Int J Comput Vis (2009) 85: 291–306
DOI 10.1007/s11263-009-0219-z

Registration of Anatomical Images Using Paths of Diffeomorphisms Parameterized with Stationary Vector Field Flows

Monica Hernandez · Matias N. Bossa · Salvador Olmos

Received: 25 January 2008 / Accepted: 27 January 2009 / Published online: 13 February 2009

Abstract Computational Anatomy aims for the study of variability in anatomical structures from images. Variability is encoded by the spatial transformations existing between anatomical images and a template selected as reference. In the absence of a more justified model for inter-subject variability, transformations are considered to belong to a convenient family of diffeomorphisms which provides a suitable mathematical setting for the analysis of anatomical variability. One of the proposed paradigms for diffeomorphic registration is the Large Deformation Diffeomorphic Metric Mapping (LDDMM). In this framework, transformations are characterized as end points of paths parameterized by time-varying flows of vector fields defined on the tangent space of a Riemannian manifold of diffeomorphisms and computed from the solution of the non-stationary transport equation associated to these flows. With this characterization, optimization in LDDMM is performed on the space of non-stationary vector field flows resulting into a time and memory consuming algorithm. Recently, an alternative characterization of paths of diffeomorphisms based on constant-time flows of vector fields has been proposed in the literature. With this parameterization, diffeomorphisms constitute solutions of stationary ODEs. In this article, the stationary parameterization is included for diffeomorphic registration in the LDDMM framework. We formulate the variational problem related to this registration scenario and derive the associated Euler-Lagrange equations. Moreover, the performance of the non-stationary vs the stationary parameterizations in real and simulated 3D-MRI brain datasets is evaluated. Compared to the non-stationary parameterization, our proposal provides similar results in terms of image matching and local differences between the diffeomorphic transformations while drastically reducing memory and time requirements.

Keywords Computational Anatomy · LDDMM · Efficient · Diffeomorphic registration · Infinite dimensional · Riemannian manifold · Stationary parameterization

M. Hernandez (✉)
Computer Science and System Department, Communication Technology Group (GTC), Aragon Institute of Engineering Research (I3A), University of Zaragoza, 50009, Zaragoza, Spain
e-mail: mhg@unizar.es

M.N. Bossa · S. Olmos
Communication Technology Group (GTC), Aragon Institute of Engineering Research (I3A), University of Zaragoza, 50009, Zaragoza, Spain

1 Introduction

Computational Anatomy aims for the study of variability in anatomical structures from images. Variability is encoded by the spatial transformations existing between the anatomical images and a template selected as reference (Grenander 1994). In the absence of a more justified physical model for inter-subject anatomical variability and under the reasonable assumption that the deformation model responsible of organ growth is related to smooth and invertible maps, transformations are usually assumed to belong to a group of diffeomorphisms (i.e. differentiable maps with differentiable inverse) endowed with the structure of Riemannian manifold. Statistics on these spaces of transformations allow modeling the anatomical variability of a population. Different models of anatomical variability have been successfully used in order to identify anatomical differences between healthy and diseased individuals or im-

prove the diagnosis of pathologies (Thompson et al. 2001; Wang et al. 2003, 2007; Csernansky et al. 2004; Qiu et al. 2007). Moreover, models of growth have been built for the assessment of the anatomical change over time (Thompson et al. 2000; Beg 2003; Miller 2004; Gerig et al. 2006; Davis et al. 2007).

Diffeomorphic registration in the large deformation setting is formulated as the problem of finding the path in the Riemannian manifold of diffeomorphisms with minimal energy that smoothly transforms the source into the target image. This problem is approached within a variational formulation from the minimization of an energy functional consisting of an image matching term that measures the similarity between the images after registration and a regularization term to favor stable numerical solutions. In the so called Large Deformation Diffeomorphic Metric Mapping paradigm (LDDMM) (Dupuis et al. 1998; Beg et al. 2005), image matching is selected as the sum of squared intensity differences and regularization is imposed on the energy associated to the length of the path.

In the last years, some modifications in the definition of the elements of the variational problem were proposed providing different algorithms for diffeomorphic registration within the LDDMM paradigm. Since Christensen et al. (1999) different authors proposed inverse consistent versions of the variational problem incorporating a source to target symmetry during registration (Joshi et al. 2004; Avants and Gee 2004; Beg and Khan 2007). Avants et al. combined landmark with image based diffeomorphic registration and introduced cross-correlation as image matching in this inverse consistent framework (Avants et al. 2006, 2008). Lorenzen et al. used information theory measures in the image matching term for multimodal diffeomorphic registration (Lorenzen et al. 2006). Garcin et al. defined a Riemannian manifold structure in the space of anatomical images in order to find geodesic paths on the manifold of diffeomorphisms as well as on the image manifold (Garcin and Younes 2006).

In all these methods, transformations are characterized as end points of paths of diffeomorphisms parameterized by time-varying flows of smooth vector fields defined on the tangent space of the Riemannian manifold of diffeomorphisms. With this representation, diffeomorphisms constitute solutions of their associated non-stationary transport equations. Optimization is performed on the space of non-stationary vector field flows. In consequence, the computational requirements of these methods linearly grow with the size of time sampling, resulting into time and memory demanding algorithms.

In order to alleviate the computational requirements of LDDMM, Younes et al. proposed to restrict optimization to time-varying flows of vector fields that fulfill the momentum conservation equation (Younes 2007). Thus, the variational problem associated to LDDMM is restricted to the space of initial momenta allowing great time and memory savings. However, the dependence of the diffeomorphisms on the initial momenta results rather complex and the more expensive although simpler implementation of original LDDMM is usually preferred in Computational Anatomy applications.

Recently, an alternative way of parameterizing paths of diffeomorphisms was proposed in Arsigny et al. (2006a). The parameterization is obtained from constant-time flows of smooth vector fields. With this representation, diffeomorphisms constitute solutions of stationary ODEs. This parameterization is closely related to the algebraic structure of the diffeomorphism group as the paths that can be parameterized using these stationary vector field flows are exactly identified with the one-parameter subgroups.

In this article, the stationary parameterization is included for diffeomorphic registration in the variational problem associated to the LDDMM framework. This restricts transformations to belong to paths identified with one-parameter subgroups of diffeomorphisms allowing great time and memory savings. We formulate the variational problem related to the registration scenario and derive the Euler-Lagrange equations associated to the minimization of the energy functional. Moreover, we evaluate the performance of the non- stationary vs the stationary parameterizations in real and simulated 3D-MRI brain datasets.

The rest of the article is divided as follows. In Sect. 2 we revisit the differential structure of the Riemannian manifold of diffeomorphisms used in the LDDMM paradigm. In Sect. 3 we present our method for diffeomorphic registration. Results in real and simulated datasets are presented in Sect. 4. Finally, Sect. 5 presents discussion and some concluding remarks.

2 Riemannian Manifolds of Diffeomorphisms

In this section we provide an overview of the fundamental aspects of Riemannian geometry regarding to infinite dimensional manifolds of diffeomorphisms used in Computational Anatomy. The study of infinite dimensional manifolds is more complicated than the case of finite dimensions. A finite dimensional differentiable manifold is an abstract mathematical space that is locally homeomorphic, and therefore topologically equivalent, to an Euclidean space, $\mathbb{R}^n$. In consequence, the local properties of the manifold can be studied in terms of linear spaces translating the methods of calculus in linear spaces to differentiable manifolds. On the other hand, infinite dimensional differentiable manifolds are locally homeomorphic to infinite dimensional metric vector spaces (namely, Frechet, Banach or Hilbert spaces), showing a complex differentiable structure.

The group of diffeomorphisms is defined from the set of homomorphisms (continuous bijective mappings with con-

tinuous inverse)

$$Diff(\Omega) := \{\varphi : \Omega \to \Omega, \varphi \text{ and } \varphi^{-1} \text{ smooth mappings}\} \quad (1)$$

together with the group operations.

Composition:

$$\mu : Diff(\Omega) \times Diff(\Omega) \to Diff(\Omega),$$
$$\mu(\varphi_1, \varphi_2) = \varphi_2 \circ \varphi_1 \quad (2)$$

Inverse:

$$\nu : Diff(\Omega) \to Diff(\Omega), \nu(\varphi) = \varphi^{-1} \quad (3)$$

where the composition operation μ provides two different group homomorphisms.

Left composition:

$$L_\varphi : Diff(\Omega) \to Diff(\Omega), L_\varphi(\rho) = \rho \circ \varphi \quad (4)$$

Right composition:

$$R_\varphi : Diff(\Omega) \to Diff(\Omega), R_\varphi(\rho) = \varphi \circ \rho \quad (5)$$

The set Ω is a compact simply connected differentiable manifold. Dealing with the non-compact case is also possible, although it results much more complicated and is out of scope of application domain of Computational Anatomy.

Manifolds of diffeomorphisms are defined from the group of diffeomorphisms $Diff(\Omega)$ by endowing with a Riemannian metric to the space of smooth vector fields in Ω. The structure of differentiable manifold defined on $Diff(\Omega)$ provides a local homeomorphism from every element φ in the manifold to the tangent space $T_\varphi(Diff(\Omega))$ at that element, which is isomorphic to the algebra of right-invariant vector fields in $Diff(\Omega)$ where the right-composition R_φ provides the canonical isomorphism between all the tangent spaces and this algebra. As in the finite dimensional case, the tangent space represents the nearest approximation of the manifold by a vector space on a neighborhood of φ. The rest of elements related to the differentiable structure (charts, tangent bundle, differentiable curves, differential maps…) can be defined analogously to the finite dimensional case (DoCarmo 1992).

Depending on the degree of differentiability of the set of diffeomorphisms (i.e. the meaning of the term "smooth" in (1)), different degrees of smoothness can be identified in the corresponding vector fields and therefore, different structures of differentiable manifold can be defined in $Diff(\Omega)$ (Schmid 2004). For example, the set of C^∞-diffeomorphisms $Diff^\infty(\Omega)$ together with the space of C^∞-vector fields in Ω is a Frechet space. As well, the set of C^k-diffeomorphisms $Diff^k(\Omega)$ with the space of C^k-vector fields is a Banach space ($k < \infty$), and the set of Sobolev H^s-diffeomorphisms $Diff^s(\Omega)$ with the space of H^s-vector fields is a Hilbert space ($s > \frac{1}{2}\dim(\Omega)$, $\dim(\Omega) \gneqq 1$) (Schmid 2004).

Dealing with C^∞-diffeomorphisms is problematic as these spaces lack of easy generalizations of inverse and implicit function theorems. Fortunately, suitable extensions for both theorems are available for both C^k and H^s diffeomorphisms. This makes both spaces appropriate candidates for differential calculus on diffeomorphisms. However, the existence of a complete scalar product makes Hilbert spaces preferable. Thus, the properties of finite dimensional vector spaces can be naturally extended to these infinite dimensional spaces. Furthermore, the theorems of existence and uniqueness of PDE solutions hold for Hilbert spaces.

$Diff^s(\Omega)$ has been widely studied in a physical context (Arnold 1989; Holm et al. 2004; Ebin and Marsden 1970) as the computation of the motion of a system in continuum mechanics can be described by a path of diffeomorphisms

$$\phi : [0, 1] \to Diff^s(\Omega), \quad t \to \phi(t) \quad (6)$$

deforming the ambient space Ω. The analogies existing between this physical problem and diffeomorphic registration allow to translate the setting for working with $Diff^s(\Omega)$ from continuum mechanics to Computational Anatomy.

Thus, in both disciplines, paths of diffeomorphisms are usually parameterized as the solution of the transport equation

$$\frac{d\phi(t)}{dt} = v(t, \phi(t)) \quad (7)$$

with initial condition $\phi(0) = id$ (group identity element), where

$$v : [0, 1] \to T(Diff^s(\Omega)),$$
$$t \to v(t, \phi(t)) \in T_{\phi(t)}(Diff^s(\Omega)) \quad (8)$$

is a time-varying flow of smooth vector fields in the tangent bundle constituted by the directional derivatives associated to the path at each point. Diffeomorphisms can be therefore computed as solutions of non- stationary ODEs. The Hilbert differentiable structure guarantees that the solution to the transport equation in $Diff^s(\Omega)$ is a path of diffeomorphisms. This solution does not exist in $Diff^\infty(\Omega)$ and $Diff^k(\Omega)$ as they are not Sobolev spaces.

The Riemannian metric in $Diff^s(\Omega)$ is constructed from the scalar product defined at the identity element, id. Thus, the scalar product of $v, w \in T_{id}(Diff^s(\Omega))$ is defined from

$$\langle v, w \rangle_{T_{id}(Diff^s(\Omega))} = \langle \mathcal{L}v, \mathcal{L}w \rangle_{L^2} \quad (9)$$

where $\mathcal{L}$ is a linear invertible differentiable operator, from which, the Riemannian metric is extended to the whole

tangent bundle by right-translation. Thus, given $v, w \in T_\varphi(Diff^s(\Omega))$

$$\langle v, w\rangle_{T_\varphi(Diff^s(\Omega))} = \langle (dR_{\varphi^{-1}})_\varphi v, (dR_{\varphi^{-1}})_\varphi w\rangle_{T_{id}(Diff^s(\Omega))} \quad (10)$$

where $(dR_{\varphi^{-1}})_\varphi$ denotes the differential map of $R_{\varphi^{-1}}$ at $T_\varphi(Diff^s(\Omega))$. From this construction, it follows that this metric is invariant under right-composition (right-invariant). However, invariance under left composition is not preserved.

The distance $d : Diff^s(\Omega) \times Diff^s(\Omega) \to \mathbb{R}^+$ associated to this Riemannian metric is given by

$$d(\varphi_1, \varphi_2) = \min\{L(\phi) \mid \phi \text{ smooth path between } id \text{ and } \varphi_2 \circ \varphi_1^{-1}\} \quad (11)$$

where $L(\phi)$ corresponds to the length of the path ϕ. With the non-stationary parameterization

$$\begin{aligned} L(\phi) &= \int_0^1 \left\| \frac{d\phi(t)}{dt} \right\|_{V(t)} dt \\ &= \int_0^1 \|(dR_{\phi(t)^{-1}})_{\phi(t)} v(t, \phi(t))\|_V dt \\ &= \int_0^1 \|v(t)\|_V dt \end{aligned} \quad (12)$$

where $V(t)$ and V denote the tangent spaces $T_{\phi(t)}(Diff^s(\Omega))$ and $T_{id}(Diff^s(\Omega))$, respectively and $\|\cdot\|_{V(t)}$ is the norm associated to the scalar product $\langle\cdot,\cdot\rangle_{V(t)}$. The kinetic energy of the path in the Riemannian manifold is defined as

$$E_k(\phi) = \int_0^1 \|v(t)\|_V^2 dt \quad (13)$$

This energy intuitively represents the cost that supposes walking on the manifold following that path. Paths minimizing this cost, and therefore minimizing their length, are called geodesic curves (DoCarmo 1992). In the case of geodesic paths starting at the identity element, this energy provides a right-invariant measure of the amount of deformation associated to the diffeomorphism $\varphi = \phi(1)$.

Geodesic paths in $Diff^s(\Omega)$ are characterized from solutions of the Euler-Poincare equation for diffeomorphisms (EPDiff equation) (Cotter and Holm 2006)

$$\begin{aligned} &\frac{d(\mathcal{L}^\dagger \mathcal{L} v(t))}{dt} + v(t) \cdot \nabla(\mathcal{L}^\dagger \mathcal{L} v(t)) \\ &\quad + \nabla^T v(t) \cdot (\mathcal{L}^\dagger \mathcal{L} v(t)) + (\mathcal{L}^\dagger \mathcal{L} v(t)) \cdot \mathrm{div}(v(t)) = 0 \end{aligned} \quad (14)$$

The EPDiff equation is equivalent to the conservation of the momentum $\mathcal{L}^\dagger \mathcal{L} v(t)$ through geodesic paths (Michor and Mumford 2006)

$$\begin{aligned} \mathcal{L}^\dagger \mathcal{L} v(t) = (D\phi^{-1}(t))^T \cdot (\mathcal{L}^\dagger \mathcal{L} v(0)) \circ \phi^{-1}(t) \\ \times |\det(D\phi^{-1}(t))| \end{aligned} \quad (15)$$

This result was also reached by Miller et al. from the extension of the Lagrangian momentum conservation in classical mechanics to diffeomorphisms (Miller et al. 2006). However, the existence of such conservation law goes back to Noether's theorem in 1918 (Noether 1918). From the momentum conservation, the right geodesic characterized by a given $v_0 \in V$ is defined as the solution of the transport equation $d\phi(t)/dt = v(t, \phi(t))$ where $v(t)$ is the unique non-stationary vector field flow that fulfills the momentum conservation equation for $v(0) = v_0$ (Holm et al. 2004; Beg and Khan 2006). By definition, the Riemannian exponential map at the identity $\exp : V \to Diff^s(\Omega)$ is identified with the point of path $\phi(t)$ at time $t = 1$.

Recently, an alternative way of parameterization of paths in $Diff^s(\Omega)$ has been proposed in the literature (Arsigny et al. 2006a). In this framework, paths are parameterized as the solution of the transport equation associated to stationary vector field flows

$$\frac{d\phi(t)}{dt} = w(\phi(t)) \quad (16)$$

with initial condition $\phi(0) = id$. With this parameterization, diffeomorphisms can be computed as solutions of stationary (also called autonomous) ODEs. The kinetic energy of a path parameterized by a stationary vector field w is defined as

$$\begin{aligned} E_k(\phi) &= \int_0^1 \|w(\phi(t))\|_{V(t)}^2 dt \\ &= \int_0^1 \|(dR_{\phi(t)^{-1}})_{\phi(t)} w(\phi(t))\|_V^2 dt = \|w\|_V^2 \end{aligned} \quad (17)$$

This parameterization is closely related to the algebraic structure defined in $Diff^s(\Omega)$ as the paths that can be parameterized using stationary vector fields are exactly the one-parameter subgroups. In this case, the vector field w is an infinitesimal generator of the subgroup and the solution of the transport equation is identified with the group exponential map $\mathrm{Exp} : V \to Diff^s(\Omega)$.

As the metric defined in $Diff^s(\Omega)$ is not bi-invariant, one-parameter subgroups are not identified with the geodesics starting at the identity element. Therefore, the group and Riemannian exponentials, exp and Exp, result into two different maps and the stationary and non- stationary parameterizations of diffeomorphisms provide elements belonging to rather different families of diffeomorphisms.

3 Diffeomorphic Registration of Anatomical Images

In the LDDMM framework, diffeomorphic registration from a template image I_0 to a target I_1 is represented by the end point $\varphi = \phi(1)$ of a path of diffeomorphisms $\phi(t)$ resulting

from the minimization of the energy functional

$$E(\varphi) = E_{\text{reg}}(\phi) + \frac{1}{\sigma^2} E_{\text{img}}(I_0 \circ \varphi^{-1}, I_1) \tag{18}$$

The term $E_{\text{reg}}(\phi)$ imposes a regularization on path energy. The term $E_{\text{img}}(I_0 \circ \varphi^{-1}, I_1)$ measures the similarity between the images after registration. Although it is usually selected to be the sum of squared intensity differences, $\|I_0 \circ \varphi^{-1} - I_1\|_{L^2}^2$, the image term can be replaced by any other energy proposed in usual registration techniques. The weighting factor $1/\sigma^2$ balances the energy contribution between regularization and matching.

In diffeomorphic registration, the space of transformations depends on the parameterization chosen to compute paths of diffeomorphisms and additional imposed constraints frequently related to geodesic properties. In this section we revisit the LDDMM method for diffeomorphic registration using the non-stationary parameterization, and formulate the method using the stationary parameterization.

3.1 Registration Using the Non-Stationary Parameterization

In the LDDMM framework, paths of diffeomorphisms are parameterized using time-varying vector field flows. Diffeomorphic registration is obtained computing the path of vector fields $v(t) \subseteq V$ resulting from the minimization of the energy functional

$$E_{I_0 \to I_1}(v) = \int_0^1 \|v(t)\|_V^2 dt + \frac{1}{\sigma^2} \|I_0 \circ \varphi^{-1} - I_1\|_{L^2}^2 \tag{19}$$

and obtaining the solution at time 1 of the non- stationary transport equation associated to $v(t)$. The regularization term is identified with the kinetic energy associated to non-stationary paths of diffeomorphisms in $Diff^s(\Omega)$ (13). In order to assure a source to target symmetry in the registration, the inverse- consistent version of this variational problem has been preferably used in the literature instead (Joshi et al. 2004)

$$E_{I_0 \leftrightarrow I_1}(v) = \int_0^1 \|v(t)\|_V^2 dt + \frac{1}{2\sigma^2} (\|I_0 \circ \phi(1)^{-1} - I_1\|_{L^2}^2 + \|I_1 \circ \phi(1) - I_0\|_{L^2}^2) \tag{20}$$

The solution to the minimization of this variational problem is approached using gradient descent techniques (Beg et al. 2005). The gradient of the energy functional is computed from the Euler-Lagrange equation associated to the minimization of the energy functional given in (21).

$$\begin{aligned}\nabla_v E(v)(t) = {} & 2v(t) - (\mathcal{L}^\dagger \mathcal{L})^{-1} \left(\frac{1}{\sigma^2} (I_0 \circ \phi(t)^{-1} - I_1 \circ \phi(1) \circ \phi(t)^{-1}) \cdot |\det(D\phi(1) \circ \phi(t)^{-1})| \cdot \nabla I_0(\phi(t)^{-1}) \right) \\ & + (\mathcal{L}^\dagger \mathcal{L})^{-1} \left(\frac{1}{\sigma^2} (I_0 \circ \phi(t)^{-1} - I_1 \circ \phi(1) \circ \phi(t)^{-1}) \cdot |\det(D\phi(t)^{-1})| \cdot \nabla I_1(\phi(1) \circ \phi(t)^{-1}) \right),\end{aligned} \tag{21}$$

$$\begin{aligned}\nabla_w E(w) = {} & 2w - (\mathcal{L}^\dagger \mathcal{L})^{-1} \left(\frac{1}{\sigma^2} (I_0 \circ \text{Exp}(w)^{-1} - I_1) \cdot \nabla (I_0 \circ \text{Exp}(w)^{-1}) \right) \\ & + (\mathcal{L}^\dagger \mathcal{L})^{-1} \left(\frac{1}{\sigma^2} (I_1 \circ \text{Exp}(w) - I_0) \cdot \nabla (I_1 \circ \text{Exp}(w)) \right)\end{aligned} \tag{22}$$

The solution to the variational problem provides a geodesic path of diffeomorphisms with minimal energy and maximum image matching at $t = 1$.

3.2 Registration Using the Stationary Parameterization

In contrast to the usual approach, our framework for diffeomorphic registration parameterizes diffeomorphisms using stationary vector field flows. Diffeomorphisms are computed as solutions of stationary ODEs associated to an element belonging to a one-parameter subgroup spanned by some infinitesimal generator $w \in V$. Thus, the diffeomorphism that connects I_0 and I_1 is represented by the group exponential map $\text{Exp}(w)$, where w is computed from the minimization of the energy functional

$$E_{I_0 \to I_1}(w) = \|w\|_V^2 + \frac{1}{\sigma^2} \|I_0 \circ \text{Exp}(w)^{-1} - I_1\|_{L^2}^2 \tag{23}$$

that in inverse consistent version is written as

$$E_{I_0 \leftrightarrow I_1}(w) = \|w\|_V^2 + \frac{1}{2\sigma^2} \big(\|I_0 \circ \text{Exp}(w)^{-1} - I_1\|_{L^2}^2 + \|I_1 \circ \text{Exp}(w) - I_0\|_{L^2}^2 \big) \tag{24}$$

The Euler-Lagrange equation associated to the minimization of the energy functional is given in (22). Its computation is detailed in the Appendix. The solution to the variational problem provides the infinitesimal generator of a one-parameter subgroup of diffeomorphisms with minimal deformation and maximum image matching at $t = 1$.

3.3 Numerical Implementation

Two different gradient descent optimization strategies have been proposed in the literature of registration in the large deformation setting (Christensen et al. 1996; Beg et al. 2005). Christensen et al. propose to solve the Euler-Lagrange equation for a given time step t_k until convergence. Then, a new template image is generated by composition of the current transformation $\phi(t_k)$ to the former template (regridding), and optimization is restarted for a new time step t_{k+1}. The resulting path of diffeomorphisms is computed from the successive composition of the transformations associated to the sequence of propagated templates. Optimization strategy in Beg et al. method, solves the Euler-Lagrange equation associated to the whole vector field flow. In contrast to Christensen's method, temporal continuity is preserved during optimization providing more stable results. Hence, in this article, we use Beg et al. optimization strategy.

Thus, the numerical implementation for finding the minimum of the energy functionals $E_{I_0 \leftrightarrow I_1}(v)$ and $E_{I_0 \leftrightarrow I_1}(w)$ proceeds within a multi-resolution strategy in a gradient descent fashion. In order to make both algorithms comparable, the same implementation criteria have been adopted in their common stages. At the coarsest resolution level, the algorithm based on the non-stationary parameterization initializes with iteration $k = 0$, $v(t) = 0_V(t)$ and $\phi(t) = id$, $\forall t$ whereas the algorithm based on the stationary parameterization initializes with $w = 0_V$, and $\varphi = id$. In the finer resolution levels, both algorithms initialize $v(t)$ and w by interpolating the vector fields resulting from the convergence in the previous resolution level using a tri-linear interpolator. Every iteration in the gradient descent consists of the steps collected in Table 1.

In both algorithms, standard line search strategies are used to estimate the step size ϵ (Nocedal and Wright 1999). The computation of the diffeomorphisms is performed by solving the corresponding transport equations ((7) and (16)) using a semi-Lagrangian numerical scheme (Staniforth and Cote 1991). The operator associated to $\langle ., . \rangle_V$ is selected to be $\mathcal{L} = \gamma Id - \alpha \nabla^2$, and $\mathcal{L}^\dagger \mathcal{L}$ is computed in the Fourier domain as in Beg et al. (2005). The convergence in each resolution level is reached if the value of ϵ in the search strategy is too small or the absolute rate of change in the energy is less than a tolerance value.

4 Results

In this section we focus on evaluating and comparing the influence of the parameterization in the performance of the registration methods associated to the variational problems $E_{I_0 \leftrightarrow I_1}(v)$ and $E_{I_0 \leftrightarrow I_1}(w)$. We provide results on real and simulated MRI brain datasets.

4.1 Evaluation Metrics

Registration performance is evaluated in terms of the final image matching and the differences between the diffeomorphic transformations. The image matching is quantified from the relative sum of squared intensity differences,

$$\mathrm{RSSD} = \frac{1}{2} \frac{\|I_0 \circ \varphi^{-1} - I_1\|_{L^2}^2 + \|I_1 \circ \varphi - I_0\|_{L^2}^2}{\|I_0 - I_1\|_{L^2}^2} \tag{25}$$

The differences between transformations are evaluated from the sum of squared differences (SSD)

$$\mathrm{SSD}(\varphi_1, \varphi_2) = \|\varphi_1 - \varphi_2\|_{L^2} \tag{26}$$

and a distance between the associated Jacobian matrices ($J = D\varphi$). In this work, we use a distance defined on the group of symmetric positive definite matrices $Sym^+(3)$ applied to the strain matrix $S = (J^T \cdot J)^{1/2}$ (Arsigny et al. 2006b)

Table 1 Algorithm for non-stationary and stationary diffeomorphic registration

(1) Compute the energy gradient from the Euler-Lagrange Equation: $\nabla_v E(v_k)$ and $\nabla_w E(w_k)$, respectively

(2) Gradient descent update: $v_k(t) = v_{k-1}(t) - \epsilon \nabla_v E(v_{k-1})(t)$ and $w_k = w_{k-1} - \epsilon \nabla_w E(w_{k-1})$, respectively

(3) Compute the inverse path of diffeomorphisms: $\phi_k^{-1}(t)$ and $\varphi_k = \mathrm{Exp}(-w_k)$, respectively

(4) Compute the direct path of diffeomorphisms: $\phi_k(t)$ and $\varphi_k = \mathrm{Exp}(w_k)$, respectively

(5) Compute the transformed images: $I_0 \circ \phi_k^{-1}(1)$, $I_1 \circ \phi_k(1)$ and $I_0 \circ \varphi_k^{-1}$, $I_1 \circ \varphi_k$, respectively

(6) Check for convergence criterion

(7) Prepare the next resolution level

$$d_{AI}(S_1, S_2) = \sqrt{\text{trace}(\log(S_1^{-1/2} S_2 S_1^{-1/2})^2)} \tag{27}$$

This is motivated by the fact that Riemannian metrics defined on $Sym^+(3)$ are being increasingly used in morphometric studies showing a great local discriminative power between transformations (Lepore et al. 2006).

In contrast to those metrics based on the logarithm of the Jacobian determinant ((Leow et al. 2006) and references therein),

$$d_{\log}(J_1, J_2) = \sqrt{(\log(\det(J_1)) - \log(\det(J_2)))^2} \tag{28}$$

Riemannian metrics use the whole strain matrix information. In the case of the distance d_{AI}, there exist many different Jacobian matrices with the same determinant ($d_{\log} = 0$) associated to different strain forces where $d_{AI} > 0$. Moreover, in the simplest case of Jacobian matrices expressing local expansion or contraction without torsion, straightforward computations show that

$$d_{\log} = |\log(RVC)| \tag{29}$$

and

$$d_{AI}(S_1, S_2) = \sqrt{3}|\log(RVC)| \tag{30}$$

where RVC denotes the relative volume change existing between J_1 and J_2, and therefore both metrics provide the same information. Furthermore, d_{AI} is a natural metric defined in the usual group structure of $Sym^+(3)$ and it is invariant under affine transformations.

4.2 Real Datasets

The population of anatomical brain images was composed by a total of 18 T1-MRI from the Hospital Clinic Barcelona, Spain. These images were acquired using a General Electric Signa Horizon CV 1.5 Tesla scan. As preprocessing steps, the images were first resampled yielding volumes of size $256 \times 256 \times 220$ with a spatial resolution of $0.9 \times 0.9 \times 0.9$ mm. Next, the skull was removed from the images (Dodgas et al. 2005). Finally, the image intensity was normalized using a histogram matching algorithm and all the images were aligned to a common coordinate system using a similarity transformation (7 dof) with the algorithms available in the Insight Toolkit (ITK).

4.3 Registration in Real Datasets

4.3.1 Quantitative Evaluation

In this experiment, we evaluate the performance of the diffeomorphic registration algorithm in real datasets. The experiment consists in the registration of one of the images randomly selected in our datasets to the rest of images. In this case, the image matching metric comprises errors in the registration due to photometric variations between the images and inaccurate matching consequence of diffeomorphic regularization constraints.

The definition of the operator $\mathcal{L}$ influences the registration results. This operator is usually related to the physical deformation model imposed on Ω. It remains an open question how to choose the best model in non-rigid registration algorithms (Modersitzki 2004). In this work we have selected $\mathcal{L} = \gamma Id - \alpha\nabla^2$ with $\gamma = 1$ and $\alpha = 0.02$ as extensively used in LDDMM literature since (Beg et al. 2005). This selection has provided smooth enough transformations with acceptable matching for our application. In future work we will study the performance of the registration algorithms with different definitions for $\mathcal{L}$.

From the rest of tunable parameters involved in the registration algorithms, the selection of $1/\sigma^2$ is the most critical. As shown in Table 2, the selection of this parameter strongly influences the final image matching (RSSD) and the extrema of the Jacobian determinant $J_{\max} = \max(\det(D\varphi^{-1}))$ and $J_{\min} = \min(\det(D\varphi^{-1}))$. From those registration results with the same relative differences, the ones with the highest minimum determinant and the lowest maximum determinant (i.e. the smoothest φ^{-1} and φ, respectively) are more

Table 2 *Registration in real datasets*. Average and standard deviation of the relative L^2 differences, RSSD, and the extrema of the Jacobian determinant, $J_{\max}$, $J_{\min}$. Up and down tables show the results obtained with $E_{I_0 \leftrightarrow I_1}(w)$ and $E_{I_0 \leftrightarrow I_1}(v)$, respectively

$1/\sigma^2$	RSSD	$J_{\min}$	$J_{\max}$
$E_{I_0 \leftrightarrow I_1}(w)$			
1.0e0	0.66 ± 0.04	0.82 ± 0.02	1.21 ± 0.03
5.0e1	0.36 ± 0.04	0.36 ± 0.14	2.25 ± 0.57
1.0e2	0.32 ± 0.03	0.29 ± 0.14	3.10 ± 1.31
5.0e2	0.30 ± 0.03	0.26 ± 0.14	3.79 ± 2.31
1.0e3	0.27 ± 0.02	0.23 ± 0.14	4.37 ± 2.75
5.0e3	0.27 ± 0.02	0.23 ± 0.13	4.15 ± 2.31
1.0e4	0.26 ± 0.02	0.23 ± 0.13	4.13 ± 2.34
5.0e4	0.26 ± 0.01	0.23 ± 0.13	4.23 ± 2.53
1.0e5	0.27 ± 0.02	0.23 ± 0.13	4.10 ± 2.18
$E_{I_0 \leftrightarrow I_1}(v)$			
1.0e0	0.66 ± 0.04	0.68 ± 0.01	1.39 ± 0.03
5.0e1	0.37 ± 0.04	0.30 ± 0.20	2.19 ± 0.56
1.0e2	0.34 ± 0.04	0.29 ± 0.12	2.89 ± 1.26
5.0e2	0.31 ± 0.03	0.25 ± 0.11	3.76 ± 2.32
1.0e3	0.29 ± 0.02	0.22 ± 0.12	4.17 ± 2.74
5.0e3	0.28 ± 0.02	0.21 ± 0.08	4.29 ± 2.62
1.0e4	0.27 ± 0.02	0.19 ± 0.10	4.35 ± 2.58
5.0e4	0.27 ± 0.02	0.19 ± 0.11	4.27 ± 2.30
1.0e5	0.27 ± 0.02	0.19 ± 0.09	4.56 ± 2.55

Table 3 *Registration in real datasets*. Average and standard deviation of the relative L^2 differences, RSSD, the absolute L^2 differences between the transformations obtained from stationary and non-stationary registration, SSD, and the distance between the associated strain matrices, d_{AI}

	RSSD	SSD	d_{AI}
$E_{I_0 \leftrightarrow I_1}(w)$	0.26 ± 0.01	0.26 ± 0.24	0.02 ± 0.02
$E_{I_0 \leftrightarrow I_1}(v)$	0.27 ± 0.02		

desirable. In the following, we select $1/\sigma^2$ equal to 5.0e4, as it guarantees the maximum average image matching in our experiments. In Table 3 we show the average and standard deviation of the RSSD metric, the SSD between the non-stationary and stationary transformations and the d_{AI} between the associated strain matrices for this selection of $1/\sigma^2$.

4.3.2 Visual Evaluation

We show a representative example of diffeomorphic registration. In order to visually assess the registration accuracy in terms of image matching, the image differences before and after registration are shown in Fig. 1. In addition, the histogram of the differences between the template, the target and the corresponding transformed images is shown in Fig. 2. Figure 3 shows the grids of the corresponding transformations. In order to visually localize the differences between the diffeomorphisms obtained with the stationary and the non-stationary parameterizations we also show the corresponding values of d_{AI}.

4.4 Simulated Datasets

The population of simulated anatomical brain images was composed by two groups of 20 images. Each group was generated from simulated diffeomorphic deformations respectively parameterized with stationary and non-stationary vector field flows applied to a template image randomly selected from the real datasets.

Stationary diffeomorphisms were simulated from the vector fields $w_1, \ldots, w_N$ resulting from the registration of the template to the rest of the real datasets using the stationary parameterization. Principal Geodesic Analysis (PGA) was performed on the covariance matrix associated to the w_i residuals using the method in Hernandez (2008). New instances of vector fields were generated from the modes of variation $u_1, \ldots, u_N$ and the corresponding eigenvalues $\lambda_1, \ldots, \lambda_N$ as $w_{\text{new}} = \Sigma_{i=1}^{N} \alpha_i \lambda_i u_i$. The parameters α_i were randomly selected from a normal distribution of zero mean and standard deviation 1. The simulated diffeomorphisms were obtained computing the corresponding group exponential map, Exp. The Jacobian determinant of the resulting simulated transformations ranged from 0.13 to 7.41.

The simulated non-stationary diffeomorphisms were generated from the non-stationary vector field flows $v_1(t), \ldots, v_N(t)$ resulting from the registration of the template to the rest of the real datasets. As the space of time-varying vector field flows is non linear, the computation of PGA was applied to the linear space of initial vector fields $v_1(0), \ldots, v_N(0)$ instead. This way, new instances of initial vector fields were generated as explained in the stationary case. The non-stationary flows associated to the simulated initial vector fields were generated via momentum conservation (15). The simulated diffeomorphisms were obtained solving the corresponding non-stationary transport equation. The Jacobian determinant of the resulting simulated transformations ranged from 0.08 to 8.17.

4.5 Registration in Simulated Datasets

4.5.1 Quantitative Evaluation

In this experiment, we have a priori knowledge of the true transformation. In this case, photometric variations between the template and target images are null. Therefore, the performance of the diffeomorphic transformation in the registration algorithm can be fully evaluated from the RSSD metric. This measure comprises the errors in the registration due to inaccurate matching consequence of diffeomorphic regularization constraints (in this case, the used parameterization). The simulated diffeomorphisms have been compared with the ones obtained via registration using the SSD metric between the transformations and the distance between the associated strain matrices. Table 4 presents the average and standard deviation of these measurements.

4.5.2 Visual Evaluation

We show two representative examples of diffeomorphic registration associated to the simulated datasets generated from the stationary and the non-stationary diffeomorphisms, respectively. In order to visually assess the registration accuracy Fig. 4 shows the image differences before and after registration. In addition, the histogram of the differences between the template, the target and the corresponding transformed images is shown in Fig. 5. Figures 6 and 7 show the grids of the corresponding transformations and the distance between the associated strain matrices.

4.6 Computational Complexity

Computational complexity is measured in terms of memory and time requirements. The optimization for diffeomorphic registration in $E_{I_0 \leftrightarrow I_1}(v)$ and $E_{I_0 \leftrightarrow I_1}(w)$ requires the storage in memory of the diffeomorphic path parameterizations $v(t)$ and w, respectively, and the corresponding energy gra-

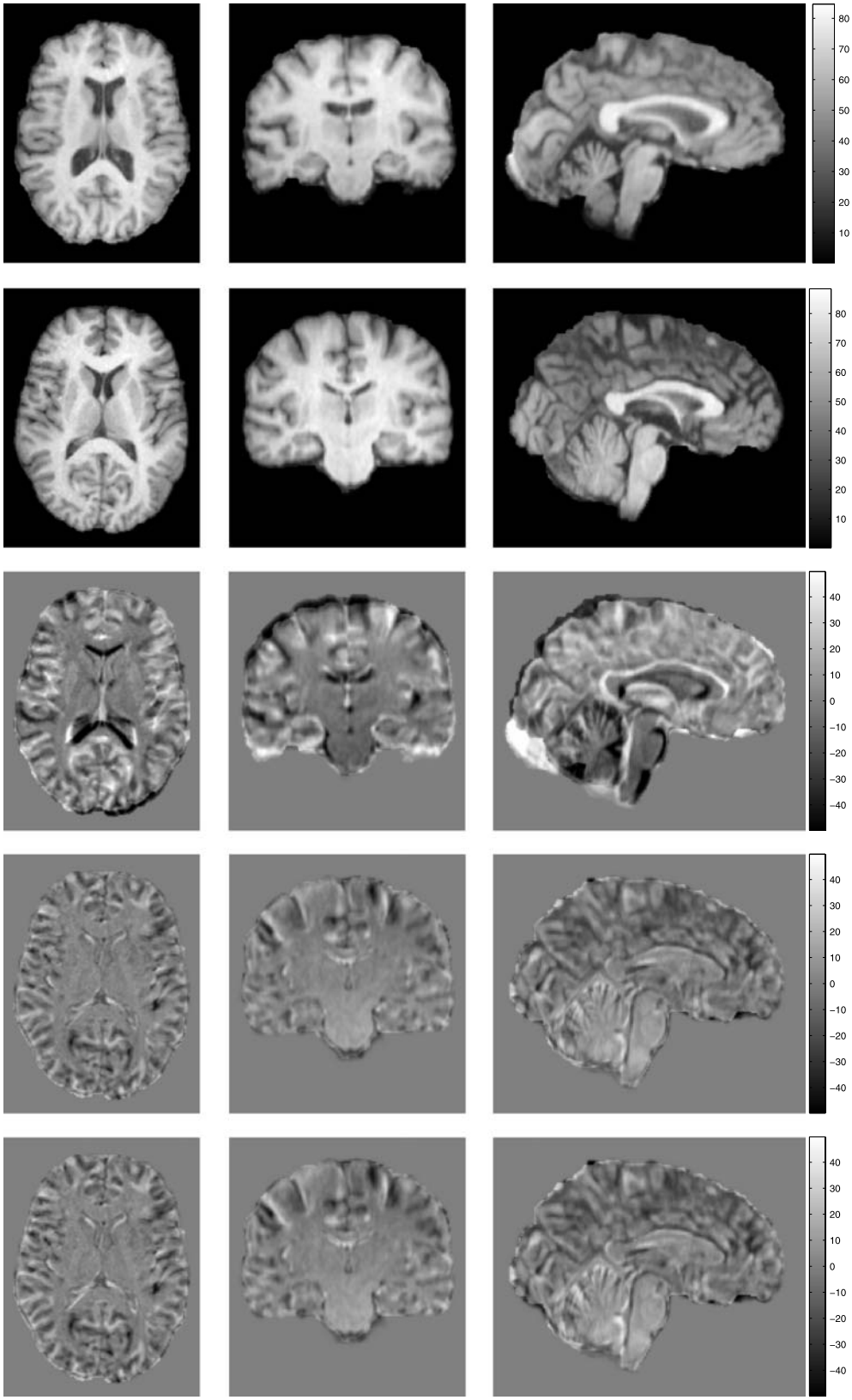

Fig. 1 *Registration in real datasets.* Illustration of axial, coronal and sagittal views of a registration experiment. The *first* and *second rows* show the template (I_0) and the target (I_1) used for registration. The *third row* shows the differences between the template and the target before registration. The *fourth* and *fifth rows* show the differences between the target and the deformed template obtained with $E_{I_0 \leftrightarrow I_1}(w)$ and $E_{I_0 \leftrightarrow I_1}(v)$, respectively. In this experiment, the image matching resulted equal to 0.26 in the case of stationary parameterization and 0.30 in the case of non-stationary parameterization

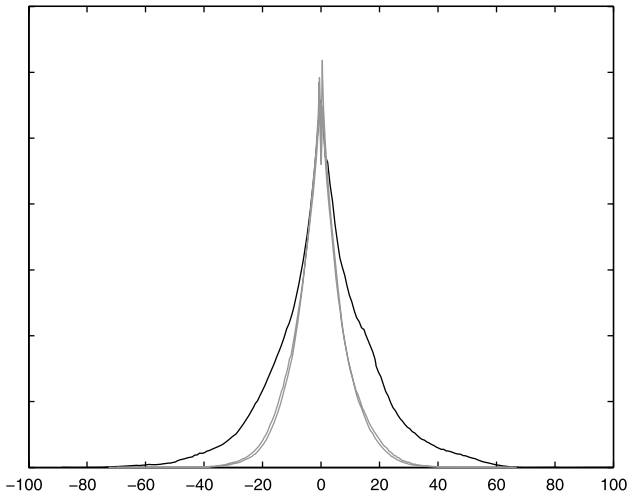

Fig. 2 *Registration in real datasets.* Histograms of the intensity differences. *Black plot* corresponds to the differences before registration ($I_0 - I_1$). *Grey plots* correspond to the differences after registration using the stationary and the non-stationary parameterizations, ($I_0 \circ \varphi^{-1} - I_1$ and $I_0 \circ \phi(1)^{-1} - I_1$). Each histogram has been plotted just taking into account differences not equal to zero

dients $\nabla_v E(v)(t)$ and $\nabla_w E(w)$, respectively. If the path of diffeomorphisms in the non-stationary parameterization is sampled into T pieces, memory and time requirements in this algorithm linearly increase with this parameter. As ex-

Table 4 *Registration in simulated datasets.* Average and standard deviation of the relative L^2 differences, RSSD, the absolute L^2 differences between the transformations obtained from registration and the ground truth transformations, SSD, and the distance between the associated strain matrices, d_{AI}. Up table shows the results obtained with the simulated datasets generated from the stationary parameterization. Down table shows the results corresponding to the non stationary parameterization

	RSSD	SSD	d_{AI}
$E_{I_0 \leftrightarrow I_1}(w)$	0.03 ± 0.00	0.48 ± 0.45	0.03 ± 0.05
$E_{I_0 \leftrightarrow I_1}(v)$	0.03 ± 0.01	0.50 ± 0.46	0.03 ± 0.05
	RSSD	SSD	d_{AI}
$E_{I_0 \leftrightarrow I_1}(w)$	0.02 ± 0.01	0.44 ± 0.35	0.04 ± 0.05
$E_{I_0 \leftrightarrow I_1}(v)$	0.03 ± 0.01	0.45 ± 0.37	0.03 ± 0.05

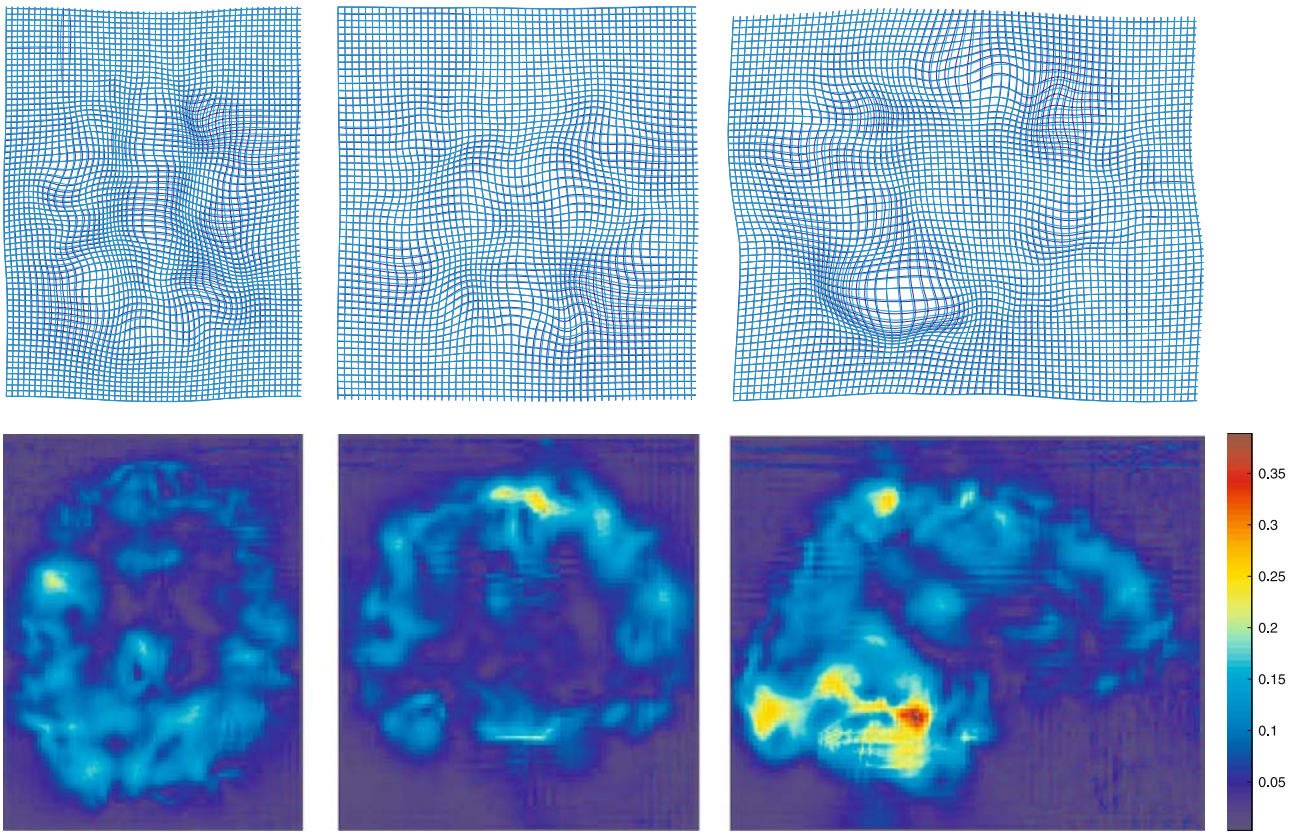

Fig. 3 *Registration in real datasets. Upper row*, projection of the diffeomorphic transformations in axial, coronal and sagittal views. *Blue grid* corresponds to the stationary and *cyan grid* to the non-stationary transformations. *Lower row*, illustration of the distance between the corresponding strain matrices, d_{AI}

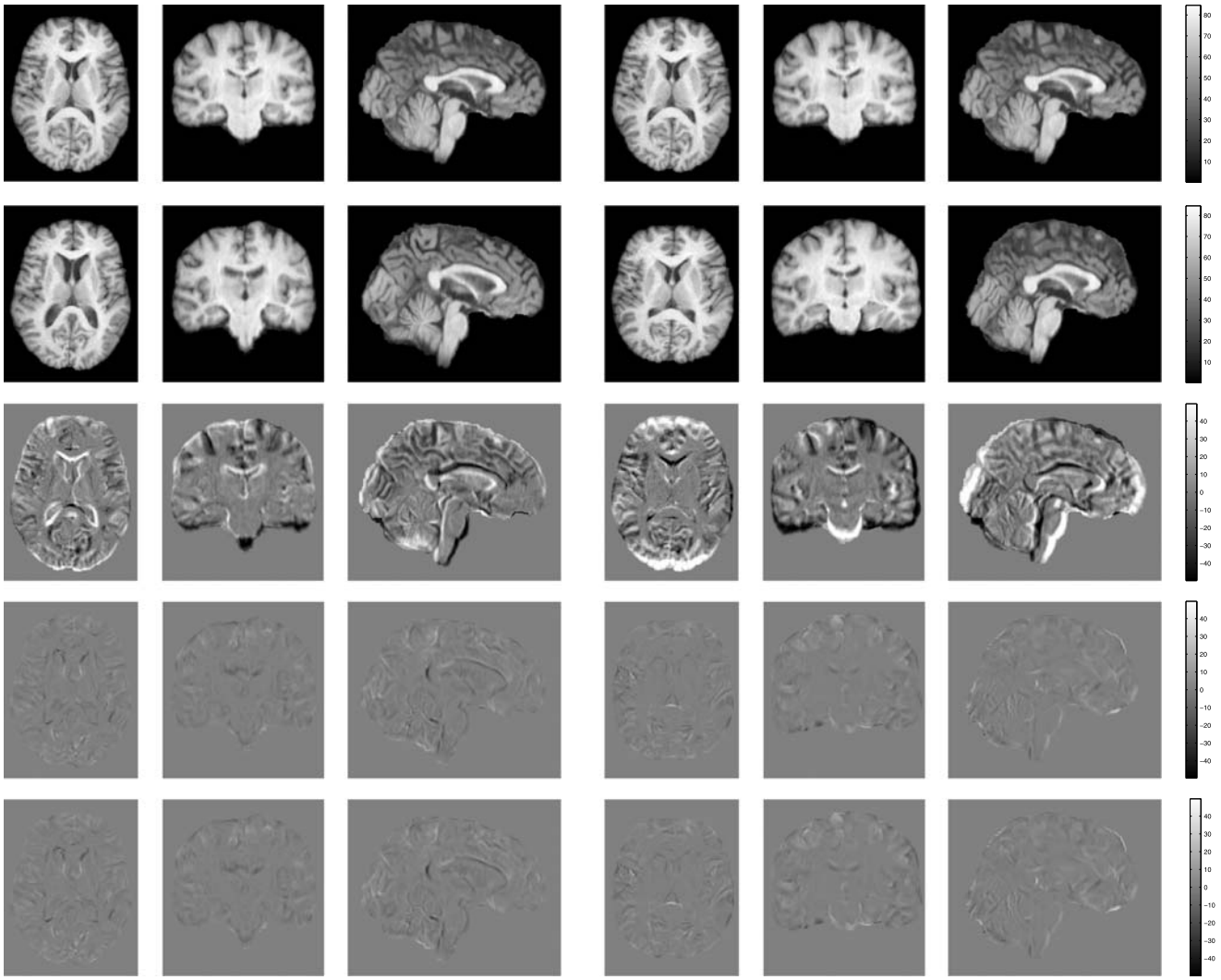

Fig. 4 *Registration in simulated datasets.* Illustration of axial, coronal and sagittal views of simulated data registration experiments generated from stationary (*left*) and non-stationary (*right*) diffeomorphisms. The *first* and *second rows* from each example show the template (I_0) and the target (I_1) used for registration. The *third row* shows the image differences before registration. The *fourth* and *fifth rows* show the differences between the target and the deformed template obtained with $E_{I_0 \leftrightarrow I_1}(w)$ and $E_{I_0 \leftrightarrow I_1}(v)$, respectively. In the experiment generated from stationary diffeomorphisms, the image matching resulted equal to 0.04 in both parameterizations. In the experiment generated from non-stationary diffeomorphisms, the image matching resulted equal to 0.01 in the case of stationary parameterization and 0.02 in the case of non-stationary parameterization

ample,[1] in a volume of size $155 \times 205 \times 170$ registration using the non-stationary parameterization in the finer resolution level required up to 1.9 GB while stationary parameterization required about 800 MB. Time requirements for a single iteration took up to 143.20 seconds using the non-stationary parameterization whereas the stationary parameterization took 10.25 seconds in a machine of 2327 MHz. This supposes a considerable time reduction of the whole registration algorithm.

[1] Code implemented in C++ based on the ITK library.

5 Discussion and Conclusions

In this article, we have presented a method for diffeomorphic registration based on the Large Deformation paradigm studied in Computational Anatomy. In contrast to traditional methods, we estimate the optimal transformation connecting two anatomical images constrained to lie on paths of diffeomorphisms parameterized by stationary vector field flows.

The performance of the stationary vs non-stationary parameterizations has been compared in a set of 18 MRI real brain datasets. Both algorithms have similar accuracy in terms of image matching (in the optimal case shown in Table 2, the average RSSD resulted to be 0.26 vs 0.27). The

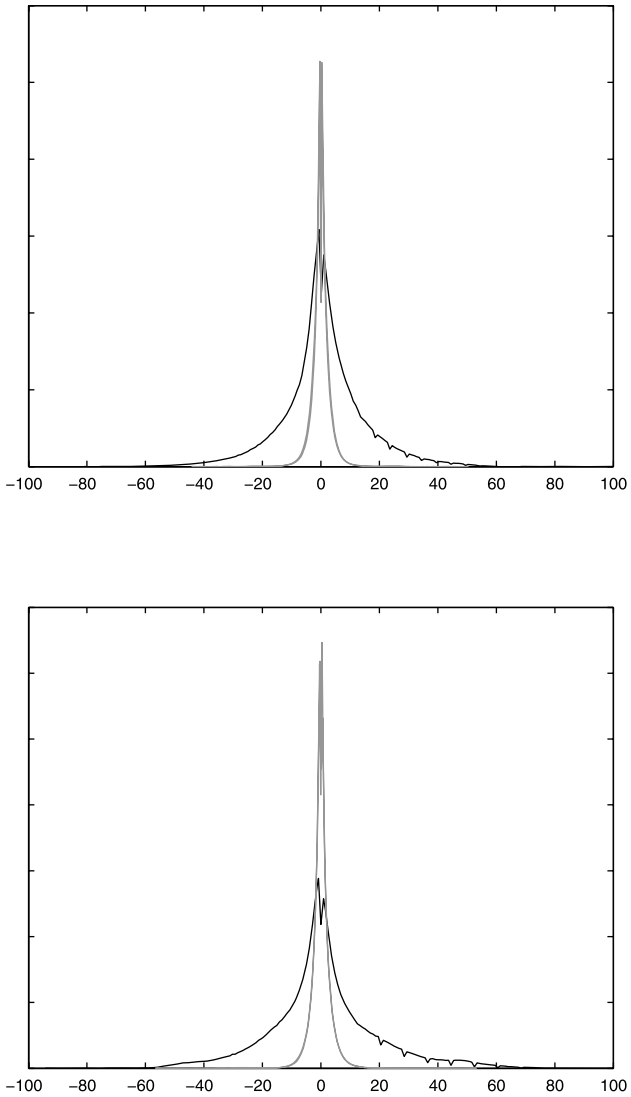

Fig. 5 *Registration in simulated datasets.* Histogram of the intensity differences. *Up figure* corresponds to the simulated example generated from a stationary diffeomorphism. *Down figure* corresponds to the example generated from a non-stationary transformation. In both figures, *black plot* corresponds to the differences before registration. *Grey plots* correspond to the differences after registration using the stationary and the non-stationary algorithms, respectively. The histogram has been plotted just taking into account differences not equal to zero

similarity between the diffeomorphic transformations was under half the voxel resolution (in average, SSD = 0.26 and $d_{AI} = 0.02$). In the representative example, the differences before and after registration were reduced in the same degree in both cases (Figs. 1 and 2). Subcortical structures presented a maximum image matching. However, image matching seemed to fail in the cortex. This may be due to the differences of intensity between images are higher in these regions of the brain. Another reason could be that the regularization constraints imposed on diffeomorphic registration makes the transformation not able to warp between structures with this high geometrical variability. The transformations presented local differences located at the cerebellum and the cortex with a maximum $RVC \leq 1.67$.[2]

In addition, the performance of both algorithms has been compared in two populations of 20 MRI datasets generated from simulated stationary and non-stationary parameterized diffeomorphisms. As pointed out by Table 4 and supported by Figs. 4 and 5, both algorithms have shown a high image matching (average RSSD less than 0.03 in all cases) regardless the parameterization used to generate the simulated diffeomorphisms. The average differences of corresponding grid points between the simulated diffeomorphisms and the ones obtained via registration resulted to be within half the voxel resolution in all cases. The average d_{AI} was less than 0.04 in all cases. In the representative examples the grids presented local differences located at the cerebellum, the ventricles, and the cortex with a maximum $RVC \leq 2.06$.

Regarding time and memory requirements, our algorithm has shown to provide a considerable reduction of the computational requirements for registration with identical accuracy results. For this reason, our algorithm may provide an alternative fast method for computing diffeomorphic registration in state of the art Computational Anatomy applications (for example, in the computation of anatomical atlases from group-wise diffeomorphic registration (Joshi et al. 2004) or temporal regression (Davis et al. 2007)). Moreover, our algorithm allows to generate elements belonging to one-parameter subgroups of diffeomorphisms where Log-Euclidean statistics on diffeomorphisms can be performed using algebraic techniques (Arsigny et al. 2006a).

As a possible limitation to the use of the stationary parameterization in Computational Anatomy, it has been shown that the set of diffeomorphisms obtained with the stationary parameterization does not span all diffeomorphisms in $Diff^s(\Omega)$. In fact, the group exponential map is not onto (Grabowski 1988). This means that there exist diffeomorphisms arbitrarily close to the identity that cannot be parameterized by stationary vector field flows and, therefore, there may exist two images where the non-stationary parameterization would provide much better registration performance than the stationary parameterization. Nevertheless, the experiments reported in this work show that, at least for MRI anatomical brain images, one can find elements from both parameterizations that provide similar and acceptable registration results in terms of image matching and transformation similarity.

As future directions, it would be interesting to explore the equivalence of both parameterizations in the registration of a wider range of datasets, different anatomies and Computational Anatomy applications.

[2]The maximum is reached if torsion free deformations provide the value 1.67.

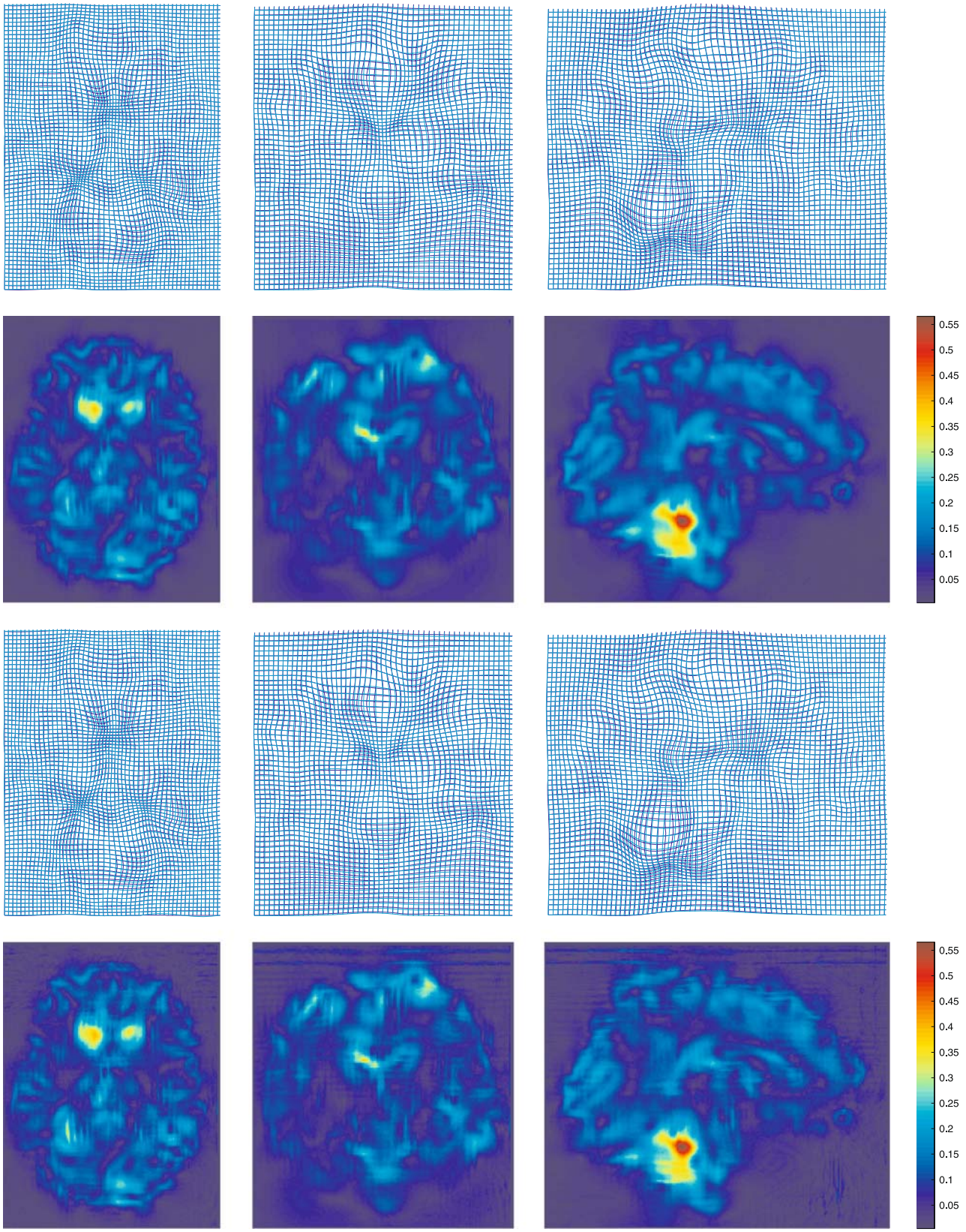

Fig. 6 *Registration in simulated datasets generated from stationary diffeomorphisms.* The *first two rows* correspond to registration results associated to the stationary parameterization and the last two rows correspond to the non-stationary parameterization. From each example, the *upper row* shows the projection of the diffeomorphic transformations in axial, coronal and sagittal views. The *blue grid* corresponds to the ground truth and the *cyan grid* to the transformation obtained from registration. The *lower row*, illustrates the distance between the corresponding strain matrices, d_{AI}

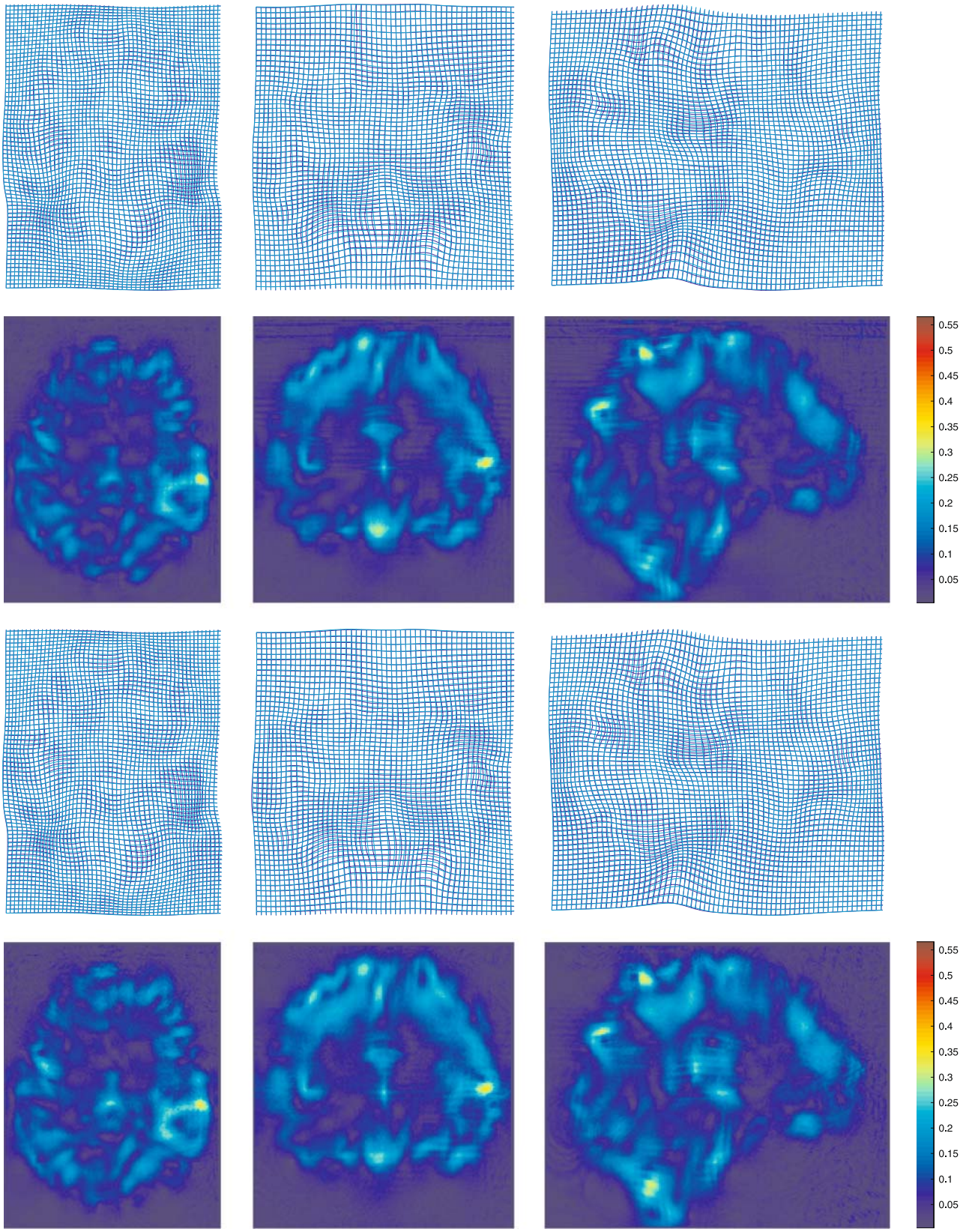

Fig. 7 *Registration in simulated datasets generated from non-stationary diffeomorphisms.* The *first two rows* correspond to registration results associated to the stationary parameterization and the *last two rows* correspond to the non-stationary parameterization. From each example, the *upper row* shows the projection of the diffeomorphic transformations in axial, coronal and sagittal views. The *blue grid* corresponds to the ground truth and the *cyan grid* to the transformation obtained from registration. The *lower row*, illustrates the distance between the corresponding strain matrices, d_{AI}

Acknowledgements The authors would like to acknowledge to Dr. C. Junque from Clinic Hospital, Barcelona, Spain for providing the MRI brain datasets. This work has been partially supported by research grants MEC TEC2006-13966-C03-02 and FIS PI04/1795. Finally, the authors would like to acknowledge to the anonymous reviewers for their useful comments and suggestions that contributed to the improvement of this manuscript.

Appendix: Euler-Lagrange equation for diffeomorphic registration

In this appendix, we present the computations to obtain the Euler-Lagrange equation for the energy functional given in (24). In general convex vector spaces, the Euler-Lagrange equation associated to a Frechet differentiable energy functional $E(w)$ is obtained from $\nabla_w E(w) = 0$. In Frechet spaces, the gradient operator relates the Frechet derivative and the Gateaux derivative (whenever both derivatives exist) by $\partial_h E(w) = \langle \nabla_w E(w), h \rangle_V$. Thus, in our case, the Euler-Lagrange equation is computed from Gateaux derivatives.

The Gateaux derivative of the energy functional along $h \in V$ is defined as the variation of $E(w)$ under the perturbation of w in the direction of h

$$\partial_h E(w) = \lim_{\epsilon \to 0} \frac{E(w + \epsilon h) - E(w)}{\epsilon} \tag{31}$$

For simplicity, we divide the computation of $\partial_h E(w)$ into $\partial_h E_1(w)$, $\partial_h E_2(w)$, and $\partial_h E_3(w)$ where

$$E_1(w) = \|w\|_V^2, \tag{32}$$

$$E_2(w) = \frac{1}{\sigma^2} \| I_0 \circ \mathrm{Exp}(w)^{-1} - I_1 \|_{L^2}^2, \tag{33}$$

and

$$E_3(w) = \frac{1}{\sigma^2} \| I_1 \circ \mathrm{Exp}(w) - I_0 \|_{L^2}^2 \tag{34}$$

Straightforward computations provide

$$\partial_h E_1(w) = 2\langle w, h \rangle_V \tag{35}$$

The chain rule allows to compute the variation of $E_2(w)$ and $E_3(w)$

$$\partial_h E_2(w) = \frac{2}{\sigma^2} \langle I_0 \circ \mathrm{Exp}(w)^{-1} - I_1, \nabla (I_0 \circ \mathrm{Exp}(w)^{-1}) \times \partial_h \mathrm{Exp}(w)^{-1} \rangle_{L^2} \tag{36}$$

$$\partial_h E_3(w) = \frac{2}{\sigma^2} \langle I_1 \circ \mathrm{Exp}(w) - I_0, \nabla (I_1 \circ \mathrm{Exp}(w)) \times \partial_h \mathrm{Exp}(w) \rangle_{L^2} \tag{37}$$

The Gateaux derivative of the exponential map is obtained using a first order approximation, $\mathrm{Exp}(w) = x + w$. The Gateaux derivative of the inverse exponential map is obtained from the fact $\mathrm{Exp}(w)^{-1} = \exp(-w)$. Thus, $\partial_h \mathrm{Exp}(w) = h$ and $\partial_h \mathrm{Exp}(w)^{-1} = -h$ and

$$\partial_h E_2(w) = -\frac{2}{\sigma^2} \langle (L^\dagger L)^{-1} ((I_0 \circ \mathrm{Exp}(w)^{-1} - I_1) \times \nabla (I_0 \circ \mathrm{Exp}(w)^{-1})), h \rangle_V \tag{38}$$

$$\partial_h E_3(w) = \frac{2}{\sigma^2} \langle (L^\dagger L)^{-1} ((I_1 \circ \mathrm{Exp}(w) - I_0) \times \nabla (I_1 \circ \mathrm{Exp}(w))), h \rangle_V \tag{39}$$

Collecting the results in (35, 38, and 39), the Euler-Lagrange equation associated to the energy functional is

$$\nabla_w E(w) = 2w + \partial_h E_2(w) + \partial_h E_3(w) \tag{40}$$

References

Arnold, V. I. (1989). *Mathematical methods of classical mechanics*. Berlin: Springer.

Arsigny, V., Commonwick, O., Pennec, X., & Ayache, N. (2006a). Statistics on diffeomorphisms in a Log-Euclidean framework. In *Lecture notes in computer science (LNCS): Vol. 4190. Proc. of the 9th international conference on medical image computing and computer assisted intervention (MICCAI'06)* (pp. 924–931). Berlin: Springer.

Arsigny, V., Pennec, X., & Ayache, N. (2006b). *Bi-invariant means in Lie groups. Application to left-invariant polyaffine transformations* (Research Report RR-5885). INRIA Sophia-Antipolis.

Avants, B., & Gee, J. C. (2004). Shape averaging with diffeomorphic flows for atlas creation. In *Proc. of the 2nd IEEE international symposium on biomedical imaging: from nano to macro (ISBI'04)* (pp. 595–598) 2004.

Avants, B., Schoenemann, P. T., & Gee, J. (2006). Lagrangian frame diffeomorphic image registration: morphometric comparison of human and chimpanzee cortex. *Medical Image Analysis*, *10*(3), 397–412.

Avants, B. B., Epstein, C. L., Grossman, M., & Gee, J. C. (2008). Symmetric dieomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, *12*, 26–41.

Beg, M. F. (2003). *Variational and computational methods for flows of diffeomorphisms in image matching and growth in computational anatomy*. Ph.D. thesis, John Hopkins University, USA.

Beg, M. F., & Khan, A. (2006). Computing an average anatomical atlas using LDDMM and geodesic shooting. In *Proc. of the 3rd IEEE international symposium on biomedical imaging: from nano to macro (ISBI'06)* (pp. 1116–1119) 2006.

Beg, M. F., & Khan, A. (2007). Symmetric data attachment terms for large deformation image registration. *IEEE Transactions on Medical Imaging*, *26*(9), 1179–1189.

Beg, M. F., Miller, M. I., Trouve, A., & Younes, L. (2005). Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International Journal Computer Vision*, *61* (2), 139–157.

Christensen, G. E. (1999). Consistent linear-elastic transformations for image matching. In *Lecture notes in computer science (LNCS). Proc. of international conference on information processing and medical imaging (IPMI'99)* (pp. 224–237). Berlin: Springer.

Christensen, G. E., Rabbitt, R. D., & Miller, M. I. (1996). Deformable templates using large deformation kinematics. *IEEE Transactions on Image Processing*, *5*(10), 1435–1447.

Cotter, C. J., & Holm, D. D. (2006). Singular solutions, momentum maps and computational anatomy. In *Proc. of the 1st international workshop on mathematical foundations of computational anatomy (MFCA'06)* (pp. 18–28) 2006.

Csernansky, J. G., Wang, L., Joshi, S. C., Ratnanather, J. T., & Miller, M. I. (2004). Computational anatomy and neuropsychiatric disease: probabilistic assessment of variation and statistical inference of group difference, hemispheric asymmetry, and time-dependent change. *Neuroimage*, *23*(1), 56–68.

Davis, B., Fletcher, P. T., Bullit, E., & Joshi, S. (2007). Population shape regression from random design data. In *Proc. of the 11th IEEE international conference on computer vision (ICCV'07)* 2007.

DoCarmo, M. P. (1992). *Riemannian geometry*. Boston: Birkhauser.

Dodgas, B., Sattuck, D. W., & Leahy, R. M. (2005). Segmentation of skull and scalp in 3D human MRI using mathematical morphology. *Human Brain Mapping*, *26*(4), 273–285.

Dupuis, P., Grenander, U., & Miller, M. (1998). Variational problems on flows of diffeomorphisms for image matching. *Quarterly of Applied Mathematics*, *3*, 587–600.

Ebin, D., & Marsden, J. (1970). Groups of diffeomorphisms and the motion of an incompressible fluid. *Annals of Mathematics*, *92*, 102–103.

Garcin, L., & Younes, L. (2006). Geodesic matching with free extremities. *Journal of Mathematical Imaging Vision*, *25*, 329–340.

Gerig, G., Davis, B., Lorenzen, P., Xu, S., Jomier, M., Piven, J., & Joshi, S. (2006). Computational anatomy to assess longitudinal trajectory of brain growth. In *Proc. of the 3rd international symposium on 3D data processing, visualization, and transmission* (pp. 1041–1047) 2006.

Grabowski, J. (1988). Free subgroups of diffeomorphism groups. *Fundamenta Mathematicae*, *131*, 103–121.

Grenander, U. (1994). *General pattern theory*. Oxford: Oxford University Press.

Hernandez, M. (2008). *Variational techniques with applications to segmentation and registration of medical images*. Ph.D. Thesis, University of Zaragoza, Spain.

Holm, D. D., Ratnanather, J. T., Trouve, A., & Younes, L. (2004). Soliton dynamics in computational anatomy. *Neuroimage*, *23*, 170–178.

Joshi, S., Davis, B., Jomier, M., & Gerig, G. (2004). Unbiased diffeomorphic atlas construction for computational anatomy. *Neuroimage*, *23*, 151–160.

Leow, A., Klunder, A. D., Jack, C. R., & Toga, A. W. (2006). Longitudinal stability of MRI for mapping brain change using tensor-based morphometry. *Neuroimage*, *31*, 627–640.

Lepore, N., Brun, C. A., Chiang, M. C., Chou, Y. Y., Dutton, R. A., Hayashi, K. M., Lopez, O. L., Aizenstein, H. J., Toga, A. W., Becker, J. T., & Thompson, P. M. (2006). Multivariate statistics of the Jacobian matrices in tensor based morphometry and their application to HIV/AIDS. In *Lecture notes in computer science (LNCS): Vol. 4190. Proc. of the 9th international conference on medical image computing and computer assisted intervention (MICCAI'06)* (pp. 191–198). Berlin: Springer.

Lorenzen, P., Prastawa, M., Davis, B., Gerig, G., Bullitt, E., & Joshi, S. (2006). Multi-modal image set registration and atlas formation. *Medical Image Analysis*, *10*, 440–451.

Michor, P. W., & Mumford, D. (2006). An overview of the Riemannian metrics on spaces of curves using the Hamiltonian approach. *Applied and Computational Harmonic Analysis*, *23*(1), 74–113.

Miller, M. I. (2004). Computational anatomy: shape, growth, and atrophy comparison via diffeomorphisms. *Neuroimage*, *23*, 19–33.

Miller, M. I., Trouve, A., & Younes, L. (2006). Geodesic shooting for computational anatomy. *Journal of Mathematical Imaging and Vision*, *24*, 209–228.

Modersitzki, J. (2004). *Numerical methods for image registration*. Oxford: Oxford University Press.

Nocedal, J., & Wright, S. J. (1999). *Numerical optimization*. New-York: Springer.

Noether, E. (1918). Invariante variationsprobleme. *Nachr. v. d. Ges. d. Wiss. zu Gottingen*, 235–257.

Qiu, A., Younes, L., Miller, M. I., & Csernansky, J. G. (2007). Parallel transport in diffeomorphisms distinguishes the time-dependent pattern of hippocampal surface deformation due to healthy aging and the dementia of the Alzheimer's type. *Neuroimage*, *40*, 68–76.

Schmid, R. (2004). Infinite dimensional Lie groups with applications to mathematical physics. *Journal of Geometry Symmetry in Physics*, *1*, 54–120.

Staniforth, A., & Cote, J. (1991). Semi-Lagrangian integration schemes for atmospheric models-a review. *Mon. Weather Rev.*, *119*, 2206–2223.

Thompson, P. M., Giedd, J. N., Woods, R. P., MacDonald, D., Evans, A. C., & Toga, A. W. (2000). Growth patterns in the developing brain detected by using continuum mechanical tensor maps. *Nature*, *404*, 190–193.

Thompson, P. M., Mega, M. S., Woods, R. P., Zoumalan, C. I., Lindshield, C. J., Blanton, R. E., Moussai, J., Holmes, C. J., Cummings, J. L., & Toga, A. W. (2001). Cortical change in Alzheimer's disease detected with a disease-specific population-based brain atlas. *Cerebral Cortex*, *11*(1), 1–16.

Wang, L., Swank, J. S., Glick, I. E., Gado, M. H., Miller, M. I., Morris, J. C., & Csernansky, J. G. (2003). Changes in hippocampal volume and shape across time distinguish dementia of the Alzheimer type from healthy aging. *Neuroimage*, *20*, 667–682.

Wang, L., Beg, F., Ratnanather, T., Ceritoglu, C., Younes, L., Morris, J. C., Csernansky, J. G., & Miller, M. I. (2007). Large deformation diffeomorphism and momentum based hippocampal shape discrimination in dementia of the Alzheimer type. *IEEE Transactions on Medical Imaging*, *26*(4), 462–470.

Younes, L. (2007). Jacobi fields in groups of diffeomorphisms and applications. *Quarterly of Applied Mathematics*, *65*, 113–134.