UNIVERSITY OF CALIFORNIA
Los Angeles

# Unbiased Nonlinear Image Registration

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Mathematics

by

**Igor Yanovsky**

2008

The dissertation of Igor Yanovsky is approved.

Luminita Vese

Paul Thompson

Andrea Bertozzi

Stanley Osher, Committee Chair

University of California, Los Angeles

2008

*To my wife Tatiana*

*and to my parents*

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# LIST OF ALGORITHMS

# NOTATIONS AND SYMBOLS

| | |
|---|---|
| $\mathbb{R}$ | the set of real numbers |
| $n$ | spatial dimension |
| $\Omega$ | a bounded domain in $\mathbb{R}^n$ |
| $\partial\Omega$ | the boundary of $\Omega$ |
| $\|\Omega\|$ | the volume of $\Omega$ in $\mathbb{R}^n$ |
| $\mathbf{x}$ | $\mathbf{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ |
| $\partial_i$ | $\partial_i = \partial_{x_i} = \dfrac{\partial}{\partial x_i}$, the partial derivative with respect to $x_i$ |
| $\nabla$ | the gradient operator, $\nabla f = (\partial_1 f, \ldots, \partial_n f)^T$ for functions $f : \mathbb{R}^n \to \mathbb{R}$ |
| $\triangle$ | the Laplace operator, $\triangle f = \sum_{i=1}^{n} \partial_{x_i x_i} f$, |
| | $\triangle \mathbf{u} = (\triangle u_1, \ldots, \triangle u_n)^T$ for vector $\mathbf{u}$ |
| $C^k(\Omega)$ | the space of differentiable functions with continuous |
| | partial derivatives up to order $k$ |
| $C_c^k(\Omega)$ | the space of functions in $C^k(\Omega)$ with compact support |
| $L^p$ | $L^p = L^p(\Omega) = \{f : \Omega \to \mathbb{R} : \int_\Omega \|f(\mathbf{x})\|^2 d\mathbf{x} < \infty\}$ |
| $\langle \cdot, \cdot \rangle$ | scalar product in $L^2(\Omega)$ |
| $\|\| \cdot \|\|_p$ | the norm on $L^p$ |
| $\|\| \cdot \|\|_F$ | the Frobenius norm |
| $\circ$ | the composition operator, $f \circ g = f(g)$ |
| $*$ | the convolution operator |
| div | the divergence operator, $\text{div}\, \mathbf{u} = \nabla \cdot \mathbf{u} = \sum_{i=1}^{n} \partial_{x_i} u_i$ |
| $\mathcal{M}_{n\times n}(\mathbb{R})$ | the set of $n \times n$ matrices with real elements |
| $\det$ | the determinant of a matrix |
| trace | the trace of a matrix $A \in \mathbb{R}^{n\times n}$, $\text{trace}(A) = \sum_{i=1}^{n} a_{ii}$ |

| | |
|---|---|
| $\mathcal{I}$ | the identity matrix |
| $\frac{D}{Dt}$ | the total time derivative |
| $I_1$ | the target image (or study, reference image) |
| $I_2$ | the source image (or template image) |
| **id** | the identity transformation |
| **g** | the transformation $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^n$ |
| $D\mathbf{g}$ | the Jacobian matrix of **g** |
| $\lvert D\mathbf{g}\rvert$ | the Jacobian determinant, or Jacobian, of deformation **g**; $\lvert D\mathbf{g}\rvert := \det(D\mathbf{g})$ |
| $dE(\mathbf{u}, \boldsymbol{\eta})$ | the Gâteaux derivative of functional $E$ at **u** with respect to a perturbation $\boldsymbol{\eta}$ |
| $\partial_{\mathbf{u}} E(\mathbf{u})$ | $\partial_{\mathbf{u}} E(\mathbf{u}) = \dfrac{\partial E(\mathbf{u})}{\partial \mathbf{u}}$ is the gradient of functional $E$, $\partial_{\mathbf{u}} E(\mathbf{u}) = \big(\partial_{u_1} E(\mathbf{u}), \ldots, \partial_{u_n} E(\mathbf{u})\big)^T$, $\partial_{\mathbf{u}} E(\mathbf{u}) = \left(\frac{\partial E(\mathbf{u})}{\partial u_1}, \ldots, \frac{\partial E(\mathbf{u})}{\partial u_n}\right)^T$ |
| **u** | displacement |
| **v** | velocity |
| **f** | force field, or body force |
| $\nu$, $\mu$ | Lame constants describing the properties of an elastic material |
| $\nu_f$, $\mu_f$ | viscosity coefficients of a fluid |
| $\rho$ | fluid density |
| $\mathcal{E}$ | strain (deformation) tensor for elasticity; $\mathcal{E} = [\varepsilon_{ij}]$, $1 \leq i, j \leq n$ |
| $\mathcal{D}$ | rate of deformation tensor for fluid flow; $\mathcal{D} = [d_{ij}]$, $1 \leq i, j \leq n$ |
| $\Sigma$ | Cauchy stress tensor; $\Sigma = [\sigma_{ij}]$, $1 \leq i, j \leq n$ |
| $\sigma$ | standard deviation |

| | |
|---|---|
| ADNI | Alzheimer's Disease Neuroimaging Initiative |
| CDF | cumulative distribution function |
| CFL | Courant-Friedrichs-Lewy (condition) |
| CT | computed tomography |
| FDR | false discovery rate |
| FFT | fast Fourier transformation |
| ICBM | International Consortium for Brain Mapping |
| KL | Kullback-Leibler (divergence) |
| MI | mutual information |
| MRI | magnetic resonance imaging |
| PDE | partial differential equation |
| PDF | probability density function |
| PET | positron emission tomography |
| ROI | region of interest |
| SKL | symmetric Kullback-Leibler (distance) |
| SSD | sum of squared differences |
| TBM | tensor-based morphometry |
| UAV | unmanned aerial vehicle |

# ACKNOWLEDGMENTS

I would like to thank my Ph.D. dissertation advisor, Professor Stanley Osher, for his advice and guidance through my graduate studies. I greatly appreciate his care and help in my life and career development.

I would like to express my gratitude to my committee member Professor Paul Thompson. Many projects I have done in collaboration with the Lab of Neuro Imaging (LONI) were done under his mentorship.

I am thankful to my committee member, Professor Luminita Vese, who I have completed numerous projects with, for her help, encouragement, and valuable suggestions in my research.

I would like to thank my committee member, Professor Andrea Bertozzi, for her valuable suggestions in my studies.

I am also grateful to exceptional Applied Mathematics faculty: Professor Chris Anderson, Professor Achi Brandt, Professor Russell Caflisch, Professor Tony Chan, Professor Bjorn Engquist, Professor Allon Percus, and Professor Eitan Tadmor. Their teachings formed an essential part of my graduate education.

I would like to acknowledge my collaborators in LONI. I am especially thankful to my colleague Alex Leow with whom we had a wonderful collaboration. I also enjoyed working with Professor Ivo Dinov and Professor Arthur Toga.

I also thank the Institute for Pure and Applied Mathematics (IPAM) and its staff for providing me with a wonderful working environment. I am especially grateful to the director of IPAM, Professor Mark Green. It had also been a pleasure working alongside such a wonderful people as Carl Hunt and Eilish Hathaway.

Additionally, I would like to thank my friends, colleagues, classmates, and office-

mates, including Jeremy Brandman, Alexandre Cunha, Bin Dong, Alexander Fridman, Boris Gutman, Lin He, Shao-Ching Huang, Sung Ha Kang, Chiu-Yen Kao, Yunho Kim, Yanina Landa, Triet Le, Tin Man Lee, Carole Le Guyader, Shingyu Leung, Tungyou Lin, Lok Ming Lui, Yu Mao, Chohong Min, Kang-Yu Ni, Neelroop Parikshak, Frederick Park, Berta Sandberg, Richard Tsai, Jinjun Xu, Jian Ye, Kensuke Yokoi, Mingqiang Zhu, Wei Zhu, and many others.

This work was supported in part by the National Institutes of Health under Grant U54 RR021813 NIH/NCRR and CCB-NIH Grant 30886. It was also supported by the National Science Foundation under NSF VIGRE Grant DMS-0601395.

Chapter 2, describing unbiased registration technique and introducing the unbiased fluid registration model, is a combination of [112, 107, 109, 60, 114]. Chapter 3, introducing the unbiased nonlinear elastic registration, is a version of [108]. Chapter 4, which introduces asymmetric unbiased model and performs calibration studies to compare registration models for tensor-based morphometry, is a combination of [110, 111, 114, 113]. Chapter 5, describing the segmentation of deformations, is a version of [115].

# VITA

| | |
|---|---|
| 2002 | B.S., Applied Mathematics, Computing Specialization<br>Summa Cum Laude<br>University of California, Los Angeles |
| 2002 | M.A., Applied Mathematics<br>University of California, Los Angeles |
| 2008 | Ph.D., Applied Mathematics<br>University of California, Los Angeles |

# PUBLICATIONS

Igor Yanovsky, Stanley Osher, Paul Thompson, Alex Leow, *Log-Unbiased Large-Deformation Image Registration*, Computer Vision Theory and Applications, vol. 1, pp. 272-279, 2007.

Igor Yanovsky, Ming-Chang Chiang, Paul Thompson, Andrea Klunder, James Becker, Simon Davis, Arthur Toga, Alex Leow, *Quantifying Deformation Using Information Theory: The Log-Unbiased Nonlinear Registration*, IEEE International Symposium on Biomedical Imaging, pp. 13-16, 2007.

Igor Yanovsky, Paul Thompson, Stanley Osher, Alex Leow, *Topology Preserving Log-Unbiased Nonlinear Image Registration: Theory and Implementation*, IEEE Conference on Computer Vision and Pattern Recognition, pp. 1-8, 2007.

Igor Yanovsky, Paul Thompson, Stanley Osher, Luminita Vese, Alex Leow, *Multiphase Segmentation of Deformation using Logarithmic Priors*, IEEE Computer Society Workshop on Image Registration and Fusion, pp. 1-6, 2007.

Alex Leow, Igor Yanovsky, Ming-Chang Chiang, Agatha Lee, Andrea Klunder, Allen Lu, James Becker, Simon Davis, Arthur Toga, Paul Thompson, *Statistical Properties of Jacobian Maps and the Realization of Unbiased Large-Deformation Nonlinear Image Registration*, IEEE Transactions on Medical Imaging, vol. 26, no. 6, pp. 822-832, 2007.

Igor Yanovsky, Paul Thompson, Andrea Klunder, Arthur Toga, Alex Leow, *Local Volume Change Maps in Nonrigid Registration: When Are Computed Changes Real?*, International Conference on Medical Image Computing and Computer Assisted Intervention, Workshop on Statistical Registration, 2007.

Igor Yanovsky, Paul Thompson, Stanley Osher, Alex Leow, *Multimodal Unbiased Image Matching via Mutual Information*, Computational Imaging VI, IS&T/SPIE Electronic Imaging, vol. 6814, 2008.

Igor Yanovsky, Paul Thompson, Stanley Osher, Xue Hua, David Shattuck, Arthur Toga, Alex Leow, *Validating Unbiased Registration on Longitudinal MRI Scans from the Alzheimer's Disease Neuroimaging Initiative (ADNI)*, IEEE International Symposium on Biomedical Imaging, pp. 1091-1094, 2008.

Igor Yanovsky, Paul Thompson, Stanley Osher, Alex Leow, *Asymmetric and Symmetric Unbiased Image Registration: Statistical Assessment of Performance*, IEEE

Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis, 2008.

ABSTRACT OF THE DISSERTATION

# Unbiased Nonlinear Image Registration

by

**Igor Yanovsky**

Doctor of Philosophy in Mathematics

University of California, Los Angeles, 2008

Professor Stanley Osher, Chair

Image registration is the process of estimating an optimal transformation between the two images. In particular, nonlinear image registration has been developing rapidly in the last decade. Different regularization techniques have been proposed, sometimes generating deformations very different from one another. In this dissertation, we present a novel unbiased nonlinear image registration technique. The unbiased framework generates theoretically and intuitively correct deformation maps, and is compatible with large-deformation models. 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 large-deformation viscous fluid registration model to computation of unbiased deformations. We show that unbiased fluid registration method generates more accurate maps compared to those generated using the viscous fluid registration model.

We also propose a large-deformation image registration model based on nonlinear elastic regularization and unbiased registration. The new model is written in a unified variational form and is minimized using gradient descent on the corresponding Euler-Lagrange equations. The unbiased nonlinear elastic registration model is com-

putationally efficient and easy to implement.

Furthermore, we examine the reproducibility and the power to detect real changes of different computational techniques. It is the first work to systematically investigate the reproducibility and variability of different registration methods in tensor-based morphometry. Specifically, we compare different matching functionals, as well as large deformation registration schemes using serial magnetic resonance imaging scans. Our results show that the unbiased methods have higher reproducibility than the conventional registration models. The unbiased methods are less likely to produce changes in the absence of any real physiological change. Moreover, they are also better in detecting biological deformations by penalizing any bias in the corresponding statistical maps.

Finally, we extend the idea of the unbiased registration to simultaneously registering and tracking deforming objects in a sequence of two or more images. A level set based Chan-Vese multiphase segmentation model is generalized to consider Jacobian fields while segmenting regions of growth and shrinkage in deformations. Deforming objects are thus classified based on magnitude of homogeneous deformation.

Numerical results are presented in each chapter.

# CHAPTER 1

# Introduction

Image registration, also known as image matching, morphing, and alignment, is the process of determining a spatial transformation, or mapping, that relates positions in one image to corresponding positions in one or more images. It can also be seen as the process of overlaying two or more images of the same scene taken at different times, from different perspectives, and by different sensors. Registration is necessary when comparing or integrating the data obtained from different measurements.

Image registration is generally classified as being rigid, affine, or nonrigid. In rigid registration, objects in images are assumed to be rotated and translated with respect to one another. Rigid body transformations preserve the distance between all points in the image and are equivalent to a change from one coordinate system to another one which differs by shift and rotation. Affine transformations, which allow for a global change of scale and shear, map parallel lines to parallel lines. Rigid and affine transformations are often referred to as linear transformations and can be conveniently represented in matrix form.

In many applications, correspondence between structures in two images cannot be achieved without some localized stretching of the images. Nonrigid transformations, which map straight lines to curves, allow local warping of image features, thus providing support for local deformations. Nonrigid registration methods can be generally subdivided into physical models and function representation based techniques. Physical models are derived from the laws of continuum mechanics and include methods

based on theories of elasticity and viscous fluid flow. Function representation based models, which originate from interpolation and approximation theory, use basis function expansions to model the deformation. Examples of basis functions include radial basis functions, $B$-splines, and wavelets [26, 50].

There are a number of reviews of rigid registration methods [48, 68, 92]. Reviews on nonrigid registration methods [26, 42, 41, 50, 63, 73, 88, 117] focus on elastic and fluid models, function expansions and splines, $B$-splines and wavelets, radial basis functions, landmark based methods, level sets registration, as well as numerical solutions to registration problems. In the next section, we list some applications of image registration.

## 1.1 Applications of Image Registration

Image registration is an important step in all image analysis tasks in which the final information is acquired from the combination of various data sources. Virtually all intelligent vision processing and understanding systems require image matching or closely related operations. Many applications of image registration are found in areas of computer vision, medical imaging, surveillance, remote sensing, environmental monitoring, and mapping.

Image registration is an important element in remote sensing. Some of the applications in this area include multispectral classification, change detection, validation and precision correction of satellite navigation data, validation of sensors, image mosaicing, weather forecasting, creating super-resolution images, integrating information into geographic information systems, environmental monitoring, as well as land use analysis and planning [33, 117].

Registering series of images acquired by an airborne sensor into a mosaic has been

developing recently. Geo-registration is becoming increasingly used in unmanned aerial vehicles (UAV) to navigate, or to geo-locate targets [65]. The fusion of multimodal video sequences has also become important for surveillance purposes, navigation and object tracking applications, as well as multichannel restoration [100].

Image matching has also been widely used in computer vision. Some of such vision applications are target localization, automatic quality control, automated image registration, object recognition, 3D scene reconstruction, vision-based autonomous navigation, image database retrieval, and signature recognition [100, 117].

In recent years, medical image registration has evolved as interdisciplinary field with many applications in functional and anatomic brain mapping, image-guided surgery, and multimodality image fusion [1, 44, 99]. Intrasubject registration is performed to align scans of the same subject at different times. When applied to serial scans of human brain, image registration offers tremendous power in detecting the earliest signs of illness, understanding normal brain development or aging, and monitoring disease progression. On the other hand, intersubject registration aligns scans of different subjects. In multi-subject studies, image matching reduces subject-specific anatomic differences by deforming individual images onto a population average brain template. Applications of image registration also include combining images of the same subject from different modalities, providing guidance during surgical procedures, among many others.

## 1.2 Registration Problem

Let $\Omega$ be an open and bounded domain in $\mathbb{R}^n$, for arbitrary $n$. Let $I_1 : \Omega \to \mathbb{R}$ and $I_2 : \Omega \to \mathbb{R}$ be the two images to be registered. The goal of image registration is to find the transformation $\mathbf{g} : \Omega \to \Omega$ that maps the source image $I_2$ into correspondence

with the target image $I_1$. The displacement field $\mathbf{u}(\mathbf{x})$ from the position $\mathbf{x}$ in the deformed image $I_2 \circ \mathbf{g}(\mathbf{x})$ back to $I_2(\mathbf{x})$ is defined in terms of the deformation $\mathbf{g}(\mathbf{x})$ by the expression $\mathbf{g}(\mathbf{x}) = \mathbf{x} - \mathbf{u}(\mathbf{x})$ at every point $\mathbf{x} \in \Omega$. 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. Thus, the problems of finding deformation $\mathbf{g}$ and displacement $\mathbf{u}$ are equivalent.

In general, nonlinear image registration models can be formulated in a variational framework. An appropriate similarity measure indicates how accurately the deformed image is aligned with the target image. To construct a deformation that is one-to-one and differentiable [21, 51, 70], a regularizing constraint must be imposed. Thus, the problem of image registration is often cast as a minimization problem with a combined cost functional $E$ consisting of an image matching functional $F$ and a regularizing constraint $R$:

$$\inf_{\mathbf{u}} \Big\{ E(\mathbf{u}) = F(\mathbf{u}) + \lambda R(\mathbf{u}) \Big\}. \tag{1.1}$$

Here, $\lambda > 0$ is a weighting parameter. In general, we expect minimizers of the energy functional $E(\mathbf{u})$ to exist. A necessary condition for a minimizer $\mathbf{u}$ is that the Gâteaux derivative $dE(\mathbf{u}, \boldsymbol{\eta})$ of $E$ vanishes for all suitable perturbations $\boldsymbol{\eta}$. This derivative is also known as the first variation of $E$ in the direction of $\boldsymbol{\eta}$:

$$dE(\mathbf{u}, \boldsymbol{\eta}) = \lim_{\epsilon \to 0} \frac{E(\mathbf{u} + \epsilon\boldsymbol{\eta}) - E(\mathbf{u})}{\epsilon} = \left. \frac{\partial E(\mathbf{u} + \epsilon\boldsymbol{\eta})}{\partial \epsilon} \right|_{\epsilon=0}. \tag{1.2}$$

It can be shown that if $E$ has a local extremum $\mathbf{u}$, its first variation at $\mathbf{u}$ must vanish:

$$dE(\mathbf{u}, \boldsymbol{\eta}) = 0. \tag{1.3}$$

Denote by $\langle \cdot, \cdot \rangle$ the scalar product in $L^2(\Omega)$. If $dE(\mathbf{u}, \boldsymbol{\eta}) = \langle \partial_{\mathbf{u}} E, \boldsymbol{\eta} \rangle$, then $\partial_{\mathbf{u}} E$ defines

the gradient of the functional $E$. Therefore, the necessary condition for a minimizer $\mathbf{u}$ becomes

$$\partial_{\mathbf{u}} E(\mathbf{u}) = 0. \tag{1.4}$$

Equation (1.4) is known as the Euler-Lagrange equation. Note that $\partial_{\mathbf{u}} E(\mathbf{u}) = \big(\partial_{u_1} E(\mathbf{u}), \ldots, \partial_{u_n} E(\mathbf{u})\big)^T$. Also, from equation (1.1), the gradient of $E$ can be written as

$$\partial_{\mathbf{u}} E(\mathbf{u}) = \partial_{\mathbf{u}} F(\mathbf{u}) + \lambda \partial_{\mathbf{u}} R(\mathbf{u}). \tag{1.5}$$

Energy functional (1.1) can be minimized using the gradient descent of the corresponding Euler-Lagrange equation to obtain

$$\frac{\partial \mathbf{u}(\mathbf{x}, t)}{\partial t} = -\partial_{\mathbf{u}} E(\mathbf{u}), \tag{1.6}$$

where $t$ is an artificial time.

Common choices of image matching functional $F$ include landmark matching, sum of squared intensity differences, cross correlation [24], and mutual information or other divergence-based or information-theoretic measures [28, 46, 86], while choices of regularization $R$ usually involve differential operators inspired by thin-plate spline theory, elasticity theory, fluid dynamics and the Euler-Poincaré equations [70, 98]. This chapter reviews some of the similarity measures and regularization constraints.

## 1.3 Registration Metrics

### 1.3.1 Landmark-Based Registration

A registration metric takes two images as input and returns a real value that indicates how well the images are aligned. One of the simplest ones is based on the distance

between corresponding pairs of landmarks that are extracted from images. Landmarks are often used for rigid-body registration [9, 68, 92]. For nonrigid registration, landmarks are often used with thin-plate splines [50].

The landmark distance measure is defined using the Euclidean norm $||\cdot||$ as

$$F_{LM}(\mathbf{u}) = \sum_{k=1}^{m} \left|\left|\mathbf{x}^{I_1,k} - \mathbf{g}(\mathbf{x}^{I_2,k})\right|\right|^2 = \sum_{k=1}^{m} \left|\left|\mathbf{x}^{I_1,k} - \left(\mathbf{x}^{I_2,k} - \mathbf{u}(\mathbf{x}^{I_2,k})\right)\right|\right|^2,$$

where $\mathbf{x}^{I_1,k}$ and $\mathbf{x}^{I_2,k}$, $1 \leq k \leq m$, are landmark locations in the target and source images, respectively.

An advantage of landmarks is that they enable the transformation to be determined in closed form. However, a large number of landmarks are needed to densely sample the deformation field. Also, the localization process introduces error.

### 1.3.2 $L^2$ distance

For intramodality registration, the most common way to define the distance between the deformed source and the target images is to use the $L^2$ norm, or the sum of squared differences (SSD). The $L^2$-norm matching functional is suitable when the images have been acquired through similar sensors and thus are expected to present the same intensity range and distribution. The $L^2$ distance between the deformed image $I_2(\mathbf{x} - \mathbf{u})$ and target image $I_1(\mathbf{x})$ is defined as

$$\begin{aligned} F_{L^2}(I_1, I_2, \mathbf{u}) &= \frac{1}{2}\left|\left|I_2(\mathbf{x} - \mathbf{u}) - I_1(\mathbf{x})\right|\right|_{L^2(\Omega)} \\ &= \frac{1}{2}\int_{\Omega} \left(I_2(\mathbf{x} - \mathbf{u}) - I_1(\mathbf{x})\right)^2 d\mathbf{x}. \end{aligned} \tag{1.7}$$

Computing the first variation of the $L^2$ similarity functional gives

$$\mathbf{f}(\mathbf{x}, \mathbf{u}(\mathbf{x})) = -\partial_{\mathbf{u}} F_{L^2}(I_1, I_2, \mathbf{u}) = \left[I_2(\mathbf{x} - \mathbf{u}(\mathbf{x})) - I_1(\mathbf{x})\right] \nabla I_2|_{\mathbf{x}-\mathbf{u}}, \tag{1.8}$$

where **f** is the force field, or the body force, which drives the source into registration with the target. The first term in the definition of **f**, namely $I_2(\mathbf{x}-\mathbf{u}) - I_1(\mathbf{x})$, is the difference in intensity between the deformed image and the target image. This term causes the field force to tend to zero in areas where the deformed source image is locally aligned with the target image. The second term $\nabla I_2|_{\mathbf{x}-\mathbf{u}}$ is the gradient of the deformed source image and has largest values at the edges of the source image. This term determines the directions of the local deformation forces applied to the source.

To use this similarity function, the images $I_1$ and $I_2$ must correspond to the same imaging modality and they may require preprocessing to equalize the intensities of the image.

### 1.3.3 Mutual Information

A concept borrowed from the information theory, mutual information measure has been successfully adopted for single-modal and multimodal image registration. The use of mutual information for image registration was first introduced in [23, 103]. One of the main advantages of using mutual information is that it can be used to align images of different modalities, without requiring knowledge of the relationship (joint intensity distribution) of the two registered images. We refer the readers to [28, 47, 105] for relevant discussions on mutual information.

To define the mutual information between the deformed image $I_2(\mathbf{x}-\mathbf{u})$ and the target image $I_1(\mathbf{x})$, we follow the notations in [47], where $p^{I_1}$ and $p^{I_2}_{\mathbf{u}}$ are used to denote the intensity distributions estimated from $I_1(\mathbf{x})$ and $I_2(\mathbf{x}-\mathbf{u})$, respectively. An estimate of their joint intensity distribution is denoted as $p^{I_1,I_2}_{\mathbf{u}}$. In this probabilistic framework, the link between two modalities is fully characterized by a joint density.

Let $i_1 = I_1(\mathbf{x})$, $i_2 = I_2(\mathbf{x}-\mathbf{u}(\mathbf{x}))$ denote intensity values at point $\mathbf{x} \in \Omega$. Given the

displacement field $\mathbf{u}$, the mutual information computed from $I_1$ and $I_2$ is provided by

$$MI_{\mathbf{u}}^{I_1,I_2} = \int_{\mathbb{R}^2} p_{\mathbf{u}}^{I_1,I_2}(i_1,i_2)\ \log \frac{p_{\mathbf{u}}^{I_1,I_2}(i_1,i_2)}{p^{I_1}(i_1)p_{\mathbf{u}}^{I_2}(i_2)}\, di_1 di_2. \tag{1.9}$$

To maximize the mutual information between $I_2(\mathbf{x}-\mathbf{u})$ and $I_1(\mathbf{x})$, the following functional is minimized:

$$F_{MI}(I_1, I_2, \mathbf{u}) = -MI_{\mathbf{u}}^{I_1,I_2}. \tag{1.10}$$

The gradient of (1.10) is derived to be

$$\partial_{\mathbf{u}} F_{MI}(\mathbf{u}) = \frac{1}{|\Omega|}\left[Q_{\mathbf{u}} * \frac{\partial \psi}{\partial \xi_2}\right](I_1(\mathbf{x}), I_2(\mathbf{x}-\mathbf{u}))\nabla I_2(\mathbf{x}-\mathbf{u}), \tag{1.11}$$

where $|\Omega|$ is the volume of $\Omega$, $Q_{\mathbf{u}}$ is defined as

$$Q_{\mathbf{u}}(i_1,i_2) = 1 + \log \frac{p_{\mathbf{u}}^{I_1,I_2}(i_1,i_2)}{p^{I_1}(i_1)p_{\mathbf{u}}^{I_2}(i_2)}, \tag{1.12}$$

and $\psi(\xi_1,\xi_2)$ is a two-dimensional Parzen windowing kernel [31, 83], which is used to estimate the joint intensity distribution from $I_2(\mathbf{x}-\mathbf{u})$ and $I_1(\mathbf{x})$. The Gaussian kernel with variance $\sigma^2$ is often used as a kernel

$$\psi(\xi_1,\xi_2) = G_\sigma(\xi_1,\xi_2) = \frac{1}{2\pi\sigma^2} e^{\frac{-(\xi_1^2+\xi_2^2)}{2\sigma^2}}. \tag{1.13}$$

In the case of the mutual information similarity measure, the force field, which drives the source image into registration with the target image, is given by

$$\mathbf{f}(\mathbf{x}, \mathbf{u}(\mathbf{x})) = -\partial_{\mathbf{u}} F_{MI}(\mathbf{u}). \tag{1.14}$$

## 1.4 Regularizations

In this section, we will review different approaches to regularizing the displacement field **u**, which include diffusion, biharmonic, linear elastic, and viscous fluid regularizations.

### 1.4.1 Diffusion Registration

Diffusion registration in $\mathbb{R}^n$ [34, 73] is based on the regularizer

$$R_{DIF}(\mathbf{u}) = \frac{1}{2}\sum_{k=1}^{n}\int_{\Omega} ||\nabla u_k||^2 d\mathbf{x}. \tag{1.15}$$

Taking the first variation of $R_{DIF}$, gives the following gradient

$$\partial_{\mathbf{u}} R_{DIF}(\mathbf{u}) = -\triangle\mathbf{u}. \tag{1.16}$$

Minimizing the functional (1.1), which consists of the diffusion regularizer (1.15) and the appropriate fidelity term, gives the Euler-Lagrange equation

$$\lambda\triangle\mathbf{u} + \mathbf{f} = 0, \tag{1.17}$$

where the expression of the force field **f** depends on the similarity measure.

Image registration based on diffusion regularization is easy to implement numerically. One of the main advantages of this model is its dependance on only one parameter. However, the transformation obtained with diffusion registration is restricted to small deformations.

### 1.4.2 Biharmonic Registration

Biharmonic registration, which is also known as curvature registration [35, 36, 73], is based on the regularizer

$$R_{BH}(\mathbf{u}) = \frac{1}{2}\sum_{k=1}^{n}\int_{\Omega}(\triangle u_k)^2 d\mathbf{x}. \tag{1.18}$$

The gradient of $R_{BH}$ is

$$\partial_{\mathbf{u}} R_{BH}(\mathbf{u}) = \triangle^2 \mathbf{u}. \tag{1.19}$$

Thus, the minimization of the biharmonic regularizer (1.18), coupled with the fidelity term, gives the following Euler-Lagrange equation

$$\lambda \triangle^2 \mathbf{u} - \mathbf{f} = 0. \tag{1.20}$$

Similar to diffusion-based registration, image matching involving biharmonic regularizer, which also depends only on one parameter, allows only small deformations.

### 1.4.3 Physical Continuum Models

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 source into registration with the target. Sections 1.4.4 and 1.4.5 briefly describe two of the most well known such models.

### 1.4.4 Linear Elastic Registration

The theory of linear elasticity is based on the notions of stress and strain. Stress is a measure of the internal force intensity developed within a body in response to external forces. It can be analyzed mathematically using the Cauchy stress tensor, which is a second rank tensor denoted by $\Sigma = [\sigma_{ij}] \in \mathbb{R}^{n\times n}$, $1 \leq i, j \leq n$, where $i$ and $j$ represent the three Cartesian directions. Stress components are either normal to the plane $\sigma_{ii}$ or within it $\sigma_{ij}$, $i \neq j$.

Strain is defined as the amount of deformation an object experiences compared to its original size and shape. The strain tensor, $\mathcal{E} = [\varepsilon_{ij}] \in \mathbb{R}^{n\times n}$ is a symmetric tensor used to quantify the strain of an object undergoing a small deformation. The diagonal coefficients $\varepsilon_{ii}$ are normal components, whereas $\varepsilon_{ij}$, $i \neq j$, are shear strains. The linear strain is defined as

$$\varepsilon_{ij}(\mathbf{u}) = \frac{1}{2}(\partial_j u_i + \partial_i u_j), \tag{1.21}$$

with the strain tensor matrix given by

$$\mathcal{E}(\mathbf{u}) = \frac{1}{2}(\nabla \mathbf{u}^t + \nabla \mathbf{u}). \tag{1.22}$$

Consider the static equilibrium of a linear elastic solid subjected to the body force vector field $\mathbf{f} = (f_1, f_2, f_3)$ in three spatial dimensions. Applying Newton's first law of motion results in the following set of differential equations which govern the stress

distribution within the solid

$$\begin{aligned}
\frac{\partial \sigma_{11}}{\partial x_1} + \frac{\partial \sigma_{12}}{\partial x_2} + \frac{\partial \sigma_{13}}{\partial x_3} + f_1 = 0, \\
\frac{\partial \sigma_{21}}{\partial x_1} + \frac{\partial \sigma_{22}}{\partial x_2} + \frac{\partial \sigma_{23}}{\partial x_3} + f_2 = 0, \\
\frac{\partial \sigma_{31}}{\partial x_1} + \frac{\partial \sigma_{32}}{\partial x_2} + \frac{\partial \sigma_{33}}{\partial x_3} + f_3 = 0.
\end{aligned} \tag{1.23}$$

For a linear elastic solid, the Cauchy stress tensor $\Sigma$ is related to the strain tensor $\mathcal{E}$ by Hooke's law

$$\Sigma = \nu(\text{trace}\,\mathcal{E})I + 2\mu\mathcal{E}, \tag{1.24}$$

where $\nu$ and $\mu$ are Lamé elastic material constants and $I$ is the identity tensor. Equivalently, the components of the stress and strain tensors are related by

$$\sigma_{ij} = \nu(\nabla \cdot \mathbf{u})\delta_{ij} + 2\mu\varepsilon_{ij}, \tag{1.25}$$

where

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases} \tag{1.26}$$

Substituting (1.21) into (1.25) and using equilibrium equations (1.23) gives the Navier-Cauchy linear elastic partial differential equation

$$\mu\triangle\mathbf{u} + (\mu + \nu)\nabla(\nabla \cdot \mathbf{u}) + \mathbf{f}(\mathbf{x}) = 0. \tag{1.27}$$

The regularization for linear elasticity could also be written in a variational framework with

$$R_{LE}(\mathbf{u}) = \frac{1}{2}\int_\Omega \left(\nu(\nabla \cdot \mathbf{u})^2 + 2\mu \sum_{i,j=1}^{n} (\varepsilon_{ij}(\mathbf{u}))^2\right) d\mathbf{x}.$$

Here, $R_{LE}$ is the linearized elastic potential. Calculating the first variation of this functional gives the following gradient

$$\partial_{\mathbf{u}} R_{LE}(\mathbf{u}) = -\mu \triangle \mathbf{u} - (\mu + \nu) \nabla (\nabla \cdot \mathbf{u}).$$

A nonlinear elastic model for nonrigid image registration was first proposed by Broit in [8]. The model was successfully used in [2, 8, 29]. However, since the Navier-Cauchy PDE (1.27) is derived assuming small angles of rotation and small linear deformations, it is only accurate for small deformations. The elastic registration model severely penalizes large displacements and is not useful in applications when large nonlinear deformations are natural.

### 1.4.5 Viscous Fluid Registration

A major shortcoming of the linear elastic approach using Navier-Cauchy equations (1.27) is that it is based on the assumption of an infinitesimally small deformation. Large deformations can not be accommodated with these linear partial differential equations. The limitations of the linear elasticity model can be overcome by a viscous fluid which allows the restoring forces to relax over time.

In the viscous fluid model, first proposed by Christensen *et al.* in [21], an Eulerian reference frame is used in describing large deformations. The Eulerian frame of reference specifies the time evolution of particle positions and velocities as observed at fixed points. Consequently, a particle located at $\mathbf{x}$ at time $t$ originated at position $\mathbf{g}(\mathbf{x}, t) = \mathbf{x} - \mathbf{u}(\mathbf{x}, t)$ at time $t_0$ ($t > t_0$), where $\mathbf{u}$ is the displacement. We let $\mathbf{v}$ denote the velocity field. The material derivative, defined by $D/Dt = \partial/\partial t + \mathbf{v} \cdot \nabla$, describes the time rate of change experienced by an element of material instantaneously at point $\mathbf{x}$ at time $t$. Hence, the Eulerian velocity field $\mathbf{v}$ is nonlinearly related to $\mathbf{u}$ and

is determined by

$$\mathbf{v} = \frac{D\mathbf{u}}{Dt} = \frac{\partial \mathbf{u}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{u}. \tag{1.28}$$

The term $\mathbf{v} \cdot \nabla \mathbf{u}$ accounts for the kinematic nonlinearities of the displacement field $\mathbf{u}$. Note that the material derivative with respect to time $t$ and partial derivative with respect to time $t$ are approximately equal for small deformations.

The laws of continuum mechanics provide the theoretical foundation for fluid flow. The rate of deformation tensor $\mathcal{D} = [d_{ij}]$, $1 \leq i, j \leq n$, is defined as

$$d_{ij} = \frac{1}{2}(\partial_j v_i + \partial_i v_j),$$

or, in vector notation,

$$\mathcal{D} = \frac{1}{2}(\nabla \mathbf{v}^t + \nabla \mathbf{v}). \tag{1.29}$$

A Navier-Poisson Newtonian fluid model is used, for which the Cauchy stress tensor $\Sigma$ is related to the rate of deformation tensor $\mathcal{D}$ by

$$\Sigma = -pI + \nu_f(\text{trace}\,\mathcal{D})I + 2\mu_f \mathcal{D}, \tag{1.30}$$

where $\nu_f$ and $\mu_f$ are the viscosity coefficients of the fluid, and $p$ is the pressure. Equivalently,

$$\sigma_{ij} = -p\delta_{ij} + \nu_f(\nabla \cdot \mathbf{v})\delta_{ij} + 2\mu_f d_{ij}, \tag{1.31}$$

In continuum mechanics applications, it is generally assumed that the mass is conserved within a control volume. However, for image deformations it is often desirable to allow local increase or reduction of mass. Assuming a mass is changed at a rate of $\eta$ per unit volume, the differential form of the mass conservation law is given as

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = \eta,$$

where $\rho$ denotes the density. The differential form of conservation of linear momentum with a mass source is given as

$$\nabla \cdot \Sigma = -\mathbf{f} + \rho \frac{d\mathbf{v}}{dt} + \eta \mathbf{v}, \tag{1.32}$$

where $\mathbf{f}$ is the force field, or the body force, and $\nabla \cdot \Sigma$ is a $n \times 1$ column vector with components $[\nabla \cdot \Sigma]_i = \sum_{j=1}^{n} (\partial \sigma_{ij} / \partial x_j)$. Substituting (1.30) into (1.32) and using (1.29) allows us to derive the equation

$$\rho \frac{d\mathbf{v}}{dt} = \mathbf{f} - \nabla p + (\nu + \mu) \nabla (\nabla \cdot \mathbf{v}) + \mu \triangle \mathbf{v} - \eta \mathbf{v}. \tag{1.33}$$

For very low Reynold's number flow, it is possible to neglect the inertial terms $\rho d\mathbf{v}/dt$ and $\eta \mathbf{v}$. Assuming there is only a small spatial variation of pressure, $\nabla p$ can also be neglected, and (1.33) simplifies to the Navier-Stokes equation for a compressible viscous fluid

$$\mu \triangle \mathbf{v} + (\mu + \nu) \nabla (\nabla \cdot \mathbf{v}) + \mathbf{f}(\mathbf{x}, \mathbf{u}) = 0. \tag{1.34}$$

Equation (1.34) describes the balance of forces acting in a given region of the fluid. The $\triangle \mathbf{v}$ term is the viscosity, which constraints the velocity field to vary smoothly. The term $\nabla(\nabla \cdot \mathbf{v})$ allows structures in the source image to change in mass. The Navier-Stokes equation of fluid flow (1.34) is identical to the Navier-Cauchy equation of linear elasticity (1.27) except that the Navier-Stokes PDE operates on velocity $\mathbf{v}$ rather than displacement $\mathbf{u}$.

Since equation (1.34) is computationally expensive to solve in practice, the authors in [28] proposed to obtain the instantaneous velocity from the convolution of $\mathbf{f}$ with Gaussian kernel $G_\sigma$ of variance $\sigma^2$:

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

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.

## 1.5 Optical Flow Methods

Optical flow methods are used to find small deformations in temporal sequences of images [52, 96]. The basic hypothesis of optical flow is to assume that the intensity of a moving object is constant with time. That is, objects in the image at time $t = t_0$ will generally still be in the image at time $t > t_0$, but will be displaced. This is represented by

$$I(\mathbf{x}, t) = I(\mathbf{x} + \delta\mathbf{x}, t + \delta t). \tag{1.35}$$

Expanding (1.35) in Taylor's series and retaining the first order terms, we obtain the optical flow equation

$$I_2(\mathbf{x}) - I_1(\mathbf{x}) = \mathbf{v} \cdot \nabla I_1(\mathbf{x}). \tag{1.36}$$

Here, $\mathbf{v}$ is the velocity. Numerically more stable expression for $\mathbf{v}$ as in [96], which also holds for small displacements, is given by

$$\mathbf{v}(\mathbf{x}) = \frac{\big(I_2(\mathbf{x}) - I_1(\mathbf{x})\big)\nabla I_1(\mathbf{x})}{(\nabla I_1(\mathbf{x}))^2 + \big(I_2(\mathbf{x}) - I_1(\mathbf{x})\big)^2}. \tag{1.37}$$

The deformation field, given by equation (1.37), is regularized.

In optical flow, $\mathbf{v}$ is considered to be a velocity because the images are two successive time frames. That is, $\mathbf{v}$ is the displacement during the time interval between the two image frames. When comparing two different images, however, there is no such temporal consideration. Hence, $\mathbf{v}$ can be thought of representing a displacement.

## 1.6 Deformation Constraints

In addition to regularization techniques described in Section 1.4, there are methods that constrain the transformation according to some desirable mathematical property [19, 50].

### 1.6.1 Inverse Consistency

The problem with many image registration techniques is that they fail to generate inverse consistent deformations. Inverse consistency can be explained by considering the transformations obtained by registering image $I_2$ to $I_1$ as $\mathbf{g}$ and the inverse one from $I_1$ to $I_2$ as $\mathbf{h}$. If the transformations $\mathbf{g}$ and $\mathbf{h}$ are consistent their composition is the identity. For almost all registration algorithms $\mathbf{g} \neq \mathbf{h}^{-1}$.

Christensen and Johnson proposed a novel inverse consistent image registration approach in [20, 54]. The authors proposed to jointly estimate the transformations $\mathbf{g} : I_2 \rightarrow I_1$ and $\mathbf{h} : I_1 \rightarrow I_2$ by defining a cost function to measure the differences between the deformed image $I_2 \circ \mathbf{g}$ and image $I_1$ and the differences between the deformed image $I_1 \circ \mathbf{h}$ and image $I_2$. In [20, 54], the $L^2$ distance is used as a similarity measure $C_{SIM}$. The transformations $\mathbf{g}$ and $\mathbf{h}$ are estimated by minimizing a symmetrized similarity function

$$\begin{aligned} C_{SIM}(I_1, I_2 \circ \mathbf{g}) + C_{SIM}(I_1 \circ \mathbf{h}, I_2) &= \int_\Omega \left| I_2(\mathbf{g}(\mathbf{x})) - I_1(\mathbf{x}) \right|^2 d\mathbf{x} \\ &+ \int_\Omega \left| I_2(\mathbf{x}) - I_1(\mathbf{h}(\mathbf{x})) \right|^2 d\mathbf{x}. \end{aligned}$$

In order to guarantee that $\mathbf{g} = \mathbf{h}^{-1}$, an inverse consistency constraint $C_{ICC}$, written

in a symmetric form, is imposed:

$$C_{ICC}(\mathbf{u},\mathbf{w}) + C_{ICC}(\mathbf{w},\mathbf{u}) = \int_\Omega \big|\big|\mathbf{g}(\mathbf{x}) - \mathbf{h}^{-1}(\mathbf{x})\big|\big|^2 d\mathbf{x} + \int_\Omega \big|\big|\mathbf{h}(\mathbf{x}) - \mathbf{g}^{-1}(\mathbf{x})\big|\big|^2 d\mathbf{x},$$

where $\mathbf{g}(\mathbf{x}) = \mathbf{x} - \mathbf{u}(\mathbf{x})$ and $\mathbf{h}(\mathbf{x}) = \mathbf{x} - \mathbf{w}(\mathbf{x})$. Here, $\mathbf{u}$ and $\mathbf{w}$ are displacement fields, which define the transformation in terms of a displacement from a location $\mathbf{x}$.

To ensure the transformations $\mathbf{g}$ and $\mathbf{h}$ are smooth, continuum mechanical models such as linear elasticity and viscous fluid can be used to regularize transformations. In [20], the authors used a linear elastic constraint of the form

$$C_{REG}(\mathbf{u}) + C_{REG}(\mathbf{w}) = \int_\Omega \big|\big|L\mathbf{u}(\mathbf{x})\big|\big|^2 d\mathbf{x} + \int_\Omega \big|\big|L\mathbf{w}(\mathbf{x})\big|\big|^2 d\mathbf{x},$$

where the linear elasticity operator $L$, as defined by the Navier-Cauchy PDE (1.27), is

$$L\mathbf{u} = -\alpha\triangle\mathbf{u} - \beta\nabla(\nabla\cdot\mathbf{u}).$$

The total cost function $C$ is hence a linear combination of symmetric quantities attributed to image similarity $C_{SIM}$, the consistency of the forward and backward transformations $C_{ICC}$, and a regularization term that is related to the energy of the deformation $C_{REG}$:

$$\begin{aligned} C(\mathbf{u},\mathbf{w}) &= C_{SIM}(I_1, I_2\circ\mathbf{g}) + C_{SIM}(I_1\circ\mathbf{h}, I_2) \\ &\quad + \rho\big(C_{ICC}(\mathbf{u},\mathbf{w}) + C_{ICC}(\mathbf{w},\mathbf{u})\big) \\ &\quad + \lambda\big(C_{REG}(\mathbf{u}) + C_{REG}(\mathbf{w})\big) \\ &= \int_\Omega \big|I_2(\mathbf{g}(\mathbf{x})) - I_1(\mathbf{x})\big|^2 d\mathbf{x} + \int_\Omega \big|I_2(\mathbf{x}) - I_1(\mathbf{h}(\mathbf{x}))\big|^2 d\mathbf{x} \\ &\quad + \rho\bigg(\int_\Omega \big|\big|\mathbf{g}(\mathbf{x}) - \mathbf{h}^{-1}(\mathbf{x})\big|\big|^2 d\mathbf{x} + \int_\Omega \big|\big|\mathbf{h}(\mathbf{x}) - \mathbf{g}^{-1}(\mathbf{x})\big|\big|^2 d\mathbf{x}\bigg) \\ &\quad + \lambda\bigg(\int_\Omega \big|\big|L\mathbf{u}(\mathbf{x})\big|\big|^2 d\mathbf{x} + \int_\Omega \big|\big|L\mathbf{w}(\mathbf{x})\big|\big|^2 d\mathbf{x}\bigg), \end{aligned} \tag{1.38}$$

where $\rho$ and $\lambda$ are positive constants.

To solve (1.38) numerically, the authors solved for $\mathbf{g}$ and $\mathbf{h}$ separately, minimizing

$$\begin{aligned} C_{\mathbf{g}} &= \int_{\Omega} \left|I_2(\mathbf{g}(\mathbf{x})) - I_1(\mathbf{x})\right|^2 d\mathbf{x} + \rho \int_{\Omega} \left|\left|\mathbf{g}(\mathbf{x}) - \mathbf{h}^{-1}(\mathbf{x})\right|\right|^2 d\mathbf{x} \\ &\quad + \lambda \int_{\Omega} \left|\left|L\mathbf{u}(\mathbf{x})\right|\right|^2 d\mathbf{x}, \\ C_{\mathbf{h}} &= \int_{\Omega} \left|I_2(\mathbf{x}) - I_1(\mathbf{h}(\mathbf{x}))\right|^2 d\mathbf{x} + \rho \int_{\Omega} \left|\left|\mathbf{h}(\mathbf{x}) - \mathbf{g}^{-1}(\mathbf{x})\right|\right|^2 d\mathbf{x} \\ &\quad + \lambda \int_{\Omega} \left|\left|L\mathbf{w}(\mathbf{x})\right|\right|^2 d\mathbf{x}. \end{aligned} \tag{1.39}$$

Equation (1.39) is essentially a two-step strategy, where the maps $\mathbf{g}$ and $\mathbf{h}$ are not estimated simultaneously. Either $\mathbf{g}$ or $\mathbf{h}$ has to be alternatively fixed, which creates a lag in numerical computation.

### 1.6.2 Topology Preservation

The topology can be preserved by ensuring that the transformation $\mathbf{g}$ satisfies two conditions [50]. First, the determinant of the Jacobian of the transformation $|D\mathbf{g}|$ is always positive. Second, the transformation is one-to-one and onto, or bijective. Continuity is implied by the existence of $|D\mathbf{g}|$.

### 1.6.3 Diffeomorphism

It is often important to compute deformations which are not only invertible, but also preserve properties such as smoothness. Hence, diffeomorphic transformations are of considerable interest. The transformation $\mathbf{g}$ is diffeomorphic if both $\mathbf{g}$ and $\mathbf{g}^{-1}$ are differentiable.

The problem with the viscous fluid method [21] is that singularities can develop when the numerical computation is performed. This can be avoided by regularizing

the velocity field. As originally proposed in [21], the transformations are modeled as arising from an evolution in time, where the motion is governed by the laws of continuum mechanics, resulting in a flow $\mathbf{g}(t)$, $t \in [0,1]$ [3, 32, 71, 72, 101]. The forward and inverse maps are uniquely defined as

$$\mathbf{g}^{-1}(\mathbf{g}(\mathbf{x},t) = \mathbf{x}, \qquad t \in [0,1].$$

This implies that the flow can be characterized by

$$\frac{\partial \mathbf{g}(\mathbf{x},t)}{\partial t} = \mathbf{v}(\mathbf{g}(\mathbf{x},t),t), \tag{1.40}$$

$$\frac{\partial \mathbf{g}^{-1}(\mathbf{y},t)}{\partial t} = D\mathbf{g}^{-1}(\mathbf{y},t)\mathbf{v}(\mathbf{y},t), \tag{1.41}$$

$$\mathbf{g}(0) = \mathbf{g}^{-1}(0) = \mathbf{id}, \tag{1.42}$$

where **id** is the identity map and $D\mathbf{g}$ is the Jacobian matrix of the transformation $\mathbf{g}$. The transformation is generated at the endpoint $\mathbf{g}(1)$ of the flow of a velocity vector field $\mathbf{v}(t)$, $t \in [0,1]$, which is described by the ordinary differential equation (1.40). This gives a path $\mathbf{g}(t)$, $t \in [0,1]$ in the space of transformations starting with $\mathbf{g}(0) = \mathbf{id}$ at $t = 0$, and terminating at the endpoint $t = 1$ of the flow, where

$$\mathbf{g}(1) = \mathbf{g}(0) + \int_0^1 \mathbf{v}(\mathbf{g}(t),t)\, dt. \tag{1.43}$$

The generated map $\mathbf{g}(1)$, simply denoted as $\mathbf{g}$ where there is no confusion, provides a matching between the two images.

In [3], the optimal transformation was estimated using the variational formulation in the space of smooth velocity vector fields $V$:

$$\min_{\mathbf{v}} \int_0^1 ||\mathbf{v}||_V^2\, dt + \lambda ||I_2 \circ \mathbf{g} - I_1||_{L^2}^2. \tag{1.44}$$

The smoothness on the velocity field is enforced by defining the norm on the space $V$ through a differential operator $L$, where $L = (-\alpha\triangle + \gamma)^p I$, with $p > 1.5$, such that $||\mathbf{v}||_V = ||L\mathbf{v}||_{L^2}$. The minimization problem (1.44) is referred to as the large deformation diffeomorphic metric mapping (LDDMM) approach.

In [3], the authors derived the Euler-Lagrange equations for the solution of the variational problem in (1.44) and presented a numerical method for obraining a diffeomorphism. This approach allows for arbitrarily large deformations as well as automatic inverse consistency. However, it is computationally expensive, as it requires integration of the velocity field in time. Moreover, this approach does not address the statistical analysis of resulting deformation maps at a voxel level, which becomes relevant in many applications.

### 1.6.4 Unbiased Deformation

The concept of unbiased deformation is introduced and discussed in this dissertation. Unbiased deformation refers to the condition when the Jacobian determinants $|D\mathbf{g}|$ of the deformation $\mathbf{g}$ recovered between a pair of images follow a log-normal distribution, with zero mean after log-transformation:

$$\frac{1}{|\Omega|}\int_{\Omega} \log|D\mathbf{g}(\mathbf{x})|d\mathbf{x} = 0. \tag{1.45}$$

In this dissertation we argue that this distribution is beneficial when recovering change in regions of homogeneous intensity, and in ensuring symmetrical results when the order of two images being registered is switched.

# CHAPTER 2

# Unbiased Large-Deformation Nonlinear Image Registration

In this chapter, we present a novel framework for constructing large deformation 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 computation of log-unbiased deformations.

## 2.1 Theory

One could not study nonlinear image registration without closely examining Jacobian maps. The Jacobian map is the determinant of the Jacobian matrix of a deformation field, and encodes the local volume difference between the source and target images. The Jacobian determinant of a diffeomorphic map is bounded below by zero but unbounded above. Thus, the statistical distribution of Jacobian values would be a better fit to a symmetric distribution if we apply the logarithmic transform.

The logarithmic transform is also advantageous since it symmetrizes the Jacobian distribution by considering halving or doubling of volume to be equally likely a priori,

i.e., assigning equal probabilities to expansions and shrinkages that are reciprocals of each other. This is a reasonable requirement as the correspondence field should be the same regardless of the order of the two images that are matched; if mappings in both directions are considered, compressions or expansions are equally likely.

Thus, the logarithmic transform is crucial in analyzing Jacobian determinant values, and in this chapter we argue that all pertinent statistical analyses should be conducted in this space. In this section, we provide rigorous mathematical analyses of the Jacobian maps and use them to construct unbiased nonlinear image registration.

### 2.1.1 Global Preservation of Density Maps

Let $\Omega$ be an open and bounded domain in $\mathbb{R}^n$. Let us assume, without loss of generality, that the volume of this domain is 1, i.e., $|\Omega| = 1$. We study smooth deformations $\mathbf{g}$ that map $\Omega$ bijectively onto itself. The inverse map of $\mathbf{g}$ is denoted as $\mathbf{g}^{-1}$ and the Jacobian matrix of $\mathbf{g}$ as $D\mathbf{g}$. The Jacobian map can thus be defined as the determinant of the Jacobian matrix

$$|D\mathbf{g}| := \det(D\mathbf{g}) = \begin{vmatrix} \partial_1 g_1 & \cdots & \partial_n g_1 \\ \vdots & \ddots & \vdots \\ \partial_1 g_n & \cdots & \partial_n g_n \end{vmatrix},$$

where $\mathbf{g} = (g_1, \ldots, g_n)$.

In volumetric studies, the determinant of the Jacobian matrix (density) applied to any given deformation $\mathbf{g}$ is an important quantity, encoding the voxelwise volume change. As $\mathbf{g}$ (and $\mathbf{g}^{-1}$) is bijective and thus globally volume-preserving, we have the following preservation of global density:

$$\begin{aligned} \int_\Omega |D\mathbf{g}(\boldsymbol{\xi})| d\boldsymbol{\xi} &= \int_\Omega d\mathbf{y} = 1, \\ \int_\Omega |D\mathbf{g}^{-1}(\boldsymbol{\xi})| d\boldsymbol{\xi} &= \int_\Omega d\mathbf{x} = 1. \end{aligned} \tag{2.1}$$

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

$$\begin{aligned} p_{\mathbf{g}}(\cdot) &= |D\mathbf{g}(\cdot)|, \\ p_{\mathbf{g}^{-1}}(\cdot) &= |D\mathbf{g}^{-1}(\cdot)|, \\ p_{\mathbf{id}}(\cdot) &= 1. \end{aligned} \tag{2.2}$$

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

$$D\mathbf{g}^{-1}(\mathbf{y}) \cdot D\mathbf{g}(\mathbf{x}) = \mathbf{id}, \tag{2.3}$$

and hence,

$$|D\mathbf{g}^{-1}(\mathbf{y})| \cdot |D\mathbf{g}(\mathbf{x})| = 1. \tag{2.4}$$

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:

$$SKL(p_{\mathbf{g}}, p_{\mathbf{id}}) = KL(p_{\mathbf{id}}, p_{\mathbf{g}}) + KL(p_{\mathbf{g}}, p_{\mathbf{id}}) \tag{2.5}$$

to measure the magnitude of any deformation $\mathbf{g}$. Here KL, the Kullback-Leibler divergence between two probability density functions $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$, is defined as

$$KL(p_1(\mathbf{x}), p_2(\mathbf{x})) = \int_{\Omega} p_1(\mathbf{x}) \log \frac{p_1(\mathbf{x})}{p_2(\mathbf{x})} d\mathbf{x} \geq 0. \tag{2.6}$$

To motivate this approach, notice that the first part of SKL measure is simply integrat-

ing the log-density over the entire computational image domain:

$$\begin{aligned}\int_\Omega \log|D\mathbf{g}(\mathbf{x})|d\mathbf{x} &= -\int_\Omega \log\frac{1}{|D\mathbf{g}(\mathbf{x})|}d\mathbf{x} = -\int_\Omega p_{\mathbf{id}}\log\frac{p_{\mathbf{id}}}{p_{\mathbf{g}}}d\mathbf{x} \\ &= -KL(p_{\mathbf{id}}, p_{\mathbf{g}}) \leq 0.\end{aligned} \tag{2.7}$$

To attach geometric meaning to the second term, we notice that the KL divergence has skew-symmetry with respect to $\mathbf{g}$ and its inverse

$$\begin{aligned}KL(p_{\mathbf{id}}, p_{\mathbf{g}^{-1}}) &= -\int_\Omega \log|D\mathbf{g}^{-1}(\mathbf{y})|d\mathbf{y} = \int_\Omega \big(\log|D\mathbf{g}(\mathbf{x})|\big)|D\mathbf{g}(\mathbf{x})|d\mathbf{x} \\ &= \int_\Omega p_{\mathbf{g}}\log\frac{p_{\mathbf{g}}}{p_{\mathbf{id}}}d\mathbf{x} = KL(p_{\mathbf{g}}, p_{\mathbf{id}}),\end{aligned} \tag{2.8}$$

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

$$\begin{aligned}KL(p_{\mathbf{id}}, p_{\mathbf{g}}) &= -\int_\Omega \log|D\mathbf{g}(\mathbf{x})|d\mathbf{x} = \int_\Omega \big(\log|D\mathbf{g}^{-1}(\mathbf{y})|\big)|D\mathbf{g}^{-1}(\mathbf{y})|\,d\mathbf{y} \\ &= \int_\Omega p_{\mathbf{g}^{-1}}\log\frac{p_{\mathbf{g}^{-1}}}{p_{\mathbf{id}}}d\mathbf{y} = KL(p_{\mathbf{g}^{-1}}, p_{\mathbf{id}}).\end{aligned} \tag{2.9}$$

To further show the close relationship between the KL divergence and Jacobian maps, we can also attach geometric meaning to the integral in equations (2.7), (2.8), (2.9). For example,

$$KL(p_{\mathbf{id}}, p_{\mathbf{g}^{-1}}) = -\int_\Omega \log|D\mathbf{g}^{-1}(\mathbf{y})|d\mathbf{y} = \int_\Omega \big(\log|D\mathbf{g}(\boldsymbol{\xi})|\big)_{\boldsymbol{\xi}=\mathbf{g}^{-1}(\mathbf{y})}d\mathbf{y}. \tag{2.10}$$

Here, the right hand side simply computes the integral of the pulled-back (by the inverse of $\mathbf{g}$) Jacobian map of $\mathbf{g}$.

To summarize, we conclude that symmetrizing KL divergence is equivalent to considering both the forward and backward mapping in image registration. As a result, the skew-symmetry in (2.8) and (2.9) is closely related to the asymmetric nature of KL

divergence. In [78], the authors proposed integrating with respect to the square root of the Jacobian determinant, in order to remove this skew-symmetry. Interestingly, this approach has an equivalent in information theory, namely, the Bhattacharyya distance $B$, another well-known measure [6]

$$B(p_{\mathbf{id}}, p_{\mathbf{g}}) = \int_{\Omega} |D\mathbf{g}(\mathbf{x})|^{\frac{1}{2}} d\mathbf{x} = \int_{\Omega} |D\mathbf{g}^{-1}(\mathbf{x})|^{\frac{1}{2}} d\mathbf{x} = B(p_{\mathbf{g}^{-1}}, p_{\mathbf{id}}). \tag{2.11}$$

Here, the Bhattacharyya distance, though not defined in the logarithmic space, is symmetrical with respect to its two arguments, as well as inverse-consistent. To further connect the KL divergence and Bhattacharyya distance, one can also consider the geodesic linking of the two probability density functions: $pdf(\cdot, t)$ parameterized by $t$

$$\begin{aligned} pdf(\mathbf{x}, t) &= \frac{p_{\mathbf{id}}(\mathbf{x})^t p_{\mathbf{g}}(\mathbf{x})^{1-t}}{N}, \\ N &= \int p_{\mathbf{id}}(\mathbf{x})^t p_{\mathbf{g}}(\mathbf{x})^{1-t} d\mathbf{x}. \end{aligned} \tag{2.12}$$

The Bhattacharyya distance corresponds to the arbitrary choice of $t = \frac{1}{2}$, while a generalization of the above leads to the Chernoff distance in information theory [17, 25].

### 2.1.2 Realizing Unbiased Deformation in the Logarithmic Space

Before developing formulations to construct unbiased deformations in the logarithmic space, we generalize equation (2.7) 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 defor-

mation **h** mapping domain $A$ in the source (with volume $a$) to domain $B$ in the target (with volume $b$), we have the following

$$\frac{1}{a}\int_A \log|D\mathbf{h}(\mathbf{x})|d\mathbf{x} \leq \log\frac{b}{a}, \tag{2.13}$$

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

$$\sum_{i=1}^{n}\log(x_i) \leq n\log(\bar{x}); \quad \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{2.14}$$

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).

### 2.1.3 Unbiased Nonlinear Image Registration in the Logarithmic Space via Symmetric KL Distance

Given equation (2.7) 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:

$$\begin{aligned}
&SKL(p_{\mathbf{g}}, p_{\mathbf{id}}) = SKL(p_{\mathbf{g}^{-1}}, p_{\mathbf{id}}) \\
&= KL(p_{\mathbf{g}}, p_{\mathbf{id}}) + KL(p_{\mathbf{g}^{-1}}, p_{\mathbf{id}}) = KL(p_{\mathbf{g}}, p_{\mathbf{id}}) + KL(p_{\mathbf{id}}, p_{\mathbf{g}}) \\
&= KL(p_{\mathbf{id}}, p_{\mathbf{g}^{-1}}) + KL(p_{\mathbf{id}}, p_{\mathbf{g}}) = KL(p_{\mathbf{id}}, p_{\mathbf{g}^{-1}}) + KL(p_{\mathbf{g}^{-1}}, p_{\mathbf{id}}) \\
&= \int_{\Omega} \big(|D\mathbf{g}(\mathbf{x})| - 1\big) \log |D\mathbf{g}(\mathbf{x})| d\mathbf{x} \\
&= \int_{\Omega} \big(|D\mathbf{g}^{-1}(\mathbf{y})| - 1\big) \log |D\mathbf{g}^{-1}(\mathbf{y})| d\mathbf{y}.
\end{aligned} \tag{2.15}$$

Given an image matching function, we argue that one achieves unbiased deformation by seeking, among all deformations minimizing this image matching functional, the deformation with minimal distance as given in (2.15).

To see why minimizing equation (2.15) leads to unbiased deformation in the logarithmic space, we observe that the integrand is always non-negative, and only evaluates to zero when $\mathbf{g}$ is volume-preserving everywhere (Jacobian of $\mathbf{g}$ 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 (2.13).

Hence, we propose to minimize the regularization functional based on the symmetric KL distance between $p_{\mathbf{g}}$ and $p_{\mathbf{id}}$

$$R_{SKL}(\mathbf{u}) = \int_{\Omega} \big(|D(\mathbf{x} - \mathbf{u}(\mathbf{x}))| - 1\big) \log |D(\mathbf{x} - \mathbf{u}(\mathbf{x}))| d\mathbf{x}. \tag{2.16}$$

Under this framework, constructing deformations can be viewed as quantifying the symmetric KL distance between the identity map and the resulting deformation (or the inverse deformation due to the equivalence in equation (2.15)). Moreover, this framework embeds statistical analyses into the construction of deformations, penalizing deformations that skew the distribution of test statistics. A second interpretation

of (2.15) is that it simply calculates the mean log-Jacobian for **g** and its inverse inside the domain, thus computing the integral effect of the test statistics on the whole image domain as well. To further link this approach to other branches of mathematics, optimization problems involving the Jacobian operator are commonly encountered in grid generation [66] and in continuum mechanics, where the Hencky tensor arises in modeling very large deformations. However, we believe that the logarithmic transform has not been formally introduced in the grid generation literature and may also be useful there.

## 2.2 Implementation

In this section, we generalize the viscous fluid registration model described in Section 1.4.5 to compute unbiased deformations. We provide a new energy functional incorporating the symmetric Kullback-Leibler distance from Section 2.1, give corresponding force field equations, and describe an algorithm for the new unbiased viscous fluid registration model (sometimes referred to as the Unbiased model).

### 2.2.1 Energy Minimization

We propose to minimize the following energy:

$$E(I_1, I_2, \mathbf{u}) = F(I_1, I_2, \mathbf{u}) + \lambda R_{SKL}(\mathbf{u}), \tag{2.17}$$

where $F$ is the similarity measure, $R_{SKL}$ is the proposed regularization term given by equation (2.16), and $\lambda > 0$ is the Lagrange multiplier. The similarity measure $F$ can assume the $L^2$-norm distance between the target image $I_1(\mathbf{x})$ and the deformed source

$I_2(\mathbf{x}-\mathbf{u})$:

$$\begin{aligned} F_{L^2}(I_1, I_2, \mathbf{u}) &= \frac{1}{2}\big|\big|I_2(\mathbf{x}-\mathbf{u}) - I_1(\mathbf{x})\big|\big|_2 \\ &= \frac{1}{2}\int_\Omega \big(I_2(\mathbf{x}-\mathbf{u}) - I_1(\mathbf{x})\big)^2 \, d\mathbf{x}, \end{aligned} \tag{2.18}$$

or a negative of the mutual information between $I_1(\mathbf{x})$ and $I_2(\mathbf{x}-\mathbf{u})$:

$$\begin{aligned} F_{MI}(I_1, I_2, \mathbf{u}) &= -MI_{\mathbf{u}}^{I_1,I_2} \\ &= -\int_{\mathbb{R}^2} p_{\mathbf{u}}^{I_1,I_2}(i_1, i_2) \, \log \frac{p_{\mathbf{u}}^{I_1,I_2}(i_1, i_2)}{p^{I_1}(i_1) p_{\mathbf{u}}^{I_2}(i_2)} \, di_1 di_2, \end{aligned} \tag{2.19}$$

which was defined in Section 1.3.3. The force field vector $\mathbf{f}$ for energy in (2.17) will be given as

$$\mathbf{f}(\mathbf{x}, \mathbf{u}(\mathbf{x}, t)) = -\partial_{\mathbf{u}} F(I_1, I_2, \mathbf{u}) - \lambda \, \partial_{\mathbf{u}} R_{SKL}(\mathbf{u}). \tag{2.20}$$

The $L^2$-based similarity measure gradient $\partial_{\mathbf{u}} F_{L^2}(I_1, I_2, \mathbf{u})$ is given in Section 1.3.2 and mutual information based similarity measure gradient $\partial_{\mathbf{u}} F_{MI}(I_1, I_2, \mathbf{u})$ is given in Section 1.3.3 and is derived in Section 2.4. Gradient of regularization term $R_{SKL}(\mathbf{u})$, namely $\partial_{\mathbf{u}} R_{SKL}(\mathbf{u})$, is derived in Section 2.2.2 for two-dimensional and three-dimensional cases.

Given the velocity field $\mathbf{v}$, the following partial differential equation can be solved to obtain the displacement field $\mathbf{u}$:

$$\frac{\partial \mathbf{u}}{\partial t} = \mathbf{v} - \mathbf{v} \cdot \nabla \mathbf{u}. \tag{2.21}$$

The instantaneous velocity as in [28] is obtained by convolving $\mathbf{f}$ with Gaussian kernel $G_\sigma$ of variance $\sigma^2$:

$$\mathbf{v} = G_\sigma * \mathbf{f}(\mathbf{x}, \mathbf{u}). \tag{2.22}$$

This equation can be solved efficiently using the Fast Fourier transform (FFT).

*Remark:* In Chapter 4, we will investigate a different regularization term $R = R_{KL}$, based on the Kullback-Leibler divergence

$$R_{KL}(\mathbf{g}) = KL(p_{\mathbf{id}}, p_{\mathbf{g}}),$$

or

$$R_{KL}(\mathbf{u}) = \int_{\Omega} -\log|D\mathbf{g}(\mathbf{x})| d\mathbf{x}.$$

In this case, the total energy for minimization will be given as

$$E(I_1, I_2, \mathbf{u}) = F(I_1, I_2, \mathbf{u}) + \lambda R_{KL}(\mathbf{u}), \tag{2.23}$$

and the corresponding force field vector **f** is

$$\mathbf{f}(\mathbf{x}, \mathbf{u}(\mathbf{x}, t)) = -\partial_{\mathbf{u}} F(I_1, I_2, \mathbf{u}) - \lambda\, \partial_{\mathbf{u}} R_{KL}(\mathbf{u}). \tag{2.24}$$

### 2.2.2 The Energy Gradients for the Unbiased Regularization

#### 2.2.2.1 Derivations of Gradient of $R(\mathbf{u})$ in Two Spatial Dimensions

In this section, we derive explicit expressions for $\partial_{\mathbf{u}} R(\mathbf{u})$ in (2.17) and (2.23), with $R = R_{SKL}$ and $R = R_{KL}$, when $\Omega \subset \mathbb{R}^2$. Let us denote the components of vector **x** to be $(x_1, x_2)$ and the components of vector **u** be $(u_1, u_2)$. We also denote $\partial_j u_i = \partial u_i / \partial x_j$.

To simplify the notation, we let $J = |D\mathbf{g}| = |D(\mathbf{x} - \mathbf{u})|$. Also, denote $L(J) = L_{KL}(J) = -\log J$, when $R = R_{KL}$, and $L(J) = L_{SKL}(J) = (J-1)\log J$, when $R = R_{SKL}$. Note that $J : \mathcal{M}_{2\times 2}(\mathbb{R}) \to \mathbb{R}$, where $\mathcal{M}_{2\times 2}(\mathbb{R})$ is the set of $2 \times 2$ matrices

with real elements, and $L : \mathbb{R} \to \mathbb{R}$. Jacobian $J$ is a function of $\partial_j u_i$, for $i, j = 1, 2$, and is given by

$$\begin{aligned} J\big(\partial_1 u_1, \partial_2 u_1, \partial_1 u_2, \partial_2 u_2\big) &= \begin{vmatrix} 1 - \partial_1 u_1 & -\partial_2 u_1 \\ -\partial_1 u_2 & 1 - \partial_2 u_2 \end{vmatrix} \\ &= (1 - \partial_1 u_1)(1 - \partial_2 u_2) - \partial_2 u_1\, \partial_1 u_2. \end{aligned}$$

We would like to minimize the functional

$$R(\mathbf{u}) = \int_\Omega L\big(\partial_1 u_1, \partial_2 u_1, \partial_1 u_2, \partial_2 u_2\big) d\mathbf{x}.$$

We derive the first Euler-Lagrange equation. For some $\eta \in C_c^1(\Omega)$, we have

$$\begin{aligned} &\frac{dR}{d\varepsilon}(u_1 + \varepsilon\eta, u_2)\big|_{\varepsilon=0} \\ &= \int_\Omega \left[\frac{dL}{dJ}\frac{\partial J}{\partial(\partial_1 u_1)}\,\partial_{x_1}\eta + \frac{dL}{dJ}\frac{\partial J}{\partial(\partial_2 u_1)}\,\partial_{x_2}\eta\right] d\mathbf{x} \\ &= -\int_\Omega \left[\frac{\partial}{\partial x_1}\left(\frac{dL}{dJ}\frac{\partial J}{\partial(\partial_1 u_1)}\right) + \frac{\partial}{\partial x_2}\left(\frac{dL}{dJ}\frac{\partial J}{\partial(\partial_2 u_1)}\right)\right]\eta\, d\mathbf{x}. \end{aligned}$$

With notation $L' = dL/dJ$, the first Euler-Lagrange equation becomes:

$$-\frac{\partial}{\partial x_1}\left(L'\frac{\partial J}{\partial(\partial_1 u_1)}\right) - \frac{\partial}{\partial x_2}\left(L'\frac{\partial J}{\partial(\partial_2 u_1)}\right) = 0.$$

Thus, minimizing the energy $R(\mathbf{u})$ with respect to $u_1$, for fixed $u_2$, yields the first component of $\partial_{\mathbf{u}} R(\mathbf{u})$:

$$\partial_{u_1} R(\mathbf{u}) = \frac{\partial}{\partial x_1}\Big(\big(1 - \partial_2 u_2\big)L'\Big) + \frac{\partial}{\partial x_2}\Big(\partial_1 u_2\, L'\Big). \tag{2.25}$$

Note that $L'_{KL}(J) = -1/J$ and $L'_{SKL}(J) = 1 + \log J - 1/J$.

Similarly, the Euler-Lagrange equation for the second component of $\partial_{\mathbf{u}} R(\mathbf{u})$ can be

found to be:

$$\partial_{u_2} R(\mathbf{u}) = \frac{\partial}{\partial x_1}\Big(\partial_2 u_1\, L'\Big) + \frac{\partial}{\partial x_2}\Big(\big(1-\partial_1 u_1\big)L'\Big). \tag{2.26}$$

#### 2.2.2.2 Derivations of Gradient of $R(\mathbf{u})$ in Three Spatial Dimensions

In this section, we derive explicit expression for $\partial_{\mathbf{u}} R(\mathbf{u})$ in (2.17) and (2.23), with $R = R_{SKL}$ and $R = R_{KL}$, when $\Omega \subset \mathbb{R}^3$. Let us denote the components of vector $\mathbf{x}$ to be $(x_1, x_2, x_3)$ and the components of vector $\mathbf{u}$ be $(u_1, u_2, u_3)$. Here, we will use the notation from Section 2.2.2.1.

Jacobian $J$ is a function of $\partial_j u_i$, for $i, j = 1, 2, 3$, and is given by

$$\begin{aligned}
&J\big(\partial_1 u_1, \partial_2 u_1, \partial_3 u_1, \partial_1 u_2, \partial_2 u_2, \partial_3 u_2, \partial_1 u_3, \partial_2 u_3, \partial_3 u_3\big) \\
&\quad = \begin{bmatrix} 1-\partial_1 u_1 & -\partial_2 u_1 & -\partial_3 u_1 \\ -\partial_1 u_2 & 1-\partial_2 u_2 & -\partial_3 u_2 \\ -\partial_1 u_3 & -\partial_2 u_3 & 1-\partial_3 u_3 \end{bmatrix} \\
&\quad = (1-\partial_1 u_1)(1-\partial_2 u_2)(1-\partial_3 u_3) - \partial_1 u_2\, \partial_2 u_3\, \partial_3 u_1 - \partial_2 u_1\, \partial_3 u_2\, \partial_1 u_3 \\
&\quad\quad -\partial_3 u_1 (1-\partial_2 u_2)\partial_1 u_3 - \partial_2 u_1\, \partial_1 u_2 (1-\partial_3 u_3) - \partial_3 u_2\, \partial_2 u_3 (1-\partial_1 u_1).
\end{aligned}$$

We would like to minimize the functional

$$R(\mathbf{u}) = \int_\Omega L\big(J(\partial_j u_i)\big) d\mathbf{x}, \quad 1 \le i, j \le 3.$$

For some $\eta$, we have

$$\begin{aligned}
&\frac{dR}{d\varepsilon}(u_1+\varepsilon\eta,u_2,u_3)\big|_{\varepsilon=0}\\
&=\int_\Omega\left[\frac{dL}{dJ}\frac{\partial J}{\partial(\partial_1 u_1)}\,\partial_{x_1}\eta+\frac{dL}{dJ}\frac{\partial J}{\partial(\partial_2 u_1)}\,\partial_{x_2}\eta+\frac{dL}{dJ}\frac{\partial J}{\partial(\partial_3 u_1)}\,\partial_{x_3}\eta\right]d\mathbf{x}\\
&=-\int_\Omega\left[\frac{\partial}{\partial x_1}\left(\frac{dL}{dJ}\frac{\partial J}{\partial(\partial_1 u_1)}\right)+\frac{\partial}{\partial x_2}\left(\frac{dL}{dJ}\frac{\partial J}{\partial(\partial_2 u_1)}\right)+\frac{\partial}{\partial x_3}\left(\frac{dL}{dJ}\frac{\partial J}{\partial(\partial_3 u_1)}\right)\right]\eta d\mathbf{x}.
\end{aligned}$$

Hence, the first Euler-Lagrange equation becomes:

$$-\frac{\partial}{\partial x_1}\left(L'\frac{\partial J}{\partial(\partial_1 u_1)}\right)-\frac{\partial}{\partial x_2}\left(L'\frac{\partial J}{\partial(\partial_2 u_1)}\right)-\frac{\partial}{\partial x_3}\left(L'\frac{\partial J}{\partial(\partial_3 u_1)}\right)=0.$$

Thus, minimizing the energy $R(\mathbf{u})$ with respect to $u_1$, for fixed $u_2$ and $u_3$, yields the first component of $\partial_{\mathbf{u}}R(\mathbf{u})$:

$$\begin{aligned}
\partial_{u_1}R(\mathbf{u})=&\frac{\partial}{\partial x_1}\Big(\big((1-\partial_2u_2)(1-\partial_3u_3)-\partial_3u_2\,\partial_2u_3\big)L'\Big)\\
&+\frac{\partial}{\partial x_2}\Big(\big(\partial_3u_2\,\partial_1u_3+\partial_1u_2(1-\partial_3u_3)\big)L'\Big)\\
&+\frac{\partial}{\partial x_3}\Big(\big(\partial_1u_2\,\partial_2u_3+(1-\partial_2u_2)\partial_1u_3\big)L'\Big).
\end{aligned}\tag{2.27}$$

Similarly, the other two Euler-Lagrange equations can be found to be:

$$\begin{aligned}
\partial_{u_2}R(\mathbf{u})=&\frac{\partial}{\partial x_1}\Big(\big(\partial_2u_3\,\partial_3u_1+\partial_2u_1(1-\partial_3u_3)\big)L'\Big)\\
&+\frac{\partial}{\partial x_2}\Big(\big((1-\partial_1u_1)(1-\partial_3u_3)-\partial_3u_1\,\partial_1u_3\big)L'\Big)\\
&+\frac{\partial}{\partial x_3}\Big(\big(\partial_2u_1\,\partial_1u_3+\partial_2u_3(1-\partial_1u_1)\big)L'\Big),
\end{aligned}\tag{2.28}$$

and

$$\begin{aligned}\partial_{u_3} R(\mathbf{u}) = &\frac{\partial}{\partial x_1}\Big(\big(\partial_2 u_1\, \partial_3 u_2 + \partial_3 u_1(1-\partial_2 u_2)\big)L'\Big)\\ &+\frac{\partial}{\partial x_2}\Big(\big(\partial_1 u_2\, \partial_3 u_1 + \partial_3 u_2(1-\partial_1 u_1)\big)L'\Big)\\ &+\frac{\partial}{\partial x_3}\Big(\big((1-\partial_1 u_1)(1-\partial_2 u_2) - \partial_2 u_1\, \partial_1 u_2\big)L'\Big).\end{aligned} \tag{2.29}$$

### 2.2.3 Numerical Discretization

#### 2.2.3.1 Discretization in 2D

To discretize equations (2.25) and (2.26), we use finite difference schemes. Let $\triangle x_1$, $\triangle x_2$ be the spatial steps, $\triangle t$ be the time step, and $(x_{1i}, x_{2j}) = (i\triangle x_1, j\triangle x_2)$ be the grid points, for $1 \le i \le M$, $1 \le j \le N$. For a function $\varphi : \Omega \to \mathbb{R}$, let $\varphi^n_{i,j} = \varphi(n\triangle t, i\triangle x_1, j\triangle x_2)$. We define the difference operators based on uniformly-spaced grid as

$$D^{x_1}\varphi^n_{i,j} = \frac{\varphi^n_{i+1,j} - \varphi^n_{i-1,j}}{2\triangle x_1}, \quad D^{x_2}\varphi^n_{i,j} = \frac{\varphi^n_{i,j+1} - \varphi^n_{i,j-1}}{2\triangle x_2}.$$

Below, we will use the following notations when it is obvious that the grid point at $(i\triangle x_1, j\triangle x_2)$ is under consideration

$$\varphi^n := \varphi^n_{i,j}, \quad D^{x_l}\varphi^n := D^{x_l}\varphi^n_{i,j}, \quad l = 1, 2.$$

Hence, in the two-dimensional case, the force field $\mathbf{f} = (f_1, f_2)$ from (2.20) and (2.24), at each grid point $(i, j)$, at $n$-th time step, is approximated by

$$\begin{aligned} f_1 &\approx -\big[\partial_{u_1} F(\mathbf{u}^n)\big] - \lambda\Big[D^{x_1}\big(\big(1 - D^{x_2}u_2^n\big)L'\big) + D^{x_2}\big(D^{x_1}u_2^n\, L'\big)\Big],\\ f_2 &\approx -\big[\partial_{u_2} F(\mathbf{u}^n)\big] - \lambda\Big[D^{x_1}\big(D^{x_2}u_1^n\, L'\big) + D^{x_2}\big(\big(1 - D^{x_1}u_1^n\big)L'\big)\Big],\end{aligned}$$

where $[\partial_{u_l} F(\mathbf{u}^n)]$, $l = 1, 2$, is a discretization of a similarity-based gradient.

#### 2.2.3.2 Discretization in 3D

To discretize equations (2.27), (2.28), and (2.29), we use finite difference schemes. Let $\triangle x_1$, $\triangle x_2$, $\triangle x_3$ be the spatial steps, $\triangle t$ be the time step, and $(x_{1i}, x_{2j}, x_{3k}) = (i\triangle x_1, j\triangle x_2, k\triangle x_3)$ be the grid points, for $1 \leq i \leq M$, $1 \leq j \leq N$, $1 \leq k \leq P$. For a function $\varphi : \Omega \to \mathbb{R}$, let $\varphi^n_{i,j,k} = \varphi(n\triangle t, i\triangle x_1, j\triangle x_2, k\triangle x_3)$. We define the difference operators based on uniformly-spaced grid as

$$\begin{aligned}
D^{x_1}\varphi^n_{i,j,k} &= \frac{\varphi^n_{i+1,j,k} - \varphi^n_{i-1,j,k}}{2\triangle x_1}, \\
D^{x_2}\varphi^n_{i,j,k} &= \frac{\varphi^n_{i,j+1,k} - \varphi^n_{i,j-1,k}}{2\triangle x_2}, \\
D^{x_3}\varphi^n_{i,j,k} &= \frac{\varphi^n_{i,j,k+1} - \varphi^n_{i,j,k-1}}{2\triangle x_3}, \\
D^{x_1x_1}\varphi^n_{i,j,k} &= \frac{\varphi^n_{i+1,j,k} - 2\varphi^n_{i,j,k} + \varphi^n_{i-1,j,k}}{\triangle x_1^2}, \\
D^{x_2x_2}\varphi^n_{i,j,k} &= \frac{\varphi^n_{i,j+1,k} - 2\varphi^n_{i,j,k} + \varphi^n_{i,j-1,k}}{\triangle x_2^2}, \\
D^{x_3x_3}\varphi^n_{i,j,k} &= \frac{\varphi^n_{i,j,k+1} - 2\varphi^n_{i,j,k} + \varphi^n_{i,j,k-1}}{\triangle x_3^2}.
\end{aligned}$$

Hence, in the three-dimensional case, the force field $\mathbf{f} = (f_1, f_2, f_3)$ from (2.20) and (2.24), at each grid point $(i, j, k)$, at $n$-th time step, is approximated by

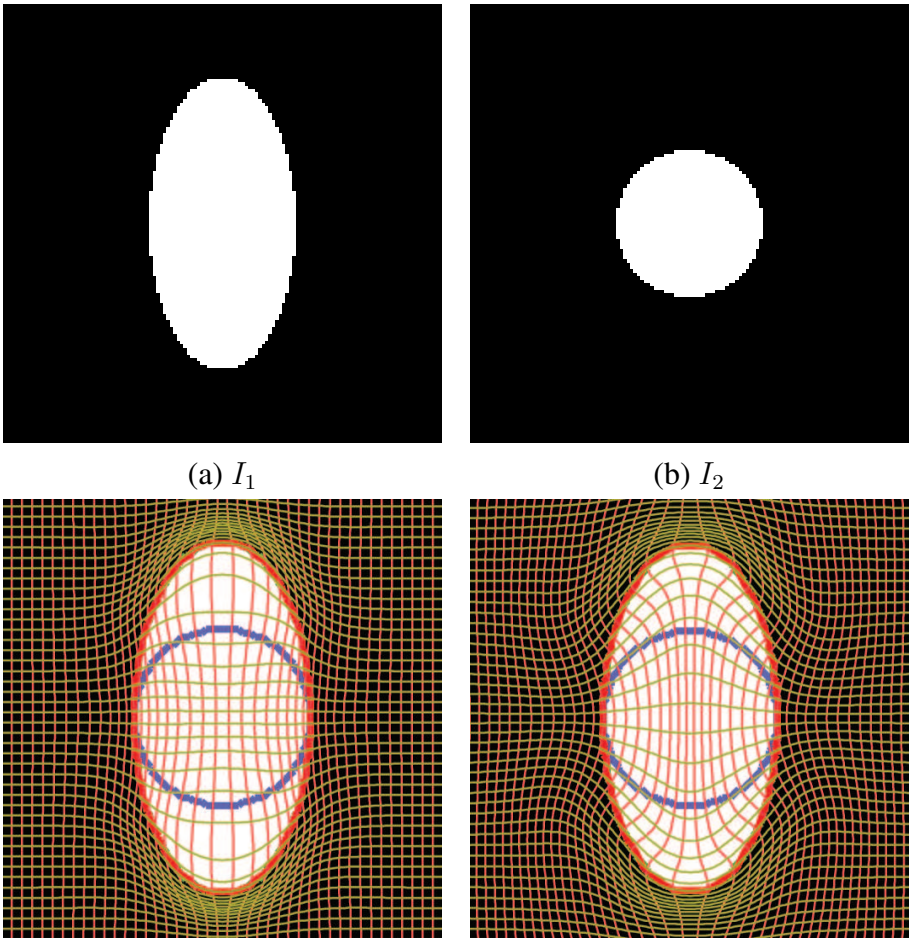
$$\begin{aligned}
f_1 \approx -\big[\partial_{u_1} F(\mathbf{u}^n)\big] \;-\; \lambda\Big[ & D^{x_1}\Big(\big((1 - D^{x_2}u_2^n)(1 - D^{x_3}u_3^n) - D^{x_3}u_2^n\, D^{x_2}u_3^n\big)L'\Big) \\
& + D^{x_2}\Big(\big(D^{x_3}u_2^n\, D^{x_1}u_3^n + D^{x_1}u_2^n(1 - D^{x_3}u_3^n)\big)L'\Big) \\
& + D^{x_3}\Big(\big(D^{x_1}u_2^n\, D^{x_2}u_3^n + (1 - D^{x_2}u_2^n)D^{x_1}u_3^n\big)L'\Big)\Big],
\end{aligned}$$

$$
\begin{aligned}
f_2 \approx -\big[\partial_{u_2} F(\mathbf{u}^n)\big] \;\; &- \;\; \lambda\Big[D^{x_1}\Big(\big(D^{x_2}u_3^n\, D^{x_3}u_1^n + D^{x_2}u_1^n(1-D^{x_3}u_3^n)\big)L'\Big) \\
&\quad + D^{x_2}\Big(\big((1-D^{x_1}u_1^n)(1-D^{x_3}u_3^n) - D^{x_3}u_1^n\, D^{x_1}u_3^n\big)L'\Big) \\
&\quad + D^{x_3}\Big(\big(D^{x_2}u_1^n\, D^{x_1}u_3^n + D^{x_2}u_3^n(1-D^{x_1}u_1^n)\big)L'\Big)\Big], \\
f_3 \approx -\big[\partial_{u_3} F(\mathbf{u}^n)\big] \;\; &- \;\; \lambda\Big[D^{x_1}\Big(\big(D^{x_2}u_1^n\, D^{x_3}u_2^n + D^{x_3}u_1^n(1-D^{x_2}u_2^n)\big)L'\Big) \\
&\quad + D^{x_2}\Big(\big(D^{x_1}u_2^n\, D^{x_3}u_1^n + D^{x_3}u_2^n(1-D^{x_1}u_1^n)\big)L'\Big) \\
&\quad + D^{x_3}\Big(\big((1-D^{x_1}u_1^n)(1-D^{x_2}u_2^n) - D^{x_2}u_1^n\, D^{x_1}u_2^n\big)L'\Big)\Big],
\end{aligned}
$$

where $[\partial_{u_l} F(\mathbf{u}^n)]$, $l = 1, 2, 3$, is a discretization of a similarity-based gradient.

### 2.2.4 Algorithm

In this section, we describe an algorithm for the unbiased fluid registration method, and in the next section, we compare the results generated using this algorithm to those obtained using the viscous fluid registration model. To obtain a fair comparison between the two methods, we do not employ re-gridding. Re-gridding is a method to relax the energy computed from the linear elasticity prior after a certain number of iterations, which allows large-deformation mappings to be recovered without any absolute penalty on the displacement field (other than via the smoothness constraint on the velocity field which is integrated to give the displacement) [21]. It is essentially a memory-less procedure, as how images are matched after each re-gridding is independent of the final deformation before the re-gridding, rendering the comparison of final Jacobian fields and cost functionals problematic. Moreover, we consider the strategy of re-gridding, through the relaxation of deformation fields over time, to be less rigorous from a theoretical standpoint, as the imposition of a regularizer can be used to secure distributional properties in the resulting statistics (e.g., symmetric log-Jacobian). For that reason, for the results presented in this chapter, velocity vector $\mathbf{v}$ was obtained from equation (2.22) rather than by solving the Navier-Stokes equation



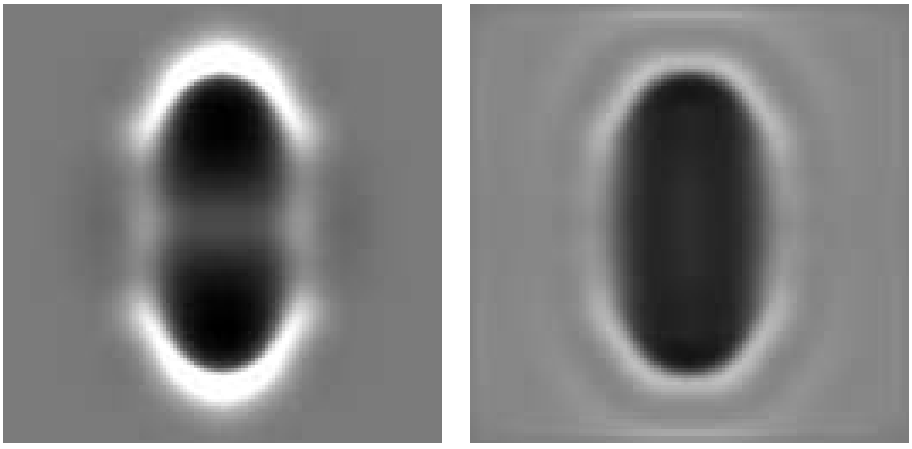
(a) $I_1$ (b) $I_2$

(c) Fluid (d) Unbiased

Figure 2.1: Disk-to-Ellipse example. (a) image $I_1$; (b) image $I_2$; (c) image $I_2$ is deformed to image $I_1$ using the viscous fluid model; (d) image $I_2$ is deformed to image $I_1$ using the Unbiased model. Yellow, blue and red contours represent the boundaries of objects in $I_1$, $I_2$, and deformed $I_2$, respectively. Note that for both methods, yellow contour is essentially invisible due to a very close match. However, the resulting grid of the Unbiased method is visually more regular.

for a compressible viscous fluid (1.34), for both the viscous fluid and the Unbiased fluid registration models. From our experience, solving equation (1.34) makes it necessary to employ re-gridding in order to achieve close matching.

We are now ready to give an algorithm for the proposed unbiased fluid registration



(a) Fluid (b) Unbiased

Figure 2.2: Disk-to-Ellipse example. Jacobian map of the deformation using (a) the viscous fluid model and (b) Unbiased model.

method.

**Algorithm 1** Unbiased Nonlinear Fluid Registration

1: Initialize $t = 0$ and $\mathbf{u}(\mathbf{x}, 0) = 0$.
2: Given $\mathbf{u}(\mathbf{x}, t)$, calculate the force field $\mathbf{f}(\mathbf{x}, \mathbf{u}(\mathbf{x}, t))$ using equation (2.20).
Numerical discretization of equation (2.20) is described in Section 2.2.3 for two-dimensional and three-dimensional cases.
Note that the viscous fluid model, described in Section 1.4.5, obtains the force field using equation (1.8).
3: Solve (2.22) for the instantaneous velocity $\mathbf{v}(\mathbf{x}, t)$.
Steps 4-6 describe the procedure for solving equation (2.21), advancing $\mathbf{u}(\mathbf{x}, t)$ in time.
4: Calculate the perturbation of the displacement field
$\mathbf{R}(\mathbf{x}) = \mathbf{v}(\mathbf{x}, t) - \mathbf{v}(\mathbf{x}, t) \cdot \nabla \mathbf{u}(\mathbf{x}, t)$.
5: Time step $\triangle t$ is calculated adaptively so that $\triangle t \cdot \max(||\mathbf{R}||_2) = \delta u$, where $\delta u$ is the maximal displacement allowed in one iteration. Results in this work are obtained with $\delta u = 0.1$.
6: Advance equation (2.21), i.e. $\partial \mathbf{u}(\mathbf{x}, t)/\partial t = \mathbf{R}(\mathbf{x})$, in time, with time step from step 4, solving for $\mathbf{u}(\mathbf{x}, t)$.
7: If the cost functional in (2.17) decreases by sufficiently small amount compared to the previous iteration, then stop.
8: Let $t := t + \triangle t$ and go to step 2.

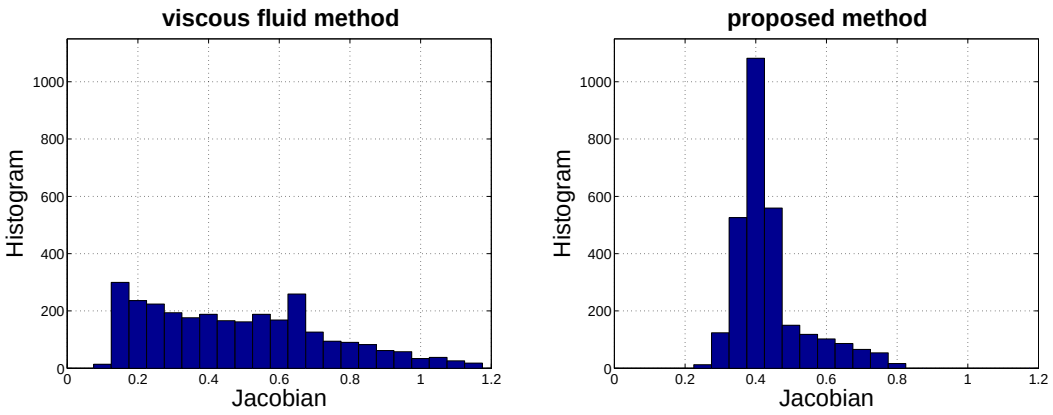



Figure 2.3: Disk-to-Ellipse example. Histograms of Jacobian values of the deformations inside the ellipse for the viscous fluid model and Unbiased model.

## 2.3 Results and Discussion

In this section, we tested the proposed Unbiased nonlinear registration model (referred to as the Unbiased model) using pairs of synthetic images, binary corpus callosum images, two-dimensional serial MRI images, as well as a set of three-dimensional serial MRI brain images. The results obtained using the Unbiased model were compared to those computed using the viscous fluid registration method. In all examples, unless otherwise specified, $L^2$-based similarity measure (2.18) was employed in equation (2.17).

### 2.3.1 Disk-to-Ellipse Example

In order to gain more insight into the effect of the regularization term in (2.16) used in Unbiased registration, we first consider matching two binary synthetic images. In Figures 2.1 through 2.4, we show the results of deforming a disk into an ellipse (both 128 by 128; $\lambda = 500$ in equation (2.20)). As seen in Figure 2.1(c,d), both the fluid registration (Christensen's) model and the Unbiased model generated a close match

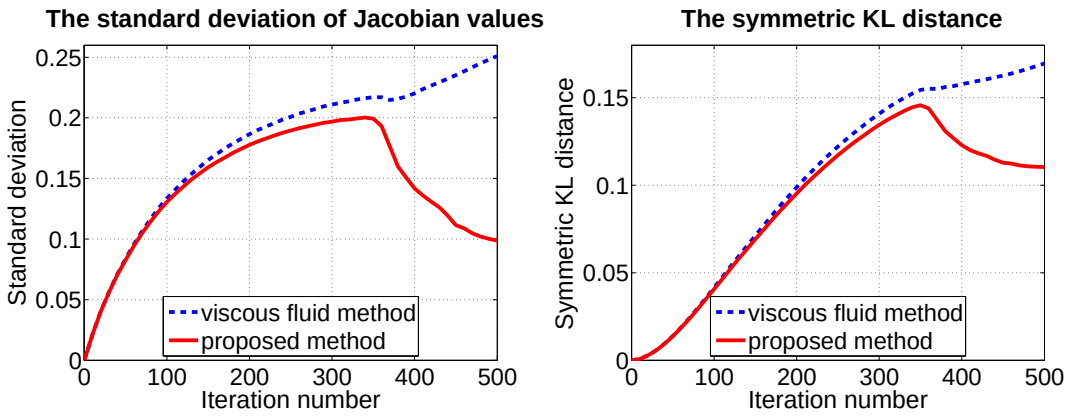
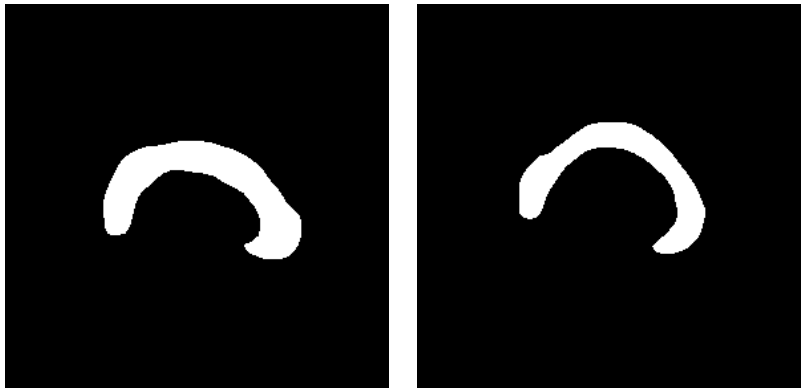
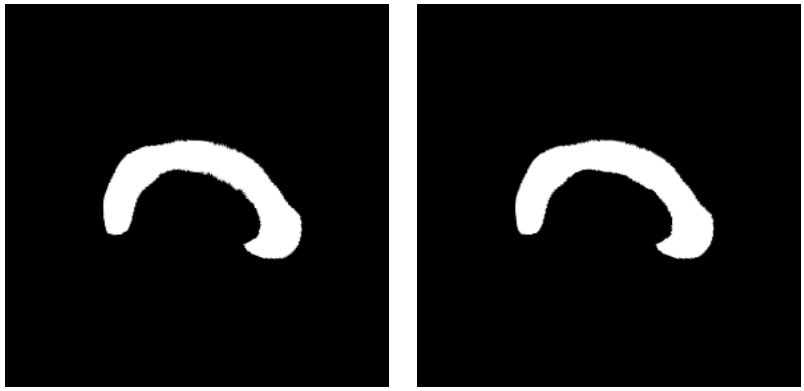



Figure 2.4: Disk-to-Ellipse example. (a) Standard deviation of Jacobian values inside the ellipse per iteration. (b) Symmetric KL distance. For the viscous fluid model (dashed blue), both standard deviation and symmetric KL distance increase while for Unbiased model (solid red), both standard deviation and symmetric KL distance stabilize.

between the deformed source and target images. Here, optimal matching was considered achieved once the overall cost functional stopped decreasing. However, as seen in Figures 2.2 and 2.3, the Unbiased method more evenly distributes deformation inside and outside an ellipse (resulting from the convex property of the logarithmic mapping in inequality (2.13)). Note the vertical stretching of the grid in the center of the ellipse for the proposed method, which is a consequence of uniform distribution of Jacobian values. In the case of a viscous fluid deformation, however, grid does not uniformly adjust to object's volume change; this is especially noticeable in the center of the ellipse. Figure 2.4(a) plots the standard deviation of the Jacobian field inside the ellipse as a function of iteration number. For the viscous fluid model, the standard deviation inside the ellipse increased with the number of iterations, while the Unbiased method yielded an optimized standard deviation as more iterations were computed. The proposed symmetric KL distance also increased for the viscous fluid method, while it was minimized for the Unbiased method as shown in Figure 2.4(b).



(a) $I_1$ (b) $I_2$

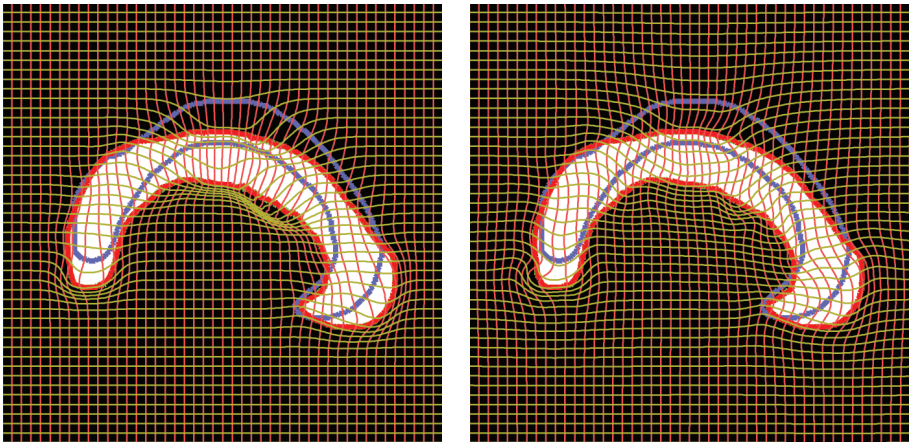


(c) $I_2 \circ \mathbf{g}$, Fluid (d) $I_2 \circ \mathbf{g}$, Unbiased

Figure 2.5: Corpus callosum example. (a) image $I_1$; (b) image $I_2$; (c) image $I_2$ is deformed to image $I_1$ using the viscous fluid model; (d) image $I_2$ is deformed to image $I_1$ using the Unbiased model.

### 2.3.2 Corpus Callosum Example

In this example, we consider matching two 2D binary images (both 289 by 289, $\lambda = 1000$ in equation (2.20)), representing midline corpus callosum contours of two control subjects (Figures 2.5 through 2.9). As seen in Figure 2.5, both the fluid registration model and the Unbiased model generated a close match between the deformed source and target images. However, Figures 2.6 and 2.7 show grid lines merge and self-cross for the viscous fluid model, which is a consequence of negative Jacobian



Fluid Unbiased

Figure 2.6: Corpus callosum example. Results obtained with the viscous fluid model and the Unbiased model. Yellow, blue and red contours represent the boundaries of corpus callosum in $I_1$, $I_2$, and deformed $I_2$, respectively. For both methods, yellow contour is essentially invisible due to a very close match. However, the resulting grid of the Unbiased method is visually more regular. Also, note the grid lines merging and self-crossing for the viscous fluid model, signifying a topology change.

values at certain places, indicating topology change. The Unbiased method, on the other hand, more evenly distributes deformation inside and outside the corpus callosum. The histograms of the Jacobian field inside the corpus callosum are shown in Figure 2.8 (notice the histogram for the Unbiased method is noticeably sharper). Figure 2.9(a) plots the standard deviation of the Jacobian field inside the corpus callosum as a function of iteration number. For the viscous fluid model, the standard deviation increased with the number of iterations, since the grid became less regular. On the other hand, the Unbiased method yielded an optimized standard deviation as more iterations were computed. The proposed symmetric KL distance also increased for viscous fluid method, while it was minimized for the Unbiased method as shown in Figure 2.9(b).

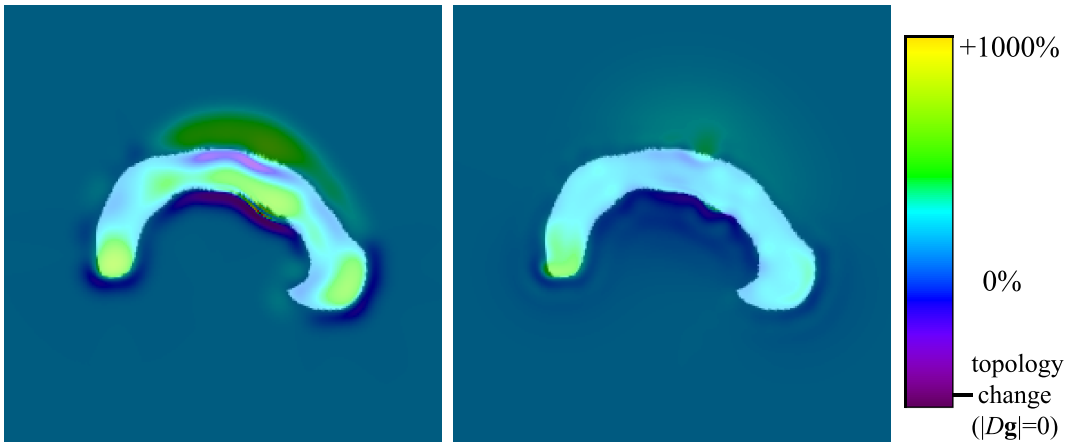



Figure 2.7: Corpus callosum example. Jacobian map of the deformation is superimposed with the deformed image for the viscous fluid model and the Unbiased model.

### 2.3.3 2D Serial MRI Example

In Figures 2.10 through 2.15, 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 equation (2.20)), where visually significant ventricle enlargement is present. Both the viscous fluid method and the Unbiased model generated a close match between the deformed source and target images (Figure 2.10(a-d)). Here, there is no reason not to evenly distribute the Jacobian field inside the ventricles, as realized using the Unbiased method. In contrast, the viscous fluid method generated a density map with extreme values along the ventricular boundary. 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 2.15, both the standard deviation inside the ventricle and the symmetric KL distance increased for the viscous fluid method, while these quantities stabilized for the Unbiased method.

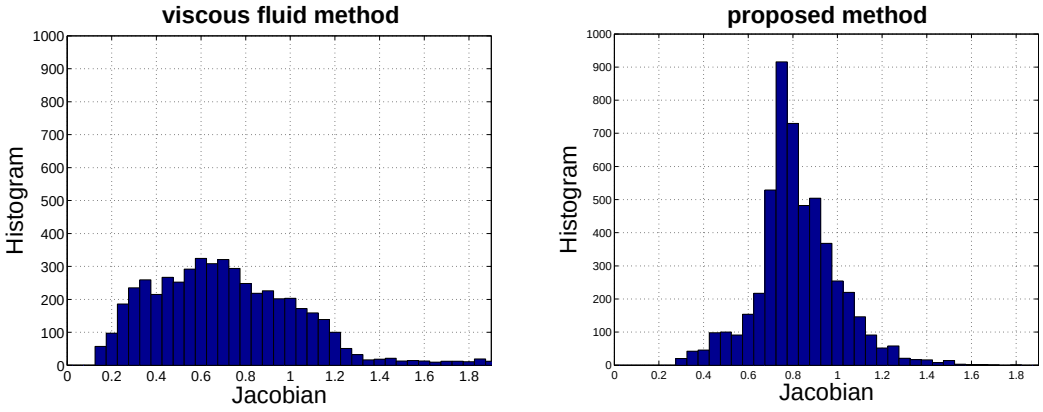



Figure 2.8: Corpus callosum example. Histograms of Jacobian values of the deformations inside corpus callosum for the viscous fluid model and the Unbiased model.

### 2.3.4 3D Serial MRI Example

In this numerical example (Figures 2.16, 2.17, 2.18, 2.19), we tested the Unbiased 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 equation (2.20). In Figure 2.18(a), the 3D Jacobian map generated using the viscous fluid 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 2.18(b), right temporal atrophy (RT) and ventricular enlargement (V) are easily visualized in the Jacobian map generated using the Unbiased method, demonstrating its advantages when recovering voxel-wise maps of local tissue change.

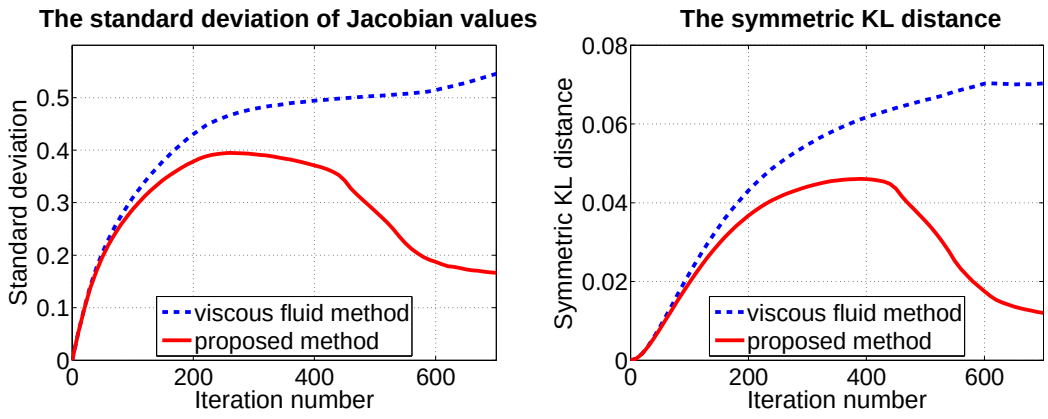
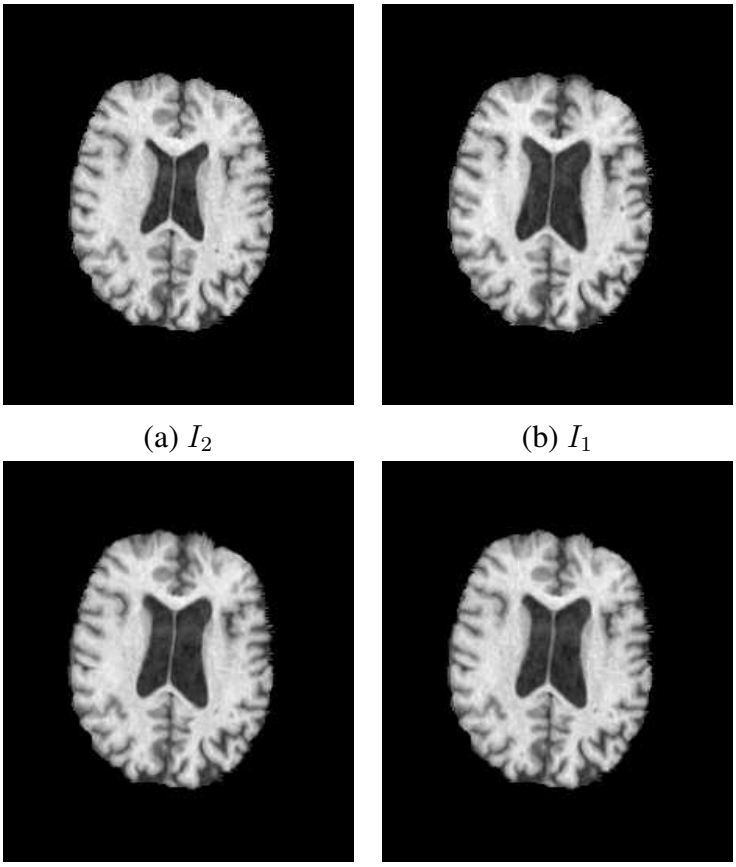



Figure 2.9: Corpus callosum example. (a) Standard deviation of Jacobian values inside corpus callosum per iteration. (b) Symmetric KL distance. For the viscous fluid model (dashed blue), both standard deviation and symmetric KL distance increase while for the Unbiased model (solid red), both standard deviation and symmetric KL distance decrease and stabilize.

### 2.3.5 Multimodal Image Matching Example

In this example, we employ mutual information (2.19) as a similarity measure in equation (2.17). In order to observe the robustness of mutual information matching and also to gain more insight into the effect of the Unbiased regularization, in Figures 2.20 and 2.21 we consider matching pairs of 2D binary slices from a set of serial MRI images (each of size 226 by 256), where visually significant ventricle enlargement is present. In both figures, source and target images are of different contrast, making the sum of squared intensity differences inapplicable as a choice of a data fidelity term. Both viscous fluid registration and Unbiased registration generated a close match between the deformed image and the target image (Figures 2.20(c,d) and 2.21(c,d)). Figures 2.20(e,f) and 2.21(e,f) show Jacobian maps of deformations. Here, there is no reason not to evenly distribute the Jacobian field inside the ventricles, as realized using the Unbiased method. In contrast, the viscous fluid method generated noisy mean

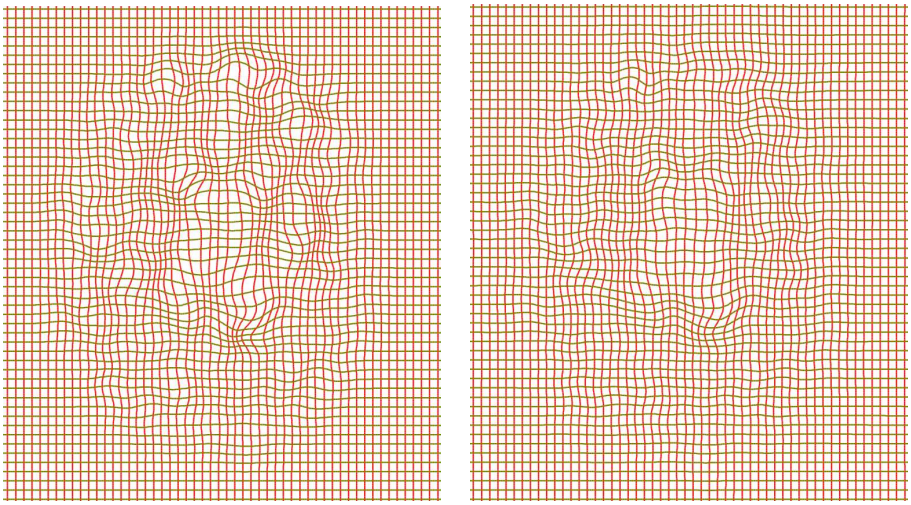


(a) $I_2$ (b) $I_1$

(c) $I_2 \circ \mathbf{g}$, Fluid (d) $I_2 \circ \mathbf{g}$, Unbiased

Figure 2.10: Serial MRI example. (a) image $I_2$; (b) image $I_1$; (c) image $I_2$ is deformed to image $I_1$ using the viscous fluid model; (d) image $I_2$ is deformed to image $I_1$ using the Unbiased model.

Jacobian maps with extreme values along the ventricular boundary. Indeed, given the overall longitudinal ventricular dilatation, we argue that the corresponding density change map should be constant inside the ventricle.



Fluid Unbiased

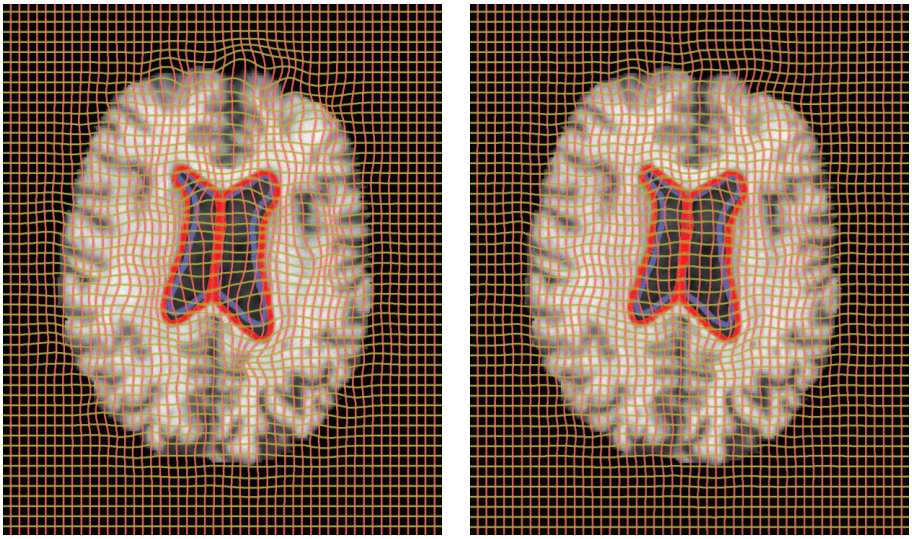
Figure 2.11: Serial MRI example. Results obtained with the viscous fluid model and the Unbiased model. The resulting grid of the Unbiased method is visually more regular.

## 2.4 Appendix: Derivation of equations for maximization of Mutual Information

In this Appendix, we derive the gradient $\partial_{\mathbf{u}} F_{MI}(\mathbf{u})$ of the mutual information matching functional in (2.19), adopting the approach of [16, 47], modeling the joint intensity distribution $p^{I_1,I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_1, i_2)$ of deformed image $I_2(\mathbf{x}-\mathbf{u})$ and image $I_1(\mathbf{x})$ as a continuous function using the Parzen windowing method. We refer the reader to Section 1.3.3 for the notations used in this Appendix.

We compute the first variation of $F_{MI}(\mathbf{u})$ by perturbing $\mathbf{u}$ in the following way

$$\frac{dF_{MI}(\mathbf{u}+\varepsilon\boldsymbol{\eta})}{d\varepsilon} = -\frac{d}{d\varepsilon}\int_{\mathbb{R}^2} p^{I_1,I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_1,i_2)\,\log\frac{p^{I_1,I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_1,i_2)}{p^{I_1}(i_1)p^{I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_2)}\,di_1di_2. \tag{2.30}$$



Fluid Unbiased

Figure 2.12: Serial MRI example. Results obtained with the viscous fluid model and the Unbiased model. The generated grids are superimposed with the deformed images. Yellow, blue, and red contours represent the boundaries of ventricles in $I_1$, $I_2$, and deformed $I_2$, respectively. Note that for both methods, yellow contour is essentially invisible due to a very close match.

Thus, we have

$$\begin{aligned}\frac{dF_{MI}(\mathbf{u}+\varepsilon\boldsymbol{\eta})}{d\varepsilon} = &-\int_{\mathbb{R}^2}\left(1+\log\frac{p^{I_1,I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_1,i_2)}{p^{I_1}(i_1)p^{I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_2)}\right)\frac{dp^{I_1,I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_1,i_2)}{d\varepsilon}\,di_1di_2\\ &+\int_{\mathbb{R}^2}\frac{p^{I_1,I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_1,i_2)}{p^{I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_2)}\frac{dp^{I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_2)}{d\varepsilon}\,di_1di_2.\end{aligned} \tag{2.31}$$

However, note that

$$\int_{\mathbb{R}} p^{I_1,I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_1,i_2)\,di_1 = p^{I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_2) \tag{2.32}$$

and

$$\int_{\mathbb{R}} p^{I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_2)\,di_2 = 1. \tag{2.33}$$

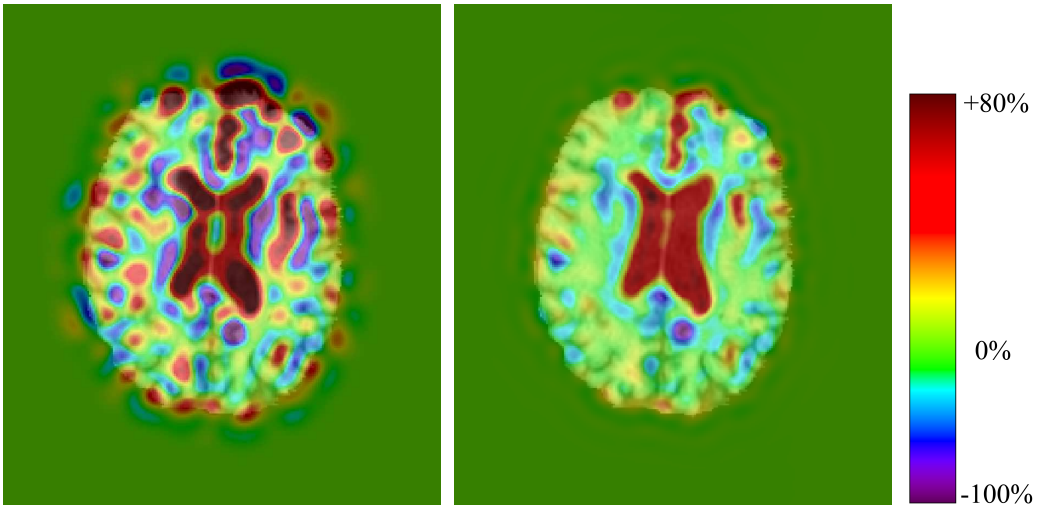



Figure 2.13: Serial MRI example. Jacobian map of the deformation is superimposed with the deformed image for the viscous fluid model and the Unbiased model.

Hence, the second term on the right hand side of the equality in (2.31) reduces to

$$\begin{aligned}
&\int_{\mathbb{R}^2} \frac{p^{I_1,I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_1,i_2)}{p^{I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_2)} \frac{dp^{I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_2)}{d\varepsilon}\, di_1 di_2 \\
&= \int_{\mathbb{R}} \frac{dp^{I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_2)}{d\varepsilon} \frac{1}{p^{I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_2)} \Big( \int_{\mathbb{R}} p^{I_1,I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_1,i_2)\, di_1 \Big)\, di_2 \\
&= \int_{\mathbb{R}} \frac{dp^{I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_2)}{d\varepsilon} \frac{1}{p^{I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_2)} p^{I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_2)\, di_2 \\
&= \frac{d}{d\varepsilon} \int_{\mathbb{R}} p^{I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_2)\, di_2 = 0.
\end{aligned} \tag{2.34}$$

Equation (2.30) becomes

$$\frac{dF_{MI}(\mathbf{u}+\varepsilon\boldsymbol{\eta})}{d\varepsilon} = -\int_{\mathbb{R}^2} \left[ \left( 1 + \log \frac{p^{I_1,I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_1,i_2)}{p^{I_1}(i_1) p^{I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_2)} \right) \frac{dp^{I_1,I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_1,i_2)}{d\varepsilon} \right] di_1 di_2. \tag{2.35}$$

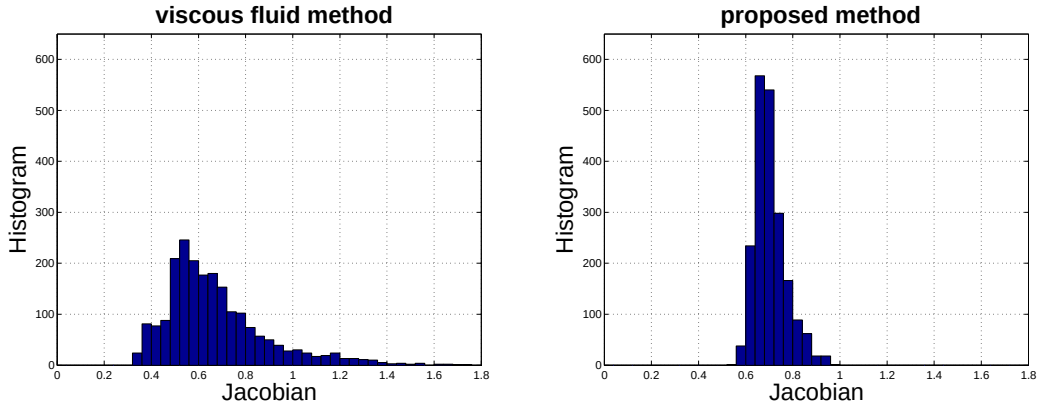



Figure 2.14: Serial MRI example. Histograms of Jacobian values of the deformations inside ventricles for the viscous fluid model and the Unbiased model.

The joint intensity distribution estimated from $I_2(\mathbf{x}-\mathbf{u})$ and $I_1(\mathbf{x})$ is given by

$$p^{I_1,I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_1,i_2)=\frac{1}{|\Omega|}\int_\Omega \psi\Big(I_1\big(\mathbf{x}\big)-i_1, I_2\big(\mathbf{x}-\mathbf{u}(\mathbf{x})-\varepsilon\boldsymbol{\eta}(\mathbf{x})\big)-i_2\Big)\,d\mathbf{x}, \tag{2.36}$$

where $|\Omega|$ is a volume of $\Omega$ and $\psi(\xi_1,\xi_2)$ is a two-dimensional Parzen windowing kernel.

The derivative of (2.36) can also be computed:

$$\begin{aligned}\frac{dp^{I_1,I_2}_{\mathbf{u}+\varepsilon\boldsymbol{\eta}}(i_1,i_2)}{d\varepsilon}=-\frac{1}{|\Omega|}\int_\Omega \frac{\partial\psi}{\partial\xi_2}\Big(I_1\big(\mathbf{x}\big)-i_1, I_2\big(\mathbf{x}-\mathbf{u}(\mathbf{x})-\varepsilon\boldsymbol{\eta}(\mathbf{x})\big)-i_2\Big)\\ \times\nabla I_2\big(\mathbf{x}-\mathbf{u}(\mathbf{x})-\varepsilon\boldsymbol{\eta}(\mathbf{x})\big)\cdot\boldsymbol{\eta}(\mathbf{x})\,d\mathbf{x}.\end{aligned} \tag{2.37}$$

Let us denote

$$Q_{\mathbf{u}}(i_1,i_2)=1+\log\frac{p^{I_1,I_2}_{\mathbf{u}}(i_1,i_2)}{p^{I_1}(i_1)p^{I_2}_{\mathbf{u}}(i_2)}. \tag{2.38}$$

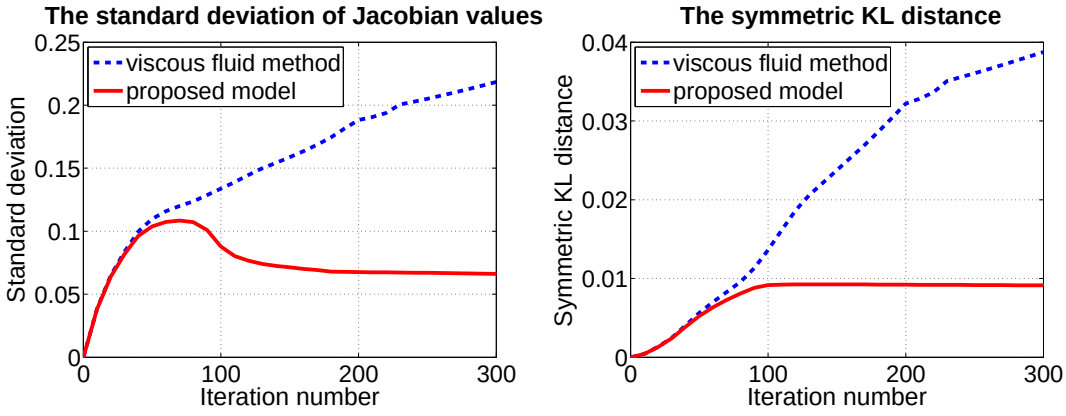



Figure 2.15: Serial MRI example. (a) Standard deviation of Jacobian values inside the ventricle per iteration. (b) Symmetric KL distance. For the viscous fluid model (dashed blue), both standard deviation and symmetric KL distance increase while for the Unbiased model (solid red), both standard deviation and symmetric KL distance stabilize.

If we let $\varepsilon = 0$, equation (2.35) gives the first variation of $F_{MI}(\mathbf{u})$:

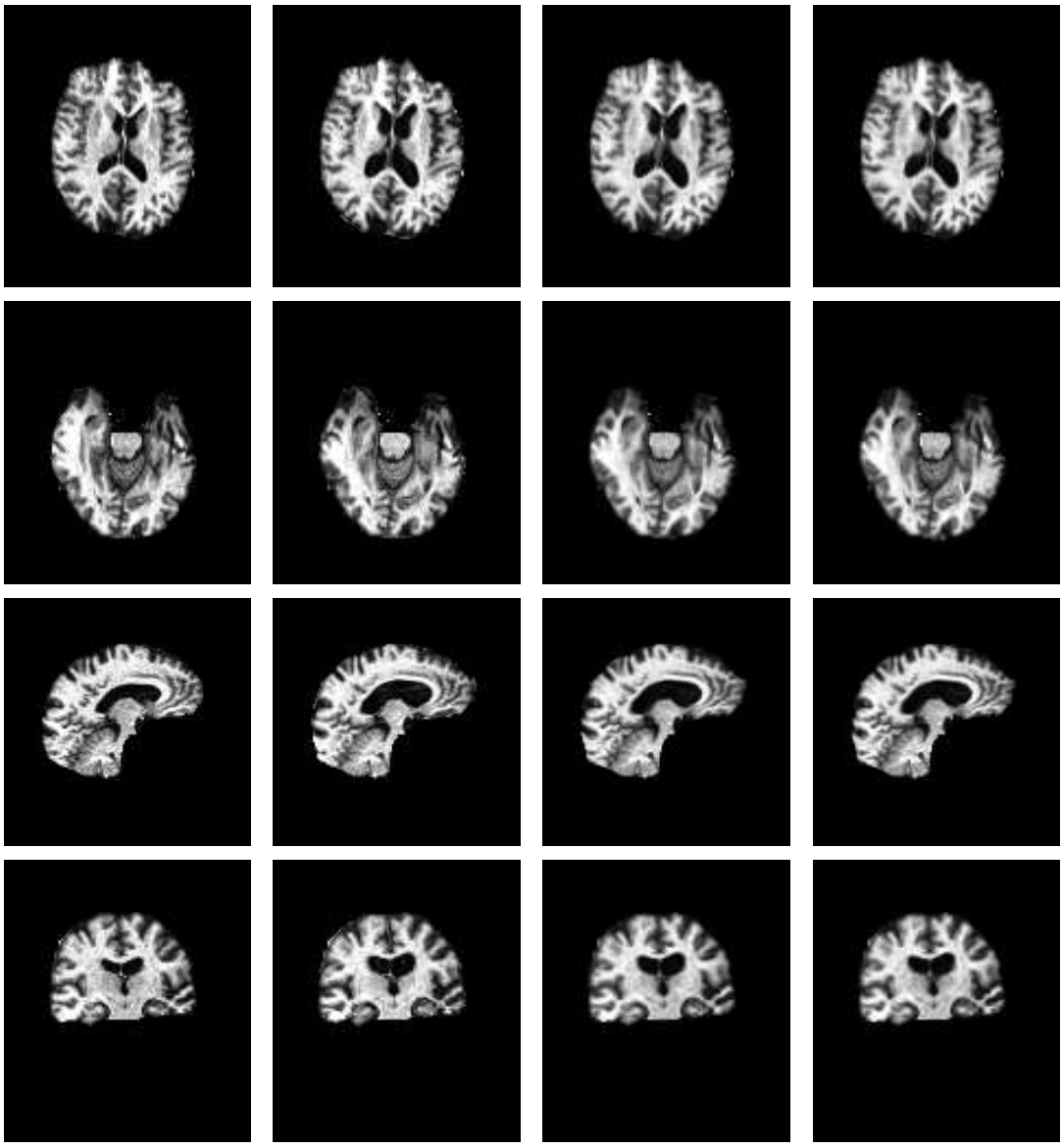
$$\begin{aligned}
&\frac{dF_{MI}(\mathbf{u}+\varepsilon\boldsymbol{\eta})}{d\varepsilon}\bigg|_{\varepsilon=0} \\
&= \int_{\mathbb{R}^2} Q_{\mathbf{u}}(i_1,i_2)\frac{1}{|\Omega|}\int_{\Omega}\frac{\partial\psi}{\partial\xi_2}\big(I_1(\mathbf{x})-i_1, I_2(\mathbf{x}-\mathbf{u}(\mathbf{x}))-i_2\big) \\
&\qquad \times\nabla I_2(\mathbf{x}-\mathbf{u}(\mathbf{x}))\cdot\boldsymbol{\eta}(\mathbf{x})\,d\mathbf{x}\,di_1di_2 \\
&= \frac{1}{|\Omega|}\int_{\Omega}\left[Q_{\mathbf{u}} * \frac{\partial\psi}{\partial\xi_2}\right]\big(I_1(\mathbf{x}), I_2(\mathbf{x}-\mathbf{u}(\mathbf{x}))\big) \\
&\times \nabla I_2(\mathbf{x}-\mathbf{u}(\mathbf{x}))\cdot\boldsymbol{\eta}(\mathbf{x})\,d\mathbf{x}.
\end{aligned} \tag{2.39}$$

Here, $*$ denotes a convolution. Thus,

$$\partial_{\mathbf{u}}F_{MI}(\mathbf{u}) = \frac{1}{|\Omega|}\left[Q_{\mathbf{u}} * \frac{\partial\psi}{\partial\xi_2}\right](I_1(\mathbf{x}), I_2(\mathbf{x}-\mathbf{u}))\nabla I_2(\mathbf{x}-\mathbf{u}). \tag{2.40}$$

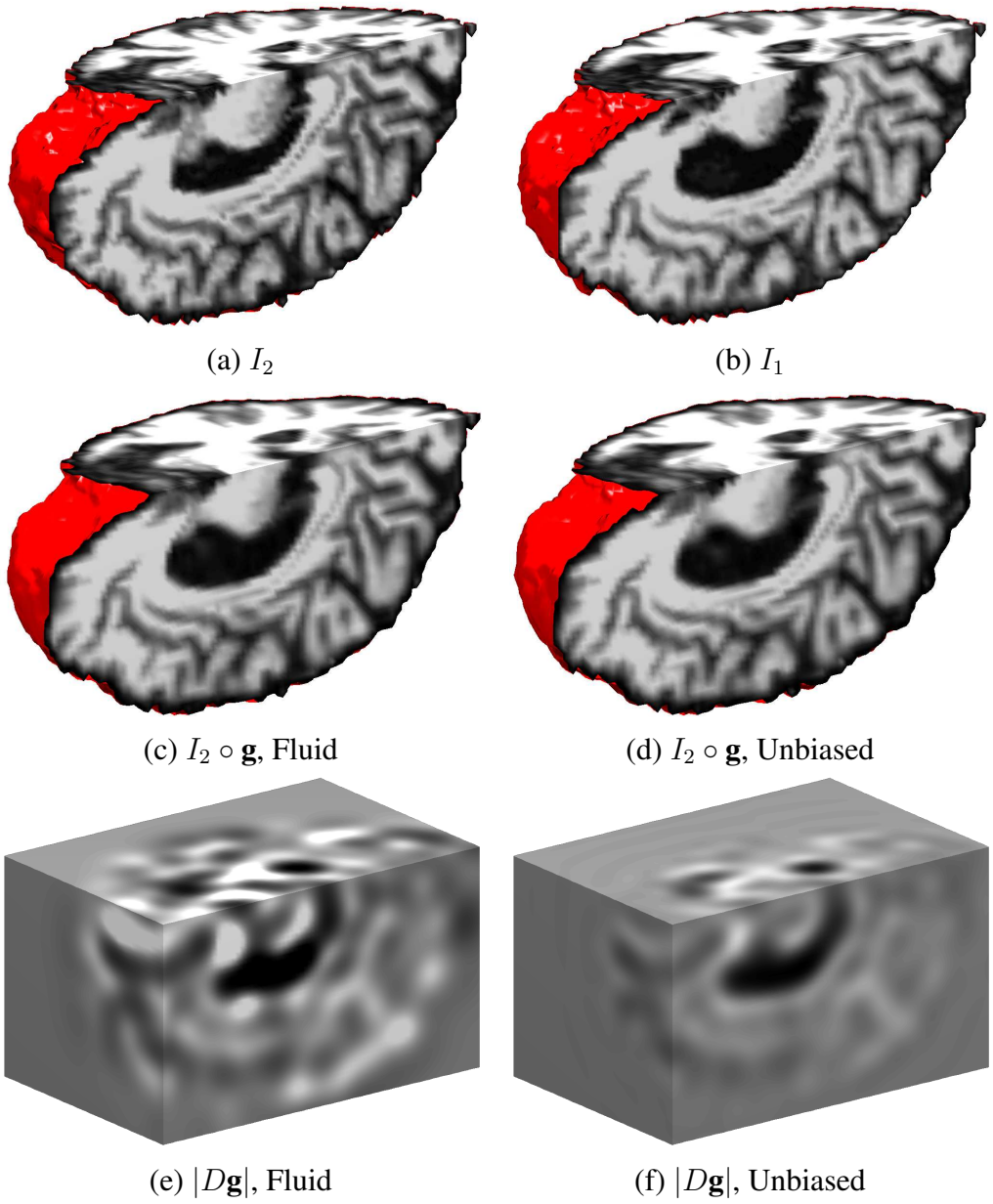
## 2.5 Conclusion

In this chapter, we applied information theory to quantify the magnitude of deformations. We examined the statistical distributions of Jacobian maps in the logarithmic space, and developed the unbiased framework for constructing image registration methods. The proposed framework yields both theoretically and intuitively correct deformation maps, and is compatible with large-deformation models. To demonstrate the power of the proposed framework, we generalized the well known large-deformation viscous fluid registration model to compute unbiased deformations. We showed that unbiased fluid registration method generates more accurate maps compared to those generated with the viscous fluid registration model.



(a) $I_2$ (time 1) (b) $I_1$ (time 3) (c) $I_2 \circ \mathbf{g}$, Fluid (d) $I_2 \circ \mathbf{g}$, Unbiased

Figure 2.16: 3D Serial MRI example. Rows depict slices in axial (rows 1 and 2), sagittal (row 3), and coronal (row 4) planes. Columns depict (a) $I_2$ (time 1); (b) $I_1$ (time 3); (c) $I_2$ deformed using the viscous fluid model; (d) $I_2$ deformed using the Unbiased model. The middle time point (time 2) is not shown.



(a) $I_2$ (b) $I_1$

(c) $I_2 \circ \mathbf{g}$, Fluid (d) $I_2 \circ \mathbf{g}$, Unbiased

(e) $|D\mathbf{g}|$, Fluid (f) $|D\mathbf{g}|$, Unbiased

Figure 2.17: 3D Serial MRI example. Volume cuts of (a) $I_2$ (time 1), (b) $I_1$ (time 3), (c) $I_2$ deformed using the viscous fluid model, and (d) $I_2$ deformed using the Unbiased model. The middle time point (time 2) is not shown. Volume cuts of Jacobian maps of deformations (time 1 to time 3) for (e) the viscous fluid model and (f) the Unbiased model. Jacobian maps of deformations from time 2 to time 3 are not shown.

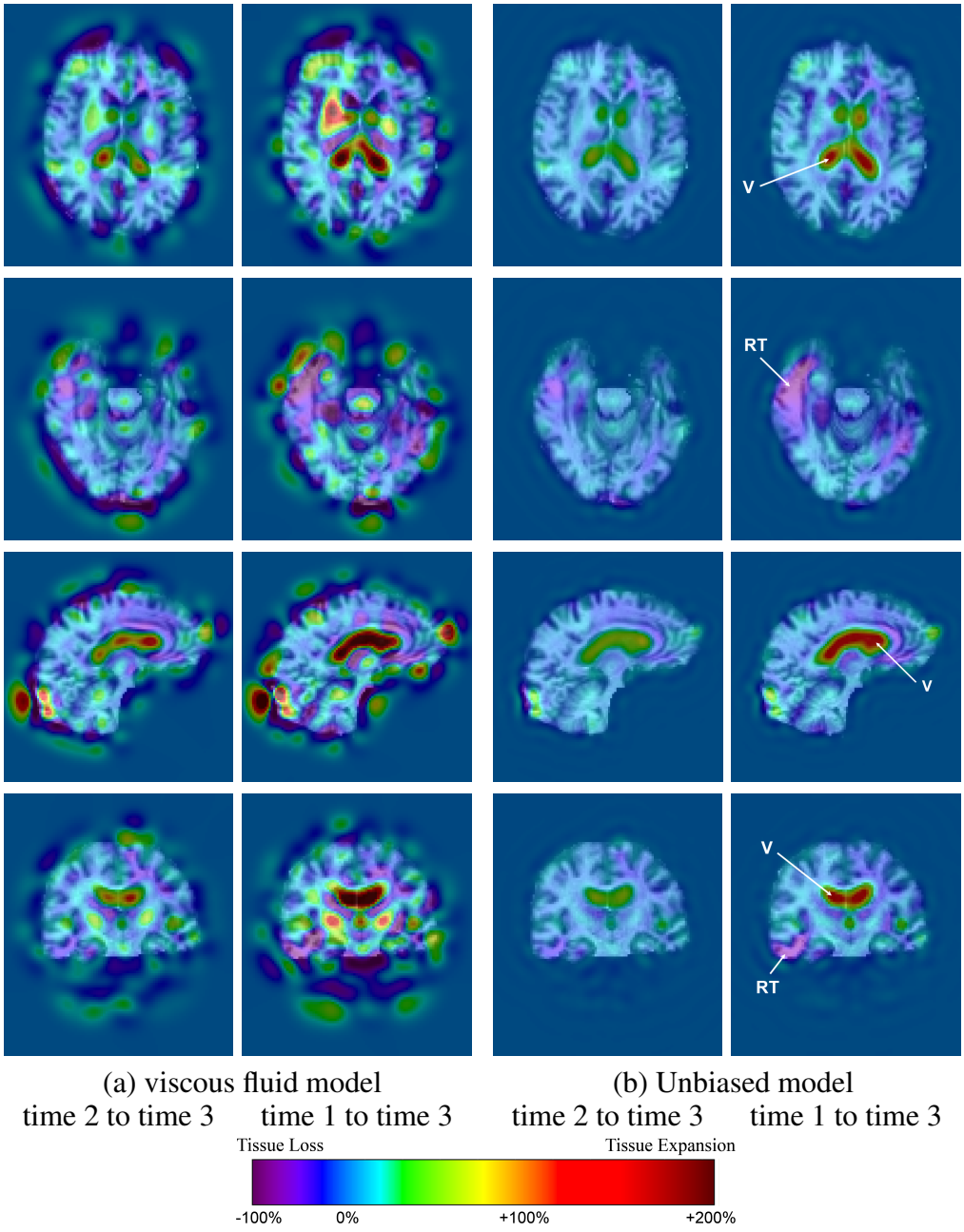



Figure 2.18: 3D Serial MRI example. Jacobian maps are superimposed with the deformed volumes for the viscous fluid model (columns 1 and 2) and the Unbiased model (columns 3 and 4). Smaller deformations (time 2 to time 3) and larger deformations (time 1 to time 3) are shown. Rows depict slices in axial (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 Unbiased method, while the viscous fluid method generated a very noisy map.

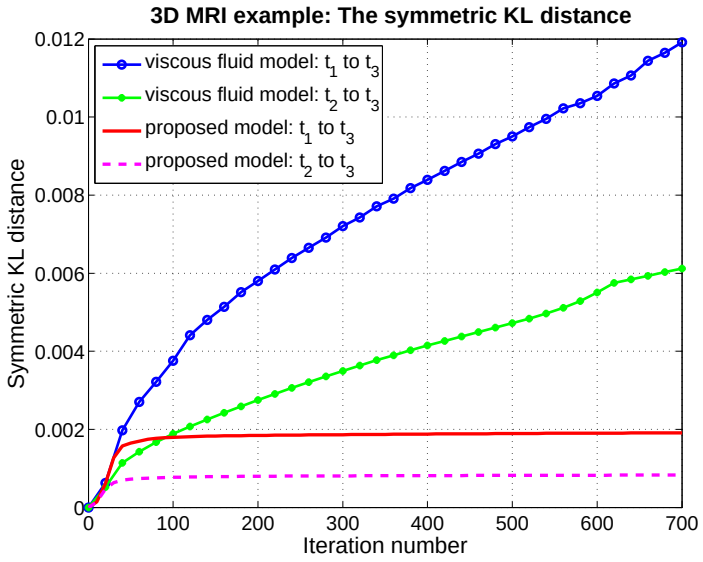
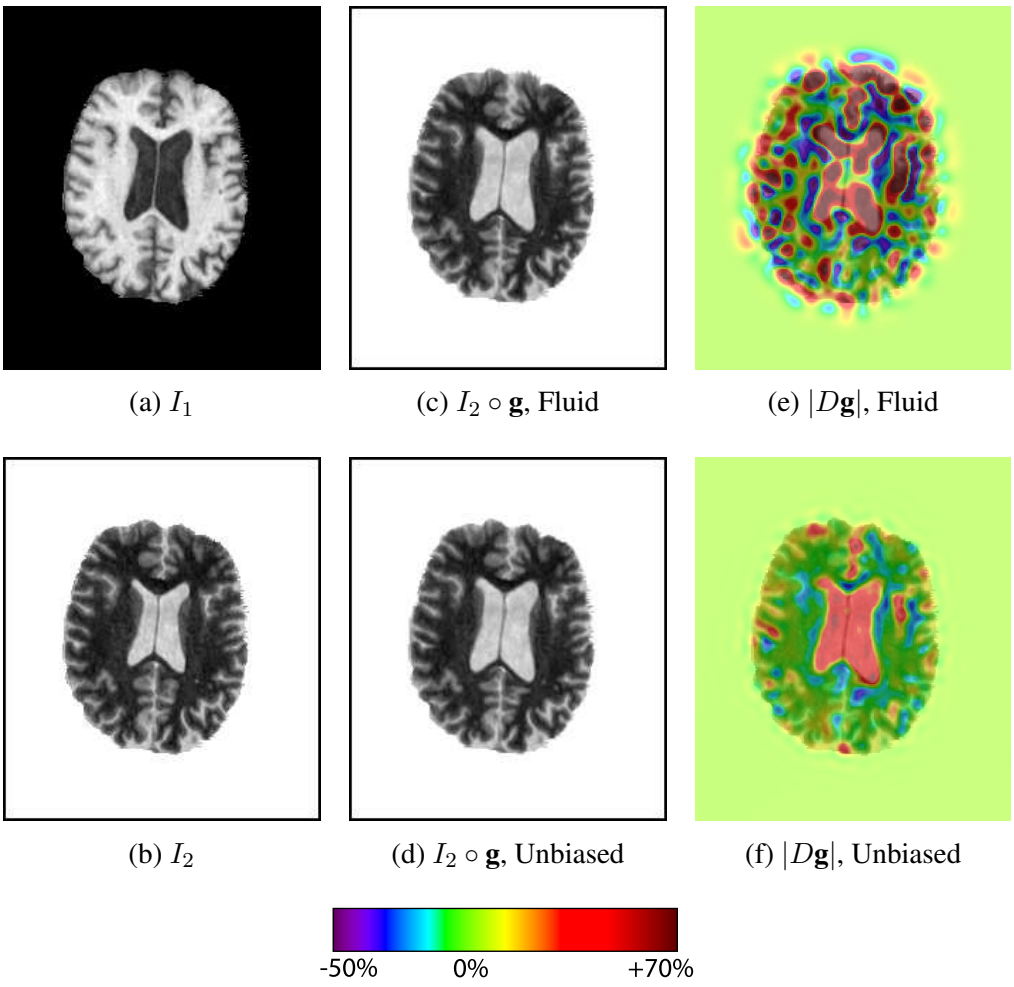



Figure 2.19: 3D Serial MRI example. Symmetric KL distance is shown for the viscous fluid and Unbiased models for a larger deformation (time 1 to time 3) and a smaller deformation (time 2 to time 3). Note that this measure is proportional to the magnitude of deformation. For the proposed method, the symmetric KL distance stabilizes.




(a) $I_1$ (c) $I_2 \circ \mathbf{g}$, Fluid (e) $|D\mathbf{g}|$, Fluid

(b) $I_2$ (d) $I_2 \circ \mathbf{g}$, Unbiased (f) $|D\mathbf{g}|$, Unbiased

Figure 2.20: Serial MRI example. (a) image $I_1$; (b) image $I_2$; (c) image $I_2$ is deformed to image $I_1$ using fluid registration coupled with mutual information; (d) image $I_2$ is deformed to image $I_1$ using the Unbiased registration coupled with mutual information. Jacobian map of the deformation is superimposed with the deformed image for (e) the viscous fluid model and (f) the Unbiased model.

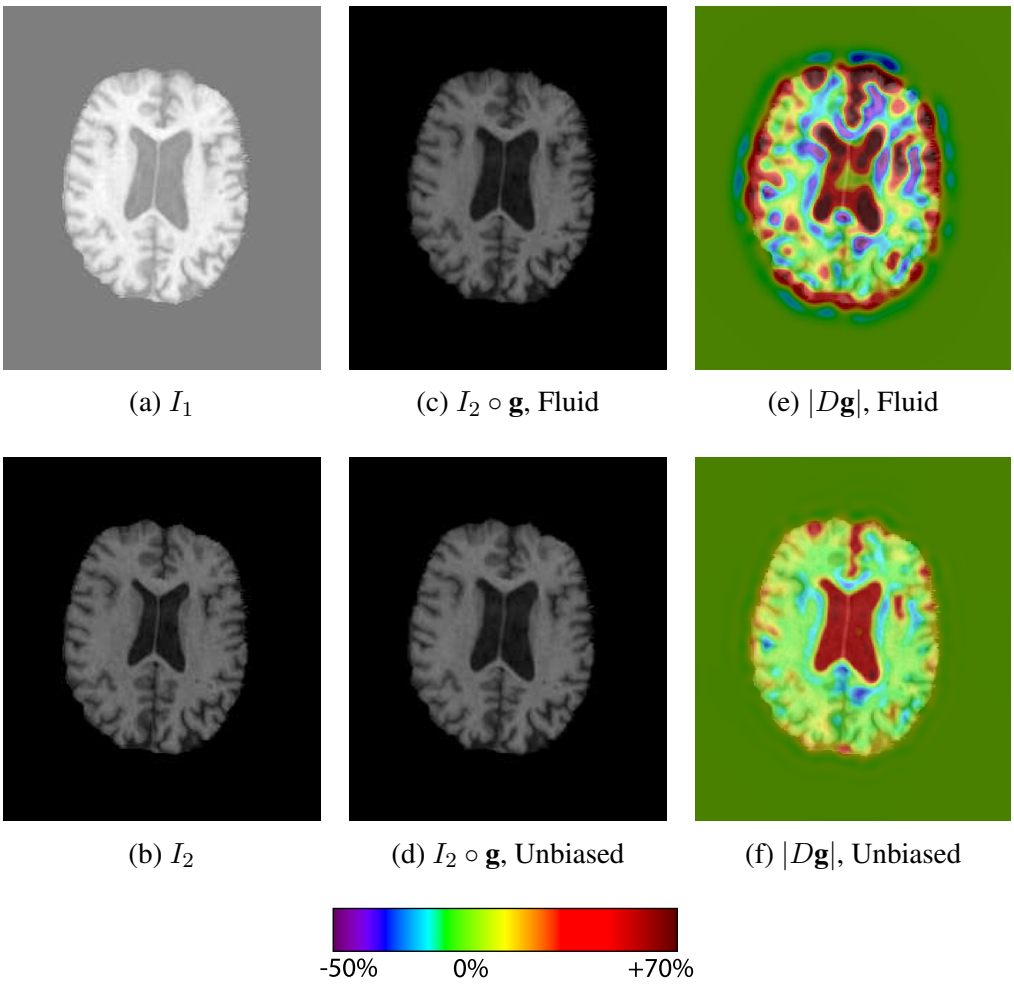



Figure 2.21: Serial MRI example. (a) image $I_1$; (b) image $I_2$; (c) image $I_2$ is deformed to image $I_1$ using fluid registration coupled with mutual information; (d) image $I_2$ is deformed to image $I_1$ using the Unbiased registration coupled with mutual information. Jacobian map of the deformation is superimposed with the deformed image for (e) the viscous fluid model and (f) the Unbiased model.

# CHAPTER 3

# Unbiased Registration via Nonlinear Elastic Regularization

In this chapter, we propose a new large-deformation nonlinear image registration model, based on nonlinear elastic regularization and unbiased registration. In contrast to the proposed unbiased fluid registration method, which was described in Chapter 2, the new model is written in a unified variational form. Both the nonlinear elastic and the unbiased functionals are simplified introducing, in the modeling, a second unknown that mimics the Jacobian matrix of the displacement vector field, removing the nonlinearity in the derivatives of the Euler-Lagrange equations. The energy functional is minimized using gradient descent and is easy to implement. The proposed numerical scheme is less sensitive to numerical constraints such as CFL condition. As a result, the new unbiased nonlinear elasticity model is computationally efficient. In this chapter, we compare the results obtained with the new unbiased nonlinear elasticity model to those obtained with unbiased fluid registration.

## 3.1 Introduction

Given two images, the source and target, the goal of image registration is to find an optimal diffeomorphic spatial transformation such that the deformed source image is aligned with the target image. In the case of non-parametric registration methods, the

problem can be formulated as a functional minimization problem whose unknown is the displacement vector field $\mathbf{u}$. Usually, the devised functional consists of a distance measure (intensity-based, correlation-based, mutual-information based [73] or metric-structure-comparison based [67]) and a regularizer that guarantees smoothness of the displacement vector field. Generally, physical arguments motivate the selection of the regularizer. Among those currently used is the linear elasticity smoother first introduced by Broit [8]. The objects to be registered are considered to be observations of the same elastic body at two different times, before and after being subjected to a deformation as mentioned in [73]. The smoother, in this case, is the linearized elastic potential of the displacement vector field. However, this model is unsuitable for problems involving large-magnitude deformations.

In [21], the authors proposed a viscous fluid model to overcome this issue (see Section 1.4.5). The deforming image is considered to be embedded in viscous fluid whose motion is governed by Navier-Stokes equations for conservation of momentum. One drawback of the viscous fluid method is the computational cost. Numerically, the image-derived force field $\mathbf{f}(\mathbf{x}, \mathbf{u}(\mathbf{x}, t))$ is first computed at time $t$. Fixing the force field $\mathbf{f}$, linear equation (1.34) is solved for $\mathbf{v}(\mathbf{x}, t)$ numerically using the successive over-relaxation (SOR) scheme. Then, an explicit Euler scheme is used to advance $\mathbf{u}$ in time. Recent works [10, 84, 85] applied Riemannian nonlinear elasticity priors to deformation velocity fields. These alternating frameworks, however, are time-consuming, which motivates the search for faster implementations (see for instance [7] or [28] in which the instantaneous velocity $\mathbf{v}$ is obtained by convolving $\mathbf{f}$ with a Gaussian kernel).

In this work, which is inspired from related works on segmentation [58] and on two-dimensional slice registration [64], we propose a novel large-deformation image registration approach. In contrast to the unbiased fluid registration model of Chapter 2, the proposed model is derived from a variational problem which is not in the form

of a two-step algorithm. For that purpose, a nonlinear elasticity regularization is introduced. As will be seen later, the computation of the Euler-Lagrange equations in this case is cumbersome. We circumvent this issue by introducing a second unknown, a matrix variable $V$, which approximates the Jacobian matrix of $\mathbf{u}$. The nonlinear elastic regularizer is now applied to $V$. The Euler-Lagrange equations are straightforwardly derived and a gradient descent method is used.

Here we use an information-theoretic approach previously introduced in [112] and described in Chapter 2. In [112], we considered a smooth deformation $\mathbf{g}$ that maps domain $\Omega$ bijectively onto itself. Consequently, $\mathbf{g}$ and $\mathbf{g}^{-1}$ are bijective and globally volume-preserving. Probability density functions can thus be associated with the deformation $\mathbf{g}$ and its inverse $\mathbf{g}^{-1}$. We proposed to quantify the magnitude of the deformation by means of computing the symmetric Kullback-Leibler distance between the probability density functions associated with the deformation and the identity mapping. In Chapter 2, the symmetric Kullback-Leibler distance, when rewritten using skew-symmetry properties, was viewed as a cost function and was combined with the viscous fluid model for registration, which lead to the unbiased fluid registration model. Unlike the unbiased fluid registration model, the unbiased nonlinear elasticity method, introduced here, allows the functional to be written "in closed form". The new model also does not require expensive Navier-Stokes solver (or its approximation) at each step as previously mentioned.

## 3.2 Method

Let $\Omega$ be an open and bounded domain in $\mathbb{R}^3$. Without loss of generality, we assume that the volume of $\Omega$ is 1, i.e. $|\Omega| = 1$. Let $I_1, I_2 : \Omega \to \mathbb{R}$ be the two images to be registered. We seek the transformation $\mathbf{g} : \Omega \to \Omega$ that maps the source image $I_2$ into correspondence with the target image $I_1$. As in Chapter 2, we will restrict this

mapping to be differentiable, one-to-one, and onto. We denote the Jacobian matrix of a deformation $\mathbf{g}$ to be $D\mathbf{g}$, with Jacobian denoted by $|D\mathbf{g}(\mathbf{x})| := \det(D\mathbf{g}(\mathbf{x}))$ (thus we will use the notation $|V| := \det(V)$ for any $3 \times 3$ matrix $V$). The displacement field $\mathbf{u}(\mathbf{x})$ from the position $\mathbf{x}$ in the deformed image $I_2 \circ \mathbf{g}(\mathbf{x})$ back to $I_2(\mathbf{x})$ is defined in terms of the deformation $\mathbf{g}(\mathbf{x})$ by the expression $\mathbf{g}(\mathbf{x}) = \mathbf{x} - \mathbf{u}(\mathbf{x})$ at every point $\mathbf{x} \in \Omega$. Thus, we consider the problems of finding $\mathbf{g}$ and $\mathbf{u}$ as equivalent.

In general, nonlinear image registration models may be formulated in a variational framework. The minimization problems often define the energy functional $E$ as a linear combination of image matching term $F$ and the regularizing term $R$

$$\inf_{\mathbf{u}} \big\{ E(\mathbf{u}) = F(\mathbf{u}) + \lambda_0 R(\mathbf{u}) \big\}.$$

Here, $\lambda_0 > 0$ is a weighting parameter.

### 3.2.1 Registration metrics

In this chapter, the matching functional $F$ takes the form of the $L^2$ norm (the sum of squared intensity differences), $F = F_{L^2}$, and the mutual information, $F = F_{MI}$.

***L*$^2$-norm:** The $L^2$-norm matching functional is suitable when the images have been acquired through similar sensors (with additive Gaussian noise) and thus are expected to present the same intensity range and distribution. The $L^2$ distance between the deformed image $I_2 \circ \mathbf{g}(\mathbf{x}) = I_2(\mathbf{x} - \mathbf{u}(\mathbf{x}))$ and target image $I_1(\mathbf{x})$ is defined as

$$F_{L^2}(\mathbf{u}) = \frac{1}{2} \int_\Omega \big( I_2(\mathbf{x} - \mathbf{u}(\mathbf{x})) - I_1(\mathbf{x}) \big)^2 \, d\mathbf{x}. \tag{3.1}$$

**Mutual Information:** Mutual information can be used to align images of different modalities, without requiring knowledge of the relationship of the two registered images. Here, the intensity distributions estimated from $I_1(\mathbf{x})$ and $I_2(\mathbf{x} - \mathbf{u}(\mathbf{x}))$ are de-

noted by $p^{I_1}$ and $p^{I_2}_{\mathbf{u}}$, respectively, and an estimate of their joint intensity distribution by $p^{I_1,I_2}_{\mathbf{u}}$. We let $i_1 = I_1(\mathbf{x})$, $i_2 = I_2(\mathbf{x} - \mathbf{u}(\mathbf{x}))$ denote intensity values at point $\mathbf{x} \in \Omega$. Given the displacement field $\mathbf{u}$, the mutual information computed from $I_1$ and $I_2$ is provided by $MI^{I_1,I_2}_{\mathbf{u}} = \int_{\mathbb{R}^2} p^{I_1,I_2}_{\mathbf{u}}(i_1, i_2)\, \log[p^{I_1,I_2}_{\mathbf{u}}(i_1, i_2)/(p^{I_1}(i_1)p^{I_2}_{\mathbf{u}}(i_2))]\, di_1 di_2$. We seek to maximize the mutual information between $I_2(\mathbf{x} - \mathbf{u}(\mathbf{x}))$ and $I_1(\mathbf{x})$, or equivalently, minimize the negative of $MI^{I_1,I_2}_{\mathbf{u}}$:

$$F_{MI}(I_1, I_2, \mathbf{u}) = -MI^{I_1,I_2}_{\mathbf{u}}. \tag{3.2}$$

### 3.2.2 Nonlinear Elastic Regularization

The theory of elasticity is based on the notion of strain. Strain is defined as the amount of deformation an object experiences compared to its original size and shape. In three spatial dimensions, the strain tensor, $\mathcal{E} = [\varepsilon_{ij}] \in \mathbb{R}^{3\times3}$, $1 \leq i, j \leq 3$, is a symmetric tensor used to quantify the strain of an object undergoing a deformation. The nonlinear strain is defined as

$$\varepsilon_{ij}(\mathbf{u}) = \frac{1}{2}\big(\partial_j u_i + \partial_i u_j + \sum_{k=1}^{3} \partial_i u_k \partial_j u_k\big),$$

with the nonlinear strain tensor matrix given by

$$\mathcal{E}(\mathbf{u}) = \frac{1}{2}\big(D\mathbf{u}^t + D\mathbf{u} + D\mathbf{u}^t D\mathbf{u}\big). \tag{3.3}$$

For Saint Venant-Kirchhoff material, the stored energy is defined as

$$W(\mathcal{E}) = \frac{\nu}{2}(\text{trace}(\mathcal{E}))^2 + \mu\text{trace}(\mathcal{E}^2),$$

where $\nu$ and $\mu$ are Lamé elastic material constants. The regularization for nonlinear elasticity becomes

$$R_E(\mathbf{u}) = \int_\Omega W(\mathcal{E}(\mathbf{u}))d\mathbf{x}.$$

The regularization term $R_E(\mathbf{u})$ can be minimized with respect to $\mathbf{u}$. However, since the regularization term is written in terms of partial derivatives of components of $\mathbf{u}$, the Euler-Lagrange equations become complicated and are computationally expensive to minimize. Instead, following earlier theoretical work [76], we minimize an approximate functional by introducing the matrix variable

$$V \approx D\mathbf{u} \tag{3.4}$$

and thus consider a new form of nonlinear elasticity regularization functional

$$R_E(\mathbf{u}, V) \;=\; \int_\Omega W(\widehat{V})\, d\mathbf{x} + \frac{\beta}{2}\int_\Omega ||V - D\mathbf{u}||_F^2\, d\mathbf{x}, \tag{3.5}$$

where $\widehat{V} = \frac{1}{2}\big(V^t + V + V^t V\big)$, $\beta$ is a positive constant, and $||\cdot||_F$ denotes the Frobenius norm. In the limit, as $\beta \to +\infty$, we obtain $V \approx D\mathbf{u}$ in the $L^2$ topology.

### 3.2.3 Unbiased Registration Constraint

In Chapter 2, we introduced an unbiased fluid image registration approach [112]. In this context, *unbiased* means that the Jacobian determinants of the deformations recovered between a pair of images follow a log-normal distribution, with zero mean after log-transformation. We argued that this distribution is beneficial when recovering change in regions of homogeneous intensity, and in ensuring symmetrical results when the order of two images being registered is switched. As derived in [112] using

information theory, the unbiased regularization term is given as

$$R_{SKL}(\mathbf{u}) = \int_{\Omega} \big(|D(\mathbf{x}-\mathbf{u}(\mathbf{x}))| - 1\big) \log |D(\mathbf{x}-\mathbf{u}(\mathbf{x}))| \, d\mathbf{x}. \tag{3.6}$$

It is important to note that $R_{SKL}$ generates inverse-consistent deformation maps. The inverse-consistent property of the unbiased technique is shown in a validation study of the unbiased fluid registration method in Chapter 4. Also, to see why minimizing equation (3.6) leads to unbiased deformation in the logarithmic space, we observe that the integrand is always non-negative, and only evaluates to zero when the deformation $\mathbf{g}$ is volume-preserving everywhere ($|D\mathbf{g}| = 1$ everywhere). Thus, by treating it as a cost, we recover zero-change by minimizing this cost when we compare images differing only in noise.

Given equation (3.4), we have $D\mathbf{g} = D(\mathbf{x}-\mathbf{u}) = \mathcal{I} - D\mathbf{u} \approx \mathcal{I} - V$, where $\mathcal{I}$ is the $3 \times 3$ identity matrix. Therefore, as in Section 3.2.2, in order to simplify the discretization of the gradient of the unbiased regularization term (3.6), we introduce

$$R_{SKL}(V) \;=\; \int_{\Omega} (|\mathcal{I} - V| - 1) \log |\mathcal{I} - V| \, d\mathbf{x}. \tag{3.7}$$

Recall that here $|\mathcal{I} - V| = \det(\mathcal{I} - V)$.

### 3.2.4 Unbiased Nonlinear Elasticity Registration

The total energy functional employed in this work, is given as a linear combination of the similarity measure $F$ (which is either $F_{L^2}$ from (3.1) or $F_{MI}$ from (3.2)), nonlinear elastic regularization $R_E$ in (3.5), and unbiased regularization $R_{SKL}$ in (3.7):

$$E(\mathbf{u}, V) \;=\; F(\mathbf{u}) + R_E(\mathbf{u}, V) + \lambda R_{SKL}(V). \tag{3.8}$$

or

$$\begin{aligned} E(\mathbf{u}, V) &= F(\mathbf{u}) + \int_\Omega W(\widehat{V}) d\mathbf{x} + \frac{\beta}{2} \int_\Omega ||V - D\mathbf{u}||_F^2 d\mathbf{x} \\ &\quad + \lambda \int_\Omega (|\mathcal{I} - V| - 1) \log |\mathcal{I} - V| \, d\mathbf{x}. \end{aligned}$$

The explicit weighting parameter is omitted in front of $R_E(\mathbf{u}, V)$, since this term is weighted by Lamé constants $\nu$ and $\mu$. Parameterizing the descent direction by an artificial time $t$, we solve the Euler-Lagrange equations in $\mathbf{u}(\mathbf{x}, t)$ and $V(\mathbf{x}, t)$ using the gradient descent method:

$$\frac{\partial \mathbf{u}}{\partial t} = -\partial E_{\mathbf{u}}(\mathbf{u}, V) = -\partial_{\mathbf{u}} F(\mathbf{u}) - \partial_{\mathbf{u}} R_E(\mathbf{u}, V), \tag{3.9}$$

$$\frac{\partial V}{\partial t} = -\partial E_V(\mathbf{u}, V) = -\partial_V R_E(\mathbf{u}, V) - \lambda \partial_V R_{SKL}(V), \tag{3.10}$$

which gives systems of three and nine equations, respectively. Explicit expressions for the gradients in equations (3.9) and (3.10) as well as the discretizations of these equations are given in Section 3.3.

*Remark:* The regularization on the deformation $\mathbf{g}$ proposed in this work can be expressed in a general form

$$R(\mathbf{g}) = \int_\Omega R_1(D\mathbf{g}) d\mathbf{x} + \int_\Omega R_2(|D\mathbf{g}|) d\mathbf{x},$$

with $|D\mathbf{g}| := \det(D\mathbf{g})$. For the minimization, an auxiliary variable can also be introduced to simplify the numerical calculations, removing the nonlinearity in the derivatives.

## 3.3 Implementation

### 3.3.1 The Energy Gradients

Computing the first variation of functional $F_{L^2}$ in equation (3.1) gives the following gradient

$$\partial_{\mathbf{u}} F_{L^2}(\mathbf{u}) = -[I_2(\mathbf{x}-\mathbf{u}(\mathbf{x})) - I_1(\mathbf{x})]\,\nabla I_2(\mathbf{x}-\mathbf{u}(\mathbf{x})).$$

The gradient of (3.2) is given by

$$\partial_{\mathbf{u}} F_{MI}(\mathbf{u}) = \frac{1}{|\Omega|}\left[Q_{\mathbf{u}} * \frac{\partial G_\sigma}{\partial \xi_2}\right](I_1(\mathbf{x}), I_2(\mathbf{x}-\mathbf{u}))\nabla I_2(\mathbf{x}-\mathbf{u}),$$

where

$$Q_{\mathbf{u}}(i_1, i_2) = 1 + \log \frac{p_{\mathbf{u}}^{I_1,I_2}(i_1,i_2)}{p^{I_1}(i_1)p_{\mathbf{u}}^{I_2}(i_2)},$$

and $G_\sigma(\xi_1, \xi_2)$ is a two-dimensional Gaussian kernel, with variance $\sigma^2$, which is used to estimate the joint intensity distribution from $I_2(\mathbf{x}-\mathbf{u})$ and $I_1(\mathbf{x})$.

Computing the first variation of functional $R_E(\mathbf{u}, V)$, in equation (3.5), with respect to $\mathbf{u}$ gives the following components of gradient $\partial_{\mathbf{u}} R_E(\mathbf{u}, V)$:

$$\partial_{u_l} R_E(\mathbf{u}, V) = \beta\big(\partial_1 v_{l1} + \partial_2 v_{l2} + \partial_3 v_{l3} - \triangle u_l\big), \quad l = 1, 2, 3.$$

The first variation of $R_E(\mathbf{u}, V)$ with respect to $V$, with $V = [v_{ij}]$, gives $\partial_V R_E(\mathbf{u}, V)$:

$$\begin{aligned}
\partial_{v_{11}} R_E(\mathbf{u}, V) &= \beta(v_{11} - \partial_1 u_1) + \nu \mathrm{c}_1(1 + v_{11}) + \mu\big(\mathrm{c}_2(1 + v_{11}) + \mathrm{c}_5 v_{12} + \mathrm{c}_6 v_{13}\big),\\
\partial_{v_{12}} R_E(\mathbf{u}, V) &= \beta(v_{12} - \partial_2 u_1) + \nu \mathrm{c}_1 v_{12} + \mu\big(\mathrm{c}_3 v_{12} + \mathrm{c}_5(1 + v_{11}) + \mathrm{c}_7 v_{13}\big),\\
\partial_{v_{13}} R_E(\mathbf{u}, V) &= \beta(v_{13} - \partial_3 u_1) + \nu \mathrm{c}_1 v_{13} + \mu\big(\mathrm{c}_4 v_{13} + \mathrm{c}_6(1 + v_{11}) + \mathrm{c}_7 v_{12}\big),\\
\partial_{v_{21}} R_E(\mathbf{u}, V) &= \beta(v_{21} - \partial_1 u_2) + \nu \mathrm{c}_1 v_{21} + \mu\big(\mathrm{c}_2 v_{21} + \mathrm{c}_5(1 + v_{22}) + \mathrm{c}_6 v_{23}\big),
\end{aligned}$$

$$
\begin{aligned}
\partial_{v_{22}} R_E(\mathbf{u}, V) &= \beta(v_{22} - \partial_2 u_2) + \nu \mathrm{c}_1(1 + v_{22}) + \mu\big(\mathrm{c}_3(1 + v_{22}) + \mathrm{c}_5 v_{21} + \mathrm{c}_7 v_{23}\big), \\
\partial_{v_{23}} R_E(\mathbf{u}, V) &= \beta(v_{23} - \partial_3 u_2) + \nu \mathrm{c}_1 v_{23} + \mu\big(\mathrm{c}_4 v_{23} + \mathrm{c}_6 v_{21} + \mathrm{c}_7(1 + v_{22})\big), \\
\partial_{v_{31}} R_E(\mathbf{u}, V) &= \beta(v_{31} - \partial_1 u_3) + \nu \mathrm{c}_1 v_{31} + \mu\big(\mathrm{c}_2 v_{31} + \mathrm{c}_5 v_{32} + \mathrm{c}_6(1 + v_{33})\big), \\
\partial_{v_{32}} R_E(\mathbf{u}, V) &= \beta(v_{32} - \partial_2 u_3) + \nu \mathrm{c}_1 v_{32} + \mu\big(\mathrm{c}_3 v_{32} + \mathrm{c}_5 v_{31} + \mathrm{c}_7(1 + v_{33})\big), \\
\partial_{v_{33}} R_E(\mathbf{u}, V) &= \beta(v_{33} - \partial_3 u_3) + \nu \mathrm{c}_1(1 + v_{33}) + \mu\big(\mathrm{c}_4(1 + v_{33}) + \mathrm{c}_6 v_{31} + \mathrm{c}_7 v_{32}\big),
\end{aligned}
$$

where

$$
\begin{aligned}
\mathrm{c}_1 &= v_{11} + v_{22} + v_{33} + \frac{1}{2}\big(v_{11}^2 + v_{21}^2 + v_{31}^2 + v_{12}^2 + v_{22}^2 + v_{32}^2 + v_{13}^2 + v_{23}^2 + v_{33}^2\big), \\
\mathrm{c}_2 &= 2v_{11} + v_{11}^2 + v_{21}^2 + v_{31}^2, \quad \mathrm{c}_5 = v_{21} + v_{12} + v_{11}v_{12} + v_{21}v_{22} + v_{31}v_{32}, \\
\mathrm{c}_3 &= 2v_{22} + v_{12}^2 + v_{22}^2 + v_{32}^2, \quad \mathrm{c}_6 = v_{31} + v_{13} + v_{11}v_{13} + v_{21}v_{23} + v_{31}v_{33}, \\
\mathrm{c}_4 &= 2v_{33} + v_{13}^2 + v_{23}^2 + v_{33}^2, \quad \mathrm{c}_7 = v_{32} + v_{23} + v_{12}v_{13} + v_{22}v_{23} + v_{32}v_{33}.
\end{aligned}
$$

We can compute the first variation of (3.7), obtaining $\partial_V R_{SKL}(V)$. We first simplify the notation, letting $J = |\mathcal{I} - V|$. Also, denote $L(J) = (J - 1)\log J$. Hence, $L'(J) = dL(J)/dJ = 1 + \log J - 1/J$. Thus,

$$
\begin{aligned}
\partial_{v_{11}} R_{SKL}(V) &= -\big((1 - v_{22})(1 - v_{33}) - v_{32}v_{23}\big)L'(J), \\
\partial_{v_{12}} R_{SKL}(V) &= -\big(v_{23}v_{31} + v_{21}(1 - v_{33})\big)L'(J), \\
\partial_{v_{13}} R_{SKL}(V) &= -\big(v_{21}v_{32} + (1 - v_{22})v_{31}\big)L'(J), \\
\partial_{v_{21}} R_{SKL}(V) &= -\big(v_{32}v_{13} + v_{12}(1 - v_{33})\big)L'(J), \\
\partial_{v_{22}} R_{SKL}(V) &= -\big((1 - v_{11})(1 - v_{33}) - v_{13}v_{31}\big)L'(J),
\end{aligned}
$$

$$
\begin{aligned}
\partial_{v_{23}} R_{SKL}(V) &= -\big(v_{12}v_{31} + v_{32}(1 - v_{11})\big)L'(J), \\
\partial_{v_{31}} R_{SKL}(V) &= -\big(v_{12}v_{23} + v_{13}(1 - v_{22})\big)L'(J), \\
\partial_{v_{32}} R_{SKL}(V) &= -\big(v_{21}v_{13} + v_{23}(1 - v_{11})\big)L'(J), \\
\partial_{v_{33}} R_{SKL}(V) &= -\big((1 - v_{11})(1 - v_{22}) - v_{12}v_{21}\big)L'(J).
\end{aligned}
$$

### 3.3.2 Numerical Discretization

Let $\triangle x_1$, $\triangle x_2$, $\triangle x_3$ be the spatial steps, $\triangle t$ be the time step, and $(x_{1i}, x_{2j}, x_{3k}) = (i\triangle x_1, j\triangle x_2, k\triangle x_3)$ be the grid points, for $1 \le i \le M$, $1 \le j \le N$, $1 \le k \le P$. For a function $\varphi : \Omega \to \mathbb{R}$, let $\varphi^n_{i,j,k} = \varphi(n\triangle t, i\triangle x_1, j\triangle x_2, k\triangle x_3)$. We define the difference operators based on uniformly-spaced grid as

$$
\begin{aligned}
D^{x_1}\varphi^n_{i,j,k} &= \frac{\varphi^n_{i+1,j,k} - \varphi^n_{i-1,j,k}}{2\triangle x_1}, \\
D^{x_2}\varphi^n_{i,j,k} &= \frac{\varphi^n_{i,j+1,k} - \varphi^n_{i,j-1,k}}{2\triangle x_2}, \\
D^{x_3}\varphi^n_{i,j,k} &= \frac{\varphi^n_{i,j,k+1} - \varphi^n_{i,j,k-1}}{2\triangle x_3}, \\
D^{x_1x_1}\varphi^n_{i,j,k} &= \frac{\varphi^n_{i+1,j,k} - 2\varphi^n_{i,j,k} + \varphi^n_{i-1,j,k}}{\triangle x_1^2}, \\
D^{x_2x_2}\varphi^n_{i,j,k} &= \frac{\varphi^n_{i,j+1,k} - 2\varphi^n_{i,j,k} + \varphi^n_{i,j-1,k}}{\triangle x_2^2}, \\
D^{x_3x_3}\varphi^n_{i,j,k} &= \frac{\varphi^n_{i,j,k+1} - 2\varphi^n_{i,j,k} + \varphi^n_{i,j,k-1}}{\triangle x_3^2}.
\end{aligned}
$$

Below, we will use the following notations when it is obvious that the grid point at $(i\triangle x_1, j\triangle x_2, k\triangle x_3)$ is under consideration

$$
\varphi^n := \varphi^n_{i,j,k}, \quad D^{x_l}\varphi^n := D^{x_l}\varphi^n_{i,j,k}, \quad D^{x_lx_l}\varphi^n := D^{x_lx_l}\varphi^n_{i,j,k}, \quad l = 1, 2, 3.
$$

To discretize equations (3.9) and (3.10), we use finite difference schemes. In order

to restrict the maximum displacement change per time step from being large, equation (3.9) is discretized using explicit scheme with adaptive time-stepping at every point $(i, j, k)$

$$\begin{aligned}
\frac{u_1^{n+1} - u_1^n}{\triangle t} &= -\big[\partial_{u_1} F(\mathbf{u}^n)\big] - \beta\big(D^{x_1} v_{11}^n + D^{x_2} v_{12}^n + D^{x_3} v_{13}^n\big) \\
&+ \beta\big(D^{x_1 x_1} u_1^n + D^{x_2 x_2} u_1^n + D^{x_3 x_3} u_1^n\big), \\
\frac{u_2^{n+1} - u_2^n}{\triangle t} &= -\big[\partial_{u_2} F(\mathbf{u}^n)\big] - \beta\big(D^{x_1} v_{21}^n + D^{x_2} v_{22}^n + D^{x_3} v_{23}^n\big) \\
&+ \beta\big(D^{x_1 x_1} u_2^n + D^{x_2 x_2} u_2^n + D^{x_3 x_3} u_2^n\big), \\
\frac{u_3^{n+1} - u_3^n}{\triangle t} &= -\big[\partial_{u_3} F(\mathbf{u}^n)\big] - \beta\big(D^{x_1} v_{31}^n + D^{x_2} v_{32}^n + D^{x_3} v_{33}^n\big) \\
&+ \beta\big(D^{x_1 x_1} u_3^n + D^{x_2 x_2} u_3^n + D^{x_3 x_3} u_3^n\big),
\end{aligned}$$

where $[\partial_{u_l} F(\mathbf{u}^n)]$, $l = 1, 2, 3$, is a discretization of a similarity-based gradient. In our numerical experiments, $\triangle x_1 = \triangle x_2 = \triangle x_3 = 1$, and $\triangle t$ is chosen so that the maximum displacement per iteration equals 0.1.

Equation (3.10) is discretized using semi-implicit scheme. This scheme is less sensitive to numerical constraints such as CFL condition. We have

$$\begin{aligned}
\frac{v_{11}^{n+1} - v_{11}^n}{\triangle t} &= \beta(D^{x_1} u_1^n - v_{11}^{n+1}) - \nu \mathrm{c}_1 (1 + v_{11}^n) - \mu\big(\mathrm{c}_2(1 + v_{11}^n) + \mathrm{c}_5 v_{12}^n + \mathrm{c}_6 v_{13}^n\big) \\
&+ \lambda\big((1 - v_{22}^n)(1 - v_{33}^n) - v_{32}^n v_{23}^n\big) L'(J), \\
\frac{v_{12}^{n+1} - v_{12}^n}{\triangle t} &= \beta(D^{x_2} u_1^n - v_{12}^{n+1}) - \nu \mathrm{c}_1 v_{12}^n - \mu\big(\mathrm{c}_3 v_{12}^n + \mathrm{c}_5(1 + v_{11}^n) + \mathrm{c}_7 v_{13}^n\big) \\
&+ \lambda\big(v_{23}^n v_{31}^n + v_{21}^n (1 - v_{33}^n)\big) L'(J), \\
\frac{v_{13}^{n+1} - v_{13}^n}{\triangle t} &= \beta(D^{x_3} u_1^n - v_{13}^{n+1}) - \nu \mathrm{c}_1 v_{13}^n - \mu\big(\mathrm{c}_4 v_{13}^n + \mathrm{c}_6(1 + v_{11}^n) + \mathrm{c}_7 v_{12}^n\big) \\
&+ \lambda\big(v_{21}^n v_{32}^n + (1 - v_{22}^n) v_{31}^n\big) L'(J),
\end{aligned}$$

$$
\begin{aligned}
\frac{v_{21}^{n+1}-v_{21}^{n}}{\triangle t} &= \beta(D^{x_1}u_2^n - v_{21}^{n+1}) - \nu \mathrm{c}_1 v_{21}^n - \mu\big(\mathrm{c}_2 v_{21}^n + \mathrm{c}_5(1+v_{22}^n) + \mathrm{c}_6 v_{23}^n\big) \\
&+ \lambda\big(v_{32}^n v_{13}^n + v_{12}^n(1-v_{33}^n)\big)L'(J), \\
\frac{v_{22}^{n+1}-v_{22}^{n}}{\triangle t} &= \beta(D^{x_2}u_2^n - v_{22}^{n+1}) - \nu \mathrm{c}_1 (1+v_{22}^n) - \mu\big(\mathrm{c}_3(1+v_{22}^n) + \mathrm{c}_5 v_{21}^n + \mathrm{c}_7 v_{23}^n\big) \\
&+ \lambda\big((1-v_{11}^n)(1-v_{33}^n) - v_{13}^n v_{31}^n\big)L'(J), \\
\frac{v_{23}^{n+1}-v_{23}^{n}}{\triangle t} &= \beta(D^{x_3}u_2^n - v_{23}^{n+1}) - \nu \mathrm{c}_1 v_{23}^n - \mu\big(\mathrm{c}_4 v_{23}^n + \mathrm{c}_6 v_{21}^n + \mathrm{c}_7(1+v_{22}^n)\big) \\
&+ \lambda\big(v_{12}^n v_{31}^n + v_{32}^n(1-v_{11}^n)\big)L'(J), \\
\frac{v_{31}^{n+1}-v_{31}^{n}}{\triangle t} &= \beta(D^{x_1}u_3^n - v_{31}^{n+1}) - \nu \mathrm{c}_1 v_{31}^n - \mu\big(\mathrm{c}_2 v_{31}^n + \mathrm{c}_5 v_{32}^n + \mathrm{c}_6(1+v_{33}^n)\big) \\
&+ \lambda\big(v_{12}^n v_{23}^n + v_{13}^n(1-v_{22}^n)\big)L'(J), \\
\frac{v_{32}^{n+1}-v_{32}^{n}}{\triangle t} &= \beta(D^{x_2}u_3^n - v_{32}^{n+1}) - \nu \mathrm{c}_1 v_{32}^n - \mu\big(\mathrm{c}_3 v_{32}^n + \mathrm{c}_5 v_{31}^n + \mathrm{c}_7(1+v_{33}^n)\big) \\
&+ \lambda\big(v_{21}^n v_{13}^n + v_{23}^n(1-v_{11}^n)\big)L'(J), \\
\frac{v_{33}^{n+1}-v_{33}^{n}}{\triangle t} &= \beta(D^{x_3}u_3^n - v_{33}^{n+1}) - \nu \mathrm{c}_1 (1+v_{33}^n) - \mu\big(\mathrm{c}_4(1+v_{33}^n) + \mathrm{c}_6 v_{31}^n + \mathrm{c}_7 v_{32}^n\big) \\
&+ \lambda\big((1-v_{11}^n)(1-v_{22}^n) - v_{12}^n v_{21}^n\big)L'(J),
\end{aligned}
$$

where $J = |\mathcal{I} - V|$, $L(J) = (J-1)\log J$, and $L'(J) = dL(J)/dJ = 1 + \log J - 1/J$.

### 3.3.3 Algorithm

We are now ready to give the algorithm for the Unbiased registration via nonlinear elastic regularization.

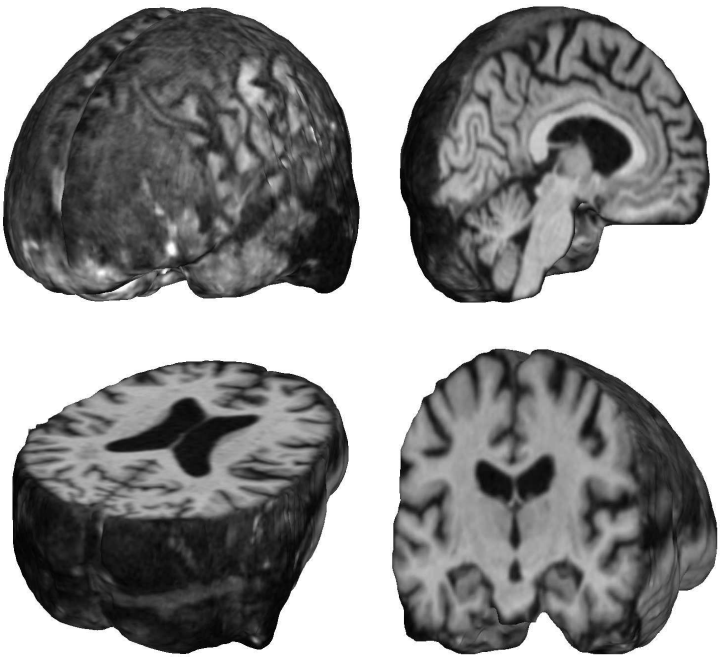


Figure 3.1: $I_1$ and $I_2$ are Serial MRI images from the ADNI follow-up dataset (images acquired one year apart). Here, volume $I_1$ is depicted as a brain volume as well as from sagittal, axial, and coronal views. Figure 3.2 shows volume $I_2$. Nonrigid registration aligns $I_2$ into correspondence with $I_1$.

**Algorithm 2** Unbiased Registration via Nonlinear Elastic Regularization

1: Initialize $t = 0$, $\mathbf{u}(\mathbf{x}, 0) = 0$, and $V(\mathbf{x}, 0) = 0$.
2: Calculate $V(\mathbf{x}, t)$ using equation (3.10), where the equation is discretized using the semi-implicit method described in Section 3.3.2.
Steps 3-5 describe the procedure for solving equation (3.9) advancing $\mathbf{u}(\mathbf{x}, t)$ in time using the explicit scheme. Numerical discretization is described in Section 3.3.2.
3: Calculate the perturbation of the displacement field $\mathbf{R}(\mathbf{x}) = -\partial E_{\mathbf{u}}(\mathbf{u}, V)$.
4: Time step $\triangle t$ is calculated adaptively so that $\triangle t \cdot \max(||\mathbf{R}||_2) = \delta u$, where $\delta u$ is the maximal displacement allowed in one iteration. Results in this chapter are obtained with $\delta u = 0.1$.
5: Advance equation (3.9), i.e. $\partial \mathbf{u}(\mathbf{x}, t)/\partial t = \mathbf{R}(\mathbf{x})$, in time, with time step from step 4, solving for $\mathbf{u}(\mathbf{x}, t)$.
6: If the cost functional in (3.8) decreases by sufficiently small amount compared to the previous iteration, then stop.
7: Let $t := t + \triangle t$ and go to step 2.

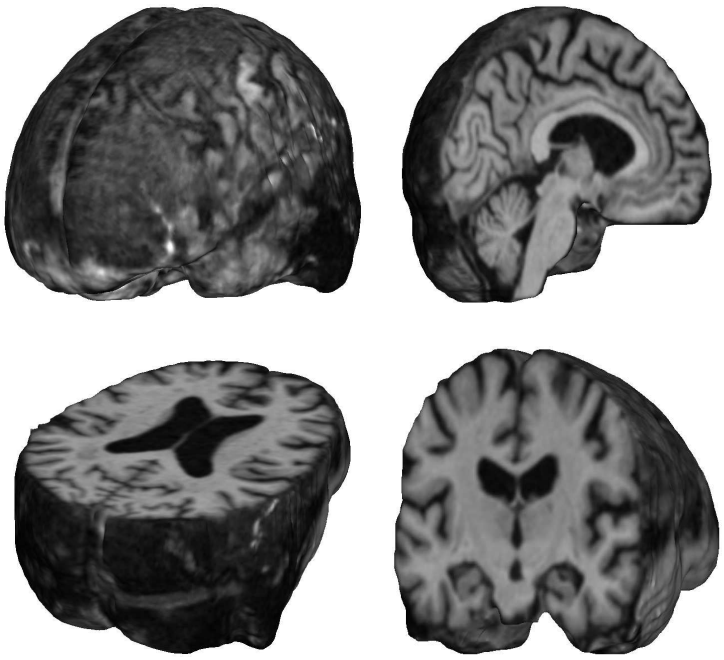


Figure 3.2: $I_1$ and $I_2$ are Serial MRI images from the ADNI follow-up dataset (images acquired one year apart). Here, volume $I_2$ is depicted as a brain volume as well as from sagittal, axial, and coronal views. Figure 3.1 shows volume $I_1$. Nonrigid registration aligns $I_2$ into correspondence with $I_1$.

## 3.4 Results and Discussion

We tested the proposed *unbiased nonlinear elastic* registration model (Algorithm 2), and compared the results to those obtained with the *unbiased fluid registration* method [112] given by Algorithm 1. Here, both methods were coupled with the $L^2$ and mutual information (MI) based similarity measures. In our experiments, we used a pair of serial MRI images ($220 \times 220 \times 220$) from the Alzheimer's Disease Neuroimaging Initiative (ADNI). Since the images were acquired one year apart, from a subject with Alzheimer's disease, real anatomical changes are present, which allows methods to be compared in the presence of true biological changes.

Figure 3.1 shows the target image $I_1$ and Figure 3.2 displays the source image $I_2$.

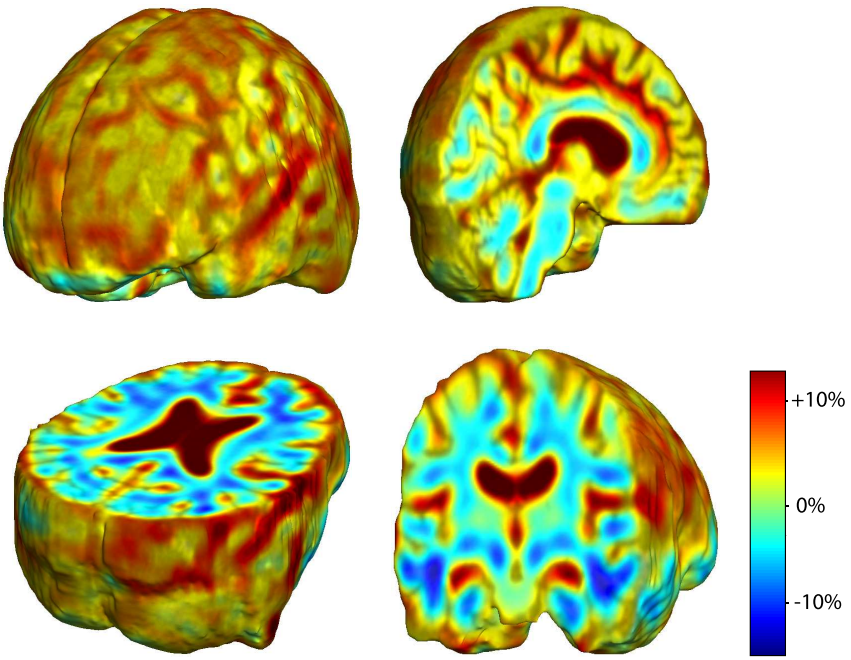



Figure 3.3: Results are obtained using *unbiased fluid registration* coupled with $L^2$ *matching*. Jacobian maps are superimposed on the brain structure.

Figures 3.3 through 3.6 show the resulting Jacobian maps obtained using unbiased fluid registration and unbiased nonlinear elastic registration, both coupled with $L^2$ and MI matching. Results generated using the fluid and nonlinear elasticity based unbiased models are similar, both suggesting a mild volume reduction in gray and white matter and ventricular enlargement that is observed in Alzheimer's disease patients. The advantages of the unbiased nonlinear elasticity model is its more locally plausible reproduction of atrophic changes in the brain and its robustness to original misalignment of brain volumes, which is especially noticeable on the brain surface. The unbiased nonlinear elasticity model coupled with $L^2$ matching generated very similar results to those obtained with the MI similarity measure, partly because difference images typically contain only noise after registration. Unbiased fluid registration method, however, is more effective in modeling the regional neuroanatomical changes, showing

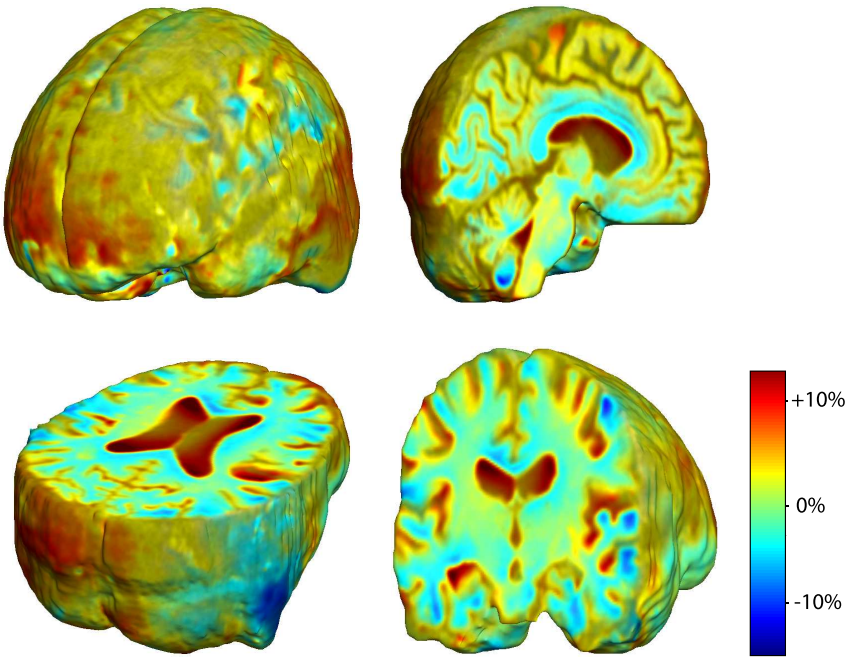



Figure 3.4: Results are obtained using *unbiased nonlinear elastic registration* coupled with $L^2$ *matching*. Jacobian maps are superimposed on the brain structure.

more clearly which parts of the volume have undergone largest tissue changes, such as ventricular enlargement as shown in Figures 3.3 through 3.6.

Figures 3.7 and 3.8 show deformed grids generated with unbiased fluid and unbiased nonlinear elastic registration models. Figure 3.9 shows the energy decrease per iteration for both models.

In Figure 3.10, we examined the inverse consistency of the mappings [20] generated using unbiased nonlinear elastic registration. Here, the deformation was computed in both directions (time 2 to time 1, and time 1 to time 2) using mutual information matching. The forward and backward Jacobian maps were concatenated (in an ideal situation, this operation should yield the identity), with the products of Jacobians having values close to 1.

The unbiased nonlinear elasticity model does not require expensive Navier-Stokes

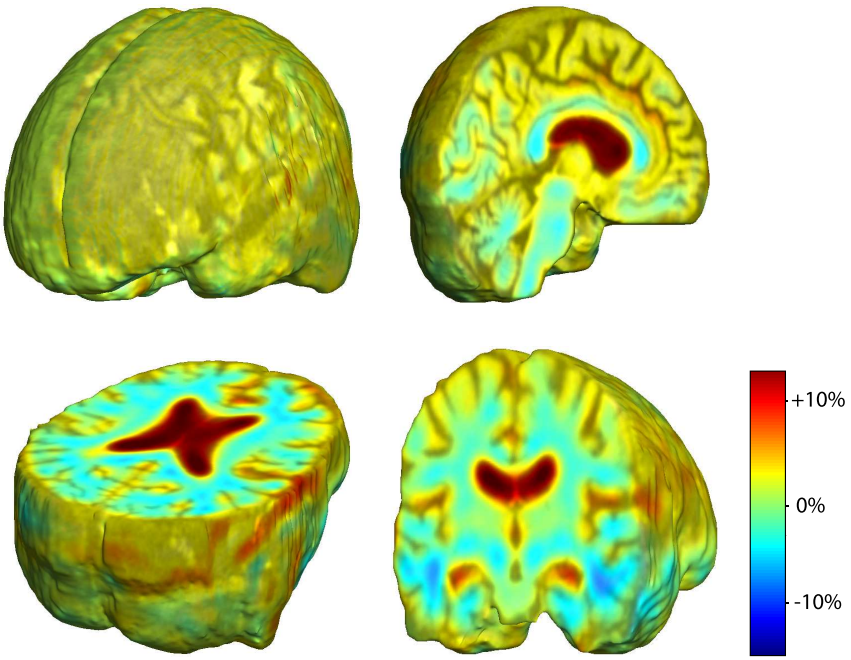



Figure 3.5: Results are obtained using *unbiased fluid registration* coupled with *mutual information matching*. Jacobian maps are superimposed on the brain structure.

solver (or its approximation), which is employed at each iteration for fluid flow models. Hence, in our experiments, each iteration of the unbiased nonlinear elasticity algorithm took less time than the unbiased fluid step. Future studies will examine the registration accuracy of the different models where ground truth is known, and will compare each model's power for detecting inter-group differences or statistical effects on rates of atrophy.

## 3.5 Conclusion

This chapter introduces a new unbiased nonlinear elastic image registration model. The nonlinear elastic and the unbiased regularization terms are simplified using the change of variables by introducing an unknown that approximates the Jacobian ma-

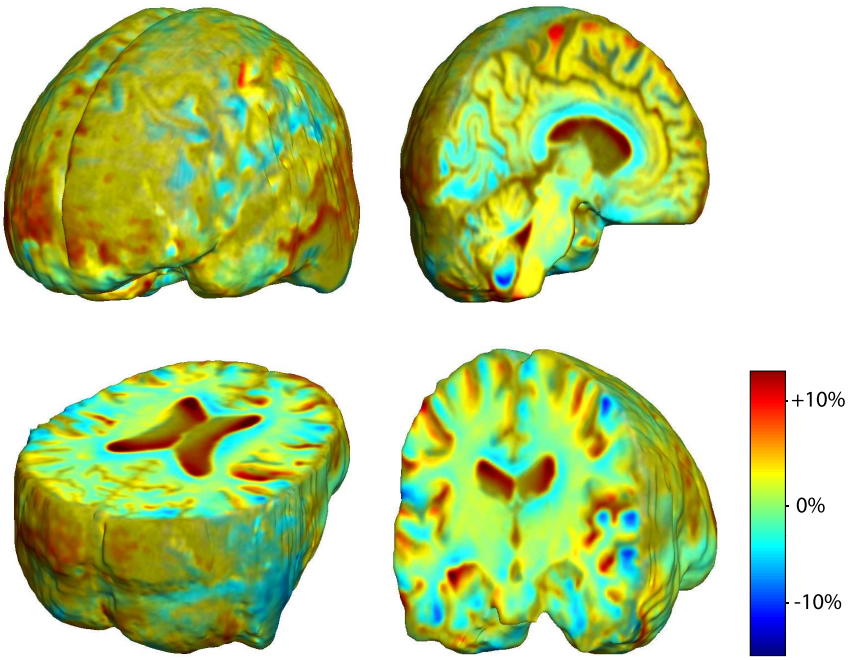
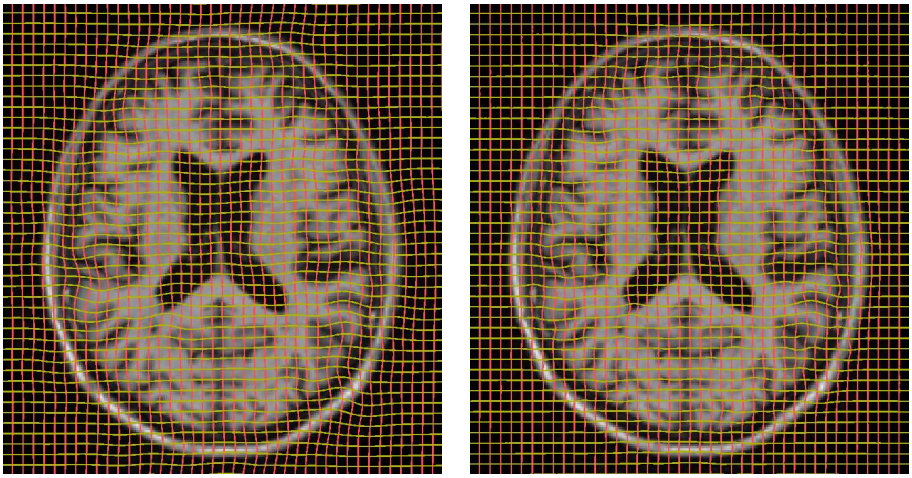


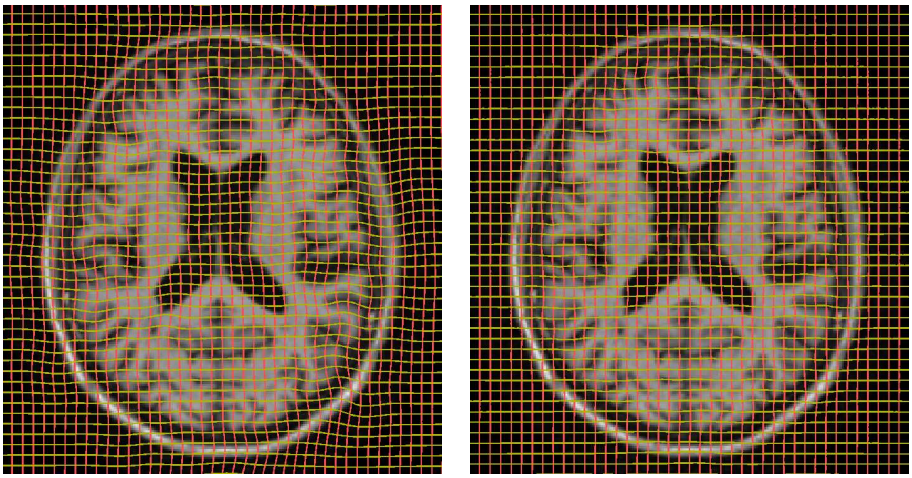

Figure 3.6: Results are obtained using *unbiased nonlinear elastic registration* coupled with *mutual information matching*. Jacobian maps are superimposed on the brain structure.

trix of the displacement field. This reduces the minimization to involve differential equations that are linear in the unknowns. In contrast to recently proposed unbiased fluid registration method, the new model is written in a unified variational form and is minimized using gradient descent. It is computationally efficient and easy to implement. The unbiased large-deformation nonlinear elastic registration was tested using volumetric serial magnetic resonance images and shown to have some advantages for medical imaging applications.



Unbiased Fluid
with $L^2$ matching

Unbiased Nonlinear Elasticity
with $L^2$ matching



Unbiased Fluid
with MI matching

Unbiased Nonlinear Elasticity
with MI matching

Figure 3.7: Results obtained using unbiased fluid registration and unbiased nonlinear elastic registration, both coupled with $L^2$ and MI matching. The generated grids are superimposed on top of 2D cross-sections of the 3D volumes.

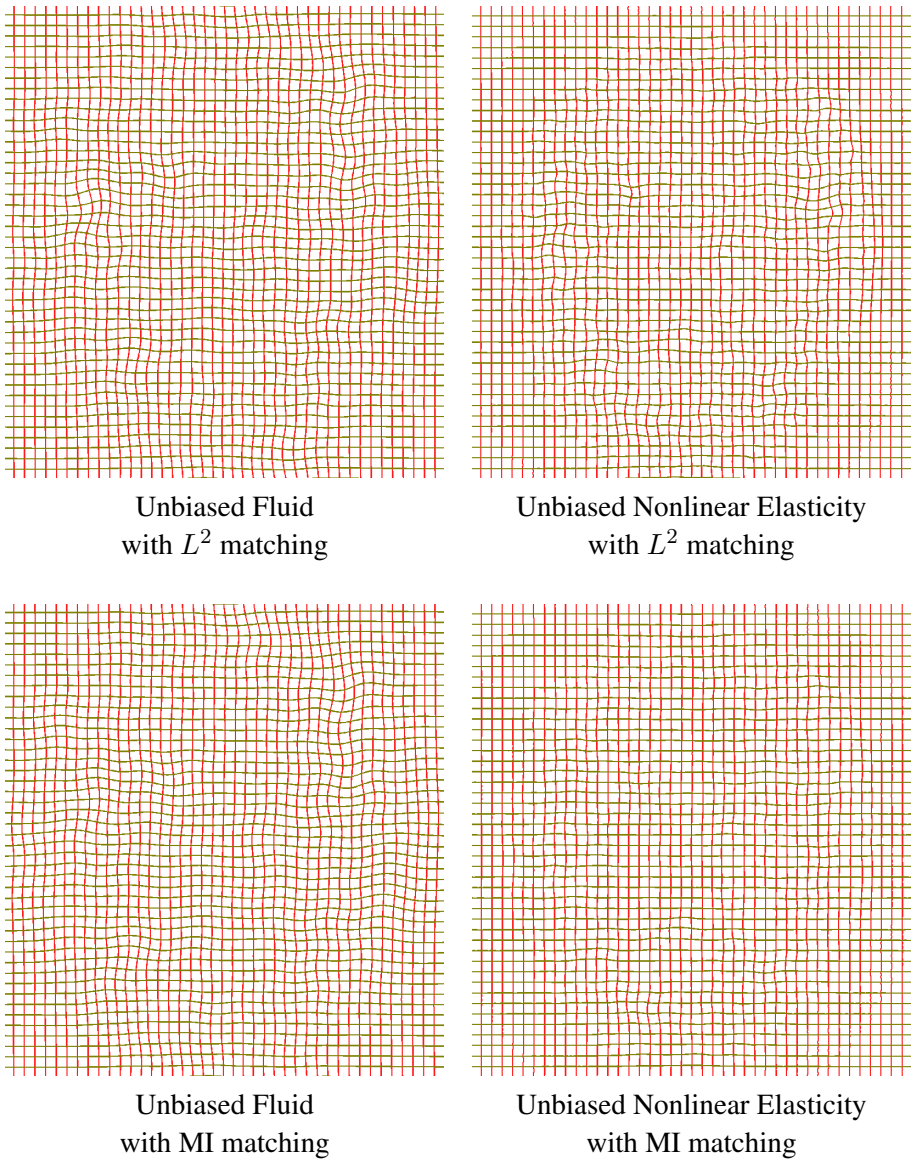
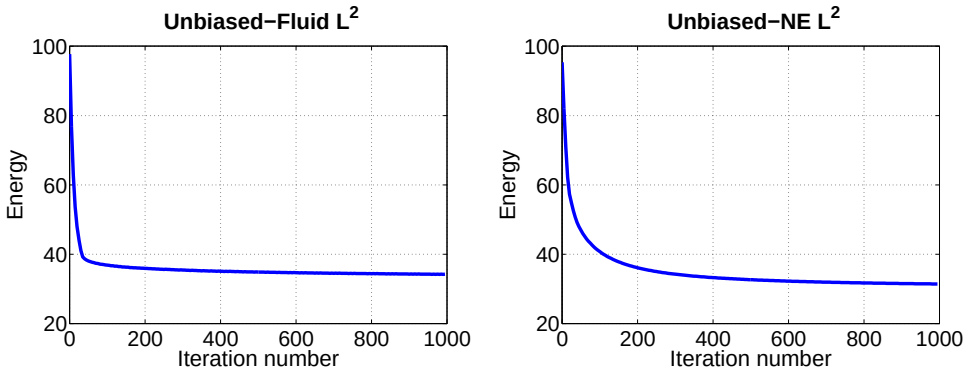


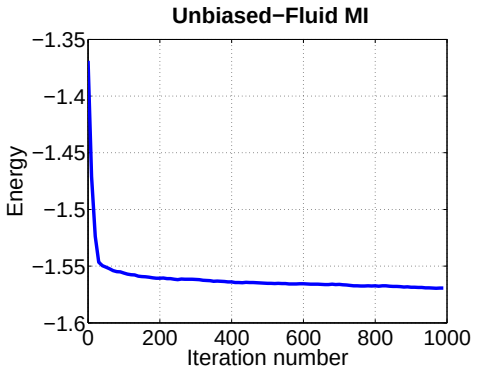

Figure 3.8: Results obtained using unbiased fluid registration and unbiased nonlinear elastic registration, both coupled with $L^2$ and MI matching. The generated grids of 2D cross-sections of the 3D volumes are shown.



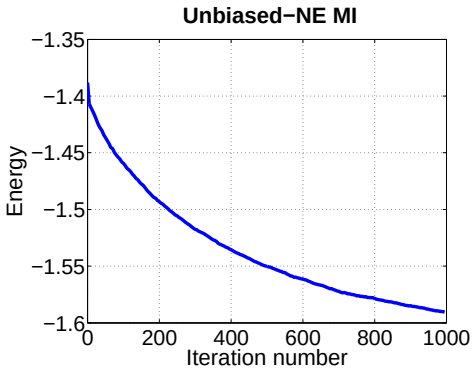

Unbiased Fluid
with $L^2$ matching




Unbiased Nonlinear Elasticity
with $L^2$ matching




Unbiased Fluid
with MI matching




Unbiased Nonlinear Elasticity
with MI matching

Figure 3.9: Energy per iteration for the unbiased fluid registration and unbiased nonlinear elastic registration, both coupled with $L^2$ and MI matching.

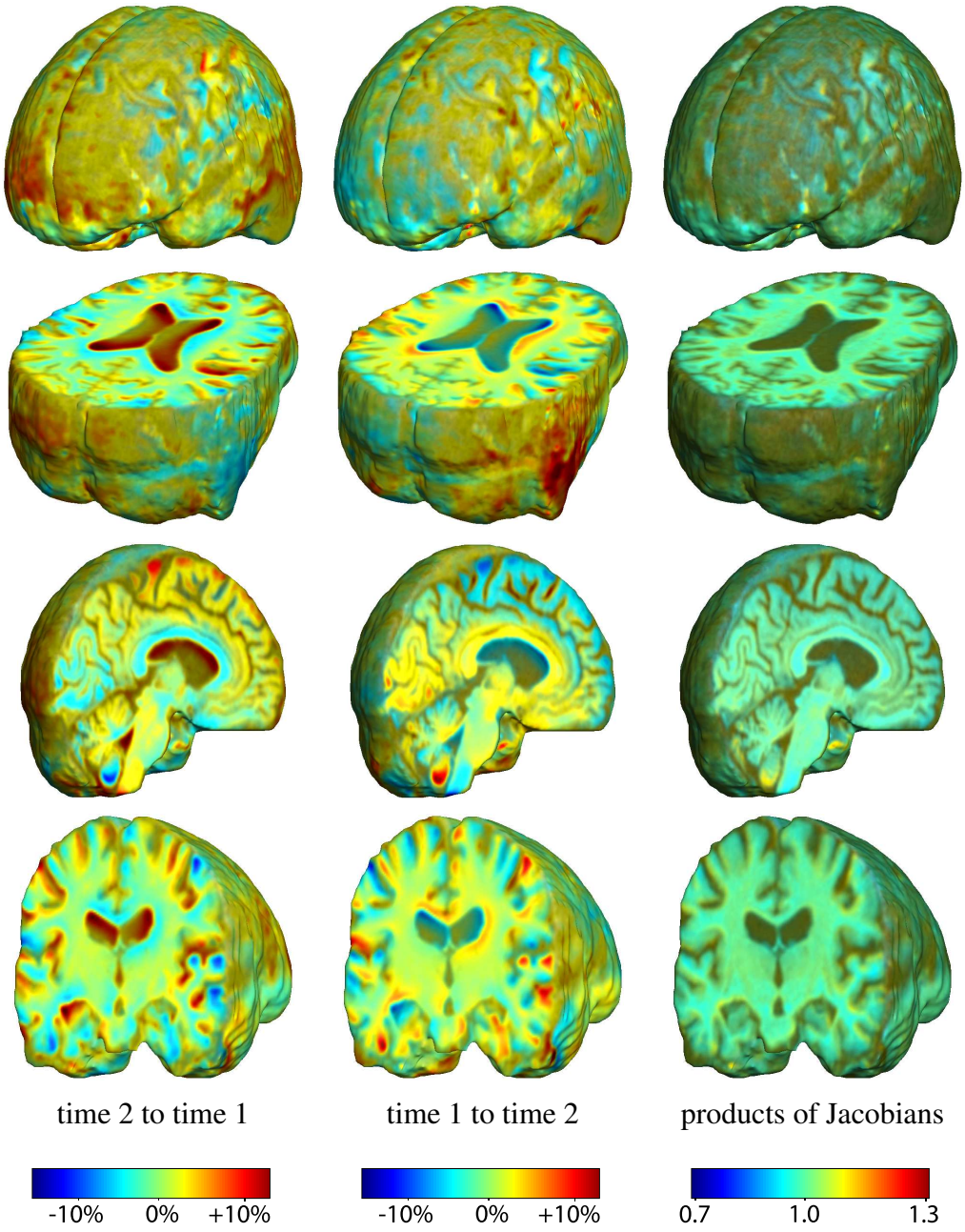



Figure 3.10: This figure examines the inverse consistency of the unbiased nonlinear elastic registration. Here, the model is coupled with mutual information matching. Jacobian maps of deformations from time 2 to time 1 (column 1) and time 1 to time 2 (column 2) are superimposed on the target volumes. The products of Jacobian maps, shown in column 3, have values close to 1, suggesting inverse consistency.

# CHAPTER 4

# Unbiased Image Registration Methods: Statistical Assessment of Performance

This chapter examines the reproducibility and the power to detect real changes of different computational techniques. It is the first work to systematically investigate the reproducibility and variability of different registration methods in tensor-based morphometry (TBM). In particular, we compare matching functionals (sum of squared differences and mutual information), as well as large deformation registration schemes (symmetric and asymmetric unbiased registration and viscous fluid registration) using serial MRI scans of ten normal elderly patients from the preparatory phase of the Alzheimer's Disease Neuroimaging Initiative (ADNI) and ten Alzheimer's subjects from the ADNI follow-up phase. Our results show that the unbiased methods, both symmetric and asymmetric, have higher reproducibility. The unbiased methods are less likely to produce changes in the absence of any real physiological change. Moreover, they are also better in detecting biological deformations by penalizing any bias in the corresponding statistical maps.

## 4.1 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, 21, 22, 24, 44, 70, 90, 99, 106]. The goal of image registration is to align, or spatially normalize, one image to another. In multi-subject studies, this reduces subject-specific anatomic differences by deforming individual images onto a population average brain template. When applied to serial scans of human brain, image registration offers tremendous power in detecting the earliest signs of illness, understanding normal brain development or aging, and monitoring disease progression [27, 30, 49, 87, 95, 97, 104]. Recently, there has been an expanding literature on various nonrigid registration techniques, with different image matching functionals, regularization schemes, and numerical implementations. In Chapter 2, we examined the statistical properties of Jacobian maps, and proposed an *unbiased* large-deformation image registration approach. In this context, unbiased means that the Jacobian determinants of the deformations recovered between a pair of images follow a log-normal distribution, with zero mean after log-transformation. We argued that this distribution is beneficial when recovering change in regions of homogeneous intensity, and in ensuring symmetrical results when the order of two images being registered is switched. This method was applied to a longitudinal MRI dataset from a single subject, showing promising results in eliminating spurious signals. We also noticed that different registration techniques, when applied to the same longitudinal dataset, may sometimes yield visually very different Jacobian maps, causing problems in interpreting local structural changes. Given this ambiguity and the increasing use of registration methods to measure brain change, more information is required concerning the baseline stability, reproducibility, and statistical properties of signals generated by different nonrigid registration techniques.

In this chapter, we introduce a novel Asymmetric Unbiased model (by contrast with the Symmetric Unbiased model). Most importantly, we aim to provide quality calibrations for different non-rigid registration techniques in TBM. In particular, we compare two common matching functionals: $L^2$, or the sum of squared intensity differences,

versus mutual information, and three regularization techniques (fluid registration versus the Asymmetric Unbiased and Symmetric Unbiased techniques). Our experiments are designed to decide which registration method is more reproducible, more reliable, and offers less artifactual variability in regions of homogeneous image intensity. The foundation of our calibrations is based on the assumption that, by scanning healthy normal human subjects twice over a 2-week period using the same protocol, serial MRI scan pairs should not show any systematic biological change. Therefore, any regional structural differences detected using TBM over such a short interval may be assumed to be errors. We apply statistical analysis to the profile of these errors, providing information on the reliability, reproducibility and variability of different registration techniques. Moreover, serial images of 10 subjects from the ADNI follow-up phase (images acquired one year apart) were analyzed in a similar fashion and compared to the ADNI baseline data. In images collected one year apart, real anatomical changes are present; neurobiological changes due to aging and dementia include widespread cell shrinkage, regional gray and white matter atrophy and expansion of fluid-filled spaces in the brain. Thus, a good computational technique should be able to differentiate between longitudinal image pairs collected for the ADNI baseline (2-week) and follow-up (1-year) phases. For details of the ADNI acquisition protocol, please refer to [53, 55, 91, 45].

The unbiased approach, which is analyzed in this chapter, couples the computation of deformations with statistical analyses on the resulting Jacobian maps. As a result, the unbiased technique ensures that deformations have intuitive axiomatic properties by penalizing any bias in the corresponding statistical maps. In the following sections, we briefly restate the formulations of the unbiased technique, define energy functionals for minimization, and perform thorough statistical analyses to demonstrate the advantages of the unbiased registration models.

## 4.2 Asymmetric and Symmetric Unbiased Large-Deformation Image Registration

In Section 2.1, we described the construction of the Unbiased Large-Deformation Image Registration. We associated three probability density functions to $\mathbf{g}$, $\mathbf{g}^{-1}$, and the identity mapping $\mathbf{id}$:

$$p_{\mathbf{g}}(\mathbf{x}) = |D\mathbf{g}(\mathbf{x})|, \quad p_{\mathbf{g}^{-1}}(\mathbf{x}) = |D\mathbf{g}^{-1}(\mathbf{x})|, \quad p_{\mathbf{id}}(\mathbf{x}) = 1.$$

By associating deformations with their corresponding global density maps, we applied information theory to quantify the magnitude of deformations. In our original approach, we chose the symmetric Kullback-Leibler (SKL) distance to quantify the deformation. In this chapter, we also investigate the Kullback-Leibler (KL) divergence as a measure of deformations. The KL divergence between two probability density functions, $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$, is defined as

$$KL(p_1(\mathbf{x}), p_2(\mathbf{x})) = \int_\Omega p_1(\mathbf{x}) \log \frac{p_1(\mathbf{x})}{p_2(\mathbf{x})} d\mathbf{x} \geq 0,$$

and the SKL distance as

$$SKL(p_1(\mathbf{x}), p_2(\mathbf{x})) = KL(p_1(\mathbf{x}), p_2(\mathbf{x})) + KL(p_2(\mathbf{x}), p_1(\mathbf{x})).$$

The Unbiased method solves for the deformation $\mathbf{g}$ (or, equivalently, for the displacement $\mathbf{u}$) minimizing the energy functional $E$, consisting of the image matching term $F$ and the regularizing term $R$ which is based on KL divergence or SKL distance. The fidelity term $F$ dependents on $I_2$ and $I_1$, as well as the displacement $\mathbf{u}$. The general

minimization problem can be written as

$$\begin{aligned} E(I_1, I_2, \mathbf{u}) &= F(I_1, I_2, \mathbf{u}) + \lambda R(\mathbf{u}), \\ &\inf_{\mathbf{u}} E(I_1, I_2, \mathbf{u}). \end{aligned} \tag{4.1}$$

Here, $\lambda > 0$ is a weighting parameter.

In this study, the matching functional $F$ takes two forms: the $L^2$ norm, given by equation (1.7), and MI (mutual information), given by equation (2.19). These functionals have each been widely used in the past for nonrigid registration, to measure the intensity agreement between the deformed image and the target image.

### 4.2.1 Asymmetric Unbiased Registration

To quantify the magnitude of deformation $\mathbf{g}$, in this chapter we introduce a new regularization term $R_{KL}$, which is an asymmetric measure between $p_{\mathbf{id}}$ and $p_{\mathbf{g}}$:

$$R_{KL}(\mathbf{g}) = KL(p_{\mathbf{id}}, p_{\mathbf{g}}).$$

This regularization term can be shown to be

$$\begin{aligned} R_{KL}(\mathbf{g}) &= \int_\Omega p_{\mathbf{id}} \log \frac{p_{\mathbf{id}}}{p_{\mathbf{g}}}\, d\mathbf{x} = \int_\Omega -\log |D\mathbf{g}(\mathbf{x})| d\mathbf{x} \\ &= \int_\Omega |D\mathbf{g}^{-1}(\mathbf{y})| \log |D\mathbf{g}^{-1}(\mathbf{y})| d\mathbf{y}. \end{aligned} \tag{4.2}$$

Thus, the energy functional in (4.1) implementing Asymmetric Unbiased registration can be written as

$$E(I_1, I_2, \mathbf{u}) = F(I_1, I_2, \mathbf{u}) - \lambda \int_\Omega \log |D(\mathbf{x} - \mathbf{u}(\mathbf{x}))| d\mathbf{x}, \tag{4.3}$$

for some distance measure $F$ between $I_2(\mathbf{x} - \mathbf{u})$ and $I_1(\mathbf{x})$.

### 4.2.2 Symmetric Unbiased Registration

The regularization functional, based on the symmetric KL distance between $p_{\mathbf{id}}$ and $p_{\mathbf{g}}$, was introduced in Chapter 2:

$$R_{SKL}(\mathbf{g}) = SKL(p_{\mathbf{id}}, p_{\mathbf{g}}).$$

The regularization term is linked to statistics on Jacobian maps as follows

$$\begin{aligned} R_{SKL}(\mathbf{g}) &= \int_{\Omega} \big(|D\mathbf{g}(\mathbf{x})| - 1\big) \log |D\mathbf{g}(\mathbf{x})| d\mathbf{x} \\ &= \int_{\Omega} \big(|D\mathbf{g}^{-1}(\mathbf{y})| - 1\big) \log |D\mathbf{g}^{-1}(\mathbf{y})| d\mathbf{y}. \end{aligned} \tag{4.4}$$

The energy functional employing Symmetric Unbiased registration can be rewritten as

$$\begin{aligned} E(I_1, I_2, \mathbf{u}) = F(I_1, I_2, \mathbf{u}) \\ +\lambda \int_{\Omega} \big(|D(\mathbf{x} - \mathbf{u}(\mathbf{x}))| - 1\big) \log |D(\mathbf{x} - \mathbf{u}(\mathbf{x}))| d\mathbf{x}, \end{aligned} \tag{4.5}$$

for some distance measure $F$. Notice that the symmetric unbiased regularizing functional (4.4) is pointwise nonnegative, while the asymmetric unbiased regularizer in (4.2) can take either positive or negative values locally.

## 4.3 Summary of Models

We summarize the methods we will be referring to in our subsequent analyses. In later discussions, minimization of the following energies

$$E(I_1, I_2, \mathbf{u}) = F_{L^2}(I_1, I_2, \mathbf{u}) + \lambda R_{KL}(\mathbf{u}) \tag{4.6}$$

$$\text{and} \quad E(I_1, I_2, \mathbf{u}) = F_{L^2}(I_1, I_2, \mathbf{u}) + \lambda R_{SKL}(\mathbf{u}) \tag{4.7}$$

will be referred to as $L^2$-Asymmetric Unbiased and $L^2$-Symmetric Unbiased models, respectively. Similarly, minimization of

$$E(I_1, I_2, \mathbf{u}) = F_{MI}(I_1, I_2, \mathbf{u}) + \lambda R_{KL}(\mathbf{u}) \tag{4.8}$$

$$\text{and} \quad E(I_1, I_2, \mathbf{u}) = F_{MI}(I_1, I_2, \mathbf{u}) + \lambda R_{SKL}(\mathbf{u}) \tag{4.9}$$

will be referred to as the MI-Asymmetric Unbiased and MI-Symmetric Unbiased models, respectively. Similar to Algorithm 1 of Chapter 2, we employ fluid regularization, defining the force field $\mathbf{f}$ as

$$\mathbf{f}(\mathbf{x}, \mathbf{u}) = -\partial_{\mathbf{u}} E(I_1, I_2, \mathbf{u}).$$

The instantaneous velocity is obtained by convolving $\mathbf{f}$ with Gaussian kernel $G_\sigma$ of variance $\sigma^2$, $\mathbf{v} = G_\sigma * \mathbf{f}(\mathbf{x}, \mathbf{u})$, and the following partial differential equation is solved to obtain the displacement field $\mathbf{u}$:

$$\frac{\partial \mathbf{u}}{\partial t} = \mathbf{v} - \mathbf{v} \cdot \nabla \mathbf{u}.$$

For this reason, the model (4.6) (or (4.7)), with $\lambda = 0$, will be referred to as the $L^2$-Fluid model. Also, the model (4.8) (or (4.9)), with $\lambda = 0$, defines the MI-Fluid model.

## 4.4 Statistical Analysis

### 4.4.1 Statistical testing on the deviation of log Jacobian maps in the absence of changes

Based on the authors' approach in [60], we observe that, given that there is no systematic structural change within two weeks, any deviation of the Jacobian map from

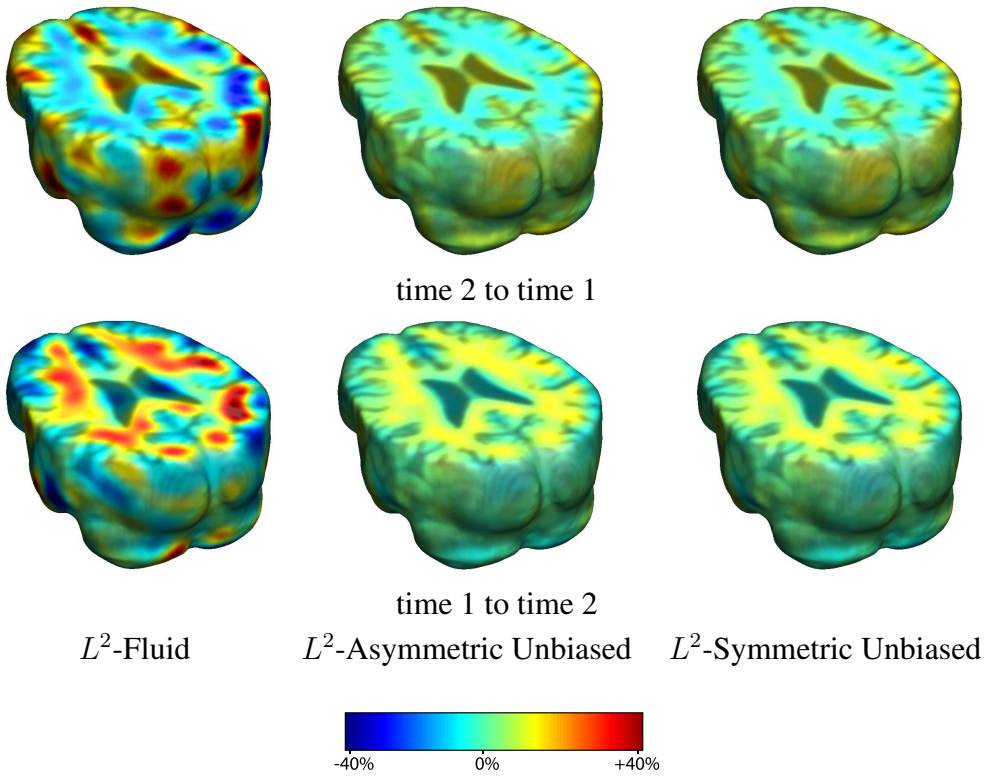



Figure 4.1: Nonrigid registration was performed on an image pair from one of the subjects from the ADNI Baseline study (serial MRI images acquired two weeks apart) using $L^2$-Fluid (column 1), $L^2$-Asymmetric Unbiased (column 2), and $L^2$-Symmetric Unbiased (column 3) registration methods. Jacobian maps of deformations from time 2 to time 1 (row 1) and time 1 to time 2 (row 2) are superimposed on the target volumes. The unbiased methods generate less noisy Jacobian maps with values closer to 1; this shows the greater stability of the approach when no volumetric change is present.

one should be considered error. Thus, we expect that a better registration technique would yield $\log|D\mathbf{g}|$ values closer to $0$ (i.e., smaller log Jacobian deviation translates into better methodology). Mathematically speaking, one way to test the performance is to consider the deviation map $dev$ of the logged (i.e., logarithmically transformed) Jacobian away from zero, defined at each voxel as

$$dev(\mathbf{x}) = \big|\log|D\mathbf{g}(\mathbf{x})|\big|. \tag{4.10}$$

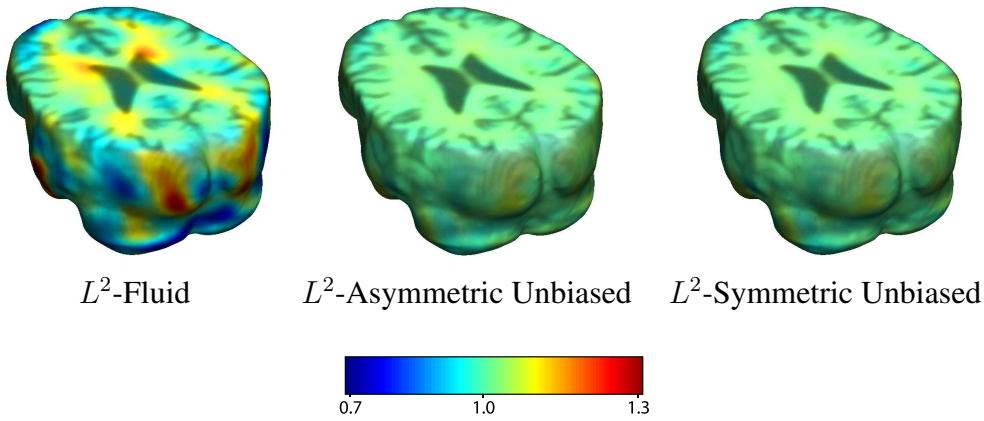
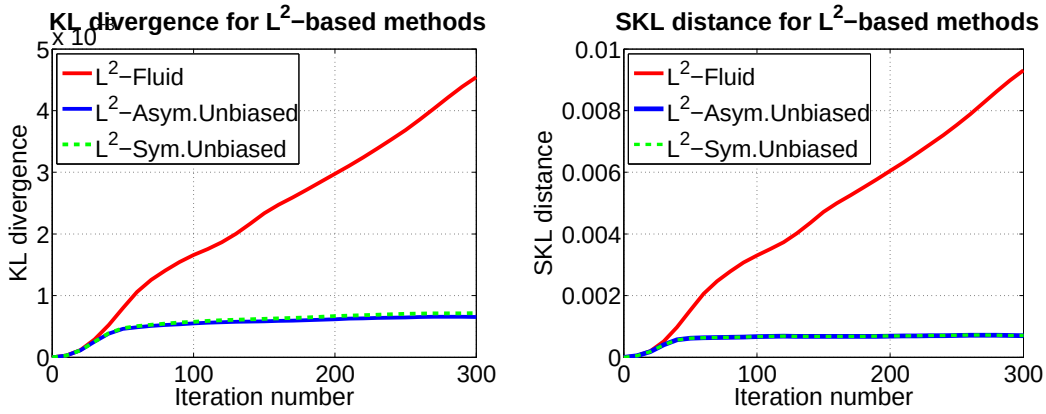



Figure 4.2: This figure examines the inverse consistency of deformation models. Products of Jacobian maps generated using all three models are shown, for forward direction (time 1 to time 2) and backward direction (time 2 to time 1). For the $L^2$-based unbiased methods, the products of the Jacobian maps are less noisy, with values closer to 1, showing better inverse consistency.

For two different registration methods $A$ and $B$, we define the voxel-wise deviation gain of $A$ over $B$ (denoted by $S^{A,B}$) as

$$S^{A,B}(\mathbf{x}) = dev^A(\mathbf{x}) - dev^B(\mathbf{x}). \tag{4.11}$$

For the ADNI baseline dataset (in which patients are scanned twice with MRI, two weeks apart), two distinct types of $t$ tests are used, a within-subject paired $t$ test and a group paired $t$ test. A within-subject paired $t$ test is conducted for each subject by pooling all voxels inside a region of interest, as defined by the ICBM whole brain mask (the ICBM brain is a standardized population average image, defined by the International Consortium for Brain Mapping [69]). This determines whether two methods differ significantly inside the whole brain (for each subject). A group paired $t$ test, on the other hand, is performed across subjects, by computing a voxel-wise $t$-map of deviation gains. In this case, to statistically compare the performance of two registration methods, we rely on the standard $t$ test on the voxel mean of $S$. To construct a




(a) KL divergence (b) SKL distance

Figure 4.3: (a) KL divergence and (b) SKL distance per iteration are shown for $L^2$-Fluid (solid red), $L^2$-Asymmetric Unbiased (solid blue), and $L^2$-Symmetric Unbiased (dashed green) methods. For $L^2$-Fluid, both KL and SKL measures increase. Even though the Asymmetric Unbiased method explicitly minimizes the KL divergence, and the Symmetric Unbiased model minimizes the SKL distance, both of the KL and SKL measures stabilize for both unbiased methods.

suitable null hypothesis, we notice that the following relation would hold, assuming $B$ outperforms $A$

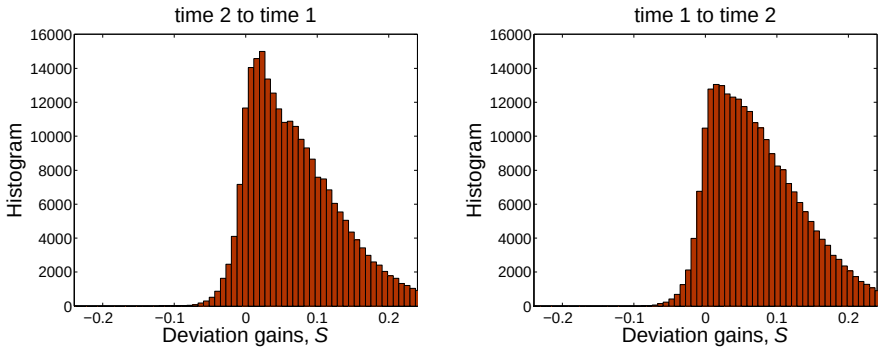
$$S^{A,B} > 0.$$

Thus, the null hypothesis in this case would be testing if the mean deviation gain is zero

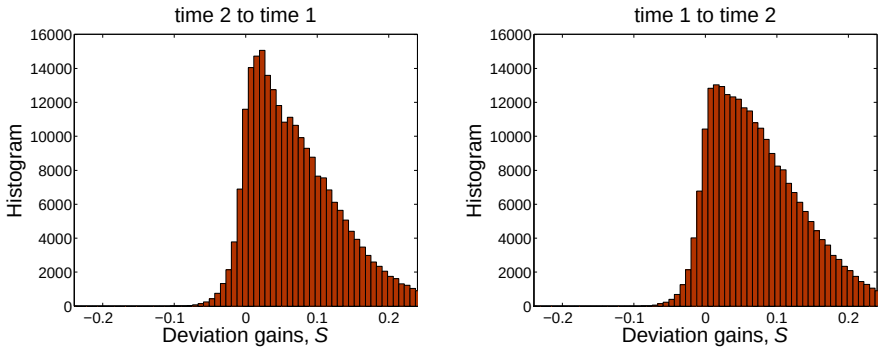
$$H_0 : \mu_{S^{A,B}} = 0.$$

To determine the ranking of $A$ and $B$, we have to consider one-sided alternative hypotheses. For example, when testing if $B$ outperforms $A$, we use the following alternative hypothesis

$$H_1 : \mu_{S^{A,B}} > 0.$$




(a) $L^2$-Fluid vs. $L^2$-Asym.Unbiased




(b) $L^2$-Fluid vs. $L^2$-Sym.Unbiased

Figure 4.4: Histograms of voxel-wise deviation gains (a) $L^2$-Fluid over $L^2$- Asymmetric Unbiased and (b) $L^2$-Fluid over $L^2$-Symmetric Unbiased for one of the subjects for the forward direction (time 2 to time 1) and backward direction (time 1 to time 2). The histograms are skewed to the right, indicating the superiority of Asymmetric Unbiased and Symmetric Unbiased registration methods over Fluid registration. A paired $t$ test shows significance ($p < 0.0001$).

The voxel-wise $T$ statistic, defined as

$$T_{S^{A,B}}(\mathbf{x}) = \frac{\sqrt{n} \cdot \overline{S^{A,B}(\mathbf{x})}}{\sigma_{S^{A,B}}(\mathbf{x})}, \tag{4.12}$$

where

$$\overline{S^{A,B}(\mathbf{x})} = \frac{\sum_i S_i^{A,B}(\mathbf{x})}{n},$$

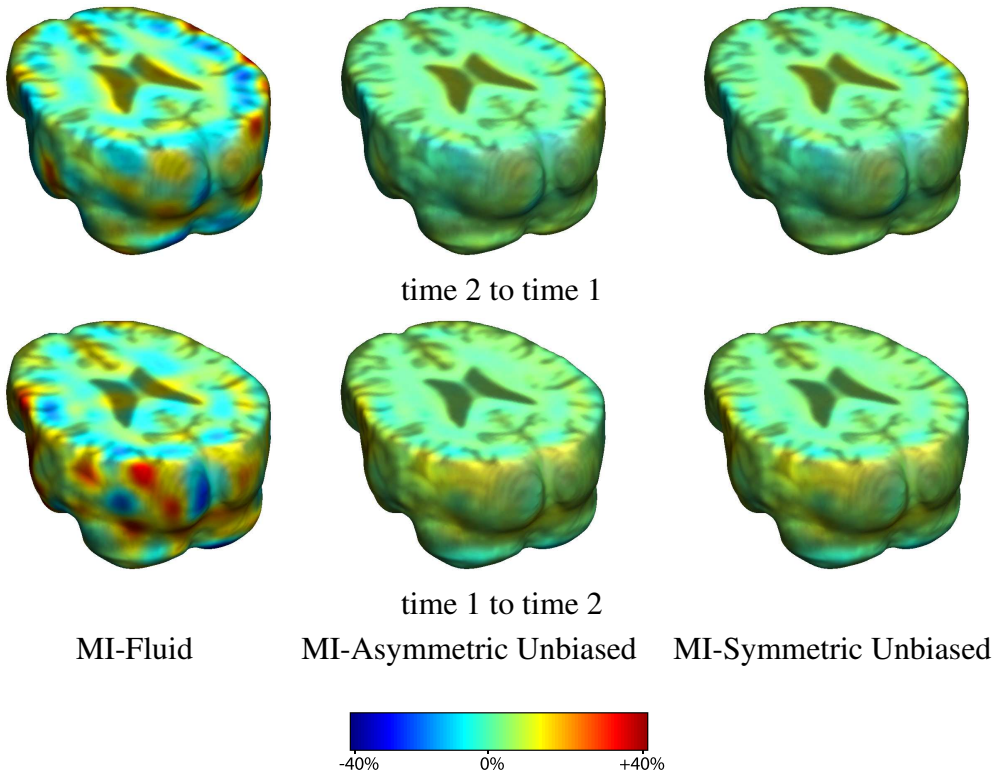



Figure 4.5: Nonrigid registration was performed on an image pair from one of the subjects from the ADNI Baseline study (serial MRI images acquired two weeks apart) using MI-Fluid (column 1), MI-Asymmetric Unbiased (column 2), and MI-Symmetric Unbiased (column 3) registration methods. Jacobian maps of deformations from time 2 to time 1 (row 1) and time 1 to time 2 (row 2) are superimposed on the target volumes. The unbiased methods generate less noisy Jacobian maps with values closer to 1; this shows the greater stability of the approach when no volumetric change is present.

and

$$\left(\sigma_{S^{A,B}}(\mathbf{x})\right)^2 = \frac{\sum_i \left(S_i^{A,B}(\mathbf{x}) - \overline{S^{A,B}(\mathbf{x})}\right)^2}{n-1},$$

thus follows the Student's $t$ distribution [37, 39, 43] under the null hypothesis and may be used to determine the $p$-value that the null hypothesis is true. If the alternative hypothesis is accepted, we confirm that sequence $B$ outperforms $A$ at point $\mathbf{x}$. Otherwise, we would rank $A$ and $B$ equally if the null hypothesis is not rejected.

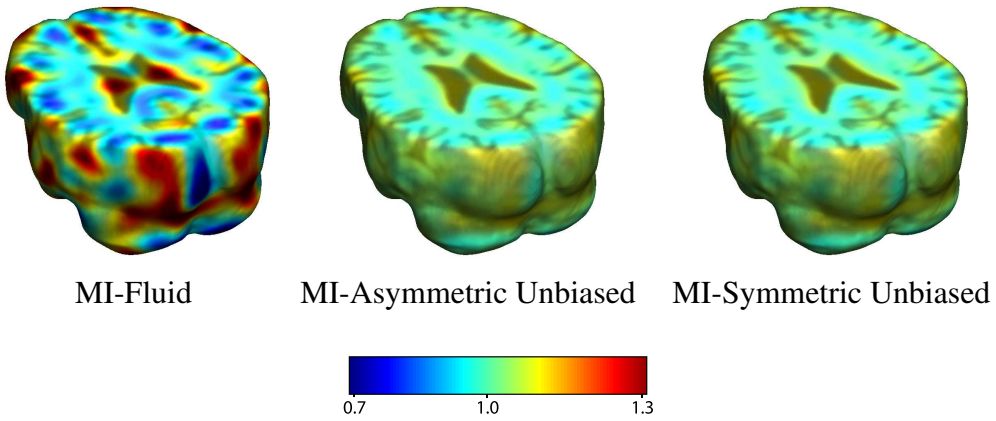



Figure 4.6: This figure examines the inverse consistency of deformation models. Products of Jacobian maps generated using all three models are shown, for forward direction (time 1 to time 2) and backward direction (time 2 to time 1). For the mutual information-based unbiased methods, the products of the Jacobian maps are less noisy, with values closer to 1, showing better inverse consistency.

### 4.4.2 Detecting Real Changes - Statistical testing on the mean log Jacobian

For both the ADNI follow-up dataset (in which patients are scanned twice with MRI, one year apart) and ADNI baseline dataset, we create a voxel-wise $t$ map using the local log Jacobian values of the ten subjects, allowing us to test the validity of the zero mean assumption. To simplify the notation, we introduce $J$ to denote $J := |D\mathbf{g}|$. The following voxel-wise $T$ statistic compared to a two-tailed Student's $t$ distribution may then be used to test the above null hypothesis

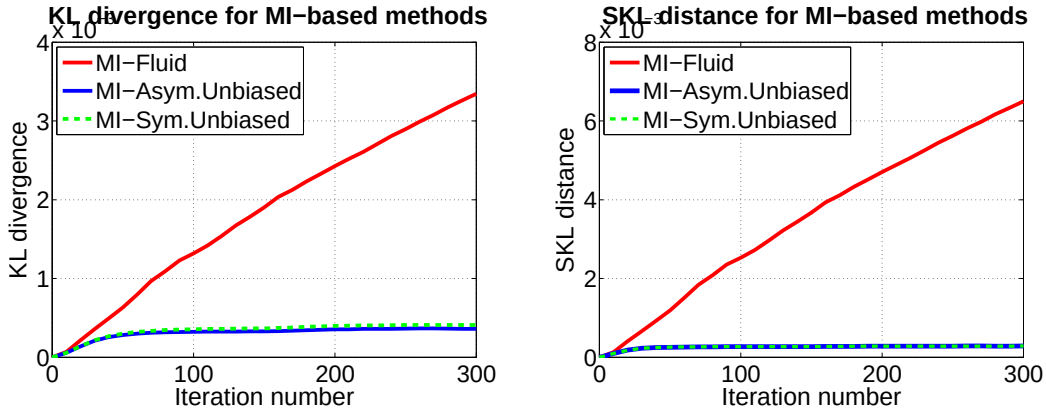
$$T_{\log J}(\mathbf{x}) = \frac{\sqrt{n} \cdot \overline{\log J(\mathbf{x})}}{\sigma_{\log J}(\mathbf{x})}, \tag{4.13}$$

where

$$\overline{\log J(\mathbf{x})} = \frac{\sum_i \log J_i(\mathbf{x})}{n},$$

and

$$\left(\sigma_{\log J}(\mathbf{x})\right)^2 = \frac{\sum_i \left(\log J_i(\mathbf{x}) - \overline{\log J(\mathbf{x})}\right)^2}{n-1}.$$



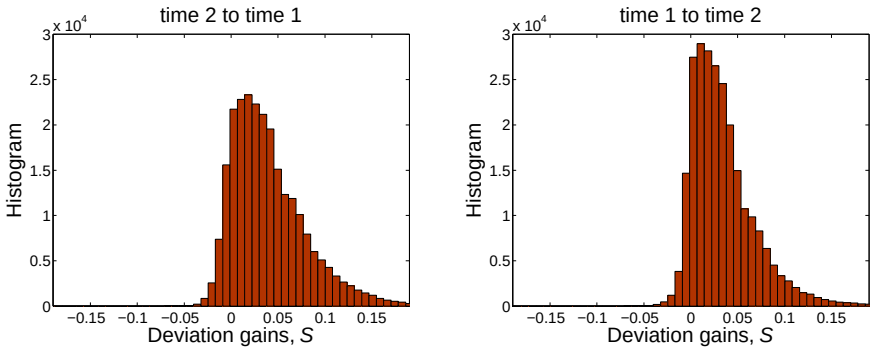

(a) KL divergence (b) SKL distance

Figure 4.7: (a) KL divergence and (b) SKL distance per iteration are shown for the MI-Fluid (solid red), MI-Asymmetric Unbiased (solid blue), and MI-Symmetric Unbiased (dashed green) methods. For MI-Fluid, both KL and SKL measures increase. Even though the Asymmetric Unbiased method explicitly minimizes the KL divergence, and the Symmetric Unbiased model minimizes the SKL distance, both the KL and SKL measures stabilize for both unbiased methods.

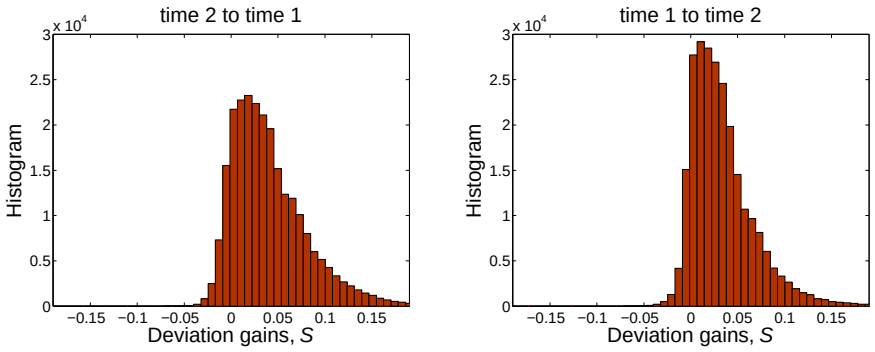
We reject the null hypothesis if the $p$ value calculated above exceeds a pre-set threshold based on a suitable confidence interval. Notice the voxel-wise variance of $\log J$ provides us with a way to assess the repeatability of a deformation method, i.e., measuring the voxel-wise spread of the given multiple observations (with higher variance corresponding to lower repeatability).

### 4.4.3 Permutation Testing to Correct Multiple Comparisons

To determine the overall global effects of different registration methods on the deviation of log Jacobian maps throughout the brain, we performed permutation tests to adjust for multiple comparisons [11, 77]. Following the analyses in [59], we resampled the observations by randomly flipping the sign of $S_i^{A,B}$ $(i = 1, 2, ..., n)$ under the null hypothesis. For each permutation, voxelwise $t$ tests are computed. We then




(a) MI-Fluid vs. MI-Asym.Unbiased




(b) MI-Fluid vs. MI-Sym.Unbiased

Figure 4.8: Histograms of voxel-wise deviation gains (a) MI-Fluid over MI- Asymmetric Unbiased and (b) MI-Fluid over MI-Symmetric Unbiased for one of the subjects, for the forward direction (time 2 to time 1) and backward direction (time 1 to time 2). The histograms are skewed to the right, indicating the superiority of Asymmetric Unbiased and Symmetric Unbiased registration methods over Fluid registration. Paired $t$ test shows significance ($p < 0.0001$).

compute the percentage of voxels inside the chosen ROI (in this case the ICBM mask) with $T$ statistics exceeding a certain threshold. The multiple comparisons corrected $p$ value may be determined by counting the number of permutations whose above-defined percentage exceeds that of the un-permuted observed data. This is comparable to 'set-level inference' in the widely-used SPM (Statistical Parametric Mapping) functional image analysis package [38]. For example, we say that sequence $B$ outperforms $A$ on the whole brain if this corrected $p$ value is smaller than 0.05 (that is, less than

**Baseline Study**

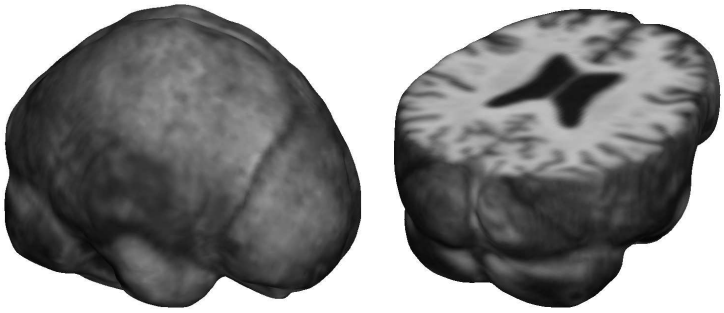


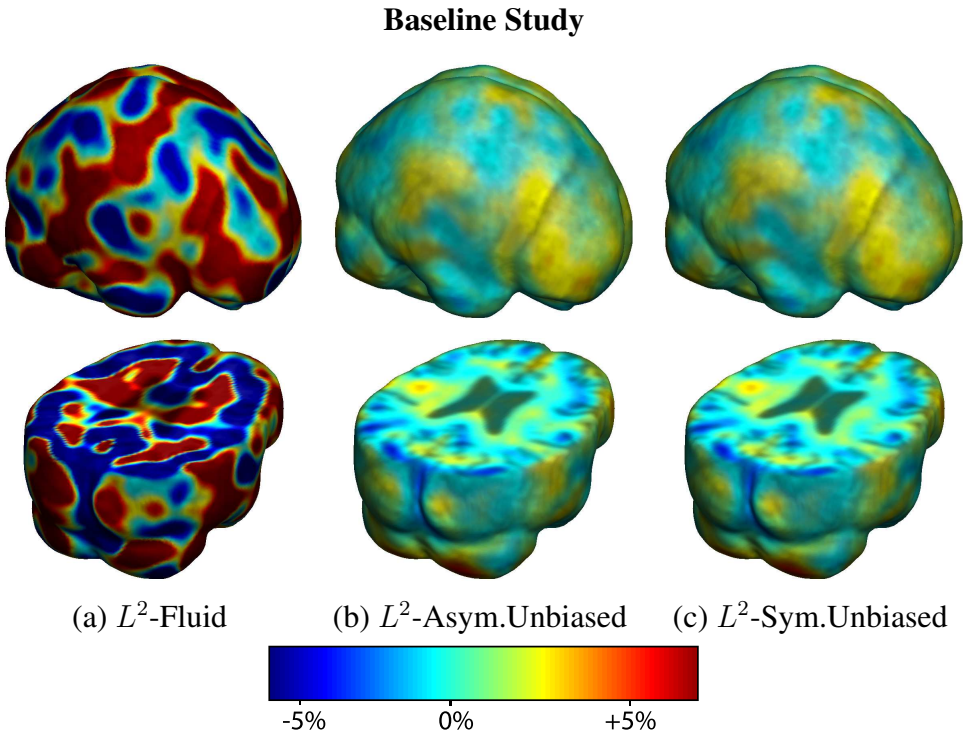
Figure 4.9: Volume from the ADNI Baseline dataset.

5% of all permutations have the above-defined percentage greater than that of the original data). In our study, with $n = 10$, all possible ($2^{10} = 1024$) permutations were considered in determining the final corrected $p$ value.

### 4.4.4 Cumulative Distribution Function (CDF)

To visually assess the global significance level of the voxel-wise $t$ tests on deviation gains and log-Jacobian values, we also employed the cumulative distribution function (CDF) plot, as in several prior studies [10, 18, 62, 74]. In brief, we plot observed cumulative probabilities against the theoretical distribution under the null hypothesis. These CDF plots are commonly created as an intermediate step, when using the false discovery rate (FDR) method to assign overall significance values to statistical maps [4, 5, 40, 93, 94, 116]. As they show the proportion of supra-threshold voxels in a statistical map, for a range of thresholds, these CDF plots (sometimes called Q-Q plots) offer a measure of the effect size in a statistical map. They also may be used to demonstrate which methodological choices influence the effect size in a method that creates statistical maps [10, 18, 62].

In the case of deviation gains $S$ of a worse technique $A$ over a better technique $B$ in




(a) $L^2$-Fluid (b) $L^2$-Asym.Unbiased (c) $L^2$-Sym.Unbiased

Figure 4.10: Nonrigid registration was performed on the ADNI Baseline study (serial MRI images acquired two weeks apart) of ten normal elderly subjects using $L^2$-Fluid (column 1), $L^2$-Asymmetric Unbiased (column 2), $L^2$-Symmetric Unbiased (column 3) registration methods. For each method, the mean of the resulting 10 Jacobian maps is superimposed on one of the brain volumes. Visually, $L^2$-Fluid generates a noisy mean map, while maps generated using $L^2$-Asymmetric Unbiased and $L^2$-Symmetric Unbiased methods are less noisy with values closer to 1. For all deformation models, regions with least stability, due to both spatial distortion and intensity inhomogeneity, are the brain stem, thalamus, and ventricles.

the ADNI baseline data, we expect a CDF curve to lie above the Null line, in the sense that a better technique exhibits less systematic changes. In the case of log-Jacobian values, a better registration technique, on the other hand, should be able to separate the CDF curves between ADNI baseline and follow-up phases (this is what we refer to as the separation of CDF curves in the presence of real physiological changes).

Table 4.1: Global $T$ statistics for all ten subjects testing whether Symmetric Unbiased registration (method B) outperforms Fluid registration (method A) when coupled with $L^2$.

| Subject # | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $\overline{S^{A,B}}$ | 0.0639 | 0.0337 | 0.0758 | 0.101 | 0.0968 |
| $\sigma^2_{S^{A,B}}$ | 0.00342 | 0.00102 | 0.00567 | 0.00906 | 0.00926 |
| $T_{S^{A,B}}$ | 542 | 524 | 499 | 525 | 499 |
| | | | | | |
| Subject # | 6 | 7 | 8 | 9 | 10 |
| $\overline{S^{A,B}}$ | 0.0581 | 0.473 | 0.583 | 0.288 | 0.238 |
| $\sigma^2_{S^{A,B}}$ | 0.00416 | 0.185 | 0.386 | 0.0618 | 0.0515 |
| $T_{S^{A,B}}$ | 447 | 546 | 465 | 575 | 520 |

## 4.5 Results

In this section, we tested the Asymmetric Unbiased and Symmetric Unbiased models and compared the results to those obtained using the Fluid registration model [21, 28]. Of note, even though Asymmetric Unbiased and Symmetric Unbiased methods minimize different energy functionals, our experiments showed that they generate very similar maps. For each regularization technique, we employed both $L^2$ and mutual information matching functionals (see equations (4.6)-(4.9)).

To obtain a fair comparison, re-gridding was not employed. Re-gridding is a method to relax the energy computed from the linear elasticity prior after a certain number of iterations, which allows large-deformation mappings to be recovered without any absolute penalty on the displacement field (other than via the smoothness constraint on the velocity field which is integrated to give the displacement) [21]. It is essentially a memory-less procedure, as how images are matched after each re-gridding is independent of the final deformation before the re-gridding, rendering the comparison of final Jacobian fields and cost functionals problematic. Moreover, we consider

Table 4.2: Global $T$ statistics for all ten subjects testing whether Symmetric Unbiased registration (method B) outperforms Fluid registration (method A) when coupled with mutual information.

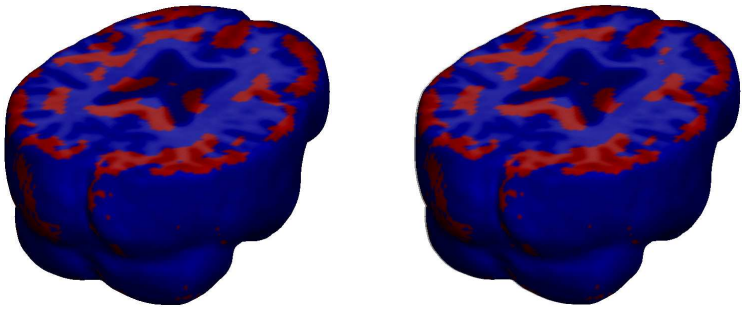
| Subject # | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $\overline{S^{A,B}}$ | 0.0697 | 0.0262 | 0.0399 | 0.0342 | 0.0379 |
| $\sigma^2_{S^{A,B}}$ | 0.00579 | 0.000774 | 0.00156 | 0.00138 | 0.00138 |
| $T_{S^{A,B}}$ | 455 | 468 | 501 | 456 | 505 |
| | | | | | |
| Subject # | 6 | 7 | 8 | 9 | 10 |
| $\overline{S^{A,B}}$ | 0.0820 | 0.0853 | 0.0774 | 0.0489 | 0.0773 |
| $\sigma^2_{S^{A,B}}$ | 0.00708 | 0.00845 | 0.00698 | 0.00232 | 0.00529 |
| $T_{S^{A,B}}$ | 484 | 460 | 460 | 504 | 527 |

the strategy of re-gridding, through the relaxation of deformation fields over time, to be less rigorous from a theoretical standpoint, as the imposition of a regularizer can be used to secure distributional properties in the resulting statistics (e.g., symmetric log-Jacobian).

Uniform values of $\lambda = 500$ in equation (4.7) and $\lambda = 1000$ in equation (4.6) were used for all deformations using $L^2$-Symmetric Unbiased and $L^2$-Asymmetric Unbiased algorithms, respectively. Also, $\lambda = 5$ in equation (4.9) and $\lambda = 10$ in equation (4.8) were used for deformations using MI-Symmetric Unbiased and MI-Asymmetric Unbiased methods. Since the Asymmetric Unbiased model quantifies only the forward deformation, the weight of the corresponding regularization functional is half the magnitude of that of the Symmetric Unbiased model, and hence, a weighting parameter twice as large should be used.

Some of the figures in this chapter were generated using BrainSuite visualization tool [89].

**Baseline Study**



$L^2$-Fluid vs. $L^2$-Asym.Unbiased  $L^2$-Fluid vs. $L^2$-Sym.Unbiased

Figure 4.11: Voxel-wise paired $t$ test for the deviation gain $S$ empirically thresholded at 2.82 ($p = 0.005$ on the voxel level with 9 degrees of freedom), showing where $L^2$-Asymmetric Unbiased and $L^2$-Symmetric Unbiased registration outperform $L^2$-Fluid registration (regions in red) with statistical significance on a voxel level. In contrast, there are no voxels with $T$ values smaller than -2.82, indicating that Fluid registration does not outperform unbiased methods at any voxel. Hence, the visualization of voxel-wise paired $t$ test with a threshold of -2.82 is omitted.

### 4.5.1 ADNI Baseline Scans

In this section, nonlinear registration was performed on a dataset that we shall refer to as the "ADNI Baseline" dataset, collected during the preparatory phase of the ADNI project, which includes serial MRI images of ten normal elderly subjects acquired two weeks apart. Each of the ten pairs of scans is represented on a $128 \times 160 \times 128$ grid. Here, the foundation of calibrations is based on the assumption that, by scanning normal control human subjects serially within a two-week period using the same MRI protocol, no systematic structural changes should be recovered.

In our first experiment, we compared methods based on $L^2$ matching ($L^2$-Fluid, $L^2$-Asymmetric Unbiased, and $L^2$-Symmetric Unbiased). Figures 4.1-4.4 show the results of registering a pair of serial MRI images for one of the subjects (subject 3). The deformation was computed in both directions (time 2 to time 1, and time 1 to time 2) using methods based on $L^2$ matching. In Figure 4.1, Jacobian maps of deformations

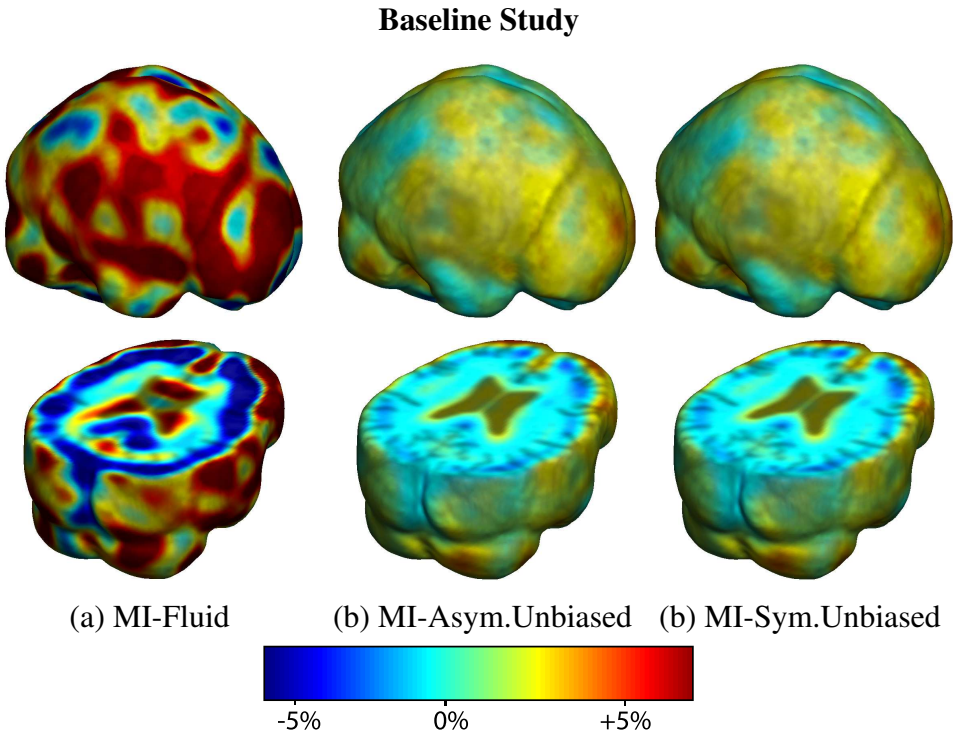
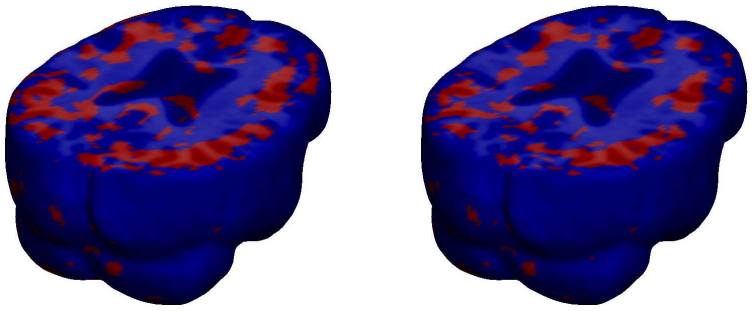



Figure 4.12: Nonrigid registration was performed on the ADNI Baseline study (serial MRI images acquired two weeks apart) of ten normal elderly subjects using MI-Fluid (column 1), MI-Asymmetric Unbiased (column 2), MI-Symmetric Unbiased (column 3) registration methods. For each method, the mean of the resulting 10 Jacobian maps is superimposed on one of the brain volumes. Visually, MI-Fluid generates a noisy mean map, while maps generated using MI-Asymmetric Unbiased and MI-Symmetric Unbiased methods are less noisy with values closer to 1. For all deformation models, regions with least stability, due to both spatial distortion and intensity inhomogeneity, are the brain stem, thalamus, and ventricles.

are superimposed on brain volumes. Both Asymmetric Unbiased and Symmetric Unbiased methods generate less noisy Jacobian maps with values closer to the identity mapping, which shows the superior stability of the Unbiased approach in the absence of physiological changes. In Figure 4.2, we visually assessed the *inverse consistency* of the mappings [20] by concatenating forward and backward Jacobian maps (in an ideal situation, this operation should yield the identity). Again, we observe noticeable visual differences between the results obtained using the unbiased methods and Fluid

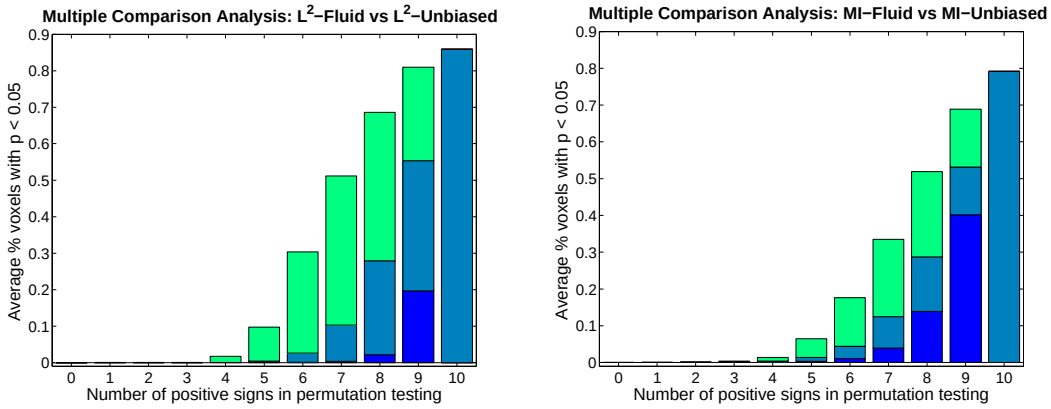



MI-Fluid vs. MI-Asym.Unbiased MI-Fluid vs. MI-Sym.Unbiased

Figure 4.13: Voxel-wise paired $t$ test for the deviation gain $S$ empirically thresholded at 2.82 ($p = 0.005$ on the voxel level with 9 degrees of freedom), showing where MI-Asymmetric Unbiased and MI-Symmetric Unbiased registration outperform MI-Fluid registration (regions in red) with statistical significance on a voxel level. In contrast, there are no voxels with $T$ values smaller than -2.82, indicating that Fluid registration does not outperform unbiased methods at any voxel. Hence, the visualization of voxel-wise paired $t$ test with a threshold of -2.82 is omitted.

registration. Figure 4.3 plots the KL divergence and SKL distance measures for each of the $L^2$-based methods. For $L^2$-Fluid method, both KL and SKL measures increase with increasing numbers of iterations. On the other hand, even though the Asymmetric Unbiased method minimizes the KL divergence and the Symmetric Unbiased model minimizes the SKL distance, these two measures stabilize for both unbiased methods. Figure 4.4 shows the histograms of voxel-wise deviation gains of $L^2$-Fluid over $L^2$-Asymmetric Unbiased as well as $L^2$-Fluid over $L^2$-Symmetric Unbiased. The histograms are skewed to the right, indicating the superiority of both unbiased registration methods over Fluid registration.

Of note, we have also considered a different deviation map, defined as $dev_2(\mathbf{x}) = \big||D\mathbf{g}(\mathbf{x})| - 1\big|$, in place of (4.10). We performed statistical analyses with this definition of deviation gain, which yielded very similar results. These results are therefore not shown here.

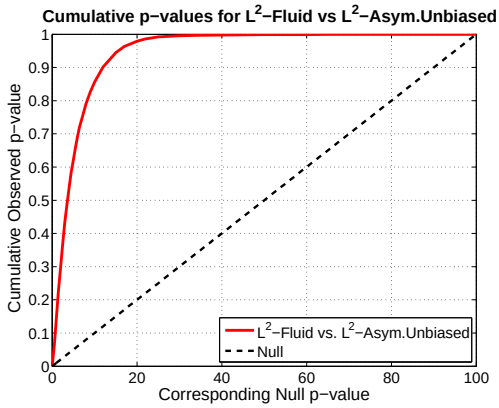



(a) $L^2$-Fluid vs. $L^2$-Symmetric Unbiased (b) MI-Fluid vs. MI-Symmetric Unbiased

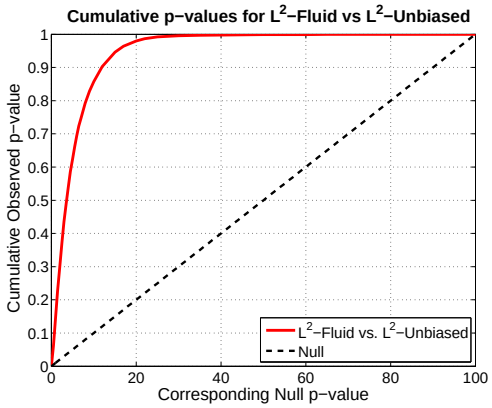
Figure 4.14: Multiple Comparison Analysis using permutation testing on the deviation gain $S$ of (a) $L^2$-Fluid over $L^2$-Symmetric Unbiased and (b) MI-Fluid over MI-Symmetric Unbiased, both for baseline ADNI dataset. Each permutation randomly assigns a positive or negative sign to each of the 10 log-Jacobian maps. Here, results are plotted with respect to the number of positive signs (from 0 to 10) with 10 positive signs indicating the observed data. Dark blue, light blue, and green colors indicate the minimum, average, and maximum percentage of voxels with $p < 0.05$ of all possible permutations with a given number of positive signs. There is only one observation for the observed data, and thus, minimum, maximum, and average values are equal for the rightmost bar. The result indicates that out of 1024 permutations, no permutation gives a greater percentage of voxels with $p < 0.05$ than the observed data does. This indicates that unbiased regularization technique outperforms Fluid methods with $p < 0.001$. Since the results obtained using Asymmetric Unbiased method are similar to those obtained using Symmetric Unbiased method, they are not shown here.

In Table 4.1, we compared $L^2$-Fluid and $L^2$-Symmetric Unbiased methods, conducting a within-subject paired $t$ test inside the ICBM mask for each of the ten subjects. In this case, $p < 0.0001$ for all subjects, indicating that the Symmetric Unbiased registration, when coupled with $L^2$ matching cost functional, produces more reproducible maps with less variability.

Figure 4.10 shows the mean Jacobian maps obtained using $L^2$-Fluid, $L^2$-Asymmetric Unbiased, and $L^2$- Symmetric Unbiased registration algorithms. Jacobian maps gen-




(a) $L^2$-Fluid vs. $L^2$-Asym.Unbiased



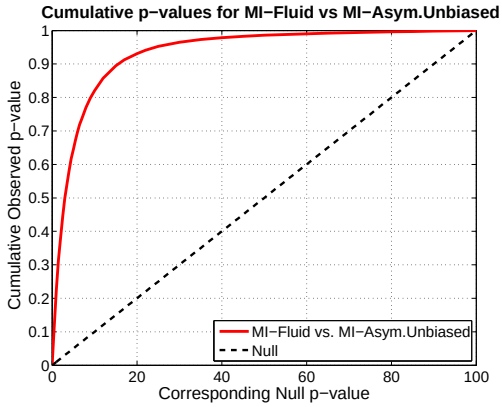

(b) $L^2$-Fluid vs. $L^2$-Sym.Unbiased

Figure 4.15: Cumulative distribution of $p$-values for the deviation gain $S$ of (a) $L^2$-Fluid over $L^2$-Asymmetric Unbiased and (b) $L^2$-Fluid over $L^2$-Symmetric Unbiased. Here, the ADNI baseline dataset is used. In both (a) and (b), the CDF line is well above the Null line ($y = x$), indicating that both asymmetric and symmetric unbiased methods outperform Fluid method (i.e. less deviation) in being less likely to exhibit structural change in the absence of biological change. Note that the interval $p \in [0, 0.05]$ is of most importance for observation.

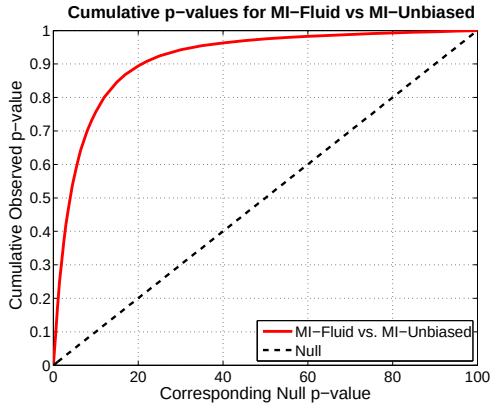
erated using unbiased models have values closer to 1, whereas $L^2$-Fluid model generated noisy mean maps. Figure 4.11, shows the results when performing 3D voxel-wise paired $t$ tests for the deviation gain of $L^2$-Fluid over $L^2$-Asymmetric Unbiased and $L^2$-Fluid over $L^2$-Symmetric Unbiased. $T$ maps for the deviation gains are empirically thresholded at 2.28 ($p = 0.005$ on the voxel level with 9 degrees of freedom) to show statistical significance.

Figure 4.14(a) shows results obtained using Multiple Comparison Analysis with permutation testing on deviation gains of $L^2$-Fluid over $L^2$-Symmetric Unbiased. The results indicate that out of 1024 permutations, no permutation yields a larger percentage of voxels with $p < 0.05$ than the observed data, which indicates that $L^2$-Symmetric Unbiased method outperforms $L^2$-Fluid with $p < 0.001$.

To emphasize the differences between the distributions of log Jacobian values for




(a) MI-Fluid vs. MI-Asym.Unbiased




(b) MI-Fluid vs. MI-Sym.Unbiased

Figure 4.16: Cumulative distribution of $p$-values for the deviation gain $S$ of (a) MI-Fluid over MI-Asymmetric Unbiased and (b) MI-Fluid over MI-Symmetric Unbiased. Here, ADNI baseline dataset is used. In both (a) and (b), the CDF line is well above the Null line, indicating that both asymmetric and symmetric unbiased methods outperform Fluid method in being less likely to exhibit structural change in the absence of biological change. Note that the interval $p \in [0, 0.05]$ is of most importance for observation.

Fluid and unbiased (both asymmetric and symmetric) methods, in Figure 4.15, we plotted the cumulative distribution function of the $p$-values in deviation gains as defined in equation (4.11). In these plots, the interval $p \in [0, 0.05]$ is of most importance for observation. For a null distribution, this cumulative plot falls along the line $y = x$ in $xy$-plane, as represented by the dashed black line. Larger upward inflections of the CDF curve near the origin are associated with significant deviation gains, indicating that both Asymmetric Unbiased and Symmetric Unbiased methods outperform Fluid method in being less likely to exhibit structural changes in the absence of systematic biological changes.

In our second experiment, we compared the performance of methods based on mutual information matching (MI-Fluid, MI-Asymmetric Unbiased, and MI- Symmetric Unbiased). As for methods based on $L^2$ matching, Figures 4.5-4.8 demonstrate MI-

Table 4.3: Global $T$ statistics for all ten subjects testing whether MI-Fluid (method B) outperforms $L^2$-Fluid (method A).

| Subject # | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $\overline{S^{A,B}}$ | 0.000997 | 0.00435 | 0.0520 | 0.0971 | 0.0721 |
| $\sigma^2_{S^{A,B}}$ | 0.00819 | 0.00121 | 0.00529 | 0.0156 | 0.0118 |
| $T_{S^{A,B}}$ | 5 | 62 | 355 | 385 | 329 |
| | | | | | |
| Subject # | 6 | 7 | 8 | 9 | 10 |
| $\overline{S^{A,B}}$ | 0.0269 | 0.445 | 0.566 | 0.274 | 0.196 |
| $\sigma^2_{S^{A,B}}$ | 0.00406 | 0.190 | 0.372 | 0.0694 | 0.0457 |
| $T_{S^{A,B}}$ | 210 | 506 | 461 | 516 | 454 |

Asymmetric Unbiased and MI-Symmetric Unbiased methods to produce inverse consistent maps with less variability. Table 4.2 shows the results of within-subject paired $t$ test for all ten subjects. The results, based on mean Jacobian maps for all subjects, show that the Unbiased regularization technique outperforms Fluid registration with confirmed statistical significance (Figures 4.12, 4.13, 4.14(b), and 4.16).

Lastly, we compared $L^2$ and mutual information cost functionals for both Fluid and Symmetric Unbiased regularization. (Since Asymmetric Unbiased and Symmetric Unbiased regularizations produce similar results, we do not show the results for the asymmetric version). We again conducted within-subject paired $t$ tests (Tables 4.3 and 4.4) as well as group paired $t$ tests (Figure 4.17) on the voxel-wise deviation gains for all voxels inside the ICBM brain mask. We showed that MI-Fluid outperforms $L^2$-Fluid with $p < 0.0001$. However, the result of the comparison of $L^2$-Symmetric Unbiased and MI-Symmetric Unbiased is inconclusive. In other words, mutual information performs better when coupled with Fluid registration, but there is no statistical difference between mutual information and $L^2$ when the Symmetric Unbiased method is used.

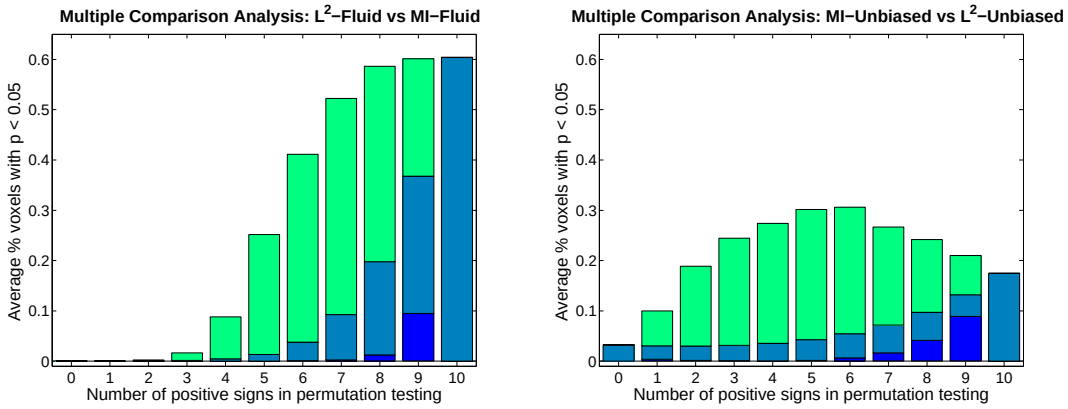
Table 4.4: Global $T$ statistics for all ten subjects testing whether $L^2$-Symmetric Unbiased (method B) outperforms MI-Symmetric Unbiased (method A).

| Subject # | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $\overline{S^{A,B}}$ | 0.00648 | 0.00314 | -0.0161 | -0.0244 | -0.0132 |
| $\sigma^2_{S^{A,B}}$ | 0.000458 | 0.000111 | 0.000424 | 0.00155 | 0.000593 |
| $T_{S^{A,B}}$ | 150 | 148 | -387 | -307 | -269 |
| | | | | | |
| Subject # | 6 | 7 | 8 | 9 | 10 |
| $\overline{S^{A,B}}$ | 0.0101 | -0.0568 | -0.0746 | -0.0352 | -0.0347 |
| $\sigma^2_{S^{A,B}}$ | 0.000262 | 0.00323 | 0.00367 | 0.00133 | 0.00121 |
| $T_{S^{A,B}}$ | 309 | -497 | -611 | -478 | -495 |

To explain this result, we postulate that by constraining the deformations less (i.e., as in Fluid registration), assuming intensity 1-to-1 correspondence (i.e., matching using $L^2$) may lead to local oscillations of the deformation maps, as minimizing $L^2$ forces a local search for the smallest intensity differences. One result of this is a Jacobian map with locally extreme values, translating into spurious signals and, in our case, less reproducibility. On the other hand, the Symmetric Unbiased method eliminates local oscillations, allowing globally better matching when intensity 1-to-1 correspondence can be assumed (i.e., when $L^2$ is applicable as a data fidelity term).

### 4.5.2 ADNI Follow-up Scans

In this section, we analyze a dataset we shall call the "ADNI Follow-up" phase dataset, which includes serial MRI images ($220 \times 220 \times 220$) of ten subjects acquired one year apart. These data were collected as part of a larger study to track degenerative brain changes in MRI in 800 subjects, ages 55 to 90, including 200 elderly controls, 400 subjects with mild cognitive impairment, and 200 patients with AD. As the images are now one year apart, real anatomical changes are present, which allows methods to be




(a) $L^2$-Fluid vs. MI-Fluid (b) MI-Sym.Unbiased vs. $L^2$-Sym.Unbiased

Figure 4.17: Multiple Comparison Analysis using permutation testing on the deviation gain $S$ of (a) $L^2$-Fluid over MI-Fluid and (b) MI-Symmetric Unbiased over $L^2$-Symmetric Unbiased, both for baseline ADNI dataset. Each permutation randomly assigns positive or negative sign to each of the 10 log-Jacobian maps. Here, results are plotted with respect to the number of positive signs (from 0 to 10) with 10 positive signs indicating the observed data. Dark blue, light blue, and green colors indicate the minimum, average, and maximum percentage of voxels with $p < 0.05$ of all possible permutations with a given number of positive signs. There is only one observation for the observed data, and thus, minimum, maximum, and average values are equal for the rightmost bar. The result in (a) indicates that out of 1024 permutations, no permutation gives a greater percentage of voxels with $p < 0.05$ than the observed data does. This indicates that MI-Fluid method outperforms $L^2$-Fluid method with $p < 0.001$. However, the comparison of MI-Symmetric Unbiased and $L^2$-Symmetric Unbiased in (b) is inconclusive. Since the results obtained using Asymmetric Unbiased method are similar to those obtained using Symmetric Unbiased method, they are not shown here.

compared in the presence of true biological changes.

In Figure 4.19, nonlinear registration was performed using Fluid, Asymmetric Unbiased, and Symmetric Unbiased methods coupled with $L^2$ matching. Visually, the Fluid method generates noisy mean Jacobian maps, while maps generated using unbiased methods suggest a volume reduction in gray matter as well as ventricular enlargement. Here, both Asymmetric Unbiased and Symmetric Unbiased methods perform equally well. Figure 4.21 displays the cumulative distribution of $p$-values for

the voxel-wise log Jacobian $t$-maps for both ADNI Baseline and ADNI Follow-up datasets. We expect a better method to separate these two CDF curves, indicating that a real biological change has occurred between the two time points. A greater separation is accomplished when Asymmetric Unbiased and Symmetric Unbiased methods are used, while the Fluid method does not differentiate between the two datasets. Similar results are obtained using mutual information based methods (Figure 4.20).

### 4.5.3 Large-Scale ADNI Study

We also performed a large-scale study using 100 pairs of serial MR images, acquired 12 months apart, from the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset. The selected sample consisted of 20 patients with Alzheimer's disease (AD), 40 individuals with mild cognitive impairment (MCI), and 40 healthy elderly controls (CTL) [61]. Here, the images were registered using Symmetric Unbiased registration coupled with $L^2$ matching. In Figure 4.22, the mean of the resulting Jacobian maps in each group is superimposed on a brain volume. The resulting statistical maps suggest widespread progressive atrophy throughout the entire brain and expansion of ventricular and CSF spaces in AD (compared to controls), and a more restricted pattern of atrophy in MCI.

## 4.6 Conclusion

In the past decade, information theory has been studied extensively in computational imaging. However, there have been few rigorous studies to date that investigate the statistical aspect of the resulting deformation fields. Different regularization techniques have been proposed, sometimes generating deformations very different from one another.

**Follow-up Study**

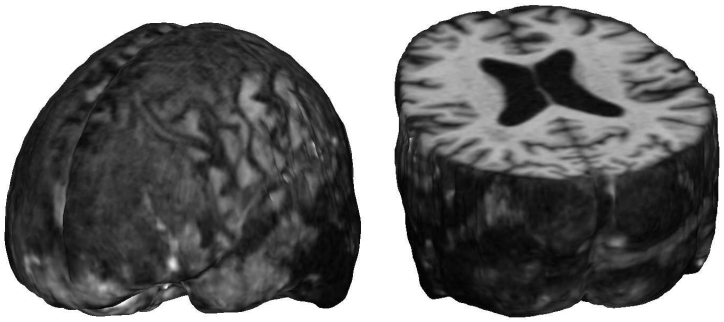
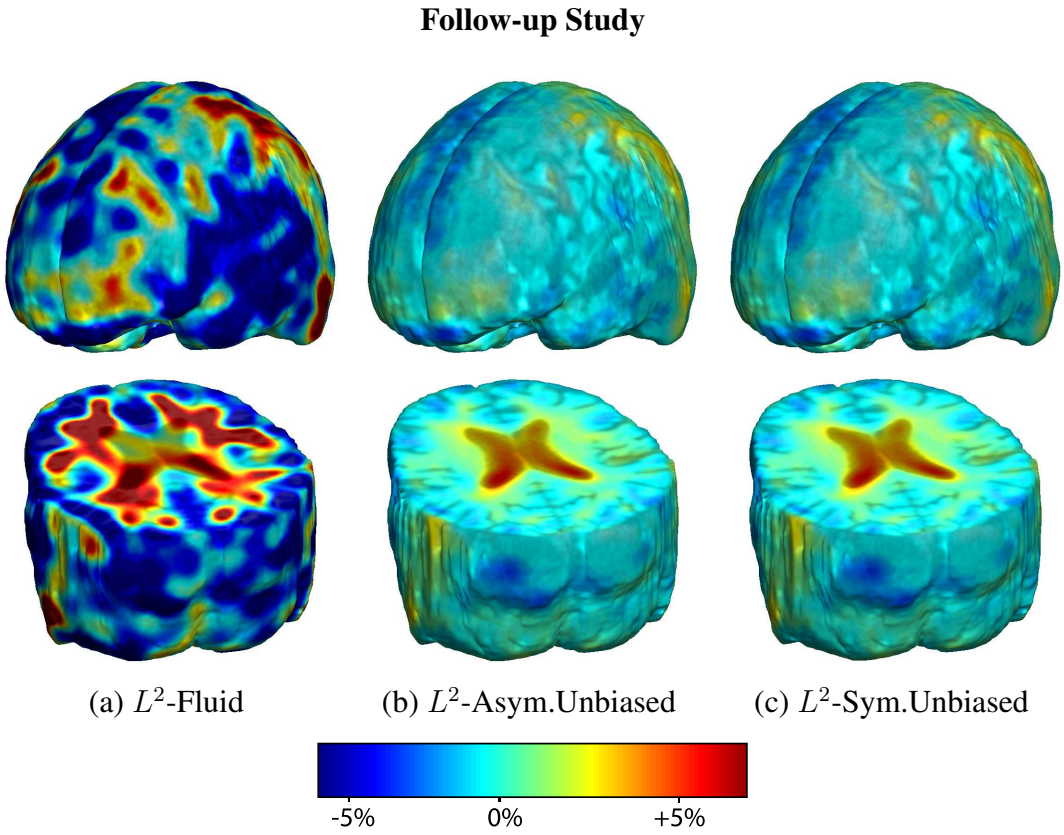


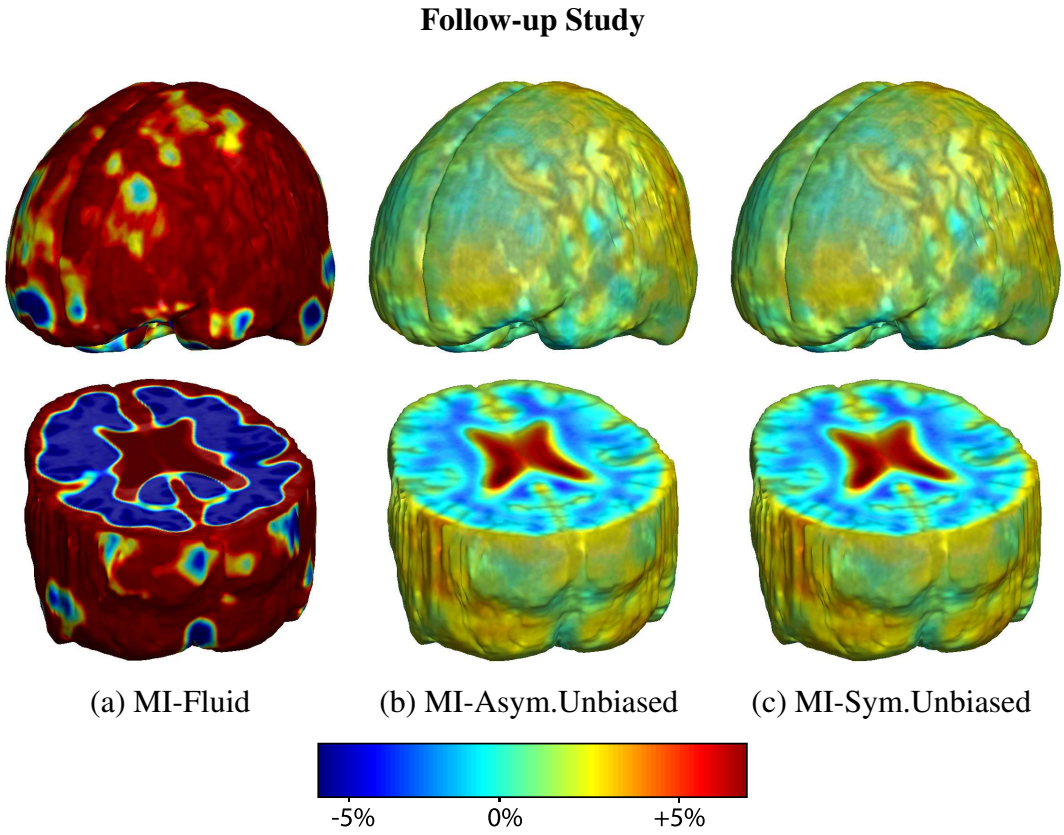
Figure 4.18: Volume from the ADNI Follow-up dataset.

In this chapter, we examined the power of different nonrigid registration models to detect changes in tensor-based morphometry, and their stability when no real changes are present. It is the first work to systematically investigate the reproducibility and variability of different registration methods in TBM. This chapter introduced a novel asymmetric unbiased registration model (the Asymmetric Unbiased model), which produces results that are very similar to those obtained using the previously introduced Symmetric Unbiased model. We showed that Asymmetric Unbiased and Symmetric Unbiased models perform significantly better than the fluid registration technique. Although various techniques have been extensively applied to detect disease effects and monitor brain changes with TBM, this work is the first calibration study to compare registration models for tensor-based morphometry. We believe our results are important, as they provide greater insight into the interpretation of TBM results in the future.




(a) $L^2$-Fluid (b) $L^2$-Asym.Unbiased (c) $L^2$-Sym.Unbiased

Figure 4.19: Nonrigid registration was performed on the ADNI Follow-up study (serial MRI images acquired 12 months apart) using $L^2$-Fluid (column 1), $L^2$-Asymmetric Unbiased (column 2), and $L^2$-Symmetric Unbiased (column 3) registration methods. For each method, the mean of the resulting 10 Jacobian maps is superimposed on one of the brain volumes. Visually, $L^2$-Fluid generates a noisy mean map, while maps generated using the $L^2$-Asymmetric Unbiased and $L^2$-Symmetric Unbiased methods suggest a volume reduction in gray matter as well as ventricular enlargement.




(a) MI-Fluid (b) MI-Asym.Unbiased (c) MI-Sym.Unbiased

Figure 4.20: Nonrigid registration was performed on the ADNI Follow-up study (serial MRI images acquired 12 months apart) using MI-Fluid (column 1), MI-Asymmetric Unbiased (column 2), and MI-Symmetric Unbiased (column 3) registration methods. For each method, the mean of the resulting 10 Jacobian maps is superimposed on one of the brain volumes. Visually, MI-Fluid generates a noisy mean map, while maps generated using the MI-Asymmetric Unbiased and MI-Symmetric Unbiased methods suggest a volume reduction in gray matter as well as ventricular enlargement.

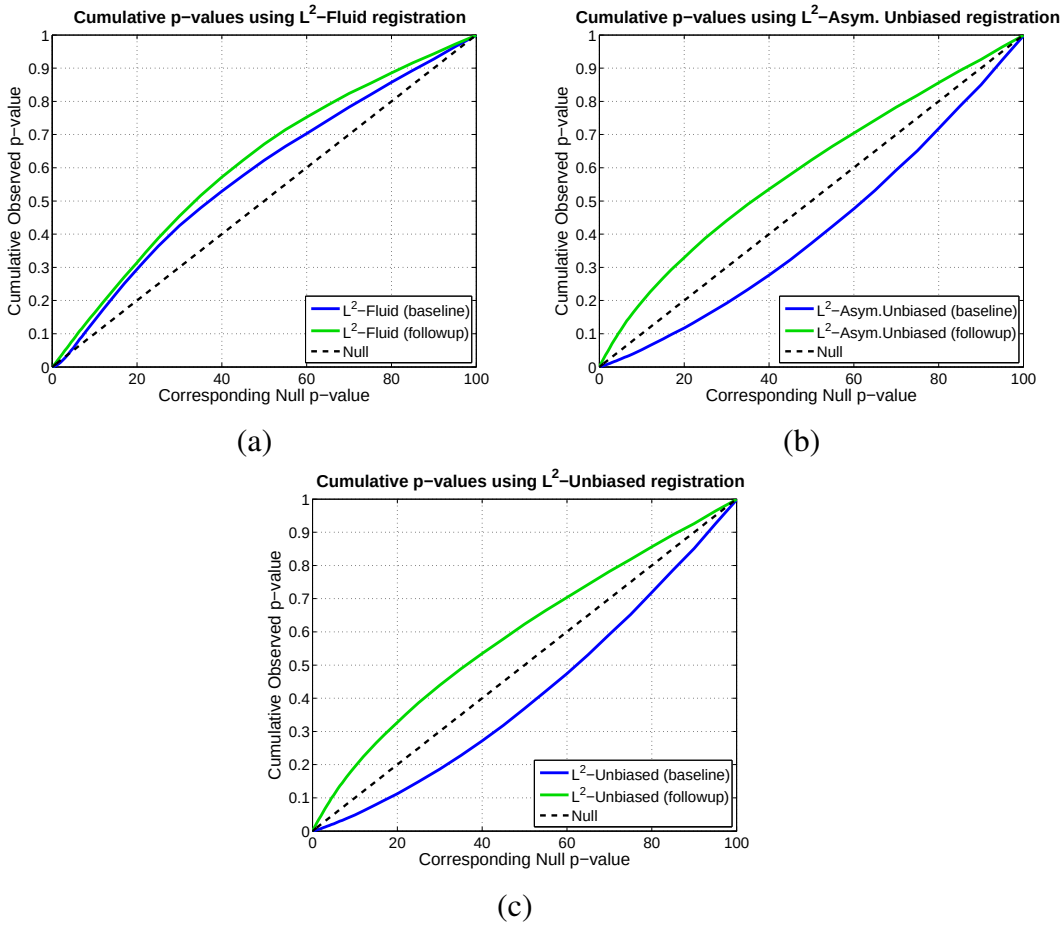



Figure 4.21: Cumulative distribution of $p$-values for the voxelwise log Jacobian $t$-maps (as defined in Equation (4.13)) for both ADNI Baseline (in blue) and Follow-up (in green) using (a) $L^2$-Fluid, (b) $L^2$-Asymmetric Unbiased, and (c) $L^2$-Symmetric Unbiased methods. Here, a better method should separate these two CDF plots (see Section 4.4.4) with the Null line in between, indicating a real biological change has occurred between these two time points. Hence, $L^2$-Asymmetric Unbiased and $L^2$-Symmetric Unbiased methods outperform $L^2$-Fluid method. Note that the interval $p \in [0, 0.05]$ is of most importance for observation.

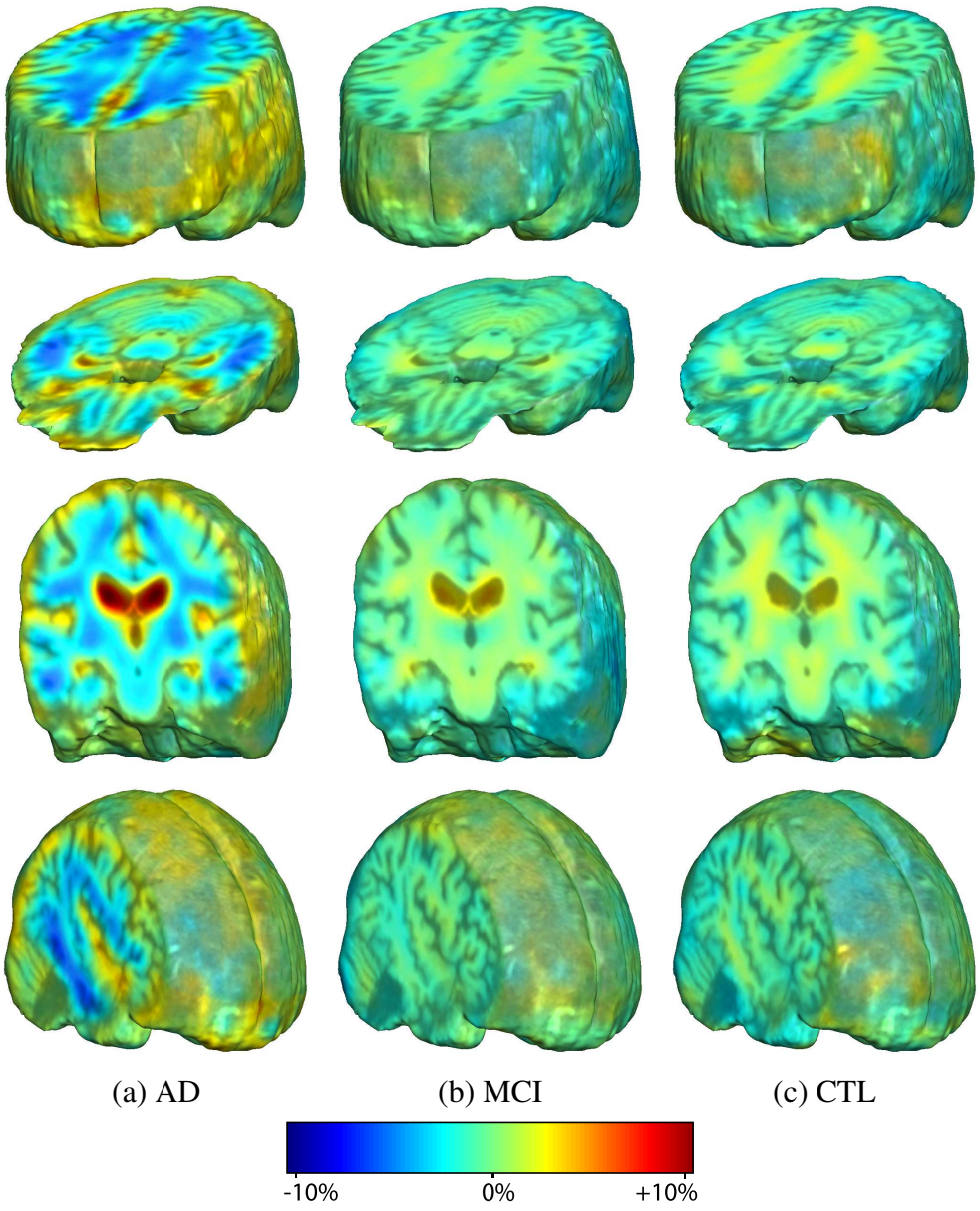



Figure 4.22: Unbiased registration with $L^2$ matching was performed on 100 pairs of serial MR images, acquired 12 months apart, from the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset. The selected sample consisted of 20 patients with Alzheimer's disease (AD), 40 individuals with mild cognitive impairment (MCI), and 40 healthy elderly controls (CTL). The mean of the resulting Jacobian maps in each group is superimposed on a brain volume.

# CHAPTER 5

# Multiphase Segmentation of Deformation using Logarithmic Priors

In Chapter 2, we described the proposed unbiased large deformation nonlinear image registration model which has been successfully used to obtain theoretically and intuitively correct deformation maps. In this chapter, we extend this idea to simultaneously registering and tracking deforming objects in a sequence of two or more images. A level set based Chan-Vese multiphase segmentation model is generalized to consider Jacobian fields while segmenting regions of growth and shrinkage in deformations. Deforming objects are thus classified based on magnitude of homogeneous deformation.

## 5.1 Introduction

Segmentation of homogeneous deformation is a challenging problem which incorporates several image processing and computer vision areas including image registration, segmentation, and tracking. The goal of deformation segmentation is to classify regions of homogeneous volume/density change based on magnitude of such change. In this work, we employ a robust image registration model for generating priors for further segmentation of deformed features.

Image registration models are used to align, or spatially normalize, one image to

match another. Presented with a choice of an image registration method, it is important to ensure that the model in consideration generates meaningful deformation fields. In general, the transformation that defines the correspondence map between the images should be diffeomorphic, to preserve the topology. As was described in Chapter 2, not all widely used image registration models generate theoretically and intuitively correct deformation fields. In Chapter 2, we proposed information theory approach to quantifying deformation and described the framework for constructing large deformation diffeomorphic image registration models.

Even though a robust nonlinear registration model can generate meaningful deformation maps, an automated segmentation process would be required to classify regions of homogeneous deformation in a sequence of images. Next, we will explain how to generalize level set based *active contour without edges* model [14] to classify regions of homogeneous deformation.

## 5.2 Segmentation of Deforming Objects

The unbiased registration model described in Chapter 2, and given by Algorithm 1, generates inverse consistent and topology preserving deformation maps. The Jacobian determinants of the deformations recovered between a pair of images using unbiased registration follow a log-normal distribution, with zero mean after log-transformation. We demonstrated that this distribution is beneficial when recovering change in regions of homogeneous intensity, and in ensuring symmetrical results when the order of two images being registered is switched.

In this section, we use the Jacobian maps for segmentation of regions of homogeneous deformation. We refer readers to prior works on variational segmentation models [12, 56, 57]. In our approach, we employ the two-phase and multiphase level

set based segmentation models [14, 15, 102] originally used for segmenting intensity images. We will explain how these models can be generalized to classify the regions of homogeneous deformation obtained with the unbiased registration model.

### 5.2.1 The Chan-Vese Intensity Based Segmentation Model

**The One Level Set (Two-Phase) Framework.** Based on the Mumford and Shah functional [75] for segmentation, Chan and Vese [14] proposed a level set method based active contour model to detect objects whose boundaries are not necessarily defined by a gradient.

Let us denote a given image by $I_0 : \Omega \rightarrow \mathbb{R}$ and suppose $C$ ($C = \partial R$) is a hypersurface representing a boundary of a region of interest $R \subset \Omega$. The Chan-Vese (CV) model minimizes the following energy:

$$\begin{aligned} F_2^{CV}(c_1, c_2, C) = &\int_{R_1=R} (I_0(\mathbf{x}) - c_1)^2 \, d\mathbf{x} \\ &+ \int_{R_2=\Omega\backslash R} (I_0(\mathbf{x}) - c_2)^2 \, d\mathbf{x} + \beta \int_{\partial R} ds, \end{aligned} \tag{5.1}$$

where $c_1, c_2$ are unknown constants, and $\beta > 0$ is the length parameter. This problem is solved using the level set method of Osher and Sethian [82]. In a level set formulation, a hypersurface $C$ is represented implicitly by the zero level set of a Lipshitz continuous function $\phi : \Omega \rightarrow \mathbb{R}$, such that:

$$\phi(\mathbf{x}) < 0 \ \text{ in } \ R, \quad \phi(\mathbf{x}) > 0 \ \text{ in } \ \Omega\backslash R.$$

The Chan-Vese functional, given in equation (5.1), can be written in the level set for-

mulation as

$$F_2^{CV}(c_1, c_2, \phi) = \int_\Omega (I_0(\mathbf{x}) - c_1)^2 (1 - H(\phi))\, d\mathbf{x} + \int_\Omega (I_0(\mathbf{x}) - c_2)^2 H(\phi)\, d\mathbf{x} + \beta \int_\Omega |\nabla H(\phi)|\, d\mathbf{x}, \tag{5.2}$$

where $H(y)$ is a heaviside function. The functional is minimized using incremental updating along the gradient descent direction of the Euler-Lagrange equation in $\phi$:

$$\frac{\partial \phi}{\partial t} = \delta(\phi)\Big[\beta \nabla \cdot \left(\frac{\nabla \phi}{|\nabla \phi|}\right) + (I_0 - c_1)^2 - (I_0 - c_2)^2\Big],$$

where $\delta(y)$ is the delta function and $t$ is an artificial time. The constants $c_1$ and $c_2$ are evaluated as

$$c_1(\phi) = \frac{\int_\Omega I_0(\mathbf{x})(1 - H(\phi(t, \mathbf{x})))d\mathbf{x}}{\int_\Omega (1 - H(\phi(t, \mathbf{x})))d\mathbf{x}},$$

$$c_2(\phi) = \frac{\int_\Omega I_0(\mathbf{x}) H(\phi(t, \mathbf{x}))d\mathbf{x}}{\int_\Omega H(\phi(t, \mathbf{x}))d\mathbf{x}}.$$

Note that $c_1$ and $c_2$ are the averages of the intensities of $I_0$ inside and outside $C$, respectively. The two-phase segmentation of the image $I_0(\mathbf{x})$ is given by

$$I(\mathbf{x}) = c_1(1 - H(\phi(\mathbf{x}))) + c_2 H(\phi(\mathbf{x})).$$

**The Two Level Set (Multiphase) Framework.** In [102], the authors generalized the one level set function *active contour without edges* model to *multiphase* framework, which incorporates two or more level set functions. A four-phase model, described in this section, allows for up to four regions to be segmented. Here, we suppose $C_1$ and $C_2$ are hypersurfaces separating an image into four disjoint regions $R_i \subset \Omega$, $1 \leq i \leq 4$. Some of these regions are allowed to be empty. The four phase model thus minimizes

the following energy:

$$
\begin{aligned}
F_4^{CV}(c_{11}, c_{12}, c_{21}, c_{22}, C_1, C_2) = \beta \int_{C_1} ds + \beta \int_{C_2} ds \\
+ \int_{R_1} (I_0(\mathbf{x}) - c_{11})^2 \, d\mathbf{x} + \int_{R_2} (I_0(\mathbf{x}) - c_{12})^2 \, d\mathbf{x} \\
+ \int_{R_3} (I_0(\mathbf{x}) - c_{21})^2 \, d\mathbf{x} + \int_{R_4} (I_0(\mathbf{x}) - c_{22})^2 \, d\mathbf{x},
\end{aligned} \tag{5.3}
$$

where $c_{11}, c_{12}, c_{21}, c_{22}$ are unknown constants, and $\beta > 0$ is the length parameter. Representing $C_1$ and $C_2$ implicity as zero level sets of functions $\phi_1$ and $\phi_2$, respectively, such that

$$
\begin{aligned}
\phi_1(\mathbf{x}) < 0, \ \phi_2(\mathbf{x}) < 0 \ \text{ in } R_1, \\
\phi_1(\mathbf{x}) < 0, \ \phi_2(\mathbf{x}) > 0 \ \text{ in } R_2, \\
\phi_1(\mathbf{x}) > 0, \ \phi_2(\mathbf{x}) < 0 \ \text{ in } R_3, \\
\phi_1(\mathbf{x}) > 0, \ \phi_2(\mathbf{x}) > 0 \ \text{ in } R_4,
\end{aligned}
$$

we can write the Chan-Vese functional, given in equation (5.3), in the level set formulation as:

$$
\begin{aligned}
F_4^{CV}(c_{11}, c_{12}, c_{21}, c_{22}, \phi_1, \phi_2) = \beta \int_{\Omega} |\nabla H(\phi_1)| \, d\mathbf{x} + \beta \int_{\Omega} |\nabla H(\phi_2)| \, d\mathbf{x} \\
+ \int_{\Omega} (I_0(\mathbf{x}) - c_{11})^2 (1 - H(\phi_1))(1 - H(\phi_2)) \, d\mathbf{x} \\
+ \int_{\Omega} (I_0(\mathbf{x}) - c_{12})^2 (1 - H(\phi_1)) H(\phi_2) \, d\mathbf{x} \\
+ \int_{\Omega} (I_0(\mathbf{x}) - c_{21})^2 H(\phi_1)(1 - H(\phi_2)) \, d\mathbf{x} \\
+ \int_{\Omega} (I_0(\mathbf{x}) - c_{22})^2 H(\phi_1) H(\phi_2) \, d\mathbf{x},
\end{aligned} \tag{5.4}
$$

where $H(y)$ is a heaviside function. This functional can be minimized using the gradient descent of the corresponding Euler-Lagrange equations for $\phi_1$ and $\phi_2$ to obtain the

following evolution equations:

$$\frac{\partial \phi_1}{\partial t} = \delta(\phi_1)\Big[\beta\,\nabla\cdot\left(\frac{\nabla\phi_1}{|\nabla\phi_1|}\right) + \big((I_0-c_{11})^2-(I_0-c_{21})^2\big)(1-H(\phi_2)) + \big((I_0-c_{12})^2-(I_0-c_{22})^2\big)H(\phi_2)\Big],$$

$$\frac{\partial \phi_2}{\partial t} = \delta(\phi_2)\Big[\beta\,\nabla\cdot\left(\frac{\nabla\phi_2}{|\nabla\phi_2|}\right) + \big((I_0-c_{11})^2-(I_0-c_{12})^2\big)(1-H(\phi_1)) + \big((I_0-c_{21})^2-(I_0-c_{22})^2\big)H(\phi_1)\Big],$$

where $\delta(y)$ is the delta function and $t$ is an artificial time. The constants $c_{11}, c_{12}, c_{21}, c_{22}$ are evaluated as

$$\begin{aligned}
c_{11}(\phi_1,\phi_2) &= \frac{\int_\Omega I_0(1-H(\phi_1))(1-H(\phi_2))d\mathbf{x}}{\int_\Omega(1-H(\phi_1))(1-H(\phi_2))d\mathbf{x}},\\
c_{12}(\phi_1,\phi_2) &= \frac{\int_\Omega I_0(1-H(\phi_1))H(\phi_2)d\mathbf{x}}{\int_\Omega(1-H(\phi_1))H(\phi_2)d\mathbf{x}},\\
c_{21}(\phi_1,\phi_2) &= \frac{\int_\Omega I_0H(\phi_1)(1-H(\phi_2))d\mathbf{x}}{\int_\Omega H(\phi_1)(1-H(\phi_2))d\mathbf{x}},\\
c_{22}(\phi_1,\phi_2) &= \frac{\int_\Omega I_0H(\phi_1)H(\phi_2)d\mathbf{x}}{\int_\Omega H(\phi_1)H(\phi_2)d\mathbf{x}}.
\end{aligned}$$

Here, $c_{11}$, $c_{12}$, $c_{21}$, and $c_{22}$ correspond to averages of intensities of $I_0$ in $R_1$, $R_2$, $R_3$, and $R_4$, respectively. The four-phase segmentation of the image $I_0(\mathbf{x})$ is given by

$$\begin{aligned}
I(\mathbf{x}) &= c_{11}(1-H(\phi_1))(1-H(\phi_2)) + c_{12}(1-H(\phi_1))H(\phi_2)\\
&\quad + c_{21}H(\phi_1)(1-H(\phi_2)) + c_{22}H(\phi_1)H(\phi_2).
\end{aligned}$$

### 5.2.2 Jacobian Based Segmentation using the CV Model

**Two-phase segmentation of deformation.** Instead of segmenting the image based on its intensity values, we propose to classify and track regions of homogeneous deformation using the Jacobian values $|D\mathbf{g}|$. As a result, the two-phase Chan-Vese functional in (5.2) applied to segmentation of deformation becomes

$$\begin{aligned} F_2(J_1, J_2, \phi) = &\int_\Omega (|D\mathbf{g}(\mathbf{x})| - J_1)^2 (1 - H(\phi))\, d\mathbf{x} \\ &+ \int_\Omega (|D\mathbf{g}(\mathbf{x})| - J_2)^2 H(\phi)\, d\mathbf{x} + \beta \int_\Omega |\nabla H(\phi)|\, d\mathbf{x}, \end{aligned} \tag{5.5}$$

where $J_1, J_2$ are unknown constants. This functional is minimized using incremental updating along the gradient descent direction of the Euler-Lagrange equation in $\phi$:

$$\frac{\partial \phi}{\partial t} = \delta(\phi)\Big[\beta \nabla \cdot \left(\frac{\nabla \phi}{|\nabla \phi|}\right) + (|D\mathbf{g}| - J_1)^2 - (|D\mathbf{g}| - J_2)^2\Big], \tag{5.6}$$

where $\delta(y)$ is the delta function and $t$ is an artificial time. The constants $J_1$ and $J_2$ are evaluated as

$$\begin{aligned} J_1(\phi) &= \frac{\int_\Omega |D\mathbf{g}(\mathbf{x})|(1 - H(\phi(t, \mathbf{x})))d\mathbf{x}}{\int_\Omega (1 - H(\phi(t, \mathbf{x})))d\mathbf{x}}, \\ J_2(\phi) &= \frac{\int_\Omega |D\mathbf{g}(\mathbf{x})| H(\phi(t, \mathbf{x}))d\mathbf{x}}{\int_\Omega H(\phi(t, \mathbf{x}))d\mathbf{x}}. \end{aligned} \tag{5.7}$$

Note that $J_1$ and $J_2$ are the averages of the Jacobian values of $\mathbf{g}$ inside and outside $C$, respectively. The two-phase segmentation of the Jacobian map $|D\mathbf{g}|$, denoted as $\mathcal{J}$, is given by

$$\mathcal{J}(\mathbf{x}) = J_1(1 - H(\phi(\mathbf{x}))) + J_2 H(\phi(\mathbf{x})). \tag{5.8}$$

**Multiphase segmentation of deformation.** The multiphase functional in equation (5.4) for homogeneous deformation segmentation becomes

$$\begin{aligned} F_4(J_{11}, J_{12}, J_{21}, J_{22}, \phi_1, \phi_2) = \beta \int_\Omega |\nabla H(\phi_1)|\, d\mathbf{x} + \beta \int_\Omega |\nabla H(\phi_2)|\, d\mathbf{x} \\ + \int_\Omega (|D\mathbf{g}(\mathbf{x})| - J_{11})^2 (1 - H(\phi_1))(1 - H(\phi_2))\, d\mathbf{x} \\ + \int_\Omega (|D\mathbf{g}(\mathbf{x})| - J_{12})^2 (1 - H(\phi_1)) H(\phi_2)\, d\mathbf{x} \\ + \int_\Omega (|D\mathbf{g}(\mathbf{x})| - J_{21})^2 H(\phi_1)(1 - H(\phi_2))\, d\mathbf{x} \\ + \int_\Omega (|D\mathbf{g}(\mathbf{x})| - J_{22})^2 H(\phi_1) H(\phi_2)\, d\mathbf{x}. \end{aligned} \tag{5.9}$$

This functional can be minimized using the gradient descent of the corresponding Euler-Lagrange equations for $\phi_1$ and $\phi_2$ to obtain the following evolution equations:

$$\begin{aligned} \frac{\partial \phi_1}{\partial t} = \delta(\phi_1) \Bigg[ \beta \nabla \cdot \left( \frac{\nabla \phi_1}{|\nabla \phi_1|} \right) \\ + \big( (|D\mathbf{g}| - J_{11})^2 - (|D\mathbf{g}| - J_{21})^2 \big)(1 - H(\phi_2)) \\ + \big( (|D\mathbf{g}| - J_{12})^2 - (|D\mathbf{g}| - J_{22})^2 \big) H(\phi_2) \Bigg], \end{aligned} \tag{5.10}$$

$$\begin{aligned} \frac{\partial \phi_2}{\partial t} = \delta(\phi_2) \Bigg[ \beta \nabla \cdot \left( \frac{\nabla \phi_2}{|\nabla \phi_2|} \right) \\ + \big( (|D\mathbf{g}| - J_{11})^2 - (|D\mathbf{g}| - J_{12})^2 \big)(1 - H(\phi_1)) \\ + \big( (|D\mathbf{g}| - J_{21})^2 - (|D\mathbf{g}| - J_{22})^2 \big) H(\phi_1) \Bigg]. \end{aligned}$$

The constants $J_{11}, J_{12}, J_{21}, J_{22}$ are evaluated as

$$\begin{aligned}
J_{11}(\phi_1, \phi_2) &= \frac{\int_\Omega |D\mathbf{g}|(1 - H(\phi_1))(1 - H(\phi_2))d\mathbf{x}}{\int_\Omega (1 - H(\phi_1))(1 - H(\phi_2))d\mathbf{x}}, \\
J_{12}(\phi_1, \phi_2) &= \frac{\int_\Omega |D\mathbf{g}|(1 - H(\phi_1))H(\phi_2)d\mathbf{x}}{\int_\Omega (1 - H(\phi_1))H(\phi_2)d\mathbf{x}}, \\
J_{21}(\phi_1, \phi_2) &= \frac{\int_\Omega |D\mathbf{g}|H(\phi_1)(1 - H(\phi_2))d\mathbf{x}}{\int_\Omega H(\phi_1)(1 - H(\phi_2))d\mathbf{x}}, \\
J_{22}(\phi_1, \phi_2) &= \frac{\int_\Omega |D\mathbf{g}|H(\phi_1)H(\phi_2)d\mathbf{x}}{\int_\Omega H(\phi_1)H(\phi_2)d\mathbf{x}}.
\end{aligned} \tag{5.11}$$

Here, $J_{11}$, $J_{12}$, $J_{21}$, and $J_{22}$ correspond to averages of Jacobian values of $\mathbf{g}$ in $R_1$, $R_2$, $R_3$, and $R_4$, respectively. The four-phase segmentation of the Jacobian map $|D\mathbf{g}|$, denoted as $\mathcal{J}$, is given by

$$\begin{aligned}
\mathcal{J}(\mathbf{x}) = c_{11}(1 - H(\phi_1))(1 - H(\phi_2)) + c_{12}(1 - H(\phi_1))H(\phi_2) \\
+ c_{21}H(\phi_1)(1 - H(\phi_2)) + c_{22}H(\phi_1)H(\phi_2).
\end{aligned} \tag{5.12}$$

### 5.2.3 Algorithm

We are now ready to give an algorithm for the proposed model on segmentation of deformation. The algorithm describes both, the two-phase and multiphase, segmentation cases.

**Algorithm 3** Segmentation of Deformation

1: Obtain Jacobian map $|D\mathbf{g}|$ using Algorithm 1.
2: Initialize $t = 0$. Initialize $\phi(0, \mathbf{x}) = \phi_0(\mathbf{x})$ for the two-phase model, or initialize $\phi_1(0, \mathbf{x}) = \phi_{1_0}(\mathbf{x})$ and $\phi_2(0, \mathbf{x}) = \phi_{2_0}(\mathbf{x})$ for the multiphase model.
3: Calculate $J_1$ and $J_2$ using (5.7) for the two-phase model, or calculate $J_{11}$, $J_{12}$, $J_{21}$, $J_{22}$ using (5.11) for the multiphase model.
4: Advance $\phi(\mathbf{x}, t)$ in time using equation (5.6) for the two-phase model, or advance $\phi_1(\mathbf{x}, t)$ and $\phi_2(\mathbf{x}, t)$ in time using equation (5.10) for the multiphase model.
5: If the cost functional in equation (5.5) for the two-phase model, or the cost functional in equation (5.9) for the multiphase model, decreases by sufficiently small amount compared to the previous iteration, then stop.
Segmented Jacobian map of the deformation is given by equation (5.8) for the two-phase model, or equation (5.12) for the multiphase model.
6: Let $t := t + \triangle t$ and go to step 3.

Of note, for providing additional flexibility, both intensity values of a target image and the Jacobian field of the deformation could be incorporated as two channels into a multichannel model described in [13]. Depending on an application, additional channels may be incorporated into the model.

## 5.3 Results

In this section, we tested the proposed segmentation of homogeneous deformation framework. In the first numerical example in Figure 5.1 we considered matching two synthetic images (each of size 256 by 256, $\lambda = 1000$ in equation (2.17) for unbiased nonlinear registration). The geometrical objects on each of these images (Figure 5.1(a,b)) are of identical intensity; however, each of these four objects undergoes a deformation of a different magnitude (Figure 5.1(d,e,f)). The upper-left object (a circle to an ellipse transformation) undergoes the biggest positive deformation (expansion) and the lower-right object (an ellipse to a circle transformation) is being contracted. Note that the square does not deform. In this example, the segmentation was done us-

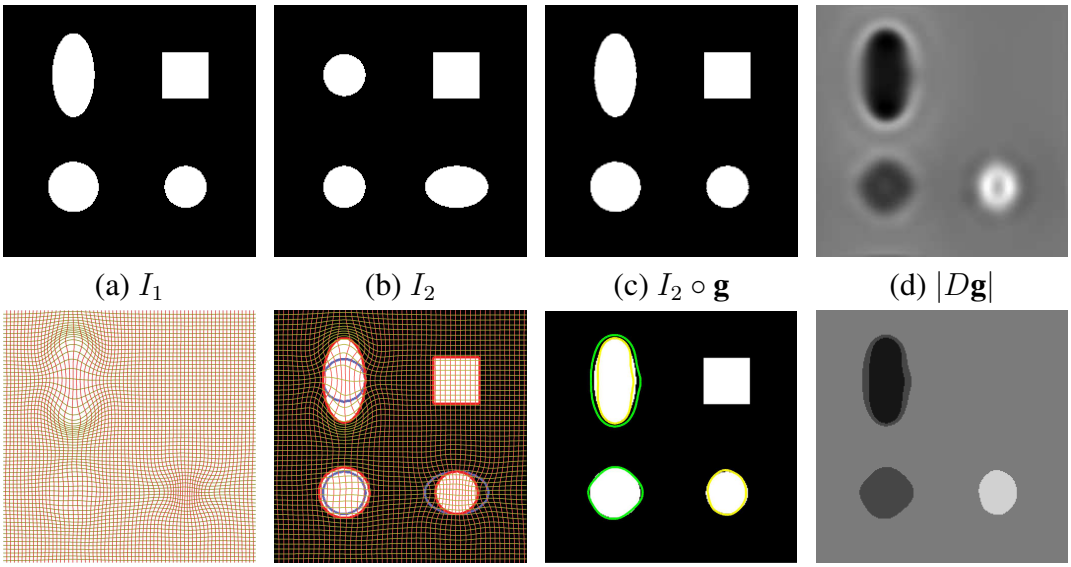
ing the four-phase segmentation model (5.9) with the length parameter $\beta = 0.02 \cdot 255^2$. The four regions of homogeneous deformation were detected (Figure 5.1(g,h)). The background and non-deforming square were classified as a single region of identity (or almost identity) deformation.

In Figure 5.2, we show the results of matching a pair of two-dimensional slices (Figure 5.2(a,b)) from a set of serial MRI images (each of size 226 by 256, $\lambda = 400$ in equation (2.17)), where visually significant ventricle enlargement is present. Here, it is desirable to distinguish the region of ventricular expansion from the rest of the image. This is successfully accomplished using the segmentation of homogeneous deformation procedure (Figure 5.2(g,h)). The four-phase segmentation model was employed in this example (with the length parameter $\beta = 0.1 \cdot 255^2$ in equation (5.9)), locating only two regions of homogeneous deformation, which is intuitively correct.

In the last numerical example (Figures 5.3 and 5.4), we tested the proposed model using a pair of three-dimensional serial MRI volumes (each of size 112x128x128) which, similar to a previous two-dimensional example, display significant ventricular growth. A fully three-dimensional computation was employed, with $\lambda = 500$ in equation (2.17), and $\beta = 0.05 \cdot 255^2$ in a two-phase segmentation model (5.5). Figure 5.4 displays the volume cuts of the two volumes matched as well as the result of segmentation in the form of a surface (zero level set of function $\phi$) of the ventricle. The two-dimensional slices of the three-dimensional volume, as well as the corresponding segmentation of deformation, are shown in Figure 5.3. The region of growth was identified and separated from the rest of the image in this example.

## 5.4 Conclusion

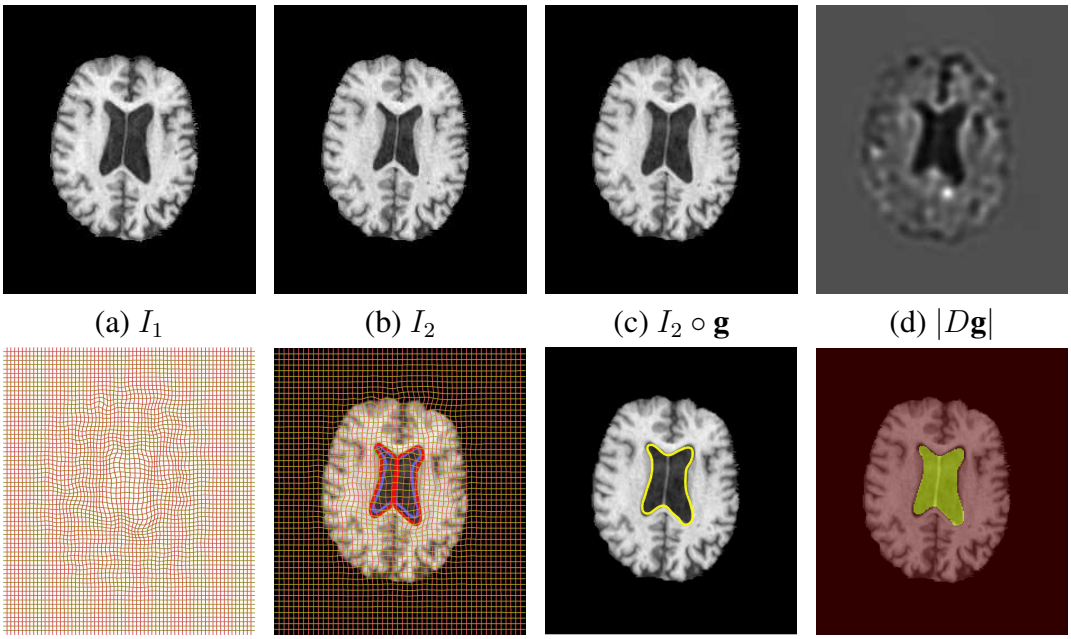
This chapter introduces a novel framework for simultaneously registering and tracking deforming objects in a sequence of two or more images. We employ unbiased non-linear image registration for generating priors for further segmentation of deformed features. A level set based active contour without edges and multiphase segmentation models are generalized to consider Jacobian fields while segmenting regions of growth and shrinkage in deformations. Deforming objects are thus classified based on magnitude of homogeneous deformation.



(a) $I_1$ (b) $I_2$ (c) $I_2 \circ \mathbf{g}$ (d) $|D\mathbf{g}|$

(e) deformed grid (f) $I_2 \circ \mathbf{g}$ and grid (g) segmentation contours (h) segmentation of $|D\mathbf{g}|$

Figure 5.1: Segmentation of deformation for a synthetic image. (a) image $I_1$; (b) image $I_2$; (c) image $I_2$ is deformed into image $I_1$. The upper-left circle in the image undergoes the largest expansion (density change), while the lower-right ellipsoid undergoes the contraction. The upper-right square does not deform. (d) The Jacobian map of the deformation. Dark and bright spots represent expanding and contracting areas, respectively. (e) The deformed grid; (f) the deformed grid and the deformed image. Here, yellow, blue, and red contours represent boundaries of objects in $I_1$, $I_2$, and deformed $I_2$, respectively. (g) Segmentation results are obtained using the four-phase (multiphase) segmentation model, which enables to find up to four regions in the image. Green and yellow contours represent the zero level sets of $\phi_1$ and $\phi_2$, respectively. The four regions of homogeneous change in density are located. (h) The segmented Jacobian map is displayed.



(a) $I_1$ (b) $I_2$ (c) $I_2 \circ \mathbf{g}$ (d) $|D\mathbf{g}|$

(e) deformed grid (f) $I_2 \circ \mathbf{g}$ and grid (g) segmentation contours (h) $I_2 \circ \mathbf{g}$ and segmentation

Figure 5.2: Segmentation of deformation for 2D serial MRI image. (a) image $I_1$; (b) image $I_2$; (c) image $I_2$ is deformed into image $I_1$. The ventricle is noticeably enlarged. (d) The Jacobian map of the deformation; (e) the deformed grid; (f) the deformed grid and the deformed image. Here, yellow, blue, and red contours represent the boundaries of objects in $I_1$, $I_2$, and deformed $I_2$, respectively. (g) Segmentation results are obtained using the four-phase (multiphase) segmentation model, which enables to find up to four regions in the image. However, since only the ventricle had undergone the deformation, the image is partitioned into two parts. (h) The deformed image is superimposed with the segmentation of the deformation.

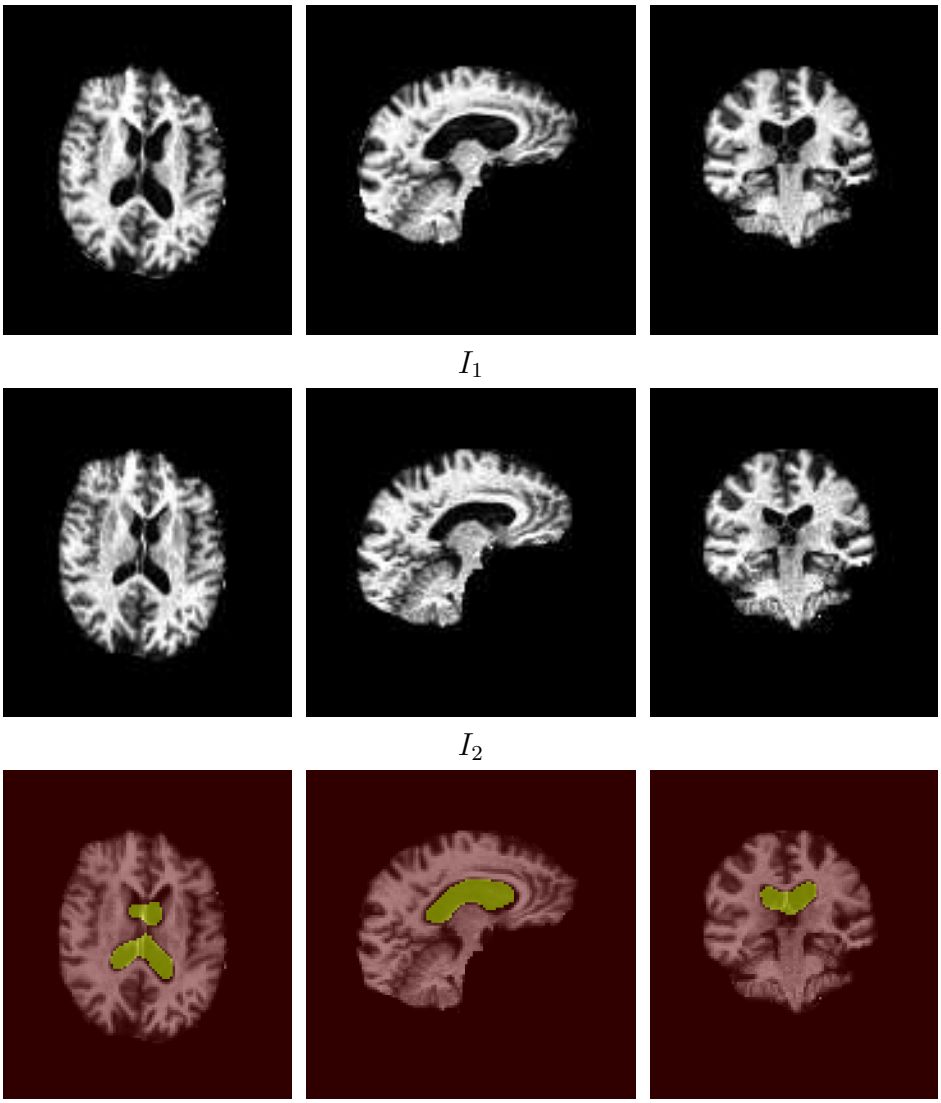



Figure 5.3: Segmentation of deformation for 3D serial MRI image. Columns depict: axial (column 1), sagittal (column 2), and coronal (column 3) slices of image $I_1$; image $I_2$; deformed image $I_2$ superimposed with the segmentation of deformation. Segmentation results are obtained using the two-phase segmentation model, which enables to separate two regions in the image. Since the ventricle underwent the largest deformation, it is separated from the rest of the image.

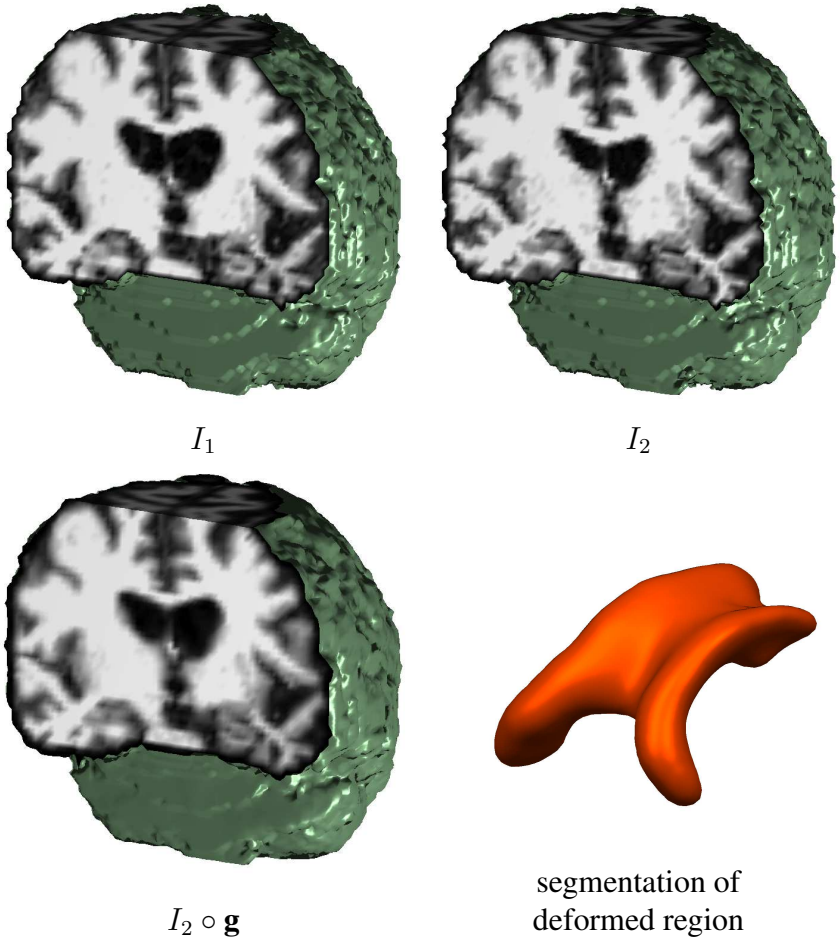



Figure 5.4: Segmentation of deformation for 3D serial MRI image. Volume cuts of image $I_1$, image $I_2$, and deformed image $I_2$ are shown. The surface (zero level set of function $\phi$) of the ventricle is shown. The ventricle surface is enlarged for better visualization.

# APPENDIX A

# Information Theory

## A.1 Shannon entropy

We provide the definition of the Shannon entropy or information entropy, which is a measure of the uncertainty of a random variable. Let $X$ be a discrete random variable with a set of outcomes $\mathcal{X}$. The probability mass function is defined as $p(x) = \text{Pr}\{X = x\}$, $x \in \mathcal{X}$. The cumulative distribution function $P(x)$ of a random variable $X$ is

$$P(x) = Pr\{X \leq x\}, \tag{A.1}$$

and we have

$$Pr\{x < X \leq z\} = P(z) - P(x). \tag{A.2}$$

For a discrete random variable $X$

$$P(x) = \sum_{\forall k \leq x} p(k). \tag{A.3}$$

The enropy $H(X)$ of a discrete random variable $X$ is defined by

$$H(X) = -\sum_{x \in \mathcal{X}} p(x) \log p(x). \tag{A.4}$$

It is sometimes convenient to denote the above quantity as $H(p)$.

The joint entropy $H(X,Y)$ of a pair of discrete random variables $(X,Y)$ with a joint distribution $p(x,y)$ is defined as

$$H(X,Y) = -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\log p(x,y), \tag{A.5}$$

and the conditional entropy $H(Y|X)$ is defined as

$$\begin{aligned} H(Y|X) &= \sum_{x\in\mathcal{X}} p(x)H(Y|X=x) && \text{(A.6)}\\ &= -\sum_{x\in\mathcal{X}} p(x)\sum_{y\in\mathcal{Y}} p(y|x)\log p(y|x) && \text{(A.7)}\\ &= -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\log p(y|x). && \text{(A.8)} \end{aligned}$$

Here, $p(y|x)$ is the conditional probability mass function. It can be shown that the joint entropy of two random variables is the entropy of one plus the conditional entropy of the other:

$$\begin{aligned} H(X,Y) &= -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\log p(x,y) && \text{(A.9)}\\ &= -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\log p(x)p(y|x) && \text{(A.10)}\\ &= -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\log p(x) - \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\log p(y|x) && \text{(A.11)}\\ &= -\sum_{x\in\mathcal{X}} p(x)\log p(x) - \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\log p(y|x) && \text{(A.12)}\\ &= H(X) + H(Y|X). && \text{(A.13)} \end{aligned}$$

## A.2 Differential entropy

The differential entropy is an extension of the idea of the Shannon entropy to continuous probability distributions. Let $X$ be a continuous random variable with continuous cumulative distribution function $P(x) = \Pr\{X \leq x\}$. Let $p(x) = P'(x)$ when the derivative is defined. If $\int p(x)dx = 1$, $p(x)$ is called the probability density function for $X$. The differential entropy $H(X)$ of a continuous random variable $X$ with density $p(x)$ is defined as

$$H(X) = -\int_D p(x) \log p(x)\, dx, \tag{A.14}$$

where $D$ is the support set of the random variable. The joint differential entropy $H(X, Y)$ of a pair of random variables $X$ and $Y$ with joint density function $p(x, y)$ is defined as

$$H(X, Y) = -\int p(x, y) \log p(x, y)\, dxdy, \tag{A.15}$$

and the conditional differential entropy $H(X|Y)$ is provided by

$$H(X|Y) = -\int p(x, y) \log p(x|y)\, dxdy. \tag{A.16}$$

## A.3 Kullback-Leibler divergence

The relative entropy or *Kullback-Leibler divergence*, which is a measure of the distance between two probability mass functions $p(x)$ and $q(x)$, is defined as

$$KL(p, q) = \sum_{x \in \mathcal{X}} p(x) \log \frac{p(x)}{q(x)}. \tag{A.17}$$

The continuous version is

$$KL(p,q) = \int p(x) \log \frac{p(x)}{q(x)}\, dx. \tag{A.18}$$

Although $KL(p,q) \geq 0$ and $KL(p,q) = 0$ if and only if $p = q$, the Kullback-Leibler divergence is not a true metric because it is not necessarily symmetric. Furthermore, $KL(\cdot,\cdot)$ need not satisfy the triangle inequality.

## A.4 Mutual information

We now introduce mutual information, which is a measure of the amount of information that one random variable contains about another random variable. Consider two random variables $X$ and $Y$ with joint probability mass function $p(x,y)$ and marginal probability mass functions $p(x)$ and $p(y)$. The mutual information $MI(X,Y)$ is the relative entropy between the joint distribution and the product of marginal distributions:

$$MI(X,Y) = KL(p(x,y), p(x)p(y)) = \sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}. \tag{A.19}$$

It can be shown that

$$\begin{aligned} MI(X,Y) &= \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} = \sum_{x,y} p(x,y) \log \frac{p(x|y)}{p(x)} && \text{(A.20)} \\ &= -\sum_{x,y} p(x,y) \log p(x) + \sum_{x,y} p(x,y) \log p(x|y) && \text{(A.21)} \\ &= H(X) - H(X|Y). && \text{(A.22)} \end{aligned}$$

Hence, the mutual information $MI(X,Y)$ is the reduction in the uncertainty of $X$ due to the knowledge of $Y$. Mutual information can be equivalently expressed as

$$MI(X,Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) = H(X) + H(Y) - H(X,Y),$$

where $H(X)$ and $H(Y)$ are the marginal entropies, $H(X|Y)$ and $H(Y|X)$ are the conditional entropies, and $H(X,Y)$ is the joint entropy of $X$ and $Y$.

The mutual information $MI(X,Y)$ between continuous random variables, $X$ and $Y$, is defined as

$$MI(X,Y) = \int p(x,y) \log \frac{p(x,y)}{p(x)p(y)} dxdy. \tag{A.23}$$

The properties of $MI(X,Y)$ in continuous case are the same as in the discrete case.

# APPENDIX B

# Level Set Method

## B.1 The Level Set Equation

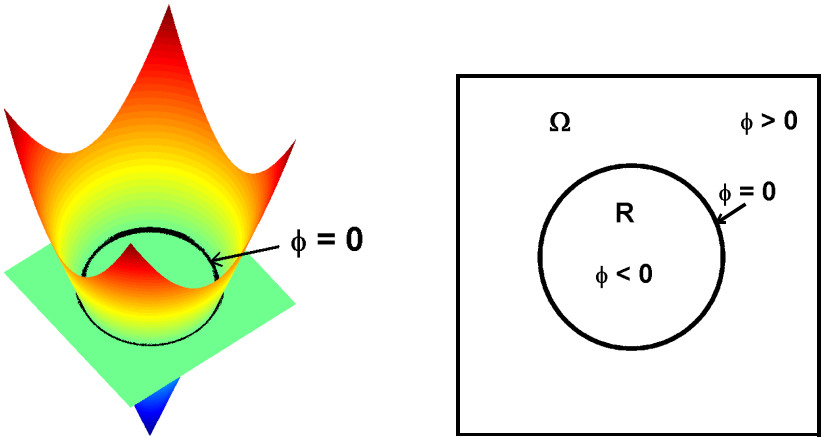
The level set method, introduced by Osher and Sethian [82], has been widely used in computational fluid dynamics, computer vision, computer graphics, image analysis, and medical imaging. For a recent survey on the level set method, we refer readers to [79, 80, 81].

Consider an interface $C$ in $\mathbb{R}^n$ of codimension one bounding an open region $R \subset \Omega$. We wish to analyze and compute its subsequent motion under a velocity field $\mathbf{v}$, where the velocity can depend on position, time, the geometry of the interface, and external physics. A moving interface $C(t)$, at time $t$, is represented implicitly by the zero level set of a Lipschitz continuous function $\phi(\mathbf{x}, t)$, such that

$$\begin{aligned} \phi(\mathbf{x}, t) &< 0 \text{ in } R(t), \\ \phi(\mathbf{x}, t) &= 0 \text{ on } \partial R(t), \\ \phi(\mathbf{x}, t) &> 0 \text{ in } \Omega \backslash R(t). \end{aligned}$$

Implicit function $\phi$ is used both to represent the interface and to evolve it. The evolution of implicit function $\phi$ is governed by the convection equation

$$\phi_t + \mathbf{v} \cdot \nabla \phi = 0. \tag{B.1}$$




(a) (b)

Figure B.1: (a) Signed distance function $\phi : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ is shown. The intersection of $\phi$ with the $x_1x_2$-plane defines the zero level set of $\phi$. (b) An interface $C = \{\mathbf{x} \mid \phi(\mathbf{x}) = 0\}$ in the $x_1x_2$-plane defines a boundary of region $R \subset \Omega$.

Equation (B.1) is sometimes referred to as the *level set equation*.

Denoting $\mathbf{n}$ to be the outward unit normal to the interface and $\boldsymbol{\tau}$ be any tangent vector to the interface, the velocity $\mathbf{v}$ can be written as $\mathbf{v} = v_N\mathbf{n} + v_T\boldsymbol{\tau}$. Here, $v_N$ and $v_T$ are the components of the velocity in the normal and tangential directions, respectively. Equation (B.1) can be written as

$$\phi_t + \left(v_N\mathbf{n} + v_T\boldsymbol{\tau}\right) \cdot \nabla\phi = 0.$$

Since $\mathbf{n}$ and $\nabla\phi$ point in the same direction, $\boldsymbol{\tau} \cdot \nabla\phi = 0$. Thus, the level set equation is equivalent to

$$\phi_t + v_N\mathbf{n} \cdot \nabla\phi = 0. \tag{B.2}$$

Also, since the unit normal to the interface can be expressed as

$$\mathbf{n} = \frac{\nabla\phi}{|\nabla\phi|},$$

we can rewrite equation (B.2) as

$$\phi_t + v_N|\nabla\phi| = 0.$$

## B.2 Geometry and Calculus Toolboxes

Geometric and calculus quantities are easily expressed in the level set notation. The mean curvature $\kappa$ of the interface is defined as the divergence of the normal $\mathbf{n}$, $\kappa = \nabla \cdot \mathbf{n}$, or

$$\kappa = \nabla \cdot \left(\frac{\nabla\phi}{|\nabla\phi|}\right).$$

The one-dimensional Heaviside function is

$$H(\phi) = \begin{cases} 0 & \text{if } \phi \leq 0 \\ 1 & \text{if } \phi > 0. \end{cases}$$

In one spatial dimension, the delta function is defined as the derivative of the one-dimensional Heaviside function

$$\delta(\phi) = H'(\phi).$$

The length of $C = \{\mathbf{x} \mid \phi(\mathbf{x}) = 0\}$ is

$$|C| = \int_{\Omega} |\nabla H(\phi)| \, d\mathbf{x} = \int_{\Omega} \delta(\phi) |\nabla \phi| \, d\mathbf{x}.$$

The area of region $R = \{\mathbf{x} \mid \phi(\mathbf{x}) < 0\} \subset \Omega$ is

$$|R| = \int_{\Omega} (1 - H(\phi)) \, d\mathbf{x}.$$

Similarly, the area of the complement of region $R$, namely $R^c$, is defined as

$$|R^c| = \int_{\Omega} H(\phi) \, d\mathbf{x}.$$

The volume integral of a function $f$ over the region $R \subset \Omega$ is defined as

$$\int_{\Omega} f(\mathbf{x})(1 - H(\phi)) \, d\mathbf{x}.$$

The mean of $f$ in $R$ is

$$\frac{\int_{\Omega} f(\mathbf{x})(1 - H(\phi)) \, d\mathbf{x}}{\int_{\Omega} (1 - H(\phi)) \, d\mathbf{x}}.$$

# REFERENCES


[1] B. Avants and J. C. Gee, "Geodesic estimation for large deformation anatomical shape averaging and interpolation," *NeuroImage*, vol. 23, suppl. 1, pp. S139–50, 2004.

[2] R. Bajcsy and S. Kovacic, "Multiresolution elastic matching," *Computer Vision Graphics and Image Processing*, vol. 46, pp. 1–21, 1989.

[3] M. Beg, M. Miller, A. Trouve, and L. Younes, "Computing large deformation metric mappings via geodesic flows of diffeomorphisms," *International Journal of Computer Vision*, vol. 61, no. 2, pp. 139–157, 2005.

[4] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: A practical and powerful approach to multiple testing," *Journal of the Royal Statistical Society*, vol. B 57, no. 1, pp. 289–300, 1995.

[5] Y. Benjamini and D. Yekutieli, "The control of the false discovery rate in multiple testing under dependency," *Annals of Statistics*, vol. 29, no. 4, pp. 1165–1188, 2001.

[6] A. Bhattacharyya, "On a measure of divergence between two statistical populations defined by their probability distributions," *Bulletin of the Calcutta Mathematical Society*, vol. 35, pp. 99–110, 1943.

[7] M. Bro-Nielsen and C. Gramkow, "Fast fluid registration of medical images," in *Visualization in Biomedical Computing*, 1996, pp. 267–276.

[8] C. Broit, "Optimal registration of deformed images," Ph.D. dissertation, University of Pennsylvania, 1981.

[9] L. Brown, "A survey of image registration techniques," *ACM Computing Surveys*, vol. 24, no. 4, pp. 325–376, 1992.

[10] C. Brun, N. Lepore, X. Pennec, Y. Chou, O. Lopez, H. Aizenstein, J. Becker, A. Toga, and P. Thompson, "Comparison of standard and Riemannian elasticity for tensor-based morphometry in HIV/AIDS," *International Conference on Medical Image Computing and Computer Assisted Intervention*, 2007.

[11] E. Bullmore, J. Suckling, S. Overmeyer, S. Rabe-Hesketh, E. Taylor, and M. Brammer, "Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain," *IEEE Transactions on Medical Imaging*, vol. 18, pp. 32–42, 1999.

[12] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International Journal of Computer Vision*, vol. 22, no. 1, pp. 61–79, 1997.

[13] T. Chan, B. Sandberg, and L. Vese, "Active contours without edges for vector-valued images," *Journal of Visual Communication and Image Representation*, vol. 11, no. 2, pp. 130–141, 2000.

[14] T. Chan and L. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.

[15] T. Chan and L. Vese, "An active contour model without edges," *Scale-Space Theories in Computer Vision, Lecture Notes in Computer Science*, vol. 1682, pp. 141–151, 1999.

[16] C. Chefd'Hotel, G. Hermosillo, and O. Faugeras, "A variational approach to multi-modal image matching," *IEEE Workshop in Variational and Level Set Methods*, pp. 21–28, 2001.

[17] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *Annals of Mathematical Statistics*, vol. 23, pp. 493–507, 1952.

[18] M. Chiang, R. Dutton, K. Hayashi, O. Lopez, H. Aizenstein, A. Toga, J. Becker, and P. Thompson, "3D pattern of brain atrophy in HIV/AIDS visualized using tensor-based morphometry," *NeuroImage*, vol. 34, pp. 44–60, 2007.

[19] G. Christensen, "Consistent linear-elastic transformations for image matching," *Lecture Notes in Computer Science*, vol. 1613, pp. 224–237, 1999.

[20] G. Christensen and H. Johnson, "Consistent image registration," *IEEE Transactions on Medical Imaging*, vol. 20, no. 7, pp. 568–582, 2001.

[21] G. Christensen, R. Rabbitt, and M. Miller, "Deformable templates using large deformation kinematics," *IEEE Transactions on Image Processing*, vol. 5, no. 10, pp. 1435–1447, 1996.

[22] M. Chung, K. Worsley, T. Paus, C. Cherif, D. Collins, J. Giedd, J. Rapoport, and A. Evans, "A unified statistical approach to deformation-based morphometry," *NeuroImage*, vol. 14, pp. 595–606, 2001.

[23] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal, "Automated multi-modality image registration based on information theory," in *Information Processing in Medical Imaging*, Y. Bizais, C. Barillot, and R. Di Paola, Eds., vol. 3. Kluwer Academic Publishers, Dordrecht, 1995, pp. 264–274.

[24] D. Collins, T. Peters, and A. Evans, "Automated 3D nonlinear deformation procedure for determination of gross morphometric variability in human brain," in *Proceedings of the International Society for Optical Engineering (SPIE) 2359*, 1994, pp. 180–190.

[25] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley-Interscience, 1991.

[26] W. Crum, T. Hartkens, and D. Hill, "Non-rigid image registration: Theory and practice," *The British Journal of Radiology*, vol. 77, pp. S140–S153, 2004.

[27] W. Crum, R. Scahill, and N. Fox, "Automated hippocampal segmentation by regional fluid registration of serial MRI: validation and application in Alzheimer's disease," *NeuroImage*, vol. 13, pp. 847–855, 2001.

[28] E. D'Agostino, F. Maes, D. Vandermeulen, and P. Suetens, "A viscous fluid model for multimodal non-rigid image registration using mutual information," *Medical Image Analysis*, vol. 7, pp. 565–575, 2003.

[29] R. Dann, J. Hoford, S. Kovacic, M. Reivich, and R. Bajcsy, "Evaluation of elastic matching system for anatomic (CT,MR) and functional (PET) cerebral images," *Journal of Computer Assisted Tomography*, vol. 13, pp. 603–611, 1989.

[30] C. Davatzikos, M. Vaillant, S. Resnick, J. Prince, S. Letovsky, and R. Bryan, "A computerized approach for morphological analysis of the corpus callosum," *Journal of Computer Assisted Tomography*, vol. 20, no. 1, pp. 88–97, 1996.

[31] R. Duda, P. Hart, and D. Stork, *Pattern Classification*. New York: Wiley-Interscience, 2001.

[32] P. Dupuis, U. Grenander, and M. Miller, "Variational problems on flows of diffeomorphisms for image matching," *Quarterly of Applied Mathematics*, vol. 56, no. 3, pp. 587–600, 1998.

[33] R. Eastman, J. Le Moigne, and N. Netanyahu, "Research issues in image registration for remote sensing," *IEEE Computer Society Workshop on Image Registration and Fusion*, May 2007.

[34] B. Fischer and J. Modersitzki, "Fast diffusion registration," *AMS Contemporary Mathematics, Inverse Problems, Image Analysis, and Medical Imaging*, vol. 313, pp. 117–129, 2002.

[35] B. Fischer and J. Modersitzki, "Curvature based image registration," *Journal of Mathematical Imaging and Vision*, vol. 18, no. 1, pp. 81–85, 2003.

[36] B. Fischer and J. Modersitzki, "A unified approach to fast image registration and a new curvature based registration technique," *Linear Algebra and its Applications, 380*, vol. 23, no. 7, pp. 107–124, 2004.

[37] R. Fisher, "Applications of "Student's" distribution," *Metron*, vol. 5, pp. 90–104, 1925.

[38] K. Friston, A. Holmes, K. Worsley, J. Poline, C. Frith, and R. Frackowiak, "Statistical parametric maps in functional imaging: A general linear approach," *Human Brain Mapping*, vol. 2, pp. 189–210, 1995.

[39] A. Gelman, J. Carlin, H. Stern, and D. Rubin, *Bayesian Data Analysis*. Chapman & Hall/CRC, 2003.

[40] C. Genovese, N. Lazar, and T. Nichols, "Thresholding of statistical maps in functional neuroimaging using the false discovery rate," *Neuroimage*, vol. 15, pp. 870–878, 2002.

[41] C. Glasbey and K. Mardia, "A review of image warping methods," *Journal of Applied Statistics*, vol. 25, pp. 155–171, 1998.

[42] A. Goshtasby, *2-D and 3-D Image Registration: for Medical, Remote Sensing, and Industrial Applications*. New York: Wiley-Interscience, 2005.

[43] W. Gosset, "The probable error of a mean," *Biometrika*, vol. 6, pp. 1–25, 1908.

[44] U. Grenander and M. Miller, "Computational anatomy: An emerging discipline," *Quarterly of Applied Mathematics*, vol. 56, pp. 617–694, 1998.

[45] J. Gunter, M. Bernstein, B. Borowski, J. Felmlee, D. Blezek, and R. Mallozzi, "Validation testing of the MRI calibration phantom for the Alzheimer's Disease Neuroimaging Initiative study," *International Society for Magnetic Resonance in Medicine 14th Scientific Meeting and Exhibition*, 2006.

[46] Y. He, A. Hamza, and H. Krim, "A generalized divergence measure for robust image registration," *IEEE Trans. Signal Process*, vol. 51, no. 5, pp. 1211–20, 2003.

[47] G. Hermosillo, C. Chefd'Hotel, and O. Faugeras, "A variational approach to multi-modal image matching," *INRIA Research report 4117*, 2001.

[48] D. Hill, P. Batchelor, M. Holden, and D. Hawkes, "Medical image registration," *Physics in Medicine and Biology*, vol. 46, pp. R1–R45, 2001.

[49] R. Hogan, L. Mark, L. Wang, S. Joshi, M. Miller, and R. Bucholz, "Mesial temporal sclerosis and temporal lobe epilepsy: MR imaging deformation-based segmentation of the hippocampus in five patients," *Radiology*, vol. 216, pp. 291–297, 2000.

[50] M. Holden, "A review of geometric transformations for nonrigid body registration," *IEEE Transactions on Medical Imaging*, vol. 27, no. 1, pp. 111–128, 2008.

[51] D. Holm, J. Ratnanather, A. Trouve, and L. Younes, "Soliton dynamics in computational anatomy," *NeuroImage*, vol. 21, 2004, suppl. 1, S170-S178.

[52] B. Horn and B. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, pp. 185–203, 1981.

[53] C. Jack, M. Bernstein, N. Fox, P. Thompson, G. Alexander, D. Harvey, B. Borowski, P. Britson, J. Whitwell, C. Ward, A. Dale, J. Felmlee, J. Gunter, D. Hill, R. Killiany, N. Schuff, S. Fox-Bosetti, C. Lin, C. Studholme, C. DeCarli, G. Krueger, H. Ward, G. Metzger, K. Scott, R. Mallozzi, D. Blezek, J. Levy, J. Debbins, A. Fleisher, M. Albert, R. Green, G. Bartzokis, G. Glover, J. Mugler, and M. Weiner, "The Alzheimer's Disease Neuroimaging Initiative (ADNI): The MR imaging protocol," *Journal of Magnetic Resonance Imaging*, vol. 27, no. 4, pp. 685–691, 2008.

[54] H. Johnson and G. Christensen, "Consistent landmark and intensity-based image registration," *IEEE Transactions on Medical Imaging*, vol. 21, no. 5, pp. 450–461, 2002.

[55] J. Jovicich, S. Czanner, D. Greve, E. Haley, A. van derKouwe, R. Gollub, D. Kennedy, F. Schmitt, G. Brown, J. MacFall, B. Fischl, and A. Dale, "Reliability in multi-site structural MRI studies: Effects of gradient non-linearity correction on phantom and human data," *NeuroImage*, vol. 30, no. 2, pp. 436–443, 2006.

[56] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.

[57] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi, "Conformal curvature flows: From phase transitions to active vision," *Archive for Rational Mechanics and Analysis*, vol. 134, no. 3, pp. 275–301, 1996.

[58] C. Le Guyader and L. Vese, "A combined segmentation and registration framework with a nonlinear elasticity smoother," *UCLA Computational and Applied Mathematics Report 08-16*, 2008.

[59] A. Leow, A. Klunder, C. Jack, A. Toga, A. Dale, M. Bernstein, P. Britson, J. Gunter, C. Ward, J. Whitwell, B. Borowski, A. Fleisher, N. Fox, D. Harvey, J. Kornak, N. Schuff, C. Studholme, G. Alexander, M. Weiner, and P. Thompson, "Longitudinal stability of MRI for mapping brain change using tensor-based morphometry," *NeuroImage*, vol. 31, no. 2, pp. 627–40, 2006.

[60] A. Leow, I. Yanovsky, M.-C. Chiang, A. Lee, A. Klunder, A. Lu, J. Becker, S. Davis, A. Toga, and P. Thompson, "Statistical properties of Jacobian maps and the realization of unbiased large-deformation nonlinear image registration," *IEEE Transactions on Medical Imaging*, vol. 26, no. 6, pp. 822–832, 2007.

[61] A. Leow, I. Yanovsky, N. Parikshak, X. Hua, S. Lee, A. Toga, C. Jack, M. Bernstein, P. Britson, J. Gunter, C. Ward, B. Borowski, L. Shaw, J. Trojanowski, A. Fleisher, D. Harvey, J. Kornak, N. Schuff, G. Alexander, M. Weiner, and P. Thompson, "Alzheimer's Disease Neuroimaging Initiative: A one-year follow up study correlating degenerative rates, biomarkers and cognition," *Submitted for publication*, 2008.

[62] N. Lepore, C. Brun, X. Pennec, Y. Chou, O. Lopez, H. Aizenstein, J. Becker, A. Toga, and P. Thompson, "Mean template for tensor-based morphometry using deformation tensors," *International Conference on Medical Image Computing and Computer Assisted Intervention*, 2007.

[63] H. Lester and S. Arridge, "A survey of hierarchical non-linear medical image registration," *Pattern Recognition*, vol. 32, no. 1, pp. 129–149, 1999.

[64] T. Lin, E.-F. Lee, I. Dinov, C. Le Guyader, P. Thompson, A. Toga, and L. Vese, "A landmark based nonlinear elasticity model for mouse atlas registration," *IEEE International Symposium on Biomedical Imaging*, pp. 788–791, 2008.

[65] Y. Lin and G. Medioni, "Map-enhanced uav image sequence registration and synchronization of multiple image sequences," *IEEE Conference on Computer Vision and Pattern Recognition*, 2007.

[66] V. Liseikin, *Grid Generation Methods*. Heidelberg: Springer-Verlag, 1999.

[67] N. Lord, J. Ho, B. Vemuri, and S. Eisenschenk, "Simultaneous registration and parcellation of bilateral hippocampal surface pairs for local asymmetry quantification," *IEEE Transactions on Medical Imaging*, vol. 26, no. 4, pp. 471–478, 2007.

[68] J. Maintz and M. Viergever, "A survey of medical image registration," *Medical Image Analysis*, vol. 2, no. 1, pp. 1–36, 1998.

[69] J. Mazziotta, A. Toga, A. Evans, P. Fox, J. Lancaster, K. Zilles, R. Woods, T. Paus, G. Simpson, B. Pike, C. Holmes, D. Collins, P. Thompson, D. MacDonald, T. Schormann, K. Amunts, N. Palomero-Gallagher, L. Parsons, K. Narr, N. Kabani, G. Le Goualher, D. Boomsma, T. Cannon, R. Kawashima, and B. Mazoyer, "A probabilistic atlas and reference system for the human brain," *Journal of the Royal Society*, vol. 356, no. 1412, pp. 1293–1322, 2001.

[70] M. Miller, "Computational anatomy: shape, growth, and atrophy comparison via diffeomorphisms," *NeuroImage*, vol. 23, suppl. 1, pp. S19–S33, 2004.

[71] M. Miller, A. Trouve, and L. Younes, "On the metrics and Euler-Lagrange equations of computational anatomy," *Annual Review of Biomedical Engineering*, vol. 4, pp. 375–405, 2002.

[72] M. Miller and L. Younes, "Group actions, homeomorphisms, and matching: A general framework," *International Journal of Computer Vision*, vol. 41, pp. 61–84, 2001.

[73] J. Modersitzki, *Numerical Methods for Image Registration (Numerical Mathematics and Scientific Computation)*. New York: Oxford University Press, 2004.

[74] J. Morra, Z. Tu, L. Apostolova, A. Green, A. Toga, and P. Thompson, "Comparison of Adaboost and support vector machines for detecting Alzheimer's disease through automated hippocampal segmentation," 2007, submitted for publication.

[75] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Communications on Pure and Applied Mathematics*, vol. 42, pp. 577–685, 1989.

[76] P. Negron-Marrero, "A numerical method for detecting singular minimizers of multidimensional problems in nonlinear elasticity," *Numerische Mathematik*, vol. 58, no. 1, pp. 135–144, 1990.

[77] T. Nichols and A. Holmes, "Nonparametric analysis of PET functional neuroimaging experiments: A primer," *Human Brain Mapping*, vol. 15, pp. 1–25, 2001.

[78] M. Nielsen, P. Johansen, A. Jackson, and B. Lautrup, "Statistical warps, a least committed model," *International Conference on Medical Image Computing and Computer Assisted Intervention*, 2001.

[79] S. Osher and R. Fedkiw, "Level set methods: An overview and some recent results," *Journal of Computational Physics*, vol. 169, pp. 463–502, 2001.

[80] S. Osher and R. Fedkiw, *Level Set Methods and Dynamic Implicit Surfaces*, ser. Applied Mathematical Sciences. New York: Springer-Verlag, 2003.

[81] S. Osher and N. Paragios, *Geometric Level Set Methods in Imaging, Vision and Graphics*. Springer Verlag, 2003.

[82] S. Osher and J. Sethian, "Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations," *Journal of Computational Physics*, vol. 79, pp. 12–49, 1988.

[83] E. Parzen, "On estimation of a probability density function and mode," *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1065–1076, 1962.

[84] X. Pennec, "Left-invariant riemannian elasticity: A distance on shape diffeomorphisms?" *International Workshop on Mathematical Foundations of Computational Anatomy*, pp. 1–13, 2006.

[85] X. Pennec, R. Stefanescu, V. Arsigny, P. Fillard, and N. Ayache, "Riemannian elasticity: A statistical regularization framework for non-linear registration," in *International Conference on Medical Image Computing and Computer Assisted Intervention*, 2005, pp. 943–950.

[86] J. Pluim, J. Maintz, and M. Viergever, "f-information measures in medical image registration," *IEEE Transactions on Medical Imaging*, vol. 23, no. 12, pp. 1508–1516, 2004.

[87] B. Ridha, J. Barnes, J. Bartlett, A. Godbolt, T. Pepple, M. Rossor, and N. Fox, "Tracking atrophy progression in familial Alzheimer's disease: a serial MRI study," *The Lancet Neurology*, vol. 5, pp. 828–834, 2006.

[88] K. Rohr, *Landmark-Based Image Analysis: Using Geometric and Intensity Models*. Boston, MA: Kluwer, 2001.

[89] D. Shattuck and R. Leahy, "Brainsuite: An automated cortical surface identification tool," *Medical Image Analysis*, vol. 6, no. 2, pp. 129–142, 2002.

[90] D. Shen and C. Davatzikos, "Very high-resolution morphometry using mass-preserving deformations and HAMMER elastic registration," *NeuroImage*, vol. 18, no. 1, pp. 28–41, 2003.

[91] J. Sled, A. Zijdenbos, and A. Evans, "A nonparametric method for automatic correction of intensity nonuniformity in MRI data," *IEEE Transactions on Medical Imaging*, vol. 17, pp. 87–97, 1998.

[92] M. Sonka and J. Fitzpatrick, *Handbook of Medical Imaging*. Bellingham, WA: SPIE Press, 2000, vol. PM80.

[93] J. Storey, "A direct approach to false discovery rates," *Journal of the Royal Statistical Society*, vol. B 64, no. 3, pp. 479–498, 2002.

[94] J. Storey, "The positive false discovery rate: A Bayesian interpretation and the q-value," *Annals of Statistics*, vol. 31, no. 6, pp. 2013–2035, 2003.

[95] C. Studholme, C. Drapaca, B. Iordanova, and V. Cardenas, "Deformation-based mapping of volume change from serial brain MRI in the presence of local tissue contrast change," *IEEE Transactions on Medical Imaging*, vol. 25, no. 5, pp. 626–639, 2006.

[96] J.-P. Thirion, "Image matching as a diffusion process: An analogy with Maxwell's demons," *Medical Image Analysis*, vol. 2, no. 3, pp. 243–260, 1998.

[97] P. Thompson, J. Giedd, R. Woods, D. MacDonald, A. Evans, and A. Toga, "Growth patterns in the developing brain detected by using continuum mechanical tensor maps," *Nature*, vol. 404, no. 6774, pp. 190–193, 2000.

[98] P. Thompson and A. Toga, "Elastic image registration and pathology detection," in *Handbook of Medical Image Processing*, I. Bankman, Ed. Academic Press, 2000.

[99] P. Thompson and A. Toga, "A framework for computational anatomy," *Computing and Visualization in Science*, vol. 5, pp. 13–34, 2002.

[100] A. Toet and E. Franken, "Fusion of visible and thermal imagery improves situational awareness," *Displays*, vol. 18, pp. 85–95, 1997.

[101] A. Trouve, "Diffeomorphisms groups and pattern matching in image analysis," *International Journal of Computer Vision*, vol. 28, no. 3, pp. 213–221, 1998.

[102] L. Vese and T. Chan, "A multiphase level set framework for image segmentation using the Mumford and Shah model," *International Journal of Computer Vision*, vol. 50, no. 3, pp. 271–293, 2002.

[103] P. Viola and W. Wells, "Alignment by maximization of mutual information," *International Conference on Computer Vision*, pp. 16–23, 1995.

[104] L. Wang, J. Swank, I. Glick, M. Gado, M. Miller, J. Morris, and J. Csernansky, "Changes in hippocampal volume and shape across time distinguish dementia of the Alzheimer type from healthy aging," *NeuroImage*, vol. 20, pp. 667–682, 2003.

[105] W. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *Medical Image Analysis*, vol. 1, no. 1, pp. 35–51, 1996.

[106] R. Woods, J. Mazziotta, and S. Cherry, "MRI-PET registration with automated algorithm," *Journal of Computer Assisted Tomography*, vol. 17, no. 4, pp. 536–546, 1993.

[107] I. Yanovsky, M.-C. Chiang, P. Thompson, A. Klunder, J. Becker, S. Davis, A. Toga, and A. Leow, "Quantifying deformation using information theory: The log-unbiased nonlinear registration," *IEEE International Symposium on Biomedical Imaging*, pp. 13–16, 2007.

[108] I. Yanovsky, C. Le Guyader, A. Leow, P. Thompson, and L. Vese, "Unbiased volumetric registration via nonlinear elastic regularization," *UCLA Computational and Applied Mathematics Report 08-27. Submitted for publication*, 2008.

[109] I. Yanovsky, S. Osher, P. Thompson, and A. Leow, "Log-unbiased large-deformation image registration," *International Conference on Computer Vision Theory and Applications*, vol. 1, pp. 272–279, 2007.

[110] I. Yanovsky, P. Thompson, A. Klunder, A. Toga, and A. Leow, "Local volume change maps in nonrigid registration: When are computed changes real?" *International Conference on Medical Image Computing and Computer Assisted Intervention*, pp. 1–8, 2007.

[111] I. Yanovsky, P. Thompson, S. Osher, X. Hua, D. Shattuck, A. Toga, and A. Leow, "Validating unbiased registration on longitudinal MRI scans from the Alzheimer's Disease Neuroimaging Initiative (ADNI)," *IEEE International Symposium on Biomedical Imaging*, pp. 1091–1094, 2008.

[112] I. Yanovsky, P. Thompson, S. Osher, and A. Leow, "Topology preserving log-unbiased nonlinear image registration: Theory and implementation," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, 2007.

[113] I. Yanovsky, P. Thompson, S. Osher, and A. Leow, "Asymmetric and symmetric unbiased image registration: Statistical assessment of performance," *IEEE Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis*, 2008.

[114] I. Yanovsky, P. Thompson, S. Osher, and A. Leow, "Multimodal unbiased image matching via mutual information," *SPIE, Computational Imaging VI*, vol. 6814, no. 1, 2008.

[115] I. Yanovsky, P. Thompson, S. Osher, L. Vese, and A. Leow, "Multiphase segmentation of deformation using logarithmic priors," *IEEE Computer Society Workshop on Image Registration and Fusion*, pp. 1–6, 2007.

[116] H. Zhu, J. Ibrahim, N. Tang, D. Rowe, H. Xuejun, R. Bansal, and B. Petersen, "A statistical analysis of brain morphology using wild bootstrapping," *IEEE Transactions on Medical Imaging*, vol. 26, no. 7, pp. 954–966, 2007.

[117] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, 2003.