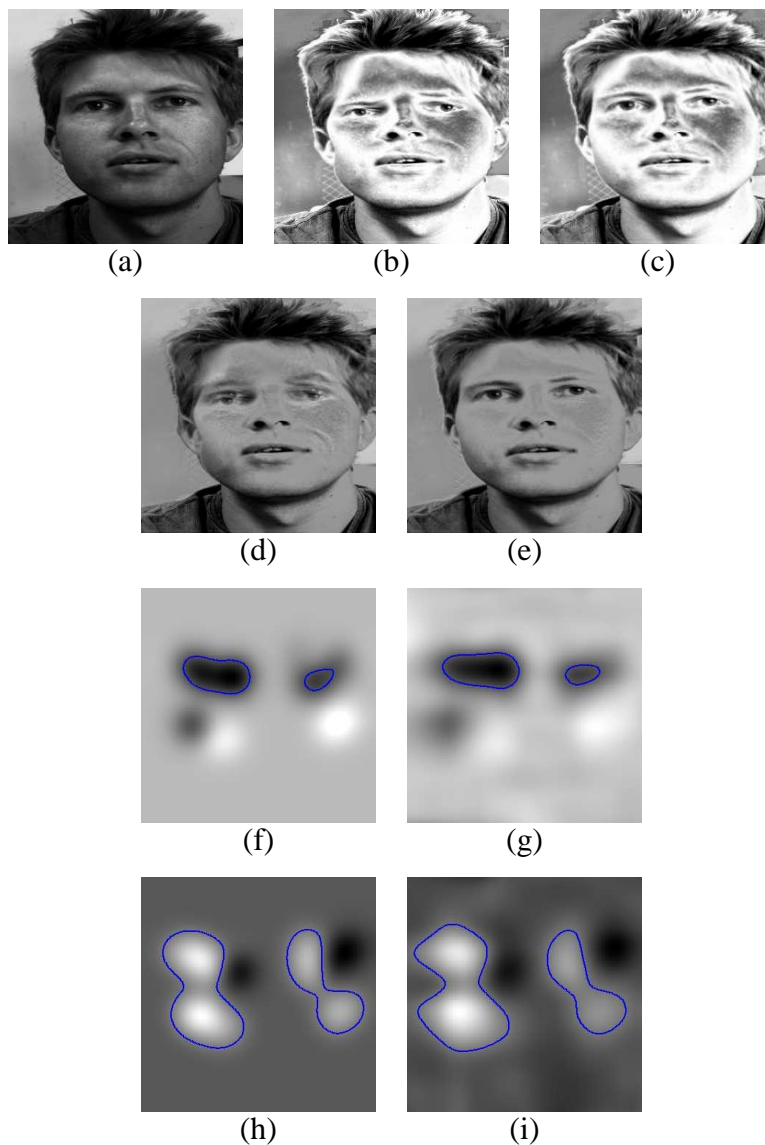


Figure 5: 2D-Matching with Gaussian-smoothing regularization : (a), (b) and (c) reference, target and registered images. (d) superposition of (a) and (b), (e) superposition of (a) and (c). (f) and (g) applied and recovered deformation fields, x-direction (level 6.00 shown). (h) and (i) applied and recovered deformation fields, y-direction (level -2.00 shown). Very similar results were obtained by maximization of **CR** and **MI**. The results from the first two rows come from the **MI** experiment, while those of the last two rows are from the **CR** test.

Acknowledgements

The authors would like to thank Javier S´anchez for fruitful discussions on geometry-driven regularization based on diffusion tensors. The MRI data used in figure 3 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. Esclari´n, M. Lefebure, and J. S´anchez. A pde model for computing the optical flow. In *Proceedings of CEDYA XVI*. Universidad de Las Palmas de Gran Canaria, 1999.
- [2] L. Alvarez, R.Deriche, J.Weickert, and X. Sanchez. Dense disparity map estimation respecting image discontinuities: A pde and scale-space based approach . Research Report 3874, INRIA, January 2000. Accepted for publication in Journal of Visual Communication and Image Representation, Special Issue on Partial Differential Equations in Image Processing, Computer Vision and Computer Graphics, Dec. 2000 (to appear).
- [3] L. Alvarez, J. Weickert, and J. Sanchez. A scale-space approach to nonlocal optical flow calculations. *Second International Conference on Scale Space, Lecture Notes in Computer Sciences*, 1682, 1999.
- [4] L. Alvarez, J. Weickert, and J. S´anchez. Reliable Estimation of Dense Optical Flow Fields with Large Displacements. Technical report, Cuadernos del Instituto Universitario de Ciencias y Tecnolog´ias Cibern´eticas, 2000. A revised version of this paper will be published in the International Journal of Computer Vision.
- [5] Y. Amit. A nonlinear variational problem for image matching. *SIAM J. Sci. Comput.*, 15(1), January 1994.

- [6] D. Bosq. *Nonparametric Statistics for Stochastic Processes*, volume 110 of *Lecture Notes in Statistics*. Springer-Verlag, 2nd edition, 1998.
- [7] Gary Christensen, MI Miller, and MW Vannier. A 3d deformable magnetic resonance textbook based on elasticity. In *Proceedings of the American Association for Artificial Intelligence, Symposium: Applications of Computer Vision in Medical Image Processing*, 1994.
- [8] G.E. Christensen, R.D. Rabbitt, and M.I. Miller. Deformable template using large deformation kinematics. *IEEE Transactions on Image Processing*, 5(10):1437–1447, 1996.
- [9] P.G. Ciarlet. *Mathematical Elasticity*, volume 1. North Holland, 1988.
- [10] R. Courant. *Calculus of Variations*. New York, 1946.
- [11] 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, December 1995.
- [12] L.C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, 1998.
- [13] T. Gaens, F. Maes D. Vandermeulen, and P. Suetens. Non-rigid multimodal image registration using mutual information. In J.van Leeuwen G. Goos, J. Hartmanis, editor, *First International Conference on Medical Image Computing and Computer-Assisted Intervention*, volume 1496 of *Lecture Notes in Computer Science*. Springer, 1998.
- [14] N. Hata, T. Dohi, S. Warfield, W. Wells III, R. Kikinis, and F. A. Jolesz. Multimodality deformable registration of pre-and intra-operative images for mri-guided brain surgery. In J.van Leeuwen G. Goos, J. Hartmanis, editor, *First International Conference on Medical Image Computing and Computer-Assisted Intervention*, volume 1496 of *Lecture Notes in Computer Science*. Springer, 1998.

- [15] S. Haykin. *Neural Networks*. Prentice Hall, 1999.
- [16] M.E. Leventon and W.E.L. Grimson. Multi-Modal Volume Registration Using Joint Intensity Distributions. In W.M. Wells, A. Colchester, and S. Delp, editors, *Medical Image Computing and Computer-Assisted Intervention-MICCAI'98*, number 1496 in Lecture Notes in Computer Science, Cambridge, MA, USA, October 1998. Springer.
- [17] J.B.A. Maintz, H.W. Meijering, and M.A. Viergever. General multimodal elastic registration based on mutual information. In *Medical Imaging 1998 - Image Processing*, volume 3338, pages 144–154. SPIE, 1998.
- [18] Chuck Meyer, Jennifer Boes, Boklye Kim, and Peyton Bland. Evaluation of control point selection in automatic, mutual information driven, 3d warping. In J. van Leeuwen G. Goos, J. Hartmanis, editor, *First International Conference on Medical Image Computing and Computer-Assisted Intervention, Proceedings*, volume 1496 of *Lecture Notes in Computer Science*, October 1998.
- [19] M.I. Miller and L. Younes. Group action, homeomorphism and matching : A general framework. To appear in special issue of IJCV, 2001.
- [20] H.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.
- [21] Esther Radmoser, Otmar Scherzer, and Joachim Weickert. Scale-space properties of regularization methods. In Mads Nielsen, P. Johansen, O.F. Olsen, and J. Weickert, editors, *Scale-Space Theories in Computer Vision*, volume 1682 of *Lecture Notes in Computer Science*, pages 211–222. Springer, September 1999.
- [22] 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.

- [23] A. Roche, G. Malandain, X. Pennec, and N. Ayache. The correlation ratio as new similarity metric for multimodal image registration. In W.M. Wells, A. Colchester, and S. Delp, editors, *Medical Image Computing and Computer-Assisted Intervention-MICCAI'98*, number 1496 in Lecture Notes in Computer Science, Cambridge, MA, USA, October 1998. Springer.
- [24] Alexis Roche, Grégoire Malandain, Xavier Pennec, and Nicholas Ayache. Multimodal image registration by maximization of the correlation ratio. Technical Report 3378, INRIA, August 1998.
- [25] D. Rückert, C. Hayes, C. Studholme, P. Summers, M. Leach, and D.J. Hawkes. Non-rigid registration of breast mr images using mutual information. In W.M. Wells, A. Colchester, and S. Delp, editors, *Medical Image Computing and Computer-Assisted Intervention-MICCAI'98*, number 1496 in Lecture Notes in Computer Science, Cambridge, MA, USA, October 1998. Springer.
- [26] M.A. Snyder. On the mathematical foundations of smoothness constraints for the determination of optical flow and for surface reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(11), November 1995.
- [27] J.P. Thirion. Image matching as a diffusion process: An analogy with maxwell's demons. *Medical Image Analysis*, 2(3):243–260, 1998.
- [28] A.N. Tikhonov and V.Y. Arsenin. *Solutions of Ill-posed Problems*. Winston and Sons, Washington, D.C., 1977.
- [29] Alain Trounev. Diffeomorphisms groups and pattern matching in image analysis. *International Journal of Computer Vision*, 28(3):213–21, 1998.
- [30] Paul Viola and William M. Wells III. Alignment by maximization of mutual information. *The International Journal of Computer Vision*, 24(2):137–154, 1997.

- [31] W.M. 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.



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

Unité de recherche INRIA Lorraine : Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - B.P. 101 - 54602 Villers lès Nancy Cedex (France)

Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)

Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot St Martin (France)

Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, B.P. 105 - 78153 Le Chesnay Cedex (France)

<http://www.inria.fr>

ISSN 0249-6399