3D IMAGE RECONSTRUCTION FROM SERIAL SECTIONS

Yang Liu

A DISSERTATION

in

Applied Mathematics and Computational Science

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2014

James C. Gee, Associate Professor of Radiology
Supervisor of Dissertation

Charles L. Epstein, Thomas A. Scott Professor of Mathematics
Graduate Group Chairperson

Dissertation committee:
James C. Gee, Associate Professor of Radiology
Charles L. Epstein, Thomas A. Scott Professor of Mathematics
Philip Gressman, Associate Professor of Mathematics
Acknowledgments

It has been a great honor for me to be associated to so many great people in my past six years at UPenn. It fills me with great joy to look back and think about all the help, guidance and encouragement I found here. Now I would like to take this great opportunity to express my sincere thankfulness to them.

First of all, I owe my deepest gratitude to my advisor, Prof. James. Gee, who has shown his extraordinary scientific talent as a scientist, and excellent guidance as a mentor. He has my admiration for his passion in research and profound insights in the field of medical image analysis. For the past years in the lab, Prof. Gee has not only offered me with time, space and research independence, but also inspiration and encouragement. My research and this thesis could not even be possible without Prof. Gee’s help and guidance.

My gratitude also goes to my thesis committee members for their timely feedback and constructional suggestions. In particular, I would like to thank Prof. Charles Epstein for offering me the opportunity to come to UPenn and experience six wonderful years. Being the chair of AMCS, Prof. Epstein has been continuously
giving me invaluable advice both as an expert in mathematics and as a mentor in graduate studies. I would also like to thank Prof. Philip Gressman who has always impressed me with his broad knowledge in all mathematical study. I have worked with Prof. Gressman on an independent study course a few years ago in which I really appreciate his insight and encouragement.

The members of PICSL have given me selfless help and countless fun. I would like to thank Prof. Paul Yushkevich for his patience in teaching me all the basic techniques as well as innovative ideas in image analysis ever since I first entered this field. Gang Song is my best friend in the lab and I am really grateful for his technical help and advice. Yan Wang, Daniel Adler, Anh Truong and Ben Kandel have shared the office with me and the discussion with them has always been inspiring. I would also like to thank Brian, Yuanjie, Phil, Sandy, John, Hongzhi, Jeff, Baohua who have always been patient and willing to answer all my questions.

The past six years at UPenn could not have been so vivid without my good friends, especially from other fellow students from AMCS program, Xiaoxian Liu, Yan Wang, Yiqing Cai and Yin Xia. I would also like to thank Hao Sun, Ying Zhang and Xiaoning Shen for sharing the those memorable days at UPenn.

The support from my family is always vital, not only to my research and career, but more importantly, to my entire life. I deeply thank my dear parents, who nurtured me with the best environment they could provide. Finally, my appreciation and love goes to my dear wife Xiao. She has given me endless love and support, as
a girlfriend, a fiance, and a wife.
ABSTRACT

3D IMAGE RECONSTRUCTION FROM SERIAL SECTIONS

Yang Liu
James Gee

Analyzing 3D anatomical structures plays an important role in modern biology and medical science. Despite the fast development of 3D medical imaging techniques such as MRI and CT, 2D imaging methods such as microscopy still excel in terms of spatial resolution and the anatomical details in fine tissue structures. This calls for an effective volumetric image reconstruction method to find the correspondence among a sequence of 2D images and recover the 3D coherence that is lost in the sampling process. In this dissertation, we introduce a novel general framework for modeling volumetric image reconstruction with a rigorous and symmetric formulation. The formulation is robust to the noise and corrupted slices by using a greater range of adaptive weights in the regularization energy. The geometrical shape difference between sections is explicitly encoded via large deformation models which provide a mathematical method to quantitatively measure the shape change. Finally, a polyaffine image transformation model is presented which combines the power of deformation and the simplicity of the affine transformation.
# Contents

1 Introduction ................................. 1
   1.1 Medical Image Analysis .................. 1
   1.2 Image Registration and Nonlinear Transformation Model ........... 2
   1.3 Volumetric Image Reconstruction from serial sections ............. 4
   1.4 Polyaffine image transformation ............ 5
   1.5 problem statement and contribution ............ 6
   1.6 Overview of chapters .................. 7

2 Mathematical Tools .......................... 9
   2.1 Geometry on Manifold .................. 9
   2.2 Lie group and Lie algebra ............... 12
      2.2.1 Lie group actions and representations .......... 15
   2.3 Mechanics and Lagrangian Framework .......... 17
   2.4 Lie-Poisson equations .................. 19

3 Image Registration .......................... 20
## Introduction

3.1 Introduction ............................................. 20

3.2 General Formulation .................................... 23

3.3 Similarity metric ....................................... 24

3.4 Transformation Models and Regularization ............ 28
  3.4.1 Linear Transformation ............................... 29
  3.4.2 Deformable Transformation .......................... 30

3.5 Diffeomorphism group and LDDMM ...................... 31
  3.5.1 Diffeomorphism ..................................... 31
  3.5.2 LDDMM framework .................................. 33
  3.5.3 Geodesic shooting ................................... 36

3.6 Optimization and Variational Framework ............... 38
  3.6.1 Numerical Schemes .................................. 40
  3.6.2 Approximation for LDDMM ........................... 41

## Volumetric image reconstruction and application to 2D histology slices

4.1 Introduction ............................................. 42
  4.1.1 Image Reconstruction ............................... 42
  4.1.2 Motivation for histology image reconstruction .... 43

4.2 Background and Literature Review ...................... 46
  4.2.1 Histological Image Data ............................. 46
  4.2.2 Literature Review ................................... 47

4.3 Method .................................................. 50
4.3.1 Method Overview ........................................ 50
4.3.2 Formulation of general image reconstruction ........ 51
4.3.3 Formulation for volumetric image reconstruction from serial 2D images 52
4.3.4 Regularization ............................................. 54
4.3.5 Rigid transforms and its optimization .................. 58
4.3.6 Deformable transforms and its optimization ........... 60
4.3.7 With External Reference ................................. 62
4.4 Experiments .................................................. 64
4.4.1 Rigid transforms ........................................... 64
4.4.2 Deformable transforms ................................... 66
4.5 Concluding remarks ....................................... 67

5 Polyaffine transformation .................................. 69
5.1 Motivation .................................................... 69
5.2 Introduction .................................................. 71
5.3 Methods ...................................................... 72
5.3.1 Diffeomorphism in Polyaffine Model .................... 73
5.3.2 Optimal velocity field via energy minimization .......... 74
5.3.3 Time-invariant approximation by extending to trajectory of local region 76
5.3.4 Series of stationary velocity field for trajectory collision ... 77
5.4 Implementation .............................................. 78
5.4.1 Computing Mask Trajectories and Their Collision ....... 78
Chapter 1

Introduction

1.1 Medical Image Analysis

The past few decades has witnessed a rapid growth of medical imaging technologies, which provide unprecedented amounts of information and have become indispensible in medical diagnostics and scientific studies. Magnetic resonance imaging (MRI), computed tomography (CT), ultrasound and microscopic imaging now generate images with high resolution and improved quality, making possible more detailed and complete characterization of the anatomy. Moreover, significant advances in computational power, storage and communication techniques have facilitated not only accessibility to large data sets but also the capability to process data with speed and accuracy.

The fast development of technology has introduced new opportunities and chal-
lenges in the field of medical image analysis. There is a growing need for more effective analysis with applications to both clinical and research medicine. Traditionally, the analysis of medical image data relied mostly on manual examination by the medical experts. However, this process becomes laborious and difficult in many cases given the increasing scale and complexity of the data.

As a result, automatic methods are getting more advantageous to modern image analysis thanks to their efficiency and scalability. Advanced mathematical theory and models have been introduced to enable systematic and quantitative methods for image processing and statistical analysis.

1.2 Image Registration and Nonlinear Transformation Model

One major question in medical image analysis is how to make comparison among multiple images. This is achieved by one fundamental technique called image registration, which aims to find corresponding points among different images via spatial transformations. Image registration enables analysis of different objects in a normalized coordinate framework so that detailed comparison of the same anatomy and statistical information can be obtained.

One key component of image registration is the transformation model which defines the how one image can be transformed into another. In medical imaging,
images are used to represent the anatomical structures of different objects, which can be considered as variables belonging to a general space of shapes. In this way, transformations are essentially actions of shape changes that map one anatomy to another. A basic transformation model like linear model provides global alignment. However, the diverse and complicated geometrical structures within anatomy, as well as the nonlinear nature of anatomical shape change call for more accurate and sophisticated representations of the transformation.

Different nonlinear transformation models are thus developed to address the challenge. Many models are derived from physical laws and one of the most popular is large-deformation mappings. Large-deformation mappings earn that reputation by guaranteeing diffeomorphism and are capable of handling long range deformation. The deformation framework also provides a rigorous way of defining the geometrical mapping and can quantitatively measure and analyze the shape change in the images. The set of diffeomorphic transformation in large-deformation mapping framework is formulated as a group where the level of smoothness is typically controlled by parameters describing the metric on the tangent space of the diffeomorphism group. This introduces a infinite-dimensional Lie group structure on the transformation group and the elegant theory of the Euler-Poincare equation indicates that the geodesic path is only determined by the initial velocity field. The smoothness of the transformation is controlled by the Lie algebra.
1.3 Volumetric Image Reconstruction from serial sections

Analyzing 3D anatomical structures plays an important role in modern biology and medical science. The fast development of many 3D medical imaging techniques such as MRI and CT has provided direct access to the 3D image representation of the tissue. However, 2D imaging methods such as optical microscopy still excel in terms of the spatial resolution and representing the anatomical details of fine tissue structures. The detailed information in these images is crucial in many studies, which require analysis at the cellular level. However, the preparation process in 2D image process often results in the loss of three dimensional geometric continuity across slices, making it difficult to do direct volumetric analysis.

Therefore building a continuous 3D volume from 2D image sections is critical in the visualization and volumetric image analysis of the underlying anatomy. This is achieved by 3D image reconstruction, which essentially aims to find the correspondence of structures across serial image sections. In this dissertation, we are interested in the development of a rigorous framework for the volumetric image reconstruction.

The past efforts of volumetric histology image reconstruction have been mainly on the development of accurate inter-slice registration schemes. One major challenge in direct image reconstruction is the irregular shape change and high inci-

dence of corruption among the 2D images in the sampling process. This often results in a poor final result due to error accumulation and propagation in the concatenation procedure. The problem is often alleviated by means of cross-slice smoothing/regularization methods. In our work, this is addressed by introducing a symmetric regularization term which includes the image difference among greater range of image sections reference and also uses an adaptive weight function directly driven by the intensity similarity of the data. This method increases the robustness of the formulation by averaging the effect of multiple transformations and lower the effect of the less informative sections.

Another issue in the previous work is that no explicit shape change is allowed between images in the reconstructed image sequence. The shape change is natural across slices in the reconstructed volume. The nonlinear transformation model, especially the large deformation model, provides a quantitative definition of the shape change. This is explicitly encoded in our total energy of the reconstructed image volume.

1.4 Polyaffine image transformation

In the sampling process, histological slices often undergo transformations that are approximately rigid or affine when restricted to local regions (Arsigny et al., 2005). This calls for an effective method to combine multiple affine transforms each of which is defined on a local region. In this dissertation, we present a polyaffine
transformation model which generates a diffeomorphism of the entire domain that encapsulates all the local transforms by a smooth and invertible displacement field. Compared with traditional weighting schemes used in combining local transforms, our new scheme guarantees that the resulting transform precisely preserves the value of each local affine transform. By introducing the trajectory of local regions instead of using regions themselves, the new approach encodes precisely each local affine transform using a diffeomorphism with one or more stationary velocity fields.

1.5 problem statement and contribution

The main focus of this dissertation lies in finding the anatomical correspondence for a sequence of images and building a continuous volume. Although there has been a large number of reconstruction strategies introduced in the medical imaging literature, only limited effort has been put into building a general and rigorous framework. We believe that a detailed analysis of the reconstruction formulation has been missing and we have been motivated to introduce one. As most of the principled reconstruction approaches rely on the formulation for the distance between images, we sum up the image distance in terms of the image intensity and shape difference. The shape change of our framework is explicitly integrated in the reconstructed volume which does not

In summary, our main contributions to the sequential image reconstruction include:
1. a rigorous unified framework to formulate 3D image reconstruction from serial sections.

2. a novel method of regularization which is symmetric and robust to noise in modeling the overall shape of the 3D structure.

3. a polyaffine image transformation framework to combine the simplicity of the affine transformation and the flexibility of the deformable transformation.

1.6 Overview of chapters

The remainder of this dissertation is organized as follows:

In Chapter 2, we review some of the mathematical tools which are used in the parameterization and measure of the diffeomorphism group of transformations. It includes some basic concepts from Lie group and algebra theory and classical mechanics.

In Chapter 3, we present a review of the image matching problem (registration) and are especially focused on the geometry of the large deformation framework. The framework is the foundation for diffeomorphic registration and motivates the model and formulation for serial image reconstruction.

In Chapter 4, we present a detailed description of the volumetric image reconstruction problem from serial sections, especially for histological data. We build a general mathematical framework both the shape change and the using the large-
deformation diffeomorphic image registration framework.

Chapter 5, we presents a polyaffine transformation model, which combines the locally affine transformations, which extends to the whole domain and introduces a fast approximation.

Chapter 6 contains a discussion of contribution of this thesis and an outline of future work.
Chapter 2

Mathematical Tools

In this chapter, the theoretical background necessary to describe the evolution of transformations in $\mathbb{R}^n$ under the action of a subgroup of diffeomorphism is presented. These concepts are essential in building the fluid-like transformation model and geometry in the image registration and image reconstruction framework.

2.1 Geometry on Manifold

Even though in the field of image analysis the manifolds we are dealing with are infinite-dimensional, it is still worthwhile to introduce the basic conceptions in the finite-dimensional case.

Intuitively, manifold is a topological space which is locally Euclidean so that we can apply some of the geometric properties of Euclidean space to some more general spaces.
Definition 2.1.1. A manifold $M$ of dimension $n$ is a topological space with a collection of coordinate charts $\mathcal{U} = \{(U_\alpha, \varphi_\alpha)\}_{\alpha \in I}$ that map open $U_\alpha \subset M$ into $\mathbb{R}^n$ via homeomorphic function $\varphi_\alpha : U_\alpha \to \varphi_\alpha(U) \subset \mathbb{R}^n$ with following property.

- $M$ is Hausdorff.
- $M = \bigcup_{\alpha \in I} U_\alpha$.
- $\{U_\alpha\}_{\alpha \in I}$ has a countable basis.

To define a smooth manifold, we require that for any two overlapping charts $(U_\alpha, \varphi_\alpha), (U_\beta, \varphi_\beta) \in \mathcal{U}, U_\alpha \cap U_\beta \neq \emptyset$, the coordinate change $\varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\alpha \cap U_\beta)$ is $C^\infty$ with $C^\infty$ inverse.

In the remainder of the text, it is assumed that all manifolds are infinitely differentiable (smooth).

The tangent space is a vector space that contains the possible directions in which one can pass while traveling on a path along a manifold through $x$. On general manifolds, there are several equivalent ways of defining tangent vectors and we choose the definition as directional derivative operators on smooth functions.

Definition 2.1.2. Let $C^\infty(x)$ be the algebra of $C^\infty$ functions whose domain of definition includes some open neighborhood of $x$, with functions identified if they agree on any neighborhood of $p$. A tangent space $T_x M$ to $M$ at $x$ is the set of all linear mappings $v : C^\infty(x) \to \mathbb{R}$ such that $\forall f, g \in C^\infty(x)$ and $\alpha, \beta \in \mathbb{R}$,

- $v(fg) = v(f)g(x) + f(x)v(g)$
\[ v(\alpha f + \beta g) = \alpha v(f) + \beta v(g) \]

A tangent vector to \( M \) at \( x \) is any \( v \in T_xM \).

**Definition 2.1.3.** Let the \( T_xM \) be the tangent space. The cotangent space \( T^*_xM \) at \( x \) is defined as the dual space of \( T_xM \). Concretely, elements of the cotangent space are linear functionals on \( T_xM \).

Every smooth map between smooth manifolds induces natural linear maps between the corresponding tangent spaces.

**Definition 2.1.4.** If there is a smooth map \( \varphi : M \to N \), then a tangent map associated to \( \varphi \) at \( x \) is a map \( d\varphi : T^*_xM \to T^*_{\varphi(x)}N \) satisfying

\[
d\varphi(v)(f) = v(f \circ \varphi) \quad v \in T_xM, \forall f \in C^\infty(x). \tag{2.1.1}\]

This tangent map is also called the pushforward and it induces a linear map \( d^*\varphi : T^*_{\varphi(x)}N \to T^*_xM \) called the pullback, satisfying

\[
d^*\varphi(w)(v) = w(d\varphi(v)) \quad \forall w \in T^*_{\varphi(x)}N, v \in T_xM, \tag{2.1.2}\]

Once tangent spaces have been introduced, one can define vector fields, which are abstractions of the velocity field of particles moving on a manifold. A vector field attaches to every point of the manifold a vector from the tangent space at that point, in a smooth manner.

**Definition 2.1.5.** A tangent bundle of a differential manifold \( M \) is a disjoint union of the tangent spaces of \( M \). That is

\[ TM = \bigcup_{x \in M} T_xM = \bigcup_{x \in M} \{ (x, y) \mid y \in T_xM \} \]
Definition 2.1.6. A vector field on a manifold is a map \( X : M \rightarrow TM \). A smooth vector field on \( M \) is a smooth map from \( M \) to \( TM \), which is often denoted as \( \mathfrak{X} \).

Definition 2.1.7. Lie bracket \([ \cdot, \cdot ] : \mathfrak{X} \times \mathfrak{X} \rightarrow \mathfrak{X}\) satisfying

\[
[X, Y](f) = X(Y(f)) - Y(X(f)), \quad \forall f \in C^\infty(M),
\]

(2.1.3)

Definition 2.1.8. A Riemannian metric on a manifold \( M \) is a smooth symmetric positive-definite covariant 2-tensor field \( g_x : T_xM \times T_xM \rightarrow \mathbb{R}, p \in M \), such that for all smooth vector fields \( X, Y \in \mathfrak{X}, x \rightarrow g_x(X(x), Y(x)) \) is a smooth function \( M \rightarrow \mathbb{R} \). The pair \((M, g)\) is called a Riemannian manifold.

2.2 Lie group and Lie algebra

Lie groups and Lie algebras are used extensively in the theoretical part of the development of the geometrical properties of the geodesic shooting method.

Definition 2.2.1. A Lie group is a smooth manifold with a group structure such that multiplication \( \mu : G \times G \rightarrow G, \mu(g, h) = gh \) and inversion \( g \rightarrow g^{-1} \) are smooth maps.

One of the fundamental properties of a Lie group is that given a local chart, one may construct an entire atlas. This is achieved by using the group multiplication.

For any given \( g \in G \), the left translation by \( g \) is the map

\[
L_g : G \rightarrow G, h \rightarrow L_g(h) := gh.
\]

(2.2.1)
In the same way, we can define right translation $R_g : G \to G, h \to R_g(h) := hg$.

We give the abstract definition of a Lie algebra and show that the tangent space at the identity $T_eG$ of any Lie group $G$ with a particular bracket, is a Lie algebra.

**Definition 2.2.2.** A *Lie algebra* is a vector space $\mathfrak{g}$ together with the binary operation $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, called the *bracket*, such that

- **bilinearity:**
  
  $[\alpha u + \beta v, w] = \alpha [u, w] + \beta [v, w], \quad [w, \alpha u + \beta v] = \alpha [w, u] + \beta [w, v], \forall u, v, w \in \mathfrak{g}, \alpha, \beta \in \mathbb{R}$

- **skew-symmetry:** $[u, v] = -[v, u]$

- **Jacobi identity:** $[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0$

An important example of Lie algebra is the vector space $\mathfrak{X}(M)$ of all smooth vector fields on a smooth manifold $M$. The bracket is the Lie bracket defined in **Definition 2.1.7**

**Definition 2.2.3.** A *Lie subalgebra* of a Lie algebra $\mathfrak{g}$ is a subspace $\mathfrak{h}$ of $\mathfrak{g}$ closed under the bracket. Clearly, this subalgebra $\mathfrak{h}$ is itself a Lie algebra.

Consider a Lie group $G$ and its tangent space at the identity $T_eG$. We are looking for a meaningful bracket $[\cdot, \cdot]_{T_eG}$ on $T_eG$. We have just learned that the vector space $\mathfrak{X}(G)$ is a Lie algebra. If we have an isomorphism between $T_eG$ and a subalgebra of $\mathfrak{X}(G)$, then we could define the bracket on $T_eG$ via this isomorphism.
\( \lambda : T_eG \to C \subset \mathfrak{X}(G) \), where \( C \) is a subalgebra of \( \mathfrak{X}(G) \), the bracket on \( T_eG \) may be defined via

\[
[\xi, \eta]_{T_eG} = [\lambda(\xi), \lambda(\eta)]_{\mathfrak{X}(G)}
\]

The map \( \lambda \) will be defined using the left translation map \( L_g \). Recall the tangent map induced by \( L_g : G \to G \) is \( dL_g \). By doing this operation for every \( g \in G \), we define a vector field

**Definition 2.2.4.** The left extension of any \( \xi \in T_eG \) is the vector field \( X^L_\xi \) given by \( X^L_\xi(g) := T_eL_g(\xi) \).

**Definition 2.2.5.** A vector field \( X : G \to TG, h \to X(h), \) is called left-invariant if \( \forall g \in G \)

\[
dL_gX(h) = X(gh)
\]

The set of left-invariant vector fields on \( G \) is denoted by \( \mathfrak{X}_L(G) \).

A vector field on \( G \) is left-invariant if and only if it equals \( X^L_\xi \) for some \( \xi \in T_eG \) and the following map is a vector space isomorphism.

The subspace \( \mathfrak{X}_L(G) \) is a Lie subalgebra of \( \mathfrak{X}(G) \).

For any Lie group, the Lie bracket on \( T_eG \) is defined using the isomorphism.

**Definition 2.2.6.** The tangent space at the identity \( T_eG \) of a Lie group \( G \), together with the Lie bracket defined above, is called the Lie algebra of \( G \).

**Definition 2.2.7.** A map \( \gamma : \mathbb{R} \to G \) is a one-parameter subgroup if it is a homomorphism of Lie groups satisfying
• \( \gamma \) is continuous

• \( \gamma(0) = e \)

• \( \gamma(t + s) = \gamma(t) \circ \gamma(s) \)

Given a vector \( \xi \in \mathfrak{g} \) and consider the left extension \( X_\xi \)

\[
\frac{dg}{dt} = X_\xi(g),
\]

\[
g(0) = e. \tag{2.2.2}
\]

**Definition 2.2.8.** Consider the left action of a Lie group \( G \) on the manifold \( M \). Let \( \xi \in \mathfrak{g} \) be a vector in the Lie algebra of \( G \) and consider the one-parameter subgroup \( \{ \exp(t\xi) : t \in \mathbb{R} \} \subset G \). The orbit of an element \( x \) with respect to the subgroup is a smooth path \( t \rightarrow (\exp(t\xi))x \). The *infinitesimal generator* associated to \( \xi \) at \( x \in M \), denoted \( \xi_M(x) \), is the tangent vector to this curve at \( x \), that is:

\[
\xi_M(x) := \frac{d}{dt} \bigg|_{t=0} (\exp(t\xi))x \in T_xM. \tag{2.2.3}
\]

The smooth vector field \( \xi_M : M \rightarrow TM, x \rightarrow \xi_M(x) \) is called the *infinitesimal generator vector field associated to \( \xi \).*

### 2.2.1 Lie group actions and representations

In the basic group representation theory, group representations describe abstract groups in terms of linear transformations of vector space. For Lie groups, there is a natural adjoint representation obtained by linearizing the action of the group
on itself by conjugation. In this way, elements of the group can be represented as linear transformations of the group’s Lie algebra which is considered to be a vector space.

**Definition 2.2.9.** The action of $G$ on itself by *conjugation*, or *inner automorphism*, is

$$G \times G \to G, (g, h) \to I_g(h) := L_g \circ R_{g^{-1}(h)} = ghg^{-1}$$

**Definition 2.2.10.** The *adjoint action* of $G$ on $\mathfrak{g}$ is defined by taking the derivative of $I_g(h)$ at group identity $h = e$,

$$\text{Ad} : G \times \mathfrak{g} \to \mathfrak{g}, (g, \xi) \to \text{Ad}_g \xi := d(I_g)_e(\xi)$$

**Definition 2.2.11.** The *co-adjoint action* of $G$ of $\mathfrak{g}^*$ is the inverse dual of the adjoint action

$$\text{Ad}^* : G \times \mathfrak{g}^* \to \mathfrak{g}^*, (g, \mu) \to \text{Ad}^*_{g^{-1}} \mu$$

where

$$(\text{Ad}^*_{g^{-1}} \mu, \xi) = (\mu, \text{Ad}_{g^{-1}} \xi), \quad \forall \mu \in \mathfrak{g}^*, \xi \in \mathfrak{g}$$

$(\cdot, \cdot) : \mathfrak{g}^* \times \mathfrak{g} \to \mathbb{R}$ is the natural pairing.

**Definition 2.2.12.** The infinitesimal generator map,

$$\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}, (\xi, \eta) \to \text{ad}_\xi(\eta) = \xi_g(\eta) = \frac{d}{dt}\text{Ad}_{\exp(t\xi)}\eta$$

is called the *adjoint action* of $\mathfrak{g}$ on itself, even it is not a group action.
It can be proved that the adjoint action on the Lie-algebra satisfies the following representation

**Definition 2.2.13.** Taking the differential of $\text{Ad}_g$ at the identity $g = e$ gives the adjoint representation of the Lie algebra.

To sum it up, we here defined

### 2.3 Mechanics and Lagrangian Framework

**Definition 2.3.1.** A Lagrangian on a manifold $M$ is defined as a smooth function $\mathcal{L} : TM \to \mathbb{R}$. For any smooth curve $\alpha(t) : [\tau_0, \tau_1] \to M$ the action functional associated with $\mathcal{L}$ is defined by

$$S(\alpha) = \int_{\tau_0}^{\tau_1} \mathcal{L}(\alpha(t), \dot{\alpha}(t))dt$$

(2.3.1)

Of special importance are smooth curves with fixed endpoints $\alpha(\tau_0) = x_0, \alpha(\tau_1) = x_2$ for which $F_{\mathcal{L}}(\alpha)$ is minimized.

This is a very important concept in mechanics and the minimizer of the functional governs the motion of the object. To minimize this functional on the paths, we need to introduce the concept of the variational derivatives.

Points in a domain $D$ represent the positions of material particles of the system. These points are labeled by $X \in \mathbb{R}^n$. A configuration denote by $g$ is an element of $\text{Diff}(D)$. The configuration space $\text{Diff}(D)$ is a group, with a group operation being the composition and the group identity being the identity map. This group acts on
$D$ in the obvious way $g \cdot X := g(X)$, where the “dot” notation is used for group action.

The location of the particles after the action of $\phi(t)$ is denoted by $x(t, X) := \phi(t) \cdot X$, also called the Lagrangian or material trajectory. The Lagrangian velocity of the system is defined by $V(t, X) := \frac{\partial}{\partial t} x(t, X)$, and the Eulerian velocity by $v(t, x) := V(t, \phi(t)^{-1} x)$, representing the velocity at time $t$ of the particle located at $x$. The Eulerian velocity $v$ can also be regarded as a time-dependent vector field $v_t \in \mathfrak{X}(D)$, where $v_t(x) := v(x, t)$.

The configuration space $\text{Diff}(D)$ is an infinite-dimensional Lie group. One specific case is when the domain $D$ is compact and without boundary and only $C^\infty$ diffeomorphisms are considered, then $\text{Diff}(D)$ is a Fréchet manifold and in fact a Fréchet Lie group.

Given a path $g(t)$ in $\text{Diff}(D)$, the corresponding Lagrangian velocity fields $V_t$ are also denoted $\dot{g}(t)$. For a given $t$, the velocity field $\dot{g}(t)$ is called the tangent vector to $\text{Diff}(D)$ at $g(t)$. The tangent space of $\text{Diff}(D)$ at $g$, denoted $T_g \text{Diff}(D)$, is the set of all tangent vectors to $\text{Diff}(D)$.

Note that $\dot{g}(x) \cdot X \in T_{g(t)}X D$, so $\cdot g(t)$ is not, in general, a vector field. However, if

The Lie algebra of $\text{Diff}(D)$ is $\mathfrak{X}(D)$ with the Lie bracket defined by

$$[u, v]_L := [X_u^L, X_v^L](e) \quad (2.3.2)$$

for all $u, v \in \mathfrak{X}(D)$. 

18
Now equip the Lie algebra $\mathfrak{g}$ of $G$ with a positive definite quadratic form directly related with the translation invariant energy of the system. In the present context, the quadratic form is equivalent to a symmetric positive bilinear metrics on $G$. To define this metric on any tangent space on $G$ and not just at the identity $e_G$, $\langle \cdot, \cdot \rangle_{e_G}$ is required to be right-invariant under right pull-backs of the metric.

### 2.4 Lie-Poisson equations

**Theorem 2.4.1.** Let $G$ be a Lie group together with a Lagrangian $\mathcal{L} : TG \to \mathbb{R}$. Then in a right trivialized coordinates $(g, \xi) := (g, g^{-1} \cdot g) := (g, (dg^{-1})_{g}) \in G \times \mathfrak{g}$ the geodesic equation on $\mathfrak{g}^{\ast}$ takes the form

$$\frac{d}{dt} \left( \frac{\delta L}{\delta \xi} \right) + \text{ad}^{\ast}_{\xi} \frac{\delta L}{\delta \xi} = 0 \tag{2.4.1}$$
Chapter 3

Image Registration

3.1 Introduction

In this chapter, we give some background and a summary of image registration. We are especially focused on the geometry in the large deformation framework. This framework is the foundation for diffeomorphic registration and motivates the model and formulation for serial image reconstruction.

What is image registration? Image registration is a procedure to find a transformation that aligns each point in one image to the corresponding point in the other, so that different images can be compared in a normalized space. In medical imaging, images are essentially representations of the anatomical shapes of real objects. Therefore, such an alignment facilitates a detailed comparison of different subjects, corresponding structures, and allows statistical analysis of information about the
same anatomy in a single data coordinate frame. Image registration has a broad range of application: e.g. analysis of temporal evolution, fusion of multimodality images (MRI, CT, DTI), inter-patients comparison (statistical analysis) and atlas construction.

Figure 3.1: image registration and matching

Figure 3.1 demonstrates an example of an alignment scenario through a two-dimensional pairwise registration.

The key components of image registration include

1. **Image similarity metrics**

   Image similarity metrics are functions that quantify the degree of similarity between intensity patterns between images. They are critical to the assessment of the quality of image registration. A variety of image similarity metrics have been proposed in the medical image community and an appropriate choice often depends on the particular application and modality of the images to be registered. The two general classes are feature-based metrics and intensity-based metrics.
2. **Image transformation parameterization and regularization**

Image transformation parameterization defines the space in which an image can be transformed into another. In general, transformation models can be classified into two types: linear and deformable transformation. The former employs a global affine/rigid alignment to match the overall shape and size while the latter enables the model to capture local variations of the objects. In most deformable models, transformation parameterization are high-dimensional and thus regularizations are required to assure the well-posedness of the problem.

3. **Optimization scheme**

The ultimate goal of image registration problem is to find the optimal transformation. This becomes a optimization problem given that the objective function and the feasible domain are set up by the similarity metric and parameterization. A variety of optimization schemes have been developed based on the formulation of the problem and one of the most widely-used classes of methods is gradient descent. In the context of functional optimization for deformable transformation, the ordinary descent method is translated into its infinite dimensional counterpart with the variational gradient generated in the Euler-Lagrange equation framework.
3.2 General Formulation

In medical image analysis, an image $I$ is represented as a real-valued (vector) function $I: \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^d$, where $n = 2$ or $3$ is the image dimension and $d$ is the number of color channels. Generally it is assumed that the maps are Lipschitz continuous, bounded and the region $\Omega$ is a regular bounded region ($C^2$). Given two images $I$ (fixed) and $J$ (moving), the goal of image registration is to find a correspondence between them by computing a transform $\phi : \Omega \rightarrow \Omega$. The transformation of the moving image $J$ under mapping $\phi$ is the pull-back image $J \circ \phi$ and the “optimal” transform should make $J \circ \phi$ close to $I$.

This is often formulated as minimization of an energy functional

$$E(\phi) = E_A(I, J, \phi) + E_R(\phi).$$

The matching energy $E_A(I, J, \phi)$ accounts for the similarity between the transformed moving image $J \circ \phi$ and the fixed image $I$. This is often measured by different image metrics, e.g. SSD (sum of squared difference)

$$E_A(I, J, \phi) = \int_{\Omega} \| J \circ \phi - I \|^2 dx.$$  \hspace{1cm} (3.2.2)

Other metrics like (CC (cross correlation) and MI (Mutual Information)) are also used to account for other types of image difference, which will be discussed in detail in section 3.3.

Also a regularization term $E_R(\phi)$ is included to ensure the regularity of the transform $\phi$, e.g. smoothness and invertability. The form of the regularization depends
on the parameterization of the model. For simple models like rigid registration, not much regularization is needed because the regularity of the transformation is ensured by its rigidity. For deformable registration, parametric models like Free-Form Deformation (FFD) (Rueckert et al., 1999) uses basis functions to reduce the dimension and ensure smoothness. For nonparametric model (Bajcsy and Kovačič, 1989; Christensen et al., 1996), the transformation is defined at each point in the domain Ω, e.g.

\[
\phi(x) = x + u(x). \quad (3.2.3)
\]

and the regularization can be enforced on the function \( u(x) \). All of these issues will be discussed in more details in section 3.4.

### 3.3 Similarity metric

Image similarity metrics are a set of functions that are designed to evaluate the quality of the image matching. During the registration process, the metric assigns a numerical score to the image pairs under the current transformation status. The choice of the image metric is one of the most critical parts of the image registration problem, because they quantitatively define the goal of the registration process and are the driving force of the transformation.

There are generally two classes of image similarity functions: feature-based and intensity-based. Feature-based similarity metrics rely on extracting key features (e.g. anatomical landmarks, edge and corner information, etc.) from images and
finding the correspondences among them. These approaches assume that a precise alignment of features is sufficient for aligning the rest of the input data. Landmarks or features are identified and constructed either manually e.g. (Maurer et al., 1995; Kremser et al., 1997) or by automatic algorithms acting on the original images before the registration process e.g. edge detection (Canny, 1986), scale analysis (Lowe, 2004).

Landmark matching are based on point sets matching with the objective function measuring the distance of the corresponding point pairs.

\[ PS(\{x_i^{(0)}\}, \{x_i^{(1)}\}) = \sum_{i=1}^{n} (x_i^{(0)} - x_i^{(1)})^2, \]  

(3.3.1)

where \(\{x_i^{(0)}\}, \{x_i^{(1)}\}\) are the two sequence of corresponding landmarks in the two images.

Intensity based similarity function relies directly on the image pixel intensities. These intensity-based methods calculate and compare various mathematical expressions using only the intensity values of the inputs. Over the past few decades, many such objective functions have been proposed. Here we give some examples of image similarity metrics used in this dissertation.

One of the simplest measures of image similarity is “Sum of Squared difference” (SSD), which is given by summing over the pointwise squared difference of corresponding pixels from different images.

The assumption of using SSD is that points from the same anatomical structures are of the similar intensity. However in practice, images often have intensity
inhomogeneities and noise and are often captured from different sources, making it inappropriate to only measure similarity in terms of absolute pointwise intensity difference.

As a result, people introduced a variety of metrics based on statistical principles to compare images to address this. Other statistical similarity measure has been used, including cross correlation, mutual information, Kullback-Leibler divergence, maximum correlation ratio, etc.

One statistical measure that has been used in medical image registration is Cross Correlation (CC) (Duda et al., 1973; Bajcsy et al., 1983). It measures image difference by taking the correlation information of intensity from patches in both images. The patches can be specified locally as the neighborhood of each point (voxel) in both images or globally as the entire images. The local case can be defined as

\[ CC_l(I, J, x) = \int_{N(x)} \frac{(I(y) - \mu_I(x))(J(y) - \mu_J(x))}{\sigma_I(x)\sigma_J(x)} dy \] (3.3.2)

where \( \mu_I(x), \sigma_I(x) \) are the mean and standard deviation of the intensity of the neighborhood \( N(x) \). With this, the total similarity energy is

\[ CC(I, J) = \int_{\Omega} CC_l(I, J, x) dx. \]

The use of CC for registration is based on the assumption that two neighborhoods (or images), are well aligned when their intensity distributions are highly correlated. Compared to SSD, this is more robust to image noise and mismeasurement even in the mono-modality case. Furthermore, cross correlation has shown to
perform well in many real-world computer vision applications where one requires robustness to unpredictable illumination, reflectance, etc. This adapts naturally to situations where locally varying intensities occur and is suitable for some multi-modality problems (Avants et al., 2008). CC depends only on estimates of the local image average and variance which may be accurately/exactly measured with relatively few samples.

Another widely-used similarity metric is mutual information (MI) (Maes et al., 1997; Wells et al., 1996). The idea was first developed in the field of the information theory. Taking two images as two signal sequences, mutual information describes how much information we know about one image if the intensity information about the other image is completely given. Given two intensity images $I$ and $J$, MI metric is defined as

$$MI(I, J) = H(I) + H(J) - H(I, J)$$ (3.3.3)

Here $H(\cdot)$ is the entropy of an image (pair) which measures the amount of disorder or uncertainty of the intensity distribution of image (Cover and Thomas, 2012). It is defined as

$$H(I) = \int_x p(I(x)) \log(p(I(x)))$$ (3.3.4)$$

$$H(I, J) = \int_x p(I(x), J(x)) \log(p(I(x), J(x)))$$

where $p(I(x), J(x))$, $p(I(x))$ and $p(J(x))$ are the probabilities of the intensities of the images in the domain $\Omega$. In the discrete case, all these are obtained by sampling the image.
Most widely-used estimators of mutual information are Parzen windowing and histogram estimation.

Compared to CC which assumes the linear dependence of the image intensity, MI does not have a specified function form of intensity dependence among images, making it more flexible to model correspondence between image intensity from different modality. However this does not come without a cost. MI need a large number of samples for reliable estimate of the statistics and the. Therefore, as localization in the MI estimate increases, its statistical reliability decreases.

3.4 Transformation Models and Regularization

The choice of transformations is defined by the characteristics of the input images and the purpose of the registration. Generally, there are two types of transformations in image registration: linear and deformable. Linear transformation is mainly used for global alignment of different images. It specifies the displacement, orientation and shape. Also it can also provide a good initialization for other more sophisticated registration methods. For linear transformation, the regularity of the transformation can be guaranteed by the overall number of degrees of the freedom, so normally no additional regularization term is required. Deformable registration methods are used to capture local variation between different images. A dense transformation field is often introduced to parameterize the pointwise displacement in the image. Because the problem is often ill-posed, regularization is needed to
restrict the smoothness and norm of the transformation.

3.4.1 Linear Transformation

The linear transformation model employs finite global parameters to align the images. It is represented as a projection matrix $A$ and a translation vector $t$:

$$
\phi(x) = Ax + t
$$

(3.4.1)

The projection matrix can be decomposed into the product of the rotation matrix $R$, scaling matrix $S$ and the shearing matrix $K$,

$$
A = R \times S \times K
$$

(3.4.2)

Depending on the choice of the projection matrix, the linear transformation can be classified into different types; two common choices in practice are rigid and affine transformation. The former only allows rotations ($A = R$) as well as translations, thus preserving the shape and size of the structures, while the latter offers additional flexibility by allowing scaling and shearing.

Thanks to the simple formulation and low dimensionality of linear transformation, they are important in the basic inter-subject studies and robust and efficient algorithms and implementations are available (Woods et al., 1998; Jenkinson and Smith, 2001). Also a good linear registration is critical as a preprocessing step in non-rigid registration to make the global transformation as small as possible.
3.4.2 Deformable Transformation

Deformable transformation models are introduced to characterize the local shape variation in the images and thus require pointwise definition of the transformation field. There are generally two types of deformable models: parametric and nonparametric.

In parametric models, different methods are employed to reduce the dimension of the transformation, e.g., introducing basis functions at control points and smoothly interpolate the movement of rest of the voxels. In this way, the smoothness of the transformation field is assured by the smoothness of the basis functions and the interpolation methods. In addition, parametric models are often computationally efficient thanks to their relatively low dimensionality.

One representative of the parametric deformable models is the Free-Form Deformation (FFD) model (Rueckert et al., 1999), which basically approximates the displacement $u(x)$ of each voxel $x=(x_1, x_2)$ by a smooth B-spline interpolation.

$$u(x) = \sum_{l=0}^{3} \sum_{m=0}^{3} B_l(x_1) B_m(x_2)$$

(3.4.3)

The nonparametric transformation models discretize the transformation field directly and define it at each pixel. The models are often inspired by the physical laws and regulate transformation with the energy derived in mechanics. Common examples include elastic models (Bajcsy and Kovačič, 1989; Gee and Bajcsy, 1998) and viscous fluid flow models (Christensen et al., 1996). Elastic models have their roots in continuum mechanics in which the regularization energy is defined on the
displacement of each point
\[
E_R(\phi) = \frac{1}{2} \int_\Omega (\mathcal{L}u(x))^2
\]  (3.4.4)

where \( \mathcal{L} \) is an differential operator which may take different forms e.g. membrane model \( L = \nabla \), Laplacian model \( L = \Delta \), Cauchy-Navier model \( L = \gamma I_d - \alpha \Delta \) (where \( \alpha \) and \( \gamma \) are constant defined by the material).

The elastic registration models work well for small deformations. However, it strongly penalize large displacements and may not be useful in applications when large nonlinear deformations are natural.

To address the challenge to characterize the large deformations, viscous flow models are thus introduced, which operate on simulated velocity to deform images. The simulated velocities are updated in small steps and the final transformation is an integration of the velocity field over time following the ordinary differential equation. More details are given in the following sections 3.5.

### 3.5 Diffeomorphism group and LDDMM

#### 3.5.1 Diffeomorphism

In medical imaging, the primary interest is to study the anatomical structures of different objects, which can be considered as variables belonging to a general space of shapes. As a result, a coordinate transformation can be considered as change between individual shapes, which accounts for the variation in anatomy. To make
image registration useful in medical study and diagnosis, the transformation must be assigned in some clinically meaningful way. On most occasions, people make the assumption that similar structures are present in both images. Therefore it is desirable that the deformation field be smooth and invertible (so that every point in one image has a corresponding point in the other). Such smooth, invertible transformations are called diffeomorphisms. Given the law of composition and identity element, all the diffeomorphisms essentially form a group of transformation \( G \):

\[
G = \{ g : g \text{ and } g^{-1} \text{ are both smooth and invertible} \} \quad (3.5.1)
\]

\( G \) acting on an image \( I \in \mathcal{I} \) defines a group action which deforms a shape in a topology-preserving way. A representative of the deformable object class, represents a topologically identifiable structure which is diffeomorphically deformed under the action of a group of smooth invertible maps i.e. an anatomical orbit. Then \( \mathcal{I} \) is composed of a set of disjoint orbits which are each homogeneous spaces under the group action of diffeomorphisms.

This theory of diffeomorphism group has been studied extensively in the world of image registration and successfully applied to large deformations models (Trouvé, 1998).
3.5.2 LDDMM framework

LDDMM (Beg et al., 2005) is one of the most well established and popular framework to parameterize the diffeomorphism in image registration. Using the algebraic and differential geometric structure of diffeomorphisms, LDDMM defines a metric that provides a well-defined notion of amount of geometric change between transformations.

In the framework, a diffeomorphism group is considered where each element \( \phi \) transformation generated by flows of smooth, time-varying velocity fields with support on \( \Omega \) for a simulated time parameter \( t \in [0, 1] \).

\[
\begin{aligned}
\frac{d}{dt}\phi_t(x) &= v_t(\phi_t(x)), \quad t \in [0, 1] \\
\phi_0(x) &= x, \\
\phi_t(x) &= x, \quad \text{for } x \in \partial \Omega
\end{aligned}
\]  

(3.5.2)

Diffeomorphic mappings between parameterized time points \( \{t_1, t_2\} \subset [0, 1] \) are obtained from Equation 3.5.2

\[
\phi_{t_2}(x) = \phi_{t_1}(x) + \int_{t_1}^{t_2} v_t(\phi_t(x))dt
\]  

(3.5.3)

By the classic theory of ODE, with certain smoothness of the velocity, we can prove the existence and the uniqueness of the solution (Dupuis et al., 1998). The introduction of the velocity fields enables large-deformation transformations to be produced while maintaining the diffeomorphism property.

For the regularization term \( E_R \), it is more convenient to define the energy in terms of the velocity \( v \) rather than on the final transformation \( \phi \). At any time
$t \in [0, 1]$, $v(\cdot, t)$ is assumed to belong to a Hilbert space $V$ with associated inner product $\langle \cdot, \cdot \rangle_V$. For $u, v \in V$, the inner product is defined using linear differential operator $L$.

$$\langle u, v \rangle_V = \langle Lu, Lv \rangle_{L^2} = \int_\Omega (Lu(x)) L^\dagger Lv(x)dx$$ \hspace{1cm} (3.5.4)

This inner product on velocity fields induces the norm

$$\|v\|_V = \sqrt{\langle v, v \rangle_V}$$ \hspace{1cm} (3.5.5)

The operator $L$ is often chosen to be a differential operator, e.g. Cauchy-Navier in \cite{Beg et al., 2005}

$$L = \alpha \Delta + \beta (\nabla \cdot ) \nabla + \gamma$$ \hspace{1cm} (3.5.6)

Let $A = L^\dagger L : V \mapsto V^*$, where $V^*$ is the dual space of $V$. Then for any $v \in V$, $Av$ is a linear functional on $V$. Interpreting $\|v\|_V^2 = (Av, v)$ as an energy, $Av$ can be thought of as the momentum associated to the velocity $v$. We assume that $V$ can be embedded in a space of smooth functions (i.e. Sobolev space), which defines a reproducing kernel Hilbert space with kernel $K = A^{-1} : V^* \mapsto V$.

$$(m, v) = \langle Km, v \rangle_V$$ \hspace{1cm} (3.5.7)

where $m$ is in the dual space $V^*$ and the bracket $(\cdot, \cdot)$ is in the standard notation for a linear functional applied to $v$. In particular, the use of smooth kernels generates matching equations that are always well-posed, and with computationally attainable solutions.
As a result, the total regularization energy can be defined by

$$E_R = \int_0^1 \|v_t\|^2 dt$$  \hspace{1cm} (3.5.8)

LDDMM framework described above essentially defines an image-to-image metric that quantitatively measures the geometric change between shapes. Intuitively, the distance between two images is given by the amount of deformation required for one image to match another. This is the foundation for many quantitative statistical study in the shape space for medical images.

The standard LDDMM setup relies on the parameterization of the entire time-variant velocity field. Inspite of its rigorous formulation, the method may suffer from high computational complexity due to the need to store the velocity field at each time step.

To address the memory challenge in LDDMM, a variety of methods have been developed to simplify the parameterization of the velocity field. (Ashburner, 2007; Vercauteren et al., 2008) proposed algorithms which employ a time-independent velocity field $v$ to generate the diffeomorphism subgroup $\Phi^v = \exp(t \cdot v)$ and the final transformation $\phi$ can be obtained by getting the This essentially defines a infinite-dimensional Lie group structure with Lie algebra being the linear space of all velocity fields.

These methods have two major advantages: 1. the introduction of a stationary velocity field significantly reduces the memory space; 2. a single velocity field also greatly simplifies the computation due to employment of the “divide and square”
algorithm (Ashburner, 2007) in the interpolation process.

Despite of the advantages of the stationary field parameterization, it has the limitation that only a small subset of the original diffeomorphism group can be generated this way. In addition, the velocity field is not necessarily the optimal solution of the original LDDMM minimization problem.

Another more rigorous formulation has been proposed in (Miller et al., 2006), which has shown that the optimal solution of the LDDMM velocity field satisfies the law of “conservation of momentum”. As a result, the final deformation field can be uniquely determined by the initial velocity (or momentum) following the Euler Poincare’s equation (Holm et al., 1998). In this formulation, the integration path of the velocity defines a geodesic path between deformations at different time points. That’s why the method is called geodesic shooting method and it will be detailed in section 3.5.3.

### 3.5.3 Geodesic shooting

In the standard LDDMM framework, the matching energy $E_A$ depends only on the final transformation $\phi$. This implies that the velocity field can always be reparameterized to minimize the regularization energy. This functional minimization property has been studied in fluid mechanics known as the Euler-Poincare equation (EPDiff) (Holm et al., 1998), which follows the law of “conservation of momentum”.
Given the initial velocity \( v_0 \in V \), or the initial momentum \( m_0 \in V^* \), the geodesic path can be generated by the following EPDiff equation, which can be interpreted as a conservation equation for the momentum \( m_t = Kv_t \),

\[
\frac{\partial m_t}{\partial t} = -\text{ad}^*_v m_t = -(Dv_t)^* m_t + \text{div}(v_t)m_t + Dm_t v_t
\]  

(3.5.9)

where \( D \) denotes the Jacobian matrix and \( \text{ad}^*_v \) is the dual of the Lie bracket of vector fields,

\[
\text{ad}_v u = -[v, u] = Dvu - Duv.
\]  

(3.5.10)

For an initial vector field \( w_0 \), \( w_t = D\phi_t(\phi_t^{-1})w_0(\phi_t^{-1}) \) is the transported vector field along the geodesic path. Then, if \( m_t \) satisfies equation (3.5.9), we have:

\[
\frac{\partial}{\partial t}(m_t, w_t) = 0
\]  

(3.5.11)

This implies that \( m_t = |D\phi_t^{-1}|(D\phi_t^{-1})^* m_0(\phi_t^{-1}), t \in [0, 1] \), which means that the flow at any point along the geodesic is completely determined (once a template is fixed) by the momentum at the origin through the geodesic equations. The initial momentum therefore provides a linear representation of the nonlinear diffeomorphic shape space in a local chart around the template, to which linear statistical analysis can be applied.

Such a point of view links geodesic formulation to a conservation of momentum law in Lagrangian coordinates and provides a powerful method for studying and modeling diffeomorphic evolution of shape. One of the most beautiful aspects of studying diffeomorphisms with this point of view is that many fundamental aspects
which can be proved in the finite dimensional case can be formally extended to retrieve well-known equations of mechanics. The quantitative deformation description allows us to focus the modeling effort on the diffeomorphism group rather than on the family of objects being deformed, and can easily be related to the problem of measuring the variability between different subjects.

3.6 Optimization and Variational Framework

Once the parameterization and the objective is set up, imageregistration is an optimization problem. In equation (3.2.1), we are trying to find the optimal solution

$$\phi^* = \arg\min_{\phi \in S} E(\phi)$$  \hspace{1cm} (3.6.1)

Assuming that the energy is sufficiently regular,

its first order variation at $\phi \in S$ in the direction of $h \in S$ is given by

$$\partial_h E(\phi) = \lim_{\varepsilon \to 0} \frac{E(\phi + \varepsilon h) - E(\phi)}{\varepsilon} = \lim_{\varepsilon \to 0} \frac{dE(\phi + \varepsilon h)}{d\varepsilon} \bigg|_{\varepsilon = 0}$$  \hspace{1cm} (3.6.2)

This is the definition of Gateaux derivative. If minimizer $\phi^*$ exists, then $\partial_h E(\phi) = 0$ must hold for every $h$.

This is called the Euler-Lagrange equations of the energy $E(\phi)$. Generally, the set of all feasible transforms forms a dense subset of a Hilbert space $H$, associated with the inner product $\langle , \rangle_H$. If mapping $h \to \partial_h E(\phi)$ is linear and continuous, the Riesz representation theorem guarantees the existene of the unique vector $\nabla_H E$ of
\( \nabla_H E \), called the gradient of \( E(\phi) \), which satisfies the equality

\[
\langle \nabla_H E, h \rangle_H = \partial_h E(\phi), \forall h \in H
\]

(3.6.3)

In the context of deformable transformation, it is not convenient to solve this equation directly. The solution is often obtained by iteratively updating the transformation field in the direction of the negative gradient of \( E \). The optimization starts from an initial deformation field \( \phi_0 \) and update the field in the direction of Sobolev gradient \( \nabla_H E(\phi) \).

\[
\begin{cases}
\frac{d\phi_s}{ds} = -\nabla_H E(\phi) \\
\phi_0 = id
\end{cases}
\]

(3.6.4)

Examples of the MI, CC are given in the appendix and the more complete derivation of different measures can be found in Hermosillo et al. (2002).

In the LDDMM framework, since the transformations are parameterized by the velocity, the derivative of the energy is given with respect to the velocity. The velocity is in the space \( V \times [0,1] \) which is in most cases a Sobolev space whose norm is define by the operator \( L \).

\( V \) is the space of function \( \Omega \to \mathbb{R} \) where and \( \mathbf{v}(\cdot, t) \in V, \forall t \in [0,1] \). The variation of mapping \( \phi^{\mathbf{v}}_{s,t} \) when \( \mathbf{v} \in L^2([0,1], V) \) is perturbed along \( h \in L^2([0,1], V) \) is given by (All details of the derivation can be found in the appendix):

\[
\partial_h \phi^{\mathbf{v}}_{s,t} = \lim_{\varepsilon \to 0} \frac{\phi^{\mathbf{v}+\varepsilon h}_{s,t} - \phi^{\mathbf{v}}_{s,t}}{\varepsilon} = D\phi^{\mathbf{v}}_{s,t} \int_s^t (D\phi^{\mathbf{v}}_{s,u})^{-1} h_u \circ (\phi^{\mathbf{v}}_{s,u}) du.
\]

(3.6.5)
Given what the derivative of the deformation field with respect to the velocity is
\[ D \phi^v_{s,t} \int_s^t (D \phi^v_{s,u})^{-1} h_u \circ (\phi^v_{s,u}) du \]

The gradient of the energy in the velocity is
\[ \nabla_v E(v) = (D \phi^v_{t \to 1})^\dagger |D \phi^v_{0 \to t}| \nabla \phi^y_{t \to 0} \quad (3.6.6) \]

### 3.6.1 Numerical Schemes

Given the variational gradient, the numerical calculation of the updates in the velocity field in LDDMM lies mostly in the spatial differentiation and the integration of the velocity field.

#### Discretization and the differential operator

In the calculation of the LDDMM framework, the discrete version of the differential operator is required for the differentiation of the velocity field. This is often implemented by the finite difference method (For example, the Laplacian in two dimensions can be approximated using the five-point stencil finite difference method). This computation can also be efficiently achieved in the fourier domain with the assumption of the boundary conditions (e.g. periodic boundaries).

#### Integration of the velocity field

The integration of the velocity field is one of the most time consuming steps in the whole computation. Semi-Lagrangian schemes (Staniforth and Côté, 1991) are adopted in the standard calculation of LDDMM algorithm. The schemes combine
the advantages of the pure Lagrangian schemes and Eulerian schemes, so it is both efficient and stable.

### 3.6.2 Approximation for LDDMM

In the original LDDMM method, the velocity field is discretized into \( N \) time steps, so it requires to store \( N \) velocity fields. In each iteration, it is required to compute also \( N \) gradient fields, \( N \) compositions for and solve \( N \) inverse problems. This makes it a very computational costly algorithm both in terms of memory space and time.

**Greedy Algorithm and Demons**

One family of greedy approximations of the LDDMM framework are the demon’s algorithms (Thirion, 1998).

In the demons algorithm, when the iterative optimization procedure uses iterations, the update \( \phi_i \), from each iteration can be viewed as the velocity fields. At iteration \( i \), the greedy optimization fixes all the velocity fields obtained before \( i \) and only optimize \( v_i \).

The advantage of such a greedy way is that this does not require to store all the velocity fields before \( i \) and only their composition is needed. The disadvantage is that such a greedy method does not put any constraint on the overall quality of \( \phi_i \). Each \( \phi_i \) is a local optimal but the entire transform path over time may be curvy.
Chapter 4

Volumetric image reconstruction and application to 2D histology slices

4.1 Introduction

4.1.1 Image Reconstruction

Image reconstruction is a general process of estimating an image object from a transformed (degraded) version of this object (Demoment, 1989). Because of its practical importance and theoretical interest, this problem has been extensively studied in many areas, such as optics, X-ray or diffraction tomography, radioastronomy, machine vision, geophysics, etc. The mathematical forms of the transfor-
The main idea of image reconstruction is that it is generally an ill-posed problem and the key to modeling lies in the regularization on the original object we are trying to reconstruct. Given a meaningful form of regularization, image reconstruction is essentially transformed to an optimization problem with the balance between the regularity of the reconstructed object and the fidelity to the observation.

In this paper, we are mainly interested in one specific type of image reconstruction problem whose goal is to restore a volumetric image from a sequence of randomly transformed (or even corrupted) histology images. The model we are looking for should not only follow the general form of the image reconstruction but also be suitable for this specific problem. In the following sections, we cover the motivation for analyzing this topic and the difficulties that arise in it.

### 4.1.2 Motivation for histology image reconstruction

Despite the fast development and wide application of various 3D imaging techniques (e.g. MRI, CT, etc) over the past years, 2D imaging methods (such as optical microscopy) still excel in generating images with much higher spatial resolution and contrast and thus provide more detailed information about tissues. The enhanced ability to visualize microscopic structures facilitates the anatomical analysis of the
tissue at cellular or even subcellular level, which is essential in many biomedical studies like identification of pathological samples.

Figure 4.1: histology image sample: high resolution of the image enables cellular level of anatomical analysis

Histological imaging methods are a broad class of imaging methods, which typically generate serial sections from a 3D volume to be studied under the microscope. It generates images with very high resolution (an example in fig 4.1). However, the preparation process often results in the loss of three dimensional geometric continuity across slices, making it difficult to directly apply volumetric analysis. Therefore it is desirable to transform the histological slices into a continuous 3D shape so that they can be viewed in a common coordinate system. This enables researchers to study the complex anatomy of the organs or structures of interest in three dimensions.

Due to the nature and high incidence of distortions and artifacts that occur during the imaging process, the reconstruction from 2D serial slices is extremely challenging in practice. The finally digitized images may suffer from local distortion, contraction, tearing or even missing parts. This necessitates an efficient automatic
volumetric reconstruction method, which should be unbiased with respect to input
data and robust to complex situations such as corrupted data and non-linear local
deformations.

The past efforts to do volumetric histology image reconstruction have been
mainly on the development of accurate inter-slice registration schemes and cross-
slice smoothing/regularization methods to alleviate the overall error accumulation
(Malandain et al., 2004). Image registration theory has a strong connection to the
reconstruction problem due to their shared interest in exploring the similarity met-
rmetric and geometric properties among images, so the advance in the field of image
registration has been inspiration for novel image reconstruction methods. On the
other hand, the coherence of the reconstructed image volume and regularization is
more specific for the image reconstruction problem in which spatial properties and
structure should be considered.

In this dissertation, we are trying to develop a novel volumetric image reconstruc-
tion framework to robustly rebuild the 3D image. The method should be general,
rigorous and robust to complex cases and balance well between the intensity and
underlying geometry difference slices.
4.2 Background and Literature Review

4.2.1 Histological Image Data

In medical imaging, histology is the study of the microscopic anatomy of cells and tissues. To obtain histological sections, the anatomical structure is first fixed using chemical fixation or frozen section. Then it is sectioned into thin sections with a constant intersection gap. Thereafter, the stained sections can be digitized by a light microscope or a high resolution transparent flat-bed scanner. The typical resolution of the image can go up to 0.2 µm in the light microscope and the thickness up to 5 - 10 µm, which offers great details of the anatomy.

The acquisition is often performed separately for each single slice, which leads to a series of consecutive and unaligned images as we can see in the figure reffig.. The imaging result thus differs from volumetric image data generated by 3D imaging methods like MRI and CT which is far less invasive and preserves the continuity. To sum up, the images generally suffer in the following aspects (Malandain et al., 2004; Simmons and Swanson, 2009):

1. Global misorientation e.g. rotation, displacement.

2. Local tissue distortion

3. Image artifacts
4.2.2 Literature Review

3D reconstruction via 2D registration

In order to find right transforms for each 2D section, one common approach (e.g. in (Ourselin et al., 2001; Guest et al., 2001)) starts from pairwise registration on consecutive slices. Then a 3D volume is constructed by concatenating the resulting transforms and warping each slice with respect to a preselected reference slice. Thanks to the fast development of the image registration field, a large variety of registration models have been applied to 2D histological alignment. In these models, different similarity metrics have been employed to guide the image reconstruction, including both feature-based (Guest et al., 2001; Kim et al., 1997) and intensity-based models (Ourselin et al., 2001; Kim et al., 1995).

In spite of broad choice of image registration schemes, interslice registration is still error-prone due to high incidence of artifacts in the sectioning process (e.g. shrinking, tearing, shearing, folding, and displacement). Therefore, direct concatenation of transformations is more likely to lead to an accumulation and propagation of the alignment error along the slice stacking dimension. Moreover, in the absence of the volume references, the reconstruction schemes may lack control on the overall shape in 3D and give a biased result. As argued at length in Malandain et al. (2004), the 3D structure is lost during sectioning and the direct stacking may result in an error referred to as z-shift (Yushkevich et al., 2006) or “banana problem” (Malandain et al., 2004).
A number of methods have been proposed to address or at least mitigate this problem. Some resort to the help of an external reference and some trying to add more coherence across a greater range of slices.

**3D reconstruction with external references**

The quality of the 3D volume reconstructed by the 2D image interslice registration may be easily improved when external information of the 3D image is available, which may come in the form such as a 3D imaging modality and help reintroduce the global information lost in sectioning.

Dauguet et al. (2007) built an initial block-face volume from block-face photographs taken during the cutting process and used it as external reference for the alignment of the histological slices. They attempt to correct deformations in the reverse order of their physical occurrence.

Adler et al. (2012) took a 3D in vivo image such as an MRI of the same tissue block before the histology imaging process and used it as the representation of ground truth volume shape. The 3D MR volume is then transformed into the histology space and a geometrically coherent global reconstruction of the histology image volume is created by registering the histological with the corresponding reference 2D slices.

Feuerstein et al. (2011) simultaneously registered the histological slices to their neighboring slices and the corresponding block-face image using a Markov random
field model. They assume this can maintain the structural homogeneity and correct the drift at the same time.

**reconstruction by cross-slice regularization**

However, reference volume or adequate prior shape information are often unavailable. In these cases, cross-slice smoothness methods have been proposed in the literature to alleviate the issues raised by the limitations of the 2D pairwise reconstruction.

*Yushkevich et al. (2006)* proposed a graph theoretic technique where each slice was rigidly registered to several neighboring slices instead of just the immediately preceding or following one. This defines a registration graph whose weights are a measure of the quality of those individual registrations. The volume can then be reconstructed by following optimal local paths in the graph.

*Ju et al. (2006)* proposed a “warp filtering” technique to achieve the overall smoothness results in the reconstructed volume. The method effectively warps each point on one section to the weighted-average location of the corresponding points on neighboring sections.

*Cifor et al. (2011)* extracted the region of the tissue of interest and then registered the contour of the region by minimizing the minflow of the volume across a range of slices.

In a global approach, *Wirtz et al. (2004)* and *Guest and Baldock (1995)* aim at
minimizing an energy functional consisting of a distance measure between consecutive sections (e.g. spring forces between corresponding points (Guest and Baldock, 1995) or pixel-wise squared differences (Wirtz et al., 2004)) and an elastic deformation potential on each section.

4.3 Method

4.3.1 Method Overview

Here we introduce a novel automatic approach for reconstructing a smooth volume by aligning serial sections, which is built on a unified framework and capable of dealing the cases with or without the external reference.

The heuristic method using pairwise registration and concatenation method only implicitly minimizes the sum of pairwise differences and certain (if any) penalty inherited from registration framework, but it is often more appealing to explicitly consider the reconstruction as a functional minimization problem. The explicit objective function and feasible domain not only give more accurate description of the problem, but also enable more mathematical and rigorous analysis.

Our general framework is equipped with a rigorous energy description of the smoothness of the reconstructed volume. This model differs from previous work (Ourselin et al., 2001; Ju et al., 2006) by explicitly including the shape difference between constructed slice pairs, which is defined using the geodesic distance on
transformation groups Miller and Younes (2001). Moreover, unlike other works (e.g. Ourselin et al. (2001)), which only measure the image difference between immediate image pairs, our method is more robust to corrupted slices by including all pairs in a neighborhood range with adapted weights. These weights are defined by both spatial kernels and an intrinsic image metric. In addition, the formulation is symmetric as no preselected reference slice is required.

To sum up, we present in this work

• An automatic reconstruction method which can be applicable to both cases with/without the 3D reference.

• rigorous mathematical framework of volumetric image reconstruction with the description of the overall smoothness of the reconstructed volume.

• symmetric regularization of the reconstructed volume which is robust to corrupted slices.

• a novel, efficient functional optimization method.

4.3.2 Formulation of general image reconstruction

In this section, we describe the formulation for general image reconstruction. We denote the original object as $I_{\text{org}}$ and the observed data as $I_{\text{obs}}$. $F$ is the filter which defines how the data is degraded before observations. In most cases, $F$ consist of both structural transformation of the original object and measurement error in the
observation.

\[ \mathcal{F}(I_{org}) = I_{obv} \quad (4.3.1) \]

The original object cannot be directly measured and must be reconstructed from the observed data. A reasonable estimation generally has two parts:

- a **data fidelity** term which defines the difference of the observed and the transformed original image.

- a **regularization** term which describes the coherence/regularity of the reconstructed image. Image reconstruction problems are generally ill-posed, because of the high dimensionality of unknowns and the instability of the inverse problem when noise is involved.

As a result, the image reconstruction problem can generally be formulated as an optimization of the energy

\[ \min_{\hat{I}} E_{\text{total}} = E_{\text{fid}}(I_{obv}, \hat{F}\hat{I}) + E_{\text{reg}}(\hat{I}) \quad (4.3.2) \]

Note that the filter \( \hat{F} \) in the formula may be unknown and estimation of it may be needed.

### 4.3.3 Formulation for volumetric image reconstruction from serial 2D images

Define \( \Omega \) as the continuous volumetric domain embedding the anatomy of interest. The ground truth of the 3D smooth volume is a function \( \mathcal{I} : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R} \) or
$\mathbb{R}^d$ (color image) where $d$ is the number of color channels. A 2D image sequence \( \{I_i\} \) at position \( \{z_i\} \) along $z$-axis, each of which has undergone some transform $\varphi_i$, \( \mathcal{I}(\cdot,z_i) = I_i(\varphi_i(\cdot)) \) And the goal is to find the $\varphi_i$ so that we can reconstruct the smooth volume.

**Data fidelity term**

The data fidelity term defines the difference of the observation and the transformed original image. In this dissertation we are mainly focused on the geometric shape change of the images in the sampling process and consider that as the only source of the error between the observation and the image to be reconstructed.

In order to quantitatively measure the difference between images in terms of the geometric shape change, we apply the ideas from image registration which defines distance between shapes by the coordinate transforms. For instance, in the LDDMM framework, the distance is defined by

\[
    D_T(\varphi) = \int_0^1 \|v_t\|^2_{v_t} dt, \\
    \frac{\partial \varphi_t}{\partial t} = v_t(\varphi_t), \quad \varphi_0 = Id, \varphi_1 = \varphi. 
\]

(4.3.3)

where the $v_t$ is a simulated velocity which generates the deformation as a flow of a time dependent vector field.

In this way the data fidelity term in the total energy is

\[
    E_{fid} = \sum_{i=1}^{N} D_T(\varphi_i), \quad \text{where } \mathcal{I}(\cdot,z_i) = I_i(\varphi_i(\cdot)) \quad (4.3.4)
\]
### 4.3.4 Regularization

**Total Variation-like Regularization**

One of the easiest ways of enforcing the regularity of a 1D signal is the total variation. Recall the definition of total variation (TV) of data \( y: [a, b] \subseteq \mathbb{R} \to \mathbb{R} \):

\[
TV(y) = \sup_{a = z_1 < z_2 < \ldots < z_N = b} \sum_{i=1}^{N-1} |y(z_i) - y(z_{i+1})|
\]

where \( z_1, \ldots, z_N \) could be any finite sequence of time points in the interval \( [a, b] \) to achieve the supremum.

With the similar idea, we can set the TV regularization in the \( z \) direction for the image \( I \). If the two end points of the region \( \Omega \) are \( t = a, t = b \), then

\[
TV(I) = \sup_{a = z_1 < z_2 < \ldots < z_N = b} \sum_{i=1}^{N-1} \text{Dist}(I(z_i), I(z_{i+1}))
\]

In our case only \( N \) sample observations are available, so the sum of the difference is finite.

Here we need define the difference between images, which follows the total energy in image registration.

\[
\text{Dist}(I, J) = \inf_{\psi \in \Phi} D_I(I, J \circ \psi) + D_G(\psi)
\]

where \( \psi \) is a transformation and \( D_I(I, J) \) is an image metric which can be any popular image metrics such as CC and MI (3.3) and \( D_G(\psi) \) is the regularization which describes the distance of a transformation. Note that this definition is not generally a distance in the strict mathematical sense because it’s not symmetric for
the two image entries. In practice, we can make it symmetric by using a symmetric formulation like in (Avants et al., 2008) or take both the forward and backward formulation. Also different from the general setup in the image registration, the “distance of transformation” can also be defined on the linear transformations.

This definition of image difference takes both the image intensity $D_I$ and image shape $D_G$ difference into account, which we believe is a more accurate way of modelling the reconstructed image volume. In this way, the total energy in the image reconstruction can be written as the sum of the fidelity term and the regularization term.

**Regularization with greater range**

The total variation of the reconstructed volume is calculated by taking the difference of the immediately neighboring sections (like the most basic finite difference scheme for approximating derivatives), which in practice is not robust to errors in the image sequence. The direct result may suffer more from error accumulation in cases where badly torn sections are present.

In order to address the challenge and make the optimization problem more robust to real noise, we apply a regularization dependent on a larger range of image sections. This is similar to nonparametric image regression (Davis et al., 2010) in the sense that the average information from different time points prove to be a more meaningful estimate. The difference is that the goal in nonparametric regression is
to estimate a sequence of fitted images \( \{ \hat{I}_i \} \) which average over the intensity while our case is to find the best transformed images \( I_i \circ \varphi_i \).

In order to better denote the energy of the shape difference in the reconstructed 3D volume, we introduce transformations \( \{ T_{i,j} \} \) for the shape change between the transformed slices \( I_i \circ \varphi_i \) and \( I_j \circ \varphi_j \).

\[
E_{reg}(\{ I_i \circ \varphi_i \}) = \sum_i \sum_{j \in \mathcal{N}(i)} \alpha_{i,j} \left[ \text{Dist}(I_i \circ \varphi_i, I_j \circ \varphi_j) \right] 
= \min_{T_{i,j}} \sum_i \sum_{j \in \mathcal{N}(i)} \alpha_{i,j} \left[ D_I(I_i \circ \varphi_i, I_j \circ \varphi_j \circ T_{i,j}) + D_g(T_{i,j}) \right]
\tag{4.3.8}
\]

Note that the energy is only a function of \( \{ \varphi_i \} \), while \( T_{i,j} \) are latent transforms of which the energy function take a minimum over all possible choices.

\( \mathcal{N}(i) \) is the neighborhood of the \( i \)th slice. \( \alpha_{i,j} \) are the weights of the slice pair \( i, j \) in the total sum. Those weights have lower values on the corrupted slices, which will be detailed in section 4.3.4. In this way, the regularization energy in (4.3.8) is trying to make \( I_i \circ \varphi_i \) close to the weighted intrinsic mean of the transformed images \( I_j \circ \varphi_j, j \in \mathcal{N}(i) \). This is illustrated in Fig 4.2.

\[
E_{total}(\{ \varphi_i \}) = \min_{T_{i,j}} \sum_i \sum_{j \in \mathcal{N}(i)} \alpha_{i,j} \left[ D_I(I_i \circ \varphi_i, I_j \circ \varphi_j \circ T_{i,j}) + D_g(T_{i,j}) \right] + \beta_i \sum_i D_T(\varphi_i).
\tag{4.3.9}
\]

**Relationship to the previous work**

Note that in the special case when \( \mathcal{N}(i) = \{ i-1, i+1 \} \) (except \( i \) on the boundary) and transforms are affine, we can prove the functional minimization is equivalent
Figure 4.2: $\mathcal{I}$ is the true continuous 3-D image. $I_i \circ \varphi_i$ and $I_j \circ \varphi_j$ are transformed images. $T_{i,j}$ is the shape difference between $I_i \circ \varphi_i$ and $I_j \circ \varphi_j$

to finding sequentially the transforms to minimize image difference between consecutive slice pairs and concatenate the resulting transforms. In this case $T_{i,j}$ will automatically go to 0 and the method reduces to known reconstruction methods that Ourselin et al. (2000); Malandain et al. (2004) adopt.

**Data-driven method for the weight**

In some cases, some 2D slices are highly distorted during scanning. They are unlikely to register to their neighbors, which may result in gross discontinuities in the reconstructed anatomy. In our model, this is dealt with by choosing the right weights $\alpha_{i,j}$. For slice $I_i$ and $I_j$, $\alpha_{i,j}$ depends on the slice position distance $|i - j|$, for which a common choice is Gaussian $\exp(-|i - j|^2/2\sigma^2)$ (constants omitted).

To accommodate the corrupted images, we precalculate an essential image similarity $CC(I_i, I_j \circ \psi_{i,j}), \forall j \in \mathcal{N}(i)$ between slices after pairwise rigid image registration $\psi_{i,j}$. 57
The weight is defined as,
\[ \alpha_{i,j} = \frac{\exp(-|i - j|^2/2\sigma^2)}{1 - CC(I_i, I_j \circ \psi_{i,j})} \] (4.3.10)

The weights are symmetric \( \alpha_{i,j} = \alpha_{j,i} \) and the method is thus independent to the direction of sequence \( \{z_i\} \).

4.3.5 Rigid transforms and its optimization

The transforms \( \{\varphi_i\}, \{T_{i,j}\} \) in our model can be linear transforms or nonparametric deformation fields. The linear model is often used to correct global misalignment of image slices and the deformable model is used to deal with local inhomogeneous stretching and distortion.

In the application to histological image reconstruction, the global misalignment is mainly due to translation and rotation of the images in the sampling process, and thus we are mainly focused on a simple model of rigid transformation.

For rigid transforms, we don’t have fidelity terms for image transformation \( \varphi_i \). However, we still takes into account the transformation difference between the slices. the distance \( D_G(T_{i,j}) \) for rigid transforms is defined as
\[ D_G(T_{i,j}) = |\theta^{T_{i,j}}|^2 + \lambda \|t^{T_{i,j}}\|^2 \] (4.3.11)
where \( \theta^{T_{i,j}} \) is the rotation angle and \( t^{T_{i,j}} \) is the translation (assuming transformation center is given) of \( T_{i,j} \).

The standard way of optimizing total energy of (4.3.9) is gradient descent with respect to the parameters defined by rotation and translation. The update of the
transformation can be applied to the images at every step and then the new gradient can be calculated.

Here we also propose an alternative greedy approximation of the energy minimization that first minimizes the term of $D_I$ by finding the best relation between $\varphi_i$ and $\varphi_j \circ T_{i,j}$. Because common choices of image intensity similarity are all right-invariant under rigid transforms, it is clear that $D_I(I_i \circ \varphi_i, I_j \circ \varphi_j \circ T_{i,j}) = D_I(I_i, I_j \circ \varphi_j \circ T_{i,j} \circ \varphi_i^{-1})$. Therefore we introduce

$$\psi_{i,j} = \varphi_j \circ T_{i,j} \circ \varphi_i^{-1} \tag{4.3.12}$$

and find the best $\psi_{i,j}$ to minimize $D_I(I_i, I_j \circ \psi_{i,j})$ using image registration. Then given a fixed $\psi_{i,j}$, we furthermore optimize $D_G(T_{i,j}) = D_G(\varphi_j^{-1} \circ \psi_{i,j} \circ \varphi_i)$ to get the optimal $\{\varphi_i\}$. This can actually take advantage of any current rigid image registration schemes.

In the greedy setup, the rotations are minimized before translations. Assuming all transforms $\{\varphi_i\}$ share a same center, one can show that the translation and rotation energy are independent and can be minimized. Moreover, the composition of transforms $\psi_{i,j}$ and $\varphi_i$ becomes the addition or linear transformation of variables $\theta^{\varphi_i}, t^{\varphi_i}$ which makes the term $D_G$ a quadratic form and thus makes minimization a linear fitting problem. For the rotation part, we need to minimize

$$\theta^{\varphi_i} = \arg \min \sum_{i,j} \alpha_{i,j} | - \theta^{\varphi_j} + \theta^{\psi_{i,j}} + \theta^{\varphi_i}|^2 \text{ given } \theta^{\psi_{i,j}} \text{ and } \alpha_{i,j} \tag{4.3.13}$$
After the rotations are applied, the translation part has a similar form and $t^{\varphi_i}$ can be solved for in the same fashion as $\theta^{\varphi_i}$.

### 4.3.6 Deformable transforms and its optimization

Linear transformation models cannot be used to deal with local stretching or distortion of images. In order to study the detailed and irregular geometrical changes of the underlying anatomy, we apply the diffeomorphic group (Miller and Younes, 2001) in which a time dependent transform $\varphi(x)$ is dense in space and parametrized by a velocity field $v(x, t)$ for $t \in [0, 1]$. Since the right-invariance of the image term does not hold anymore in the case of general diffeomorphic transforms, we develop an iterative greedy scheme to minimize the energy in (4.3.9).

The process of finding the best transformation in the reconstruction is similar to the idea of diffeomorphism template building (Joshi et al., 2004) for each slice’s neighborhood, where the template by definition is a reconstructed slice which is closest to other images in the neighborhood in terms of image distance. The major difference between our volumeric reconstruction and template building is that each object in the image group (within the neighborhood) in our case is not fixed in the reconstruction process.

As a result, this motivates an iterative approach which involves two major steps in the optimization process like the “map and reduce” fashion. In the “map” step, we are trying to find the correspondence between $i$th slice and each slice in its
neighborhood $\mathcal{N}(i)$. In this way, we can obtain an update on the correspondence between the transformed image $I_i \circ \varphi_i$ to its transformed neighbors $I_j \circ \varphi_j, j \in \mathcal{N}(i)$. In the “reduce” step, the update in the correspondence transformation $T_{i,j}$ is collected and used to update the current $\varphi_i$.

The alternating computation between the registration and the combination of image transformation effectively decreases the total energy $E_{total}$ in each step. In addition we don’t need to finish the final registration process before combining transformation.

The details of the method is described as following:

For the initial step all $\{\varphi_i\}$ are set equal to the identity. For each iteration step $k$, we optimize $D_I$ and $D_G$ alternatively:

1. Register each slice $i$ to each one of its neighbors in $\mathcal{N}(i)$ where the registration can be done by the shooting method 3.5.3. Compute the gradient of image similarity $D_I(I_i \circ \varphi_i^{(k)}, I_j \circ \varphi_j^{(k)} \circ T_{i,j})$ with respect to $T_{i,j}$:

$$v_{T_{i,j}}|_{t=0} = \frac{d}{dT_{i,j}} D_I(I_i \circ \varphi_i^{(k)}, I_j \circ \varphi_j^{(k)} \circ T_{i,j}), \forall j \in \mathcal{N}(i).$$

2. Update $\varphi_i^{(k)}$ to minimize $\sum_{j \in \mathcal{N}(i)} \alpha_{i,j} D_G(T_{i,j}^{(k)}) = \sum_{j \in \mathcal{N}(i)} \alpha_{i,j} \int_0^1 \|v_{T_{i,j}}\|^2 V dt$ while keeping $D_I$ by fixing all the composition $T_{i,j} \circ (\varphi_i^{(k)})^{-1}$. This is done by optimizing over $T_{i,j}$ and $\varphi_i$ to minimize $\sum_{j \in \mathcal{N}(i)} \alpha_{i,j} D_G(T_{i,j})$ while keeping $T_{i,j} \circ \varphi_i^{-1}, \forall j \in \mathcal{N}(i)$ fixed from step 1. This leads to the velocity $v_{\varphi_i}$’s update at time $t = 0$

$$v_{\varphi_i} = \arg\min_v \alpha_{i,j} \|v_{T_{i,j}} - v\|^2_V$$
\[ \varphi_i^{(k+1)} = \varphi_i^{(k)} \circ (\Delta t \cdot \mathbf{v}^{(k)}), \]

where \( \Delta t \) is the step size in the update. This process is iterated until convergence.

In Fig 4.3, we give the illustration of how the deformations are combined given the velocity at the end.

**Figure 4.3:** This update of the deformations given the correspondence between images

### 4.3.7 With External Reference

In some cases, an external image reference is available which may be obtained by invivo 3D imaging techniques (Malandain et al., 2004; Adler et al., 2012) or block-face image (Feuerstein et al., 2011; Ourselin et al., 2001). In addition, there are
publicly available atlases such as Waxholm space for mouse brain (Johnson et al., 2010) which can be applied to specific types of data. 3D image references provide valuable overall shape information for the reconstructed image which can be used as a prior in the formulation of image reconstruction. In this way, the image reconstruction result can generally be improved in terms of the coherence across sections and thus recover from the “banana effect” or “z-shift” (Malandain et al., 2004).

In our framework, integrating the external reference information is accomplished by simply adding an extra regularization term on the reconstructed image to the total energy.

$$E_{ref} = \sum_{i}^{N} \text{Dist}(I_i \circ \varphi_i, J_i)$$  (4.3.14)

where $J_i$ is the $i$th corresponding section in the reference volume, which needs to be determined in prior.

In order to make use of the reference volume, we first find the correct correspondence between the histology slices and slices in the MRI volume. This is accomplished by 3D affine image registration which account for global positioning and shape differences between the MRI template and the roughly constructed 3D histology volume. The MRI volume is then transformed and resampled in the space of histology and the histology volume can be corrected. Then each histology slice is registered to its an estimation of its corresponding MRI slice and its neighbors. During the registration, the transformation is done only to the histology slices and not to the MRI slices. This is because there is no tissue loss or distortion in the
MRI slices.

Given the correspondence between the reference slice and the histology slice, we can apply the same process as in the case where no reference is present. The only difference is the registration is not only between the neighbors but also between the histology slice and the reference slice.

4.4 Experiments

In order to quantitatively evaluate our method, we apply the performance measures commonly used for optical flow evaluation by the vision community. In detail, we compute the absolute endpoint error (EE) and the relative angular error (AE) between the ground truth and resulting deformation fields as well as the interpolation error (IE) and normalized IE (NE) between the ground truth and deformed images.

4.4.1 Rigid transforms

First we evaluate our algorithm on the synthetic data generated from the Waxholm histology mouse brain atlas Johnson et al. (2010). The volume is downsampled to 46 slices, each of which has dimension 128x128 pixels and spacing 0.086x0.086mm. We generate random rigid transforms with rotation angles between \([-\pi/6, \pi/6]\) and translations between \([-1/20, 1/20]\) of the image dimension for each slice. 5 slices are manually replaced as corrupted ones into the sequence. In all experiments, ANTS toolkit is used with cross-correlation metric for image registration between the
slices. The result is shown in Fig 4.4, which demonstrates that our method is robust to corrupted slices.

![Image](image1.png)

**Figure 4.4:** Rigidly reconstructing a stack of randomly perturbed histology slices (a) into a smooth rigid volume (b). In (d), when corrupted slices are present, our reconstruction method gives more robust result (left) than that from registration of only consecutive slices (right). (c) shows the angle error for reconstruction with corrupted slices. Our method performs well on middle slices while showing some fluctuations at the ends. This may be because those data are harder to register.

Next we apply our method to the real C57BL/6 mouse brain data with 51 slices each with resolution downsampled to 225x141 and spacing 0.054x0.054mm. The result is shown in Fig.4.5 and the reconstructed volume is smooth visually.
4.4.2 Deformable transforms

A synthesized data of a sequence of ellipses is created to evaluate the deformable method. The 20 concentric ellipses have the same short axis but different long axis generated by a linear function $(18 + 0.5z)$ with uniform random noise $\sigma = 4$. Fig. 4.6 shows that the volume is smoother after deformable registration. While the consecutive pair registration generates result that all slices deformed towards the reference (Fig. 4.6 (e)(f)), our method preserves some overall shape by allowing the shape change between transformed slices.

We apply deformable reconstruction to real mouse brain data. The data has dimension 871x539 and spacing 0.018x0.018mm and the slices are aligned by rigid transforms before deformable reconstruction. The result is shown in Fig. 4.7 and the volume is smoother after the reconstruction.
Figure 4.6: (a) is the coronal view of the stacking of the 20 ellipses before reconstruction and (b) is its 3-D rendering. (c) and (d) are results after proposed reconstruction. (e) and (f) are results after consecutive pairwise registration (with reference in the middle).

Figure 4.7: Smooth diffeomorphic reconstruction applied to real mouse data.

4.5 Concluding remarks

In this work, we have presented a general framework for volume reconstruction from serial image data. A smoothness energy is introduced as a weighted sum of both image intensity and underlying shape difference between image pairs. The results on synthetic and real data indicate that the reconstructed volumes are visually smooth and robust to the corrupted data. The method outperforms the common
consecutive pair registration.

The framework is general in that different choices of transform and similarity metric can be applied. We have discussed the rigid and diffeomorphic cases (with certain distance) and have used cross-correlation metric in experiments. To further constrain the global reconstruction, we could use an external reference to regularize the overall shape when it is available. This could be integrated to the total energy by adding terms of the image distance between the transformed image slices and the corresponding reference slices. Furthermore, this whole framework has natural extensions to higher dimensions like 3D longitudinal study which may require a 4D reconstructions from 3D scans.
Chapter 5

Polyaffine transformation

5.1 Motivation

In the sampling process, histological slices often undergo transformations that are approximately rigid or affine when restricted to local regions (Arsigny et al., 2005). An example of this is illustrated in figure 5.1, which includes different subregions moving in different directions in the sampling process. Each transformation on the subregion is roughly rigid and can be approximated by a different local affine transformation.

As shown in Chapter registration, rigid and affine transformation models have a simple form and a small number of degrees of freedom, but lack the flexibility to model complex local transformations. In order to combine the simplicity of the affine transformations and the power of deformable model, we are motivated
Figure 5.1: Histological image sections example given in (Arsigny et al., 2005)
to develop tools to find meaningful diffeomorphic transformations over the entire domain while preserving the rigidity of the local regions.

5.2 Introduction

The transform models applied in image registration have a wide span of degrees of freedom. The transform can be as simple as an affine transform (Jenkinson and Smith, 2001), which is a linear function defined on the whole image domain and only requires 12 scalar parameters for an image of three dimensions. In contrast, one can also use a deformation field as the transform in a nonrigid image registration, which has an arbitrary number of degrees of freedom at the cost of expensive computation and difficult optimization. Between these extremes of parameterization, many other transforms have been studied, such as B-Splines used in free-form deformation (Rueckert et al., 1999), Geodesic Interpolating Splines (). These transforms are capable of describing a wide range of non-rigid transforms while using fewer parameters than a dense displacement field.

The polyaffine transform is a parameterization for deformable maps that fills the gap between a global affine transform and a deformation field transform. It exploits the prior knowledge that for many applications of medical image registration, the underlying anatomical structure is comprised of multiple local regions. The polyaffine transform is a mathematical framework that fuses local affine transforms through on non-rigid transform.
The key concept in polyaffine is to construct a velocity field over time by fusing multiple regions with different affine velocity to generate an invertible transform. This is often achieved by interpolation of velocity outside the local regions on which each affine transformation is defined. In our work, we reformulate the weight function used to fuse affine velocities by minimizing the energy defined in the large deformation framework. As we show, this also provides a more meaningful and visually superior result.

Furthermore, a time-invariant velocity approximation is proposed in Arsigny et al. (2009). However, as we are going to discuss in Section 5.3.3, it cannot guarantee that the resulting final transform gives the exact same value of the input transforms in each local region.

5.3 Methods

Consider a set of $K$ affine transforms, $T_i(x) = A_i x + b_i$, each defined on a local region $M_i$. We want to integrate these local transforms into one transform $\phi(x)$ defined on the whole image domain $\Omega$. For simplicity, these $K$ local regions are subregions of $\Omega$ and do not overlap with each other. The polyaffine problem can be formulated as finding a feasible solution $\phi(x) : \Omega \mapsto \Omega$, satisfying the constraints

$$\phi(x) = A_i x + b_i, \quad \forall x \in M_i \quad (5.3.1)$$

A direct way to construct such a transformation $\phi$ is by simply averaging each
local affine transformation using a weight function as in \( \phi(x) = \sum^K_{i=1} \omega_i(x)T_i(x) \).

A popular choice of weighting function is defined by radial based functions such as isotropic Gaussian. Then the \( K \) weights \( \omega_i(x) \) computed at the location \( X \) are further normalized such that \( \sum_i \omega_i(x) = 1 \).

This transform is smooth. However, one significant drawback is that it is not invertible in general (Arsigny et al., 2009). To obtain invertibility, i.e. making \( \phi \) a diffeomorphism, the velocity field \( v_t(x) \) was introduced for constructing \( \phi \) in Arsigny et al. (2009).

### 5.3.1 Diffeomorphism in Polyaffine Model

A diffeomorphism \( \phi \) can be obtained by solving the system of ordinary differential equations over a time variable \( t \).

\[
\frac{\partial \phi_t}{\partial t} = v(\phi_t, t). \tag{5.3.2}
\]

At time \( t = 1 \), the diffeomorphism \( \phi_1 \) is obtained by integrating the velocity field \( v_t \). The affine transform \( T(x) = Ax + b \) is an example of such a diffeomorphism.

Using homogeneous coordinates, an affine transform can be generated by a velocity field \( v \) (Arsigny et al., 2009).

\[
v_t = Lx + u, \text{ with } L \begin{bmatrix} \log \left( \begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix} \right) \end{bmatrix}
\tag{5.3.3}
\]

Arsigny et al. (2005) proposed to first construct a velocity field \( v \) by fusing the
$K$ local affine velocity fields:

$$v(x) = \sum_i \omega_i(x),$$

(5.3.4)

where $v_i(x) = L_i x + u_i$, $w_i(x) \propto \exp\left(-\frac{\text{dist}(x, M_i)}{\sigma^2}\right)$.

Introducing the velocity field $v$ as the generator for the transformation guarantees the invertibility of the final transform. This, however, adds a new constraint in finding a feasible solution to equation (5.3.1). Instead of requiring the resulting transform $\phi$ to match each local affine transform $T_i(x)$ within $M_i$ at $t_1$, it requires the velocity $v$ to match each transform at all times from $t = 0$ to 1. Thus we can specialize the problem of equation (5.3.1) into a more restricted one:

$$v(\phi(x, t), t) = v_i(\phi(x, t), t) \text{ for } x \in M_i$$

(5.3.5)

The solution of the new equation (5.3.5) ensures the whole temporal trajectory for all $x$ in $M_i$ matches $T_i$. Indeed, given the uniqueness of the solution to the ODE (5.3.2), we have $\phi(x, t) = T_i(x, t)$ for $x \in M_i$ since $T_i(x) = A_i x + b_i$ is the solution when $v(\phi(x, t), t) = v_i(\phi(x, t), t)$. The trajectory of $M_i$ can thus be tracked over the time $\phi(M_i, t) = \{\phi(x, t) : x \in M_i\}$.

### 5.3.2 Optimal velocity field via energy minimization

A common practice for interpolating the velocity field in the region outside the mask is to use a weight function $\omega_i(x)$ based on the distance between the point $x$ to the region $M_i$. The weight is computed using a Gaussian function or other decreasing functions and are further normalized by scale so that $\sum_i \omega_i = 1$. 

74
However, the choice of the weight function is adhoc and a more rigorous approach can be adopted based on the theory of diffeomorphic registration in 3.5. The desired velocity field over the whole region minimizes the total energy

\[ E_R(v) = \int_0^1 \|v\|^2_V dt \]  

(5.3.6)

where the norm \( \| \cdot \|_V \) is often defined by the differential operator \( L \) such that \( \|v\|^2 = \langle Lv, Lv \rangle_{L^2} \).

Equation (5.3.3) provides a natural way to find velocity field to generate the final affine transformation over the whole space. However in the system of polyaffine transformations, each velocity field is only defined on a subregion, so we need to track the change of the mask regions \( M_i \) over time. The final transformation for the points in the region can be computed by integrating the velocity over time.

Given the trajectory of each mask region \( M_i \) over time, the minimization of the energy in (5.3.6) is confined to the domain of \( \Omega \setminus \bigcup_i \phi(M_i, t) \), which is equivalent to finding the solution to a sequence of Dirichlet problems

\[
\begin{cases}
L^\dagger Lv = 0 & x \in \Omega \setminus \bigcup_i \phi(M_i, t) \\
v|_{\partial \phi(M_i, t)} = v_i, & v|_{\partial \Omega} = 0.
\end{cases}
\]

(5.3.7)

The solution the velocity field over the entire region \( \Omega \) for \( t \in [0, 1] \). The final transformation over the entire field can be generated by integration.
5.3.3 Time-invariant approximation by extending to trajectory of local region

The region $M_i$ and velocity field $v_i$ changes over time, so the minimization of the energy in (5.3.6) requires the solution to equation (5.3.7) at every time step, which proves to be inefficient in computation.

Arsigny et al. (2009) proposed to use a stationary velocity field to approximate the final transformation in Arsigny et al. (2005). This was originally designed to reduce the computation time in the integration of $\phi$ from the velocity field $v$ by an efficient recursive scaling-and-squaring method (Arsigny et al., 2009) in the case when $v$ is stationary. In our case this has extra benefit in reducing the number of partial differential equations (5.3.7) to solve.

However the original fast algorithm cannot guarantee that the velocity $v(\phi(x,t),t)$ is still dominated by $T_i$ when a point $x \in M_i$ at $t = 0$ moves to a new location $\phi(x,t)$ at time $t$. Thus it could not preserve the trajectory of each local affine region. This makes it an non-admissible solution to equation (5.3.5).

In order to preserve the nice property of stationary velocity, we propose a novel way to define a stationary velocity field which still satisfies equation (5.3.5). Define the trajectory of the region $M_i$ from time $t_1$ to $t_2$ as notation $M_i|_{t_1}^{t_2} = \cup_{\tau=t_1}^{t_2} \phi(M_i, \tau)$.

Without loss of generality, we first assume that these trajectories of the original subdomains do not overlap in the spatial domain $\Omega$. Define a new stationary velocity
which is a solution to the following equation.

\[
\begin{align*}
L^\dagger Lv &= 0 \\
v|_{(M^*_i, t)} &= v_i, \quad v|_{\partial \Omega} = 0
\end{align*}
\] (5.3.8)

By introducing the region trajectory \( M^*_i|_0 \), the proposed stationary velocity gives the same \( \phi \) with time-varying version within the trajectory over time.

When comparing the proposed stationary velocity in (5.3.8) and its generalized time-varying version in (5.3.7), one should notice that \( v(x, t) \) in (5.3.7) is defined on the spatial-temporal domain and \( v(x) \) in (5.3.7) is a “squeezed version” that collapses the region trajectory \( M^*_i \) along the temporal axis into the spatial domain.

### 5.3.4 Series of stationary velocity field for trajectory collision

One critical assumption in eliminating \( t \) from Equation is that all trajectories \( M^*_i|_0 \) do not overlap in the spatial domain, \( M^*_i|_0 \cap M^*_j|_0 = \emptyset \). Otherwise it will be ambiguous to determine which local affine transform should be used when a spatial point belongs to multiple trajectories.

This non-collision assumption is nevertheless not true in general. Consider two local affine transforms \( T_1 \) and \( T_2 \) defined on region \( M_1 \) and \( M_2 \) respectively. The end of trajectory of \( M_1 \), \( T_1(M_1) \) overlaps with the region of \( M_2 \) at \( t = 0 \). In this case, for any point \( y \in T_1(M_1) \cap M_2 \), it is ambiguous to define its weight \( v(x) \).

Our solution to this dilemma is to find a period from time \( t_1 \) to \( t_2 \) so that the
trajectories within this period are not overlapped. By the time $t = 1$ when $M_i$ moves to $T_1(M_i)$. In general we need at any time $\tau_1, \tau_2 \in [t_1, t_2]$, no local regions are overlapped, $\phi(M_i, \tau_1) \cap \phi(M_j, \tau_2) = \emptyset, \forall i \neq j$. When these trajectories are disjoint in the spatial-temporal domain, such $[t_1, t_2]$ is feasible. Thus it is possible to break the time from 0 to 1 into a sequence of $C + 1$ time points $[t_0, \cdots, t_C]$, such that

$$t_0 = 0 \text{ and } t_{k+1} = \max_{\tau} \{ \tau : M_i^*|_{t_k} \cap M_j^*|_{t_k}, \forall i \neq j \}$$

(5.3.9)

For each non-collision period $[t_{k-1}, t_k]$, a stationary velocity $v^k$ are defined in the same way as in (5.3.8). For each stationary velocity $v^k$, its transform $\phi^k$ is computed using the efficient scale-and-square method in Arsigny et al. (2009). The final transform $\phi$ is the concatenation of these $C$ diffeomorphism transforms which is still a diffeomorphism.

### 5.4 Implementation

#### 5.4.1 Computing Mask Trajectories and Their Collision

One important implementation detail is how to efficiently compute each region trajectory $M_i^*$ given its predefined velocity $v_i(x) = L_ix + u_i$. This is trivial when $M_i$ is just a single point. However, when $M_i$ is a region, there could be multiple trajectories. When a region can be efficiently parameterized as a polygon in 2D (or polyhedron in 3D), one can track the trajectory of each vertex and compute the
collision of any two polygons. Here we instead choose a simpler approximation by using a point set to represent an arbitrary region.

For a region $M$, a point set $\{p_a\}$ is uniformly sampled inside using sampling diameter $d$. The region $M$ transformed at time $t$, $T(M,t)$ is also approximated by the dilation of the point set $\{T(p_a,t)\}$. The condition of collision detection in equation (5.3.9) can be achieved.

5.4.2 Solving Elliptic PDE

As shown in the formulation of equation (5.3.8), the minimization of the energy of the velocity field is equivalent to solving the PDE with dirichlet boundary condition.

In our problem, we apply the finite element method (FEM) (Strang and Fix, 1973) to solve the elliptic PDE. FEM generally divides a continuous domain into a set of discrete subdomains and approximates solution space using a subspace with piecewise polynomial basis functions. Within each subdomain, simple element equations are introduced to locally approximate the original complex PDE. Then the element equations are combined into a global system of equations for the final calculation.

One advantage of FEM is that the formulation is based on rigorous theory of error estimation, which suits well for our original minimization purpose. In addition, the meshing in FEM is flexible so it can model complex geometry of the domain of the PDE. As an example, figure shows a generic meshing of the domain.
Figure 5.2: A sample triangular meshing for FEM. The sizes of the triangles can be adaptive.
5.5 Results

We evaluated the accuracy of our proposed approach using synthetic experiments. Two rigid transforms were defined on two polygon regions (one is a square and the other one is a trapezoid).

We compare the result between the time-dependent method and time-independent method and also use method in Arsigny et al. (2009) as the baseline. Both the cases with collisions and without are included.

In practice, the choice of operator \( L \) depends on the understanding of the problem. For simplicity, in this evaluation, we only use the \( L = \nabla \) for all cases which induces a Laplace equation for velocity field in the domain,

\[
\Delta v = 0.
\] (5.5.1)

5.5.1 Result without Trajectory Collisions

As shown in figure 5.3, the test case involves two rigid transformations whose trajectories are not overlapping. In figure 5.4 and 5.5, the triangular mesh and the velocity fields are shown for both the time-dependent and stationary methods. The final transformation generated by both time-dependent and stationary velocity fields can preserve the rigidity of the masked region nicely, as shown in figure 5.6.

Also the time cost for both methods is included in table 5.1. If solving the differential equation is the most time-consuming step in the whole scheme, then using a stationary velocity field is much faster than the time-dependent method.
Figure 5.3: Input transformations defined on two subdomains. Two input affine transformations are shown in red and black. The starting and the ending locations are indicated by the boxes in solid and dashed lines respectively.

<table>
<thead>
<tr>
<th>#time steps</th>
<th>time-dependent method</th>
<th>stationary method</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>50.47</td>
<td>6.45</td>
</tr>
<tr>
<td>20</td>
<td>100.42</td>
<td>6.47</td>
</tr>
<tr>
<td>100</td>
<td>490.32</td>
<td>6.72</td>
</tr>
</tbody>
</table>

Table 5.1: The time difference between the two different methods
Figure 5.4: First row: the mesh of the velocity field at $t = 0$ and $t = 1$ respectively.

Second row: the velocity field at $t = 0$ and $t = 1$ respectively.
Figure 5.5: Result given by the stationary velocity field. First row: the mesh for domain of the stationary velocity field. Second row: velocity field.
Figure 5.6: The Comparison between the final transformation result of the entire domain given by the stationary and time-dependent velocity fields. First row: result by time-dependent approach. Second row: result by stationary approach.
5.5.2 Result with Trajectory Collisions

For the case with trajectory collision, we cannot generate the final transformation with a single stationary velocity field. So the entire trajectory of each transformation is divided into several subtrajectories for the stationary method.

Figure 5.7: Input transformations defined on the two subdomains. The trajectories $M_1^*$ and $M_2^*$ overlap.

5.6 Discussion

In this work we presented a novel approach to constructing a polyaffine transform which can precisely preserve each affine transform using one diffeomorphism over the whole region. The polyaffine problem is formulated as finding a feasible solution with a new constraint that preserves the affine trajectories of each local region. The
Figure 5.8: First row: the mesh of the velocity field at $t = 0$ and $t = 1$ respectively.

Second row: the velocity field at $t = 0$ and $t = 1$ respectively.
Figure 5.9: Result given by the stationary velocity field. First row: the mesh for domain of the two stationary velocity fields before and after collision of the trajectories. Second row: two velocity fields before and after collision of the trajectories.
Figure 5.10: The Comparison between the final transformation result of the entire domain given by the stationary and time-dependent velocity fields. First row: result by time-dependent approach. Second row: result by stationary approach.
Table 5.2: The time difference between the two different methods

<table>
<thead>
<tr>
<th>#time steps</th>
<th>time-dependent method</th>
<th>stationary method</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>50.23</td>
<td>12.75</td>
</tr>
<tr>
<td>20</td>
<td>101.23</td>
<td>13.17</td>
</tr>
<tr>
<td>100</td>
<td>493.26</td>
<td>13.72</td>
</tr>
</tbody>
</table>

natural solution uses a time-varying velocity field by minimizing the energy of the transformation at domain outside the moving subregion.

Our approach instead uses a composition of one or more diffeomorphisms of stationary velocities and is thus both efficient in computation and accurate in preserving the new constraints.

In contrast, our new approach preserves the affine velocity for any time $t$, and therefore preserves the affine transform. Depending on the values of input affine transformation their regions, our approach requires one or more stationary velocity fields to preserve different affine velocities in each region over time. The number of stationary velocity fields are determined by the collision detection of region trajectories.
Chapter 6

Conclusion and Future works

6.1 Conclusion

In this dissertation, we present a novel framework to reconstruct 3D image from serial 2D image sections. This framework follows the standard setup in general image reconstruction. Our contribution is in devising an objective function that is unbiased to any reference slice and robust to noises and corrupted slices in the serial sections. In addition, the geometrical shape difference between sections is explicitly encoded via large deformation models which provide a mathematical method to quantitatively measure the shape change. The experiments have shown improvements over the past methods.

Our second contribution is in presenting an algorithm for constructing diffeomorphism of the entire domain by combining multiple affine transformations defined
in subdomains. The algorithm applies the interpolation following the variational minimization principle. In addition, the calculation time can be greatly reduced with a stationary velocity field by extending the prescribed velocity field on the local regions to trajectories of the local regions.

6.2 Future direction

First, there are a few unanswered questions regarding the parameter choices in the model. The choice of bandwidth and weight kernel should depend on the resolution of the model in $z$ direction and the assumption on the smoothness of the original volume. This regularity over the volume may be learnt from the past data where the manually inspected reconstructed volume is available.

In our setup, intensity-based similarity objective functions are employed in the image reconstruction for measuring appearance difference between images. However, feature-based methods may have better performance if accurate feature correspondence is available in the image dataset. It would be valuable to compare the performance of both intensity and feature-based methods on the data sets.

In addition, a few variations of the transformation model can be considered, which may provide a simpler formulation and reduce the total computational cost. The original LDDMM framework is very expensive in terms of both computational time and memory space. As we have mentioned in 3.3, models such as DARTEL (Ashburner, 2007) or log-Euclidean method (Vercauteren et al., 2008) can
save both memory and time. In theory, these methods only include subsets of the
diffeomorphism group, so more experiments are required to test whether they are
good approximation in practice.

In terms of practice concern, the preparation process may result in topological
changes such as tearings in the images, which cannot be recovered by diffeomorphic
transformations. This calls for an effective method in the preprocessing step to
detect or even correct this type of changes. There has been limited effort on the tear
correction by using information of the neighboring slices and it would be valuable
to integrate that into the framework.

Lastly, the reconstruction model we apply depends on the assumption that the
image sections are planar objects, which may not be a accurate approximation in
some cases. It is of great value to allow off-plane registration/reconstruction of the
slices, which is more accurate representation of the true anatomy.
Appendix A

Euler Lagrange equation for LDDMM

This appendix derives the Euler-Lagrange equation that is used to construct a gradient descent algorithm for the large-deformations diffeomorphic metric mapping problem. See Chapter “refer to the right chapter” for appropriate context and notation. This derivation was published in Beg thing and Beg thing. cite theorem provides a detailed derivation of the Euler-Lagrange equation. This theorem depends on the following lemma, which describes the effect of a perturbation of a velocity field on a deformation. Lemma B-0.2 (Beg (2003); Beg et al. (2005)). The variation of the mapping $v$

$V$ is the space of function $\Omega \rightarrow \mathbb{R}$ where and $v(\cdot, t) \in V, \forall t \in [0, 1]$. The variation of mapping $\varphi^\gamma_{s,t}$ when $v \in L^2([0, 1], V)$ is perturbed along $h \in L^2([0, 1], V)$ is given
by:

\[
\partial_h \varphi^v_{s,t} = \lim_{\varepsilon \to 0} \frac{\varphi^v_{s,t} + \varepsilon h - \varphi^v_{s,t}}{\varepsilon} = D\varphi^v_{s,t} \int_s^t (D\varphi^v_{s,u})^{-1} h_u \circ (\varphi^v_{s,u}) du. \tag{A.0.1}
\]

Proof. We provide a proof under the assumption that the derivative with respect to \( \varepsilon \) in Eq. (A.0.1) exists and proceed to its identification. The proof of existence can be carried on by standard ordinary differential equations (ODE) arguments. We have

\[
d_t \varphi^v_{s,t} + \varepsilon h = v_t \circ \varphi^v_{s,t} + \varepsilon h_t \circ \varphi^v_{s,t}.
\]

Computing the differential in \( \varepsilon \) at \( \varepsilon = 0 \) yields

\[
d_t \partial_h \varphi^v_{s,t} = D\varphi^v_{s,t} v \partial_h \varphi^v_{s,t} + h_t \circ \varphi^v_{s,t} \tag{A.0.2}
\]

Thus \( \partial_h \varphi^v_{s,t} \) is the solution of a non-homogeneous differential equation. We can find the solution from finding the homogeneous solution and use variation of a constant given the initial condition \( \partial_h \varphi^v_{s,s} = 0 \).

\[
\frac{d}{dt} D\varphi^v_{s,t} = D\varphi^v_{s,t} v D\varphi^v_{s,t}
\]

which simply comes from computing the space differential of \( d\varphi^v_{s,t}/dt = v_t \circ \varphi^v_{s,t} \) the homogeneous equation associated to.

So here we get the derivative of \( \varphi^v \) w.r.t. the velocity \( v \).
A.0.1 uniform derivative for the Energy to the flow

The total energy usually includes two parts: similarity and regularization.

\[ E = E_S + E_R \]  

(A.0.3)

In LDDMM, \( E_S \) does not explicitly depend on \( v \) but \( \varphi_v \), while the \( E_R \) depends on \( v \) directly. The derivative of \( E_R \) w.r.t. is relatively easy. So here mainly consider the derivative on the similarity term. Suppose we have an inner product \( \langle \cdot, \cdot \rangle \) for example \( L^2(\Omega) \). If

\[ \langle \nabla_{\varphi} E_S, h_{\varphi} \rangle_{L^2(\Omega)} = \lim_{\varepsilon \to 0} \frac{E_S(\varphi + \varepsilon h_{\varphi}) - E_S(\varphi)}{\varepsilon} \]  

(A.0.4)

holds for all test function \( h_{\varphi} \in L^2(\Omega) \), then we define \( \nabla_{\varphi} E_S \) the derivative of \( E_S \) w.r.t. \( \varphi \) in the \( L^2 \) sense, and if we have the derivative of \( \varphi \)

In this way, we could define the derivative of the energy to the velocity

\[ \langle \nabla_v E_S, h_v \rangle_{L^2(\Omega \times [0,1])} = \lim_{\varepsilon \to 0} \frac{E_S(\varphi^{v+\varepsilon v}) - E_S(\varphi^v)}{\varepsilon} \]

(A.0.5)

= \langle \nabla_{\varphi} E_S, \int_0^1 D\varphi_{t \to 1}^v h_v \circ \varphi_{0 \to t}^v dt \rangle_{L^2(\Omega)}

= \langle \nabla_{\varphi} E_S \circ \varphi_{0 \to t}^v, \int_0^1 |D\varphi_{0 \to t}^v| D\varphi_{t \to 1}^v h_v dt \rangle_{L^2(\Omega)}

= \langle (D\varphi_{t \to 1}^v)^\dagger |D\varphi_{0 \to t}^v| \nabla_{\varphi} E_S \circ \varphi_{0 \to t}^v, h_v \rangle_{L^2(\Omega \times [0,1])}

Then we could substitute different \( E_S \) into the formula above. Here \( I \) is the fixed image and \( J \) is the moving image.
Case 0: SSD (sum of squared difference)

\[ E_S = \int_\Omega |I - J \circ \varphi|^2 dx \]

\[ \nabla \varphi E_S = (I - J \circ \varphi) \nabla J(\varphi) \]

\[ \nabla \varphi E_S = (D\varphi_{t \to t+1})^\dagger |D\varphi_{0 \to t}| \nabla \varphi E_S \circ \varphi_{t \to 0} \]

\[ = (D\varphi_{t \to t+1})^\dagger |D\varphi_{0 \to t}| [(I - J \circ \varphi_{0 \to 1}) \nabla K(\varphi_{0 \to 1})] \circ \varphi_{t \to 0} \]

\[ = (D\varphi_{t \to t+1})^\dagger |D\varphi_{0 \to t}| [(I \circ \varphi_{t \to 0} - J \circ \varphi_{t \to 1}) \nabla J(\varphi_{t \to 1})] \]

\[ = |D\varphi_{0 \to t}| [(I \circ \varphi_{t \to 0} - J \circ \varphi_{t \to 1}) \nabla (J \circ \varphi_{t \to 1})] \]

The last equality in (A.0.6) holds by combining

Of course, this could be easily adapted to the symmetric domain case. Let’s say \( v_1(t) \) and \( v_2(t) \), \( t \in [0, 0.5] \)

\[ E_S = \int_\Omega |I \circ \varphi^{v_1} - J \circ \varphi^{v_2}|^2 dx \]

\[ \nabla v_1 E_S = |D\varphi_{0 \to t}^{v_1}| [(I \circ \varphi_{t \to 0.5}^{v_1} - J \circ \varphi_{t \to 0}^{v_2}) \nabla (I \circ \varphi_{t \to 0.5}^{v_1})] \]

\[ \nabla v_2 E_S = |D\varphi_{0 \to t}^{v_2}| [(I \circ \varphi_{t \to 0}^{v_1} - J \circ \varphi_{t \to 0.5}^{v_2}) \nabla (J \circ \varphi_{t \to 0.5}^{v_2})] \]

normalized cross correlation

\[ E_S = \int_\Omega CC^2(x) dx \]

\[ CC(x) = \nabla \varphi E_S = (I - J \circ \varphi) \nabla J(\varphi) \]

\[ \nabla \varphi E_S = (D\varphi_{t \to t+1})^\dagger |D\varphi_{0 \to t}^\varphi| \nabla \varphi E_S \circ \varphi_{t \to 0} \]

\[ = (D\varphi_{t \to t+1})^\dagger |D\varphi_{0 \to t}^\varphi| [(I - J \circ \varphi_{0 \to 1}) \nabla K(\varphi_{0 \to 1})] \circ \varphi_{t \to 0} \]

\[ = (D\varphi_{t \to t+1})^\dagger |D\varphi_{0 \to t}^\varphi| [(I \circ \varphi_{t \to 0} - J \circ \varphi_{t \to 1}) \nabla J(\varphi_{t \to 1})] \]

\[ = |D\varphi_{0 \to t}^\varphi| [(I \circ \varphi_{t \to 0} - J \circ \varphi_{t \to 1}) \nabla (J \circ \varphi_{t \to 1})] \]

97
Appendix B

Efficient adjoint calculation

Given the initial velocity $v_0 \in V$, or equivalently, the initial momentum $m_0 \in V^*$, we see in section that the geodesic path $\phi(t)$ is determined by the EPDiff equation.

We denote the deformed moving image by $J_t(x) = J \circ \varphi_t^{-1}(x)$ and it evolves via the equation

$$\frac{\partial J_t}{\partial t} + v_t \cdot \nabla J_t = 0. \tag{B.0.1}$$

This follows directly from the material derivatives in continuum mechanics.

Vialard et al. (2012) proposed an optimal control theory approach to calculate the gradient of the total energy of image matching.

$$\begin{aligned}
\frac{\partial J_t}{\partial t} + \nabla J_t \cdot v_t &= 0 \\
\frac{\partial m_t}{\partial t} + \ad_{v_t}^* m_t &= 0 \\
v_t - K \ast m_t &= 0
\end{aligned} \tag{B.0.2}$$
Theorem B.0.1. The gradient of $E$ is given by

$$\partial_{m_0} E = K \star m_0 - \dot{m}_0.$$  \hspace{1cm} (B.0.3)

**Proof.** In order to prove Theorem B.0.1, we use the standard method of a time dependent Lagrange multiplier to derive the adjoint equations. Let $\hat{J}, \hat{m}, \hat{v}$ be the adjoint variables for respectively $J, m, v$ (the variable $\hat{m}$ is in fact redundant). We then extend the functional $E$ with the Lagrange multipliers to get

$$\hat{E} = E + \int_0^1 \langle \hat{m}, \delta(\dot{m} + \text{ad}_v^* m) \rangle_{L^2} + \int_0^1 \langle \hat{J}, \dot{J} + \nabla \cdot v \rangle_{L^2} dt + \int_0^1 \langle \hat{v}, K \star m - v \rangle_{L^2} \hspace{1cm} (B.0.4)$$

First, we use the definition of the adjoint action

$$\text{ad}_v(m) = -[v, m]$$

to derive the variation on the term $\langle \hat{m}, \delta(\dot{m} + \text{ad}_v^* m) \rangle_{L^2}$

$$\delta \langle \hat{m}, \dot{m} + \text{ad}_v^* m \rangle_{L^2}$$

$$= \langle \dot{m}, \delta \dot{m} \rangle_{L^2} + \delta \langle \text{ad}_v \dot{m}, m \rangle_{L^2}$$

$$= \langle \dot{m}, \delta \dot{m} \rangle_{L^2} - \delta \langle [v, \dot{m}], m \rangle_{L^2}$$

$$= \langle \dot{m}, \delta \dot{m} \rangle_{L^2} - \langle \text{ad}_v \delta v, m \rangle_{L^2} - \langle [v, \dot{m}], \delta m \rangle_{L^2}$$

$$= \langle \dot{m}, \delta \dot{m} \rangle_{L^2} - \langle \text{ad}_v \delta v, m \rangle_{L^2} - \langle [v, \dot{m}], \delta m \rangle_{L^2}$$

$$= \langle \dot{m}, \delta \dot{m} \rangle_{L^2} - \langle \text{ad}_v \delta v, m \rangle_{L^2} - \langle [v, \dot{m}], \delta m \rangle_{L^2}$$
Then recall the integration by part formula

\[ \int_{\Omega} f \nabla \cdot v dx + \int_{\Omega} (\nabla f)^T v dx = \int_{\partial\Omega} n^T v ds \]  

(B.0.6)

with \( n \) denoting the normalized outward-pointing normal vector to the boundary \( \partial\Omega \). In our situation, we assume that the kernel \( K \) vanishes on \( \partial\Omega \) such that for any vector field \( v \) satisfying the last equation of

\[ \langle f, \nabla \cdot (vg) \rangle_{L^2(\Omega)} = -\langle v \cdot \nabla f, g \rangle_{L^2(\Omega)} \]  

(B.0.7)

For the general things here we denote the similarity function \( E_S = S(I,J) \), where \( I \) and \( J \) are fixed and moving image respectively. We now take variations of the augmented functional using the previous formula to get.

\[ \delta E = 2\lambda \langle K \ast m_0, \delta m_0 \rangle_{L^2} + \langle \frac{\partial S(I,J)}{\partial J}, \delta J_1 \rangle \]

\[ + \int_0^1 \{ -\langle \text{ad}_m^* m, \delta v \rangle_{L^2} - \langle [v, \dot{m}], \delta m \rangle_{L^2} - \langle \dot{m}, \delta m \rangle_{L^2} \} dt + \langle \dot{m}_1, \delta m_1 \rangle_{L^2} - \langle \dot{m}_0, \delta m_0 \rangle_{L^2} \]

\[ + \int_0^1 \{ \langle -\dot{J} - \nabla \cdot (J v), \delta J \rangle_{L^2} + \langle \dot{J} \nabla J, \delta v \rangle_{L^2} \} dt + \langle \dot{J}_1, \delta J_1 \rangle_{L^2} - \langle \dot{J}_0, \delta J_0 \rangle_{L^2} \]

\[ + \int_0^1 \{ \langle K \ast \dot{v}, \delta m \rangle_{L^2} - \langle \dot{v}, \delta v \rangle_{L^2} \} dt \]  

(B.0.8)

In this way we can get the adjoint equation in terms of \( \dot{J}, \dot{m}, \dot{v} \)

\[ \begin{align*}
- \dot{m} + \text{ad}_v \dot{m} + K \ast \dot{v} &= 0 \\
\dot{J} + \nabla \cdot (J \dot{v}) &= 0 \\
- \text{ad}_m^* m + J \nabla J - \dot{v} &= 0
\end{align*} \]  

(B.0.9)
subject to

\[
\begin{align*}
\dot{J}_1 + \frac{\partial S(I, J)}{\partial J} &= 0 \\
\dot{m}_1 &= 0
\end{align*}
\]

(B.0.10)
Bibliography


J. Dauguet, T. Delzescaux, F. Condé, J.-F. Mangin, N. Ayache, P. Hantraye, and V. Frouin. Three-dimensional reconstruction of stained histological slices and


of histologic images of a whole rat brain for three-dimensional reconstruction.

