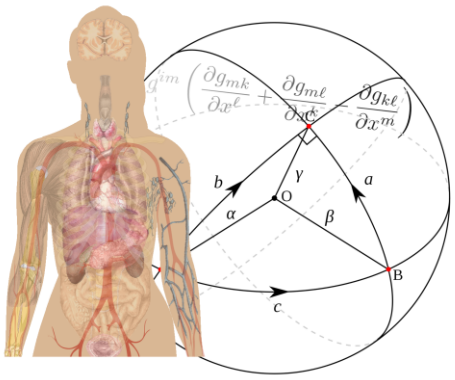


MFCA 2013

4th MICCAI workshop on

Mathematical Foundations of of Computational Anatomy

Mathematical Foundations of Computational Anatomy

Geometrical and Statistical Methods for Biological Shape Variability Modeling

September 22th, 2013, Nagoya, Japan.

http://www.inria.fr/sophia/asclepios/events/MFCA113
http://hal.inria.fr/MFCA/

Editors:
Xavier Pennec (Asclepios, INRIA Sophia-Antipolis, France)
Sarang Joshi (SCI, University of Utah, USA)
Mads Nielsen (University of Copenhagen, DK)
Tom Fletcher (SCI, University of Utah, USA)
Stanley Durrleman (ARAMIS, Inria / ICM, Paris, FR)
Stefan Sommer (University of Copenhagen, DK)

informatics mathematics

MICCAI

Preface

Computational anatomy is an emerging discipline at the interface of geometry, statistics and image analysis which aims at modeling and analyzing the biological shape of tissues and organs. The goal is to estimate representative organ anatomies across diseases, populations, species or ages, to model the organ development across time (growth or aging), to establish their variability, and to correlate this variability information with other functional, genetic or structural information.

The **Mathematical Foundations of Computational Anatomy (MFCA)** workshop aims at fostering the interactions between the mathematical community around shapes and the MICCAI community in view of computational anatomy applications. It targets more particularly researchers investigating the combination of statistical and geometrical aspects in the modeling of the variability of biological shapes. The workshop is a forum for the exchange of the theoretical ideas and aims at being a source of inspiration for new methodological developments in computational anatomy. A special emphasis is put on theoretical developments, applications and results being welcomed as illustrations. Following the first edition of this workshop in 2006[1], second edition in New-York in 2008[2], the third edition in Toronto in 2011[3], the forth edition was held in Nagoya Japan on September 22 2013[4].

Contributions were solicited in Riemannian and group theoretical methods, advanced statistics on deformations and shapes, metrics for computational anatomy, statistics of surfaces, time-evolving geometric processes, stratified spaces, and related subjects. 12 papers were selected and organized in 4 sessions.

August 2011

Xavier Pennec
General Chair
MFCA'13

[1] http://www.inria.fr/sophia/asclepios/events/MFCA06/
[2] http://www.inria.fr/sophia/asclepios/events/MFCA08/
[3] http://www.inria.fr/sophia/asclepios/events/MFCA11/
[4] http://www.inria.fr/sophia/asclepios/events/MFCA13/

Organization

Workshop Chairs

Xavier Pennec	(INRIA Sophia-Antipolis, France)
Sarang Joshi	(SCI, University of Utah, USA)
Mads Nielsen	(University of Copenhagen, Denmark)
Tom Fletcher	(SCI, University of Utah, USA)
Stanley Durrleman	(ARAMIS, Inria / ICM, Paris, FR)
Stefan Sommer	(University of Copenhagen, DK)

Program Committee

Stéphanie Allassoniere	(Ecole Polytechnique, France)
Rachid Deriche	(INRIA, France)
Ian L. Dryden	(University of Nottingham, UK)
Aasa Feragen	(University of Copenhagen, DK)
Luc Florac	(Eindhoven U. of Technology, NL)
Ghassan Hamarneh	(Simon Fraser U., CA)
Darryl Holm	(Imperial College London, UK)
Susan Holmes	(Stanford U., USA)
Steve Marron	(UNC Chapel Hill, USA)
Stephen Marsland	(Massey University, NZ)
Yoshitaka Masutani	(U. of Tokyo Hosp, JP)
Marc Niethammer	(UNC Chapel Hill, USA)
Salvador Olmos	(U. of Saragossa, SP)
Bruno Pelletier	(U. Rennes, FR)
Jerry Prince	(Johns Hopkins U., USA)
Daniel Rueckert	(Imperial Coll. London, UK)
Kaleem Siddiqi	(McGill U., CA)
Anuj Srivastava	(Florida State U., USA)
Martin Styner	(UNC Chapel Hill, USA)
Alain Trouv	(ENS Cachan, FR)
Carole Twining	(U. of Manchester, UK)
Baba Vemuri	(U. of Florida, USA)
Francois Xavier Vialard	(Dauphine U., FR)
William M. Wells III	(MIT & Harvard, USA)

Table of Contents

LDDMM and scale

Brain Morphometry

Shape and Image Registration

Poster session

Riemannian metrics for statistics on shapes: Parallel transport and scale invariance

Marc Niethammer[1] and François-Xavier Vialard[2]

[1] UNC Chapel Hill
mn@cs.unc.edu
[2] Université Paris-Dauphine, Ceremade UMR CNRS 7534
vialard@ceremade.dauphine.fr

Abstract. To be able to statistically compare evolutions of image time-series data requires a method to express these evolutions in a common coordinate system. This requires a mechanism to transport evolutions between coordinate systems: e.g., parallel transport has been used for large-displacement diffeomorphic metric mapping (LDDMM) approaches. A common purpose to study evolutions is to assess local tissue growth or decay as observed in the context of neurodevelopment or neurodegeneration. Hence, preserving this information under transport is important to allow for faithful statistical analysis in the common coordinate system. Most basically, we require scale invariance. Here, we show that a scale invariant metric does not exist in the LDDMM setting. We illustrate the impact of this non-invariance on parallel transport. We also propose a new class of Riemannian metrics on shapes which preserves the variation of a global indicator such as volume under parallel transport.

Keywords: parallel transport, scale invariance, Riemannian metrics on shapes

1 Introduction

Classical image registration deals with the spatial alignment of pairs of images. It is one of the most fundamental problems in medical image analysis. In particular, for population-studies image registration is an indispensable tool, as it allows to align image information to a common coordinate system for localized comparisons. Recently, studies for example on Alzheimer's disease (ADNI), osteoarthritis (OAI), and brain development (NIHPD) have acquired large volumes of longitudinal imaging data. However, computational methods to adequately analyze such longitudinal data are still in their infancy. Analyzing longitudinal image data is challenging: not only is a method for spatial alignment to a common coordinate system required, but also the temporal aspect of a longitudinal image change needs to be expressed in this common coordinate system. Arguably, the theoretically most advanced existing methods to address these problems have been methods grounded in the theory of large-displacement-diffeomorphic metric mapping (LDDMM) [2]. Indeed, LDDMM provides a convenient Riemannian

setting [15] for image registration. Other Riemannian metrics have been developed in the past years [16, 11, 12] sometimes due to the simple calculation of geodesics. Thus, tools from Riemannian geometry can be used to perform statistics on shape deformations [14, 9, 6]. In particular, parallel transport under the Levi-Civita connection gives a method to transport small longitudinal evolutions between two different images. The use of parallel transport in computational anatomy has been introduced in [17], including a numerical method based on Jacobi fields for its computation. Other numerical methods for parallel transport have been successfully developed in [4]. Note that parallel transport is path-dependent. For shapes it has a strong relation to shape correspondence [13] making it a promising candidate to transport longitudinal information. Alternatively, the adjoint [7] and the co-adjoint [3] actions on the tangent space have been proposed to transport tangent information.

In this paper, we show that these methods preserve properties which may be undesirable for computational anatomy. Further, we explore the design of a Riemannian metric conserving quantities such as absolute or relative volume variation. As a case in point, consider Alzheimer's disease where the decay of the hippocampus is an important biomarker. Hence, preserving the relative volume variation when transporting longitudinal change to an analysis space is desirable. However, only specific metrics result in such a parallel transport – volume variations will be distorted under parallel transport using an "unsuitable" metric. In particular, this is the case for LDDMM, which is not scale-invariant.

Sec. 2 illustrates shortcomings of parallel and co-adjoint transport for LDDMM. This is not a shortcoming of a *particular* LDDMM metric, but holds for all as a non-degenerate scale-invariant metric does not exist for LDDMM (see Sec. 3). Consequentially, we introduce a new model decomposing volume and shape variation in Sec. 4 as an example of a Riemannian metric addressing some of the shortcomings of LDDMM. Sec. 5 illustrates behavioral differences between LDDMM and the shape/volume-decomposed model. The paper concludes with a summary of results and an outlook on future work in Sec. 6.

2 Motivating examples

To illustrate the behavior of different types of transport under the LDDMM model we consider a uniformly expanding or contracting n-sphere of radius r, S_n, with uniformly-distributed momentum. Due to the spherical symmetry this allows us to explicitly compute expressions for co-adjoint and parallel transport. Specifically, we define the momentum at radius 1 as $m_1 = c\frac{x}{\|x\|}\delta\{\|x\| - 1\}$, where $c \in \mathbb{R}$ is a given constant and $\delta\{x\}$ denotes the Dirac delta function. Uniform scaling to radius r is described by the map $\Phi^{-1}(x) = \frac{1}{r}x$, which is in the coordinate system of the sphere of radius r. We note that the local volume-change of the n-sphere, $|D\Phi^{-1}|$ with respect to the unit-sphere is given by

$$|D\Phi^{-1}| = \frac{vol(S_n(r))}{vol(S_n(1))} = \left(\frac{1}{r}\right)^d, \tag{1}$$

where d is the space dimension and $vol(S_n(r))$ denotes the volume of $S_n(r)$.

Co-adjoint and parallel transport respectively preserve the dual pairing and the riemannian norm, so that due to the symmetry of momentums and spheres, this property completely determines the transported momentums in both cases. Sec. 2.1 derives the co-adjoint and Sec. 2.2 the parallel transport for this sphere under contraction and expansion. Sec. 2.3 demonstrates their differences.

2.1 Co-adjoint transport

In what follows, n will denote the unit normal to the sphere and v is vector field defined on the whole domain.

Definition: We can define the co-adjoint transport to the momentum-velocity pairing $\langle m_1, v\rangle := \int_{S_n(1)} \langle n, v(x)\rangle \; dS(x)$. Then (with $g^{-1} = \Phi$)

$$\langle Ad_{g^{-1}}{}^*(m_1), v\rangle = \int_{S_n(1)} \langle n, r^{-1} v(rx)\rangle \; dS(x) = \int_{S_n(r)} \langle n, v(y)\rangle r^{-d} \; dS(y). \quad (2)$$

Hence, the co-adjoint transport of the momentum is given by

$$m_r = m_1 \circ \Phi^{-1} |D\Phi^{-1}| = c \frac{\frac{1}{r}x}{\|\frac{1}{r}x\|} \delta\{\|\frac{1}{r}x\| - 1\} \left(\frac{1}{r}\right)^d = c \frac{x}{\|x\|} \delta\{\|x\| - r\} \left(\frac{1}{r}\right)^d.$$

Velocity computation: The velocity is the momentum convolved with kernel K:

$$v_r(x) = K * m_r(x) = \left(\frac{1}{r}\right)^d c \int_{S_n(r)} K(x-y) \frac{y}{\|y\|} \; dS(y). \quad (3)$$

Since we assume a perfectly symmetric distribution it is sufficient to evaluate the velocity at one location on the circle, e.g., at re_1, i.e., we need to compute

$$\tilde{v}_r := v_r(re_1) \cdot e_1 = \left(\frac{1}{r}\right)^d c \underbrace{\int_{S_n(r)} K(re_1 - y) y_1 \; dS}_{:= q(r)} = \frac{c q(r)}{r^d}, \quad (4)$$

where we made use of the fact that $v_r(re_1)$ will only have a velocity component in the e_1 direction due to symmetry and e_1 is the first canonical unit vector.

Note that we computed co-adjoint transport with respect to the uniform scaling map Φ, which is not an element of the diffeomorphisms group since it does not converge to the identity at infinity. Another natural map that could be used for co-adjoint transport is the optimal diffeomorphism obtained by solving the LDDMM functional, but this reduces to parallel transport developed in the next section.

2.2 Parallel transport

The geodesic between the sphere of radius 1 and radius r will possess the same symmetry in its shape evolution. The momentum will also be radial and constant on the sphere. It can also be checked that the parallel transport (along the geodesic) of m_1 to radius r will be λm_r for a real λ that we have to determine. To this end, we use the conservation of the norm under parallel transport, i.e.,

$$\int_{S_n(1)} \langle v_1(x), m_1(x) \rangle \, dS = \int_{S_n(r)} \langle v_r(x), m_r(x) \rangle \, dS. \tag{5}$$

In the e_1 direction we can then write $\tilde{v}_1 \tilde{m}_1 = r^{d-1} \tilde{v}_r \tilde{m}_r$, where $\tilde{m}_r := m_r(re_1) \cdot e_1$. But according to our assumption: $\tilde{m}_r = \lambda \tilde{m}_1$ and we obtain $\tilde{v}_1 = r^{d-1} \lambda \tilde{v}_r$, and $\tilde{v}_r = \lambda c q(r)$, which yields

$$|\lambda| = \sqrt{\frac{\tilde{v}_1}{r^{d-1} c q(r)}} \text{ and finally } \tilde{v}_r = \lambda c q(r) = sign(c) \sqrt{\frac{\tilde{v}_1 c q(r)}{r^{d-1}}} \, . \tag{6}$$

We note that this is up to a multiplicative constant, the square root of the co-adjoint transport. Hence, we can expect a drastically different behavior for the two types of transport.

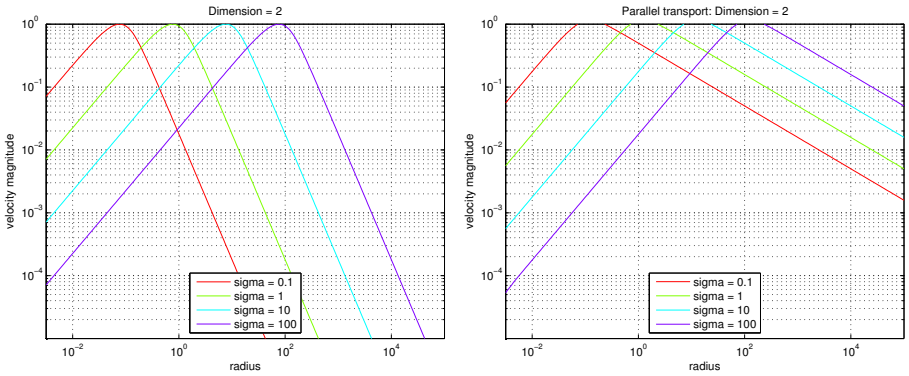
2.3 Simulations

We solve the equations for parallel (6) and co-adjoint (4) transport numerically for different kernels and different radii. In particular, we use an isotropic Gaussian kernel of the form $K(x) = c_g e^{-\frac{x^T x}{2\sigma^2}}$, where the normalization constant c_g can be subsumed into c (and will therefore be disregarded in what follows). In our experiments we computed all integrations in polar coordinates (2D) and spherical coordinates (3D) respectively. Fig. 1 shows numerical results for S_1 (i.e., the two-dimensional case) for the scaling map. We observe the following:

1) The velocity is kernel-size dependent.
2) Depending on the relation of the size of the object to the kernel-size, velocity may either increase or decrease with increased radius.
3) The radius for which the maximal velocity is obtained roughly coincides with the standard deviation of the Gaussian kernel.
4) The velocity versus radius plots are asymmetric.
5) Velocities converge to zero as $r \to 0^+$ and $r \to \infty$.
6) Parallel transport and co-adjoint transport show similar trends however with different asymptotes for the velocity.

The same conclusions hold in the 3D case, albeit with different slopes than in 2D (figures not shown).

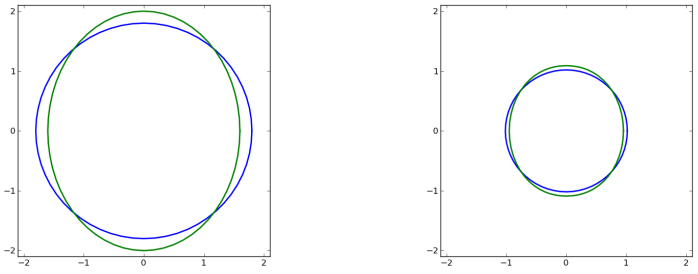
To illustrate the effect of parallel transport on shapes (represented as a group of points) we compute the geodesic evolution between a circle and an ellipse with small anisotropy using LDDMM. The resulting initial momentum is then parallel



Co-adjoint transport Parallel transport

Fig. 1. Co-adjoint and parallel transport for a 2D circular example for varying kernel sizes. Double-logarithmic plot (velocity magnitude over radius). A clear dependence of the velocity on the kernel size is observed. Maximal velocity normalized to 1 for comparison. Results were obtained using recursive adaptive Simpson quadrature.

transported along a geodesic mapping the initial circle to a smaller circle. Fig. 2 shows the used shapes and the result of evaluating the exponential map at time 1. From a geometric point of view it would be desirable to retain the anisotropy of the resulting ellipse at $t = 1$. However, LDDMM-based parallel transport clearly distorts the geometry and results in a much more circular shape: the ratio between the biggest and smallest axes decreases from 1.25 to 1.18.



Longitudinal evolution (geodesic) Transported evolution

Fig. 2. Left: 60 points on a circle are matched via a geodesic onto an ellipse with small anisotropy. Right: the transported evolution on a smaller circle: The anisotropy is distorted.

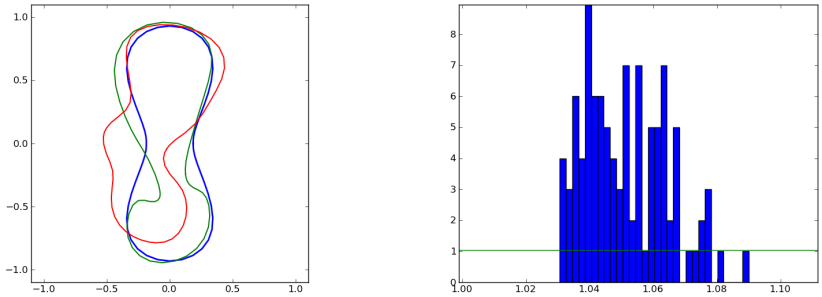


Fig. 3. Left: two random shapes in red and green and the blue template. Right: histogram of transported volume changes.

Last, we show an experiment illustrating the scaling issue when using parallel transport for population studies under the LDDMM metric: consider a population of closed curves drawn from a Gaussian model on the initial momentum around a template shape as shown in Fig. 3. Each shape of the population undergoes a small longitudinal change which is a uniform scaling centered on its barycenter. The population is separated into two groups, the first group has a scale evolution of 1.04 and the second group of 1.06. Hence, when looking at volume variation *alone*, the population clearly separates into two groups using a Gaussian Mixture Model (GMM) for instance. However, GMM completely fails to distinguish between the population when applied to the transported volume change. The histogram of the transported volume changes is given in Fig. 3.

These experiments motivate the need for a metric invariant to scale. Ideally, this should be accomplished by LDDMM to allow building on all its theory. Unfortunately, this cannot be achieved as described in the following Sec. 3.

3 Scale invariance and LDDMM geodesic flow

In Sec. 2 we observed that co-adjoint and parallel transport may exhibit counter-intuitive behavior under scaling using a Gaussian kernel. In this section, we show that the non-linear scaling effect is unavoidable when working with LDDMM, whatever the choice of the right-invariant metric.

We consider a group G of diffeomorphisms of the Euclidean space $\mathbb{R}^n$ which may or may not contain a group of diffeomorphisms denoted by G_0 that will represent the scaling transformations for instance. The first attempt to have scale invariance in the LDDMM framework is to ask whether we can design a kernel that defines a metric producing a global invariance of the flow of geodesics. Let us assume that G_0 contains the group of scaling transformations: $x \in \mathbb{R}^n \mapsto \lambda x$ for $\lambda \in \mathbb{R}_+^*$, but it may include more transformations such as translations and

rotations. The LDDMM framework is built on a group with a right-invariant distance that acts on the left on the space of shapes. It then induces a Riemannian metric on each orbit. A priori, requiring scale invariance for the induced metric is less demanding than asking it on the group itself. However, let us first explore the case of scale invariance on the group:

Definition 1 *The geodesic flow is invariant under G_0 if for any geodesics $t \mapsto \phi(t)$ the curve $t \mapsto g_0\,\phi(t)\,g_0^{-1}$ is also a geodesic.*

Remark 31 *This definition implies the following fact: if (q_0, q_1) are two objects connected by the geodesic $q(t)$ then $g_0 \cdot q(t)$ is a geodesic connecting $(g_0 \cdot q_0, g_0 \cdot q_1)$.*

Theorem 32 *There does not exist any smooth right-invariant metric on the group of diffeomorphisms of $\mathbb{R}^n$ for which the geodesic flow is invariant under G_0 that contains rotations, translations and the scaling transformations.*

Proof. Let us consider a geodesic path $\phi(t)$ whose vector field is denoted by $v(t)$. By Ad invariance the vector field associated with $g_0\phi(t)g_0^{-1}$ is $u(t) := \mathrm{Ad}_{g_0}(v(t))$. In addition to that, we know that $\phi(t)$ is a geodesic on the group if and only if it satisfies the Euler-Poincaré equation which is

$$\dot{m}(t) + \mathrm{ad}^*_{v(t)}\, m(t) = 0\,, \tag{7}$$

where $m(t) = Lv(t)$ (or $v(t) = Km(t)$). Equivalently, we have

$$\dot{v}(t) + K\,\mathrm{ad}^*_{v(t)}\, Lv(t) = 0, \;\; \mathrm{Ad}_{g_0}(\dot{v}(t)) + K\,\mathrm{ad}^*_{u(t)}\, Lu(t) = 0.$$

This implies that

$$\mathrm{Ad}^{-1}_{g_0}(K\,\mathrm{ad}^*_{\mathrm{Ad}_{g_0}\, v(t)}\, L\,\mathrm{Ad}_{g_0}\, v(t)) = K\,\mathrm{ad}^*_{v(t)}\, Lv(t)\,, \tag{8}$$

which is, taking the dual pairing with $\omega \in V^*$ and using the Ad invariance:

$$\begin{aligned}
&\left(\mathrm{ad}_{\mathrm{Ad}_{g_0}\, v(t)}\, K\,\mathrm{Ad}^{-1*}_{g_0}\,\omega, L\,\mathrm{Ad}_{g_0}\, v(t)\right) = \left(\mathrm{ad}_{v(t)}\, K\omega, Lv(t)\right)\,.\\
&\mathrm{ad}_{\mathrm{Ad}_{g_0}\, v(t)}\, K\,\mathrm{Ad}^{-1*}_{g_0}\,\omega = \mathrm{Ad}_{g_0}\left(\mathrm{ad}_{v(t)}\,\mathrm{Ad}^{-1}_{g_0}\, K\,\mathrm{Ad}^{-1*}_{g_0}\,\omega\right)\,.
\end{aligned}$$

Therefore, we have:

$$\left(\mathrm{ad}_{v(t)}\,\tilde{K}\omega, \tilde{L}v(t)\right) = \left(\mathrm{ad}_{v(t)}\, K\omega, Lv(t)\right)\,, \tag{9}$$

where $\tilde{K} = \mathrm{Ad}^{-1}_{g_0}\, K\,\mathrm{Ad}^{-1*}_{g_0}$ and $\tilde{L} = \tilde{K}^{-1}$. Hence, the Levi-Civita connections respectively associated with the right-invariant metrics K and $\tilde{K}$ are the same.

Let g_2 and $\tilde{g}_2$ denote the induced Riemannian metric on the space of two landmarks. The two metrics have the same geodesic equations and the same parallel transport. Since the two metrics are invariant w.r.t. translations and rotations, each kernel induces a metric on the landmark space on the real line that have the same geodesic equations. Let $L_{2,1}$ denote the manifold of 2 landmarks on $\mathbb{R}$, which is of dimension 2 and let $x \in L_{2,1}$ be a given point.

Let us first notice that the sectional curvature of the metric is given in page 80 of [5],

$$\kappa(r) = \frac{\gamma(0) - \gamma(r)}{\gamma(0) + \gamma(r)}\gamma''(r) - \frac{2\gamma(0) - \gamma(r)}{(\gamma(0) + \gamma(r))^2}[\gamma'(r)]^2\,, \tag{10}$$

where r denotes the distance between the two landmarks. The second term is non positive and a simple argument will show that the first term is negative at some point: The function γ is positive and decreasing and cannot be constant (due to the positivity of the kernel); this implies that γ is not concave (otherwise its limit in $+\infty$ would be $-\infty$) so that there exists $x_0 \in \mathbb{R}_+^*$ such that $\gamma''(x_0) < 0$. Therefore, at this point x_0, $\kappa(x_0) < 0$.

Up to a scaling factor, we can assume that the two metrics coincide at a given vector in $v \in T_x L_{2,1}$. Since parallel transports coincide, it implies that the two metrics coincide on the image of v under the holomy group of v. In order to conclude that the two metrics are equal, we note that the holonomy group of $L_{2,1}$ has no non-trivial invariant subspaces: If it were the case, the metric would be decomposable by the De Rham decomposition theorem as a product of two one-dimensional spaces. As a consequence, the curvature should vanish everywhere.

This equality (up to a scaling factor $\alpha > 0$) evaluated on Dirac distributions gives:

$$\left(\mathrm{Ad}_\lambda^{-1*}\, \delta_x^{p_x}, K\, \mathrm{Ad}_\lambda^{-1*}\, \delta_y^{p_y}\right) = \alpha \left(\delta_x^{p_x}, K \delta_y^{p_y}\right) . \tag{11}$$

Finally, using the fact that $\mathrm{Ad}_{g_0}^{-1*}\, \delta_x^{p_x} = \delta_{\lambda x}^{p_x/\lambda}$ when $g_0 : x \mapsto \lambda x$, we get:

$$\lambda^2 \langle p_x, K(x,y) p_y \rangle = \alpha \langle p_x, K(\lambda x, \lambda y) p_y \rangle \,, \tag{12}$$

and therefore: $\lambda^2 K(x,y) = \alpha K(\lambda x, \lambda y)$ for all $\lambda \in \mathbb{R}_+^*$. In particular, letting $\lambda \to 0$, we obtain, using the continuity of K, $K(0,0) = 0$ which is not a positive symmetric matrix. This is a contradiction. □

Remark 33 *If the kernel were not required to be continuous, a possible solution would be the Dirac kernel $\delta_{x,y}$. Such metrics are known to be degenerate [1].*

Hence, it is necessary to go beyond the standard LDDMM framework to obtain scale invariance. Our interest in scale invariance was only motivated by the study of parallel transport and its global and local effects. In Sec. 4, we characterize Riemannian metrics with invariance of a global indicator under parallel transport.

4 Designing Riemannian metrics

4.1 Decomposition theorem

Let us assume that we aim at preserving the volume variation in longitudinal evolutions, which is of interest in the case of Alzheimer's disease. In more mathematical words, the volume variation must be preserved under parallel transport. We will restrict ourselves to the space of shapes that are described as embeddings of the unit circle S_1 in $\mathbb{R}^2$ or sphere embeddings in $\mathbb{R}^3$. We would like to distinguish between local shape variation and global volume change. Although the two quantities are strongly linked in general situations, it seems quite reasonable to assume a uniform volume change. A natural approach to distinguish between volume variation and shape (up to scaling) variation is to decompose the space of shapes (in the spirit of [10]) using the following map:

$$\mathrm{Emb}(S_1, \mathbb{R}^2) \mapsto \mathbb{R}_+^* \times \mathrm{Emb}_1(S_1, \mathbb{R}^2); \quad s \to (\mathrm{vol}(s), P(s)) \,,$$

where vol is the surface delimited by the closed curve s and P is a chosen projection on the space of unit surface embeddings denoted by $\mathrm{Emb}_1(S_1, \mathbb{R}^2)$. A product between the standard Euclidean metric on $\mathbb{R}_+$ and a metric on the space $\mathrm{Emb}_1(S_1, \mathbb{R}^2)$ gives a (Riemannian) metric on the shape of space that meets our requirement: Namely, that the volume variation is invariant w.r.t. parallel transport. Using the de Rham decomposition theorem, it turns out that this is the only possible sort of metric that fulfills this invariance condition. The following result is a particular case of the De Rham decomposition theorem and we give a proof for completeness (see [8] for a proof in the general case).

Theorem 41 *Let g be a Riemannian metric on a connected Riemannian manifold M and a surjective function $f : M \mapsto \mathbb{R}$ such that $df(x) \neq 0$ for all $x \in M$ and which is invariant under parallel transport, i.e. $\nabla df = 0$, then (M, g) can be decomposed into a direct product of Riemannian metrics as follows:*

$$(M, g) = (\mathbb{R}, dt^2) \times (M_0, g_0) \tag{13}$$

where g_0 is a Riemannian metric on the submanifold $M_0 := f^{-1}(\{0\})$.

Proof. Let us first introduce the notation V being the unit length vector field associated with df via the metric g. In other notations, one has $V = df^\sharp$. In particular, if $Y \in T_{x_0}M_0$ then $\langle V, Y\rangle = df(Y) = 0$. It is easy to prove that the following mapping is a global diffeomorphism:

$$\Psi : \mathbb{R} \times M_0 \mapsto M; \quad (t, x_0) \mapsto \exp_{x_0}(tV(x_0)), \tag{14}$$

where exp denotes the Riemannian exponential. Indeed, since

$$\nabla_X V = 0 \ \forall X \in \chi(M), \tag{15}$$

we get $\nabla_V V = 0$ so that V is a geodesic vector field. In addition, for every vector field X, we have $R(V, X)V = \nabla_V \nabla_X V - \nabla_X \nabla_V V - \nabla_{[V,X]}V = 0$ so that the Jacobi field equation for a vector $J(x_0) \in T_{x_0}M$ reduces to $\frac{D^2J}{dt^2} + R(V, J)V = \frac{D^2J}{dt^2} = 0$.

Hence, the map (14) is a local diffeomorphism, which is obviously an injection so that this is a global diffeomorphism. Now, let us denote by X, Y two vector fields on M_0 trivially extended on M via the diffeomorphism Ψ. Namely, for each $t \in \mathbb{R}^*$, one defines $(t, x) \mapsto \psi_*(t)X(x)$ and $(t, x) \mapsto \psi_*(t)Y(x)$ the natural extensions of X and Y. One has $\psi_*(t)[X, Y] = [\psi_*(t)X, \psi_*(t)Y] = 0$. By construction, we also have $V(t, x_0) = \psi_*(t)V(0, x_0)$ so that this implies that $[X, V] = [Y, V] = 0$. Using $[X, V] = 0 = \nabla_X V - \nabla_V X$, and equation (15), we get $\nabla_V X = 0$. In particular, the equality

$$V \cdot g(X, Y) = g(\nabla_V X, Y) + g(X, \nabla_V Y) = 0$$

holds which means that the pull-back of the metric g by Ψ is $dt^2 + g_0$. □

Remark 42
- *This result is valid in finite dimension and might remain valid in a smooth infinite dimensional context. In applications however, shapes are approximated in high-dimensional spaces and the theorem does apply. Proving a convergence theorem (when the dimension increases) goes beyond the scope of the paper.*
- *The theorem can be generalized to k functions. The condition on the differentials would be that the family $(df_1(x), \ldots, df_k(x))$ is linearly independent.*

4.2 An induced Riemannian metric

To design a metric following the theorem, we therefore need a Riemannian metric on the space of unit volume shapes. To this end, one can use the restriction of any Riemannian metric on the space of volume preserving embeddings. Since we based our discussion on the LDDMM metric, we consider the Riemannian metric induced by LDDMM on that submanifold and we now present the geodesic computation for the LDDMM metric with a volume constraint.

We denote by V an admissible RKHS of vector fields (see [15]). Let $q \in \mathrm{Emb}_1(S_1, \mathbb{R}^2)$ be an embedding of surface of volume 1, we consider the set of vector fields $V_q := \{v \in V \,|\, d\,\mathrm{vol}_q(v(q)) = 0\}$. Since V is a RKHS of admissible vector fields, $v \to d\,\mathrm{vol}_q(v(q))$ is a continuous linear form and its kernel V_q is a closed subspace of V. Note that the notation $d\,\mathrm{vol}_q$ stands for the differential of the volume at point q which is a linear form on the tangent space $T_q\,\mathrm{Emb}(S_1, \mathbb{R}^2)$. We denote by π_q the orthogonal projection on $\mathrm{Ker}\, d\,\mathrm{vol}_q$ for the L^2 scalar product on $L^2(S_1, \mathbb{R}^2) \simeq T_q\,\mathrm{Emb}(S_1, \mathbb{R}^2)$.

A geodesic between two elements $q_0, q_1 \in \mathrm{Emb}_1(S_1, \mathbb{R}^2)$ is a solution of

$$\inf \int_0^1 \|v(t)\|^2_{V_{q(t)}}\, dt\,, \tag{16}$$

under the constraints $\dot{q} = v(t)(q)$ where $v(t) \in V_{q(t)}$ and $q(0) = q_0$ and $q(1) = q_1$.

Proposition 1 *The minimization problem* (16) *can be recast into:*

$$\inf \int_0^1 \|v(t)\|^2_V\, dt\,, \tag{17}$$

under the constraint $\dot{q} = \pi_q(v(t)(q))$ *where* $v(t) \in V$ *and* $q(0) = q_0$ *and* $q(1) = q_1$.

Proof. Clearly, Problem (16) is contained in Problem (17), since $v(t) \in V_{q(t)}$ implies $\pi_q(v(t)(q(t))) = v(t)(q(t))$. Let $v(t)$ be an optimal solution of Problem (17), then denoting $\pi_q : V \mapsto V_q$ the orthogonal projection, we get:

$$\int_0^1 \|\pi_q(v(t))\|^2_V\, dt \leq \int_0^1 \|v(t)\|^2_V\, dt\,.$$

Therefore, we have $\pi_q(v(t)) = v(t)$ which is a solution of Problem (16). □

Using an optimal control approach, the optimal solutions of Problem (17) are given by the solutions of the Hamiltonian equations

$$\begin{cases} \dot{q} = \partial_p H(p,q) \\ \dot{p} = -\partial_q H(p,q)\,, \end{cases} \tag{18}$$

where the Hamiltonian function is given by $H(p,q) = \frac{1}{2}\langle \pi(p), K(q)\pi(p)\rangle$; $K(q)$ is the kernel matrix associated to the LDDMM metric at point q.

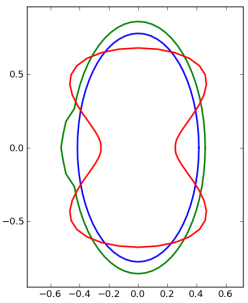
4.3 The choice of projection

A geodesic on the space of shapes can then be decomposed into a straight line on the volume axis and a geodesic on the submanifold of unit volume shapes. In the previous section, we have defined the metric on the space of unit volume shapes. It remains to define the volume geodesics. As mentioned in remark 42, the range of choice of projection may be large and it is natural to impose additional assumptions such as invariance w.r.t. translations, i.e. $P(T(m)) = T(P(m))$ for every translations T in $\mathbb{R}^n$. This is still not sufficiently constrained to uniquely determine the metric. Scaling invariance around the barycenter of the shape defined by $m(s) := \int_S c(s)ds = 0$ uniquely defines the projection: scale invariance means $P(\lambda m) = P(m)$ for every $\lambda \in \mathbb{R}_+^*$ for a centered (at 0) shape m. This is the first metric we will consider in the experiments.

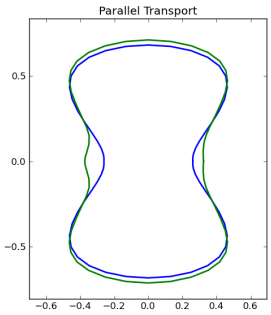
The notion of scale invariance also depends on the definition of the center of the shape which may be unnatural for some shapes. In order to avoid such a bias, we also propose to define the projection using the gradient flow of the volume with respect to a given metric, for instance the LDDMM metric: indeed, if f is a real function defined on a manifold M with no critical points, then the vector field ∇f is non-vanishing on M and defines the volume geodesics. However, the gradient is defined by the choice of a metric on the tangent space: for instance, an LDDMM type of metric which provides spatial correlation. This defines the second metric in our experiments.

5 Experimental Results

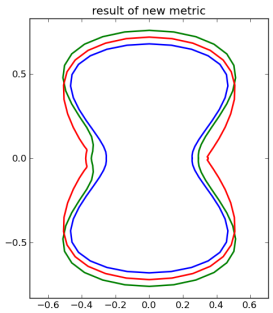
We compare LDDMM and the volume/shape-decoupled model represented by the two metrics introduced in Sec. 4.3. We use the Schild's ladder method to compute parallel transport. Fig. 4 illustrates the effect of the non-preservation of volume variation with the standard LDDMM metric even if template and target volumes and scales are equivalent. This shows that volume variation transport is already affected by shape deformations at the same scale. We use 60 landmarks and a Gaussian kernel of standard deviation 0.1 for the simulation. The volume variation for the LDDMM transported evolution is 1.06 whereas the initial data shows a volume variation of 1.104. By construction, for the new metrics, the volume variation is the same for the transported evolution. However, since the projections are different, the two different final curves in red and green are distinct. The last experiment illustrates the difference between the two new metrics, where the second metric uses a Gaussian kernel of width 0.01 in the definition of the volume gradient. We perform parallel transport of the longitudinal evolution shown in Fig. 4 (upper left) on the blue curve (bottom right). The transported evolutions exhibit very different behavior: the green curve is the transport using the scale invariant metric and shows that in some parts of the shape there is no local growth, whereas the other metric (represented by the red curve) offers a more uniform growth pattern on the shape.



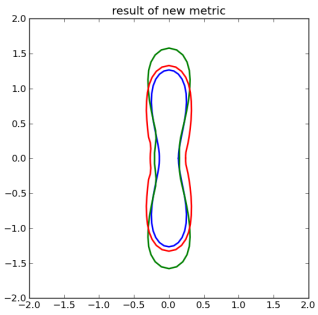
Longitudinal evolution, target shape in red (a scaling composed with a bump).



Transported evolution LDDMM with Gaussian kernel $\sigma = 0.1$.



Transported evolutions under the scale invariant metric in green and the other metric in red.



Difference between the two metrics.

Fig. 4. Examples of parallel transport under the new metrics.

6 Conclusions and Future Work

This paper explored the behavior of parallel transport for the LDDMM registration model. We showed that LDDMM is never scale invariant and does not conserve global properties such as absolute or relative volume changes. To achieve preservation of global properties we developed a new set of Riemannian metrics and demonstrated their behavior in comparison to the standard LDDMM model. While this paper so far only scratched the surface of metric design to achieve desired properties under parallel transport it raises fundamental issues for the analysis of longitudinal shape and image data when moving beyond global indicators. Future work will consist in estimating the statistical gain (e.g., w.r.t. LDDMM) when using the proposed metrics on a particular data set of biomedical shapes where a global indicator already achieves good performance.

References

1. Bauer, M., Bruveris, M., Harms, P., Michor, P.: Geodesic distance for right invariant Sobolev metrics of fractional order on the diffeomorphism group. Annals of Global Analysis and Geometry pp. 1–17 (2012)
2. Beg, M., Miller, M., Trouvé, A., Younes, L.: Computing large deformation metric mappings via geodesic flows of diffeomorphisms. IJCV 61(2), 139–157 (2005)
3. Fiot, J.B., Risser, L., Cohen, L.D., Fripp, J., Vialard, F.X.: Local vs global descriptors of hippocampus shape evolution for Alzheimer's longitudinal population analysis. In: STIA. pp. 13–24 (2012)
4. Lorenzi, M., Ayache, N., Pennec, X.: Schilds ladder for the parallel transport of deformations in time series of images. In: IPMI. pp. 463–474. Springer (2011)
5. Micheli, M.: The Differential Geometry of Landmark Shape Manifolds: Metrics, Geodesics, and Curvature. Ph.D. thesis, Division of Applied Mathematics, Brown University (2008)
6. Muralidharan, P., Fletcher, P.: Sasaki metrics for analysis of longitudinal data on manifolds. In: CVPR. pp. 1027–1034. IEEE (2012)
7. Rao, A., Chandrashekara, R., Sanchez-Ortiz, G., Mohiaddin, R., Aljabar, P., Hajnal, J., Puri, B.K., Rueckert, D.: Spatial transformation of motion and deformation fields using nonrigid registration. IEEE TMI 23(9), 1065–1076 (2004)
8. Sakai, T.: Riemannian geometry. Transl. from the Japanese by Takashi Sakai. Translations of Mathematical Monographs. 149. Providence, RI: American Mathematical Society (AMS). 358 p. (1996)
9. Singh, N., Fletcher, P., Preston, J., Ha, L., King, R., Marron, J., Wiener, M., Joshi, S.: Multivariate statistical analysis of deformation momenta relating anatomical shape to neuropsychological measures. MICCAI pp. 529–537 (2010)
10. Sundaramoorthi, G., Mennucci, A., Soatto, S., Yezzi, A.: A new geometric metric in the space of curves, and applications to tracking deforming objects by prediction and filtering. SIAM Journal on Imaging Sciences 4(1), 109–145 (2011)
11. Sundaramoorthi, G., Mennucci, A., Soatto, S., Yezzi, A.: A new geometric metric in the space of curves, and applications to tracking deforming objects by prediction and filtering. SIAM J.also Imaging Sciences (2011)
12. Sundaramoorthi, G., Yezzi, A.J., Mennucci, A.C.: Properties of Sobolev-type metrics in the space of curves. Interfaces and Free Boundaries, European Mathematical Society , 10(4), 423–445 (,2008)
13. Twining, C., Marsland, S., Taylor, C.: Metrics, connections, and correspondence: the setting for groupwise shape analysis. In: Energy Minimization Methods in Computer Vision and Pattern Recognition. pp. 399–412. Springer (2011)
14. Vaillant, M., Miller, M., Younes, L., Trouvé, A., et al.: Statistics on diffeomorphisms via tangent space representations. NeuroImage 23(1), 161 (2004)
15. Younes, L.: Shapes and diffeomorphisms, vol. 171. Springer (2010)
16. Younes, L., Michor, P., Shah, J., Mumford, D.: A metric on shape space with explicit geodesics. Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. 19(1), 25–57 (2008)
17. Younes, L., Qiu, A., Winslow, R., Miller, M.: Transport of relational structures in groups of diffeomorphisms. Journal of mathematical imaging and vision 32(1), 41–56 (2008)

Symmetries in LDDMM with higher-order momentum distributions

Henry O. Jacobs

Mathematics Department, Imperial College London, 180 Queen's Gate Road, London UK, SW7 2AZ

Abstract. In the *landmark large deformation diffeomorphic metric mapping* (landmark-LDDMM) formulation for image registration, we consider the motion of particles which locally translate image data. We then lift the motion of the particles to obtain a motion on the entire image. However, it is certainly possible to consider particles which also apply local rotations, scalings, and sheerings. These locally linear transformations manifest as a result of a non-trivial first derivative of a diffeomorphism. In this article, we seek to understand a structurally augmented particle which applies these local transformations. Moreover, we seek to go beyond locally linear transformations by considering the algebra behind the k-th order Taylor expansions of a diffeomorphism, a.k.a. the k-jet of a diffeomorphism. The particles which result from understanding the algebra of k-jets permit the use of higher-order local deformations. Additionally, these particles have internal symmetries which lead to conserved momenta when solving for geodesics between images. Knowledge of these structures provide both a sanity check and guide for future implementations of the LDDMM formalism.

1 Introduction

In the *Large Deformation Diffeomorphic Metric Mapping* (LDDMM) formulation of image registration, we begin by considering an image on a manifold M which we transform via the diffeomorphism group, $\mathrm{Diff}(M)$. As a finite dimensional representation of $\mathrm{Diff}(M)$ we consider the space of Landmarks, $Q := \{(\mathbf{x}_1, \dots, \mathbf{x}_N) \in M^N \mid \mathbf{x}_i \neq \mathbf{x}_j \text{ when } i \neq j\}$. Given a trajectory $q(t) \in Q$ we can construct a diffeomorphism $\varphi \in \mathrm{Diff}(M)$ by integrating a time dependent ODE obtained through a horizontal lift from TQ to $T\,\mathrm{Diff}(M)$. In particular, LDDMM is fueled by a natural lift from geodesics on Q into geodesics on $\mathrm{Diff}(M)$. However, this version of LDDMM only allows for local translations of image data (see Figure 1). If we desire to consider higher-order local transformations (such as shown in Figures 2 and 3) we should augment our particles with extra structure. This augmentation is precisely what is done in [SNDP13]. In this paper we explore this extra structure further by finding that the space in which these particles exist is a principal bundle, $\pi^{(k)} : Q^{(k)} \to Q$.[1] In other

[1] For example, if $M = \mathbb{R}^n$, we find $Q^{(1)} = \mathrm{GL}(n) \times \mathbb{R}^n$.

words, the particles of [SNDP13] have extra structure, analogous to the gauge of a Yang-Mills particle. As a result of Noether's theorem, this extra symmetry will yield conserved quantities. Identifying such conserved quantities can guide future implementations of the LDDMM formalism. In particular, integrators which conserve Noetherian momenta are typically very stable, and capable of accuracy over large time-steps [HLW02]. Finally, we seek to study this symmetry using coordinate-free notions so as to avoid restricting ourselves to $\mathbb{R}^n$. This is particularly important in medical imaging wherein one often deals with the topology of the 2-sphere [GVI04,KSKL13].

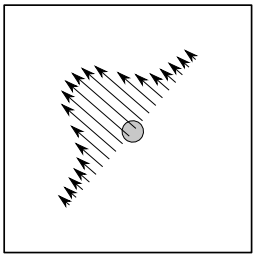


Fig. 1. A 0^{th}-order jet

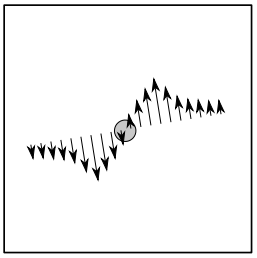


Fig. 2. A 1^{st}-order jet

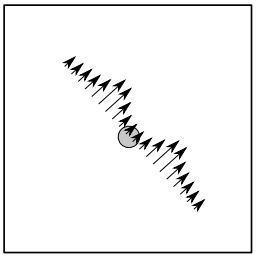


Fig. 3. A 2^{nd}-order jet

2 Background

Consider LDDMM [FIAL05,YILL05] and the deformable template model [Gre94]. In this framework one seeks to minimize a cost functional on $\mathrm{Diff}(M)$ which provides some notion of distance between images. It is customary to use particles (or "landmarks") and interpolate the particle velocities in order to obtain smooth vector fields on all of M. In certain cases, the diffeomorphisms obtained by integration of these fields satisfy a set of geodesic equations (see [MD10, §5.4] or [HSS09, Ch. 15]). However, local translations of data do not encompass all the local transformations of local image data. By "local image data" we mean the data of an image which is contained in the germ of a single point in M. Such considerations are articulated in [FRVK96] wherein the authors consider the truncated Taylor expansion of gray-scale image data about a single point. There the authors sought to understand how the heat equation evolves this Taylor expansion in order to design multi-scale filtering techniques. The spirit of [FRVK96] is very close to what will be explained here, where we will consider Taylor expansions of diffeomorphisms (a.k.a. "jets"). It is notable that the k^{th}-order jet of a diffeomorphism is precisely the (finite-dimensional) object required to advect the Taylor expansions of image data; thus, this paper can be seen as an LDDMM analog of [FRVK96]. In particular, we will be investigating the version of landmarck LDDMM proposed in [SNDP13], in which the particles exhibit higher order data. This higher order data is equivalent to the time-derivative of a k-jet

of a diffeomorphism. For the case $k = 1$, we obtain particles which are capable of locally scaling and rotating certain regions (these are called "locally affine transformations" in [SNDP13]). These more sophisticated particles have structures analogous to the gauge symmetry of Yang-Mills particles [CMR01,YM54]. In order to understand this, we will invoke the theory of symplectic reduction by symmetry [MW74,AM78]. In particular, we will perform subgroup reduction [CMR01,MMO$^+$07] to understand these new particles.

3 Geodesic flows on Diff(M)

Let M be a manifold and let $\mathrm{Diff}(M)$ be the group of smooth diffeomorphisms on M. The Lie algebra of $\mathrm{Diff}(M)$ is the vector space of smooth vector field on M, denoted $\mathfrak{X}(M)$, equipped with the *Jacobi-Lie bracket*.[2] In fact, any tangent vector $v_\varphi \in T_\varphi \mathrm{Diff}(M)$ can be viewed as a composition $v \circ \varphi$ for some $v \in \mathfrak{X}(M)$. We can construct a right invariant metric by choosing an inner product $\langle \cdot, \cdot \rangle_{\mathfrak{X}(M)} : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathbb{R}$. Given this inner-product, the corresponding Riemannian metric on $\mathrm{Diff}(M)$ is given by

$$\langle v_\varphi, u_\varphi \rangle_{\mathrm{Diff}(M)} := \langle \rho(v_\varphi), \rho(u_\varphi) \rangle_{\mathfrak{X}(M)} \tag{1}$$

where $\rho : v_\varphi \in T\,\mathrm{Diff}(M) \mapsto v_\varphi \circ \varphi^{-1} \in \mathfrak{X}(M)$ is the right Maurer-Cartan form.

The LDDMM formalism involves computing geodesics on $\mathrm{Diff}(M)$ with respect to a metric of the form (1). In particular this means solving the *Euler-Lagrange equations* with respect to the Lagrangian $L : T\,\mathrm{Diff}(M) \to \mathbb{R}$ given by

$$L(v_\varphi) := \frac{1}{2} \langle v_\varphi, v_\varphi \rangle_{\mathrm{Diff}(M)}. \tag{2}$$

Note that for any $\psi \in \mathrm{Diff}(M)$ we can act on the vector $v_\varphi \in T_\varphi \mathrm{Diff}(M)$ by the right action $v_\varphi \mapsto v_\varphi \circ \psi$. Under this action we obtain the following proposition [HSS09,MD10].

Proposition 1. *The Lagrangian of* (2) *is invariant with respect to the right action of* $\mathrm{Diff}(M)$ *on* $T\,\mathrm{Diff}(M)$ *given by composition.*

While $\mathrm{Diff}(M)$ is interesting, we will instead be leveraging smaller symmetry groups via the following corollary.

Corollary 1. *Let* $G \subset \mathrm{Diff}(M)$ *be a subgroup. Then* L *is* G*-invariant.*

4 Higher-order isotropy subgroups

Consider the space of n-tuples of non-overlapping points in M denoted by

$$Q := \{(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) \in M^n \mid i \neq j \implies \mathbf{x}_i \neq \mathbf{x}_j\}.$$

[2] We will either assume M is compact, or if $M = \mathbb{R}^d$ we will only consider vector-fields in the Schwartz space of all vector fields.

Given any $z \in Q$ we can consider the subgroup of $\mathrm{Diff}(M)$ given by $G_z := \{\psi \in \mathrm{Diff}(M) \mid \psi(z) = z\}$ where $\psi(z) \in M^n$ is short-hand for $(\psi(\mathbf{z}_1), \ldots, \psi(\mathbf{z}_n))$ with $z = (\mathbf{z}_1, \ldots, \mathbf{z}_n) \in M^n$. The homogenous space $\mathrm{Diff}(M)/G_z \equiv Q$ and landmark LDDMM can be seen as a reduction by symmetry with respect to G_z. However, we wish to consider the version of LDDMM proposed in [SNDP13] wherein higher-order momentum distributions are considered. To obtain particles with orientation and shape, we consider the subgroup $G_z^{(1)} = \{\psi \in G_z \mid T_z\psi = T_z id\}$ where $id \in \mathrm{Diff}(M)$ is the identity. In local coordinates $x^1, \ldots, x^d$ for M we see that $\psi \in G_z^{(1)}$ if and only if $\frac{\partial \psi^i}{\partial x^j}\Big|_{x=z_k} = \delta^i_j$ for $k = 1, \ldots, N$.

We find that the homogenous space $\mathrm{Diff}(M)/G_z^{(1)}$ is a non-overlapping subset of the N copies of the frame bundle of M [KMS99, Chapter 4]. This gives 1st-order particles new qualities, such as shape and orientation, and allows the landmark LDDMM formalism to express localized transformations such as local rotations and scalings. However, these locally linear transformations are only the beginning. We may consider "higher-order" objects as well. This requires a "higher-order" notion of isotropy, which in turn requires a "higher-order" notion of the tangent functor T.

4.1 Higher order tangent functors

Given two curves $a(t), b(t) \in M$ we write $a(\cdot) \sim_k b(\cdot)$ if the kth-order time derivatives at $t = 0$ are identical. In fact, $\sim_k$ is an equivalence class on the space of curves on M, and we denote the equivalence class of an arbitrary curve $a(\cdot)$ by $[a]_k$. For $k = 0$ the quotient-space induced by this equivalence is M itself. For $k = 1$ the quotient space is TM. For arbitrary $k \in \mathbb{N}$ we call the quotient space $T^{(k)}M$. In effect, $T^{(k)}M$ consists of points in M equipped with velocities, accelerations, and other higher order time-derivatives up to order k.

Finally, given any $\varphi \in \mathrm{Diff}(M)$ the map $T^{(k)}\varphi : T^{(k)}M \to T^{(k)}M$ is the unique map which sends the equivalence class $[a]_k$ consisting of the position $a(0)$, velocity $\frac{d}{dt}\big|_{t=0} a(t)$, acceleration $\frac{d^2}{dt^2}\Big|_{t=0} a(t)$, $\ldots$, $\frac{d^k}{dt^k}\Big|_{t=0} a(t)$ to the equivalence class $[\varphi \circ a]_k$ consisting of the position $\varphi(a(0))$, velocity $\frac{d}{dt}\big|_{t=0} \varphi(\dot{a}(t))$, acceleration $\frac{d^2}{dt^2}\Big|_{t=0} \varphi(a(t))$, $\ldots$, $\frac{d^k}{dt^k}\Big|_{t=0} \varphi(a(t))$. In other words, $T^{(k)}\varphi([a]_k) := [\varphi \circ a]_k$. It is simple to observe that $T^{(k)}\varphi : T^{(k)}M \to T^{(k)}M$ is a fiber-bundle diffeomorphism for each $\varphi \in \mathrm{Diff}(M)$. Moreover, $T^{(k)}$ is truly a functor, in the sense that $T^{(k)}(\varphi_1 \circ \varphi_2) = T^{(k)}\varphi_1 \circ T^{(k)}\varphi_2$ for any two $\varphi_1, \varphi_2 \in \mathrm{Diff}(M)$. Equipped with the functor $T^{(k)}$ we can define a notion of higher order isotropy.

4.2 Higher order isotropy subgroups

Again, choose a fixed $z \in Q$ and define $G_z^{(k)} := \{\psi \in G_z \mid T_z^{(k)}\psi = T_z^{(k)} id\}$. We may verify that $G_z^{(k)}$ is subgroup of G_z since for any two $\psi_1, \psi_2 \in G_z^{(k)}$ we observe

$$T_z^{(k)}(\psi_1 \circ \psi_2) = T^{(k)}_{\psi_2(z)}\psi_1 \cdot T_z^{(k)}\psi_2 = T_z^{(k)}\psi_1 \circ T_z id = T_z id.$$

In local coordinates $x^1, \dots, x^d$ for M we see that $\psi \in G_z^{(k)}$ if and only if

$$\left.\frac{\partial \psi^i}{\partial x^j}\right|_{x=z_k} = \delta^i_j \quad , \quad \left.\frac{\partial^{|\alpha|}\psi^i}{\partial x^\alpha}\right|_{x=z_k} = 0 \quad , \quad 1 < |\alpha| \leq k \quad , \quad k = 1, \dots, N.$$

In particular, the notion of a jet becomes relevant.

Definition 1. *The data which consists of local partial derivatives up to k-th order of $\varphi \in \mathrm{Diff}(M)$ about the points of $z \in Q$ is called the* k-jet of φ with source z *and denoted by $j^k_z(\varphi)$. We denote the set of jets sourced at z by $\mathcal{J}^k_z(\mathrm{Diff}(M))$. We call $z' = \varphi(z)$ the target of the jet $j^k_z(\varphi)$ and we can consider the set of jets with target z'.*

The above definition allows us to use the group structure of $\mathrm{Diff}(M)$ to induce a groupoid structure on jets. Namely, given two diffeomorphism $\varphi_1, \varphi_2 \in \mathrm{Diff}(M)$ such that $z' = \varphi_1(z)$ we define the product $j^k_{z'}(\varphi_2) \cdot j^k_z(\varphi_1) = j^k_z(\varphi_2 \circ \varphi_1)$, a jet with source z and target $\varphi_2(z')$.[3] Finally, the jet-functor allows us to define $G_z^{(k)}$ as $G_z^{(k)} = \{\psi \in G_z \mid j^k_z(\psi) = j^k_z(id)\}$.

Before we go further we must state the following observation.

Proposition 2. *For $z \in Q$, the group $G_z^{(k)}$ is a normal subgroup of G_z and the quotient space $G_z/G_z^{(k)}$ is itself a group.*

Proof. Let $\varphi \in G_z$ and let $\psi \in G_z^{(k)}$. We see that

$$T^{(k)}_z(\varphi \circ \psi \circ \varphi^{-1}) = T^{(k)}_{\psi(\varphi^{-1}(z))}\varphi \circ T^{(k)}_{\varphi^{-1}(z)}\psi \circ T^{(k)}_z\varphi^{-1}.$$

However $\varphi^{-1}(z) = z$, $\psi(z) = z$, and $T^{(k)}_z\psi = T^{(k)}_z(id)$ so that

$$T^{(k)}_z(\varphi \circ \psi \circ \varphi^{-1}) = T^{(k)}_z\varphi \circ T^{(k)}_z\psi \circ T^{(k)}_z\varphi^{-1} = T^{(k)}_z(id).$$

Thus $\varphi \circ \psi \circ \varphi^{-1} \in G_z^{(k)}$. As $\varphi \in G_z$ and $\psi \in G_z^{(k)}$ were chosen arbitrarily we see that $G_z^{(k)} \subset G_z$ is normal. Normality implies that the quotient $G_z/G_z^{(k)}$ is itself a Lie group.

It is not difficult to verify that $G_z/G_z^{(k)}$ is the set of k-jets of elements in G_z with source z. Then multiplication in $G_z/G_z^{(k)}$ is given by $j^k_z(\psi_1)\cdot j^k_z(\psi_2) = j^k_z(\psi_1 \circ \psi_2)$ for $\psi_1, \psi_2 \in G_z$. In the case that $k = 1$ we observe $G_z/G_z^{(1)} = \mathrm{GL}(d)^n$. In the case that $k = 2$ it is important to note that the set of constants $c^i_{jk} = \left.\frac{\partial^2 \varphi^i}{\partial x^j \partial x^j}\right|_{z_k}$ form a contravariant rank 1 covariant rank 2 tensors which satisfies $c^i_{jk} = c^i_{kj}$. We call the vector space of such tensors S^1_2 and we observe that $G_z/G_z^{(2)}$ is an n-fold cartesian product of the centered semi-direct product group $\mathrm{GL}(d) \ltimes S^1_2$ [CJ13].

Finally, we denote the homogenous space $Q^{(k)} := \mathrm{Diff}(M)/G_z^{(k)}$. The space $Q^{(k)}$ manifests in the LDDMM formalism of [SNDP13] by adding structure to the landmarks. Just as $Q^{(1)}$ was a right $\mathrm{GL}(d)$-principal bundle over Q, we will find $Q^{(k)}$ is a right $G_z/G_z^{(k)}$-principal bundle over Q.

[3] We refer the reader to [JRD13, §4] or [KMS99, Ch 4] for more information on jets and jet groupoids.

5 Reduction by symmetry at 0-momenta

In this section, our aim is to express $G_z/G_z^{(k)}$ as a symmetry group for a reduced Hamiltonian system on $T^*Q^{(k)}$. We will discuss reduction theory on the Hamiltonian side as performed in [AM78,MR99,Sin01]. Of course, the reduction theory of Lagrangian mechanics is well understood [CMR01] and we will use the notion of a Lagrangian momentum map to simplify certain computations.

We shall assume that the reader is familiar with the transition from Lagrangian mechanics to Hamiltonian mechanics via the Legendre transformation. In our case, this transformation yields a Hamiltonian $H : T^* \operatorname{Diff}(M) \to \mathbb{R}$. As H admits $\operatorname{Diff}(M)$-symmetry inherited from L, we obtain a Noether theorem. To investigate this, we consider the momentum map for the cases to be considered in this paper.

Proposition 3. *Given a Lie subgroup $G \subset \operatorname{Diff}(M)$ with Lie algebra $\mathfrak{g}$, there exists a natural action of $g \in G$ on $T^* \operatorname{Diff}(M)$ given by $\langle T^*R_g(p), v_\varphi\rangle = \langle p, TR_g v\rangle$. The momentum map $J : T^* \operatorname{Diff}(M) \to \mathfrak{g}^*$ induced by this action is given by $J(p_\varphi) = T^*\varphi \cdot p_\varphi|_{\mathfrak{g}}$*

Proof. By the definition of the momentum map [AM78] for each $\xi \in \mathfrak{g}$ and $p_\varphi \in T^*_\varphi \operatorname{Diff}(M)$ it must be the case that $\langle J(p_\varphi), \xi\rangle = \langle p_\varphi, T\varphi\cdot\xi\rangle = \langle T^*\varphi\cdot p_\varphi, \xi\rangle$. By construction $T^*\varphi \cdot p_\varphi \in (\mathfrak{X}(M))^*$. We observe that the momentum $J(p_\varphi)$ is merely the restriction of $T^*\varphi \cdot p_\varphi$ to the subspace of $\mathfrak{X}(M)$ given by $\mathfrak{g}$.

Noether's theorem states that J is constant in time along solutions of Hamilton's equations. Therefore, if $J = 0$ at $t = 0$ then $J = 0$ for all time. Moreover, we know that our $G_z^{(k)}$ symmetry allows us to reduce our equations of motion to the space $J^{-1}(0)/G_z^{(k)} \equiv T^*Q^{(k)}$ [MMO+07, Theorem 2.2.2]. We can therefore obtain a class of solutions to our equations of motion on $T^* \operatorname{Diff}(M)$ by solving a set of Hamilton's equations on $T^*Q^{(k)}$ with respect to a reduced Hamiltonian $H^{(k)} \in C^\infty(T^*Q^{(k)})$. Note that $H(T^*R_\psi \cdot p_\varphi) = H(p_\varphi)$ for any $\psi \in G_z^{(k)}$, and so H maps the entire $G_z^{(k)}$ equivalence class of a $p_\varphi \in T^* \operatorname{Diff}(M)$ to a single element. In particular, for any $p_\varphi \in J^{-1}(0)$ we set p equal to the $G_z^{(k)}$ equivalence class of p_φ and define $H^{(k)}(p) = H(p_\varphi)$. This implicitly defines the reduced Hamiltonian $H^{(k)} : J^{-1}(0)/G_z^{(k)} \equiv T^*Q^{(k)} \to \mathbb{R}$.

On the Lagrangian side we may define the Lagrangian momentum map $J_L : T \operatorname{Diff}(M) \to \mathfrak{g}_z^{(k)}$ given by composing J with the Legendre transform [AM78, Corollary 4.2.14]. Traversing a parallel path on the Lagrangian side will lead us to the reduced phase space $TQ^{(k)} \equiv J_L^{-1}(0)/G_z^{(k)}$. The reduced Lagrangian is defined by noting that for any $v_\varphi \in T \operatorname{Diff}(M)$, the $G_z^{(k)}$ symmetry of L implies that L sends the entire $G_z^{(k)}$-equivalence class of v_φ to a single number. Thus there exists a function $L^{(k)} : TQ^{(k)} \equiv J_L^{-1}(0)/G_z^{(k)} \to \mathbb{R}$ defined by $L^{(k)}(j_z^k(v_\varphi)) = L(v_\varphi)$ for any $v_\varphi \in J^{-1}(0)$. We can then solve geodesic equations on $TQ^{(k)}$ to obtain geodesics on $\operatorname{Diff}(M)$ via reconstruction.

Existing implementations of landmark LDDMM implicitly use this reduction to construct diffeomorphism by lifting paths in $Q^{(0)}$ to paths $\mathrm{Diff}(M)$ [BGBHR11,MD10]. In the context of particle methods for incompressible fluids, this correspondence is described explicitly in [JRD13]. In particular, the approach outlined in [SNDP13] suggests lifting paths in $Q^{(k)}$ to paths in $\mathrm{Diff}(M)$, and provides an implementation for $k = 1$. These more sophisticated particles contain extra structure, which opens the potential for extra symmetries. In particular, the fact that the system on $TQ^{(k)}$ (or $T^*Q^{(k)}$) is obtained by a $G_z^{(k)}$ reduction of systems with G_z symmetry has a consequence.

Theorem 1. *The reduced Hamiltonian and Lagrangian systems on $T^*Q^{(k)}$ and $TQ^{(k)}$ respectively have $G_z/G_z^{(k)}$ symmetry.*

Proof. Elements of $G_z/G_z^{(k)}$ are represented by $j_z^k(\psi)$ for some $\psi \in G_z$. We observe the natural action on $Q^{(k)}$ is given by $j_z^k(\varphi)\cdot j_z^k(\psi) := j_z^k(\varphi\circ\psi)$. We can lift this action to $TQ^{(k)}$ in the natural way by viewing a vector $v_\varphi \in T_\varphi \mathrm{Diff}(M)$ as the tangent of a curve $\varphi_t \in \mathrm{Diff}(M)$ with $\varphi_0 = \varphi$. In any case, we define the k-jet of v_φ sourced at z to be

$$j_z^k(v_\varphi) = \left.\frac{d}{dt}\right|_{t=0} j_z^k(\varphi_t) \in T(\mathcal{J}_z^k(\mathrm{Diff}(M))).$$

We then observe

$$j_z^k(v_\varphi)\cdot j_z^k(\psi) = \left.\frac{d}{dt}\right|_{t=0} j_z^k(\varphi_t\circ\psi) = \left.\frac{d}{dt}\right|_{t=0} j_z^k(\varphi_t) = j_z^k(v_\varphi).$$

Under this tangent lifted action we find

$$L^{(k)}(j_z^k(v_\varphi)\cdot j_z^k(\psi)) = L^{(k)}(j_z^k(v_\varphi\circ\psi)) = L(v_\varphi\circ\psi) = L(v_\varphi) = L^{(k)}(j_z^k(v_\varphi)).$$

Thus $L^{(k)}$ is $G_z/G_z^{(k)}$. As $H^{(k)}$ is merely the Hamiltonian associated to the Lagrangian $L^{(k)}$ it must be the case that $H^{(k)}$ also inherits this symmetry.

By Noether's theorem we find the following

Corollary 2. *Let $J : T^*Q^{(k)} \to \mathfrak{g}^*$ be the momentum map associated to the right action of $G = G_z/G_z^{(k)}$ on $T^*Q^{(k)}$. Then J is conserved along solutions to Hamilton's equations with respect to the Hamiltonian $H^{(k)} \in C^\infty(T^*Q^{(k)})$.*

6 Examples in $\mathbb{R}^d$

From this point on, let $M = \mathbb{R}^d$ and choose an inner product on $\mathfrak{X}(\mathbb{R}^d)$ given by the expression

$$\langle u, v\rangle_{\mathfrak{X}(\mathbb{R}^d)} = \int u(\mathbf{x})\cdot[\mathbb{I}(v)](\mathbf{x})d\mathbf{x}$$

where $\mathbb{I} : \mathfrak{X}(M) \to \mathfrak{X}(M)$ is a $\mathrm{SE}(d)$ invariant psuedo-differential operator with a C^k kernel given by $K : \mathbb{R}^d \to \mathbb{R}$. For example, we could consider $\mathbb{I} = (1 - \frac{1}{k+1}\Delta)^{k+1}$.

6.1 0^{th}-order particles in $\mathbb{R}^d$

Let us consider a single particle with initial condition $\mathbf{0} \in \mathbb{R}^d$. In this case we consider the isotropy group G_0, and the Lagrangian momentum map is given by

$$\langle J_L(v_\varphi), \xi \rangle = \int [\mathbb{I}(\rho(v_\varphi))] \cdot \xi(\mathbf{x}) d\mathbf{x} \quad \forall \xi \in \mathfrak{g}_0.$$

We see that $J_L(v_\varphi) = 0$ if and only if $\mathbb{I}(\rho(v_\varphi))$ is a constant vector-field times a dirac-delta centered at $\mathbf{0}$. That is to say

$$\langle J_L(v_\varphi), \xi \rangle = 0 \iff \int [\mathbb{I}(\rho(v_\varphi))] \cdot u(\mathbf{x}) d\mathbf{x} = a^i \cdot u^i(\mathbf{0})$$

for some constants $a^1, \dots, a^d \in \mathbb{R}$. By inspection, the statement holds if and only if

$$\rho(v_\varphi)(\mathbf{x}) = \boldsymbol{e}_i a^i K(\mathbf{x}) \tag{3}$$

where $\boldsymbol{e}_i$ is the ith basis vector of $\mathbb{R}^d$. For example, if $\mathbb{I} = \lim_{k\to\infty}(1 - \frac{1}{k}\Delta)^k$ then $K(\mathbf{x}) = \exp(-\|\mathbf{x}\|^2/2)$ [MMar]. Moreover, $J_L(v_\varphi) = 0$ if and only if $\rho(v_\varphi)(\mathbf{x}) = \boldsymbol{a} \exp(-\|\boldsymbol{x}\|^2/2)$ for some $\boldsymbol{a} \in \mathbb{R}^d$. As $G_0^{(k)} = G_0$ we find that the symmetry group for 0^{th}-order particles is the trivial group $G_0/G_0 = \{e\}$. In other words, 0^{th}-order particles admit trivial internal symmetry, and the reduced configuration manifold is simply $\mathbb{R}^d$.

6.2 1^{st}-order particles in $\mathbb{R}^d$

In this case we consider the isotropy group $G_0^{(1)}$, and the Lagrangian momentum map is given by

$$J_L(v_\varphi) \cdot \xi = \int [\mathbb{I}(\rho(v_\varphi))] \cdot \xi(\boldsymbol{x}) d\boldsymbol{x} \quad \forall \xi \in \mathfrak{g}_0^{(1)}.$$

We see that $J_L(v_\varphi) = 0$ if and only if $\mathbb{I}(\rho(v_\varphi))$ is the sum of a velocity field of the form (3) plus a second vector field which satisfies the derivative reproducing property

$$\int \mathbb{I}(\rho(v_\varphi))(\boldsymbol{x}) u(\boldsymbol{x}) d\boldsymbol{x} = -b_j^i \partial_j u^i(\mathbf{0})$$

for some set of constants b_j^i. If K is differentiable then a simple integration by parts argument reveals that

$$\rho(v_\varphi)(\boldsymbol{x}) = \boldsymbol{e}_i (a^i K(\boldsymbol{x}) + b_j^i \partial_j K(\boldsymbol{x})) \tag{4}$$

for some set of real numbers $a^i, b_j^i \in \mathbb{R}$ [SNDP13, c.f equation (4.1)]. For example, if $\mathbb{I} = \lim_{k\to\infty}(1 - \frac{1}{k}\Delta)^k$ then $K(\boldsymbol{x}) = \exp\left(-\|\boldsymbol{x}\|^2/2\right)$ and $J_L(v_\varphi) = 0$ if and only if

$$\rho(v_\varphi)(\boldsymbol{x}) = \boldsymbol{e}_i a^i \exp\left(\frac{-\|\boldsymbol{x}\|^2}{2}\right) + \boldsymbol{e}_i b_j^i x^j \exp\left(\frac{-\|\boldsymbol{x}\|^2}{2}\right).$$

These are the affine particles mentioned in [SNDP13]. A schematic of one of these particles is shown in Figure 2. It is simple to verify that the reduced configuration manifold is $Q^{(1)} = \mathbb{R}^d \times \mathrm{GL}(d)$.

6.3 Symmetries for 1st-order particles in $\mathbb{R}^d$

The symmetry group for a single 1st-order particle is $G_0^{(1)}/G_0 = \mathrm{GL}(d)$. We can verify this by noting that for each $\psi \in G_0$, the jet $j_0^1(\psi)$ is described by the partial derivative $\frac{\partial \psi^i}{\partial x^j}\Big|_{x=0}$. Moreover, by the chain rule we find

$$\frac{\partial}{\partial x_j}\bigg|_{x=0}(\psi_1^i \circ \psi_2) = \frac{\partial \psi_1^i}{\partial x^k}\frac{\partial \psi_2^k}{\partial x^j}\bigg|_{x=0}$$

which is the coordinate expression for the definition of the jet-groupoid composition $j_0^1(\psi_1)\cdot j_0^1(\psi_2) = j_0^1(\psi_1 \circ \psi_2)$. It is notable that the group of scalings and rotations is contained in $\mathrm{GL}(d)$. These transformations were leveraged in [SNDP13] to create "locally affine" transformations.

Finally, $\mathrm{GL}(d)$ acts on the $\mathrm{GL}(d)$ component of $Q^{(1)} = \mathbb{R}^d \times \mathrm{GL}(d)$ by right matrix multiplication. This makes $Q^{(1)}$ a trivial right $\mathrm{GL}(d)$-principal bundle over $\mathbb{R}^d$.

6.4 2nd order particles in $\mathbb{R}^d$ (and beyond)

In this case we consider the isotropy group $G_0^{(2)}$. We see that $J_L(v_\varphi) = 0$ if and only if $\mathbb{I}(\rho(v_\varphi))$ is the sum of a velocity field of the form (4) plus a second vector field which satisfies the second-derivative reproducing property

$$\int \mathbb{I}(\rho(v_\varphi))(\boldsymbol{x})u(\boldsymbol{x}) = c^i_{jk}\partial_j\partial_k u^i(\mathbf{0})$$

for some set of constants c^i_{jk}. If K is twice-differentiable, then a second integration by parts reveals

$$\rho(v_\varphi)(\boldsymbol{x}) = \boldsymbol{e}_i(a^i K(\boldsymbol{x}) + b^i_j\partial_j K(\boldsymbol{x}) + c^i_{jk}\partial_j\partial_k K(\boldsymbol{x})). \tag{5}$$

For example, if $\mathbb{I} = \lim_{k\to\infty}(1-\frac{1}{k}\Delta)^k$ then $K(\boldsymbol{x}) = \exp\left(-\|\boldsymbol{x}\|^2/2\right)$ and $J_L(v_\varphi) = 0$ if and only if

$$\begin{aligned}\rho(v_\varphi)(\boldsymbol{x}) = \boldsymbol{e}_i a^i \exp\left(\frac{-\|\boldsymbol{x}\|^2}{2}\right) + \boldsymbol{e}_i b^i_j x^j \exp\left(\frac{-\|\boldsymbol{x}\|^2}{2}\right)\\ + \boldsymbol{e}_i c^i_{jk}(x^j x^k - \delta^k_j)\exp\left(\frac{-\|\boldsymbol{x}\|^2}{2}\right).\end{aligned}$$

We see that the collection of constants $\{c^i_{jk}\}$ transform as (and therefore must be equal to) contravariant rank 1 covariant rank 2 tensors. Moreover, we see observe the symmetry $c^i_{jk} = c^i_{kj}$. We will denote the set of such tensors by S^1_2

so that the reduced configuration space is $Q^{(2)} = \mathbb{R}^d \times \mathrm{GL}(d) \times S^1_2$. See Figure 3 for a schematic of one of these particles.

At this point, we may deduce that for reduction by $G^{(k)}_z$, the set $J_L^{-1}(0)$ consists of vectors $v_\varphi \in T\,\mathrm{Diff}(M)$ which satisfy $\rho(v_\varphi)(x) = \sum_{|\alpha|\leq k} \boldsymbol{e}_i c^i_\alpha \partial_\alpha K$ for some series of $c_\alpha \in \mathbb{R}$, where the dummy index α is a multi-index. The reduced configuration space is $Q^{(k)} = \mathbb{R}^d \times \mathrm{GL}(d) \times S^1_2 \times \cdots \times S^1_k$.

6.5 Symmetries for 2nd-order particles in $\mathbb{R}^d$

In the case for a single second order particle in $\mathbb{R}^d$, the symmetry group is $G^{(2)}_0/G_0$. We see that the 2-jet of a $\psi \in G_0$ is described by the numbers $\frac{\partial \psi^i}{\partial x^j}(\mathbf{0})$ and $\frac{\partial^2 \psi^i}{\partial x^k \partial x^j}(\mathbf{0})$. By representing the 2-jets concretely as partial derivatives, we arrive at the following description for $G_z/G^{(2)}_z$.

Proposition 4. *The group $G^{(2)}_0/G_0$ is isomorphic to the centered semi-direct product $\mathrm{GL}(d) \ltimes S^1_2$ where S^1_2 is the vector space of rank $(1,2)$ tensors on $\mathbb{R}^d$ which are symmetric in the lower indices.*

Proof. For any $\psi, \varphi \in G_0$ we find that

$$\frac{\partial}{\partial x^j}\bigg|_{x=0} (\varphi^i \circ \psi) = \frac{\partial \varphi^i}{\partial x^k} \cdot \frac{\partial \psi^k}{\partial x^j}\bigg|_{x=0}$$

and

$$\begin{aligned}\frac{\partial^2}{\partial x^k \partial x^j}\bigg|_{x=0} (\varphi^i \circ \psi) &= \frac{\partial}{\partial x^k}\bigg|_{x=0} \left(\frac{\partial \varphi^i}{\partial x^\ell}(\psi(x)) \cdot \frac{\partial \psi^\ell}{\partial x^j}(x) \right) \\ &= \left(\frac{\partial^2 \varphi^i}{\partial x^\ell \partial x^m} \frac{\partial \psi^\ell}{\partial x^j} \frac{\partial \psi^m}{\partial x^k} + \frac{\partial \varphi^i}{\partial x^j} \frac{\partial^2 \psi^j}{\partial x^j \partial x^k} \right)\bigg|_{x=0}\end{aligned}$$

Observe that $\frac{\partial^2 \partial}{\partial x^j \partial x^k}\Big|_{x=0}$ is a contravariant rank 1 covariant rank 2 tensor, symmetric in the covariant indices. In other words, it is an element of S^1_2. Moreover, the above formulas match the composition rule for a centered semi-direct product $\mathrm{GL}(d) \ltimes S^1_2$ where we use the natural left action of $\mathrm{GL}(d)$ on S^1_2 given by

$$(b \cdot c)^i_{jk} = b^i_\ell \cdot c^\ell_{jk} \qquad b \in \mathrm{GL}(d), c \in S^1_2$$

and the natural right action given by

$$(c \cdot b)^i_{jk} = c^i_{\ell m} b^\ell_j b^m_k \qquad b \in \mathrm{GL}(d), c \in S^1_2.$$

In particular, the group composition is given by $(b, c) \cdot (\tilde{b}, \tilde{c}) = (b \cdot \tilde{b}, b \cdot \tilde{c} + c \cdot \tilde{b})$. For more a verification that this is a well defined group we refer the reader to [CJ13].

Given this symmetry group, we find that $Q^{(2)} = \mathbb{R}^d \times \mathrm{GL}(d) \ltimes S^1_2$ is a (trivial) right $\mathrm{GL}(d) \ltimes S^1_2$-principal bundle over $Q = \mathbb{R}^d$.

7 Conclusion

In this paper we have identified a family of isotropy groups which can be used to perform reduction by symmetry of geodesic equations on $\mathrm{Diff}(M)$. We then observed that the reduced configuration spaces consisted of particles with extra group symmetries much like Yang-Mills particles. This extra structure was interpreted as a "localized transformation" and corresponds to the higher-order structures described in [SNDP13]. Computations for $M = \mathbb{R}^d$ were performed, and the appropriate velocity fields matched those described in [SNDP13]. Finally, the symmetry groups for two classes of higher-order particles were computed to be $\mathrm{GL}(d)$ and $\mathrm{GL}(d) \ltimes S^1_2$. This extra structure can be leveraged to provide greater accuracy and flexibility in existing implementations of landmark LDDMM.

Acknowledgments

The author would like to thank Sarang Joshi for initiating this article as well as introducing him to [SNDP13]. We also wish to thank Tudor S. Ratiu for helpful discussions on jets as well as David Meier and Darryl D. Holm for clarifying the Clebsch approach to deriving equations of motion. Finally, the author would like to thank the reviewers for improving the accessibility of the paper and introducing him to [FRVK96]. H.O.J. is supported by European Research Council Advanced Grant 267382 FCCA.

References

[AM78] Ralph Abraham and Jerrold E. Marsden, *Foundations of Mechanics*, Benjamin/Cummings Publishing Co. Inc. Advanced Book Program, Reading, Mass., 1978, Second edition, revised and enlarged, with the assistance of Tudor Ratiu and Richard Cushman. Reprinted by AMS Chelsea, 2008. MR 515141 (81e:58025)

[BGBHR11] M. Bruveris, F. Gay-Balmaz, D.D. Holm, and T.S. Ratiu, *The momentum map representation of images.*, J. Nonlinear Sci. **21** (2011), no. 1, 115–150 (English).

[CJ13] L Colombo and H O Jacobs, *Lagrangian mechanics on centered semi-direct products*, to appear in the Fields Institute Communications, arXiv:1303.3883, March 2013.

[CMR01] Hernán Cendra, Jerrold E. Marsden, and Tudor S. Ratiu, *Lagrangian reduction by stages*, Mem. Amer. Math. Soc. **152** (2001), no. 722. MR 1840979 (2002c:37081)

[FIAL05] Beg M F, Miller M I, Trouve A, and Younes L, *Computing large deformation metric mappings via geodesic flows of diffeomorphisms*, IJCV **61** (2005), 139–157.

[FRVK96] L Florack, B T H Romeny, M Viergever, and J Koenderink, *The Guassian scale-space paradigm and the multiscale local jet*, International Journal of Computer Vision **18** (1996), 61–75.

[Gre94] Ulf Grenander, *General pattern theory: A mathematical study of regular structures*, Oxford University Press, 1994.

[GVI04] J Glaunès, M Vaillant, and Miller M I, *Landmark matching via large deformation diffeomorphisms on the sphere*, Journal of mathematical imaging and vision **20** (2004), no. 1–2, 179–200.

[HLW02] E Hairer, C Lubich, and G Wanner, *Geometric numerical integration, structure preserving algorithms for ordinary differential equations*, Series in Computational Mathematics, vol. 31, Springer Verlag, 2002.

[HSS09] D D Holm, T Schmah, and C Stoica, *Geometric mechanics and symmetry*, Oxford University Press, 2009.

[JRD13] H O Jacobs, T S Ratiu, and M Desbrun, *On the coupling between an ideal fluid and immersed particles*, arXiv:1208.6561, 2013.

[KMS99] I Kolář, P W Michor, and J Slovák, *Natural operations in differential geometry*, Springer Verlag, 1999.

[KSKL13] Sebastian Kurtek, Anuj Srivastava, Eric Klassen, and Hamid Laga, *Landmark-guided elastic shape analysis of spherically-parameterized surfaces*, Computer Graphics Forum **32** (2013), no. 2pt4, 429–438.

[MD10] David Mumford and Agnès Desolneux, *Pattern theory*, Applying Mathematics, A K Peters Ltd., Natick, MA, 2010, The stochastic analysis of real-world signals. MR 2723182 (2011m:94045)

[MMO+07] J E Marsden, G Misiolek, J P Ortega, M Perlmutter, and T S Ratiu, *Hamiltonian reduction by stages*, Lecture Notes in Mathematics, vol. 1913, Springer, 2007.

[MMar] D Mumford and P W Michor, *On Euler's equation and 'EPDiff'*, Journal of Geometric Mechanics (to appear), arXiv:1209.6576 [math.AP].

[MR99] J E Marsden and T S Ratiu, *Introduction to mechanics and symmetry*, 2nd ed., Texts in Applied Mathematics, vol. 17, Springer Verlag, 1999.

[MW74] J E Marsden and A Weinstein, *Reduction of symplectic manifolds with symmetry*, Reports on Mathematical Physics **5** (1974), 121–130.

[Sin01] Stephanie Frank Singer, *Symmetry in mechanics*, Birkhäuser Boston Inc., Boston, MA, 2001, A gentle, modern introduction. MR 1816059 (2002b:53128)

[SNDP13] S. Sommer, M. Nielsen, S. Darkner, and X. Pennec, *Higher-order momentum distributions and locally affine LDDMM registration*, SIAM Journal on Imaging Sciences **6** (2013), no. 1, 341–367.

[YILL05] Cao Y, Miller M I, Winslow R L, and Younes L, *Large deformation diffeomorphic metric mapping of vector fields*, IEEE Transactions on Medical Imaging (2005), 1216–1230.

[YM54] C. N. Yang and R. L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, Physical Review **96** (1954), 191–195.

Combining Thickness Information with Surface Tensor-based Morphometry for the 3D Statistical Analysis of the Corpus Callosum

Liang Xu[a], Olivier Collignon[d], Gang Wang[a,e], Yue Kang[c], Franco Leporé[f], Jie Shi[a], Yi Lao[c], Anand Joshi[g], Natasha Leporé[b,c,*], Yalin Wang[a,*]

[a] *School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Tempe, AZ, USA*
[b] *Department of Radiology, Children's Hospital Los Angeles, CA, USA*
[c] *Departments of Radiology & Biomedical Engineering, University of Southern California, CA, USA*
[d] *Center for Mind/Brain Science, University of Trento, Trento, Italy*
[e] *School of Computer Science and Technology, Ludong University, China*
[f] *Department of Psychology, University of Montreal, Montreal, QC, Canada*
[g] *Signal and Image Processing Institute, Brain and Creativity Institute, University of Southern California, Los Angeles, CA, USA*
[*] *Equal senior authors.*

Abstract. We propose a novel framework to capture a complete set of 3D morphological differences in the corpus callosum (CC) between two groups of subjects. The CCs are segmented from whole brain T1-weighted magnetic resonance images and modeled as 3D tetrahedral meshes. The callosal surface is divided into superior and inferior patches on which we compute a volumetric harmonic field by solving the Laplace's equation with Dirichlet boundary conditions. We adopt a refined tetrahedral mesh to compute the Laplacian operator, so our computation can achieve sub-voxel accuracy. Thickness is estimated by tracing the streamlines in the harmonic field. We combine areal changes found using surface tensor-based morphometry and thickness information into a vector at each vertex to be used as a metric for the statistical analysis. Group differences are assessed on this combined measure through Hotelling's T^2 test. The method is applied to statistically compare three groups consisting of: congenitally blind (CB), late blind (LB; onset > 8 years old) and sighted (SC) subjects. Our results reveal significant differences in several regions of the CC between both blind groups and the sighted group, and to a lesser extent between the LB and CB groups. These results demonstrate the crucial role of visual deprivation during the developmental period in reshaping the structural architecture of the CC.

1 Introduction

The corpus callosum (CC) is one of the most highly studied subcortical structure in post-processing analyses of magnetic resonance images. This is in part due to its involvement in numerous disorders that affect the brain. The splenium of the

CC carries fibers that connect visuo-spatial areas of the brain, and the isthmus is also involved in visuo-spatial processing, as it contains fibers connecting the posterior parietal areas, which fuse multimodality sensory information [1]. The CC undergoes extensive myelinization during development until adolescence, and waves of peak growth rates can be observed in the CC's of children of different ages [2]. Hence, studying the respective impact of congenitally (CB) versus lately acquired blindness (LB) on the anatomy of the CC provides a unique model to probe how experience at different developmental periods shapes the structural organization of the brain.

On the processing side, in T1-weighted magnetic resonance images (MRI), its high contrast difference from surrounding structures make accurate callosal segmentations straightforward for both manual and automatic methods. Additionally, its functional differentiation along an elongated sagittal axis has allowed researchers to focus on 2D analysis of the mid-sagittal section, allowing for simpler and faster numerical tools. Even so, it is clear that a 3D structural analysis can help visualization and may pick up some important information about the 3D structure of the CC that is discarded by 2D process.

We propose a novel 3D pipeline for the 3D analysis of the CC. While most studies have focussed on 2D representations of this structure, Wang et al. [3] compared the 3D CC of premature neonates to that of term-born controls. In that work, a surface grid was generated on the CC, and callosal thickness was computed as the distance from a medial axis. Statistical significance was assessed at each vertex on a vector containing the thickness and the deformation tensors from a multivariate tensor-based morphometry analysis (mTBM). The deformation tensors represent changes in area on the surface. However, for concave callosal surfaces, the medial axis is not well-defined and does not always have a biologically meaningful interpretation. Here we propose a new thickness computation to be combined with the standard mTBM analysis as in [3]. Given 3D tetrahedral meshes of the CC, we use the mesh based Laplacian operator to compute a harmonic field. The thickness is computed from the streamlines of the harmonic field. The estimated callosal thickness is well-defined, and may reflect the intrinsic 3D geometrical structure better than thickness derived from a medial axis and facilitate consistent cross-subject comparisons.

In the field of computational anatomy, tensor-based morphometry (TBM) [4, 5] and more recently its multivariate extension, mTBM [6, 7], have been used extensively to detect regional differences in surface and volume brain morphology between groups of subjects. Here we focus more specifically on mTBM on the callosal surface. Prior work [6] combining mTBM with other statistics such as the radial distance has significantly improved statistical power. Intuitively, thickness and mTBM are complementary, as thickness describes distances roughly along the surface's normal direction, while mTBM detects surface dissimilarities, including differences in the surface metric tensor induced by the particular surface parameterization. So we argue that a combination of thickness and mTBM will offer a more complete set of surface statistics for callosal morphometry and hypothesize that they may boost statistical power to detect disease effects.

In this paper, we propose a combined multivariate morphometry statistics to study callosal differences associated with congenital-onset versus late-onset blindness. Our pipeline is applied on a data set consisting of: 14 congenitally blind (CB), 10 late blind (LB; onset > 8 years old) and 20 sighted control (SC) subjects. Prior 2D TBM analyses of the corpus callosum [7] revealed reductions in the ithmus and the splenium of the corpus callosum in early but not late blind compared to sighted controls. Comparisons of the early and late blind groups did not find any significant changes, though we hypothesize that they may be detected by our more powerful method. Additionally, [7] observed changes in the frontal lobes, though those were not reflected in the 2D analyses of the callosal regions connecting areas of the frontal lobes.

There are three main contributions in this paper. First, we propose an efficient method to compute the harmonic field with a tetrahedral mesh. Prior work on voxel-based brain thickness analysis [8, 9] relied on a three-dimensional cubic voxel grid to solve partial differential equations (PDE) in the potential field. However, due to the restrictions on the grid resolution which cannot precisely characterize the curved cortical surfaces in MR images, the measurement accuracy from this method is low and sensitive to noise. Our approach overcomes the defect of the limited grid resolution by adopting a high quality, adaptive tetrahedral mesh [10] and a finite element based Laplacian operator [11]. Compared with prior work [8, 9], our PDE solving computation can achieve sub-voxel accuracy. Second, we propose a multivariate statistics by combining the callosal thickness computed from our new method and mTBM. Lastly, through multiple comparison, we identify statistically significant areas on CC between the CB and LB groups. This discovery may help further our understanding of brain plasticity and in the long term, improve the effectiveness of rehabilitation techniques for blind individuals.

2 3D Callosal Thickness Computation with Harmonic Field

2.1 Solving Laplace's Equation with Volumetric Laplacian Operators

Laplace's equation $\Delta f = \nabla^2 f = 0$ in 3D Cartesian coordinates takes the form:

$$(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}) f(x, y, z) = 0.$$

f is called harmonic if it satisfies the Laplace's equation with Dirichlet boundary conditions. The computed function is called the *harmonic field*. Assume there are two boundaries, B_0 and B_1, the harmonic field is computed by solving for the harmonic function $f_M : M \to \mathbb{R}$, such that

$$\begin{cases} \Delta f_M(p) = 0 \; \forall p \notin B_0 \cup B_1 \\ f_M(p) = 0 \quad \forall p \in B_0 \\ f_M(p) = 1 \quad \forall p \in B_1 \end{cases} \tag{1}$$

Eqn. 1 has been used to estimate the thickness of cerebral cortex [8, 12–14] and CC thickness on the mid-sagittal section [9]. Here we propose a finite element approach to solve Eqn. 1 and achieve sub-voxel accuracy on the boundaries. Compared with prior voxel-based scheme [8, 9, 12–14], our new work may overcome the numerical inaccuracy due to the limited resolution of 3D grid.

Practically, we use tetrahedra to represent the volume data. Suppose K is a simplicial complex, and $g : |K| \to \mathbb{R}^3$ a function that embeds $|K|$ in $\mathbb{R}^3$, then (K, g) is called a mesh. For a 3-simplex, it is a tetrahedral mesh, Te, and for a 2-simplex, it is a triangular mesh, Tr. Clearly, the boundary of a tetrahedral mesh is a triangular mesh, $Tr = \partial Te$.

Since the conventional harmonic energy is equivalent to the discrete harmonic energy [11] on a mesh, one may use the discrete Laplacian operator to minimize the harmonic energy. Here we adopt the Laplacian operator defined with tetrahedral mesh [11], as well as the following definitions.

Definition 1. *Suppose that edge $\{u, v\}$ is shared by n tetrahedra; thus it lies against n dihedral angles, θ_i, $i = 1, \ldots, n$. Denote l_i as the length of each edge that edge $\{u, v\}$ lies against in the domain manifold M, one can define the parameters $k_{u,v} = \frac{1}{12} \sum_{i=1}^{n} l_i \cot(\theta_i)$.*

Definition 2. *The* piecewise Laplacian *is the linear operator $\Delta_{PL} : C^{PL} \to C^{PL}$ on the space of piecewise linear functions on K, defined by the formula $\Delta_{PL}(f) = \sum_{\{u,v\} \in K} k(u, v)(f(v) - f(u))$.*

Definition 3. *Given a tetrahedral mesh, the* graph weight matrix *is defined as $S_{u,v} = \begin{cases} k_{u,v} & \exists e_{u,v} \\ 0 & \neg \exists e_{u,v} \end{cases}$, where $k_{u,v}$ is defined in Definition 2 [11]. Clearly, S is a sparse matrix and can be decomposed as $S = \begin{pmatrix} W_{VV} & W_{V\partial V} \\ W_{\partial VV} & W_{\partial V\partial V} \end{pmatrix}$, where V and ∂V represent the set of internal vertices and boundary vertices, respectively.*

Definition 4. *Under Dirichlet boundary conditions, the Laplacian matrix is $L_p = D_{VV} + D_{V\partial V} - S_{VV}$, where the diagonal matrix $D_{V\partial V} = diag(S_{V\partial V}\mathbf{e}_i)$, $\mathbf{e}_i$ is the ith column vector in an identity matrix, i.e., $(D_{V\partial V})_{ii}$ = the sum of i-th row in $W_{V\partial V}$.*

Definition 5. *With the discrete Laplacian operator definition, we compute the harmonic field with Dirichlet boundary conditions,*

$$L_p x = c, \tag{2}$$

where x is a $|u| \times 1$ vector ($|u|$ is the number of internal vertices). Note x only contains unknown function values on internal vertices, i.e. W_{VV}, as shown in the definition of L_p; and constant vector c is computed by $c_u = f_l^T W_{\partial VV} = \Sigma_{[v,w] \in M} f_l k_{v,w}$, where f_l is the specified function value on boundary vertices.

2.2 Thickness Profile Generation with the Harmonic Fields

Eqn. 2 is the discretized version of Eqn. 1. After computing the harmonic field f by solving x for internal vertices in Eqn. 2, we can compute the streamlines

to connect the two surfaces [8, 9]. Computationally, we construct a streamline as a parametric curves $u(s)$ with arc length parameter s. The thickness is defined as the total arc length of the streamline that traverses the CC from superior to inferior (or, alternatively, from inferior to superior) patches. Formally, we solve the following ordinary differential equation to construct the streamlines:

$$\begin{cases} u'(s) = \pm \frac{\nabla f(u(s))}{|\nabla f(u(s))|} \\ u(0) = \mathbf{x} \end{cases} \tag{3}$$

where $\mathbf{x}$ is a point on the starting surface patch and the streamline stops when it intersects the other surface patch. u' takes different sign based on the starting surface patch. Solving for Eqn. 1 using B_0 as either the superior or inferior surface, and B_1 as the other surface, we can compute the thickness at each point on superior and inferior surfaces, respectively.

3 Multivariate Morphometric Feature Computation

Alg. 1 and Fig. 1 illustrate the algorithm pipeline of our multivariate morphometry computation. In the following, we explain each step in details.

Step 1. Tetrahedral mesh and triangular mesh generation. Our meshes are generated by an adaptively sized tetrahedral mesh modeling method [10]. The method produces meshes conforming to the voxelized regions in the image by minimizing an energy function consisting of a smoothing term, a fidelity term and an elasticity term. Fig. 1(a) shows the binary image of a segmented corpus collosum and (b) shows its tetrahedral mesh. The boundary of the tetrahedral mesh gives a surface triangular mesh for the callosal surface (Fig. 1(c)).

Step 2. Surface registration and surface decomposition. The goal is to register CC surfaces and decompose them into two boundaries for thickness analysis. Given the long and thin structure of a CC surface, existing area-preserving spherical mapping based subcortical algorithms [15] may produce much distortion. For an accurate surface registration and decomposition, we adopt a holomorphic 1-form based method [6]. First, given a callosal surface, we label two consistent landmark curves at the caudal and rostral endpoints. They are biologically valid and consistent landmarks across subjects as shown in Fig. 1(d) (yellow lines). We call this process topological optimization. Given the callosal horizonal tube-like shape, these landmarks curves can be automatically detected by checking the extreme points along the first principal direction of the geometric moments of the surface. Secondly, we conformally map the callosal surface onto a rectangular planar domain with a holomorphic 1-form based conformal parameterization algorithm as in [6]. Fig. 1(e) illustrates the conformal parameterization by texture mapping the checkboard back to the surface. Finally, given two callosal surfaces S_1 and S_2 and their parameterizations, $\tau_1 : S_1 \to \mathbb{R}^2$ and $\tau_2 : S_2 \to \mathbb{R}^2$, we find a harmonic map $\tau : \mathbb{R}^2 \to \mathbb{R}^2$ between the parameter domains, such that:

$$\tau \circ \tau_1(S_1) = \tau_2(S_2), \tau \circ \tau_1(\partial S_1) = \tau_2(\partial S_2), \Delta\tau = 0.$$

Algorithm 1 Multivariate Morphometry of 3D Corpus Callosum.

Input: Binary image of segmented corpus callosum
Output: Morphometry features for each boundary vertex, including thickness and deformation tensors.
1. Build tetrahedral mesh from the binary image; build triangular mesh by computing the boundary of the tetrahedral mesh;
2. Register surfaces via holomorphic 1-form method [6]; decompose a surface into superior and inferior patches by tracing iso-parametric curves;
3. Compute callosal thickness using the harmonic field;
4. Compute deformation tensors; construct the multivariate morphometry features by combining mTBM and thickness feature.

On the registered surfaces, we generate two iso-parametric curves which pass the extreme points on two lateral sides. By cutting along these two curves and removing their attached triangles (also the tetrahedra in the tetrahedral mesh), we produce superior and inferior surface patches, which are used for the callosal thickness computation. (Fig. 1(f) shows the segmented superior and inferior surface patches (superior patch in blue and inferior patch in yellow).

Step 3. Callosal thickness computation. Similar to prior CC morphometric approaches with 2D Laplace's equations [9, 16], our approach computes the 3D Laplace's equation with an efficient discrete Laplacian operator. The details of the algorithm is described in Sec. 2. Fig. 1(g) shows how to compute the streamlines from the harmonic fields. We generate level set surfaces of the harmonic fields and the stream lines are computed by tracing their normal directions. (h) shows the color maps of the computed thickness profile on the callosal surfaces.

Step 4. Multivariate morphometry feature computation. Our complete multivariate morphometry feature consists of deformation tensors in log-Euclidean space and callosal thickness. Given two triangles, $[v_1, v_2, v_3]$ and $[w_1, w_2, w_3]$, first, we isometrically embed them onto the plane $\mathbb{R}^2$; the planar coordinates of the vertices of v_i, w_j are denoted using the same symbols v_i, w_j. Then we explicitly compute the Jacobian matrix J,

$$J = [w_3 - w_1, w_2 - w_1][v_3 - v_1, v_2 - v_1]^{-1}.$$

The deformation tensor can be defined as $S = (J^T J)^{\frac{1}{2}}$. Instead of analyzing shape change based on the eigenvalues of the deformation tensor, a new family of metrics, the "Log-Euclidean metrics" [17] is used in multivariate tensor-based morphometry (mTBM). This conversion makes computations on tensors easier to perform and statistical parameters can then be computed easily using the standard formulae for Euclidean spaces.

To compute group differences with multivariate morphometry features, we then apply Hotelling's T^2 test on sets of multivariate morphometry values. Given two groups of $n \times 1$-dimensional vectors, $S_i, i = 1, 2, , p, T_j, j = 1, 2, , q$, we use the Mahalanobis distance M to measure the group mean difference,

$$M = \frac{N_S N_T}{N_S + N_T}(\bar{S} - \bar{T})\Sigma^{-1}(\bar{S} - \bar{T})$$

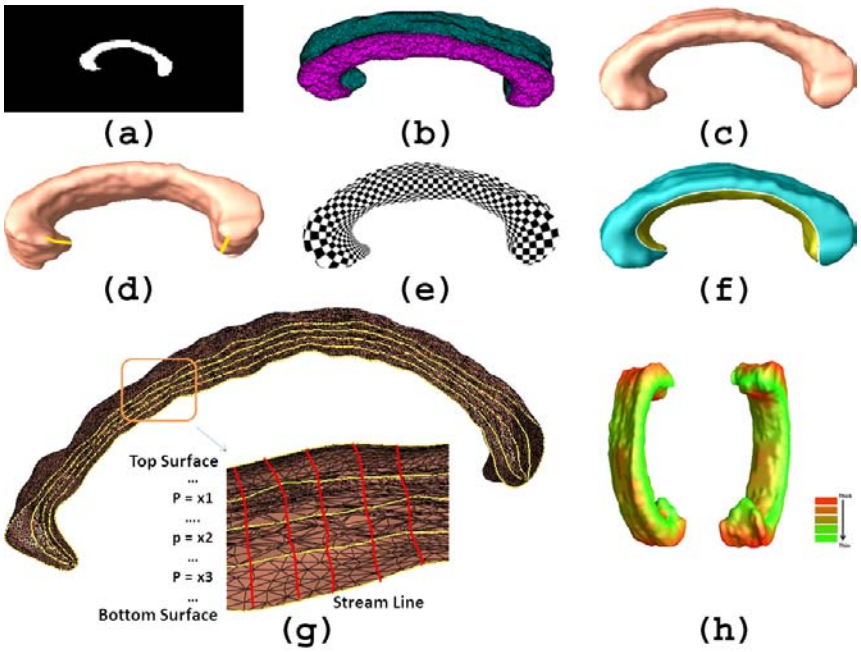


Fig. 1: Algorithm Pipeline illustrated by intermediate results.

where N_S and N_T are the number of subjects in the two groups, $\bar{S}$ and $\bar{T}$ are the means of the two groups and Σ is the combined covariance matrix of the two groups [18]. Since the statistic M is univariate, our analysis does not introduce any bias because of the increase in the number of variables.

4 Experimental Results

Our data set consists of 14 CB, 10 LB and 20 SC adult subjects scanned on a 3T MPRAGE Siemens Tim Trio MRI Scanner quality. Images were aligned and scaled to the ICBM-53 brain template (International Consortium for Brain Mapping) with the FLIRT software [19], using a 9-parameter linear transformations (3 translations, 3 rotations and 3 scales). Then we manually segmented the CCs with Insight Toolkit's SNAP program [20]. Tracings were performed in the registered template space by a trained investigator (Y.K.) and the results were checked by an experienced neuroscientist (F.L.). We consulted neuroanatomical references of the corpus callosum to help guide the placement of the contours. Fig. 2 shows some segmented results.

Then we apply Alg. 1 on the obtained binary images. Specifically, we generate tetrahedral meshes [10], compute conformal grids on their surface, register surfaces with constrained harmonic map and segment them into superior and inferior patches. We then estimate mTBM and the thickness (THK) at each vertex between the two patches using the harmonic field. Similar to the practice in [18], we also linearly covary the multivariate statistics at each pixel with subject age

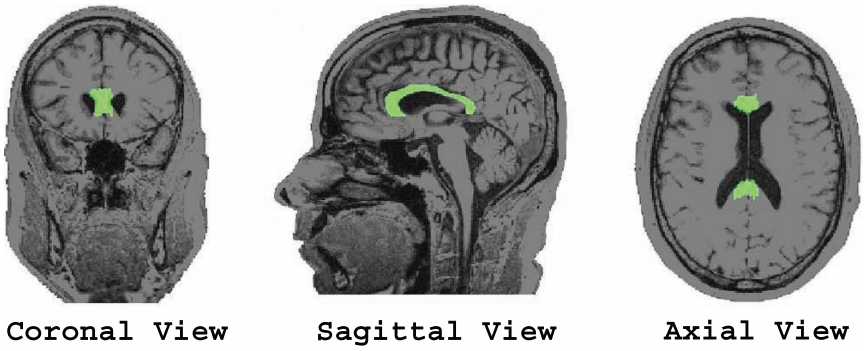


Fig. 2: Manually segmented callosal structure overlayed on an MRI image.

and gender information. The covaried statistics are used for a group difference study. For the group difference test, we run a permutation test with 5000 random assignments of subjects to groups to estimate the statistical significance (p-maps) in surface morphometry [6].

Fig. 3 shows uncorrected p-maps for three group difference studies. To explore whether our multivariate statistics provide extra power when combining thickness with mTBM, in each experiment, we also conducted three additional statistical tests using the thickness and different tensor-based statistics derived from the Jacobian matrix. The other statistics we studied are: (1) the thickness (THK) itself; (2) the determinant of Jacobian matrix; and (3) the mTBM. For statistics (1) and (2), we applied a Student's t-test to compute the group mean difference at each surface point. In case (3) and for our new combined measure, we used Hotelling's T^2 statistics to compute the group mean difference. In all sets of results, we detected significant areas around splenium areas for the combined measure. The CB also show significant changes in the body of the CC.

All group difference p-maps were corrected for multiple comparisons using the false discovery rate method (FDR) [21]. The FDR method determines the *critical p-value*, which is the highest threshold p-value that controls the FDR at the given threshold, e.g. 5%. To rank which clinical measures were most strongly associated with callosal morphology, we created cumulative distribution function (CDF) plots of the resulting uncorrected p-values. The critical p-value, which is the highest non-zero point at which the CDF plot intersects the $y = 20x$ line, represents the highest statistical threshold for which at most 5% false positive are expected in the map. If there is no such intersection point (other than the origin), there is no evidence to reject the null hypothesis. Also, steeper CDFs show stronger effect sizes. FDR results are shown in Fig. 4. All measures are significant for the CB vs. controls, while only our new combined measure falls above the $y = 20x$ line in the case of LB vs. controls.

Note that our results are consistent with previous work [7] and with the hypothesis that splenium regions should be affected in all blind groups, but more so in the CB. Our results are also consistent with another DTI tractography study [22], which found fractional anisotropy was significantly reduced in the

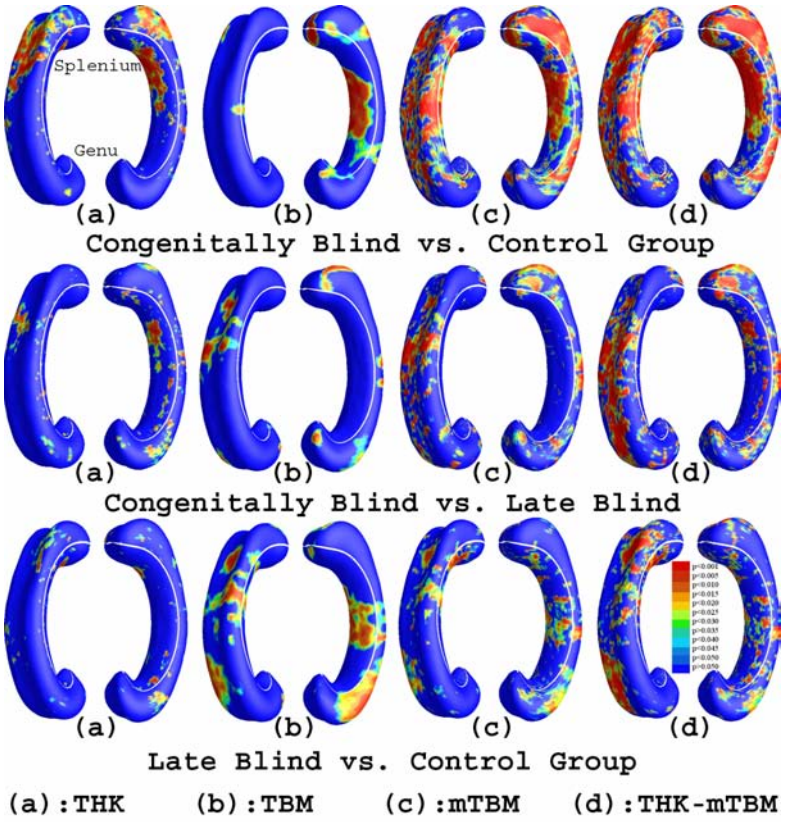


Fig. 3: Comparison of p-maps with four statistics on 3 different group difference studies. Non-blue colors show vertices with statistical differences, at the 0.05 level, uncorrected. The combined multivariate statistics outperformed all three individual statistics (the critical p-values for these maps are shown in Fig. 4).

splenium of CB subjects. The splenium is primarily composed of fibers connecting the visuo-spatial areas of the brain. The differences seen here may be due to reduced myelination of these fibers in the absence of visual input.

In addition, with our novel multivariate statistics, we found changes in the body of the CC in the LB group while no such difference was detected in [7]. In the late blind subjects, the process of myelination is relatively advanced, so that the structure of the corpus callosum may not be that strongly influenced by the loss of visual perception. Our new discovery, generally consistent with this understanding, may provide additional insights to the myelination and cortical plasticity process. More importantly, these results also suggest that the newly proposed multivariate morphometry has more detection power in terms of effect size, likely because it captures callosal thickness and more directional and rotational information when measuring geometric differences.

In future, we will combine and correlate our multivariate statistical framework with other MRI imaging systems, such as cortical morphometry and diffu-

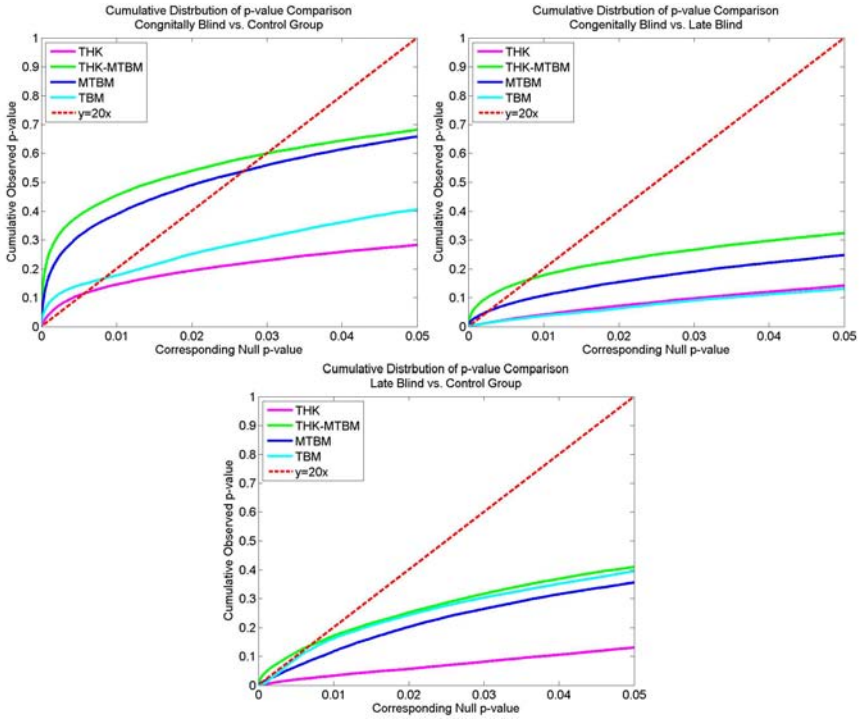


Fig. 4: The cumulative distributions of the p-values for difference detected between three diagnostic groups (CB, LB and SC) for all four statistics. The critical p-values are the intersection points of the curves and the $y = 20x$ line. The new multivariate statistics achieved the highest critical p-values in all 3 comparisons.

sion tensor imaging (DTI) tractography, to advance our understanding of blindness and improve the effectiveness of rehabilitation techniques and life quality for blind individuals.

References

1. Hofer, S., Frahm, J.: Topography of the human corpus callosum revisited–comprehensive fiber tractography using diffusion tensor magnetic resonance imaging. Neuroimage **32**(3) (Sep 2006) 989–994
2. Hua, X., Leow, A.D., Levitt, J.G., Caplan, R., Thompson, P.M., Toga, A.W.: Detecting brain growth patterns in normal children using tensor-based morphometry. Hum Brain Mapp **30**(1) (Jan 2009) 209–219
3. Wang, Y., Panigrahy, A., Shi, J., Ceschin, R., Nie, Z., Nelson, M.D., Leporé, N.: 3D vs. 2D surface shape analysis of the corpus callosum in premature neonates. In: MICCAI workshop on Paediatric and Perinatal Imaging, Nice, France (2012)
4. Davatzikos, C.: Spatial normalization of 3D brain images using deformable models. J Comput Assist Tomogr **20**(4) (1996) 656–665
5. Chung, M.K., Worsley, K.J., Paus, T., Cherif, C., Collins, D.L., Giedd, J.N., Rapoport, J.L., Evans, A.C.: A unified statistical approach to deformation-based morphometry. Neuroimage **14**(3) (Sep 2001) 595–606

6. Wang, Y., Song, Y., Rajagopalan, P., An, T., Liu, K., Chou, Y.Y., Gutman, B., Toga, A.W., Thompson, P.M.: Surface-based TBM boosts power to detect disease effects on the brain: An N=804 ADNI study. Neuroimage **56**(4) (2011) 1993–2010
7. Leporé, N., Voss, P., Lepore, F., Chou, Y.Y., Fortin, M., Gougoux, F., Lee, A.D., Brun, C., Lassonde, M., Madsen, S.K., Toga, A.W., Thompson, P.M.: Brain structure changes visualized in early- and late-onset blind subjects. Neuroimage **49**(1) (Jan 2010) 134–140
8. Jones, S.E., Buchbinder, B.R., Aharon, I.: Three-dimensional mapping of cortical thickness using Laplace's equation. Hum Brain Mapp **11**(1) (Sep 2000) 12–32
9. Adamson, C.L., Wood, A.G., Chen, J., Barton, S., Reutens, D.C., Pantelis, C., Velakoulis, D., Walterfang, M.: Thickness profile generation for the corpus callosum using Laplace's equation. Hum Brain Mapp **32**(12) (Dec 2011) 2131–2140
10. Lederman, C., Joshi, A., Dinov, I., Vese, L., Toga, A., Van Horn, J.D.: The generation of tetrahedral mesh models for neuroanatomical MRI. Neuroimage **55**(1) (Mar 2011) 153–164
11. Wang, Y., Gu, X., Chan, T.F., Thompson, P.M., Yau, S.T.: Volumetric harmonic brain mapping. In: Biomedical Imaging: From Nano to Macro, 2004. ISBI 2004. IEEE International Symposium on. (Apr. 2004) 1275–1278
12. Schmitt, O., Bohme, M.: A robust transcortical profile scanner for generating 2-D traverses in histological sections of richly curved cortical courses. Neuroimage **16**(4) (Aug 2002) 1103–1119
13. Yezzi, A.J., Prince, J.L.: An Eulerian PDE approach for computing tissue thickness. IEEE Trans Med Imaging **22**(10) (Oct 2003) 1332–1339
14. Hutton, C., De Vita, E., Ashburner, J., Deichmann, R., Turner, R.: Voxel-based cortical thickness measurements in MRI. Neuroimage **40**(4) (May 2008) 1701–1710
15. Styner, M., Lieberman, J.A., McClure, R.K., Weinberger, D.R., Jones, D.W., Gerig, G.: Morphometric analysis of lateral ventricles in schizophrenia and healthy controls regarding genetic and disease-specific factors. Proc. Natl. Acad. Sci. U. S. A. **102**(13) (2005) 4872–4877
16. Herron, T.J., Kang, X., Woods, D.L.: Automated measurement of the human corpus callosum using MRI. Front Neuroinform **6** (2012) 25
17. Arsigny, V., Fillard, P., Pennec, X., Ayache, N.: Log-Euclidean metrics for fast and simple calculus on diffusion tensors. Magn. Reson. Med. **56**(2) (2006) 411–421
18. Leporé, N., Brun, C., Chou, Y.Y., Chiang, M.C., Dutton, R.A., Hayashi, K.M., Luders, E., Lopez, O.L., Aizenstein, H.J., Toga, A.W., Becker, J.T., Thompson, P.M.: Generalized tensor-based morphometry of HIV/AIDS using multivariate statistics on deformation tensors. IEEE Trans. Med. Imag. **27**(1) (2008) 129–141
19. Jenkinson, M., Smith, S.: A global optimisation method for robust affine registration of brain images. Med Image Anal **5**(2) (Jun 2001) 143–156
20. Yushkevich, P.A., Piven, J., Hazlett, H.C., Smith, R.G., Ho, S., Gee, J.C., Gerig, G.: User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage **31**(3) (Jul 2006) 1116–1128
21. Benjamini, Y., Hochberg, Y.: Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society. Series B (Methodological) **57**(1) (1995) 289–300
22. Yu, C., Shu, N., Li, J., Qin, W., Jiang, T., Li, K.: Plasticity of the corticospinal tract in early blindness revealed by quantitative analysis of fractional anisotropy based on diffusion tensor tractography. Neuroimage **36**(2) (Jun 2007) 411–417

A right-invariant Riemannian distance on $\mathrm{GL}^+(n)$ and hypothesis testing on Jacobian matrices.

Ernesto Zacur, Matias Bossa, and Salvador Olmos, for the Alzheimer's Disease Neuroimaging Initiative**

GTC, I3A, Universidad de Zaragoza
Maria de Luna 1, 50018 Zaragoza, Spain
{zacur,bossa,olmos}@unizar.es

Abstract. Tensor-based morphometry (TBM) studies encode the anatomical information in spatial deformations, which are locally characterized by Jacobian matrices. Current methods perform statistical analysis on incomplete features of the Jacobian matrices, such as their determinants or the Cauchy-Green deformation tensor. In this work we propose the use of a right-invariant Riemannian distance on $\mathrm{GL}^+(n)$, providing more information about the local deformation than previous approaches.

1 Introduction

Tensor-based morphometry (TBM) is a methodology to analyze anatomical information encoded by the spatial transformations that map a reference template to a set of images. The spatial mappings are estimated by means of non-rigid registration. Afterwards, voxel-wise statistical analysis is performed on the spatial derivatives of the deformations (Jacobian matrices, hereinafter denoted by $\mathbf{J}(\mathbf{x})$ for each location $\mathbf{x}$ of the template domain).

The simplest and still most widely used approach for TBM is based on the Jacobian determinant. This feature has two main advantages. Firstly, it has an intuitive interpretation because it represents the local volume change. Secondly, standard univariate statistical analysis can be easily performed on Jacobian determinants (or their logarithms) [1, 2]. The main limitation of the Jacobian determinant is that it only provides a coarse description of the deformation, because it only quantifies local volume change. To overcome this limitation, a multivariate TBM has been proposed based on the Cauchy-Green deformation tensor, $\mathbf{C} = \mathbf{J}^T\mathbf{J}$ which provides more complete description of the local deformation.

Most of the available statistical tools are well defined for Euclidean data. However, so far there are no many statistical tools for manifold-valued data,

** Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.ucla.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.ucla.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

such as Cauchy-Green deformation tensors or Jacobian matrices. Statistics on Lie groups were analyzed in [3, 4] defining principal geodesic analysis. Analysis of diffusion tensor images (DTI) promoted the development of statistical tools on symmetric positive definite (SPD) tensors [5–8]. Some recent works deal with principal geodesics analysis using both, intrinsic statistics and the tangent plane approximation [9] and regression on manifold data [10–12]. A multivariate extension of the Hotelling's T^2 test has been proposed for analyzing longitudinal data [13].

In these previous works the definition of an appropriate distance between manifold elements plays an important role. The statistical tools for TBM studies use Jacobian matrices as input data which belong to the general linear group $\mathrm{GL}^+(n)$. A contribution of this work is the use of distances between Jacobian matrices for TBM studies. A further requirement for a TBM study is that statistical results should be independent of the template choice. Under a transitive non-rigid registration process, we will show that the use of right-invariant distances provides a sufficient condition for holding the template invariance requirement.

The aim of this paper is twofold. Firstly, to introduce a distance between Jacobian matrices for TBM studies. This is formulated as a right-invariant Riemannian distance on $\mathrm{GL}^+(n)$. The second aim is to illustrate results on a synthetic spatial deformation study and a MRI brain image study using three different right-invariant distances: d_{DET} which is a distance based on Jacobian determinants; d_{AFF} which is based on Cauchy-Green deformation tensors; and d_{RI} which is a Riemannian metric on $\mathrm{GL}^+(n)$. Voxel-wise hypothesis testing is performed by means of the Cramér test [14, 15] for several reasons: it is computed from the set of distances between observations solely and it can be used for both univariate and multivariate data.

2 Background on univariate/multivariate TBM

Let $\boldsymbol{\Phi} = [\phi^1, \phi^2, \cdots, \phi^n]^T \colon \Omega \to \Upsilon$ be an invertible, orientation preserving and differentiable spatial mapping (a diffeomorphism), where Ω and Υ are simply connected subsets of $\mathbb{R}^n$. Up to first order $\boldsymbol{\Phi}(\mathbf{x}+\mathrm{d}\mathbf{x}) = \boldsymbol{\Phi}(\mathbf{x}) + \mathbf{J}(\mathbf{x})\mathrm{d}\mathbf{x} + O(\mathrm{d}\mathbf{x}^2)$, where $\mathbf{J}$ is a field of linear transformations belonging to $\mathrm{GL}^+(n)$ ($\mathbf{J} \colon \Omega \to \mathrm{GL}^+(n)$). Every element $\mathbf{J}(\mathbf{x})$ is an $n \times n$ matrix with positive determinant. The set of these matrices together with the matrix-matrix multiplication is a matrix group and therefore a Lie group [16]. The n^2 elements of the matrix $\mathbf{J}(\mathbf{x}) = (\mathcal{D}\boldsymbol{\Phi})\big|_{\mathbf{x}}$ are the spatial derivatives of the deformation, such that the element $\mathbf{J}^{i,j}(\mathbf{x}) = \partial_j \phi^i\big|_{\mathbf{x}}$, *i.e.* the derivative of the i-th component of $\boldsymbol{\Phi}$ along the j-th coordinate.

Let T and A be two images representing similar contents. Registration of T and A is formulated as finding the spatial transformation $\boldsymbol{\Phi} \colon \mathrm{dom}(\mathsf{T}) \to \mathrm{dom}(\mathsf{A})$ such that $\boldsymbol{\Phi} \star \mathsf{T} \sim \mathsf{A}$, where $\star$ denotes the action on the image $(\boldsymbol{\Phi} \star \mathsf{T})(\mathbf{x}) = \mathsf{T}(\boldsymbol{\Phi}^{-1}(\mathbf{x}))$, and $(\cdot \sim \cdot)$ refers to an equivalence relation defined typically by a matching energy. Once the matching is found, both, the target image A and the

deformed image $\boldsymbol{\Phi} \star \mathsf{T}$ look spatially similar and a correspondence is obtained between all points of the domains dom(T) and dom(A).

In order to asses statistical differences between two groups of images $\mathcal{A} = \{\mathsf{A}_a\}$ and $\mathcal{B} = \{\mathsf{A}_b\}$ at each location a hypothesis test is performed on the corresponding Jacobian matrices. Just to fix notation, let $\boldsymbol{\Phi}_\mu$ be the mapping registering the template T to the image A_μ such that $\boldsymbol{\Phi}_\mu \star \mathsf{T} \sim \mathsf{A}_\mu$, and let $\mathbf{J}_\mu(\mathbf{x})$ be its corresponding Jacobian matrix at the point $\mathbf{x} \in \text{dom}(\mathsf{T})$. From now on, the subindexes a or a' run over the elements of the set $\mathcal{A}$, b or b' runs over the set $\mathcal{B}$, and, in general, indexes μ and ν run over any instance.

In order to perform statistics on Jacobian matrices an appropiate distance function $d(\cdot,\cdot)\colon \text{GL}^+(n) \times \text{GL}^+(n) \to \mathbb{R}^+$ must be employed.

2.1 Invariance with respect to the template

The template define the anatomical coordinates where to perform the statistical analysis. It is desired and expected that the methodology provides mainly the same result for any template choice allowing to make general statements about the anatomical location of the findings. Let T and W be two possible templates. Let $\boldsymbol{\Psi}$ be the mapping relating W and T such that, $\boldsymbol{\Psi} \star \mathsf{W} \sim \mathsf{T}$. Every point $\mathbf{y} \in \text{dom}(\mathsf{W})$ is in correspondence with a point $\mathbf{x} = \boldsymbol{\Psi}(\mathbf{y}) \in \text{dom}(\mathsf{T})$. Let $\tilde{\boldsymbol{\Phi}}_\mu$ be the spatial deformation to register W to A_μ. In this work it is assumed that the registration process is transitive [17, 18] and therefore $\tilde{\boldsymbol{\Phi}}_\mu = \boldsymbol{\Phi}_\mu \circ \boldsymbol{\Psi}$. Although this assumption is rarely achieved by current registration methods it is a key ingredient to describe the findings in anatomical coordinates irrespectively of the selected template.

The Jacobian matrix field derived from $\tilde{\boldsymbol{\Phi}}_\mu$ is given by $\tilde{\mathbf{J}}_\mu(\mathbf{y}) = \mathbf{J}_\mu(\boldsymbol{\Psi}(\mathbf{y}))\mathbf{P}(\mathbf{y})$, where $\mathbf{P}(\mathbf{y}) = (\mathcal{D}\boldsymbol{\Psi})|_{\mathbf{y}}$. For a statistic based on distances, a sufficient condition to achieve template invariance is to use a distance which fulfills

$$d(\tilde{\mathbf{J}}_\mu(\mathbf{y}), \tilde{\mathbf{J}}_\nu(\mathbf{y})) = d(\mathbf{J}_\mu(\mathbf{x})\mathbf{P}(\mathbf{y}), \mathbf{J}_\nu(\mathbf{x})\mathbf{P}(\mathbf{y})) = d(\mathbf{J}_\mu(\mathbf{x}), \mathbf{J}_\nu(\mathbf{x}))$$

which is accomplished by a right-invariant distance, *i.e.*

$$d(\mathbf{J}_\mu\mathbf{P}, \mathbf{J}_\nu\mathbf{P}) = d(\mathbf{J}_\mu, \mathbf{J}_\nu) \tag{1}$$

for any $\mathbf{J}_\mu, \mathbf{J}_\nu, \mathbf{P} \in \text{GL}^+(n)$.

2.2 Jacobian determinant

The most widely used tests to assess group differences are based on the determinant of Jacobian matrices. These determinants belong to the group of positive numbers under multiplication. This is a Lie group and an invariant distance is

$$d_{DET}(\mathbf{J}_\mu, \mathbf{J}_\nu) = |\log(\det(\mathbf{J}_\mu)) - \log(\det(\mathbf{J}_\nu))|\,. \tag{2}$$

The determinant of a Jacobian matrix can only quantify local volume changes. Note that d_{DET} does not satisfy the coincidence axiom which can be seen by the fact that $d_{DET}(\mathbf{J}, \mathbf{LJ}) = d_{DET}(\mathbf{J}, \mathbf{JL}) = 0$ for any matrix $\mathbf{L}$ with $\det(\mathbf{L}) = 1$.

2.3 Deformation tensor

In the area of continuum mechanics a commonly used feature to measure a local deformation at the point $\mathbf{x}$ is the called Cauchy-Green deformation tensor [19] $\mathbf{C}(\mathbf{x}) = \mathbf{J}^T(\mathbf{x})\mathbf{J}(\mathbf{x})$. The tensor $\mathbf{C}$ is a symmetric positive definite (SPD) matrix and it is related to the amount of anisotropic deformation up to a rotation. Under a change of template, Jacobian matrices transform as $\mathbf{J} \mapsto \mathbf{JP}$ and therefore Cauchy-Green deformation tensors transform as $\mathbf{C} \mapsto \mathbf{P}^T\mathbf{CP}$.

In [5, 6] a distance between SPD matrices was proposed[1]: $d_{SPD}(\mathbf{C}_\mu, \mathbf{C}_\nu) = \left\| \operatorname{logm}\left((\mathbf{C}_\mu)^{-1/2}(\mathbf{C}_\nu)(\mathbf{C}_\mu)^{-1/2}\right)\right\|_F$ and for for any $\mathbf{P}$ with positive determinant satisfies $d_{SPD}(\mathbf{C}_\mu, \mathbf{C}_\nu) = d_{SPD}(\mathbf{P}^T\mathbf{C}_\mu\mathbf{P}, \mathbf{P}^T\mathbf{C}_\nu\mathbf{P})$. The distance d_{SPD} defines a distance between Jacobian matrices

$$d_{AFF}(\mathbf{J}_\mu, \mathbf{J}_\nu) = \left\| \operatorname{logm}\left((\mathbf{J}_\mu^T\mathbf{J}_\mu)^{-1/2}(\mathbf{J}_\nu^T\mathbf{J}_\nu)(\mathbf{J}_\mu^T\mathbf{J}_\mu)^{-1/2}\right)\right\|_F \tag{3}$$

satisfying $d_{AFF}(\mathbf{J}_\mu, \mathbf{J}_\nu) = d_{AFF}(\mathbf{J}_\mu\mathbf{P}, \mathbf{J}_\nu\mathbf{P})$ and therefore fulfills the template invariance requirement.

Likewise d_{DET}, the distance d_{AFF} does not satisfy the coincidence axiom, specifically $d_{AFF}(\mathbf{J}, \mathbf{RJ}) = 0$ for any rotation $\mathbf{R}$.

3 A metric on $\mathbf{GL}^+(n)$

The two previous distances fulfill the template invariance requirement. However they do not satisfy the coincidence axiom and accordingly some differences may not be measured. To overcome this drawback a right-invariant Riemannian metric on the space of Jacobian matrices is presented below.

3.1 Invariant Riemannian distances on Lie groups

Let $\mathcal{M}$ be a differentiable manifold and $T_z\mathcal{M}$ its tangent space at the element $z \in \mathcal{M}$. A Riemannian metric $(\mathcal{M}, \langle u, v\rangle_z)$ on $\mathcal{M}$ is a smooth assignment of inner products to every tangent space where $z \in \mathcal{M}$ and $u, v \in T_z\mathcal{M}$ [20]. Using this assignment, the length of a curve segment $\gamma : [t_0, t_1] \subset \mathbb{R} \to \mathcal{M}$ is defined as

$$\operatorname{Length}(\gamma; t_0, t_1) = \int_{t_0}^{t_1} \langle \dot\gamma_{(s)}, \dot\gamma_{(s)}\rangle_{\gamma(s)}^{1/2} \mathrm{d}s\,.$$

A geodesic segment between two elements z and w belonging to $\mathcal{M}$ is an arc-length parameterized curve segment which locally minimizes the length. The Riemannian distance between z and w is the length of the shortest geodesic segment connecting both elements.

Geodesics can be uniquely described by an initial point and an initial velocity. This description is related to the Riemannian exponential function $\operatorname{Exp}_z(v)$

[1] $\operatorname{logm}(\cdot)$ denotes the matrix logarithm (the inverse of the matrix exponential, denoted as $\operatorname{expm}(\cdot)$) and $\|\cdot\|_F$ is the Frobenius norm.

where $v \in T_z\mathcal{M}$ [21]. The curve generated by $\mathrm{Exp}_z(tv)$ is a geodesic and the length of the segment from $t = 0$ to $t = 1$ is equal to $\langle v, v\rangle_z^{1/2}$.

The set $\mathrm{GL}^+(n)$ of all $n \times n$ matrices with positive determinant can be endowed with a differentiable manifold structure [16]. Furthermore, it has a Lie group structure (defined by the matrix-matrix product) and their elements, curves, tangent spaces and velocities can be translated by its group action [22].

A right-invariant metric is a Riemannian metric which naturally arises in Lie groups. Under this metric geodesics and distances remain invariant under right-translations. The metric can be defined as an inner product at a single tangent space (usually at $T_I\mathcal{M}$, called the group algebra) and propagated from right to the whole group

$$\langle u, v\rangle_z = \langle T_z\boldsymbol{R}_y u, T_z\boldsymbol{R}_y v\rangle_{zy}\,,$$

where $T_z\boldsymbol{R}_y$ is the tangent lift of the $\boldsymbol{R}_y$ operator and right-translates a velocity from $T_z\mathcal{M}$ to $T_{zy}\mathcal{M}$. Additionally the following invariance property of the Riemannian exponential function holds:

$$\mathrm{Exp}_z(v) = \mathrm{Exp}_I(T_z\boldsymbol{R}_{z^{-1}}v)z\,. \tag{4}$$

For a matrix group the tangent lift of the right translation takes the form $T\boldsymbol{R}_z v \mapsto \mathbf{VZ}$, where $\mathbf{V}$ and $\mathbf{Z}$ are a matrix representation of v and z.

3.2 Riemannian right-invariant distance on $\mathrm{GL}^+(n)$

A closed form solution for the invariant Riemannian exponential on $\mathrm{GL}^+(n)$ for the case where $\langle \mathbf{U}_1, \mathbf{U}_2\rangle_{\mathbf{I}} = \mathrm{trace}(\mathbf{U}_1^T\mathbf{U}_2)$, being $\mathbf{U}_1$ and $\mathbf{U}_2$ elements of $T_{\mathbf{I}}\mathrm{GL}^+(n)$, was given in[2] [23]:

$$\begin{aligned}\mathrm{Exp}_{\mathbf{Q}}(\mathbf{V}) &= \mathrm{Exp}_{\mathbf{I}}(\mathbf{U})\mathbf{Q} \\ &= \mathrm{expm}\left(\mathbf{U} - \mathbf{U}^T\right)\mathrm{expm}\left(\mathbf{U}^T\right)\mathbf{Q}\,,\end{aligned} \tag{5}$$

where $\mathbf{V} \in T_{\mathbf{Q}}\mathrm{GL}^+(n)$ and $\mathbf{U} = \mathbf{VQ}^{-1} \in T_{\mathbf{I}}\mathrm{GL}^+(n)$.

From the right-invariant Riemannian metric on $\mathrm{GL}^+(n)$ the following distance is induced:

$$d_{RI}(\mathbf{J}_\mu, \mathbf{J}_\nu) = \langle \mathbf{U}^*, \mathbf{U}^*\rangle_{\mathbf{J}_\mu}^{1/2}\,,$$

where $\mathbf{U}^*$ is the smallest initial velocity satisfying $\mathrm{Exp}_{\mathbf{I}}(\mathbf{U})\mathbf{J}_\mu = \mathbf{J}_\nu$. The distance inherits the right-invariance from the metric, $d_{RI}(\mathbf{J}_\mu, \mathbf{J}_\nu) = d_{RI}(\mathbf{J}_\mu\mathbf{P}, \mathbf{J}_\nu\mathbf{P})$ for any $\mathbf{P}$ in $\mathrm{GL}^+(n)$, and therefore the invariance under the template holds. In addition, d_{RI} fulfills the coincidence axiom, *i.e.* $d_{RI}(\mathbf{J}_\mu, \mathbf{J}_\nu) = 0$ if and only if $\mathbf{J}_\mu = \mathbf{J}_\nu$.

[2] Actually, in [23] the closed-form solution is given for a left-invariant Riemannian metric, however using the identity $\mathrm{Exp}_{\mathbf{I}}^{right}(\mathbf{U}) = (\mathrm{Exp}_{\mathbf{I}}^{left}(-\mathbf{U}))^{-1}$ it is possible to express in closed-form its right-invariant version.

In this work it is proposed to compute the Riemannian distance solving the following problem:

$$\underset{\mathbf{U}\in T_{\mathbf{I}}\mathrm{GL}^+(n)}{\text{minimize}} \quad \left\|\mathrm{Exp}_{\mathbf{I}}(\mathbf{U})\mathbf{J}_\mu - \mathbf{J}_\nu\right\|_F^2 \quad , \tag{L1}$$

where, for simplicity, the Frobenius norm was chosen.

The Riemannian exponential function under a right-invariant metric is a surjective mapping to the identity component of the group. As the group $\mathrm{GL}^+(n)$ consists of only one connected component, then the distance is well defined for any pair of elements.

Using the right-invariance property in eq. (4), it is possible to reformulate the problem (L1) as

$$\underset{\mathbf{U}\in\mathfrak{gl}(n)}{\text{minimize}} \quad \left\|\mathrm{Exp}_{\mathbf{I}}(\mathbf{U}) - \mathbf{J}_\nu\mathbf{J}_\mu^{-1}\right\|_F^2 \quad , \tag{L2}$$

where $\mathfrak{gl}(n) \equiv \mathbb{R}^{n\times n}$ is the Lie algebra of $\mathrm{GL}^+(n)$.

A gradient descent procedure can be used to solve (L2) where the gradient with respect to $\mathbf{U}$ is needed. Let $\mathbf{Q} \in \mathrm{GL}^+(n)$ and $E(\mathbf{U};\mathbf{Q}) = \left\|\mathrm{Exp}_{\mathbf{I}}(\mathbf{U}) - \mathbf{Q}\right\|_F^2 = \left\|\mathrm{expm}(\mathbf{U}-\mathbf{U}^T)\mathrm{expm}(\mathbf{U}^T) - \mathbf{Q}\right\|_F^2$ be the objective function. Its derivative with respect to $\mathbf{U}$ results in

$$\mathcal{D}_{\mathbf{U}}E(\mathbf{U};\mathbf{Q}) = 2\left(\overline{\mathrm{Exp}_{\mathbf{I}}(\mathbf{U})} - \overline{\mathbf{Q}}\right)^T \mathcal{D}_{\mathbf{U}}\mathrm{Exp}_{\mathbf{I}}(\mathbf{U})\,, \tag{6}$$

where $\overline{\mathbf{M}}$ denotes the n^2 dimensional vector resulting by the stacking of the columns of $\mathbf{M}$. $\mathcal{D}$ is the Fréchet derivative operator: for a matrix function of matrix argument $\mathbf{F}\colon \mathbb{R}^{m\times n} \to \mathbb{R}^{p\times q}$, $\mathcal{D}_{\mathbf{M}}\mathbf{F}(\mathbf{M})$ is a $(pq)\times(mn)$ linear operator fulfilling $[\mathcal{D}_{\mathbf{M}}\mathbf{F}(\mathbf{M})]^{i,j} = \partial_{\overline{\mathbf{M}}^j}\overline{\mathbf{F}(\mathbf{M})}^i$. Some rules to compute matrix functions derivatives are in [24, Chapter 13].

Deriving the expression in eq. (5) it is obtained

$$\begin{aligned}\mathcal{D}_{\mathbf{U}}\mathrm{Exp}_{\mathbf{I}}(\mathbf{U}) = &\left(\mathbf{I}_n \otimes \mathrm{expm}(\mathbf{U}-\mathbf{U}^T)\right)\mathrm{dexpm}(\mathbf{U}^T)\mathbf{K}_{nn} + \\ &+ \left(\mathrm{expm}(\mathbf{U})\otimes\mathbf{I}_n\right)\mathrm{dexpm}(\mathbf{U}-\mathbf{U}^T)\left(\mathbf{I}_{n^2} - \mathbf{K}_{nn}\right)\end{aligned} \tag{7}$$

where $\mathbf{I}_m$ is the $m\times m$ identity matrix, $\mathbf{K}_{nn}$ is the commutation matrix [24] defined by $\mathbf{K}_{nn}\overline{\mathbf{X}} = \overline{\mathbf{X}^T}$ for an $n\times n$ matrix $\mathbf{X}$ and $\mathrm{dexpm}(\cdot)$ is the Fréchet derivative of the matrix exponential function (see Appendix A).

The optimization problem to compute d_{RI} is presented as an iterative descent optimization with a backtracking line-search along the direction $-\mathcal{D}_{\mathbf{U}}E(\mathbf{U};\mathbf{Q})$ given by eq. (6).

Once the optimal initial velocity $\mathbf{U}^*$ is obtained, the length of the geodesic segment is $\langle\mathbf{U}^*,\mathbf{U}^*\rangle_{\mathbf{I}}^{1/2} = \left\|\mathbf{U}^*\right\|_F$. If its Riemannian exponentiation generate the shortest curve segment between $\mathbf{I}$ and $\mathbf{J}_\nu\mathbf{J}_\mu^{-1}$, then $d_{RI}(\mathbf{J}_\mu,\mathbf{J}_\nu) = d_{RI}(\mathbf{I},\mathbf{J}_\nu\mathbf{J}_\mu^{-1}) = \left\|\mathbf{U}\right\|_F$.

While the existence of a zero minimizer is guaranteed, a drawback of this formulation is the non-uniqueness of a velocity satisfying $\mathrm{Exp}_{\mathbf{I}}(\mathbf{U}) = \mathbf{Q}$. There may exist different initial velocities which generate geodesics segments between $\mathbf{J}_\mu$ and $\mathbf{J}_\nu$.

4 Cramér test

Three different distance functions over Jacobian matrices have been considered in previous sections: d_{DET}, d_{AFF} and d_{RI}. Even though many statistical analysis tools can be defined from the set of distances among observations, this work focused on the Cramér two-sample test [14, 15]. This test was selected because its statistic depends only on the inter-point distances

$$\sigma(\{D_{\mu\nu}\}) = \frac{n_a n_b}{n_a + n_b}\left(\frac{1}{n_a n_b}\sum_{a=1}^{n_a}\sum_{b=1}^{n_b} D_{ab} - \frac{1}{2n_a^2}\sum_{a=1}^{n_a}\sum_{a'=1}^{n_a} D_{aa'} - \frac{1}{2n_b^2}\sum_{b=1}^{n_b}\sum_{b'=1}^{n_b} D_{b'b}\right)$$

where $D_{\mu\nu} = d(\mathbf{J}_\mu, \mathbf{J}_\nu)$ is the inter-element distance and n_a and n_b is the cardinality of the set $\mathcal{A}$ and $\mathcal{B}$ respectively. Accordingly, this test can be directly used on both univariate and multivariate data.

Statistical significance can be assessed by means of random permutation tests. The distribution of the statistic under the null hypothesis is empirically estimated from random permutations of the group labels. The null hypothesis is that both groups instances are drawn from the same distribution and it is rejected for large enough values of σ. The p-value is assessed as the proportion of the permutations having a σ value larger or equal than the σ value without relabeling.

In TBM studies hypothesis tests are assessed at each image location and therefore multiple comparison correction is an important issue. In this work the *false discovery rate* (FDR) [25] criterion was used for the correction of the proportion of false-positives among the rejected null hypotheses.

5 Results

5.1 Synthetic study

A synthetic study was designed to generate a set of controlled deformations which will allow to illustrate the results of the three right-invariant distances in a TBM application environment.

In order to facilitate the visualization of the results the experiment was performed generating deformations on a 2D image. Two sets of 50 deformation fields were generated to deform a template. The first set was designed to represent the anatomical variability within a 'control' group, while the second set aimed at representing a 'pathological' group with group-driven anatomical differences and intra-group variability. Representative examples of the deformation maps

are illustrated in Fig. 1. Fig. 1 also shows 5 deformed contours from 'control' and 'pathological' groups depicting the intra-group variabilities.

Deformation instances for 'control' group $\mathcal{A}$ were modeled as smooth and invertible random deformation fields which were obtained by integration of a spatially correlated stationary velocity field. Instances for 'pathological' group $\mathcal{B}$ were modeled as the composition of a common group-driven deformation and random deformations similar to ones used for the 'control' group. The group-driven deformation was designed as an invertible mapping producing the following changes on the 'subcortical structures': a clockwise rotation of the structure in the 'left hemisphere' of 15 degrees; a counter-clockwise rotation of the structure in the 'right hemisphere' of 15 degrees and a subsequent anisotropical scaling with factors 0.7 and $1/0.7$ along the horizontal and vertical directions respectively. Note that the group-driven deformation preserves the volume of the 'subcortical structures' but the surrounding regions suffer more complex deformation due to the continuity of the deformations. This effect can be seen in the zoomed panel of Fig. 1.

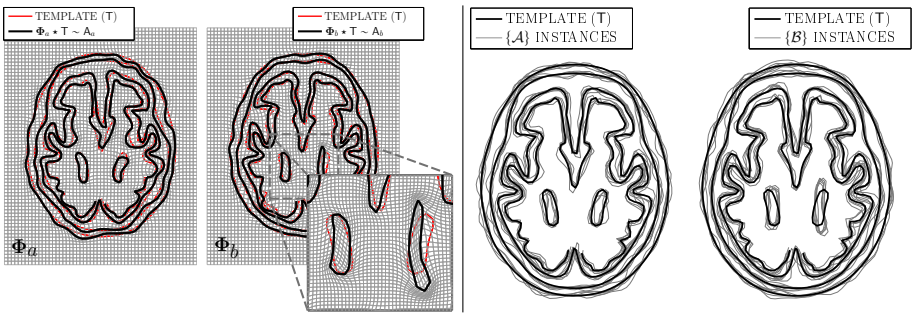


Fig. 1. Synthetic data for 'control' group and 'pathological' group. Left: illustrative examples of the deformations map $\boldsymbol{\Phi}_a$ and $\boldsymbol{\Phi}_b$ mapping the template to a given instance. Right: five contours depicting the intra-group anatomical variability.

TBM studies were performed using the previous defined distances: d_{DET}, d_{AFF} and d_{RI}. At each location Cramér tests were computed and a set of 100,000 random permutations were performed in order to assess the p-value maps. In order to correct for multiple comparison, FDR criterion was used.

Fig. 2 shows the FDR corrected p-value maps corresponding to the Cramér tests. The distance d_{DET} was not able to detect statistically significant volume changes in the interior of the 'subcortical structures', because there was no significant local volume changes in those regions. However, significant volume differences were found at outer regions surrounding these structures. This behavior shows that the deformations driven by a rotation of a structure surrounded by a static region mainly generates volume changes outside the structure.

The statistical map using d_{AFF} shows significant differences in the interior of the 'right subcortical structure', because there was an anisotropic scaling. However, no significant differences were found in the interior of the 'left subcortical structure' because the deformations were mainly driven by a rotation.

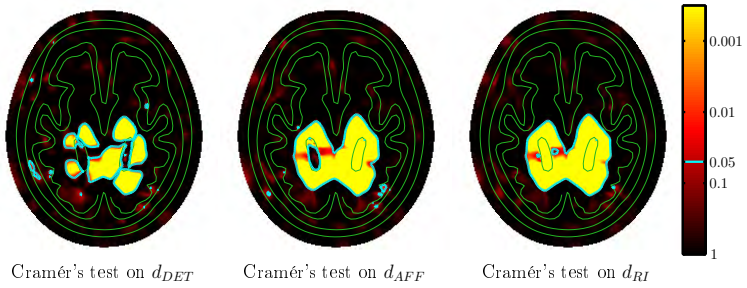


Fig. 2. Statistical map of FDR corrected p-values for Cramér test based on different distances. Template contours are illustrated in green for localization purposes. Cyan contours show the boundary of the regions with significant differences with the criterion of $p_{FDR} \leqslant 0.05$.

Regarding the statistical map using d_{RI}, it can be seen that significant differences were found in the interior of both 'subcortical structures'.

5.2 ADNI dataset

A set of T1-MRI brain images was selected from ADNI database (adni.loni.ucla.edu). Twenty elderly control subjects (denoted here as NOR group) and twenty AD patients (denoted as AD group) were selected from the database. As the brain atrophy is affected by factors such as age and gender, subjects were selected to be gender-matched and within a narrow age interval (72 ± 1 years). Clinical scores, such as MiniMental State Examination (MMSE) [26] or the memory score in the Clinical Dementia Ratio (CDR) [27], were significantly different between both groups under a Student's t-test (MMSE: 29.3 ± 1.0 for NOR group while 23.3 ± 1.8 for AD group; CDR: 0.0 ± 0.0 for NOR group while 0.9 ± 0.3 for AD group).

The MRI template T was built from 40 elderly control subjects as described in [2]. The deformations fields $\boldsymbol{\Phi}_\mu$ were estimated using a SVF diffeomorphic registration [28, 2] between the MRI template T and each image instance.

Voxel-wise Cramér tests were performed using the three right-invariant distances d_{DET}, d_{AFF} and d_{RI}. Critical values were estimated by means of a permutation test using 100,000 permutations. After, the p-values maps were corrected with FDR criterion. Fig. 3 shows a coronal and sagittal illustrative slice of the corrected p-value map for each method.

The three statistical maps are in agreement with the pathophysiological knowledge of Alzheimer's disease. As it was expected, the number of voxels with significant differences in the d_{DET} map was much smaller than in d_{AFF} and d_{RI} maps.

6 Conclusion

Previous literature on TBM focused on selecting different features or descriptors of the Jacobian matrix on which to perform either univariate or multivariate

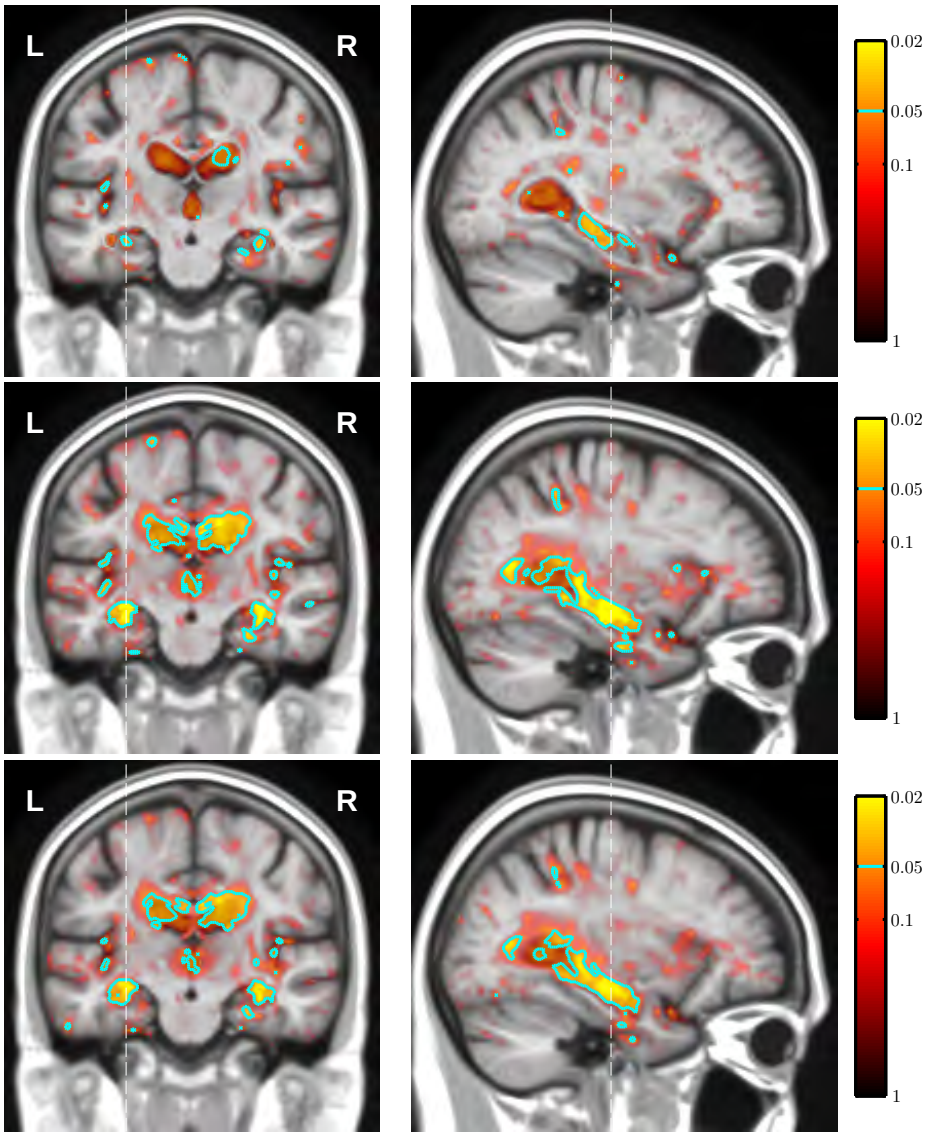


Fig. 3. Illustrative coronal (left) and sagittal (right) view of FDR-corrected p-value maps of the Cramér test on three right-invariant distances: d_{DET} (top row), d_{AFF} (middle row) and d_{RI} (bottom row). L and R denote left and right hemisphere respectively. The dashed vertical lines show the location of the sagittal and coronal slices.

statistics. In this work different distance functions over Jacobian matrices were analyzed. From this distance perspective, the template invariance requirement led to the use of right-invariant distances.

Three different right-invariant distances on $\mathrm{GL}^+(n)$ were considered in this work. The first two distances have been previously used in the literature: the distance between log of Jacobian determinants d_{DET} and a distance between

deformation tensors d_{AFF}. While the distance d_{DET} only measures local volume changes, d_{AFF} additionally quantifies anisotropic scalings and shearings but no local rotations. The third one is a right-invariant Riemannian distance between Jacobian matrices d_{RI}, and it has not been previously used on TBM. The Riemannian distance d_{RI} quantifies the complete local deformation. This behavior was clearly illustrated in the results from the synthetic data.

A Fréchet derivative of the matrix exponential

To compute the Fréchet derivative $\mathcal{D}_{\mathbf{U}}\mathrm{Exp}_{\mathbf{I}}(\mathbf{U})$ given in eq. (7) an expression for $\mathrm{dexpm}(\mathbf{M}) \equiv \mathcal{D}_{\mathbf{M}}\mathrm{expm}(\mathbf{M})$ is required. This is the linear operator containing the derivatives of each element of $\mathrm{expm}(\mathbf{M})$ with respect to a perturbation on each element of $\mathbf{M}$ and results in an $n^2 \times n^2$ matrix.

There are different attempts to compute $\mathcal{D}_{\mathbf{M}}\mathrm{expm}(\mathbf{M})$ [29–31]. For our computations we used the approach given in [32], where for an analytic matrix function $\mathbf{F}(\mathbf{M})$

$$\mathbf{F}\left(\left(\begin{array}{c|c}\mathbf{M} & \mathbf{P}\\ \hline \mathbf{0} & \mathbf{M}\end{array}\right)\right) = \left(\begin{array}{c|c}\mathbf{F}(\mathbf{M}) & \mathrm{d}_r\mathbf{F}(\mathbf{M}+r\mathbf{P})\\ \hline \mathbf{0} & \mathbf{F}(\mathbf{M})\end{array}\right)$$

and if $\mathbf{P}^j$ is the j-th canonical perturbation, then the vectorization of the upper-right submatrix $\overline{\mathrm{d}_r\mathbf{F}(\mathbf{M}+r\mathbf{P}^j)}$ is the j-th column of the matrix $\mathcal{D}_{\mathbf{M}}\mathbf{F}(\mathbf{M})$.

Acknowledgements

This work was funded by research grants TEC2009-14587-C03-01 from CICYT, AMIT project CEN-20101014 from CENIT program, CIM project IPT-2011-1638-900000 from INNPACTO program, Spain. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904).

References

1. Davatzikos, C., *et. al*: Voxel-based morphometry using the RAVENS maps: methods and validation using simulated longitudinal atrophy. NeuroImage **14**(6) (2001) 1361–1369
2. Bossa, M., Zacur, E., Olmos, S., ADNI: Tensor-based morphometry with stationary velocity field diffeomorphic registration: application to ADNI. NeuroImage **51**(3) (2010) 956–969
3. Fletcher, P.T.: Statistical Variability in Nonlinear Spaces: Application to Shape Analysis and DT-MRI. PhD thesis, University of North Carolina (2004)
4. Boisvert, J., *et. al*: Geometric variability of the scoliotic spine using statistics on articulated shape models. IEEE Trans. Med. Imaging **27**(4) (2008) 557–568
5. Moakher, M.: A differential geometric approach to the geometric mean of symmetric positive-definite matrices. SIAM J. Matrix Anal. Appl. **26**(3) (2005) 735–747
6. Batchelor, P.G., *et. al*: A rigorous framework for diffusion tensor calculus. Magn. Reson. Med. **53**(1) (2005) 221–225
7. Pennec, X., Fillard, P., Ayache, N.: A Riemannian framework for tensor computing. Int. J. Comput. Vis. **66**(1) (2006) 41–66
8. Fletcher, P.T., Joshi, S.: Riemannian geometry for the statistical analysis of diffusion tensor data. Signal Processing **87**(2) (2007) 250–262

9. Sommer, S., *et. al*: Manifold valued statistics, exact principal geodesic analysis and the effect of linear approximations. In: Europ. Conf. on Comp. Vision (2010) 43–56
10. Machado, L., Silva Leite, F., Krakowski, K.: Higher-order smoothing splines versus least squares problems on Riemannian manifolds. J. Dynam. Control Systems **16**(1) (2010) 121–148
11. Fletcher, P.T.: Geodesic regression on Riemannian manifolds. In: MFCA Workshop of MICCAI (2011) 86
12. Hinkle, J., *et. al*: Polynomial regression on Riemannian manifolds. arXiv preprint (2012)
13. Muralidharan, P., Fletcher, P.T.: Sasaki metrics for analysis of longitudinal data on manifolds. In: IEEE Conference on CVPR. (2012) 1027–1034
14. Baringhaus, L., Franz, C.: On a new multivariate two-sample test. J. Multivariate Anal. **88**(1) (2004) 190–206
15. Székely, G.J., Rizzo, M.L.: Testing for equal distributions in high dimension. InterStat **11**(5) (2004)
16. Tapp, K.: Matrix groups for undergraduates. AMS (2005)
17. Christensen, G.E., Johnson, H.J.: Invertibility and transitivity analysis for nonrigid image registration. Journal of Electronic Imaging **12**(1) (2003) 106–117
18. Škrinjar, O., Bistoquet, A., Tagare, H.: Symmetric and transitive registration of image sequences. Journal of Biomedical Imaging **2008** (2008) 14
19. Pennec, X.: Left-invariant Riemannian elasticity: a distance on shape diffeomorphisms? In: MFCA Workshop of MICCAI (2006) 1–13
20. Do Carmo, M.P.: Riemannian geometry. Birkhauser (1992)
21. Absil, P.A., Mahony, R., Sepulchre, R.: Optimization algorithms on matrix manifolds. Princeton Univ. Press (2009)
22. Hall, B.C.: Lie groups, Lie algebras, and representations: an elementary introduction. Springer (2003)
23. Andruchow, E., *et. al*: The left invariant metric in the general linear group. arXiv preprint (2011)
24. Abadir, K.M., Magnus, J.R.: Matrix algebra. Cambridge Univ. Press (2005)
25. Benjamini, Y., Hochberg, Y.: Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat Soc. Ser. B. **57** (1995) 289–300
26. Cockrell, J.R., Folstein, M.F.: Mini-mental state examination (MMSE). Psychopharmacol. Bull. **24**(4) (1988) 689–692
27. Morris, J.C.: The clinical dementia rating (CDR): current version and scoring rules. Neurology **43**(11) (Nov 1993) 2412–2414
28. Hernandez, M., Bossa, M., Olmos, S.: Registration of anatomical images using paths of diffeomorphisms parameterized with stationary vector field flows. Int. J. Comput. Vis. **85**(3) (2009) 291–306
29. Al-Mohy, A.H., Higham, N.J.: Computing the Fréchet derivative of the matrix exponential, with an application to condition number estimation. SIAM J. Matrix Anal. Appl. **30**(4) (2009) 1639–1657
30. Najfeld, I., Havel, T.F.: Derivatives of the matrix exponential and their computation. Adv. in Appl. Math. **16**(3) (1995) 321–375
31. Li, C., Sheng, Y., Wang, M.: An effective method to compute Fréchet derivative of matrix exponential and its error analysis. J. Inform. Comput. Sci. **7** (2010) 1854–1859
32. Mathias, R.: A chain rule for matrix functions and applications. SIAM J. Matrix Anal. Appl. **17**(3) (1996) 610–620

Stratified Voxel-Based Morphometry (sVBM)

M. Jorge Cardoso[1], Ivor Simpson[2], Marc Modat[1], and Sebastien Ourselin[1,2]

[1] Centre for Medical Image Computing (CMIC), University College London, UK,
[2] Dementia Research Centre (DRC), University College London, UK

Abstract. In neuroimaging, voxel based morphometry (VBM) has been a valuable tool of identifying brain-wide differences between populations. One of the key elements of VBM is to define a space for voxelwise comparisons. However, errors in this mapping to common space and variations of brain morphology, both natural and pathologic can result in false positives. In this work we explore a new framework, where a spatially varying morphological similarity graph is created between pairs of images. This graph is then used to stratify natural and pathological variability in a VBM-like setting. In contrast to VBM, which describes the group differences on an average brain morphology, sVBM describes how different brain morphologies are independently affected by pathology. Due to its pairwise nature, this technique provides smoother and better localised differences between populations, possibly providing novel insights into the homogeneity of pathological effects for different brain morphologies.

1 Introduction

The characterisation of morphometric differences between healthy and pathological populations is one of the cornerstones of medical imaging. Brain morphometry has been extensively studied in diverse populations, ranging from dementia (e.g. Alzheimer's and Huntington's disease) and schizophrenia, to autism and even normal ageing.

The morphometric characterisation of populations explores differences in volume, mass, shape, cortical thickness and tissue density. Due to the advent of high resolution anatomical imaging and increased computational power, voxel-based techniques have become the *de facto* tool for morphometric analysis. These techniques range from the classic voxel-based morphometry [1], characterising differences in tissue density, to tensor- [2] and deformation-based [3] morphometry, characterising differences in the mapping tensor and deformation parameters respectively. Voxel based morphometry (VBM) involves a voxel-wise comparison of the local concentration of grey matter between two groups of subjects when mapped to a common space. In order to reduce the bias towards the choice of template, a mean shape/appearance space, known as the groupwise space, is often used for the comparison[4]. Despite their seeming optimality for morphometric comparison, groupwise spaces suffer from three main problems: (1) their construction is highly dependent on the choice of image similarity metric and regularisation [5]; (2) the mapping errors to the groupwise space can result

Fig. 1. The "Apples and Oranges" painting by Paul Cézanne, illustrating the discrimination complexity between different but similar objects.

in morphometric mismatch, a problem which has generated wide criticism [6, 7]; (3) and groupwise spaces mix naturally occourring morphologies (e.g. sulcal patterns, brain shape) with pathological (e.g. atrophy) effects. This work focuses on the latter effect.

Some of the above effects have been studied to some degree. Lepore *et al.* [8] and Koikkalainen *et al.* [9] have looked at multi-template TBM but have not explored the stratified space statistics. Other techniques based on pattern-based, feature-based morphometry, patch-based morphometry have also been used to overcome some of the VBM/TBM limitations by attempting to improve the power to detect pathological differences at the population level. However, these techniques do not explore how pathology affects each different brain morphology independently.

By mixing pathological effects with different morphologies, the VBM analysis is to some degree "comparing apples to oranges", i.e. when two items or groups of items cannot or should not be meaningfully compared (see Fig. 1). The use of a groupwise space for morphometric comparison has two underlying assumptions: first, that pathology affects all morphologies in a similar manner, i.e. it assumes that the pathological and natural variability in the high dimensional space of all brain morphologies are linearly separable, even though it is well known that different morphologies have different connectivity and functional activation patterns [10]; second, that the point estimate of the mean and variance of the population tissue densities assumes a monomodal distribution and accurately summarises the full population (see Fig. 2-right), i.e. VBM combines all modulated tissue densities in a groupwise space and then obtains a point estimate of the difference in tissue density between populations. However, certain phenotypes might have protective or deleterious effects on pathology, indicating that the tissue density differences might not be monomodal.

Some groups have explored the idea of stratifying populations in different subgroups for the purpose of segmentation [11, 12]. However, this stratification

normally attempts to separate the groups according to their pathological status, not taking into account the fact that there are different non-pathology-related morphological subgroups. Subsequently, further stratification into different local morphological subgroups is necessary. In the limit, this population stratification process considers each subject's morphological subgroup as a stratum. Interestingly, this limit situation can be interpreted as the same problem solved by the multi-atlas propagation and label fusion community [13, 14], where the morphological similarity between subjects is used as a hub for label propagation. This stratification strategy was used in [15] for the purpose of classification, producing one of the highest reported accuracy results for AD prediction.

This work explores a novel approach for stratified voxel-based morphometry (sVBM). We re-interpret the question behind VBM as a morphology specific process. This approach is built on the Geodesic Information Flows (GIF) implicit-graph construct proposed by Cardoso *et al.* [16], which stratifies the subjects in a population with the use of an implicit spatially variant graph (see Fig. 2-right). This stratified space is then used for morphometric comparison. Overall, this work attempts to answer the question: how is each brain morphology affected by pathology?

2 Methods

This section will first introduce the mathematical framework and the undirected graph for geodesic information flow (GIF), followed by the introduction of the morphological distance metric between images. The GIF framework will then be used for voxel-based tissue density comparison in order to separate naturally-occurring from pathological effects. These stratified effects are then combined in a groupwise space for population analysis.

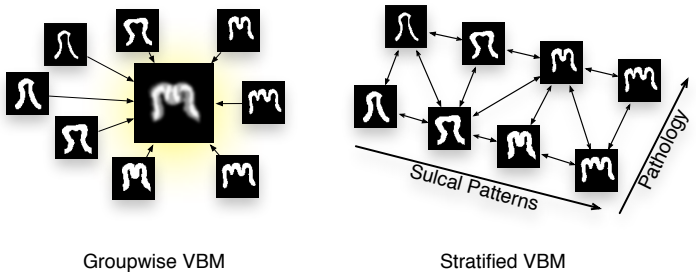


Fig. 2. Left) Representation of how the tissue density flows between subjects in a groupwise approach, representing the average shape and appearance of the population. Right) In a pairwise approach, a spatially varying implicit graph is constructed. At a certain voxel location, the neighbourhood of each subject is dependent on the morphological similarity.

2.1 The geodesic information flows framework

Let a set Y of N images be the full set of observed T1-weighted MRI data with the i-th image of this set denoted by Y_i. Each image Y_i is a vector of size L_i, with the voxel $\boldsymbol{v}$ denoted by $Y_i(\boldsymbol{v})$. In this framework, all images in the database need to be independently non-rigidly mapped to every other image (pairwise registration). The mapping between image Y_i and image Y_j at voxel $\boldsymbol{v}$ is denoted by $T_{i\rightarrow j}(\boldsymbol{v})$. As the framework requires a symmetric, inverse consistent and diffeomorphic coordinate transformation, we use a symmetric variant of a non-rigid free-form registration algorithm as described in [17].

As in [16], let D be a set of distance matrices characterising the morphological similarity between images, with the i-th matrix of this set denoted by D_i. Here, D_i will be an $L_i \times (N-1)$ matrix describing the distance between the image Y_i and each one of the remaining $(N-1)$ images at every voxel position $\boldsymbol{v}$. More specifically, $D_{ij}(\boldsymbol{v})$ will contain the distance between the $\boldsymbol{v}$-th voxel of image Y_i and its corresponding location in image Y_j. Under this assumption, a heat kernel decay function $W_{ij}(\boldsymbol{v})$ is then used as a weight to diffuse information between images. This kernel is defined as $W_{ij}(\boldsymbol{v}) = \exp(-\frac{D_{ij}(\boldsymbol{v})}{t})$, with t being the temperature of the heat kernel that will determine the distance and the speed of information diffusion and $D_{ij}(\boldsymbol{v})$ is a morphological distance defined as in Cardoso *et al.* [16]. Here, t is set to 1 and its optimisation is out of the scope of this work. Note that the graph is undirected only if the distances are a semi-metric (subadditivity is not required). In short, the morphological neighbourhood of a voxel in a certain subject will be restricted to the most similar subjects from each group. In this work, α (as defined in [16]) will be set to 0.9, reducing the influence of the co-ordinate mapping in the morphological similarity metric. This information will be later introduce through Jacobian modulation. Also, no truncation is applied to $W_{ij}(\boldsymbol{v})$, making it spann the full space of morphologies.

2.2 Tissue Segmentation

All N_i images are segmented into their 3 constituent tissue classes: white-matter (WM), grey-matter (GM) and cerebrospinal-fluid (CSF). The tissue segmentation is obtained using AdaPT [18], an EM-based probabilistic segmentation algorithm. This algorithm reduces the bias introduced in the tissue segmentation by the choice of tissue priors through a patient-specific prior sampling procedure. Here, $S_i^c(\boldsymbol{v})$ represents the probability for voxel $\boldsymbol{v}$ in image i to belong to tissue class c. All images are skull-stripped using STEPS [19].

2.3 Stratified VBM

In VBM framework, all tissue segmentations are mapped to a groupwise space and commonly modulated by the Jacobian determinants of the transformation. The modulated segmentation of image i at voxel $\boldsymbol{v}$ in the groupwise space for tissue X is denoted by $M^c_{i\leftarrow \mathrm{GW}}(\boldsymbol{v}) = S_i^c(T_{i\leftarrow \mathrm{GW}}(\boldsymbol{v})) \times |\mathrm{Jac}(T_{i\leftarrow \mathrm{GW}}(\boldsymbol{v}))|$. The

global affine transformation between the images was removed from the Jacobian estimation in order to reduce the influence of head size.

Unlike VBM, which compares all the segmentations in a groupwise space, sVBM compares the segmentations in the space of every image in the database. For a subject i, the segmentation of tissue type C of all other subjects $j \in N \setminus \{i\}$ is mapped to the space of i and modulated by the Jacobian determinants. The modulated segmentations are denoted by $M^c_{i \leftarrow j}$. The same process is repeated for all $i \in N$. Thus, $N \times (N-1)$ modulated probability images are generated per tissue type. Each modulated probability is smoothed by a $2mm$ isotropic full width at half maximum (FWHM) Gaussian filter. This filtering is intrinsically smaller than in VBM as the pairwise registration and the weighting scheme reduces the registration errors.

As previously mentioned, $W_{ij}(\boldsymbol{v})$ characterises the morphological similarity between image i and j at voxel $\boldsymbol{v}$. Thus, $W_{ij}(\boldsymbol{v})$ can be used to weight the modulated tissue densities by the morphological similarity between subjects.

Let the N subjects be divided into two sub-populations of interest: e.g. healthy-controls (N^{HC}), and diseased subjects (N^P). Within the space of a specific subject i, the difference in modulated tissue densities between the two groups at voxel $\boldsymbol{v}$ and for tissue X, here denoted by $D^c_i(\boldsymbol{v})$, will be given by:

$$D^c_i(\boldsymbol{v}) = \frac{\sum_{\forall j \in N^P \setminus \{i\}} W_{ij}(\boldsymbol{v}) \times M^c_{i \leftarrow j}(\boldsymbol{v})}{\sum_{\forall j \in N^P \setminus \{i\}} W_{ij}(\boldsymbol{v})} - \frac{\sum_{\forall j \in N^{HC} \setminus \{i\}} W_{ij}(\boldsymbol{v}) \times M^c_{i \leftarrow j}(\boldsymbol{v})}{\sum_{\forall j \in N^{HC} \setminus \{i\}} W_{ij}(\boldsymbol{v})}$$

These density difference maps are obtained for every subject i and tissue c, resulting in $N \times C$ modulated tissue difference maps between the two populations. Each one of these maps, D^c_i, provides voxel-wise information about the difference in the density of tissue type c between the two populations, weighted by the local similarity between each subject j and the subject i. Note that the current implementation does not take into account the presence of covariates like age and gender, as the testing data is balanced. However, this can be easily added through a weighted least-squares GLM, using W_{ij} as weights.

2.4 Groupwise projection

The values of D^c_i already characterise the differences between the populations for each of the subjects' morphological cluster. Thus, one can now map them to the groupwise space without entangling the pathological with the naturally-occurring variability. For all subjects and tissue types, the tissue density differences are then mapped to the groupwise space, in a similar fashion to the VBM approach. However, in contrast to VBM, where only a point estimate of the populations tissue density difference is obtained at each voxel, in sVBM, a non-parametric distribution of tissue density differences for the different morphologies and between the diseased and heathy populations is obtained at each voxel $\boldsymbol{v}$ of the groupwise space. The vector of tissue density differences at voxel $\boldsymbol{v}$ in the groupwise space is denoted by $D^c_{GW}(\boldsymbol{v}) = \{D^c_1(T_{1 \leftarrow \mathrm{GW}}(\boldsymbol{v})), \ldots, D^c_N(T_{N \leftarrow \mathrm{GW}}(\boldsymbol{v}))\}$.

2.5 Sampling the null distribution

In order to test the significance of the findings, a non-parametric strategy similar to permutation testing is employed. Permutation methods (also known as randomisation methods) can be used for inference and thresholding of statistic maps when the null distribution is not known due to noise, population sample and the use of non-standard statistics. Ideally for sVBM, the null distribution should be constructed in the space of each patient's morphology in order to provide information about the null distribution of density differences for each subject's morphological cluster. For a relatively large number N, this process would be computationally very time consuming. Furthermore, the multiple null distributions on the space of each subject's morphology would have to be propagated to the groupwise space using $T_{i \leftarrow \mathrm{GW}}(\boldsymbol{v})$, breaking the null distribution's independence criterion for each voxel. Instead, an approximation of the true null distribution of D^c_{GW} is obtained by permuting the labels of $M^c_{i \leftarrow \mathrm{GW}}$ [20]. A $6mm$ isotropic FWHM Gaussian filter smoothing is necessary here to correct for the problem of low variance [21] and also in order to account for minor mapping errors to the groupwise space. In this work, 20,000 random label samples were used per voxel to build the null distribution $\mathscr{H}_0$. An example of the null distribution for a voxel in the parahippocampal gyrus can be seen in purple in Fig. 3. The probability that a sample of D^c_{GW} comes from $\mathscr{H}_0$ will thus be given by $\int f(D^c_{GW}) \times f(\mathscr{H}_0)$, with $f(D^c_{GW})$ and $f(\mathscr{H}_0)$ being the probability density function of D^c_{GW} and $\mathscr{H}_0$ respectively.

3 Data

Due to the computational complexity of the pairwise registration strategy, only 100 ADNI2 (`http://adni.loni.ucla.edu`) data sets (50 healthy-controls and

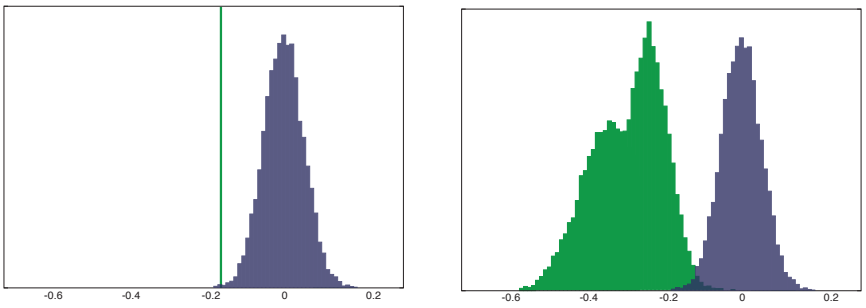


Fig. 3. The null distribution $\mathscr{H}_0$ (in purple) for a voxel in the parahippocampal gyrus. Left) The green line represents the voxel's GM tissue density difference; Right) A smoothed probability density function of the samples in D^c_{GW} (in green) suggesting a non-gaussian distribution of pathological effects for the different brain morphologies.

50 AD subjects) were used for validation. The 50 subjects of each group were chosen as the 50 first subjects ranked by scan number. The list of subjects used in this work will be made available on the author's website at the time of publication. No age and gender statistical differences were found between the two populations.

4 Validation

All subjects were affinely mapped to a template's physical space without resampling by updating the header information. All these affinely-aligned images were then non-rigidly mapped to each other in order to obtain $T_{i\to j}(\boldsymbol{v})$, $W_{i\to j}(\boldsymbol{v})$ and $M_{i\to j}(\boldsymbol{v})$, requiring 9900 registrations (100×99). All density difference maps D^c_{GW} and segmentations were also mapped to a mean deformation space, here called the groupwise space. The segmentations were modulated by the Jacobian determinant and smoothed with a 6mm FWHM Gaussian filter ($M^c_{i\leftarrow \mathrm{GW}}$), similar to VBM, accounting for possible mapping errors. These modulated segmentations were then used to sample $\mathscr{H}_0$ and the p-values (P-map) of the sVBM's population differences map. For comparison purposes, the same smoothed modulated segmentations were also used to estimate a regular VBM-based difference map between the AD and HC populations and its associated p-values. All p-maps are presented uncorrected for multiple comparisons and masked to regions with density above 0.1. Results for the GM are shown in Fig. 4.

5 Discussion

The proposed method shows a well localised and anatomically plausible pattern of atrophy for AD when compared to HC. By stratifying the naturally-occurring image morphologies from the estimation of pathological effects, the mean density differences between the populations become shaper and more pronounced. More interesting than the accuracy of density difference localisation are the characteristics of the non-parametric distribution D_{GW} (see Fig. 4-bottom). The mean, std, 10th and 90th percentiles of the density difference distribution D_{GW} at each voxel are presented in Fig. 4. Note that the least pathologically-affected brain morphologies (90th percentile) still suffer from atrophy in the temporal lobe and thalami area. However, the most pathologically-affected morphologies (10th percentile) suffer from atrophy in nearly all brain regions. The skewness and standard deviation (std) of D_{GW} suggest that the disease process does not affect all morphologies in a similar manner, with some specific local brain morphologies showing almost no pathological effects (see Fig. 4-bottom right).

6 Conclusion

The proposed work presents a novel technique for morphometric characterisation of population differences. In contrast to VBM, which provides an answer to

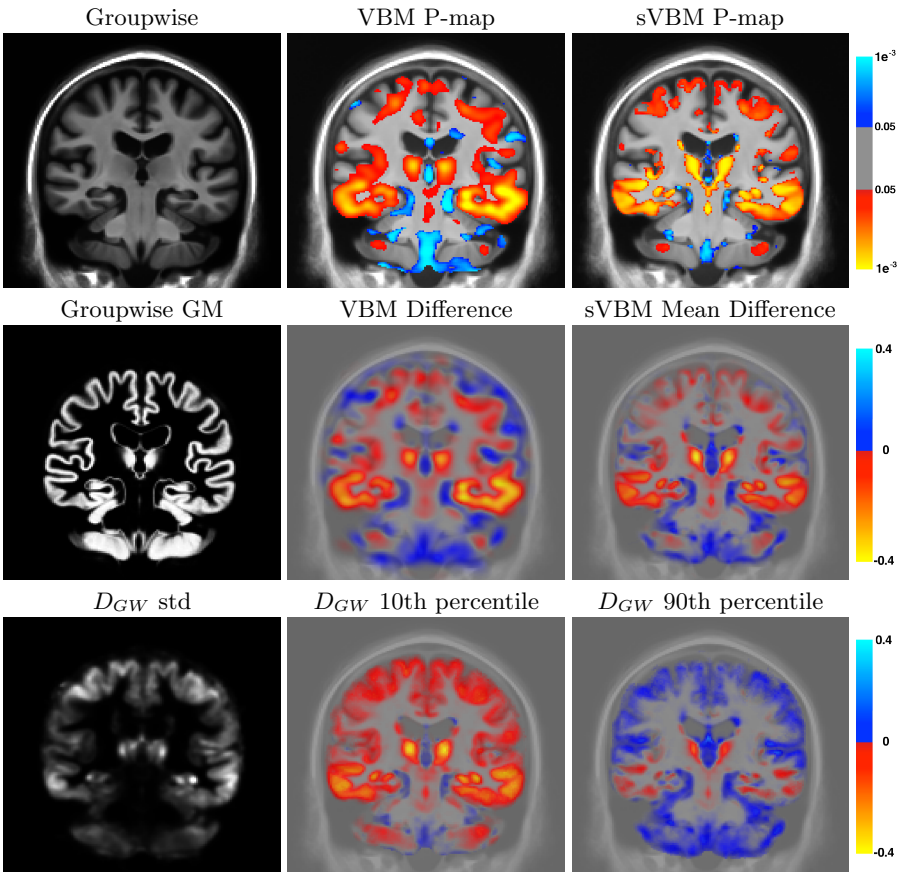


Fig. 4. The first row shows the groupwise mean image, and the positive and negative VBM and sVBM p-maps thresholded at 0.05, showing sharper atrophy localisation. The red and blue colours represent atrophy and expansion respectively. Note the improved localisation accuracy of the pathological differences in sVBM. The second row shows the mean GM segmentation, the VBM GM density difference and the sVBM GM density difference mean, both between the AD and HC populations. The last row shows the std and the 10th and 90th percentile of the density difference distribution D_{GW}.

the question "How does pathology affect an average brain morphology?", the proposed work attempts to answer the question "How does pathology affect each individual brain morphology?". By answering a different question, sVBM could provide a complementary and possibly richer source of information about neurodegeneration.

Acknowledgements

The Dementia Research Centre is an Alzheimer's Research Trust Co-ordinating centre and has also received equipment founded by the Alzheimer's Research Trust. MJC, MM and SO are funded by both the EPSRC (EP/H046410/1) and the CBRC Strategic Investment Award (Ref. 168).

References

1. Ashburner, J., Friston, K.J.: Voxel-based morphometry - The methods. NeuroImage **11**(6) (2000) 805–821
2. Woods, R.P.: Characterizing volume and surface deformations in an atlas framework: theory, applications, and implementation. NeuroImage **18**(3) (March 2003) 769–788
3. Ashburner, J., Hutton, C., Frackowiak, R., Johnsrude, I., Price, C., Friston, K.: Identifying global anatomical differences: deformation-based morphometry. Human Brain Mapping **6**(5-6) (1998) 348–357
4. Ashburner, J., Friston, K.J.: Unified segmentation. NeuroImage **26**(3) (2005)
5. Yeo, B.T.T., Sabuncu, M.R., Desikan, R.S.R., Fischl, B., Golland, P.: Effects of registration regularization and atlas sharpness on segmentation accuracy. MedIA **12**(5) (October 2008) 603–615
6. Bookstein, F.L.: "Voxel-Based Morphometry" Should Not Be Used with Imperfectly Registered Images. NeuroImage **14**(6) (December 2001) 1454–1462
7. Davatzikos, C.: Why voxel-based morphometric analysis should be used with great caution when characterizing group differences. NeuroImage **23**(1) (September 2004) 17–20
8. Leporé, N., Brun, C., Chou, Y.Y., Lee, A., Barysheva, M., De Zubicaray, G.I., Meredith, M., Macmahon, K., Wright, M., Toga, A., Thompson, P.M.: Multi-Atlas Tensor-Based Morphometry and its Application to a Genetic Study of 92 Twins. 2nd MICCAI Workshop ... (October 2008) 48–55
9. Koikkalainen, J., Lötjönen, J., Thurfjell, L., Rueckert, D., Waldemar, G., Soininen, H., Alzheimer's Disease Neuroimaging Initiative: Multi-template tensor-based morphometry: application to analysis of Alzheimer's disease. NeuroImage **56**(3) (June 2011) 1134–1144
10. Amiez, C., Neveu, R., Warrot, D., Petrides, M., Knoblauch, K., Procyk, E.: The Location of Feedback-Related Activity in the Midcingulate Cortex Is Predicted by Local Morphology. Journal of Neuroscience **33**(5) (January 2013) 2217–2228
11. Sabuncu, M.R., Balci, S.K., Shenton, M.E., Golland, P.: Image-driven population analysis through mixture modeling. IEEE TMI **28**(9) (September 2009) 1473–1487
12. Ribbens, A., Hermans, J., Maes, F., Vandermeulen, D., Suetens, P.: SPARC: unified framework for automatic segmentation, probabilistic atlas construction, registration and clustering of brain MR images. In: IEEE International Symposium on Biomedical Imaging, IEEE (2010) 856–859
13. Rohlfing, T., Russakoff, D.B., Maurer Jr, C.R.: Performance-based classifier combination in atlas-based image segmentation using expectation-maximization parameter estimation. IEEE TMI **23**(8) (August 2004) 983–994
14. Warfield, S.K., Zou, K.H., Wells III, W.M.: Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. IEEE TMI **23**(7) (July 2004) 903–921

15. Coupé, P., Eskildsen, S.F., Manjón, J.V., Fonov, V.S.: Scoring by nonlocal image patch estimator for early detection of Alzheimer's disease. NeuroImage: Clinical (2012)
16. Cardoso, M.J., Wolz, R., Modat, M., Fox, N., Rueckert, D., Ourselin, S.: Geodesic Information Flows. In Ayache, N., Delingette, H., Golland, P., Mori, K., eds.: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012. Springer Berlin / Heidelberg, Berlin, Heidelberg (2012) 262–270
17. Modat, M., Cardoso, M.J., Daga, P., Cash, D., Fox, N.C., Ourselin, S.: Inverse-Consistent Symmetric Free Form Deformation. In: WBIR. (July 2012)
18. Cardoso, M.J., Melbourne, A., Kendall, G.S., Modat, M., Robertson, N.J., Marlow, N., Ourselin, S.: AdaPT: An adaptive preterm segmentation algorithm for neonatal brain MRI. NeuroImage (August 2012)
19. Jorge Cardoso, M., Leung, K.K., Modat, M., Keihaninejad, S., Cash, D., Barnes, J., Fox, N.C., Ourselin, S., for the Alzheimer's Disease Neuroimaging Initiative: STEPS: Similarity and Truth Estimation for Propagated Segmentations and its application to hippocampal segmentation and brain parcelation. MedIA (March 2013)
20. Nichols, T.E., Holmes, A.P.: Nonparametric permutation tests for functional neuroimaging: a primer with examples. Human Brain Mapping **15**(1) (January 2002) 1–25
21. Ridgway, G.R., Omar, R., Ourselin, S., Hill, D.L.G., Warren, J.D., Fox, N.C.: Issues with threshold masking in voxel-based morphometry of atrophied brains. NeuroImage **44**(1) (2009) 99–111

Surface Shape Matching and Analysis using Intrinsic Coordinate Parameterizations

Shantanu H. Joshi[1⋆], Jie Shi[2], Yalin Wang[2], Katherine L. Narr[1,3], Arthur W. Toga[1], and Roger P. Woods[1,3]

[1] Laboratory of Neuro Imaging, University of California Los Angeles, CA, USA
[2] School of Computer Sci. and Engineering, Arizona State University, Tempe, USA
[3] Department of Psychiatry and Biobehavioral Sciences, University of California Los Angeles, CA, USA

Abstract. This paper presents a geometric method for parameterization, matching, and analysis of surface shapes. Surfaces are parameterized and represented by intrinsic coordinate maps derived from the conformal structure of the shape. This parameterization is invariant to rigid transformations of the shape, as well as angle-preserving parameterizations of the surface. Shape matching between coordinate maps of two surfaces is achieved by i) deforming the isothermal curves of the intrinsic parameterization under a nonlinear transformation, and ii) locally reparameterizing the isothermal curves to yield invariant diffeomorphic matchings. Experimental results are shown in synthetic data as well as neuroanatomical shapes such as the hippocampus and the cortex.

1 Introduction

Over the past decade, we have seen diverse approaches and applications, both in the continuous and the discrete setting for geometric shape modeling. The different approaches for surface modeling and shape analysis are usually fine tuned for the problem in hand. For computer aided design (CAD) applications, the computer graphics community is interested in parameterized or even parametric surface modeling [1]. The spectral shape approach [2] has been originally proposed for computer vision applications such as 3D shape retrieval and recognition, although it has also been recently applied to anatomical shapes [3–5]. Similarly, Fourier-based modeling approaches such as SPHARM and its variations [6–8] have been proposed for brain subcortical structures. A slightly different approach models the medial information in elongated shapes and represents it using medial representations [9, 10] and its variants, M-reps [11], or continuous M-reps [12]. The LDDMM approach [13] relies on high- (usually infinite) dimensional embeddings of shapes in ambient manifolds and finds an optimal matching between objects through momentum fields. Another interesting recent approach [14] represents and models vector fields on surfaces and considers a

⋆ This work was supported in part by the National Institutes of Health (NIH) through grants NIH/NIBIB P41-EB015922, and P41-RR013642.

space of such surfaces. A slightly different approach [15] uses distributions or currents for modeling shapes of surfaces. There is also the widely used landmark based approach [16, 17] with a well defined statistical formulation.

It is often useful to directly analyze parameterized representations of surface shapes, where quantities such as magnitudes and directions of deformations have convenient physical interpretations. Thus, parameterization-based approaches to shape surface modeling are still considered beneficial and will continue to be developed. Recently, Auzias et al. [18] proposed a model-based parameterization of the cortical surface, where sulci and gyri are constrained to a specific coordinate system. In our work, we represent surfaces using intrinsic curves and allow a continuous deformation based matching of the surfaces via the curves. Surfaces have been previously approximated by curves (for example, see NURBS [1]) in the field of computer graphics, strictly for the purpose of surface editing and deformations, and more recently for facial recognition and analysis [19].

In this paper, we propose an approach that defines an intrinsic coordinate parameterization on the surface which is defined by the surface conformal structure, and present a novel approach for matching the underlying coordinate mappings across different shapes. This method is general and can be applied to a wide variety of shapes. Our approach directly represents the coordinate system by isothermal curves and deforms the coordinate systems via invariant mappings between them. As shown in this paper, this idea has important implications for analyzing anatomical form. This paper is organized as follows: Sec. 2 introduces the surface shape representation, Sec. 3 outlines a method for elastically matching the coordinate systems across shapes, followed by experimental results (Sec. 4) on synthetic data, hippocampal shapes, and the cortex.

2 Surface Representation via Intrinsic Coordinate Maps

We consider a smooth, two-dimensional, genus zero, oriented surface $\mathcal{S}$, and assume that it admits a smooth parameterization $\mathbf{x} : \mathcal{D} \rightarrow \mathbb{R}^3$. The domain $\mathcal{D}$ can either be a unit plane or a sphere $\mathbb{S}^2$, although for the purpose of this paper, we will assume $\mathcal{D} \equiv (u, v) \subset \mathbb{R}^2$. Specifically $\mathcal{D} \equiv [0, 1] \times [0, 1]$. Further we also assume that the parameterization $\mathbf{x}$ is isothermal. This means that the local coordinates of the parameterization admit a conformal mapping from a neighborhood of the surface $\mathcal{S}$ to the Euclidean plane. Then the coordinate curves of the parameterization $\mathbf{x}$ are also isothermal. This condition implies that

$$\langle \frac{\partial \mathbf{x}}{\partial u}, \frac{\partial \mathbf{x}}{\partial u} \rangle = \langle \frac{\partial \mathbf{x}}{\partial v}, \frac{\partial \mathbf{x}}{\partial v} \rangle, \text{ and } \langle \frac{\partial \mathbf{x}}{\partial u}, \frac{\partial \mathbf{x}}{\partial v} \rangle = 0, \tag{1}$$

thus ensuring that the isothermal curves are orthogonal to each other. We denote the isothermal curves by $\psi_u(v = \text{constant})$, and $\psi_v(u = \text{constant})$.

Theorem 1. *Given a pair of orthogonal vector fields $(\mathbf{w}_u, \mathbf{w}_v)$ at a point $p \in \mathcal{S}$, where $\frac{d\psi_u}{dt}(0) = \mathbf{w}_u$, $\frac{d\psi_v}{dt}(0) = \mathbf{w}_v$, and $\psi_u(0) = \psi_v(0) = p$, there always exists a parameterization $\mathbf{x}$ in the neighborhood of p such that the coordinate curves of $\mathbf{x}$ are the integral curves of $\mathbf{w}_u$ and $\mathbf{w}_v$.*

The proof of Theorem 1 due to Korn-Lichtenstein [20, 21] uses the notion of the first fundamental form of surfaces, and the reader is referred to Varolin [22] for a modern proof. The fundamental idea here is that one can always recover a surface patch locally using the vector fields on the isothermal curves. Further, the entire surface can be reconstructed using a collection of such overlapping local patches. We thus represent the surface $\mathcal{S}$ by the coordinate map given by the collection of curves,

$$\Psi \equiv (\{\psi_u\}, \{\psi_v\}), \forall u, v \in [0,1], \text{where} \tag{2}$$

$$\psi_u, \psi_v : [0, 2\pi] \rightarrow \mathbb{R}^3 \Big| \int_0^{2\pi} \dot{\psi}_{u(t)} dt = \int_0^{2\pi} \dot{\psi}_{v(t)} dt = 0. \tag{3}$$

We list a few properties of the set Ψ below.

1. Since Ψ is defined intrinsically on the surface, it is invariant to rigid motions (translations and rotations) of the surface.
2. Any angle-preserving mapping of the original surface generates a new Ψ that will consist of a new family of isothermal curves. The original Ψ can be recovered up to a diffeomorphism of the plane or the sphere.
3. The coordinate map Ψ can alternately be generated as a solution of the Laplace's equation $\Delta f = 0$ for any harmonic function $f : \mathcal{S} \rightarrow \mathbb{R}$.

2.1 Construction of Isothermal Curves using Conformal Mapping

Any regular surface admits an isothermal coordinate system [23]. A collection of isothermal coordinate systems under a particular orientation defines a Riemannian structure on the surface. Furthermore, such a surface is locally conformal to a plane. Then the coordinate functions of any two transition or overlapping charts on this surface are holomorphic. We will exploit this idea to construct isothermal curves on surfaces. Particularly, in our work, isothermal curves are computed from a global conformal mapping of the surface with a holomorphic 1-form based method [24]. As discussed earlier, holomorphic functions and the associated differential forms can be generalized to Riemann surfaces by using the notion of conformal structure. For example, a *holomorphic 1-form* ω is a complex differential form, such that in each local frame, (u_α, v_α), $z_\alpha = u_\alpha + \mathbf{i}v_\alpha$, where $\mathbf{i} = \sqrt{-1}$, the parametric representation is $\omega = f(z_\alpha)dz_\alpha$, where $f(z_\alpha)$ is a holomorphic function. On a different chart $\{U_\beta, \phi_\beta\}$, with another local frame , (u_β, v_β), $z_\beta = u_\beta + \mathbf{i}v_\beta$, where $\mathbf{i} = \sqrt{-1}$, we have $\omega = f_\beta(z_\beta)dz_\beta = f_\beta(z_\beta(z_\alpha))\frac{dz_\beta}{dz_\alpha}dz_\alpha$, where $f_\beta \frac{dz_\beta}{dz_\alpha}$ is still a holomorphic function. By Hodge theory, one may compute holomorphic 1-forms [24] on simply connected (such as the cortex) or multiply connected (such as the hippocampus) surfaces, where the computed holomorphic 1-forms induce a surface

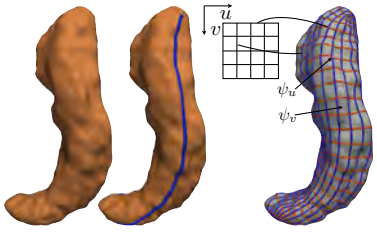


Fig. 1: Example of coordinate curves on a hippocampal surface.

conformal parameterization on the Euclidean domain. For both hippocampal and cortical surfaces, the holomorphic 1-form induces a special system of curves on a surface, the so-called *conformal net*. Horizontal trajectories are the curves that are mapped to iso-v lines in the parameter domain. Similarly, vertical trajectories are the curves that are mapped to iso-u lines in the parameter domain. The horizontal and vertical trajectories form a web on the surface. Because the holomorphic 1-form induces a conformal parameterization, these curves become isothermal curves defined in the prior section. For a tube-like shape or a genus zero surface with one open boundary, the trajectory connectivity structure is trivial (as there is no zero point). Our work uses these trajectories as the curves defined on surface to represent an intrinsic coordinate system. Fig. 1 shows an example of a conformal net on a hippocampal surface.

For some anatomical shapes, we label two consistent landmark curves at the anterior and the posterior end. For example, in case of the hippocampus, we represent its anterior junction with the amygdala, and its posterior limit as it turns into the white matter of the fornix. These are biologically valid and consistent landmarks across subjects. We call this process *topological optimization*. Given the hippocampal tube-like shape, the landmarks curves can be automatically identified by finding the extreme points by searching along the first principal direction of geometric moments of the surface [25]. By cutting along these two landmark curves, a hippocampal surface is modeled by a cylinder-like surface, i.e. a multiply connected surface. Alternately, the cortex is represented by a genus zero surface with one open boundary by cutting along the boundary of the corpus callosum.

3 Elastic Shape Matching between Surfaces

Given two surfaces $\mathcal{S}_1$ and $\mathcal{S}_2$, and their coordinate maps Ψ_1 and Ψ_2, we will find an optimal mapping by deforming the coordinate maps from one surface to the other. Since the map Ψ is not scale invariant, we normalize the scale as follows. First in a neighborhood of a point $p \in \mathcal{S}$, we have the local curve parameterizations $\psi_u : t \in [0, 2\pi) \rightarrow \mathbb{R}^3$, and $\psi_v : t \in [0, 2\pi) \rightarrow \mathbb{R}^3$, that are used to define the local surface parameterization as $\frac{\partial \mathbf{x}}{\partial u} = \frac{d\psi_u}{dt}\frac{dt}{du}, \frac{\partial \mathbf{x}}{\partial v} = \frac{d\psi_v}{dt}\frac{dt}{dv}$. Then the scale normalization is achieved using,

$$\tilde{\Psi} = \frac{\Psi}{\sqrt{\int_{\mathcal{D}} ||\frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v}|| du\ dv}} \tag{4}$$

With a slight abuse of notation, we will refer to the scale-normalized map $\tilde{\Psi}$ as Ψ throughout the paper. For rotational alignment, we will first compute the centroids of the set of any one of the coordinate curves. For example, for surface $\mathcal{S}_1$, we compute $\sigma_u^1 = \int_0^1 \psi_u du, \forall u \in [0, 1]$. Here $\sigma_u^1 \in \mathbb{R}^3, \forall u \in [0, 1]$. Similarly, for surface $\mathcal{S}_2$, we compute $\sigma_u^2 = \int_0^1 \psi_u du, \forall u \in [0, 1]$. Then the optimal rotation ($\hat{O} \in SO(3)$) is initialized by carrying out a singular value decomposition $\hat{O} =$

$ADB^T = \int_0^1 \sigma_u^1 \sigma_u^2 du$, where A and B are left and right unitary matrices, and D is a matrix given by

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & |A||B| \end{bmatrix}. \tag{5}$$

Our approach is as follows. We simultaneously obtain a global as well as a locally optimal matching of the surface via its isothermal curves. Instead of explicitly imposing a Riemannian metric on the underlying geometrical structure, as done in [14, 26], we propose a distance function on the coordinate curves directly. Figure 2 shows the schematic of the workflow.

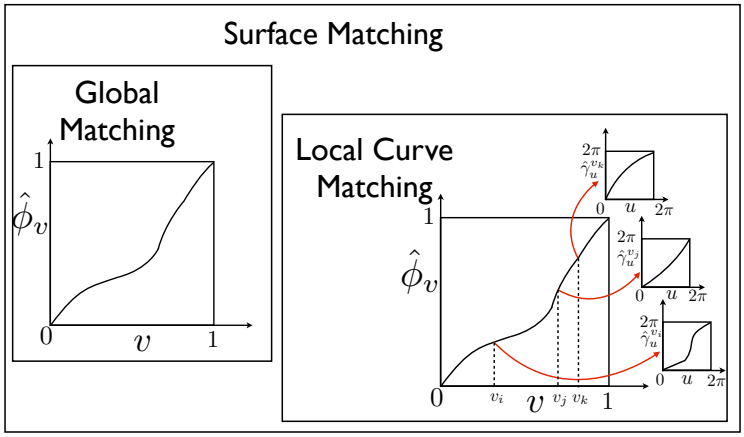


Fig. 2: Illustration of the global and local surface matching procedure. For each $v \in [0, 1]$, there is a local diffeomorphism of the isothermal curve.

3.1 Global Matching between Surfaces

In order to simplify the global matching, we observe that the isothermal curves ψ_v can also be written as functions of u, since ψ_u can be reconstructed from ψ_v owing to orthogonality. We will thus simplify the notation in Eqn. 3 and by slight abuse of notation, denote the surface by the collection,

$$\Psi \equiv \{\psi_v(u)\}, \text{where } u \in [0, 2\pi), \tag{6}$$

$$\psi_v : [0, 2\pi] \to \mathbb{R}^3 \Big| \int_0^{2\pi} \dot{\psi}_v(u) du = 0, \forall v \in [0, 1]. \tag{7}$$

Let the coordinate maps Ψ_1, and Ψ_2 be given by $\{\psi_v^1(u)\}$, and $\{\psi_v^2(u)\}$ respectively. Then we define the matching problem as the minimizer,

$$d_{\mathcal{S}}(\Psi_1, \Psi_2) = \min_{\phi_v} \int_0^1 d_\psi(\psi_v^1, \hat{O}\ \psi_{\phi_v}^2)^2 [1 + \dot{\phi}_v] dv, \tag{8}$$

where $\phi_v : [0, 1] \to [0, 1]$ is a diffeomorphism, and $v \in [0, 1]$ is the index of the isothermal curve. Furthermore d_ψ is the distance between isothermal curves v,

and ϕ_v for a given v, and is defined in Sec. 3.2. Ideally, we would like the distance $d_{\mathcal{S}}$ to be inverse-consistent and invariant to reparameterizations by the index of the isothermal curve s.

Proposition 1. *The distance $d_{\mathcal{S}}$ is invariant to reparameterization by the index function of the isothermal curves.*

Proof. In order to check for invariance due to reparameterization, we evaluate Eqn. 8 under an arbitrary reparameterization of the time parameter $\lambda(v) : [0,1] \to [0,1]$. Without loss of generality, we set $\hat{O} = I$, an identity matrix Thus Eqn. 8 becomes

$$d_{\mathcal{S}}(\Psi_1(\lambda(v)), \Psi_2(\lambda(v))) = \underset{\phi_{\lambda(v)}}{\operatorname{argmin}} \int_0^1 d_\psi\big(\psi^1_{\lambda(v)}, \psi^2_{\phi(\lambda(v))}\big)^2 [1 + \dot{\phi}(\lambda(v))]\dot{\lambda}(v)dv. \tag{9}$$

By performing a change of variables $\lambda(v) = g$, and letting $\lambda(0) = 0, \lambda(1) = 1$ without loss of generality, we get $dg = \dot{\lambda}(v)dv$. Also we note that $\frac{d}{dv}\phi(\lambda(v)) = \dot{\phi}(\lambda(v))\dot{\lambda}(v)dv$. Thus the reparameterized distance becomes

$$\underset{\phi_g}{\operatorname{argmin}} \int_0^1 d_\psi\big(\psi^1_g, \psi^2_{\phi_g}\big)^2 [1 + \dot{\phi}_g]dg, \tag{10}$$

which is same as Eqn. 8.

Proposition 2. *Given that the distance between the isothermal curves d_ψ is symmetric, then the distance $d_{\mathcal{S}}$ is an inverse-consistent (symmetric) distance between coordinate maps.*

Proof. In order to check for inverse-consistency, we need to verify that $d_{\mathcal{S}}(\Psi_1, \Psi_2)$ under the optimal time warp $\hat{\phi}_v$ is same as $d_{\mathcal{S}}(\Psi_2, \Psi_1)$ under the optimal time warp $\hat{\phi}_v^{-1}$. Again, without loss of generality, we set $\hat{O} = I$, an identity matrix. Noting that $\frac{d}{dv}\phi_v^{-1} = \frac{1}{\dot{\phi}(\phi_v^{-1})}$ under the inverse function theorem, we rewrite Eqn. 8 as

$$d_{\mathcal{S}}(\Psi_2, \Psi_1) = \min_{\phi_v^{-1}} \int_0^1 d_\psi\big(\psi^2_v, \psi^1_{\phi_v^{-1}}\big)^2 \big[1 + \frac{1}{\dot{\phi}_v(\phi_v^{-1})}\big]dv. \tag{11}$$

Again by change of variables we let $\phi_v^{-1} = t$. Thus $dv = \dot{\phi}_v dt$. Therefore the above distance becomes

$$d_{\mathcal{S}}(\Psi_2, \Psi_1) = \min_{\phi_v^{-1}} \int_0^1 d_\psi(\psi^1_v, \psi^2_{\phi_v})^2 [\dot{\phi}(t) + 1]dt, \tag{12}$$

which is same as Eqn. 8.

3.2 Local Curve Matching of Coordinate Parameterizations

The objective function (Eqn. 8) for global matching involves the curve distance between isothermal curves, and further requires it to be symmetric. To obtain an

efficient, invariant and symmetric mapping between coordinate curves ψ_v^1, and ψ_v^2, we consider the following distance function:

$$d_\psi(\psi_v^1, \psi_v^2) = \underset{r, O \in SO(3), \gamma_u \in \mathbb{S}^1}{\operatorname{argmin}} \frac{1}{2} \int_0^{2\pi} \big[\big\langle \psi_v^1 - r \cdot O(\psi_v^2 \circ \gamma_u), \psi_v^1 - r \cdot O(\psi_v^2 \circ \gamma_u) \big\rangle + \big\langle \psi_v^2 - r \cdot O(\psi_v^1 \circ \gamma_u^{-1}), \psi_v^2 - r \cdot O(\psi_v^1 \circ \gamma_u^{-1}) \big\rangle \big] du, \tag{13}$$

where the operation $r \cdot \psi_v^1 = \psi_v^1(s - r)|_{\text{mod} 2\pi}$, simply shifts the starting point of the curve, $O \in SO(3)$ is a local rotation, and $\gamma : \mathbb{S}^1 \to \mathbb{S}^1$ is a local reparameterization. For a discretized version of the curve, the shift in origin r is found exhaustively, and the optimal rotation is found using singular value decomposition as described before. Finally, given an optimal shift and rotation, we solve Eqn. 13 using dynamic programming. Eq. 13 is symmetric, which also causes the distance defined in Eq. 8 to be symmetric. It should be noted that the Eq. 8 jointly optimizes the mapping in both u and v directions via the global parameterization ϕ_v in Eqn. 8 and the local parameterization γ_u in Eqn. 13.

4 Results

4.1 Synthetic Data

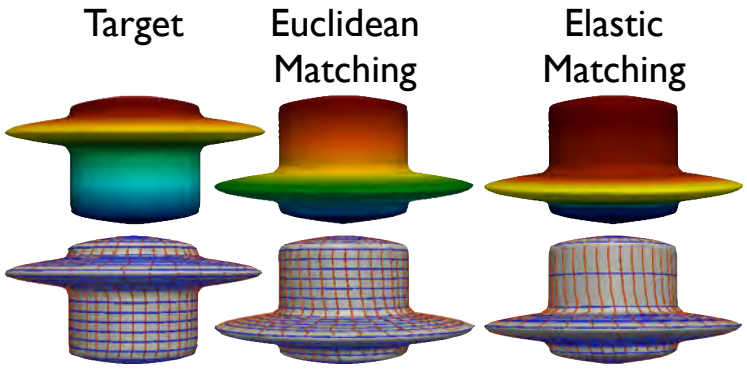


Fig. 3: Comparison of Euclidean and elastic surface matching for synthetic data. Top row shows the matching features in color. Bottom row shows the corresponding parameterizations.

The synthetic data consists of two cylindrical shapes that differ by the locations of the protrusions along the length. We use the same method as the hippocampal surface to induce isothermal curves on them. Although this is a manufactured example, the alignment resulting from shape matching provides insights about the behavior of the algorithm. Fig. 3 (top row) shows the target and the source shapes color coded according to the parameterization in the v direction. The middle column shows the Euclidean matching of the source to the target. This is achieved by setting $\phi_v = id, \forall v \in [0, 1]$, and $\gamma_u = id$ for each v.

The last column shows the elastic shape matching of the source by minimizing Eq. 8 and Eq. 13. The bottom row shows the respective parameterizations resulting from the matching. In the last column, we observe that the optimal matching is achieved by shifting the bump downward, while preserving the features (shown by color) on the bump. Thus the corresponding parameterization (last row, last column) shows large stretching before the bump, and compression immediately after. We also show a smooth deformation between the target to the source for

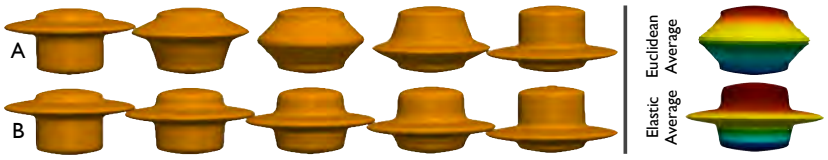


Fig. 4: Left panel: Smooth deformations between the target and the source shapes for A. Euclidean matching, and B. Elastic matching. Right panel: Average shapes using Euclidean and elastic matching.

both Euclidean matching (Fig. 4 A) and the elastic matching (Fig. 4 B). The right panel in Fig. 4 also shows the average shapes. The average shape for the elastic matching preserves the bump as well as the parameterization (shown by color) compared to the Euclidean case.

4.2 Cortical Registration

Here, we show an example of matching intrinsic coordinate maps between two cortical surfaces, extracted using the MNI protocol [27]. Fig. 5 shows comparisons of Euclidean and elastic matching of the coordinate maps for the cortical surfaces. It should be noted that no sulcal curves were traced on the surfaces, and that the shape matching is landmark-free. The circles highlight where there is a noticeable improvement in cortical homology due to the elastic matching. While this example is for demonstration purposes only, and the nature of the cortical correspondence warrants careful validation, our goal here was to illustrate the potential for a simple elastic curve registration based method to solve a general surface matching problem.

4.3 Hippocampal shape analysis

For the purpose of hippocampal morphometry, the imaging data consisted of high resolution T1 MRI images of 40 healthy subjects (15 Males / 25 Females) in the age range $18 - 80$ years with the mean age of 44 ± 20 years. The hippocampal surface meshes were segmented using Freesurfer [28] and the intrinsic coordinate maps were generated using the procedure in Sec. 2.1. Fig. 6 shows the shape alignment between pairs of target and the source hippocampi from this dataset. Each row of Fig. 6 shows the source hippocampus followed by the aligned target due to Euclidean matching, and the aligned target shape due to elastic matching. All shapes are color-coded according to the underlying parameterization, and the structural alignment can be visualized by similar colors. Visually, the elastic

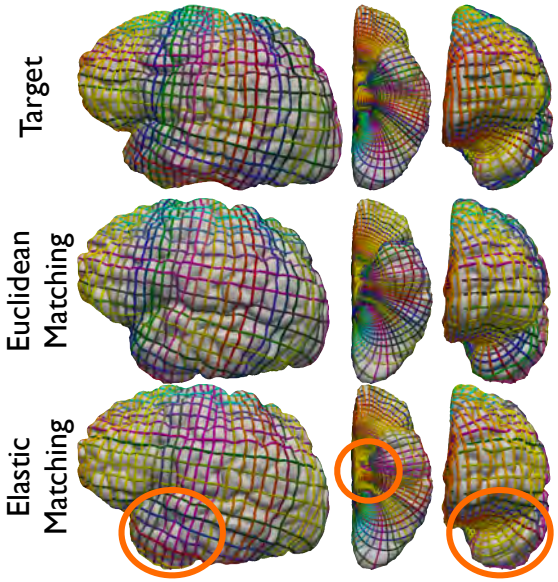


Fig. 5: Cortical surface mapping between the target shape (top) and the source shape using Euclidean (middle) and elastic matching (bottom). Also shown are the coordinate curves colored according to the parameterization. Circles denote improvement in registration due to elastic matching.

matching yields a better alignment as observed in the anterior (top), central, and the posterior regions.

We then performed statistical analysis of the morphology of the hippocampus to model the effect of age on the shape. We selected an arbitrary shape in the dataset as the template and aligned all the surfaces to it using the elastic shape matching. Then all the aligned surfaces were grouped together, and an average shape of the hippocampal surfaces was computed for both the left and the right hemisphere in the brain. Finally for each aligned surface, we computed the radial distance from each vertex on the surface to the medial axis. This distance was used as the shape variable in the statistical analysis. We used a general linear model for inferring the effect of age after covarying for gender on the hippocampal morphology. Fig. 7 shows the statistically significant p-values (after correcting for multiple comparisons using the false discovery rate yielding a cut-off value $p = 0.01$) overlaid on three different views of the averages of the left and the right hippocampus. These shapes were uniformly scaled, so the results show contractions exclusively due to shape with the progression of age. This result is in agreement with the established knowledge that the brain volume and also the hippocampal volume atrophies with age. Finally we also performed supervised classification to determine the accuracy of hippocampal shape as a predictor for age. We defined the two classes as subjects with age ≤ 30 years, and those with age > 30 years. We then computed the distance given in Eqn. 8

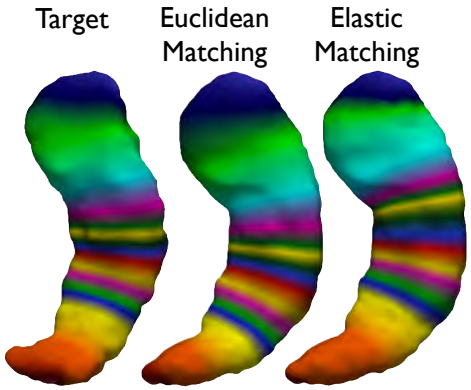


Fig. 6: Shape matching for hippocampal surfaces, colored according to the parameterization. Each row shows a different matched pair of hippocampal shapes.

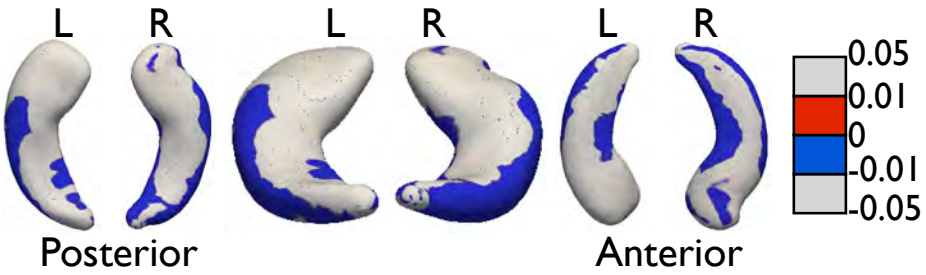


Fig. 7: Significant age effects (corrected for multiple comparisons) on hippocampal shape shown by p-values after applying a general linear model (GLM) covarying for gender. The colormap is shown at the bottom.

between the individual shapes to the mean and used it as a classifier in a support vector machine (SVM) framework. When computing distances, we computed the shape distances in the following ways i) based on exhaustive matching of the conformal nets, and ii) based on partial matching ($\pm$ 20 coordinate curves from the center) of the conformal nets from individuals to the mean shape. Fig. 8 shows the accuracy for predicting the age based on hippocampal shape for the two strategies. We observed that the average accuracy improved from 57.5% to 62.5% when we restricted the matching to the subset of the conformal nets. This may be because the classification of age may be sensitive to morphological changes near the center of the shape (Also see Fig. 7).

5 Discussion

We presented a geometric approach for mapping intrinsic coordinate systems across shapes. The coordinate systems are derived from the holomorphic 1-form that exploits the conformal structure of the surface. Bidirectional coordinate

57.5%	Age < 30	Age > 30
Age < 30	11	8
Age > 30	9	12

(a)

62.5%	Age < 30	Age > 30
Age < 30	11	6
Age > 30	9	14

(b)

Fig. 8: Percent Accuracy and confusion matrix for age prediction from hippocampal shape after training a 10-fold SVM classifier. Percent accuracy differs depending upon (a) Exhaustive matching of conformal nets from individual to mean, and (b) Partial matching of conformal nets from individual to mean.

matching is achieved jointly using both a global diffeomorphic warping function as well as a local diffeomorphic function invariant to local coordinate reparameterizations. This achieves both global shape alignment while optimizing local shape correspondences. In the future, we will investigate methods for performing statistical shape analysis of surfaces via the coordinate map representation, as well as perform detailed validation and comparisons with other approaches such as the spectral methods for subcortical shapes, as well as fluid-based registration approaches for cortical surfaces.

References

1. Piegl, L.: On NURBS: a survey. Computer Graphics and Applications, IEEE **11**(1) (1991) 55–71
2. Reuter, M., Wolter, F.E., Peinecke, N.: Laplace–Beltrami spectra as Shape-DNA of surfaces and solids. Computer-Aided Design **38**(4) (2006) 342–366
3. Qiu, A., Bitouk, D., Miller, M.I.: Smooth functional and structural maps on the neocortex via orthonormal bases of the Laplace-Beltrami operator. IEEE Transactions on Medical Imaging **25**(10) (2006) 1296–1306
4. Reuter, M., Wolter, F.E., Shenton, M., Niethammer, M.: Laplace–Beltrami eigenvalues and topological features of eigenfunctions for statistical shape analysis. Computer-Aided Design **41**(10) (2009) 739–755
5. Shi, Y., Morra, J., Thompson, P., Toga, A.: Inverse-consistent surface mapping with Laplace-Beltrami eigen-features. In: Information Processing in Medical Imaging, Springer (2009) 467–478
6. Styner, M., Oguz, I., Xu, S., Brechbühler, C., Pantazis, D., Levitt, J.J., Shenton, M.E., Gerig, G.: Framework for the statistical shape analysis of brain structures using SPHARM-PDM. The Insight Journal (1071) (2006) 1–20
7. Chung, M.K., Robbins, S.M., Dalton, K.M., Davidson, R.J., Alexander, A.L., Evans, A.C.: Cortical thickness analysis in autism with heat kernel smoothing. NeuroImage **25**(4) (2005) 1256–1265
8. Shen, L., Makedon, F.: Spherical mapping for processing of 3d closed surfaces. Image and Vision Computing **24**(7) (2006) 743–761
9. Joshi, S., Pizer, S., Fletcher, P.T., Yushkevich, P., Thall, A., Marron, J.: Multiscale deformable model segmentation and statistical shape analysis using medial descriptions. IEEE Transactions on Medical Imaging **21**(5) (2002) 538–550

10. Bouix, S., Pruessner, J.C., Louis Collins, D., Siddiqi, K.: Hippocampal shape analysis using medial surfaces. Neuroimage **25**(4) (2005) 1077–1089
11. Pizer, S.M., Fletcher, P.T., Joshi, S., Thall, A., Chen, J.Z., Fridman, Y., Fritsch, D.S., Gash, A.G., Glotzer, J.M., Jiroutek, M.R., et al.: Deformable M-reps for 3D medical image segmentation. International Journal of Computer Vision **55**(2) (2003) 85–106
12. Yushkevich, P.A., Zhang, H., Gee, J.C.: Continuous medial representation for anatomical structures. IEEE Transactions on Med. Imaging **25**(12) (2006) 1547–1564
13. Beg, M.F., Miller, M.I., Trouvé, A., Younes, L.: Computing large deformation metric mappings via geodesic flows of diffeomorphisms. International Journal of Computer Vision **61**(2) (2005) 139–157
14. Kurtek, S., Klassen, E., Ding, Z., Jacobson, S.W., Jacobson, J.L., Avison, M.J., Srivastava, A.: Parameterization-invariant shape comparisons of anatomical surfaces. IEEE Transactions on Med. Imaging **30**(3) (Mar 2011) 849–858
15. Durrleman, S., Pennec, X., Trouvé, A., Ayache, N.: Statistical models of sets of curves and surfaces based on currents. Medical Image Analysis **13**(5) (2009) 793–808
16. Dryden, I., Mardia, K.: Statistical analysis of shape. John Wiley & Son (1998)
17. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J., et al.: Active shape models-their training and application. Computer Vision and Image Understanding **61**(1) (1995) 38–59
18. Auzias, G., Lefevre, J., Le Troter, A., Fisher, C., Perrot, M., Regis, J., Coulon, O.: Model-driven harmonic parameterization of the cortical surface: Hip-hop. IEEE Transactions on Medical Imaging **32**(5) (2013) 873–887
19. Srivastava, A., Samir, C., Joshi, S.H., Daoudi, M.: Elastic shape models for face analysis using curvilinear coordinates. Journal of Mathematical Imaging and Vision **33**(2) (2009) 253–265
20. Osserman, R.: Zwei anwendungen der methode der sukzessiven approximationen. Festschrift fur H.A. Schwarz, Springer (1914) 215–229
21. Lichtenstein, L.: Zur theorie der konformen abbildung: Konforme abbildung nichtanalytischer, singularitätenfreier flächenstücke auf ebene gebiete. Bull. Acad. Sci. Cracovie, Cl. Sci. Math. Nat. Ser. A (1916) 192–217
22. Varolin, D.: Riemann surfaces by way of complex analytic geometry. Volume 125. American Mathematical Soc. (2011)
23. Donaldson, S.K.: Riemann surfaces. Oxford University Press Oxford (2011)
24. Wang, Y., Lui, L.M., Gu, X., Hayashi, K.M., Chan, T.F., Toga, A.W., Thompson, P.M., Yau, S.T.: Brain surface conformal parameterization using Riemann surface structure. IEEE Transactions Med. Imag. **26**(6) (June 2007) 853–865
25. Elad, M., Milanfar, P., Golub, G.H.: Shape from moments-an estimation theory perspective. IEEE Transactions on Signal Processing **52**(7) (2004) 1814–1829
26. Pennec, X.: Left-invariant Riemannian elasticity: a distance on shape diffeomorphisms? In: 1st MICCAI Workshop on Mathematical Foundations of Computational Anatomy: Geometrical, Statistical and Registration Methods for Modeling Biological Shape Variability. (2006) 1–13
27. Holmes, C., MacDonald, D., Sled, J., Toga, A., Evans, A.: Cortical peeling: CSF/grey/white matter boundaries visualized by nesting isosurfaces. Visualization in Biomedical Computing (1996) 99–104
28. Dale, A.M., Fischl, B., Sereno, M.I.: Cortical surface-based analysis. I. segmentation and surface reconstruction. NeuroImage **9**(2) (1999) 179–94

A Relaxed Problem of Registration Based on the Saint Venant-Kirchhoff Material Stored Energy for the Mapping of Mouse Brain Gene Expression Data to a Neuroanatomical Mouse Atlas

Ratiba Derfoul[1,2] and Carole Le Guyader[1]
{ratiba.derfoul,carole.le-guyader}@insa-rouen.fr

[1] Laboratoire de Mathématiques, INSA de Rouen, Saint-Étienne-du-Rouvray, France
[2] IFP Énergies nouvelles, Rueil-Malmaison, France

Abstract. In this paper, we address the issue of designing a theoretically well-motivated registration model capable of handling large deformations. Motivated by the fine properties of the Saint Venant-Kirchhoff material, we propose a variational model combining a measure of dissimilarity and a regularizer based on the stored energy of such a material. We prove the existence of generalized solutions of this problem. We then describe and analyze a numerical method of resolution based on the introduction of an associated decoupled problem under inequality constraint in which an auxiliary variable simulates the Jacobian matrix of the deformation field. A theoretical result is established and we investigate the efficiency of the proposed matching model for the registration of mouse brain gene expression data to a neuroanatomical mouse atlas.

Keywords: Image registration, calculus of variations, nonlinear elasticity, Saint Venant-Kirchhoff materials, relaxed problem, decoupled problem, theoretical convergence, mouse atlas

1 Introduction

Given two images called Template and Reference, registration consists in determining an optimal diffeomorphic transformation φ such that the deformed Template image is aligned with the Reference. This technique has gained much interest in clinical studies among others, when comparing an image to a database or for volumetric purposes. For images of the same modality, the goal of registration is to correlate the geometrical features and the intensity level distribution of the Reference and those of the Template. When the images have been acquired through different mecanisms and have different modalities, registration aims to correlate both images while preserving the modality of the Template image.
There are forward and backward registrations. In this work, we adopt the Eulerian framework to find a backward transformation Φ such that the grid points y

in the deformed Template originate from non-grid points $x = \Phi(y)$.
For an overview of existing parametric and non-parametric registration methods (including landmarks, L^2 similarity, linear elasticity, linear diffusion, splines, viscous fluid model, etc.), we refer the reader to [17]. In the case of non-parametric methods (our framework), the transformation is not restricted to a parameterizable set and the problem is phrased in terms of minimization of functional (with unknown the deformation vector field φ) including a distance measure criterion and a smoother on the deformation vector field φ. Physical arguments often motivate the way the smoother is devised. Classical regularizers such as linear elasticity (see [4]) are not suitable for problems involving large deformations since assuming small strains and the validity of Hooke's law. The scope of the proposed work is thus to devise a theoretically well-motivated registration model in a variational formulation, authorizing large and smooth deformations, while keeping the deformation map topology-preserving.
For problems involving large deformations but by a different approach from ours, we refer the reader to [5] in which Christensen et al. propose a viscous fluid model (not in a variational form). The objects to be matched are viewed as fluids evolving in accordance with the fluid dynamic Navier-Stokes equations. More precisely, given the force field f, the deforming image is considered to be embedded in a viscous fluid whose motion is governed by Navier-Stokes equations for conservation of momentum:

$$\mu \triangle v(x,t) + (\nu + \mu)\nabla\ (\nabla \cdot v(x,t)) = f(x, u(x,t)), \tag{1}$$

$$v(x,t) = \partial_t u(x,t) + \nabla u(x,t) \cdot v(x,t). \tag{2}$$

Equation (2), defining material derivative of the displacement field u, nonlinearly relates the velocity v and the displacement vector field. Constants μ and ν are viscosity coefficients of the fluid. One drawback of this method is the computational cost. Numerically, the image-derived force field $f(x, u(x,t))$ is first computed at time t. Fixing the force field f, linear equation (1) is solved for $v(x,t)$ numerically using the successive over-relaxation scheme. Then an explicit Euler scheme is used to advance u in time. These elements motivated us to propose an alternative framework requiring faster implementation.
We now depict the mathematical setting. Let us emphasize that the focus of the paper is on the mathematical presentation and well-posedness of a nonlinear elasticity-based registration model. Hence, the computational results are currently restricted to mouse brain gene expression data. Later work may go to a larger range of medical images with comparisons to reference methods.
Let Ω be a connected bounded open subset of $\mathbb{R}^2$ with Lipschitz boundary $\partial\Omega$ representing the reference configuration. Let us denote by $R : \bar{\Omega} \to \mathbb{R}$ the Reference image and by $T : \bar{\Omega} \to \mathbb{R}$ the Template image. It is assumed that T and R are both defined on the open and bounded domain Ω in the plane, a rectangle in general. Also, for theoretical and numerical purposes, we assume that T is compactly supported on Ω to ensure that $T \circ \varphi$ is always defined and we assume that T is Lipschitz continuous with Lipschitz constant $K > 0$. Also, R is supposed to be sufficiently smooth.

Let $\varphi : \bar{\Omega} \to \mathbb{R}^2$ be the deformation of the reference configuration. A deformation is a smooth enough mapping that is orientation-preserving and injective, except possibly on $\partial\Omega$. We also denote by u the associated displacement such that $\varphi = Id + u$. The deformation gradient is $\nabla\varphi = I_2 + \nabla u$, $\bar{\Omega} \to M_2(\mathbb{R})$, the set $M_2(\mathbb{R})$ being the set of all real square matrices of order 2. Thus the idea is to find a smooth deformation field φ such that the deformed Template matches the Reference. The model is phrased as a functional minimization problem whose unknown is φ. It combines a distance measure criterion chosen to be the L^2-norm of the difference between the deformed Template and the Reference, and a smoother on the deformation field. An alternative to intensity-based measures would consist in using information-theoretic-based matching measures such as mutual information (see [17]), suitable when dealing with images of different modalities or in comparing, for instance, the normalized gradient fields of both images (see [8] and [11]).
The proposed matching criterion is complemented by a regularizer on the deformation field φ. To allow large deformations, we introduce a nonlinear-elasticity-based smoother, the theory of linear elasticity being unsuitable in this case since assuming small strains and the validity of Hooke's law. We propose to view the shapes to be warped as isotropic, homogeneous, hyperelastic materials and more precisely as Saint Venant-Kirchhoff materials. Note that rubber, filled elastomers, biological tissues are often modeled within the hyperelastic framework, which motivates our modelling.
For the sake of completeness, we would like to refer the reader to previous works related to registration based on nonlinear elasticity principles. In [8], Droske and Rumpf address the issue of non-rigid registration of multi-modal image data. The matching criterion includes first order derivatives of the deformation and is complemented by a nonlinear elastic regularization based on a polyconvex stored energy function, which is different from our proposed approach. We also mention the combined segmentation/registration model introduced by Le Guyader and Vese ([12]) in which the shapes to be matched are viewed as Ciarlet-Geymonat materials, the works [2] and [16] for a variational registration method for large deformations (Large Deformation Diffeomorphic Metric Mapping - LDDMM), and refer to [20] for a much related work that also uses nonlinear elasticity regularization but implemented by the finite element method.

Before depicting the mathematical material, we review some fundamental concepts and notations (see [6] for further details). We recall that the right Cauchy-Green strain tensor is defined by $C = \nabla\varphi^T\nabla\varphi = F^TF \in \mathcal{S}^2$ with $\mathcal{S}^2 = \left\{A \in M_2(\mathbb{R}), A = A^T\right\}$, set of all real symmetric matrices of order 2. Physically, the right Cauchy-Green tensor can be interpreted as a quantifier of the square of local change in distances due to deformation. The Green- Saint Venant strain tensor is defined by $E = \frac{1}{2}\left(\nabla u + \nabla u^T + \nabla u^T\nabla u\right)$. Associated with a given deformation φ, it is a measure of the deviation between φ and a rigid deformation. We also need the following notations: $A : B = \mathrm{tr}A^TB$ the matrix inner product in $\mathbb{R}^2$ and $||A|| = \sqrt{A : A}$, matrix norm in $\mathbb{R}^2$ (Frobenius

norm). The stored energy of an isotropic, homogeneous, hyperelastic material, if the reference configuration is a natural state, is of the form:

$$W(F) = \widehat{W}(E) = \frac{\lambda}{2} \left(\operatorname{tr} E\right)^2 + \mu \operatorname{tr} E^2 + o\left(||E||^2\right), \; F^T F = I + 2E. \quad (3)$$

The stored energy function of a Saint Venant-Kirchhoff material is defined by $W_{SVK}(F) = \widehat{W}(E) = \frac{\lambda}{2} \left(\operatorname{tr} E\right)^2 + \mu \operatorname{tr} E^2$. The Saint Venant-Kirchhoff material is thus the simplest one that agrees with expansion (3). Moreover, as suggested in [12] or [14], from a numerical viewpoint, this modelling can generate high-magnitude deformations. These two arguments motivate the choice of the model we propose. At last, to ensure that the distribution of the deformation Jacobian determinants does not exhibit too large contractions or too large expansions, we propose to complement the model by a term controlling that the Jacobian determinant remains close to 1. In definitive, we propose to consider the following minimization problem:

$$\inf \left\{ I(\varphi) = \int_\Omega f(x, \varphi(x), \nabla\varphi(x))\, dx \; : \; \varphi \in \mathrm{Id} + W_0^{1,4}(\Omega, \mathbb{R}^2) \right\}, \text{ with} \quad (4)$$

$f(x, \varphi, \xi) = \frac{\nu}{2} \left(T(\varphi) - R\right)^2 + W_{SVK}(\xi) + \mu \left(\det\xi - 1\right)^2 = \frac{\nu}{2} \left(T(\varphi) - R\right)^2 + W(\xi)$. Also, $\varphi \in \mathrm{Id} + W_0^{1,4}(\Omega, \mathbb{R}^2)$ means that $\varphi = \mathrm{Id}$ on $\partial\Omega$ and $\varphi \in W^{1,4}(\Omega, \mathbb{R}^2)$. $W^{1,4}(\Omega, \mathbb{R}^2)$ denotes the Sobolev space of functions $\varphi \in L^4(\Omega, \mathbb{R}^2)$ with distributional derivatives up to order 1 which also belong to $L^4(\Omega)$. (We justify later that $W^{1,4}(\Omega, \mathbb{R}^2)$ is a suitable functional space for the considered problem).

2 Theoretical Results

2.1 Introduction of the Relaxed Problem

Function f in (4) fails to be quasiconvex, which raises a drawback of theoretical nature since we cannot obtain the weak lower semi-continuity of the functional. The idea is thus to replace problem (4) by the so-called associated relaxed problem (QP) formulated in terms of the quasiconvex envelope of f. Even though the original f is not quasiconvex and therefore in general the infimum of (4) is not attained, one has $\inf(4) = \inf(QP)$ and with some extra coercivity condition, the infimum of (QP) is reached. To the best of our knowledge, this technique of relaxation has never been used in the context of image registration before.
In the sequel, we start by establishing the explicit expression of the quasiconvex envelope of f and derive the associated relaxed problem.

Proposition 1. *The relaxed problem associated to (4) is defined by:*

$$\inf \left\{ \bar{I}(\varphi) = \int_\Omega Qf(x, \varphi(x), \nabla\varphi(x))\, dx \; : \; \varphi \in Id + W_0^{1,4}(\Omega, \mathbb{R}^2) \right\}, \quad (5)$$

with Qf *given by:* $Qf(x,\varphi,\xi) = \begin{cases} \dfrac{\nu}{2}\,(T(\varphi)-R)^2 + W(\xi) \;\; \textit{if } ||\xi||^2 \geq 2\dfrac{\lambda+\mu}{\lambda+2\mu}, \\ \dfrac{\nu}{2}\,(T(\varphi)-R)^2 + \Psi(det\xi) \;\; \textit{if } ||\xi||^2 < 2\dfrac{\lambda+\mu}{\lambda+2\mu}, \end{cases}$

and Ψ *the convex mapping such that* $\Psi : t \mapsto -\dfrac{\mu}{2}\,t^2 + \mu\,(t-1)^2 + \dfrac{\mu(\lambda+\mu)}{2(\lambda+2\mu)}$.

Proof. By definition (see Chapter 9, p. 432 of [7]), for almost every $x \in \Omega$ and for every $(\varphi,\xi) \in \mathbb{R}^2 \times \mathbb{R}^4$, the quasiconvex envelope of f with respect to the last variable is defined by:

$$Qf(x,\varphi,\xi) = \inf\left\{ \frac{1}{\text{meas}(D)} \int_D f(x,\varphi,\xi+\nabla\Phi(y))\,dy \;:\; \Phi \in W_0^{1,\infty}(D,\mathbb{R}^2) \right\},$$

$D \subset \mathbb{R}^2$ being a bounded open set. Consequently, in our case, $Qf(x,\varphi,\xi) = \dfrac{\nu}{2}\,(T(\varphi)-R)^2 + QW(\xi)$.
After some intermediate computations, one has $W(\xi) = \beta\,\left(||\xi||^2-\alpha\right)^2 + \Psi(\det\xi)$ with $\alpha = 2\dfrac{\lambda+\mu}{\lambda+2\mu}$ and $\beta = \dfrac{\lambda+2\mu}{8}$. From Theorem 3.1, p. 35 of [3], the result is straightforward. This judicious rewriting of $W(\xi)$ allows to see that $W^{1,4}(\Omega,\mathbb{R}^2)$ is a suitable functional space for φ: from generalized Hölder's inequality, if $\varphi \in W^{1,4}(\Omega,\mathbb{R}^2)$, then $\det\nabla\varphi \in L^2(\Omega)$.

It now remains to be proved that the infimum of (QP) is attained and that if $\bar{\varphi}$ is a solution of (5), then there exists a minimizing sequence $\{\varphi_\nu\}$ of (4) such that φ_ν weakly converges to $\bar{\varphi}$ and $I(\varphi_\nu) \to \bar{I}(\bar{\varphi})$. The solutions of (5) are considered as generalized solutions of (4), in the sense of weak convergence.

2.2 Existence of Minimizers for the Relaxed Problem and Relaxation Theorem

We now state the main theoretical result.

Theorem 1. *The infimum of (5) is attained. Let then* $\bar{\varphi} \in W^{1,4}(\Omega,\mathbb{R}^2)$ *be a minimizer of the relaxed problem (5). Then there exists a sequence* $\{\varphi_\nu\}_{\nu=1}^{\infty} \subset \bar{\varphi} + W_0^{1,4}(\Omega,\mathbb{R}^2)$ *such that* $\varphi_\nu \to \bar{\varphi}$ *in* $L^4(\Omega,\mathbb{R}^2)$ *as* $\nu \to \infty$ *and* $I(\varphi_\nu) \to \bar{I}(\bar{\varphi})$ *as* $\nu \to \infty$. *Moreover, the following holds:* $\varphi_\nu \rightharpoonup \bar{\varphi}$ *in* $W^{1,4}(\Omega,\mathbb{R}^2)$ *as* $\nu \to \infty$.

Proof. The proof of existence of minimizers of (5) is based on Theorem 8.29, p. 404, Chapter 8 of [7] (theorem due to Acerbi-Fusco [1] and Marcellini [15]) and rests upon the derivation of the following coerciveness and continuity double inequality:

$$C_3\,||\xi||^4 - C_4 \leq Qf(x,\varphi,\xi) \leq \left(\beta + \frac{\mu}{2}\right)||\xi||^4 + 4K^2||\varphi||^2 + C_2,$$

with C_3 a positive constant and with C_2 and C_4 two non-negative constants.
The proof of the second part of the theorem is based on Theorem 9.8, p. 432, Chapter 9 of [7].

We now propose a numerical method for the resolution of the relaxed problem. It is motivated by a related prior work by Negrón Marrero ([18]).

3 A Numerical Method of Resolution

3.1 Description and Analysis of the Proposed Numerical Method

In [18], Negrón Marrero describes and analyzes a numerical method that detects singular minimizers and avoids the Lavrentiev phenomenon for three dimensional problems in nonlinear elasticity. This method consists in decoupling the function φ from its gradient and in formulating a related decoupled problem under inequality constraint. In the same spirit, we introduce an auxiliary variable V simulating the Jacobian deformation field $\nabla\varphi$ (-the underlying idea being to remove the nonlinearity in the derivatives of the deformation-) and derive a functional minimization problem phrased in terms of the two variables φ and V. Nevertheless, our approach is different from the one in [18] in several points: in [18], the author focuses on the decoupled discretized problem (discretized with the finite element method - the paper provides neither numerical applications, nor details of the implementation) for which the existence of minimizers is guaranteed, while we consider the continuous problem. Also, the author assumes that the finite element approximations satisfy some convergence hypotheses. Moreover, in our case, less regularity is required for the formulation of the inequality constraint (see in particular the remark in the proof of Theorem 2).
The decoupled problem is thus defined by means of the following functional:

$$\bar{I}(\varphi, V) = \frac{\nu}{2}\int_{\Omega} \left(T(\varphi(x)) - R(x)\right)^2 \, dx + \int_{\Omega} \mathbb{W}(V)\, dx, \tag{6}$$

with

$$\begin{cases} \mathbb{W}(V) = W(V) \ \text{ if } ||V||^2 \geq \alpha = 2\dfrac{\lambda+\mu}{\lambda+2\mu} \quad \text{and} \\ \mathbb{W}(V) = \Psi(\det V) \ \text{ si } ||V||^2 < \alpha = 2\dfrac{\lambda+\mu}{\lambda+2\mu}. \end{cases}$$

Let us now denote by $\widehat{\mathcal{W}}$ the functional space defined by $\widehat{\mathcal{W}} = \mathrm{Id} + W_0^{1,2}(\Omega, \mathbb{R}^2)$ and by $\widehat{\chi}$, $\widehat{\chi} = \left\{V \in L^4(\Omega, M_2), \ \det V \in L^2(\Omega)\right\}$. The decoupled problem consists in minimizing (6) on $\widehat{\mathcal{W}} \times \widehat{\chi}$ under the constraint $||\nabla\varphi - V||^2_{L^2(\Omega,M_2)} \leq \varepsilon$, with $\varepsilon > 0$ and $\varepsilon \in]0, \varepsilon_0]$, $\varepsilon_0 > 0$ fixed. The idea is of course to let ε go to 0. Then the following theorem holds.

Theorem 2. *Let (ε_j) be a sequence such that $\lim\limits_{j\to+\infty} \varepsilon_j = 0$. Let also $(\varphi_k(\varepsilon_j), V_k(\varepsilon_j))$ be a minimizing sequence of the decoupled problem under inequality constraint defined with $\varepsilon = \varepsilon_j$. Then there exist a subsequence denoted by $\left(\varphi_{N(\varepsilon_{\Psi\circ g(j)})}(\varepsilon_{\Psi\circ g(j)}), V_{N(\varepsilon_{\Psi\circ g(j)})}(\varepsilon_{\Psi\circ g(j)})\right)$ of $(\varphi_k(\varepsilon_j), V_k(\varepsilon_j))$ and a minimizer $\bar{\varphi}$ of $\bar{I}(\varphi)$ ($\bar{\varphi} \in Id + W_0^{1,4}(\Omega, \mathbb{R}^2)$) such that:*

$$\lim_{j\to+\infty} \bar{I}\left(\varphi_{N(\varepsilon_{\Psi\circ g(j)})}(\varepsilon_{\Psi\circ g(j)}), V_{N(\varepsilon_{\Psi\circ g(j)})}(\varepsilon_{\Psi\circ g(j)})\right) = \bar{I}(\bar{\varphi}).$$

Proof. The proof is rather long so we only give the broad lines for the sake of conciseness. Let $\varepsilon > 0$ be given, $\varepsilon \in]0, \varepsilon_0]$, $\varepsilon_0 > 0$ fixed. There exists $\widehat{\varphi_\varepsilon} \in \mathcal{W} = \mathrm{Id} + W_0^{1,4}(\Omega, \mathbb{R}^2)$ such that:

$$\inf_{(\varphi,V)\in\widehat{\mathcal{W}}\times\widehat{\chi}} \bar{I}(\varphi, V) \leq \bar{I}(\widehat{\varphi_\varepsilon}, \nabla\widehat{\varphi_\varepsilon}) = \bar{I}(\widehat{\varphi_\varepsilon}) < \inf_{\varphi\in\mathcal{W}} \bar{I}(\varphi) + \varepsilon \leq \inf_{\varphi\in\mathcal{W}} \bar{I}(\varphi) + \varepsilon_0.$$

since $W^{1,4}(\Omega, \mathbb{R}^2) \subset W^{1,2}(\Omega, \mathbb{R}^2)$. Consequently,

$$\inf_{(\varphi,V)\in\widehat{\mathcal{W}}\times\widehat{\chi}} \bar{I}(\varphi, V) \leq \inf_{\varphi\in\mathcal{W}} \bar{I}(\varphi) + \varepsilon.$$

The second part of the proof consists in taking a sequence (ε_j) such that $\lim_{j\to+\infty} \varepsilon_j = 0$. We then consider a minimizing sequence denoted by $(\varphi_k(\varepsilon_j), V_k(\varepsilon_j))$ for the decoupled problem with $\varepsilon = \varepsilon_j$. We prove that there exists $N(\varepsilon_j) \in \mathbb{N}$ such that

$$\bar{I}\left(\varphi_{N(\varepsilon_j)}(\varepsilon_j), V_{N(\varepsilon_j)}(\varepsilon_j)\right) \leq \inf_{\varphi\in\mathcal{W}} \bar{I}(\varphi) + 2\varepsilon_j \leq \inf_{\varphi\in\mathcal{W}} \bar{I}(\varphi) + 2\varepsilon_0 < +\infty.$$

Using compactness arguments (among which Rellich-Kondrachov compact embedding theorem) and Theorem 1.14 from [7] (that states that if $u_\nu \rightharpoonup u$ in $W^{1,2}(\Omega, \mathbb{R}^2)$ then $\det \nabla u_\nu \rightharpoonup \det \nabla u$ in the sense of distributions), we demonstrate the existence of a subsequence $\left(\varphi_{N(\varepsilon_{\Psi\circ g(j)})}(\varepsilon_{\Psi\circ g(j)}), V_{N(\varepsilon_{\Psi\circ g(j)})}(\varepsilon_{\Psi\circ g(j)})\right)$ such that:

$$\inf_{\varphi\in\mathcal{W}} \bar{I}(\varphi) \leq \bar{I}(\bar{\varphi}) = \bar{I}(\bar{\varphi}, \nabla\bar{\varphi}) \leq \liminf_{j\to+\infty} \bar{I}\left(\varphi_{N(\varepsilon_{\Psi\circ g(j)})}(\varepsilon_{\Psi\circ g(j)}), V_{N(\varepsilon_{\Psi\circ g(j)})}(\varepsilon_{\Psi\circ g(j)})\right).$$

We remark that we gain some regularity: indeed, $\varphi_{N(\varepsilon_{\Psi\circ g(j)})}(\varepsilon_{\Psi\circ g(j)})$ is only $W^{1,2}(\Omega, \mathbb{R}^2)$ owing to the definition of the inequality constraint but when passing to the limit when $j \to +\infty$, we prove that $\varphi_{N(\varepsilon_{\Psi\circ g(j)})}(\varepsilon_{\Psi\circ g(j)}) \rightharpoonup \bar{\varphi}$ in $W^{1,2}(\Omega, \mathbb{R}^2)$ with $\bar{\varphi} \in W^{1,4}(\Omega, \mathbb{R}^2)$.

Inspired by this theoretical result, we now turn to the discretization of the considered problem.

3.2 Numerical Scheme

The algorithm requires the evaluation of the Template T at $\varphi(x)$. We thus assume that T is a smooth mapping that has been obtained by interpolating the image data provided on the grid. As an additional convention, T is supposed to vanish outside the domain, *i.e.*, $T(x) = 0$ if $x \notin \Omega$. As suggested by Modersitzki in [17], Chapter 3, subsection 3.6.1, for the interpolation stage we apply a multiscale interpolation technique which includes a weighting parameter controlling smoothness versus data proximity. Also, for the sake of optimization, a multilevel representation of the data is adopted (see Chapter 3, section 3.7 of [17]).
We now focus on the discretization of the problem. In this purpose, we denote by H_ϵ the regularized one-dimensional Heaviside function defined by $H_\epsilon : z \mapsto$

$\frac{1}{2}\left(1+\frac{2}{\pi}\text{Arctan}\frac{z}{\epsilon}\right)$.
As seen in subsection 3.1, we overcome the difficulty of minimizing the original relaxed functional (5) by introducing an auxiliary variable V mimicking $\nabla\varphi$. The nonlinear-elasticity-based regularizer is thus applied to V and no longer to $\nabla\varphi$. We thus propose to minimize:

$$\inf_{\varphi,V} \bar{\mathcal{J}}_\epsilon(\varphi,V) + \frac{\alpha'}{2}\int_\Omega ||V-\nabla\varphi||^2\,dx \ \text{, with} \tag{7}$$

$$\begin{aligned}\bar{\mathcal{J}}_\epsilon(\varphi,V) =& \frac{\nu}{2}\int_\Omega \left(T(\varphi)-R\right)^2\,dx + \int_\Omega W(V)\,H_\epsilon\left(||V||^2 - 2\frac{\lambda+\mu}{\lambda+2\mu}\right)\,dx \\ &+ \int_\Omega \Psi(\det V)\,H_\epsilon\left(2\frac{\lambda+\mu}{\lambda+2\mu} - ||V||^2\right)\,dx,\end{aligned} \tag{8}$$

where α' is a positive constant, big enough to ensure that V and $\nabla\varphi$ are sufficiently close in the sense of the L^2-norm.

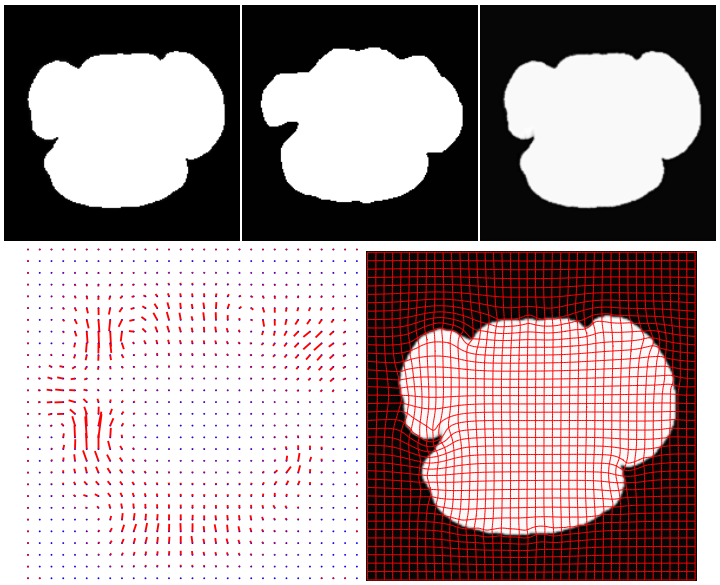
The system of optimality conditions is obtained. Numerically, the Euler-Lagrange equations in φ and V are solved using a gradient descent method, parameterizing the descent direction by an artificial time $t \geq 0$. Systems of 2 and 4 equations are obtained (solved by implicit and semi-implicit finite difference schemes - in particular, the unconditional stability of the numerical scheme in φ for fixed V is ensured), equipped with the boundary conditions $\varphi = \text{Id}$ on $\partial\Omega$.

3.3 Numerical Simulations and Discussion

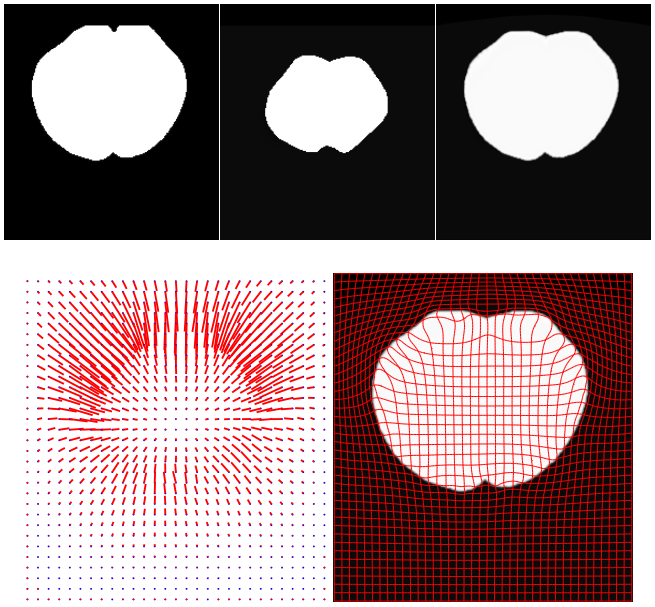
We conclude the paper by presenting some medical illustrations on real data. The method is proposed for mapping a 2D slice of mouse brain gene expression data (template T) to its corresponding 2D slice of the mouse brain atlas in order to facilitate the integration of anatomic, genetic and physiologic observations from multiple subjects in a common space. Since genetic mutations and knock-out strains of mice provide critical models for a variety of human diseases, such linkage between genetic information and anatomical structure is important. The data are provided by the Center for Computational Biology, UCLA. The mouse atlas acquired from the LONI database was pre-segmented. The gene expression data were segmented manually to facilitate data processing in other applications. Some studies have developed algorithms for automatically segmenting the brain area of gene expression data. The non-brain regions have been removed to produce better matching.
The deformation must remain physically and mechanically meaningful, and reflect material properties: self-penetration of the matter (indicating that the transformation is not injective) should be prohibited. In the studied functional, there is no term preventing the Jacobian determinant from tending to 0^+. An alternative to the straight penalization of the Jacobian determinant is proposed in [5]. Their fluid model is complemented by a regridding technique ensuring positivity of the Jacobian determinant. We also refer the reader to [13]: the given

topology-preserving model is motivated by a mathematical characterization of topology preservation for a deformation field mapping two subsets of $\mathbb{Z}^2$. Note that in the present work, we did not need to resort to such methods.

The model has been tested on 4 pairs, all of size 200×200 pixels (Figs. 1, 2) (except the second one of size 220×200). For the first two tests, the data are artificially obtained from the real data used later, which enables us to compare the results of the proposed method with some of those presented in [14]. The application related to Pair 1 (Fig. 1, first row) is similar to the one performed in [14]. As suggested in [14], classical non-parametric methods such as elastic registration, linear diffusion registration, biharmonic-regularizer-based registration (see [17] for further details) fail to correctly correlate the two images, some artifacts being visible on the deformed Template (under the left ear for the observer, where the deformations are the largest). Our method qualitatively performs as the one in [14] and produces a smooth deformation field, but unlike the one in [14], is theoretically well-motivated. Also, the algorithm has been optimized; first, by capitalizing on LAPACK and Basic Linear Algebra Subprogram (BLAS) routines. Second, by parallelizing the code (we focused on the OpenMP Application Program Interface). For all the applications, the ranges of the parameters are the same (see sub-captions of Figs. 1, 2). Parameter ν balancing the fidelity term is between 1.4 and 2, the Lamé coefficient λ is set to 10 (it has no physical meaning but is related to Poisson's ratio, measure of Poisson's effect which can be regarded as the ability of a material compressed in one direction to expand in the other (two) direction(s) - this choice of λ is not physically inconsistent), while the coefficient μ is between 3000 and 5000. Parameter α' is always set to 60000. This choice gives satisfactory results. For the two remaining pairs (Figs. 2), although the Template and the Reference are of different modalities, we still use the standard L^2 distance as measure for simplicity of calculation. The algorithm produces a smooth deformation between the gene and the atlas data with a regular distribution of the Jacobian determinants, without requiring regridding steps. Also, contrary to [12] in which the deformation is mainly present near the boundaries of the brain region, here the deformation is present both near the boundary and inside the brain region. If we keep comparing our results with the ones obtained by Le Guyader and Vese in [12], for Pair 2, 4 regridding steps were necessary and the range of the Jacobian determinants was $[0.005, 28.94]$ (in [12]). For Pair 4, 3 regridding steps were necessary and the range of the Jacobian determinants was $[0.008, 10.2]$. The range of the Jacobian determinants is also larger in [14] for the same applications. In our proposed method, the Jacobian determinant thus remains closer to 1. Nevertheless, this constraint can be weakened by downplaying the role of parameter μ. The proposed model thus easily handles large deformations without requiring prohibitive computational times. In our future studies, we will examine the registration accuracy of the model when ground truth is known and will adapt it to the 3D case.

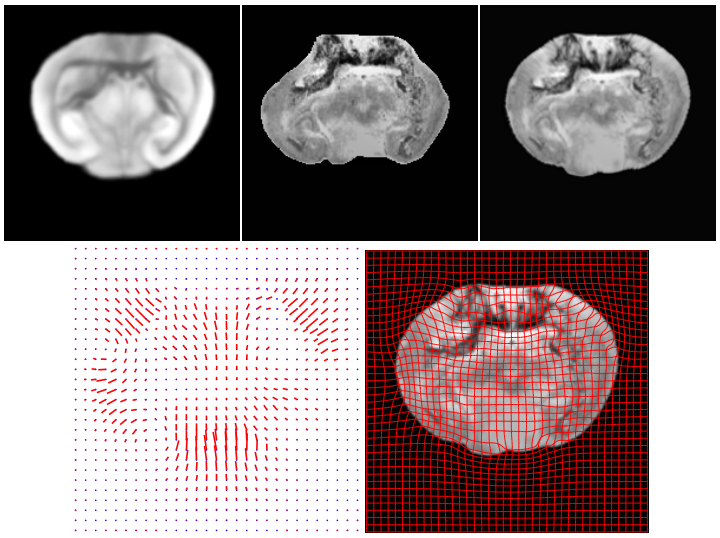


(a) Pair 1: $\nu = 2$, $\lambda = 10$, $\mu = 3000$, $\alpha' = 60000$, $\min\det\nabla\varphi = 0.09$, $\max\det\nabla\varphi = 2.47$, execution time: 1.8 s. From left to right, top to bottom: Reference R; Template T (mouse atlas and gene expression data); deformed Template; distortion map drawing the vectors from the grid points from the Reference image to non grid points after registration every 7 rows and columns; deformed grid.

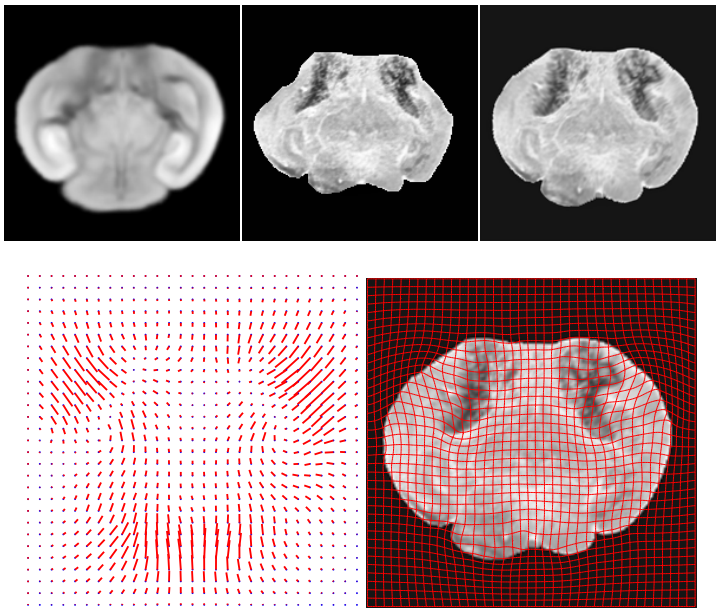


(b) Pair 2: $\nu = 1.8$, $\lambda = 10$, $\mu = 4000$, $\alpha' = 60000$, $\min\det\nabla\varphi = 0.02$, $\max\det\nabla\varphi = 2.21$, execution time: 8.4 s. From left to right, top to bottom: Reference R; Template T (mouse atlas and gene expression data); deformed Template; distortion map drawing the vectors from the grid points from the Reference image to non grid points after registration every 7 rows and columns; deformed grid.

Fig. 1: From top to bottom: Pair 1 till Pair 2.



(a) Pair 3: $\nu = 2$, $\lambda = 10$, $\mu = 3000$, $\alpha' = 60000$, $\min \det \nabla\varphi = 0.008$, $\max \det \nabla\varphi = 1.96$, execution time: 1.9 s. From left to right, top to bottom: Reference R; Template T (mouse atlas and gene expression data); deformed Template; distortion map drawing the vectors from the grid points from the Reference image to non grid points after registration every 7 rows and columns; deformed grid.



(b) Pair 4: $\nu = 1.4$, $\lambda = 10$, $\mu = 5000$, $\alpha' = 60000$, $\min \det \nabla\varphi = 0.02$, $\max \det \nabla\varphi = 1.56$, execution time: 9.5 s. From left to right, top to bottom: Reference R; Template T (mouse atlas and gene expression data); deformed Template; distortion map drawing the vectors from the grid points from the Reference image to non grid points after registration every 7 rows and columns; deformed grid.

Fig. 2: From top to bottom: Pair 3 till Pair 4.

References

1. Acerbi, E., Fusco, N.: Semicontinuity Problems in the Calculus of Variations. Arch. Rational Mech. Anal. 86, 125–145 (1984)
2. Beg, F., Miller, M., Trouvé, A., Younes, L.: Computing Large Deformation Metric Mappings Via Geodesic Flows of Diffeomorphisms. IJCV 61(2), 139–157 (2005)
3. Bousselsal, M.: Étude de Quelques Problèmes de Calcul des Variations Liés à la Mécanique. PhD thesis, Applied Mathematics, University of Metz, France (1993)
4. Broit, C.: Registration of Deformed Images. PhD thesis, Computer and Information Science, University of Pennsylvania, 1981.
5. Christensen, G.E., Rabbitt, R.D., Miller, M.I.: Deformable Templates Using Large Deformation Kinematics. IEEE Trans. Image Process. 5(10), 1435–1447 (1996)
6. Ciarlet, P.G.: Elasticité Tridimensionnelle. Masson (1985)
7. Dacorogna, B.: Direct Methods in the Calculus of Variations, Second Edition. Springer (2008)
8. Droske, M., Rumpf, M.: A Variational Approach to Non-Rigid Morphological Registration. SIAM J. Appl. Math. 64(2), 668–687 (2004)
9. Fischer, B., Modersitzki, J.: Fast Diffusion Registration. AMS Contemporary Mathematics, Inverse Problems, Image Analysis, and Medical Imaging 313, 11–129 (2002)
10. Fischer, B., Modersitzki, J.: Curvature based image registration. JMIV 18(1), 81–85 (2003)
11. Haber, E., Modersitzki, J.: Intensity Gradient Based Registration and Fusion for Multi-Modal Images. In Barillot, C., Haynor, D.R., Hellier, P. (eds.) MICCAI 2006. LNCS, vol. 3216, pp. 591–598. Springer, Heidelberg (2006)
12. Le Guyader, C., Vese, L.A.: A Combined Segmentation and Registration Framework with a Nonlinear Elasticity Smoother. Comput. Vis. Image Underst. 115(12), 1689–1709 (2011)
13. Le Guyader, C., Apprato, D., Gout, C.: On the Construction of Topology-Preserving Deformation Fields. IEEE T. Image Process. 21(4), 1587–1599 (2012)
14. Lin, T., Le Guyader, C., Dinov, I., Thompson, P., Toga, A., Vese, L.A.: Gene Expression Data to Mouse Atlas Registration Using a Nonlinear Elasticity Smoother and Landmark Points Constraints. J. Sci. Comput. 50, 586–609 (2012)
15. Marcellini, P.: Approximation of Quasiconvex Functions and Lower Semicontinuity of Multiple Integrals. Manuscripta Math. 51, 1–28 (1985)
16. Miller, M., Trouvé, A., Younes, L.: On the Metrics and Euler-Lagrange Equations of Computational Anatomy. Annu. Rev. B. Eng. 4, 375–405 (2002)
17. Modersitzki, J.: FAIR: Flexible Algorithms for Image Registration. SIAM, Philadelphia (2009)
18. Negrón Marrero, P. V.: A numerical method for detecting singular minimizers of multidimensional problems in nonlinear elasticity. Numerische Mathematik 58, 135–144 (1990)
19. Nocedal, J., Wright, S. J.: Numerical Optimization. Springer, New York (1999)
20. Rabbitt, R.D., Weiss, J.A., Christensen, G.E., Miller, M.I.: Mapping of Hyperelastic Deformable Templates Using the Finite Element Method. In: Proceedings SPIE, vol. 2573, pp. 252–265 (1995)

Computing Diffeomorphic Paths with Applications to Cardiac Motion Analysis

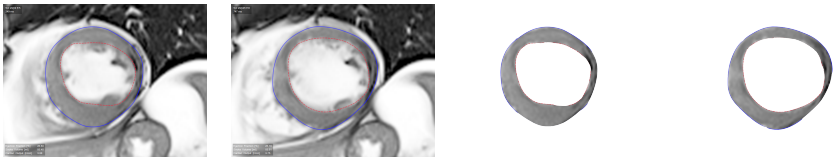
Dohyung Seo[1], Jeffrey Ho[2], Jay H. Traverse[3], John Forder[4]
, and Baba C. Vemuri[2]

[1] Department of Electrical and Computer Engineering, University of Florida
[2] Department of Computer and Information Sciences and Engineering
University of Florida
[3] Minneapolis Heart Institute Foundation at Abbott Northwestern Hospital
[4] Department of Biomedical Engineering, University of Florida

Abstract. This paper presents an application of a recently introduced novel framework for computing the diffeomorphic path between two given diffeomorphisms computed from two pairs of image frames in a motion sequence [1]. The specific application we address here is that of cardiac motion analysis. The framework involves a two-step algorithm wherein we first project the given pair of diffeomorphisms onto the space of densities – defined by quotienting out the volume-preserving diffeomorphisms considered as the nuisance parameters and then compute the geodesic path between the two points in this space. The second step lifts the aforementioned geodesic path back into the space of diffeomorphisms. The lifting problem is formulated as a quadratic programming problem with bilinear constraints and solved using the augmented Lagrangian method. Unlike the LDDMM-based techniques, this approach yields a path joining the given pair of diffeomorphisms that becomes a geodesic path upon quotienting out the volume-preserving diffeomorphisms. We use this path as the main feature to discriminate between ischemia patients with and without stem cell treatment. Such a discriminatory power can be very useful in the quantification of differences between controls and patients undergoing the treatment.

1 Introduction

While substantial advances in diagnosis and treatment have been made in the past half century, heart disease unfortunately is still a major cause of morbidity and mortality around the world. Among the novel therapeutic approaches that are currently under active investigation, stem cell-based therapy (e.g., [10]) holds a great promise for the future with revolutionary potential for significant improvement in patient prognosis and disease outlook. This paper proposes an application of a novel computational approach recently introduced in [1] to cardiac motion analysis with the aim of providing clinical researchers image processing tools that can extract subtle and useful features from the MR images for high-level clinical tasks such as patient classification and outcome assessment.



(a) (b) (c) (d)

Fig. 1. Top row (a) and (b):the left ventricular myocardium (LVM) at end-systole (ES) and end-diastole (ED). (c) and (d): LVM segmentations of (a) and (b), respectively.

The key idea here is to use smooth paths in the space of diffeomorphisms (of the image domain) as the characterization of the continuous variation in shape and volume of the patient's myocardium across cardiac cycles. More specifically, given a sequence of MR images of a cardiac cycle (Fig.1), we compute a diffeomorphic map between each image and a reference image, the image at the beginning of the cycle. These discretely sampled diffeormorphisms are interpolated to form a smooth path of diffeomorphisms from which useful high-level features can be extracted directly. Unfortunately, the diffeomorphisms are often noisy because the myocardium constantly undergoes small and random movement due to the pressure exerted from its surroundings. While these small random perturbations should be treated as nuisance parameters and modelled with diffeomorphisms close to the identity, there is an important equivalence among them: two small perturbations with the same myocardium volume should have the same clinical significance because it implies the same volume of ejecta in each case. This particular insight naturally leads to the idea of treating volume-preserving diffeomorphisms as the nuisance parameters, and the desired diffeomorphic paths should be smooth paths in an appropriate space that has "taken out" the effects of these nuisance parameters.

The above viewpoint admits a lucid and transparent formulation using the idea of quotient space. Specifically, let Ω denote a 3-D (image) domain with the volume form μ, and $\mathbf{Diff}(\Omega), \mathbf{Diff}(\Omega)_\mu$ the infinite-dimensional group of diffeomorphisms of Ω and its infinite-dimensional subgroup of volume-preserving diffeomorphisms, respectively. Mathematically, the quotient space $\mathbf{Diff}(\Omega)/\mathbf{Diff}(\Omega)_\mu$ is the space that has factored (taken) out the effect of $\mathbf{Diff}(\Omega)_\mu$, and there is a canonical projection map $\mathbf{Diff}(\Omega) \rightarrow \mathbf{Diff}(\Omega)/\mathbf{Diff}(\Omega)_\mu$. The quotient space has been studied in [6], and it can be identified with the space of density functions $\mathbf{Den}(\Omega)$ on Ω that can be canonically embedded as a convex subset of a sphere in the Hilbert space (See [6] for details). While $\mathbf{Diff}(\Omega)$ is a complicated space, the reduction $\mathbf{Diff}(\Omega) \rightarrow \mathbf{Diff}(\Omega)/\mathbf{Diff}(\Omega)_\mu = \mathbf{Den}(\Omega)$ provides us with the space $\mathbf{Den}(\Omega)$ of known and computationally tractable geometry. In particular, one straightforward way to exploit this reduction for computing diffeomorphic paths in $\mathbf{Diff}(\Omega)$ is to first compute a geodesic path in $\mathbf{Den}(\Omega)$ and then lift the path back into $\mathbf{Diff}(\Omega)$. Specifically, given the diffeomorphism ρ computed using any registration technique, our method will compute a path in the space of diffeomorphisms connecting ρ with the identity diffeomorphism (see next section for details) using a two-step algorithm that first projects ρ and the

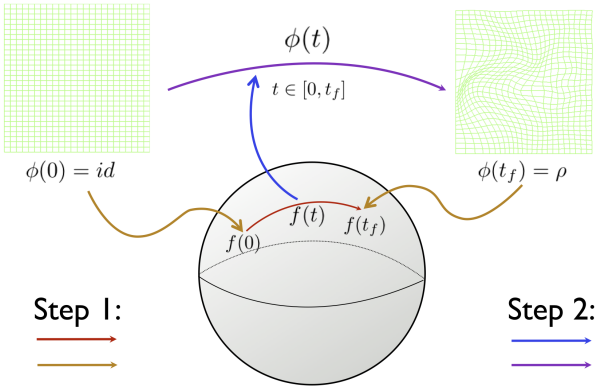


Fig. 2. Illustation of the two main steps in our algorithm for computing diffeomorphic paths. $f(0)$, $f(t_f)$ and $f(t)$ represent projected points of the two given diffeomorphisms and the great circle connecting the two projected points on the sphere, respectively.

identity onto the sphere and computes the unique geodesic joining the two projected points. This is followed by the second step that lifts the geodesic path on the density sphere back to $\mathbf{Diff}(\Omega)$, that is formulated as a quadratic programming problem with bilinear constraints which can be solved using the numerical algorithm proposed in [8] based on the augmented Lagrangian method. While the diffeomorphic path $\phi(t)$ computed by our method is not a geodesic path in $\mathbf{Diff}(\Omega)$, it nevertheless acquires a geodesic interpretation in the sense that by quotienting out the volume-preserving diffeomorphisms, the path $\phi(t)$ is indeed a geodesic path in the quotient space. Fig.2 provides a diagrammatic illustration of this algorithm.

In this paper, we apply our framework to cardiac motion analysis and the classification of ischemia patients. In particular, we show that features extracted from diffeomorphic paths are useful in classifying ischemia patients undergoing stem cell treatment and without treatment. Note that the classification problem does not require us to perform the second step defined above, namely, the lifting step. However, it is the lifting step that explicitly produces an interpolating path between the two given diffeomorphisms, and this step is necessary for visualizing the diffeomorphic path as well as for evaluating the quality of interpolation between the given time sequence of data that are needed in the validation step to assess the accuracy of the computed path.

We structure this paper somewhat differently by first presenting the details of our proposed algorithm and the experimental results on cardiac motion analysis in the following three sections. Related work will be discussed after the presentation of our framework and results, and the discussion will be centered on the relations between our work and that of Large Deformation Diffeomorphic Metric Mapping framework (LDDMM) [2–5] and iDiff framework introduced very recently in [7]. We end this paper with a short conclusion and a plan for future work.

2 Theory and Algorithm

This section presents the aforementioned two-step algorithm for estimating a path of diffeomorphisms connecting a given pair of diffeomorphisms ρ_1, ρ_2. For simplicity of exposition, we will assume that ρ_1 is **id**, the identity diffeomorphism and the necessary modification for general case is straightforward and simply involves a group operation applied to the two diffeomorphisms as $\rho_1^{-1} \circ \rho_1 = \mathbf{id}$ and $\rho = \rho_1^{-1} \circ \rho_2$.

The spaces of densities and diffeomorphisms: Our aim is to compute a diffeomorphic path, $\phi(t)$ between id and ρ such that $\phi(0) = id$ and $\phi(t = t_f) = \rho$ where, $0 \leq t \leq t_f$, and the image domain Ω will be considered as a compact subset of $\mathbb{R}^3$ with a given volume form $\mu = dxdydz$.

The projection map $\mathbf{\Phi}$ that maps $\mathbf{Diff}(\Omega)$ to the quotient space $\mathbf{Den}(\Omega)$ is given by,

$$\mathbf{\Phi} : \phi \rightarrow f = \sqrt{|D\phi|}, \tag{1}$$

where $|D\phi|$ is the determinant of the Jacobian of ϕ with respect to the volume form μ. We note that $\mathbf{\Phi}(\mathbf{id}) = 1$, the constant function one. The space of densities can be identified with an infinite-dimensional sphere of radius $\mu(\Omega)$, the volume of Ω,

$$\int_M \mathbf{\Phi}^2(\phi) d\mu = \mu(\Omega). \tag{2}$$

By scaling $\mathbf{\Phi}$ with $\sqrt{\mu(\Omega)}$, we can further identify $\mathbf{Den}(\Omega)$ with an infinite-dimensional unit sphere. More details are given in [6]. Once the two projected points on the sphere are identified, the unique geodesic path joining the two points f_1, f_2 can be readily determined using the formula [9].

$$f(t) = \frac{1}{\sin(\theta)} \left[\sin(\theta - t) f_1 + \sin(t) f_2 \right], \tag{3}$$

where $f_1 = \sqrt{|D\mathbf{id}|}$, $f_2 = \sqrt{|D\rho|}$, and θ is the angle between the two points on the sphere.

We remark that if the two given diffeomorphisms are already volume-preserving, then by definition, they project to the same point in $\mathbf{Den}(\Omega)$. Hence, their geodesic in $\mathbf{Den}(\Omega)$ is degenerate and consists of just one point, and the lifted diffeomorphic path in $\mathbf{Diff}(\Omega)$ is then a path consisting of only volume-preserving diffeomorphisms.

Given the geodesic path $f(t)$ on $\mathbf{Den}(\Omega)$ and representing diffeomorphic paths using 3-D deformation vector fields, $\phi(t) = (x + U(x, y, z, t), y + V(x, y, z, t), z + W(x, y, z, t))$, the diffeomorphic path $\phi(t) \in \mathbf{Diff}(\Omega)$ is required to satisfy the constraint,

$$\mathbf{\Phi}(\phi(t)) = f(t). \tag{4}$$

Since,

$$|D\phi(t)| = \begin{vmatrix} 1 + U(t)_x & U(t)_y & U(t)_z \\ V(t)_x & 1 + V(t)_y & V(t)_z \\ W(t)_x & W(t)_y & 1 + W(t)_z \end{vmatrix}, \tag{5}$$

we have

$$f(t)^2 = |D\phi(t)|/\mu(\Omega). \tag{6}$$

Eq.(5) gives the determinant of the Jacobian of ϕ, and $U(t)_i$, $V(t)_i$, and $W(t)_i$ denote the first-order derivatives with respect to $i \in \{x, y, z\}$. Our next step is to recover the deformation fields, $\phi(t)$ from Eq.(6). We remark that the proposed method does not explicitly perform spatial regularization on diffeomorphic paths, Instead, the diffeomorphic paths will be computed by lifting using the Jacobian constraints.

Path Lifting: The constraint provided by Eq.(6) does not determine a unique solution, and geometrically, this corresponds to the fact that Eq.(6) only requires $\phi(t)$ to lie on a $\mathbf{Diff}(\Omega)_\mu$-orbit in $\mathbf{Diff}(\Omega)$ parameterized by the point $f(t)$ in the sphere. Therefore, we will compute the lifted path in $\mathbf{Diff}(\Omega)$ using L^2-based regularization: L^2 smoothness of the deformation vector fields over time as the main regularization criterion. We note that while other regularization schemes are possible, e.g., using the Sobolev norm, we have chosen the L^2 regularization because of its computational simplicity. This leads to the following quadratic programming problem with bilinear constraints:

$$\begin{aligned} &\min \int \left|\frac{dU(x,y,z,t)}{dt}\right|^2 + \left|\frac{dV(x,y,z,t)}{dt}\right|^2 + \left|\frac{dW(x,y,z,t)}{dt}\right|^2 d\mu dt \\ &s.t.\ \ |D\phi(t)| = f(t)^2\mu(\Omega). \end{aligned} \tag{7}$$

We remark that the continuity of the computed diffeomorphic path is always assured because its projection on the sphere is a continuous path (in fact, a smooth geodesic path). Specifically, the domain of the optimization problem specified in Equation 7 can be taken to be the set $\mathcal{C}$ of all C^1-diffeomorphic paths[5] that project to the given geodesic on the sphere (i.e., satisfying the constraint Equation 4). The set $\mathcal{C}$ is nonempty because of the connectedness of both the fiber (the group of volume-preserving diffeomorphisms) and the base curve (the geodesic path on sphere). We also note that our diffeomorphic path is always computed for two given diffeomorphisms (**id** and ρ defined above) and not among several diffeomorphisms.

[5] Or some suitable completion of the set of all C^1-diffeomorphic paths under an appropriate Soblev norm.

3 Optimization

While Eq.(7) is formulated for continuous variables x, y, z and t, in practice, we have to work with discrete pixels and time, and its discrete version is given by,

$$\min_{U^t_{ijk}, V^t_{ijk}, W^t_{ijk}} \sum_{i,j,k,t} \left[\left| \frac{U^{t+1}_{ijk} - U^{t-1}_{ijk}}{2\delta x} \right|^2 + \left| \frac{V^{t+1}_{ijk} - V^{t-1}_{ijk}}{2\delta y} \right|^2 + \left| \frac{W^{t+1}_{ijk} - W^{t-1}_{ijk}}{2\delta z} \right|^2 \right] \delta t \delta x \delta y \delta z$$
$$s.t. \quad \begin{vmatrix} 1 + U^t_{x_{ijk}} & U^t_{y_{ijk}} & U(t)^t_{z_{ijk}} \\ V^t_{x_{ijk}} & 1 + V^t_{y_{ijk}} & V^t_{z_{ijk}} \\ W^t_{x_{ijk}} & W^t_{y_{ijk}} & 1 + W^t_{z_{ijk}} \end{vmatrix} - \mu(\Omega)(f^t_{ijk})^2 = 0, \tag{8}$$

where the superscript t and subscript i, j, k denote the discrete time and voxel indices respectively, and the subscripts x, y and z in the constraints denote the first-order derivatives w.r.t. x, y and z respectively. We set δx, δy and δz to be one, and therefore the volume $\mu(\Omega)$ is the image size.

Geometrically, the determinant of the Jacobian $|D\phi(t)|$ is the ratio of the change of volume elements by $\phi(t)$ at domain points and time t. In our numerical scheme, $|D\phi(t)|$ is formulated as the volume of a hexahedral cell composed with six-neighbour voxels which are displaced by the deformation vector fields. This 3-D formulation is adopted in Eq.(1) as well, and efficient ways to compute the volume are discussed in [11]. In [1], $|D\phi(t)|$ in 2-D was formulated as a quadrilateral composed with four neighbouring pixels, and the optimization problem was formulated as a quadratic programming problem with bilinear constraints. Now, in this 3-D problem, the formulation leads to the same optimization problem but with trilinear constraints: As shown in [12], considering $|D\phi(t)|$ to be the volume of a hexahedral cell and vectorizing all deformation vector fields, the optimization problem is formulated explicitly as follows:

$$\begin{aligned} &\min \frac{1}{2}\mathbf{U}^\intercal \mathbf{R}\mathbf{U} + \frac{1}{2}\mathbf{V}^\intercal \mathbf{R}\mathbf{V} + \frac{1}{2}\mathbf{W}^\intercal \mathbf{R}\mathbf{W} + \mathbf{B}_u^\intercal \mathbf{U} + \mathbf{B}_v^\intercal \mathbf{V} + \mathbf{B}_w^\intercal \mathbf{W} \\ &s.t.\ \mathbf{h} = \mathbf{c} + \mathbf{C}_V\mathbf{V} + \mathbf{C}_W\mathbf{W} + (\mathbf{C}_U + \mathbf{D}_U(\mathbf{V}) + \mathbf{D}_U(\mathbf{W}) + \mathbf{E}_U(\mathbf{V}, \mathbf{W}))\mathbf{U} = 0 \\ &\quad or\ \mathbf{h} = \mathbf{c} + \mathbf{C}_U\mathbf{U} + \mathbf{C}_W\mathbf{W} + (\mathbf{C}_V + \mathbf{D}_V(\mathbf{U}) + \mathbf{D}_V(\mathbf{W}) + \mathbf{E}_V(\mathbf{U}, \mathbf{W}))\mathbf{V} = 0 \\ &\quad or\ \mathbf{h} = \mathbf{c} + \mathbf{C}_U\mathbf{U} + \mathbf{C}_V\mathbf{V} + (\mathbf{C}_W + \mathbf{D}_W(\mathbf{U}) + \mathbf{D}_W(\mathbf{V}) + \mathbf{E}_W(\mathbf{U}, \mathbf{V}))\mathbf{W} = 0, \end{aligned} \tag{9}$$

Let us suppose that the size of domain is $M \times N \times L$. Then, in the above, $\mathbf{U}^\intercal = (\mathbf{U}^\intercal(t = 2), ..., \mathbf{U}^\intercal(t = T - 1))$, and $\mathbf{B}_u^\intercal = (\mathbf{U}^\intercal(t = 1), \mathbf{0}^\intercal, \mathbf{U}^\intercal(t = T))$ where $\mathbf{0}$ is a vector of zeros with length $M \times N \times L \times (T - 4)$ if we have a time sequence $t \in [1, T]$ including two boundary conditions namely, $\phi(0) = id$ and $\phi(T) = \rho$. $\mathbf{V}$ ($\mathbf{W}$) and $\mathbf{B}_v$ ($\mathbf{B}_w$) are defined in the same way. In Eq.(9), $\mathbf{R}$ is the matrix for the quadratic component and its size is $M \times N \times L \times (T - 2)$. $\mathbf{C}_U\mathbf{U}$, $\mathbf{C}_V\mathbf{V}$ and $\mathbf{C}_W\mathbf{W}$ correspond to the linear terms of $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{V}$ in the volume formulation of the hexahedral cells, respectively. $\mathbf{D}$'s and $\mathbf{E}$'s correspond to the bilinear and trilinear terms after factoring out $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{W}$, respectively. Finally, $\mathbf{c}$ is the vectorization of $\mu(\Omega)(f^t_{ijk})^2$ in Eq.(8) plus a vector of ones and $\mu(\Omega) = (M-1)(N-1)(L-1)$. More details are provided in Appendix E of [12]

We solve the optimization using the augmented Lagrangian method with penalty [8] by iteratively solving for the three blocks of variables $\mathbf{U}, \mathbf{V}, \mathbf{W}$, fixing

two of these while optimizing the rest. For instance, with fixed $\mathbf{V}$ and $\mathbf{W}$, the problem is given by,

$$\begin{aligned} &\min \frac{1}{2}\mathbf{U}^\intercal \mathbf{R}\mathbf{U} + \mathbf{B}_u^\intercal \mathbf{U} + \lambda^\intercal \mathbf{h} + \frac{1}{2}c||\mathbf{h}||^2 \\ &s.t.\ \ \mathbf{U} \in \mathbb{R}^m, \end{aligned} \tag{10}$$

where, $\lambda \in \mathbb{R}^m$ is the Lagrange multiplier, $c \to \infty$ and $m = M \times N \times L \times (T-2)$. This is an unconstrained optimization problem and can be rewritten more clearly as,

$$\begin{aligned} &\min \frac{1}{2}\mathbf{U}^\intercal(\mathbf{R} + c\mathbf{H}_u^\intercal \mathbf{H}_u)\mathbf{U} + (\mathbf{B}_u + \frac{1}{2}c((\mathbf{G}_u^\intercal \mathbf{H}_u)^\intercal + \mathbf{H}_u^\intercal \mathbf{G}_u) + (\lambda^\intercal \mathbf{H}_u)^\intercal)^\intercal \mathbf{U} \\ &s.t.\ \mathbf{U} \in \mathbb{R}^m, \end{aligned} \tag{11}$$

where, $\mathbf{H}_u = \mathbf{C}_U + \mathbf{D}_U(\mathbf{V}) + \mathbf{D}_U(\mathbf{W}) + \mathbf{E}_U$ and $\mathbf{G}_u = \mathbf{c} + \mathbf{C}_V\mathbf{V} + \mathbf{C}_W\mathbf{W}$. That is, $\mathbf{h} = \mathbf{H}_u\mathbf{U} + \mathbf{G}_u$.

4 Experiments

We have acquired cardiac MR scans of patients with acute myocardial infraction (AMI), and they were randomized into one of two groups: control and treatment. The treatment group received endothelial progenitor stem cell (EPCs) treatment, while the control received a cell-free infusion. Cardiac MRI exams were performed at baseline which was prior to EPC therapy, and following six months post-treatment [10]. Generally, cardiac analysis involves measuring the average wall motion, or wall thickening reported for each of the individual segments of hearts. We propose a patient classification method (between control and treatment groups) by measuring changes of diffeomorphic paths over six months. These diffeomorphic paths are the diastolic filling motion of the left ventricular myocardium (LVM) computed using our method. For the experiments, we have expert-segmented LVM's from the cardiac MR scans, and examples of LVM and its segmentation are shown in Fig.1.

Data Acquisition: The cardiac data was acquired along the short-axis of the heart using a cine multislice sequence using a Siemens Avanto 1.5T whole body MRI scanner. Twenty-five phases were collected throughout the cardiac cycle using a TrueFISP (a balanced coherent gradient sequence using Fast Imaging with Steady state free Precession), with an echo time of 1.13ms and an apparent repetition time of 71.82ms. Data was collected using a total of 161 phases encoded into a matrix of 256×151, yielding an in-plane resolution of $1.71875 \times 1.71875\text{mm}^2$. Slice thickness was 7mm, with a slice separation of 10mm. Coverage of the myocardium extended from above the valve plane at the base, to below the apex throughout the cardiac cycle.

In the first experiment, we compare the interpolated paths between end-systole (ES) and end-diastole (ED) using our method and LDDMM, given the ground truth. For this comparison, we choose eighteen intermediate time points

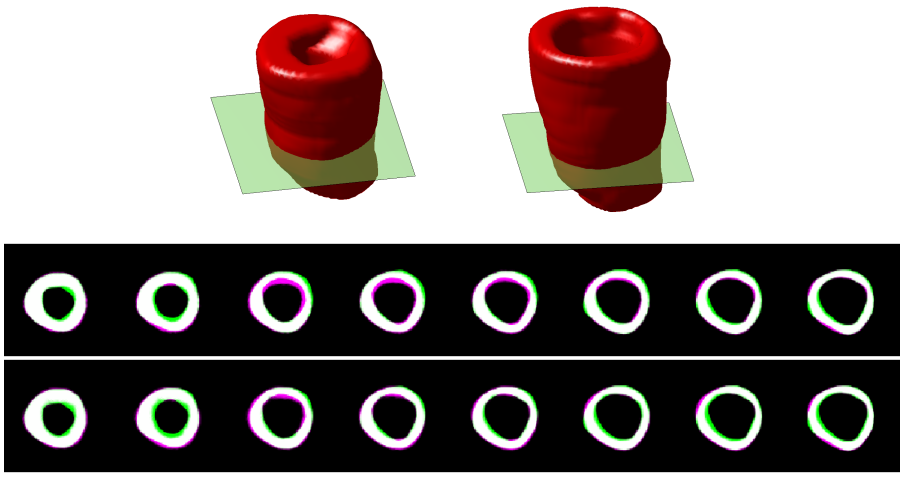


Fig. 3. Comparison between the ground truth and estimated temporal changes of the cardiac chamber from ES to ED along the short-axis in 3-D. A 2-D slice shown in green color plane is chosen to depict the changes over the cardiac cycle in the bottom two rows. Top: 3-D shapes of LVM at ES (left) and ED (right). Middle and bottom: computed changes using LDDMM and our method respectively. The white region represents the intersection of the ground truth and estimates, and green and magenta represent voxels belonging to the ground truth and estimates respectively.

along the interpolated path at which the comparison is performed. This comparison is repeated for five different patients. A visual comparison is provided in Fig,3. The images in the top row in Fig.3 represent 3-D shapes of LVM at ES (left) and ED (right). For visual comparison, first, we choose a short-axis which is presented as green planes in the top images and compare the ground truth of the temporal change of this 2-D slice (from 3-D) with computed ones using LDDMM and our method respectively. The middle and bottom rows in Fig.3 show the results from LDDMM and our method respectively. The white region represents the intersection of the ground truth and estimates from the applied algorithm, and green and magenta represent voxels which belong to the ground truth and estimates. It is hard to tell which method estimates the ground truth with higher quality (even though we can recognize that in the middle frames the shapes from our method fit better to the ground truth than the ones from LDDMM). However, quantitative comparison shows clear difference between our method and LDDMM: We report the Dice coefficients between the ground truth and the interpolated LVM images obtained using our method and the LDDMM method. The results tabulated in Table 1 show that the proposed method outperforms LDDMM using the Dice coefficient as the evaluation metric.

Ischemia Patient Classification: In the second experiment, we classify the two groups of ischemia patients described in the previous sections. The stem-cell therapy used in the treated patient group is supposed to improve the diastolic filling motion of the left ventricular myocardium (LVM); therefore, an approach

Table 1. The Dice coefficients of the estimated intermediate cardiac frames using our method and the LDDMM.

	Average	Median	Variance
Our method	0.912	0.908	0.0009
LDDMM	0.901	0.897	0.0018

for classification between the two groups is to measure how the diastolic filling motion of LVM changes over six months with/without EPC therapy. The cycle of LVM diastolic filling motion is from end-systole (ES) to end-diastole (ED), and we take the images of segmented LVM at ES and ED as the I_s's and I_t's. We then compute the diffeomorphic path $\phi(t)$ of nine subjects in the control group and 21 in the treated group, and all paths have ten intermediate time points between ES and ED. Next, we express $U(t)$,$V(t)$, and $W(t)$ of $\phi(t) = (x+U(t), y+V(t), z+W(t)$ in a discrete sine and cosine (DSC) basis. We collect all the DSC coefficients of U,V and W into a single vector for each patient, and the motion change of LVM over six months for each patient is measured using the L_2 distance of DSC coefficients between baseline scan and 6-months scan. We use this as the feature for classification. In this experiment, we use the Support Vector Machine (SVM) as the classifier with polynomial kernel of degree three. The number of DSC basis elements for each component of $\phi(t)$ is five, and we use a leave-one-out cross validation. Table 2 shows the classification results using our method and LDDMM, and it is evident that our method significantly outperforms LDDMM. Note that the LDDMM-based method results in lower classification rates because of the inherent hard parameter selection problem associated with it. In contrast, our method does not have any such parameter selection issue as evident from our quadratic programming formulation.

Table 2. Classification scores using Support Vector Machine

Validation method	Classification score			
	The proposed method		LDDMM	
	Control	Treatment	Control	Treatment
Leave-one-out	77.78%	85.71%	66.67%	69.56%

5 Discussion

One of the most intensively studied frameworks for computing diffeomorphic path is the Large Deformation Differomorphic Metric Mapping framework (LDDMM) [2–5]. In LDDMM, the diffeomorphic path ϕ is obtained by solving the following equation

$$\frac{d}{dt}\phi(t, \cdot) = v(t, \phi(t, \cdot)), \tag{12}$$

where $v(t, \cdot)$ is a time-varying velocity field defined on Ω. For computing a diffeomorphic path joining two diffeomorphisms (e.g., **id** and ρ), ϕ is required to satisfy

the boundary condition $\phi(1, \cdot) = \rho$. In the registration context, LDDMM's objective function almost always has two terms: the data term that matches images and the regularization term that is based on the magnitude of the time-varying vector field v measured in an appropriate norm. Therefore, the objective function requires at least one coupling constant to modulate the contributions of the two terms. In contrast, our method detaches the registration process from the process that computes the diffeomorphic path. Consequently, it does not require a coupling constant whose determination is often highly nontrivial, and the experimental comparisons presented above can be partially attributed to the (potentially) non-optimal coupling constants used for LDDMM. Furthermore, it is also highly nontrivial to modify the LDDMM framework for minimizing the effect of volume-preserving diffeomorphisms (as nuisance parameters). In fact, the diffeomoprhic paths computed by LDDMM and our method are generally different, and projecting LDDMM's diffeomorphic path onto the quotient space $\mathbf{Diff}(\Omega)/\mathbf{Diff}(\Omega)_\mu$ in general does not produce a geodesic on $\mathbf{Den}(\Omega)$. The inability of LDDMM to deal with the nuisance parameters at least in part explains the difference in classification performance reported in Table 2.

The space of diffeomorphisms $\mathbf{Diff}(\Omega)$ is infinite-dimensional, which readily explains the difficulty of computing diffeomorphic paths. Using the projection map $\mathbf{\Phi}$ (Equation 1), the tangent space of $\mathbf{Diff}(\Omega)$ at ρ has the following natural decomposition (given a Riemannian metric on $\mathbf{Diff}(\Omega)$)

$$\mathbf{T}_\rho\mathbf{Diff}(\Omega) = \mathbf{T}_{\mathbf{id}}\mathbf{Diff}(\Omega)_\mu \oplus \mathbf{T}_{\mathbf{\Phi}(\rho)}\mathbf{Den}(\Omega), \tag{13}$$

where $\mathbf{T}_{\mathbf{id}}\mathbf{Diff}(\Omega)_\mu$ is the tangent space of $\mathbf{Diff}(\Omega)_\mu$ at the identity $\mathbf{id}$ and $\mathbf{T}_{\mathbf{\Phi}(\rho)}\mathbf{Den}(\Omega)$ is the tangent space of $\mathbf{Den}(\Omega)$ at $\mathbf{\Phi}(\rho)$. Direct minimization of the norm of the time-dependent velocity field v as in LDDMM requires considering the entire tangent space $\mathbf{T}_\rho\mathbf{Diff}(\Omega)$, and it is often fraught with problems, not the least because there are no guarantees on the quality of the solution. Compared with LDDMM, the minimization specified in Equation 7 requires only $\mathbf{T}_{\mathbf{id}}\mathbf{Diff}(\Omega)_\mu$-component as the $\mathbf{T}_{\mathbf{\Phi}(\rho)}\mathbf{Den}(\Omega)$-component is fixed by the Jacobian constraint. Therefore, our method reduces the dimension by half, and although still infinite-dimensional, the reduction provides a more restricted domain for optimization that can be easier to analyze and optimize.

Recently, after submitting our work, we were made aware of a novel framework that shares a significant conceptual similarity with ours. The paper [7] introduces the notion of irrotational diffeomorphisms **iDiff** based on Brenier's polar factorization theorem for diffeomorphisms: an arbitrary diffeormorphism $\varphi = S \circ \psi$ in $\mathbb{R}^n$ can be uniquely factorized as the composition of a volume-preserving diffeomorphism S and an irrorational diffeomorphism $\psi \in \mathbf{iDiff}(\mathbb{R}^n)$. In particular, $|D\varphi| = |D\psi|$, and $\mathbf{\Phi}(\varphi) = \mathbf{\Phi}(\psi)$. Using the language of quotient space, the result in [7] can be interpreted as an explicit construction of a section of the projection $\mathbf{\Phi}$ in Equation 1. In other words, the projection map $\mathbf{\Phi}$ restricted to the irrotational diffeomorphisms **iDiff** is bijective onto the image of $\mathbf{\Phi}$. With this interpretation, the appearance of the square-root of the deter-

minant of the Jacobian in the following main formula (Equation 17 in [7])

$$P : \mathbf{iDiff}(\mathbb{R}^d) \to L^2(\mathbb{R}^d) \ , P(\psi) = 2\sqrt{|D\psi| - 1} = 2(\mathbf{\Phi}(\psi) - 1) \qquad (14)$$

becomes less surprising. Furthermore, at tangent space-level, the factorization corresponds to the Helmholtz-Hodge decomposition of the vector fields that decompose the tangent space $\mathbf{T_{id}Diff}(\Omega)$ as the direct sum of divergence-free vector fields and their orthogonal complement, which corresponds to the decomposition in Equation 13 above with $\mathbf{T_{id}Diff}(\Omega)_\mu$ identified with divergence-free vector fields. However, an important and subtle difference is that [7] works with diffeomorphisms on $\mathbb{R}^n$ with compact support while we work with diffeomorphisms on a connected compact domain in $\mathbb{R}^n$. The compactness of the domain implies that it has finite volume, and this translates into the fact that the image of $\mathbf{\Phi}$ is on a sphere of finite radius (Equation 2) with nonzero curvature. The above map $P : \mathbf{iDiff}(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ can be interpreted geometrically as spherical chords centered at $\mathbf{\Phi}(\mathbf{id})$, and due to the curvature of the sphere, the image of P is never an open neighborhood of zero in $L^2(\mathbb{R}^d)$ when the domain has finite volume. However, the flatness result in [7] can be readily seen (as a limit) from the following heuristic argument[6].

Let $n = 2$ and $D_1 \subset D_2 \subset ... \subset \mathbb{R}^2$ denote a nested sequence of domains where $D_k = \{(x, y) \in \mathbb{R}^2 \| \ |x| \leq k, |y| \leq k \}$. Each pair $(D_k, \mu), k = 1, 2, ...$ is a compact domain in $\mathbb{R}^2$ with area (volume) $4k^2$. Correspondingly, we will denote $\mathbf{iDiff}_1 \subset \mathbf{iDiff}_2 ... \subset \mathbf{iDiff}(\mathbb{R}^2)$ the nested sequence of subsets of $\mathbf{iDiff}(\mathbb{R}^2)$, where $\mathbf{iDiff}_n$ contains irrotational diffeomorphisms with support in D_n. The set of irrotational diffeomorphisms with compact support is then the union of $\mathbf{iDiff}_n$:

$$\mathbf{iDiff}(\mathbb{R}^2) = \bigcup_{n=1}^{\infty} \mathbf{iDiff}_n.$$

Therefore, under the projection map $\mathbf{\Phi}$, $\mathbf{iDiff}_k$ maps to a convex set on the sphere of radius $4k^2$. Recall that the curvature of a sphere with radius $4k^2$ is $\frac{1}{4k^2}$, and this shows that as $k \to \infty$, $\mathbf{\Phi}$ maps $\mathbf{iDiff}(\mathbb{R}^2)$ to a sphere with infinite radius, i.e., a space with zero curvature.

6 Conclusions

In this paper, we have applied the novel framework introduced in [1] for computing a diffeomorphic path interpolating two given diffeomorphisms from a cardiac cycle. We have argued that the volume-preserving diffeomorphisms should be considered as nuisance parameters, and clinically relevant diffeomorphic paths should be induced from smooth paths in the corresponding quotient space, taking out the effect of the nuisance parameters. The diffeomorphic path computed by our method enjoys the property that its projection onto the quotient space is indeed a geodesic. The proposed algorithm has two main steps. First, it projects

[6] A more detailed comparison of these two approaches will be presented in a forthcoming paper

the two given diffeomorphisms computed from any image registration process onto the quotient space that has been shown to be a convex subset of the infinite-dimensional sphere, and it computes the geodesic between these two projected points in closed form. Second, it lifts this geodesic path on the sphere back to the space of diffeomorphisms using an L_2-regularization that results in a quadratic programming problem with trilinear constraints for 3-D problems that can be efficiently solved using the augmented Lagranigian method.

We have applied our algorithm to two cardiac motion analysis experiments. In the first experiment, the result has shown that the proposed method performs better than LDDMM in estimating the diffeomorphic path between end-systole (ES) and end-diastole (ED). The result from the second experiment has demonstrated the features extracted from the diffeomorphic paths are relevant and useful for patient classification (treated and untreated groups). Our future research will focus on developing more statistical analysis methods in the quotient space described in this paper.

Acknowledgement This research was in part funded by the NIH grant NS066340 to BCV.

References

1. Seo, D., Ho, J., Vemuri, B.C.: Computing Diffeomorphic Paths for Large Motion Interpolation. *Proc.* CVPR, 1227-1232, (2013).
2. L. Younes: Shapes and Diffeomorphisms. Springer, Heidelberg (2010).
3. Beg, M.F., Miller, M.I., Trouv, A., Younes, L.: Computing Large Deformation Metric Mappings vis Geodesic Flows of Diffeomorphisms. IJCV **61**(2), 139-157, (2005).
4. Younes, L., Arrate, F., Miller, M.I.: Evolutions Equations in Computational Anatomy. Neuroimages, 45, 40-50, (2009).
5. Sommer, S., Nielsen, M., Lauze, F., Pennec, X.: A Multi-scale Kernel Bundle for LDDMM: Towards Sparse Deformation Description Across Space and Scales. IPMI 2011, LNCS, vol. 6081, pp. 624-635, (2011).
6. Khesin, B., Lenells, J., Misiolek, G., Preston, S.C.: Geometry of Diffeomorphism Groups, Complete Integrability and Geometric Statistics. arXiv:1105.0643 (2011).
7. Hinkle, J., Joshi, S.: IDiff: Irrotational Diffeomorphisms for Computational Anatomy. In: Proc. IMPI, 754-765, (2013).
8. Carotenuto, L., Raiconi, G.: On the Minimization of Quadratic Functions with Bilinear Constraints via Augmented Lagrangians. Journal of optimization theory and applications, **55**, 1, 23-36, (Oct. 1987).
9. Srivastava, A., Jermyn, I., Joshi, S.H.: Riemannian Analysis of Probability Density Functions with Application in Vision. In: Proc. CVPR, 18-23, (2007).
10. Petersen, J.W., Forder, J.R., Thomas, J.D., Moye, L.A., Lawson, M., Loghin, C., Traverse, J.H., Baraniuk, S., Silva, G., Pepine, C.J.: Quantification of Myocardial Segmental Function in Acute and Chronic Ischemic Heart Disease and Implications for Cardiovascular Cell Therapy Trials: a Review from the NHLBI-Cardiovascular Cell Therapy Research Network, JACC Cardiovasc Imaging, **4**(6), 71-9, (2011).
11. Grandy, J.: Efficient computation of volume of hexaheral cells, No. UCRL-ID-128886, Lawrence Livermore National Lab., CA (United States), (1997).
12. Seo, D.: A Geometric Approach to Image Matching and Synthesis of Diffeomorphic Paths. Ph.D dissertation, University of Florida, (2013).

The Shape Collapse Problem in Image Registration

Oguz C. Durumeric[1], Ipek Oguz[2], and Gary E. Christensen[2,3]

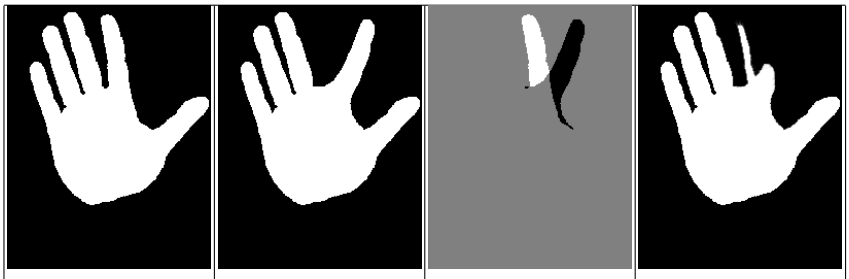
Departments of [1]Mathematics, [2] Electrical and Computer Engineering and [3] Radiation Oncology, The University of Iowa

Abstract. This paper investigates the shape collapse problem in non-rigid image registration. The shape collapse problem is the situation when an appendage of a deforming object does not overlap with the target shape and collapses to a set of zero measure during the registration process. The dual problem occurs when a new appendage grows out of the object to match the target shape. In both cases, the estimated correspondence between the source and target objects is often undesirable. The shape collapse problem is caused by deforming the moving image in the gradient direction of the similarity cost and affects both small and large deformation registration algorithms. Minimizing a registration cost function by following the similarity-cost gradient drives the registration to a local energy minimum and does not permit an increase in energy to ultimately reach a lower energy state. Furthermore, once an object collapses locally, it has zero measure under the similarity cost in this region and is permanently stuck in a local minimum. This paper presents a criterion for detecting image regions that will collapse if the similarity cost gradient direction is followed during optimization. This criterion is based on the skeletal points of the moving image in the symmetric difference of the original two binary images. Experimental results are presented that demonstrate that the shape collapse problem can be detected before registration.

Keywords: image registration, skeleton, shape collapse, collapse points, predicting regions of collapse

1 Introduction

A common problem that affects both small and large-deformation image registration algorithms is when a part of an object collapses during the registration process. The problem can be illustrated by registering images of hands in which four of five fingers overlap (see Fig. 1). The shape collapse occurs during registration when the index finger in the source image collapses into a thin set. At the same time, a new finger grows out of the palm of the hand to match the finger in the target hand. The biologically relevant transformation in this case would be instead to shift the finger to match its pose in the target image. This solution is difficult to reach with greedy optimization algorithms often used in image registration, because collapsing the index finger in the source image to a



(a) Moving image | (b) Target image | (c) Difference before registration | (d) Deformed moving image

Fig. 1. Example of an undesirable shape collapse during image registration.

thin set reduces the registration similarity cost immediately, as opposed to rotating the finger which may increase the cost initially. Note that Fig. 1-d shows the final result for a small deformation image registration and only a partial deformation for a large deformation registration. Additionally, once a region of source image collapses during the registration process, it remains collapsed, i.e. a collapsed region in the source image corresponds to a local minimum of the registration cost function.

We are not aware of others that have studied the shape collapse problem in image registration. The distance function from smooth submanifolds and its cut loci have been thoroughly studied by many in differential geometry. See [2] for basic introduction and a review. The notion of skeleton generalizes the notion of cut locus when the boundary is not differentiable. Skeletonization and medial representations have been thoroughly studied; a good overview can be found in [3].

In this paper we present a novel method for a priori detection of object points where shape collapse is likely to occur and predicting the collapse loci. We provide several examples where our algorithm successfully predicts collapse loci. Detecting collapse loci is important for the alignment of sulci in brain image registration, which is similar to the alignment of fingers in the hand example discussed above. In this case, once we detect collapsing sulci, we can devise solutions to this problem by, for instance, enforcing the sulcus to retract into or sliding along the cortex rather than collapsing. This may also be used as a potential method to detect different folding patterns between the brains. The approach to handle collapsing is different based on the application domain, and as such, is beyond the scope of this paper.

2 Mathematical Formulation

The ideas presented in this paper are independent of space dimension. For example, the area-decreasing discussion below becomes volume-decreasing in 3-D. Likewise, the definition of skeleton below allows for curves and sheets in 3-D.

Let Ω be a subset of $\mathbb{R}^n$, typically with $n = 2$ or 3. Let $I_i : \Omega \to \mathbb{R}$, $i = 1, 2$ be two binary images to be registered, where I_1 denotes the moving image and I_2 denotes the target image. We denote the foreground of I_i by $V_i \subset \Omega$ for $i = 1, 2$. We assume that the V_i is compact. For the sake of simplicity, we will assume that the boundary of V_i is piecewise C^1 closed curve when $n = 2$. In the $n = 3$ case, the boundary of V_i is assumed to be a piecewise-smooth surface that is a union of finitely many surfaces, curves and points (e.g., the boundary of a cube).

The collapse problem occurs when V_1 has an appendage which is not included in V_2, and the growth problem occurs when V_2 has an appendage which is not included in V_1. The collapse and growth behaviors look different if they are observed only from the point of view of the foreground. However, the growth of the foreground can be studied as a certain type of collapsing of the background.

Definition 1. *Let V be a subset of a metric space (Ω, d), and $B_r(p)$ denotes the open metric balls $\{x \in \Omega : d(x, p) < r\}$. A closed ball $\overline{B} \subset V$ is called a maximal ball of V, if for every closed ball $\overline{B}'$, $\overline{B} \subseteq \overline{B}' \subseteq V$, one has $\overline{B} = \overline{B}'$. The set $\left\{q \in V : \exists r > 0, \overline{B}_r(q) \text{ is maximal ball of } V\right\}$ is defined to be the skeleton $S(V)$ of V by maximal balls [4].*

Suppose Ω is a metric space. Let p be a point of Ω, and let V be any subset of Ω. The interior of V, denoted $\mathrm{Int}(V)$, is the union of all open subsets of Ω contained in V. The exterior of V, denoted $\mathrm{Ext}(V)$, is the union of all open subsets of Ω contained in $\Omega - V$. The boundary of V, denoted by ∂V, is the set of all points of Ω that are in neither $\mathrm{Int}(V)$ nor $\mathrm{Ext}(V)$.

We first discuss the collapsing behavior in the simple case of $V_2 = \varnothing$ and no regularization. Since we have a binary image, a greedy algorithm for the similarity-cost gradient follows the direction that decreases the area in the fastest way. We study this area-decreasing flow on the interior of V_1, and the distance function $f : V_1 \to [0, \infty)$ to the boundary ∂V_1, given by $f(x) = dist(x, \partial V_1) = \inf\{|x - y| : y \in \partial V_1\}$. For a piecewise C^1 boundary, f is C^1 except on a set (containing the skeleton and the boundary) of measure zero, and its gradient ∇f (when it exists) is perpendicular to the level sets of f.

The fastest area decreasing occurs by deforming the interior of V_1 along ∇f until the skeleton is reached at that direction, since following the gradient one goes into the interior deeper in the fastest way. The superlevel sets $sl_r = \{x \in V_1 : f(x) \geq r\}$ are the stages of this deformation in the continuous category, as justified below.

For every $q \in S(V_1)$, there is a unique maximal ball $\overline{B}_r(q)$ of V_1 centered at q, and the set of points along the boundary ∂V_1 associated to q is $A_\partial(q) := \overline{B}_r(q) \cap \partial V_1$. The associated set $A_\partial(q)$ is nonempty. It usually has two or more points, but it can be one point, such as the focal point of a boundary curve (or a surface for $n = 3$) at a strict local maximum of curvature (or principal curvature for $n = 3$). For every $p \in \partial V_1$, we can also define the associated set $A_S(p) = \{q \in S(V_1) : p \in A_\partial(q)\}$, which is the set of points along the skeleton associated with p. If ∂V_1 is C^2 about p, then $A_S(p)$ contains one point. Let $L(p, q)$ denote the line segment with end points p and q, and $L^\circ(p, q) = L(p, q) - \{p, q\}$. For

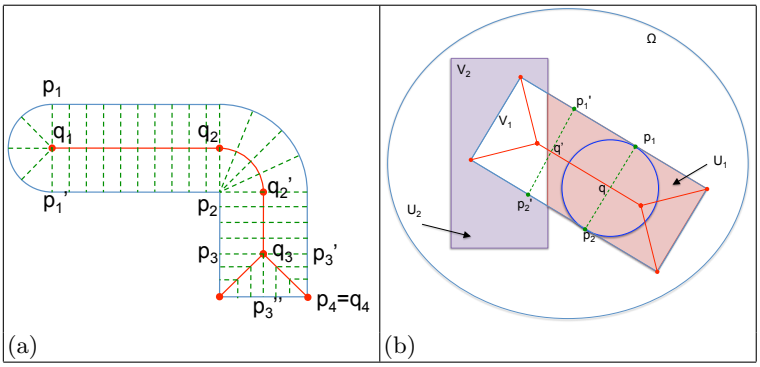


Fig. 2. *(a)* Possible configurations of the skeletal and boundary points. A skeletal point can correspond to one (q_4), two (q_2), three (q_3) or more (q_1) boundary points. Similarly, a boundary point can have one (p_4) or many (p_2) associated skeletal points. *(b)* Collapsing and non-collapsing points. p_1 is a collapsing point because it belongs to the same maximal ball $B_r(q)$ as p_2, and both p_1 and p_2 are in the non-overlap region U. In contrast, p_1' is not a collapsing point because it only shares a maximal ball ($B_r(q')$) with p_2', but p_2' is not in U.

simplicity, we will use 2-D examples to illustrate these concepts throughout this paper, even though the extension to n-D using line segments is straightforward. Figure 2 illustrates the possible configurations of boundary and skeletal points and their associations, as well as the notation used.

Proposition 1. *The set of line segments* $\{L^\circ(p,q) : q \in S(V_1)$ *and* $p \in A_\partial(q)\}$ *is a partition of* $V_1 - (\partial V_1 \cup S(V_1))$.

Proof. We want to show that every point $a \in V_1$ belongs to a $L(p,q)$ where $p \in \partial V_1$, $q \in S(V_1)$ and they are associated with each other. The point a has a closest point $p \in \partial V_1$ at a distance $f(a) = c$ and hence $\overline{B}_c(a) \subseteq V_1$. Then $\overline{B}_c(a) \subseteq \overline{B}_r(q) \subseteq V_1$ for a maximal ball $\overline{B}_r(q)$. If $p \in \overline{B}_c(a) \subseteq \overline{B}_r(q)$ is an interior point of $\overline{B}_r(q)$ then it cannot be on ∂V_1. Hence $\partial\overline{B}_r(q)$ and $\partial\overline{B}_c(a)$ are tangential at p, and thus $a \in L(p,q)$. Therefore V_1 is the union of all $L(p,q)$ with $q \in S(V_1)$ and $p \in A_\partial(q)$.

For $p_i \in \partial V_1$ and $q_i \in S(V_1)$ with $p_i \in A_\partial(q_i)$, for $i = 1, 2$, if $L(p_1,q_1) \cap L(p_2,q_2) \neq \varnothing$ then either $p_1 = p_2$ or $q_1 = q_2$, i.e. these segments can only intersect along the boundary or the skeleton unless they are identical. Because, an intersection at the interior of the line segments would contradict the triangle inequality and p_i being the closest point of ∂V_i to q_i. □

By Proposition 1 and its notation, if $a \in L^\circ(p,q)$, then p is the unique closest point of ∂V_1 to a. This justifies that $f^{-1}(r)$ is the set of all points a_{pq} on $L(p,q)$ with $\|a_{pq} - p\| = r$ as long as $\|q - p\| \geq r$, as q varies through all of $S(V_1)$ and p through all of $A_\partial(q)$. The fastest area decreasing flow on the interior of V_1 is along the line segments $L^\circ(p,q)$ with unit speed until the skeleton is reached

along each segment. The flow is not definable along the skeleton. However, away from the skeleton, sl_r are the stages of this deformation, and $L(p,q)$ is parallel to $\nabla f(a_{pq})$ if it exists.

We remark that this flow needs to be considered on the interior of V_1, since it is possible that $A_S(p)$ may contain more than one point if ∂V_1 is not differentiable at $p \in \partial V_1$ (e.g. p_2 in Fig. 2-a). In that case, the flow is not defined at p, but it is well defined on the union of $\{L^\circ(p,q) : \forall q \in A_S(p)\}$. Rather than letting ∂V_1 follow the flow, we use the boundary of sl_r, namely $f^{-1}(r)$. Using sl_r has the additional advantage of describing the area reduction process accurately around the skeleton. This is because $S(sl_r) = S(V_1) \cap sl_r$ and the flow along the $L(p,q)$ reaches different parts of the skeleton at different times and stops. Therefore it is possible that sl_r may break into components. In this case, the flow is not continuous at the skeleton; however, the area-reduction is continuous.

In the continuous area-reduction case and with $V_2 = \varnothing$, the above process (without regularization) will reduce V_1 to $\varnothing$ since it will remove the skeleton along the way. By using small (but not zero) regularization and discrete image registration, we see that the skeleton actually remains. The deforming forces towards a point q on the skeleton are from different directions along $L(p,q)$ for several $p \in A_\partial(q)$. The regularization will reduce the effect of these forces by averaging and slow down the area reduction. Since the skeleton has measure zero and the deformations we use for the image registration are discrete, removing the skeleton has no or little gain in the cost function, and therefore a skeleton remains. Since the deformations are not going exactly along sl_r, the skeleton reached through image registration cannot be expected to be same as $S(V_1)$, but it is a very good approximation since the above mentioned factors are not in effect until the flow comes near $S(V_1)$.

The next step is to consider the case in which the target image contains a foreground object, i.e., $V_2 \neq \varnothing$. We begin with some definitions.

Definition 2. *Let V_1 and V_2 be subsets of Ω representing the foreground objects of images I_1 and I_2, respectively. The overlap of the foreground objects is denoted as $W = V_1 \cap V_2$ and the symmetric difference region (non-overlap) $U = U_1 \cup U_2$ where $U_1 = V_1 - V_2 = \{x \in V_1 : x \notin V_2\}$ and $U_2 = V_2 - V_1$.*

Definition 3. *A point $p_1 \in \partial V_1$ is a collapsing point for $q \in S(V_1)$, if $\exists p_2 \in \partial V_1$ such that $p_1 \neq p_2$, $\{p_1, p_2\} \subset A_\partial(q)$, and $L(p_1,q) \cup L(p_2,q) \subset U_1$. The notation* CLPS($q$) *denotes all collapsing points (if any) along ∂V_1 associated with $q \in S(V_1)$ and* CLPS(V_1) *denotes the union of* CLPS(q) *for all $q \in S(V_1)$.*

Figure 2-b illustrates these definitions. In this example, p_1 is a collapsing point because there exists a p_2 satisfies all three conditions: $p_1 \neq p_2$, $\{p_1, p_2\} \subset A_\partial(q)$, and $L(p_1,q) \cup L(p_2,q) \subset U_1$. In contrast, p_1' is not a collapsing point because the only point in ∂S_1 that shares a maximal ball with it is p_2', and $p_2' \notin U_1$. During the registration of the image of S_1 into the image of S_2, p_1 p_2 will move towards q on the skeleton of S_1 whereas p_1' will move towards S_2 and p_2' will stay in S_2.

Regions of foreground reduction of the moving image occur on U_1 and regions of foreground growth occurs on U_2. We can study U_1 and U_2 separately, since the deformation algorithms are local. We start by discussing the area-decreasing flow away from skeleton $S(V_1)$ in the continuous case without regularization again. If $p_1 \in \partial V_1 \subset V_1$ but $p_1 \notin U_1$, that is $p_1 \in V_2$, then it is already in the target set, there will be no area-decreasing flow near p_1. If $p_1 \in U_1 \cap \partial V_1$, then the area-decreasing flow on the union of $\{L^\circ(p,q) : \forall q \in A_S(p)\}$ starts as described before for p near p_1, and continues along $L^\circ(p,q)$ until it comes close to either (i) a $q \in A_S(p)$ when $L^\circ(p,q)$ is away from the overlap W, or (ii) the flow comes close to a point of ∂V_2 provided that $L^\circ(p,q)$ enters V_2 at most once.

For the cases when $p_1 \in U_1 \cap \partial V_1$, $q \in A_S(p_1)$ and $L(p_1,q) \subset U_1$, but p_1 is not a collapsing point due to the rest of $A_\partial(q)$ being in W, there will be one-sided flows towards q at the start. However, since these one-sided flows will alter the skeleton of the remaining set, the flow needs to be readjusted according to the new skeleton, a case that will not happen when $V_2 = \varnothing$. All flows near W will have this adjustment.

At the collapsing points (nonempty) $\mathrm{CLPS}(q) \subset A_\partial(q)$, we expect that the area decreasing flow to reduce the area tending towards q from the several directions $\{L(p,q) : \forall p \in \mathrm{CLPS}(q)\}$, and some type of collapse occur at q.

If there exists an open set Z bounded away from W such that for all $q \in Z \cap S(V_1)$ one has $\mathrm{CLPS}(q) = A_\partial(q)$ and for all $p \in \mathrm{CLPS}(q)$, $L(p,q) \subset Z$ then the area decreasing procedure in Z will behave the same as $V_2 = \varnothing$. Understanding how this works allows us to predict the outcome with a small regularization factor. A part of a skeleton will be reached through image registration for the collapsing of U_1 within Z, it cannot be expected to be same as a part of $S(V_1)$, but it is a good approximation since the above mentioned factors when $V_2 = \varnothing$ are not in effect until the flow comes near this particular portion of $S(V_1)$.

The dual problem of growth in the foreground of the moving image is a reduction in the background. Hence, we can predict the behavior of the growth as well by using the same proceedure. Revisiting Fig. 1, we can now compare the collapse and growth of the index finger. The collapse of the index finger U_1 is mostly sideways flowing towards the vertical part of the skeleton $S(V_1)$ through U_1, by becoming thinner and thinner. But the growth of the finger in U_2 is a reduction of a part of the complement of V_1 (not V_2!). It has to follow the level sets of the distance function to ∂V_1 on the complement of V_1. These levels sets restricted to U_2 are parallel to the index finger in the moving image and therefore U_2 grows in the perpendicular direction to these level sets. The growth in U_2 keeps the maximal width since there is no skeleton of the background of the moving image about that area.

3 Discussion

Figure 3 presents a simple 2D registration example that illustrates the shape collapse problem for parts of the foreground and background of the moving image. The moving image (Fig. 3-a) is a 128x128 pixel, binary image that has

value 0 for the background and 1 for the foreground. The target image (Fig. 3-b) differs from the moving image by a translation of the lower rectangle appendage. The difference image (Fig. 3-c) demonstrates the amount of overlap between the appendage of the moving and target image.

Figure 3 shows the skeletal points for the foreground (d) and background (e) of the moving image. The skeletal points of the moving shape are used to predict the location where a potential collapse of the moving image will occur. Note that the designation of foreground and background objects are for convenience of presentation and could have just been referred to as the white and black object of the moving image. Regions of shape collapse only occur in the non-overlap region between the moving and target images. Fig. 3-g and 3-h shows the skeletal points masked by the region of overlap since these are the only skeletal points that we are concerned with.

As discussed in Section 2, we will detect the object boundary points where the region of collapse may happen when the images are registered. The set of boundary points where the shape will collapse can only occur in the region of non-overlap between the moving and target image. The foreground and background boundary points that occur in the symmetric difference region correspond to a super set of the boundary collapse points and are shown in Fig.3-f and 3-i. The foreground and background boundary collapse points are shown in Fig. 3-j and 3-k. The collapse points are defined as the boundary points in the symmetric difference region that correspond to the skeletal points in the symmetric difference region. The predicted foreground and background collapse points are shown superimposed on the moving image in Fig. 3-l. The skeletal points associated with boundary collapse points are a rough estimate of where the moving image will collapse when registered to the target image, as shown in Figure 4.

Fig. 4 shows that the foreground and the background of the moving image collapsed following the predictions of the boundary collapse points. The foreground collapses to a thin strip of white foreground (square callout) and the background collapses to a thin strip of black background (circle callout). Note that the regions of collapse shown in Fig. 4 are small and hard to see. It is because regions of collapse are hard to see that they are often ignored.

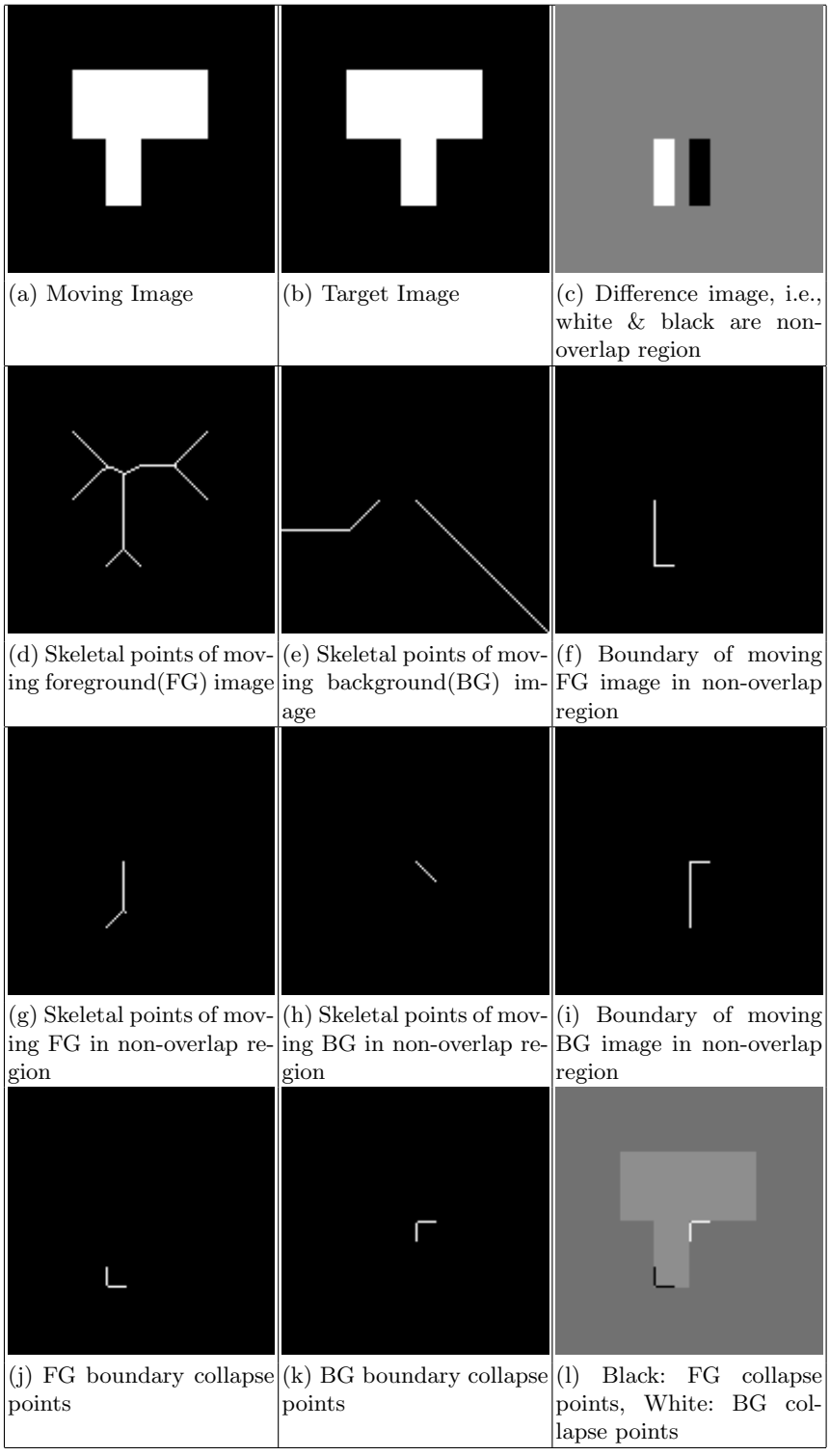
Figure 5 shows the result of registering the moving image to the target image for the example in Fig. 3. The image registration algorithm[1] minimized the objective function $C(u)$ which consists of the weighted sum of a similarity cost (sum of squared intensity differences (SSD)) and a regularization term (membrane model), as described by the following equations:

$$C(u) = \frac{\alpha}{2}\int_{\Omega}\left(I_1(x+u(x)) - I_2(x)\right)^2 dx + \frac{\beta}{2}\int_{\Omega}||Lu||^2 dx \quad \text{Registration Cost}$$

$$\delta C(u(x)) = \alpha\left(I_1(x+u(x)) - I_2(x)\right)\cdot \nabla I_1|_{x+u(x)} + \beta L^2 u \quad \text{Gradient of Cost}$$

$$u^{(i+1)}(x) = u^{(i)}(x) - \Delta\delta C(u(x)) \quad \text{Gradient Descent}$$

The gradient descent used a step size of $\Delta = 0.95$, a weight of $\alpha = 1$ for the SSD term and a weight of $\beta = 0.1$ for the membrane regularization term.



(a) Moving Image | (b) Target Image | (c) Difference image, i.e., white & black are non-overlap region

(d) Skeletal points of moving foreground(FG) image | (e) Skeletal points of moving background(BG) image | (f) Boundary of moving FG image in non-overlap region

(g) Skeletal points of moving FG in non-overlap region | (h) Skeletal points of moving BG in non-overlap region | (i) Boundary of moving BG image in non-overlap region

(j) FG boundary collapse points | (k) BG boundary collapse points | (l) Black: FG collapse points, White: BG collapse points

Fig. 3. Detection of boundary collapse points for overlapping appendage experiment. Both the foreground and background collapses in this example.

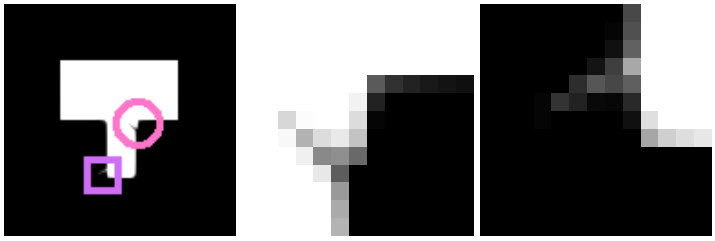
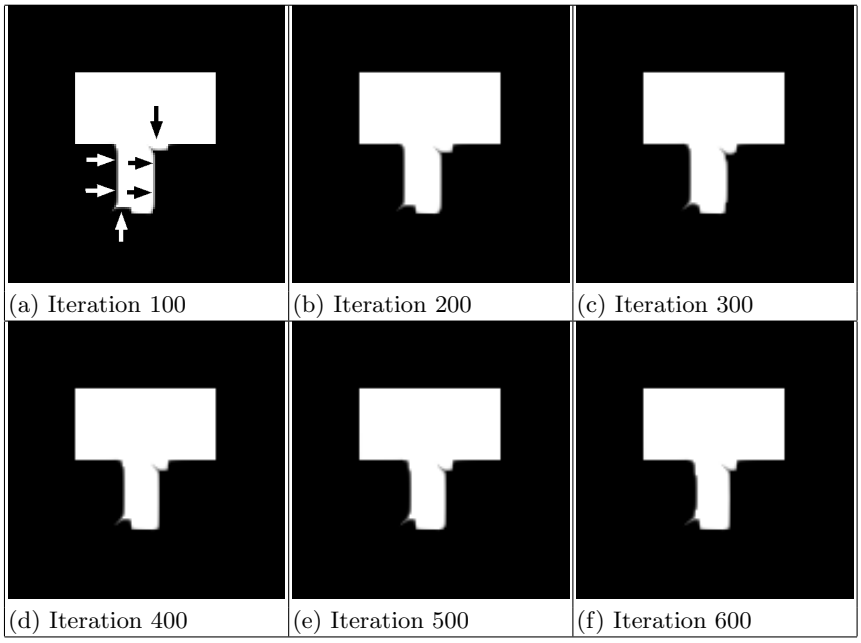


Fig. 4. The result of transforming the moving image into the shape of the target image shown in the top row of Figure 3. The middle panel magnifies the collapse region of the background (circle inset). The right panel magnifies the collapse region of foreground (square inset).

The operator L in the regularization term is a linear differential, self-adjoint operator and corresponds to $L^2u = \nabla^2 u$ for the membrane regularization model and ∇^2 is the Laplacian. For this experiment, the displacement field $u(x)$ was represented as a 128x128 vector field. Bilinear interpolation was used to compute the deformation of the moving image. 1000 iterations of gradient descent were used to minimize the registration cost.

Figure 5 illustrates the time progression of how using a greedy optimization method causes the regions of the moving image to collapse. The arrows in Figure 5-a show the direction of the force generated by the gradient of the SSD cost function. The white arrows show that there is a force pushing right and a force pushing up. The combination of these two forces causes the collapse of the foreground object at the lower left corner of the object. The black arrows show a force pushing right and another pushing down. The result of these two forces causes the background of moving image to collapse at the inside corner of the object. Note that once a region of the image collapses, there is no way for a greedy optimization method to "un-collapse" the region. The reason for this is that the collapsed region has zero measure in the SSD cost function and therefore has little to no cost. Thus, there is no incentive (i.e., no gradient force) for the greedy optimization method to retract or otherwise fix the collapsed region.

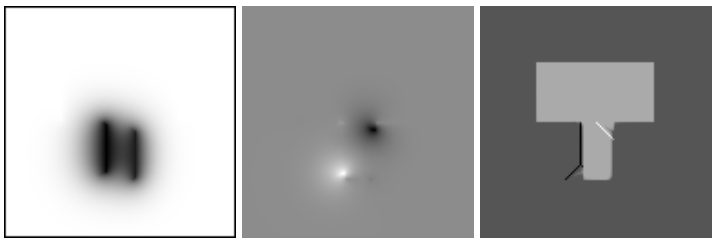
Figure 6 shows the x- and y-displacement fields associated with this registration experiment. Notice that the x-displacement field is not symmetric in the vertical direction, i.e., the left black region appears to be shifted up compared to the black region on the right. The vertical asymmetry of the x-displacement field results from the shape collapse of the moving image. The asymmetry of the x-displacement field corresponds to the nonzero regions of the y-displacement field in the regions of image collapse. The nonzero regions of the y-displacement field indicate a problem with the estimated transformation for this experiment. It is reasonable to assume that the vertical appendage of the moving image only needs to be shifted to the right to match the corresponding structure in the target image. Thus, any nonzero displacement in the y-direction indicates a counter-intuitive registration result.



(a) Iteration 100 (b) Iteration 200 (c) Iteration 300

(d) Iteration 400 (e) Iteration 500 (f) Iteration 600

Fig. 5. Progression of deformation from moving to target image. The arrows show in panel (a) the direction and location of the deformation. The white arrows show where the foreground of the moving image collapses and the black arrows show where the background of the moving image collapses.



(a) (b) (c)

Fig. 6. Displacement fields for the image registration of overlapping appendage experiment. (a) X-displacement, (b) Y-displacement, (c) foreground and background skeletal points overlaid on deformed image predict regions of collapse

The displacement fields shown in Figure 6 demonstrate that the regions of collapse should not be ignored because they lead to poor correspondence in the regions of collapse. Ignoring the poor correspondence in collapsed regions may have a major impact on the conclusions drawn in the applications using such transformations. These include mapping brain function, computing average

shapes, computing shape statistics, computing mechanical properties of the lung, adaptive radiotherapy planning, and computing cumulative dose in radiation therapy to mention a few application areas.

Figure 6-c shows the skeletal points of the foreground and background objects from the original moving image within the symmetric difference region, superimposed on the registration result. Notice that the skeletal points do a good job of predicting the locations of the moving image collapse. The skeletal points do not predict the actual locations of the moving image collapse due to the shape change of the moving image during registration and the effect of regularization.

The minimum value of the Jacobian for the estimated transformation for this experiment was 0.0853. The fact that the minimum value of the Jacobian was positive indicates that the estimated transformation remained a diffeomorphism throughout the optimization procedure. This fact illustrates that the collapse problem occurs even when the transformation is a diffeomorphism.

Figure 7 shows three additional examples of predicting collapse points. The first row of this figure illustrates collapse points occurring when a subregion of an object retracts. The second row illustrates collapse points occurring when a subregion of an object expands. The final row shows the cortical ribbon extracted from a clinical MR image with $1 \times 1\text{mm}^2$ resolution. The cortical segmentation was adapted to limit the example to illustrate four retraction and two expansion regions around sulci. This example is important to show the scale of the expected effect size of the collapsing problem for neuroimaging applications.

4 Conclusions

This paper presented the shape collapse problem for image registration which is a problem for both small and large deformation image registration algorithms. A mathematical justification was presented for why the collapsing problem occurs for binary images that each contain a single corresponding object. Collapsing boundary points were defined mathematically and were shown to predict before the registration, which boundary points in the moving image will collapse during an iterative image registration procedure that minimize the registration cost at each iteration. Furthermore, the foreground and background skeletons of the moving image in the symmetric difference region were shown to closely predict the collapse loci in the moving image.

References

1. Christensen, G.E.: Deformable Shape Models for Anatomy. Ph.D. thesis, Washington University, St. Louis, MO. (Aug 1994)
2. Do Carmo, M.: Riemannian Geometry. Birkhauser, Boston (1992)
3. Siddiqi, K., Pizer, S.: Medial Representations: Mathematics, Algorithms and Applications. Springer Publishers (2008)
4. Sonka, M., Hlavac, V., Boyle, R.: Image Processing, Analysis, and Machine Vision. Thomson Engineering, Toronto, Canada, 3rd edn. (2008)

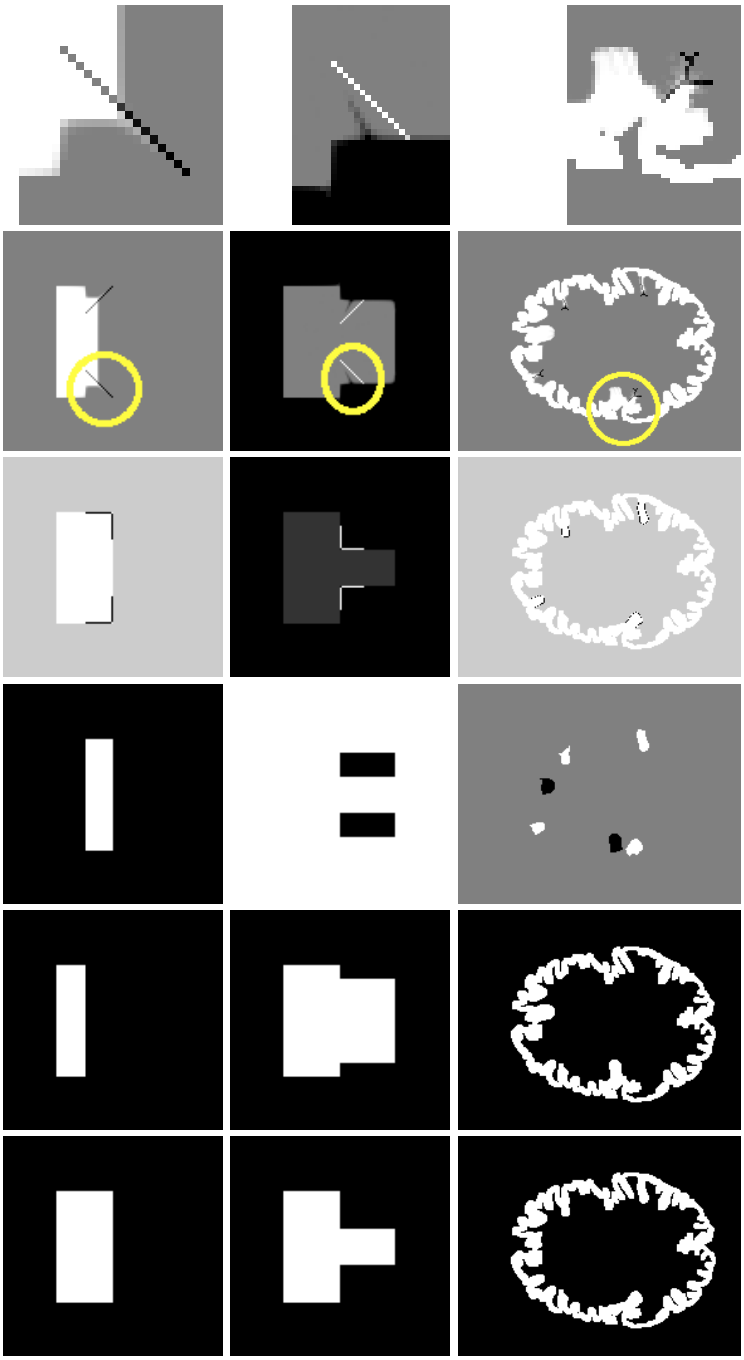


Fig. 7. Examples. Each row of this figure shows an additional example of collapsing point detection. The columns from left to right show the moving image, target image, difference image before registration, deformed collapse points on the moving image, and the foreground and background skeletal points overlaid on the deformed image.

Sparse Gaussian graphical model of spatial distribution of anatomical landmarks – whole torso model building with training datasets of partial imaging ranges

Shouhei Hanaoka[1], Yoshitaka Masutani[1,2], Mitsutaka Nemoto[1], Yukihiro Nomura[1], Soichiro Miki[3], Takeharu Yoshikawa[3], Naoto Hayashi[3], Kuni Ohtomo[1,2]

[1] Department of Radiology, [2] Division of Radiology and Biomedical Engineering, Graduate School of Medicine and [3] Department of Computational Diagnostic Radiology and Preventive Medicine, The University of Tokyo Hospital, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan

hanaoka-tky@umin.ac.jp

Abstract. A method of building a point distribution model for multiple anatomical landmarks from training datasets with various different imaging ranges is presented. To cope with the missing data problem caused by partial imaging ranges of training datasets, the MissGLasso algorithm is applied to build the model. The joint probability distribution of logarithmic distances between all landmark pairs is approximated as a sparse Gaussian graphical model which has a sparse precision matrix calculated by the graphical lasso method. Additionally, EM algorithm is also utilized to apply graphical lasso to training sample vectors which include many missing values. To evaluate models built by the proposed method, the models were embedded into multiple-landmark detection systems and the detection sensitivities were compared between models trained with partial and entire imaging range datasets. The overall sensitivities of landmark detection using these two models were 72.78 and 75.02%, respectively. Because the detection sensitivity for most of the landmarks differs little between the two models, it can be concluded that the MissGLasso algorithm could effectively handle missing values in training landmark distribution models.

Keywords: Sparse Gaussian graphical model, Landmark, Graphical lasso

1 Introduction

The automatic detection of anatomical landmark positions often plays a key role in various medical image analyses, such as organ segmentation, interindividual or intermodality image registration or computer-aided lesion detection. Various point distribution models (PDMs) of the spatial distribution of multiple anatomical landmarks have been used to detect a series of landmarks [1][2][3]. For example, in [1] Seifert et al. used predefined spatial constraints between specific pairs of landmarks in detecting 19 body trunk landmarks. In [2], Potesil et al. proposed the pictorial structure

model in which a graph structure connects 22 landmarks from cervical to pelvic structures. Although the graph structure itself is arbitrarily predefined, the spatial constraint between two connected landmarks is learned from training datasets. On the other hand, in [3] Hanaoka et al. built a statistical PDM on 173 landmark positions and applied it to a landmark detection task. Their target imaging range was neck-to-pelvis and the head was not included. In all these methods, PDMs play an important role by providing prior knowledge on the human body structure. Therefore, it is suggested that the quality of PDM will strongly affect the overall performance of landmark detection applications.

Theoretically, a landmark-based PDM for the whole human body can be built in the same manner as the methods mentioned above. Such a whole body PDM will be more versatile than partial-body PDMs, that is, it can be applied to a wider range of medical image processing and to various organs. However, to the best of our knowledge, no research on building a PDM for the whole body by means of statistical modeling has been reported. It is considered that the main reason for this is the difficulty of preparing a sufficient amount of whole-body training datasets. Medical image examinations, such as computed tomography (CT) or magnetic resonance imaging (MRI), are rarely performed for the whole body because of limited imaging machine resources or excess radiation exposure. If only partial-body datasets with various imaging ranges (e.g. chest only, abdomen only, etc.) are available as training datasets, the usual statistical model estimation methods will not be applicable. Instead, an alternative method that can handle missing values in the training datasets will be required. One of the aims of this study is to overcome this problem by applying an EM-algorithm-based missing value imputation method [4] in estimating the statistical landmark distribution model.

The other aim of this study is to apply the sparse Gaussian graphical model [5] to landmark distribution modeling. In particular, we focused on estimating the precision matrix (inverse covariance matrix) on the interlandmark distances between landmark pairs as a sparse matrix. We chose the interlandmark distances as the model variables, instead of the landmark coordinates themselves, because of their desirable features such as spatial rotation invariance and robustness to local deformations. Generally speaking, it can be expected that a certain distribution can be statistically modeled better with a sparse precision matrix if most of the variable pairs are conditionally independent. This is not the case for the interlandmark distances, however, because most of the distances have strong positive correlations with each other owing to the scale factor. A large person has larger interlandmark distances, and vice versa. Nevertheless, we found that the distances become much less correlated after the scale factor is normalized appropriately. Thus, we assumed that the precision matrix can be estimated as a sparse matrix if the effect of the scale factor is removed in advance. On the basis of this assumption, we developed a novel method to estimate the sparse precision matrix while handling the scale factor separately.

For these two purposes, we propose a method for building a sparse Gaussian graphical model of interlandmark distances from training datasets with insufficient imaging ranges. The method is based on the MissGLasso method, proposed by Städler et al. [4], which consists of two algorithms: (1) an expectation-maximization

(EM) algorithm to handle missing data and (2) the graphical lasso method [5] for estimating the precision matrix with l_1 norm regularization. We modified the original MissGLasso method so that both the scale factor and the scale-normalized interlandmark logarithmic distance set can be modeled simultaneously. The proposed method was evaluated with 78 chest-to-pelvis CT volume datasets. For comparison, two models were built: (1) a statistical landmark distribution model built with intentionally cropped training volumes and (2) a model built with the whole imaging ranges. Then, the landmark detection performances using these models were compared to evaluate the applicability of the proposed method to partial volume training datasets.

2 Methods

2.1 Landmark distribution model based on interlandmark distances

In this study, the spatial landmark distribution is modeled as a joint probability function whose variables are the interlandmark logarithmic distances between all landmark pairs. Here, the reason why logarithmic distances are used is that it enables the proposed algorithm to handle the scale factor as a linear factor. Additionally, since the domain of log-distances is not restricted to positive but can have negative values, they may be more suitable to be modeled with Gaussian distribution.

Suppose that the number of landmarks is L and that their coordinates are $\mathbf{x}_i,\ i \in \{1,2,3,\dots,L\}$. Then, the logarithmic distances are defined as $d_{i,j} = \log|\mathbf{x}_i - \mathbf{x}_j|, 1 \leq i < j \leq L$. Let the concatenated vector of all $d_{i,j}$ be $\mathbf{d} = \left(d_{1,1}d_{1,2}\cdots d_{i,j}\cdots d_{L-1,L}\right)^t$. Suppose that the mean vector of $\mathbf{d}$ is estimated as $\bar{\mathbf{d}}$ and the precision matrix is estimated as $\mathbf{K}$, then the probability function $p(\mathbf{X})$ for any landmark position set $\mathbf{X} = \{\mathbf{x}_i\}$ can be estimated as a multidimensional Gaussian distribution as follows:

$$p(\mathbf{X}) = \sqrt{2\pi}^{-M} \cdot \sqrt{|\mathbf{K}|} \cdot \exp\left\{-\frac{1}{2}\left(\mathbf{d} - \bar{\mathbf{d}}\right)^t \mathbf{K}\left(\mathbf{d} - \bar{\mathbf{d}}\right)\right\}. \quad (1)$$

where $M = {}_L\mathrm{C}_2$ is the dimension number of $\mathbf{d}$. Therefore, only the mean vector $\bar{\mathbf{d}}$ and precision matrix $\mathbf{K}$ must be determined to define a landmark distribution model. $\bar{\mathbf{d}}$ and $\mathbf{K}$ can be estimated from manually input ground-truth landmark positions in the training datasets. The main topic of this study is how to adequately estimate $\bar{\mathbf{d}}$ and $\mathbf{K}$ where the number of training datasets N is far less than the model dimension ${}_L\mathrm{C}_2$ and many missing values are included.

Note that the distribution (1) is translation- and rotation-invariant because only the interlandmark distances are considered. Moreover, the logarithmic distance vector $\mathbf{d}$ is altered linearly by a scale transformation; that is, when a scale transformation with scale factor α is applied to the landmark position set (i.e., $\mathbf{x}_i \mapsto \alpha \cdot \mathbf{x}_i, \forall \mathrm{i}$), the vector $\mathbf{d}$ will be translated as $\mathbf{d} \mapsto \mathbf{d} + \log\alpha \cdot \mathbf{1}$. This is a simple translation parallel to the vector $\mathbf{1} = (1\ 1\ \dots\ 1)^{\mathrm{t}}$.

2.2 Graphical lasso with scale factor normalization

Graphical lasso [5] is a method of estimating a sparse precision matrix $\mathbf{\Theta}$ from a given sample covariance matrix $\mathbf{S}$ by maximizing the l_1 norm-penalized log likelihood as follows:

$$\log\det\mathbf{\Theta} - \mathrm{trace}(\mathbf{S\Theta}) - \rho\|\mathbf{\Theta}\|_1. \tag{2}$$

Here, ρ is a parameter that controls the sparsity of $\mathbf{\Theta}$.

The sample covariance matrix $\mathbf{S}$ can be calculated from the set of the vector $\mathbf{d}$ in the given training datasets. However, as described above, $\mathbf{S}$ is not suitable for the graphical lasso estimation because of the scale factor. To overcome this problem, $\mathbf{S}$ is divided into two: the covariance matrix of the scale-normalized log-distances $\mathbf{S}_{\mathrm{nor}}$ and the variance of the scale factor γ. After this division, $\mathbf{S}_{\mathrm{nor}}$ is processed by the graphical lasso method to estimate the precision matrix.

The division of $\mathbf{S}$ is performed as follows. Firstly, the scale factor of a given vector $\mathbf{d}$ is defined as $\mathrm{scale}(\mathbf{d}) = \left(\mathbf{d} - \bar{\mathbf{d}}\right) \cdot \left(\frac{1}{M} \cdot \mathbf{1}\right)$. Here, $\bar{\mathbf{d}}$ is the sample mean of $\mathbf{d}$ and $\mathrm{scale}(\mathbf{d})$ is the average of M elements of the mean-subtracted vector $\left(\mathbf{d} - \bar{\mathbf{d}}\right)$. The variance of the scale factors is calculated as the variance of all $\mathrm{scale}(\mathbf{d})$ from the datasets. Secondly, the scale-normalized log-distance is defined as $\mathbf{d}_{\mathrm{nor}} = \mathbf{d} - \bar{\mathbf{d}} - \mathrm{scale}(\mathbf{d}) \cdot \left(\frac{1}{M} \cdot \mathbf{1}\right)$. Then the covariance matrix $\mathbf{S}_{\mathrm{nor}}$ is calculated from all $\mathbf{d}_{\mathrm{nor}}$ extracted from training datasets.

The precision matrix of $\mathbf{d}_{\mathrm{nor}}$, namely $\mathbf{\Theta}_{\mathrm{nor}}$, is estimated using the graphical lasso algorithm by maximizing (2) with the given $\mathbf{S}_{\mathrm{nor}}$. Note that $\mathbf{K}$ (the precision matrix of $\mathbf{d}$) can be estimated from $\mathbf{\Theta}_{\mathrm{nor}}$ as follows: Assuming that the scale factor $\mathrm{scale}(\mathbf{d})$ and the normalized distance vector $\mathbf{d}_{\mathrm{nor}}$ are statistically independent, a simple addition rule of covariance matrices $\mathbf{K}^{-1} = \mathbf{\Theta}_{\mathrm{nor}}^{-1} + \gamma \cdot \mathbf{1}\mathbf{1}^t$ is satisfied. ($\gamma \cdot \mathbf{1}\mathbf{1}^t$ corresponds to the covariance matrix of the scale factor.) Then, from the Sherman-Morrison formula,

$$\mathbf{K} = (\mathbf{\Theta}_{\mathrm{nor}}^{-1} + \gamma \cdot \mathbf{1}\mathbf{1}^t)^{-1} = \mathbf{\Theta}_{\mathrm{nor}} - \gamma/(1 + \gamma \cdot \mathbf{1}^t\mathbf{\Theta}_{\mathrm{nor}}\mathbf{1}) \cdot \mathbf{\Theta}_{\mathrm{nor}}\mathbf{1}\mathbf{1}^t\mathbf{\Theta}_{\mathrm{nor}} \tag{3}$$

is satisfied. Note that using this equation, the probability calculation in Eq. (1) can be speeded up by taking advantage of high sparsity of the matrix $\mathbf{\Theta}_{\mathrm{nor}}$.

A modification was introduced to the original graphical lasso method in this study. It is to change the l_1-norm-penalizing weights for each variable according to its standard deviation. In graphical lasso it is possible to change the sparsity control factor ρ for each element of $\mathbf{\Theta}$ [5]. This is performed by modifying Eq. (2) as follows:

$$\log\det\mathbf{\Theta} - \mathrm{trace}(\mathbf{S\Theta}) - \|\mathbf{\Theta} * \mathbf{P}\|_1 \tag{4}$$

where $\mathbf{P} = \{\rho_{jk}\}$ with $\rho_{jk} = \rho_{kj}$, and * indicates component-wise multiplication. On this basis, we determined the element-wise l_1-norm-penalizing weight as follows:

$$\rho_{jk} = \rho \cdot f\left(\sigma_j\right) \cdot f(\sigma_k) \tag{5}$$

$$f(x) = \begin{cases} x & \text{if } \sigma_{threshold} \leq x \\ \sigma_{threshold} & \text{if } x < \sigma_{threshold} \end{cases}$$

where σ_i is the standard deviation of the i-th element of $\mathbf{d}$, and $\sigma_{threshold}$ is a parameter to be determined in advance. The motivation for introducing this weight is to avoid the excessive effect of variables with very large variances. We found that this weight modification is very useful for improving the result.

2.3 EM algorithm and MissGLasso

Most of the methods of estimating the distribution from given sample data require that no values are missing from the given sample. However, this is not the case when the imaging range is limited and some landmarks are out of range. One element in $\mathbf{d}$ (i.e., one interlandmark distance) will be "missed" if either of the two landmarks is out of the imaging range. Therefore, a significant number of elements in $\mathbf{d}$ from a partial volume training dataset may be missed. One frequently used solution for such a "missing value" problem is to apply the EM algorithm. In this study we utilized the MissGLasso method, which is a fusion of the EM algorithm and the graphical lasso method.

Details of the MissGLasso method are available in [4]. In brief, the model parameter $\mathbf{\Theta}_{\text{nor}}$ is iteratively updated in MissGLasso by alternately applying E and M steps. (Fig. 1, left) In the E step, the covariance matrix and mean vector are estimated from both the observed (nonmissed) values and recently estimated $\mathbf{\Theta}_{\text{nor}}$. Using the updated covariance matrix and mean vector, the precision matrix $\mathbf{\Theta}_{\text{nor}}$ is estimated and updated in the following M step. In MissGLasso, the M step is virtually the graphical lasso method itself. In this study, we iterated the EM algorithm 10 times to find an optimal $\mathbf{\Theta}_{\text{nor}}$ from partial volume training datasets.

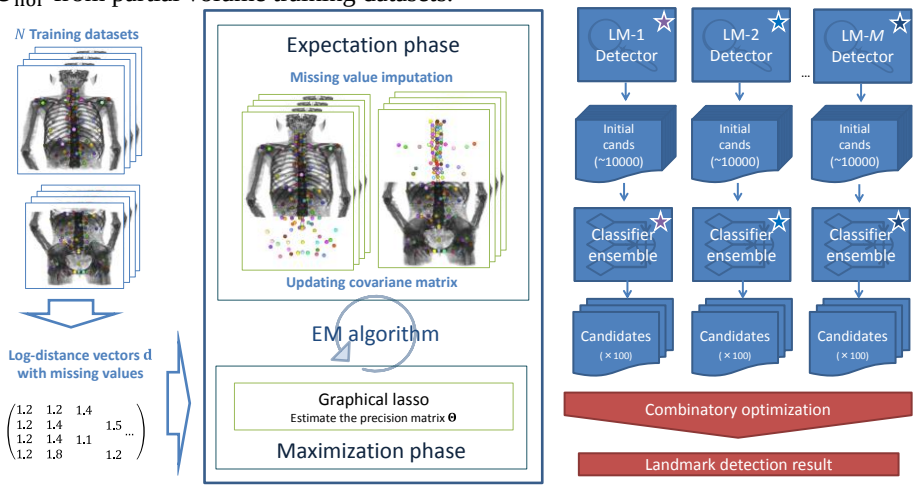


Fig. 1. Outline of proposed method. (Left) MissGLasso-based model estimation. (Right) Landmark detection system used in evaluation.

2.4 Evaluation

The positions of a total of 138 (64 chest, 37 abdominal, 37 pelvic) landmarks in the human torso were modeled and detected in this study. We compared the landmark detection performances using two different models. The *GLasso* model was trained with ground truth landmark positions from full-range (chest-to-pelvis) CT volumes. On the other hand, the *MissGLasso* model was trained with not all the ground truth positions. Instead, the pelvic landmarks were removed from half of the training datasets and the chest landmarks were removed from the other half. Thus, these two halves simulate chest-to-abdomen and abdomen-to-pelvis imaging ranges, respectively.

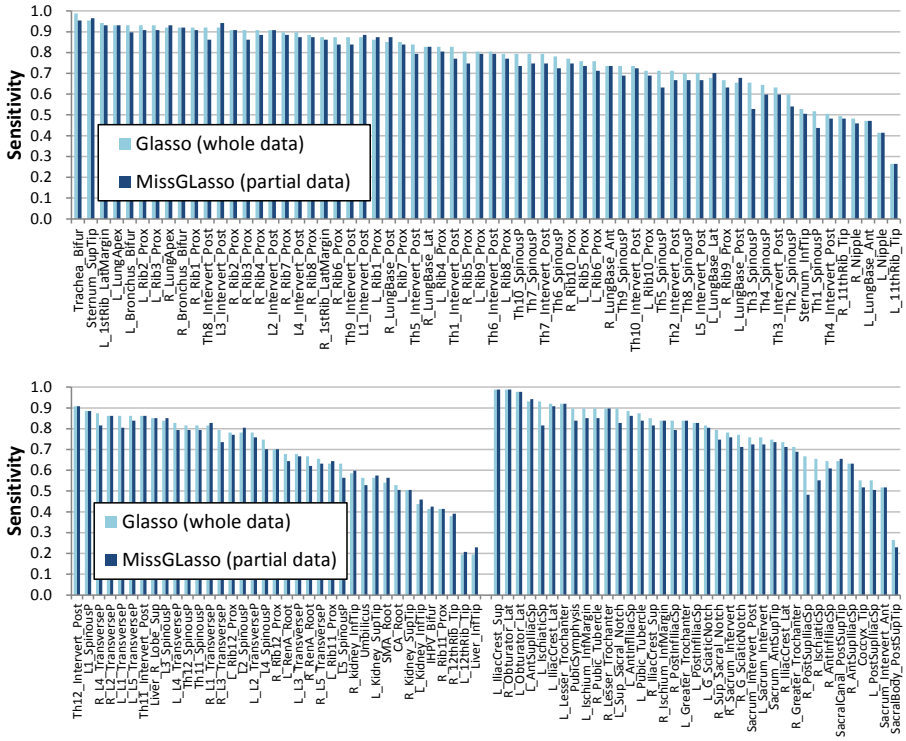


Fig. 2. Result of landmark detection performance. (Top) Sensitivities for thoracic landmarks. (Bottom, from the left) Sensitivities for abdominal and pelvic landmarks. All landmarks are sorted by the sensitivities using the *GLasso* model.

A total 87 of CT datasets of healthy or diseased human torsos without an intravenous contrast agent were used in the evaluation. The model training was performed with 4-fold cross-validation; 87 datasets were divided into 4 groups, and the datasets in each group were tested with a model that was trained using the other 3 groups. For *MissGLasso* models, these training datasets were merged and divided again into chest-to-abdomen and abdomen-to-pelvis imaging range simulation subgroups. The values of ρ and $\sigma_{\text{threshold}}$ were determined empirically as 0.75 and 0.2, respectively.

To evaluate the suitability of the estimated models, a multiple-landmark detection method proposed in [3] (Fig. 1, right) was applied to the test datasets and the detection performances using *GLasso* and *MissGLasso* models were compared. The outline of the landmark detection method used is as follows. Firstly, each of the target landmarks is detected independently by a landmark-specific detector that is composed of an appearance-model-based initial detector and a classifier ensemble. Each detector searches possible landmark positions within the given CT volume and then outputs a number of candidate points (100 candidates in this study). Then, the following combinatory optimization algorithm chooses the most suitable combination of all landmark positions from the lists of candidates. In the latter phase, a landmark distribution model [Eq. (1)] is used as prior knowledge to seek the best combination. The optimization problem is formulated by maximum *a posteriori* (MAP) estimation and is solved by Gibbs sampling and simulated annealing methods. Each detection result was evaluated as successful if the detected point was within 2 cm from the manually input ground truth point.

Both the model estimation and landmark detection system were implemented on and experimented with a computer with Intel Xeon E5-2640 CPU×2, 64 GB RAM and nVidia Tesla K20m GPGPU. It took approximately 2 h to estimate one *MissGLasso* model and 30 min for one landmark detection task.

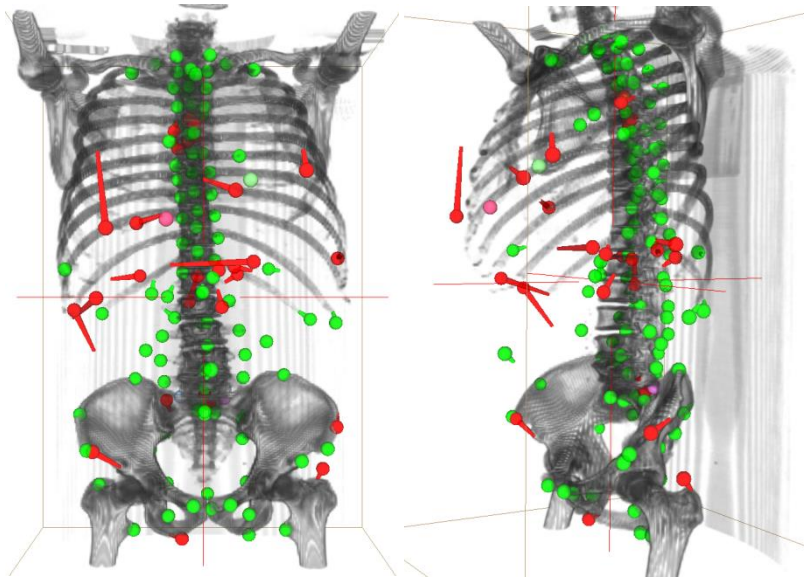


Fig. 3. An example of landmark detection result. (Left) a frontal view. (Right) a left anterior oblique view. Green and red pins represent successfully and wrongly detected landmarks, respectively. The pin heads represent the detected positions, whereas the pin tails are the true landmark positions. Most of wrongly detected landmarks were those defined on soft-tissue structures.

3 Results

The overall result is shown in Fig. 2. The averages and standard deviations of sensitivities were 75.02% ± 0.17 and 72.78% ± 0.17 with *GLasso* and *MissGLasso*, respectively. Within the detection failures of both results (24.98% and 27.22%, respectively), 9.92% were due to the detection phase (i.e., no true position was outputted by the detector as a candidate). Thus, the failures due to the combinatory optimization phase using the models were roughly estimated as 15.06% and 17.30%, respectively.

Because the detection sensitivities for most of the landmarks differ little between *GLasso* and *MissGLasso*, it can be concluded that the MissGLasso algorithm could effectively handle missing values in the training datasets. The detection performance was similar even for landmarks with very poor sensitivity.

A performance comparison with a related study recently reported by Liu et al. [8] was performed on 5 landmarks (all landmarks shared by their study and ours). The mean errors of our MissGLasso setting were 4.4, 8.4, 7.8, 10.8, 40.6 (mm) for trachea bifurcation, left and right lung tops, liver top and bottom, respectively. In Liu et al., the corresponding errors were 2.5, 2.6, 3.2, 2.5 and 6.4 (mm), respectively. Though our result showed less accuracy than [8], we believe that our results for 4 landmarks, aside from the bottom of the liver, are not so poor if ambiguities of landmark positions are taken into account. Moreover, it may be improved by any appropriate post-process (e.g., [7]).

4 Discussion

A method of estimating a landmark point distribution model from training datasets with various partial imaging ranges was presented. The deterioration of the landmark detection performance was minimal using the model estimated with only chest-to-abdomen and abdomen-to-pelvis image datasets. Therefore, it was suggested that the proposed method can be a key technique for building landmark distribution models with a wider range, or ultimately, a whole body landmark PDM.

Application of our method to build a whole-body PDM is not limited to landmark detection. Possible applications include estimation or imputation of structures unseen in a given image (due to imaging range or nature of the modality). For example, the model may estimate the body height from the landmark positions of the pelvis only.

The main motivation of introducing sparsity in this study is to handle the High-Dimensional Low Sample Size (HDLSS) problem better. Regularization is a critical issue for HDLSS. L1-norm regularization is a state-of-the-art method and its superiority against classical L2-norm regularization has been suggested by many studies [9]. Our precision matrices were very sparse, having approx. 0.1% of non-zero elements.

This study has several limitations. Firstly, the proposed method was not compared with other precision matrix estimation methods such as Tikhonov regularization. However, the landmark detection system evaluates the distribution probability (Eq. (1)) about 15 million times per detection task under our current implementation and experimental conditions. The speed of this calculation completely depends upon the

sparsity of the precision matrix in this study. Thus it is expected that landmark detection with a dense precision matrix is virtually impossible. Although in [6] Hanaoka et al. proposed an effective way of calculating Eq. (1) without explicitly considering the entire precision matrix, it instead depends on explicit training sample vectors and cannot be applied to a problem with missing values.

Secondly, the detection sensitivity is not sufficient for some landmarks. For example, the sensitivities of 42 landmarks out of 158 were less than 70% with the criterion of 2 cm from the ground truth. Improvement of the detection performance is an issue in the future. Possible solutions include the landmark position fine-tuning algorithm proposed by Nemoto et al [7] or parameter tuning for each landmark detector [2].

As a future work, we are planning to apply the proposed method to building a multimodality landmark PDM. For example, suppose that you have a number of CT datasets with manually input landmark positions, and you also have other MRI datasets with landmark positions input. If the two datasets share some of the landmarks, the proposed method can build a combined landmark distribution model from both datasets. Such a combined model will be useful in various ways. For instance, using a CT-MRI combined model, the position of a certain landmark that is only visible in MRI can be estimated in any given CT images. Our final goal is to represent all landmark positions in the whole body and in any medical images in a single landmark distribution model.

Acknowledgement

This study is a part of the research project "Computational Anatomy for Computer-aided Diagnosis and therapy: Frontiers of Medical Image Sciences", supported by a grant-in-aid for scientific research on innovative areas MEXT, Japan.

References

1. Seifert, S., Barbu, A., Zhou, S. K., Liu, D., Feulner, J., Huber, M., Suehling, M., Cavallaro, A., Comaniciu, D.: Hierarchical parsing and semantic navigation of full body CT data. In Proc. of SPIE 7259 (2009)
2. Potesil V., Kadir T., Platsch G., Brady M.: Improved Anatomical Landmark Localization in Medical Images Using Dense Matching of Graphical Models. In Proc. of BMVC 1-10 (2010)
3. Hanaoka S., Masutani Y., Nemoto M., Nomura Y., Yoshikawa T., Hayashi N., Yoshioka N. and Ohtomo K.: An improved multiple anatomical landmark detection method with combinatorial optimization and Madaboost-based candidate likelihood determination. Int. J. CARS, 7 Suppl.1, S330-331 (2012)
4. Städler N, Bühlmann P.: Missing values: sparse inverse covariance estimation and an extension to sparse regression. Statistics and Computing 22, 219-235 (2012)
5. Friedman J, Hastie T, Tibshirani R: Sparse inverse covariance estimation with the graphical Lasso. Biostatistics 9, 432–441 (2008)
6. Hanaoka S., Masutani Y., Nemoto M., Nomura Y., Yoshikawa T., Hayashi N., Yoshioka N. and Ohtomo K.: Probabilistic Modeling of Landmark Distances and Structure for

Anomaly-proof Landmark Detection. In: Proceedings of the Third International Workshop on Mathematical Foundations of Computational Anatomy 159-169 (2011)
7. Nemoto M., Masutani Y., Hanaoka S., Nomura Y., Miki S., Yoshikawa T., Hayashi N. and Ohtomo K.: Coarse-to-fine localization of anatomical landmarks in CT images based on multi-scale local appearance and rotation-invariant spatial landmark distribution model. In Proc. of SPIE 8669 (2013)
8. Liu D, Zhou S. K. Anatomical Landmark Detection Using Nearest Neighbor Matching and Submodular Optimization. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012. Lecture Notes in Computer Science Volume 7512, pp 393-401 (2012)
9. Honorio J., Samaras D.: Multi-task learning of Gaussian graphical models. Proc. of the 27th International Conference on Machine Learning, Haifa, Israel, 2010.

Cortical Shape Analysis using the Anisotropic Global Point Signature

Anand A Joshi[1,3], Syed Ashrafulla[1], David W Shattuck[2], Hanna Damasio[3] and Richard M Leahy[1] *

[1] Signal and Image Processing Institute, University of Southern California, Los Angeles, CA
[2] Laboratory of Neuro Imaging, University of California, Los Angeles, CA
[3] Brain and Creativity Institute, University of Southern California, Los Angeles, CA

Abstract. We present a novel shape representation that characterizes the shape of a surface in terms of a coordinate system based on the eigensystem of the anisotropic Laplace-Beltrami operator. In contrast to the existing techniques, our representation can capture developable transformations and is therefore useful for analysis of cortical folding patterns. This representation has desirable properties including stability, uniqueness and invariance to scaling as well as isometric transformations. Additionally, the resulting shape space has a standard Euclidean metric simplifying shape analysis. We also present an approach that provides a fast and accurate computational method for solving the eigensystem using a finite element formulation. We demonstrate the utility of this representation for two brain shape analysis applications: quantifying symmetries in shape between the two cortical hemispheres and finding variance of cortical surface shapes across populations.

1 Introduction

Quantification, matching and comparison of cortical shapes are challenging problems with wide utility [17,12]. Most of the traditional approaches for analyzing brain shapes are deformation-based. Quantitative analysis of anatomical shape differences is performed with these approaches by analyzing the deformation required to warp a subject brain to a template brain. For example, tensor-based morphometry [8] analyzes local linear approximations (the deformation tensors) of the deformation field. Alternate methods such as deformation-based morphometry [1] and pattern-based morphometry [3] use different aspects of the deformation field.

While quantification of shape differences by analysis of the deformation field is a plausible approach, it suffers from a number of disadvantages. The results of these methods depend to a large extent on the image registration method used. Only regions where registration works well – typically subcortical structures – tend to show high statistical power [8]. Using the deformation field as a shape

* This work was supported by NIH grants P41 EB015922 and R01 NS074980.

descriptor often magnifies the effects of registration errors. In addition, it is not clear if the shape details are indeed encoded in the deformation tensor as this tensor is a local linear approximation of the deformation field. This is especially the case for large shape differences that require large deformations. Also, there is no ground truth deformation for the purpose of aligning one brain to another; in other words, registration may provide only one of multiple equally accurate deformation fields from one brain image to another. Finally, the deformation field does not define a shape space on the cortex in the sense that sulcal and gyral shapes are not directly encoded in the deformation field.

Recent approaches for brain shape analysis are based on spectral geometry, in which the shape of a manifold is characterized by the eigenspectrum of a differential operator defined on the manifold. An invariant representation of a 2D surface can be generated using the Global Point Signature (GPS) representation, which is based on the eigensystem of the isotropic Laplace-Beltrami operator defined on that surface [13,14]. Methods using GPS are not directly applicable

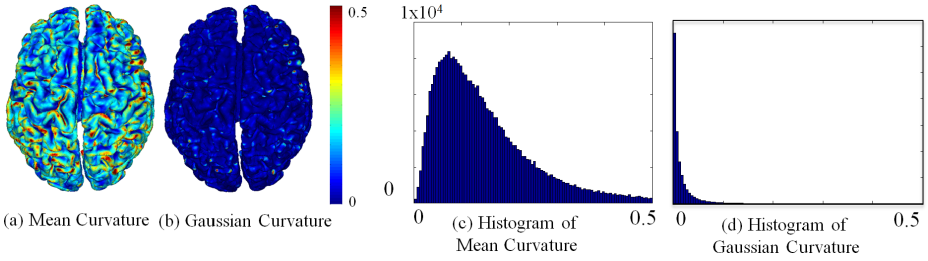


Fig. 1. Absolute value of mean and Gaussian curvature of a cortical surface.

to cortical shape analysis because the isotropic Laplace-Beltrami operator only captures the intrinsic geometry of the surface. The majority of the curvature information of the cortex is in the mean curvature, as seen in figure 1, which is extrinsic to the surface. 3D shape descriptors such as spherical harmonics [7] are not convenient for cortical shape analysis since such descriptors require an impractically large number of basis functions and do not efficiently encode shape information related to elastic deformations of shapes. This paper presents an approach for shape analysis of 2D surface patches using an anisotropic version of the Laplace-Beltrami operator as described in the next section.

2 Materials and Methods

2.1 AGPS Shape Representation

We assume as input an anatomically labeled cortical surface representation such as obtained by BrainSuite [15,6]. Motivated by spectral theory [14,13,5], we model a surface S representing a cortical region as an inhomogeneous vibrating membrane. Its harmonic behavior is thus governed by the 2D anisotropic

Helmholtz equation where the mean curvature $\kappa(s)$ is used to introduce anisotropy:

$$\begin{cases} \nabla \cdot \kappa(s) \nabla \Phi(s) & = \lambda \Phi(s) \\ \frac{\partial \Phi(s)}{\partial \boldsymbol{n}}|_{\partial S} & = 0 \end{cases}, \forall s \in S, \tag{1}$$

where ∇ denotes a gradient operator defined in the geometry of the surface, $\Phi(s)$ represents an eigenfunction with eigenvalue λ, ∂S is the boundary of the surface patch and $\boldsymbol{n}$ is the normal to the surface. We use the eigenfunctions and eigenvalues to define the Anisotropic Global Point Signature (AGPS) embedding of the surface S in the spectral domain by the map:

$$\text{AGPS}(s) = \left(\frac{1}{\sqrt{\lambda_1}} \Phi_1(s), \frac{1}{\sqrt{\lambda_2}} \Phi_2(s), \frac{1}{\sqrt{\lambda_3}} \Phi_3(s), \ldots \right), \; p \in S \tag{2}$$

where $\Phi_1, \Phi_2, \ldots$ are eigenfunctions with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots$ arranged in ascending order. Thus, each point of the manifold is embedded into an infinite-dimensional space. The importance of modeling anisotropy in the representation is illustrated in figure 2. A developable transformation (bending)

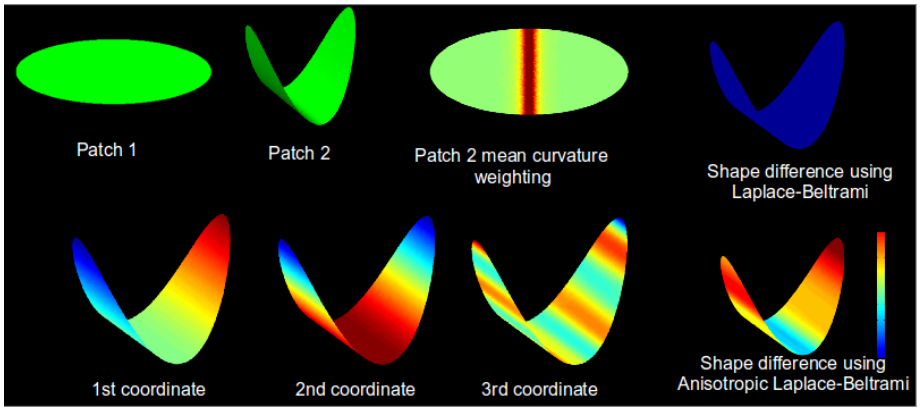


Fig. 2. Introduction of an anisotropic term helps characterize developable transformations. Use of the isotropic Laplace-Beltrami operator does not find any shape differences between the two elliptical patches as the transformation between the two is developable (top right). Use of an anisotropic operator with a mean curvature anisotropy helps in capturing the shape differences (bottom right). The resulting AGPS coordinates are also shown in the bottom row.

applied to an elliptical surface does not affect the intrinsic geometry of the surface and therefore the Gaussian curvature remains unchanged. However, the extrinsic geometry is altered by this transformation resulting in changes in mean curvature. It can be seen that the bending does not change the isotropic Laplace-Beltrami eigenspectrum and therefore the shape change is not detected by GPS. However, the anisotropy introduced via the mean curvature allows the AGPS to successfully capture the change in the shape.

The AGPS embedding presented in equation (2) has many favorable properties. First, its coordinates are isometry invariant as they depend only on the derivatives and curvature, which in turn are dependent only on the shape. Second, scaling the manifold by a factor α results only in scaling mean curvature by $1/\alpha$. Therefore, we can obtain scale-invariance if desired by normalizing the eigenvalues; however, for the specific application of cortical shape representation, we do not want scale-invariance. Third, changes of the manifold's shape result in continuous changes in the spectrum so the representation is stable. Fourth, in the embedding space, the inner product is given by the anisotropic Green's function due to the identity $G(s_1, s_2) = \sum_k \frac{\Phi_k(s_1)\Phi_k(s_2)}{\lambda_k}$, $s_1, s_2 \in S$ [13]. As a result, the AGPS representation encodes anisotropic diffusion distances [9] on the surface. In addition, both local and global shape information is represented in the embedding. Finally, in this infinite-dimensional shape space, the metric is Euclidean, allowing standard ℓ^2 space analysis.

This invariant spectral geometric representation of surfaces has interesting physical interpretations. The surfaces can be modeled as vibrating membranes and the vibrations are damped proportionally to the mean curvature at each point. The anisotropic Laplace-Beltrami eigenspectrum corresponds to the modes of vibrations of this membrane (Fig. 2 (bottom)). Thus the AGPS representation encodes information about the modes of vibration of membranes (surface patches) as the basis for shape modeling. The AGPS shape representation intuitively encodes curvature characteristic at and around the points on the manifold. Perturbations at a point in a shape lead to local changes in curvature around that point which are captured in higher-order AGPS coordinates. On the other hand global shape changes lead to curvature changes everywhere in the shape which are captured by lower-order AGPS coordinates. Due to this association between AGPS coordinates and the spatial extent of shape changes, AGPS-based comparisons provide a natural description of changes in shape at different scales.

2.2 Numerical Implementation

To solve equation (1) we first compute the anisotropy term represented by mean curvature using the method described in [10]. Next, we use a finite element method (FEM) to discretize the anisotropic Helmholtz equation (1). We discretize the derivative operators using FEM directly in the geometry of the surface mesh, and therefore we do not need to explicitly compute the Riemannian metric coefficients as is often done if the surfaces are mapped to a plane or sphere [17]. We choose linear FEMs for functions and Galerkin's formulation [16] for robustness to tessellation errors. Let $\Phi(s) = \sum_i \phi_i e_i(s)$ be an eigenfunction and $N(s) = \sum_i \eta_i e_i(s)$ be a 'test function', each represented as weighted sums of linear elements $e_i(s)$. The eigenvalue problem from equation (1) then becomes:

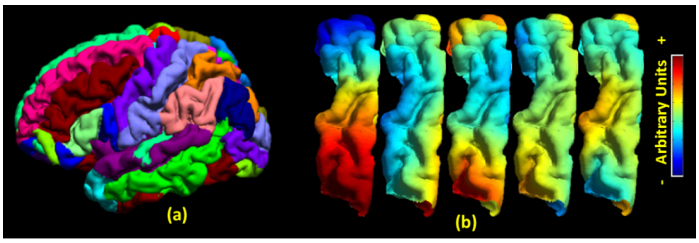


Fig. 3. An AGPS example: (a) automatically generated parcellations of a cortical surface; (b) first five color-coded AGPS coordinates of the left superior-frontal gyrus.

$$
\begin{aligned}
(\nabla \cdot \kappa(s)\nabla)\,\Phi(s) &= \lambda\Phi(s) \\
\Longrightarrow \int_S (\nabla \cdot \kappa(s)\nabla\Phi(s))\,N(s)ds &= \lambda \int_S \Phi(s)N(s)ds \\
\Longrightarrow -\int_S \kappa(s)\nabla\Phi(s)\nabla\eta N(s)ds &= \lambda \int_S \Phi(s)N(s)ds
\end{aligned}
$$

where the latter follows using integration by parts and Neumann boundary conditions in equation (1). Substituting the FEM into this equation, we get:

$$
\begin{aligned}
-\sum_i \sum_j \phi_i \eta_j \kappa_{ij} \int \nabla e_i(s)\nabla e_j(s)ds &= \lambda \sum_i \sum_j \phi_i \eta_j \int e_i(s)e_j(s)ds \\
\Longrightarrow\ KS\Phi &= -\lambda M\Phi \qquad (3)
\end{aligned}
$$

where $\kappa_{ij} = \frac{\kappa_i + \kappa_j}{2}$ is the average of curvatures calculated at points i and j, K is a matrix with i^{th} row and j^{th} column given by κ_{ij}, and Φ is a column vector with i^{th} entry given by ϕ_i. For a triangulated surface mesh with linear elements, the element-wise matrix is given by $M_{el} = \frac{A_{el}}{12}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$ and the element-wise stiffness matrix is given by $S_{el} = D_x D_x + D_y D_y$ where D_x and D_y are discretizations of derivatives in the x and y directions, respectively. The mass and stiffness matrices M and S are obtained from the corresponding element-wise matrices M_{el} and S_{el} respectively by finite element matrix assembly procedures as described in [16]. The matrix equation (3) is a generalized eigenvalue problem that can be solved using standard methods such as the QZ method in the Matlab function `eigs`. For this analysis, we chose to approximate the infinite-dimensional AGPS by its first seven coordinates based on the spread of the eigenvalue spectrum. One example of the computed AGPS coordinates for the surface patch representing the left superior frontal gyrus is shown in figure 3.

2.3 Brain Shape Analysis using AGPS

In order to illustrate the potential of the proposed representation for cortical shape analysis, we apply AGPS for two group studies of $N = 24$ subjects: (1)

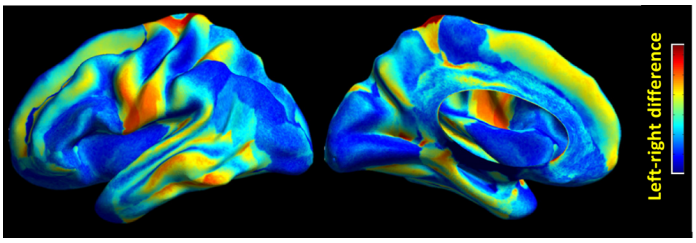


Fig. 4. Left-to-right hemisphere average shape difference, based on gyral AGPS. The color-coded overlay shows the degree of symmetry (blue) and asymmetry (red).

asymmetry analysis and (2) variability analysis. For this purpose, we generated the AGPS coordinates for each vertex using the method described in the previous section. While it is possible to get a full description of shape change at different scales by analyzing coordinate-wise AGPS differences, we summarize the shape difference with ℓ^2-norm since the shape space admits an Euclidean metric.

To map symmetry between the left and right hemispheres of subjects, we first define AGPS_L and AGPS_R as the AGPS representations of left and right cortical hemispheres of a subject and transfer them to the common atlas space for comparison. We register the atlas's right hemisphere R to the atlas's left hemisphere L forming a correspondence denoted by $\Psi : R \to L$. With this correspondence, we can then compute mean AGPS distance between hemispheres at each vertex s in the atlas's right hemisphere by $D(s) = \|\mathrm{AGPS}_L(\Psi(s)) - \mathrm{AGPS}_R(s)\|_2$. To find group asymmetry, we average $D(s)$ over all the subjects.

Shape variability on the cortex can be found by estimating the population variance of the AGPS coordinates. We compute the AGPS coordinates for each subject in the native space and then transfer these coordinates to the standard atlas space. We then estimate the population variance at each vertex s in the atlas space by $\sigma^2(s) = \frac{1}{N}\sum_{n=1}^{N}\left(\mathrm{AGPS}_n(s) - \frac{1}{N}\sum_{m=1}^{N}\mathrm{AGPS}_m(s)\right)^2$. Note that we use the correspondence established using surface registration for comparing the shapes in the atlas domain, but we do not use the deformation field as a shape descriptor due to the reasons discussed in section 1.

3 Results

3.1 Asymmetry across hemispheres

After parcellation, the left and right hemispheres of a subject's brain contain homologous regions that may differ in shape. This hemispheric brain asymmetry is possibly related to functional lateralization due to evolutionary, hereditary and developmental factors [18]. In order to map cortical asymmetry, we use the procedure described in section 2.3 with results shown in figure 4. The most asymmetric regions are in the inferior sector of the pre- and post-central gyrus, the mid portion of the middle temporal gyrus and posterior portion of the inferior temporal gyrus, and, to a lesser extent in the mesial sector of the superior frontal

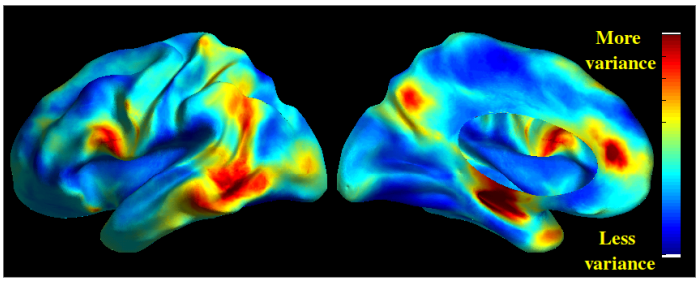


Fig. 5. Population variance of cortical shape. The color-coded overlay is the variance of the AGPS representation, plotted on an inflated representation of the cortex.

gyrus. One slightly surprising result is that there is minimal asymmetry at the end of the Sylvian fissure and needs further investigation.

3.2 Variability in Shapes

In addition to analyzing shape variability between cortical hemispheres across the subject population, we can analyze the variability in cortical shape over a population. In figure 5 on the left (lateral aspect of the hemisphere) we see high variability in the posterior sector of the middle and inferior temporal gyri and the inferior parietal lobule hugging the posterior end of the Sylvian fissure, as well as in the frontal operculum; on the right (mesial aspect of the hemisphere) the areas of maximal variability are found in the pre-cuneus, the anterior sector of the cingulate gyrus and in the anterior sector of the parahippocampal gyrus (site of the maximal variance). Most of these regions are in the association cortex.

4 Discussion and Conclusion

This paper presents a new invariant shape representation, AGPS, that captures differences in surface shapes due to developable transformations, a critical class of transformations in the analysis of cortical folds. The shape space generated by this representation is ℓ^2, readily allowing the use of existing standard statistical techniques for shape analysis. We illustrated the benefits of AGPS in quantifying shape differences across hemispheres and shape variation across subjects.

The question of whether a surface is unique (within an isometry) given its AGPS coordinates is related to the existence of Bonnet surfaces, i.e. surfaces that are not completely defined by their metric and mean curvature [2]. However, it is unlikely that cortical surfaces would suffer from such an ambiguity. Another possible ambiguity is in the order of eigenfunctions of the anisotropic Laplace-Beltrami operator in cases where there are repeated eigenvalues, which can occur if there are certain symmetries in the shape. An algorithm for resolving this ambiguity can be found in [4]. For the purpose of brain shape analysis, we did not encounter any ambiguity possibly due to shapes not having axes of symmetry.

Many aspects of the preprocessing can affect the AGPS representations of cortical regions, thus introducing sources of error in cortical shape analysis.

More specifically, it is unclear whether different parcellation schemes can result in different shape analysis results. We note that recent registration and labeling methods can parcellate the brain with accuracy approaching manual labeling [11]. In the future, we plan to explore the effect of parcellation decisions on cortical shape analysis using AGPS.

The presented AGPS representation and analysis methods can be applied for applications other than cortical shape analysis. AGPS representations can help in various computer vision applications requiring shape analysis where the intrinsic geometry does not fully capture the shape. In addition, this representation can also be extended to 3 or more dimensions using approaches presented in this paper for 2D surfaces and in previous work for 1D curves [5].

References

1. Ashburner, J., Hutton, C., Frackowiak, R., Johnsrude, I., Price, C., Friston, K.: Identifying global anatomical differences: deformation-based morphometry. Human Brain Mapping 6(5-6), 348–357 (1998)
2. Bobenko, A., Eitner, U.: Bonnet surfaces and Painlevé equations. Journal fur die Reine und Angewandte Mathematik 499, 47–80 (1998)
3. Gaonkar, B., Pohl, K., Davatzikos, C.: Pattern based morphometry. Proc MICCAI pp. 459–466 (2011)
4. Jain, V., Zhang, H.: Robust 3D shape correspondence in the spectral domain. In: SMI 2006. pp. 19–19. IEEE (2006)
5. Joshi, A., Ashrafulla, S., Shattuck, D., Damasio, H., Leahy, R.: An invariant shape representation using the anisotropic Helmholtz equation. MICCAI 2012 pp. 607–614 (2012)
6. Joshi, A., Shattuck, D., Leahy, R.: A method for automated cortical surface registration and labeling. In: Proceedings of the 5th international conference on Biomedical Image Registration. pp. 180–189. Springer-Verlag (2012)
7. Kazhdan, M., Funkhouser, T., Rusinkiewicz, S.: Rotation invariant spherical harmonic representation of 3D shape descriptors. In: Proc Eurographics/ACM SIGGRAPH. pp. 156–164. Eurographics Association (2003)
8. Lepore, N., Brun, C., Chou, Y., Chiang, M., Dutton, R., Hayashi, K., Luders, E., Lopez, O., Aizenstein, H., Toga, A., et al.: Generalized tensor-based morphometry of HIV/AIDS using multivariate statistics on deformation tensors. IEEE Trans Med Image 27(1), 129–141 (2008)
9. Lipman, Y., Rustamov, R., Funkhouser, T.: Biharmonic distance. ACM Transactions on Graphics (TOG) 29(3), 27 (2010)
10. Meyer, M., Desbrun, M., Schröder, P., Barr, A.: Discrete differential-geometry operators for triangulated 2-manifolds. Visualization and Math 3(7), 34–57 (2002)
11. Pantazis, D., Joshi, A., Jiang, J., Shattuck, D., Bernstein, L., Damasio, H., Leahy, R.: Comparison of landmark-based and automatic methods for cortical surface registration. NeuroImage 49(3), 2479–2493 (2010)
12. Rettmann, M.E., Prince, J.L., Resnick, S.M.: Analysis of sulcal shape changes associated with aging. Proc Human Brain Mapping (2003)
13. Reuter, M.: Hierarchical shape segmentation and registration via topological features of Laplace-Beltrami eigenfunctions. IJCV 89(2), 287–308 (Sep 2010)
14. Rustamov, R.: Laplace-Beltrami eigenfunctions for deformation invariant shape representation. In: Proc 5th Euro Symp Geom Proc. pp. 225–233 (2007)

15. Shattuck, D.W., Leahy, R.M.: BrainSuite: An automated cortical surface identification tool. Medical Image Analysis 8(2), 129–142 (2002)
16. Smith, G.D.: Numerical solution of partial differential equations: Finite difference methods. Oxford : Clarendon Press (1985)
17. Thompson, P., Woods, R., Mega, M., Toga, A.: Mathematical/computational challenges in creating deformable and probabilistic atlases of the human brain. Human Brain Mapping 9(2), 81–92 (2000)
18. Toga, A., Thompson, P.: Mapping brain asymmetry. Nature Reviews Neuroscience 4(1), 37–48 (2003)

Author Index