Left-Invariant Riemannian Elasticity: a distance on shape diffeomorphisms ?

Xavier Pennec

Left-Invariant Riemannian Elasticity: a distance on shape diffeomorphisms?

X. Pennec

INRIA Sophia - Projet Epidaure, 2004 Route des Lucioles BP 93
06902 Sophia Antipolis Cedex, France
Xavier.Pennec@sophia.inria.fr

Abstract. In inter-subject registration, one often lacks a good model of the transformation variability to choose the optimal regularization. Some works attempt to model the variability in a statistical way, but the re-introduction in a registration algorithm is not easy. In [1], we interpreted the elastic energy as the distance of the Green-St Venant strain tensor to the identity. By changing the Euclidean metric for a more suitable Riemannian one, we defined a consistent statistical framework to quantify the amount of deformation. In particular, the mean and the covariance matrix of the strain tensor could be efficiently computed from a population of non-linear transformations and introduced as parameters in a Mahalanobis distance to measure the statistical deviation from the observed variability. This statistical Riemannian elasticity was able to handle anisotropic deformations but its isotropic stationary version was locally inverse-consistent. In this paper, we investigate how to modify the Riemannian elasticity to make it globally inverse consistent. This allows to define a left-invariant "distance" between shape diffeomorphisms that we call the left-invariant Riemannian elasticity. Such a closed form energy on diffeomorphisms can optimize it directly without relying on a time and memory consuming numerical optimization of the geodesic path.

1 Introduction

Most non-linear image registration algorithms optimize a criterion including an image intensity similarity and a regularization term. Many image similarity criteria are now available, ranging from the simple sum of squared intensity differences to robust information theory based measures. In inter-subject registration, the main problem is not really the intensity similarity measure but rather the regularization criterion. Some authors used physical models like elasticity or fluid models [2, 3]. For efficiency reasons, other authors proposed to use non-physical but efficient regularization methods like Gaussian filtering [4–6]. This type of regularization was then extended to more general isotropic vectorial filters [7], and to non-stationary regularization criteria in order to take into account some anatomical information about the tissue types [8, 9].

However, since we do not have in general a model of the deformation of organs across subjects, no regularization criterion is obviously more justified than the others. We could think of of relating the anatomy of two different subjects by building a model of the organ growth: inverting the model from the first subject to a sufficiently early stage and growing toward the second subject image would allow to relate the two anatomies. However, such a computational model is out of reach now, and most of the existing work in the literature rather try to capture the organ variability from a statistical point of view on a representative population of subjects (see e.g. [10–12]). Although the image databases are now large enough to be representative of the organ variability, the problem remains of how to use this information to better guide inter-subject registration.

Ashburner at al observed in [13] that, *as the structural variability is often greater in certain directions [14], some form of a tensor field describing normal variability in each direction may be appropriate. A data representation of this form, together with a canonical brain template and associated error variance image, would allow anatomical comparisons to be made against the normal population.* This is in essence what we proposed with the Statistical Riemannian Elasticity [1]: an integrated framework to compute statistics on deformations and reintroduce them in the registration procedure, based on the field of strain tensors. The basic idea is to interpret the elastic energy as a distance in the space of positive definite symmetric matrices (tensors). By changing the classical Euclidean metric for a more suitable one, namely a log-Euclidean one in [1], we defined a natural framework for computing statistics on the strain tensor. A related idea was already present in [13] with a regularization prior based on a log-Gaussian distribution of the singular values of the Jacobian matrix of the transformation. Our key contribution in [1] was to consider the strain tensor instead of the Jacobian of the transformation. This allows to easily extend such an isotropic and stationary prior to anisotropic and non stationary ones.

In this paper, the goal is to better understand the link between Riemannian elasticity and invariant metrics on groups of diffeomorphisms, as used for instance in [15, 16]. We also reformulate the derivation of the whole theory to better stress the link with classical mechanics. We first detail how the standard elastic regularization can be optimized in a gradient descent based registration algorithm. Then, we introduce in Section 3 the Riemannian elasticity energy by changing the Euclidean distance on the strain Tensor to the identity by a log-Euclidean Riemannian distance. The simplest distances are the isotropic ones: the energy expression turns out to be very similar to the classical elastic energy while being locally inverse-consistent. One can also include non-stationary and anisotropic statistics on the strain tensors observed in a population by taking the Mahalanobis distance on the logarithmic strain tensor (statistical Riemannian elasticity). The gradients of these Riemannian elastic criteria needed to implement a practical registration algorithm are detailed in Section 4. In Section 5, we modify the spatial integration of the isotropic Riemannian elasticity in order to make it globally inverse-consistent. This leads to a left- (or right-) invariant energy on shape diffeomorphisms that can be optimized directly without having to find the geodesics through an optimization process as in standard diffeomorphic matching algorithm.

2 Standard elastic regularization

Let $I(x)$ and $J(x)$ the intensity functions of two images and $\Phi(x)$ be a non-linear space transformation assumed to be diffeomorphic with a positive Jacobian everywhere. We denote by $\{e_\alpha\}$ a set of orthonormal vectors (a basis) of the three-dimensional space, and by $\partial_\alpha \Phi$ the directional derivatives of the transformation along the spaces axis α. The general registration method is to optimize an energy of the type: $C(\Phi) = Sim(\text{Images}, \Phi) + Reg(\Phi)$. Starting from an initial transformation Φ_0, a first order gradient descent methods computes the gradient of the energy $\nabla C(\Phi)$, and update the transformation using: $\Phi_{t+1} = \Phi_t - \eta \nabla C(\Phi_t)$. From a computational points of view, this Lagrangian framework can be advantageously changed into a Eulerian framework to better conserve the diffeomorphic nature of the mappings [9]. In the following, we do not focus on the optimization of the similarity criterion (see e.g. [5, 6]), but rather on the computation of the gradient of the regularization. We assume Neumann boundary conditions on transformations and an invariant integration domain ($\Phi(\Omega) = \Omega$), so that we can drop the integration domain to simplify notations.

2.1 Elastic deformations

In continuum mechanics [17], one characterizes the deformation of an infinitesimal volume element in the Lagrangian framework using the Cauchy-Green tensor $\Sigma = \nabla\Phi^{\mathrm{T}}\,\nabla\Phi = \sum_\alpha \partial_\alpha\Phi\,\partial_\alpha\Phi^{\mathrm{T}}$. This symmetric matrix is positive definite if the transformation is diffeomorphic, and measures the local amount of non-rigidity. Let $\nabla\Phi = V\,S\,R^{\mathrm{T}}$ be a singular value decomposition of the transformation Jacobian (R and V are two rotation matrices and S is the diagonal matrix of the positive singular values). The Cauchy-Green tensor $\Sigma = R\,S^2\,R^{\mathrm{T}}$ is equal to the identity if and only if the transformation is locally a rigid transformation. Eigenvalues between 0 and 1 indicate a local compression of the material along the associated eigenvector, while a value above 1 indicates an expansion.

To quantify the deformation, one usually prefers the related Green-St Venant strain tensor $E = \frac{1}{2}(\Sigma - \mathrm{Id})$, whose eigenvalues are null for no deformation. This tensor is often expressed using the displacement field: $E = \frac{1}{2}(\nabla U + \nabla U^{\mathrm{T}} + \nabla U^{\mathrm{T}}\,\nabla U)$ (dropping the quadratic term leads to the linear elasticity). Assuming an isotropic material and a linear Hooks law to relate strain and stress tensors, one can show that the motion equations derive from the *St Venant-Kirchoff elasticity* energy [17]:

$$Reg_{SVKE}(\Phi) = \int \mu \mathrm{Tr}(E^2) + \frac{\lambda}{2}\mathrm{Tr}(E)^2 = \int \frac{\mu}{4}\mathrm{Tr}\left((\Sigma - \mathrm{Id})^2\right) + \frac{\lambda}{8}\mathrm{Tr}(\Sigma - \mathrm{Id})^2$$

2.2 Optimizing the elasticity

To minimize this energy in a registration algorithm, we need its gradient. Since $\partial_u \Sigma = \sum_\alpha \left(\partial_\alpha\Phi\,\partial_\alpha u^{\mathrm{T}} + \partial_\alpha u\,\partial_\alpha\Phi^{\mathrm{T}}\right)$, the derivative of the elastic energy in the direction (i.e. displacement field) u is:

$$\begin{aligned}\partial_u Reg_{SVKE}(\Phi) &= \int \tfrac{\mu}{2}\mathrm{Tr}((\Sigma - \mathrm{Id})\,\partial_u\Sigma) + \tfrac{\lambda}{4}\mathrm{Tr}(\Sigma - \mathrm{Id})\,\mathrm{Tr}(\partial_u\Sigma)\\ &= \textstyle\sum_\alpha \int \langle\, \mu\,(\Sigma - \mathrm{Id})\,\partial_\alpha\Phi \mid \partial_\alpha u\,\rangle + \tfrac{\lambda}{2}\mathrm{Tr}(\Sigma - \mathrm{Id})\,\langle\, \partial_\alpha\Phi \mid \partial_\alpha u\,\rangle\end{aligned}$$

Using an integration by part with homogeneous Neumann boundary conditions [6], we have $\int \langle\, v \mid \partial_\alpha u\,\rangle = -\int \langle\, \partial_\alpha v \mid u\,\rangle$, so that the gradient is finally:

$$\nabla Reg_{SVKE}(\Phi) = -\textstyle\sum_\alpha \partial_\alpha\big(Z\,\partial_\alpha\Phi\big) \quad \text{with} \quad Z = \mu(\Sigma - \mathrm{Id}) + \tfrac{\lambda}{2}\mathrm{Tr}(\Sigma - \mathrm{Id})\,\mathrm{Id}$$

Here, Z is the derivative of the density of energy at each point with respect to the strain tensor Σ and is known as the 2nd Piola-Kirchoff tensor. The 3rd order tensor $Z\partial_\alpha\Phi$ is the first Piola-Kirchoff tensor and corresponds to the derivative of the density of energy with respect to the Jacobian of the transformation.

3 Log-Euclidean Riemannian elasticity

In the standard elasticity theory, the deviation of the positive definite symmetric matrix Σ (the strain tensor) from the identity (the rigidity) is measured using the Euclidean matrix distance $\mathrm{dist}^2_{Eucl}(\Sigma, \mathrm{Id}) = \mathrm{Tr}((\Sigma - \mathrm{Id})^2)$. However, it has been argued in recent works that the Euclidean metric is not a good metric for the tensor space because positive definite symmetric matrices only constitute a cone in the Euclidean matrix space. Thus, the tensor space is not complete (null or negative eigenvalues are at a finite distance). For instance, an expansion of a factor $\sqrt{2}$ in each direction (leading to $\Sigma = 2\,\mathrm{Id}$) is at the same Euclidean distance from the identity than the "black hole" transformation $\Phi(x) = 0$ (which has a non physical null strain tensor). In non-linear registration, this asymmetry of the regularization leads to different results if we look for the forward or the backward transformation: this is the inverse-consistency problem [18].

3.1 A Log-Euclidean metric on the strain tensor

To solve the problems of the Euclidean tensor computing, affine-invariant Riemannian metrics were recently proposed [19–22]. Using these metrics, symmetric matrices with null eigenvalues are basically at an infinite distance from any tensor, and the notion of mean value corresponds to a geometric mean, even if it has to be computed iteratively. More recently, [23] proposed Log-Euclidean metrics, which exhibit the same properties while being much easier to compute. As these metrics simply consist in taking a standard Euclidean metric after a (matrix) logarithm, and since they correspond to the previous ones as long as the reference point is the identity, we relied on the later in the original definition of the Riemannian elasticity [1]. However, the Riemannian Elasticity principle can be generalized to any Riemannian metric on the tensor space without any restriction. We will see that the full linear invariance properties of affine-invariant metrics will prove to be necessary in Section 5 to properly define a left-invariant energy on shape diffeomorphisms.

In the log-Euclidean Riemannian framework, the deviation between the tensor Σ and the identity is the tangent vector $\log(\Sigma) - \log(\mathrm{Id}) = \log(\Sigma)$. Interestingly, this tensor is known in continuum mechanics as the logarithmic or Hencky strain tensor [24], and is used for modeling very large deformations. [25]. It is considered as the natural strain tensor for many materials, but its use was hampered for a long time because of its computational complexity [26].

For registration, the basic idea is to replace the elastic energy with a regularization that measures the amount of logarithmic strain by taking a Riemannian distance between Σ and Id. With a log-Euclidean metric, this give the *log-Euclidean Riemannian elasticity*:

$$Reg_{LERE}(\Phi) = \frac{1}{4}\int \mathrm{dist}^2_{Log}(\Sigma, \mathrm{Id}) = \frac{1}{4}\int \mathrm{dist}^2_{Eucl}(\log(\Sigma), \log(\mathrm{Id}))^2 = \frac{1}{4}\int \|\log(\Sigma)\|^2$$

3.2 Isotropic Log-Euclidean Riemannian elasticity

The simplest metric on a logarithmic strain tensor $W = \log(\Sigma)$ is $\|W\|^2 = \mathrm{Tr}(W^2)$. More generally, any metric is given by a bilinear form $G(W_1, W_2)$ on the space of symmetric matrices, and is uniquely specified by the quadratic form $\|W\|^2 = G(W, W)$. A metric is isotropic if $\|W\|^2 = \|R\,W\,R^{\mathrm{T}}\|^2$ for any rotation R. This means that it only depends on the eigenvalues of W, or equivalently on the matrix invariants $\mathrm{Tr}(W)$, $\mathrm{Tr}(W^2)$ and $\mathrm{Tr}(W^3)$. However, as the form is quadratic in W, we are left only with $\mathrm{Tr}(W)^2$ and $\mathrm{Tr}(W^2)$ that can be weighted arbitrarily, e.g. by μ and $\lambda/2$ (with $n.\lambda > -2\mu$ where n is the dimension of the space to ensure the positive definiteness of the metric). Finally, the *isotropic log-Euclidean Riemannian elasticity (ILERE)* energy has the form:

$$Reg_{ILERE}(\Phi) = \int \frac{\mu}{4}\mathrm{Tr}\left((\log(\Sigma))^2\right) + \frac{\lambda}{8}\mathrm{Tr}(\log(\Sigma))^2$$

We retrieve the classical form of the isotropic elastic energy with Lamé coefficients, but with the logarithmic strain tensor. This form was expected as the St Venant-Kirchoff energy was also derived for isotropic materials.

3.3 Incorporating deformation statistics

In the context of inter-subject or atlas-to-image registration, we do not know a priori the deformability of the material. Moreover, we don't expect it to be isotropic nor stationary. An interesting idea is to learn the local deformability characteristics from a population of typical transformations $\Phi_i(x)$.

Based on the statistical framework presented in [22], we considered in [1] the strain tensor as a random variable in the Riemannian space of tensors. We defined the *a priori* deformability $\bar{\Sigma}(x)$ as the Riemannian mean of deformation tensors $\Sigma_i(x) = \nabla\Phi_i^{\mathrm{T}}\,\nabla\Phi_i$. A related idea was suggested directly on the Jacobian matrix of the transformation $\nabla\Phi$ in [27], but using a general matrix instead of a symmetric one raises important computational and theoretical problems. With the Log-Euclidean metric on strain tensors, the statistics are quite simple since we have a closed form for the mean value:

$$\bar{\Sigma}(x) = \exp(\bar{W}(x)) \quad \text{with} \quad \bar{W}(x) = \frac{1}{N}\sum_i \log(\Sigma_i(x))$$

This mean deformability $\bar{\Sigma}$ is not so easy to understand. If the reference image is optimally centered with respect to the data, one could expect the mean deformation to be null ($\bar{W} = 0$). However, this equation specifies $n(n+1)/2$ scalar components while there are only n free scalar components at each point of a displacement field. Thus, it seems at the first glance that $n(n-1)/2$ scalar components (e.g. off diagonal terms of $\bar{W}$) could not be prescribed to zero. More powerful tools from the singularity theory are probably necessary to definitely conclude on that point.

Going one step further, we can compute the covariance matrix of the random process $\mathrm{Cov}(\Sigma_i(x))$ at each point. Let us decompose the symmetric tensor $W = \log(\Sigma)$ into a vector $\mathrm{Vect}(W)^{\mathrm{T}} = (w_{11}, w_{22}, w_{33}, \sqrt{2}w_{12}, \sqrt{2}w_{13}, \sqrt{2}w_{23})$ that gathers all the tensor components in an orthonormal basis. In this coordinate system, we define the covariance matrix $\mathrm{Cov} = \frac{1}{N}\sum \mathrm{Vect}(W_i - \bar{W})\,\mathrm{Vect}(W_i - \bar{W})^{\mathrm{T}}$.

To adapt the metric on strain tensors to these first and second order moments of the random deformation process, a well known and simple tool is the Mahalanobis distance, so that we finally define the *statistical Log-Euclidean Riemannian elasticity (SLERE)* energy as:

$$Reg_{SLERE}(\Phi) = \frac{1}{4}\int \mu^2_{(\bar{W},Cov)}(\log(\Sigma(x))) = \frac{1}{4}\int \mathrm{Vect}(W-\bar{W})\mathrm{Cov}^{(-1)}\mathrm{Vect}(W-\bar{W})^{\mathrm{T}}$$

As we are using a Mahalanobis distance, this least-squares criterion can be seen as the log-likelihood of a Gaussian process on strain tensor fields: we are implicitly modeling the a-priori probability of the deformation. In a registration framework, this point of view is particularly interesting as it opens the way to use Bayesian estimation methods for non-linear registration.

4 Optimizing the Riemannian elasticity

To use the logarithmic elasticity energies as regularization criteria in the registration framework, we have to compute their gradient. Let us consider the isotropic Riemannian elasticity first. Thanks to the properties of the differential of the log (see appendix A), we have $\mathrm{Tr}(\partial_V \log(\Sigma)) = \mathrm{Tr}(\Sigma^{(-1)}\, V)$ and $\langle\, \partial_V \log(\Sigma) \mid W \,\rangle = \langle\, \partial_W \log(\Sigma) \mid V \,\rangle$. Thus, using $V = \partial_u \Sigma = \sum_\alpha \left(\partial_\alpha u\, \partial_\alpha \Phi^{\mathrm{T}} + \partial_\alpha \Phi\, \partial_\alpha u^{\mathrm{T}}\right)$ and $W = \log(\Sigma)$, we can write the directional derivative of the criterion:

$$\begin{aligned}
\partial_u Reg_{ILERE}(\Phi) &= \int \tfrac{\mu}{2}\,\langle\, W \mid \partial_V \log(\Sigma) \,\rangle + \tfrac{\lambda}{4}\mathrm{Tr}(W)\,\mathrm{Tr}(\partial_V \log(\Sigma)) \\
&= \int \tfrac{\mu}{2}\,\langle\, \partial_W \log(\Sigma) \mid V \,\rangle + \tfrac{\lambda}{4}\mathrm{Tr}(W)\,\mathrm{Tr}(\Sigma^{(-1)}\, V) \\
&= \textstyle\sum_\alpha \int \mu\,\langle\, \partial_W \log(\Sigma)\, \partial_\alpha \Phi \mid \partial_\alpha u \,\rangle + \tfrac{\lambda}{2}\mathrm{Tr}(W)\ \langle\, \Sigma^{(-1)}\, \partial_\alpha \Phi \mid \partial_\alpha u \,\rangle
\end{aligned}$$

Integrating by part with homogeneous Neumann boundary conditions, we end up with the gradient:

$$\nabla Reg_{IRE}(\Phi) = -\textstyle\sum_\alpha \partial_\alpha (Z\, \partial_\alpha \Phi) \quad \text{with} \quad Z = \mu\, \partial_W \log(\Sigma) + \frac{\lambda}{2}\mathrm{Tr}(W)\, \Sigma^{(-1)} \quad (1)$$

The same formula still holds for the general statistical Riemannian elasticity with $Z = \partial_X \log(\Sigma)$ where X is the symmetric matrix defined by $\mathrm{Vect}(X) =$

$\text{Cov}^{(-1)}\,\text{Vect}(\log(\Sigma) - \bar{W})$. Thus, we may write the gradient of all (St-Venant-Kirchoff, Isotropic Riemannian and Statistical Riemannian) elastic energies as:

$$\nabla Reg(\Phi) = -\textstyle\sum_\alpha \partial_\alpha (Z\, \partial_\alpha \Phi) \tag{2}$$

and only the 2nd Piola-Kirchoff tensor Z differs:

$$Z_{SVKE} = \mu(\Sigma - \text{ Id}) + \tfrac{\lambda}{2}\text{Tr}(\Sigma - \text{ Id})\ \text{Id} \tag{3}$$
$$Z_{ILERE} = \mu\, \partial_W \log(\Sigma) + \tfrac{\lambda}{2}\text{Tr}(\log(\Sigma))\ \Sigma^{(-1)} \tag{4}$$
$$Z_{SLERE} = \partial_X \log(\Sigma) \quad \text{with} \quad \text{Vect}(X) = \text{Cov}^{(-1)}\,\text{Vect}(\log(\Sigma) - \bar{W}) \tag{5}$$

4.1 Practical implementation

A simple and easily parallelisable implementation is the following. First, one computes the image of the gradient of the transformation, or more particularly the directional derivatives, for instance using finite differences. $\partial_\alpha \Phi(x) = (\Phi(x + \tau_\alpha\, e_\alpha) - \Phi(x - \tau_\alpha\, e_\alpha))/2\tau_\alpha$, where τ_α is the voxel size in the direction α. This operation is not computationally expensive, but requires to access the value of the transformation field at neighboring points, which can be time consuming due to systematic memory page faults in large images.

Then, we process these 3 vectors completely locally to compute 3 new vectors $v_\alpha = Z(\partial_\alpha \Phi)$. This operation is computationally more expensive but is memory efficient as the resulting vectors can replace the old directional derivatives. Finally, the gradient of the criterion $\nabla E = \sum_\alpha \partial_\alpha v_\alpha$ may be computed using finite differences on the resulting image. $\nabla E(x) = \sum_\alpha (v_\alpha(x + \tau_\alpha\, e_\alpha) - v_\alpha(x - \tau_\alpha\, e_\alpha))/2\tau_\alpha$. Once again, this is not computationally expensive, but it requires intensive memory accesses.

The only additional cost for the Riemannian Elasticity is the computation of the logarithm $W = \log(\Sigma)$ and its directional derivative $\partial_W \log(\Sigma)$. This would probably be prohibitive if we had to rely on numerical approximation methods. Fortunately, we were able to compute an explicit and very simple and efficient closed-form expression that only requires the diagonalization of Σ (see appendix A). Experiments performed in [1] showed that optimizing the isotropic Riemannian elasticity was only 3 time longer than optimizing the standard elasticity.

5 Left Invariant Riemannian Elasticity

Let us now investigate the invariance properties in view of relating the Riemannian elasticity to metrics on diffeomorphisms. Since $\nabla(\Phi^{(-1)}) \circ \Phi = (\nabla\Phi)^{(-1)}$, the isotropic logarithmic distance of a strain tensor to the identity is locally inverse-consistent. We have indeed $\text{Tr}\left(\log(\Sigma_\Phi)^2\right) = \text{Tr}\left(\log(\Sigma_{\Phi^{(-1)}} \circ \Phi)^2\right)$ and $\text{Tr}\left(\log(\Sigma_\Phi)\right) = \text{Tr}\left(\log(\Sigma_{\Phi^{(-1)}} \circ \Phi)\right)$. This means that, locally, a scaling of a factor 2 at the same distance from the identity than a scaling of 0.5. However, this property does not hold globally due to the change of the volume element during the change of variable $y = \Phi(x)$:

$$\begin{aligned} Reg_{IRE}(\Phi^{(-1)}) &= \int \tfrac{\mu}{4} \mathrm{Tr}\left((\log(\Sigma_{\Phi^{(-1)}}(y)))^2 \right) + \tfrac{\lambda}{8} \mathrm{Tr}(\log(\Sigma_{\Phi^{(-1)}}(y)))^2 .dy \\ &= \int \tfrac{\mu}{4} \mathrm{Tr}\left((\log(\Sigma_{\Phi}(x)))^2 \right) + \tfrac{\lambda}{8} \mathrm{Tr}(\log(\Sigma_{\Phi}(x)))^2 . \sqrt{|\Sigma_{\Phi}(x)|} .dx \end{aligned}$$

5.1 Inverse Consistent Riemannian Elasticity

Following an idea suggested in [28], we can integrate with a volume element which is the geometric mean between the one in the original space and the one in the arrival space, i.e.: $\sqrt{|\nabla \Phi(x)|}.dx = |\Sigma(x)|^{1/4}.dx$. If f is a locally inverse consistent functional (i.e. such that $f(\Phi^{(-1)}) \circ \Phi = f(\Phi)$), then the integral value $F(\Phi) = \int f(\Phi).\sqrt{|\nabla \Phi|}$ is also inverse consistent. Indeed, the change of variable $y = \Phi(x)$ induces $dy = |\nabla \Phi(x)|.dx$, but since $|\nabla(\Phi^{(-1)}) \circ \Phi| = |\nabla \Phi|^{(-1)}$, we have:

$$F(\Phi^{(-1)}) = \int f(\Phi^{(-1)})(y).\sqrt{|\nabla(\Phi^{(-1)})(y)|}.dy = \int f(\Phi)(x).\sqrt{|\nabla(\Phi)(x)|}.dx = F(\Phi)$$

As the log-Euclidean distance of a strain tensor to the identity is locally inverse consistent, we thus obtain a globally inverse consistent (isotropic) Riemannian elasticity with:

$$Reg_{ICRE}(\Phi) = \int \left\{ \frac{\mu}{4} \mathrm{Tr}\left(\log(\Sigma)^2 \right) + \frac{\lambda}{8} \mathrm{Tr}(\log(\Sigma))^2 \right\} .|\Sigma|^{1/4}.dx$$

Another formulation may be obtained using the change of variable $y = \Phi(x)$ and will turn out to be generalizable to a left-invariant energy:

$$Reg_{ICRE}(\Phi) = \int \left\| \log(\Sigma \circ \Phi^{(-1)}) \right\|^2 .|\Sigma \circ \Phi^{(-1)}|^{-1/4} \tag{6}$$

In this formula, the norm $\|.\|$ refers to an isotropic norm on symmetric matrices.

The derivative of this new criterion can be deduced from $\partial_u Reg_{IRE}$ using:

$$\partial_u \det(\Sigma)^{1/4} = \tfrac{1}{4} \mathrm{Tr}(\Sigma^{(-1)}.\partial_u \Sigma). \det(\Sigma)^{1/4} = \tfrac{1}{2} \textstyle\sum_\alpha \langle\, \Sigma^{(-1)}.\partial_\alpha \Phi \mid \partial_\alpha u \rangle \det(\Sigma)^{1/4}$$

We have one again $\partial_u Reg_{ICRE}(\Phi) = \sum_\alpha \int \langle\, Z.\partial_\alpha \Phi \mid \partial_\alpha u \rangle$, with

$$Z_{ICRE} = \left(Z_{IRE} + \frac{1}{2} \|W\|^2 .\Sigma^{(-1)} \right) \det(\Sigma)^{1/4}$$

Thus, we have obtained an inverse invariant energy on diffeomorphisms which allows us to optimize directly their regularity in registration processes without having to integrate numerically along the transformation trajectory for computing the length of geodesics, as for the invariant metrics on diffeomorphisms proposed in [15, 16].

5.2 Left-invariant Riemannian elasticity

This energy is positive and null only is the transformation is locally rigid everywhere. It can be turned into a left- (or right-) invariant "distance" by left- (resp. right) translation. Let us investigate the left-invariant "distance" (The right-invariant distance is automatically given by $\operatorname{dist}_R(\Phi, \Psi) = \operatorname{dist}_L(\Phi^{(-1)}, \Psi^{(-1)})$):

$$\operatorname{dist}^2_L(\Phi, \Psi) = Reg_{ICRE}(\Phi^{(-1)} \circ \Psi) = \int \|\log(\Sigma_{\Phi^{(-1)} \circ \Psi})\|^2 . |\Phi^{(-1)} \circ \Psi|^{1/4}$$

Thanks to the inverse invariance, the "distance" is symmetric. It is null if and only if the two diffeomorphisms differ by a local rotation everywhere. However, to show that this is really a left-invariant distance on diffeomorphisms of rigid shapes, the triangular inequality remains to be established. Moreover, we suspect that we obtain an extrinsic distance and not a Riemannian one.

The expression of the left-invariant distance can be worked out to see how much it differs from the previously proposed statistical Riemannian elasticity. We first notice that $\nabla(\Phi^{(-1)} \circ \Psi) = \nabla\Psi . \nabla(\Phi^{(-1)}) \circ \Psi = \nabla\Psi . (\nabla\Phi)^{(-1)} \circ (\Phi^{(-1)} \circ \Psi)$. Using the singular value decomposition $\nabla\Phi = U.S.V^{\mathrm{T}}$, there exists a rotation $R = V.U^{\mathrm{T}}$ at each point such that $R.\nabla\Phi = \Sigma_\Phi^{1/2}$. Thus, we have:

$$\Sigma_{\Phi^{(-1)} \circ \Psi} = R^{\mathrm{T}} . \left(\Sigma_\Phi^{-1/2} \circ (\Phi^{(-1)} \circ \Psi) \right) . \Sigma_\Psi . \left(\Sigma_\Phi^{-1/2} \circ (\Phi^{(-1)} \circ \Psi) \right) . R$$

But thanks to $\log(R^{\mathrm{T}}.\Sigma.R) = R^{\mathrm{T}}.\log(\Sigma).R$ and to the isotropy of the norm on symmetric matrices, the rotation R disappear in the distance. Finally, using the change of variable $y = \Psi(x)$, we end up with

$$\operatorname{dist}^2_L(\Phi, \Psi) = \int \left\| \log \left((\Sigma_\Phi^{-1/2} \circ \Phi^{(-1)}) . (\Sigma_\Psi \circ \Psi^{(-1)}) . (\Sigma_\Phi^{-1/2} \circ \Phi^{(-1)}) \right) \right\|^2 . \\ \det(\Sigma_\Psi \circ \Psi^{(-1)})^{-1/4} . \det(\Sigma_\Phi \circ \Phi^{(-1)})^{-1/4}$$

Besides symmetric corrections for the volume element, one recognize here the affine-invariant distance on symmetric matrices instead of the log-Euclidean one as we originally proposed for the statistical Riemannian elasticity. Using the resampled tensor fields $\hat{\Sigma}_\phi = \Sigma_\Phi \circ \Phi^{(-1)}$ and $\hat{\Sigma}_\psi = \Sigma_\Psi \circ \Psi^{(-1)}$, we finally obtain:

$$\operatorname{dist}^2_L(\Phi, \Psi) = \int \operatorname{dist}^2_{Aff} \left(\hat{\Sigma}_\Phi \, , \, \hat{\Sigma}_\Psi \right) . \det(\hat{\Sigma}_\Psi)^{-1/4} . \det(\hat{\Sigma}_\Phi)^{-1/4} \tag{7}$$

Other simple formulations of the left (and of the right) invariant "distance" are possible, and we are currently analyzing them to find out the more intuitive ones. Following the statistical framework of [29], computing the derivatives will allow determining the barycentric equation of the Fréchet "mean diffeomorphisms" according to these "metrics", and a gradient descent algorithm to obtain them. Then, we hope to be able to compute second order moment and to define a kind of Mahalanobis distance (including local anisotropy and non-stationarity) on shape diffeomorphisms.

6 Discussion

Riemannian elasticity is an integrated framework to compute the statistics on deformations and re-introduce them as constraints in non-linear registration algorithms. This framework is based on the interpretation of the elastic energy as a Euclidean distance between the Cauchy-Green strain tensor and the identity (i.e. the local rigidity). By providing the space of tensors with a more suitable Riemannian metric, for instance a Log-Euclidean one, we can define proper statistics on deformations, like the mean and the covariance matrix. Taking these measurements into account in a statistical (i.e. a Mahalanobis) distance, we end-up with the statistical Riemannian elasticity regularization criterion. This criterion can also be viewed as the log-likelihood of the deformation probability, which opens the way to Bayesian deformable image registration algorithms.

We investigated in this paper the theoretical properties of the isotropic and stationary version and we showed that it was possible to obtain an inverse-consistent criterion by modifying the spatial integration measure. It is remarkable that this allows to define a left or right invariant energy between two diffeomorphisms without having to optimize for the geodesic path between them. However, many questions are left open. For instance, it remains to be established that our energy is a distance, and if it is Riemannian or extrinsic. Determining the geodesics (if they exist) would also be very interesting to better understand the properties of these energies. This would probably help also in generalizing the statistical Riemannian elasticity in a consistent way, in order to measure and take into account anisotropic and non-stationary behavior of the deformations. On a more theoretical point of view, it would be interesting to make the link between our approach and the Brownian warps of [28, 30].

References

1. X. Pennec, R. Stefanescu, V. Arsigny, P. Fillard, and N. Ayache. Riemannian elasticity: A statistical regularization framework for non-linear registration. In *Proc. of MICCAI 2005, Part II*, volume 3750 of *LNCS*, pages 943–950, Palm Springs, CA, USA, October 26-29, 2005. Springer Verlag.
2. R. Bajcsy and S. Kovačič. Multiresolution elastic matching. *Computer Vision, Graphics and Image Processing*, 46:1–21, 1989.
3. G. E. Christensen, S. C. Joshi, and M. I. Miller. Volumetric transformation of brain anatomy. *IEEE Trans. on Medical Imaging*, 16(6):864–877, December 1997.
4. J.-P. Thirion. Image matching as a diffusion process: an analogy with maxwell's demons. *Medical Image Analysis*, 2(3), 1998.
5. X. Pennec, P. Cachier, and N. Ayache. Understanding the "demon's algorithm": 3D non-rigid registration by gradient descent. In *Proc. of MICCAI'99*, volume 1679 of *LNCS*, pages 597–605, Cambridge, UK, September 1999. Springer Verlag.
6. J. Modersitzki. *Numerical Methods for Image Registration*. Numerical Mathematics and Scientific Computations. Oxford University Press, 2004.
7. P. Cachier and N. Ayache. Isotropic energies, filters and splines for vectorial regularization. *J. of Math. Imaging and Vision*, 20(3):251–265, May 2004.

8. H. Lester, S.R. Arridge, K.M. Jansons, L. Lemieux, J.V. Hajnal, and A. Oatridge. Non-linear registration with the variable viscosity fluid algorithm. In *Proc of IPMI;99*, pages 238–251, 1999.
9. R. Stefanescu, X. Pennec, and N. Ayache. Grid powered nonlinear image registration with locally adaptive regularization. *Medical Image Analysis*, 8(3):325–342, September 2004.
10. P.M. Thompson, M.S. Mega, K.L. Narr, E.R. Sowell, R.E. Blanton, and A.W. Toga. Brain image analysis and atlas construction. In M. Fitzpatrick and M. Sonka, editors, *Handbook of Medical Image Proc. and Analysis*, chapter 17. SPIE, 2000.
11. D. Rueckert, A.F. Frangi, and J.A. Schnabel. Automatic construction of 3D statistical deformation models of the brain using non-rigid registration. *IEEE TMI*, 22:1014–1025, 2003.
12. P. Fillard, V. Arsigny, X. Pennec, P. Thompson, and N. Ayache. Extrapolation of sparse tensor fields: Application to the modeling of brain variability. In *Proc. of IPMI'05*, volume 3565 of *LNCS*, pages 27–38, Glenwood springs, Colorado, USA, July 2005. Springer.
13. J. Ashburner, J.L.R. Andersson, and K.J. Friston. High-dimensional image registration using symmetric priors. *NeuroImage*, 9:619–628, 1999.
14. P.M. Thompson, C. Schwartz, R. T. Lin, A. A. Khan, and A.W. Toga. 3D statistical analysis of sulcal variability in the human brain. *J. Neurosci*, 16(13):4261–4274, 1996.
15. S.C. Joshi and M.I Miller. Landmark matching via large deformation diffeomorphisms. *IEEE Transactions on Image Processing*, 9(8):1357–1370, 2000.
16. M.F. Beg, M.I. Miller, A. Trouvé, and L. Younes. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *Int. Journal of Computer Vision*, 61(2):139–157, 2005.
17. P.G. Ciarlet. *Mathematical elasticity Vol. 1: Three-dimensionnal elasticity.* Elsevier Science B.V., 1988.
18. GE Christensen and HJ. Johnson. Consistent image registration. *IEEE Trans Med Imaging*, 20(7):568–82, July 2001.
19. P. Batchelor, M. Moakher, D. Atkinson, F. Calamante, and A. Connelly. A rigorous framework for diffusion tensor calculus. *Mag. Res. in Med.*, 53:221–225, 2005.
20. P.T. Fletcher and S.C. Joshi. Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors. In *Proc. of CVAMIA and MMBIA Workshops, Prague, Czech Republic, May 15, 2004*, LNCS 3117, pages 87–98. Springer, 2004.
21. Ch. Lenglet, M. Rousson, R. Deriche, and O. Faugeras. Statistics on the manifold of multivariate normal distributions: Theory and application to diffusion tensor MRI processingy. *Journal of Mathematical Imaging and Vision*, 2006. To appear. A preliminary version is available as INRIA Research Report RR-5242, 2004.
22. X. Pennec, P. Fillard, and N. Ayache. A Riemannian framework for tensor computing. *International Journal of Computer Vision*, 66(1):41–66, January 2006.
23. V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Fast and simple calculus on tensors in the log-Euclidean framework. In *Proc. of MICCAI 2005, Part I*, volume 3749 of *LNCS*, pages 115–122, Palm Springs, CA, USA, October 26-29, 2005. Springer Verlag.
24. H. Hencky. Uber die form des elastizitatsgesetzes bei ideal elastischen stoffen. *Zeit. Techn. Phys.*, 9:241–247, 1928.
25. P. Roué. *Mécanique des grandes transformations.* Number 25 in Mathématiques & Applications. Springer, 1997.
26. A.D. Freed. Natural strain. *Journal of Engineering Materials & Technology*, 117:379–385, 1995.

27. R.P. Woods. Characterizing volume and surface deformations in an atlas framework: theory, applications, and implementation. *NeuroImage*, 18(3):769–788, 2003.
28. M. Nielsen, P. Johansen, A.D. Jackson, and B. Lautrup. Brownian warps: A least committed prior for non-rigid registration. In *Proc of MICCAI'02, Part II*, volume 2489 of *LNCS*, pages 557 – 564, 2002.
29. Xavier Pennec. Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements. *Journal of Mathematical Imaging and Vision*, 2006. To appear (avalailable at dx.doi.org/10.1007/s10851-006-6228-4).
30. B. Markussen. A statistical approach to large deformation diffeomorphisms. *2004 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'04)*, 12:181, 2004.

A Appendix: tensor derivatives

A.1 Exponential of a tensor

Let $W = R\,S\,R^{\mathrm{T}}$ be a diagonalization of a symmetric matrix. We can write any power of W in the same basis: $W^k = R\,S^k\,R^{\mathrm{T}}$. Thus, the rotation matrices can be factored out in the series defining the matrix exponential, so that the exponential is applied directly to the eigenvalues:

$$\exp(W) = \sum_{k=0}^{+\infty} \frac{W^k}{k!} = R\,\mathrm{DIAG}(\exp(s_i))\,R^{\mathrm{T}}$$

This series converges for any symmetric matrix argument, and it is easy to see that its inverse is well defined for any positive definite symmetric matrix $\Sigma = R\,\mathrm{DIAG}(a_i)\,R^{\mathrm{T}}$. This is the function: $\log(\Sigma) = R\,(\mathrm{DIAG}(\log(a_i)))\,R^{\mathrm{T}}$. It is important to notice that there is no series expansion which is converging for all arguments, like for the exponential.

A.2 Differential of the exponential

The matrix exponential and logarithm realize a one-to-one mapping between the space of symmetric matrices to the the space of tensors. Moreover, one can show that this mapping is diffeomorphic, since the differential has no singularities. Using the Taylor expansion $(W + \varepsilon V)^k = W^k + \varepsilon \sum_{i=0}^{k-1} W^i\,V\,W^{k-i-1} + O(\varepsilon^2)$ for $k \geq 1$, we obtain by identification the directional derivative $\partial_V \exp(W)$ by gathering the first order terms in ε in the series $\exp(W + \varepsilon V) = \sum_{k=0}^{+\infty} (W + \varepsilon V)^k / k!$:

$$\partial_V \exp(W) = (d\exp(W))(V) = \sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{i=0}^{k-1} W^i\,V\,W^{k-i-1} \tag{8}$$

For simplifying the differential, we can see that using the diagonalization $W = RS\,R^{\mathrm{T}}$ in the series gives:

$$\partial_V \exp(W) = R\,\partial_{(R^{\mathrm{T}}\,V\,R)} \exp(S)\,R^{\mathrm{T}}$$

Thus, we are left with the computation of $\partial_V \exp(S)$ for S diagonal. As $[S^l V S^{k-l-1}]_{ij} = s_i^l \, v_{ij} \, s_j^{k-l-1}$, we have $[\partial_V \exp(S)]_{ij} = \left\{ \sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{l=0}^{k-1} s_i^l \, s_j^{k-l-1} \right\} v_{ij} = q_{ij} \, v_{ij}$ with

$$
\begin{aligned}
q_{ij} &= \sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{l=0}^{k-1} s_i^l \, s_j^{k-l-1} = \sum_{k=1}^{+\infty} \frac{1}{k!} \frac{s_i^k - s_j^k}{s_i - s_j} = \frac{\exp(s_i) - \exp(s_j)}{s_i - s_j} \\
&= \exp(s_j) \left(1 + \frac{(s_i - s_j)}{2} + \frac{(s_i - s_j)^2}{6} + O\left((s_i - s_j)^3 \right) \right)
\end{aligned}
$$

The last Taylor expansion shows that this formula is computationally well posed. Moreover, we have $q_{ij} \geq 1 > 0$, so that we can conclude that $d \exp(S)$ is a diagonal linear form that is always invertible: the exponential is a diffeomorphism.

A.3 Differential of the logarithm

To compute the differential of the logarithm function, we do not have a series that we could perturb like for the exponential, but we can simply inverse the differential of the exponential as a linear form: as $\exp(\log(\Sigma)) = \Sigma$, we have $(d \log(\Sigma)(V) = (d \exp(\log(\Sigma))^{(-1)} V$. Using $D = \exp(S)$, the inverse is easily expressed for a diagonal matrix: $[(d \exp(S))^{(-1)} \, V]_{ij} = v_{ij}/q_{ij}$. Thus we have:

$$
[\partial_V \log(D)]_{ij} = v_{ij} \frac{\log(d_i) - \log(d_j)}{d_i - d_j}
$$

Notice that

$$
q_{ij}^{(-1)} = \frac{\log(d_i) - \log(d_j)}{d_i - d_j} = \frac{1}{d_j} \left(1 - \frac{d_i - d_j}{2 \, d_j} + \frac{(d_i - d_j)^2}{3 \, d_j^2} + O\left((d_i - d_j)^3 \right) \right)
$$

so that the formula is numerically stable. Finally, using the identity $\log(\Sigma) = R^{\mathrm{T}} \, \log(R \, \Sigma \, R^{\mathrm{T}}) \, R$ for any rotation R, we have:

$$
\partial_V \log(R \, D \, R^{\mathrm{T}}) = R \, \left(\partial_{R^{\mathrm{T}} \, V \, R} \log(D) \, \right) R^{\mathrm{T}}
$$

That way, we may compute the differential at any point $\Sigma = R \, D \, R^{\mathrm{T}}$.

A.4 Remarkable identities

$$
\partial_{\log(\Sigma)} \log(\Sigma) = \Sigma^{(-1)} \, \log(\Sigma) = \log(\Sigma) \, \Sigma^{(-1)} \tag{9}
$$

$$
\langle \, \partial_V \log(\Sigma) \mid W \, \rangle = \langle \, \partial_W \log(\Sigma) \mid V \, \rangle \tag{10}
$$

$$
\partial_{\log(\Sigma)} \log(\Sigma) = R \, \left(\partial_{\log(D)} \log(D) \right) R^{\mathrm{T}} = R \, \mathrm{Diag}(\log(d_i)/d_i) \, R^{\mathrm{T}} = \Sigma^{(-1)} \, \log(\Sigma)
$$

$$
\begin{aligned}
\langle \, \partial_V \log(\Sigma) \mid R \, D \, R^{\mathrm{T}} \, \rangle &= \mathrm{Tr}\left((\partial_{R^{\mathrm{T}} \, V \, R} \log(D) \,) \, R^{\mathrm{T}} \, W \, R \right) \\
&= [R^{\mathrm{T}} \, V \, R]_{ij} \, \frac{\log(d_i) - \log(d_j)}{d_i - d_j} \, [R^{\mathrm{T}} \, W \, R]_{ij} \\
&= \mathrm{Tr}\left((\partial_{R^{\mathrm{T}} \, W \, R} \log(D) \,) \, R^{\mathrm{T}} \, V \, R \right) = \langle \, \partial_W \log(\Sigma) \mid V \, \rangle
\end{aligned}
$$