Large Deformation Unbiased Diffeomorphic Nonlinear Image Registration: Theory and Implementation

UCLA CAM Report 06-71, 2006

Igor Yanovsky 1 Paul M. Thompson 2 Stanley Osher 1 Alex D. Leow 2, 3

1 Department of Mathematics, University of California, Los Angeles, CA 90095
2 Laboratory of Neuro Imaging, UCLA David Geffen School of Medicine, Los Angeles, CA 90095
3 Neuropsychiatric Institute, UCLA David Geffen School of Medicine, Los Angeles, CA 90095

Abstract

In this paper, we present a novel framework for constructing large deformation log-unbiased image registration models that generate theoretically and intuitively correct deformation maps. Such registration models do not rely on regridding and are inherently topology preserving. We apply information theory to quantify the magnitude of deformations and examine the statistical distributions of Jacobian maps in the logarithmic space. To demonstrate the power of the proposed framework, we generalize the well known viscous fluid registration model to compute log-unbiased deformations. We tested the proposed method using a pair of binary corpus callosum images, a pair of two-dimensional serial MRI images, and a set of three-dimensional serial MRI brain images. We compared our results to those computed using the viscous fluid registration method, and demonstrated that the proposed method is advantageous when recovering voxel-wise maps of local tissue change.

Index Terms

Nonlinear image registration, information theory, mutual information, log-unbiased deformation, biomedical imaging, energy minimization, variational methods.

1. Introduction

In recent years, computational neuroimaging has become an exciting interdisciplinary field with many applications in functional and anatomic brain mapping, image-guided surgery, and multimodality image fusion [1, 8, 14]. The goal of image registration is to align, or spatially normalize, one image to match another. In general, the transformation that defines the correspondence map between the images should be diffeomorphic, to preserve the topology.

To construct a deformation that is one-to-one and differentiable [4, 10, 11], we must impose a regularizing constraint. Thus, the problem of image registration is often cast as a minimization problem with a combined cost functional consisting of an image matching functional and a regularizing constraint on the deformation. Common choices of image matching functional include squared intensity difference, cross correlation [5], and (normalized) mutual information or other divergence-based or information-theoretic measures [6, 9, 12], while choices of regularization usually involve differential operators inspired by thin-plate spline theory, elasticity theory, fluid dynamics and the Euler-Poincaré equations [11, 13].

2. Background

Let us denote the template image as $T(\vec{x})$ and the study image as $S(\vec{x})$, which are images defined on the spatial domain $\Omega$. The problem of image registration is to find a displacement field $\vec{u}(\vec{x})$ at each point $\vec{x}$ such that the deformed template $T^*(\vec{x}) = T(\vec{x} - \vec{u})$ is in some sense close to $S(\vec{x})$. The term displacement is used because it can be viewed as how a point in the deformed template is moved away from its original location. The most common way to define the distance between the deformed template and the study is the $L^2$ norm

$$E_{L^2}(T, S, \vec{u}) = \frac{1}{2} \int_\Omega |T(\vec{x} - \vec{u}) - S(\vec{x})|^2 d\vec{x}, \quad (1)$$

which is also known as a Gaussian sensor model. This distance can be minimized using the gradient descent of the corresponding Euler-Lagrange equation to obtain

$$\frac{\partial \vec{u}(\vec{x}, t)}{\partial t} = -\vec{f}(\vec{x}, \vec{u}(\vec{x}, t)), \quad (2)$$
where $t$ is an artificial time and

$$f(\vec{x}, \vec{u}(\vec{x}, t)) = -T(\vec{x} - \vec{u}) - S(\vec{x}) \nabla T|_{\vec{x} - \vec{u}}$$

(3)
is the force field or the body force, which drives the template into registration with the study. The first term in the definition of $f$, namely $\nabla T|_{\vec{x} - \vec{u}}$, is the gradient of the deformed template and has largest values at the edges of the template. This term determines the directions of the local deformation forces applied to the template. The second term $T(\vec{x} - \vec{u}) - S(\vec{x})$ is the difference in intensity between the deformed template and the study. This term causes the field force to tend to zero in areas where the deformed template is locally aligned with the study.

This optimization problem, however, is known to be ill-posed. In particular, the displacement field $\vec{u}$ is not unique, and the regularization on $\vec{u}$ is required to make the problem be well-posed. In the next section, we will review different approaches to regularizing the displacement field.

3. Previous Work

An important observation, which stimulated the development of intensity-based nonlinear image registration algorithms, was the connection of the image data with a physically deforming system in three dimensions. Physical continuum models consider the deforming image to be embedded in a three-dimensional deformable medium, which can be either an elastic material or a viscous fluid. The medium is subjected to certain distributed internal forces, which reconfigure the medium and eventually drive the template into registration with the study. In this section, we briefly describe two of the most well known such models.

3.1. Elastic Registration

In [2, 3, 7], the authors noticed the similarity between image deformation and deformation of elastic plates. For linear elastic solids, the force field $\vec{f}$ is proportional to the displacement field $\vec{u}$. The spatial transformation satisfies the partial differential equation

$$\mu \nabla^2 \vec{u} + (\mu + \nu) \nabla (\nabla \cdot \vec{u}) = \vec{f}(\vec{x}, \vec{u}),$$

(4)where $\mu$ and $\nu$ are Lamé constants, describing the properties of the material.

The linear elastic equation (4) is derived assuming small angles of rotation and small linear deformations. As a result, the elastic registration model severely penalizes large displacements and is not useful in applications when large nonlinear deformations are natural.

3.2. Viscous Fluid Registration

In elastic deformation, particles are usually tracked by their initial coordinates in the Lagrangian frame of reference. In the viscous fluid model, first proposed by Christensen et al. [4], an Eulerian reference frame is used instead. The Eulerian frame of reference specifies the time evolution of particle positions and velocities as observed at fixed points. Consequently, a particle located at $\vec{x}$ at time $t$ originated at position $\vec{h}(\vec{x}, t) = \vec{x} - \vec{u}(\vec{x}, t)$ at time $t_0$ ($t > t_0$), where $\vec{u}$ is the displacement. We let $\vec{v}$ denote the velocity field. The material derivative, defined by $D/Dt = \partial / \partial t + \vec{v} \cdot \nabla$, describes the time rate of change experienced by an element of material instantaneously at point $\vec{x}$ at time $t$. Hence, the Eulerian velocity field $\vec{v}$ is nonlinearly related to $\vec{u}$ and is determined by

$$\vec{v} = \frac{D\vec{u}}{Dt} = \frac{\partial \vec{u}}{\partial t} + \vec{v} \cdot \nabla \vec{u}.$$ (5)

Given the velocity field $\vec{v}$, equation (5) can be solved to obtain the displacement field $\vec{u}$. In [4], the authors considered the deforming template image to be embedded in a viscous fluid whose motion is governed by Navier-Stokes equation for conservation of momentum. Some simplification of the momentum conservation equation resulted in the following equation for $\vec{v}$:

$$\mu \Delta \vec{v} + (\mu + \nu) \nabla (\nabla \cdot \vec{v}) = \vec{f}(\vec{x}, \vec{u}).$$

(6)
The $\Delta \vec{v}$ term is the viscosity, which constrains the velocity field to vary smoothly. The term $\nabla (\nabla \cdot \vec{v})$ allows structures in the template to change in mass. The nonlinear deformation force field $\vec{f}$ used in this model is defined as in equation (3).

Since equation (6) is computationally expensive to solve in practice, the authors in [6] proposed to obtain the instantaneous velocity from the convolution of $\vec{f}$ (up to a sign) with a Gaussian kernel $G_\sigma$, with variance $\sigma$:

$$\vec{v} = G_\sigma * (-\vec{f}(\vec{x}, \vec{u})).$$

(7)

Even though the viscous fluid registration model allows large deformations, numerical implementation of this model may not produce diffeomorphic and topology preserving maps unless regridding is used.

4. Theory

One could not study nonlinear image registration without closely examining Jacobian maps. In this section, we provide rigorous mathematical analyses of the Jacobian maps and use them to construct unbiased nonlinear image registration.

4.1. Global Preservation of Density Maps

We study smooth deformations $\vec{h}$ that map a computational domain $\Omega$ bijectively onto itself. Let us assume, without loss of generality, that the volume of this domain is 1,
It is defined as the determinant of the Jacobian matrix
\[ |D\vec{h}(\xi)|d\xi = \int_{\Omega} d\vec{y} = 1, \]
\[ |D\vec{h}^{-1}(\xi)|d\xi = \int_{\Omega} d\vec{x} = 1. \]  

Given global preservation of density maps, we can associate three probability density functions to \( \vec{h}, \vec{h}^{-1} \), and the identity map \( id \):

\[ P_h(\cdot) = |D\vec{h}(\cdot)|, \]
\[ P_{h^{-1}}(\cdot) = |D\vec{h}^{-1}(\cdot)|, \]
\[ P_{id}(\cdot) = 1. \]  

Differentiating the identity \( \vec{h}^{-1}(\vec{h}(\vec{x})) = \vec{x} \) on both sides and setting \( \vec{y} = \vec{h}(\vec{x}) \), we obtain

\[ D\vec{h}^{-1}(\vec{y}) \cdot D\vec{h}(\vec{x}) = id, \]

and hence,

\[ |D\vec{h}^{-1}(\vec{y})| \cdot |D\vec{h}(\vec{x})| = 1. \]

By associating deformations with their corresponding global density maps, we can now apply information theory to quantify the magnitude of deformations. In our approach, we choose the symmetric Kullback-Leibler \( (sKL) \) distance:

\[ sKL(P_h, P_{id}) = KL(P_{id}, P_h) + KL(P_h, P_{id}) \]

(12)

To measure the magnitude of any deformation \( \vec{h} \). Here \( KL \), the Kullback-Leibler distance between two probability density functions \( X \) and \( Y \), is defined as

\[ KL(X,Y) = \int_{\Omega} X \log \frac{X}{Y} d\vec{x} \geq 0. \]

(13)

To motivate this approach, notice that the first part of \( sKL \) measure is simply integrating the log-density over the entire computational image domain:

\[ \int_{\Omega} \log |D\vec{h}(\vec{x})|d\vec{x} = \int_{\Omega} \log \frac{1}{|D\vec{h}(\vec{x})|} d\vec{x} \]
\[ = -\int_{\Omega} P_{id} \log \frac{P_{id}}{P_h} d\vec{x} \]
\[ = -KL(P_{id}, P_h) \leq 0. \]

(14)

Figure 1. Corpus callosum example. (a) image T; (b) image S; (c) image T is deformed to image S using Christensen’s model; (d) image T is deformed to image S using the proposed model.

Figure 2. Corpus callosum example. Results obtained with Christensen’s model and the proposed model. Blue, yellow and red contours represent the boundaries of corpus callosum in T, S, and deformed T, respectively. For both methods, yellow contour is essentially invisible due to a very close match. However, the resulting grid of the proposed method is visually more regular. Also, note the grid lines merging and self-crossing for Christensen’s model, signifying a topology change.

To attach geometric meaning to the second term, we notice that the \( KL \) distance has skew-symmetry with respect to \( \vec{h} \) and its inverse

\[ KL(P_{id}, P_{h^{-1}}) = -\int_{\Omega} \log |D\vec{h}^{-1}(\vec{y})|d\vec{y} \]
\[ = \int_{\Omega} (\log |D\vec{h}(\vec{x})|) |D\vec{h}(\vec{x})|d\vec{x} \]
\[ = \int_{\Omega} P_h \log \frac{P_h}{P_{id}} d\vec{x} \]
\[ = KL(P_h, P_{id}). \]

(15)
where the second equality was obtained using a change of variables, \( \tilde{y} = \tilde{h}(\tilde{x}) \). Similarly, we have

\[
KL(P_{id}, P_h) = KL(P_{h^{-1}}, P_{id}).
\]  

(16)

### 4.2. Unbiased Deformation in the Logarithmic Space

Before developing formulations to construct unbiased deformations in the logarithmic space, we generalize equation (14) to the case of mapping regions of interest (ROI). Assuming we have a priori knowledge that one ROI is mapped to another, we would like to recover a mapping that is unbiased in the logarithmic space. Intuitively, without further knowledge other than overall ROI matching, the resulting Jacobian map should take a constant value inside the ROI. This can be achieved using the proposed formulations. Indeed, given any deformation \( \tilde{g} \), we have the following

\[
\frac{1}{a} \int_A \log |D\tilde{g}\tilde{y}| d\tilde{x} \leq \log \frac{b}{a},
\]  

(17)

with equality obtained if and only if the Jacobian map of \( \tilde{g} \) takes a constant value (i.e., \( b/a \)). This generalization can be shown by observing that the logarithmic mapping is a convex mapping:

\[
\sum_{i=1}^{n} \log(x_i) \leq n \log(\bar{x}); \quad \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i.
\]  

(18)

With the above generalization, one can see that, assuming the only constraint is an ROI deformation from \( A \) to \( B \), the unbiased mapping under the logarithmic operation has an evenly distributed Jacobian field, which is also intuitively correct (as there is no reason to assume non-uniformity of the Jacobian field inside the ROI).

Given equation (14) and its generalization, we propose to quantify the distance between any given deformation and the identity map by computing the symmetric \( KL \) distance through their density functions. Due to the above mentioned skew-symmetry, this distance takes the following several equivalent forms:

\[
sKL(P_h, P_{id}) = sKL(P_{h^{-1}}, P_{id})
\]

\[
= KL(P_h, P_{id}) + KL(P_{h^{-1}}, P_{id})
\]

\[
= KL(P_{id}, P_h) + KL(P_{id}, P_{h^{-1}})
\]

\[
= KL(P_{id}, P_h) + KL(P_{h^{-1}}, P_{id})
\]

(19)

To see why minimizing equation (19) leads to unbiased deformation in the logarithmic space, we observe that the integrand is always non-negative, and only evaluates to zero when \( \tilde{h} \) is volume-preserving everywhere (Jacobian of \( \tilde{h} \) is 1 everywhere). Thus, by treating it as a cost, we recover zero-change by minimizing this cost when we compare images differing only in noise. Also, this approach is unbiased for mapping ROIs in the logarithmic space, due to the inequality in (17).

---

Figure 3. Corpus callosum example. Jacobian map of the deformation is superimposed with the deformed image for Christensen’s model and the proposed model.

Figure 4. Corpus callosum example. Histograms of Jacobian values of the deformations inside corpus callosum for Christensen’s model and the proposed model.

Figure 5. Corpus callosum example. (a) Standard deviation of Jacobian values inside corpus callosum per iteration. (b) Symmetric \( KL \) distance. For Christensen’s model (dashed blue), both standard deviation and symmetric \( KL \) distance increase while for the proposed model (solid red), both standard deviation and symmetric \( KL \) distance decrease and stabilize.
5. Implementation

In this section, we generalize the viscous fluid registration model (referred to as Christensen’s method) described in section 3.2. We provide a new energy functional incorporating the symmetric Kullback-Leibler distance in section 4, give corresponding force field equations, and describe an algorithm for the new log-unbiased viscous fluid registration model.

Rather than minimizing energy functional defined in equation (1), we propose to minimize the following energy:

\[
E(T, S, \vec{u}) = \frac{1}{2} \int_{\Omega} |T(x) - \vec{u}(x)|^2 \, d\vec{x} + \lambda \int_{\Omega} (|D\vec{h}(\vec{x})| - 1) \log |D\vec{h}(\vec{x})| \, d\vec{x},
\]

(20)

where \( \lambda > 0 \) is the Lagrange multiplier. Without loss of generality, we will consider the case where \( \Omega \subset \mathbb{R}^2 \). The equations that follow can be easily extended to three spatial dimensions. Denoting \( \vec{x} = (x_1, x_2) \) and \( \vec{h} = (h_1, h_2) \), we can write the Jacobian matrix as \( D\vec{h} = (\frac{\partial h_i}{\partial x_j}) \), where \( i, j = 1, 2 \). Defining function \( L \) as \( L(y) = (y - 1) \log y \), the force field vector \( \vec{f} \) for energy in (20) will be given as

\[
\vec{f}(\vec{x}, \vec{u}(\vec{x}, t)) = -[T(\vec{x} - \vec{u}) - S(\vec{x})] \nabla T |_{\vec{x} - \vec{u}} - \lambda \left[ -\frac{\partial}{\partial x_1} \left( \frac{\partial h_2}{\partial x_2} L' \right) + \frac{\partial}{\partial x_2} \left( \frac{\partial h_1}{\partial x_2} L' \right) \right],
\]

(21)

where \( L' = L'(|D\vec{h}|) = 1 + \log(|D\vec{h}|) - 1/|D\vec{h}| \). We are now ready to give an algorithm for the proposed method.

Algorithm. Log-Unbiased Nonlinear Image Registration

Initialize \( t = 0 \) and \( \vec{u}(\vec{x}, 0) = 0 \).

1) Given \( \vec{u}(\vec{x}, t) \), calculate the force field \( \vec{f}(\vec{x}, \vec{u}(\vec{x}, t)) \) using equation (21).

   Note that the viscous fluid model, described in section 3.2, obtains the force field using equation (3).

2) Solve (7) for the instantaneous velocity \( \vec{v}(\vec{x}, t) \).

Steps 3-5 describe the procedure for solving equation (5), advancing \( \vec{u}(\vec{x}, t) \) in time.

3) Calculate the perturbation of the displacement field \( \vec{R}(\vec{x}) = \vec{v}(\vec{x}, t) - \vec{u}(\vec{x}, t) \).

4) Time step \( dt \) is calculated adaptively so that \( dt \cdot \max(||\vec{R}||) = \Delta t \), where \( \Delta t \) is the maximal displacement allowed in one iteration. Results in this paper are obtained with \( \Delta t = 0.1 \).

5) Advance equation (5), i.e., \( \partial \vec{u}(\vec{x}, t)/\partial t = \vec{R}(\vec{x}) \), in time, with time step from step 4, solving for \( \vec{u}(\vec{x}, t) \).

6) If the cost functional in (20) decreases by sufficiently small amount compared to the previous iteration, then stop.

7) Let \( t = t + dt \) and go to step 1.

To obtain a fair comparison between the proposed and the viscous fluid method, re-gridding was not employed. Re-gridding is essentially a memoryless procedure, as how images are matched after each re-gridding is independent of the deformation before the re-gridding, rendering the comparison of final Jacobian fields and cost functionals problematic. Moreover, the strategy of re-gridding, through the relaxation of deformation over time, is less rigorous from a theoretical standpoint. For that reason, for the results presented in this paper, velocity vector \( \vec{v} \) was obtained from equation (7) rather than by solving the simplification of the momentum conservation equation (6), for both the proposed and Christensen’s models. From our experience, solving equation (6) makes it necessary to employ re-gridding in order to achieve close matching.

6. Results and Discussion

In this section, we tested the proposed nonlinear registration model and compared the results to those obtained with the fluid registration (Christensen’s) model. In order to gain more insight into the effect of the symmetric \( KL \) distance term in (19), we first consider matching two 2D binary images (both 289 by 289, \( \lambda = 1000 \) in (21)), representing midline corpus callosum contours of two control subjects (Figures 1 through 5). As seen in Figure 1, both the
Christensen proposed

Figure 7. Serial MRI example. Results obtained with Christensen’s model and the proposed model. The generated grids are superimposed with the deformed images (row 1) and are shown separately (row 2). Blue, yellow and red contours represent the boundaries of ventricles in $T$, $S$, and deformed $T$, respectively. Note that for both methods, yellow contour is essentially invisible due to a very close match. However, the resulting grid of the proposed method is visually more regular.

Figure 8. Serial MRI example. Jacobian map of the deformation is superimposed with the deformed image for Christensen’s model and the proposed model.

Figure 9. Serial MRI example. Histograms of Jacobian values of the deformations inside ventricles for Christensen’s model and the proposed model.

Figure 10. Serial MRI example. (a) Standard deviation of Jacobian values inside the ventricle per iteration. (b) Symmetric KL distance. For Christensen’s model (dashed blue), both standard deviation and symmetric KL distance increase while for the proposed model (solid red), both standard deviation and symmetric KL distance stabilize.

fluid registration model and the proposed model generated a close match between the deformed image and the study. However, Figures 2 and 3 show grid lines merge and self-cross for Christensen’s model, which is a consequence of negative Jacobian values at certain places, indicating topology change. The proposed method, on the other hand, more evenly distributes deformation inside and outside the corpus callosum (resulting from the convex property of the logarithmic mapping in inequality (17)). The histograms of the Jacobian field inside the corpus callosum are shown in Figure 4 (notice the histogram for the proposed method is noticeably sharper). Figure 5(a) plots the standard deviation of the Jacobian field inside the corpus callosum as a function of iteration number. For Christensen’s model, the standard deviation increased with the number of iterations, since the grid became less regular. On the other hand, the proposed method yielded an optimized standard deviation as more iterations were computed. The proposed symmetric KL distance also increased for Christensen’s method, while it was minimized for the proposed method as shown in Figure 5(b).

In Figures 6 through 10, we show the results of matching a pair of 2D slices from a set of serial MRI images (each of size 226 by 256; $\lambda = 400$ in (21)), where visually significant ventricle enlargement is present. Both Christensen’s method and the proposed model generated a close match between the deformed image and the study (Figure 6(a-d)). Here, there is no reason not to evenly distribute the Jacobian field inside the ventricles, as realized using the proposed method. In contrast, Christensen’s method generated a density map with extreme values along the ventricular bound-
ary. Indeed, given the overall longitudinal ventricular dilatation, we argue that the corresponding density change map should be constant inside the ventricle. As seen in Figure 10, both the standard deviation inside the ventricle and the symmetric $KL$ distance increased for Christensen’s method, while these quantities stabilized for the proposed method.

In the last numerical example (Figures 11 through 15), we tested the proposed model using a set of three 3D Serial MRI volumes obtained from a patient with right-side
semantic dementia (temporal atrophy). In this example, the initial scan was obtained in 02/1993, with the two follow-up scans, in 02/1996 and 08/1999, referred to as time 1, time 2, and time 3, respectively (each volume was downsampled to 112x128x128). A fully three-dimensional computation was employed, with $\lambda = 500$ in the generalization of equation (21) to three spatial dimensions. In Figure 14(a), the 3D Jacobian map generated using the Christensen’s method is visually very noisy with extreme values along the boundaries of the brain as well as in the background, masking the real change over the right temporal area. In contrast, as shown in Figure 14(b), right temporal atrophy (RT) and ventricular enlargement (V) are easily visualized in the Jacobian map generated using the proposed method, demonstrating its theoretical and practical advantages. Future work will investigate whether statistical differences in groups of structural brain images are detected with higher power using the proposed method, which avoids bias in the distribution of the Jacobian in homogeneous brain regions.

Acknowledgements

This work was supported by Grants U54 RR021813 NIH/NCCR, Grant U01 AG024904, and Grants R21 RR019771, EB01651, AG016570, NS049194 to PT.

The authors would like to thank James Becker and Simon Davis at the University of Pittsburgh for providing the MRI dataset.

References

Figure 14. 3D Serial MRI example. Jacobian maps are superimposed with the deformed volumes for Christensen’s model (columns 1 and 2) and the proposed model (columns 3 and 4). Smaller deformations (time 2 to time 3) and larger deformations (time 1 to time 3) are shown. Rows depict slices in axial (rows 1 and 2), sagittal (row 3), and coronal (row 4) planes. Right temporal atrophy (RT) and ventricular enlargement (V) are easily visualized in the Jacobian map generated using the proposed method, while Christensen’s method generated a very noisy map.
Figure 15. 3D Serial MRI example. Symmetric KL distance is shown for Christensen’s and proposed models for a larger deformation (time 1 to time 3) and a smaller deformation (time 2 to time 3). Note that this measure is proportional to the magnitude of deformation. For the proposed method, the symmetric KL distance stabilizes.

Contact Information

Igor Yanovsky is with the Department of Mathematics, UCLA, 405 Hilgard Avenue, Los Angeles, CA 90095-1555, USA
Web: http://www.math.ucla.edu/~yanovsky/

Paul Thompson is with the Laboratory of Neuro Imaging, Department of Neurology, 635 Charles E. Young Drive South, UCLA David Geffen School of Medicine, Los Angeles, CA 90095-7332, USA
Web: http://www.loni.ucla.edu/~thompson/

Stanley Osher is with the Department of Mathematics, UCLA, 405 Hilgard Avenue, Los Angeles, CA 90095-1555, USA
Web: http://www.math.ucla.edu/~sjo/

Alex Leow is with the Neuropsychiatric Institute and the Laboratory of Neuro Imaging, Department of Neurology, 635 Charles E. Young Drive South, UCLA David Geffen School of Medicine, Los Angeles, CA 90095-7332, USA
Web: http://www.loni.ucla.edu/~aleow/